Partial Differential Equations

Lawrence C. Evans

Graduate Studies in Mathematics

Volume 19

American Mathematical Society

CONTENTS

PART III: THEORY FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

APPENDICES

PREFACE

I present in this book a wide-ranging survey of many important topics in the theory of partial differential equations (PDE), with particular emphasis on various modern approaches. I have made a huge number of editorial decisions about what to keep and what to toss out, and can only claim that this selection seems to me about right. I of course include the usual formulas for solutions of the usual linear PDE, but also devote large amounts of exposition to energy methods within Sobolev spaces, to the calculus of variations, to conservation laws, etc.

My general working principles in the writing have been these:

a. PDE theory is (mostly) not restricted to two independent variables. Many texts describe PDE as if functions of the two variables (x, y) or (x, t) were all that matter. This emphasis seems to me misleading, as modern discoveries concerning many types of equations, both linear and nonlinear, have allowed for the rigorous treatment of these in any number of dimensions. I also find it unsatisfactory to "classify" partial differential equations: this is possible in two variables, but creates the false impression that there is some kind of general and useful classification scheme available in general.

b. Many interesting equations are nonlinear. My view is that overall we know too much about linear PDE and too little about nonlinear PDE. I have accordingly introduced nonlinear concepts early in the text and have tried hard to emphasize everywhere nonlinear analogues of the linear theory.

c. Understanding generalized solutions is fundamental. Many of the partial differential equations we study, especially nonlinear first-order equations, do not in general possess smooth solutions. It is therefore essential to devise some kind of proper notion of generalized or weak solution. This is an important but subtle undertaking, and much of the hardest material in this book concerns the uniqueness of appropriately defined weak solutions.

d. PDE theory is not a branch of functional analysis. Whereas certain classes of equations can profitably be viewed as generating abstract operators between Banach spaces, the insistence on an overly abstract viewpoint, and consequent ignoring of deep calculus and measure theoretic estimates, is ultimately limiting.

e. Notation is a nightmare. I have really tried to introduce consistent notation, which works for all the important classes of equations studied. This attempt is sometimes at variance with notational conventions within a subarea.

f. Good theory is (almost) as useful as exact formulas. I incorporate this principle into the overall organization of the text, which is subdivided into three parts, roughly mimicking the historical development of PDE theory itself. Part I concerns the search for explicit formulas for solutions, and Part II the abandoning of this quest in favor of general theory asserting the existence and other properties of solutions for linear equations. Part III is the mostly modern endeavor of fashioning general theory for important classes of nonlinear PDE.

Let me also explicitly comment here that I intend the development within each section to be rigorous and complete (exceptions being the frankly heuristic treatment of asymptotics in §4.5 and an occasional reference to a research paper). This means that even locally within each chapter the topics do not necessarily progress logically from "easy" to "hard" concepts. There are many difficult proofs and computations early on, but as compensation many easier ideas later. The student should certainly omit on first reading some of the more arcane proofs.

I wish next to emphasize that this is a *textbook*, and not a reference book. I have tried everywhere to present the essential ideas in the clearest possible settings, and therefore have almost never established sharp versions of any of the theorems. Research articles and advanced monographs, many of them listed in the Bibliography, provide such precision and generality.

My goal has rather been to explain, as best I can, the many fundamental ideas of the subject within fairly simple contexts.

I have greatly profited from the comments and thoughtful suggestions of many of my colleagues, friends and students, in particular: S. Antman, J. Bang, X. Chen, A. Chorin, M. Christ, J. Cima, P. Colella, J. Cooper, M. Crandall, B. Driver, M. Feldman, M. Fitzpatrick, R. Gariepy, J. Goldstein, D. Gomes, O. Hald, W. Han, W. Hrusa, T. Ilmanen, I. Ishii, I. Israel, R. Jerrard, C. Jones, B. Kawohl, S. Koike, J. Lewis, T.-P. Liu, H. Lopes, J. McLaughlin, K. Miller, J. Morford, J. Neu, M. Portilheiro, J. Ralston, F. Rezakhanlou, W. Schlag, D. Serre, P. Souganidis, J. Strain, W. Strauss, M. Struwe, R. Temam, B. Tvedt, J.-L. Vazquez, M. Weinstein, P. Wolfe, and Y. Zheng.

I especially thank Tai-Ping Liu for many years ago writing out for me the first draft of what is now Chapter 11.

I am extremely grateful for the suggestions and lists of mistakes from earlier drafts of this book sent me by many readers, and I encourage others to send me their comments, at evans@math.berkeley.edu. I have come to realize that I must be more than slightly mad to try to write a book of this length and complexity, but I am not yet crazy enough to think that I have made no mistakes. **I will therefore maintain a listing of errors which come to light, and will make this accessible through the math.berkeley.edu homepage.**

Faye Yeager at UC Berkeley has done a really magnificent job typing and updating these notes, and Jaya Nagendra heroically typed an earlier version at the University of Maryland. My deepest thanks to both.

I have been supported by the NSF during much of the writing, most recently under grant DMS-9424342.

LCE
August, 1997
Berkeley

Chapter 1

INTRODUCTION

This chapter surveys the principal theoretical issues concerning the solving of partial differential equations.

To follow the subsequent discussion, the reader should first of all turn to Appendix A and look over the notation presented there, particularly the multiindex notation for partial derivatives.

1.1. PARTIAL DIFFERENTIAL EQUATIONS

A *partial differential equation (PDE)* is an equation involving an unknown function of two or more variables and certain of its partial derivatives.

Using the notation explained in Appendix A, we can write out symbolically a typical PDE, as follows. Fix an integer $k \geq 1$ and let U denote an open subset of $\mathbb{R}^n$.

DEFINITION. *An expression of the form*

$$F(D^k u(x), D^{k-1}u(x), \dots, Du(x), u(x), x) = 0 \quad (x \in U) \tag{1}$$

is called a k$^{\text{th}}$-order partial differential equation, *where*

$$F : \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times \cdots \times \mathbb{R}^n \times \mathbb{R} \times U \to \mathbb{R}$$

is given, and

$$u : U \to \mathbb{R}$$

is the unknown.

We *solve* the PDE if we find all u verifying (1), possibly only among those functions satisfying certain auxiliary boundary conditions on some part Γ of ∂U. By finding the solutions we mean, ideally, obtaining simple, explicit solutions, or, failing that, deducing the existence and other properties of solutions.

DEFINITIONS.

(i) *The partial differential equation* (1) *is called* linear *if it has the form*

$$\sum_{|\alpha|\leq k} a_\alpha(x) D^\alpha u = f(x)$$

for given functions a_α $(|\alpha| \leq k), f$. *This linear PDE is* homogeneous *if* $f \equiv 0$.

(ii) *The PDE* (1) *is* semilinear *if it has the form*

$$\sum_{|\alpha|=k} a_\alpha(x) D^\alpha u + a_0(D^{k-1}u, \ldots, Du, u, x) = 0.$$

(iii) *The PDE* (1) *is* quasilinear *if it has the form*

$$\sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \ldots, Du, u, x) D^\alpha u + a_0(D^{k-1}u, \ldots, Du, u, x) = 0.$$

(iv) *The PDE* (1) *is* fully nonlinear *if it depends nonlinearly upon the highest order derivatives.*

A *system* of partial differential equations is, informally speaking, a collection of several PDE for several unknown functions.

DEFINITION. *An expression of the form*

$$\mathbf{F}(D^k\mathbf{u}(x), D^{k-1}\mathbf{u}(x), \ldots, D\mathbf{u}(x), \mathbf{u}(x), x) = \mathbf{0} \quad (x \in U) \tag{2}$$

is called a k^{th}-order system of partial differential equations, *where*

$$\mathbf{F} : \mathbb{R}^{mn^k} \times \mathbb{R}^{mn^{k-1}} \times \cdots \times \mathbb{R}^{mn} \times \mathbb{R}^m \times U \to \mathbb{R}^m$$

is given and

$$\mathbf{u} : U \to \mathbb{R}^m, \ \mathbf{u} = (u^1, \ldots, u^m)$$

is the unknown.

Here we are supposing that the system comprises the same number m of scalar equations as unknowns $(u^1, \ldots, u^m)$. This is the most common circumstance, although other systems may have fewer or more equations than unknowns.

Systems are classified in the obvious way as being linear, semilinear, etc.

Remark. We use "PDE" as an abbreviation for both "partial differential equation" and "partial differential equation*s*". □

1.2. EXAMPLES

There is no general theory known concerning the solvability of all partial differential equations. Such a theory is extremely unlikely to exist, given the rich variety of physical, geometric, and probabilistic phenomena which can be modeled by PDE. Instead, research focuses on various particular partial differential equations that are important for applications within and outside of mathematics, with the hope that insight from the origins of these PDE can give clues as to their solutions.

Following is a list of many specific partial differential equations of interest in current research. This listing is intended merely to familiarize the reader with the names and forms of various famous PDE. To display most clearly the mathematical structure of these equations, we have mostly set relevant physical constants to unity. We will later discuss the origin and interpretation of many of these PDE.

Throughout $x \in U$, where U is an open subset of $\mathbb{R}^n$, and $t \geq 0$. Also $Du = D_x u = (u_{x_1}, \ldots, u_{x_n})$ denotes the gradient of u with respect to the spatial variable $x = (x_1, \ldots, x_n)$.

1.2.1. Single partial differential equations.

a. Linear equations.

1. *Laplace's equation*

$$\Delta u = \sum_{i=1}^{n} u_{x_i x_i} = 0.$$

2. *Helmholtz's (or eigenvalue) equation*

$$-\Delta u = \lambda u.$$

3. *Linear transport equation*

$$u_t + \sum_{i=1}^{n} b^i u_{x_i} = 0.$$

4. *Liouville's equation*

$$u_t - \sum_{i=1}^{n} (b^i u)_{x_i} = 0.$$

5. *Heat (or diffusion) equation*

$$u_t - \Delta u = 0.$$

6. *Schrödinger's equation*

$$iu_t + \Delta u = 0.$$

7. *Kolmogorov's equation*

$$u_t - \sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} = 0.$$

8. *Fokker–Planck equation*

$$u_t - \sum_{i,j=1}^{n} (a^{ij} u)_{x_i x_j} - \sum_{i=1}^{n} (b^i u)_{x_i} = 0.$$

9. *Wave equation*

$$u_{tt} - \Delta u = 0.$$

10. *Telegraph equation*

$$u_{tt} + du_t - u_{xx} = 0.$$

11. *General wave equation*

$$u_{tt} - \sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} = 0.$$

12. *Airy's equation*

$$u_t + u_{xxx} = 0.$$

13. *Beam equation*

$$u_t + u_{xxxx} = 0.$$

b. Nonlinear equations.

1. *Eikonal equation*

$$|Du| = 1.$$

2. *Nonlinear Poisson equation*

$$-\Delta u = f(u).$$

3. *p-Laplacian equation*

$$\operatorname{div}(|Du|^{p-2}Du) = 0.$$

4. *Minimal surface equation*

$$\operatorname{div}\left(\frac{Du}{(1+|Du|^2)^{1/2}}\right) = 0.$$

5. *Monge–Ampère equation*

$$\det(D^2u) = f.$$

6. *Hamilton–Jacobi equation*

$$u_t + H(Du, x) = 0.$$

7. *Scalar conservation law*

$$u_t + \operatorname{div}\mathbf{F}(u) = 0.$$

8. *Inviscid Burgers' equation*

$$u_t + uu_x = 0.$$

9. *Scalar reaction-diffusion equation*

$$u_t - \Delta u = f(u).$$

10. *Porous medium equation*

$$u_t - \Delta(u^\gamma) = 0.$$

11. *Nonlinear wave equations*

$$u_{tt} - \Delta u = f(u),$$
$$u_{tt} - \operatorname{div}\mathbf{a}(Du) = 0.$$

12. *Korteweg–deVries (KdV) equation*

$$u_t + uu_x + u_{xxx} = 0.$$

1.2.2. Systems of partial differential equations.

a. Linear systems.

1. *Equilibrium equations of linear elasticity*

$$\mu\Delta\mathbf{u} + (\lambda + \mu)D(\operatorname{div}\mathbf{u}) = \mathbf{0}.$$

2. *Evolution equations of linear elasticity*

$$\mathbf{u}_{tt} - \mu\Delta\mathbf{u} - (\lambda + \mu)D(\operatorname{div}\mathbf{u}) = \mathbf{0}.$$

3. *Maxwell's equations*

$$\begin{cases} \mathbf{E}_t = \operatorname{curl}\mathbf{B} \\ \mathbf{B}_t = -\operatorname{curl}\mathbf{E} \\ \operatorname{div}\mathbf{B} = \operatorname{div}\mathbf{E} = 0. \end{cases}$$

b. Nonlinear systems.

1. *System of conservation laws*

$$\mathbf{u}_t + \operatorname{div}\mathbf{F}(\mathbf{u}) = \mathbf{0}.$$

2. *Reaction-diffusion system*

$$\mathbf{u}_t - \Delta\mathbf{u} = \mathbf{f}(\mathbf{u}).$$

3. *Euler's equations for incompressible, inviscid flow*

$$\begin{cases} \mathbf{u}_t + \mathbf{u}\cdot D\mathbf{u} = -Dp \\ \operatorname{div}\mathbf{u} = 0. \end{cases}$$

4. *Navier–Stokes equations for incompressible, viscous flow*

$$\begin{cases} \mathbf{u}_t + \mathbf{u}\cdot D\mathbf{u} - \Delta\mathbf{u} = -Dp \\ \operatorname{div}\mathbf{u} = 0. \end{cases}$$

See Zwillinger [**ZW**] for a much more extensive listing of interesting PDE.

1.3. STRATEGIES FOR STUDYING PDE

As explained in §1.1 our goal is the discovery of ways to solve partial differential equations of various sorts, but—as should now be clear in view of the many diverse examples set forth in §1.2—this is no easy task. And indeed the very question of what it means to "solve" a given PDE can be subtle, depending in large part on the particular structure of the problem at hand.

1.3.1. Well-posed problems, classical solutions.

The informal notion of a well-posed problem captures many of the desirable features of what it means to solve a PDE. We say that a given problem for a partial differential equation is *well-posed* if

(a) the problem in fact has a solution;

(b) this solution is unique;

and

(c) the solution depends continuously on the data given in the problem.

The last condition is particularly important for problems arising from physical applications: we would prefer that our (unique) solution changes only a little when the conditions specifying the problem change a little. (For many problems, on the other hand, uniqueness is not to be expected. In these cases the primary mathematical tasks are to classify and characterize the solutions.)

Now clearly it would be desirable to "solve" PDE in such a way that (a)–(c) hold. But notice that we still have not carefully defined what we mean by a "solution". Should we ask, for example, that a "solution" u must be real analytic or at least infinitely differentiable? This might be desirable, but perhaps we are asking too much. Maybe it would be wiser to require a solution of a PDE of order k to be at least k times continuously differentiable. Then at least all the derivatives which appear in the statement of the PDE will exist and be continuous, although maybe certain higher derivatives will not exist. Let us informally call a solution with this much smoothness a *classical* solution of the PDE: this is certainly the most obvious notion of solution.

So by solving a partial differential equation in the classical sense we mean if possible to write down a formula for a classical solution satisfying (a)–(c) above, or at least to show such a solution exists, and to deduce various of its properties.

1.3.2. Weak solutions and regularity.

But can we achieve this? The answer is that certain specific partial differential equations (e.g. Laplace's equation) can be solved in the classical sense, but many others, if not most others, cannot. Consider for instance the scalar conservation law

$$u_t + F(u)_x = 0.$$

We will see in §3.4 that this PDE governs various one-dimensional phenomena involving fluid dynamics, and in particular models the formation and propagation of shock waves. Now a shock wave is a curve of discontinuity of the solution u; and so if we wish to study conservation laws, and recover the underlying physics, we must surely allow for solutions u which are not continuously differentiable or even continuous. In general, as we shall see, the conservation law has no classical solutions, but *is* well-posed if we allow for properly defined *generalized* or *weak solutions*.

This is all to say that we may be forced by the structure of the particular equation to abandon the search for smooth, classical solutions. We must instead, while still hoping to achieve the well-posedness conditions (a)–(c), investigate a wider class of candidates for solutions. And in fact, even for those PDE which turn out to be classically solvable, it is often most expedient initially to search for some appropriate kind of weak solution.

The point is this: if from the outset we demand that our solutions be very regular, say k-times continuously differentiable, then we are usually going to have a really hard time finding them, as our proofs must then necessarily include possibly intricate demonstrations that the functions we are building are in fact smooth enough. A far more reasonable strategy is to consider as separate the *existence* and the *smoothness* (or *regularity*) problems. The idea is to define for a given PDE a reasonably wide notion of a *weak solution*, with the expectation that since we are not asking too much by way of smoothness of this weak solution, it may be easier to establish its existence, uniqueness, and continuous dependence on the given data. Thus, to repeat, it is often wise to aim at proving well-posedness in some appropriate class of weak or generalized solutions.

Now, as noted above, for various partial differential equations this is the best that can be done. For other equations we can hope that our weak solution may turn out after all to be smooth enough to qualify as a classical solution. This leads to the question of *regularity* of weak solutions. As we will see, it is often the case that the existence of weak solutions depends upon rather simple estimates plus ideas of functional analysis, whereas the regularity of the weak solutions, when true, usually rests upon many intricate calculus estimates.

Let me explicitly note here that once we are past Part I (Chapters 2–4), our efforts will be largely devoted to proving mathematically the existence of solutions to various sorts of partial differential equations, and not so much to deriving formulas for these solutions. This may seem wasted or misguided effort, but in fact mathematicians are like theologians: we regard existence as the prime attribute of what we study. But unlike most theologians, we need not always rely upon faith alone.

1.3.3. Typical difficulties.

Following are some vague but general principles, which may be useful to keep in mind:

(1) Nonlinear equations are more difficult than linear equations; and, indeed, the more the nonlinearity affects the higher derivatives, the more difficult the PDE is.
(2) Higher-order PDE are more difficult than lower-order PDE.
(3) Systems are harder than single equations.
(4) Partial differential equations entailing many independent variables are harder than PDE entailing few independent variables.
(5) For most partial differential equations it is not possible to write out explicit formulas for solutions.

None of these assertions is without important exceptions.

1.4. OVERVIEW

This textbook is divided into three major Parts.

PART I: Representation Formulas for Solutions

Here we identify those important partial differential equations for which in certain circumstances explicit or more-or-less explicit formulas can be had for solutions. The general progression of the exposition is from direct formulas for certain linear equations, to far less concrete representation formulas, of a sort, for various nonlinear PDE.

Chapter 2 is a detailed study of four exactly solvable partial differential equations: the linear transport equation, Laplace's equation, the heat equation, and the wave equation. These PDE, which serve as archetypes for the more complicated equations introduced later, admit directly computable solutions, at least in the case that there is no domain whose boundary geometry complicates matters. The explicit formulas are augmented by various indirect, but easy and attractive, "energy"-type arguments, which serve as motivation for the developments in Chapters 6, 7 and thereafter.

Chapter 3 continues the theme of searching for explicit formulas, now for general first-order nonlinear PDE. The key insight is that such PDE can, locally at least, be transformed into systems of ordinary differential equations (ODE), the characteristic equations. We stipulate that once the problem becomes "only" the question of integrating a system of ODE, it is in principle solved, sometimes quite explicitly. The derivation of the characteristic equations given in the text is very simple and does not require any geometric insights. It is in truth so easy to derive the characteristic equations that no real purpose is had by dealing with the quasilinear case first.

We introduce also the Hopf–Lax formula for Hamilton–Jacobi equations (§3.3) and the Lax–Oleinik formula for scalar conservation laws (§3.4). (Some knowledge of measure theory is useful here, but is not essential.) These sections provide an early acquaintance with the global theory of these important nonlinear PDE, and so motivate the later Chapters 10 and 11.

Chapter 4 is a grab bag of techniques for explicitly (or kind of explicitly) solving various linear and nonlinear partial differential equations, and the reader should study only whatever seems interesting. The section on the Fourier transform is, however, essential. The Cauchy-Kovalevskaya Theorem appears at the very end. Although this is basically the only general existence theorem in the subject, and thus logically should perhaps be regarded as central, in practice these power series methods are not so prevalent.

PART II: Theory for Linear Partial Differential Equations

Next we abandon the search for explicit formulas and instead rely on functional analysis and relatively easy "energy" estimates to prove the existence of weak solutions to various linear PDE. We investigate also the uniqueness and regularity of such solutions, and deduce various other properties.

Chapter 5 is an introduction to Sobolev spaces, the proper setting for the study of many linear and nonlinear partial differential equations via energy methods. This is a hard chapter, the real worth of which is only later revealed, and requires some basic knowledge of Lebesgue measure theory. However, the requirements are not really so great, and the review in Appendix E should suffice. In my opinion there is no particular advantage in considering only the Sobolev spaces with exponent $p = 2$, and indeed insisting upon this obscures the two central inequalities, those of Gagliardo–Nirenberg–Sobolev (§5.6.1) and of Morrey (§5.6.2).

In Chapter 6 we vastly generalize our knowledge of Laplace's equation to other second-order elliptic equations. Here we work through a rather complete treatment of existence, uniqueness and regularity theory for solutions, including the maximum principle, and also a reasonable introduction to the study of eigenvalues, including a discussion of the principal eigenvalue for nonselfadjoint operators.

Chapter 7 expands the energy methods to a variety of linear partial differential equations characterizing evolutions in time. We broaden our earlier investigation of the heat equation to general second-order parabolic PDE, and of the wave equation to general second-order hyperbolic PDE. We study as well linear first-order hyperbolic systems, with the aim of motivating the developments concerning nonlinear systems of conservation laws in Chapter 11. The concluding section 7.4 presents the alternative functional analytic method of semigroups for building solutions.

(Missing from this long Part II on linear partial differential equations is any discussion of distribution theory or potential theory. These are important topics, but for our purposes seem dispensable, even in a book of such length. These omissions do not slow us up much, and make room for more nonlinear theory.)

PART III: Theory for Nonlinear Partial Differential Equations

This section parallels for nonlinear PDE the development in Part II, but is far less unified in its approach, as the various types of nonlinearity must be treated in quite different ways.

Chapter 8 commences the general study of nonlinear partial differential equations with an extensive discussion of the calculus of variations. Here we set forth a careful derivation of the direct method for deducing the existence of minimizers, and discuss also a variety of variational systems and constrained problems, as well as minimax methods. Variational theory is the most useful and accessible of the methods for nonlinear PDE, and so this chapter is fundamental.

Chapter 9 is, rather like Chapter 4 before, a gathering of assorted other techniques of use for nonlinear elliptic and parabolic partial differential equations. We encounter here monotonicity and fixed-point methods, and a variety of other devices, mostly involving the maximum principle. We study as well certain nice aspects of nonlinear semigroup theory, to complement the linear semigroup theory from Chapter 7.

Chapter 10 is an introduction to the modern theory of Hamilton–Jacobi PDE, and in particular to the notion of "viscosity solutions". We encounter also the connections with the optimal control of ODE, through dynamic programming.

Chapter 11 picks up from Chapter 3 the discussion of conservation laws, now systems of conservation laws. Unlike the general theoretical developments in Chapters 5–9, for which Sobolev spaces provide the proper abstract framework, we are forced to employ here direct linear algebra and calculus computations. We pay particular attention to the solution of Riemann's problem and to entropy criteria.

Appendices A–E provide for the reader's convenience some background material, with selected proofs, on inequalities, linear functional analysis, measure theory, etc.

The Bibliography primarily provides a listing of interesting PDE books to consult for further information. Since this is a textbook , and not a reference monograph, I have mostly not attempted to track down and document the original sources for the myriads of ideas and methods we will encounter. The mathematical literature for partial differential equations is truly vast, but the books cited in the Bibliography should at least provide a starting point for locating the primary sources.

1.5. PROBLEMS

1. Classify each of the partial differential equations in §1.2 as follows:
   (a) Is the PDE linear, semilinear, quasilinear or fully nonlinear?
   (b) What is the order of the PDE?

The next exercises provide some practice with the multiindex notation introduced in Appendix A.

2. Prove the *Multinomial Theorem*

$$(x_1 + \ldots + x_n)^k = \sum_{|\alpha|=k} \binom{|\alpha|}{\alpha} x^\alpha,$$

   where $\binom{|\alpha|}{\alpha} := \frac{|\alpha|!}{\alpha!}$, $\alpha! = \alpha_1!\alpha_2!\ldots\alpha_n!$, and $x^\alpha = x_1^{\alpha_1}\ldots x_n^{\alpha_n}$. The sum is taken over all multiindices $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $|\alpha| = k$.

3. Prove *Leibniz' formula*

$$D^\alpha(uv) = \sum_{\beta\leq\alpha} \binom{\alpha}{\beta} D^\beta u D^{\alpha-\beta} v,$$

   where $u, v : \mathbb{R}^n \to \mathbb{R}$ are smooth, $\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!}$, and $\beta \leq \alpha$ means $\beta_i \leq \alpha_i$ $(i = 1, \ldots, n)$.

4. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is smooth. Prove

$$f(x) = \sum_{|\alpha|\leq k} \frac{1}{\alpha!} D^\alpha f(0) x^\alpha + O(|x|^{k+1}) \quad \text{as } x \to 0$$

for each $k = 1, 2, \ldots$. This is *Taylor's formula* in multiindex notation. (Hint: Fix $x \in \mathbb{R}^n$ and consider the function of one variable $g(t) := f(tx)$.)

Chapter 2

FOUR IMPORTANT LINEAR PARTIAL DIFFERENTIAL EQUATIONS

In this chapter we introduce four fundamental linear partial differential equations for which various explicit formulas for solutions are available. These are

the transport equation	$u_t + b \cdot Du = 0$	(§2.1),
Laplace's equation	$\Delta u = 0$	(§2.2),
the heat equation	$u_t - \Delta u = 0$	(§2.3),
the wave equation	$u_{tt} - \Delta u = 0$	(§2.4).

Before going further, the reader should review the discussions of inequalities, integration by parts, Green's formulas, convolutions, etc. in Appendices B and C, and later refer back to these as necessary.

2.1. TRANSPORT EQUATION

Probably the simplest partial differential equation of all is the *transport equation* with constant coefficients. This is the PDE

$$u_t + b \cdot Du = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty), \tag{1}$$

where b is a fixed vector in $\mathbb{R}^n$, $b = (b_1, \dots, b_n)$, and $u : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ is the unknown, $u = u(x,t)$. Here $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ denotes a typical point in space, and $t \geq 0$ denotes a typical time. We write $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$ for the gradient of u with respect to the spatial variables x.

Which functions u solve (1)? To answer, let us suppose for the moment we are given some smooth solution u and try to compute it. To do so, we first must recognize that the partial differential equation (1) asserts that a particular directional derivative of u vanishes. We exploit this insight by fixing any point $(x,t) \in \mathbb{R}^n \times (0,\infty)$ and defining

$$z(s) := u(x + sb, t + s) \quad (s \in \mathbb{R}).$$

We then calculate

$$\dot{z}(s) = Du(x + sb, t + s) \cdot b + u_t(x + sb, t + s) = 0 \quad \left(\cdot = \frac{d}{ds} \right),$$

the second equality holding owing to (1). Thus $z(\cdot)$ is a constant function of s, and consequently for each point (x,t), u is constant on the line through (x,t) with the direction $(b,1) \in \mathbb{R}^{n+1}$. Hence if we know the value of u at any point on each such line, we know its value everywhere in $\mathbb{R}^n \times (0,\infty)$.

2.1.1. Initial-value problem.

For definiteness therefore, let us consider the initial-value problem

$$\begin{cases} u_t + b \cdot Du = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{2}$$

Here $b \in \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$ are known, and the problem is to compute u. Given (x,t) as above, the line through (x,t) with direction $(b,1)$ is represented parametrically by $(x + sb, t + s)$ $(s \in \mathbb{R})$. This line hits the plane $\Gamma := \mathbb{R}^n \times \{t = 0\}$ when $s = -t$, at the point $(x - tb, 0)$. Since u is constant on the line and $u(x - tb, 0) = g(x - tb)$, we deduce

$$u(x,t) = g(x - tb) \quad (x \in \mathbb{R}^n, t \geq 0). \tag{3}$$

So, *if* (2) has a sufficiently regular solution u, it must certainly be given by (3). And conversely, it is easy to check directly that if g is C^1, then u defined by (3) is indeed a solution of (2).

Remark. If g is not C^1, then there is obviously no C^1 solution of (2). But even in this case formula (3) certainly provides a strong, and in fact the only reasonable, candidate for a solution. We may thus informally declare $u(x,t) = g(x-tb)$ $(x \in \mathbb{R}^n,\ t \geq 0)$ to be a *weak solution* of (2), even should g not be C^1. This all makes sense even if g, and thus u, are discontinuous. Such a notion, that a nonsmooth or even discontinuous function may sometimes solve a PDE, will come up again later when we study nonlinear transport phenomena in §3.4. □

2.1.2. Nonhomogeneous problem.

Next let us look at the associated nonhomogeneous problem

$$\text{(4)} \qquad \begin{cases} u_t + b \cdot Du = f & \text{in } \mathbb{R}^n \times (0,\infty) \\ \qquad\quad u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

As before fix $(x,t) \in \mathbb{R}^{n+1}$ and, inspired by the calculation above, set $z(s) := u(x+sb, t+s)$ for $s \in \mathbb{R}$. Then

$$\dot{z}(s) = Du(x+sb, t+s) \cdot b + u_t(x+sb, t+s) = f(x+sb, t+s).$$

Consequently

$$\begin{aligned} u(x,t) - g(x-bt) = z(0) - z(-t) &= \int_{-t}^{0} \dot{z}(s)\, ds \\ &= \int_{-t}^{0} f(x+sb, t+s)\, ds \\ &= \int_{0}^{t} f(x+(s-t)b, s)\, ds; \end{aligned}$$

and so

$$\text{(5)} \qquad u(x,t) = g(x-tb) + \int_0^t f(x+(s-t)b, s)\, ds \quad (x \in \mathbb{R}^n,\ t \geq 0)$$

solves the initial-value problem (4).

We will later employ this formula to solve the one-dimensional wave equation, in §2.4.1.

Remark. Observe that we have derived our solutions (3), (5) by in effect converting the partial differential equations into ordinary differential equations. This procedure is a special case of the *method of characteristics*, developed later in §3.2. □

2.2. LAPLACE'S EQUATION

Among the most important of all partial differential equations are undoubtedly *Laplace's equation*

(1)
$$\Delta u = 0$$

and *Poisson's equation*

(2)
$$-\Delta u = f.^{*}$$

In both (1) and (2), $x \in U$ and the unknown is $u : \bar{U} \to \mathbb{R}$, $u = u(x)$, where $U \subset \mathbb{R}^n$ is a given open set. In (2) the function $f : U \to \mathbb{R}$ is also given. Remember from §A.3 that the *Laplacian* of u is $\Delta u = \sum_{i=1}^{n} u_{x_i x_i}$.

DEFINITION. *A C^2 function u satisfying (1) is called a* harmonic *function.*

Physical interpretation. Laplace's equation comes up in a wide variety of physical contexts. In a typical interpretation u denotes the density of some quantity (e.g. a chemical concentration) in equilibrium. Then if V is any smooth subregion within U, the net flux of u through ∂V is zero:

$$\int_{\partial V} \mathbf{F} \cdot \nu \, dS = 0,$$

$\mathbf{F}$ denoting the flux density and ν the unit outer normal field. In view of the Gauss–Green Theorem (§C.2), we have

$$\int_V \operatorname{div} \mathbf{F} \, dx = \int_{\partial V} \mathbf{F} \cdot \nu \, dS = 0,$$

and so

(3)
$$\operatorname{div} \mathbf{F} = 0 \quad \text{in } U,$$

since V was arbitrary. In many instances it is physically reasonable to assume the flux $\mathbf{F}$ is proportional to the gradient Du, but points in the opposite direction (since the flow is from regions of higher to lower concentration). Thus

(4)
$$\mathbf{F} = -aDu \quad (a > 0).$$

*I prefer to write (2) with the minus sign, to be consistent with the notation for general second-order elliptic operators in Chapter 6.

Substituting into (3), we obtain Laplace's equation

$$\operatorname{div}(Du) = \Delta u = 0.$$

If u denotes the

$$\begin{cases} \text{chemical concentration} \\ \text{temperature} \\ \text{electrostatic potential,} \end{cases}$$

equation (4) is

$$\begin{cases} \text{Fick's law of diffusion} \\ \text{Fourier's law of heat conduction} \\ \text{Ohm's law of electrical conduction.} \end{cases}$$

See Feynman–Leighton–Sands [**F-L-S,** Chapter 12] for a discussion of the ubiquity of Laplace's equation in mathematical physics. Laplace's equation arises as well in the study of analytic functions and the probabilistic investigation of Brownian motion. □

2.2.1. Fundamental solution.

a. Derivation of fundamental solution.

One good strategy for investigating any partial differential equation is first to identify some explicit solutions and then, provided the PDE is linear, to assemble more complicated solutions out of the specific ones previously noted. Furthermore, in looking for explicit solutions it is often wise to restrict attention to classes of functions with certain symmetry properties. Since Laplace's equation is invariant under rotations (Problem 2), it consequently seems advisable to search first for *radial* solutions, that is, functions of $r = |x|$.

Let us therefore attempt to find a solution u of Laplace's equation (1) in $U = \mathbb{R}^n$, having the form

$$u(x) = v(r),$$

where $r = |x| = (x_1^2 + \cdots + x_n^2)^{1/2}$ and v is to be selected (if possible) so that $\Delta u = 0$ holds. First note for $i = 1, \ldots, n$ that

$$\frac{\partial r}{\partial x_i} = \frac{1}{2}\left(x_1^2 + \cdots + x_n^2\right)^{-1/2} 2x_i = \frac{x_i}{r} \quad (x \neq 0).$$

We thus have

$$u_{x_i} = v'(r)\frac{x_i}{r}, \quad u_{x_i x_i} = v''(r)\frac{x_i^2}{r^2} + v'(r)\left(\frac{1}{r} - \frac{x_i^2}{r^3}\right)$$

for $i = 1, \dots, n$, and so

$$\Delta u = v''(r) + \frac{n-1}{r} v'(r).$$

Hence $\Delta u = 0$ if and only if

(5) $$v'' + \frac{n-1}{r} v' = 0.$$

If $v' \neq 0$, we deduce

$$\log(v')' = \frac{v''}{v'} = \frac{1-n}{r},$$

and hence $v'(r) = \frac{a}{r^{n-1}}$ for some constant a. Consequently if $r > 0$, we have

$$v(r) = \begin{cases} b \log r + c & (n = 2) \\ \frac{b}{r^{n-2}} + c & (n \geq 3), \end{cases}$$

where b and c are constants.

These considerations motivate the following

DEFINITION. *The function*

(6) $$\Phi(x) := \begin{cases} -\frac{1}{2\pi} \log |x| & (n = 2) \\ \frac{1}{n(n-2)\alpha(n)} \frac{1}{|x|^{n-2}} & (n \geq 3), \end{cases}$$

defined for $x \in \mathbb{R}^n$, $x \neq 0$, *is the* fundamental solution of Laplace's equation.

The reason for the particular choices of the constants in (6) will be apparent in a moment. (Recall from §A.2 that $\alpha(n)$ denotes the volume of the unit ball in $\mathbb{R}^n$.)

We will sometimes slightly abuse notation and write $\Phi(x) = \Phi(|x|)$ to emphasize that the fundamental solution is radial. Observe also that we have the estimates

(7) $$|D\Phi(x)| \leq \frac{C}{|x|^{n-1}}, \ |D^2\Phi(x)| \leq \frac{C}{|x|^n} \quad (x \neq 0)$$

for some constant $C > 0$.

b. Poisson's equation.

By construction the function $x \mapsto \Phi(x)$ is harmonic for $x \neq 0$. If we shift the origin to a new point y, the PDE (1) is unchanged; and so $x \mapsto \Phi(x-y)$ is also harmonic as a function of x, $x \neq y$. Let us now take $f : \mathbb{R}^n \to \mathbb{R}$ and note that the mapping $x \mapsto \Phi(x-y)f(y)$ $(x \neq y)$ is harmonic for each point

$y \in \mathbb{R}^n$, and thus so is the sum of finitely many such expressions built for different points y.

This reasoning might suggest that the convolution

$$
u(x) = \int_{\mathbb{R}^n} \Phi(x-y) f(y)\, dy
$$

(8)
$$
= \begin{cases} -\frac{1}{2\pi} \int_{\mathbb{R}^2} \log(|x-y|) f(y)\, dy & (n=2) \\ \frac{1}{n(n-2)\alpha(n)} \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-2}}\, dy & (n \geq 3) \end{cases}
$$

will solve Laplace's equation (1). *However, this is wrong*: we *cannot* just compute

(9)
$$
\Delta u(x) = \int_{\mathbb{R}^n} \Delta_x \Phi(x-y) f(y)\, dy = 0.
$$

Indeed, as intimated by estimate (7), $D^2\Phi(x-y)$ is *not* summable near the singularity at $y = x$, and so the differentiation under the integral sign above is unjustified (and incorrect). We must proceed more carefully in calculating Δu.

Let us for simplicity now assume $f \in C_c^2(\mathbb{R}^n)$; that is, f is twice continuously differentiable, with compact support.

THEOREM 1 (Solving Poisson's equation). *Define u by* (8). *Then*

(i) $u \in C^2(\mathbb{R}^n)$

and

(ii) $-\Delta u = f$ *in* $\mathbb{R}^n$.

We consequently see that (8) provides us with a formula for a solution of Poisson's equation (2) in $\mathbb{R}^n$.

Proof. 1. We have

$$
u(x) = \int_{\mathbb{R}^n} \Phi(x-y) f(y)\, dy = \int_{\mathbb{R}^n} \Phi(y) f(x-y)\, dy;
$$

hence

$$
\frac{u(x+he_i) - u(x)}{h} = \int_{\mathbb{R}^n} \Phi(y) \left[\frac{f(x+he_i-y) - f(x-y)}{h} \right] dy,
$$

where $h \neq 0$ and $e_i = (0, \ldots, 1, \ldots, 0)$, the 1 in the i^{th}-slot. But

$$
\frac{f(x+he_i-y) - f(x-y)}{h} \to \frac{\partial f}{\partial x_i}(x-y)
$$

uniformly on $\mathbb{R}^n$ as $h \to 0$, and thus

$$\frac{\partial u}{\partial x_i}(x) = \int_{\mathbb{R}^n} \Phi(y) \frac{\partial f}{\partial x_i}(x-y)\, dy \quad (i = 1, \ldots, n).$$

Similarly

$$\frac{\partial^2 u}{\partial x_i \partial x_j}(x) = \int_{\mathbb{R}^n} \Phi(y) \frac{\partial^2 f}{\partial x_i \partial x_j}(x-y)\, dy \quad (i, j = 1, \ldots, n). \tag{10}$$

As the expression on the right hand side of (10) is continuous in the variable x, we see $u \in C^2(\mathbb{R}^n)$.

2. Since Φ blows up at 0, we will need for subsequent calculations to isolate this singularity inside a small ball. So fix $\varepsilon > 0$. Then

$$\begin{aligned} \Delta u(x) &= \int_{B(0,\varepsilon)} \Phi(y) \Delta_x f(x-y)\, dy + \int_{\mathbb{R}^n - B(0,\varepsilon)} \Phi(y) \Delta_x f(x-y)\, dy \\ &=: I_\varepsilon + J_\varepsilon. \end{aligned} \tag{11}$$

Now

$$|I_\varepsilon| \le C \|D^2 f\|_{L^\infty(\mathbb{R}^n)} \int_{B(0,\varepsilon)} |\Phi(y)|\, dy \le \begin{cases} C\varepsilon^2 |\log \varepsilon| & (n = 2) \\ C\varepsilon^2 & (n \ge 3). \end{cases} \tag{12}$$

An integration by parts (see §C.2) yields

$$\begin{aligned} J_\varepsilon &= \int_{\mathbb{R}^n - B(0,\varepsilon)} \Phi(y) \Delta_y f(x-y)\, dy \\ &= -\int_{\mathbb{R}^n - B(0,\varepsilon)} D\Phi(y) \cdot D_y f(x-y)\, dy \\ &\quad + \int_{\partial B(0,\varepsilon)} \Phi(y) \frac{\partial f}{\partial \nu}(x-y)\, dS(y) \\ &=: K_\varepsilon + L_\varepsilon, \end{aligned} \tag{13}$$

ν denoting the *inward* pointing unit normal along $\partial B(0, \varepsilon)$. We readily check

$$|L_\varepsilon| \le \|Df\|_{L^\infty(\mathbb{R}^n)} \int_{\partial B(0,\varepsilon)} |\Phi(y)|\, dS(y) \le \begin{cases} C\varepsilon |\log \varepsilon| & (n = 2) \\ C\varepsilon & (n \ge 3). \end{cases} \tag{14}$$

3. We continue by integrating by parts once again in the term K_ε, to discover

$$\begin{aligned} K_\varepsilon &= \int_{\mathbb{R}^n - B(0,\varepsilon)} \Delta\Phi(y) f(x-y)\, dy - \int_{\partial B(0,\varepsilon)} \frac{\partial \Phi}{\partial \nu}(y) f(x-y)\, dS(y) \\ &= -\int_{\partial B(0,\varepsilon)} \frac{\partial \Phi}{\partial \nu}(y) f(x-y)\, dS(y), \end{aligned}$$

since Φ is harmonic away from the origin. Now $D\Phi(y) = \frac{-1}{n\alpha(n)} \frac{y}{|y|^n}$ $(y \neq 0)$ and $\nu = \frac{-y}{|y|} = -\frac{y}{\varepsilon}$ on $\partial B(0,\varepsilon)$. Consequently $\frac{\partial \Phi}{\partial \nu}(y) = \nu \cdot D\Phi(y) = \frac{1}{n\alpha(n)\varepsilon^{n-1}}$ on $\partial B(0,\varepsilon)$. Since $n\alpha(n)\varepsilon^{n-1}$ is the surface area of the sphere $\partial B(0,\varepsilon)$, we have

$$
\begin{aligned}
K_\varepsilon &= -\frac{1}{n\alpha(n)\varepsilon^{n-1}} \int_{\partial B(0,\varepsilon)} f(x-y)\, dS(y) \\
&= -⨍_{\partial B(x,\varepsilon)} f(y)\, dS(y) \to -f(x) \quad \text{as } \varepsilon \to 0.
\end{aligned} \tag{15}
$$

(Remember from §A.3 that a slash through an integral denotes an average.)

4. Combining now (11)–(15) and letting $\varepsilon \to 0$, we find $-\Delta u(x) = f(x)$, as asserted. □

Remarks. (i) We sometimes write

$$-\Delta \Phi = \delta_0 \quad \text{in } \mathbb{R}^n,$$

δ_0 denoting the Dirac measure on $\mathbb{R}^n$ giving unit mass to the point 0. Adopting this notation, we may formally compute:

$$
\begin{aligned}
-\Delta u(x) &= \int_{\mathbb{R}^n} -\Delta_x \Phi(x-y)\, f(y) dy \\
&= \int_{\mathbb{R}^n} \delta_x f(y)\, dy = f(x) \quad (x \in \mathbb{R}^n),
\end{aligned}
$$

in accordance with Theorem 1. This corrects the erroneous calculation (9).

(ii) Theorem 1 is in fact valid under far less stringent smoothness requirements for f: see Gilbarg–Trudinger [**G-T**]. □

2.2.2. Mean-value formulas.

Consider now an open set $U \subset \mathbb{R}^n$ and suppose u is a harmonic function within U. We next derive the important *mean-value formulas*, which declare that $u(x)$ equals both the average of u over the sphere $\partial B(x,r)$ and the average of u over the entire ball $B(x,r)$, provided $B(x,r) \subset U$. These implicit formulas involving u generate a remarkable number of consequences, as we will momentarily see.

THEOREM 2 (Mean-value formulas for Laplace's equation). *If $u \in C^2(U)$ is harmonic, then*

$$u(x) = ⨍_{\partial B(x,r)} u\, dS = ⨍_{B(x,r)} u\, dy \tag{16}$$

for each ball $B(x,r) \subset U$.

Proof. 1. Set

$$\phi(r) := \fint_{\partial B(x,r)} u(y)\,dS(y) = \fint_{\partial B(0,1)} u(x+rz)\,dS(z).$$

Then

$$\phi'(r) = \fint_{\partial B(0,1)} Du(x+rz)\cdot z\,dS(z),$$

and consequently, using Green's formulas from §C.2, we compute

$$\begin{aligned}\phi'(r) &= \fint_{\partial B(x,r)} Du(y)\cdot \frac{y-x}{r}\,dS(y)\\ &= \fint_{\partial B(x,r)} \frac{\partial u}{\partial \nu}\,dS(y)\\ &= \frac{r}{n}\fint_{B(x,r)} \Delta u(y)\,dy = 0.\end{aligned}$$

Hence ϕ is constant, and so

$$\phi(r) = \lim_{t\to 0}\phi(t) = \lim_{t\to 0}\fint_{\partial B(x,t)} u(y)\,dS(y) = u(x).$$

2. Observe next that our employing polar coordinates, as in §C.3, gives

$$\begin{aligned}\int_{B(x,r)} u dy &= \int_0^r \left(\int_{\partial B(x,s)} u\,dS\right) ds\\ &= u(x)\int_0^r n\alpha(n)s^{n-1}ds = \alpha(n)r^n u(x).\end{aligned}$$

□

THEOREM 3 (Converse to mean-value property). *If $u \in C^2(U)$ satisfies*

$$u(x) = \fint_{\partial B(x,r)} u\,dS$$

for each ball $B(x,r) \subset U$, then u is harmonic.

Proof. If $\Delta u \not\equiv 0$, there exists some ball $B(x,r) \subset U$ such that, say, $\Delta u > 0$ within $B(x,r)$. But then for ϕ as above,

$$0 = \phi'(r) = \frac{r}{n}\fint_{B(x,r)} \Delta u(y)\,dy > 0,$$

a contradiction. □

2.2.3. Properties of harmonic functions.

We now present a sequence of interesting deductions about harmonic functions, all based upon the mean-value formulas. Assume for the following that $U \subset \mathbb{R}^n$ is open and bounded.

a. Strong maximum principle, uniqueness.

THEOREM 4 (Strong maximum principle). *Suppose* $u \in C^2(U) \cap C(\bar{U})$ *is harmonic within* U.

(i) *Then*

$$\max_{\bar{U}} u = \max_{\partial U} u.$$

(ii) *Furthermore, if* U *is connected and there exists a point* $x_0 \in U$ *such that*

$$u(x_0) = \max_{\bar{U}} u,$$

then

$$u \textit{ is constant within } U.$$

Assertion (i) is the *maximum principle* for Laplace's equation and (ii) is the *strong maximum principle.* Replacing u by $-u$, we recover also similar assertions with "min" replacing "max".

Proof. Suppose there exists a point $x_0 \in U$ with $u(x_0) = M := \max_{\bar{U}} u$. Then for $0 < r < \operatorname{dist}(x_0, \partial U)$, the mean-value property asserts

$$M = u(x_0) = ⨍_{B(x_0,r)} u \, dy \leq M.$$

As equality holds only if $u \equiv M$ within $B(x_0, r)$, we see $u(y) = M$ for all $y \in B(x, r)$. Hence the set $\{x \in U \mid u(x) = M\}$ is both open and relatively closed in U, and thus equals U if U is connected. This proves assertion (ii), from which (i) follows. □

Remark. The strong maximum principle asserts in particular that if U is connected and $u \in C^2(U) \cap C(\bar{U})$ satisfies

$$\begin{cases} \Delta u = 0 & \text{in } U \\ u = g & \text{on } \partial U, \end{cases}$$

where $g \geq 0$, then u is positive *everywhere* in U if g is positive *somewhere* on ∂U. □

An important application of the maximum principle is establishing the uniqueness of solutions to certain boundary-value problems for Poisson's equation.

THEOREM 5 (Uniqueness). *Let $g \in C(\partial U)$, $f \in C(U)$. Then there exists at most one solution $u \in C^2(U) \cap C(\bar{U})$ of the boundary-value problem*

$$\begin{cases} -\Delta u = f & \text{in } U \\ \quad u = g & \text{on } \partial U. \end{cases} \tag{17}$$

Proof. If u and $\tilde{u}$ both satisfy (17), apply Theorem 4 to the harmonic functions $w := \pm(u - \tilde{u})$. □

b. Regularity.

Now we prove that if $u \in C^2$ is harmonic, then necessarily $u \in C^\infty$. Thus *harmonic functions are automatically infinitely differentiable.* This sort of assertion is called a *regularity* theorem. The interesting point is that the algebraic structure of Laplace's equation $\Delta u = \sum_{i=1}^n u_{x_i x_i} = 0$ leads to the analytic deduction that all the partial derivatives of u exist, even those which do not appear in the PDE.

THEOREM 6 (Smoothness). *If $u \in C(U)$ satisfies the mean-value property* (16) *for each ball $B(x,r) \subset U$, then*

$$u \in C^\infty(U).$$

Note carefully that u may not be smooth, or even continuous, up to ∂U.

Proof. Let η be a standard mollifier, as described in §C.4, and recall that η is a radial function. Set $u^\varepsilon := \eta_\varepsilon * u$ in $U_\varepsilon = \{x \in U \mid \operatorname{dist}(x, \partial U) > \varepsilon\}$. As shown in §C.4, $u^\varepsilon \in C^\infty(U_\varepsilon)$.

We will prove u is smooth by demonstrating that in fact $u \equiv u^\varepsilon$ on U_ε. Indeed if $x \in U_\varepsilon$, then

$$\begin{aligned} u^\varepsilon(x) &= \int_U \eta_\varepsilon(x-y) u(y)\, dy \\ &= \frac{1}{\varepsilon^n} \int_{B(x,\varepsilon)} \eta\left(\frac{|x-y|}{\varepsilon}\right) u(y)\, dy \\ &= \frac{1}{\varepsilon^n} \int_0^\varepsilon \eta\left(\frac{r}{\varepsilon}\right) \left(\int_{\partial B(x,r)} u\, dS \right) dr \\ &= \frac{1}{\varepsilon^n} u(x) \int_0^\varepsilon \eta\left(\frac{r}{\varepsilon}\right) n\alpha(n) r^{n-1} dr \quad \text{by (16)} \\ &= u(x) \int_{B(0,\varepsilon)} \eta_\varepsilon\, dy = u(x). \end{aligned}$$

Thus $u^\varepsilon \equiv u$ in U_ε, and so $u \in C^\infty(U_\varepsilon)$ for each $\varepsilon > 0$. □

c. Local estimates for harmonic functions.

Next we employ the mean-value formulas to derive careful estimates on the various partial derivatives of a harmonic function. The precise structure of these estimates will be needed below, when we prove analyticity.

THEOREM 7 (Estimates on derivatives). *Assume u is harmonic in U. Then*

$$|D^\alpha u(x_0)| \leq \frac{C_k}{r^{n+k}} \|u\|_{L^1(B(x_0,r))} \tag{18}$$

for each ball $B(x_0, r) \subset U$ and each multiindex α of order $|\alpha| = k$.

Here

$$C_0 = \frac{1}{\alpha(n)}, \; C_k = \frac{(2^{n+1}nk)^k}{\alpha(n)} \quad (k = 1, \dots). \tag{19}$$

Proof. 1. We establish (18), (19) by induction on k, the case $k = 0$ being immediate from the mean-value formula (16). For $k = 1$, we note upon differentiating Laplace's equation that u_{x_i} $(i = 1, \dots, n)$ is harmonic. Consequently

$$\begin{aligned} |u_{x_i}(x_0)| &= \left| \fint_{B(x_0,r/2)} u_{x_i}\, dx \right| \\ &= \left| \frac{2^n}{\alpha(n) r^n} \int_{\partial B(x_0,r/2)} u \nu_i \, dS \right| \\ &\leq \frac{2n}{r} \|u\|_{L^\infty(\partial B(x_0,\frac{r}{2}))}. \end{aligned} \tag{20}$$

Now if $x \in \partial B(x_0, r/2)$, then $B(x, r/2) \subset B(x_0, r) \subset U$, and so

$$|u(x)| \leq \frac{1}{\alpha(n)} \left(\frac{2}{r}\right)^n \|u\|_{L^1(B(x_0,r))}$$

by (18), (19) for $k = 0$. Combining the inequalities above, we deduce

$$|D^\alpha u(x_0)| \leq \frac{2^{n+1} n}{\alpha(n)} \frac{1}{r^{n+1}} \|u\|_{L^1(B(x_0,r))}$$

if $|\alpha| = 1$. This verifies (18), (19) for $k = 1$.

2. Assume now $k \geq 2$ and (18), (19) is valid for all balls in U and each multiindex of order less than or equal to $k - 1$. Fix $B(x_0, r) \subset U$ and let α

be a multiindex with $|\alpha| = k$. Then $D^\alpha u = (D^\beta u)_{x_i}$ for some $i \in \{1, \dots, n\}$, $|\beta| = k - 1$. By calculations similar to those in (20), we establish that

$$|D^\alpha u(x_0)| \leq \frac{nk}{r} \|D^\beta u\|_{L^\infty(\partial B(x_0, \frac{r}{k}))}.$$

If $x \in \partial B(x_0, \frac{r}{k})$, then $B(x, \frac{k-1}{k}r) \subset B(x_0, r) \subset U$. Thus (18), (19) for $k - 1$ imply

$$|D^\beta u(x)| \leq \frac{(2^{n+1}n(k-1))^{k-1}}{\alpha(n)\left(\frac{k-1}{k}r\right)^{n+k-1}} \|u\|_{L^1(B(x_0,r))}.$$

Combining the two previous estimates yields the bound

$$|D^\alpha u(x_0)| \leq \frac{(2^{n+1}nk)^k}{\alpha(n) r^{n+k}} \|u\|_{L^1(B(x_0,r))}. \tag{21}$$

This confirms (18), (19) for $|\alpha| = k$. □

d. Liouville's Theorem.

Next we see that there are no nontrivial bounded harmonic functions on all of $\mathbb{R}^n$.

THEOREM 8 (Liouville's Theorem). *Suppose $u : \mathbb{R}^n \to \mathbb{R}$ is harmonic and bounded. Then u is constant.*

Proof. Fix $x_0 \in \mathbb{R}^n$, $r > 0$, and apply Theorem 7 on $B(x_0, r)$:

$$\begin{aligned} |Du(x_0)| &\leq \frac{C_1}{r^{n+1}} \|u\|_{L^1(B(x_0,r))} \\ &\leq \frac{C_1 \alpha(n)}{r} \|u\|_{L^\infty(\mathbb{R}^n)} \to 0, \end{aligned}$$

as $r \to \infty$. Thus $Du \equiv 0$, and so u is constant. □

THEOREM 9 (Representation formula). *Let $f \in C_c^2(\mathbb{R}^n)$, $n \geq 3$. Then any bounded solution of*

$$-\Delta u = f \quad \text{in } \mathbb{R}^n$$

has the form

$$u(x) = \int_{\mathbb{R}^n} \Phi(x - y) f(y)\,dy + C \quad (x \in \mathbb{R}^n)$$

for some constant C.

Proof. Since $\Phi(x) \to 0$ as $|x| \to \infty$ for $n \geq 3$, $\tilde{u}(x) := \int_{\mathbb{R}^n} \Phi(x-y)f(y)dy$ is a bounded solution of $-\Delta u = f$ in $\mathbb{R}^n$. If u is another solution, $w := u - \tilde{u}$ is constant, according to Liouville's Theorem. □

Remark. If $n = 2$, $\Phi(x) = -\frac{1}{2\pi}\log|x|$ is unbounded as $|x| \to \infty$, and so may be $\int_{\mathbb{R}^2} \Phi(x-y)f(y)\,dy$. □

e. Analyticity.

Next we refine Theorem 6:

THEOREM 10 (Analyticity). *Assume u is harmonic in U. Then u is analytic in U.*

Proof. 1. Fix any point $x_0 \in U$. We must show u can be represented by a convergent power series in some neighborhood of x_0.

Let $r := \frac{1}{4}\operatorname{dist}(x_0, \partial U)$. Then $M := \frac{1}{\alpha(n)r^n}\|u\|_{L^1(B(x_0,2r))} < \infty$.

2. Since $B(x,r) \subset B(x_0,2r) \subset U$ for each $x \in B(x_0,r)$, Theorem 7 provides the bound

$$\|D^\alpha u\|_{L^\infty(B(x_0,r))} \leq M\left(\frac{2^{n+1}n}{r}\right)^{|\alpha|}|\alpha|^{|\alpha|}.$$

Now Stirling's formula ([**RD**, §8.22]) asserts $\lim_{k\to\infty}\frac{k^{k+\frac{1}{2}}}{k!e^k} = \frac{1}{(2\pi)^{1/2}}$. Hence

$$|\alpha|^{|\alpha|} \leq Ce^{|\alpha|}|\alpha|!$$

for some constant C and all multiindices α. Furthermore, the Multinomial Theorem implies

$$n^k = (1+\cdots+1)^k = \sum_{|\alpha|=k}\frac{|\alpha|!}{\alpha!};$$

whence

$$|\alpha|! \leq n^{|\alpha|}\alpha!.$$

Combining the previous inequalities now yields

$$\|D^\alpha u\|_{L^\infty(B(x_0,r))} \leq CM\left(\frac{2^{n+1}n^2e}{r}\right)^{|\alpha|}\alpha!. \tag{22}$$

3. The Taylor series for u at x_0 is

$$\sum_\alpha \frac{D^\alpha u(x_0)}{\alpha!}(x-x_0)^\alpha,$$

the sum taken over all multiindices. We assert this power series converges, provided

$$|x - x_0| < \frac{r}{2^{n+2}n^3 e}. \tag{23}$$

To verify this, let us compute for each N the remainder term:

$$\begin{aligned} R_N(x) &:= u(x) - \sum_{k=0}^{N-1} \sum_{|\alpha|=k} \frac{D^\alpha u(x_0)(x-x_0)^\alpha}{\alpha!} \\ &= \sum_{|\alpha|=N} \frac{D^\alpha u(x_0 + t(x-x_0))(x-x_0)^\alpha}{\alpha!} \end{aligned}$$

for some $0 \leq t \leq 1$, t depending on x. We establish this formula by writing out the first N terms and the error in the Taylor expansion about 0 for the function of one variable $g(t) := u(x_0 + t(x - x_0))$, at $t = 1$. Employing (22), (23), we can estimate

$$\begin{aligned} |R_N(x)| &\leq CM \sum_{|\alpha|=N} \left(\frac{2^{n+1}n^2 e}{r}\right)^N \left(\frac{r}{2^{n+2}n^3 e}\right)^N \\ &\leq CMn^N \frac{1}{(2n)^N} = \frac{CM}{2^N} \\ &\to 0 \quad \text{as } N \to \infty. \end{aligned}$$

□

See §4.6.2 for more on analytic functions and partial differential equations.

f. Harnack's inequality.

Recall from §A.2 that we write $V \subset\subset U$ to mean $V \subset \bar{V} \subset U$ and $\bar{V}$ is compact.

THEOREM 11 (Harnack's inequality). *For each connected open set $V \subset\subset U$, there exists a positive constant C, depending only on V, such that*

$$\sup_V u \leq C \inf_V u$$

for all nonnegative harmonic functions u in U.

Thus in particular

$$\frac{1}{C} u(y) \leq u(x) \leq Cu(y)$$

for all points $x, y \in V$. These inequalities assert that *the values of a nonnegative harmonic function within V are all comparable*: u cannot be very small (or very large) at any point of V unless u is very small (or very large) everywhere in V. The intuitive idea is that since V is a positive distance away from ∂U, there is "room for the averaging effects of Laplace's equation to occur".

Proof. Let $r := \frac{1}{4}\operatorname{dist}(V, \partial U)$. Choose $x, y \in V$, $|x - y| \leq r$. Then

$$\begin{aligned} u(x) &= ⨍_{B(x,2r)} u\,dz \geq \frac{1}{\alpha(n)2^n r^n}\int_{B(y,r)} u\,dz \\ &= \frac{1}{2^n} ⨍_{B(y,r)} u\,dz = \frac{1}{2^n}u(y). \end{aligned}$$

Thus $2^n u(y) \geq u(x) \geq \frac{1}{2^n}u(y)$ if $x, y \in V$, $|x - y| \leq r$.

Since V is connected and $\bar{V}$ is compact, we can cover $\bar{V}$ by a chain of finitely many balls $\{B_i\}_{i=1}^N$, each of which has radius r and $B_i \cap B_{i-1} \neq \emptyset$ for $i = 2, \ldots, N$. Then

$$u(x) \geq \frac{1}{2^{nN}}u(y)$$

for all $x, y \in V$. □

2.2.4. Green's function.

Assume now $U \subset \mathbb{R}^n$ is open, bounded, and ∂U is C^1. We propose next to obtain a general representation formula for the solution of Poisson's equation

$$-\Delta u = f \quad \text{in } U,$$

subject to the prescribed boundary condition

$$u = g \quad \text{on } \partial U.$$

a. Derivation of Green's function.

Suppose first of all $u \in C^2(\bar{U})$ is an arbitrary function. Fix $x \in U$, choose $\varepsilon > 0$ so small that $B(x, \varepsilon) \subset U$, and apply Green's formula from §C.2 on the region $V_\varepsilon := U - B(x, \varepsilon)$ to $u(y)$ and $\Phi(y - x)$. We thereby compute

$$\begin{aligned} &\int_{V_\varepsilon} u(y)\Delta\Phi(y-x) - \Phi(y-x)\Delta u(y)\,dy \\ &\quad = \int_{\partial V_\varepsilon} u(y)\frac{\partial\Phi}{\partial\nu}(y-x) - \Phi(y-x)\frac{\partial u}{\partial\nu}(y)\,dS(y), \end{aligned} \tag{24}$$

ν denoting the outer unit normal vector on ∂V_ε. Recall next $\Delta\Phi(x-y)=0$ for $x \neq y$. We observe also

$$\Big|\int_{\partial B(x,\varepsilon)} \Phi(y-x)\frac{\partial u}{\partial \nu}(y)\,dS(y)\Big| \leq C\varepsilon^{n-1}\max_{\partial B(0,\varepsilon)}|\Phi| = o(1)$$

as $\varepsilon \to 0$. Furthermore the calculations in the proof of Theorem 1 show

$$\int_{\partial B(x,\varepsilon)} u(y)\frac{\partial \Phi}{\partial \nu}(y-x)\,dS(y) = \fint_{\partial B(x,\varepsilon)} u(y)\,dS(y) \to u(x)$$

as $\varepsilon \to 0$. Hence our sending $\varepsilon \to 0$ in (24) yields the formula:

$$\begin{aligned} u(x) = &\int_{\partial U} \Phi(y-x)\frac{\partial u}{\partial \nu}(y) - u(y)\frac{\partial \Phi}{\partial \nu}(y-x)\,dS(y) \\ &- \int_U \Phi(y-x)\Delta u(y)\,dy. \end{aligned} \tag{25}$$

This identity is valid for any point $x \in U$ and any function $u \in C^2(\bar{U})$.

Now formula (25) would permit us to solve for $u(x)$ *if* we knew the values of Δu within U and the values of $u, \partial u/\partial\nu$ along ∂U. However for our application to Poisson's equation with prescribed boundary values for u, the normal derivative $\partial u/\partial \nu$ along ∂U is unknown to us. We must therefore somehow modify (25) to remove this term.

The idea is now to introduce for fixed x a *corrector* function $\phi^x = \phi^x(y)$, solving the boundary-value problem:

$$\begin{cases} \Delta \phi^x = 0 & \text{in } U \\ \phi^x = \Phi(y-x) & \text{on } \partial U. \end{cases} \tag{26}$$

Let us apply Green's formula once more, now to compute

$$\begin{aligned} -\int_U \phi^x(y)\Delta u(y)\,dy &= \int_{\partial U} u(y)\frac{\partial \phi^x}{\partial \nu}(y) - \phi^x(y)\frac{\partial u}{\partial \nu}(y)\,dS(y) \\ &= \int_{\partial U} u(y)\frac{\partial \phi^x}{\partial \nu}(y) - \Phi(y-x)\frac{\partial u}{\partial \nu}(y)\,dS(y). \end{aligned} \tag{27}$$

We introduce next this

DEFINITION. Green's function *for the region* U *is*

$$G(x,y) := \Phi(y-x) - \phi^x(y) \quad (x, y \in U,\ x \neq y).$$

Adopting this terminology and adding (27) to (25), we find

$$(28)\quad u(x) = -\int_{\partial U} u(y)\frac{\partial G}{\partial \nu}(x,y)\,dS(y) - \int_U G(x,y)\Delta u(y)\,dy \quad (x \in U),$$

where

$$\frac{\partial G}{\partial \nu}(x,y) = D_y G(x,y)\cdot \boldsymbol{\nu}(y)$$

is the outer normal derivative of G with respect to the variable y. Observe that the term $\partial u/\partial \nu$ does not appear in equation (28): we introduced the corrector ϕ^x precisely to achieve this.

Suppose now $u \in C^2(\bar{U})$ solves the boundary-value problem

$$(29)\quad \begin{cases} -\Delta u = f & \text{in } U \\ u = g & \text{on } \partial U, \end{cases}$$

for given continuous functions f, g. Plugging into (28), we obtain

THEOREM 12 (Representation formula using Green's function). *If $u \in C^2(\bar{U})$ solves problem* (29), *then*

$$(30)\quad u(x) = -\int_{\partial U} g(y)\frac{\partial G}{\partial \nu}(x,y)\,dS(y) + \int_U f(y)G(x,y)\,dy \quad (x \in U).$$

Here we have a formula for the solution of the boundary-value problem (29), provided we can construct Green's function G for the given domain U. This is in general a difficult matter, and can be done only when U has simple geometry. Subsequent subsections identify some special cases for which an explicit calculation of G is possible.

Remark. Fix $x \in U$. Then regarding G as a function of y, we may symbolically write

$$\begin{cases} -\Delta G = \delta_x & \text{in } U \\ G = 0 & \text{on } \partial U, \end{cases}$$

δ_x denoting the Dirac measure giving unit mass to the point x. □

Before moving on to specific examples, let us record the general assertion that G is symmetric in the variables x and y:

THEOREM 13 (Symmetry of Green's function). *For all $x, y \in U$, $x \neq y$, we have*

$$G(y,x) = G(x,y).$$

Proof. Fix $x, y \in U$, $x \neq y$. Write

$$v(z) := G(x,z),\ w(z) := G(y,z) \quad (z \in U).$$

Then $\Delta v(z) = 0$ $(z \neq x)$, $\Delta w(z) = 0$ $(z \neq y)$ and $w = v = 0$ on ∂U. Thus our applying Green's identity on $V := U - [B(x,\varepsilon) \cup B(y,\varepsilon)]$ for sufficiently small $\varepsilon > 0$ yields

$$\text{(31)} \qquad \int_{\partial B(x,\varepsilon)} \frac{\partial v}{\partial \nu} w - \frac{\partial w}{\partial \nu} v \, dS(z) = \int_{\partial B(y,\varepsilon)} \frac{\partial w}{\partial \nu} v - \frac{\partial v}{\partial \nu} w \, dS(z),$$

$\boldsymbol{\nu}$ denoting the inward pointing unit vector field on $\partial B(x,\varepsilon) \cup \partial B(y,\varepsilon)$. Now w is smooth near x; whence

$$\left| \int_{\partial B(x,\varepsilon)} \frac{\partial w}{\partial \nu} v \, dS \right| \leq C\varepsilon^{n-1} \sup_{\partial B(x,\varepsilon)} |v| = o(1) \quad \text{as } \varepsilon \to 0.$$

On the other hand, $v(z) = \Phi(z - x) - \phi^x(z)$, where ϕ^x is smooth in U. Thus

$$\lim_{\varepsilon \to 0} \int_{\partial B(x,\varepsilon)} \frac{\partial v}{\partial \nu} w \, dS = \lim_{\varepsilon \to 0} \int_{\partial B(x,\varepsilon)} \frac{\partial \Phi}{\partial \nu}(x - z) w(z) \, dS = w(x),$$

by calculations as in the proof of Theorem 1. Thus the left-hand side of (31) converges to $w(x)$ as $\varepsilon \to 0$. Likewise the right hand side converges to $v(y)$. Consequently

$$G(y,x) = w(x) = v(y) = G(x,y).$$

□

b. Green's function for a half-space.

In this and the next subsection we will build Green's functions for two regions with simple geometry, namely the half-space $\mathbb{R}^n_+$ and the unit ball $B(0,1)$. Everything depends upon our explicitly solving the corrector problem (26) in these regions, and this in turn depends upon some clever geometric reflection tricks.

First let us consider the half-space

$$\mathbb{R}^n_+ = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_n > 0\}.$$

Although this region is unbounded, and so the calculations in the previous section do not directly apply, we will attempt nevertheless to build Green's function using the ideas developed before. Later of course we must check directly that the corresponding representation formula is valid.

DEFINITION. *If* $x = (x_1, \ldots, x_{n-1}, x_n) \in \mathbb{R}^n_+$, *its* reflection *in the plane* $\partial\mathbb{R}^n_+$ *is the point*

$$\tilde{x} = (x_1, \ldots, x_{n-1}, -x_n).$$

We will solve problem (26) for the half-space by setting

$$\phi^x(y) := \Phi(y - \tilde{x}) = \Phi(y_1 - x_1, \ldots, y_{n-1} - x_{n-1}, y_n + x_n) \quad (x, y \in \mathbb{R}^n_+).$$

The idea is that the corrector ϕ^x is built from Φ by "reflecting the singularity" from $x \in \mathbb{R}^n_+$ to $\tilde{x} \notin \mathbb{R}^n_+$. We note

$$\phi^x(y) = \Phi(y - x) \quad \text{if } y \in \partial\mathbb{R}^n_+,$$

and thus

$$\begin{cases} \Delta\phi^x = 0 & \text{in } \mathbb{R}^n_+ \\ \phi^x = \Phi(y - x) & \text{on } \partial\mathbb{R}^n_+, \end{cases}$$

as required.

DEFINITION. Green's function for the half-space $\mathbb{R}^n_+$ *is*

$$G(x, y) := \Phi(y - x) - \Phi(y - \tilde{x}) \quad (x, y \in \mathbb{R}^n_+, \; x \neq y).$$

Then

$$\begin{aligned} \frac{\partial G}{\partial y_n}(x, y) &= \frac{\partial \Phi}{\partial y_n}(y - x) - \frac{\partial \Phi}{\partial y_n}(y - \tilde{x}) \\ &= \frac{-1}{n\alpha(n)} \left[\frac{y_n - x_n}{|y - x|^n} - \frac{y_n + x_n}{|y - \tilde{x}|^n} \right]. \end{aligned}$$

Consequently if $y \in \partial\mathbb{R}^n_+$,

$$\frac{\partial G}{\partial \nu}(x, y) = -\frac{\partial G}{\partial y_n}(x, y) = -\frac{-2x_n}{n\alpha(n)} \frac{1}{|x - y|^n}.$$

Suppose now u solves the boundary-value problem

$$(32) \qquad \begin{cases} \Delta u = 0 & \text{in } \mathbb{R}^n_+ \\ u = g & \text{on } \partial\mathbb{R}^n_+. \end{cases}$$

Then from (30) we expect

$$(33) \qquad u(x) = \frac{2x_n}{n\alpha(n)} \int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|x - y|^n} dy \quad (x \in \mathbb{R}^n_+)$$

to be a representation formula for our solution. The function

$$K(x, y) := \frac{2x_n}{n\alpha(n)} \frac{1}{|x - y|^n} \quad (x \in \mathbb{R}^n_+, y \in \partial\mathbb{R}^n_+)$$

is *Poisson's kernel* for $\mathbb{R}^n_+$, and (33) is *Poisson's formula.*

We must now check directly that formula (33) does indeed provide us with a solution of the boundary-value problem (32).

THEOREM 14 (Poisson's formula for half-space). *Assume* $g \in C(\mathbb{R}^{n-1}) \cap L^\infty(\mathbb{R}^{n-1})$, *and define* u *by* (33). *Then*

(i) $u \in C^\infty(\mathbb{R}^n_+) \cap L^\infty(\mathbb{R}^n_+)$,

(ii) $\Delta u = 0 \quad in\ \mathbb{R}^n_+$,

and

(iii) $\lim_{\substack{x \to x^0 \\ x \in \mathbb{R}^n_+}} u(x) = g(x^0) \qquad$ *for each point* $x^0 \in \partial\mathbb{R}^n_+$.

Proof. 1. For each fixed x, the mapping $y \mapsto G(x, y)$ is harmonic, except for $y = x$. As $G(x, y) = G(y, x)$ according to Theorem 13, $x \mapsto G(x, y)$ is harmonic, except for $x = y$. Thus $x \mapsto -\frac{\partial G}{\partial y_n}(x, y) = K(x, y)$ is harmonic for $x \in \mathbb{R}^n_+$, $y \in \partial\mathbb{R}^n_+$.

2. A direct calculation, the details of which we omit, verifies

$$1 = \int_{\partial\mathbb{R}^n_+} K(x, y)\, dy \tag{34}$$

for each $x \in \mathbb{R}^n_+$. As g is bounded, u defined by (33) is likewise bounded. Since $x \mapsto K(x, y)$ is smooth for $x \neq y$, we easily verify as well $u \in C^\infty(\mathbb{R}^n_+)$, with

$$\Delta u(x) = \int_{\partial\mathbb{R}^n_+} \Delta_x K(x, y) g(y)\, dy = 0 \quad (x \in \mathbb{R}^n_+).$$

3. Now fix $x^0 \in \partial\mathbb{R}^n_+$, $\varepsilon > 0$. Choose $\delta > 0$ so small that

$$|g(y) - g(x^0)| < \varepsilon \quad \text{if } |y - x^0| < \delta,\ y \in \partial\mathbb{R}^n_+. \tag{35}$$

Then if $|x - x^0| < \frac{\delta}{2}$, $x \in \mathbb{R}^n_+$,

$$\begin{aligned} |u(x) - g(x^0)| &= \left| \int_{\partial\mathbb{R}^n_+} K(x, y)[g(y) - g(x^0)]\, dy \right| \\ &\leq \int_{\partial\mathbb{R}^n_+ \cap B(x^0, \delta)} K(x, y) |g(y) - g(x^0)|\, dy \\ &\quad + \int_{\partial\mathbb{R}^n_+ - B(x^0, \delta)} K(x, y) |g(y) - g(x^0)|\, dy \\ &=: I + J. \end{aligned} \tag{36}$$

Now (34), (35) imply

$$I \leq \varepsilon \int_{\partial\mathbb{R}^n_+} K(x, y)\, dy = \varepsilon.$$

Furthermore if $|x - x^0| \leq \frac{\delta}{2}$ and $|y - x^0| \geq \delta$, we have

$$|y - x^0| \leq |y - x| + \frac{\delta}{2} \leq |y - x| + \frac{1}{2}|y - x^0|;$$

and so $|y - x| \geq \frac{1}{2}|y - x^0|$. Thus

$$\begin{aligned} J &\leq 2\|g\|_{L^\infty} \int_{\partial\mathbb{R}^n_+ - B(x^0,\delta)} K(x,y)\,dy \\ &\leq \frac{2^{n+2}\|g\|_{L^\infty} x_n}{n\alpha(n)} \int_{\partial\mathbb{R}^n_+ - B(x^0,\delta)} |y - x^0|^{-n}\,dy \\ &\to 0 \quad \text{as } x_n \to 0^+. \end{aligned}$$

Combining this calculation with estimate (36), we deduce $|u(x) - g(x^0)| \leq 2\varepsilon$, provided $|x - x^0|$ is sufficiently small. □

c. Green's function for a ball.

To construct Green's function for the unit ball $B(0,1)$ we will again employ a kind of reflection, this time through the sphere $\partial B(0,1)$.

DEFINITION. *If $x \in \mathbb{R}^n - \{0\}$, the point*

$$\tilde{x} = \frac{x}{|x|^2}$$

is called the point dual *to x with respect to $\partial B(0,1)$. The mapping $x \mapsto \tilde{x}$ is* inversion *through the unit sphere $\partial B(0,1)$.*

We now employ inversion through the sphere to compute Green's function for the unit ball $U = B^0(0,1)$. Fix $x \in B^0(0,1)$. Remember that we must find a corrector function $\phi^x = \phi^x(y)$ solving

$$(37) \qquad \begin{cases} \Delta\phi^x = 0 & \text{in } B^0(0,1) \\ \phi^x = \Phi(y - x) & \text{on } \partial B(0,1); \end{cases}$$

then Green's function will be

$$(38) \qquad G(x,y) = \Phi(y - x) - \phi^x(y).$$

The idea now is to "invert the singularity" from $x \in B^0(0,1)$ to $\tilde{x} \notin B(0,1)$. Assume for the moment $n \geq 3$. Now the mapping $y \mapsto \Phi(y - \tilde{x})$ is harmonic for $y \neq \tilde{x}$. Thus $y \mapsto |x|^{2-n}\Phi(y - \tilde{x})$ is harmonic for $y \neq \tilde{x}$, and so

$$(39) \qquad \phi^x(y) := \Phi(|x|(y - \tilde{x}))$$

is harmonic in U. Furthermore, if $y \in \partial B(0,1)$ and $x \neq 0$,

$$|x|^2|y-\tilde{x}|^2 = |x|^2\left(|y|^2 - \frac{2y\cdot x}{|x|^2} + \frac{1}{|x|^2}\right)$$
$$= |x|^2 - 2y\cdot x + 1 = |x-y|^2.$$

Thus $(|x||y-\tilde{x}|)^{-(n-2)} = |x-y|^{-(n-2)}$. Consequently

(40) $$\phi^x(y) = \Phi(y-x) \quad (y \in \partial B(0,1)),$$

as required.

DEFINITION. Green's function for the unit ball *is*

(41) $$G(x,y) := \Phi(y-x) - \Phi(|x|(y-\tilde{x})) \quad (x,y \in B(0,1),\ x \neq y).$$

The same formula is valid for $n=2$ as well.

Assume now u solves the boundary-value problem

(42) $$\begin{cases} \Delta u = 0 & \text{in } B^0(0,1) \\ u = g & \text{in } \partial B(0,1). \end{cases}$$

Then using (30), we see

(43) $$u(x) = -\int_{\partial B(0,1)} g(y)\frac{\partial G}{\partial \nu}(x,y)\,dS(y).$$

According to formula (41),

$$\frac{\partial G}{\partial y_i}(x,y) = \frac{\partial \Phi}{\partial y_i}(y-x) - \frac{\partial}{\partial y_i}\Phi(|x|(y-\tilde{x})).$$

But

$$\frac{\partial \Phi}{\partial y_i}(x-y) = \frac{1}{n\alpha(n)}\frac{x_i - y_i}{|x-y|^n},$$

and furthermore

$$\frac{\partial \Phi}{\partial y_i}(|x|(y-\tilde{x})) = \frac{-1}{n\alpha(n)}\frac{y_i|x|^2 - x_i}{(|x||y-\tilde{x}|)^n} = -\frac{1}{n\alpha(n)}\frac{y_i|x|^2 - x_i}{|x-y|^n}$$

if $y \in \partial B(0,1)$. Accordingly

$$\frac{\partial G}{\partial \nu}(x,y) = \sum_{i=1}^n y_i \frac{\partial G}{\partial y_i}(x,y)$$
$$= \frac{-1}{n\alpha(n)}\frac{1}{|x-y|^n}\sum_{i=1}^n y_i((y_i - x_i) - y_i|x|^2 + x_i)$$
$$= \frac{-1}{n\alpha(n)}\frac{1-|x|^2}{|x-y|^n}.$$

Hence formula (43) yields the representation formula

$$u(x) = \frac{1-|x|^2}{n\alpha(n)} \int_{\partial B(0,1)} \frac{g(y)}{|x-y|^n}\, dS(y).$$

Suppose now instead of (42) u solves the boundary-value problem

$$\begin{cases} \Delta u = 0 & \text{in } B^0(0,r) \\ u = g & \text{on } \partial B(0,r) \end{cases} \tag{44}$$

for $r > 0$. Then $\tilde{u}(x) = u(rx)$ solves (42), with $\tilde{g}(x) = g(rx)$ replacing g. We change variables to obtain *Poisson's formula*

$$u(x) = \frac{r^2-|x|^2}{n\alpha(n)r} \int_{\partial B(0,r)} \frac{g(y)}{|x-y|^n}\, dS(y) \quad (x \in B^0(0,r)). \tag{45}$$

The function

$$K(x,y) := \frac{r^2-|x|^2}{n\alpha(n)r} \frac{1}{|x-y|^n} \quad (x \in B^0(0,r),\ y \in \partial B(0,r))$$

is *Poisson's kernel* for the ball $B(0,r)$.

We have established (45) under the assumption that a smooth solution of (44) exists. We next assert that this formula in fact gives a solution:

THEOREM 15 (Poisson's formula for ball). *Assume $g \in C(\partial B(0,r))$ and define u by* (45). *Then*

(i) $u \in C^\infty(B^0(0,r))$,

(ii) $\Delta u = 0$ *in* $B^0(0,r)$,

and

(iii) $\displaystyle\lim_{\substack{x\to x^0 \\ x\in B^0(0,r)}} u(x) = g(x^0)$ *for each point* $x^0 \in \partial B(0,r)$.

The proof is similar to that for Theorem 14, and is left as an exercise.

2.2.5. Energy methods.

Most of our analysis of harmonic functions thus far has depended upon fairly explicit representation formulas entailing the fundamental solution, Green's functions, etc. In this concluding section we illustrate some "energy" methods, which is to say techniques involving the L^2-norms of various expressions. These ideas foreshadow latter theoretical developments in Parts II and III.

a. Uniqueness.

Consider first the boundary-value problem

$$\begin{cases} -\Delta u = f & \text{in } U \\ u = g & \text{on } \partial U. \end{cases} \tag{46}$$

We have already employed the maximum principle in §2.2.3 to show uniqueness, but now set forth a simple alternative proof. Assume U is open, bounded, and ∂U is C^1.

THEOREM 16 (Uniqueness). *There exists at most one solution* $u \in C^2(\bar{U})$ *of* (46).

Proof. Assume $\tilde{u}$ is another solution and set $w := u - \tilde{u}$. Then $\Delta w = 0$ in U, and so an integration by parts shows

$$0 = -\int_U w\Delta w \, dx = \int_U |Dw|^2 dx.$$

Thus $Dw \equiv 0$ in U, and, since $w = 0$ on ∂U, we deduce $w = u - \tilde{u} \equiv 0$ in U. □

b. Dirichlet's principle.

Next let us demonstrate that a solution of the boundary-value problem (46) for Poisson's equation can be characterized as the minimizer of an appropriate functional. For this, we define the *energy* functional

$$I[w] = \int_U \frac{1}{2}|Dw|^2 - wf \, dx,$$

w belonging to the *admissible set*

$$\mathcal{A} = \{w \in C^2(\bar{U}) \mid w = g \text{ on } \partial U\}.$$

THEOREM 17 (Dirichlet's principle). *Assume* $u \in C^2(\bar{U})$ *solves* (46). *Then*

$$I[u] = \min_{w \in \mathcal{A}} I[w]. \tag{47}$$

Conversely, if $u \in \mathcal{A}$ *satisfies* (47), *then* u *solves the boundary-value problem* (46).

In other words if $u \in \mathcal{A}$, the PDE $-\Delta u = f$ is equivalent to the statement that u minimizes the energy $I[\cdot]$.

Proof. 1. Choose $w \in \mathcal{A}$. Then (46) implies

$$0 = \int_U (-\Delta u - f)(u - w)\, dx.$$

An integration by parts yields

$$0 = \int_U Du \cdot D(u - w) - f(u - w)\, dx,$$

and there is no boundary term since $u - w = g - g = 0$ on ∂U. Hence

$$\begin{aligned}\int_U |Du|^2 - uf\, dx &= \int_U Du \cdot Dw - wf\, dx \\ &\leq \int_U \frac{1}{2}|Du|^2\, dx + \int_U \frac{1}{2}|Dw|^2 - wf\, dx,\end{aligned}$$

where we employed the estimates

$$|Du \cdot Dw| \leq |Du||Dw| \leq \frac{1}{2}|Du|^2 + \frac{1}{2}|Dw|^2,$$

following from the Cauchy–Schwarz and Cauchy inequalities (§B.2). Rearranging, we conclude

$$I[u] \leq I[w] \quad (w \in \mathcal{A}). \tag{48}$$

Since $u \in \mathcal{A}$, (47) follows from (48).

2. Now, conversely, suppose (47) holds. Fix any $v \in C_c^\infty(U)$ and write

$$i(\tau) := I[u + \tau v] \quad (\tau \in \mathbb{R}).$$

Since $u + \tau v \in \mathcal{A}$ for each τ, the scalar function $i(\cdot)$ has a minimum at zero, and thus

$$i'(0) = 0 \quad \left(' = \frac{d}{d\tau}\right),$$

provided this derivative exists. But

$$\begin{aligned} i(\tau) &= \int_U \frac{1}{2}|Du + \tau Dv|^2 - (u + \tau v)f\, dx \\ &= \int_U \frac{1}{2}|Du|^2 + \tau Du \cdot Dv + \frac{\tau^2}{2}|Dv|^2 - (u + \tau v)f\, dx.\end{aligned}$$

Consequently

$$0 = i'(0) = \int_U Du \cdot Dv - vf\, dx = \int_U (-\Delta u - f)v\, dx.$$

This identity is valid for each function $v \in C_c^\infty(U)$ and so $-\Delta u = f$ in U. □

Dirichlet's principle is an instance of the *calculus of variations* applied to Laplace's equation. See Chapter 8 for more.

2.3. HEAT EQUATION

Next we study the *heat equation*

$$(1) \qquad u_t - \Delta u = 0$$

and the *nonhomogeneous heat equation*

$$(2) \qquad u_t - \Delta u = f,$$

subject to appropriate initial and boundary conditions. Here $t > 0$ and $x \in U$, where $U \subset \mathbb{R}^n$ is open. The unknown is $u : \bar{U} \times [0, \infty) \to \mathbb{R}$, $u = u(x,t)$, and the Laplacian Δ is taken with respect to the spatial variables $x = (x_1, \ldots, x_n)$: $\Delta u = \Delta_x u = \sum_{i=1}^n u_{x_i x_i}$. In (2) the function $f : U \times [0, \infty) \to \mathbb{R}$ is given.

A guiding principle is that any assertion about harmonic functions yields an analogous (but more complicated) statement about solutions of the heat equation. Accordingly our development will largely parallel the corresponding theory for Laplace's equation.

Physical interpretation. The heat equation, also known as the *diffusion equation*, describes in typical applications the evolution in time of the density u of some quantity such as heat, chemical concentration, etc. If $V \subset U$ is any smooth subregion, the rate of change of the total quantity within V equals the negative of the net flux through ∂V:

$$\frac{d}{dt} \int_V u \, dx = - \int_{\partial V} \mathbf{F} \cdot \nu \, dS,$$

$\mathbf{F}$ being the flux density. Thus

$$(3) \qquad u_t = -\operatorname{div} \mathbf{F},$$

as V was arbitrary. In many situations $\mathbf{F}$ is proportional to the gradient of u, but points in the opposite direction (since the flow is from regions of higher to lower concentration):

$$\mathbf{F} = -aDu \quad (a > 0).$$

Substituting into (3), we obtain the PDE

$$u_t = a \operatorname{div}(Du) = a\Delta u,$$

which for $a = 1$ is the heat equation.

The heat equation appears as well in the study of Brownian motion. □

2.3.1. Fundamental solution.

a. Derivation of the fundamental solution.

As noted in §2.2.1 an important first step in studying any PDE is often to come up with some specific solutions.

We observe that the heat equation involves one derivative with respect to the time variable t, but two derivatives with respect to the space variables x_i $(i = 1, \dots, n)$. Consequently we see that if u solves (1), then so does $u(\lambda x, \lambda^2 t)$ for $\lambda \in \mathbb{R}$. This scaling indicates the ratio $\frac{r^2}{t}$ $(r = |x|)$ is important for the heat equation and suggests that we search for a solution of (1) having the form $u(x,t) = v(\frac{r^2}{t}) = v(\frac{|x|^2}{t})$ $(t > 0,\ x \in \mathbb{R}^n)$, for some function v as yet undetermined.

Although this approach eventually leads to what we want (see Problem 11), it is quicker to seek a solution u having the special structure

$$u(x,t) = \frac{1}{t^\alpha} v\Big(\frac{x}{t^\beta}\Big) \quad (x \in \mathbb{R}^n,\ t > 0), \tag{4}$$

where the constants α, β and the function $v : \mathbb{R}^n \to \mathbb{R}$ must be found. We come to (4) if we look for a solution u of the heat equation invariant under the *dilation scaling*

$$u(x,t) \mapsto \lambda^\alpha u(\lambda^\beta x, \lambda t).$$

That is, we ask

$$u(x,t) = \lambda^\alpha u(\lambda^\beta x, \lambda t)$$

for all $\lambda > 0$, $x \in \mathbb{R}^n$, $t > 0$. Setting $\lambda = t^{-1}$, we derive (4) for $v(y) := u(y,1)$.

Let us insert (4) into (1), and thereafter compute

$$\alpha t^{-(\alpha+1)} v(y) + \beta t^{-(\alpha+1)} y \cdot Dv(y) + t^{-(\alpha+2\beta)} \Delta v(y) = 0 \tag{5}$$

for $y := t^{-\beta} x$. In order to transform (5) into an expression involving the variable y alone, we take $\beta = \frac{1}{2}$. Then the terms with t are identical, and so (5) reduces to

$$\alpha v + \frac{1}{2} y \cdot Dv + \Delta v = 0. \tag{6}$$

We simplify further by guessing v to be radial; that is, $v(y) = w(|y|)$ for some $w : \mathbb{R} \to \mathbb{R}$. Thereupon (6) becomes

$$\alpha w + \frac{1}{2} r w' + w'' + \frac{n-1}{r} w' = 0,$$

for $r = |y|$, $' = \frac{d}{dr}$. Now if we set $\alpha = \frac{n}{2}$, this simplifies to read

$$(r^{n-1}w')' + \frac{1}{2}(r^n w)' = 0.$$

Thus

$$r^{n-1}w' + \frac{1}{2}r^n w = a$$

for some constant a. Assuming $\lim_{r\to\infty} w,\ w' = 0$, we conclude $a = 0$; whence

$$w' = -\frac{1}{2}rw.$$

But then for some constant b

$$w = be^{-\frac{r^2}{4}}. \tag{7}$$

Combining (4), (7) and our choices for α, β, we conclude that $\frac{b}{t^{n/2}}e^{-\frac{|x|^2}{4t}}$ solves the heat equation (1).

This computation motivates the following

DEFINITION. *The function*

$$\Phi(x,t) := \begin{cases} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} & (x \in \mathbb{R}^n,\ t > 0) \\ 0 & (x \in \mathbb{R}^n,\ t < 0) \end{cases}$$

is called the fundamental solution of the heat equation.

Notice that Φ is singular at the point $(0,0)$. We will sometimes write $\Phi(x,t) = \Phi(|x|,t)$ to emphasize that the fundamental solution is radial in the variable x. The choice of the normalizing constant $(4\pi)^{-n/2}$ is dictated by the following

LEMMA (Integral of fundamental solution). *For each time $t > 0$,*

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

Proof. We calculate

$$\begin{aligned} \int_{\mathbb{R}^n} \Phi(x,t)\,dx &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x|^2}{4t}}\,dx \\ &= \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} e^{-|z|^2} dz \\ &= \frac{1}{\pi^{n/2}} \prod_{i=1}^{n} \int_{-\infty}^{\infty} e^{-z_i^2}\,dz_i = 1. \end{aligned}$$

□

A different derivation of the fundamental solution of the heat equation appears in §4.3.2.

b. Initial-value problem.

We now employ Φ to fashion a solution to the *initial-value* (or *Cauchy*) *problem*

$$\begin{cases} u_t - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{8}$$

Let us note the function $(x,t) \mapsto \Phi(x,t)$ solves the heat equation away from the singularity at $(0,0)$, and thus so does $(x,t) \mapsto \Phi(x-y,t)$ for each fixed $y \in \mathbb{R}^n$. Consequently the convolution

$$\begin{aligned} u(x,t) &= \int_{\mathbb{R}^n} \Phi(x-y,t)g(y)\,dy \\ &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\,dy \quad (x \in \mathbb{R}^n,\ t>0) \end{aligned} \tag{9}$$

should also be a solution.

THEOREM 1 (Solution of initial-value problem). *Assume* $g \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, *and define* u *by* (9). *Then*

(i) $u \in C^\infty(\mathbb{R}^n \times (0,\infty))$,

(ii) $u_t(x,t) - \Delta u(x,t) = 0 \quad (x \in \mathbb{R}^n,\ t>0)$,

and

(iii) $\displaystyle\lim_{\substack{(x,t)\to(x^0,0) \\ x\in\mathbb{R}^n,\ t>0}} u(x,t) = g(x^0)$ *for each point* $x^0 \in \mathbb{R}^n$.

Proof. 1. Since the function $\frac{1}{t^{n/2}}e^{-\frac{|x|^2}{4t}}$ is infinitely differentiable, with uniformly bounded derivatives of all orders, on $\mathbb{R}^n \times [\delta,\infty)$ for each $\delta > 0$, we see that $u \in C^\infty(\mathbb{R}^n \times (0,\infty))$. Furthermore

$$\begin{aligned} u_t(x,t) - \Delta u(x,t) &= \int_{\mathbb{R}^n} [(\Phi_t - \Delta_x\Phi)(x-y,t)]g(y)\,dy \\ &= 0 \quad (x \in \mathbb{R}^n,\ t>0), \end{aligned} \tag{10}$$

since Φ itself solves the heat equation.

2. Fix $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$. Choose $\delta > 0$ such that

$$|g(y) - g(x^0)| < \varepsilon \quad \text{if } |y - x^0| < \delta,\ y \in \mathbb{R}^n. \tag{11}$$

Then if $|x - x^0| < \frac{\delta}{2}$, we have, according to the lemma,

$$\begin{aligned}|u(x,t) - g(x^0)| &= \left|\int_{\mathbb{R}^n} \Phi(x-y,t)[g(y) - g(x^0)]\,dy\right| \\ &\le \int_{B(x^0,\delta)} \Phi(x-y,t)|g(y) - g(x^0)|\,dy \\ &\quad + \int_{\mathbb{R}^n - B(x^0,\delta)} \Phi(x-y,t)|g(y) - g(x^0)|\,dy \\ &=: I + J.\end{aligned}$$

Now

$$I \le \varepsilon \int_{\mathbb{R}^n} \Phi(x-y,t)\,dy = \varepsilon,$$

owing to (11) and the lemma. Furthermore, if $|x - x^0| \le \frac{\delta}{2}$ and $|y - x^0| \ge \delta$, then

$$|y - x^0| \le |y - x| + \frac{\delta}{2} \le |y - x| + \frac{1}{2}|y - x^0|.$$

Thus $|y - x| \ge \frac{1}{2}|y - x^0|$. Consequently

$$\begin{aligned}J &\le 2\|g\|_{L^\infty} \int_{\mathbb{R}^n - B(x^0,\delta)} \Phi(x-y,t)\,dy \\ &\le \frac{C}{t^{n/2}} \int_{\mathbb{R}^n - B(x^0,\delta)} e^{-\frac{|x-y|^2}{4t}}\,dy \\ &\le \frac{C}{t^{n/2}} \int_{\mathbb{R}^n - B(x^0,\delta)} e^{-\frac{|y-x^0|^2}{16t}}\,dy \\ &= \frac{C}{t^{n/2}} \int_\delta^\infty e^{-\frac{r^2}{16t}} r^{n-1}\,dr \to 0 \quad \text{as } t \to 0^+.\end{aligned}$$

Hence if $|x - x^0| < \frac{\delta}{2}$ and $t > 0$ is small enough, $|u(x,t) - g(x^0)| < 2\varepsilon$. □

Remarks. (i) In view of Theorem 1 we sometimes write

$$\begin{cases} \Phi_t - \Delta\Phi = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ \Phi = \delta_0 & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases}$$

δ_0 denoting the Dirac measure on $\mathbb{R}^n$ giving unit mass to the point 0.

(ii) Notice that if g is bounded, continuous, $g \ge 0$, $g \not\equiv 0$, then

$$u(x,t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\,dy$$

is in fact positive for *all* points $x \in \mathbb{R}^n$ and times $t > 0$. We interpret this observation by saying the heat equation forces *infinite propagation speed* for disturbances. If the initial temperature is nonnegative and is positive somewhere, the temperature at any later time (no matter how small) is everywhere positive. (We will learn in §2.4.3 that the wave equation in contrast supports finite propagation speed for disturbances.) □

c. Nonhomogeneous problem.

Now let us turn our attention to the *nonhomogeneous* initial-value problem

$$\begin{cases} u_t - \Delta u = f & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = 0 & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{12}$$

How can we produce a formula for the solution? If we recall the motivation leading up to (9), we should note further that the mapping $(x,t) \mapsto \Phi(x-y,t-s)$ is a solution of the heat equation (for given $y \in \mathbb{R}^n$, $0 < s < t$). Now for fixed s, the function

$$u = u(x,t;s) = \int_{\mathbb{R}^n} \Phi(x-y,t-s) f(y,s)\, dy$$

solves

$$\begin{cases} u_t(\cdot;s) - \Delta u(\cdot;s) = 0 & \text{in } \mathbb{R}^n \times (s,\infty) \\ u(\cdot;s) = f(\cdot,s) & \text{on } \mathbb{R}^n \times \{t=s\}, \end{cases} \tag{12_s}$$

which is just an initial-value problem of the form (8), with the starting time $t = 0$ replaced by $t = s$, and g replaced by $f(\cdot,s)$. Thus $u(\cdot;s)$ is certainly not a solution of (12).

However *Duhamel's principle** asserts that we can build a solution of (12) out of the solutions of (12_s), by integrating with respect to s. The idea is to consider

$$u(x,t) = \int_0^t u(x,t;s)\, ds \quad (x \in \mathbb{R}^n,\ t \geq 0).$$

Rewriting, we have

$$\begin{aligned} u(x,t) &= \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s) f(y,s)\, dyds \\ &= \int_0^t \frac{1}{(4\pi(t-s))^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} f(y,s)\, dyds, \end{aligned} \tag{13}$$

for $x \in \mathbb{R}^n$, $t > 0$.

To confirm that formula (13) works, let us for simplicity assume $f \in C_1^2(\mathbb{R}^n \times [0,\infty))$ and f has compact support.

*Duhamel's principle has wide applicability to linear ODE and PDE, and does not depend on the specific structure of the heat equation. It yields, for example, the solution of the nonhomogeneous transport equation, obtained by different means in §2.1.2. We will invoke Duhamel's principle for the wave equation in §2.4.2.

THEOREM 2 (Solution of nonhomogeneous problem). *Define* u *by* (13). *Then*

(i) $u \in C_1^2(\mathbb{R}^n \times (0,\infty))$,

(ii) $u_t(x,t) - \Delta u(x,t) = f(x,t) \quad (x \in \mathbb{R}^n,\ t > 0)$,

and

(iii) $\displaystyle\lim_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n,\ t>0}} u(x,t) = 0$ *for each point* $x^0 \in \mathbb{R}^n$.

Proof. 1. Since Φ has a singularity at $(0,0)$, we cannot directly justify differentiating under the integral sign. We instead proceed somewhat as in the proof of Theorem 1 in §2.2.1.

First we change variables, to write

$$u(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s) f(x-y,t-s)\, dyds.$$

As $f \in C_1^2(\mathbb{R}^n \times [0,\infty))$ has compact support and $\Phi = \Phi(y,s)$ is smooth near $s = t > 0$, we compute

$$\begin{aligned} u_t(x,t) = &\int_0^t \int_{\mathbb{R}^n} \Phi(y,s) f_t(x-y,t-s)\, dyds \\ &+ \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy \end{aligned}$$

and

$$\frac{\partial^2 u}{\partial x_i \partial x_j}(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s) \frac{\partial^2}{\partial x_i \partial x_j} f(x-y,t-s)\, dyds \quad (i,j = 1,\dots,n).$$

Thus u_t, $D_x^2 u$, and likewise $u, D_x u$, belong to $C(\mathbb{R}^n \times (0,\infty))$.

2. We then calculate

(14)

$$\begin{aligned} u_t(x,t) - \Delta u(x,t) &= \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)[(\frac{\partial}{\partial t} - \Delta_x) f(x-y,t-s)]\, dyds \\ &\quad + \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy \\ &= \int_\varepsilon^t \int_{\mathbb{R}^n} \Phi(y,s)[(-\frac{\partial}{\partial s} - \Delta_y) f(x-y,t-s)]\, dyds \\ &\quad + \int_0^\varepsilon \int_{\mathbb{R}^n} \Phi(y,s)[(-\frac{\partial}{\partial s} - \Delta_y) f(x-y,t-s)]\, dyds \\ &\quad + \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy. \\ &=: I_\varepsilon + J_\varepsilon + K. \end{aligned}$$

Now

$$|J_\varepsilon| \leq (\|f_t\|_{L^\infty} + \|D^2 f\|_{L^\infty}) \int_0^\varepsilon \int_{\mathbb{R}^n} \Phi(y,s)\, dyds \leq \varepsilon C, \tag{15}$$

by the lemma. Integrating by parts, we also find

$$\begin{aligned} I_\varepsilon = & \int_\varepsilon^t \int_{\mathbb{R}^n} [(\frac{\partial}{\partial s} - \Delta_y)\Phi(y,s)] f(x-y, t-s)\, dyds \\ & + \int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y, t-\varepsilon)\, dy \\ & - \int_{\mathbb{R}^n} \Phi(y,t) f(x-y, 0)\, dy \\ = & \int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y,t-\varepsilon)\, dy - K, \end{aligned} \tag{16}$$

since Φ solves the heat equation. Combining (14)–(16), we ascertain

$$\begin{aligned} u_t(x,t) - \Delta u(x,t) &= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y, t-\varepsilon)\, dy \\ &= f(x,t) \quad (x \in \mathbb{R}^n,\, t > 0), \end{aligned}$$

the limit as $\varepsilon \to 0$ being computed as in the proof of Theorem 1. Finally note $\|u(\cdot,t)\|_{L^\infty} \leq t\|f\|_{L^\infty} \to 0$. □

Remark. We can of course combine Theorems 1 and 2 to discover that

$$u(x,t) = \int_{\mathbb{R}^n} \Phi(x-y,t) g(y)\, dy + \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s) f(y,s)\, dyds \tag{17}$$

is, under the hypotheses on g and f as above, a solution of

$$\begin{cases} u_t - \Delta u = f & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{18}$$

□

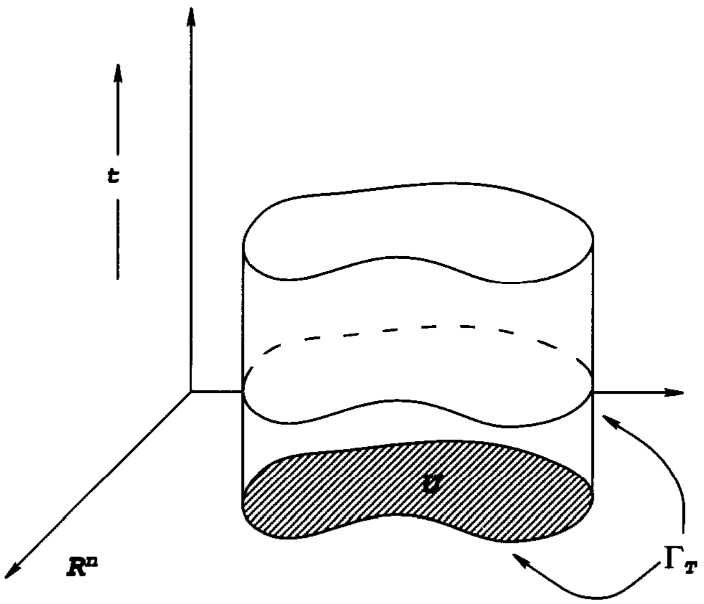
2.3.2. Mean-value formula.

First we recall some useful notation from §A.2. Assume $U \subset \mathbb{R}^n$ is open and bounded, and fix a time $T > 0$.

DEFINITIONS.

(i) *We define the* parabolic cylinder

$$U_T := U \times (0,T].$$



The region U_T

(ii) *The* parabolic boundary *of* U_T *is*

$$\Gamma_T := \bar{U}_T - U_T.$$

We interpret U_T as being the *parabolic interior* of $\bar{U} \times [0,T]$: note carefully that U_T includes the top $U \times \{t = T\}$. The parabolic boundary Γ_T comprises the bottom and vertical sides of $U \times [0,T]$, but not the top.

We want next to derive a kind of analogue to the mean-value property for harmonic functions, as discussed in §2.2.2. There is no such simple formula. However let us observe that for fixed x the spheres $\partial B(x,r)$ are level sets of the fundamental solution $\Phi(x-y)$ for Laplace's equation. This suggests that perhaps for fixed (x,t) the level sets of fundamental solution $\Phi(x-y,t-s)$ for the heat equation may be relevant.

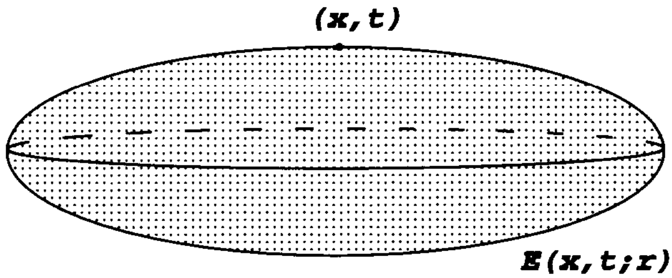
DEFINITION. *For fixed $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, $r > 0$, we define*

$$E(x,t;r) := \left\{ (y,s) \in \mathbb{R}^{n+1} \mid s \leq t,\ \Phi(x-y,t-s) \geq \frac{1}{r^n} \right\}.$$

This is a region in space-time, the boundary of which is a level set of $\Phi(x-y,t-s)$. Note that the point (x,t) is at the center of the top. $E(x,t;r)$ is sometimes called a "heat ball".

THEOREM 3 (A mean-value property for the heat equation). *Let $u \in C_1^2(U_T)$ solve the heat equation. Then*

$$u(x,t) = \frac{1}{4r^n} \iint_{E(x,t;r)} u(y,s) \frac{|x-y|^2}{(t-s)^2}\, dyds \tag{19}$$



A "heat ball"

for each $E(x,t;r) \subset U_T$.

Formula (19) is a sort of analogue for the heat equation of the mean-value formulas for Laplace's equation. Observe that the right hand side involves only $u(y,s)$ for times $s \leq t$. This is reasonable, as the value $u(x,t)$ should not depend upon future times.

Proof. We may as well assume upon translating the space and time coordinates that $x = 0$ and $t = 0$. Write $E(r) = E(0,0;r)$ and set

$$\text{(20)} \qquad \begin{aligned} \phi(r) &:= \frac{1}{r^n} \iint_{E(r)} u(y,s) \frac{|y|^2}{s^2}\, dyds \\ &= \iint_{E(1)} u(ry, r^2 s) \frac{|y|^2}{s^2}\, dyds. \end{aligned}$$

We compute

$$\begin{aligned} \phi'(r) &= \iint_{E(1)} \sum_{i=1}^n u_{y_i} y_i \frac{|y|^2}{s^2} + 2r u_s \frac{|y|^2}{s}\, dyds \\ &= \frac{1}{r^{n+1}} \iint_{E(r)} \sum_{i=1}^n u_{y_i} y_i \frac{|y|^2}{s^2} + 2 u_s \frac{|y|^2}{s}\, dyds \\ &=: A + B. \end{aligned}$$

Also, let us introduce the useful function

$$\text{(21)} \qquad \psi := -\frac{n}{2}\log(-4\pi s) + \frac{|y|^2}{4s} + n \log r,$$

and observe $\psi = 0$ on $\partial E(r)$, since $\Phi(y,-s) = r^{-n}$ on $\partial E(r)$. We utilize (21) to write

$$\begin{aligned} B &= \frac{1}{r^{n+1}} \iint_{E(r)} 4u_s \sum_{i=1}^n y_i \psi_{y_i}\, dyds \\ &= -\frac{1}{r^{n+1}} \iint_{E(r)} 4n u_s \psi + 4 \sum_{i=1}^n u_{s y_i} y_i \psi\, dyds; \end{aligned}$$

there is no boundary term since $\psi = 0$ on $\partial E(r)$. Integrating by parts with respect to s, we discover

$$\begin{aligned} B &= \frac{1}{r^{n+1}} \iint_{E(r)} -4nu_s\psi + 4\sum_{i=1}^n u_{y_i} y_i \psi_s \, dyds \\ &= \frac{1}{r^{n+1}} \iint_{E(r)} -4nu_s\psi + 4\sum_{i=1}^n u_{y_i} y_i \left(-\frac{n}{2s} - \frac{|y|^2}{4s^2}\right) dyds \\ &= \frac{1}{r^{n+1}} \iint_{E(r)} -4nu_s\psi - \frac{2n}{s}\sum_{i=1}^n u_{y_i} y_i \, dyds - A. \end{aligned}$$

Consequently, since u solves the heat equation,

$$\begin{aligned} \phi'(r) &= A + B \\ &= \frac{1}{r^{n+1}} \iint_{E(r)} -4n\Delta u\psi - \frac{2n}{s}\sum_{i=1}^n u_{y_i} y_i \, dyds \\ &= \sum_{i=1}^n \frac{1}{r^{n+1}} \iint_{E(r)} 4nu_{y_i}\psi_{y_i} - \frac{2n}{s} u_{y_i} y_i \, dyds \\ &= 0, \quad \text{according to (21).} \end{aligned}$$

Thus ϕ is constant, and therefore

$$\phi(r) = \lim_{t\to 0} \phi(t) = u(0,0)\left(\lim_{t\to 0} \frac{1}{t^n} \iint_{E(t)} \frac{|y|^2}{s^2} dyds\right) = 4u(0,0),$$

as

$$\frac{1}{t^n} \iint_{E(t)} \frac{|y|^2}{s^2} dyds = \iint_{E(1)} \frac{|y|^2}{s^2} dyds = 4.$$

We omit the details of this last computation. □

2.3.3. Properties of solutions.

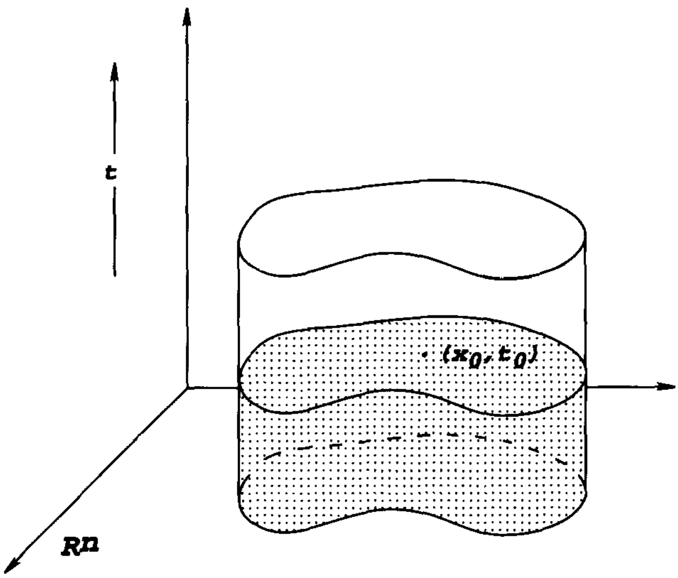
a. Strong maximum principle, uniqueness.

First we employ the mean-value property to give a quick proof of the strong maximum principle.

THEOREM 4 (Strong maximum principle for the heat equation). *Assume $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ solves the heat equation in U_T.*

(i) *Then*

$$\max_{\bar{U}_T} u = \max_{\Gamma_T} u.$$



Strong maximum principle for the heat equation

(ii) *Furthermore, if U is connected and there exists a point $(x_0, t_0) \in U_T$ such that*

$$u(x_0, t_0) = \max_{\bar{U}_T} u,$$

then

$$u \text{ is constant in } \bar{U}_{t_0}.$$

Assertion (i) is the *maximum principle* for the heat equation and (ii) is the *strong maximum principle.* Similar assertions are valid with "min" replacing "max".

Remark. *So if u attains its maximum (or minimum) at an interior point, then u is constant at all earlier times.* This accords with our strong intuitive interpretation of the variable t as denoting time: the solution will be constant on the time interval $[0, t_0]$ provided the initial and boundary conditions are constant. However, the solution may change at times $t > t_0$, provided the boundary conditions alter after t_0. The solution will however not respond to changes in boundary conditions until these changes happen.

Take note that whereas all this is obvious on intuitive, physical grounds, such insights do not constitute a proof. The task is to *deduce* such behavior from the PDE. □

Proof. 1. Suppose there exists a point $(x_0, t_0) \in U_T$ with $u(x_0, t_0) = M := \max_{\bar{U}_T} u$. Then for all sufficiently small $r > 0$, $E(x_0, t_0; r) \subset U_T$; and we employ the mean-value property to deduce

$$M = u(x_0, t_0) = \frac{1}{4r^n} \iint_{E(x_0,t_0;r)} u(y,s) \frac{|x_0 - y|^2}{(t_0 - s)^2}\, dyds \leq M,$$

since

$$1 = \frac{1}{4r^n} \iint_{E(x_0,t_0;r)} \frac{|x_0 - y|^2}{(t_0 - s)^2}\, dyds.$$

Equality holds only if u is identically equal to M within $E(x_0, t_0; r)$. Consequently

$$u(y,s) = M \quad \text{for all } (y,s) \in E(x_0, t_0; r).$$

Draw any line segment L in U_T connecting (x_0, t_0) with some other point $(y_0, s_0) \in U_T$, with $s_0 < t_0$. Consider

$$r_0 := \min\{s \geq s_0 \mid u(x,t) = M \text{ for all points } (x,t) \in L,\ s \leq t \leq t_0\}.$$

Since u is continuous, the minimum is attained. Assume $r_0 > s_0$. Then $u(z_0, r_0) = M$ for some point (z_0, r_0) on $L \cap U_T$ and so $u \equiv M$ on $E(z_0, r_0; r)$ for all sufficiently small $r > 0$. Since $E(z_0, r_0; r)$ contains $L \cap \{r_0 - \sigma \leq t \leq r_0\}$ for some small $\sigma > 0$, we have a contradiction. Thus $r_0 = s_0$, and hence $u \equiv M$ on L.

2. Now fix any point $x \in U$ and any time $0 \leq t < t_0$. There exist points $\{x_0, x_1, \ldots, x_m = x\}$ such that the line segments in $\mathbb{R}^n$ connecting x_{i-1} to x_i lie in U for $i = 1, \ldots, m$. (This follows since the set of points in U which can be so connected to x_0 by a polygonal path is nonempty, open and relatively closed in U.) Select times $t_0 > t_1 > \cdots > t_m = t$. Then the line segments in $\mathbb{R}^{n+1}$ connecting (x_{i-1}, t_{i-1}) to (x_i, t_i) $(i = 1, \ldots, m)$ lie in U_T. According to Step 1, $u \equiv M$ on each such segment and so $u(x,t) = M$. □

Remark. The strong maximum principle implies that if U is connected and $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ satisfies

$$\begin{cases} u_t - \Delta u = 0 & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g & \text{on } U \times \{t = 0\} \end{cases}$$

where $g \geq 0$, then u is positive *everywhere* within U_T if g is positive *somewhere* on U. This is another illustration of infinite propagation speed for disturbances. □

An important application of the maximum principle is the following uniqueness assertion.

THEOREM 5 (Uniqueness on bounded domains). *Let $g \in C(\Gamma_T)$, $f \in C(U_T)$. Then there exists at most one solution $u \in C_1^2(U_T) \cap C(\bar{U}_T)$ of the initial/boundary-value problem*

$$(22) \qquad \begin{cases} u_t - \Delta u = f & \text{in } U_T \\ u = g & \text{on } \Gamma_T. \end{cases}$$

Proof. If u and $\tilde{u}$ are two solutions of (22), apply Theorem 4 to $w := \pm(u - \tilde{u})$. □

We next extend our uniqueness assertion to the *Cauchy problem*, that is, the initial value problem for $U = \mathbb{R}^n$. As we are no longer on a bounded region, we must introduce some control on the behavior of solutions for large $|x|$.

THEOREM 6 (Maximum principle for the Cauchy problem). *Suppose $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ solves*

$$(23) \qquad \begin{cases} u_t - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\} \end{cases}$$

and satisfies the growth estimate

$$(24) \qquad u(x,t) \le Ae^{a|x|^2} \quad (x \in \mathbb{R}^n, 0 \le t \le T)$$

for constants $A, a > 0$. Then

$$\sup_{\mathbb{R}^n \times [0,T]} u = \sup_{\mathbb{R}^n} g.$$

Proof. 1. First assume

$$(25) \qquad 4aT < 1;$$

in which case

$$(26) \qquad 4a(T + \varepsilon) < 1$$

for some $\varepsilon > 0$. Fix $y \in \mathbb{R}^n$, $\mu > 0$, and define

$$v(x,t) := u(x,t) - \frac{\mu}{(T + \varepsilon - t)^{n/2}} e^{\frac{|x-y|^2}{4(T+\varepsilon-t)}} \quad (x \in \mathbb{R}^n,\ t > 0).$$

A direct calculation (cf. §2.3.1) shows

$$v_t - \Delta v = 0 \quad \text{in } \mathbb{R}^n \times (0,T].$$

Fix $r > 0$ and set $U := B^0(y,r)$, $U_T = B^0(y,r) \times (0,T]$. Then according to Theorem 4,

$$\max_{\bar{U}_T} v = \max_{\Gamma_T} v. \tag{27}$$

2. Now if $x \in \mathbb{R}^n$,

$$\begin{aligned} v(x,0) &= u(x,0) - \frac{\mu}{(T+\varepsilon)^{n/2}} e^{\frac{|x-y|^2}{4(T+\varepsilon)}} \\ &\leq u(x,0) = g(x); \end{aligned} \tag{28}$$

and if $|x-y| = r$, $0 \leq t \leq T$, then

$$\begin{aligned} v(x,t) &= u(x,t) - \frac{\mu}{(T+\varepsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\varepsilon-t)}} \\ &\leq Ae^{a|x|^2} - \frac{\mu}{(T+\varepsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\varepsilon-t)}} \qquad \text{by (24)} \\ &\leq Ae^{a(|y|+r)^2} - \frac{\mu}{(T+\varepsilon)^{n/2}} e^{\frac{r^2}{4(T+\varepsilon)}}. \end{aligned}$$

Now according to (26), $\frac{1}{4(T+\varepsilon)} = a+\gamma$ for some $\gamma > 0$. Thus we may continue the calculation above to find

$$v(x,t) \leq Ae^{a(|y|+r)^2} - \mu(4(a+\gamma))^{n/2} e^{(a+\gamma)r^2} \leq \sup_{\mathbb{R}^n} g, \tag{29}$$

for r selected sufficiently large. Thus (27)–(29) imply

$$v(y,t) \leq \sup_{\mathbb{R}^n} g$$

for all $y \in \mathbb{R}^n$, $0 \leq t \leq T$, provided (25) is valid. Let $\mu \to 0$.

3. In the general case that (25) fails, we repeatedly apply the result above on the time intervals $[0,T_1]$, $[T_1, 2T_1,]$, etc., for $T_1 = \frac{1}{8a}$. □

THEOREM 7 (Uniqueness for Cauchy problem). *Let $g \in C(\mathbb{R}^n)$, $f \in C(\mathbb{R}^n \times [0,T])$. Then there exists at most one solution $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ of the initial-value problem*

$$\begin{cases} u_t - \Delta u = f & \text{in } \mathbb{R}^n \times (0,T) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\} \end{cases} \tag{30}$$

satisfying the growth estimate

$$|u(x,t)| \leq Ae^{a|x|^2} \quad (x \in \mathbb{R}^n,\ 0 \leq t \leq T) \tag{31}$$

for constants A, $a > 0$.

Proof. If u and $\tilde{u}$ both satisfy (30), (31), we apply Theorem 6 to $w := \pm(u - \tilde{u})$. □

Remark. There are in fact infinitely many solutions of

$$\begin{cases} u_t - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}; \end{cases} \tag{32}$$

see for instance John [**J**, Chapter 7]. Each of the solutions besides $u \equiv 0$ grows very rapidly as $|x| \to \infty$.

There is an interesting point here: although $u \equiv 0$ is certainly the "physically correct" solution of (32), this initial-value problem in fact admits other, "nonphysical" solutions. Theorem 7 provides a criterion which excludes the "wrong" solutions. We will encounter somewhat analogous situations in our study of Hamilton–Jacobi equations and conservation laws, in Chapters 3, 10 and 11. □

b. Regularity.

We next demonstrate that solutions of the heat equation are automatically smooth.

THEOREM 8 (Smoothness). *Suppose $u \in C^2_1(U_T)$ solves the heat equation in U_T. Then*

$$u \in C^\infty(U_T).$$

This regularity assertion is valid even if u attains nonsmooth boundary values on Γ_T.

Proof. 1. Recall from §A.2 that we write

$$C(x,t;r) = \{(y,s) \mid |x-y| \le r,\ t - r^2 \le s \le t\}$$

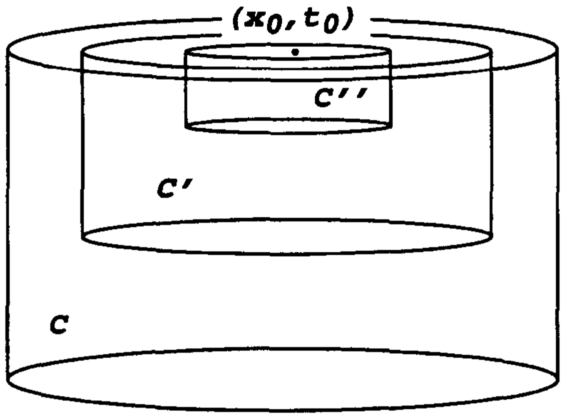
to denote the closed circular cylinder of radius r, height r^2, and top center point (x,t).

Fix $(x_0,t_0) \in U_T$ and choose $r > 0$ so small that $C := C(x_0,t_0;r) \subset U_T$. Define also the smaller cylinders $C' := C(x_0,t_0;\frac{3}{4}r)$, $C'' := C(x_0,t_0;\frac{1}{2}r)$, which have the same top center point (x_0,t_0).

Choose a smooth cutoff function $\zeta = \zeta(x,t)$ such that

$$\begin{cases} 0 \le \zeta \le 1,\ \zeta \equiv 1 \text{ on } C', \\ \zeta \equiv 0 \text{ near the parabolic boundary of } C. \end{cases}$$

Extend $\zeta \equiv 0$ in $(\mathbb{R}^n \times [0,t_0]) - C$.



2. Assume temporarily that $u \in C^\infty(U_T)$ and set

$$v(x,t) := \zeta(x,t)u(x,t) \quad (x \in \mathbb{R}^n,\, 0 \le t \le t_0).$$

Then

$$v_t = \zeta u_t + \zeta_t u, \;\; \Delta v = \zeta \Delta u + 2D\zeta \cdot Du + u\Delta\zeta.$$

Consequently

(33) $$v = 0 \quad \text{on } \mathbb{R}^n \times \{t = 0\},$$

and

(34) $$v_t - \Delta v = \zeta_t u - 2D\zeta \cdot Du - u\Delta\zeta =: \tilde{f}$$

in $\mathbb{R}^n \times (0, t_0)$. Now set

$$\tilde{v}(x,t) := \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\tilde{f}(y,s)\, dyds.$$

According to Theorem 2

(35) $$\begin{cases} \tilde{v}_t - \Delta\tilde{v} = \tilde{f} & \text{in } \mathbb{R}^n \times (0, t_0) \\ \tilde{v} = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Since $|v|, |\tilde{v}| \le A$ for some constant A, Theorem 7 implies $v \equiv \tilde{v}$; that is,

(36) $$v(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\tilde{f}(y,s)\, dyds.$$

Now suppose $(x,t) \in C''$. As $\zeta \equiv 0$ off the cylinder C, (34) and (36) imply

$$u(x,t) = \iint_C \Phi(x-y, t-s)[(\zeta_s(y,s) - \Delta\zeta(y,s))u(y,s) - 2D\zeta(y,s) \cdot Du(y,s)]\, dyds.$$

Note in this expression that the expression in the square brackets vanishes in some region *near* the singularity of Φ. Integrate the last term by parts:

$$
\begin{aligned}
(37) \qquad u(x,t) = \iint_C [&\Phi(x-y,t-s)(\zeta_s(y,s)+\Delta\zeta(y,s)) \\
&+ 2D_y\Phi(x-y,t-s)\cdot D\zeta(y,s)]u(y,s)\,dyds.
\end{aligned}
$$

We have proved this formula assuming $u \in C^\infty$. If u satisfies only the hypotheses of the theorem, we derive (37) with $u^\varepsilon = \eta_\varepsilon * u$ replacing u, η_ε being the standard mollifier in the variables x and t, and let $\varepsilon \to 0$.

3. Formula (37) has the form

$$
(38) \qquad u(x,t) = \iint_C K(x,t,y,s)u(y,s)\,dyds \quad ((x,t) \in C''),
$$

where

$$
K(x,t,y,s) = 0 \quad \text{for all points } (y,s) \in C',
$$

since $\zeta \equiv 1$ on C'. Note also K is smooth on $C - C'$. In view of expression (38), we see u is C^∞ within $C'' = C(x_0,t_0;\frac{1}{2}r)$. □

c. Local estimates for solutions of the heat equation.

Next we record some estimates on the derivatives of solutions to the heat equation, paying attention to the differences between derivatives with respect to x_i $(i = 1,\dots,n)$ and with respect to t.

THEOREM 9 (Estimates on derivatives). *There exists for each pair of integers $k,l = 0,1,\dots,$ a constant $C_{k,l}$ such that*

$$
\max_{C(x,t;\,r/2)} |D_x^k D_t^l u| \le \frac{C_{kl}}{r^{k+2l+n+2}} \|u\|_{L^1(C(x,t;r))}
$$

for all cylinders $C(x,t;r/2) \subset C(x,t;r) \subset U_T$, and all solutions u of the heat equation in U_T.

Proof. 1. Fix some point in U_T. Upon shifting the coordinates, we may as well assume the point is $(0,0)$. Suppose first that the cylinder $C(1) :=$ $C(0,0;1)$ lies in U_T. Let $C\left(\frac{1}{2}\right) := C\left(0,0;\frac{1}{2}\right)$. Then, as in the proof of Theorem 8,

$$
u(x,t) = \iint_{C(1)} K(x,t,y,s)u(y,s)\,dyds \quad ((x,t) \in C(\tfrac{1}{2}))
$$

for some smooth function K. Consequently

$$
\begin{aligned}
(39)\qquad |D_x^k D_t^l u(x,t)| &\le \iint_{C(1)} |D_t^l D_x^k K(x,t,y,s)||u(y,s)|\,dyds \\
&\le C_{kl}\|u\|_{L^1(C(1))}
\end{aligned}
$$

for some constant C_{kl}.

2. Now suppose the cylinder $C(r) := C(0,0;r)$ lies in U_T. Let $C(r/2) = C(0,0;r/2)$. We rescale by defining

$$v(x,t) := u(rx, r^2 t).$$

Then $v_t - \Delta v = 0$ in the cylinder $C(1)$. According to (39),

$$|D_x^k D_t^l v(x,t)| \le C_{kl}\|v\|_{L^1(C(1))} \quad ((x,t)\in C(\tfrac{1}{2})).$$

But $D_x^k D_t^l v(x,t) = r^{2l+k} D_x^k D_t^l u(rx, r^2 t)$ and $\|v\|_{L^1(C(1))} = \frac{1}{r^{n+2}}\|u\|_{L^1(C(r))}$. Therefore

$$\max_{C(r/2)} |D_x^k D_t^l u| \le \frac{C_{kl}}{r^{2l+k+n+2}}\|u\|_{L^1(C(r))}.$$

□

Remark. If u solves the heat equation within U_T, then for each fixed time $0 < t \le T$, the mapping $x \mapsto u(x,t)$ is analytic. (See Mikhailov [**M**].) However the mapping $t \mapsto u(x,t)$ is not in general analytic. □

2.3.4. Energy methods.

a. Uniqueness.

Let us investigate again the initial/boundary-value problem

$$
(40)\qquad \begin{cases} u_t - \Delta u = f & \text{in } U_T \\ \quad\quad\;\; u = g & \text{on } \Gamma_T. \end{cases}
$$

We earlier invoked the maximum principle to show uniqueness, and now—by analogy with §2.2.5—provide an alternative argument based upon integration by parts. We assume as usual that $U \subset \mathbb{R}^n$ is open, bounded and that ∂U is C^1. The terminal time $T > 0$ is given.

THEOREM 10 (Uniqueness). *There exists at most one solution $u \in C_1^2(\bar{U}_T)$ of* (40).

Proof. 1. If $\tilde{u}$ is another solution, $w := u - \tilde{u}$ solves

$$\begin{cases} w_t - \Delta w = 0 & \text{in } U_T \\ w = 0 & \text{on } \Gamma_T. \end{cases} \tag{41}$$

2. Set

$$e(t) := \int_U w^2(x,t)\,dx \quad (0 \leq t \leq T).$$

Then

$$\begin{aligned} \dot{e}(t) &= 2\int_U w w_t\,dx \quad \left(\cdot = \frac{d}{dt}\right) \\ &= 2\int_U w\Delta w\,dx \\ &= -2\int_U |Dw|^2 dx \leq 0, \end{aligned}$$

and so

$$e(t) \leq e(0) = 0 \quad (0 \leq t \leq T).$$

Consequently $w = u - \tilde{u} \equiv 0$ in U_T. □

Observe that the foregoing is a time-dependent variant of the proof of Theorem 16 in §2.2.5.

b. Backwards uniqueness.

A rather more subtle question concerns uniqueness *backwards in time* for the heat equation. For this, suppose u and $\tilde{u}$ are both smooth solutions of the heat equation in U_T, with the same boundary conditions on ∂U:

$$\begin{cases} u_t - \Delta u = 0 & \text{in } U_T \\ u = g & \text{on } \partial U \times [0,T], \end{cases} \tag{42}$$

$$\begin{cases} \tilde{u}_t - \Delta \tilde{u} = 0 & \text{in } U_T \\ \tilde{u} = g & \text{on } \partial U \times [0,T], \end{cases} \tag{43}$$

for some function g. Note carefully that we are *not* supposing $u = \tilde{u}$ at time $t = 0$.

THEOREM 11 (Backwards uniqueness). *Suppose* $u, \tilde{u} \in C^2(\bar{U}_T)$ *solve* (42), (43). *If*

$$u(x,T) = \tilde{u}(x,T) \quad (x \in U),$$

then

$$u \equiv \tilde{u} \quad \textit{within } U_T.$$

In other words, if two temperature distributions on U agree at some time $T > 0$, and have had the same boundary values for times $0 \leq t \leq T$, then these temperatures must have been identically equal within U at all earlier times. This is not at all obvious.

Proof. 1. Write $w := u - \tilde{u}$ and, as in the proof of Theorem 10, set

$$e(t) := \int_U w^2(x,t)\,dx \quad (0 \leq t \leq T).$$

As before

$$\dot{e}(t) = -2\int_U |Dw|^2 dx \quad \left(\dot{} = \frac{d}{dt}\right). \tag{44}$$

Furthermore

$$\begin{aligned} \ddot{e}(t) &= -4\int_U Dw \cdot Dw_t\,dx \\ &= 4\int_U \Delta w w_t\,dx \\ &= 4\int_U (\Delta w)^2 dx \quad \text{by (41).} \end{aligned} \tag{45}$$

Now since $w = 0$ on ∂U,

$$\begin{aligned} \int_U |Dw|^2 dx &= -\int_U w\Delta w\,dx \\ &\leq \left(\int_U w^2 dx\right)^{1/2} \left(\int_U (\Delta w)^2 dx\right)^{1/2}. \end{aligned}$$

Thus (44) and (45) imply

$$\begin{aligned} (\dot{e}(t))^2 &= 4\left(\int_U |Dw|^2 dx\right)^2 \\ &\leq \left(\int_U w^2 dx\right)\left(4\int_U (\Delta w)^2 dx\right) \\ &= e(t)\ddot{e}(t). \end{aligned}$$

Hence

$$\ddot{e}(t)e(t) \geq (\dot{e}(t))^2 \quad (0 \leq t \leq T). \tag{46}$$

2. Now if $e(t) = 0$ for all $0 \leq t \leq T$, we are done. Otherwise there exists an interval $[t_1, t_2] \subset [0, T]$, with

$$e(t) > 0 \quad \text{for } t_1 \leq t < t_2, \quad e(t_2) = 0. \tag{47}$$

3. Now write

$$f(t) := \log e(t) \quad (t_1 \leq t < t_2). \tag{48}$$

Then

$$\ddot{f}(t) = \frac{\ddot{e}(t)}{e(t)} - \frac{\dot{e}(t)^2}{e(t)^2} \geq 0 \quad \text{by (46);}$$

and so f is convex on the interval (t_1, t_2). Consequently if $0 < \tau < 1$, $t_1 < t < t_2$, we have

$$f((1-\tau)t_1 + \tau t) \leq (1-\tau)f(t_1) + \tau f(t).$$

Recalling (48), we deduce

$$e((1-\tau)t_1 + \tau t) \leq e(t_1)^{1-\tau} e(t)^{\tau},$$

and so

$$0 \leq e((1-\tau)t_1 + \tau t_2) \leq e(t_1)^{1-\tau} e(t_2)^{\tau} \quad (0 < \tau < 1).$$

But in view of (47) this inequality implies $e(t) = 0$ for all times $t_1 \leq t \leq t_2$, a contradiction. □

2.4. WAVE EQUATION

In this section we investigate the *wave equation*

$$u_{tt} - \Delta u = 0 \tag{1}$$

and the *nonhomogeneous wave equation*

$$u_{tt} - \Delta = f, \tag{2}$$

subject to appropriate initial and boundary conditions. Here $t > 0$ and $x \in U$, where $U \subset \mathbb{R}^n$ is open. The unknown is $u : \bar{U} \times [0, \infty) \to \mathbb{R}$, $u = u(x,t)$, and the Laplacian Δ is taken with respect to the spatial variables

$x = (x_1, \ldots, x_n)$. In (2) the function $f : U \times [0, \infty) \to \mathbb{R}$ is given. A common abbreviation is to write

$$\Box u = u_{tt} - \Delta u.$$

We shall discover that solutions of the wave equation behave quite differently than solutions of Laplace's equation or the heat equation. For example, these solutions are generally not C^∞, exhibit finite speed of propagation, etc.

Physical interpretation. The wave equation is a simplified model for a vibrating string ($n = 1$), membrane ($n = 2$), or elastic solid ($n = 3$). In these physical interpretations $u(x, t)$ represents the displacement in some direction of the point x at time $t \geq 0$.

Let V represent any smooth subregion of U. The acceleration within V is then

$$\frac{d^2}{dt^2} \int_V u \, dx = \int_V u_{tt} \, dx$$

and the net contact force is

$$-\int_{\partial V} \mathbf{F} \cdot \nu \, dS,$$

where $\mathbf{F}$ denotes the force acting on V through ∂V and the mass density is taken to be unity. Newton's law asserts the mass times the acceleration equals the net force:

$$\int_V u_{tt} \, dx = -\int_{\partial V} \mathbf{F} \cdot \nu \, dS.$$

This identity obtains for each subregion V and so

$$u_{tt} = -\operatorname{div} \mathbf{F}.$$

For elastic bodies, $\mathbf{F}$ is a function of the displacement gradient Du; whence

$$u_{tt} + \operatorname{div} \mathbf{F}(Du) = 0.$$

For small Du, the linearization $\mathbf{F}(Du) \approx -aDu$ is often appropriate; and so

$$u_{tt} - a\Delta u = 0.$$

This is the wave equation if $a = 1$. □

This physical interpretation strongly suggests it will be mathematically appropriate to specify *two* initial conditions, on the *displacement* u and the *velocity* u_t, at time $t = 0$.

2.4.1. Solution by spherical means.

We began §§2.2.1 and 2.3.1 by searching for certain scaling invariant solutions of Laplace's equation and the heat equation. For the wave equation however we will instead present the (reasonably) elegant method of solving (1) first for $n = 1$ directly and then for $n \geq 2$ by the method of spherical means.

a. Solution for n = 1, d'Alembert's formula.

We first focus our attention on the initial-value problem for the one-dimensional wave equation in all of $\mathbb{R}$:

$$\text{(3)} \qquad \begin{cases} u_{tt} - u_{xx} = 0 & \text{in } \mathbb{R} \times (0, \infty) \\ u = g, \ u_t = h & \text{on } \mathbb{R} \times \{t = 0\}, \end{cases}$$

where g, h are given. We desire to derive a formula for u in terms of g and h.

Let us first note the PDE in (3) can be "factored", to read

$$\text{(4)} \qquad \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) u = u_{tt} - u_{xx} = 0.$$

Write

$$\text{(5)} \qquad v(x,t) := \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) u(x,t).$$

Then (4) says

$$v_t(x,t) + v_x(x,t) = 0 \quad (x \in \mathbb{R}, \ t > 0).$$

This is a transport equation with constant coefficients. Applying formula (3) from §2.1.1 (with $n = 1$, $b = 1$), we find

$$\text{(6)} \qquad v(x,t) = a(x - t)$$

for $a(x) := v(x, 0)$. Combining now (4)–(6), we obtain

$$u_t(x,t) - u_x(x,t) = a(x - t) \quad \text{in } \mathbb{R} \times (0, \infty).$$

This is a nonhomogeneous transport equation; and so formula (5) from §2.1.2 (with $n = 1$, $b = -1$, $f(x,t) = a(x - t)$) implies

$$\text{(7)} \qquad \begin{aligned} u(x,t) &= \int_0^t a(x + (t - s) - s)\, ds + b(x + t) \\ &= \frac{1}{2} \int_{x-t}^{x+t} a(y)\, dy + b(x + t), \end{aligned}$$

where we have $b(x) := u(x, 0)$.

We lastly invoke the initial conditions in (3) to compute a and b. The first initial condition in (3) gives

$$b(x) = g(x) \quad (x \in \mathbb{R});$$

whereas the second initial condition and (5) imply

$$a(x) = v(x, 0) = u_t(x, 0) - u_x(x, 0) = h(x) - g'(x) \quad (x \in \mathbb{R}).$$

Our substituting into (7) now yields

$$u(x, t) = \frac{1}{2}\int_{x-t}^{x+t} h(y) - g'(y)\, dy + g(x + t).$$

Hence

$$(8) \qquad u(x, t) = \frac{1}{2}[g(x + t) + g(x - t)] + \frac{1}{2}\int_{x-t}^{x+t} h(y)\, dy \quad (x \in \mathbb{R},\ t \geq 0).$$

This is *d'Alembert's formula.*

We have derived formula (8) assuming u is a (sufficiently smooth) solution of (3). We need to check that this really is a solution.

THEOREM 1 (Solution of wave equation, $n = 1$). *Assume $g \in C^2(\mathbb{R})$, $h \in C^1(\mathbb{R})$, and define u by d'Alembert's formula* (8). *Then*

(i) $u \in C^2(\mathbb{R}^n \times [0, \infty))$,

(ii) $u_{tt} - u_{xx} = 0$ *in* $\mathbb{R} \times (0, \infty)$,

and

(iii) $\displaystyle\lim_{\substack{(x,t)\to(x^0,0)\\ t>0}} u(x, t) = g(x^0), \quad \lim_{\substack{(x,t)\to(x^0,0)\\ t>0}} u_t(x, t) = h(x^0)$

for each point $x^0 \in \mathbb{R}$.

The proof is a straightforward calculation.

Remarks. (i) In view of (8), our solution u has the form

$$u(x, t) = F(x + t) + G(x - t)$$

for appropriate functions F and G. Conversely any function of this form solves $u_{tt} - u_{xx} = 0$. Hence the general solution of the one-dimensional wave equation is a sum of the general solution of $u_t - u_x = 0$ and the general solution of $u_t + u_x = 0$. This is a consequence of the factorization (4).

(ii) We see from (8) that if $g \in C^k$ and $h \in C^{k-1}$, then $u \in C^k$, but is not in general smoother. Thus the wave equation does *not* cause instantaneous smoothing of the initial data, as does the heat equation. □

A reflection method. To illustrate a further application of d'Alembert's formula, let us next consider this initial/boundary-value problem on the half-line $\mathbb{R}_+ = \{x > 0\}$:

$$(9) \qquad \begin{cases} u_{tt} - u_{xx} = 0 & \text{in } \mathbb{R}_+ \times (0,\infty) \\ u = g,\ u_t = h & \text{on } \mathbb{R}_+ \times \{t = 0\} \\ u = 0 & \text{on } \{x = 0\} \times (0,\infty), \end{cases}$$

where g, h are given, with $g(0) = h(0) = 0$.

We convert (9) into the form (3) by extending u, g, h to all of $\mathbb{R}$ by *odd reflection.* That is, we set

$$\tilde{u}(x,t) := \begin{cases} u(x,t) & (x \geq 0,\ t \geq 0) \\ -u(-x,t) & (x \leq 0,\ t \geq 0), \end{cases}$$

$$\tilde{g}(x) := \begin{cases} g(x) & (x \geq 0) \\ -g(-x) & (x \leq 0), \end{cases}$$

$$\tilde{h}(x) := \begin{cases} h(x) & (x \geq 0) \\ -h(-x) & (x \leq 0). \end{cases}$$

Then (9) becomes

$$\begin{cases} \tilde{u}_{tt} = \tilde{u}_{xx} & \text{in } \mathbb{R} \times (0,\infty) \\ \tilde{u} = \tilde{g},\ \tilde{u}_t = \tilde{h} & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases}$$

Hence d'Alembert's formula (8) implies

$$\tilde{u}(x,t) = \frac{1}{2}[\tilde{g}(x+t) + \tilde{g}(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} \tilde{h}(y)\,dy.$$

Recalling the definitions of $\tilde{u}, \tilde{g}, \tilde{h}$ above, we can transform this expression to read for $x \geq 0$, $t \geq 0$:

$$(10) \quad u(x,t) = \begin{cases} \frac{1}{2}[g(x+t) + g(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} h(y)\,dy & \text{if } x \geq t \geq 0 \\ \frac{1}{2}[g(x+t) - g(t-x)] + \frac{1}{2}\int_{-x+t}^{x+t} h(y)\,dy & \text{if } 0 \leq x \leq t. \end{cases}$$

If $h \equiv 0$, we can understand formula (10) as saying that an initial displacement g splits into two parts, one moving to the right with speed one and the other to the left with speed one. The latter then reflects off the point $x = 0$, where the vibrating string is held fixed. □

b. Spherical means.

Now suppose $n \geq 2$, $m \geq 2$, and $u \in C^m(\mathbb{R}^n \times [0,\infty))$ solves the initial-value problem

$$\text{(11)} \qquad \begin{cases} u_{tt} - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g, \; u_t = h & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

We intend to derive an explicit formula for u in terms of g, h. The plan will be to study first the average of u over certain spheres. These averages, taken as functions of the time t and the radius r, turn out to solve the Euler–Poisson–Darboux equation, a PDE which we can for odd n convert into the ordinary one-dimensional wave equation. Applying d'Alembert's formula, or more precisely its variant (10), eventually leads us to a formula for the solution.

Notation. (i) Let $x \in \mathbb{R}^n$, $t > 0$, $r > 0$. Define

$$\text{(12)} \qquad U(x;r,t) := ⨍_{\partial B(x,r)} u(y,t)\, dS(y),$$

the average of $u(\cdot,t)$ over the sphere $\partial B(x,r)$.

(ii) Similarly,

$$\text{(13)} \qquad \begin{cases} G(x;r) := ⨍_{\partial B(x,r)} g(y)\, dS(y) \\ H(x;r) := ⨍_{\partial B(x,r)} h(y)\, dS(y). \end{cases}$$

For fixed x, we hereafter regard U as a function of r and t, and discover a partial differential equation U solves:

LEMMA 1 (Euler–Poisson–Darboux equation). *Fix $x \in \mathbb{R}^n$, and let u satisfy* (11). *Then $U \in C^m(\bar{\mathbb{R}}_+ \times [0,\infty))$ and*

$$\text{(14)} \qquad \begin{cases} U_{tt} - U_{rr} - \frac{n-1}{r} U_r = 0 & \text{in } \mathbb{R}_+ \times (0,\infty) \\ U = G, \; U_t = H & \text{on } \mathbb{R}_+ \times \{t=0\}. \end{cases}$$

The partial differential equation in (14) is the *Euler–Poisson–Darboux equation.* (Note that the term $U_{rr} + \frac{n-1}{r} U_r$ is the radial part of the Laplacian Δ in polar coordinates.)

Proof. 1. As in the proof of Theorem 2 in §2.2.2 we compute for $r > 0$

$$U_r(x;r,t) = \frac{r}{n} ⨍_{B(x,r)} \Delta u(y,t)\, dy. \tag{15}$$

From this equality we deduce $\lim_{r\to 0^+} U_r(x;r,t) = 0$. We next differentiate (15), to discover after some computations that

$$U_{rr}(x;r,t) = ⨍_{\partial B(x,r)} \Delta u\, dS + \left(\frac{1}{n} - 1\right) ⨍_{B(x,r)} \Delta u\, dy. \tag{16}$$

Thus $\lim_{r\to 0^+} U_{rr}(x;r,t) = \frac{1}{n}\Delta u(x,t)$. Using formula (16) we can similarly compute U_{rrr}, etc., and so verify that $U \in C^m(\bar{\mathbb{R}}_+ \times [0,\infty))$.

2. Continuing the calculation above, we see from (15) that

$$\begin{aligned} U_r &= \frac{r}{n} ⨍_{B(x,r)} u_{tt}\, dy \quad \text{by (11)} \\ &= \frac{1}{n\alpha(n)} \frac{1}{r^{n-1}} \int_{B(x,r)} u_{tt}\, dy. \end{aligned}$$

Thus

$$r^{n-1} U_r = \frac{1}{n\alpha(n)} \int_{B(x,r)} u_{tt}\, dy,$$

and so

$$\begin{aligned} \left(r^{n-1} U_r\right)_r &= \frac{1}{n\alpha(n)} \int_{\partial B(x,r)} u_{tt}\, dS \\ &= r^{n-1} ⨍_{\partial B(x,r)} u_{tt}\, dS = r^{n-1} U_{tt}. \end{aligned}$$

□

c. Solution for n = 3, 2, Kirchhoff's and Poisson's formulas.

The overall plan in the ensuing subsections will be to transform the Euler–Poisson–Darboux equation (14) into the usual one-dimensional wave equation. As the full procedure is rather complicated, we pause here to handle the simpler cases $n = 3, 2$, in that order.

Solution for n = 3. Let us therefore hereafter take $n = 3$, and suppose $u \in C^2(\mathbb{R}^3 \times [0,\infty))$ solves the initial-value problem (11). We recall the definitions (12), (13) of U, G, H, and then set

$$\tilde{U} := rU, \tag{17}$$

$$\tilde{G} := rG, \ \tilde{H} := rH. \tag{18}$$

We now assert that $\tilde{U}$ solves

$$\begin{cases} \tilde{U}_{tt} - \tilde{U}_{rr} = 0 & \text{in } \mathbb{R}_+ \times (0,\infty) \\ \tilde{U} = \tilde{G}, \ \tilde{U}_t = \tilde{H} & \text{on } \mathbb{R}_+ \times \{t=0\} \\ \tilde{U} = 0 & \text{on } \{r=0\} \times (0,\infty). \end{cases} \tag{19}$$

Indeed

$$\begin{aligned} \tilde{U}_{tt} &= rU_{tt} \\ &= r\left[U_{rr} + \frac{2}{r}U_r\right] \quad \text{by (14), with } n=3 \\ &= rU_{rr} + 2U_r = (U + rU_r)_r \\ &= \tilde{U}_{rr}. \end{aligned}$$

Applying formula (10) to (19), we find for $0 \le r \le t$

$$\tilde{U}(x;r,t) = \frac{1}{2}[\tilde{G}(r+t) - \tilde{G}(t-r)] + \frac{1}{2}\int_{-r+t}^{r+t} \tilde{H}(y)\,dy. \tag{20}$$

Since (12) implies $u(x,t) = \lim_{r\to 0^+} U(x;r,t)$, we conclude from (17), (18), (20) that

$$\begin{aligned} u(x,t) &= \lim_{r\to 0^+} \frac{\tilde{U}(x;r,t)}{r} \\ &= \lim_{r\to 0^+} \left[\frac{\tilde{G}(t+r) - \tilde{G}(t-r)}{2r} + \frac{1}{2r}\int_{t-r}^{t+r} \tilde{H}(y)\,dy\right] \\ &= \tilde{G}'(t) + \tilde{H}(t). \end{aligned}$$

Owing then to (13), we deduce

$$u(x,t) = \frac{\partial}{\partial t}\left(t ⨍_{\partial B(x,t)} g\,dS\right) + t ⨍_{\partial B(x,t)} h\,dS. \tag{21}$$

But

$$⨍_{\partial B(x,t)} g(y)\,dS(y) = ⨍_{\partial B(0,1)} g(x+tz)\,dS(z);$$

and so

$$\begin{aligned} \frac{\partial}{\partial t}\left(⨍_{\partial B(x,t)} g\,dS\right) &= ⨍_{\partial B(0,1)} Dg(x+tz)\cdot z\,dS(z) \\ &= ⨍_{\partial B(x,t)} Dg(y)\cdot\left(\frac{y-x}{t}\right)dS(y). \end{aligned}$$

Returning to (21), we therefore conclude

$$(22)\quad u(x,t) = \fint_{\partial B(x,t)} th(y) + g(y) + Dg(y)\cdot(y-x)\,dS(y) \quad (x \in \mathbb{R}^3,\ t > 0).$$

This is *Kirchhoff's formula* for the solution of the initial-value problem (11) in three dimensions.

Solution for n = 2. No transformation like (17) works to convert the Euler–Poisson–Darboux equation into the one-dimensional wave equation when $n = 2$. Instead we will take the initial-value problem (11) for $n = 2$ and simply regard it as a problem for $n = 3$, in which the third spatial variable x_3 does not appear.

Indeed, assuming $u \in C^2(\mathbb{R}^2 \times [0,\infty))$ solves (11) for $n = 2$, let us write

$$(23)\qquad \bar{u}(x_1, x_2, x_3, t) := u(x_1, x_2, t).$$

Then (11) implies

$$(24)\qquad \begin{cases} \bar{u}_{tt} - \Delta \bar{u} = 0 & \text{in } \mathbb{R}^3 \times (0,\infty) \\ \bar{u} = \bar{g},\ \bar{u}_t = \bar{h} & \text{on } \mathbb{R}^3 \times \{t = 0\}, \end{cases}$$

for

$$\bar{g}(x_1, x_2, x_3) := g(x_1, x_2),\ \bar{h}(x_1, x_2, x_3) := h(x_1, x_2).$$

If we write $x = (x_1, x_2) \in \mathbb{R}^2$ and $\bar{x} = (x_1, x_2, 0) \in \mathbb{R}^3$, then (24) and Kirchhoff's formula (in the form (21)) imply

$$(25)\qquad \begin{aligned} u(x,t) &= \bar{u}(\bar{x}, t) \\ &= \frac{\partial}{\partial t}\left(t \fint_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} \right) + t \fint_{\partial \bar{B}(\bar{x},t)} \bar{h}\, d\bar{S}, \end{aligned}$$

where $\bar{B}(\bar{x}, t)$ denotes the ball in $\mathbb{R}^3$ with center $\bar{x}$, radius $t > 0$, and $d\bar{S}$ denotes two-dimensional surface measure on $\partial \bar{B}(\bar{x}, t)$. We simplify (25) by observing

$$\begin{aligned} \fint_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} &= \frac{1}{4\pi t^2} \int_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} \\ &= \frac{2}{4\pi t^2} \int_{B(x,t)} g(y)(1 + |D\gamma(y)|^2)^{1/2} dy, \end{aligned}$$

where $\gamma(y) = (t^2 - |y - x|^2)^{1/2}$ for $y \in B(x,t)$. The factor "2" enters since $\partial \bar{B}(\bar{x}, t)$ consists of two hemispheres. Observe that $(1 + |D\gamma|^2)^{1/2} =$

$t(t^2 - |y-x|^2)^{-1/2}$. Therefore

$$\fint_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} = \frac{1}{2\pi t} \int_{B(x,t)} \frac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy$$
$$= \frac{t}{2} \fint_{B(x,t)} \frac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy.$$

Consequently formula (25) becomes

$$u(x,t) = \frac{1}{2}\frac{\partial}{\partial t}\left(t^2 \fint_{B(x,t)} \frac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy \right) + \frac{t^2}{2} \fint_{B(x,t)} \frac{h(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy. \tag{26}$$

But

$$t^2 \fint_{B(x,t)} \frac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy = t \fint_{B(0,1)} \frac{g(x+tz)}{(1 - |z|^2)^{1/2}}\, dz,$$

and so

$$\frac{\partial}{\partial t}\left(t^2 \fint_{B(x,t)} \frac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy \right)$$
$$= \fint_{B(0,1)} \frac{g(x+tz)}{(1 - |z|^2)^{1/2}}\, dz + t \fint_{B(0,1)} \frac{Dg(x+tz)\cdot z}{(1 - |z|^2)^{1/2}}\, dz$$
$$= t \fint_{B(x,t)} \frac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\, dy + t \fint_{B(x,t)} \frac{Dg(y)\cdot(y-x)}{(t^2 - |y-x|^2)^{1/2}}\, dy.$$

Hence we can rewrite (26) and obtain the relation

$$u(x,t) = \frac{1}{2} \fint_{B(x,t)} \frac{tg(y) + t^2 h(y) + tDg(y)\cdot(y-x)}{(t^2 - |y-x|^2)^{1/2}}\, dy \tag{27}$$

for $x \in \mathbb{R}^2$, $t > 0$. This is *Poisson's formula* for the solution of the initial-value problem (11) in two dimensions.

This trick of solving the problem for $n = 3$ first and then dropping to $n = 2$ is the *method of descent*.

d. Solution for odd n.

In this subsection we solve the Euler–Poisson–Darboux PDE for odd $n \geq 3$. We first record some technical facts.

LEMMA 2 (Some useful identities). *Let $\phi : \mathbb{R} \to \mathbb{R}$ be C^{k+1}. Then for $k = 1, 2, \ldots$:*

(i) $\left(\frac{d^2}{dr^2}\right)\left(\frac{1}{r}\frac{d}{dr}\right)^{k-1}\left(r^{2k-1}\phi(r)\right) = \left(\frac{1}{r}\frac{d}{dr}\right)^{k}\left(r^{2k}\frac{d\phi}{dr}(r)\right)$,

(ii) $\left(\frac{1}{r}\frac{d}{dr}\right)^{k-1}\left(r^{2k-1}\phi(r)\right) = \sum_{j=0}^{k-1}\beta_j^k r^{j+1}\frac{d^j\phi}{dr^j}(r)$,
where the constants β_j^k $(j = 0, \ldots, k-1)$ are independent of ϕ.

Furthermore,
(iii) $\beta_0^k = 1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2k-1)$.

The proof by induction is left as an exercise.

Now assume

$$n \geq 3 \text{ is an odd integer}$$

and set

$$n = 2k + 1 \quad (k \geq 1).$$

Henceforth suppose $u \in C^{k+1}(\mathbb{R}^n \times [0, \infty))$ solves the initial-value problem (11). Then the function U defined by (12) is C^{k+1}.

Notation. We write

$$\begin{cases} \tilde{U}(r,t) := \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}U(x;r,t)\right) \\ \tilde{G}(r) := \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}G(x;r)\right) & (r > 0,\ t \geq 0) \\ \tilde{H}(r) := \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}H(x;r)\right). \end{cases} \tag{28}$$

Then

$$\tilde{U}(r,0) = \tilde{G}(r),\ \tilde{U}_t(r,0) = \tilde{H}(r). \tag{29}$$

Next we combine Lemma 1 and the identities provided by Lemma 2 to demonstrate that the transformation (28) of U into $\tilde{U}$ in effect converts the Euler–Poisson–Darboux equation into the wave equation.

LEMMA 3 ($\tilde{U}$ solves the one-dimensional wave equation). *We have*

$$\begin{cases} \tilde{U}_{tt} - \tilde{U}_{rr} = 0 & \text{in } \mathbb{R}_+ \times (0,\infty) \\ \tilde{U} = \tilde{G},\ \tilde{U}_t = \tilde{H} & \text{on } \mathbb{R}_+ \times \{t = 0\} \\ \tilde{U} = 0 & \text{on } \{r = 0\} \times (0,\infty). \end{cases}$$

Proof. If $r > 0$,

$$
\begin{aligned}
\tilde{U}_{rr} &= \left(\frac{\partial^2}{\partial r^2}\right)\left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}(r^{2k-1}U) \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k}(r^{2k}U_r) \quad \text{by Lemma 2,(i)} \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}[r^{2k-1}U_{rr} + 2kr^{2k-2}U_r] \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left[r^{2k-1}\left(U_{rr} + \frac{n-1}{r}U_r\right)\right] \quad (n = 2k+1) \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}(r^{2k-1}U_{tt}) = \tilde{U}_{tt},
\end{aligned}
$$

the next-to-last equality holding according to (14). Using Lemma 2,(ii) we conclude as well that $\tilde{U} = 0$ on $\{r = 0\}$. □

In view of Lemma 3, (29), and formula (10), we conclude for $0 \le r \le t$ that

$$
\tilde{U}(r,t) = \frac{1}{2}[\tilde{G}(r+t) - \tilde{G}(t-r)] + \frac{1}{2}\int_{t-r}^{t+r} \tilde{H}(y)\,dy \tag{30}
$$

for all $r \in \mathbb{R}$, $t \ge 0$. But recall $u(x,t) = \lim_{r\to 0} U(x;r,t)$. Furthermore Lemma 2,(ii) asserts

$$
\begin{aligned}
\tilde{U}(r,t) &= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}(r^{2k-1}U(x;r,t)) \\
&= \sum_{j=0}^{k-1} \beta_j^k r^{j+1} \frac{\partial^j}{\partial r^j} U(x;r,t);
\end{aligned}
$$

and so

$$
\lim_{r\to 0} \frac{\tilde{U}(r,t)}{\beta_0^k r} = \lim_{r\to 0} U(x;r,t) = u(x,t).
$$

Thus (30) implies

$$
\begin{aligned}
u(x,t) &= \frac{1}{\beta_0^k} \lim_{r\to 0} \left[\frac{\tilde{G}(t+r) - \tilde{G}(t-r)}{2r} + \frac{1}{2r}\int_{t-r}^{t+r} \tilde{H}(y)\,dy\right] \\
&= \frac{1}{\beta_0^k}[\tilde{G}'(t) + \tilde{H}(t)].
\end{aligned}
$$

Finally then, since $n = 2k+1$, (30) and Lemma 2,(iii) yield this representation formula:

$$(31)\quad \begin{cases} u(x,t) = \dfrac{1}{\gamma_n}\left[\left(\dfrac{\partial}{\partial t}\right)\left(\dfrac{1}{t}\dfrac{\partial}{\partial t}\right)^{\frac{n-3}{2}}\left(t^{n-2}⨍_{\partial B(x,t)} g\,dS\right)\right. \\ \qquad\qquad \left. + \left(\dfrac{1}{t}\dfrac{d}{dt}\right)^{\frac{n-3}{2}}\left(t^{n-2}⨍_{\partial B(x,t)} h\,dS\right)\right] \\ \text{where } n \text{ is } \textit{odd} \text{ and } \gamma_n = 1\cdot 3\cdot 5\cdots\cdots(n-2), \end{cases}$$

for $x \in \mathbb{R}^n$, $t > 0$.

We note that $\gamma_3 = 1$, and so (31) agrees for $n = 3$ with (21) and thus with Kirchhoff's formula (22).

It remains to check that formula (31) really provides a solution of (11).

THEOREM 2 (Solution of wave equation in odd dimensions). *Assume n is an odd integer, $n \geq 3$, and suppose also $g \in C^{m+1}(\mathbb{R}^n)$, $h \in C^m(\mathbb{R}^n)$, for $m = \frac{n+1}{2}$. Define u by* (31). *Then*

(i) $u \in C^2(\mathbb{R}^n \times [0,\infty))$,

(ii) $u_{tt} - \Delta u = 0 \quad$ *in* $\mathbb{R}^n \times (0,\infty)$,

and

(iii) $\displaystyle\lim_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n,\ t>0}} u(x,t) = g(x^0), \quad \lim_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n,\ t>0}} u_t(x,t) = h(x^0)$

for each point $x^0 \in \mathbb{R}^n$.

Proof. 1. Suppose first $g \equiv 0$; so that

$$(32)\qquad u(x,t) = \frac{1}{\gamma_n}\left(\frac{1}{t}\frac{d}{dt}\right)^{\frac{n-3}{2}}\left(t^{n-2}H(x;t)\right).$$

Then Lemma 2,(i) lets us compute

$$u_{tt} = \frac{1}{\gamma_n}\left(\frac{1}{t}\frac{d}{dt}\right)^{\frac{n-1}{2}}\left(t^{n-1}H_t\right).$$

From the calculation in the proof of Theorem 2 in §2.2.2, we see as well that

$$H_t = \frac{t}{n}⨍_{B(x,t)} \Delta h\,dy.$$

Consequently

$$\begin{aligned} u_{tt} &= \frac{1}{n\alpha(n)\gamma_n}\left(\frac{1}{t}\frac{d}{dt}\right)^{\frac{n-1}{2}}\left(\int_{B(x,t)} \Delta h\,dy\right) \\ &= \frac{1}{n\alpha(n)\gamma_n}\left(\frac{1}{t}\frac{d}{dt}\right)^{\frac{n-3}{2}}\left(\frac{1}{t}\int_{\partial B(x,t)} \Delta h\,dS\right). \end{aligned}$$

On the other hand,

$$\begin{aligned}\Delta H(x;t) &= \Delta_x ⨍_{\partial B(0,t)} h(x+y)\,dS(y)\\ &= ⨍_{\partial B(x,t)} \Delta h\,dS.\end{aligned}$$

Consequently (32) and the calculations above imply $u_{tt} = \Delta u$ in $\mathbb{R}^n \times (0,\infty)$.

A similar computation works if $h \equiv 0$.

2. We leave it as an exercise to confirm, using Lemma 2,(ii)-(iii), that u takes on the correct initial conditions. □

Remarks. (i) Notice that to compute $u(x,t)$ we need only have information on g, h and their derivatives on the sphere $\partial B(x,t)$, and not on the entire ball $B(x,t)$.

(ii) Comparing formula (31) with d'Alembert's formula (8) ($n = 1$), we observe that the latter does not involve the derivatives of g. This suggests that for $n > 1$, a solution of the wave equation (11) need not for times $t > 0$ be as smooth as its initial value g: irregularities in g may focus at times $t > 0$, thereby causing u to be less regular. (We will see later in §2.4.3 that the "energy norm" of u does *not* deteriorate for $t > 0$.)

(iii) Once again (as in the case $n = 1$) we see the phenomenon of finite propagation speed of the initial disturbance.

(iv) A completely different derivation of formula (31) (using the heat equation!) is in §4.3.2. □

e. Solution for even n.

Assume now

$$n \text{ is an even integer.}$$

Suppose u is a C^m solution of (11), $m = \frac{n+2}{2}$. We want to fashion a representation formula like (31) for u. The trick, as above for $n = 2$, is to note

$$\bar{u}(x_1,\dots,x_{n+1},t) := u(x_1,\dots,x_n,t) \tag{33}$$

solves the wave equation in $\mathbb{R}^{n+1} \times (0,\infty)$, with the initial conditions

$$\bar{u} = \bar{g},\ \bar{u}_t = \bar{h} \quad \text{on } \mathbb{R}^{n+1} \times \{t = 0\},$$

where

$$(34)\qquad \begin{cases} \bar{g}(x_1,\dots,x_{n+1}) := g(x_1,\dots,x_n) \\ \bar{h}(x_1,\dots,x_{n+1}) := h(x_1,\dots,x_n). \end{cases}$$

As $n+1$ is odd, we may employ (31) (with $n+1$ replacing n) to secure a representation formula for $\bar{u}$ in terms of $\bar{g}, \bar{h}$. But then (33) and (34) yield at once a formula for u in terms of g, h. This is again the method of descent.

To carry out the details, let us fix $x \in \mathbb{R}^n$, $t > 0$, and write $\bar{x} = (x_1, \dots, x_n, 0) \in \mathbb{R}^{n+1}$. Then (31), with $n+1$ replacing n, gives

$$(35)\qquad \begin{aligned} u(x,t) = \frac{1}{\gamma_{n+1}} \Bigg[& \frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-2}{2}} \left(t^{n-1} ⨍_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} \right) \\ & + \left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-2}{2}} \left(t^{n-1} ⨍_{\partial \bar{B}(\bar{x},t)} \bar{h}\, d\bar{S} \right) \Bigg], \end{aligned}$$

$\bar{B}(\bar{x},t)$ denoting the ball in $\mathbb{R}^{n+1}$ with center $\bar{x}$ and radius t, and $d\bar{S}$ n-dimensional surface measure on $\partial\bar{B}(\bar{x},t)$. Now

$$(36)\qquad ⨍_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} = \frac{1}{(n+1)\alpha(n+1)t^n} \int_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S}.$$

Note that $\partial\bar{B}(x,t) \cap \{y_{n+1} \geq 0\}$ is the graph of the function $\gamma(y) := (t^2 - |y-x|^2)^{1/2}$ for $y \in B(x,t) \subset \mathbb{R}^n$. Likewise $\partial\bar{B}(x,t) \cap \{y_{n+1} \leq 0\}$ is the graph of $-\gamma$. Thus (36) implies

$$(37)\qquad ⨍_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} = \frac{2}{(n+1)\alpha(n+1)t^n} \int_{B(x,t)} g(y)(1+|D\gamma(y)|^2)^{1/2} dy,$$

the factor "2" entering because $\partial\bar{B}(\bar{x},t)$ comprises two hemispheres. Note that $(1+|D\gamma(y)|^2)^{1/2} = t(t^2 - |y-x|^2)^{-1/2}$. Our substituting this into (37) yields

$$\begin{aligned} ⨍_{\partial \bar{B}(\bar{x},t)} \bar{g}\, d\bar{S} &= \frac{2}{(n+1)\alpha(n+1)t^{n-1}} \int_{B(x,t)} \frac{g(y)}{(t^2-|y-x|^2)^{1/2}}\, dy \\ &= \frac{2t\alpha(n)}{(n+1)\alpha(n+1)} ⨍_{B(x,t)} \frac{g(y)}{(t^2-|y-x|^2)^{1/2}} dy. \end{aligned}$$

We insert this formula and the similar one with h in place of g into (37), and find

$$\begin{aligned} u(x,t) = \\ \frac{1}{\gamma_{n+1}} \frac{2\alpha(n)}{(n+1)\alpha(n+1)} \Bigg[& \frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-2}{2}} \left(t^n ⨍_{B(x,t)} \frac{g(y)}{(t^2-|y-x|^2)^{1/2}}\, dy \right) \\ & + \left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-2}{2}} \left(t^n ⨍_{B(x,t)} \frac{h(y)}{(t^2-|y-x|^2)^{1/2}}\, dy \right) \Bigg]. \end{aligned}$$

Since $\gamma_{n+1} = 1 \cdot 3 \cdot 5 \cdots (n-1)$ and $\alpha(n) = \frac{\pi^{n/2}}{\Gamma(\frac{n+2}{2})}$, we may compute $\gamma_n = 2 \cdot 4 \cdots (n-2) \cdot n$.

Hence the resulting representation formula for even n is:

$$
(38)\quad \begin{cases} u(x,t) = \dfrac{1}{\gamma_n}\left[\left(\dfrac{\partial}{\partial t}\right)\left(\dfrac{1}{t}\dfrac{\partial}{\partial t}\right)^{\frac{n-2}{2}}\left(t^n \displaystyle⨍_{B(x,t)} \dfrac{g(y)}{(t^2 - |y-x|^2)^{1/2}}\,dy\right)\right. \\ \qquad\qquad \left. + \left(\dfrac{1}{t}\dfrac{\partial}{\partial t}\right)^{\frac{n-2}{2}}\left(t^n \displaystyle⨍_{B(x,t)} \dfrac{h(y)}{(t^2 - |y-x|^2)^{1/2}}\,dy\right)\right], \\ \text{where } n \text{ is } \textit{even} \text{ and } \gamma_n = 2 \cdot 4 \cdots (n-2) \cdot n, \end{cases}
$$

for $x \in \mathbb{R}^n,\ t > 0$.

Since $\gamma_2 = 2$, this agrees with Poisson's formula (27) if $n = 2$.

THEOREM 3 (Solution of wave equation in even dimensions). *Assume n is an even integer, $n \geq 2$, and suppose also $g \in C^{m+1}(\mathbb{R}^n)$, $h \in C^m(\mathbb{R}^n)$, for $m = \frac{n+2}{2}$. Define u by (38). Then*

(i) $u \in C^2(\mathbb{R}^n \times [0,\infty))$,

(ii) $u_{tt} - \Delta u = 0$ *in* $\mathbb{R}^n \times (0,\infty)$,

and

(iii) $\displaystyle\lim_{\substack{(x,t)\to(x^0,0) \\ x\in\mathbb{R}^n,\ t>0}} u(x,t) = g(x^0), \quad \lim_{\substack{(x,t)\to(x^0,0) \\ x\in\mathbb{R}^n,\ t>0}} u_t(x,t) = h(x^0)$

for each point $x^0 \in \mathbb{R}^n$.

This follows from Theorem 2.

Remarks. (i) Observe, in contrast to formula (31), that to compute $u(x,t)$ for even n we need information on $u = g$, $u_t = h$ on all of $B(x,t)$, and not just on $\partial B(x,t)$.

(ii) Comparing (31) and (38) we observe that if n is odd and $n \geq 3$, the data g and h at a given point $x \in \mathbb{R}^n$ affect the solution u only on the *boundary* $\{(y,t) \mid t > 0,\ |x-y| = t\}$ of the cone $C = \{(y,t) \mid t > 0,\ |x-y| < t\}$. On the other hand, if n is even the data g and h affect u within all of C. In other words, *a "disturbance" originating at x propagates along a sharp wavefront in odd dimensions, but in even dimensions continues to have effects even after the leading edge of the wavefront passes.* This is *Huygens' principle.* □

2.4.2. Nonhomogeneous problem.

We next investigate the initial-value problem for the nonhomogeneous wave equation

$$\text{(39)}\qquad \begin{cases} u_{tt} - \Delta u = f & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = 0,\ u_t = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Motivated by Duhamel's principle (introduced earlier in §2.3.1), we define $u = u(x,t;s)$ to be the solution of

$$(40_s)\qquad \begin{cases} u_{tt}(\cdot;s) - \Delta u(\cdot;s) = 0 & \text{in } \mathbb{R}^n \times (s,\infty) \\ u(\cdot;s) = 0,\ u_t(\cdot;s) = f(\cdot,s) & \text{on } \mathbb{R}^n \times \{t = s\}. \end{cases}$$

Now set

$$\text{(41)}\qquad u(x,t) := \int_0^t u(x,t;s)\,ds \quad (x \in \mathbb{R}^n, t \geq 0).$$

Duhamel's principle asserts this is a solution of

$$\text{(42)}\qquad \begin{cases} u_{tt} - \Delta u = f & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = 0,\ u_t = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

THEOREM 4 (Solution of nonhomogeneous wave equation). *Assume* $n \geq 2$ *and* $f \in C^{[n/2]+1}(\mathbb{R}^n \times [0,\infty))$. *Define* u *by* (41). *Then*

(i) $u \in C^2(\mathbb{R}^n \times [0,\infty))$,

(ii) $u_{tt} - \Delta u = f$ *in* $\mathbb{R}^n \times (0,\infty)$,

and

(iii) $\lim_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n,\ t>0}} u(x,t) = 0, \quad \lim_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n,\ t>0}} u_t(x,t) = 0$ *for each point* $x^0 \in \mathbb{R}^n$.

Proof. 1. If n is odd, $\left[\frac{n}{2}\right] + 1 = \frac{n+1}{2}$. According to Theorem 2 $u(\cdot,\cdot;s) \in C^2(\mathbb{R}^n \times [0,\infty))$ for each $s \geq 0$, and so $u \in C^2(\mathbb{R}^n \times [0,\infty))$. If n is even, $\left[\frac{n}{2}\right] + 1 = \frac{n+2}{2}$. Hence $u \in C^2(\mathbb{R}^n \times [0,\infty))$, according to Theorem 3.

2. We then compute:

$$u_t(x,t) = u(x,t;t) + \int_0^t u_t(x,t;s)\,ds = \int_0^t u_t(x,t;s)\,ds,$$

$$\begin{aligned} u_{tt}(x,t) &= u_t(x,t;t) + \int_0^t u_{tt}(x,t;s)\,ds \\ &= f(x,t) + \int_0^t u_{tt}(x,t;s)\,ds. \end{aligned}$$

Furthermore

$$\Delta u(x,t) = \int_0^t \Delta u(x,t;s)\,ds = \int_0^t u_{tt}(x,t;s)\,ds.$$

Thus

$$u_{tt}(x,t) - \Delta u(x,t) = f(x,t) \quad (x \in \mathbb{R}^n, t > 0),$$

and clearly $u(x,0) = u_t(x,0) = 0$ for $x \in \mathbb{R}^n$. □

Remark. The solution of the general nonhomogeneous problem is consequently the sum of the solution of (11) (given by formulas (8), (31) or (38)) and the solution of (42) (given by (41)). □

Examples. (i) Let us work out explicitly how to solve (42) for $n = 1$. In this case d'Alembert's formula (8) gives

$$u(x,t;s) = \frac{1}{2}\int_{x-t+s}^{x+t-s} f(y,s)\,dy, \;\; u(x,t) = \frac{1}{2}\int_0^t \int_{x-t+s}^{x+t-s} f(y,s)\,dyds.$$

That is,

$$u(x,t) = \frac{1}{2}\int_0^t \int_{x-s}^{x+s} f(y,t-s)\,dyds \quad (x \in \mathbb{R},\; t \geq 0). \tag{43}$$

(ii) For $n = 3$, Kirchhoff's formula (22) implies

$$u(x,t;s) = (t-s) ⨍_{\partial B(x,t-s)} f(y,s)\,dS;$$

so that

$$\begin{aligned} u(x,t) &= \int_0^t (t-s)\left(⨍_{\partial B(x,t-s)} f(y,s)\,dS \right) ds \\ &= \frac{1}{4\pi}\int_0^t \int_{\partial B(x,t-s)} \frac{f(y,s)}{(t-s)}\,dSds \\ &= \frac{1}{4\pi}\int_0^t \int_{\partial B(x,r)} \frac{f(y,t-r)}{r}\,dSdr. \end{aligned}$$

Therefore

$$u(x,t) = \frac{1}{4\pi}\int_{B(x,t)} \frac{f(y,t-|y-x|)}{|y-x|}\,dy \quad (x \in \mathbb{R}^3,\; t \geq 0) \tag{44}$$

solves (42) for $n = 3$. The integrand on the right is called a *retarded potential*. □

2.4.3. Energy methods.

The explicit formulas (31) and (38) demonstrate the necessity of making more and more smoothness assumptions upon the data g and h to ensure the existence of a C^2 solution of the wave equation for larger and larger n. This suggests that perhaps some other way of measuring the size and smoothness of functions may be more appropriate. Indeed we will see in this section that the wave equation is nicely behaved (for all n) with respect to certain integral "energy" norms.

a. Uniqueness.

Let $U \subset \mathbb{R}^n$ be a bounded, open set with a smooth boundary ∂U, and as usual set $U_T = U \times (0, T]$, $\Gamma_T = \bar{U}_T - U_T$, where $T > 0$.

We are interested in the initial/boundary-value problem

$$
(45) \qquad \begin{cases} u_{tt} - \Delta u = f & \text{in } U_T \\ u = g & \text{on } \Gamma_T \\ u_t = h & \text{on } U \times \{t = 0\}. \end{cases}
$$

THEOREM 5 (Uniqueness for wave equation). *There exists at most one function $u \in C^2(\bar{U}_T)$ solving* (45).

Proof. If $\tilde{u}$ is another such solution, then $w := u - \tilde{u}$ solves

$$
\begin{cases} w_{tt} - \Delta w = 0 & \text{in } U_T \\ w = 0 & \text{on } \Gamma_T \\ w_t = 0 & \text{on } U \times \{t = 0\}. \end{cases}
$$

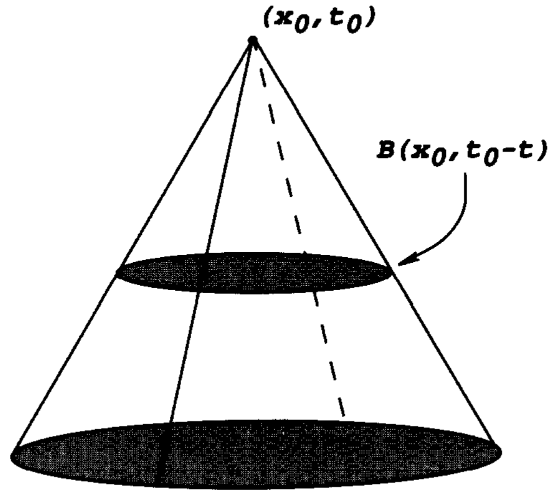
Define the "energy"

$$
e(t) := \frac{1}{2} \int_U w_t^2(x,t) + |Dw(x,t)|^2 dx \quad (0 \le t \le T).
$$

We compute

$$
\begin{aligned} \dot{e}(t) &= \int_U w_t w_{tt} + Dw \cdot Dw_t \, dx \qquad \left(\cdot = \frac{d}{dt} \right) \\ &= \int_U w_t (w_{tt} - \Delta w) dx = 0. \end{aligned}
$$

There is no boundary term since $w = 0$, and hence $w_t = 0$, on $\partial U \times [0, T]$. Thus for all $0 \le t \le T$, $e(t) = e(0) = 0$, and so $w_t, Dw \equiv 0$ within U_T. Since $w \equiv 0$ on $U \times \{t = 0\}$, we conclude $w = u - \tilde{u} \equiv 0$ in U_T. □



Cone of dependence

b. Domain of dependence.

As another illustration of energy methods, let us examine again the domain of dependence of solutions to the wave equation in all of space. For this, suppose $u \in C^2$ solves

$$u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty).$$

Fix $x_0 \in \mathbb{R}^n$, $t_0 > 0$ and consider the cone

$$C = \{(x,t) \mid 0 \leq t \leq t_0, |x - x_0| \leq t_0 - t\}.$$

THEOREM 6 (Finite propagation speed). *If $u \equiv u_t \equiv 0$ on $B(x_0, t_0)$, then $u \equiv 0$ within the cone C.*

In particular, we see that any "disturbance" originating outside $B(x_0, t_0)$ has no effect on the solution within C, and consequently has finite propagation speed. We already know this from the representation formulas (31) and (38), at least assuming $g = u$ and $h = u_t$ on $\mathbb{R}^n \times \{t = 0\}$ are sufficiently smooth. The point is that energy methods provide a *much* simpler proof.

Proof. Define

$$e(t) := \frac{1}{2} \int_{B(x_0, t_0 - t)} u_t^2(x,t) + |Du(x,t)|^2 dx \quad (0 \leq t \leq t_0).$$

Then

$$\begin{aligned}
\dot{e}(t) &= \int_{B(x_0,t_0-t)} u_t u_{tt} + Du \cdot Du_t \, dx - \frac{1}{2}\int_{\partial B(x_0,t_0-t)} u_t^2 + |Du|^2 dS \\
&= \int_{B(x_0,t_0-t)} u_t(u_{tt} - \Delta u)\, dx \\
&\quad + \int_{\partial B(x_0,t_0-t)} \frac{\partial u}{\partial \nu} u_t \, dS - \frac{1}{2}\int_{\partial B(x_0,t_0-t)} u_t^2 + |Du|^2 dS \\
&= \int_{\partial B(x_0,t_0-t)} \frac{\partial u}{\partial \nu} u_t - \frac{1}{2}u_t^2 - \frac{1}{2}|Du|^2 dS.
\end{aligned} \tag{46}$$

Now

$$\left|\frac{\partial u}{\partial \nu} u_t\right| \leq |u_t||Du| \leq \frac{1}{2}u_t^2 + \frac{1}{2}|Du|^2, \tag{47}$$

by the Cauchy–Schwarz and Cauchy inequalities (§B.2). Inserting (47) into (46), we find $\dot{e}(t) \leq 0$; and so $e(t) \leq e(0) = 0$ for all $0 \leq t \leq t_0$. Thus u_t, $Du \equiv 0$, and consequently $u \equiv 0$ within the cone C. □

A generalization of this proof to more complicated geometry appears later, in §7.2.4.

2.5. PROBLEMS

In the following exercises, all given functions are assumed smooth, unless otherwise stated.

1. Write down an explicit formula for a function u solving the initial-value problem

$$\begin{cases} u_t + b \cdot Du + cu = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

   Here $c \in \mathbb{R}$ and $b \in \mathbb{R}^n$ are constants.

2. Prove that Laplace's equation $\Delta u = 0$ is rotation invariant; that is, if O is an orthogonal $n \times n$ matrix and we define

$$v(x) := u(Ox) \quad (x \in \mathbb{R}^n),$$

   then $\Delta v = 0$.

3. Modify the proof of the mean value formulas to show for $n \geq 3$ that

$$u(0) = ⨍_{\partial B(0,r)} g\, dS + \frac{1}{n(n-2)\alpha(n)} \int_{B(0,r)} \left(\frac{1}{|x|^{n-2}} - \frac{1}{r^{n-2}}\right) f\, dx,$$

provided

$$\begin{cases} -\Delta u = f & \text{in } B^0(0,r) \\ u = g & \text{on } \partial B(0,r). \end{cases}$$

4. We say $v \in C^2(\bar{U})$ is *subharmonic* if

$$-\Delta v \leq 0 \quad \text{in } U.$$

(a) Prove for subharmonic v that

$$v(x) \leq ⨍_{B(x,r)} v\, dy \quad \text{for all } B(x,r) \subset U.$$

(b) Prove that therefore $\max_{\bar{U}} v = \max_{\partial U} v$.

(c) Let $\phi : \mathbb{R} \to \mathbb{R}$ be smooth and convex. Assume u is harmonic and $v := \phi(u)$. Prove v is subharmonic.

(d) Prove $v := |Du|^2$ is subharmonic, whenever u is harmonic.

5. Prove that there exists a constant C, depending only on n, such that

$$\max_{B(0,1)} |u| \leq C(\max_{\partial B(0,1)} |g| + \max_{B(0,1)} |f|)$$

whenever u is a smooth solution of

$$\begin{cases} -\Delta u = f & \text{in } B^0(0,1) \\ u = g & \text{on } \partial B(0,1). \end{cases}$$

6. Use Poisson's formula for the ball to prove

$$r^{n-2}\frac{r-|x|}{(r+|x|)^{n-1}}u(0) \leq u(x) \leq r^{n-2}\frac{r+|x|}{(r-|x|)^{n-1}}u(0)$$

whenever u is positive and harmonic in $B^0(0,r)$. This is an explicit form of Harnack's inequality.

7. Prove Theorem 15 in §2.2.4. (Hint: Since $u \equiv 1$ solves (44) for $g \equiv 1$, the theory automatically implies

$$\int_{\partial B(0,1)} K(x,y)\, dS(y) = 1$$

for each $x \in B^0(0,1)$.)

8. Let u be the solution of

$$\begin{cases} \Delta u = 0 & \text{in } \mathbb{R}^n_+ \\ u = g & \text{on } \partial\mathbb{R}^n_+ \end{cases}$$

given by Poisson's formula for the half-space. Assume g is bounded and $g(x) = |x|$ for $x \in \partial\mathbb{R}^n_+$, $|x| \leq 1$. Show Du is *not* bounded near $x = 0$. (Hint: Estimate $\frac{u(\lambda e_n)-u(0)}{\lambda}$.)

9. Let U^+ denote the open half-ball $\{x \in \mathbb{R}^n \mid |x| < 1,\ x_n > 0\}$. Assume $u \in C(\bar{U}^+)$ is harmonic in U^+, with $u = 0$ on $\partial U^+ \cap \{x_n = 0\}$. Set

$$v(x) := \begin{cases} u(x) & \text{if } x_n \geq 0 \\ -u(x_1, \dots, x_{n-1}, -x_n) & \text{if } x_n < 0 \end{cases}$$

for $x \in U = B^0(0,1)$. Prove v is harmonic in U.

10. Suppose u is smooth and solves $u_t - \Delta u = 0$ in $\mathbb{R}^n \times (0,\infty)$.
    (i) Show $u_\lambda(x,t) := u(\lambda x, \lambda^2 t)$ also solves the heat equation for each $\lambda \in \mathbb{R}$.
    (ii) Use (i) to show $v(x,t) := x \cdot Du(x,t) + 2tu_t(x,t)$ solves the heat equation as well.

11. Assume $n = 1$ and $u(x,t) = v\left(\frac{x^2}{t}\right)$.
    (a) Show

$$u_t = u_{xx}$$

if and only if

$$(*) \qquad 4zv''(z) + (2+z)v'(z) = 0 \quad (z > 0).$$

    (b) Show that the general solution of $(*)$ is

$$v(z) = c\int_0^z e^{-s/4}s^{-1/2}ds + d.$$

    (c) Differentiate $v\left(\frac{x^2}{t}\right)$ with respect to x and select the constant c properly, so as to obtain the fundamental solution Φ for $n = 1$.

12. Write down an explicit formula for a solution of

$$\begin{cases} u_t - \Delta u + cu = f & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases}$$

where $c \in \mathbb{R}$.

13. Given $g : [0,\infty) \to \mathbb{R}$, with $g(0) = 0$, derive the formula

$$u(x,t) = \frac{x}{\sqrt{4\pi}}\int_0^t \frac{1}{(t-s)^{3/2}} e^{\frac{-x^2}{4(t-s)}} g(s)\,ds$$

for a solution of the initial/boundary-value problem

$$\begin{cases} u_t - u_{xx} = 0 & \text{in } \mathbb{R}_+ \times (0,\infty) \\ u = 0 & \text{on } \mathbb{R}_+ \times \{t=0\}, \\ u = g & \text{on } \{x=0\} \times [0,\infty). \end{cases}$$

(Hint: Let $v(x,t) := u(x,t) - g(t)$ and extend v to $\{x < 0\}$ by odd reflection.)

14. We say $v \in C_1^2(U_T)$ is a *subsolution* of the heat equation if

$$v_t - \Delta v \leq 0 \quad \text{in } U_T.$$

(a) Prove for a subsolution v that

$$v(x,t) \leq \frac{1}{4r^n} \iint_{E(x,t;r)} v(y,s) \frac{|x-y|^2}{(t-s)^2}\, dyds$$

for all $E(x,t;r) \subset U_T$.

(b) Prove that therefore $\max_{\bar{U}_T} v = \max_{\Gamma_T} v$.

(c) Let $\phi : \mathbb{R} \to \mathbb{R}$ be smooth and convex. Assume u solves the heat equation and $v := \phi(u)$. Prove v is a subsolution.

(d) Prove $v := |Du|^2 + u_t^2$ is a subsolution, whenever u solves the heat equation.

15. (a) Show the general solution of the PDE $u_{xy} = 0$ is

$$u(x,y) = F(x) + G(y)$$

for arbitrary functions F, G.

(b) Using the change of variables $\xi = x + t, \quad \eta = x - t$, show $u_{tt} - u_{xx} = 0$ if and only if $u_{\xi\eta} = 0$.

(c) Use (a) and (b) to rederive d'Alembert's formula.

16. Assume $\mathbf{E} = (E^1, E^2, E^3)$ and $\mathbf{B} = (B^1, B^2, B^3)$ solve Maxwell's equations (§1.2.2). Show

$$u_{tt} - \Delta u = 0$$

where $u = E^i$ or B^i $(i = 1,2,3)$.

17. (Equipartition of energy). Let $u \in C^2(\mathbb{R} \times [0,\infty))$ solve the initial-value problem for the wave equation in one dimension:

$$\begin{cases} u_{tt} - u_{xx} = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g, u_t = h & \text{on } \mathbb{R} \times \{t=0\}. \end{cases}$$

Suppose g, h have compact support. The *kinetic energy* is $k(t) := \frac{1}{2}\int_{-\infty}^{\infty} u_t^2(x,t)\,dx$ and the *potential energy* is $p(t) := \frac{1}{2}\int_{-\infty}^{\infty} u_x^2(x,t)\,dx$. Prove

(i) $k(t) + p(t)$ is constant in t,

(ii) $k(t) = p(t)$ for all large enough times t.

18. Let u solve

$$\begin{cases} u_{tt} - \Delta u = 0 & \text{in } \mathbb{R}^3 \times (0,\infty) \\ u = g, u_t = h & \text{on } \mathbb{R}^3 \times \{t = 0\}, \end{cases}$$

where g, h are smooth and have compact support. Show there exists a constant C such that

$$|u(x,t)| \leq C/t \quad (x \in \mathbb{R}^3,\ t > 0).$$

2.6. REFERENCES

Section 2.2 A good source for more on Laplace's and Poisson's equations is Gilbarg–Trudinger [**G-T,** Chapters 2-4]. The proof of analyticity is from Mikhailov [**M**]. J. Cooper helped me with Green's functions.

Section 2.3 See John [**J,** Chapter 7] or Friedman [**FR2**] for further information concerning the heat equation. Theorem 3 is due to N. Watson [**W**], as is the proof of Theorem 4. Theorem 6 is taken from John [**J**], and Theorem 8 follows Mikhailov [**M**]. Theorem 11 is from Payne [**PA,** §2.3].

Section 2.4 See Antman [**A**] for a careful derivation of the one-dimensional wave equation as a model for a vibrating string. The solution of the wave equation presented here follows Folland [**F1**], Strauss [**ST**].

Section 2.5 J. Goldstein suggested Problem 17.

Chapter 3

NONLINEAR FIRST-ORDER PDE

In this chapter we investigate general nonlinear first-order partial differential equations of the form

$$F(Du, u, x) = 0,$$

where $x \in U$ and U is an open subset of $\mathbb{R}^n$. Here

$$F : \mathbb{R}^n \times \mathbb{R} \times \bar{U} \to \mathbb{R}$$

is given, and $u : \bar{U} \to \mathbb{R}$ is the unknown, $u = u(x)$.

Notation. Let us write

$$F = F(p, z, x) = F(p_1, \ldots, p_n, z, x_1, \ldots, x_n)$$

for $p \in \mathbb{R}^n$, $z \in \mathbb{R}$, $x \in U$. Thus "p" is the name of the variable for which we substitute the gradient $Du(x)$, and "z" is the variable for which we substitute $u(x)$. We also assume hereafter that F is smooth, and set

$$\begin{cases} D_pF = (F_{p_1}, \ldots, F_{p_n}) \\ D_zF = F_z \\ D_xF = (F_{x_1}, \ldots, F_{x_n}). \end{cases}$$

□

We are concerned with discovering solutions u of the PDE $F(Du, u, x) = 0$ in U, usually subject to the boundary condition

$$u = g \quad \text{on } \Gamma,$$

where Γ is some given subset of ∂U and $g : \Gamma \to \mathbb{R}$ is prescribed.

Nonlinear first-order partial differential equations arise in a variety of physical theories, primarily in dynamics (to generate canonical transformations), continuum mechanics (to record conservation of mass, momentum, energy, etc.) and optics (to describe wavefronts). Although the strong nonlinearity generally precludes our deriving any simple formulas for solutions, we can, remarkably, often employ calculus to glean fairly detailed information about solutions. Such techniques, discussed in §§3.1 and 3.2, are typically only local. In §§3.3 and 3.4 we will for the important cases of Hamilton–Jacobi equations and conservation laws derive certain global representation formulas for appropriately defined weak solutions.

3.1. COMPLETE INTEGRALS, ENVELOPES

3.1.1. Complete integrals.

We begin our analysis of the nonlinear first-order PDE

$$F(Du, u, x) = 0 \tag{1}$$

by describing some simple classes of solutions and then learning how to build from them more complicated solutions.

Suppose first $A \subset \mathbb{R}^n$ is an open set. Assume for each parameter $a = (a_1, \dots, a_n) \in A$ we have a C^2 solution $u = u(x; a)$ of the PDE (1).

Notation. We write

$$(D_a u, D^2_{xa} u) := \begin{pmatrix} u_{a_1} & u_{x_1 a_1} & \cdots & u_{x_n a_1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{a_n} & u_{x_1 a_n} & \cdots & u_{x_n a_n} \end{pmatrix}_{n \times (n+1)}. \tag{2}$$

□

DEFINITION. *A C^2 function $u = u(x; a)$ is called a* complete integral *in $U \times A$ provided*

$$u(x; a) \textit{ solves the PDE (1) for each } a \in A \tag{i}$$

and

$$\operatorname{rank}(D_a u, D^2_{xa} u) = n \quad (x \in U,\ a \in A). \tag{ii}$$

Remark. Condition (ii) ensures $u(x;a)$ "depends on all the n independent parameters $a_1, \dots, a_n$". To see this, suppose $B \subset \mathbb{R}^{n-1}$ is open, and for each $b \in B$ assume $v = v(x;b)$ $(x \in U)$ is a solution of (1). Suppose also there exists a C^1 mapping $\psi : A \to B$, $\psi = (\psi^1, \dots, \psi^{n-1})$, such that

$$u(x;a) = v(x;\psi(a)) \quad (x \in U,\ a \in A). \tag{3}$$

That is, we are supposing the function $u(x;a)$ "really depends only on the $n-1$ parameters $b_1, \dots, b_{n-1}$". But then

$$u_{x_i a_j}(x;a) = \sum_{k=1}^{n-1} v_{x_i b_k}(x;\psi(a))\psi^k_{a_j}(a) \quad (i,j = 1, \dots, n).$$

Consequently

$$\det(D^2_{xa}u) = \sum_{k_1,\dots,k_n=1}^{n-1} v_{x_1 b_{k_1}} \cdots v_{x_n b_{k_n}} \det \begin{pmatrix} \psi^{k_1}_{a_1} & \dots & \psi^{k_1}_{a_n} \\ & \ddots & \\ \psi^{k_n}_{a_1} & \dots & \psi^{k_n}_{a_n} \end{pmatrix} = 0,$$

since for each choice of $k_1, \dots, k_n \in \{1, \dots, n-1\}$, at least two columns in the corresponding matrix are equal. As

$$u_{a_j}(x;a) = \sum_{k=1}^{n-1} v_{b_k}(x;\psi(a))\psi^k_{a_j}(a) \quad (j = 1, \dots, n),$$

a similar argument shows the determinant of each $n \times n$ submatrix of $(D_a u, D^2_{xa}u)$ equals zero, and thus this matrix has rank strictly less than n. □

Example 1. *Clairaut's equation* from differential geometry is the PDE

$$x \cdot Du + f(Du) = u, \tag{4}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is given. A complete integral is

$$u(x;a) = a \cdot x + f(a) \quad (x \in U) \tag{5}$$

for $a \in \mathbb{R}^n$. □

Example 2. The *eikonal* equation* from geometric optics is the PDE

$$|Du| = 1. \tag{6}$$

A complete integral is

$$u(x;a,b) = a \cdot x + b \quad (x \in U) \tag{7}$$

for $x \in U$, $a \in \partial B(0,1)$, $b \in \mathbb{R}$. □

*εικών = image (Greek)

Example 3. The *Hamilton–Jacobi equation* from mechanics is in its simplest form the partial differential equation

$$u_t + H(Du) = 0, \tag{8}$$

where $H : \mathbb{R}^n \to \mathbb{R}$. Here u depends on $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $t \in \mathbb{R}$. As before we have set $t = x_{n+1}$ and written $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$. A complete integral is

$$u(x, t; a, b) = a \cdot x - tH(a) + b \quad (x \in \mathbb{R}^n,\ t \geq 0) \tag{9}$$

where $a \in \mathbb{R}^n$, $b \in \mathbb{R}$.

□

3.1.2. New solutions from envelopes.

We next demonstrate how to build more complicated solutions of our nonlinear first-order PDE (1), solutions which depend on an arbitrary *function* of $n-1$ variables, and not just on n parameters. We will construct these new solutions as envelopes of complete integrals or, more generally, of other m-parameter families of solutions.

DEFINITION. *Let $u = u(x; a)$ be a C^1 function of $x \in U$, $a \in A$, where $U \subset \mathbb{R}^n$ and $A \subset \mathbb{R}^m$ are open sets. Consider the vector equation*

$$D_a u(x; a) = 0 \quad (x \in U,\ a \in A). \tag{10}$$

Suppose that we can solve (10) *for the parameter a as a C^1 function of x,*

$$a = \phi(x); \tag{11}$$

thus

$$D_a u(x; \phi(x)) = 0 \quad (x \in U). \tag{12}$$

We then call

$$v(x) := u(x; \phi(x)) \quad (x \in U) \tag{13}$$

the envelope *of the functions $\{u(\cdot\,; a)\}_{a \in A}$.*

By forming envelopes we can build new solutions of our nonlinear first-order partial differential equation:

THEOREM 1 (Construction of new solutions). *Suppose for each $a \in A$ as above that $u = u(\cdot; a)$ solves the partial differential equation* (1). *Assume further that the envelope v, defined by* (12) *and* (13) *above, exists and is a C^1 function. Then v solves* (1) *as well.*

The envelope v defined above is sometimes called a *singular integral* of (1).

Proof. We have $v(x) = u(x; \phi(x))$; and so for $i = 1, \ldots, n$

$$\begin{aligned} v_{x_i}(x) &= u_{x_i}(x; \phi(x)) + \sum_{j=1}^{n} u_{a_j}(x, \phi(x))\phi^j_{x_i}(x) \\ &= u_{x_i}(x; \phi(x)), \text{ according to (12).} \end{aligned}$$

Hence for each $x \in U$,

$$F(Dv(x), v(x), x) = F(Du(x; \phi(x)), u(x; \phi(x)), x) = 0.$$

□

Remark. The geometric idea is that for each $x \in U$, the graph of v is tangent to the graph of $u(\cdot\,; a)$ for $a = \phi(x)$. Thus $Dv = D_x u(\cdot\,; a)$ at x, for $a = \phi(x)$. □

Example 4. Consider the PDE

$$u^2(1 + |Du|^2) = 1. \tag{14}$$

A complete integral is

$$u(x, a) = \pm(1 - |x - a|^2)^{1/2} \quad (|x - a| < 1).$$

We compute

$$D_a u = \frac{\mp(x - a)}{(1 - |x - a|^2)^{1/2}} = 0$$

provided $a = \phi(x) = x$. Thus $v \equiv \pm 1$ are singular integrals of (14). □

To generate still more solutions of the PDE (1) from a complete integral, we vary the above construction. Choose any open set $A' \subset \mathbb{R}^{n-1}$ and any C^1 function $h : A' \to \mathbb{R}$, so that the graph of h lies within A. Let us write

$$a = (a_1, \ldots, a_n) = (a', a_n) \quad \text{for } a' = (a_1, \ldots, a_{n-1}) \in \mathbb{R}^{n-1}.$$

DEFINITION. *The* general integral *(depending on h) is the envelope $v' = v'(x)$ of the functions*

$$u'(x;a') = u(x;a',h(a')) \quad (x \in U,\ a' \in A'),$$

provided this envelope exists and is C^1.

In other words, in computing the envelope we are now restricting only to parameters a of the form $a = (a', h(a'))$, for some explicit choice of the function h.

Remarks. (i) Thus from a complete integral, which depends upon n arbitrary constants $a_1, \dots, a_n$, we build (whenever the foregoing construction works) a solution depending on an arbitrary function h of $n-1$ variables.

(ii) It is tempting to believe that once we can find as above a solution of (1) depending on an arbitrary function h we have found all the solutions of (1). This need not be so, however. Suppose our PDE has the structure

$$F(Du,u,x) = F_1(Du,u,x)F_2(Du,u,x) = 0.$$

If $u_1(x,a)$ is a complete integral of the PDE $F_1(Du,u,x) = 0$, and we succeed in finding a general integral corresponding to any function h, we will still have missed all the solutions of the PDE $F_2(Du,u,x) = 0$. □

Example 5. An alternative form for a complete integral of the eikonal equation $|Du| = 1$ for $n = 2$ is

(15) $$u(x;a) = x_1 \cos a_1 + x_2 \sin a_1 + a_2 \quad (x, a \in \mathbb{R}^2).$$

We set $h \equiv 0$, so that

$$u'(x;a_1) = x_1 \cos a_1 + x_2 \sin a_1$$

represents the subfamily of planar solutions of $|Du| = 1$, whose graphs pass through the point $(0,0,0) \in \mathbb{R}^3$. We then compute the envelope by writing

$$D_{a_1}u' = -x_1 \sin a_1 + x_2 \cos a_1 = 0.$$

Thus $a_1 = \arctan \frac{x_2}{x_1}$, and consequently

$$u'(x) = x_1 \cos(\arctan \frac{x_2}{x_1}) + x_2 \sin(\arctan \frac{x_2}{x_1}) = \pm|x| \quad (x \in \mathbb{R}^2)$$

solves $|Du| = 1$ for $x \neq 0$. □

Example 6. Let $H(p) = |p|^2$, $h \equiv 0$ in Example 3 above. Then

$$u'(x,t;a) = x \cdot a - t|a|^2.$$

We calculate the envelope by setting $D_a u' = x - 2ta = 0$. Hence $a = \frac{x}{2t}$, and so

$$u'(x,t) = x \cdot \frac{x}{2t} - t\left|\frac{x}{2t}\right|^2 = \frac{|x|^2}{4t} \quad (x \in \mathbb{R}^n,\ t > 0)$$

solves the Hamilton–Jacobi equation $u_t + |Du|^2 = 0$. □

3.2. CHARACTERISTICS

3.2.1. Derivation of characteristic ODE.

We return to our basic nonlinear first-order PDE

$$F(Du, u, x) = 0 \quad \text{in } U, \tag{1}$$

subject now to the boundary condition

$$u = g \quad \text{on } \Gamma, \tag{2}$$

where $\Gamma \subseteq \partial U$ and $g : \Gamma \to \mathbb{R}$ are given. We hereafter suppose that F, g are smooth functions.

We develop next the method of *characteristics*, which solves (1), (2) by converting the PDE into an appropriate system of ODE. This is the plan. Suppose u solves (1), (2) and fix any point $x \in U$. We would like to calculate $u(x)$ by finding some curve lying within U, connecting x with a point $x^0 \in \Gamma$ and along which we can compute u. Since (2) says $u = g$ on Γ, we know the value of u at the one end x^0. We hope then to be able to calculate u all along the curve, and so in particular at x.

Finding the characteristic ODE. How can we choose the curve so all this will work? Let us suppose it is described parametrically by the function $\mathbf{x}(s) = (x^1(s), \dots, x^n(s))$, the parameter s lying in some subinterval of $\mathbb{R}$. Assuming u is a C^2 solution of (1), we define also

$$z(s) := u(\mathbf{x}(s)). \tag{3}$$

In addition, set

$$\mathbf{p}(s) := Du(\mathbf{x}(s)); \tag{4}$$

that is, $\mathbf{p}(s) = (p^1(s), \dots, p^n(s))$, where

$$p^i(s) = u_{x_i}(\mathbf{x}(s)) \quad (i = 1, \dots, n). \tag{5}$$

So $z(\cdot)$ gives the values of u along the curve and $\mathbf{p}(\cdot)$ records the values of the gradient Du. We must choose the function $\mathbf{x}(\cdot)$ in such a way that we can compute $z(\cdot)$ and $\mathbf{p}(\cdot)$.

For this, first differentiate (5):

$$\dot{p}^i(s) = \sum_{j=1}^{n} u_{x_i x_j}(\mathbf{x}(s))\dot{x}^j(s) \quad \left(\cdot = \frac{d}{ds}\right). \tag{6}$$

This expression is not too promising, since it involves the second derivatives of u. On the other hand, we can also differentiate the PDE (1) with respect to x_i:

$$(7) \quad \sum_{j=1}^{n} \frac{\partial F}{\partial p_j}(Du,u,x)u_{x_j x_i} + \frac{\partial F}{\partial z}(Du,u,x)u_{x_i} + \frac{\partial F}{\partial x_i}(Du,u,x) = 0.$$

We are able to employ this identity to get rid of the "dangerous" second derivative terms in (6), provided we first set

$$(8) \qquad \dot{x}^j(s) = \frac{\partial F}{\partial p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)) \quad (j = 1, \dots, n).$$

Assuming now (8) holds, we evaluate (7) at $x = \mathbf{x}(s)$, obtaining thereby from (3), (4) the identity:

$$\sum_{j=1}^{n} \frac{\partial F}{\partial p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s))u_{x_i x_j}(\mathbf{x}(s)) + \frac{\partial F}{\partial z}(\mathbf{p}(s), z(s), \mathbf{x}(s))p^i(s) + \frac{\partial F}{\partial x_i}(\mathbf{p}(s), z(s), \mathbf{x}(s)) = 0.$$

Substitute this expression and (8) into (6):

$$(9) \qquad \begin{aligned} \dot{p}^i(s) = &-\frac{\partial F}{\partial x_i}(\mathbf{p}(s), z(s), \mathbf{x}(s)) \\ &- \frac{\partial F}{\partial z}(\mathbf{p}(s), z(s), \mathbf{x}(s))p^i(s) \quad (i = 1, \dots, n). \end{aligned}$$

Finally we differentiate (3):

$$(10) \quad \dot{z}(s) = \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(\mathbf{x}(s))\dot{x}^j(s) = \sum_{j=1}^{n} p^j(s) \frac{\partial F}{\partial p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)),$$

the second equality holding by (5) and (8).

We summarize by rewriting equations (8)–(10) in vector notation:

$$(11) \quad \begin{cases} \text{(a) } \dot{\mathbf{p}}(s) = -D_x F(\mathbf{p}(s), z(s), \mathbf{x}(s)) - D_z F(\mathbf{p}(s), z(s), \mathbf{x}(s))\mathbf{p}(s) \\ \text{(b) } \dot{z}(s) = D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) \\ \text{(c) } \dot{\mathbf{x}}(s) = D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)). \end{cases}$$

This important system of $2n + 1$ first-order ODE comprises the *characteristic equations* of the nonlinear first-order PDE (1). The functions $\mathbf{p}(\cdot) = (p^1(\cdot), \dots, p^n(\cdot))$, $z(\cdot)$, $\mathbf{x}(\cdot) = (x^1(\cdot), \dots, x^n(\cdot))$ are called the *characteristics*. We will sometimes refer to $\mathbf{x}(\cdot)$ as the *projected characteristic*: it is the projection of the full characteristics $(\mathbf{p}(\cdot), z(\cdot), \mathbf{x}(\cdot)) \subset \mathbb{R}^{2n+1}$ onto the physical region $U \subset \mathbb{R}^n$.

We have proved:

THEOREM 1 (Structure of characteristic ODE). *Let $u \in C^2(U)$ solve the nonlinear, first-order partial differential equation* (1) *in U. Assume $\mathbf{x}(\cdot)$ solves the* ODE (11)(c), *where $\mathbf{p}(\cdot) = Du(\mathbf{x}(\cdot))$, $z(\cdot) = u(\mathbf{x}(\cdot))$. Then $\mathbf{p}(\cdot)$ solves the* ODE (11)(a) *and $z(\cdot)$ solves the* ODE (11)(b), *for those s such that $\mathbf{x}(s) \in U$.*

We still need to discover appropriate initial conditions for the system of ODE (11), in order that this theorem be useful. We accomplish this in §3.2.3 below.

Remark. The characteristic ODE are truly remarkable in that they form a closed system of equations for $\mathbf{x}(\cdot)$, $z(\cdot) = u(\mathbf{x}(\cdot))$, and $\mathbf{p}(\cdot) = Du(\mathbf{x}(\cdot))$, whenever u is a smooth solution of the general nonlinear PDE (1). The key step in the derivation is our setting $\dot{\mathbf{x}} = D_pF$, so that—as explained above—the terms involving second derivatives drop out. We thereby obtain *closure*, and in particular are not forced to introduce ODE for the second and higher derivatives of u. □

3.2.2. Examples.

Before continuing our investigation of the characteristic equations (11), we pause to consider some special cases for which the structure of these equations is especially simple. We illustrate as well how we can sometimes actually solve the characteristic ODE and thereby explicitly compute solutions of certain first-order PDE, subject to appropriate boundary conditions.

a. F linear.

Consider first the situation that our PDE (1) is linear and homogeneous, and thus has the form

$$F(Du, u, x) = \mathbf{b}(x) \cdot Du(x) + c(x)u(x) = 0 \quad (x \in U). \tag{12}$$

Then $F(p, z, x) = \mathbf{b}(x) \cdot p + c(x)z$, and so

$$D_pF = \mathbf{b}(x). \tag{13}$$

In this circumstance equation (11)(c) becomes

$$\mathbf{x}(s) = \mathbf{b}(\mathbf{x}(s)), \tag{14}$$

an ODE involving only the function $\mathbf{x}(\cdot)$. Furthermore equation (11)(b) becomes

$$\dot{z}(s) = \mathbf{b}(\mathbf{x}(s)) \cdot \mathbf{p}(s). \tag{15}$$

Since $\mathbf{p}(\cdot) = Du(\mathbf{x}(\cdot))$, equation (12) simplifies (15), yielding

$$\dot{z}(s) = -c(\mathbf{x}(s))z(s). \tag{16}$$

This ODE is linear in $z(\cdot)$, once we know the function $\mathbf{x}(\cdot)$ by solving (14). In summary,

$$\begin{cases} \text{(a)} & \dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s)) \\ \text{(b)} & \dot{z}(s) = -c(\mathbf{x}(s))z(s) \end{cases} \tag{17}$$

comprise the characteristic equations for the linear, first-order PDE (12). (We will see later that the equation for $\mathbf{p}(\cdot)$ is not needed.) □

Example 1. We demonstrate the utility of equations (17) by explicitly solving the problem

$$\begin{cases} x_1 u_{x_2} - x_2 u_{x_1} = u & \text{in } U \\ u = g & \text{on } \Gamma, \end{cases} \tag{18}$$

where U is the quadrant $\{x_1 > 0, x_2 > 0\}$ and $\Gamma = \{x_1 > 0, x_2 = 0\} \subseteq \partial U$. The PDE in (18) is of the form (12), for $\mathbf{b} = (-x_2, x_1)$ and $c = -1$. Thus the equations (17) read

$$\begin{cases} \dot{x}^1 = -x^2, \; \dot{x}^2 = x^1 \\ \dot{z} = z. \end{cases} \tag{19}$$

Accordingly we have

$$\begin{cases} x^1(s) = x^0 \cos s, \; x^2(s) = x^0 \sin s \\ z(s) = z^0 e^s = g(x^0)\, e^s, \end{cases}$$

where $x^0 \geq 0$, $0 \leq s \leq \frac{\pi}{2}$. Fix a point $(x_1, x_2) \in U$. We select $s > 0$, $x^0 > 0$ so that $(x_1, x_2) = (x^1(s), x^2(s)) = (x^0 \cos s, x^0 \sin s)$. That is, $x^0 = (x_1^2 + x_2^2)^{1/2}$, $s = \arctan\left(\frac{x_2}{x_1}\right)$. Therefore

$$\begin{aligned} u(x_1, x_2) &= u(x^1(s), x^2(s)) = z(s) = g(x^0)\, e^s \\ &= g((x_1^2 + x_2^2)^{1/2})\, e^{\arctan\left(\frac{x_2}{x_1}\right)}. \end{aligned}$$

□

b. F quasilinear.

The partial differential equation (1) is quasilinear should it have the form

$$F(Du,u,x) = \mathbf{b}(x,u(x)) \cdot Du(x) + c(x,u(x)) = 0. \tag{20}$$

In this circumstance $F(p,z,x) = \mathbf{b}(x,z) \cdot p + c(x,z)$; whence

$$D_pF = \mathbf{b}(x,z).$$

Hence equation (11)(c) reads

$$\dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s), z(s)),$$

and (11)(b) becomes

$$\begin{aligned} \dot{z}(s) &= \mathbf{b}(\mathbf{x}(s), z(s)) \cdot \mathbf{p}(s) \\ &= -c(\mathbf{x}(s), z(s)), \qquad \text{by (20).} \end{aligned}$$

Consequently

$$\begin{cases} \text{(a)} & \dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s), z(s)) \\ \text{(b)} & \dot{z}(s) = -c(\mathbf{x}(s), z(s)) \end{cases} \tag{21}$$

are the characteristic equations for the quasilinear first-order PDE (20). (Once again the equation for $\mathbf{p}(\cdot)$ is not needed.) □

Example 2. The characteristic ODE (21) are in general difficult to solve, and so we work out in this example the simpler case of a boundary-value problem for a semilinear PDE:

$$\begin{cases} u_{x_1} + u_{x_2} = u^2 & \text{in } U \\ \qquad\qquad u = g & \text{on } \Gamma. \end{cases} \tag{22}$$

Now U is the half-space $\{x_2 > 0\}$ and $\Gamma = \{x_2 = 0\} = \partial U$. Here $\mathbf{b} = (1,1)$ and $c = -z^2$. Then (21) becomes

$$\begin{cases} \dot{x}^1 = 1, \ \dot{x}^2 = 1 \\ \dot{z} = z^2. \end{cases}$$

Consequently

$$\begin{cases} x^1(s) = x^0 + s, \ x^2(s) = s \\ z(s) = \frac{z^0}{1-sz^0} = \frac{g(x^0)}{1-sg(x^0)}, \end{cases}$$

where $x^0 \in \mathbb{R}$, $s \geq 0$, provided the denominator is not zero.

Fix a point $(x_1, x_2) \in U$. We select $s > 0$ and $x^0 \in \mathbb{R}$ so that $(x_1, x_2) = (x^1(s), x^2(s)) = (x^0 + s, s)$; that is, $x^0 = x_1 - x_2$, $s = x_2$. Then

$$\begin{aligned} u(x_1, x_2) = u(x^1(s), x^2(s)) = z(s) &= \frac{g(x^0)}{1 - sg(x^0)} \\ &= \frac{g(x_1 - x_2)}{1 - x_2 g(x_1 - x_2)}. \end{aligned}$$

This solution of course makes sense only if $1 - x_2 g(x_1 - x_2) \neq 0$. □

c. F fully nonlinear.

In the general case, the full characteristic equations (11) must be integrated, if possible.

Example 3. Consider the fully nonlinear problem

$$(23) \qquad \begin{cases} u_{x_1} u_{x_2} = u & \text{in } U \\ u = x_2^2 & \text{on } \Gamma, \end{cases}$$

where $U = \{x_1 > 0\}$, $\Gamma = \{x_1 = 0\} = \partial U$. Here $F(p, z, x) = p_1 p_2 - z$, and hence the characteristic ODE (11) become

$$\begin{cases} \dot{p}^1 = p^1, \ \dot{p}^2 = p^2 \\ \dot{z} = 2p^1 p^2 \\ \dot{x}^1 = p^2, \ \dot{x}^2 = p^1. \end{cases}$$

We integrate these equations to find

$$\begin{cases} x^1(s) = p_2^0(e^s - 1), \ x^2(s) = x^0 + p_1^0(e^s - 1) \\ z(s) = z^0 + p_1^0 p_2^0 (e^{2s} - 1) \\ p^1(s) = p_1^0 e^s \, p^2(s) = p_2^0 e^s, \end{cases}$$

where $x^0 \in \mathbb{R}$, $s \in \mathbb{R}$, and $z^0 = (x^0)^2$.

We must determine $p^0 = (p_1^0, p_2^0)$. Since $u = x_2^2$ on Γ, $p_2^0 = u_{x_2}(0, x^0) = 2x^0$. Furthermore the PDE $u_{x_1} u_{x_2} = u$ itself implies $p_1^0 p_2^0 = z^0 = (x^0)^2$, and so $p_1^0 = \frac{x^0}{2}$. Consequently the formulas above become

$$\begin{cases} x^1(s) = 2x^0(e^s - 1), x^2(s) = \frac{x^0}{2}(e^s + 1) \\ z(s) = (x^0)^2 e^{2s} \\ p^1(s) = \frac{x^0}{2} e^s, \ p^2(s) = 2x^0 e^s. \end{cases}$$

Fix a point $(x_1, x_2) \in U$. Select s and x^0 so that $(x_1, x_2) = (x^1(s), x^2(s)) = (2x^0(e^s - 1), \frac{x^0}{2}(e^s + 1))$. This equality implies $x^0 = \frac{4x_2 - x_1}{4}$, $e^s = \frac{x_1 + 4x_2}{4x_2 - x_1}$; and so

$$\begin{aligned} u(x_1, x_2) = u(x^1(s), x^2(s)) = z(s) &= (x^0)^2 e^{2s} \\ &= \frac{(x_1 + 4x_2)^2}{16}. \end{aligned}$$

□

3.2.3. Boundary conditions.

We return now to developing the general theory.

a. Straightening the boundary.

We intend in the section following to invoke the characteristic ODE (11) actually to solve the boundary-value problem (1), (2), at least in a small region near an appropriate portion Γ of ∂U. In order to simplify the relevant calculations, it is convenient first to change variables, so as to "flatten out" part of the boundary ∂U. To accomplish this, we first fix any point $x^0 \in \partial U$. Then utilizing the notation from §C.1, we find smooth mappings $\mathbf{\Phi}, \mathbf{\Psi} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\mathbf{\Psi} = \mathbf{\Phi}^{-1}$ and $\mathbf{\Phi}$ straightens out ∂U near x^0. (See the illustration in §C.1.)

Given any function $u : U \to \mathbb{R}$, let us write $V := \mathbf{\Phi}(U)$ and set

$$v(y) := u(\mathbf{\Psi}(y)) \quad (y \in V). \tag{24}$$

Then

$$u(x) = v(\mathbf{\Phi}(x)) \quad (x \in U). \tag{25}$$

Now suppose that u is a C^1 solution of our boundary-value problem (1), (2) in U. What PDE does v then satisfy in V?

According to (25), we see

$$u_{x_i}(x) = \sum_{k=1}^{n} v_{y_k}(\mathbf{\Phi}(x))\Phi^k_{x_i}(x) \quad (i = 1, \dots, n);$$

that is,

$$Du(x) = Dv(y)D\mathbf{\Phi}(x).$$

Thus (1) implies

$$\begin{aligned} 0 &= F(Du(x), u(x), x) \\ &= F(Dv(y)D\mathbf{\Phi}(\mathbf{\Psi}(y)), v(y), \mathbf{\Psi}(y)). \end{aligned} \tag{26}$$

This is an expression having the form

$$G(Dv(y), v(y), y) = 0 \quad \text{in } V.$$

In addition $v = h$ on Δ, where $\Delta := \mathbf{\Phi}(\Gamma)$ and $h(y) := g(\mathbf{\Psi}(y))$.

In summary, our problem (1), (2) transforms to read

$$\begin{cases} G(Dv, v, y) = 0 & \text{in } V \\ \qquad\quad v = h & \text{on } \Delta, \end{cases} \tag{27}$$

for G, h as above. The point is that if we change variables to straighten out the boundary near x^0, the boundary-value problem (1), (2) converts into a problem having the same form.

b. Compatibility conditions on boundary data.

In view of the foregoing computations, if we are given a point $x^0 \in \Gamma$ we may as well assume from the outset that Γ is flat near x^0, lying in the plane $\{x_n = 0\}$.

We intend now to utilize the characteristic ODE to construct a solution (1), (2), at least near x^0, and for this we must discover appropriate initial conditions

$$\mathbf{p}(0) = p^0, \; z(0) = z^0, \; \mathbf{x}(0) = x^0. \tag{28}$$

Now clearly if the curve $\mathbf{x}(\cdot)$ passes through x^0, we should insist that

$$z^0 = g(x^0). \tag{29}$$

What should we require concerning $\mathbf{p}(0) = p^0$? Since (2) implies $u(x_1 \dots x_{n-1}, 0) = g(x_1 \dots x_{n-1})$ near x^0, we may differentiate to find

$$u_{x_i}(x^0) = g_{x_i}(x^0) \quad (i = 1, \dots, n-1).$$

As we also want the PDE (1) to hold, we should therefore insist $p^0 = (p_1^0, \dots, p_n^0)$ satisfies these relations:

$$\begin{cases} p_i^0 = g_{x_i}(x^0) & (i = 1, \dots, n-1) \\ F(p^0, z^0, x^0) = 0. \end{cases} \tag{30}$$

These identities provide n equations for the n quantities $p^0 = (p_1^0, \dots, p_n^0)$.

We call (29) and (30) the *compatibility conditions.* A triple $(p^0, z^0, x^0) \in \mathbb{R}^{2n+1}$ verifying (29), (30) is *admissible.* Note z^0 is uniquely determined by the boundary condition and our choice of the point x^0, but a vector p^0 satisfying (30) may not exist or may not be unique.

c. Noncharacteristic boundary data.

So now assume as above that $x^0 \in \Gamma$, that Γ near x^0 lies in the plane $\{x_n = 0\}$, and that the triple (p^0, z^0, x^0) is admissible. We are planning to construct a solution u of (1), (2) in U near x^0 by integrating the characteristic ODE (11). So far we have ascertained $\mathbf{x}(0) = x^0$, $z(0) = z^0$, $\mathbf{p}(0) = p^0$ are appropriate boundary conditions for the characteristic ODE, with $\mathbf{x}(\cdot)$ intersecting Γ at x^0. But we will need in fact to solve these ODE for *nearby* initial points as well, and must consequently now ask if we can somehow appropriately perturb (p^0, z^0, x^0), keeping the compatibility conditions.

In other words, given a point $y = (y_1, \ldots, y_{n-1}, 0) \in \Gamma$, with y close to x^0, we intend to solve the characteristic ODE

$$(31)\quad \begin{cases} \text{(a) } \dot{\mathbf{p}}(s) = -D_xF(\mathbf{p}(s), z(s), \mathbf{x}(s)) - D_zF(\mathbf{p}(s), z(s), \mathbf{x}(s))\mathbf{p}(s) \\ \text{(b) } \dot{z}(s) = D_pF(\mathbf{p}(s), z(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) \\ \text{(c) } \dot{\mathbf{x}}(s) = D_pF(\mathbf{p}(s), z(s), \mathbf{x}(s)), \end{cases}$$

with the initial conditions

$$(32)\quad \mathbf{p}(0) = \mathbf{q}(y),\ z(0) = g(y),\ \mathbf{x}(0) = y.$$

Our task then is to find a function $\mathbf{q}(\cdot) = (q^1(\cdot), \ldots, q^n(\cdot))$, so that

$$(33)\quad \mathbf{q}(x^0) = p^0$$

and $(\mathbf{q}(y), g(y), y)$ is admissible; that is, the compatibility conditions

$$(34)\quad \begin{cases} q^i(y) = g_{x_i}(y) & (i = 1, \ldots, n-1) \\ F(\mathbf{q}(y), g(y), y) = 0 \end{cases}$$

hold for all $y \in \Gamma$ close to x^0.

LEMMA 1 (Noncharacteristic boundary conditions). *There exists a unique solution* $\mathbf{q}(\cdot)$ *of* (33), (34) *for all* $y \in \Gamma$ *sufficiently close to* x^0, *provided*

$$(35)\quad F_{p_n}(p^0, z^0, x^0) \neq 0.$$

We say the admissible triple (p^0, z^0, x^0) is *noncharacteristic* if (35) holds. We henceforth assume this condition.

Proof. To simplify notation, let us now temporarily write $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. We apply the Implicit Function Theorem (§C.6) to the mapping

$$\mathbf{G} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n,\ \mathbf{G}(p, y) = (G^1(p, y), \ldots, G^n(p, y)),$$

where

$$\begin{cases} G^i(p, y) = p_i - g_{x_i}(y) & (i = 1, \ldots, n-1) \\ G^n(p, y) = F(p, g(y), y). \end{cases}$$

Now $\mathbf{G}(p^0, x^0) = 0$, according to (29), (30). Also

$$D_p\mathbf{G}(p^0, x^0) = \begin{pmatrix} 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \\ F_{p_1}(p^0, z^0, x^0) & \ldots & & F_{p_n}(p^0, z^0, x^0) \end{pmatrix}_{n \times n},$$

and thus

$$\det D_p\mathbf{G}(p^0, x^0) = F_{p_n}(p^0, z^0, x^0) \neq 0,$$

in view of the noncharacteristic condition (35). The Implicit Function Theorem thus ensures we can uniquely solve the identity $G(p, y) = 0$ for $p = \mathbf{q}(y)$, provided y is close enough to x^0. □

Remark. If Γ is not flat near x^0, the condition that Γ be noncharacteristic reads

$$D_pF(p^0, z^0, x^0) \cdot \boldsymbol{\nu}(x_0) \neq 0, \tag{36}$$

$\boldsymbol{\nu}(x^0)$ denoting the outward unit normal to ∂U at x^0. □

3.2.4. Local solution.

Remember that our aim is to use the characteristic ODE to build a solution u of (1), (2), at least near Γ. So as before we select a point $x^0 \in \Gamma$ and, as shown in §3.2.3, may as well assume that near x^0 the surface Γ is flat, lying in the plane $\{x_n = 0\}$. Suppose further that (p^0, z^0, x^0) is an admissible triple of boundary data, which is noncharacteristic. According to Lemma 1 there is a function $\mathbf{q}(\cdot)$ so that $p^0 = \mathbf{q}(x^0)$ and the triple $(\mathbf{q}(y), g(y), y)$ is admissible, for all y sufficiently close to x^0.

Given any such point $y = (y_1, \dots, y_{n-1}, 0)$, we solve the characteristic ODE (31), subject to initial conditions (32).

Notation. Let us write

$$\begin{cases} \mathbf{p}(s) = \mathbf{p}(y,s) = \mathbf{p}(y_1, \dots, y_{n-1}, s) \\ z(s) = z(y,s) = z(y_1, \dots, y_{n-1}, s) \\ \mathbf{x}(s) = \mathbf{x}(y,s) = \mathbf{x}(y_1, \dots, y_{n-1}, s) \end{cases}$$

to display the dependence of the solution of (31), (32) on s and y. □

LEMMA 2 (Local invertibility). *Assume we have the noncharacteristic condition $F_{p_n}(p^0, z^0, x^0) \neq 0$. Then there exist an open interval $I \subset \mathbb{R}$ containing 0, a neighborhood W of x^0 in $\Gamma \subset \mathbb{R}^{n-1}$, and a neighborhood V of x^0 in $\mathbb{R}^n$, such that for each $x \in V$ there exist unique $s \in I$, $y \in W$ such that*

$$x = \mathbf{x}(y,s).$$

The mappings $x \mapsto s, y$ are C^2.

Proof. We have $\mathbf{x}(x^0, 0) = x^0$. Consequently the Inverse Function Theorem (§C.5) gives the result, provided $\det D\mathbf{x}(x^0, 0) \neq 0$. Now

$$\mathbf{x}(y, 0) = (y, 0) \quad (y \in \Gamma),$$

and so if $i = 1, \dots, n-1$,

$$\frac{\partial x^j}{\partial y_i}(x^0, 0) = \begin{cases} \delta_{ij} & (j = 1, \dots n-1) \\ 0 & (j = n). \end{cases}$$

Furthermore equation (31)(c) implies

$$\frac{\partial x^j}{\partial s}(x^0,0) = F_{p_j}(p^0,z^0,x^0).$$

Thus

$$Dx(x^0,0) = \begin{pmatrix} 1 & & 0 & F_{p_1}(p^0,z^0,x^0) \\ & \ddots & & \vdots \\ 0 & & 1 & \vdots \\ 0 & \cdots & 0 & F_{p_n}(p^0,z^0,x^0) \end{pmatrix}_{n\times n};$$

whence $\det D\mathbf{x}(x^0,0) \neq 0$ follows from the noncharacteristic condition (35). □

In view of Lemma 2 for each $x \in V$, we can locally uniquely solve the equation

$$\text{(37)} \qquad \begin{cases} x = \mathbf{x}(y,s), \\ \text{for } \; y = \mathbf{y}(x), \;\; s = s(x). \end{cases}$$

Finally, let us define

$$\text{(38)} \qquad \begin{cases} u(x) := z(\mathbf{y}(x), s(x)) \\ \mathbf{p}(x) := \mathbf{p}(\mathbf{y}(x), s(x)) \end{cases}$$

for $x \in V$ and s, y as in (37).

We come finally to our principal assertion, namely, that we can locally weave together the solutions of the characteristic ODE into a solution of the PDE.

THEOREM 2 (Local Existence Theorem). *The function u defined above is C^2 and solves the* PDE

$$F(Du(x), u(x), x) = 0 \quad (x \in V),$$

with the boundary condition

$$u(x) = g(x) \quad (x \in \Gamma \cap V).$$

Proof. 1. First of all, fix $y \in \Gamma$ close to x^0 and, as above, solve the characteristic ODE (31), (32) for $\mathbf{p}(s) = \mathbf{p}(y,s)$, $z(s) = z(y,s)$, and $\mathbf{x}(s) = \mathbf{x}(y,s)$.

2. We assert that if $y \in \Gamma$ is sufficiently close to x^0, then

$$\text{(39)} \qquad f(y,s) := F(\mathbf{p}(y,s), z(y,s), \mathbf{x}(y,s)) = 0 \quad (s \in \mathbb{R}).$$

To see this, note

$$f(y,0) = F(\mathbf{p}(y,0), z(y,0), \mathbf{x}(y,0)) \\ = F(\mathbf{q}(y), g(y), y) = 0, \tag{40}$$

by the compatibility condition (34). Furthermore

$$\begin{aligned}\frac{\partial f}{\partial s}(y,s) &= \sum_{j=1}^{n} \frac{\partial F}{\partial p_j}\dot{p}^j + \frac{\partial F}{\partial z}\dot{z} + \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\dot{x}^j \\ &= \sum_{j=1}^{n} \frac{\partial F}{\partial p_j}\left(-\frac{\partial F}{\partial x_j} - \frac{\partial F}{\partial z}p^j\right) + \frac{\partial F}{\partial z}\left(\sum_{j=1}^{n} \frac{\partial F}{\partial p_j}p^j\right) \\ &\quad + \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\left(\frac{\partial F}{\partial p_j}\right), \qquad \text{according to (31)} \\ &= 0.\end{aligned}$$

This calculation and (40) prove (39).

3. In view of Lemma 2 and (37)–(39), we have

$$F(\mathbf{p}(x), u(x), x) = 0 \quad (x \in V).$$

To conclude, we must therefore show

$$\mathbf{p}(x) = Du(x) \quad (x \in V). \tag{41}$$

4. In order to prove (41), let us first demonstrate for $s \in I$, $y \in W$ that

$$\frac{\partial z}{\partial s}(y,s) = \sum_{j=1}^{n} p^j(y,s)\frac{\partial x^j}{\partial s}(y,s) \tag{42}$$

and

$$\frac{\partial z}{\partial y_i}(y,s) = \sum_{j=1}^{n} p^j(y,s)\frac{\partial x^j}{\partial y_i}(y,s) \quad (i = 1, \dots, n-1). \tag{43}$$

These formulas are obviously consistent with the equality (41) and will later help us prove it. The identity (42) results at once from the characteristic ODE (31)(b),(c). To establish (43), fix $y \in \Gamma$, $i \in \{1, \dots, n-1\}$, and set

$$r^i(s) := \frac{\partial z}{\partial y_i}(y,s) - \sum_{j=1}^{n} p^j(y,s)\frac{\partial x^j}{\partial y_i}(y,s). \tag{44}$$

We first note $r^i(0) = g_{x_i}(y) - q^i(y) = 0$ according to the compatibility condition (34). In addition, we can compute

$$\dot{r}^i(s) = \frac{\partial^2 z}{\partial y_i \partial s} - \sum_{j=1}^n \left[\frac{\partial p^j}{\partial s} \frac{\partial x^j}{\partial y_i} + p^j \frac{\partial^2 x^j}{\partial y_i \partial s} \right]. \tag{45}$$

In order to simplify this expression, let us first differentiate the identity (42) with respect to y_i:

$$\frac{\partial^2 z}{\partial s \partial y_i} = \sum_{j=1}^n \left[\frac{\partial p^j}{\partial y_i} \frac{\partial x^j}{\partial s} + p^j \frac{\partial^2 x^j}{\partial s \partial y_i} \right]. \tag{46}$$

Substituting (46) into (45), we discover

$$\begin{aligned} \dot{r}^i(s) &= \sum_{j=1}^n \left[\frac{\partial p^j}{\partial y_i} \frac{\partial x^j}{\partial s} - \frac{\partial p^j}{\partial s} \frac{\partial x^j}{\partial y_i} \right] \\ &= \sum_{j=1}^n \left[\frac{\partial p^j}{\partial y_i} \left(\frac{\partial F}{\partial p_j} \right) - \left(-\frac{\partial F}{\partial x_j} - \frac{\partial F}{\partial z} p^j \right) \frac{\partial x^j}{\partial y_i} \right] \quad \text{by (31)(a).} \end{aligned} \tag{47}$$

Now differentiate (39) with respect to y_i:

$$\sum_{j=1}^n \frac{\partial F}{\partial p_j} \frac{\partial p^j}{\partial y_i} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial y_i} + \sum_{j=1}^n \frac{\partial F}{\partial x_j} \frac{\partial x^j}{\partial y_i} = 0.$$

We employ this identity in (47), thereby obtaining

$$\dot{r}^i(s) = \frac{\partial F}{\partial z} \left[\sum_{j=1}^n p^j \frac{\partial x^j}{\partial y_i} - \frac{\partial z}{\partial y_i} \right] = -\frac{\partial F}{\partial z} r^i(s). \tag{48}$$

Hence $r^i(\cdot)$ solves the linear ODE (48), with the initial condition $r^i(0) = 0$. Consequently $r^i(s) = 0$ ($s \in \mathbb{R}$, $i = 1, \ldots, n-1$); and so identity (43) is verified.

5. We finally employ (42), (43) in proving (41). Indeed, if $j = 1, \ldots, n$,

$$\begin{aligned} \frac{\partial u}{\partial x_j} &= \frac{\partial z}{\partial s} \frac{\partial s}{\partial x_j} + \sum_{i=1}^{n-1} \frac{\partial z}{\partial y_i} \frac{\partial y^i}{\partial x_j} \quad \text{by (38)} \\ &= \left(\sum_{k=1}^n p^k \frac{\partial x^k}{\partial s} \right) \frac{\partial s}{\partial x_j} + \sum_{i=1}^{n-1} \left(\sum_{k=1}^n p^k \frac{\partial x^k}{\partial y_i} \right) \frac{\partial y^i}{\partial x_j} \quad \text{by (42), (43)} \\ &= \sum_{k=1}^n p^k \left(\frac{\partial x^k}{\partial s} \frac{\partial s}{\partial x_j} + \sum_{i=1}^{n-1} \frac{\partial x^k}{\partial y_i} \frac{\partial y^i}{\partial x_j} \right) \\ &= \sum_{k=1}^n p^k \frac{\partial x^k}{\partial x_j} = \sum_{k=1}^n p^k \delta_{jk} = p^j. \end{aligned}$$

This assertion at last establishes (41), and so finishes up the proof. □

3.2.5. Applications.

We turn now to various special cases, to see how the local existence theory simplifies in these circumstances.

a. F linear.

Recall that a linear, homogeneous, first-order PDE has the form

$$F(Du, u, x) = \mathbf{b}(x) \cdot Du(x) + c(x)u(x) = 0 \quad (x \in U). \tag{49}$$

Our noncharacteristic assumption (36) at a point $x^0 \in \Gamma$ as above becomes

$$\mathbf{b}(x^0) \cdot \boldsymbol{\nu}(x^0) \neq 0, \tag{50}$$

and thus does not involve z^0 or p^0 at all. Furthermore if we specify the boundary condition

$$u = g \quad \text{on } \Gamma, \tag{51}$$

we can uniquely solve equation (34) for $\mathbf{q}(y)$ if $y \in \Gamma$ is near x^0. Thus we can apply the Local Existence Theorem 2 to construct a unique solution of (49), (51) in some neighborhood V containing x^0. Note carefully that although we have utilized the full characteristic equations (31) in the proof of Theorem 2, once we know the solution exists, we can use the reduced equations (17) (which do not involve $\mathbf{p}(\cdot)$) to compute the solution. Observe also the projected characteristics $\mathbf{x}(\cdot)$ emanating from distinct points on Γ cannot cross, owing to uniqueness of solutions of the initial-value problem for the ODE (17)(a).

Example 4. Suppose the trajectories of the ODE

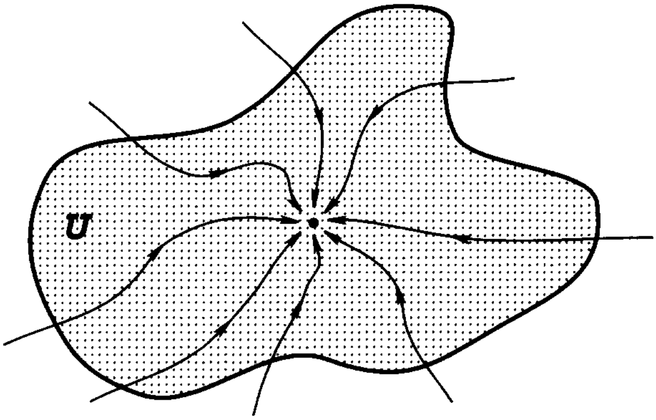
$$\dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s)) \tag{52}$$

are as drawn for Case 1.

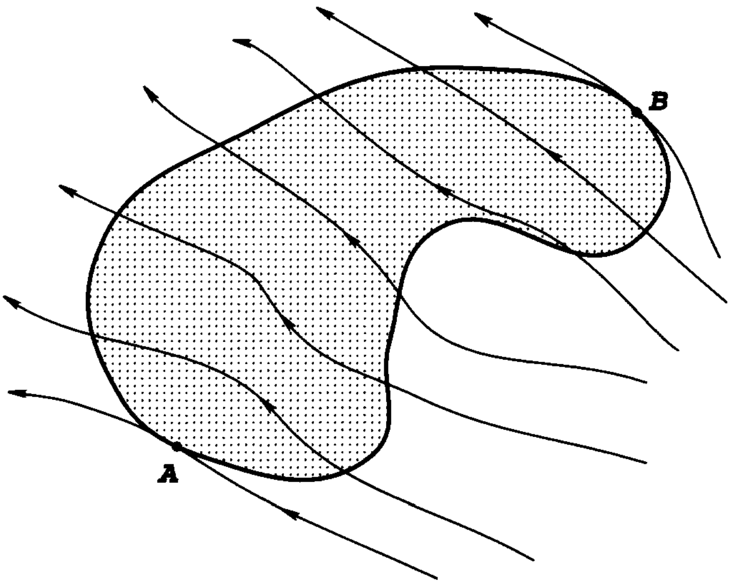
We are thus assuming the vector field $\mathbf{b}$ vanishes within U only at one point, which we will take to be the origin 0, and $\mathbf{b} \cdot \boldsymbol{\nu} < 0$ on $\Gamma := \partial U$. Can we solve the linear boundary-value problem

$$\begin{cases} \mathbf{b} \cdot Du = 0 & \text{in } U \\ \quad\;\; u = g & \text{on } \Gamma\ ? \end{cases} \tag{53}$$

Invoking Theorem 2 we see that there exists a unique solution u defined *near* Γ, and indeed that $u(\mathbf{x}(s)) \equiv u(\mathbf{x}(0)) = g(x^0)$ for each solution of the



Case 1: flow to an attracting point



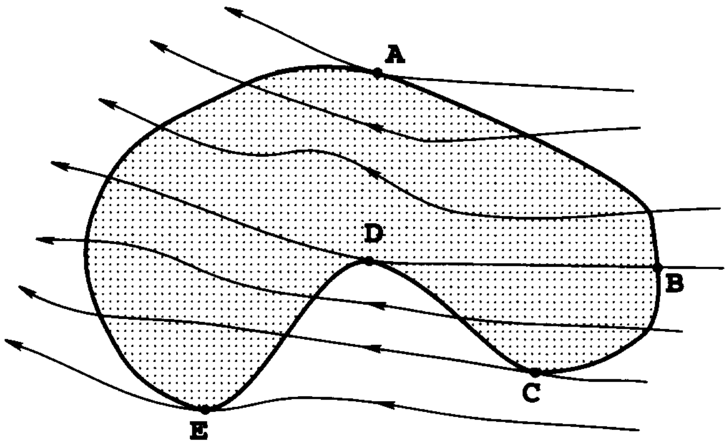
Case 2: flow across a domain

ODE (52), with the initial condition $\mathbf{x}(0) = x^0 \in \Gamma$. However, this solution cannot be smoothly continued to all of U (unless g is constant): any smooth solution of (53) is constant on trajectories of (52), and thus takes on different values near $x = 0$.

On the other hand, now suppose the trajectories of the ODE (52) look like the illustration for Case 2. We are consequently now assuming that each trajectory of the ODE (except those through the characteristic points A, B) enters U precisely once, somewhere through the set

$$\Gamma := \{x \in \partial U \mid \mathbf{b}(x) \cdot \boldsymbol{\nu}(x) < 0\},$$

and exits U precisely once. In this circumstance we can find a smooth solution of (53) by setting u to be constant along each flow line.



Case 3: flow with characteristic points

Assume finally the flow looks like Case 3. We can now define u to be constant along trajectories, but then u will be discontinuous (unless $g(B) = g(D)$).

Note that the point D is characteristic and that the local existence theory fails near D. □

b. F quasilinear.

Should F be quasilinear, the PDE (1) becomes

$$F(Du, u, x) = \mathbf{b}(x, u) \cdot Du + c(x, u) = 0. \tag{54}$$

The noncharacteristic assumption (36) at a point $x^0 \in \Gamma$ reads $\mathbf{b}(x^0, z^0) \cdot \boldsymbol{\nu}(x^0) \neq 0$, where $z^0 = g(x^0)$. As in the preceding example, if we specify the boundary condition

$$u = g \quad \text{on } \Gamma, \tag{55}$$

we can uniquely solve the equations (34) for $\mathbf{q}(y)$ if $y \in \Gamma$ near x^0. Thus Theorem 2 yields the existence of a unique solution of (54), (55) in some neighborhood V of x^0. We can compute this solution in V using the reduced characteristic equations (21), which do not explicitly involve $\mathbf{p}(\cdot)$.

In contrast to the linear case, however, *it is possible that the projected characteristics emanating from distinct points in* Γ *may intersect outside* V; such an occurrence usually signals the failure of our local solution to exist within all of U.

Example 5 (Characteristics for conservation laws). As an instance of a quasilinear first-order PDE, we turn now to the *scalar conservation law*

$$\begin{aligned} G(Du, u_t, u, x, t) &= u_t + \operatorname{div} \mathbf{F}(u) \\ &= u_t + \mathbf{F}'(u) \cdot Du = 0 \end{aligned} \tag{56}$$

in $U = \mathbb{R}^n \times (0, \infty)$, subject to the initial condition

$$u = g \quad \text{on } \Gamma = \mathbb{R}^n \times \{t = 0\}. \tag{57}$$

Here $\mathbf{F} : \mathbb{R} \to \mathbb{R}^n$, $\mathbf{F} = (F^1, \dots, F^n)$, and, as usual, we have set $t = x_{n+1}$. Also, "div" denotes the divergence with respect to the spatial variables $(x_1, \dots, x_n)$, and $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$.

Since the direction $t = x_{n+1}$ plays a special role, we appropriately modify our notation. Writing now $q = (p, p_{n+1})$ and $y = (x, t)$, we have

$$G(q, z, y) = p_{n+1} + \mathbf{F}'(z) \cdot p,$$

and consequently

$$D_q G = (\mathbf{F}'(z), 1), \; D_y G = 0, \; D_z G = \mathbf{F}''(z) \cdot p.$$

Clearly the noncharacteristic condition (35) is satisfied at each point $y^0 = (x^0, 0) \in \Gamma$. Furthermore equation (21)(a) becomes

$$\begin{cases} \dot{x}^i(s) = F^{i\prime}(z(s)) \quad (i = 1, \dots, n) \\ \dot{x}^{n+1}(s) = 1. \end{cases} \tag{58}$$

Hence $x^{n+1}(s) = s$, in agreement with our having written $x_{n+1} = t$ above. In other words, we can identify the parameter s with the time t.

Equation (21)(b) reads $\dot{z}(s) = 0$. Consequently

$$z(s) = z^0 = g(x^0); \tag{59}$$

and (58) implies

$$\mathbf{x}(s) = \mathbf{F}'(g(x^0))s + x^0. \tag{60}$$

Thus the projected characteristic $\mathbf{y}(s) = (\mathbf{x}(s), s) = (\mathbf{F}'(g(x^0))s + x^0, s)$ $(s \geq 0)$ is a straight line, along which u is constant.

Crossing characteristics. But suppose now we apply the same reasoning to a different initial point $z^0 \in \Gamma$, where $g(x^0) \neq g(z^0)$. *The projected characteristics may possibly then intersect at some time* $t > 0$. Since Theorem 1 tells us $u \equiv g(x^0)$ on the projected characteristic through x^0 and $u \equiv g(z^0)$ on the projected characteristic through z^0, an apparent contradiction arises. The resolution is that *the initial-value problem (56), (57) does not in general have a smooth solution, existing for all times* $t > 0$. □

We will discuss in §3.4 the interesting possibility of extending the local solution (guaranteed to exist for short times by Theorem 2) to all times $t > 0$, as a kind of "weak" or "generalized" solution.

Remark. Let us also note we can eliminate s from equations (59), (60) to obtain an implicit formula for u. Indeed given $x \in \mathbb{R}^n$ and $t > 0$, we see that since $s = t$,

$$\begin{aligned} u(\mathbf{x}(t), t) = z(t) &= g(\mathbf{x}(t) - t\mathbf{F}'(z^0)) \\ &= g(\mathbf{x}(t) - t\mathbf{F}'(u(\mathbf{x}(t), t))). \end{aligned}$$

Hence

$$u = g(x - t\mathbf{F}'(u)). \tag{61}$$

This implicit formula for u as a function of x and t is a nonlinear analogue of equation (3) in §2.1. It is easy to check (61) does indeed give a solution, provided

$$1 + tDg(x - t\mathbf{F}'(u)) \cdot \mathbf{F}''(u) \neq 0.$$

In particular if $n = 1$, we require

$$1 + tg'(x - tF'(u))F''(u) \neq 0.$$

Note that if $F'' > 0$, but $g' < 0$, then this will definitely be false at some time $t > 0$. This failure of the implicit formula (61) reflects also the failure of the characteristic method. □

c. F fully nonlinear.

The form of the full characteristic equations can be quite complicated for fully nonlinear first-order PDE, but sometimes a remarkable mathematical structure emerges.

Example 6 (Characteristics for the Hamilton–Jacobi equation). We look now at the general Hamilton–Jacobi PDE

$$G(Du, u_t, u, x, t) = u_t + H(Du, x) = 0, \tag{62}$$

where $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$. Then writing $q = (p, p_{n+1})$, $y = (x, t)$, we have

$$G(q, z, y) = p_{n+1} + H(p, x);$$

and so

$$D_q G = (D_p H(p, x), 1), \ D_y G = (D_x H(p, x), 0), \ D_z G = 0.$$

Thus equation (11)(c) becomes

$$\begin{cases} \dot{x}^i(s) = \frac{\partial H}{\partial p_i}(\mathbf{p}(s), \mathbf{x}(s)) \quad (i = 1, \dots, n) \\ \dot{x}^{n+1}(s) = 1. \end{cases} \tag{63}$$

In particular we can identify the parameter s with the time t.

Equation (11)(a) for the case at hand reads

$$\begin{cases} \dot{p}^i(s) = -\frac{\partial H}{\partial x_i}(\mathbf{p}(s), \mathbf{x}(s)) \quad (i = 1, \dots, n) \\ \dot{p}^{n+1}(s) = 0; \end{cases}$$

the equation (11)(b) is

$$\begin{aligned} \dot{z}(s) &= D_p H(\mathbf{p}(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) + p^{n+1}(s) \\ &= D_p H(\mathbf{p}(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) - H(\mathbf{p}(s), \mathbf{x}(s)). \end{aligned}$$

In summary, the characteristic equations for the Hamilton–Jacobi equation are:

$$(64) \qquad \begin{cases} \text{(a)} \quad \dot{\mathbf{p}}(s) = -D_x H(\mathbf{p}(s), \mathbf{x}(s)) \\ \text{(b)} \quad \dot{z}(s) = D_p H(\mathbf{p}(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) - H(\mathbf{p}(s), \mathbf{x}(s)) \\ \text{(c)} \quad \dot{\mathbf{x}}(s) = D_p H(\mathbf{p}(s), \mathbf{x}(s)) \end{cases}$$

for $\mathbf{p}(\cdot) = (p^1(\cdot), \dots, p^n(\cdot))$, $z(\cdot)$, and $\mathbf{x}(\cdot) = (x^1(\cdot), \dots, x^n(\cdot))$.

The first and third of these equalities,

$$(65) \qquad \begin{cases} \dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}), \end{cases}$$

are called *Hamilton's equations*. We will discuss these ODE and their relationship to the Hamilton–Jacobi equation in much more detail, just below in §3.3. Observe that the equation for $z(\cdot)$ is trivial, once $\mathbf{x}(\cdot)$ and $\mathbf{p}(\cdot)$ have been found by solving Hamilton's equations. □

As for conservation laws (Example 5), the initial-value problem for the Hamilton–Jacobi equation does not in general have a smooth solution u lasting for all times $t > 0$.

3.3. INTRODUCTION TO HAMILTON–JACOBI EQUATIONS

In this section we study in some detail the initial-value problem for the Hamilton–Jacobi equation:

$$(1) \qquad \begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Here $u : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ is the unknown, $u = u(x,t)$, and $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$. We are given the *Hamiltonian* $H : \mathbb{R}^n \to \mathbb{R}$ and the initial function $g : \mathbb{R}^n \to \mathbb{R}$.

Our goal is to find a formula for an appropriate weak or generalized solution, existing for all times $t > 0$, even after the method of characteristics has failed.

3.3.1. Calculus of variations, Hamilton's ODE.

Remember from §3.2.5 that two of the characteristic equations associated with the Hamilton–Jacobi PDE

$$u_t + H(Du, x) = 0$$

are Hamilton's ODE

$$\begin{cases} \dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}), \end{cases}$$

which arise in the classical calculus of variations and in mechanics. (Note the x-dependence in H here.) In this section we recall the derivation of these ODE from a variational principle. We will then discover in §3.3.2 that this discussion contains a clue as to how to build a weak solution of the initial-value problem (1).

a. The calculus of variations.

Assume that $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a given smooth function, hereafter called the *Lagrangian*.

Notation. We write

$$L = L(q, x) = L(q_1, \dots, q_n, x_1, \dots, x_n) \quad (q, x \in \mathbb{R}^n)$$

and

$$\begin{cases} D_q L = (L_{q_1} \cdots L_{q_n}) \\ D_x L = (L_{x_1} \cdots L_{x_n}). \end{cases}$$

Thus in the formula (2) below "q" is the name of the variable for which we substitute $\dot{\mathbf{w}}(s)$, and "x" is the variable for which we substitute $\mathbf{w}(s)$.

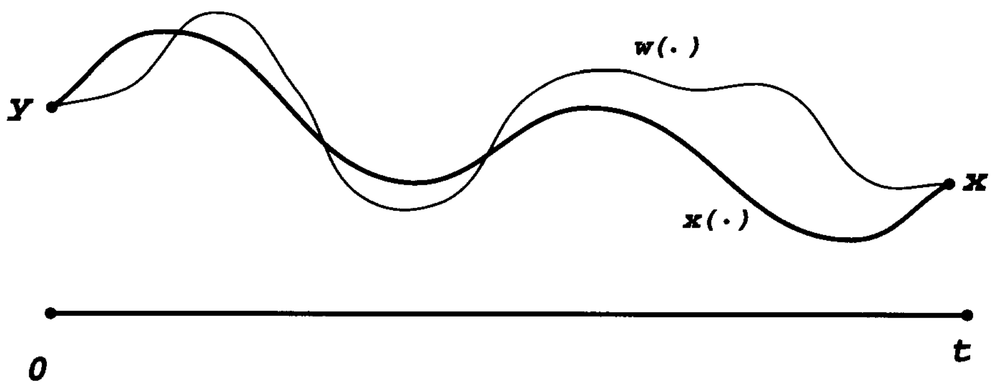
□

Now fix two points $x, y \in \mathbb{R}^n$ and a time $t > 0$. We introduce then the *action* functional

$$I[\mathbf{w}(\cdot)] = \int_0^t L(\dot{\mathbf{w}}(s), \mathbf{w}(s))\, ds \quad \left(\dot{} = \frac{d}{ds}\right), \tag{2}$$

defined for functions $\mathbf{w}(\cdot) = (w^1(\cdot), w^2(\cdot), \dots, w^n(\cdot))$ belonging to the *admissible class*

$$\mathcal{A} = \{\mathbf{w}(\cdot) \in C^2([0, t]; \mathbb{R}^n) \mid \mathbf{w}(0) = y, \mathbf{w}(t) = x\}.$$



A problem in the calculus of variations

Thus a C^2 curve $\mathbf{w}(\cdot)$ lies in $\mathcal{A}$ if it starts at the point y at time 0, and reaches the point x at time t. A basic problem in the *calculus of variations* is then to find a curve $\mathbf{x}(\cdot) \in \mathcal{A}$ satisfying

$$I[\mathbf{x}(\cdot)] = \min_{\mathbf{w}(\cdot)\in\mathcal{A}} I[\mathbf{w}(\cdot)]. \tag{3}$$

That is, we are asking for a function $\mathbf{x}(\cdot)$ which minimizes the functional $I[\cdot]$ among all admissible candidates $\mathbf{w}(\cdot) \in \mathcal{A}$.

We assume next that there in fact exists a function $\mathbf{x}(\cdot) \in \mathcal{A}$ satisfying our calculus of variations problem, and will deduce some of its properties.

THEOREM 1 (Euler–Lagrange equations). *The function* $\mathbf{x}(\cdot)$ *solves the system of* Euler–Lagrange equations

$$-\frac{d}{ds}\left(D_qL(\dot{\mathbf{x}}(s),\mathbf{x}(s))\right) + D_xL(\dot{\mathbf{x}}(s),\mathbf{x}(s)) = 0 \quad (0 \le s \le t). \tag{4}$$

This is a vector equation, consisting of n coupled second-order equations.

Proof. 1. Choose a smooth function $\mathbf{v} : [0,t] \to \mathbb{R}^n$, $\mathbf{v} = (v^1,\dots,v^n)$, satisfying

$$\mathbf{v}(0) = \mathbf{v}(t) = 0, \tag{5}$$

and define for $\tau \in \mathbb{R}$

$$\mathbf{w}(\cdot) := \mathbf{x}(\cdot) + \tau\mathbf{v}(\cdot). \tag{6}$$

Then $\mathbf{w}(\cdot) \in \mathcal{A}$ and so

$$I[\mathbf{x}(\cdot)] \le I[\mathbf{w}(\cdot)].$$

Thus the real-valued function

$$i(\tau) := I[\mathbf{x}(\cdot) + \tau\mathbf{v}(\cdot)]$$

has a minimum at $\tau = 0$, and consequently

$$i'(0) = 0 \quad \left(' = \frac{d}{d\tau} \right), \tag{7}$$

provided $i'(0)$ exists.

2. We explicitly compute this derivative. Observe

$$i(\tau) = \int_0^t L(\dot{\mathbf{x}}(s) + \tau\dot{\mathbf{v}}(s), \mathbf{x}(s) + \tau\mathbf{v}(s))\, ds,$$

and so

$$i'(\tau) = \int_0^t \sum_{i=1}^n L_{q_i}(\dot{\mathbf{x}} + \tau\dot{\mathbf{v}}, x + \tau\mathbf{v})\dot{v}^i + L_{x_i}(\dot{\mathbf{x}} + \tau\dot{\mathbf{v}}, x + \tau\mathbf{v})v^i\, ds.$$

Set $\tau = 0$ and remember (7):

$$0 = i'(0) = \int_0^t \sum_{i=1}^n L_{q_i}(\dot{\mathbf{x}}, \mathbf{x})\dot{v}^i + L_{x_i}(\dot{\mathbf{x}}, \mathbf{x})v^i\, ds.$$

We recall (5) and then integrate by parts in the first term inside the integral, to discover

$$0 = \sum_{i=1}^n \int_0^t \left[-\frac{d}{ds}\left(L_{q_i}(\dot{\mathbf{x}}, \mathbf{x})\right) + L_{x_i}(\dot{\mathbf{x}}, \mathbf{x}) \right] v^i\, ds.$$

This identity is valid for all smooth functions $\mathbf{v} = (v^1, \dots, v^n)$ satisfying the boundary conditions (5), and so

$$-\frac{d}{ds}\left(L_{q_i}(\dot{\mathbf{x}}, \mathbf{x})\right) + L_{x_i}(\dot{\mathbf{x}}, \mathbf{x}) = 0$$

for $0 \le s \le t$, $i = 1, \dots, n$. □

Remark. We have just demonstrated that any minimizer $\mathbf{x}(\cdot) \in \mathcal{A}$ of $I[\cdot]$ solves the Euler–Lagrange system of ODE. It is of course possible that a curve $\mathbf{x}(\cdot) \in \mathcal{A}$ may solve the Euler–Lagrange equations without necessarily being a minimizer: in this case we say $\mathbf{x}(\cdot)$ is a *critical point* of $I[\cdot]$. So every minimizer is a critical point, but a critical point need not be a minimizer. □

Example. If $L(q,x) = \frac{1}{2}m|q|^2 - \phi(x)$, where $m > 0$, the corresponding Euler–Lagrange equation is

$$m\ddot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s))$$

for $\mathbf{f} := -D\phi$. This is Newton's law for the motion of a particle of mass m moving in the force field $\mathbf{f}$ generated by the potential ϕ. (See Feynman–Leighton–Sands [**F-L-S,** Chapter 19].) □

b. Hamilton's ODE.

We now convert the Euler–Lagrange equations, a system of n second-order ODE, into Hamilton's equations, a system of $2n$ first-order ODE. We hereafter assume the C^2 function $\mathbf{x}(\cdot)$ is a critical point of the action functional, and thus solves the Euler–Lagrange equations (4).

First we set

$$\mathbf{p}(s) := D_q L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) \quad (0 \le s \le t); \tag{8}$$

$\mathbf{p}(\cdot)$ is called the *generalized momentum* corresponding to the *position* $\mathbf{x}(\cdot)$ and *velocity* $\dot{\mathbf{x}}(\cdot)$. We next make this important hypothesis:

$$\begin{cases} \text{Suppose for all } x, p \in \mathbb{R}^n \text{ that the equation} \\ \qquad p = D_q L(q,x) \\ \text{can be uniquely solved for } q \text{ as a smooth} \\ \text{function of } p \text{ and } x,\ q = \mathbf{q}(p,x). \end{cases} \tag{9}$$

We will examine this assumption in more detail later: see §3.3.2.

DEFINITION. *The* Hamiltonian H *associated with the Lagrangian* L *is*

$$H(p,x) := p \cdot \mathbf{q}(p,x) - L(\mathbf{q}(p,x), x) \quad (p, x \in \mathbb{R}^n),$$

where the function $\mathbf{q}(\cdot,\cdot)$ *is defined implicitly by* (9).

Example (continued). The Hamiltonian corresponding to the Lagrangian $L(q,x) = \frac{1}{2}m|q|^2 - \phi(x)$ is

$$H(p,x) = \frac{1}{2m}|p|^2 + \phi(x).$$

The Hamiltonian is thus the total energy, the sum of the kinetic and potential energies; the Lagrangian is the difference between the kinetic and potential energies. □

Next we rewrite the Euler–Lagrange equations in terms of $\mathbf{p}(\cdot), \mathbf{x}(\cdot)$:

THEOREM 2 (Derivation of Hamilton's ODE). *The functions* $\mathbf{x}(\cdot)$ *and* $\mathbf{p}(\cdot)$ *satisfy* Hamilton's equations*:*

$$\begin{cases} \dot{\mathbf{x}}(s) = D_pH(\mathbf{p}(s), \mathbf{x}(s)) \\ \dot{\mathbf{p}}(s) = -D_xH(\mathbf{p}(s), \mathbf{x}(s)) \end{cases} \tag{10}$$

for $0 \leq s \leq t$. *Furthermore,*

$$\textit{the mapping } s \mapsto H(\mathbf{p}(s), \mathbf{x}(s)) \textit{ is constant.}$$

Remark. The equations (10) comprise a coupled system of $2n$ first-order ODE for $\mathbf{x}(\cdot) = (x^1(\cdot), \dots, x^n(\cdot))$ and $\mathbf{p}(\cdot) = (p^1(\cdot), \dots, p^n(\cdot))$ (defined by (8)). □

Proof. First note from (8) and (9) that $\dot{\mathbf{x}}(s) = \mathbf{q}(\mathbf{p}(s), \mathbf{x}(s))$.

Let us hereafter write $\mathbf{q}(\cdot) = (q^1(\cdot), \dots, q^n(\cdot))$. We then compute for $i = 1, \dots, n$:

$$\begin{aligned} \frac{\partial H}{\partial x_i}(p, x) &= \sum_{k=1}^n p_k \frac{\partial q^k}{\partial x_i}(p, x) - \frac{\partial L}{\partial q_k}(q, x) \frac{\partial q^k}{\partial x_i}(p, x) - \frac{\partial L}{\partial x_i}(q, x) \\ &= -\frac{\partial L}{\partial x_i}(q, x) \quad \text{by (9)}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial H}{\partial p_i}(p, x) &= q^i(p, x) + \sum_{k=1}^n p_k \frac{\partial q^k}{\partial p_i}(p, x) - \frac{\partial L}{\partial q_k}(q, x) \frac{\partial q^k}{\partial p_i}(p, x) \\ &= q^i(p, x), \quad \text{again by (9)}. \end{aligned}$$

Thus

$$\frac{\partial H}{\partial p_i}(\mathbf{p}(s), \mathbf{x}(s)) = q^i(\mathbf{p}(s), \mathbf{x}(s)) = \dot{x}^i(s);$$

and likewise

$$\begin{aligned} \frac{\partial H}{\partial x_i}(\mathbf{p}(s), \mathbf{x}(s)) &= -\frac{\partial L}{\partial x_i}(\mathbf{q}(\mathbf{p}(s), \mathbf{x}(s)), \mathbf{x}(s)) = -\frac{\partial L}{\partial x_i}(\dot{\mathbf{x}}(s), \mathbf{x}(s)) \\ &= -\frac{d}{ds}\left(\frac{\partial L}{\partial q_i}(\dot{\mathbf{x}}(s), \mathbf{x}(s))\right) \text{ according to (4)} \\ &= -\dot{p}^i(s). \end{aligned}$$

Finally, observe

$$\begin{aligned} \frac{d}{ds} H(\mathbf{p}(s), \mathbf{x}(s)) &= \sum_{i=1}^n \frac{\partial H}{\partial p_i} \dot{p}^i + \frac{\partial H}{\partial x_i} \dot{x}^i \\ &= \sum_{i=1}^n \frac{\partial H}{\partial p_i} \left(-\frac{\partial H}{\partial x_i}\right) + \frac{\partial H}{\partial x_i} \left(\frac{\partial H}{\partial p_i}\right) = 0. \end{aligned}$$

□

Remark. See Arnold [**AR,** Chapter 9] for more on Hamilton's ODE and Hamilton–Jacobi PDE in classical mechanics. We are employing here different notation than is customary in mechanics: our notation is better overall for PDE theory. □

3.3.2. Legendre transform, Hopf–Lax formula.

Now let us try to find a connection between the Hamilton–Jacobi PDE and the calculus of variations problem (2)–(4). To simplify further, we also drop the x-dependence in the Hamiltonian, so that afterwards $H = H(p)$. We start by reexamining the definition of the Hamiltonian in §3.3.1.

a. Legendre transform.

We hereafter suppose the Lagrangian $L : \mathbb{R}^n \to \mathbb{R}$ satisfies these conditions:

$$\text{the mapping } q \mapsto L(q) \text{ is convex} \tag{11}$$

and

$$\lim_{q\to\infty} \frac{L(q)}{|q|} = +\infty. \tag{12}$$

The convexity implies L is continuous.

DEFINITION. *The* Legendre transform *of* L *is*

$$L^*(p) = \sup_{q\in\mathbb{R}^n} \{p \cdot q - L(q)\} \quad (p \in \mathbb{R}^n). \tag{13}$$

Motivation for Legendre transform. Why do we make this definition? For some insight let us note in view of (12) that the "sup" in (13) is really a "max"; that is, there exists some $q^* \in \mathbb{R}^n$ for which

$$L^*(p) = p \cdot q^* - L(q^*)$$

and the mapping $q \mapsto p \cdot q - L(q)$ has a maximum at $q = q^*$. But then $p = DL(q^*)$, provided L is differentiable at q^*. Hence the equation $p = DL(q)$ is solvable (although perhaps not uniquely) for q in terms of p, $q^* = \mathbf{q}(p)$. Therefore

$$L^*(p) = p \cdot \mathbf{q}(p) - L(\mathbf{q}(p)).$$

However, this is almost exactly the definition of the Hamiltonian H associated with L in §3.3.1 (where, recall, we are now assuming the variable x does not appear). We consequently henceforth write

$$H = L^*. \tag{14}$$

Thus (13) tells us how to obtain the Hamiltonian H from the Lagrangian L. □

Now we ask the converse question: given H, how do we compute L?

THEOREM 3 (Convex duality of Hamiltonian and Lagrangian). *Assume L satisfies* (11), (12) *and define H by* (13), (14).

(i) *Then*

$$\text{the mapping } p \mapsto H(p) \text{ is convex}$$

and

$$\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty.$$

(ii) *Furthermore*

$$L = H^*. \tag{15}$$

Remark. Thus H is the Legendre transform of L, and vice versa:

$$L = H^*, \ H = L^*.$$

We say H and L are *dual* convex functions. □

Proof. 1. For each fixed q, the function $p \mapsto p \cdot q - L(q)$ is linear, and consequently the mapping

$$p \mapsto H(p) = L^*(p) = \sup_{q\in\mathbb{R}^n} \{p \cdot q - L(q)\}$$

is convex. Indeed, if $0 \le \tau \le 1$, $p, \hat{p} \in \mathbb{R}^n$,

$$\begin{aligned} H(\tau p + (1-\tau)\hat{p}) &= \sup_q \{(\tau p + (1-\tau)\hat{p}) \cdot q - L(q)\} \\ &\le \tau \sup_q \{p \cdot q - L(q)\} \\ &\quad\ + (1-\tau) \sup_q \{\hat{p} \cdot q - L(q)\} \\ &= \tau H(p) + (1-\tau) H(\hat{p}). \end{aligned}$$

2. Fix any $\lambda > 0$, $p \neq 0$. Then

$$\begin{aligned} H(p) &= \sup_{q\in\mathbb{R}^n} \{p \cdot q - L(q)\} \\ &\ge \lambda|p| - L(\lambda\frac{p}{|p|}) \quad (q = \lambda\frac{p}{|p|}) \\ &\ge \lambda|p| - \max_{B(0,\lambda)} L. \end{aligned}$$

Thus $\liminf_{|p|\to\infty} \frac{H(p)}{|p|} \ge \lambda$ for all $\lambda > 0$.

3. In view of (14)

$$H(p) + L(q) \geq p \cdot q$$

for all $p, q \in \mathbb{R}^n$, and consequently

$$L(q) \geq \sup_{p \in \mathbb{R}^n} \{p \cdot q - H(p)\} = H^*(q).$$

On the other hand

$$\begin{aligned} H^*(q) &= \sup_{p \in \mathbb{R}^n} \{p \cdot q - \sup_{r \in \mathbb{R}^n} \{p \cdot r - L(r)\}\} \\ &= \sup_{p \in \mathbb{R}^n} \inf_{r \in \mathbb{R}^n} \{p \cdot (q - r) + L(r)\}. \end{aligned} \tag{16}$$

Now since $q \mapsto L(q)$ is convex, according to §B.1 there exists $s \in \mathbb{R}^n$ such that

$$L(r) \geq L(q) + s \cdot (r - q) \quad (r \in \mathbb{R}^n).$$

(If L is differentiable at q, take $s = DL(q)$.) Taking $p = s$ in (16), we compute

$$H^*(q) \geq \inf_{r \in \mathbb{R}^n} \{s \cdot (q - r) + L(r)\} = L(q).$$

□

b. Hopf–Lax formula.

Let us now return to the initial-value problem (1). Recall that the calculus of variations problem with Lagrangian L, discussed in §3.3.1, led to Hamilton's ODE for the associated Hamiltonian H. Since these ODE are also the characteristic equations of the Hamilton–Jacobi PDE, we conjecture there is probably a direct connection between this PDE and the calculus of variations.

So if $x \in \mathbb{R}^n$ and $t > 0$ are given, we should presumably try to minimize the action

$$\int_0^t L(\dot{\mathbf{w}}(s))\, ds$$

over functions $\mathbf{w} : [0, t] \to \mathbb{R}^n$ satisfying $\mathbf{w}(t) = x$. But what should we take for $\mathbf{w}(0)$? As we must somehow take into account the initial condition for our PDE, let us try modifying the action to include the function g evaluated at $\mathbf{w}(0)$:

$$\int_0^t L(\dot{\mathbf{w}}(s))\, ds + g(\mathbf{w}(0)).$$

Next let us construct a candidate for a solution to the initial-value problem (1), in terms of a variational principle entailing this modified action. We accordingly set

$$u(x,t) := \inf\left\{\int_0^t L(\dot{\mathbf{w}}(s))\,ds + g(y) \mid \mathbf{w}(0)=y, \mathbf{w}(t)=x\right\}, \tag{17}$$

the infimum taken over all C^1 functions $\mathbf{w}(\cdot)$ with $\mathbf{w}(t)=x$. (Better justification for this guess will be provided much later, in Chapter 10.)

We propose now to investigate the sense in which u so defined by (17) actually solves the initial-value problem for the Hamilton–Jacobi PDE:

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{18}$$

Recall we are assuming H is smooth,

$$\begin{cases} H \text{ is convex and} \\ \lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty. \end{cases} \tag{19}$$

We henceforth suppose also

$$g : \mathbb{R}^n \to \mathbb{R} \quad \text{is Lipschitz continuous;} \tag{20}$$

this means $\mathrm{Lip}(g) := \sup_{\substack{x,y\in\mathbb{R}^n \\ x\neq y}} \left\{\frac{|g(x)-g(y)|}{|x-y|}\right\} < \infty$.

First we note formula (17) can be simplified:

THEOREM 4 (Hopf–Lax formula). *If $x \in \mathbb{R}^n$ and $t>0$, then the solution $u=u(x,t)$ of the minimization problem* (17) *is*

$$u(x,t) = \min_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + g(y)\right\}. \tag{21}$$

DEFINITION. *We call the expression on the right hand side of* (21) *the* Hopf–Lax formula.

Proof. 1. Fix any $y \in \mathbb{R}^n$ and define $\mathbf{w}(s) := y + \frac{s}{t}(x-y)$ $(0 \le s \le t)$. Then the definition (17) of u implies

$$u(x,t) \le \int_0^t L(\dot{\mathbf{w}}(s))\,ds + g(y) = tL\left(\frac{x-y}{t}\right) + g(y),$$

and so

$$u(x,t) \leq \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\}.$$

2. On the other hand, if $\mathbf{w}(\cdot)$ is any C^1 function satisfying $\mathbf{w}(t) = x$, we have

$$L\left(\frac{1}{t}\int_0^t \dot{\mathbf{w}}(s)\,ds\right) \leq \frac{1}{t}\int_0^t L(\dot{\mathbf{w}}(s))\,ds$$

by Jensen's inequality (§B.1). Thus if we write $y = \mathbf{w}(0)$, we find

$$tL\left(\frac{x-y}{t}\right) + g(y) \leq \int_0^t L(\dot{\mathbf{w}}(s))\,ds + g(y);$$

and consequently

$$\inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \leq u(x,t).$$

3. We have so far shown

$$u(x,t) = \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\},$$

and leave it as an exercise to prove the infimum above is really a minimum. □

We now commence a study of various properties of the function u defined by the Hopf–Lax formula (21). Our ultimate goal is showing this formula provides a reasonable weak solution of the initial-value problem (18) for the Hamilton–Jacobi equation.

First, we record some preliminary observations.

LEMMA 1 (A functional identity). *For each $x \in \mathbb{R}^n$ and $0 \leq s < t$, we have*

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \right\}. \tag{22}$$

In other words, to compute $u(\cdot,t)$, we can calculate u at time s and then use $u(\cdot,s)$ as the initial condition on the remaining time interval $[s,t]$.

Proof. 1. Fix $y \in \mathbb{R}^n$, $0 < s < t$ and choose $z \in \mathbb{R}^n$ so that

$$u(y,s) = sL\left(\frac{y-z}{s}\right) + g(z). \tag{23}$$

Now since L is convex and $\frac{x-z}{t} = \left(1 - \frac{s}{t}\right)\frac{x-y}{t-s} + \frac{s}{t}\frac{y-z}{s}$, we have

$$L\left(\frac{x-z}{t}\right) \le \left(1-\frac{s}{t}\right)L\left(\frac{x-y}{t-s}\right) + \frac{s}{t}L\left(\frac{y-z}{s}\right).$$

Thus

$$\begin{aligned} u(x,t) \le tL\left(\frac{x-z}{t}\right) + g(z) &\le (t-s)L\left(\frac{x-y}{t-s}\right) + sL\left(\frac{y-z}{s}\right) + g(z) \\ &= (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s), \end{aligned}$$

by (23). This inequality is true for each $y \in \mathbb{R}^n$. Therefore, since $y \mapsto u(y,s)$ is continuous (according to Lemma 2 below), we have

$$u(x,t) \le \min_{y\in\mathbb{R}^n}\left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\}. \tag{24}$$

2. Now choose w such that

$$u(x,t) = tL\left(\frac{x-w}{t}\right) + g(w), \tag{25}$$

and set $y := \frac{s}{t}x + \left(1-\frac{s}{t}\right)w$. Then $\frac{x-y}{t-s} = \frac{x-w}{t} = \frac{y-w}{s}$. Consequently

$$\begin{aligned} &(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \\ &\qquad \le (t-s)L\left(\frac{x-w}{t}\right) + sL\left(\frac{y-w}{s}\right) + g(w) \\ &\qquad = tL\left(\frac{x-w}{t}\right) + g(w) = u(x,t), \end{aligned}$$

by (25). Hence

$$\min_{y\in\mathbb{R}^n}\left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\} \le u(x,t). \tag{26}$$

□

LEMMA 2 (Lipschitz continuity). *The function u is Lipschitz continuous in $\mathbb{R}^n \times [0,\infty)$, and*

$$u = g \quad \text{on } \mathbb{R}^n \times \{t=0\}.$$

Proof. 1. Fix $t > 0$, $x, \hat{x} \in \mathbb{R}^n$. Choose $y \in \mathbb{R}^n$ such that

$$tL\left(\frac{x-y}{t}\right) + g(y) = u(x,t). \tag{27}$$

Then

$$\begin{aligned} u(\hat{x},t) - u(x,t) &= \inf_z \left\{ tL\left(\frac{\hat{x}-z}{t}\right) + g(z)\right\} - tL\left(\frac{x-y}{t}\right) - g(y) \\ &\leq g(\hat{x} - x + y) - g(y) \leq \operatorname{Lip}(g)|\hat{x} - x|. \end{aligned}$$

Hence

$$u(\hat{x},t) - u(x,t) \leq \operatorname{Lip}(g)|\hat{x} - x|;$$

and, interchanging the roles of $\hat{x}$ and x, we find

$$|u(x,t) - u(\hat{x},t)| \leq \operatorname{Lip}(g)|x - \hat{x}|. \tag{28}$$

2. Now select $x \in \mathbb{R}^n$, $t > 0$. Choosing $y = x$ in (21), we discover

$$u(x,t) \leq tL(0) + g(x). \tag{29}$$

Furthermore,

$$\begin{aligned} u(x,t) &= \min_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y)\right\} \\ &\geq g(x) + \min_{y\in\mathbb{R}^n} \left\{ -\operatorname{Lip}(g)|x-y| + tL\left(\frac{x-y}{t}\right)\right\} \\ &= g(x) - t\max_{z\in\mathbb{R}^n}\{\operatorname{Lip}(g)|z| - L(z)\} \quad \left(z = \frac{x-y}{t}\right) \\ &= g(x) - t \max_{w\in B(0,\operatorname{Lip}(g))} \max_{z\in\mathbb{R}^n}\{w\cdot z - L(z)\} \\ &= g(x) - t \max_{B(0,\operatorname{Lip}(g))} H. \end{aligned}$$

This inequality and (29) imply

$$|u(x,t) - g(x)| \leq Ct$$

for

$$C := \max(|L(0)|, \max_{B(0,\operatorname{Lip}(g))} |H|). \tag{30}$$

3. Finally select $x \in \mathbb{R}^n$, $0 < \hat{t} < t$. Then $\operatorname{Lip}(u(\cdot,t)) \leq \operatorname{Lip}(g)$ by (28) above. Consequently Lemma 1 and calculations like those employed in step 2 above imply

$$|u(x,t) - u(x,\hat{t})| \leq C|t - \hat{t}|$$

for the constant C defined by (30). □

Now Rademacher's Theorem (which we will prove later, in §5.8.3) asserts that a Lipschitz function is differentiable almost everywhere. Consequently in view of Lemma 2 our function u defined by the Hopf–Lax formula (21) is differentiable for a.e. $(x,t) \in \mathbb{R}^n \times (0,\infty)$. The next theorem asserts u in fact solves the Hamilton–Jacobi PDE wherever u is differentiable.

THEOREM 5 (Solving the Hamilton–Jacobi equation). *Suppose $x \in \mathbb{R}^n$, $t > 0$, and u defined by the Hopf-Lax formula* (21) *is differentiable at a point $(x,t) \in \mathbb{R}^n \times (0,\infty)$. Then*

$$u_t(x,t) + H(Du(x,t)) = 0.$$

Proof. 1. Fix $q \in \mathbb{R}^n$, $h > 0$. Owing to Lemma 1,

$$\begin{aligned} u(x+hq,t+h) &= \min_{y\in\mathbb{R}^n}\left\{hL\left(\frac{x+hq-y}{h}\right)+u(y,t)\right\} \\ &\leq hL(q)+u(x,t). \end{aligned}$$

Hence

$$\frac{u(x+hq,t+h)-u(x,t)}{h} \leq L(q).$$

Let $h \to 0^+$, to compute

$$q \cdot Du(x,t) + u_t(x,t) \leq L(q).$$

This inequality is valid for all $q \in \mathbb{R}^n$, and so

$$(31) \qquad u_t(x,t) + H(Du(x,t)) = u_t(x,t) + \max_{q\in\mathbb{R}^n}\{q\cdot Du(x,t) - L(q)\} \leq 0.$$

The first equality holds since $H = L^*$.

2. Now choose z such that $u(x,t) = tL\left(\frac{x-z}{t}\right) + g(z)$. Fix $h > 0$ and set $s = t-h$, $y = \frac{s}{t}x + \left(1-\frac{s}{t}\right)z$. Then $\frac{x-z}{t} = \frac{y-z}{s}$, and thus

$$\begin{aligned} u(x,t)-u(y,s) &\geq tL\left(\frac{x-z}{t}\right)+g(z)-\left[sL\left(\frac{y-z}{s}\right)+g(z)\right] \\ &= (t-s)L\left(\frac{x-z}{t}\right). \end{aligned}$$

That is,

$$\frac{u(x,t)-u(\left(1-\frac{h}{t}\right)x+\frac{h}{t}z,t-h)}{h} \geq L\left(\frac{x-z}{t}\right).$$

Let $h \to 0^+$ to compute

$$\frac{x-z}{t} \cdot Du(x,t) + u_t(x,t) \geq L\left(\frac{x-z}{t}\right).$$

Consequently

$$\begin{aligned} u_t(x,t) + H(Du(x,t)) &= u_t(x,t) + \max_{q \in \mathbb{R}^n} \{q \cdot Du(x,t) - L(q)\} \\ &\geq u_t(x,t) + \frac{x-z}{t} \cdot Du(x,t) - L\left(\frac{x-z}{t}\right) \\ &\geq 0. \end{aligned}$$

This inequality and (31) complete the proof. □

We summarize:

THEOREM 6 (Hopf–Lax formula as solution). *The function u defined by the Hopf–Lax formula* (21) *is Lipschitz continuous, is differentiable a.e. in $\mathbb{R}^n \times (0,\infty)$, and solves the initial-value problem*

$$\text{(32)} \qquad \begin{cases} u_t + H(Du) = 0 & \text{a.e. in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

3.3.3. Weak solutions, uniqueness.

a. Semiconcavity.

In view of Theorem 6 above it may seem reasonable to define a weak solution of the initial-value problem (18) to be a Lipschitz function which agrees with g on $\mathbb{R}^n \times \{t=0\}$, and solves the PDE a.e. on $\mathbb{R}^n \times (0,\infty)$. However, this turns out to be an inadequate definition, *as such weak solutions would not in general be unique.*

Example. Consider the initial-value problem

$$\text{(33)} \qquad \begin{cases} u_t + |u_x|^2 = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = 0 & \text{on } \mathbb{R} \times \{t=0\}. \end{cases}$$

One obvious solution is

$$u_1(x,t) \equiv 0.$$

However the function

$$u_2(x,t) := \begin{cases} 0 & \text{if } |x| \geq t \\ x-t & \text{if } 0 \leq x \leq t \\ -x-t & \text{if } -t \leq x \leq 0 \end{cases}$$

is Lipschitz continuous and also solves the PDE a.e. (everywhere, in fact, except on the lines $x = 0, \pm t$). It is easy to see that actually there are infinitely many Lipschitz functions satisfying (33). □

This example shows we must presumably require more of a weak solution than merely that it satisfy the PDE a.e. We will look to the Hopf–Lax formula (21) for a further clue as to what is needed to ensure uniqueness. The following lemma demonstrates that u inherits a kind of "one-sided" second-derivative estimate from the initial function g.

LEMMA 3 (Semiconcavity). *Suppose there exists a constant C such that*

$$g(x+z) - 2g(x) + g(x-z) \leq C|z|^2 \tag{34}$$

for all $x, z \in \mathbb{R}^n$. Define u by the Hopf–Lax formula (21). *Then*

$$u(x+z,t) - 2u(x,t) + u(x-z,t) \leq C|z|^2$$

for all $x, z \in \mathbb{R}^n$, $t > 0$.

Remark. We say g is *semiconcave* provided (34) holds. It is easy to check (34) is valid if g is C^2 and $\sup_{\mathbb{R}^n} |D^2 g| < \infty$. Note that g is semiconcave if and only if the mapping $x \mapsto g(x) - \frac{C}{2}|x|^2$ is concave for some constant C. □

Proof. Choose $y \in \mathbb{R}^n$ so that $u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y)$. Then, putting $y+z$ and $y-z$ in the Hopf–Lax formulas for $u(x+z,t)$ and $u(x-z,t)$, we find

$$\begin{aligned} &u(x+z,t) - 2u(x,t) + u(x-z,t) \\ &\quad\leq \left[tL\left(\frac{x-y}{t}\right) + g(y+z)\right] - 2\left[tL\left(\frac{x-y}{t}\right) + g(y)\right] \\ &\qquad + \left[tL\left(\frac{x-y}{t}\right) + g(y-z)\right] \\ &\quad= g(y+z) - 2g(y) + g(y-z) \\ &\quad\leq C|z|^2, \quad \text{by (34).} \end{aligned}$$

□

As a semiconcavity condition for u will turn out to be important, we pause to identify some other circumstances under which it is valid. We will no longer assume g to be semiconcave, but will suppose the Hamiltonian H to be uniformly convex.

DEFINITION. *A C^2 convex function $H : \mathbb{R}^n \to \mathbb{R}$ is called* uniformly convex *(with constant $\theta > 0$) if*

$$\sum_{i,j=1}^{n} H_{p_i p_j}(p)\xi_i \xi_j \geq \theta |\xi|^2 \quad \textit{for all } p, \xi \in \mathbb{R}^n. \tag{35}$$

We now prove that even if g is not semiconcave, the uniform convexity of H forces u to become semiconcave for times $t > 0$: this is a kind of mild regularizing effect for the Hopf–Lax solution of the initial-value problem (18).

LEMMA 4 (Semiconcavity again). *Suppose that H is uniformly convex (with constant θ) and u is defined by the Hopf–Lax formula* (21). *Then*

$$u(x+z,t) - 2u(x,t) + u(x-z,t) \leq \frac{1}{\theta t}|z|^2$$

for all $x, z \in \mathbb{R}^n$, $t > 0$.

Proof. 1. We note first using Taylor's formula that (35) implies

$$H\left(\frac{p_1+p_2}{2}\right) \leq \frac{1}{2}H(p_1) + \frac{1}{2}H(p_2) - \frac{\theta}{8}|p_1 - p_2|^2. \tag{36}$$

Next we claim that for the Lagrangian L we have the estimate

$$\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) \leq L\left(\frac{q_1+q_2}{2}\right) + \frac{1}{8\theta}|q_1 - q_2|^2 \tag{37}$$

for all $q_1, q_2 \in \mathbb{R}^n$. Verification is left as an exercise.

2. Now choose y so that $u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y)$. Then using the same value of y in the Hopf–Lax formulas for $u(x+z,t)$ and $u(x-z,t)$, we calculate

$$\begin{aligned}
u(x+z,t) &- 2u(x,t) + u(x-z,t) \\
&\leq \left[tL\left(\frac{x+z-y}{t}\right) + g(y)\right] - 2\left[tL\left(\frac{x-y}{t}\right) + g(y)\right] \\
&\quad + \left[tL\left(\frac{x-z-y}{t}\right) + g(y)\right] \\
&= 2t\left[\frac{1}{2}L\left(\frac{x+z-y}{t}\right) + \frac{1}{2}L\left(\frac{x-z-y}{t}\right) - L\left(\frac{x-y}{t}\right)\right] \\
&\leq 2t\frac{1}{8\theta}\left|\frac{2z}{t}\right|^2 \leq \frac{1}{\theta t}|z|^2,
\end{aligned}$$

the next-to-last inequality following from (37). □

b. Weak solutions, uniqueness.

In this section we show that semiconcavity conditions of the sorts discovered for the Hopf–Lax solution u in Lemmas 3 and 4 can be utilized as uniqueness criteria.

DEFINITION. *We say that a Lipschitz continuous function* $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ *is a* weak solution *of the initial-value problem:*

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\} \end{cases} \tag{38}$$

provided

(a) $u(x,0) = g(x) \quad (x \in \mathbb{R}^n)$,

(b) $u_t(x,t) + H(Du(x,t)) = 0$ *for a.e.* $(x,t) \in \mathbb{R}^n \times (0,\infty)$,

and

(c) $u(x+z,t) - 2u(x,t) + u(x-z,t) \le C\left(1+\frac{1}{t}\right)|z|^2$

for some constant $C \ge 0$ *and all* $x, z \in \mathbb{R}^n$, $t > 0$.

Next we prove that a weak solution of (38) is unique, the key point being that this uniqueness assertion follows from the *inequality* condition (c).

THEOREM 7 (Uniqueness of weak solutions). *Assume H is C^2 and satisfies* (19)*, and g satisfies* (20)*. Then there exists at most one weak solution of the initial-value problem* (38).

Proof*. 1. Suppose that u and $\tilde{u}$ are two weak solutions of (38) and write $w := u - \tilde{u}$.

Observe now at any point (y,s) where both u and $\tilde{u}$ are differentiable and solve our PDE, we have

$$\begin{aligned} w_t(y,s) &= u_t(y,s) - \tilde{u}_t(y,s) \\ &= -H(Du(y,s)) + H(D\tilde{u}(y,s)) \\ &= -\int_0^1 \frac{d}{dr} H(rDu(y,s) + (1-r)D\tilde{u}(y,s))\,dr \\ &= -\int_0^1 DH(rDu(y,s) + (1-r)D\tilde{u}(y,s))\,dr \cdot (Du(y,s) - D\tilde{u}(y,s)) \\ &=: -\mathbf{b}(y,s)\cdot Dw(y,s). \end{aligned}$$

Consequently

$$w_t + \mathbf{b}\cdot Dw = 0 \quad \text{a.e.} \tag{39}$$

*Omit on first reading.

2. Write $v := \phi(w) \geq 0$, where $\phi : \mathbb{R} \to [0, \infty)$ is a smooth function to be selected later. We multiply (39) by $\phi'(w)$ to discover

$$v_t + b \cdot Dv = 0 \quad \text{a.e.} \tag{40}$$

3. Now choose $\varepsilon > 0$ and define $u^\varepsilon := \eta_\varepsilon * u$, $\tilde{u}^\varepsilon := \eta_\varepsilon * \tilde{u}$, where η_ε is the standard mollifier in the x and t variables. Then according to §C.4

$$|Du^\varepsilon| \leq \mathrm{Lip}(u), \ |D\tilde{u}^\varepsilon| \leq \mathrm{Lip}(\tilde{u}), \tag{41}$$

and

$$Du^\varepsilon \to Du, \ D\tilde{u}^\varepsilon \to D\tilde{u} \quad \text{a.e., as } \epsilon \to 0. \tag{42}$$

Furthermore inequality (c) in the definition of weak solution implies

$$D^2u^\varepsilon, D^2\tilde{u}^\varepsilon \leq C\left(1 + \frac{1}{s}\right) I \tag{43}$$

for an appropriate constant C and all $\varepsilon > 0$, $y \in \mathbb{R}^n$, $s > 2\varepsilon$. Verification is left as an exercise.

4. Write

$$\mathbf{b}_\varepsilon(y, s) := \int_0^1 DH(rDu^\varepsilon(y, s) + (1 - r)D\tilde{u}^\varepsilon(y, s))\, dr. \tag{44}$$

Then (40) becomes

$$v_t + \mathbf{b}_\varepsilon \cdot Dv = (\mathbf{b}_\varepsilon - \mathbf{b}) \cdot Dv \quad \text{a.e.;}$$

hence

$$v_t + \operatorname{div}(v\mathbf{b}_\varepsilon) = (\operatorname{div} \mathbf{b}_\varepsilon)v + (\mathbf{b}_\varepsilon - \mathbf{b}) \cdot Dv \quad \text{a.e.} \tag{45}$$

5. Now

$$\begin{aligned}\operatorname{div} \mathbf{b}_\varepsilon &= \int_0^1 \sum_{k,l=1}^n H_{p_k p_l}(rDu^\varepsilon + (1 - r)D\tilde{u}^\varepsilon)(ru^\varepsilon_{x_l x_k} + (1 - r)\tilde{u}^\varepsilon_{x_l x_k})\, dr \\ &\leq C\left(1 + \frac{1}{s}\right) \end{aligned} \tag{46}$$

for some constant C, in view of (41), (43). Here we note that H convex implies $D^2H \geq 0$.

6. Fix $x_0 \in \mathbb{R}^n$, $t_0 > 0$, and set

$$R := \max\{|DH(p)| \mid |p| \leq \max(\mathrm{Lip}(u), \mathrm{Lip}(\tilde{u}))\}. \tag{47}$$

Define also the cone

$$C := \{(x,t) \mid 0 \leq t \leq t_0, |x - x_0| \leq R(t_0 - t)\}.$$

Next write

$$e(t) = \int_{B(x_0, R(t_0-t))} v(x,t)\, dx$$

and compute for a.e. $t > 0$:

$$\begin{aligned}
\dot{e}(t) &= \int_{B(x_0,R(t_0-t))} v_t\, dx - R \int_{\partial B(x_0,R(t_0-t))} v\, dS \\
&= \int_{B(x_0,R(t_0-t))} -\operatorname{div}(v\mathbf{b}_\varepsilon) + (\operatorname{div}\mathbf{b}_\varepsilon)v + (\mathbf{b}_\varepsilon - \mathbf{b}) \cdot Dv\, dx \\
&\quad - R \int_{\partial B(x_0,R(t_0-t))} v\, dS \quad \text{by (45)} \\
&= -\int_{\partial B(x_0,R(t_0-t))} v(\mathbf{b}_\varepsilon \cdot \nu + R)\, dS \\
&\quad + \int_{B(x_0,R(t_0-t))} (\operatorname{div}\mathbf{b}_\varepsilon)v + (\mathbf{b}_\varepsilon - \mathbf{b}) \cdot Dv\, dx \\
&\leq \int_{B(x_0,R(t_0-t))} (\operatorname{div}\mathbf{b}_\varepsilon)v + (\mathbf{b}_\varepsilon - \mathbf{b}) \cdot Dv\, dx \quad \text{by (41), (44)} \\
&\leq C\left(1 + \frac{1}{t}\right) e(t) + \int_{B(x_0,R(t_0-t))} (\mathbf{b}_\varepsilon - \mathbf{b}) \cdot Dv\, dx
\end{aligned}$$

by (46). The last term on the right hand side goes to zero as $\varepsilon \to 0$, for a.e. $t_0 > 0$, according to (41), (42) and the Dominated Convergence Theorem. Thus

$$\dot{e}(t) \leq C\left(1 + \frac{1}{t}\right) e(t) \quad \text{for a.e. } 0 < t < t_0. \tag{48}$$

7. Fix $0 < \varepsilon < r < t$ and choose the function $\phi(z)$ to equal zero if

$$|z| \leq \varepsilon[\mathrm{Lip}(u) + \mathrm{Lip}(\tilde{u})]$$

and to be positive otherwise. Since $u = \tilde{u}$ on $\mathbb{R}^n \times \{t = 0\}$,

$$v = \phi(w) = \phi(u - \tilde{u}) = 0 \quad \text{at } \{t = \varepsilon\}.$$

Thus $e(\varepsilon) = 0$. Consequently Gronwall's inequality (see §B.2) and (48) imply

$$e(r) \le e(\varepsilon)e^{\int_\varepsilon^r C\left(1+\frac{1}{s}\right)ds} = 0.$$

Hence

$$|u - \tilde{u}| \le \varepsilon[\mathrm{Lip}(u) + \mathrm{Lip}(\tilde{u})] \quad \text{on } B(x_0, R(t_0 - r)).$$

This inequality is valid for all $\varepsilon > 0$, and so $u \equiv \tilde{u}$ in $B(x_0, R(t_0 - r))$. Therefore, in particular, $u(x_0, t_0) = \tilde{u}(x_0, t_0)$. □

In light of Lemmas 3, 4 and Theorem 7, we have

THEOREM 8 (Hopf–Lax formula as weak solution). *Suppose H is C^2 and satisfies* (19)*, and g satisfies* (20)*. If either g is semiconcave or H is uniformly convex, then*

$$u(x,t) = \min_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + g(y)\right\}$$

is the unique weak solution of the initial-value problem (38) *for the Hamilton–Jacobi equation.*

Examples. (i) Consider the initial-value problem:

$$(49)\qquad \begin{cases} u_t + \frac{1}{2}|Du|^2 = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = |x| & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Here $H(p) = \frac{1}{2}|p|^2$ and so $L(q) = \frac{1}{2}|q|^2$. The Hopf–Lax formula for the unique, weak solution of (49) is

$$(50)\qquad u(x,t) = \min_{y\in\mathbb{R}^n}\left\{\frac{|x-y|^2}{2t} + |y|\right\}.$$

Assume $|x| > t$. Then

$$D_y\left(\frac{|x-y|^2}{2t} + |y|\right) = \frac{y-x}{t} + \frac{y}{|y|} \quad (y \ne 0);$$

and this expression equals zero if $x = y + \frac{y}{|y|}t$, $y = (|x| - t)\frac{x}{|x|} \ne 0$. Thus $u(x,t) = |x| - \frac{t}{2}$ if $|x| > t$. If $|x| \le t$, the minimum in (50) is attained at $y = 0$. Consequently

$$u(x,t) = \begin{cases} |x| - t/2 & \text{if } |x| \ge t \\ \frac{|x|^2}{2t} & \text{if } |x| \le t. \end{cases}$$

Observe that the solution becomes semiconcave at times $t > 0$, even though the initial function $g(x) = |x|$ is not semiconcave. This accords with Lemma 4.

(ii) We next examine the problem with reversed initial conditions:

$$\begin{cases} u_t + \frac{1}{2}|Du|^2 = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = -|x| & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{51}$$

Then

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ \frac{|x-y|^2}{2t} - |y| \right\}.$$

Now

$$D_y \left(\frac{|x-y|^2}{2t} - |y| \right) = \frac{y-x}{t} - \frac{y}{|y|} \quad (y \neq 0),$$

and this equals zero if $x = y - \frac{y}{|y|}t$, $y = (|x| + t)\frac{x}{|x|}$. Thus

$$u(x,t) = -|x| - \frac{t}{2} \quad (x \in \mathbb{R}^n,\ t \geq 0).$$

The initial function $g(x) = -|x|$ is semiconcave, and the solution remains so for times $t > 0$. □

In Chapter 10 we will again study Hamilton–Jacobi PDE and discover another notion of weak solution, which is applicable even if H is not convex.

3.4. INTRODUCTION TO CONSERVATION LAWS

In this section we investigate the initial-value problem for scalar conservation laws in one space dimension:

$$\begin{cases} u_t + F(u)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{1}$$

Here $F : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are given and $u : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ is the unknown, $u = u(x,t)$. As noted in §3.2, the method of characteristics demonstrates that there does not in general exist a smooth solution of (1), existing for all times $t > 0$. By analogy with the developments in §3.3.5, we therefore look for some sort of weak or generalized solution.

3.4.1. Shocks, entropy condition.

a. Distribution solutions; Rankine–Hugoniot condition.

We open our discussion by noting that since we cannot in general find a smooth solution of (1), we must devise some way to interpret a less regular function u as somehow "solving" this initial-value problem. But as it stands, the PDE does not even make sense unless u is differentiable. However, observe that if we *temporarily* assume u is smooth, we can as follows rewrite, so that the resulting expression does not directly involve the derivatives of u. The idea is to multiply the PDE in (1) by a smooth function v and then to integrate by parts, thereby transferring the derivatives onto v.

More precisely, assume

$$(2)\qquad \begin{cases} v : \mathbb{R} \times [0, \infty) \to \mathbb{R} \text{ is smooth, with} \\ \text{compact support.} \end{cases}$$

We call v a *test function.* Now multiply the PDE $u_t + F(u)_x = 0$ by v and integrate by parts:

$$(3)\qquad \begin{aligned} 0 &= \int_0^\infty \int_{-\infty}^\infty (u_t + F(u)_x)\, v \, dx dt \\ &= - \int_0^\infty \int_{-\infty}^\infty u v_t \; dx dt - \int_{-\infty}^\infty uv\, dx|_{t=0} \\ &\quad - \int_0^\infty \int_{-\infty}^\infty F(u) v_x \, dx dt. \end{aligned}$$

In view of the initial condition $u = g$ on $\mathbb{R} \times \{t = 0\}$, we thereby obtain the identity

$$(4)\qquad \int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \; dx dt + \int_{-\infty}^\infty gv \; dx|_{t=0} = 0.$$

We derived this equality supposing u to be a smooth solution of (1), but the resulting formula has meaning even if u is only bounded.

DEFINITION. *We say that $u \in L^\infty(\mathbb{R} \times (0, \infty))$ is an* integral solution *of* (1), *provided equality* (4) *holds for each test function v satisfying* (2).

Suppose then that we have an integral solution of (1). What can we deduce about this solution from the identities (4)?

We partially answer this question by looking at a situation for which u, although not continuous, has a particularly simple structure. Let us in fact suppose in some open region $V \subset \mathbb{R} \times (0, \infty)$ that u is smooth on either side of a smooth curve C. Let V_l be that part of V on the left of the curve and V_r that part on the right. We assume that u is an integral solution of (1), and that u and its first derivatives are uniformly continuous in V_l and in V_r.

First of all, choose a test function v with compact support in V_l. Then (4) becomes

$$(5) \qquad 0 = \int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \, dxdt = -\int_0^\infty \int_{-\infty}^\infty [u_t + F(u)_x] v \, dxdt,$$

the integration by parts being justified since u is C^1 in V_l and v vanishes near the boundary of V_l. The identity (5) holds for all test functions v with compact support in V_l, and so

$$(6) \qquad u_t + F(u)_x = 0 \quad \text{in } V_l.$$

Likewise,

$$(7) \qquad u_t + F(u)_x = 0 \quad \text{in } V_r.$$

Now select a test function v with compact support in V, but which does not necessarily vanish along the curve C. Again employing (4), we deduce

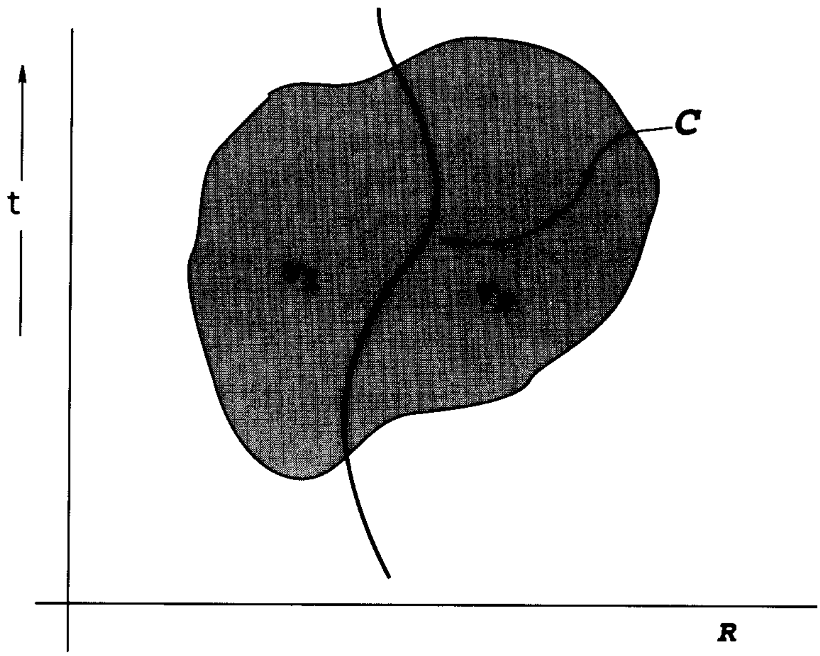
$$(8) \qquad \begin{aligned} 0 &= \int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \, dxdt \\ &= \iint_{V_l} uv_t + F(u)v_x \, dxdt \\ &\quad + \iint_{V_r} uv_t + F(u)v_x \, dxdt. \end{aligned}$$

Now since v has compact support within V, we have

$$(9) \qquad \begin{aligned} \iint_{V_l} uv_t + F(u)v_x \, dxdt &= -\iint_{V_l} [u_t + F(u)_x] v \, dxdt \\ &\quad + \int_C (u_l \nu^2 + F(u_l)\nu^1) v \, dl \\ &= \int_C (u_l \nu^2 + F(u_l)\nu^1) v \, dl \end{aligned}$$

in view of (6). Here $\nu = (\nu^1, \nu^2)$ is the unit normal to the curve C, pointing from V_l into V_r, and the subscript "l" denotes the limit from the left. Similarly, (7) implies

$$\iint_{V_r} uv_t + F(u)v_x \, dxdt = -\int_C (u_r \nu^2 + F(u_r)\nu^1) v \, dl,$$



Rankine-Hugoniot condition

the subscript "r" denoting the limit from the right. Adding this identity to (9) and recalling (8) gives us:

$$\int_C [(F(u_l) - F(u_r))\nu^1 + (u_l - u_r)\nu^2] v \, dl = 0.$$

This equality holds for all test functions v as above, and so

$$(F(u_l) - F(u_r))\nu^1 + (u_l - u_r)\nu^2 = 0 \quad \text{along } C. \tag{10}$$

Now suppose C is represented parametrically as $\{(x,t) \mid x = s(t)\}$ for some smooth function $s(\cdot) : [0,\infty) \to \mathbb{R}$. We can then take $\boldsymbol{\nu} = (\nu^1, \nu^2) = (1 + \dot{s}^2)^{-1/2}(1, -\dot{s})$. Consequently (10) implies

$$F(u_l) - F(u_r) = \dot{s}(u_l - u_r) \tag{11}$$

in V, along the curve C.

Notation.

$$\begin{cases} [[u]] = u_l - u_r = \textit{jump} \text{ in } u \text{ across the curve } C \\ [[F(u)]] = F(u_l) - F(u_r) = \textit{jump} \text{ in } F(u) \\ \sigma = \dot{s} = \textit{speed} \text{ of the curve } C. \end{cases}$$

□

Let us then rewrite (11) as the identity

$$[[F(u)]] = \sigma[[u]] \tag{12}$$

along the discontinuity curve. This is the *Rankine–Hugoniot condition.* Observe that the speed σ and the values $u_l, u_r, F(u_l), F(u_r)$ will generally vary along the curve C. The point is that even though these quantities may change, the expressions $[[F(u)]] = F(u_l) - F(u_r)$ and $\sigma[[u]] = \dot{s}(u_l - u_r)$ must always exactly balance.

Example 1 (Shock waves). Let us consider the initial-value problem for *Burgers' equation*:

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times \{t = 0\}, \end{cases} \tag{13}$$

with the initial data

$$g(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 1 - x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x \geq 1. \end{cases} \tag{14}$$

According to the characteristic equations (cf. §3.2.5) any smooth solution u of (13), (14), takes the constant value $z^0 = g(x^0)$ along the projected characteristic

$$\mathbf{y}(s) = (g(x^0)s + x^0, s) \quad (s \geq 0)$$

for each $x^0 \in \mathbb{R}$. Thus

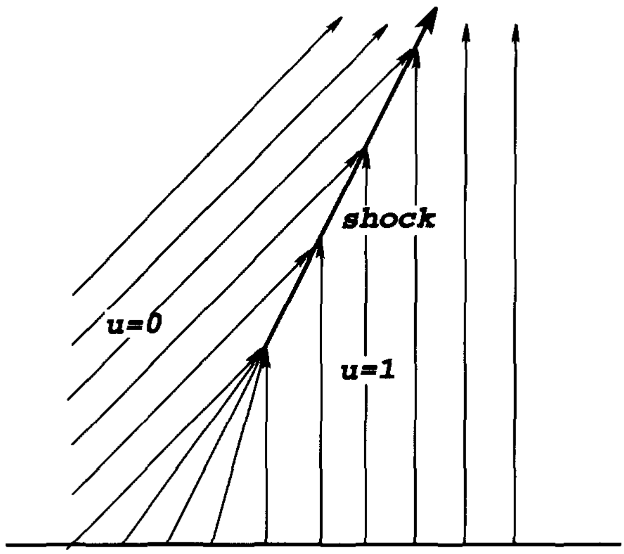
$$u(x,t) := \begin{cases} 1 & \text{if } x \leq t,\ 0 \leq t \leq 1 \\ \frac{1-x}{1-t} & \text{if } t \leq x \leq 1,\ 0 \leq t \leq 1 \\ 0 & \text{if } x \geq 1,\ 0 \leq t \leq 1. \end{cases}$$

Observe that for $t \geq 1$ this method breaks down, since the projected characteristics then cross. So how should we define u for $t \geq 1$?

Let us set $s(t) = \frac{1+t}{2}$, and write

$$u(x,t) := \begin{cases} 1 & \text{if } x < s(t) \\ 0 & \text{if } s(t) < x \end{cases}$$

if $t \geq 1$. Now along the curve parameterized by $s(\cdot)$, $u_l = 1$, we have $u_r = 0$, $F(u_l) = \frac{1}{2}(u_l)^2 = \frac{1}{2}$, $F(u_r) = 0$. Thus $[[F(u)]] = \frac{1}{2} = \sigma[[u]]$, as required by the Rankine–Hugoniot condition (12). □



Formation of a shock

b. Shocks, entropy condition.

We try now to solve a similar problem by the same techniques.

Example 2 (Rarefaction waves and nonphysical shocks). Again consider the initial-value problem (13), for which now we take

$$g(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0. \end{cases} \tag{15}$$

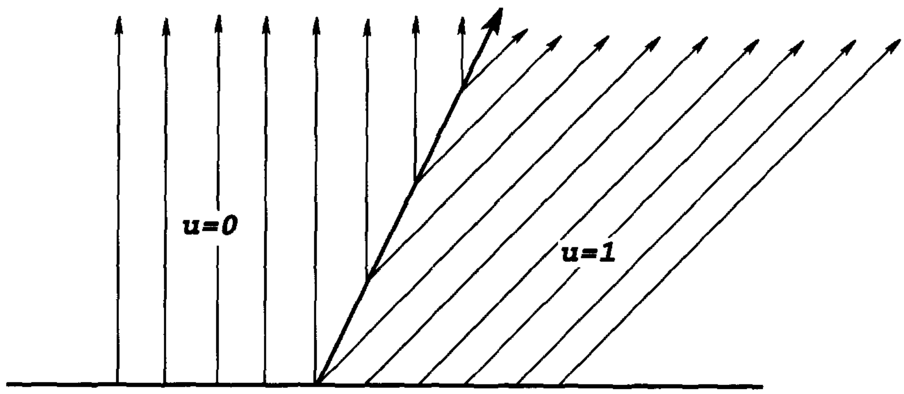
The method of characteristics this time does not lead to any ambiguity in defining u, but does fail to provide any information within the wedge $\{0 < x < t\}$. To illustrate this lack of knowledge, let us first set

$$u_1(x,t) := \begin{cases} 0 & \text{if } x < \frac{t}{2} \\ 1 & \text{if } x > \frac{t}{2}. \end{cases}$$

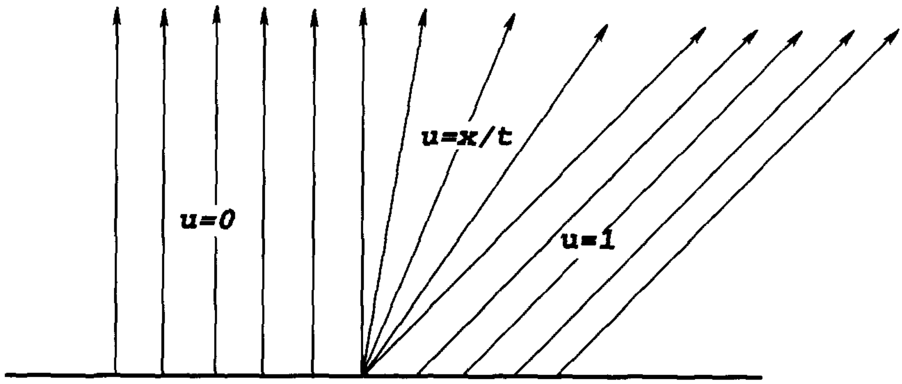
It is easy to check that the Rankine–Hugoniot condition holds and, indeed, that u is an integral solution of (13), (15). However, we can create another such solution by writing

$$u_2(x,t) := \begin{cases} 1 & \text{if } x > t \\ \frac{x}{t} & \text{if } 0 < x < t \\ 0 & \text{if } x < 0. \end{cases}$$

The function u_2, called a *rarefaction wave*, is also a continuous integral solution of (13), (15). □



A "nonphysical" shock



Rarefaction wave

Thus we see that *integral solutions are not in general unique.* Presumably the class of integral solutions includes various "nonphysical" solutions, which we want somehow to exclude. Can we find some further criterion which ensures uniqueness?

Entropy condition. Let us recall from §3.2.5 that for the general scalar conservation law of the form

$$u_t + F(u)_x = 0,$$

the solution u, whenever smooth, takes the constant value $z^0 = g(x^0)$ along the projected characteristic

$$\mathbf{y}(s) = (F'(g(x^0))s + x^0, s) \quad (s \geq 0). \tag{16}$$

Now we know that typically we will encounter the crossing of characteristics, and resultant discontinuities in the solution, if we move *forward* in time. However, we can hope that if we start at some point in $\mathbb{R}^n \times (0, \infty)$ and go *backwards* in time along a characteristic, we will not cross any others. In other words, let us consider the class of, say, piecewise-smooth integral solutions of (1) with the property that if we move backwards in t along any characteristic, we will not encounter any lines of discontinuity for u.

So now suppose at some point on a curve C of discontinuities that u has distinct left and right limits, u_l and u_r, and that a characteristic from the left and a characteristic from the right hit C at this point. Then in view of (16) we deduce

$$F'(u_l) > \sigma > F'(u_r). \tag{17}$$

These inequalities are called the *entropy condition* (from a rough analogy with the thermodynamic principle that physical entropy cannot decrease as time goes forward). A curve of discontinuity for u is called a *shock* provided both the Rankine–Hugoniot identity (12) and the entropy inequalities (17) hold.

Let us further interpret the entropy condition under the additional assumption that

$$F \text{ is uniformly convex.} \tag{18}$$

This means $F'' \geq \theta > 0$ for some constant θ. Thus in particular F' is strictly increasing. Then (17) is equivalent to our requiring the inequality

$$u_l > u_r \tag{19}$$

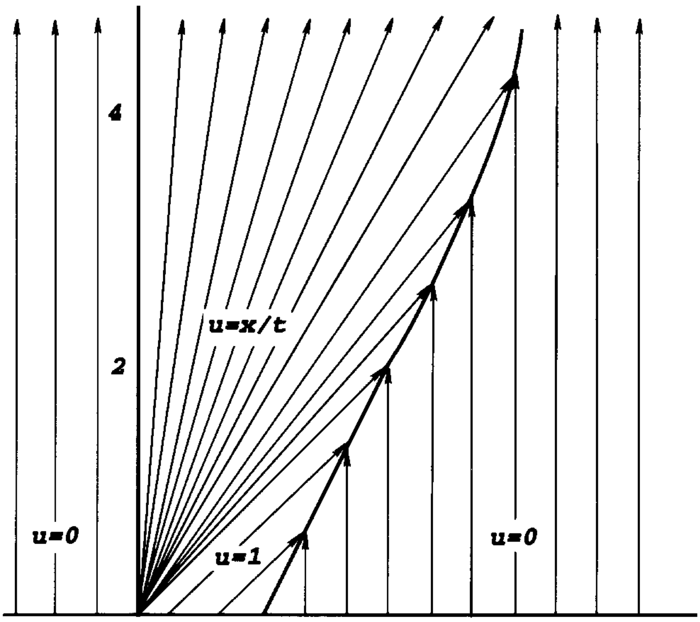
along any shock curve. □

Example 3. We again return to Burgers' equation (13), now for the initial function

$$g(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x > 1. \end{cases} \tag{20}$$

For $0 \leq t \leq 2$, we may combine the analysis in Examples 1 and 2 above to find

$$u(x,t) := \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{t} & \text{if } 0 < x < t \\ 1 & \text{if } t < x < 1 + \frac{t}{2} \\ 0 & \text{if } x > 1 + \frac{t}{2} \end{cases} \qquad (0 \leq t \leq 2). \tag{21}$$



For times $t \geq 2$, we expect the shock wave parameterized by $s(\cdot)$ to continue, with $u = x/t$ to the left of $s(\cdot)$, $u = 0$ to the right. This is compatible with the entropy condition (19). We calculate the behavior of the shock curve by applying the Rankine–Hugoniot jump condition (12). Now

$$[[u]] = \frac{s(t)}{t}, \quad [[F(u)]] = \frac{1}{2}\left(\frac{s(t)}{t}\right)^2, \quad \sigma = \dot{s}(t)$$

along the shock curve for $t \geq 0$. Thus (12) implies

$$s(t) = \frac{s(t)}{2t} \quad (t \geq 2).$$

Additionally $s(2) = 2$, and so we can solve this ODE to find $s(t) = (2t)^{1/2}$ $(t \geq 2)$. Hence we may augment (21) by setting

$$u(x,t) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{t} & \text{if } 0 < x < (2t)^{1/2} \\ 0 & \text{if } x > (2t)^{1/2} \end{cases} \quad (t \geq 2).$$

See the illustration. □

3.4.2. Lax–Oleinik formula.

We now try to obtain a formula for an appropriate weak solution of the initial-value problem (1), assuming as above that the flux function F is uniformly convex. With no loss of generality we may as well also take

$$F(0) = 0. \tag{22}$$

As motivation, suppose now $g \in L^\infty(\mathbb{R})$ and define

$$h(x) := \int_0^x g(y)\,dy \quad (x \in \mathbb{R}). \tag{23}$$

Recall the Hopf–Lax formula from §3.3 and set

$$w(x,t) := \min_{y\in\mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\} \quad (x \in \mathbb{R},\ t > 0), \tag{24}$$

where

$$L = F^*. \tag{25}$$

Thus w is the unique, weak solution of this initial-value problem for the Hamilton–Jacobi equation:

$$\begin{cases} w_t + F(w_x) = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ w = h & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{26}$$

For the moment assume w is smooth. We now differentiate the PDE and its initial condition with respect to x, to deduce

$$\begin{cases} w_{xt} + F(w_x)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ w_x = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases}$$

Hence if we set $u = w_x$, we discover u solves problem (1).

The foregoing computation is only formal, as we know that w defined by (24) is not in general smooth. But recall from §3.3 that w is in fact differentiable a.e. Consequently

$$u(x,t) := \frac{\partial}{\partial x}\left[\min_{y\in\mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\} \right] \tag{27}$$

is defined for a.e. (x,t) and is presumably a leading candidate for some sort of weak solution of the initial-value problem (1). Our intention henceforth is to justify this expectation.

First, we will need to rewrite the expression (27) into a more useful form.

Notation. Since F is uniformly convex, F' is strictly increasing and onto. Write

$$G := (F')^{-1} \tag{28}$$

for the inverse of F'. □

THEOREM 1 (Lax–Oleinik formula). *Assume $F : \mathbb{R} \to \mathbb{R}$ is smooth, uniformly convex, and $g \in L^\infty(\mathbb{R})$.*

(i) *For each time $t > 0$, there exists for all but at most countably many values of $x \in \mathbb{R}$ a unique point $y(x,t)$ such that*

$$\min_{y\in\mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\} = tL\left(\frac{x-y(x,t)}{t}\right) + h(y(x,t)).$$

(ii) *The mapping $x \mapsto y(x,t)$ is nondecreasing.*

(iii) *For each time $t > 0$, the function u defined by (27) is*

$$u(x,t) = G\left(\frac{x-y(x,t)}{t}\right) \tag{29}$$

for a.e. x. In particular, the formula (29) holds for a.e. $(x,t) \in \mathbb{R}^n \times (0,\infty)$.

DEFINITION. *We call equation* (29) *the* Lax–Oleinik formula *for the solution* (1)*, where h is defined by* (23)*, L by* (25).

Proof. 1. First, we note

$$L(q) = \max_{p\in\mathbb{R}} (qp - F(p)) = qp^* - F(p^*),$$

where $F'(p^*) = q$. But then $p^* = G(q)$ according to (28), and so

$$L(q) = qG(q) - F(G(q)) \quad (q \in \mathbb{R})$$

(cf. §3.3.1). In particular, L is C^2. Furthermore

$$L'(q) = G(q) + qG'(q) - F'(G(q))G'(q) = G(q) \tag{30}$$

by (28); and $L''(q) = G'(q) > 0$. This and (22) imply L is nonnegative and strictly convex.

2. Fix $t > 0$, $x_1 < x_2$. As in §3.3 there exists at least one point $y_1 \in \mathbb{R}$ such that

$$\left\{ tL\left(\frac{x_1-y_1}{t}\right) + h(y_1) \right\} = \min_{y\in\mathbb{R}} \left\{ tL\left(\frac{x_1-y}{t}\right) + h(y) \right\}. \tag{31}$$

We next claim

$$tL\left(\frac{x_2-y_1}{t}\right) + h(y_1) < tL\left(\frac{x_2-y}{t}\right) + h(y) \quad \text{if } y < y_1. \tag{32}$$

To see this, we calculate $x_2 - y_1 = \tau(x_1 - y_1) + (1-\tau)(x_2 - y)$ and $x_1 - y = (1-\tau)(x_1 - y_1) + \tau(x_2 - y)$ for

$$0 < \tau := \frac{y_1 - y}{x_2 - x_1 + y_1 - y} < 1.$$

Since $L'' > 0$, we thus have

$$L\left(\frac{x_2 - y_1}{t}\right) < \tau L\left(\frac{x_1 - y_1}{t}\right) + (1-\tau) L\left(\frac{x_2 - y}{t}\right),$$
$$L\left(\frac{x_1 - y}{t}\right) < (1-\tau) L\left(\frac{x_1 - y_1}{t}\right) + \tau L\left(\frac{x_2 - y}{t}\right);$$

and hence

$$L\left(\frac{x_2 - y_1}{t}\right) + L\left(\frac{x_1 - y}{t}\right) < L\left(\frac{x_1 - y_1}{t}\right) + L\left(\frac{x_2 - y}{t}\right). \tag{33}$$

Now notice from (31) that

$$tL\left(\frac{x_1 - y_1}{t}\right) + h(y_1) \le tL\left(\frac{x_1 - y}{t}\right) + h(y).$$

We multiply (33) by t, add $h(y_1) + h(y)$ to both sides, and add the resulting expression to the above inequality to obtain (32).

3. In view of (31), in computing the minimum of $tL\left(\frac{x_2 - y}{t}\right) + h(y)$ we need only consider those $y \ge y_1$, where y_1 satisfies (31). Now for each $x \in \mathbb{R}$ and $t > 0$, define the point $y(x,t)$ to equal the smallest of those points y giving the minimum of $tL\left(\frac{x-y}{t}\right) + h(y)$. Then the mapping $x \mapsto y(x,t)$ is nondecreasing and is thus continuous for all but at most countably many x. At a point x of continuity of $y(\cdot, t)$, $y(x,t)$ is the unique value of y yielding the minimum.

4. According to the theory developed in §3.3 for each fixed $t > 0$, the mapping

$$\begin{aligned} x \mapsto w(x,t) &:= \min_{y \in \mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\} \\ &= tL\left(\frac{x - y(x,t)}{t}\right) + h(y(x,t)) \end{aligned}$$

is differentiable a.e. Furthermore the mapping $x \mapsto y(x,t)$ is monotone and consequently differentiable a.e. as well. Thus given $t > 0$, for a.e. x the mappings $x \mapsto L(\frac{x - y(x,t)}{t})$ and so also $x \mapsto h(y(x,t))$ are differentiable as well.

Consequently formula (27) becomes

$$u(x,t) = \frac{\partial}{\partial x}\left[tL\left(\frac{x-y(x,t)}{t}\right) + h(y(x,t))\right]$$
$$= L'\left(\frac{x-y(x,t)}{t}\right)(1-y_x(x,t)) + \frac{\partial}{\partial x}h(y(x,t)).$$

But since $y \mapsto tL\left(\frac{x-y}{t}\right) + h(y)$ has a minimum at $y = y(x,t)$, the mapping $z \mapsto tL\left(\frac{x-y(z,t)}{t}\right) + h(y(z,t))$ has a minimum at $z = x$. Therefore

$$-L'\left(\frac{x-y(x,t)}{t}\right)y_x(x,t) + \frac{\partial}{\partial x}h(y(x,t)) = 0,$$

and hence

$$u(x,t) = L'\left(\frac{x-y(x,t)}{t}\right) = G\left(\frac{x-y(x,t)}{t}\right),$$

according to (30). □

We now investigate the precise sense in which formula (29) provides us with a solution of the initial-value problem (1).

THEOREM 2 (Lax–Oleinik formula as integral solution). *Under the assumptions of Theorem* 1, *the function* u *defined by* (29) *is an integral solution of the initial-value problem* (1).

Proof. As above, define

$$w(x,t) = \min_{y\in\mathbb{R}}\left\{tL\left(\frac{x-y}{t}\right) + h(y)\right\} \quad (x \in \mathbb{R},\ t > 0).$$

Then Theorem 6 in §3.3.2 tells us w is Lipschitz continuous, is differentiable for a.e. (x,t), and solves

$$\text{(34)} \qquad \begin{cases} w_t + F(w_x) = 0 & \text{a.e. in } \mathbb{R}\times(0,\infty) \\ w = h & \text{on } \mathbb{R}\times\{t=0\}. \end{cases}$$

Choose any test function v satisfying (2). Multiply the PDE $w_t + F(w_x) = 0$ by v_x and integrate over $\mathbb{R}\times(0,\infty)$:

$$\text{(35)} \qquad 0 = \int_0^\infty\int_{-\infty}^\infty [w_t + F(w_x)]\, v_x \, dxdt.$$

Observe

$$\int_0^\infty \int_{-\infty}^\infty w_t v_x \, dxdt = -\int_0^\infty \int_{-\infty}^\infty w v_{tx} \, dxdt - \int_{-\infty}^\infty w v_x \, dx|_{t=0}$$
$$= \int_0^\infty \int_{-\infty}^\infty w_x v_t \, dxdt + \int_{-\infty}^\infty w_x \, dx|_{t=0}.$$

These integrations by parts are valid since the mapping $x \mapsto w(x,t)$ is Lipschitz continuous, and thus absolutely continuous, for each time $t > 0$. Likewise $t \mapsto w(x,t)$ is absolutely continuous for each $x \in \mathbb{R}$. Now $w(x,0) = h(x) = \int_0^x g(y)\,dy$, and so $w_x(x,0) = g(x)$ for a.e. x. Consequently

$$\int_0^\infty \int_{-\infty}^\infty w_t v_x \, dxdt = \int_0^\infty \int_{-\infty}^\infty w_x v_t \, dxdt + \int_{-\infty}^\infty g v \, dx|_{t=0}.$$

Substitute this identity into (35) and recall $u = w_x$ a.e., to derive the integral identity (4). □

3.4.3. Weak solutions, uniqueness.

a. Entropy condition revisited.

We have already seen in §3.4.1 that integral solutions of (1) are not generally unique. Since we believe the Lax–Oleinik formula does in fact provide the "correct" solution of this initial-value problem, we must see if it satisfies some appropriate form of the entropy condition discussed in §3.4.1. This is not straightforward, however, since it is not usually the case that the function u defined by the Lax–Oleinik formula is smooth, or even piecewise smooth.

We identify now a kind of "one-sided" derivative estimate for the function u defined by the Lax–Oleinik formula (27). This estimate—which is an analogue for conservation laws of the semiconcavity estimate from Lemmas 3, 4 in §3.3.3 for Hamilton–Jacobi equations—will turn out to be a uniqueness criterion.

LEMMA (A one-sided jump estimate). *Under the assumptions of Theorem 1, there exists a constant C such that the function u defined by the Lax–Oleinik formula (29) satisfies the inequality*

$$u(x+z,t) - u(x,t) \leq \frac{C}{t} z \tag{36}$$

for all $t > 0$ and $x, z \in \mathbb{R}$, $z > 0$.

DEFINITION. *We call inequality* (36) *the* entropy condition.

It follows from (36) that for $t > 0$ the function $x \mapsto u(x,t) - \frac{C}{t}x$ is nonincreasing, and consequently has left and right hand limits at each point. Thus also $x \mapsto u(x,t)$ has left and right hand limits at each point, with $u_l(x,t) \geq u_r(x,t)$. In particular, the original form of the entropy condition (19) holds at any point of discontinuity.

Proof. We know from §3.3 that in computing the minimum in (29) we need only consider those y such that $\left|\frac{x-y}{t}\right| \leq C$ for some constant C; verification is left to the reader. Consequently we may assume, upon redefining G if necessary off some bounded interval, that G is Lipschitz continuous.

2. As $G = (F')^{-1}$ and $y(\cdot,t)$ are nondecreasing, we have

$$\begin{aligned} u(x,t) &= G\left(\frac{x - y(x,t)}{t}\right) \\ &\geq G\left(\frac{x - y(x+z,t)}{t}\right) \quad \text{for } z > 0 \\ &\geq G\left(\frac{x + z - y(x+z,t)}{t}\right) - \frac{\operatorname{Lip}(G)z}{t} \\ &= u(x+z,t) - \frac{\operatorname{Lip}(G)z}{t}. \end{aligned}$$

□

b. Weak solutions, uniqueness.

We now establish the important assertion that an integral solution which satisfies the entropy condition is unique.

DEFINITION. *We say that a function* $u \in L^\infty(\mathbb{R} \times (0,\infty))$ *is an* entropy solution *of the initial-value problem*

$$(37) \qquad \begin{cases} u_t + F(u)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\} \end{cases}$$

provided

$$\text{(i)} \qquad \int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \, dxdt + \int_{-\infty}^\infty gv \, dx|_{t=0} = 0$$

for all test functions $v : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ *with compact support, and*

$$\text{(ii)} \qquad u(x+z,t) - u(x,t) \leq C\left(1 + \frac{1}{t}\right)z$$

for some constant $C \geq 0$ *and a.e.* $x, z \in \mathbb{R}$, $t > 0$, *with* $z > 0$.

THEOREM 3 (Uniqueness of entropy solutions). *Assume F is convex and smooth. Then there exists—up to a set of measure zero—at most one entropy solution of* (37).

Proof*. 1. Assume that u and $\tilde{u}$ are two entropy solutions of (37), and write $w := u - \tilde{u}$. Observe for any point (x, t) that

$$\begin{aligned} F(u(x,t)) - F(\tilde{u}(x,t)) &= \int_0^1 \frac{d}{dr} F(ru(x,t) + (1-r)\tilde{u}(x,t))\,dr \\ &= \int_0^1 F'(ru(x,t) + (1-r)\tilde{u}(x,t))\,dr\,(u(x,t) - \tilde{u}(x,t)) \\ &=: b(x,t)w(x,t). \end{aligned}$$

Consequently if v is a test function as above,

$$\begin{aligned} 0 &= \int_0^\infty \int_{-\infty}^\infty (u - \tilde{u})v_t + [F(u) - F(\tilde{u})]v_x\,dxdt \\ &= \int_0^\infty \int_{-\infty}^\infty w[v_t + bv_x]\,dxdt. \end{aligned} \tag{38}$$

2. Now take $\varepsilon > 0$ and define $u^\varepsilon = \eta_\varepsilon * u$, $\tilde{u}^\varepsilon = \eta_\varepsilon * \tilde{u}$, where η_ε is the standard mollifier in the x and t variables. Then according to §C.4

$$\|u^\varepsilon\|_{L^\infty} \le \|u\|_{L^\infty},\ \|\tilde{u}^\varepsilon\|_{L^\infty} \le \|\tilde{u}\|_{L^\infty}, \tag{39}$$

$$u^\varepsilon \to u,\ \tilde{u}^\varepsilon \to \tilde{u} \quad \text{a.e., as } \epsilon \to 0. \tag{40}$$

Furthermore the entropy inequality (ii) implies

$$u^\varepsilon_x(x,t),\ \tilde{u}^\varepsilon_x(x,t) \le C\left(1 + \frac{1}{t}\right) \tag{41}$$

for an appropriate constant C and all $\varepsilon > 0$, $x \in \mathbb{R}$, $t > 0$.

3. Write

$$b_\varepsilon(x,t) := \int_0^1 F'(ru^\varepsilon(x,t) + (1-r)\tilde{u}^\varepsilon(x,t))\,dr.$$

Then (38) becomes

$$0 = \int_0^\infty \int_{-\infty}^\infty w[v_t + b_\varepsilon v_x]\,dxdt + \int_0^\infty \int_{-\infty}^\infty w[b - b_\varepsilon]v_x\,dxdt. \tag{42}$$

*Omit on first reading.

4. Now select $T > 0$ and any smooth function $\psi : \mathbb{R} \times (0,T) \to \mathbb{R}$ with compact support. We choose v to be the solution of the following terminal-value problem for a linear transport equation:

$$\begin{cases} v_t^\varepsilon + b_\varepsilon v_x^\varepsilon = \psi & \text{in } \mathbb{R} \times (0,T) \\ v = 0 & \text{on } \mathbb{R} \times \{t = T\}. \end{cases} \tag{43}$$

Let us solve (43) by the method of characteristics. For this, fix $x \in \mathbb{R}$, $0 \le t \le T$, and denote by $x_\varepsilon(\cdot)$ the solution of the ODE

$$\begin{cases} \dot{x}_\varepsilon(s) = b_\varepsilon(x_\varepsilon(s), s) \quad (s \ge t) \\ x_\varepsilon(t) = x, \end{cases} \tag{44}$$

and set

$$v^\varepsilon(x,t) := -\int_t^T \psi(x_\varepsilon(s), s)\, ds \quad (x \in \mathbb{R},\ 0 \le t \le T). \tag{45}$$

Then v^ε is smooth and is the unique solution of (43). Since $|b_\varepsilon|$ is bounded and ψ has compact support, v^ε has compact support in $\mathbb{R} \times [0,T)$.

5. We now claim that for each $s > 0$, there exists a constant C_s such that

$$|v_x^\varepsilon| \le C_s \quad \text{on } \mathbb{R} \times (s,T). \tag{46}$$

To prove this, first note that if $0 < s \le t \le T$, then

$$\begin{aligned} b_{\varepsilon,x}(x,t) &= \int_0^1 F''(ru^\varepsilon + (1-r)\tilde{u}^\varepsilon)(ru_x^\varepsilon + (1-r)\tilde{u}_x^\varepsilon)\, dr \\ &\le \frac{C}{t} \le \frac{C}{s} \end{aligned} \tag{47}$$

by (41), since F is convex.

Next, differentiate the PDE in (43) with respect to x:

$$v_{tx}^\varepsilon + b_\varepsilon v_{xx}^\varepsilon + b_{\varepsilon,x} v_x^\varepsilon = \psi_x. \tag{48}$$

Now set $a(x,t) := e^{\lambda t} v_x^\varepsilon(x,t)$, for

$$\lambda = \frac{C}{s} + 1. \tag{49}$$

Then

$$\begin{aligned} a_t + b_\varepsilon a_x &= \lambda a + e^{\lambda t}[v_{xt}^\varepsilon + b_\varepsilon v_{xx}] \\ &= \lambda a + e^{\lambda t}[-b_{\varepsilon,x} v_x^\varepsilon + \psi_x] \quad \text{by (48)} \\ &= [\lambda - b_{\varepsilon,x}]a + e^{\lambda t}\psi_x. \end{aligned} \tag{50}$$

Since v^ε has compact support, a attains a nonnegative maximum over $\mathbb{R} \times [s, T]$ at some finite point (x_0, t_0). If $t_0 = T$, then $v_x = 0$. If $0 \leq t_0 < T$, then

$$a_t(x_0, t_0) \leq 0, \ a_x(x_0, t_0) = 0.$$

Consequently equation (50) gives

$$[\lambda - b_{\varepsilon,x}]a + e^{\lambda t_0}\psi_x \leq 0 \quad \text{at } (x_0, t_0). \tag{51}$$

But since $b_{\varepsilon,x} \leq \frac{C}{s}$ and λ is given by (49), inequality (51) implies

$$a(x_0, t_0) \leq -e^{\lambda t_0}\psi_x \leq e^{\lambda T}\|\psi_x\|_{L^\infty}.$$

A similar argument shows

$$a(x_1, t_1) \geq -e^{\lambda T}\|\psi_x\|_{L^\infty}$$

at any point (x_1, t_1) where a attains a nonpositive minimum. These two estimates and the definition of a imply (46).

6. We will need one more inequality, namely

$$\int_{-\infty}^{\infty} |v_x^\varepsilon(x, t)|\, dx \leq D \tag{52}$$

for all $0 \leq t \leq \tau$ and some constant D, provided τ is small enough.

To prove this, choose $\tau > 0$ so small that $\psi = 0$ on $\mathbb{R} \times (0, \tau)$. Then if $0 \leq t \leq \tau$, we see from (45) that v is constant along the characteristic curve $x_\varepsilon(\cdot)$ (solving (44)) for $t \leq s \leq \tau$. Select any partition $x_0 < x_1 < \cdots < x_N$. Then $y_0 < y_1 < \cdots < y_N$, where $y_i := x_i(s)$ $(i = 1, \ldots, N)$ for

$$\begin{cases} \dot{x}_\varepsilon(s) = b_\varepsilon(x_\varepsilon(s), s) & (t \leq s \leq \tau) \\ x_\varepsilon(t) = x_i. \end{cases}$$

As v^ε is constant along each characteristic curve $x_i(\cdot)$, we have

$$\begin{aligned} \sum_{i=1}^{N} |v^\varepsilon(x_i, t) - v^\varepsilon(x_{i-1}, t)| &= \sum_{i=1}^{N} |v^\varepsilon(y_i, \tau) - v^\varepsilon(y_{i-1}, \tau)| \\ &\leq \operatorname{var} v^\varepsilon(\cdot, \tau), \end{aligned}$$

"var" denoting variation with respect to x. Taking the supremum over all such partitions, we find

$$\int_{-\infty}^{\infty} |v_x^\varepsilon(x, t)|\, dx = \operatorname{var} v^\varepsilon(\cdot, t) \leq \operatorname{var} v^\varepsilon(\cdot, \tau) = \int_{-\infty}^{\infty} |v_x^\varepsilon(x, \tau)|\, dx \leq C,$$

since v^ε has constant support and estimate (41) is valid for $s = \tau$.

7. Now, at last, we complete the proof by setting $v = v^\varepsilon$ in (42) and substituting, using (43):

$$\begin{aligned}
\int_0^\infty \int_{-\infty}^\infty w\psi \, dxdt &= \int_0^\infty \int_{-\infty}^\infty w[b_\varepsilon - b]v_x^\varepsilon \, dxdt \\
&= \int_\tau^T \int_{-\infty}^\infty w[b_\varepsilon - b]v_x^\varepsilon \, dxdt \\
&\quad + \int_0^\tau \int_{-\infty}^\infty w[b_\varepsilon - b]v_x^\varepsilon \, dxdt \\
&=: I_\tau^\varepsilon + J_\tau^\varepsilon.
\end{aligned}$$

Then in view of (40), (46), and the Dominated Convergence Theorem,

$$I_\tau^\varepsilon \to 0 \quad \text{as } \varepsilon \to 0$$

for each $\tau > 0$. On the other hand, if $0 < \tau < T$, we see

$$|J_\tau^\varepsilon| \leq \tau C \max_{0 \leq t \leq \tau} \int_{-\infty}^\infty |v_x^\varepsilon| \, dx \leq \tau C, \qquad \text{by (52).}$$

Thus

$$\int_0^\infty \int_{-\infty}^\infty w\psi \, dxdt = 0$$

for all smooth functions ψ as above, and so $w = u - \tilde{u} = 0$ a.e. □

3.4.4. Riemann's problem.

The initial-value problem (1) with the piecewise-constant initial function

$$(53) \qquad g(x) = \begin{cases} u_l & \text{if } x < 0 \\ u_r & \text{if } x > 0 \end{cases}$$

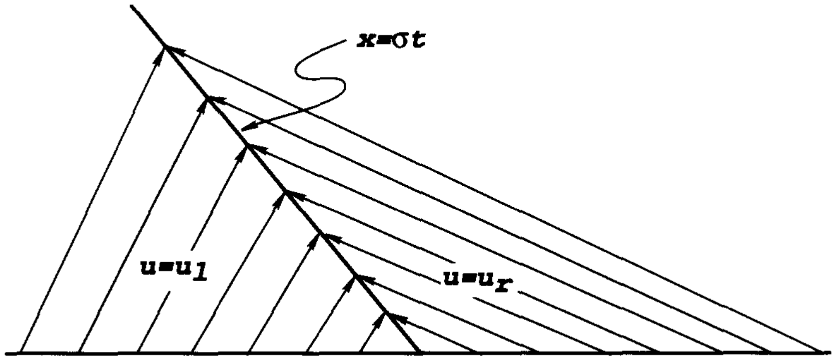
is called *Riemann's problem* for the scalar conservation law (1). Here $u_l, u_r \in \mathbb{R}$ are the left and right *initial states*, $u_l \neq u_r$

We continue to assume F is uniformly convex and C^2, and as before we write $G = (F')^{-1}$.

THEOREM 4 (Solution of Riemann's problem).

(i) *If $u_l > u_r$, the unique entropy solution of the Riemann problem* (1), (53) *is*

$$(54) \qquad u(x,t) := \begin{cases} u_l & \text{if } \frac{x}{t} < \sigma \\ u_r & \text{if } \frac{x}{t} > \sigma \end{cases} \qquad (x \in \mathbb{R},\ t > 0),$$



Shock wave solving Riemann's problem for $u_l > u_r$

where

$$\sigma := \frac{F(u_l) - F(u_r)}{u_l - u_r}. \tag{55}$$

(ii) *If $u_l < u_r$, the unique entropy solution of the Riemann problem* (1), (53) *is*

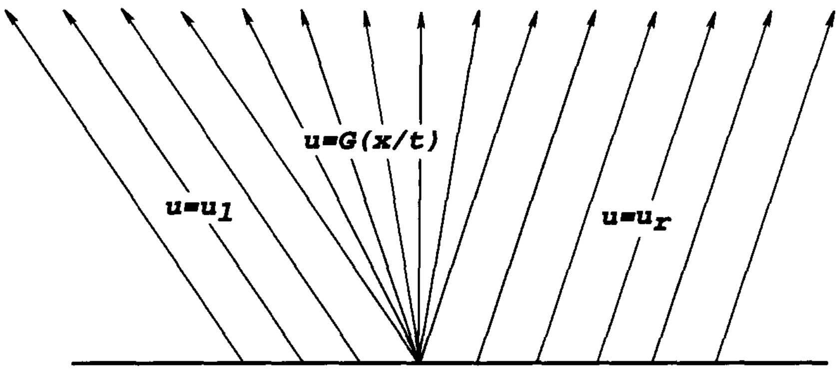
$$u(x,t) := \begin{cases} u_l & \text{if } \frac{x}{t} < F'(u_l) \\ G\left(\frac{x}{t}\right) & \text{if } F'(u_l) < \frac{x}{t} < F'(u_r) \\ u_r & \text{if } \frac{x}{t} > F'(u_r) \end{cases} \quad (x \in \mathbb{R},\ t > 0). \tag{56}$$

Remarks. (i) In the first case the states u_l and u_r are separated by a *shock wave* with constant speed σ. In the second case the states u_l and u_r are separated by a *rarefaction wave.*

(ii) We know from the theory set forth in §§3.4.2–3.4.3 that the Lax–Oleinik formula must generate these solutions, and it is an interesting exercise to verify this directly. We will instead construct the functions (54), (56) from first principles and verify they are in fact entropy solutions. By uniqueness, then, they must agree with Lax–Oleinik formulas. This is a nice illustration of the power of the uniqueness assertion, Theorem 3. □

Proof. 1. Assume $u_l > u_r$. Clearly u defined by (54), (55) is then an integral solution of our PDE. In particular since $\sigma = [[F(u)]]/[[u]]$, the Rankine–Hugoniot condition holds. Furthermore note

$$F'(u_r) < \sigma = \frac{F(u_l) - F(u_r)}{u_l - u_r} = \fint_{u_r}^{u_l} F'(r)dr < F'(u_l)$$



Rarefaction wave solving Riemann's problem for $u_l < u_r$

in accordance with (17). Since $u_l > u_r$, the entropy condition holds as well. Uniqueness follows from Theorem 3.

2. Assume now that $u_l < u_r$. We must first check that u defined by (56) solves the conservation law in the region $\{F'(u_l) < \frac{x}{t} < F'(u_r)\}$. To verify this, let us ask the general question as to when a function u of the form

$$u(x,t) = v\left(\frac{x}{t}\right)$$

solves (1). We compute

$$\begin{aligned} u_t + F(u)_x &= u_t + F'(u)u_x \\ &= -v'\left(\frac{x}{t}\right)\frac{x}{t^2} + F'(v)v'\left(\frac{x}{t}\right)\frac{1}{t} \\ &= v'\left(\frac{x}{t}\right)\frac{1}{t}\left[F'(v) - \frac{x}{t}\right]. \end{aligned}$$

Thus, assuming v' never vanishes, we find $F'\left(v(\frac{x}{t})\right) = \frac{x}{t}$. Hence

$$u(x,t) = v\left(\frac{x}{t}\right) = G\left(\frac{x}{t}\right)$$

solves the conservation law. Now $v(\frac{x}{t}) = u_l$ provided $\frac{x}{t} = F'(u_l)$; and similarly $v(\frac{x}{t}) = u_r$ if $\frac{x}{t} = F'(u_r)$.

As a consequence we see that the rarefaction wave u defined by (56) is continuous in $\mathbb{R} \times (0,\infty)$, and is a solution of the PDE $u_t + F(u)_x = 0$ in each of its regions of definition. It is easy to check that u is thus an integral solution of (1), (53). Furthermore, since as noted in §3.4.3 we may as well assume G is Lipschitz continuous, we have

$$u(x+z,t) - u(x,t) = G\left(\frac{x+z}{t}\right) - G\left(\frac{x}{t}\right) \leq \frac{\text{Lip}(G)z}{t}$$

if $F'(u_l)t < x < x+z < F'(u_r)t$. This inequality implies that u also satisfies the entropy condition. Uniqueness is once more a consequence of Theorem 3.

□

3.4.5. Long time behavior.

a. Decay in sup-norm.

We now employ the Lax–Oleinik formula (29) to study the behavior of our entropy solution u of (1) as $t \to \infty$. We assume below that F is smooth, uniformly convex, $F(0) = 0$, and g is bounded and summable.

THEOREM 5 (Asymptotics in L^∞-norm). *There exists a constant C such that*

$$|u(x,t)| \leq \frac{C}{t^{1/2}} \tag{57}$$

for all $x \in \mathbb{R}$, $t > 0$.

Proof. 1. Set

$$\sigma := F'(0); \tag{58}$$

then

$$G(\sigma) = 0, \tag{59}$$

and therefore

$$L(\sigma) = \sigma G(\sigma) - F(G(\sigma)) = 0, \ L'(\sigma) = 0. \tag{60}$$

2. In view of (60) and the uniform convexity of L,

$$\begin{aligned} tL\left(\frac{x-y}{t}\right) &= tL\left(\frac{x-y-\sigma t}{t} + \sigma\right) \\ &\geq t\left[L(\sigma) + L'(\sigma)\left(\frac{x-y-\sigma t}{t}\right) + \theta\left(\frac{x-y-\sigma t}{t}\right)^2\right] \\ &= \theta\frac{|x-y-\sigma t|^2}{t} \end{aligned} \tag{61}$$

for some constant $\theta > 0$. Since $h = \int_0^x g\,dy$ is bounded by $M := \|g\|_{L^1}$, we see from (61) that

$$tL\left(\frac{x-y}{t}\right) + h(y) \geq \theta\frac{|x-y-\sigma t|^2}{t} - M.$$

On the other hand,

$$tL\left(\frac{x-(x-\sigma t)}{t}\right)+h(x-\sigma t)\leq M.$$

Thus at the minimizing point $y(x,t)$ we have

$$\theta\frac{|x-y(x,t)-\sigma t|^2}{t}\leq 2M;$$

and so

$$\left|\frac{x-y(x,t)}{t}-\sigma\right|\leq\frac{C}{t^{1/2}} \tag{62}$$

for some constant C.

3. But since $G(\sigma)=0$, for any $x\in\mathbb{R}$, $t>0$ we have

$$\begin{aligned}|u(x,t)|&=\left|G\left(\frac{x-y(x,t)}{t}\right)\right|\\&=\left|G\left(\frac{x-y(x,t)}{t}-\sigma+\sigma\right)-G(\sigma)\right|\\&\leq\operatorname{Lip}(G)\left|\frac{x-y(x,t)}{t}-\sigma\right|\leq\frac{C}{t^{1/2}},\end{aligned}$$

according to (62). □

Example 3 in §3.4.1 shows this $t^{-1/2}$ decay rate to be optimal.

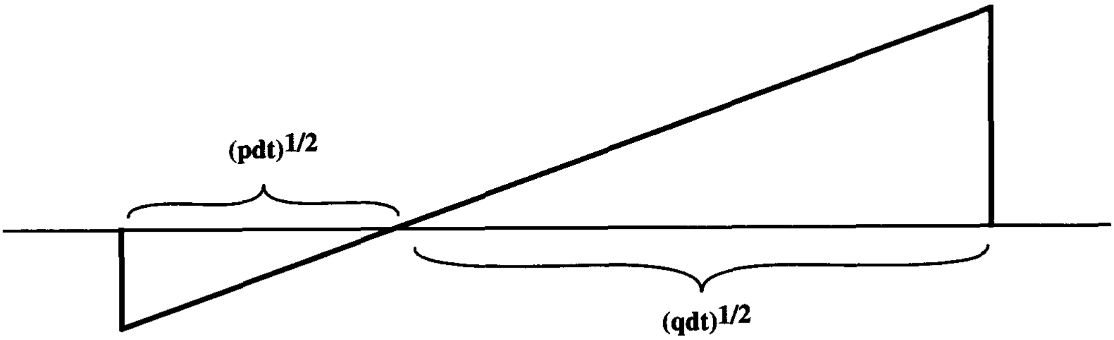
b. Decay to N-wave.

Estimate (57) asserts that the L^∞-norm of u goes to zero as $t\to\infty$. On the other hand we note from Example 3 in §3.4.1 that the L^1-norm of u need not go to zero; indeed, the integral of u over $\mathbb{R}$ is conserved (Problem 13). We instead show here that u evolves in L^1 into a simple shape, assuming now that

$$g \text{ has compact support.}$$

Given constants p,q,d,σ, with $p,q\geq 0$, $d>0$, we define the corresponding *N-wave* to be the function

$$N(x,t):=\begin{cases}\frac{1}{d}\left(\frac{x}{t}-\sigma\right) & \text{if } -(pdt)^{1/2}<x-\sigma t<(qdt)^{1/2}\\ 0 & \text{otherwise.}\end{cases} \tag{63}$$



N-**wave**

The constant σ is the *velocity* of the N-wave.

Now define σ by (58), set

$$d := F''(0) > 0, \tag{64}$$

and also write

$$p := -2 \min_{y\in\mathbb{R}} \int_{-\infty}^{y} g\,dx, \quad q := 2 \max_{y\in\mathbb{R}} \int_{y}^{\infty} g\,dx. \tag{65}$$

Note $p, q \geq 0$ and

$$G'(\sigma) = \frac{1}{d}. \tag{66}$$

THEOREM 6 (Asymptotics in L^1-norm). *Assume that $p, q > 0$. Then there exists a constant C such that*

$$\int_{-\infty}^{\infty} |u(\cdot,t) - N(\cdot,t)|\,dx \leq \frac{C}{t^{1/2}} \tag{67}$$

for all $t > 0$.

Proof. 1. From estimate (62) in the proof of Theorem 5 we have

$$\left| \frac{(x-\sigma t) - y(x,t)}{t} \right| \leq \frac{C}{t^{1/2}}. \tag{68}$$

Now

$$\begin{aligned} u(x,t) &= G\left(\frac{x - y(x,t)}{t}\right) \\ &= G\left(\frac{(x-\sigma t) - y(x,t)}{t} + \sigma\right) \\ &= G(\sigma) + G'(\sigma)\left(\frac{(x-\sigma t) - y(x,t)}{t}\right) \\ &\quad + O\left(\left|\frac{(x-\sigma t) - y(x,t)}{t}\right|^2\right). \end{aligned}$$

Consequently (59), (66) and (68) imply

$$\left| u(x,t) - \frac{1}{d}\frac{(x-\sigma t) - y(x,t)}{t} \right| \leq \frac{C}{t}. \tag{69}$$

2. Since g has compact support, we may assume for some constant $R > 0$ that $g \equiv 0$ on $\mathbb{R} \cap \{|x| \geq R\}$. Therefore

$$h(x) = \begin{cases} h_- & \text{if } x \leq -R \\ h_+ & \text{if } x \geq R, \end{cases}$$

for constants $h_\pm$. A calculation shows

$$\min_{\mathbb{R}} h = -\frac{p}{2} + h_- = -\frac{q}{2} + h_+. \tag{70}$$

We next set

$$\varepsilon = \varepsilon(t) := \frac{A}{t^{1/2}} \quad (t > 0), \tag{71}$$

the constant A to be selected later.

3. We now claim that if A is sufficiently large, then

$$u(x,t) = 0 \quad \text{for } x - \sigma t < -R - (pd(1+\varepsilon)t)^{1/2} \tag{72}$$

and

$$u(x,t) = 0 \quad \text{for } x - \sigma t > R + (qd(1+\varepsilon)t)^{1/2}. \tag{73}$$

In fact, since (64) implies

$$L''(\sigma) = \frac{1}{d},$$

we deduce from (61) and (62) that

$$tL\left(\frac{x-y}{t}\right) = \frac{1}{d}\frac{|(x-\sigma t) - y|^2}{2t} + O\left(t^{-1/2}\right) \quad \text{as } t \to \infty.$$

Hence

$$tL\left(\frac{x-y}{t}\right) + h(y) = \frac{1}{d}\frac{|(x-\sigma t) - y|^2}{2t} + h(y) + O\left(t^{-1/2}\right). \tag{74}$$

Assume

$$x - \sigma t < -R - (pd(1+\varepsilon)t)^{1/2}. \tag{75}$$

Then $h(x - \sigma t) = h_-$ and so

$$tL\left(\frac{(x-(x-\sigma t))}{t}\right) + h(x-\sigma t) = tL(\sigma) + h_- = h_-.$$

Now if $y \leq -R$, then

$$tL\left(\frac{x-y}{t}\right) + h(y) \geq h_-,$$

since $L \geq 0$. On the other hand if $y \geq -R$, we employ (74) and (70) to estimate

$$\begin{aligned} tL\left(\frac{x-y}{t}\right) + h(y) &\geq \frac{1}{d}\frac{|(x-\sigma t)-y|^2}{2t} - \frac{p}{2} + h_- + O\left(t^{-1/2}\right) \\ &\geq \frac{pd(1+\varepsilon)t}{2dt} - \frac{p}{2} + h_- + O\left(t^{-1/2}\right) \quad \text{by (75)} \\ &= \frac{p}{2}\frac{A}{t^{\frac{1}{2}}} + h_- + O\left(t^{-1/2}\right) \quad \text{by (71)} \\ &\geq h_-, \end{aligned}$$

provided A is large enough.

We conclude that (75) forces $y(x,t) = x - \sigma t$, and so $u(x,t) = G(\sigma) = 0$. This establishes assertion (72), and the proof of (73) is analogous.

4. Next we assert for A and t large enough that

$$(76) \qquad y(x,t) \geq -R \quad \text{if } x - \sigma t = R - (pd(1-\varepsilon)t)^{1/2}.$$

To see this, notice that $y(x,t) \leq -R$ implies as above that

$$tL\left(\frac{x-y}{t}\right) + h(y) \geq h_-.$$

Select now a point z such that $h(z) = \min h = -\frac{p}{2} + h_-$ and $|z| \leq R$. Then we can as before invoke (74) to estimate

$$\begin{aligned} tL\left(\frac{x-z}{t}\right) + h(z) &\leq \frac{1}{d}\frac{|(x-\sigma t)-z|^2}{2t} - \frac{p}{2} + h_- + O\left(t^{-1/2}\right) \\ &\leq \frac{pd(1-\varepsilon)t}{2dt} - \frac{p}{2} + h_- + O\left(t^{-1/2}\right) \\ &= -\frac{p}{2}\frac{A}{t^{\frac{1}{2}}} + h_- + O\left(t^{-1/2}\right) < h_-, \end{aligned}$$

for A large enough. This proves (76) and a similar argument establishes that

$$(77) \qquad y(x,t) \leq R \quad \text{if } x - \sigma t = -R + (qd(1-\varepsilon)t)^{1/2}.$$

5. Remember from the proof of Theorem 1 in §3.4.2 that the mapping $x \mapsto y(x,t)$ is nondecreasing. Hence (69), (76) and (77) imply for large t that

$$(78) \qquad \begin{cases} \left|u(x,t) - \frac{1}{d}\left(\frac{x}{t} - \sigma\right)\right| \leq \frac{C}{t} \quad \text{if} \\ R - (pd(1-\varepsilon)t)^{1/2} < x - \sigma t < -R + (qd(1-\varepsilon)t)^{1/2}. \end{cases}$$

6. Owing to Theorem 5, we have $|u| = O(t^{-\frac{1}{2}})$ and by definition $|N| = O(t^{-\frac{1}{2}})$. In addition (71) implies $((1 \pm \varepsilon)t)^{\frac{1}{2}} - t^{\frac{1}{2}} = O(1)$. Using these bounds along with (72), (73) and (78), we estimate

$$\int_{-\infty}^{\infty} |u(x,t) - N(x,t)|\, dx = O\left(t^{-1/2}\right),$$

as desired. □

Example 3 (continued). Observe we have $p = 0$, $q = 2$, $\sigma = 0$, $d = 1$ in Example 3 of §3.4.1. In this case

$$N(x,t) = \begin{cases} \frac{x}{t} & \text{if } 0 < x < (2t)^{1/2} \\ 0 & \text{otherwise,} \end{cases}$$

and so in fact $u \equiv N$ for times $t \geq 2$. □

We will study *systems* of conservation laws in Chapter 11.

3.5. PROBLEMS

1. Prove

$$u(x,t,a,b) = a \cdot x - tH(a) + b \quad (a \in \mathbb{R}^n, b \in \mathbb{R})$$

is a complete integral of the Hamilton–Jacobi equation

$$u_t + H(Du) = 0.$$

2.

(a) Write down the characteristic equations for the PDE

$$(*) \qquad u_t + b \cdot Du = f \quad \text{in } \mathbb{R}^n \times (0,\infty),$$

where $b \in \mathbb{R}^n$, $f = f(x,t)$.

(b) Use the characteristic ODE to solve (∗) subject to the initial condition

$$u = g \quad \text{on } \mathbb{R}^n \times \{t = 0\}.$$

Make sure your answer agrees with formula (5) in §2.1.2.

3. Solve using characteristics:

(a) $x_1 u_{x_1} + x_2 u_{x_2} = 2u, \quad u(x_1, 1) = g(x_1).$

(b) $u u_{x_1} + u_{x_2} = 1, \quad u(x_1, x_1) = \frac{1}{2} x_1.$

(c) $x_1 u_{x_1} + 2x_2 u_{x_2} = 3u, \quad u(x_1, x_2, 0) = g(x_1, x_2).$

4. Verify that formula (61) in §3.2.5 provides an implicit solution of the scalar conservation law.

5. Write $L = H^*$, if $H : \mathbb{R}^n \to \mathbb{R}$ is convex.

(a) Let $H(p) = \frac{1}{r}|p|^r$, for $1 < r < \infty$. Show

$$L(q) = \frac{1}{s}|q|^s, \quad \text{where } \frac{1}{r} + \frac{1}{s} = 1.$$

(b) Let $H(p) = \frac{1}{2}\sum_{i,j=1}^n a_{ij} p_i p_j + \sum_{i=1}^n b_i p_i$, where $A = ((a_{ij}))$ is a symmetric, positive definite matrix, $b \in \mathbb{R}^n$. Compute $L(q)$.

6. Let $H : \mathbb{R}^n \to \mathbb{R}$ be convex. We say q belongs to the *subdifferential* of H at p, written

$$q \in \partial H(p),$$

if

$$H(r) \geq H(p) + q \cdot (r - p) \quad \text{for all } r \in \mathbb{R}^n.$$

Prove $q \in \partial H(p)$ if and only if $p \in \partial L(q)$ if and only if $p \cdot q = H(p) + L(q)$, where $L = H^*$.

7. Prove that the Hopf–Lax formula reads

$$\begin{aligned} u(x,t) &= \min_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \\ &= \min_{y \in B(x,Rt)} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \end{aligned}$$

for $R = \sup_{\mathbb{R}^n} |DH(Dg)|$, $H = L^*$. (This proves *finite propagation speed* for a Hamilton–Jacobi PDE with convex Hamiltonian and Lipschitz continuous initial function g. Hint: Use the previous problem.)

8. Let E be a closed subset of $\mathbb{R}^n$. Show that *if* the Hopf–Lax formula could be applied to the initial-value problem

$$\begin{cases} u_t + |Du|^2 = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = \begin{cases} 0 & x \in E \\ +\infty & x \notin E \end{cases} & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases}$$

it would give the solution

$$u(x,t) = \frac{1}{4t}\operatorname{dist}(x,E)^2.$$

9. Fill in all details for the proof of Lemma 4 in §3.3.3.
10. Assume u^1, u^2 are two weak solutions of the initial-value problems

$$\begin{cases} u^i_t + H(Du^i) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u^i = g^i & \text{on } \mathbb{R}^n \times \{t=0\} \ (i=1,2), \end{cases}$$

for H as in §3.3. Prove the L^∞-contraction inequality

$$\sup_{\mathbb{R}} |u^1(\cdot,t) - u^2(\cdot,t)| \le \sup_{\mathbb{R}} |g^1 - g^2| \quad (t>0).$$

11. Show that

$$u(x,t) = \begin{cases} -\frac{2}{3}\left(t + \sqrt{3x+t^2}\right) & \text{if } 4x+t^2 > 0 \\ 0 & \text{if } 4x+t^2 < 0 \end{cases}$$

is an (unbounded) entropy solution of $u_t + \left(\frac{u^2}{2}\right)_x = 0$.

12. Assume $u(x+z) - u(x) \le Ez$ for all $z > 0$. Let $u^\epsilon = \eta_\epsilon * u$, and show

$$u^\epsilon_x \le E.$$

13. Assume $F(0) = 0$, u is a continuous integral solution of the conservation law

$$\begin{cases} u_t + F(u)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times \{t=0\}, \end{cases}$$

and u has compact support in $\mathbb{R} \times [0,\infty]$. Prove

$$\int_{-\infty}^{\infty} u(\cdot,t)\,dx = \int_{-\infty}^{\infty} g\,dx$$

for all $t > 0$.

14. Compute explicitly the unique entropy solution of

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times \{t=0\}, \end{cases}$$

for

$$g(x) = \begin{cases} 1 & \text{if } x < -1 \\ 0 & \text{if } -1 < x < 0 \\ 2 & \text{if } 0 < x < 1 \\ 0 & \text{if } x > 1. \end{cases}$$

Draw a picture documenting your answer, being sure to illustrate what happens for all times $t > 0$.

3.6. REFERENCES

Section 3.1 A nice reference for this material is Courant–Hilbert [**C-H**, Chapter 2].

Section 3.2 This derivation of the characteristic differential equations is found in Carathéodory [**C**]. The proof of Theorem 2 follows Garabedian [**G**, Chapter 2], John [**J**, Chapter 1], etc. Chester [**CH**] and Sneddon [**SN**] are also good texts for more on first-order PDE. Example 3 in §3.2.2 is from Zwillinger [**ZW**].

Section 3.3 See Lions [**LI**], Rund [**RU**] and Benton [**BE**] for more on Hamilton–Jacobi PDE. The uniqueness proof, which is due to A. Douglis, is from [**BE**].

Section 3.4 See Lax [**LA**] and Smoller [**S**, Chapters 15,16] (from which I took the proof of Theorem 3, due to O. Oleinik). Theorems 5 and 6 are from DiPerna [**DP**] and I am indebted to M. Struwe for help with the proofs. A good overall reference on nonlinear waves is Whitham [**WH**].

OTHER WAYS TO REPRESENT SOLUTIONS

This chapter collects together a wide variety of techniques that are sometimes useful for finding certain more-or-less explicit solutions to various partial differential equations, or at least representation formulas for solutions.

4.1. SEPARATION OF VARIABLES

The method of *separation of variables* tries to construct a solution u to a given partial differential equation as some sort of combination of functions of fewer variables. In other words, the idea is to guess that u can be written as, say, a sum or product of as yet undetermined constituent functions, to plug this guess into the PDE, and finally to choose the simpler functions to ensure u really is a solution. This technique is best understood in examples.

Example 1. Let $U \subset \mathbb{R}^n$ be a bounded, open set with smooth boundary. We consider the initial/boundary-value problem for the heat equation

$$
\begin{cases} u_t - \Delta u = 0 & \text{in } U \times (0,\infty) \\ u = 0 & \text{on } \partial U \times [0,\infty) \\ u = g & \text{on } U \times \{t=0\}, \end{cases} \tag{1}
$$

where $g : U \to \mathbb{R}$ is given. We conjecture there exists a solution having the multiplicative form

$$
u(x,t) = v(t)w(x) \quad (x \in U,\ t \geq 0); \tag{2}
$$

that is, we look for a solution of (1) with the variables $x = (x_1, \ldots, x_n) \in U$ "separated" from the variable $t \in [0,T]$.

Will this work? To find out, we compute

$$
u_t(x,t) = v'(t)w(x),\ \Delta u(x,t) = v(t)\Delta w(x).
$$

Hence

$$
0 = u_t(x,t) - \Delta u(x,t) = v'(t)w(x) - v(t)\Delta w(x)
$$

if and only if

$$
\frac{v'(t)}{v(t)} = \frac{\Delta w(x)}{w(x)} \tag{3}
$$

for all $x \in U$ and $t > 0$ such that $w(x), v(t) \neq 0$. Now observe the left hand side of (3) depends only on t and the right hand side depends only on x. This is impossible unless each is constant, say

$$
\frac{v'(t)}{v(t)} = \mu = \frac{\Delta w(x)}{w(x)} \quad (t \geq 0,\ x \in U).
$$

Then

$$
v' = \mu v, \tag{4}
$$

$$
\Delta w = \mu w. \tag{5}
$$

We must solve these equations for the unknowns w, v and μ.

Notice first that if μ is known, the solution of (4) is $v = de^{\mu t}$ for an arbitrary constant d. Consequently we need only investigate equation (5).

We say that λ is an *eigenvalue* of the operator $-\Delta$ on U (subject to zero boundary conditions) provided there exists a function w, not identically equal to zero, solving

$$\begin{cases} -\Delta w = \lambda w & \text{in } U \\ w = 0 & \text{on } \partial U. \end{cases}$$

The function w is a corresponding *eigenfunction.* (See Chapter 6 for the theory of eigenvalues, eigenfunctions.)

If λ is an eigenvalue and w is a related eigenfunction, we set $\mu = -\lambda$ above, to find

$$u = de^{-\lambda t} w \tag{6}$$

solves

$$\begin{cases} u_t - \Delta u = 0 & \text{in } U \times (0, \infty) \\ u = 0 & \text{on } \partial U \times [0, \infty), \end{cases} \tag{7}$$

with the initial condition $u(\cdot, 0) = dw$. Thus the function u defined by (6) solves problem (1), provided $g = dw$. More generally, if $\lambda_1, \dots, \lambda_m$ are eigenvalues, $w_1, \dots, w_m$ corresponding eigenfunctions, and $d_1, \dots, d_m$ are constants, then

$$u = \sum_{k=1}^{m} d_k e^{-\lambda_k t} w_k \tag{8}$$

solves (7), with the initial condition $u(\cdot, 0) = \sum_{k=1}^{m} d_k w_k$. If we can find $m, w_1, \dots,$ etc. such that $\sum_{k=1}^{m} d_k w_k = g$, we are done.

We can hope to generalize further by trying to find a countable sequence $\lambda_1, \dots$ of eigenvalues with corresponding eigenfunctions $w_1, \dots,$ so that

$$\sum_{k=1}^{\infty} d_k w_k = g \quad \text{in } U \tag{9}$$

for appropriate constants $d_1, \dots$. Then presumably

$$u = \sum_{k=1}^{\infty} d_k e^{-\lambda_k t} w_k \tag{10}$$

will be the solution of the initial-value problem (1).

This is an attractive representation formula for the solution, but depends upon (a) our being able to find eigenvalues, eigenfunctions and constants satisfying (9), and (b) our verifying that the series in (10) converges in some appropriate sense. We will discuss these matters further in Chapters 6, 7, within the context of Galerkin approximations. □

Remark. Take note that only our solution (6) is determined by separation of variables; the more complicated forms (8) and (10) depend upon the linearity of the heat equation. □

Example 2. Let us next apply the separation of variables technique to discover a solution of the *porous medium equation*

$$u_t - \Delta(u^\gamma) = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty), \tag{11}$$

where $u \geq 0$ and $\gamma > 1$ is a constant. The expression (11) is a nonlinear diffusion equation, in which the rate of diffusion of some density u depends upon u itself. This PDE describes flow in porous media, thin-film lubrication, and a variety of other phenomena.

As in the previous example, we seek a solution of the form

$$u(x,t) = v(t)w(x) \quad (x \in \mathbb{R}^n,\ t \geq 0). \tag{12}$$

Inserting into (11), we discover that

$$\frac{v'(t)}{v(t)^\gamma} = \mu = \frac{\Delta w^\gamma(x)}{w(x)} \tag{13}$$

for some constant μ and all $x \in \mathbb{R}^n$, $t \geq 0$, such that $w(x), v(t) \neq 0$.

We solve the ODE for v and find

$$v = ((1-\gamma)\mu t + \lambda)^{\frac{1}{1-\gamma}},$$

for some constant λ, which we will take to be positive. To discover w, we must then solve the PDE

$$\Delta(w^\gamma) = \mu w. \tag{14}$$

Let us now guess that

$$w = |x|^\alpha,$$

for some constant α that must be determined. Then

$$\mu w - \Delta(w^\gamma) = \mu|x|^\alpha - \alpha\gamma(\alpha\gamma + n - 2)|x|^{\alpha\gamma-2}. \tag{15}$$

So in order that (14) hold in $\mathbb{R}^n$, we should first require that $\alpha = \alpha\gamma - 2$, and hence

$$\alpha = \frac{2}{\gamma - 1}. \tag{16}$$

Returning to (15), we see that we must further set

$$\mu = \alpha\gamma(\alpha\gamma + n - 2) > 0. \tag{17}$$

In summary then, for each $\lambda > 0$ the function

$$u = ((1-\gamma)\mu t + \lambda)^{\frac{1}{1-\gamma}}|x|^\alpha$$

solves the porous medium equation (11), the parameters α, μ defined by (16), (17). □

Remark. Observe that since $\gamma > 1$, this solution blows up for $x \neq 0$ as $t \to t_*$, for $t_* := \frac{\lambda}{(\gamma-1)\mu}$. Physically, a huge amount of mass "diffuses in from infinity" in finite time. See §4.2.2 for another, better behaved solution of the porous medium equation, and see §9.4.1 for more on blow-up phenomena for nonlinear diffusion equations. □

In the previous example separation of variables worked owing to the homogeneity of the nonlinearity, which is compatible with functions u having the multiplicative form (12). In other circumstances it is profitable to look for a solution in which the variables are separated additively:

Example 3. Let us turn once again to the Hamilton–Jacobi equation

$$u_t + H(Du) = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty) \tag{18}$$

and look for a solution u having the form

$$u(x,t) = w(x) + v(t) \quad (x \in \mathbb{R}^n,\ t \geq 0).$$

Then

$$0 = u_t(x,t) + H(Du(x,t)) = v'(t) + H(Dw(x))$$

if and only if

$$H(Dw(x)) = \mu = -v'(t) \quad (x \in \mathbb{R}^n, t > 0)$$

for some constant μ. Consequently if

$$H(Dw) = \mu$$

for some $\mu \in \mathbb{R}$, then

$$u(x,t) = w(x) - \mu t + b$$

will for any constant b solve $u_t + H(Du) = 0$. In particular, if we choose $w(x) = a \cdot x$ for some $a \in \mathbb{R}^n$ and set $\mu = H(a)$, we discover the solution

$$u = a \cdot x - H(a)t + b$$

already noted in §3.1.

□

4.2. SIMILARITY SOLUTIONS

When investigating partial differential equations it is often profitable to look for specific solutions u, the form of which reflects various symmetries in the structure of the PDE. We have already seen this idea in our derivation of the fundamental solutions for Laplace's and the heat equations in §2.2.1 and §2.3.1, and our discovery of rarefaction waves for conservation laws in §3.4.4. Following are some other applications of this important method.

4.2.1. Plane and traveling waves, solitons.

Consider first a partial differential equation involving the two variables $x \in \mathbb{R}$, $t \in \mathbb{R}$. A solution u of the form

$$(1) \qquad u(x,t) = v(x - \sigma t) \quad (x \in \mathbb{R},\ t \in \mathbb{R})$$

is called a *traveling wave* (with *speed* σ and *profile* v). More generally, a solution u of a PDE in the $n+1$ variables $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $t \in \mathbb{R}$ having the form

$$(2) \qquad u(x,t) = v(y \cdot x - \sigma t) \quad (x \in \mathbb{R}^n,\ t \in \mathbb{R})$$

is called a *plane wave* (with *wavefront* normal to $y \in \mathbb{R}^n$, *velocity* $\frac{\sigma}{|y|}$, and *profile* v).

a. Exponential solutions.

In view of the Fourier transform (discussed later, in §4.3.1), it is particularly enlightening when studying linear partial differential equations to consider complex-valued plane wave solutions of the form

$$(3) \qquad u(x,t) = e^{i(y \cdot x + \omega t)},$$

where $\omega \in \mathbb{C}$ and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, ω being the *frequency* and $\{y_i\}_{i=1}^n$ the *wave numbers.* We will next substitute trial solutions of the form (3) into various linear PDE, paying particular attention to the relationship between y and ω forced by the structure of the equation.

(i) Heat equation. If u is given by (3), we compute

$$u_t - \Delta u = (i\omega + |y|^2)u = 0,$$

provided $\omega = i|y|^2$. Hence

$$u = e^{iy \cdot x - |y|^2 t}$$

solves the heat equation for each $y \in \mathbb{R}^n$. Taking real and imaginary parts, we discover further that $e^{-|y|^2 t}\cos(y \cdot x)$ and $e^{-|y|^2 t}\sin(y \cdot x)$ are solutions as well. Notice in this example that since ω is purely imaginary, there results a real, negative exponential term $e^{-|y|^2 t}$ in the formulas, which corresponds to *dissipation.*

(ii) Wave equation. Upon our substituting (3) into the wave equation, we discover

$$u_{tt} - \Delta u = (-\omega^2 + |y|^2)u = 0,$$

provided $\omega = \pm|y|$. Consequently

$$u = e^{i(y \cdot x \pm |y| t)}$$

solves the wave equation, as do the pair of functions $\cos(y \cdot x \pm |y|t)$ and $\sin(y \cdot x \pm |y|t)$. Since ω is real, there are no dissipation effects in these solutions.

(iii) Dispersive equations. We now let $n = 1$ and substitute $u = e^{i(yx+\omega t)}$ into *Airy's equation*

$$u_t + u_{xxx} = 0.$$

We calculate

$$u_t + u_{xxx} = i(\omega - y^3)u = 0,$$

whenever $\omega = y^3$. Thus

$$u = e^{i(yx+y^3 t)}$$

solves Airy's equation, and once again as ω is real there is no dissipation. Notice however that the velocity of propagation is y^2, which depends nonlinearly upon the frequency of the initial value e^{iyx}. Thus waves of different frequencies propagate at different velocities: the PDE creates *dispersion.*

Likewise, if $n \geq 1$ and we substitute $u = e^{i(y \cdot x + \omega t)}$ into *Schrödinger's equation*

$$iu_t + \Delta u = 0,$$

we compute

$$iu_t + \Delta u = -(\omega + |y|^2)u = 0.$$

Consequently $\omega = -|y|^2$, and

$$u = e^{i(y \cdot x - |y|^2 t)}.$$

Again, the solution displays dispersion. □

b. Solitons.

We consider next the *Korteweg–de Vries* (KdV) equation in the form

$$u_t + 6uu_x + u_{xxx} = 0 \quad \text{in } \mathbb{R} \times (0, \infty), \tag{4}$$

this nonlinear dispersive equation being a model for surface waves in water. We seek a traveling wave solution having the structure

$$u(x,t) = v(x - \sigma t) \quad (x \in \mathbb{R},\ t > 0). \tag{5}$$

Then u solves the KdV equation (4), provided v satisfies the ODE

$$-\sigma v' + 6vv' + v''' = 0 \quad \left(' = \frac{d}{ds}\right). \tag{6}$$

We integrate (6) by first noting

$$-\sigma v + 3v^2 + v'' = a, \tag{7}$$

a denoting some constant. Multiply this equality by v' to obtain

$$-\sigma vv' + 3v^2v' + v''v' = av',$$

and so deduce

$$\frac{(v')^2}{2} = -v^3 + \frac{\sigma}{2}v^2 + av + b \tag{8}$$

where b is another arbitrary constant.

We investigate (8) by looking now only for solutions v which satisfy $v, v', v'' \to 0$ as $s \to \pm\infty$. (In which case the function u having the form (5) is called a *solitary* wave.) Then (7), (8) imply $a = b = 0$. Equation (8) thereupon simplifies to read

$$\frac{(v')^2}{2} = v^2\left(-v + \frac{\sigma}{2}\right).$$

Hence $v' = \pm v(\sigma - 2v)^{1/2}$.

We take the minus sign above for computational convenience, and obtain then this implicit formula for v:

$$s = -\int_0^{v(s)} \frac{dz}{z(\sigma - 2z)^{1/2}} + c, \tag{9}$$

for some constant c. Now substitute $z = \frac{\sigma}{2}\operatorname{sech}^2\theta$. It follows that $\frac{dz}{d\theta} = -\sigma\operatorname{sech}^2\theta\tanh\theta$ and $z(\sigma-2z)^{1/2} = \frac{\sigma^{3/2}}{2}\operatorname{sech}^2\theta\tanh\theta$. Hence (9) becomes

$$(10)\qquad s = \frac{2}{\sqrt{\sigma}}\theta + c,$$

where θ is implicitly given by the relation

$$(11)\qquad \frac{\sigma}{2}\operatorname{sech}^2\theta = v(s).$$

We lastly combine (10) and (11), to compute

$$v(s) = \frac{\sigma}{2}\operatorname{sech}^2\left(\frac{\sqrt{\sigma}}{2}(s-c)\right) \quad (s\in\mathbb{R}).$$

Conversely, it is routine to check v so defined actually solves the ODE (6). The upshot is that

$$u(x,t) = \frac{\sigma}{2}\operatorname{sech}^2\left(\frac{\sqrt{\sigma}}{2}(x-\sigma t-c)\right) \quad (x\in\mathbb{R},\ t\geq 0)$$

is a solution of the KdV equation for each $c\in\mathbb{R}$, $\sigma>0$. A solution of this form is called a *soliton*. Note the velocity of the soliton depends upon its height. □

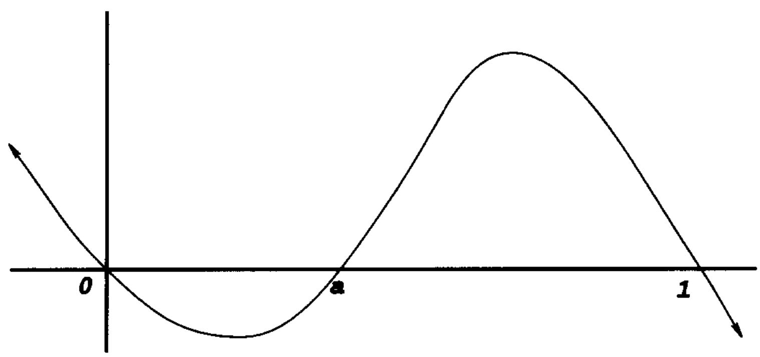
Remark. The KdV equation is in fact utterly remarkable, in that it is *completely integrable*, which means that in principle the exact solution can be computed for essentially arbitrary initial data. The relevant techniques are however beyond the scope of this book: see Drazin [**D**] for more information. □

c. Traveling waves for a bistable equation.

Consider next the scalar reaction-diffusion equation

$$(12)\qquad u_t - u_{xx} = f(u) \quad \text{in } \mathbb{R}\times(0,\infty),$$

where $f:\mathbb{R}\to\mathbb{R}$ has a "cubic-like" shape.



Graph of the function f

We assume, more precisely, f is smooth and verifies

$$\text{(13)} \qquad \begin{cases} \text{(a)} \ \ f(0) = f(a) = f(1) = 0 \\ \text{(b)} \ \ f < 0 \text{ on } (0,a), \ f > 0 \text{ on } (a,1) \\ \text{(c)} \ \ f'(0) < 0, \ f'(1) < 0 \\ \text{(d)} \ \ \int_0^1 f(z)\,dz > 0 \end{cases}$$

for some point $0 < a < 1$.

We look for a traveling wave solution of the form

$$\text{(14)} \qquad u(x,t) = v(x - \sigma t),$$

the profile v and velocity σ to be determined, such that

$$u \to 0 \quad \text{as } x \to -\infty, \qquad u \to 1 \quad \text{as } x \to +\infty.$$

Now since $f' < 0$ at $z = 0, 1$, the constants 0 and 1 are stable solutions of the PDE (and since $f' \geq 0$ at $z = a$, the constant a is an unstable solution). So we want our traveling wave (14) to interpolate between the two stable states $z = 0, 1$ at $x = \mp\infty$.

Plugging (14) into (12), we see v must satisfy the ordinary differential equation

$$\text{(15)} \qquad v'' + \sigma v' + f(v) = 0 \quad \left(' = \frac{d}{ds} \right),$$

subject to the conditions

$$\text{(16)} \qquad \lim_{s \to +\infty} v(s) = 1, \ \lim_{s \to -\infty} v(s) = 0, \ \lim_{s \to \pm\infty} v'(s) = 0.$$

We outline now (without complete proofs) a *phase plane analysis* of the ODE problem (15), (16). We begin by setting

$$w := v'.$$

Then (15), (16) transform into the autonomous first-order system:

$$\text{(17)} \qquad \begin{cases} v' = w \\ w' = -\sigma w - f(v), \end{cases}$$

with

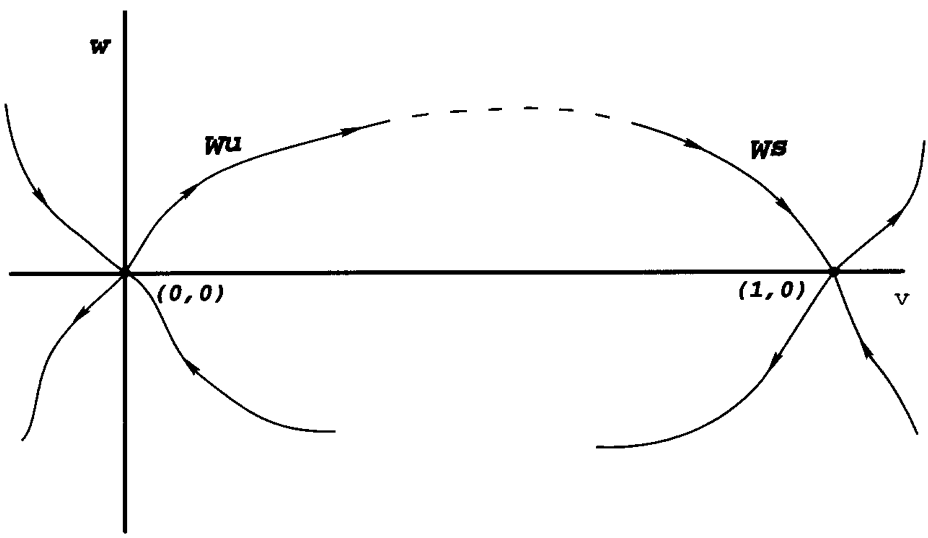
$$\text{(18)} \qquad \lim_{s \to \infty} (v,w) = (1,0), \ \lim_{s \to -\infty} (v,w) = (0,0).$$

Now $(0,0)$ and $(1,0)$ are critical points for the system (17), and the eigenvalues of the corresponding linearizations are

$$(19) \qquad \lambda_0^{\pm} = \frac{-\sigma \pm (\sigma^2 - 4f'(0))^{1/2}}{2}, \quad \lambda_1^{\pm} = \frac{-\sigma \pm (\sigma^2 - 4f'(1))^{1/2}}{2}.$$

In view of (13)(c), $\lambda_0^{\pm}$, $\lambda_1^{\pm}$ are real, with differing sign, and thus $(0,0)$ and $(1,0)$ are saddle points for the flow (17). Consequently an "unstable curve" W^u leaves $(0,0)$ and a "stable curve" W^s approaches $(1,0)$, as drawn. Furthermore, by calculating eigenvectors corresponding to (19) we see

$$(20) \qquad \begin{cases} W^u \text{ is tangent to the line } w = \lambda_0^+ v \text{ at } (0,0) \\ W^s \text{ is tangent to the line } w = \lambda_1^-(v-1) \text{ at } (1,0). \end{cases}$$



Stable and unstable curves

Note that $\lambda_0^{\pm}, \lambda_1^{\pm}, W^u$ and W^s depend upon the parameter σ. Our intention is to find $\sigma < 0$ so that

$$(21) \qquad W^u = W^s \quad \text{in the region } \{v > 0, w > 0\}.$$

Then we will have a solution of (17), (18), whose path in the phase plane is a heteroclinic orbit connecting $(0,0)$ to $(1,0)$.

To establish (21), we fix now a small number $\varepsilon > 0$ and let L denote the vertical line through the point $(a + \varepsilon, 0)$. We claim

$$(22) \qquad W^s \cap L \neq \emptyset, \ W^u \cap L \neq \emptyset$$

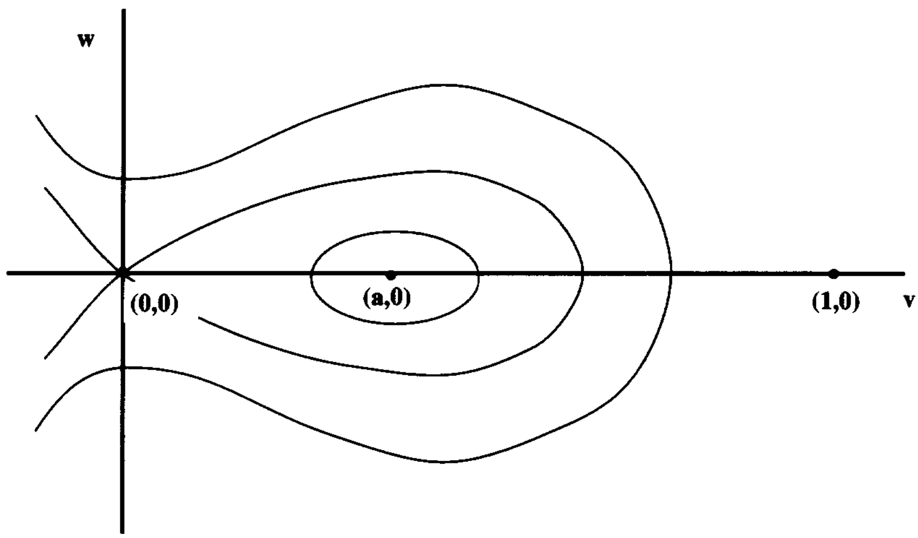
if $\sigma < 0$. To check this assertion, define

$$E(v,w) := \frac{w^2}{2} + \int_0^v f(z)\,dz \quad (v, w \in \mathbb{R})$$

and compute

$$\begin{aligned}\frac{d}{dt}E(v(t),w(t)) &= w(t)w'(t) + f(v(t))v'(t)\\ &= -\sigma w^2(t) \quad \text{by (17).}\end{aligned}$$

As $\sigma < 0$, we see that E is nondecreasing along trajectories of the ODE (17). Note also the level sets of E have the shapes illustrated.



Level curves of E

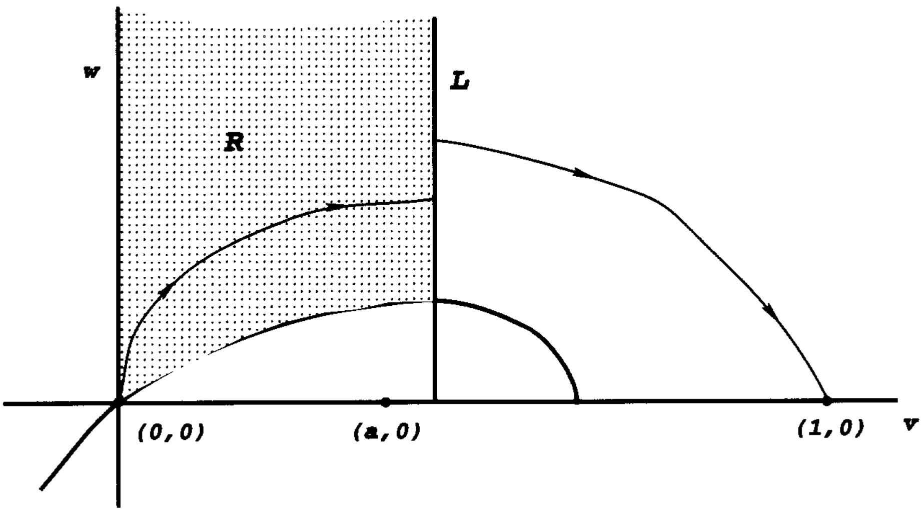
Consider next the region R, as drawn. The unstable curve enters R from $(0,0)$ and cannot exit through the bottom, top or left hand side. Using (17) we deduce that W^u must exit R through the line L, at a point $(a+\varepsilon, w_0(\sigma))$. Similarly we argue W^s must hit L at a point $(a+\varepsilon, w_1(\sigma))$. This verifies claim (22).

We next observe

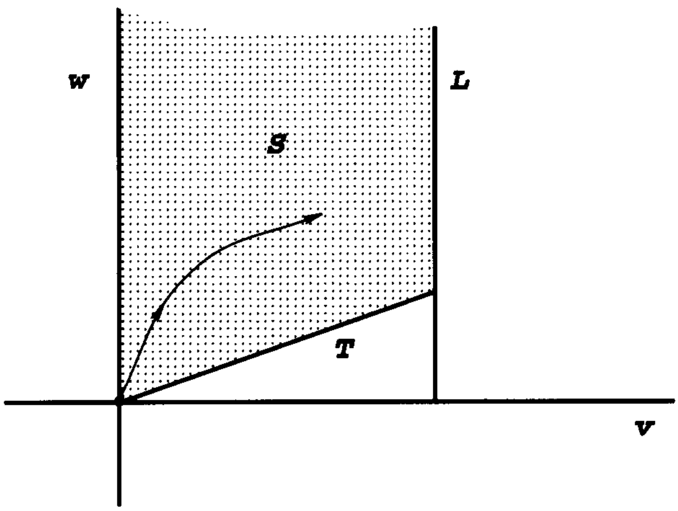
$$w_0(0) < w_1(0); \tag{23}$$

this follows since trajectories of (17) for $\sigma = 0$ are contained in level sets of E. We assert further that

$$w_0(\sigma) > w_1(\sigma) \tag{24}$$



The region R



The region S

provided $\sigma < 0$ and $|\sigma|$ is large enough. To see this, fix $\beta > 0$ and consider the region S, as drawn.

Now along the line segment $T := \{0 \leq v \leq a + \varepsilon, w = \beta v\}$, we have

$$\frac{w'}{v'} = \frac{-\sigma w - f(v)}{w} = -\sigma - \frac{f(v)}{\beta v}.$$

Since $\left|\frac{f(v)}{v}\right|$ is bounded for $0 \leq v \leq a + \varepsilon$, we see

$$\frac{w'}{v'} \geq -\sigma - \frac{C}{\beta} > \beta \quad \text{on } T, \tag{25}$$

provided $\sigma < 0$ and $|\sigma|$ is large enough.

The calculation (25) shows that W^u cannot exit S through the line segment T, and so $w_0(\sigma) \geq \beta(a + \varepsilon)$ if $\sigma = \sigma(\beta)$ is sufficiently negative. On the other hand, $w_1(\sigma) \leq w_1(0)$ for all $\sigma \leq 0$. Thus we see that (23) will follow once we choose β large enough and then σ sufficiently negative.

Since w_0 and w_1 depend smoothly on σ, we deduce from (22) and (23) that there exists $\sigma < 0$ with

$$w_0(\sigma) = w_1(\sigma). \tag{26}$$

For this velocity σ there consequently exists a solution of the ODE (17), (18). Hence we have found for our reaction-diffusion PDE (12) a traveling wave of the form (14). □

Remark. A more refined analysis demonstrates that the velocity σ verifying (26) is unique. Hence given the nonlinearity f satisfying hypotheses (13), there exists a *unique* velocity for which there is a corresponding traveling wave. Compare this assertion with Example 2 above, where we found soliton traveling waves for each given velocity. □

4.2.2. Similarity under scaling.

We next illustrate the possibility of finding other types of "similarity" solutions to PDE.

Example (A scaling invariant solution). Consider again the *porous medium equation*

$$u_t - \Delta(u^\gamma) = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \tag{27}$$

where $u \geq 0$ and $\gamma > 1$ is a constant.

As in our earlier derivation of the fundamental solution of the heat equation in §2.3.1, let us look for a solution u having the form

$$u(x,t) = \frac{1}{t^\alpha} v\left(\frac{x}{t^\beta}\right) \quad (x \in \mathbb{R}^n,\ t > 0), \tag{28}$$

where the constants α, β and the function $v : \mathbb{R}^n \to \mathbb{R}$ must be determined. Remember that we come upon (28) if we seek a solution u of (27) invariant under the *dilation scaling*

$$u(x,t) \mapsto \lambda^\alpha u(\lambda^\beta x, \lambda t);$$

so that

$$u(x,t) = \lambda^\alpha u(\lambda^\beta x, \lambda t)$$

for all $\lambda > 0$, $x \in \mathbb{R}^n$, $t > 0$. Setting $\lambda = t^{-1}$, we obtain (28) for $v(y) := u(y,1)$.

We insert (28) into (27), and discover

$$\alpha t^{-(\alpha+1)} v(y) + \beta t^{-(\alpha+1)} y \cdot Dv(y) + t^{-(\alpha\gamma+2\beta)} \Delta(v^\gamma)(y) = 0 \tag{29}$$

for $y = t^{-\beta}x$. In order to convert (29) into an expression involving the variable y alone, let us require

$$\alpha + 1 = \alpha\gamma + 2\beta. \tag{30}$$

Then (29) reduces to

$$\alpha v + \beta y \cdot Dv + \Delta(v^\gamma) = 0. \tag{31}$$

At this point we have effected a reduction from $n+1$ to n variables. We simplify further by supposing v is radial; that is, $v(y) = w(|y|)$ for some $w : \mathbb{R} \to \mathbb{R}$. Then (31) becomes

$$\alpha w + \beta r w' + (w^\gamma)'' + \frac{n-1}{r}(w^\gamma)' = 0, \tag{32}$$

where $r = |y|$, $' = \frac{d}{dr}$. Now if we set

$$\alpha = n\beta, \tag{33}$$

(32) thereupon simplifies to read

$$(r^{n-1}(w^\gamma)')' + \beta(r^n w)' = 0.$$

Thus

$$r^{n-1}(w^\gamma)' + \beta r^n w = a$$

for some constant a. Assuming $\lim_{r\to\infty} w,\ w' = 0$, we conclude $a = 0$; whence

$$(w^\gamma)' = -\beta r w.$$

But then

$$(w^{\gamma-1})' = -\frac{(\gamma-1)}{\gamma}\beta r.$$

Consequently

$$w^{\gamma-1} = b - \frac{\gamma-1}{2\gamma}\beta r^2,$$

b a constant; and so

$$w = \left(b - \frac{\gamma-1}{2\gamma}\beta r^2\right)^{+\ \frac{1}{\gamma-1}}, \tag{34}$$

where we took the positive part of the right hand side of (34) to ensure $w \geq 0$. Recalling $v(y) = w(r)$ and (28), we obtain

$$u(x,t) = \frac{1}{t^\alpha}\left(b - \frac{\gamma-1}{2\gamma}\beta\frac{|x|^2}{t^{2\beta}}\right)^{+\ \frac{1}{\gamma-1}} \quad (x \in \mathbb{R}^n,\ t > 0), \tag{35}$$

where, from (30), (33),

$$\alpha = \frac{n}{n(\gamma-1)+2}, \quad \beta = \frac{1}{n(\gamma-1)+2}. \tag{36}$$

The formulas (35), (36) are *Barenblatt's solution* to the porous medium equation. □

Remarks. Observe that Barenblatt's solution has compact support for each time $t > 0$. This is a general feature for (appropriately defined) weak, non-negative solutions of the porous medium equation with compactly supported initial data. The nonlinear parabolic PDE (27) becomes degenerate wherever $u = 0$, and so the set $\{u > 0\}$ moves with *finite propagation speed.* Consequently the porous medium equation (27) is often regarded as a better model of diffusive spreading than the linear heat equation (which predicts infinite propagation speed). □

4.3. TRANSFORM METHODS

In this section we develop some of the theory of Fourier and Laplace transforms, which provides extremely powerful tools for converting certain linear partial differential equations into either algebraic equations or else differential equations involving fewer variables.

4.3.1. Fourier transform.

In this section all functions are complex-valued, and $\bar{\ }$ denotes the complex conjugate.

a. Definitions and elementary properties.

Definition of Fourier transform on $\mathbf{L^1}$. If $u \in L^1(\mathbb{R}^n)$, we define its *Fourier transform*

(1) $$\hat{u}(y) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot y} u(x)\, dx \quad (y \in \mathbb{R}^n)$$

and its *inverse Fourier transform*

(2) $$\check{u}(y) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot y} u(x)\, dx \quad (y \in \mathbb{R}^n).$$

Since $|e^{\pm ix\cdot y}| = 1$ and $u \in L^1(\mathbb{R}^n)$, these integrals converge for each $y \in \mathbb{R}^n$.

We intend now to extend definitions (1), (2) to functions $u \in L^2(\mathbb{R}^n)$.

THEOREM 1 (Plancherel's Theorem). *Assume $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\hat{u}, \check{u} \in L^2(\mathbb{R}^n)$ and*

(3) $$\|\hat{u}\|_{L^2(\mathbb{R}^n)} = \|\check{u}\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}.$$

Proof. 1. First we note that if $v, w \in L^1(\mathbb{R}^n)$, then $\hat{v}, \hat{w} \in L^\infty(\mathbb{R}^n)$. Also

(4) $$\int_{\mathbb{R}^n} v(x)\hat{w}(x)\, dx = \int_{\mathbb{R}^n} \hat{v}(y) w(y)\, dy,$$

since both expressions equal $\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-ix\cdot y} v(x) w(y)\, dxdy$. Furthermore, as we will explicitly compute below in Example 1 of §4.3.2,

$$\int_{\mathbb{R}^n} e^{ix\cdot y - t|x|^2} dx = \left(\frac{\pi}{t}\right)^{n/2} e^{-\frac{|y|^2}{4t}} \quad (t > 0).$$

Consequently if $\varepsilon > 0$ and $v_\varepsilon(x) := e^{-\varepsilon|x|^2}$, we have $\hat{v}_\varepsilon(y) = \frac{e^{-\frac{|y|^2}{4\varepsilon}}}{(2\varepsilon)^{n/2}}$. Thus (4) implies for each $\varepsilon > 0$ that

(5) $$\int_{\mathbb{R}^n} \hat{w}(y) e^{-\varepsilon|y|^2} dy = \frac{1}{(2\varepsilon)^{n/2}} \int_{\mathbb{R}^n} w(x) e^{-\frac{|x|^2}{4\varepsilon}} dx.$$

2. Now take $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and set $v(x) := \bar{u}(-x)$. Let $w := u * v \in L^1(\mathbb{R}^n) \cap C(\mathbb{R}^n)$ and check (cf. Theorem 2 below) that

$$\hat{w} = (2\pi)^{n/2} \hat{u}\hat{v} \in L^\infty(\mathbb{R}^n).$$

But

$$\hat{v}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot y}\bar{u}(-x)\, dx = \bar{\hat{u}}(y);$$

and so $\hat{w} = (2\pi)^{n/2}|\hat{u}|^2$.

Now w is continuous and thus

$$\lim_{\varepsilon\to 0} \frac{1}{(2\varepsilon)^{n/2}} \int_{\mathbb{R}^n} w(x)e^{-\frac{|x|^2}{4\varepsilon}}\, dx = (2\pi)^{n/2} w(0),$$

where we employed the lemma from §2.3.1. Since $\hat{w} = (2\pi)^{n/2}|\hat{u}|^2 \geq 0$, we deduce upon sending $\varepsilon \to 0^+$ in (5) that $\hat{w}$ is summable, with

$$\int_{\mathbb{R}^n} \hat{w}(y)\, dy = (2\pi)^{n/2} w(0).$$

Hence

$$\int_{\mathbb{R}^n} |\hat{u}|^2 dy = w(0) = \int_{\mathbb{R}^n} u(x)v(-x)\, dx = \int_{\mathbb{R}^n} |u|^2 dx.$$

The proof for $\check{u}$ is similar. □

Definition of Fourier transform on $\mathbf{L^2}$. In view of the equality (3) we can define the Fourier transforms of a function $u \in L^2(\mathbb{R}^n)$ as follows. Choose a sequence $\{u_k\}_{k=1}^\infty \subset L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ with

$$u_k \to u \quad \text{in } L^2(\mathbb{R}^n).$$

According to (3), $\|\hat{u}_k - \hat{u}_j\|_{L^2(\mathbb{R}^n)} = \|\widehat{u_k - u_j}\|_{L^2(\mathbb{R}^n)} = \|u_k - u_j\|_{L^2(\mathbb{R}^n)}$, and thus $\{\hat{u}_k\}_{k=1}^\infty$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. This sequence consequently converges to a limit, which we define to be $\hat{u}$:

$$\hat{u}_k \to \hat{u} \quad \text{in } L^2(\mathbb{R}^n).$$

The definition of $\hat{u}$ does not depend upon the choice of approximating sequence $\{\hat{u}_k\}_{k=1}^\infty$. We similarly define $\check{u}$.

Next we record some useful formulas.

THEOREM 2 (Properties of Fourier transform). *Assume $u, v \in L^2(\mathbb{R}^n)$. Then*

(i) $\int_{\mathbb{R}^n} u\bar{v}dx = \int_{\mathbb{R}^n} \hat{u}\bar{\hat{v}}dy$,

(ii) $\widehat{D^\alpha u} = (iy)^\alpha \hat{u}$ *for each multiindex α such that $D^\alpha u \in L^2(\mathbb{R}^n)$*,

(iii) $(u * v)\hat{} = (2\pi)^{n/2}\hat{u}\hat{v}$,

(iv) $u = (\hat{u})\check{}$.

Proof. 1. Let $u, v \in L^2(\mathbb{R}^n)$ and $\alpha \in \mathbb{C}$. Then

$$\|u + \alpha v\|^2_{L^2(\mathbb{R}^n)} = \|\hat{u} + \widehat{\alpha v}\|^2_{L^2(\mathbb{R}^n)}.$$

Expanding, we deduce

$$\int_{\mathbb{R}^n} |u|^2 + |\alpha v|^2 + \bar{u}(\alpha v) + u(\bar{\alpha}\bar{v})\, dx = \int_{\mathbb{R}^n} |\hat{u}|^2 + |\widehat{\alpha v}|^2 + \bar{\hat{u}}(\widehat{\alpha v}) + \hat{u}(\bar{\alpha}\bar{\hat{v}})\, dy;$$

and so according to Theorem 1,

$$\int_{\mathbb{R}^n} \alpha \bar{u} v + \bar{\alpha} u \bar{v}\, dx = \int_{\mathbb{R}^n} \alpha \bar{\hat{u}} \hat{v} + \bar{\alpha} \hat{u} \bar{\hat{v}}\, dy.$$

Take $\alpha = 1, i$ and combine the resulting equalities to deduce

$$\int_{\mathbb{R}^n} u \bar{v}\, dx = \int_{\mathbb{R}^n} \hat{u} \bar{\hat{v}}\, dy.$$

This proves (i).

2. If u is smooth and has compact support, we calculate

$$\begin{aligned} \widehat{D^\alpha u}(y) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot y} D^\alpha u(x)\, dx \\ &= \frac{(-1)^{|\alpha|}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} D^\alpha_x (e^{-ix\cdot y}) u(x)\, dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot y} (iy)^\alpha u(x)\, dx = (iy)^\alpha \hat{u}(y). \end{aligned}$$

By approximation the same formula is true if $D^\alpha u \in L^2(\mathbb{R}^n)$.

3. We compute for $u, v \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $y \in \mathbb{R}^n$ that

$$\begin{aligned} (\widehat{u * v})(y) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot y} \int_{\mathbb{R}^n} u(z) v(x - z)\, dz dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-iz\cdot y} u(z) \left(\int_{\mathbb{R}^n} e^{-i(x-z)\cdot y} v(x - z) dx \right) dz \\ &= \int_{\mathbb{R}^n} e^{-iz\cdot y} u(z)\, dz\ \hat{v}(y) = (2\pi)^{n/2} \hat{u}(y) \hat{v}(y). \end{aligned}$$

4. Fix $z \in \mathbb{R}^n$, $\varepsilon > 0$ and write $v_\varepsilon(x) := e^{ix\cdot z - \varepsilon|x|^2}$. Then

$$\hat{v}_\varepsilon(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot(y-z) - \varepsilon|x|^2} dx = \frac{1}{(2\varepsilon)^{n/2}} e^{-\frac{|x-z|^2}{4\varepsilon}},$$

where we followed computations from the proof of Theorem 1. Utilizing formula (4), we deduce for $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ that

$$\int_{\mathbb{R}^n} \hat{u}(y) e^{iz\cdot y - \varepsilon|y|^2} dy = \frac{1}{(2\varepsilon)^{n/2}} \int_{\mathbb{R}^n} u(x) e^{-\frac{|x-z|^2}{4\varepsilon}} dz.$$

The expression on the right converges to $(2\pi)^{n/2}u(z)$ as $\varepsilon \to 0^+$, for each Lebesgue point of u. Thus

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{u}(y) e^{iz\cdot y} dy = u(z) \quad \text{for a.e. } z.$$

This proves (iv). □

b. Applications.

The Fourier transform is an especially powerful technique for studying linear, constant-coefficient partial differential equations.

Example 1 (Bessel potentials). We investigate first the PDE

$$-\Delta u + u = f \quad \text{in } \mathbb{R}^n, \tag{6}$$

where $f \in L^2(\mathbb{R}^n)$. To find an explicit formula for u, we take the Fourier transform, recalling Theorem 2,(ii) to obtain

$$(1 + |y|^2)\hat{u}(y) = \hat{f}(y) \quad (y \in \mathbb{R}^n). \tag{7}$$

The effect of the Fourier transform has been to convert the PDE (6) into the algebraic equation (7), the solution of which is trivial:

$$\hat{u} = \frac{\hat{f}}{1 + |y|^2}.$$

Thus

$$u = \left(\frac{\hat{f}}{1 + |y|^2} \right)^{\vee}, \tag{8}$$

and so the only real problem is to rewrite the right hand side of (8) into a more explicit form.

Invoking Theorem 2,(iii), we see

$$u = \frac{f * B}{(2\pi)^{n/2}}, \tag{9}$$

where

$$\hat{B} = \frac{1}{1+|y|^2}. \tag{10}$$

We solve for B as follows. Since $\frac{1}{a} = \int_0^\infty e^{-ta}dt$ for each $a > 0$, we have $\frac{1}{1+|y|^2} = \int_0^\infty e^{-t(1+|y|^2)}dt$. Thus

$$B = \left(\frac{1}{1+|y|^2}\right)^{\vee} = \frac{1}{(2\pi)^{n/2}} \int_0^\infty e^{-t} \left(\int_{\mathbb{R}^n} e^{ix\cdot y - t|y|^2} dy\right) dt. \tag{11}$$

Now if $a, b \in \mathbb{R}$, $b > 0$, and we set $z = b^{1/2}x - \frac{a}{2b^{1/2}}i$, we find

$$\int_{-\infty}^{\infty} e^{iax-bx^2} dx = \frac{e^{-a^2/4b}}{b^{1/2}} \int_\Gamma e^{-z^2} dz,$$

Γ denoting the contour $\left\{\operatorname{Im}(z) = -\frac{a}{2b^{1/2}}\right\}$ in the complex plane. Deforming Γ into the real axis, we compute $\int_\Gamma e^{-z^2}dz = \int_{-\infty}^\infty e^{-x^2}dx = \pi^{1/2}$; and hence

$$\int_{-\infty}^{\infty} e^{iax-bx^2} dx = e^{-a^2/4b} \left(\frac{\pi}{b}\right)^{1/2}. \tag{12}$$

Thus

$$\int_{\mathbb{R}^n} e^{ix\cdot y - t|y|^2} dy = \prod_{j=1}^n \int_{-\infty}^\infty e^{ix_j y_j - t y_j^2} dy_j = \left(\frac{\pi}{t}\right)^{n/2} e^{-\frac{|x|^2}{4t}} \tag{13}$$

by (12). Consequently, we conclude from (11), (13) that

$$B(x) = \frac{1}{2^{n/2}} \int_0^\infty \frac{e^{-t-\frac{|x|^2}{4t}}}{t^{n/2}} dt \quad (x \in \mathbb{R}^n). \tag{14}$$

B is called a *Bessel potential.* Employing (9), we derive then the formula

$$u(x) = \frac{1}{(4\pi)^{n/2}} \int_0^\infty \int_{\mathbb{R}^n} \frac{e^{-t-\frac{|x-y|^2}{4t}}}{t^{n/2}} f(y)\, dydt \quad (x \in \mathbb{R}^n) \tag{15}$$

for the solution of (6). □

Example 2 (Fundamental solution of heat equation). Consider again the initial-value problem for the heat equation

$$\text{(16)} \qquad \begin{cases} u_t - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

We establish a new method for solving (16) by computing $\hat{u}$, the Fourier transform of u in *the spatial variables x only.* Thus

$$\begin{cases} \hat{u}_t + |y|^2\hat{u} = 0 & \text{for } t > 0 \\ \hat{u} = \hat{g} & \text{for } t = 0; \end{cases}$$

whence

$$\hat{u} = e^{-t|y|^2}\hat{g}.$$

Consequently $u = \left(e^{-t|y|^2}\hat{g}\right)^{\vee}$, and therefore

$$\text{(17)} \qquad u = \frac{g * F}{(2\pi)^{n/2}},$$

where $\hat{F} = e^{-t|y|^2}$. But then

$$F = \left(e^{-t|y|^2}\right)^{\vee} = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot y - t|y|^2}\,dy = \frac{1}{(2t)^{n/2}} e^{-\frac{|x|^2}{4t}}$$

by (13). Invoking (17), we compute

$$\text{(18)} \qquad u(x,t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\,dy \quad (x \in \mathbb{R}^n,\ t > 0),$$

in agreement with §2.3.1. The Fourier transform has provided us with a new derivation of the fundamental solution of the heat equation. □

Example 3 (Fundamental solution of Schrödinger's equation). Let us next look at the initial-value problem for Schrödinger's equation

$$\text{(19)} \qquad \begin{cases} iu_t + \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

Here u and g are complex-valued.

If we formally replace t by it on the right hand side of (18), we obtain the formula

$$\text{(20)} \qquad u(x,t) = \frac{1}{(4\pi it)^{n/2}} \int_{\mathbb{R}^n} e^{\frac{i|x-y|^2}{4t}} g(y)\,dy \quad (x \in \mathbb{R}^n,\ t > 0),$$

where we interpret $i^{\frac{1}{2}}$ as $e^{\frac{i\pi}{4}}$. This expression clearly makes sense for all times $t > 0$, provided $g \in L^1(\mathbb{R}^n)$. Furthermore if $|y|^2 g \in L^1(\mathbb{R}^n)$, we can check by a direct calculation that u solves $iu_t + \Delta u = 0$ in $\mathbb{R}^n \times (0, \infty)$. (We will not discuss here the sense in which $u(\cdot, t) \to g$ as $t \to 0^+$, but see §4.5.3 below and Problem 5.)

Let us next rewrite formula (20) as

$$u(x,t) = \frac{e^{\frac{i|x|^2}{4t}}}{(4\pi i t)^{n/2}} \int_{\mathbb{R}^n} e^{\frac{-ix\cdot y}{2t}} e^{\frac{i|y|^2}{4t}} g(y)\, dy.$$

Since $|e^{\frac{i|x|^2}{4t}}, e^{\frac{i|y|^2}{4t}}| = 1$, we can check as in Theorem 1 that if $g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then

$$\|u(\cdot,t)\|_{L^2(\mathbb{R}^n)} = \|g\|_{L^2(\mathbb{R}^n)} \quad (t > 0). \tag{21}$$

Hence the mapping $g \mapsto u(\cdot, t)$ preserves the L^2-norm. Therefore we can extend formula (20) to functions $g \in L^2(\mathbb{R}^n)$, in the same way that we extended the definition of Fourier transform. □

Remark. We call

$$\Psi(x,t) := \frac{1}{(4\pi i t)^{n/2}} e^{\frac{i|x|^2}{4t}} \quad (x \in \mathbb{R}^n,\ t \neq 0) \tag{22}$$

the *fundamental solution* of Schrödinger's equation. Note that formula (20), $u = g * \Psi$, makes sense for all times $t \neq 0$, even $t < 0$. Thus we in fact have solved

$$\begin{cases} iu_t + \Delta u = 0 & \text{in } \mathbb{R}^n \times (-\infty, \infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{23}$$

In particular, Schrödinger's equation is *reversible in time*, whereas the heat equation is not (in spite of Theorem 11 in §2.3.4). □

Example 4 (Wave equation). We next analyze the initial-value problem for the wave equation

$$\begin{cases} u_{tt} - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ u = g,\ u_t = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{24}$$

where for simplicity we suppose the initial velocity to be zero. Take as before $\hat{u}$ to be the Fourier transform of u in the variable $x \in \mathbb{R}^n$. Then

$$\begin{cases} \hat{u}_{tt} + |y|^2 \hat{u} = 0 & \text{for } t > 0 \\ \hat{u} = \hat{g},\ \hat{u}_t = 0 & \text{for } t = 0. \end{cases} \tag{25}$$

This is an ODE for each fixed $y \in \mathbb{R}^n$. We look for a solution having the form $\hat{u} = \beta e^{t\gamma}$ ($\beta, \gamma \in \mathbb{C}$). Plugging into (25) gives $\gamma^2 + |y|^2 = 0$ and so $\gamma = \pm i|y|$. Remembering the initial conditions from (25), we deduce

$$\hat{u} = \frac{\hat{g}}{2}(e^{it|y|} + e^{-it|y|}).$$

Inverting, we find

$$u(x,t) = \left[\frac{\hat{g}}{2}(e^{it|y|} + e^{-it|y|})\right]^{\vee};$$

and consequently

$$u(x,t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{\hat{g}(y)}{2}(e^{i(x\cdot y + t|y|)} + e^{i(x\cdot y - t|y|)})\, dy \tag{26}$$

for $x \in \mathbb{R}^n$, $t \geq 0$. We will further analyze this formula in certain asymptotic limits later, in §4.5.3. See also Example 1(ii) in §4.2.1. □

Example 5 (Telegraph equation). The initial-value problem for the one-dimensional *telegraph equation* is

$$\begin{cases} u_{tt} + 2du_t - u_{xx} = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g, \ u_t = h & \text{on } \mathbb{R} \times \{t = 0\}, \end{cases} \tag{27}$$

for $d > 0$, the term "$2du_t$" representing a physical damping of wave propagation. As before

$$\begin{cases} \hat{u}_{tt} + 2d\hat{u}_t + |y|^2\hat{u} = 0 & \text{for } t > 0 \\ \hat{u} = \hat{g}, \ \hat{u}_t = \hat{h} & \text{for } t = 0. \end{cases} \tag{28}$$

We again seek a solution of the form $\hat{u} = \beta e^{t\gamma}$ ($\beta, \gamma \in \mathbb{C}$). Plugging into (28), we deduce $\gamma^2 + 2d\gamma + |y|^2 = 0$; whence $\gamma = -d \pm (d^2 - |y|^2)^{1/2}$. Consequently

$$\hat{u}(y,t) = \begin{cases} e^{-dt}(\beta_1(y)e^{\gamma(y)t} + \beta_2(y)e^{-\gamma(y)t}) & \text{if } |y| \leq d \\ e^{-dt}(\beta_1(y)e^{i\delta(y)t} + \beta_2(y)e^{-i\delta(y)t}) & \text{if } |y| \geq d \end{cases}$$

for $\gamma(y) := (d^2 - |y|^2)^{1/2} (|y| \leq d)$, $\delta(y) := (|y|^2 - d^2)^{1/2}$ $(|y| \geq d)$, where $\beta_1(y)$ and $\beta_2(y)$ are selected so that

$$\hat{g}(y) = \beta_1(y) + \beta_2(y)$$

and

$$\hat{h}(y) = \begin{cases} \beta_1(y)(\gamma(y) - d) + \beta_2(y)(-\gamma(y) - d) & \text{if } |y| \leq d \\ \beta_1(y)(i\delta(y) - d) + \beta_2(y)(-i\delta(y) - d) & \text{if } |y| \geq d. \end{cases}$$

We thereby obtain the representation formula:

$$\begin{aligned} u(x,t) = {} & \frac{e^{-dt}}{(2\pi)^{n/2}} \int_{\{|y|\leq d\}} \beta_1(y)e^{ixy+\gamma(y)t} + \beta_2(y)e^{ixy-\gamma(y)t}\, dy \\ & + \frac{e^{-dt}}{(2\pi)^{n/2}} \int_{\{|y|\geq d\}} \beta_1(y)e^{i(xy+\delta(y)t)} + \beta_2(y)e^{i(xy-\delta(y)t)}\, dy. \end{aligned}$$

Notice the terms e^{-dt}, which correspond to damping as $t \to \infty$. □

4.3.2. Laplace transform.

Remember that we write $R_+ = (0, \infty)$.

DEFINITION. *If* $u \in L^1(\mathbb{R}_+)$, *we define its* Laplace transform *to be*

$$u^\#(s) := \int_0^\infty e^{-st} u(t)\,dt \quad (s \geq 0). \tag{29}$$

Whereas the Fourier transform is most appropriate for functions defined on all of $\mathbb{R}$ (or $\mathbb{R}^n$), the Laplace transform is useful for functions defined only on $\mathbb{R}_+$. In practice this means that for a partial differential equation involving time, it may be useful to perform a Laplace transform in t, holding the space variables x fixed. (This is the opposite of the technique from Examples 2–5 above.)

Example 1 (Resolvents and Laplace transform). Consider again the heat equation

$$\begin{cases} v_t - \Delta v = 0 & \text{in } U \times (0, \infty) \\ \quad\;\; v = f & \text{on } U \times \{t = 0\}, \end{cases} \tag{30}$$

and perform a Laplace transform with respect to time:

$$v^\#(x, s) = \int_0^\infty e^{-st} v(x, t)\,dt \quad (s > 0).$$

What PDE does $v^\#$ satisfy? We compute

$$\begin{aligned} \Delta v^\#(x, s) &= \int_0^\infty e^{-st} \Delta v(x, t)\,dt = \int_0^\infty e^{-st} v_t(x, t)\,dt \\ &= s \int_0^\infty e^{-st} v(x, t)\,dt + e^{-st} v\big|_{t=0}^{t=\infty} = s v^\#(x, s) - f(x). \end{aligned}$$

Think now of $s > 0$ being fixed, and write $u(x) := v^\#(x, s)$. Then

$$-\Delta u + su = f \quad \text{in } U. \tag{31}$$

Thus the solution of the resolvent equation (31) with right hand side f is the Laplace transform of the solution of the heat equation (30) with initial data f. (If $U = \mathbb{R}^n$ and $s = 1$, we could now represent v in terms of the fundamental solution, to rederive formula (15).) □

The connection between the resolvent equation and the Laplace transform will be made clearer by the discussion in §7.4 of semigroup theory.

Example 2 (Wave equation from the heat equation). Next we employ some Laplace transform ideas to provide a new derivation of the solution for the wave equation (cf. §2.4.1), based—surprisingly—upon the heat equation.

Suppose u is a bounded, smooth solution of the initial-value problem:

$$\begin{cases} u_{tt} - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g, \ u_t = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{32}$$

where n is odd and g is smooth, with compact support. We extend u to negative times by writing

$$u(x,t) = u(x,-t) \quad \text{if } x \in \mathbb{R}^n, \ t < 0. \tag{33}$$

Then

$$u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}.$$

Next define

$$v(x,t) := \frac{1}{(4\pi t)^{1/2}} \int_{-\infty}^{\infty} e^{-s^2/4t} u(x,s)\, ds \quad (x \in \mathbb{R}^n, \ t > 0). \tag{34}$$

Hence

$$\lim_{t \to 0} v = g \quad \text{uniformly on } \mathbb{R}^n.$$

In addition

$$\begin{aligned} \Delta v(x,t) &= \frac{1}{(4\pi t)^{1/2}} \int_{-\infty}^{\infty} e^{-s^2/4t} \Delta u(x,s)\, ds \\ &= \frac{1}{(4\pi t)^{1/2}} \int_{-\infty}^{\infty} e^{-s^2/4t} u_{ss}(x,s)\, ds \\ &= \frac{1}{(4\pi t)^{1/2}} \int_{-\infty}^{\infty} \frac{s}{2t} e^{-s^2/4t} u_s(x,s)\, ds \\ &= \frac{1}{(4\pi t)^{1/2}} \int_{-\infty}^{\infty} \left(\frac{s^2}{4t^2} - \frac{1}{2t} \right) e^{-s^2/4t} u(x,s)\, ds = v_t(x,t). \end{aligned}$$

Consequently v solves this initial-value problem for the heat equation:

$$\begin{cases} v_t - \Delta v = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ \qquad v = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

As v is bounded, we deduce from §2.3 that

$$v(x,t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\, dy. \tag{35}$$

We equate (34) with (35), recall (33), and set $\lambda = \frac{1}{4t}$, thereby obtaining the identity

$$\int_0^\infty u(x,s)e^{-\lambda s^2}\,ds = \frac{1}{2}\left(\frac{\lambda}{\pi}\right)^{\frac{n-1}{2}}\int_{\mathbb{R}^n} e^{-\lambda|x-y|^2}g(y)\,dy.$$

Thus

$$\int_0^\infty u(x,s)e^{-\lambda s^2}\,ds = \frac{n\alpha(n)}{2}\left(\frac{\lambda}{\pi}\right)^{\frac{n-1}{2}}\int_0^\infty e^{-\lambda r^2}r^{n-1}G(x;r)\,dr, \tag{36}$$

for all $\lambda > 0$, where

$$G(x;r) = ⨍_{\partial B(x,r)} g(y)\,dS(y). \tag{37}$$

We will solve (36), (37) for u. To do so, we write $n = 2k+1$ and note $-\frac{1}{2r}\frac{d}{dr}(e^{-\lambda r^2}) = \lambda e^{-\lambda r^2}$. Hence

$$\begin{aligned}
\lambda^{\frac{n-1}{2}}\int_0^\infty e^{-\lambda r^2}r^{n-1}G(x;r)\,dr &= \int_0^\infty \lambda^k e^{-\lambda r^2}r^{2k}G(x;r)\,dr \\
&= \frac{(-1)^k}{2^k}\int_0^\infty \left[\left(\frac{1}{r}\frac{d}{dr}\right)^k (e^{-\lambda r^2})\right] r^{2k}G(x;r)\,dr \\
&= \frac{1}{2^k}\int_0^\infty r\left[\left(\frac{1}{r}\frac{\partial}{\partial r}\right)^k (r^{2k-1}G(x;r))\right] e^{-\lambda r^2}\,dr,
\end{aligned}$$

where we integrated by parts k times for the last equality.

Owing to (36) (with r replacing s in the expression on the left), we deduce

$$\int_0^\infty u(x,r)e^{-\lambda r^2}dr = \frac{n\alpha(n)}{\pi^{\frac{n-1}{2}}2^{k+1}}\int_0^\infty r\left[\left(\frac{1}{r}\frac{\partial}{\partial r}\right)^k (r^{2k-1}G(x;r))\right] e^{-\lambda r^2}dr.$$

Upon substituting $\tau = r^2$ we see that each side above, taken as a function of λ, is a Laplace transform. As two Laplace transforms agree only if the original functions were identical, we deduce

$$u(x,t) = \frac{n\alpha(n)}{\pi^k 2^{k+1}}\,t\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^k (t^{2k-1}G(x,t)). \tag{38}$$

Now $n = 2k+1$ and $\alpha(n) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} = \frac{\pi^{k+\frac{1}{2}}}{\Gamma(\frac{n}{2}+1)}$. Since $\Gamma\left(\frac{1}{2}\right) = \pi^{1/2}$ and $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$ (cf. [**RD,** Chapter 8]), we can compute

$$\frac{n\alpha(n)}{\pi^k 2^{k+1}} = \frac{n\pi^{1/2}}{2^{k+1}\Gamma\left(\frac{n}{2}+1\right)} = \frac{1}{(n-2)(n-4)\cdots 5\cdot 3} = \frac{1}{\gamma_n}.$$

We insert this deduction into (38) and simplify:

$$(39)\qquad u(x,t) = \frac{1}{\gamma_n}\frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-3}{2}}\left(t^{n-2}\fint_{\partial B(x,t)} g\,dS\right)\quad (x\in\mathbb{R}^n,\ t>0).$$

This is formula (31) in §2.4.1 (for $h \equiv 0$). □

4.4. CONVERTING NONLINEAR INTO LINEAR PDE

In this section we describe several techniques which are sometime useful for converting certain nonlinear equations into linear equations.

4.4.1. Hopf–Cole transformation.

a. A parabolic PDE with quadratic nonlinearity.

We consider first of all an initial-value problem for a quasilinear parabolic equation:

$$(1)\qquad \begin{cases} u_t - a\Delta u + b|Du|^2 = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases}$$

where $a > 0$. This sort of nonlinear PDE arises in stochastic optimal control theory.

Assuming for the moment u is a smooth solution of (1), we set

$$w := \phi(u),$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a smooth function, as yet unspecified. We will try to choose ϕ so that w solves a linear equation. We have

$$w_t = \phi'(u)u_t,\ \Delta w = \phi'(u)\Delta u + \phi''(u)|Du|^2;$$

and consequently (1) implies

$$\begin{aligned} w_t = \phi'(u)u_t &= \phi'(u)[a\Delta u - b|Du|^2] \\ &= a\Delta w - [a\phi''(u) + b\phi'(u)]|Du|^2 \\ &= a\Delta w, \end{aligned}$$

provided we choose ϕ to satisfy $a\phi'' + b\phi' = 0$. We solve this differential equation by setting $\phi = e^{\frac{-bz}{a}}$. Thus we see that if u solves (1), then

$$(2)\qquad w = e^{\frac{-bu}{a}}$$

solves this initial-value problem for the heat equation (with conductivity a):

$$(3)\qquad \begin{cases} w_t - a\Delta w = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ w = e^{\frac{-bg}{a}} & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

Formula (2) is the *Hopf–Cole transformation.*

Now the unique bounded solution of (3) is

$$w(x,t) = \frac{1}{(4\pi a t)^{n/2}} \int_{\mathbb{R}^n} e^{\frac{-|x-y|^2}{4at}} e^{\frac{-b}{a}g(y)}\,dy \quad (x \in \mathbb{R}^n,\ t > 0);$$

and, since (2) implies

$$u = -\frac{a}{b}\log w,$$

we obtain thereby the explicit formula

$$(4)\quad u(x,t) = -\frac{a}{b}\log\left(\frac{1}{(4\pi a t)^{n/2}} \int_{\mathbb{R}^n} e^{\frac{-|x-y|^2}{4at} - \frac{b}{a}g(y)}\,dy\right) \quad (x \in \mathbb{R}^n,\ t > 0)$$

for a solution of quasilinear initial-value problem (1).

b. Burgers' equation with viscosity.

As a further application, we examine now for $n = 1$ the initial-value problem for the *viscous Burgers' equation*:

$$(5)\qquad \begin{cases} u_t - a u_{xx} + u u_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times (0,\infty). \end{cases}$$

If we set

$$(6)\qquad w(x,t) := \int_{-\infty}^{x} u(y,t)\,dy$$

and

$$(7)\qquad h(x) := \int_{-\infty}^{x} g(y)\,dy$$

(cf. §3.4), we have

$$(8)\qquad \begin{cases} w_t - a w_{xx} + \frac{1}{2}w_x^2 = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ w = h & \text{on } \mathbb{R} \times \{t=0\}. \end{cases}$$

This is an equation of the form (1) for $n = 1$, $b = \frac{1}{2}$; and so (4) provides the formula

$$(9)\qquad w(x,t) = -2a\log\left(\frac{1}{(4\pi a t)^{1/2}} \int_{\mathbb{R}} e^{\frac{-|x-y|^2}{4at} - \frac{h(y)}{2a}}\,dy\right).$$

But then since $u = w_x$, we find upon differentiating (9) that

$$(10)\qquad u(x,t) = \frac{\int_{-\infty}^{\infty} \frac{x-y}{t} e^{\frac{-|x-y|^2}{4at} - \frac{h(y)}{2a}}\,dy}{\int_{-\infty}^{\infty} e^{\frac{-|x-y|^2}{4at} - \frac{h(y)}{2a}}\,dy} \quad (x \in \mathbb{R},\ t > 0)$$

is a solution of problem (5), where h is defined by (7). We will scrutinize this formula further in §4.5.2.

4.4.2. Potential functions.

Another technique is to utilize a *potential function* to convert a nonlinear system of PDE into a single linear PDE. We consider as an example *Euler's equations* for inviscid, incompressible fluid flow:

$$(11)\qquad \begin{cases} \text{(a)}\ \ \mathbf{u}_t + \mathbf{u}\cdot D\mathbf{u} = -Dp + \mathbf{f} & \text{in } \mathbb{R}^3 \times (0,\infty) \\ \text{(b)}\qquad\qquad \operatorname{div}\mathbf{u} = 0 & \text{in } \mathbb{R}^3 \times (0,\infty) \\ \text{(c)}\qquad\qquad\quad \mathbf{u} = \mathbf{g} & \text{on } \mathbb{R}^3 \times \{t=0\}. \end{cases}$$

Here the unknowns are the *velocity* field $\mathbf{u} = (u^1, u^2, u^3)$ and the scalar *pressure* p; the external force $\mathbf{f} = (f^1, f^2, f^3)$ and initial velocity $\mathbf{g} = (g^1, g^2, g^3)$ are given. Here D as usual denotes the gradient in the spatial variables $x = (x_1, x_2, x_3)$. The vector equation 11(a) means

$$u^i_t + \sum_{j=1}^{3} u^j u^i_{x_j} = -p_{x_i} + f^i \quad (i = 1,2,3).$$

We will assume

$$(12)\qquad \operatorname{div}\mathbf{g} = 0.$$

If furthermore there exists a scalar function $h : \mathbb{R}^3 \times (0,\infty) \to \mathbb{R}$ such that

$$(13)\qquad \mathbf{f} = Dh,$$

we say that the external force is derived from the *potential* h.

We will try to find a solution $(\mathbf{u}, p)$ of (11) for which the velocity field $\mathbf{u}$ is also derived from a potential, say

$$(14)\qquad \mathbf{u} = Dv.$$

Our flow will then be irrotational, as $\operatorname{curl}\mathbf{u} \equiv 0$. Now equation (11)(b) says

$$(15)\qquad 0 = \operatorname{div}\mathbf{u} = \Delta v$$

and so v must be harmonic as a function of x, for each time $t > 0$. Thus if we can find a smooth function v satisfying (15) and $Dv(\cdot, 0) = g$, we can then recover $\mathbf{u}$ from v by (14).

How do we compute the pressure p? Let us observe that if $\mathbf{u} = Dv$, then $\mathbf{u}\cdot D\mathbf{u} = \frac{1}{2}D(|Dv|^2)$. Consequently (11)(a) reads $D\left(v_t + \frac{1}{2}|Dv|^2\right) = D(-p+h)$, in view of (13). Therefore we may take

$$(16)\qquad v_t + \frac{1}{2}|Dv|^2 + p = h.$$

This is *Bernoulli's law.* But now we can employ (16) to compute p, since v and h are already known.

4.4.3. Hodograph and Legendre transforms.

a. Hodograph transform.

The *hodograph transform* is a technique for converting certain quasilinear systems of PDE into linear systems, by reversing the roles of the dependent and independent variables. As this method is most easily understood by an example, we investigate here the equations of steady, two-dimensional, irrotational fluid flow:

$$(17)\qquad \begin{cases} \text{(a)}\ (\sigma^2(\mathbf{u})-(u^1)^2)u^1_{x_1}-u^1u^2(u^1_{x_2}+u^2_{x_1}) \\ \qquad\qquad\qquad\qquad +(\sigma^2(\mathbf{u})-(u^2)^2)u^2_{x_2}=0 \\ \text{(b)}\ u^1_{x_2}-u^2_{x_1}=0 \end{cases}$$

in $\mathbb{R}^2$. The unknown is the *velocity* field $\mathbf{u}=(u^1,u^2)$, and the function $\sigma(\cdot):\mathbb{R}^2\to\mathbb{R}$, the local *sound speed*, is given.

The system (17) is quasilinear. Let us now, however, no longer regard u^1 and u^2 as functions of x_1 and x_2:

$$(18)\qquad u^1=u^1(x_1,x_2),\ u^2=u^2(x_1,x_2),$$

but rather regard x^1 and x^2 as functions of u_1 and u_2:

$$(19)\qquad x^1=x^1(u_1,u_2),\ x^2=x^2(u_1,u_2).$$

We have exchanged sub- and superscripts in the notation to emphasize the interchange between independent and dependent variables.

According to the Inverse Function Theorem (§C.5) we can, locally at least, invert equations (18) to yield (19), provided

$$(20)\qquad J=\frac{\partial(u^1,u^2)}{\partial(x_1,x_2)}=u^1_{x_1}u^2_{x_2}-u^1_{x_2}u^2_{x_1}\neq 0$$

in some region of $\mathbb{R}^2$. Assuming now (20) holds, we calculate

$$(21)\qquad \begin{cases} u^2_{x_2}=Jx^1_{u_1},\ u^2_{x_1}=-Jx^2_{u_1} \\ u^1_{x_2}=-Jx^1_{u_2},\ u^1_{x_1}=Jx^2_{u_2}. \end{cases}$$

We insert (21) into (17), to discover

$$(22)\qquad \begin{cases} \text{(a)}\ (\sigma^2(u)-u_1^2)x^2_{u_2}+u_1u_2(x^1_{u_2}+x^2_{u_1})+(\sigma^2(u)-u_2^2)x^1_{u_1}=0 \\ \text{(b)}\ x^1_{u_2}-x^2_{u_1}=0. \end{cases}$$

This is a *linear* system for $\mathbf{x}=(x^1,x^2)$, as a function of $u=(u_1,u_2)$.

Remark. We can utilize the method of potential functions (§4.4.2) to simplify (22) further. Indeed, equation (22)(b) suggests that we look for a single function $z = z(u)$ such that

$$\begin{cases} x^1 = z_{u_1} \\ x^2 = z_{u_2}. \end{cases}$$

Then (22)(a) transforms into the linear, second-order PDE

$$(23) \qquad (\sigma^2(u) - u_1^2) z_{u_2 u_2} + 2u_1 u_2 z_{u_1 u_2} + (\sigma^2(u) - u_2^2) z_{u_1 u_1} = 0.$$

□

b. Legendre transform.

A technique closely related to the hodograph transform is the classical *Legendre transform*, a version of which we have already encountered before, in §3.3. The idea is to regard the components of the gradient of a solution as new independent variables.

Once again an example is instructive. We investigate the *minimal surface equation* (cf. Example 4 in §8.1.2)

$$\operatorname{div}\left(\frac{Du}{(1+|Du|^2)^{1/2}}\right) = 0,$$

which for $n = 2$ may be rewritten as

$$(24) \qquad (1 + u_{x_2}^2) u_{x_1 x_1} - 2u_{x_1} u_{x_2} u_{x_1 x_2} + (1 + u_{x_1}^2) u_{x_2 x_2} = 0.$$

Let us now assume that at least in some region of $\mathbb{R}^2$, we can invert the relations

$$(25) \qquad p^1 = u_{x_1}(x_1, x_2), \ p^2 = u_{x_2}(x_1, x_2),$$

to solve for

$$(26) \qquad x^1 = x^1(p_1, p_2), \ x^2 = x^2(p_1, p_2).$$

The Inverse Function Theorem assures us we can do so in a neighborhood of any point where

$$(27) \qquad J = \det D^2 u \neq 0.$$

Now define

$$v(p) := \mathbf{x}(p) \cdot p - u(\mathbf{x}(p)), \tag{28}$$

where $\mathbf{x} = (x^1, x^2)$ is given by (26), $p = (p_1, p_2)$. We discover after some calculations that

$$\begin{cases} u_{x_1x_1} = Jv_{p_2p_2} \\ u_{x_1x_2} = -Jv_{p_1p_2} \\ u_{x_2x_2} = Jv_{p_1p_1}. \end{cases} \tag{29}$$

Upon substituting the identities (29) into (24), we derive for v the *linear* equation

$$(1 + p_2^2)v_{p_2p_2} + 2p_1p_2v_{p_1p_2} + (1 + p_1^2)v_{p_1p_1} = 0. \tag{30}$$

Remark. The hodograph and Legendre transform techniques for obtaining linear out of nonlinear PDE are in practice tricky to use, as it is usually not possible to transform given boundary conditions very easily. □

4.5. ASYMPTOTICS

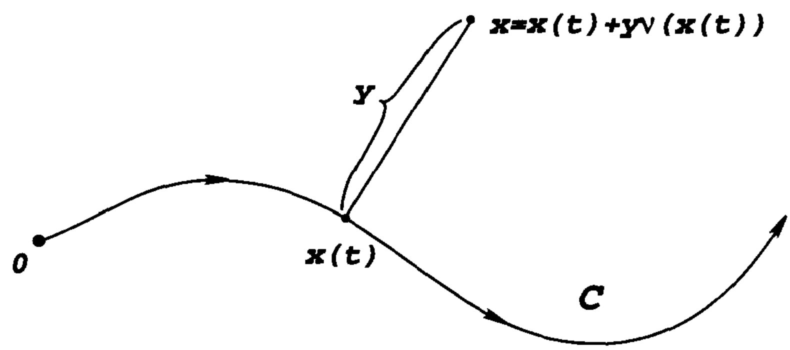
It is often the case that even when explicit representation formulas can be had for solutions of partial differential equations, these are too complicated to be of much immediate use. In such circumstances it sometimes becomes profitable to study the formulas in various asymptotic limits, whereupon simplifications often appear.

Following are several rather complicated examples, illustrating typical issues involved in asymptotics for PDE. The results in this section are explained only heuristically, mostly without formal proofs.

4.5.1. Singular perturbations.

A *singular perturbation* is a modification of a given PDE by adding a small multiple ε times a higher order term. In accordance with the informal principle that the behavior of solutions is governed primarily by the highest order terms, a solution u^ε of the perturbed problem will often behave analytically quite differently from a solution u of the original equation.

Example 1 (Transport and small diffusion). We illustrate this idea by studying formally the effects of small diffusion upon the transport of dye within a moving fluid in $\mathbb{R}^2$.



Flow of dye without diffusion

Suppose we are given a smooth vector field $\mathbf{b} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{b} = (b^1, b^2)$, representing the steady fluid velocity. Assume dye has been continuously injected at unit rate into the fluid at the origin, and let $u(x)$ represent the density of dye at the point $x \in \mathbb{R}^2$, $x = (x_1, x_2)$. Then, formally at least as we shall see,

$$\operatorname{div}(u\mathbf{b}) = \delta_0 \quad \text{in } \mathbb{R}^2, \tag{1}$$

where δ_0 is the Dirac measure on $\mathbb{R}^2$ giving unit mass to the point 0. This PDE implies that the dye density is transported with the fluid motion at points $x \neq 0$.

Consider now for $\varepsilon > 0$ the singular perturbation:

$$-\varepsilon \Delta u^\varepsilon + \operatorname{div}(u^\varepsilon \mathbf{b}) = \delta_0 \quad \text{in } \mathbb{R}^2. \tag{2}$$

The new term "$\varepsilon\Delta$" represents a small, isotropic diffusion of the dye within the background fluid motion. We are interested in understanding in an approximate way the structure of the solution u^ε of (2) and, in particular, describing if and how u^ε approximates u for small $\varepsilon > 0$.

a. Analysis of problem (1).

We turn our attention first to the unperturbed PDE (1). Consider the characteristic ODE

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{b}(\mathbf{x}(t)) & (t \geq 0) \\ \mathbf{x}(0) = 0, \end{cases} \tag{3}$$

the solution $\mathbf{x}(t) = (x^1(t), x^2(t))$ of which we assume to trace out a curve C, as drawn.

Given a point $x \in \mathbb{R}^2$ near C, we write

$$x = \mathbf{x}(t) + y\boldsymbol{\nu}(\mathbf{x}(t)), \tag{4}$$

where $\boldsymbol{\nu} = (\nu^1, \nu^2)$ is the (upward pointing) unit normal to C, $y \in \mathbb{R}$, and t is the time required for the solution of the ODE (3) to reach the point $\mathbf{x}(t)$ along C closest to x. We hereafter regard (t, y) as providing a new coordinate system near the curve C; so that $x = (x^1(y,t), x^2(y,t))$.

Using (3) and (4), we compute

$$\frac{\partial(x^1, x^2)}{\partial(t,y)} = \det \begin{pmatrix} \frac{\partial x^1}{\partial t} & \frac{\partial x^1}{\partial y} \\ \frac{\partial x^2}{\partial t} & \frac{\partial x^2}{\partial y} \end{pmatrix} = \det \begin{pmatrix} b^1 + y\dot{\nu}^1 & \nu^1 \\ b^2 + y\dot{\nu}^2 & \nu^2 \end{pmatrix}.$$

Let us write $\sigma = |\mathbf{b}|$, $\boldsymbol{\nu} = (-b^2, b^1)/\sigma$ and $\dot{\boldsymbol{\nu}} = -\sigma\kappa\boldsymbol{\tau} = -\kappa\mathbf{b}$ (where σ = speed, κ = curvature, $\boldsymbol{\tau} = \frac{\mathbf{b}}{\sigma}$ = unit tangent). We then simplify, to obtain

$$\frac{\partial(x^1, x^2)}{\partial(t,y)} = \sigma(1 - \kappa y). \tag{5}$$

Return now to the PDE (1), which we rewrite to read

$$\mathbf{b} \cdot Du + (\operatorname{div} \mathbf{b})u = \delta_0 \quad \text{in } \mathbb{R}^2. \tag{6}$$

As in §3.2 we see $u \equiv 0$ off the curve C. Let us next guess u has the form

$$u(x) = \rho(t)\delta(y) \tag{7}$$

in the (t, y)-coordinates, δ denoting the Dirac measure on $\mathbb{R}$ giving unit mass to the origin.

What is $\rho(t)$? To compute it, take R to be a small, smooth region in the (x_1, x_2)-plane, with boundary intersecting the curve C at the points $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$, $0 < t_1 < t_2$. Let R' denote the corresponding region in the (t, y)-plane. Then using (5) we calculate

$$\int_R u\,dx = \int_{R'} \rho(t)\delta(y)\sigma(t)(1 - \kappa y)\,dydt = \int_{t_1}^{t_2} \rho(t)\sigma(t)\,dt.$$

Now $\int_R u\,dx$ represents the total amount of dye within the region R, which is to say, the total amount released between times t_1 and t_2. This is simply $t_2 - t_1$. Thus

$$\int_{t_1}^{t_2} \rho(t)\sigma(t)\,dt = t_2 - t_1.$$

This identity holds for all $0 < t_1 < t_2$, and so $\rho(t) = \sigma(t)^{-1}$. Hence (7) says

$$u(x,t) = \delta(y)/\sigma(t) \tag{8}$$

is a solution of (1), for $\sigma(t) := |\mathbf{b}(\mathbf{x}(t))|$, $t \geq 0$. In other words, u represents the density along the curve C of the dye, whose concentration varies inversely with the speed of the fluid.

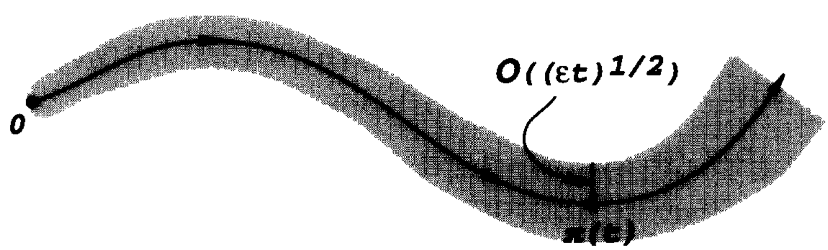
We can further confirm this formula as follows. Let $v \in C_c^\infty(\mathbb{R}^2)$. Then using (5), we compute

$$\begin{aligned}\int_{\mathbb{R}^2} Dv \cdot \mathbf{b} u \, dx &= \int_{\mathbb{R}^2} Dv \cdot \mathbf{b} \frac{\delta(y)}{\sigma(t)} \sigma(t)(1 - \kappa y) \, dy dt \\ &= \int_0^\infty Dv(\mathbf{x}(t)) \cdot \mathbf{b}(\mathbf{x}(t)) \, dt \\ &= \int_0^\infty \frac{d}{dt} v(\mathbf{x}(t)) \, dt = -v(0).\end{aligned}$$

Hence we may indeed interpret u defined by (8) as a weak solution of the unperturbed PDE (1).

b. Analysis of problem (2) for $0 < \varepsilon << 1$.

We look now at the perturbed problem (2). We expect that at time $t > 0$, the diffusing dye will fill a ball of radius approximately $O((\varepsilon t)^{1/2})$ about the point $\mathbf{x}(t)$. The dye will thus be mostly concentrated in a plume as drawn, about the curve C.



Flow of dye with diffusion

We wish to understand the structure of the solution u^ε of (2) within this plume, as $\varepsilon \to 0$.

Since the width is presumably of order $O(\varepsilon^{1/2})$ for times $0 < t_1 \leq t \leq t_2$ and the total mass of dye corresponding to the same time interval is $t_2 - t_1$, we expect u^ε to be of order $O(\varepsilon^{-1/2})$ along C. This suggests that for us to understand the asymptotics as $\varepsilon \to 0$, we should turn our attention to the *rescaled variables*

(9) $$z := \varepsilon^{-1/2} y, \; v^\varepsilon := \varepsilon^{1/2} u^\varepsilon,$$

the powers of ε selected so that $z, v^\varepsilon = O(1)$.

We must therefore rewrite the PDE (2) in terms of the new variables t, z and v^ε. For this, we need first to study the structure of the velocity field $\mathbf{b}$ along the curve C. Let us therefore write

$$\mathbf{b} = \sigma(t)\boldsymbol{\tau} + \{\alpha(t)\boldsymbol{\tau} + \beta(t)\boldsymbol{\nu}\}y + O(y^2), \tag{10}$$

where, as noted before, $\boldsymbol{\tau} = \mathbf{b}/\sigma$ is a unit tangent vector to C. Now for any smooth function w:

$$w_t = w_{x_1}\frac{\partial x^1}{\partial t} + w_{x_2}\frac{\partial x^2}{\partial t} = w_{x_1}\sigma(1-\kappa y)\tau^1 + w_{x_2}\sigma(1-\kappa y)\tau^2$$

and

$$w_y = w_{x_1}\frac{\partial x^1}{\partial y} + w_{x_2}\frac{\partial x^2}{\partial y} = w_{x_1}\nu^1 + w_{x_2}\nu^2.$$

Thus

$$\begin{cases} w_{x_1} = \dfrac{w_t\nu^2 - w_y\sigma(1-\kappa y)\tau^2}{\sigma(1-\kappa y)}, \\ w_{x_2} = \dfrac{-w_t\nu^1 + w_y\sigma(1-\kappa y)\tau^1}{\sigma(1-\kappa y)}. \end{cases} \tag{11}$$

Therefore using (10), (11) (for $w = u^\varepsilon$), we can compute:

$$\begin{aligned} \mathbf{b}\cdot Du^\varepsilon &= [\sigma\boldsymbol{\tau} + (\alpha\boldsymbol{\tau} + \beta\boldsymbol{\nu})y + O(y^2)]\cdot Du^\varepsilon \\ &= \frac{u_t^\varepsilon}{(1-\kappa y)} + \frac{\alpha y u_t^\varepsilon}{\sigma(1-\kappa y)} + \beta y u_y^\varepsilon + O(y^2|Du^\varepsilon|). \end{aligned}$$

Since $v^\varepsilon = \varepsilon^{1/2}u^\varepsilon$, $z = \varepsilon^{-1/2}y$, we can rewrite the foregoing as

$$\mathbf{b}\cdot Dv^\varepsilon = v_t^\varepsilon + \beta z v_z^\varepsilon + O(\varepsilon^{1/2}). \tag{12}$$

Similarly, we calculate using (10) and (11) (for $w = b^1, b^2$) that

$$\begin{aligned} \operatorname{div}\mathbf{b} &= \frac{1}{\sigma(1-\kappa y)}[b_t^1\nu^2 - b_t^2\nu^1] + [\tau^1 b_y^2 - \tau^2 b_y^1] \\ &= \frac{1}{\sigma(1-\kappa y)}[(\sigma\tau^1)_t\nu^2 - (\sigma\tau^2)_t\nu^1] \\ &\quad + [\tau^1(\alpha\tau^2 + \beta\nu^2) - \tau^2(\alpha\tau^1 + \beta\nu^1)] + O(y) \\ &= \left(\frac{\dot\sigma}{\sigma} + \beta\right) + O(y). \end{aligned}$$

Here we used the identity $\dot{\boldsymbol{\tau}} = \sigma\kappa\boldsymbol{\nu}$. It follows that

$$(\operatorname{div}\mathbf{b})v^\varepsilon = \left(\frac{\dot\sigma}{\sigma} + \beta\right)v^\varepsilon + O(\varepsilon^{1/2}). \tag{13}$$

In addition, a similar heuristic argument, the details of which we omit, suggests that

$$\varepsilon \Delta u^\varepsilon = v^\varepsilon_{zz} + O(\varepsilon^{1/2}). \tag{14}$$

Combining now (12)–(14) and recalling (2), we at last deduce v^ε satisfies

$$v^\varepsilon_t - v^\varepsilon_{zz} + (\beta z v^\varepsilon)_z + \frac{\dot{\sigma}}{\sigma} v^\varepsilon = O(\varepsilon^{1/2}). \tag{15}$$

We suppose now that as $\varepsilon \to 0$, the functions v^ε converge in some sense to a limit:

$$v^\varepsilon \to v \quad \text{in } \mathbb{R}^2. \tag{16}$$

Then presumably from (15) we will have

$$v_t - v_{zz} + (\beta z v)_z + \frac{\dot{\sigma}}{\sigma} v = 0 \quad \text{in } \mathbb{R} \times (0, \infty). \tag{17}$$

We therefore expect

$$u^\varepsilon = \varepsilon^{-1/2} v^\varepsilon = \varepsilon^{-1/2}(v + o(1)), \tag{18}$$

with v solving (17). The PDE (17) is consequently a *parabolic approximation* (in the variables t, $z = \varepsilon^{-1/2} y$) to our elliptic equation (2). The proper initial condition should be

$$v = \frac{\delta(z)}{\sigma(0)} \quad \text{on } \mathbb{R} \times \{t = 0\}. \tag{19}$$

We will see in Problem 6 that an explicit solution of (18), (19) can be found, in terms of the solution of an ODE involving β.

4.5.2. Laplace's method.

Laplace's method concerns the asymptotics as $\varepsilon \to 0$ of integrals involving expressions of the form $e^{-I/\varepsilon}$, I denoting some given function.

Example 2 (Vanishing viscosity method for Burgers' equation). We next investigate the limit as $\varepsilon \to 0$ of the solution u^ε of the initial-value problem for the viscous Burgers' equation

$$\begin{cases} u^\varepsilon_t + u^\varepsilon u^\varepsilon_x - \varepsilon u^\varepsilon_{xx} = 0 & \text{in } \mathbb{R} \times (0, \infty) \\ u^\varepsilon = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{20}$$

Remembering formula (10) from §4.4.1, we note

$$u^\varepsilon(x,t) = \frac{\int_{-\infty}^{\infty} \frac{x-y}{t} e^{\frac{-K(x,y,t)}{2\varepsilon}}\,dy}{\int_{-\infty}^{\infty} e^{\frac{-K(x,y,t)}{2\varepsilon}}\,dy}, \tag{21}$$

for

$$K(x,y,t) := \frac{|x-y|^2}{2t} + h(y) \quad (x, y \in \mathbb{R},\ t > 0), \tag{22}$$

where h is an antiderivative of g.

What happens to u^ε as $\varepsilon \to 0$? Mathematically the term "εu_{xx}" in (20) makes the partial differential equation act somewhat like the heat equation, in that the solution u^ε is infinitely differentiable in $\mathbb{R}^n \times (0,\infty)$, in spite of the nonlinearity. This follows from the explicit formula (21). On the other hand, an obvious guess is that the solutions u^ε should converge as $\varepsilon \to 0$ to a solution u of the conservation law

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{23}$$

Physically, we regard the term "$\varepsilon u^\varepsilon_{xx}$" as imposing an "artificial viscosity" effect, which we are now sending to zero. We expect that this *vanishing viscosity* technique should allow us to recover the correct entropy solution u of (23), which may have discontinuities across shock waves, as the limit of the solutions u^ε of (20), which are smooth.

We must understand the limiting behavior of the expression on the right hand side of (21), as $\varepsilon \to 0$.

LEMMA (Asymptotics). *Suppose that $k, l : \mathbb{R} \to \mathbb{R}$ are continuous functions, that l grows at most linearly and that k grows at least quadratically. Assume also there exists a unique point $y_0 \in \mathbb{R}$ such that*

$$k(y_0) = \min_{y \in \mathbb{R}} k(y).$$

Then

$$\lim_{\varepsilon \to 0} \frac{\int_{-\infty}^{\infty} l(y) e^{\frac{-k(y)}{\varepsilon}}\,dy}{\int_{-\infty}^{\infty} e^{\frac{-k(y)}{\varepsilon}}\,dy} = l(y_0). \tag{24}$$

Proof. Write $k_0 = k(y_0)$. Then the function

$$\mu_\varepsilon(y) := \frac{e^{\frac{k_0 - k(y)}{\varepsilon}}}{\int_{-\infty}^{\infty} e^{\frac{k_0 - k(z)}{\varepsilon}}\, dz} \quad (y \in \mathbb{R})$$

satisfies

$$\text{(25)} \qquad \begin{cases} \mu_\varepsilon \geq 0, \ \int_{-\infty}^{\infty} \mu_\varepsilon(y)\, dy = 1, \\ \mu_\varepsilon(y) \to 0 \text{ exponentially fast for } y \neq y_0, \text{ as } \varepsilon \to 0. \end{cases}$$

Consequently

$$\lim_{\varepsilon \to 0} \frac{\int_{-\infty}^{\infty} l(y) e^{\frac{-k(y)}{\varepsilon}}\, dy}{\int_{-\infty}^{\infty} e^{\frac{-k(y)}{\varepsilon}}\, dy} = \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} l(y)\mu_\varepsilon(y)\, dy = l(y_0).$$

□

Return now to (21), (22). We observe $K(x, y, t) = tL\left(\frac{x-y}{t}\right) + h(y)$, where $L = F^*$ for $F(z) = \frac{z^2}{2}$. According to the analysis in §3.4, for each time $t > 0$ the mapping $y \mapsto K(x, y, t)$ attains its minimum at a unique point $y = y(x, t)$ for all but at most countably many points x. But then the lemma implies

$$\text{(26)} \qquad \lim_{\varepsilon \to 0} u^\varepsilon(x, t) = \frac{x - y(x, t)}{t} = G\left(\frac{x - y(x, t)}{t}\right) = u(x, t)$$

for $G := (F')^{-1}$.

The final equality in (26) is the Lax–Oleinik formula for the unique entropy solution of the initial-value problem (23). It is a powerful endorsement of the methods from §3.4 that this formula has reappeared in the context of vanishing viscosity. (See also Problem 3.) □

We will later discuss the vanishing viscosity method for symmetric hyperbolic systems in §7.3.2, for Hamilton–Jacobi equations in §10.1, and for systems of conservation laws in §11.4.

4.5.3. Geometric optics, stationary phase.

This section investigates the behavior of certain highly oscillatory solutions of the wave equation. We begin with some crude, but instructive, calculations.

a. Geometric optics.

Example 3 (Oscillating solutions). Let us once more turn our attention to the wave equation

$$u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty), \tag{27}$$

and we now regard the solution u as taking complex values. We fix $\varepsilon > 0$ and seek a solution $u = u^\varepsilon$ of (27) having the form

$$u^\varepsilon(x,t) = e^{\frac{ip^\varepsilon(x,t)}{\varepsilon}} a^\varepsilon(x,t) \quad (x \in \mathbb{R}^n, t \geq 0), \tag{28}$$

the real-valued function p^ε representing the *phase*, and the real-valued function a^ε representing the *amplitude*. The proposed form (28) for the solution is called the *geometric optics ansatz**. The idea is that highly oscillatory solutions of the wave equation can be understood by studying a PDE for the phase function in the limit as $\varepsilon \to 0$. Following is a formal demonstration.

Substituting (28) into (27), we find after some computations that

$$\begin{aligned} 0 = u^\varepsilon_{tt} - \Delta u^\varepsilon &= e^{ip^\varepsilon/\varepsilon}\left(\frac{ip^\varepsilon_{tt}}{\varepsilon}a^\varepsilon - \left(\frac{p^\varepsilon_t}{\varepsilon}\right)^2 a^\varepsilon + 2\frac{ip^\varepsilon_t a^\varepsilon_t}{\varepsilon} + a^\varepsilon_{tt}\right) \\ &\quad - e^{ip^\varepsilon/\varepsilon}\left(\frac{i\Delta p^\varepsilon}{\varepsilon}a^\varepsilon - \frac{|Dp^\varepsilon|^2}{\varepsilon^2}a^\varepsilon + \frac{2iDp^\varepsilon \cdot Da^\varepsilon}{\varepsilon} + \Delta a^\varepsilon\right). \end{aligned}$$

We cancel the term $e^{ip^\varepsilon/\varepsilon}$ and take the real part of the resulting expression, to find

$$a^\varepsilon((p^\varepsilon_t)^2 - |Dp^\varepsilon|^2) = \varepsilon^2(a^\varepsilon_{tt} - \Delta a^\varepsilon). \tag{29}$$

Now *if* as $\varepsilon \to 0$

$$p^\varepsilon \to p, \ a^\varepsilon \to a \neq 0 \tag{30}$$

in some sense, then presumably from (29) it follows that

$$p_t \pm |Dp| = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty). \tag{31}$$

We may informally regard the straight line characteristics of these Hamilton–Jacobi PDE as rays along which the solution (28) concentrates in the high-frequency limit as $\varepsilon \to 0$.

More generally, let us consider the second-order hyperbolic PDE

$$u_{tt} - \sum_{k,l=1}^{n} a^{kl}(x)u_{x_k x_l} = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty) \tag{32}$$

**ansatz* = formulation (German).

with $a^{kl} = a^{lk}$ $(k, l = 1, \dots, n)$. We again look for a complex-valued solution $u = u^\varepsilon$ of the form (28), and calculate

$$\begin{aligned} 0 &= u^\varepsilon_{tt} - \sum_{k,l=1}^n a^{kl} u^\varepsilon_{x_k x_l} \\ &= e^{ip^\varepsilon/\varepsilon}\left(\frac{ip^\varepsilon_{tt}}{\varepsilon}a^\varepsilon - \left(\frac{p^\varepsilon_t}{\varepsilon}\right)^2 a^\varepsilon + \frac{2ip^\varepsilon_t a^\varepsilon_t}{\varepsilon} + a^\varepsilon_{tt}\right) \\ &\quad - e^{ip^\varepsilon/\varepsilon}\left(\sum_{k,l=1}^n a^{kl}\left(\frac{ip^\varepsilon_{x_k x_l}}{\varepsilon}a^\varepsilon - \frac{p^\varepsilon_{x_k}p^\varepsilon_{x_l}}{\varepsilon^2}a^\varepsilon + \frac{2ip^\varepsilon_{x_k}a^\varepsilon_{x_l}}{\varepsilon} + a^\varepsilon_{x_k x_l}\right)\right). \end{aligned}$$

We once again cancel $e^{ip^\varepsilon/\varepsilon}$ and take real parts to find

$$a^\varepsilon\left((p^\varepsilon_t)^2 - \sum_{k,l=1}^n a^{kl} p^\varepsilon_{x_k} p^\varepsilon_{x_l}\right) = \varepsilon^2\left(a^\varepsilon_{tt} - \sum_{k,l=1}^n a^{kl} a^\varepsilon_{x_k x_l}\right).$$

Hence if (30) holds in some sense, we may then expect

$$\text{(33)} \qquad p_t \pm \left(\sum_{k,l=1}^n a^{kl} p_{x_k} p_{x_l}\right)^{1/2} = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty).$$

See below and also §4.6.1, §7.2.4 for further elaboration of these ideas. □

b. Stationary phase.

The foregoing example suggests that the Hamilton–Jacobi PDE (31) somehow "controls the high frequency asymptotics for the wave equation". However the range of validity of the geometric optics ansatz is highly uncertain in the preceding strictly formal computations. To understand more clearly the behavior of the solution, we employ next the method of *stationary phase*, which is a variant of Laplace's method had by replacing the -1 in the exponent (cf. §4.5.2) with i.

Example 4 (Stationary phase for the wave equation). Look again at the initial-value problem for the wave equation

$$\text{(34)} \qquad \begin{cases} u^\varepsilon_{tt} - \Delta u^\varepsilon = 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ u^\varepsilon = g^\varepsilon, \ u^\varepsilon_t = 0 & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases}$$

where we hereafter assume g^ε to have the rapidly oscillating structure

$$\text{(35)} \qquad g^\varepsilon(x) = a(x) e^{\frac{ip(x)}{\varepsilon}} \quad (x \in \mathbb{R}^n).$$

Here $\varepsilon > 0$, $a, p \in C_c^\infty(\mathbb{R}^n)$, and we suppose

$$Dp \neq 0 \quad \text{on the support of } a. \tag{36}$$

Utilizing formula (26) from §4.3.2, we can write

$$u^\varepsilon(x,t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{\hat{g}^\varepsilon(y)}{2} (e^{i(x\cdot y + t|y|)} + e^{i(x\cdot y - t|y|)})\, dy \quad (x \in \mathbb{R}^n,\ t > 0).$$

Invoking (35), we see

$$u^\varepsilon(x,t) = \frac{1}{2}(I_+^\varepsilon(x,t) + I_-^\varepsilon(x,t)), \tag{37}$$

where

$$I_\pm^\varepsilon(x,t) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a(z) e^{i((x-z)\cdot y \pm t|y| + \frac{p(z)}{\varepsilon})}\, dydz.$$

Changing variables gives

$$I_\pm^\varepsilon(x,t) = \frac{1}{(2\pi\varepsilon)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a(z) e^{\frac{i}{\varepsilon}\phi_\pm(x,y,z,t)}\, dydz, \tag{38}$$

for

$$\phi_\pm(x,y,z,t) := (x-z)\cdot y \pm t|y| + p(z). \tag{39}$$

We want to study the asymptotics of $I_\pm^\varepsilon$ as $\varepsilon \to 0$. Let us pause in this example and develop some general machinery, which we will later apply to (38), (39). □

Example 4 motivates our considering general integral expressions of the form

$$I_\varepsilon := \int_{\mathbb{R}^n} e^{i\frac{\phi(y)}{\varepsilon}} a(y)\, dy \quad (x \in \mathbb{R}^n), \tag{40}$$

where a, ϕ are smooth functions, a has compact support, and $\varepsilon > 0$. We wish to understand the limiting behavior of I_ε as $\varepsilon \to 0$.

We first examine the special case that ϕ is linear in y:

LEMMA 1 (Asymptotics for linear terms). *Let $a \in C_c^\infty(\mathbb{R}^n)$ and $p \in \mathbb{R}^n$, $p \neq 0$. Then for $m = 1, 2, \ldots$*

$$\int_{\mathbb{R}^n} e^{\frac{i}{\varepsilon}p\cdot y} a(y)\, dy = O(\varepsilon^m) \quad \textit{as } \varepsilon \to 0.$$

Proof. Without loss we may assume $p = (p_1, \dots, p_n)$, $p_1 \neq 0$. Then for $m = 1, 2, \dots$

$$\begin{aligned}\int_{\mathbb{R}^n} e^{\frac{i}{\varepsilon}p\cdot y}a(y)\,dy &= \left(\frac{\varepsilon}{ip_1}\right)^m \int_{\mathbb{R}^m} \frac{\partial^m}{\partial y_1^m}(e^{\frac{i}{\varepsilon}p\cdot y})a(y)\,dy \\ &= \left(\frac{-\varepsilon}{ip_1}\right)^m \int_{\mathbb{R}^m} e^{\frac{i}{\varepsilon}p\cdot y}\frac{\partial^m}{\partial y_1^m}a(y)\,dy = O(\varepsilon^m).\end{aligned}$$

□

Next we suppose ϕ is quadratic in y:

LEMMA 2 (Asymptotics for quadratic terms). *Let $a \in C_c^\infty(\mathbb{R}^n)$ and suppose A is a real, nonsingular, symmetric matrix. Then*

$$(41)\qquad \frac{1}{(2\pi\varepsilon)^{n/2}}\int_{\mathbb{R}^n} e^{\frac{i}{2\varepsilon}y\cdot Ay}a(y)\,dy = \frac{e^{i\frac{\pi}{4}\operatorname{sgn} A}}{|\det A|^{1/2}}(a(0) + O(\varepsilon)) \qquad \textit{as } \varepsilon \to 0.$$

Here sgn A, the *signature* of A, denotes the number of positive eigenvalues of A minus the number of negative eigenvalues.

Proof. 1. First we claim for each $\phi \in C_c^\infty(\mathbb{R}^n)$ that

$$(42)\qquad \begin{aligned}\lim_{\delta\to 0^+}&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{ix\cdot Ax-\delta|x|^2-ix\cdot y}\phi(y)\,dxdy \\ &= \frac{\pi^{n/2}}{|\det A|^{1/2}}e^{i\frac{\pi}{4}\operatorname{sgn} A}\int_{\mathbb{R}^n} e^{-\frac{i}{4}y\cdot A^{-1}y}\phi(y)\,dy.\end{aligned}$$

To confirm this, we start by assuming A is diagonal:

$$(43)\qquad A = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \quad (\lambda_k \neq 0,\ k = 1, \dots, n).$$

Now for fixed $y, \lambda \in \mathbb{R}$ and $\delta > 0$, we have

$$\begin{aligned}\int_{\mathbb{R}} e^{i\lambda x^2-\delta x^2-ixy}dx &= e^{\frac{y^2}{4(i\lambda-\delta)}}\int_{\mathbb{R}} e^{(i\lambda-\delta)\left(x-\frac{iy}{2(i\lambda-\delta)}\right)^2}dx \\ &= \frac{e^{\frac{y^2}{4(i\lambda-\delta)}}}{(\delta-i\lambda)^{1/2}}\int_\Gamma e^{-z^2}dz,\end{aligned}$$

where $\Gamma = \{z = (\delta-i\lambda)^{1/2}(x-\frac{iy}{2(i\lambda-\delta)}) \mid x \in \mathbb{R}\}$ and we take $\operatorname{Re}(\delta-i\lambda)^{1/2} > 0$. Thus Γ is a line in the complex plane, which intersects the x-axis at an angle less than $\frac{\pi}{4}$. We consequently may deform the integral over Γ into the integral along the real axis: see Problem 7. Hence $\int_\Gamma e^{-z^2}dz = \int_{\mathbb{R}} e^{-x^2}dx = \pi^{1/2}$, and thus

$$\int_{\mathbb{R}} e^{i\lambda x^2-\delta x^2-ixy}dx = \frac{\pi^{1/2}}{(\delta - i\lambda)^{1/2}} e^{\frac{y^2}{4(i\lambda-\delta)}}.$$

Since A has the diagonal form (43), we consequently deduce

$$\begin{aligned} J_\delta(y) &:= \int_{\mathbb{R}^n} e^{ix\cdot Ax-\delta|x|^2-ix\cdot y}dx \\ &= \prod_{k=1}^n \int_{\mathbb{R}} e^{i\lambda_k x_k^2-\delta x_k^2-ix_k y_k}dx_k = \pi^{n/2}\prod_{k=1}^n \frac{e^{\frac{y_k^2}{4(i\lambda_k-\delta)}}}{(\delta-i\lambda_k)^{1/2}}. \end{aligned}$$

2. Now let $\phi \in C_c^\infty(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n}\phi(y)J_\delta(y)\,dy = \pi^{n/2}\int_{\mathbb{R}^n}\phi(y)\prod_{k=1}^n \frac{e^{\frac{y_k^2}{4(i\lambda_k-\delta)}}}{(\delta-i\lambda_k)^{1/2}}\,dy.$$

Applying the Dominated Convergence Theorem, we deduce

$$\lim_{\delta\to 0}\int_{\mathbb{R}^n}\phi(y)J_\delta(y)\,dy = \pi^{n/2}\int_{\mathbb{R}^n}\phi(y)\prod_{k=1}^n \frac{e^{-\frac{iy_k^2}{4\lambda_k}}}{(-i\lambda_k)^{1/2}}\,dy. \tag{44}$$

Recall we are supposing $\operatorname{Re}(-i\lambda_k)^{1/2} > 0$. Thus if $\lambda_k > 0$, $(-i\lambda_k)^{1/2} = |\lambda_k|^{1/2}e^{-\frac{i\pi}{4}}$. If instead $\lambda_k < 0$, then $(-i\lambda_k)^{1/2} = |\lambda_k|^{1/2}e^{\frac{i\pi}{4}}$. Therefore

$$\prod_{k=1}^\infty(-i\lambda_k)^{1/2} = |\det A|^{1/2}e^{-\frac{i\pi}{4}\operatorname{sgn} A},$$

and so (44) gives (42), provided A is diagonal.

If A is not diagonal, we rotate to new coordinates to diagonalize A, and again verify (42).

3. Let us now write $\alpha_\varepsilon(y) := e^{\frac{i}{2\varepsilon}y\cdot Ay}$. Then if $a \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} a(x)\alpha_\varepsilon(x)\,dx = \int_{\mathbb{R}^n}\hat{a}(y)\hat{\alpha}_\varepsilon(-y)\,dy.$$

According to (42) (with $\frac{1}{2\varepsilon}A$ replacing A):

$$\begin{aligned} \hat{\alpha}_\varepsilon(y) &= \frac{\varepsilon^{n/2}}{|\det A|^{1/2}}e^{\frac{i\pi}{4}\operatorname{sgn} A}e^{-\frac{i\varepsilon}{2}y\cdot A^{-1}y} \\ &= \frac{\varepsilon^{n/2}}{|\det A|^{1/2}}e^{\frac{i\pi}{4}\operatorname{sgn} A}(1+O(\varepsilon|y|^2)), \end{aligned}$$

$\hat{\alpha}_\varepsilon$ interpreted as in (42). Consequently

$$\frac{1}{\varepsilon^{n/2}} \int_{\mathbb{R}^n} a(x)\alpha_\varepsilon(x)\,dx = \frac{e^{\frac{i\pi}{4}\operatorname{sgn} A}}{|\det A|^{1/2}} \int_{\mathbb{R}^n} \hat{a}(y)(1 + O(\varepsilon|y|^2))\,dy.$$

But $\int_{\mathbb{R}^n} \hat{a}(y)\,dy = (2\pi)^{\frac{n}{2}} a(0)$ and $\int_{\mathbb{R}^n} \hat{a}(y)|y|^2 dy < \infty$. Formula (41) follows. □

For a general phase function ϕ, we will employ the following result to change variables and thereby convert locally to one of the earlier cases.

LEMMA 3 (Changing coordinates). *Assume* $\phi : \mathbb{R}^n \to \mathbb{R}$ *is smooth.*

(i) *Suppose that*

$$D\phi(0) \neq 0.$$

Then there exists a smooth function $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$ *such that*

$$(45) \qquad \begin{cases} \mathbf{\Phi}(0) = 0, \ D\mathbf{\Phi}(0) = I, \ \textit{and} \\ \phi(\mathbf{\Phi}(x)) = \phi(0) + D\phi(0) \cdot x \quad \textit{for } |x| \textit{ small.} \end{cases}$$

(ii) (Morse Lemma) *Suppose instead that*

$$D\phi(0) = 0, \ \det D^2\phi(0) \neq 0.$$

Then there exists a smooth function $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$ *such that*

$$(46) \qquad \begin{cases} \mathbf{\Phi}(0) = 0, \ D\mathbf{\Phi}(0) = I, \ \textit{and} \\ \phi(\mathbf{\Phi}(x)) = \phi(0) + \frac{1}{2}x \cdot D^2\phi(0)x \quad \textit{for } |x| \textit{ small.} \end{cases}$$

In other words, we can change variables near 0 to make ϕ affine in case (i), quadratic in case (ii).

Proof. 1. Assume $\mathbf{r}_n := D\phi(0) \neq 0$. Then there exist vectors $\mathbf{r}_1, \dots, \mathbf{r}_{n-1}$ so that $\{\mathbf{r}_k\}_{k=1}^n$ is an orthogonal basis of $\mathbb{R}^n$. Define $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$\mathbf{f}(x, y) := (\mathbf{r}_1 \cdot (y - x), \dots, \mathbf{r}_{n-1} \cdot (y - x), \phi(y) - \phi(0) - D\phi(0) \cdot x).$$

Therefore

$$D_y\mathbf{f}(0,0) = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{pmatrix},$$

the $\{\mathbf{r}_k\}_{k=1}^n$ regarded as row vectors, and so $\det D_y\mathbf{f}(0,0) \neq 0$. The Implicit Function Theorem (§C.6) implies we can find $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\mathbf{\Phi}(0) = 0$ and

$$\mathbf{f}(x, \mathbf{\Phi}(x)) = 0 \ \text{ for } |x| \text{ small.}$$

In particular,

$$\text{(47)} \qquad \begin{cases} \phi(\mathbf{\Phi}(x)) = \phi(0) + D\phi(0)\cdot x \\ \mathbf{r}_k \cdot (\mathbf{\Phi}(x) - x) = 0 \quad (k = 1, \dots, n-1). \end{cases}$$

Differentiating with respect to x, we deduce as well

$$(D\mathbf{\Phi}(0) - I)\mathbf{r}_k = 0 \quad (k = 1, \dots, n)$$

and so $D\mathbf{\Phi}(0) = I$. This proves assertion (i).

2. Fix $x \in \mathbb{R}^n$. Then $\psi(t) := \phi(tx)$ satisfies

$$\psi(1) = \psi(0) + \psi'(0) + \int_0^1 (1-t)\psi''(t)\,dt.$$

Thus if $D\phi(0) = 0$, we have

$$\text{(48)} \qquad \phi(x) = \phi(0) + \frac{1}{2} x \cdot \mathbf{A}(x) x$$

for the symmetric matrix

$$\mathbf{A}(x) := 2\int_0^1 (1-t) D^2\phi(tx)\,dt.$$

Observe $\mathbf{A}(0) = D^2\phi(0)$. Let us hereafter suppose $D^2\phi(0)$ is nonsingular, and so the same is true for $\mathbf{A}(x)$, provided $|x|$ is small. Furthermore, we may assume upon rotating to new coordinates if necessary that

$$\mathbf{A}(0) = D^2\phi(0) \quad \text{is diagonal.}$$

3. We now claim that there exists for each $m \in \{0, 1, \dots, n\}$ a smooth mapping $\mathbf{\Phi}_m : \mathbb{R}^n \to \mathbb{R}^n$, such that

$$\text{(49)} \qquad \begin{cases} \mathbf{\Phi}_m(0) = 0, \ D\mathbf{\Phi}_m(0) = I, \ \text{and} \\ \phi(\mathbf{\Phi}_m(x)) = \phi(0) + \frac{1}{2}\sum_{i=1}^m \phi_{x_i x_i}(0) x_i^2 + \frac{1}{2}\sum_{i,j=m+1}^n a_m^{ij}(x) x_i x_j, \end{cases}$$

for $|x|$ small, where $\mathbf{A}_m = ((a_m^{ij}))$ is smooth and symmetric.

Observe in particular that (49) implies

$$\text{(50)} \qquad a_m^{ij}(0) = \phi_{x_i x_j}(0) \quad (i, j = m+1, \dots, n),$$

and thus $a_m^{m+1,m+1}(x) \neq 0$ for $|x|$ sufficiently small.

4. Assertion (49) for $m = 0$ is (48) with $\mathbf{A}_0 = \mathbf{A}$ and $\mathbf{\Phi}_0$ the identity mapping. So next assume by induction that (49) holds for some $m \in \{0, \dots, n-1\}$ and write

$$\phi_m(x) := \phi(\mathbf{\Phi}_m(x)).$$

Then

$$\text{(51)} \quad \phi_m(x) = \phi(0) + \frac{1}{2}\sum_{i=1}^{m} \phi_{x_i x_i}(0)x_i^2 + \frac{1}{2}\sum_{i,j=m+1}^{n} a_m^{ij}(x)x_i x_j \quad \text{for } |x| \text{ small.}$$

Define a mapping $\mathbf{\Pi}_{m+1} : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{\Pi}_{m+1}(y) = x$, by writing

$$\mathbf{\Pi}_{m+1}(y) :=$$
$$\left(y_1, \dots, y_m, \left(\frac{a_m^{m+1,m+1}(y)}{\phi_{x_{m+1}x_{m+1}}(0)}\right)^{\frac{1}{2}}\left[y_{m+1} + \sum_{j=m+2}^{n} \frac{y_j a_m^{m+1,j}(y)}{a_m^{m+1,m+1}(y)}\right], \dots, y_n\right)$$

for small $|y|$. It follows then from (51) that

$$\phi_m(y) = \phi(0) + \frac{1}{2}\sum_{i=1}^{m+1} \phi_{x_i x_i}(0)x_i^2 + \frac{1}{2}\sum_{i,j=m+2}^{n} b_{m+1}^{ij}(y)x_i x_j,$$

where

$$b_{m+1}^{ij}(y) := \begin{cases} a_m^{ij}(y) - \frac{a_m^{m+1,i}(y)a_m^{m+1,j}(y)}{a_m^{m+1,m+1}(y)} & i,j = m+2, \dots, n \\ 0 & \text{otherwise.} \end{cases}$$

Since $D^2\phi(0)$ is diagonal, (50) implies

$$\mathbf{\Pi}_{m+1}(0) = 0, \; D\mathbf{\Pi}_{m+1}(0) = I.$$

Consequently we can define for small $|x|$ the inverse mapping $\mathbf{\Xi}_{m+1} := \mathbf{\Pi}_{m+1}^{-1}$, $y = \mathbf{\Xi}_{m+1}(x)$. Therefore

$$\phi_m(\mathbf{\Xi}_{m+1}(x)) = \phi(0) + \frac{1}{2}\sum_{i=1}^{m+1} \phi_{x_i x_i}(0)x_i^2 + \frac{1}{2}\sum_{i,j=m+2}^{n} a_{m+1}^{ij}(x)x_i x_j,$$

for $\mathbf{A}_{m+1} := \mathbf{B}_{m+1} \circ \mathbf{\Xi}_{m+1}$. This is statement (49), with $m+1$ replacing m and with $\mathbf{\Phi}_{m+1} := \mathbf{\Phi}_m \circ \mathbf{\Xi}_{m+1}$.

The case $m = n$ is assertion (ii) of the theorem. □

The stationary phase method. We can at last combine the information gleaned in Lemmas 1–3 to explain informally the stationary phase technique for deriving the asymptotics of

$$I_\varepsilon = \int_{\mathbb{R}^n} e^{\frac{i\phi(y)}{\varepsilon}} a(y)\, dy$$

as $\varepsilon \to 0$. We will assume

$$\text{(52)} \qquad \begin{cases} D\phi \text{ vanishes within the support of } a \\ \text{only at the points } y_1, \dots, y_N, \end{cases}$$

and furthermore

$$\text{(53)} \qquad D^2\phi(y_k) \text{ is nonsingular} \quad (k = 1, \dots, N).$$

Fix $\delta > 0$ so small the balls $\{B(y_k, \delta)\}_{k=1}^N$ are disjoint. Then for $m = 1, \dots,$

$$\left| \int_{\mathbb{R}^n - \cup_{i=1}^N B(y,\delta)} e^{\frac{i\phi(y)}{\varepsilon}} a(y)\, dy \right| = O(\varepsilon^m) \quad \text{as } \varepsilon \to 0.$$

This follows since we can employ Lemma 3,(i) to change variables near any point $y \notin \bigcup_{i=1}^N B(y, \delta)$ to make ϕ affine, with nonvanishing gradient. Thus Lemma 1 and a partition of unity argument gives the stated estimate.

On the other hand if $\delta > 0$ is small enough, we can employ Lemma 3,(ii) to compute

$$\begin{aligned} \int_{B(y_k,\delta)} e^{i\phi(y)} a(y)\, dy &= \int_{\Phi^{-1}(B(y_k,\delta))} e^{i\phi(\Phi(x))} a(\Phi(x)) |\det D\Phi(x)|\, dx \\ &= e^{\frac{i\phi(y_k)}{\varepsilon}} \int_{\Phi^{-1}(B(y_k,\delta))} e^{\frac{i}{2\varepsilon}(x-y_k)\cdot D^2\phi(y_k)(x-y_k)} a(\Phi(x)) \\ &\qquad\qquad |\det D\Phi(x)|\, dx \\ &= e^{\frac{i\phi(y_k)}{\varepsilon}} \frac{(2\pi\varepsilon)^{n/2}}{|\det D^2\phi(y_k)|^{1/2}} e^{\frac{i\pi}{4} \operatorname{sgn}(D^2\phi(y_k))} (a(y_k) + O(\varepsilon)), \end{aligned}$$

according to Lemma 2. We thereby obtain the asymptotic formula

$$\text{(54)} \qquad I_\varepsilon = (2\pi\varepsilon)^{n/2} \sum_{k=1}^N \frac{e^{\frac{i\phi(y_k)}{\varepsilon}}}{|\det D^2\phi(y_k)|^{1/2}} e^{\frac{i\pi}{4} \operatorname{sgn}(D^2\phi(y_k))} (a(y_k) + O(\varepsilon)),$$

as $\varepsilon \to 0$.

Example 4 (Stationary phase for the wave equation, continued). We can now apply the foregoing theory to (38), which states

$$I_+^\varepsilon(x,t) = \frac{1}{(2\pi\varepsilon)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} a(z)e^{\frac{i}{\varepsilon}\phi_+(x,y,z,t)}\,dydz.$$

This is of the form (40), with (x,t) replacing x and (y,z) replacing y.

Define for fixed $x \in \mathbb{R}^n$, $t > 0$, the set where the mapping $(y,z) \mapsto \phi_+(x,y,z,t)$ is stationary:

$$S^+ := \{(y,z) \mid D_{y,z}\phi_+(x,y,z,t) = 0\}.$$

Recall from (39) that $\phi_+(x,y,z,t) = (x-z)\cdot y + t|y| + p(z)$; and so

$$\begin{cases} D_y\phi_+ = (x-z) + t\frac{y}{|y|} & (y \neq 0),\\ D_z\phi_+ = -y + Dp(z). \end{cases}$$

Consequently

$$S^+ = \left\{(y,z) \mid x = z - \frac{tDp(z)}{|Dp(z)|},\ y = Dp(z)\right\}, \tag{55}$$

and we here and henceforth assume that $y = Dp(z) \neq 0$ if $(y,z) \in S^+$.

Now if $y \neq 0$,

$$D^2_{y,z}\phi_+ = \begin{pmatrix} D^2_y\phi_+ & D^2_{yz}\phi_+ \\ D^2_{yz}\phi_+ & D^2_z\phi_+ \end{pmatrix}_{2n\times 2n} = \begin{pmatrix} \frac{t}{|y|}\mathbf{P}(y) & -I \\ -I & D^2p \end{pmatrix},$$

for $\mathbf{P}(y) := I - \frac{y\otimes y}{|y|^2}$. We have $y = Dp(z)$ on the stationary set S^+, and so

$$\det(D^2_{y,z}\phi_+) = (-1)^n \det\left(I - \frac{t}{|Dp|}D^2p\,\mathbf{P}(Dp)\right).$$

Now the symmetric matrix $\mathbf{\Sigma}(z) := \frac{\mathbf{P}(Dp)}{|Dp|}D^2p\mathbf{P}(Dp)$ has n real eigenvalues $\lambda_1(z),\dots,\lambda_n(z)$. Since $\mathbf{\Sigma}(z)Dp = 0$, we may take $\lambda_n(z) = 0$. The other eigenvalues $\lambda_1(z),\dots,\lambda_{n-1}(z)$ turn out to be the principal curvatures $\kappa_1(z),\dots,\kappa_{n-1}(z)$ of the level surface of p passing through z. Since $\mathbf{P}^2 = \mathbf{P}$, the nonzero eigenvalues of $\frac{1}{|Dp|}D^2p\,\mathbf{P}(Dp)$ are the nonzero principal curvatures. Thus on S^+,

$$\det(D^2_{y,z}\phi_+) = (-1)^n\prod_{i=1}^{n-1}(1 - t\kappa_i(z)). \tag{56}$$

We apply the stationary phase estimates for $x_0 \in \mathbb{R}^n$ and small $t_0 > 0$. If t_0 is small enough, we can invoke the Implicit Function Theorem to solve uniquely the expressions

$$x_0 = z - t_0 \frac{Dp(z)}{|Dp(z)|}, \ y = Dp(z)$$

for $y_0 = y(x_0, t_0)$, $z_0 = z(x_0, t_0)$. Thus the asymptotic formula (54) (with $2n$ replacing n) implies

$$\begin{aligned} I_+^\varepsilon(x_0, t_0) &= \frac{1}{(2\pi\varepsilon)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a(z) e^{\frac{i}{\varepsilon}\phi_+(x_0,y,z,t_0)}\, dydz \\ &= \frac{e^{\frac{i\pi}{4}\operatorname{sgn}(D^2_{y,z}\phi_+)}}{|\det D^2_{yz}\phi_+|^{1/2}} e^{i\varepsilon\phi_+}[a(z_0) + O(\varepsilon)] \quad \text{as } \varepsilon \to 0, \end{aligned} \tag{57}$$

ϕ_+ and $D^2_{yz}\phi_+$ evaluated at (x_0, y_0, z_0, t_0). Recall further that (56) gives us an explicit function for $\det(D^2_{yz}\phi_+)$. A similar asymptotic formula holds for $I_-^\varepsilon(x_0, t_0)$. Since $u^\varepsilon(x_0, t_0) = \frac{1}{2}(I_+^\varepsilon(x_0, t_0) + I_-^\varepsilon(x_0, t_0))$, we derive detailed information concerning the limits as $\varepsilon \to 0$, at least for small times $t_0 > 0$.

□

Remark (Optics and stationary phase). It remains to discuss briefly the connections between the formal geometric optics and the stationary phase approaches. Recall that the former brought us to the two Hamilton-Jacobi equations

$$p_t \pm |Dp| = 0 \tag{58}$$

for the phase function of $u^\varepsilon = a_\varepsilon e^{\frac{ip\varepsilon}{\varepsilon}} = (a + o(1))e^{\frac{ip+o(1)}{\varepsilon}}$. Now the characteristic equations for the PDE $p_t - |Dp| = 0$ are

$$\begin{cases} \dot{\mathbf{x}}(s) = -\frac{\mathbf{p}(s)}{|\mathbf{p}(s)|} \\ \dot{\mathbf{p}}(s) = 0, \end{cases} \tag{59}$$

as previously discussed in §3.2.2. In particular given a point $x \in \mathbb{R}^n$, $t > 0$, where t is small, the projected characteristic $\mathbf{x}(\cdot)$ is a straight line, starting at the unique point z satisfying

$$z = x + t\frac{Dp(z)}{|Dp(z)|}.$$

But this relation is precisely what determines the stationary set S^+ above. Likewise, the characteristics of the partial differential equation $p_t + |Dp| = 0$ determine the stationary set S^- for ϕ_-. □

4.5.4. Homogenization.

Homogenization theory studies the effects of high-frequency oscillations in the coefficients upon solutions of PDE. In the simplest setting we are given a partial differential equation with two natural *length scales*, a macroscopic scale of order 1 and a microscopic scale of order ε, the latter measuring the period of the oscillations. For fixed, but small, $\varepsilon > 0$ the solution u^ε of the PDE will in general be complicated, having different behaviors on the two length scales.

Homogenization theory studies the limiting behavior $u^\varepsilon \to u$ as $\varepsilon \to 0$. The idea is that in this limit the high-frequency effects will "average out", and there will be a simpler, effective limiting PDE that u solves. One of the difficulties is even to guess the form of the limiting partial differential equation, and for this formal, multiscale expansions in ε may be useful.

Example 5 (Periodic homogenization of an elliptic equation). This example assumes some familiarity with the theory of divergence-structure, second-order elliptic PDE, as developed later in Chapter 6.

Let U denote an open, bounded subset of $\mathbb{R}^n$, with smooth boundary ∂U, and consider this boundary-value problem for a divergence structure PDE:

$$\begin{cases} -\sum_{i,j=1}^{n} \left(a^{ij}\left(\frac{x}{\varepsilon}\right) u^\varepsilon_{x_i}\right)_{x_j} = f & \text{in } U \\ \qquad\qquad\qquad u^\varepsilon = 0 & \text{in } \partial U. \end{cases} \tag{60}$$

Here $f : U \to \mathbb{R}$ is given, as are the coefficients a^{ij} $(i, j = 1, \ldots, n)$. We will assume the uniform ellipticity condition

$$\sum_{i,j=1}^{n} a^{ij}(y)\xi_i\xi_j \geq \theta|\xi|^2$$

for some constant $\theta > 0$ and all $y, \xi \in \mathbb{R}^n$. We suppose also

$$\text{the mapping } y \mapsto a^{ij}(y) \text{ is } Q\text{-periodic} \quad (y \in \mathbb{R}^n), \tag{61}$$

Q denoting the unit cube in $\mathbb{R}^n$. Thus the coefficients $a^{ij}\left(\frac{x}{\varepsilon}\right)$ in (60) are rapidly oscillating in x for small $\varepsilon > 0$, and we inquire as to the effect this has upon the solution u^ε. (In applications u^ε represents, say, the electric field within a nonisotropic body having small-scale, periodic structure.)

In the following heuristic discussion let us assume

$$u^\varepsilon \to u \quad \text{as } \varepsilon \to 0 \tag{62}$$

in some suitable sense and try to determine an equation which u satisfies. The trick is to suppose u^ε admits the following *two-scale* expansion:

$$u^\varepsilon(x) = u_0(x, x/\varepsilon) + \varepsilon u_1(x, x/\varepsilon) + \varepsilon^2 u_2(x, x/\varepsilon) + \dots, \tag{63}$$

where $u_i : U \times Q \to \mathbb{R}$ $(i = 0, 1, \dots)$, $u_i = u_i(x, y)$. We are thus thinking of the terms u_i as being both functions of the macroscopic variable x and periodic functions of the microscopic variable $y = \frac{x}{\varepsilon}$. The plan is to plug (63) into (60), and to determine thereby u_0, u_1, etc. We are primarily interested in $u = u_0$.

Now if $v(x) = w(x, x/\varepsilon)$ for some function $w = w(x, y)$, then $\frac{\partial}{\partial x_i} v = \left(\frac{\partial}{\partial x_i} + \frac{1}{\varepsilon}\frac{\partial}{\partial y_i}\right) w$, $i = 1, \dots, n$. Thus, writing

$$Lv = -\sum_{i,j=1}^n (a^{ij}(x/\varepsilon) v_{x_i})_{x_j},$$

we have

$$L = \frac{1}{\varepsilon^2} L_1 + \frac{1}{\varepsilon} L_2 + L_3, \tag{64}$$

where

$$\begin{cases} \text{(a)} \ \ L_1 w := -\sum_{i,j=1}^n (a^{ij}(y) w_{y_i})_{y_j}, \\ \text{(b)} \ \ L_2 w := -\sum_{i,j=1}^n (a^{ij}(y) w_{x_i})_{y_j} + (a^{ij}(y) w_{y_i})_{x_j}, \\ \text{(c)} \ \ L_3 w := -\sum_{i,j=1}^n (a^{ij}(y) w_{x_i})_{x_j}. \end{cases} \tag{65}$$

Next plug the expansion (63) into the PDE $Lu^\varepsilon = f$, and utilize the decomposition (64), (65) to find

$$\begin{aligned} \frac{1}{\varepsilon^2} L_1 u_0 + &\frac{1}{\varepsilon}(L_1 u_1 + L_2 u_0) + (L_1 u_2 + L_2 u_1 + L_3 u_0) \\ &+ \{\text{terms involving } \varepsilon, \varepsilon^2, \dots\} = f. \end{aligned}$$

Equating like powers of ε, we deduce

$$\begin{cases} \text{(a)} \ \ L_1 u_0 = 0, \\ \text{(b)} \ \ L_1 u_1 + L_2 u_0 = 0, \\ \text{(c)} \ \ L_1 u_2 + L_2 u_1 + L_3 u_0 = f, \quad \text{etc.} \end{cases} \tag{66}$$

We examine these PDE to deduce information concerning u_0, u_1, u_2. Now in view of (65)(a), (66)(a) for each fixed x, $u_0(x, y)$ solves $L_1 u_0 = 0$

and is Q-periodic. It turns out that the only such solutions are constant in y. Thus in fact

$$u_0 = u(x) \text{ depends only on } x. \tag{67}$$

Next employ (67), (65)(b), (66)(b) to discover

$$L_1 u_1 = \sum_{i,j=1}^{n} a^{ij}(y)_{y_j} u_{x_i}. \tag{68}$$

We can as follows separate variables to represent u_1 more simply. For $i = 1, \dots, n$, let $\chi^i = \chi^i(y)$ solve

$$\begin{cases} L_1 \chi^i = -\sum_{j=1}^{n} a^{ij}(y)_{y_j} & \text{in } Q \\ \chi^i \quad Q\text{-periodic.} \end{cases} \tag{69}$$

As the right hand side of the PDE in (69) has integral zero over Q, this problem has a solution χ^i (unique up to an additive constant). Here we are applying the *Fredholm alternative*: see Chapter 6.

Using (69) we obtain

$$u_1(x,y) = -\sum_{i=1}^{n} \chi^i(y) u_{x_i}(x) + \tilde{u}_1(x), \tag{70}$$

$\tilde{u}_1$ denoting an arbitrary function of x alone.

Finally let us recall (66)(c):

$$L_1 u_2 = f - L_2 u_1 - L_3 u_0. \tag{71}$$

In view of (65)(c) this PDE will have a Q-periodic solution (in the variable y) only if the integral of the right hand side over Q is zero. Thus we require

$$\int_Q L_2 u_1 + L_3 u_0 \, dy = \int_Q f \, dy = f(x). \tag{72}$$

Owing to (65)(b) and (70),

$$\begin{aligned} \int_Q L_2 u_1 \, dy &= -\sum_{j,k=1}^{n} \left(\int_Q a^{jk}(y) u_{1,y_k} dy \right)_{x_j} \\ &= \sum_{i,j,k=1}^{n} \left(\int_Q a^{jk}(y) \chi^i_{y_k}(y) \, dy \right) u_{x_i x_j}. \end{aligned}$$

Since $u = u_0$, this calculation and (72) imply

$$-\sum_{i,j=1}^{n}\left(\int_Q a^{ij}(y) - \sum_{k=1}^{n} a^{jk}(y)\chi^i_{y_k}(y)\,dy\right) u_{x_i x_j}(x) = f(x).$$

That is,

$$(73)\qquad \begin{cases} -\sum_{i,j=1}^{n} \bar{a}^{ij} u_{x_i x_j} = f & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases}$$

where

$$(74)\qquad \bar{a}^{ij} := \int_Q a^{ij}(y) - \sum_{k=1}^{n} a^{jk}(y)\chi^i_{y_k}(y)\,dy \quad (i,j = 1,\dots,n)$$

are the *homogenized coefficients*, and χ^i solves the *corrector problem* (69) $(i = 1,\dots,n)$. Thus we expect $u^\varepsilon \to u$ as $\varepsilon \to 0$, and u to solve the limit problem (73). □

This example clearly illustrates the power of the multiscale expansion method. It is not at all readily apparent that the high-frequency oscillations in the coefficients of (60) lead to a constant coefficient PDE of the precise form (73), (74).

4.6. POWER SERIES

We discuss in this final section solving boundary-value problems for partial differential equations by looking for solutions expressed as power series.

4.6.1. Noncharacteristic surfaces.

We begin with some fairly general comments concerning the solvability of the k^{th}-order quasilinear PDE

$$(1)\qquad \sum_{|\alpha|=k} a_\alpha(D^{k-1}u,\dots,u,x)D^\alpha u + a_0(D^{k-1}u,\dots,u,x) = 0$$

in some open region $U \subset \mathbb{R}^n$. Let us assume that Γ is a smooth, $(n-1)$-dimensional hypersurface in U, the unit normal to which at any point $x^0 \in \Gamma$ is $\nu(x^0) = \nu = (\nu_1,\dots,\nu_n)$.

Notation. *The* j^{th} normal derivative *of* u *at* $x^0 \in \Gamma$ *is*

$$\frac{\partial^j u}{\partial \nu^j} := \sum_{|\alpha|=j} D^\alpha u\, \nu^\alpha = \sum_{\alpha_1+\cdots+\alpha_n=j} \frac{\partial^j u}{\partial^{\alpha_1}_{x_1}\dots\partial^{\alpha_n}_{x_n}} \nu_1^{\alpha_1}\dots\nu_n^{\alpha_n}.$$ □

Now let $g_0, \dots, g_{k-1} : \Gamma \to \mathbb{R}$ be k given functions. The *Cauchy problem* is then to find a function u solving the PDE (1), subject to the boundary conditions

$$u = g_0,\ \frac{\partial u}{\partial \nu} = g_1, \dots,\ \frac{\partial^{k-1} u}{\partial \nu^{k-1}} = g_{k-1} \quad \text{on } \Gamma. \tag{2}$$

We say that the equations (2) prescribe the *Cauchy data* $g_0, \dots, g_{k-1}$ on Γ.

We now pose a basic question:

$$\left\{ \begin{array}{l} \text{Assuming } u \text{ is a smooth solution of the PDE (1),} \\ \text{do conditions (2) allow us to compute } \textit{all} \text{ the partial} \\ \text{derivatives of } u \text{ along } \Gamma? \end{array} \right. \tag{3}$$

This must certainly be so, if we are ever going to be able to calculate the terms of a power series representation formula for u.

a. Flat boundaries.

We examine first the special circumstance that $U = \mathbb{R}^n$ and Γ is the plane $\{x_n = 0\}$. In this situation we can take $\nu = e_n$, and so the Cauchy conditions (2) read

$$u = g_0,\ \frac{\partial u}{\partial x_n} = g_1, \dots,\ \frac{\partial^{k-1} u}{\partial x_n^{k-1}} = g_{k-1} \quad \text{on } \{x_n = 0\}. \tag{4}$$

Which further partial derivatives of u can we compute along the plane $\Gamma = \{x_n = 0\}$? First, notice that since $u = g_0$ on all of Γ we can differentiate tangentially, that is, with respect to x_i $(i = 1, \dots, n-1)$, to find

$$\frac{\partial u}{\partial x_i} = \frac{\partial g_0}{\partial x_i} \quad (i = 1, \dots, n-1).$$

Since we also know from (4) that

$$\frac{\partial u}{\partial x_n} = g_1,$$

we can determine the full gradient Du along $\Gamma = \{x_n = 0\}$. Similarly, we have

$$\left\{ \begin{array}{l} \frac{\partial^2 u}{\partial x_n \partial x_i} = \frac{\partial g_1}{\partial x_i} \quad (i = 1, \dots, n-1) \\ \frac{\partial^2 u}{\partial x_n^2} = g_2, \end{array} \right.$$

and hence we can compute D^2u on Γ. Next, we see

$$\frac{\partial^3 u}{\partial x_i \partial x_j \partial x_m} = \begin{cases} \frac{\partial^3 g_0}{\partial x_i \partial x_j \partial x_m} & \text{if} \quad i,j,m = 1,\dots,n-1 \\ \frac{\partial^2 g_1}{\partial x_i \partial x_j} & \text{if} \quad i,j = 1,\dots,n-1; \; m = n \\ \frac{\partial g_2}{\partial x_i} & \text{if} \quad i = 1,\dots,n-1; \; j = m = n \\ g_3 & \text{if} \quad i = j = m = n \end{cases}$$

along Γ, and so we can compute D^3u there. Continuing, it is straightforward to check that employing the Cauchy conditions (4) we can compute $u, Du, \dots, D^{k-1}u$ on Γ.

Difficulties will arise, however, when we try to calculate $D^k u$. In this circumstance it is not hard to verify that we can determine each partial derivative of u of order k along $\Gamma = \{x_n = 0\}$ from the Cauchy data (4), except for the k^{th}-order normal derivative

$$\frac{\partial^k u}{\partial x_n^k}.$$

Here, at last, we turn to PDE (1) for help. We observe from (1) that *if the coefficient* $a_{(0,\dots,0,k)}$ *is nonzero*, we can then solve for

$$\frac{\partial^k u}{\partial x_n^k} = -\frac{1}{a_{(0,\dots,0,k)}} \Bigg[\sum_{\substack{|\alpha|=k \\ \alpha \neq (0,\dots,0,k)}} a_\alpha D^\alpha u + a_0 \Bigg], \tag{5}$$

with the coefficients $a_\alpha (|\alpha| = k)$ and a_0 evaluated at $(D^{k-1}u, \dots, u, x)$ along Γ. Now in view of the remarks above, everything on the right hand side of equality (5) can be calculated in terms of the Cauchy data along the plane Γ, and thus we have a formula for the missing k^{th} partial derivative. Consequently we can in fact compute all of $D^k u$ on Γ, provided

$$a_{(0,\dots,0,k)} \neq 0. \tag{6}$$

We say that the plane $\Gamma = \{x_n = 0\}$ is *noncharacteristic* for the PDE (1), if the function $a_{(0,\dots,k)}$ is nonzero for all values of its arguments.

Can we calculate still higher partial derivatives? Assuming the noncharacteristic condition (6), we observe that we can now augment our list (4) of Cauchy data with the new equality

$$\frac{\partial^k u}{\partial x_n^k} = g_k \quad \text{on } \Gamma = \{x_n = 0\}, \tag{7}$$

g_k denoting the right hand side of (5). But then we can, as before, compute all of $D^{k+1}u$ along Γ, except for the term

$$\frac{\partial^{k+1} u}{\partial x_n^{k+1}}.$$

Again we employ the PDE (1). We differentiate (1) with respect to x_n, evaluate the resulting expression on the plane Γ, and rearrange to find

$$\frac{\partial^{k+1} u}{\partial x_n^{k+1}} = \frac{1}{a_{(0,\dots,0,k)}}\{\cdots\},$$

the dots denoting the sum of various expressions, each of which can be computed along Γ in terms of $g_0, \dots, g_k$. Consequently we can ascertain all of $D^{k+1}u$ on Γ, and an induction verifies that in fact we can compute all the partial derivatives of u on the plane Γ.

b. General surfaces.

We now propose to generalize the results obtained above to the general case that Γ is a smooth hypersurface with normal vector field $\boldsymbol{\nu}$.

DEFINITION. *We say the surface* Γ *is* noncharacteristic *for the partial differential equation* (1) *provided*

$$\sum_{|\alpha|=k} a_\alpha \boldsymbol{\nu}^\alpha \neq 0 \quad on\ \Gamma, \tag{8}$$

for all values of the arguments of the coefficients a_α $(|\alpha| = k)$.

THEOREM 1 (Cauchy data and noncharacteristic surfaces). *Assume that* Γ *is noncharacteristic for the* PDE (1). *Then if* u *is a smooth solution of* (1) *and* u *satisfies the Cauchy conditions* (2), *we can uniquely compute all the partial derivatives of* u *along* Γ *in terms of* Γ, *the functions* $g_0, \dots, g_{k-1}$, *and the coefficients* a_α $(|\alpha| = k), a_0$.

Proof. We will reduce to the special case considered above.

For this, let us choose any point $x^0 \in \Gamma$ and recall §C.1 to find smooth maps $\mathbf{\Phi}, \mathbf{\Psi} : \mathbb{R}^n \to \mathbb{R}^n$, so that $\mathbf{\Psi} = \mathbf{\Phi}^{-1}$ and

$$\mathbf{\Phi}(\Gamma \cap B(x^0, r)) \subset \{y_n = 0\}$$

for some $r > 0$. Define

$$v(y) := u(\mathbf{\Psi}(y));$$

so that

$$u(x) = v(\Phi(x)). \tag{9}$$

It is relatively easy to check now v satisfies a quasilinear partial differential equation having the form

$$\sum_{|\alpha|=k} b_\alpha D^\alpha v + b_0 = 0. \tag{10}$$

2. We claim

$$b_{(0,\dots,0,k)} \neq 0 \quad \text{on } \{y_n = 0\}. \tag{11}$$

Indeed from (9) we see that for any multiindex α with $|\alpha| = k$, we have

$$D^\alpha u = \frac{\partial^k v}{\partial y_n^k}(D\Phi^n)^\alpha + \left\{ \text{terms not involving } \frac{\partial^k v}{\partial y_n^k} \right\}.$$

Thus from (1) it follows that

$$\begin{aligned} 0 &= \sum_{|\alpha|=k} a_\alpha D^\alpha u + a_0 \\ &= \sum_{|\alpha|=k} a_\alpha (D\Phi^n)^\alpha \frac{\partial^k v}{\partial y_n^k} + \left\{ \text{terms not involving } \frac{\partial^k v}{\partial y_n^k} \right\}; \end{aligned}$$

and so

$$b_{(0,\dots,0,k)} = \sum_{|\alpha|=k} a_\alpha (D\Phi^n)^\alpha.$$

But $D\Phi^n$ is parallel to $\boldsymbol{\nu}$ on Γ. Consequently $b_{(0,\dots,k)}$ is a nonzero multiple of the term

$$\sum_{|\alpha|=k} a_\alpha \nu^\alpha \neq 0.$$

This verifies the claim (11).

3. Let us now define the functions $h_0, h_1, \dots, h_{k-1} : \mathbb{R}^{n-1} \to \mathbb{R}$ by

$$v = h_0, \ \frac{\partial v}{\partial y_n} = h_1, \dots, \ \frac{\partial^{k-1} v}{\partial y_n^k} = h_{k-1} \quad \text{on } \{y_n = 0\}. \tag{12}$$

Thus we can compute $h_0, \dots, h_{k-1}$ near $y = 0$ in terms of $\mathbf{\Phi}$ and the functions $g_0, \dots, g_{k-1}$. But then, using (11) and the special case discussed above, we see that we can calculate all of the partial derivatives of v on $\{y_n = 0\}$ near $y = 0$.

And finally, upon recalling (9), we at last observe we can compute all the partial derivatives of u on Γ near x^0. □

Remark. It is sometimes convenient to recast the noncharacteristic condition (8) into a somewhat different form, by representing Γ as the zero set of a function $w : \mathbb{R}^n \to \mathbb{R}$. So assume that we are given a function w with

$$\Gamma = \{w = 0\}$$

and $Dw \neq 0$ on Γ. Then $\boldsymbol{\nu} = \pm\frac{Dw}{|Dw|}$ on Γ , and so the noncharacteristic condition (8) becomes

$$\sum_{|\alpha|=k} a_\alpha (Dw)^\alpha \neq 0 \quad \text{on } \Gamma. \tag{13}$$

□

4.6.2. Real analytic functions.

We review in this section the representation of real-valued functions by power series.

DEFINITION. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called (*real*)* analytic *near x_0 if there exist $r > 0$ and constants $\{f_\alpha\}$ such that*

$$f(x) = \sum_\alpha f_\alpha (x - x_0)^\alpha \quad (|x - x_0| < r),$$

the sum taken over all multiindices α.

Remarks. (i) Remember that we write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, for the multiindex $\alpha = (\alpha_1, \ldots, \alpha_n)$.

(ii) If f is analytic near x_0, then f is C^∞ near x_0. Furthermore the constants f_α are computed as $f_\alpha = \frac{D^\alpha f(x_0)}{\alpha!}$, where $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$. Thus f equals its Taylor expansion about x_0:

$$f(x) = \sum_\alpha \frac{1}{\alpha!} D^\alpha f(x_0)(x - x_0)^\alpha \quad (|x - x_0| < r).$$

To simplify, we hereafter take $x_0 = 0$. □

Example. If $r > 0$, set

$$f(x) := \frac{r}{r - (x_1 + \cdots + x_n)} \qquad \text{for } |x| < r/\sqrt{n}.$$

Then

$$\begin{aligned} f(x) &= \frac{1}{1 - \left(\frac{x_1 + \cdots + x_n}{r}\right)} = \sum_{k=0}^\infty \left(\frac{x_1 + \cdots + x_n}{r}\right)^k \\ &= \sum_{k=0}^\infty \frac{1}{r^k} \sum_{|\alpha|=k} \binom{|\alpha|}{\alpha} x^\alpha = \sum_\alpha \frac{|\alpha|!}{r^{|\alpha|}\alpha!} x^\alpha. \end{aligned}$$

We employed the Multinomial Theorem for the third equality above and recalled that $\binom{|\alpha|}{\alpha} = \frac{|\alpha|!}{\alpha!}$. This power series is absolutely convergent for $|x| < r/\sqrt{n}$. Indeed,

$$\sum_\alpha \frac{|\alpha|!}{r^{|\alpha|}\alpha!}|x^\alpha| = \sum_{k=0}^\infty \left(\frac{|x_1| + \cdots + |x_n|}{r}\right)^k < \infty,$$

since $|x_1| + \cdots + |x_n| \leq |x|\sqrt{n} < r$. □

We will see momentarily that the simple power series illustrated in this example is rather important, since we can use it to majorize, and so confirm the convergence of, other power series.

DEFINITION. *Let*

$$f = \sum_\alpha f_\alpha x^\alpha, \; g = \sum_\alpha g_\alpha x^\alpha$$

be two power series. We say g majorizes f, *written*

$$g \gg f,$$

provided

$$g_\alpha \geq |f_\alpha| \quad \text{for all multiindices } \alpha.$$

LEMMA (Majorants).

(i) *If* $g \gg f$ *and* g *converges for* $|x| < r$, *then* f *also converges for* $|x| < r$.

(ii) *If* $f = \sum_\alpha f_\alpha x^\alpha$ *converges for* $|x| < r$ *and* $0 < s\sqrt{n} < r$, *then* f *has a majorant for* $|x| < s\sqrt{n}$.

Proof. 1. To verify assertion (i), we check

$$\sum_\alpha |f_\alpha x^\alpha| \leq \sum_\alpha g_\alpha |x_1|^{\alpha_1} \cdots |x_n|^{\alpha_n} < \infty \quad \text{if } |x| < r.$$

2. Let $0 < s\sqrt{n} < r$ and set $y := s(1, \ldots, 1)$. Then $|y| = s\sqrt{n} < r$ and so $\sum_\alpha f_\alpha y^\alpha$ converges. Thus there exists a constant C such that

$$|f_\alpha y^\alpha| \leq C \quad \text{for each multiindex } \alpha.$$

In particular,

$$|f_\alpha| \leq \frac{C}{y_1^{\alpha_1} \cdots y_n^{\alpha_n}} = \frac{C}{s^{|\alpha|}} \leq C\frac{|\alpha|!}{s^{|\alpha|}\alpha!}.$$

But then

$$g(x) := \frac{Cs}{s - (x_1 + \cdots + x_n)} = C\sum_\alpha \frac{|\alpha|!}{s^{|\alpha|}\alpha!}x^\alpha$$

majorizes f for $|x| < s\sqrt{n}$. □

Remark. We will later need to extend our notation to vector-valued series. So given power series $\{f^k\}_{k=1}^m, \{g^k\}_{k=1}^m$, we set $\mathbf{f} = (f^1, \dots, f^m)$, $\mathbf{g} = (g^1, \dots, g^m)$, and write

$$\mathbf{g} \gg \mathbf{f}$$

to mean

$$g^k \gg f^k \quad (k = 1, \dots, m).$$

□

4.6.3. Cauchy-Kovalevskaya Theorem.

We turn now to our primary task of building a power series solution for the k^{th}-order quasilinear partial differential equation (1), with analytic Cauchy data (2) specified on an analytic, noncharacteristic hypersurface Γ.

a. Reduction to a first-order system.

We intend to construct a solution u as a power series, but must first transform the boundary-value problem (1), (2) into a more convenient form.

First of all, upon flattening out the boundary by an analytic mapping (as in §4.6.1), we can reduce to the situation that $\Gamma \subset \{x_n = 0\}$. Additionally, by subtracting off appropriate analytic functions, we may assume the Cauchy data are identically zero. Consequently we may assume without loss that our problem reads:

$$\text{(14)} \quad \begin{cases} \sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \dots, u, x) D^\alpha u \\ \qquad\qquad + a_0(D^{k-1}u, \dots, u, x) = 0 \quad \text{for } |x| < r \\ \qquad\qquad u = \frac{\partial u}{\partial x_n} = \cdots = \frac{\partial^{k-1} u}{\partial x_n^{k-1}} = 0 \quad \text{for } |x'| < r, x_n = 0, \end{cases}$$

$r > 0$ to be found. As usual we write $x' = (x_1, \dots, x_{n-1})$.

Finally we transform to a first-order *system*. To do so we introduce the function

$$\mathbf{u} := \left(u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \dots, \frac{\partial^{k-1} u}{\partial x_n^{k-1}}\right),$$

the components of which are all the partial derivatives of u of order less than k. Let m hereafter denote the number of components of $\mathbf{u}$ by m, so that $\mathbf{u} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \dots, u^m)$. Observe from the boundary condition in (14) that $\mathbf{u} = 0$ for $|x'| < r, x_n = 0$.

Now for $k \in \{1, \dots, m-1\}$, we can compute $u^k_{x_n}$ in terms of $\{\mathbf{u}_{x_j}\}_{j=1}^{n-1}$. Furthermore in view of the noncharacteristic condition $a_{(0,\dots,0,k)} \neq 0$ near

0, we can utilize the PDE in (14) also to solve for $u^m_{x_n}$ in terms of $\mathbf{u}$ and $\{\mathbf{u}_{x_j}\}_{j=1}^{n-1}$.

Employing these relations, we can consequently transform (14) into a boundary-value problem for a first-order system for $\mathbf{u}$, the coefficients of which are analytic functions. This system is of the general form:

$$(15)\qquad \begin{cases} \mathbf{u}_{x_n} = \sum_{j=1}^{n-1} \mathbf{B}_j(\mathbf{u}, x')\mathbf{u}_{x_j} + \mathbf{c}(\mathbf{u}, x') & \text{for } |x| < r \\ \mathbf{u} = 0 & \text{for } |x'| < r,\ x_n = 0, \end{cases}$$

where we are given the analytic functions $\mathbf{B}_j : \mathbb{R}^m \times \mathbb{R}^{n-1} \to \mathbb{M}^{m\times m}$ $(j = 1, \dots, n-1)$ and $\mathbf{c} : \mathbb{R}^m \times \mathbb{R}^{n-1} \to \mathbb{R}^m$. We will write $\mathbf{B}_j = ((b_j^{kl}))$ and $\mathbf{c} = (c^1, \dots, c^m)$. Carefully note that we have assumed $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$ do not depend on x_n. We can always reduce to this situation by introducing if necessary a new component u^{m+1} of the unknown $\mathbf{u}$, with $u^{m+1} \equiv x_n$.

In particular, the components of the system of partial differential equations in (15) read

$$(16)\qquad u^k_{x_n} = \sum_{j=1}^{n-1}\sum_{l=1}^{m} b_j^{kl}(\mathbf{u}, x')u^l_{x_j} + c^k(\mathbf{u}, x') \quad (k = 1, \dots, m).$$

b. Power series for solutions.

Having reduced to the special form (15), we can now expand $\mathbf{u}$ into a power series, and, more importantly, verify that this series converges near 0.

THEOREM 2 (Cauchy-Kovalevskaya Theorem). *Assume* $\{\mathbf{B}_j\}_{j=1}^{n-1}$ *and* $\mathbf{c}$ *are real analytic functions. Then there exist* $r > 0$ *and a real analytic function*

$$(17)\qquad \mathbf{u} = \sum_\alpha \mathbf{u}_\alpha x^\alpha$$

solving the boundary-value problem (15).

Proof. 1. We must compute the coefficients

$$(18)\qquad \mathbf{u}_\alpha = \frac{D^\alpha \mathbf{u}(0)}{\alpha!}$$

in terms of $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$, and then show that the power series (18) so obtained in fact converges if $|x| < r$ and $r > 0$ is small enough.

2. As the functions $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$ are analytic, we can write

$$\mathbf{B}_j(z,x') = \sum_{\gamma,\delta} \mathbf{B}_{j,\gamma,\delta} z^\gamma x^\delta \quad (j=1,\ldots,n-1) \tag{19}$$

and

$$\mathbf{c}(z,x') = \sum_{\gamma,\delta} \mathbf{c}_{\gamma,\delta} z^\gamma x^\delta, \tag{20}$$

these power series convergent if $|z|+|x'|<s$ for some small $s>0$. Thus

$$\mathbf{B}_{j,\gamma,\delta} = \frac{D_z^\gamma D_x^\delta \mathbf{B}_j(0,0)}{(\gamma+\delta)!}, \quad \mathbf{c}_{\gamma,\delta} = \frac{D_z^\gamma D_x^\delta \mathbf{c}(0,0)}{(\gamma+\delta)!} \tag{21}$$

for $j=1,\ldots,n-1$ and all multiindices γ,δ.

3. Since $\mathbf{u} \equiv 0$ on $\{x_n = 0\}$, we have

$$\mathbf{u}_\alpha = \frac{D^\alpha \mathbf{u}(0)}{\alpha!} = 0 \quad \text{for all multiindices } \alpha \text{ with } \alpha_n = 0. \tag{22}$$

Now fix $i \in \{1,\ldots,n-1\}$ and differentiate (16) with respect to x_i:

$$u^k_{x_n x_i} = \sum_{j=1}^{n-1}\sum_{l=1}^{m}\left(b_j^{kl} u^l_{x_i x_j} + b^{kl}_{j,x_i} u^l_{x_j} + \sum_{p=1}^m b^{kl}_{j,z_p} u^p_{x_i} u^l_{x_j}\right) + c^k_{x_i} + \sum_{p=1}^m c^k_{z_p} u^p_{x_i}.$$

In view of (22), we conclude $u^k_{x_n x_i}(0) = c^k_{x_i}(0,0)$.

If α is a multiindex having the form $\alpha = (\alpha_1,\ldots,\alpha_{n-1},1) = (\alpha',1)$, we likewise prove by induction that

$$D^\alpha u^k(0) = D^{\alpha'} c^k(0,0).$$

Next suppose $\alpha = (\alpha',2)$. Then

$$\begin{aligned} D^\alpha u^k &= D^{\alpha'}(u^k_{x_n})_{x_n} \\ &= D^{\alpha'}\left(\sum_{j=1}^{n-1}\sum_{l=1}^m b_j^{kl} u^l_{x_j} + c^k \right)_{x_n} \quad \text{by (16)} \\ &= D^{\alpha'}\left(\sum_{j=1}^{n-1}\sum_{l=1}^m (b_j^{kl} u^l_{x_j x_n} + \sum_{p=1}^m b^{kl}_{j,z_p} u^p_{x_n} u^l_{x_j}) + \sum_{p=1}^m c^k_{z_p} u^p_{x_n} \right). \end{aligned}$$

Thus

$$D^\alpha u^k(0) = D^{\alpha'}\left(\sum_{j=1}^{n-1}\sum_{l=1}^m b_j^{kl} u^l_{x_j x_n} + \sum_{p=1}^m c^k_{z_p} u^p_{x_n} \right)\Bigg|_{x=\mathbf{u}=0}.$$

The expression on the right hand side can be worked out to be a polynomial *with nonnegative coefficients* involving various derivatives of $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$, and the derivatives $D^\beta \mathbf{u}$, where $\beta_n \leq 1$.

More generally, for each multiindex α and each $k \in \{1, \dots, m\}$, we compute

$$D^\alpha u^k(0) = p_\alpha^k(\dots, D_z^\gamma D_x^\delta \mathbf{B}_j, \dots, D_z^\gamma D_x^\delta \mathbf{c}, \dots, D^\beta \mathbf{u}, \dots)\big|_{x=\mathbf{u}=0},$$

where p_α^k denotes some polynomial with nonnegative coefficients.

Recalling (18)–(21), we deduce for each α, k that

$$u_\alpha^k = q_\alpha^k(\dots, \mathbf{B}_{j,\gamma,\delta}, \dots, \mathbf{c}_{\gamma,\delta}, \dots, \mathbf{u}_\beta, \dots), \tag{23}$$

where

$$q_\alpha^k \text{ is a polynomial } \textit{with nonnegative coefficients} \tag{24}$$

and

$$\beta_n \leq \alpha_n - 1 \text{ for each multiindex } \beta \text{ on the right hand side of (23).} \tag{25}$$

4. Thus far we have merely demonstrated that if there is a smooth solution of (15), then we can compute all of its derivatives at 0 in terms of known quantities. This of course we already know from the discussion in §4.6.1, since the plane $\{x_n = 0\}$ is noncharacteristic.

We now intend to employ (22)–(25) and the *method of majorants* to show the power series (17) actually converges if $|x| < r$ and r is small. For this, let us first suppose

$$\mathbf{B}_j^* \gg \mathbf{B}_j \quad (j = 1, \dots, n-1) \tag{26}$$

and

$$\mathbf{c}^* \gg \mathbf{c}, \tag{27}$$

where

$$\mathbf{B}_j^* := \sum_{\gamma,\delta} \mathbf{B}_{j,\gamma,\delta}^* z^\gamma x^\delta \quad (j = 1, \dots, n-1)$$

and

$$\mathbf{c}^* := \sum_{\gamma,\delta} \mathbf{c}_{\gamma,\delta}^* z^\gamma x^\delta,$$

these power series convergent for $|z| + |x'| < s$. Then

$$0 \leq |\mathbf{B}_{j,\gamma,\delta}| \leq \mathbf{B}^*_{j,\gamma,\delta}, \; 0 \leq |\mathbf{c}_{\gamma,\delta}| \leq \mathbf{c}^*_{\gamma,\delta} \tag{28}$$

for all j, γ, δ.

We consider next the new boundary-value problem

$$\begin{cases} \mathbf{u}^*_{x_n} = \sum_{j=1}^{n-1} \mathbf{B}^*_j(\mathbf{u}^*, x')\mathbf{u}^*_{x_j} + \mathbf{c}^*(\mathbf{u}^*, x') & \text{for } |x| < r \\ \mathbf{u}^* = 0 & \text{for } |x'| < r, \; x_n = 0, \end{cases} \tag{29}$$

and, as above, look for a solution having the form

$$\mathbf{u}^* = \sum_\alpha \mathbf{u}^*_\alpha x^\alpha, \tag{30}$$

where

$$\mathbf{u}^*_\alpha = \frac{D^\alpha \mathbf{u}^*(0)}{\alpha!}. \tag{31}$$

5. We claim

$$0 \leq |u^k_\alpha| \leq u^{k*}_\alpha \quad \text{for each multiindex } \alpha.$$

The proof is by induction. The general step follows since

$$\begin{aligned} |u^k_\alpha| &= |q^k_\alpha(\dots, \mathbf{B}_{j,\gamma,\delta}, \dots, \mathbf{c}_{\gamma,\delta}, \dots, \mathbf{u}_\beta, \dots)| \quad \text{by (23)} \\ &\leq q^k_\alpha(\dots, |\mathbf{B}_{j,\gamma,\delta}|, \dots, |\mathbf{c}_{\gamma,\delta}|, \dots, |\mathbf{u}_\beta|, \dots) \quad \text{by (24)} \\ &\leq q^k_\alpha(\dots, \mathbf{B}^*_{j,\gamma,\delta}, \dots, \mathbf{c}^*_{\gamma,\delta}, \dots, \mathbf{u}^*_\beta, \dots) \quad \text{by (24), (28) and induction} \\ &= u^{k*}_\alpha. \end{aligned}$$

Thus

$$\mathbf{u}^* \gg \mathbf{u}, \tag{32}$$

and so it suffices to prove that the power series (30) converges near zero.

6. As demonstrated in the proof of assertion (ii) of the lemma in §4.6.2, if we choose

$$\mathbf{B}^*_j := \frac{Cr}{r - (x_1 + \cdots + x_{n-1}) - (z_1 + \cdots + z_m)} \begin{pmatrix} 1 & \dots & 1 \\ & \ddots & \\ 1 & \dots & 1 \end{pmatrix}$$

for $j = 1, \ldots, n-1$, and

$$\mathbf{c}^* := \frac{Cr}{r - (x_1 + \cdots + x_{n-1}) - (z_1 + \cdots + z_m)}(1, \ldots, 1),$$

then (26), (27) will hold if C is large enough, $r > 0$ is small enough, and $|x'| + |z| < r$.

Hence the problem (29) reads

$$\begin{cases} \mathbf{u}^*_{x_n} = \dfrac{Cr}{r - (x_1 + \cdots + x_{n-1}) - (u^{1*} + \cdots + u^{m*})} \left(\sum_{j,l} \mathbf{u}^{l*}_{x_j} + 1\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } |x| < r \\ \qquad\qquad\qquad\qquad\qquad \mathbf{u}^* = 0 \quad \text{for } |x'| < r,\ x_n = 0. \end{cases}$$

However, this problem has an explicit solution, namely

$$\mathbf{u}^* = v^*(1, \ldots, 1), \tag{33}$$

for

$$\begin{aligned} v^*(x) :=& \frac{1}{mn}(r - (x_1 + \cdots + x_{n-1}) \\ & - [(r - (x_1 + \cdots + x_{n-1}))^2 - 2mnCrx_n]^{1/2}). \end{aligned} \tag{34}$$

This expression is analytic for $|x| < r$, provided $r > 0$ is sufficiently small. Thus $\mathbf{u}^*$ defined by (33) necessarily has the form (30), (31), the power series (30) converging for $|x| < r$. As $\mathbf{u}^* \gg \mathbf{u}$, the power series (17) converges as well for $|x| < r$.

This defines the analytic function $\mathbf{u}$ near 0. Since the Taylor expansions of the analytic functions $\mathbf{u}_{x_n}$ and $\sum_{j=1}^{n-1} \mathbf{B}_j(\mathbf{u}, x)\mathbf{u}_{x_j} + \mathbf{c}(\mathbf{u}, x)$ agree at 0, they agree as well throughout the region $|x| < r$. □

Remark. The Cauchy-Kovalevskaya Theorem is valid also for fully nonlinear, analytic PDE: see Folland [**F1**] □

4.7. PROBLEMS

1. Use separation of variables to find a nontrivial solution u of the PDE

$$u_{x_1}^2 u_{x_1x_1} + 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2}^2 u_{x_2x_2} = 0 \quad \text{in } \mathbb{R}^2.$$

2. Consider Laplace's equation $\Delta u = 0$ in $\mathbb{R}^2$, taken with the Cauchy data

$$u = 0,\ \frac{\partial u}{\partial x_2} = \frac{1}{n}\sin(nx_1) \quad \text{on } \{x_2 = 0\}.$$

Employ separation of variables to derive the solution

$$u = \tfrac{1}{n^2}\sin(nx_1)\sinh(nx_2).$$

What happens to u as $n \to \infty$? Is the Cauchy problem for Laplace's equation well-posed? (This example is due to Hadamard.)

3. Consider the viscous conservation law

$$(*) \qquad u_t + F(u)_x - au_{xx} = 0 \quad \text{in } \mathbb{R} \times (0,\infty),$$

where $a > 0$ and F is uniformly convex.

(i) Show u solves $(*)$ if $u(x,t) = v(x-\sigma t)$ and v is defined implicitly by the formula

$$s = \int_c^{v(s)} \frac{a}{F(z) - \sigma z + b}\, dz \quad (s \in \mathbb{R}),$$

where b and c are constants.

(ii) Demonstrate that we can find a traveling wave satisfying

$$\lim_{s\to-\infty} v(s) = u_l, \quad \lim_{s\to\infty} v(s) = u_r$$

for $u_l > u_r$, if and only if

$$\sigma = \frac{F(u_l) - F(u_r)}{u_l - u_r}.$$

(iii) Let u^ε denote the above traveling wave solution of $(*)$ for $a = \varepsilon$, with $u^\varepsilon(0,0) = \frac{u_l+u_r}{2}$. Compute $\lim_{\varepsilon\to 0} u^\varepsilon$ and explain your answer.

4. Prove that if u is the solution of problem (23) for Schrödinger's equation in §4.3 given by formula (20), then

$$\|u(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} \le \frac{1}{(4\pi|t|)^{n/2}}\|g\|_{L^1(\mathbb{R}^n)}$$

for each $t \neq 0$.

5. Utilize Lemma 2 in §4.5.3 to discuss the sense in which u defined by formula (20) in §4.3.1 converges to the initial data g as $t \to 0^+$.

6. Show that we can construct an explicit solution of the initial-value problem (18), (19) from Example 1 in §4.5.1, having the form

$$v(z,t) = \frac{1}{\sigma(t)}\frac{1}{(\pi\gamma(t))^{1/2}} e^{-z^2/\gamma(t)} \quad (z \in \mathbb{R},\ t > 0),$$

the function $\gamma(t)$ to be found. Substitute into the PDE and determine an ODE γ should satisfy. What is the initial condition for this ODE?

7. Justify in the proof of Lemma 2 in §4.5.3 the transformation of the integral of e^{-z^2} over the line Γ to the integral over the real axis.

8. Let $n = 1$ and suppose that u^ε solves the problem

$$\begin{cases} -(a(\frac{x}{\varepsilon})u_x^\varepsilon)_x = f & \text{in } (0,1) \\ u^\varepsilon(0) = u^\varepsilon(1) = 0, \end{cases}$$

where a is a smooth, positive function, which is 1-periodic. Assume also that $f \in L^2(0,1)$.

(i) Show that $u^\varepsilon \rightharpoonup u$ weakly in $H_0^1(0,1)$, where u solves

$$\begin{cases} -\bar{a}u_{xx} = f & \text{in } (0,1) \\ u(0) = u(1) = 0, \end{cases}$$

for $\bar{a} := (\int_0^1 a(y)^{-1}\,dy)^{-1}$.

(ii) Check that this answer agrees with the conclusions (73), (74) in §4.5.4.

(This problem requires knowledge of energy estimates, Sobolev spaces, etc. from Chapters 5, 6.)

9. Show that the line $\{t = 0\}$ is characteristic for the heat equation $u_t = u_{xx}$. Show there does not exist an analytic solution u of the heat equation in $\mathbb{R} \times \mathbb{R}$, with $u = \frac{1}{1+x^2}$ on $\{t = 0\}$. (Hint: Assume there is an analytic solution, compute its coefficients, and show that the resulting power series diverges except at $(0,0)$. This example is due to Kovalevskaya.)

4.8. REFERENCES

Section 4.1 See for instance Pinsky [**P**], Strauss [**ST**], Thoe–Zachmanoglou [**T-Z**], Weinberger [**WE**], etc. for more on separation of variables.

Section 4.2 C. Jones provided the discussion of traveling waves for the bistable equation, and J.-L. Vazquez showed me the derivation of Barenblatt's solution. P. Olver's book [**O**] explains much more about symmetry methods for PDE.

Section 4.3 Stein–Weiss [**S-W**], Stein [**SE**], Hörmander [**H1**], Rauch [**R**], Treves [**T**], etc. provide much more information concerning Fourier transform techniques. M. Weinstein helped me with Schrödinger's equation. Example 5 is from Pinsky [**P**], and the solution of the wave equation in §4.3.2 is from Pinsky–Taylor [**P-T**].

Section 4.4 See Courant–Hilbert [**C-H**] for more on the hodograph and Legendre transforms.

Section 4.5 J. Neu contributed §4.5.1. Section 4.5.3 is based upon some classroom lectures of J. Ralston, following Hörmander [**H2**]. The discussion of homogenization in §4.5.4 follows Bensoussan–Lions–Papanicolaou [**B-L-P**].

Section 4.6 See Folland [**F1,** Chapter 1], John [**J,** Chapter 3], DiBenedetto [**DB,** Chapter 1].

Section 4.7 Problem 1 is due to Aronsson, and Problem 9 is from Mikhailov [**M**].

Chapter 5

SOBOLEV SPACES

This chapter mostly develops the theory of *Sobolev spaces*, which turn out often to be the proper setting in which to apply ideas of functional analysis to glean information concerning partial differential equations. The following material is often subtle, and will seem largely unmotivated, but ultimately will prove extremely useful.

Since we have in mind eventual applications to rather wide classes of partial differential equations, it is worth sketching out here our overall point of view. Our intention, broadly put, will be later to take various specific PDE and to recast them abstractly as operators acting on appropriate linear spaces. We can symbolically write this as

$$A : X \to Y,$$

where the operator A encodes the structure of the partial differential equations, including possibly boundary conditions, etc., and X, Y are spaces of functions. The great advantage is that once our PDE problem has been suitably interpreted in this form, we can often employ the general and elegant principles of functional analysis (Appendix D) to study the solvability of various equations involving A. We will later see that the really hard work is not so much the invocation of functional analysis, but rather finding the "right" spaces X, Y and the "right" abstract operators A. Sobolev spaces are designed precisely to make all this work out properly, and so these are usually the proper choices for X, Y.

We will utilize Sobolev spaces for studying linear elliptic, parabolic and hyperbolic PDE in Chapters 6–7, and for studying nonlinear elliptic and parabolic equations in Chapters 8–9 .

The reader may wish to look over some of the terminology for functional analysis in Appendix D before going further.

5.1. HÖLDER SPACES

Before turning to Sobolev spaces, we first discuss the simpler *Hölder spaces.*

Assume $U \subset \mathbb{R}^n$ is open and $0 < \gamma \leq 1$. We have previously considered the class of Lipschitz continuous functions $u : U \to \mathbb{R}$, which by definition satisfy the estimate

$$|u(x) - u(y)| \leq C|x - y| \quad (x, y \in U) \tag{1}$$

for some constant C. Now (1) of course implies u is continuous, and more importantly provides a uniform modulus of continuity. It turns out to be useful to consider also functions u satisfying a variant of (1), namely

$$|u(x) - u(y)| \leq C|x - y|^\gamma \quad (x, y \in U) \tag{2}$$

for some constant C. Such a function is said to be *Hölder continuous with exponent* γ.

DEFINITIONS. (i) *If* $u : U \to \mathbb{R}$ *is bounded and continuous, we write*

$$\|u\|_{C(\bar{U})} := \sup_{x \in U} |u(x)|.$$

(ii) *The* γ^{th}-Hölder seminorm *of* $u : U \to \mathbb{R}$ *is*

$$[u]_{C^{0,\gamma}(\bar{U})} := \sup_{\substack{x,y \in U \\ x \neq y}} \left\{ \frac{|u(x) - u(y)|}{|x - y|^\gamma} \right\},$$

and the γ^{th}-Hölder norm *is*

$$\|u\|_{C^{0,\gamma}(\bar{U})} := \|u\|_{C(\bar{U})} + [u]_{C^{0,\gamma}(\bar{U})}.$$

DEFINITION. *The* Hölder space

$$C^{k,\gamma}(\bar{U})$$

consists of all functions $u \in C^k(\bar{U})$ *for which the norm*

$$\|u\|_{C^{k,\gamma}(\bar{U})} := \sum_{|\alpha|\leq k} \|D^\alpha u\|_{C(\bar{U})} + \sum_{|\alpha|=k} [D^\alpha u]_{C^{0,\gamma}(\bar{U})} \tag{3}$$

is finite.

So the space $C^{k,\gamma}(\bar{U})$ consists of those functions u that are k-times continuously differentiable and whose k^{th}-partial derivatives are Hölder continuous with exponent γ. Such functions are well-behaved, and furthermore the space $C^{k,\gamma}(\bar{U})$ itself possesses a good mathematical structure:

THEOREM 1 (Hölder spaces as function spaces). *The space of functions* $C^{k,\gamma}(\bar{U})$ *is a Banach space.*

The proof is left as an exercise (Problem 1), but let us pause here to make clear what is being asserted. Recall from §D.1 that if X denotes a real linear space, then a mapping $\| \quad \| : X \to [0,\infty)$ is called a *norm* provided

(i) $\|u+v\| \leq \|u\| + \|v\|$ for all $u, v \in X$,
(ii) $\|\lambda u\| = |\lambda|\|u\|$ for all $u \in X$, $\lambda \in \mathbb{R}$,
(iii) $\|u\| = 0$ if and only if $u = 0$.

A norm provides us with a notion of convergence: we say a sequence $\{u_k\}_{k=1}^\infty \subset X$ *converges* to $u \in X$, written $u_k \to u$, if $\lim_{k\to\infty} \|u_k - u\| = 0$. A *Banach space* is then a normed linear space which is *complete*, that is, within which each Cauchy sequence converges.

So in Theorem 1 we are stating that if we take on the linear space $C^{k,\gamma}(\bar{U})$ the norm $\|\cdot\| = \|\cdot\|_{C^{k,\gamma}(\bar{U})}$, defined by (3), then $\|\cdot\|$ verifies properties (i)–(iii) above, and in addition each Cauchy sequence converges.

5.2. SOBOLEV SPACES

The Hölder spaces introduced in §5.1 are unfortunately not often suitable settings for elementary PDE theory, as we usually cannot make good enough analytic estimates to demonstrate that the solutions we construct actually belong to such spaces. What are needed rather are some other kinds of spaces, containing less smooth functions. In practice we must strike a balance, by designing spaces comprising functions which have some, but not too great, smoothness properties.

5.2.1. Weak derivatives.

We start off by substantially weakening the notion of partial derivatives.

Notation. Let $C_c^\infty(U)$ denote the space of infinitely differentiable functions $\phi : U \to \mathbb{R}$, with compact support in U. We will call a function ϕ belonging to $C_c^\infty(U)$ a *test function.* □

Motivation for definition of weak derivative. Assume we are given a function $u \in C^1(U)$. Then if $\phi \in C_c^\infty(U)$, we see from the integration by parts formula that

$$\int_U u\phi_{x_i}\,dx = -\int_U u_{x_i}\phi\,dx \quad (i = 1, \dots, n). \tag{1}$$

There are no boundary terms, since ϕ has compact support in U and thus vanishes near ∂U. More generally now, if k is a positive integer, $u \in C^k(U)$, and $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multiindex of order $|\alpha| = \alpha_1 + \cdots + \alpha_n = k$, then

$$\int_U uD^\alpha\phi\,dx = (-1)^{|\alpha|}\int_U D^\alpha u\phi\,dx. \tag{2}$$

This equality holds since

$$D^\alpha\phi = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}\phi$$

and we can apply formula (1) $|\alpha|$ times.

We next examine formula (2), valid for $u \in C^k(U)$, and ask whether some variant of it might still be true even if u is not k times continuously differentiable. Now the left hand side of (2) makes sense if u is only locally summable: the problem is rather that if u is not C^k, then the expression "$D^\alpha u$" on the right hand side of (2) has no obvious meaning. We resolve this difficulty by asking if there exists a locally summable function v for which formula (2) is valid, with v replacing $D^\alpha u$:

DEFINITION. *Suppose $u, v \in L^1_{\text{loc}}(U)$, and α is a multiindex. We say that v is the α^{th}*-weak partial derivative of u, *written*

$$D^\alpha u = v,$$

provided

$$\int_U u D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_U v \phi \, dx \tag{3}$$

for all test functions $\phi \in C_c^\infty(U)$.

In other words, if we are given u and if there happens to exist a function v which verifies (3) for all ϕ, we say that $D^\alpha u = v$ in the weak sense. If there does not exist such a function v, then u does not possess a weak α^{th}-partial derivative.

LEMMA (Uniqueness of weak derivatives). *A weak* α^{th}*-partial derivative of* u*, if it exists, is uniquely defined up to a set of measure zero.*

Proof. Assume that $v, \tilde{v} \in L^1_{\text{loc}}(U)$ satisfy

$$\int_U u D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_U v \phi \, dx = (-1)^{|\alpha|} \int_U \tilde{v} \phi \, dx$$

for all $\phi \in C_c^\infty(U)$. Then

$$\int_U (v - \tilde{v}) \phi \, dx = 0 \tag{4}$$

for all $\phi \in C_c^\infty(U)$; whence $v - \tilde{v} = 0$ a.e. □

Example 1. Let $n = 1$, $U = (0, 2)$, and

$$u(x) = \begin{cases} x & \text{if } 0 < x \leq 1 \\ 1 & \text{if } 1 \leq x < 2. \end{cases}$$

Define

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

Let us show $u' = v$ in the weak sense. To see this, choose any $\phi \in C_c^\infty(U)$. We must demonstrate

$$\int_0^2 u \phi' \, dx = -\int_0^2 v \phi \, dx.$$

But we easily calculate

$$\begin{aligned} \int_0^2 u \phi' \, dx &= \int_0^1 x \phi' \, dx + \int_1^2 \phi' \, dx \\ &= -\int_0^1 \phi \, dx + \phi(1) - \phi(1) = -\int_0^2 v \phi \, dx, \end{aligned}$$

as required. □

Example 2. Let $n = 1$, $U = (0, 2)$, and

$$u(x) = \begin{cases} x & \text{if } 0 < x \leq 1 \\ 2 & \text{if } 1 < x < 2. \end{cases}$$

We assert u' does not exist in the weak sense. To check this, we must show there does not exist any function $v \in L^1_{\mathrm{loc}}(U)$ satisfying

$$\int_0^2 u\phi' \, dx = -\int_0^2 v\phi \, dx \tag{5}$$

for all $\phi \in C_c^\infty(U)$. Suppose, to the contrary, (5) were valid for some v and all ϕ. Then

$$\begin{aligned} -\int_0^2 v\phi \, dx = \int_0^2 u\phi' \, dx &= \int_0^1 x\phi' \, dx + 2\int_1^2 \phi' \, dx \\ &= -\int_0^1 \phi \, dx - \phi(1). \end{aligned} \tag{6}$$

Choose a sequence $\{\phi_m\}_{m=1}^\infty$ of smooth functions satisfying

$$0 \leq \phi_m \leq 1, \quad \phi_m(1) = 1, \quad \phi_m(x) \to 0 \text{ for all } x \neq 1.$$

Replacing ϕ by ϕ_m in (6) and sending $m \to \infty$, we discover

$$1 = \lim_{m\to\infty} \phi_m(1) = \lim_{m\to\infty} \left[\int_0^2 v\phi_m \, dx - \int_0^1 \phi_m \, dx \right] = 0,$$

a contradiction. □

More sophisticated examples appear in the next section.

5.2.2. Definition of Sobolev spaces.

Fix $1 \leq p \leq \infty$ and let k be a nonnegative integer. We define now certain function spaces, whose members have weak derivatives of various orders lying in various L^p spaces.

DEFINITION. *The* Sobolev space

$$W^{k,p}(U)$$

consists of all locally summable functions $u : U \to \mathbb{R}$ such that for each multiindex α with $|\alpha| \leq k$, $D^\alpha u$ exists in the weak sense and belongs to $L^p(U)$.

Remarks. (i) If $p = 2$, we usually write

$$H^k(U) = W^{k,2}(U) \quad (k = 0, 1, \dots).$$

The letter H is used, since—as we will see—$H^k(U)$ is a Hilbert space. Note that $H^0(U) = L^2(U)$.

(ii) We henceforth identify functions in $W^{k,p}(U)$ which agree a.e. □

DEFINITION. *If $u \in W^{k,p}(U)$, we define its* norm *to be*

$$\|u\|_{W^{k,p}(U)} := \begin{cases} \left(\sum_{|\alpha| \le k} \int_U |D^\alpha u|^p \, dx\right)^{1/p} & (1 \le p < \infty) \\ \sum_{|\alpha| \le k} \operatorname{ess\,sup}_U |D^\alpha u| & (p = \infty). \end{cases}$$

DEFINITIONS. (i) *Let $\{u_m\}_{m=1}^\infty$, $u \in W^{k,p}(U)$. We say u_m* converges *to u in $W^{k,p}(U)$, written*

$$u_m \to u \quad \text{in } W^{k,p}(U),$$

provided

$$\lim_{m \to \infty} \|u_m - u\|_{W^{k,p}(U)} = 0.$$

(ii) *We write*

$$u_m \to u \quad \text{in } W^{k,p}_{\text{loc}}(U),$$

to mean

$$u_m \to u \quad \text{in } W^{k,p}(V)$$

for each $V \subset\subset U$.

DEFINITION. *We denote by*

$$W_0^{k,p}(U)$$

the closure of $C_c^\infty(U)$ in $W^{k,p}(U)$.

Thus $u \in W_0^{k,p}(U)$ if and only if there exist functions $u_m \in C_c^\infty(U)$ such that $u_m \to u$ in $W^{k,p}(U)$. We interpret $W_0^{k,p}(U)$ as comprising those functions $u \in W^{k,p}(U)$ such that

$$\text{“}D^\alpha u = 0 \text{ on } \partial U\text{” for all } |\alpha| \le k - 1.$$

This will all be made clearer with the discussion of traces in §5.5.

Notation. It is customary to write

$$H_0^k(U) = W_0^{k,2}(U).$$

□

We will see in the exercises that if $n = 1$ and U is an open interval in $\mathbb{R}^1$, then $u \in W^{1,p}(U)$ if and only if u equals a.e. an absolutely continuous function whose ordinary derivative (which exists a.e.) belongs to $L^p(U)$. Such a simple characterization is however only available for $n = 1$. In general a function can belong to a Sobolev space, and yet be discontinuous and/or unbounded.

Example 3. Take $U = B^0(0,1)$, the open unit ball in $\mathbb{R}^n$, and

$$u(x) = |x|^{-\alpha} \quad (x \in U,\ x \neq 0).$$

For which values of $\alpha > 0, n, p$ does u belong to $W^{1,p}(U)$? To answer, note first u is smooth away from 0, with

$$u_{x_i}(x) = \frac{-\alpha x_i}{|x|^{\alpha+2}} \quad (x \neq 0),$$

and so

$$|Du(x)| = \frac{|\alpha|}{|x|^{\alpha+1}} \quad (x \neq 0).$$

Let $\phi \in C_c^\infty(U)$ and fix $\varepsilon > 0$. Then

$$\int_{U-B(0,\varepsilon)} u\phi_{x_i}\,dx = -\int_{U-B(0,\varepsilon)} u_{x_i}\phi\,dx + \int_{\partial B(0,\varepsilon)} u\phi\nu^i\,dS,$$

$\boldsymbol{\nu} = (\nu^1, \dots, \nu^n)$ denoting the inward pointing normal on $\partial B(0,\varepsilon)$. Now if $\alpha + 1 < n$, $|Du(x)| \in L^1(U)$. In this case

$$\left|\int_{\partial B(0,\varepsilon)} u\phi\nu^i\,dS\right| \leq \|\phi\|_{L^\infty} \int_{\partial B(0,\varepsilon)} \varepsilon^{-\alpha}\,dS \leq C\varepsilon^{n-1-\alpha} \to 0.$$

Thus

$$\int_U u\phi_{x_i}\,dx = -\int_U u_{x_i}\phi\,dx$$

for all $\phi \in C_c^\infty(U)$, provided $0 \leq \alpha < n-1$. Furthermore $|Du(x)| = \frac{\alpha}{|x|^{\alpha+1}} \in L^p(U)$ if and only if $(\alpha+1)p < n$. Consequently $u \in W^{1,p}(U)$ if and only if $\alpha < \frac{n-p}{p}$. In particular $u \notin W^{1,p}(U)$ for each $p \geq n$. □

Example 4. Let $\{r_k\}_{k=1}^{\infty}$ be a countable, dense subset of $U = B^0(0,1)$. Write

$$u(x) = \sum_{k=1}^{\infty} \frac{1}{2^k}|x - r_k|^{-\alpha} \quad (x \in U).$$

Then $u \in W^{1,p}(U)$ if and only if $\alpha < \frac{n-p}{p}$. If $0 < \alpha < \frac{n-p}{p}$, we see that u belongs to $W^{1,p}(U)$ and yet is unbounded on each open subset of U. □

This last example illustrates a fundamental fact of life, that although a function u belonging to a Sobolev space possesses certain smoothness properties, it can still be rather badly behaved in other ways.

5.2.3. Elementary properties.

Next we verify certain properties of weak derivatives. Note very carefully that whereas these various rules are obviously true for smooth functions, functions in Sobolev space are not necessarily smooth: we must always rely solely upon the definition of weak derivatives.

THEOREM 1 (Properties of weak derivatives). *Assume $u, v \in W^{k,p}(U)$, $|\alpha| \le k$. Then*

(i) *$D^\alpha u \in W^{k-|\alpha|,p}(U)$ and $D^\beta(D^\alpha u) = D^\alpha(D^\beta u) = D^{\alpha+\beta}u$ for all multiindices α, β with $|\alpha| + |\beta| \le k$.*

(ii) *For each $\lambda, \mu \in \mathbb{R}$, $\lambda u + \mu v \in W^{k,p}(U)$ and $D^\alpha(\lambda u + \mu v) = \lambda D^\alpha u + \mu D^\alpha v$, $|\alpha| \le k$.*

(iii) *If V is an open subset of U, then $u \in W^{k,p}(V)$.*

(iv) *If $\zeta \in C_c^\infty(U)$, then $\zeta u \in W^{k,p}(U)$ and*

$$D^\alpha(\zeta u) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} D^\beta \zeta D^{\alpha-\beta} u \quad \textit{(Leibniz' formula)}, \tag{7}$$

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

Proof. 1. To prove (i), first fix $\phi \in C_c^\infty(U)$. Then $D^\beta\phi \in C_c^\infty(U)$, and so

$$\begin{aligned}\int_U D^\alpha u D^\beta \phi \, dx &= (-1)^{|\alpha|} \int_U u D^{\alpha+\beta}\phi \, dx \\ &= (-1)^{|\alpha|}(-1)^{|\alpha+\beta|} \int_U D^{\alpha+\beta} u \phi \, dx \\ &= (-1)^{|\beta|} \int_U D^{\alpha+\beta} u \phi \, dx.\end{aligned}$$

Thus $D^\beta(D^\alpha u) = D^{\alpha+\beta}u$ in the weak sense.

2. Assertions (ii) and (iii) are easy, and the proofs are omitted.

3. We prove (7) by induction on $|\alpha|$. Suppose first $|\alpha| = 1$. Choose any $\phi \in C_c^\infty(U)$. Then

$$\int_U \zeta u D^\alpha \phi \, dx = \int_U u D^\alpha(\zeta\phi) - u(D^\alpha \zeta)\phi \, dx$$
$$= -\int_U (\zeta D^\alpha u + u D^\alpha \zeta)\phi \, dx.$$

Thus $D^\alpha(\zeta u) = \zeta D^\alpha u + u D^\alpha \zeta$, as required.

Next assume $l < k$ and formula (7) is valid for all $|\alpha| \leq l$ and all functions ζ. Choose a multiindex α with $|\alpha| = l+1$. Then $\alpha = \beta + \gamma$ for some $|\beta| = l$, $|\gamma| = 1$. Then for ϕ as above,

$$\int_U \zeta u D^\alpha \phi \, dx = \int_U \zeta u D^\beta(D^\gamma \phi)\, dx$$
$$= (-1)^{|\beta|} \int_U \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^\sigma \zeta D^{\beta-\sigma} u D^\gamma \phi \, dx$$

(by the induction assumption)

$$= (-1)^{|\beta|+|\gamma|} \int_U \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^\gamma (D^\sigma \zeta D^{\beta-\sigma} u)\phi \, dx$$

(by the induction assumption again)

$$= (-1)^{|\alpha|} \int_U \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} [D^\rho \zeta D^{\alpha-\rho} u + D^\sigma \zeta D^{\alpha-\sigma} u]\phi \; dx$$

(where $\rho = \sigma + \gamma$)

$$= (-1)^{|\alpha|} \int_U \left[\sum_{\sigma \leq \alpha} \binom{\alpha}{\sigma} D^\sigma \zeta D^{\alpha-\sigma} u) \right] \phi \; dx,$$

since

$$\binom{\beta}{\sigma - \gamma} + \binom{\beta}{\sigma} = \binom{\alpha}{\sigma}.$$

□

Not only do many of the usual rules of calculus apply to weak derivatives, but the Sobolev spaces themselves have a good mathematical structure:

THEOREM 2 (Sobolev spaces as function spaces). *For each $k = 1, \dots$ and $1 \le p \le \infty$, the Sobolev space $W^{k,p}(U)$ is a Banach space.*

Proof. 1. Let us first of all check that $\|u\|_{W^{k,p}(U)}$ is a norm. (See the discussion at the end of §5.1, or refer to §D.1, for definitions.) Clearly

$$\|\lambda u\|_{W^{k,p}(U)} = |\lambda| \|u\|_{W^{k,p}(U)},$$

and

$$\|u\|_{W^{k,p}(U)} = 0 \text{ if and only if } u = 0 \text{ a.e.}$$

Next assume $u, v \in W^{k,p}(U)$. Then if $1 \le p < \infty$, Minkowski's inequality (§B.2) implies

$$\begin{aligned}
\|u+v\|_{W^{k,p}(U)} &= \Big(\sum_{|\alpha| \le k} \|D^\alpha u + D^\alpha v\|^p_{L^p(U)} \Big)^{1/p} \\
&\le \Big(\sum_{|\alpha| \le k} (\|D^\alpha u\|_{L^p(U)} + \|D^\alpha v\|_{L^p(U)})^p \Big)^{1/p} \\
&\le \Big(\sum_{|\alpha| \le k} \|D^\alpha u\|^p_{L^p(U)} \Big)^{1/p} + \Big(\sum_{|\alpha| \le k} \|D^\alpha v\|^p_{L^p(U)} \Big)^{1/p} \\
&= \|u\|_{W^{k,p}(U)} + \|v\|_{W^{k,p}(U)}.
\end{aligned}$$

2. It remains to show that $W^{k,p}(U)$ is complete. So assume $\{u_m\}_{m=1}^\infty$ is a Cauchy sequence in $W^{k,p}(U)$. Then for each $|\alpha| \le k$, $\{D^\alpha u_m\}_{m=1}^\infty$ is a Cauchy sequence in $L^p(U)$. Since $L^p(U)$ is complete, there exist functions $u_\alpha \in L^p(U)$ such that

$$D^\alpha u_m \to u_\alpha \quad \text{in } L^p(U)$$

for each $|\alpha| \le k$. In particular,

$$u_m \to u_{(0,\dots,0)} =: u \quad \text{in } L^p(U).$$

3. We now claim

$$u \in W^{k,p}(U), \ D^\alpha u = u_\alpha \quad (|\alpha| \le k). \tag{8}$$

To verify this assertion, fix $\phi \in C_c^\infty(U)$. Then

$$\begin{aligned}
\int_U u D^\alpha \phi \, dx &= \lim_{m\to\infty} \int_U u_m D^\alpha \phi \, dx \\
&= \lim_{m\to\infty} (-1)^{|\alpha|} \int_U D^\alpha u_m \phi \, dx \\
&= (-1)^{|\alpha|} \int_U u_\alpha \phi \, dx.
\end{aligned}$$

Thus (8) is valid. Since therefore $D^\alpha u_m \to D^\alpha u$ in $L^p(U)$ for all $|\alpha| \le k$, we see that $u_m \to u$ in $W^{k,p}(U)$, as required. □

5.3. APPROXIMATION

5.3.1. Interior approximation by smooth functions.

It is awkward to return continually to the definition of weak derivatives. In order to study the deeper properties of Sobolev spaces, we therefore need to develop some systematic procedures for approximating a function in a Sobolev space by smooth functions. The method of mollifiers, developed in §C.4, provides the tool.

Fix a positive integer k and $1 \le p < \infty$. Remember that $U_\varepsilon = \{x \in U \mid \operatorname{dist}(x, \partial U) > \varepsilon\}$.

THEOREM 1 (Local approximation by smooth functions). *Assume $u \in W^{k,p}(U)$ for some $1 \le p < \infty$, and set*

$$u^\varepsilon = \eta_\varepsilon * u \quad \textit{in } U_\varepsilon.$$

Then

(i) $u^\varepsilon \in C^\infty(U_\varepsilon)$ *for each* $\varepsilon > 0$,

and

(ii) $u^\varepsilon \to u$ *in* $W^{k,p}_{\mathrm{loc}}(U)$, *as* $\varepsilon \to 0$.

Proof. 1. Assertion (i) is proved in §C.4.

2. We next claim that if $|\alpha| \le k$, then

$$D^\alpha u^\varepsilon = \eta_\varepsilon * D^\alpha u \quad \text{in } U_\varepsilon; \tag{1}$$

that is, *the ordinary α^{th}-partial derivative of the smooth function u^ε is the ε-mollification of the α^{th}-weak partial derivative of u.* To confirm this, we compute for $x \in U_\varepsilon$

$$\begin{aligned} D^\alpha u^\varepsilon(x) &= D^\alpha \int_U \eta_\varepsilon(x-y)u(y)\,dy \\ &= \int_U D^\alpha_x \eta_\varepsilon(x-y)u(y)\,dy \\ &= (-1)^{|\alpha|} \int_U D^\alpha_y \eta_\varepsilon(x-y)u(y)\,dy. \end{aligned}$$

Now for fixed $x \in U_\varepsilon$ the function $\phi(y) := \eta_\varepsilon(x-y)$ belongs to $C^\infty_c(U)$. Consequently the definition of the α^{th}-weak partial derivative implies:

$$\int_U D^\alpha_y \eta_\varepsilon(x-y)u(y)\,dy = (-1)^{|\alpha|}\int_U \eta_\varepsilon(x-y)D^\alpha u(y)\,dy.$$

Thus

$$\begin{aligned} D^\alpha u^\varepsilon(x) &= (-1)^{|\alpha|+|\alpha|}\int_U \eta_\varepsilon(x-y)D^\alpha u(y)\,dy \\ &= [\eta_\varepsilon * D^\alpha u]\,(x). \end{aligned}$$

This establishes (1).

3. Now choose an open set $V \subset\subset U$. In view of (1) and §C.4, $D^\alpha u^\varepsilon \to D^\alpha u$ in $L^p(V)$ as $\varepsilon \to 0$, for each $|\alpha| \leq k$. Consequently

$$\|u^\varepsilon - u\|^p_{W^{k,p}(V)} = \sum_{|\alpha|\leq k} \|D^\alpha u^\varepsilon - D^\alpha u\|^p_{L^p(V)} \to 0$$

as $\varepsilon \to 0$. This proves assertion (ii). □

5.3.2. Approximation by smooth functions.

Next we show that we can find smooth functions which approximate in $W^{k,p}(U)$, and not just in $W^{k,p}_{\text{loc}}(U)$. Notice in the following that we make no assumptions about the smoothness of ∂U.

THEOREM 2 (Global approximation by smooth functions). *Assume U is bounded, and suppose as well that $u \in W^{k,p}(U)$ for some $1 \leq p < \infty$. Then there exist functions $u_m \in C^\infty(U) \cap W^{k,p}(U)$ such that*

$$u_m \to u \quad \textit{in } W^{k,p}(U).$$

Remark. Note carefully that we do not assert $u_m \in C^\infty(\bar{U})$ (but see Theorem 3 below). □

Proof. 1. We have $U = \bigcup_{i=1}^\infty U_i$, where

$$U_i := \{x \in U \mid \operatorname{dist}(x, \partial U) > 1/i\} \quad (i = 1, 2, \dots).$$

Write $V_i := U_{i+3} - \bar{U}_{i+1}$.

Choose also any open set $V_0 \subset\subset U$ so that $U = \bigcup_{i=0}^\infty V_i$. Now let $\{\zeta_i\}_{i=0}^\infty$ be a smooth partition of unity subordinate to the open sets $\{V_i\}_{i=0}^\infty$; that is, suppose

$$(2) \qquad \begin{cases} 0 \leq \zeta_i \leq 1, & \zeta_i \in C_c^\infty(V_i) \\ \sum_{i=0}^\infty \zeta_i = 1 & \text{on } U. \end{cases}$$

Next, choose any function $u \in W^{k,p}(U)$. According to Theorem 1(iv) in §5.2, $\zeta_i u \in W^{k,p}(U)$ and $\operatorname{spt}(\zeta_i u) \subset V_i$.

2. Fix $\delta > 0$. Choose then $\varepsilon_i > 0$ so small that $u^i := \eta_{\varepsilon_i} * (\zeta_i u)$ satisfies

$$
(3) \qquad \begin{cases} \|u^i - \zeta_i u\|_{W^{k,p}(U)} \leq \frac{\delta}{2^{i+1}} & (i = 0, 1, \dots) \\ \operatorname{spt} u^i \subset W_i & (i = 1, \dots), \end{cases}
$$

for $W_i := U_{i+4} - \bar{U}_i \supset V_i$ $(i = 1, \dots)$.

3. Write $v := \sum_{i=0}^{\infty} u^i$. This function belongs to $C^\infty(U)$, since for each open set $V \subset\subset U$ there are at most finitely many nonzero terms in the sum. Since $u = \sum_{i=0}^{\infty} \zeta_i u$, we have for each $V \subset\subset U$

$$
\begin{aligned}
\|v - u\|_{W^{k,p}(V)} &\leq \sum_{i=0}^{\infty} \|u^i - \zeta_i u\|_{W^{k,p}(U)} \\
&\leq \delta \sum_{i=0}^{\infty} \frac{1}{2^{i+1}} \quad \text{by (3)} \\
&= \delta.
\end{aligned}
$$

Take the supremum over sets $V \subset\subset U$, to conclude $\|v - u\|_{W^{k,p}(U)} \leq \delta$. □

5.3.3. Global approximation by smooth functions.

We now ask when it is possible to approximate a given function $u \in W^{k,p}(U)$ by functions belonging to $C^\infty(\bar{U})$, and not just $C^\infty(U)$. Such an approximation requires some condition to exclude ∂U being wild geometrically.

THEOREM 3 (Global approximation by functions smooth up to the boundary). *Assume U is bounded and ∂U is C^1. Suppose $u \in W^{k,p}(U)$ for some $1 \leq p < \infty$. Then there exist functions $u_m \in C^\infty(\bar{U})$ such that*

$$
u_m \to u \quad \text{in } W^{k,p}(U).
$$

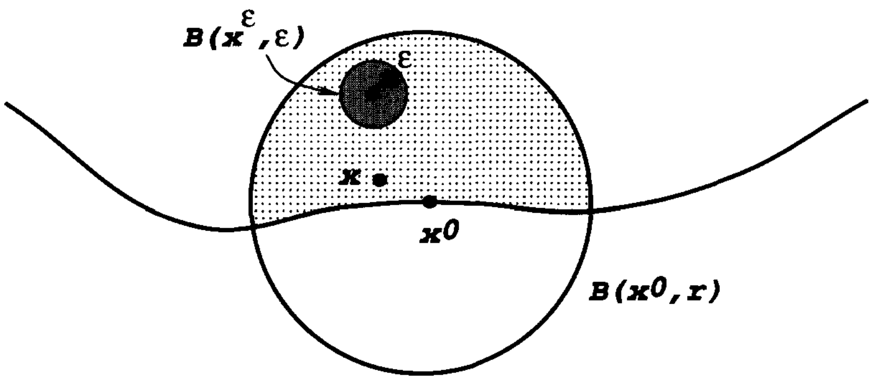
Proof. 1. Fix any point $x^0 \in \partial U$. As ∂U is C^1, there exist, according to §C.1, a radius $r > 0$ and a C^1 function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that—upon relabeling the coordinate axes if necessary—we have

$$
U \cap B(x^0, r) = \{x \in B(x^0, r) \mid x_n > \gamma(x_1, \dots, x_{n-1})\}.
$$

Set $V := U \cap B(x^0, r/2)$.

2. Define the shifted point

$$
x^\varepsilon := x + \lambda \varepsilon e_n \quad (x \in V,\ \varepsilon > 0),
$$



and observe that for some fixed, sufficiently large number $\lambda > 0$ the ball $B(x^\varepsilon, \varepsilon)$ lies in $U \cap B(x^0, r)$ for all $x \in V$ and all small $\varepsilon > 0$.

Now define $u_\varepsilon(x) := u(x^\varepsilon)$ $(x \in V)$; this is the function u translated a distance $\lambda\varepsilon$ in the e_n direction. Next write $v^\varepsilon = \eta_\varepsilon * u_\varepsilon$. The idea is that we have moved up enough so that "there is room to mollify within U". Clearly $v^\varepsilon \in C^\infty(\bar{V})$.

3. We now claim

$$v^\varepsilon \to u \quad \text{in } W^{k,p}(V). \tag{4}$$

To confirm this, take α to be any multiindex with $|\alpha| \leq k$. Then

$$\|D^\alpha v^\varepsilon - D^\alpha u\|_{L^p(V)} \leq \|D^\alpha v^\varepsilon - D^\alpha u_\varepsilon\|_{L^p(V)} + \|D^\alpha u_\varepsilon - D^\alpha u\|_{L^p(V)}.$$

The second term on the right hand side goes to zero with ε, since translation is continuous in the L^p-norm; and the first term also vanishes in the limit, by reasoning similar to that in the proof of Theorem 1.

4. Select $\delta > 0$. Since ∂U is compact, we can find finitely many points $x_i^0 \in \partial U$, radii $r_i > 0$, corresponding sets $V_i = U \cap B\left(x_i^0, \frac{r_i}{2}\right)$, and functions $v_i \in C^\infty(\bar{V}_i)$ $(i = 1, \dots, N)$ such that $\partial U \subset \bigcup_{i=1}^N B^0\left(x_i^0, \frac{r_i}{2}\right)$, and

$$\|v_i - u\|_{W^{k,p}(V_i)} \leq \delta. \tag{5}$$

Take an open set $V_0 \subset\subset U$ such that $U \subset \bigcup_{i=0}^N V_i$ and select, using Theorem 1, a function $v_0 \in C^\infty(\bar{V}_0)$ satisfying

$$\|v_0 - u\|_{W^{k,p}(V_0)} \leq \delta. \tag{6}$$

5. Now let $\{\zeta_i\}_{i=0}^N$ be a smooth partition of unity subordinate to the open sets $\{V_i\}_{i=0}^N$ in U. Define $v := \sum_{i=0}^N \zeta_i v_i$. Then clearly $v \in C^\infty(\bar{U})$. In

addition, since $u = \sum_{i=0}^N \zeta_i u$, we see using Theorem 1 in §5.2.3 that for each $|\alpha| \leq k$:

$$\begin{aligned}\|D^\alpha v - D^\alpha u\|_{L^p(U)} &\leq \sum_{i=0}^N \|D^\alpha(\zeta_i v_i) - D^\alpha(\zeta_i u)\|_{L^p(V_i)} \\ &\leq C\sum_{i=0}^N \|v_i - u\|_{W^{k,p}(V_i)} = CN\delta,\end{aligned}$$

according to (5) and (6). □

5.4. EXTENSIONS

Our goal next is to extend functions in the Sobolev space $W^{1,p}(U)$ to become functions in the Sobolev space $W^{1,p}(\mathbb{R}^n)$. This can be subtle. Observe for instance that our extending $u \in W^{1,p}(U)$ to be zero in $\mathbb{R}^n - U$ will not in general work, as we may thereby create such a bad discontinuity along ∂U that the extended function no longer has a weak first partial derivative. We must instead invent a way to extend u which "preserves the weak derivatives across ∂U".

Suppose $1 \leq p \leq \infty$.

THEOREM 1 (Extension Theorem). *Assume U is bounded and ∂U is C^1. Select a bounded open set V such that $U \subset\subset V$. Then there exists a bounded linear operator*

(1) $$E : W^{1,p}(U) \to W^{1,p}(\mathbb{R}^n)$$

such that for each $u \in W^{1,p}(U)$:

(i) *$Eu = u$ a.e. in U,*

(ii) *Eu has support within V,*

and

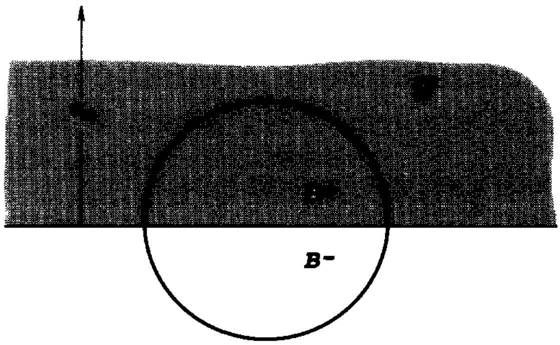
(iii)
$$\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(U)},$$

the constant C depending only on p, U, and V.

DEFINITION. *We call Eu an* extension *of u to $\mathbb{R}^n$.*

Proof. 1. Fix $x^0 \in \partial U$ and suppose first

(2) ∂U is flat near x^0, lying in the plane $\{x_n = 0\}$.



A half-ball at the boundary

Then we may assume there exists an open ball B, with center x^0 and radius r, such that

$$\begin{cases} B^+ := B \cap \{x_n \geq 0\} \subset \bar{U} \\ B^- := B \cap \{x_n \leq 0\} \subset \mathbb{R}^n - U. \end{cases}$$

2. Temporarily suppose also $u \in C^\infty(\bar{U})$. We define then

(3)
$$\bar{u}(x) := \begin{cases} u(x) & \text{if } x \in B^+ \\ -3u(x_1, \ldots, x_{n-1}, -x_n) + 4u(x_1, \ldots, x_{n-1}, -\frac{x_n}{2}) & \text{if } x \in B^-. \end{cases}$$

This is called a *higher-order reflection* of u from B^+ to B^-.

3. We claim

(4)
$$\bar{u} \in C^1(B).$$

To check this, let us write $u^- := \bar{u}|_{B^-}$, $u^+ := \bar{u}|_{B^+}$. We demonstrate first

(5)
$$u^-_{x_n} = u^+_{x_n} \quad \text{on } \{x_n = 0\}.$$

Indeed according to (3),

$$\frac{\partial u^-}{\partial x_n}(x) = 3\frac{\partial u}{\partial x_n}(x_1, \ldots, x_{n-1}, -x_n) - 2\frac{\partial u}{\partial x_n}(x_1, \ldots, x_{n-1}, -\frac{x_n}{2})$$

and so

$$u^-_{x_n}\big|_{\{x_n=0\}} = u^+_{x_n}\big|_{\{x_n=0\}}.$$

This confirms (5). Now since $u^+ = u^-$ on $\{x_n = 0\}$, we see as well that

(6)
$$u^-_{x_i}|_{\{x_n=0\}} = u^+_{x_i}|_{\{x_n=0\}}$$

for $i = 1, \ldots, n-1$. But then (5) and (6) together imply

$$D^\alpha u^-|_{\{x_n=0\}} = D^\alpha u^+|_{\{x_n=0\}}$$

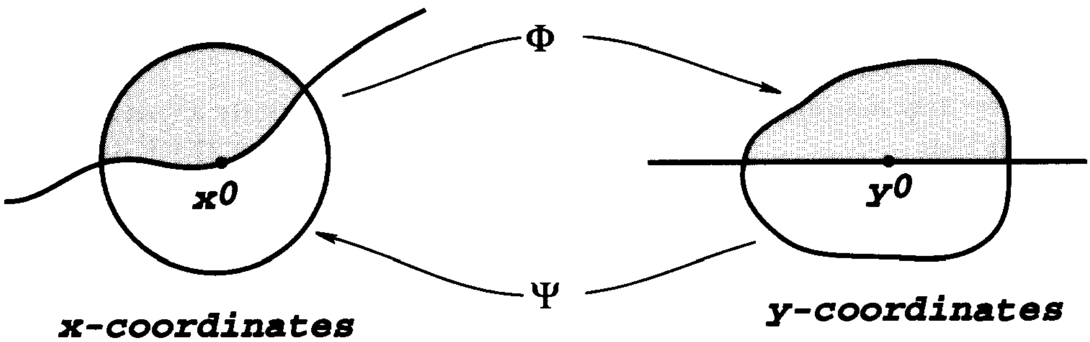
for each $|\alpha| \leq 1$, and so (4) follows.

4. Using these calculations we readily check as well

$$\|\bar{u}\|_{W^{1,p}(B)} \leq C\|u\|_{W^{1,p}(B^+)} \tag{7}$$

for some constant C which does not depend on u.

5. Let us next consider the situation that ∂U is not necessarily flat near x^0. Then utilizing the notation and terminology from §C.1, we can find a C^1 mapping $\mathbf{\Phi}$, with inverse $\mathbf{\Psi}$, such that $\mathbf{\Phi}$ "straightens out ∂U near x^0".



Straightening out the boundary

We write $y = \mathbf{\Phi}(x)$, $x = \mathbf{\Psi}(y)$, $u'(y) := u(\mathbf{\Psi}(y))$. Choose a small ball B as drawn before. Then utilizing steps 1–3 above we extend u' from B^+ to a function $\bar{u}'$ defined on all of B, such that $\bar{u}'$ is C^1 and we have the estimate

$$\|\bar{u}'\|_{W^{1,p}(B)} \leq C\|u'\|_{W^{1,p}(B^+)}.$$

Let $W := \mathbf{\Psi}(B)$. Then converting back to the x-variables, we obtain an extension $\bar{u}$ of u to W, with

$$\|\bar{u}\|_{W^{1,p}(W)} \leq C\|u\|_{W^{1,p}(U)}. \tag{8}$$

6. Since ∂U is compact, there exist finitely many points $x_i^0 \in \partial U$, open sets W_i, and extensions $\bar{u}_i$ of u to W_i $(i = 1, \dots, N)$, as above, such that $\Gamma \subset \bigcup_{i=1}^N W_i$. Take $W_0 \subset\subset U$ so that $U \subset \bigcup_{i=0}^N W_i$, and let $\{\zeta_i\}_{i=0}^N$ be an associated partition of unity. Write $\bar{u} := \sum_{i=0}^N \zeta_i \bar{u}_i$, where $\bar{u}_0 = u$. Then utilizing estimate (8) (with u_i in place of u, $\bar{u}_i$ in place of $\bar{u}$) we obtain the bound

$$\|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(U)} \tag{9}$$

for some constant C, depending on U, p, n, etc., but not on u. Furthermore we can arrange for the support of $\bar{u}$ to lie within $V \supset\supset U$.

7. We henceforth write $Eu := \bar{u}$, and observe that the mapping $u \mapsto Eu$ is linear.

Recall that the construction so far assumed $u \in C^\infty(\bar{U})$. Suppose now $u \in W^{1,p}(U)$, and choose $u_m \in C^\infty(\bar{U})$ converging to u in $W^{1,p}(U)$. Estimate (9) and the linearity of E imply

$$\|Eu_m - Eu_l\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u_m - u_l\|_{W^{1,p}(U)}.$$

Thus $\{Eu_m\}_{m=1}^\infty$ is a Cauchy sequence and so converges to $\bar{u} =: Eu$. This extension, which does not depend on the particular choice of the approximating sequence $\{u_m\}_{m=1}^\infty$, satisfies the conclusions of the theorem. □

Remarks. (i) Assume now that ∂U is C^2. Then the extension operator E constructed above is also a bounded linear operator from $W^{2,p}(U)$ to $W^{2,p}(\mathbb{R}^n)$. To see this, note first in steps 3, 4 of the proof that although $\bar{u}$ is not in general C^2, it does belong to $W^{2,p}(B)$. We also have the bound

$$\|\bar{u}\|_{W^{2,p}(B)} \leq C\|u\|_{W^{2,p}(B^+)},$$

which follows from the definition (3). As before, we consequently derive the estimate

$$\|Eu\|_{W^{2,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{2,p}(U)}, \tag{10}$$

provided ∂U is C^2, the constants C depending only only on U, V, n and p.

We will need these observations later.

(ii) The above construction does *not* provide us with an extension for the Sobolev spaces $W^{k,p}(U)$, if $k > 2$. This requires a more complicated higher-order reflection technique. □

5.5. TRACES

Next we discuss the possibility of assigning "boundary values" along ∂U to a function $u \in W^{1,p}(U)$, assuming that ∂U is C^1. Now if $u \in C(\bar{U})$, then clearly u has values on ∂U in the usual sense. The problem is that a typical function $u \in W^{1,p}(U)$ is not in general continuous and, even worse, is only defined a.e. in U. Since ∂U has n-dimensional Lebesgue measure zero, there is no direct meaning we can give to the expression "u restricted to ∂U". The notion of a *trace operator* resolves this problem.

For this section we take $1 \leq p < \infty$.

THEOREM 1 (Trace Theorem). *Assume U is bounded and ∂U is C^1. Then there exists a bounded linear operator*

$$T : W^{1,p}(U) \to L^p(\partial U)$$

such that

(i) $Tu = u|_{\partial U}$ *if* $u \in W^{1,p}(U) \cap C(\bar{U})$

and

(ii)

$$\|Tu\|_{L^p(\partial U)} \leq C\|u\|_{W^{1,p}(U)},$$

for each $u \in W^{1,p}(U)$, with the constant C depending only on p and U.

DEFINITION. *We call Tu the* trace *of u on ∂U.*

Proof. 1. Assume first $u \in C^1(\bar{U})$. As in the first part of the proof of Theorem 1 in §5.4 let us also initially suppose $x^0 \in \partial U$ and ∂U is flat near x^0, lying in the plane $\{x_n = 0\}$. Choose an open ball B as in the previous proof and let $\hat{B}$ denote the concentric ball with radius $r/2$.

Select $\zeta \in C_c^\infty(B)$, with $\zeta \geq 0$ in B, $\zeta \equiv 1$ on $\hat{B}$. Denote by Γ that portion of ∂U within $\hat{B}$. Set $x' = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} = \{x_n = 0\}$.

Then

$$\begin{aligned} \int_\Gamma |u|^p \, dx' &\leq \int_{\{x_n=0\}} \zeta |u|^p \, dx' = -\int_{B^+} (\zeta |u|^p)_{x_n} \, dx \\ &= -\int_{B^+} |u|^p \zeta_{x_n} + p|u|^{p-1}(\operatorname{sgn} u) u_{x_n} \zeta \, dx \qquad (1) \\ &\leq C \int_{B^+} |u|^p + |Du|^p \, dx, \end{aligned}$$

where we employed Young's inequality, from §B.2.

2. If $x^0 \in \partial U$, but ∂U is not flat near x^0, we as usual straighten out the boundary near x^0 to obtain the setting above. Applying estimate (1) and changing variables, we obtain the bound

$$\int_\Gamma |u|^p dS \leq C \int_U |u|^p + |Du|^p \, dx,$$

where Γ is some open subset of ∂U containing x^0.

3. Since ∂U is compact, there exist finitely many points $x_i^0 \in \partial U$ and open subsets $\Gamma_i \subset \partial U$ $(i = 1, \ldots, N)$ such that $\partial U = \bigcup_{i=1}^N \Gamma_i$ and

$$\|u\|_{L^p(\Gamma_i)} \leq C\|u\|_{W^{1,p}(U)} \quad (i = 1, \ldots, N).$$

Consequently, if we write

$$Tu := u|_{\partial U},$$

then

$$\|Tu\|_{L^p(\partial U)} \le C\|u\|_{W^{1,p}(U)} \tag{2}$$

for some appropriate constant C, which does not depend on u.

4. Inequality (2) holds for $u \in C^1(\bar{U})$. Assume now $u \in W^{1,p}(U)$. Then there exist functions $u_m \in C^\infty(\bar{U})$ converging to u in $W^{1,p}(U)$. According to (2) we have

$$\|Tu_m - Tu_l\|_{L^p(\partial U)} \le C\|u_m - u_l\|_{W^{1,p}(U)}; \tag{3}$$

so that $\{Tu_m\}_{m=1}^\infty$ is a Cauchy sequence in $L^p(\partial U)$. We define

$$Tu := \lim_{m\to\infty} Tu_m,$$

the limit taken in $L^p(\partial U)$. According to (3) this definition does not depend on the particular choice of smooth functions approximating u.

Finally if $u \in W^{1,p}(U) \cap C(\bar{U})$, we note that the functions $u_m \in C^\infty(\bar{U})$ constructed in the proof of Theorem 3 in §5.3.3 converge uniformly to u on $\bar{U}$. Hence $Tu = u|_{\partial U}$. □

We next examine more closely what it means for a function to have zero trace.

THEOREM 2 (Trace-zero functions in $W^{1,p}$). *Assume U is bounded and ∂U is C^1. Suppose furthermore that $u \in W^{1,p}(U)$. Then*

$$u \in W_0^{1,p}(U) \quad \textit{if and only if} \quad Tu = 0 \textit{ on } \partial U. \tag{4}$$

Proof*. 1. Suppose first $u \in W_0^{1,p}(U)$. Then by definition there exist functions $u_m \in C_c^\infty(U)$ such that

$$u_m \to u \quad \text{in } W^{1,p}(U).$$

As $Tu_m = 0$ on ∂U $(m = 1, \dots)$ and $T : W^{1,p}(U) \to L^p(\partial U)$ is a bounded linear operator, we deduce $Tu = 0$ on ∂U.

2. The converse statement is more difficult. Let us assume that

$$Tu = 0 \quad \text{on } \partial U. \tag{5}$$

*Omit on first reading.

Using partitions of unity and flattening out ∂U as usual, we may as well assume

$$\text{(6)} \qquad \begin{cases} u \in W^{1,p}(\mathbb{R}^n_+), \; u \text{ has compact support in } \bar{\mathbb{R}}^n_+, \\ Tu = 0 \text{ on } \partial\mathbb{R}^n_+ = \mathbb{R}^{n-1}. \end{cases}$$

Then since $Tu = 0$ on $\mathbb{R}^{n-1}$, there exist functions $u_m \in C^1(\bar{\mathbb{R}}^n_+)$ such that

$$\text{(7)} \qquad u_m \to u \quad \text{in } W^{1,p}(\mathbb{R}^n_+)$$

and

$$\text{(8)} \qquad Tu_m = u_m|_{\mathbb{R}^{n-1}} \to 0 \quad \text{in } L^p(\mathbb{R}^{n-1}).$$

Now if $x' \in \mathbb{R}^{n-1}$, $x_n \geq 0$, we have

$$|u_m(x', x_n)| \leq |u_m(x', 0)| + \int_0^{x_n} |u_{m,x_n}(x', t)|\, dt.$$

Thus

$$\int_{\mathbb{R}^{n-1}} |u_m(x', x_n)|^p\, dx' \leq C\left(\int_{\mathbb{R}^{n-1}} |u_m(x', 0)|^p\, dx' + x_n^{p-1}\int_0^{x_n}\int_{\mathbb{R}^{n-1}} |Du_m(x', t)|^p\, dx'dt\right).$$

Letting $m \to \infty$ and recalling (7), (8), we deduce:

$$\text{(9)} \qquad \int_{\mathbb{R}^{n-1}} |u(x', x_n)|^p\, dx' \leq Cx_n^{p-1}\int_0^{x_n}\int_{\mathbb{R}^{n-1}} |Du|^p\, dx'dt$$

for a.e. $x_n > 0$.

3. Next let $\zeta \in C^\infty(\mathbb{R})$ satisfy

$$\zeta \equiv 1 \text{ on } [0, 1], \; \zeta \equiv 0 \text{ on } \mathbb{R} - [0, 2], \; 0 \leq \zeta \leq 1,$$

and write

$$\begin{cases} \zeta_m(x) := \zeta(mx_n) \quad (x \in \mathbb{R}^n_+) \\ w_m := u(x)(1 - \zeta_m). \end{cases}$$

Then

$$\begin{cases} w_{m,x_n} = u_{x_n}(1 - \zeta_m) - mu\zeta' \\ D_{x'}w_m = D_{x'}u(1 - \zeta_m). \end{cases}$$

Consequently

$$\text{(10)} \qquad \begin{aligned} \int_{\mathbb{R}^n_+} |Dw_m - Du|^p\, dx &\leq C\int_{\mathbb{R}^n_+} |\zeta_m|^p |Du|^p\, dx \\ &\quad + Cm^p\int_0^{2/m}\int_{\mathbb{R}^{n-1}} |u|^p\, dx'dt \\ &=: A + B. \end{aligned}$$

Now

$$\text{(11)} \qquad A \to 0 \quad \text{as } m \to \infty,$$

since $\zeta_m \neq 0$ only if $0 \leq x_n \leq 2/m$. To estimate the term B, we utilize inequality (9):

$$\text{(12)} \qquad \begin{aligned} B &\leq Cm^p \left(\int_0^{2/m} t^{p-1} dt \right) \left(\int_0^{2/m} \int_{\mathbb{R}^{n-1}} |Du|^p \, dx' dx_n \right) \\ &\leq C \int_0^{2/m} \int_{\mathbb{R}^{n-1}} |Du|^p \, dx' dx_n \to 0 \quad \text{as } m \to 0. \end{aligned}$$

Employing (10)–(12), we deduce $Dw_m \to Du$ in $L^p(\mathbb{R}^n_+)$. Since clearly $w_m \to u$ in $L^p(\mathbb{R}^n_+)$, we conclude

$$w_m \to u \quad \text{in } W^{1,p}(\mathbb{R}^n_+).$$

But $w_m = 0$ if $0 < x_n < 1/m$. We can therefore mollify the w_m to produce functions $u_m \in C_c^\infty(\mathbb{R}^n_+)$ such that $u_m \to u$ in $W^{1,p}(\mathbb{R}^n_+)$. Hence $u \in W_0^{1,p}(\mathbb{R}^n_+)$. □

5.6. SOBOLEV INEQUALITIES

Our goal in this section is to discover embeddings of various Sobolev spaces into others. The crucial analytic tools here will be certain so-called "Sobolev-type inequalities", which we will prove below for smooth functions. These will then establish the estimates for arbitrary functions in the various relevant Sobolev spaces, since—as we saw in §5.2—smooth functions are dense.

To clarify the presentation we will consider first only the Sobolev space $W^{1,p}(U)$ and ask the following basic question: *If a function u belongs to $W^{1,p}(U)$, does u automatically belong to certain other spaces?* The answer will be "yes", but which other spaces depends upon whether

$$\text{(1)} \qquad 1 \leq p < n,$$

$$\text{(2)} \qquad p = n,$$

$$\text{(3)} \qquad n < p \leq \infty.$$

We study case (1) in §5.6.1, case (3) in §5.6.2, and the borderline case (2) only later in §5.8.1.

5.6.1. Gagliardo–Nirenberg–Sobolev inequality.

For this section let us assume

$$1 \le p < n \tag{4}$$

and first ask whether we can establish an estimate of the form

$$\|u\|_{L^q(\mathbb{R}^n)} \le C\|Du\|_{L^p(\mathbb{R}^n)}, \tag{5}$$

for certain constants $C > 0$, $1 \le q < \infty$ and all functions $u \in C_c^\infty(\mathbb{R}^n)$. The point is that the constants C and q should not depend on u.

Motivation. Let us first demonstrate that *if* any inequality of the form (5) holds, then the number q cannot be arbitrary, but must in fact have a very specific form. For this, choose a function $u \in C_c^\infty(\mathbb{R}^n)$, $u \not\equiv 0$, and define for $\lambda > 0$ the rescaled function

$$u_\lambda(x) := u(\lambda x) \quad (x \in \mathbb{R}^n).$$

Applying (5) to u_λ, we find:

$$\|u_\lambda\|_{L^q(\mathbb{R}^n)} \le C\|Du_\lambda\|_{L^p(\mathbb{R}^n)}. \tag{6}$$

Now

$$\int_{\mathbb{R}^n} |u_\lambda|^q dx = \int_{\mathbb{R}^n} |u(\lambda x)|^q dx = \frac{1}{\lambda^n}\int_{\mathbb{R}^n} |u(y)|^q dy,$$

and

$$\int_{\mathbb{R}^n} |Du_\lambda|^p dx = \lambda^p \int_{\mathbb{R}^n} |Du(\lambda x)|^p dx = \frac{\lambda^p}{\lambda^n}\int_{\mathbb{R}^n} |Du(y)|^p dy.$$

Inserting these equalities into (6), we discover

$$\frac{1}{\lambda^{n/q}}\|u\|_{L^q(\mathbb{R}^n)} \le C\frac{\lambda}{\lambda^{n/p}}\|Du\|_{L^p(\mathbb{R}^n)},$$

and so

$$\|u\|_{L^q(\mathbb{R}^n)} \le C\lambda^{1-\frac{n}{p}+\frac{n}{q}}\|Du\|_{L^p(\mathbb{R}^n)}. \tag{7}$$

But then if $1 - \frac{n}{p} + \frac{n}{q} \ne 0$, we can upon sending λ to either 0 or ∞ in (7) obtain a contradiction. Thus *if* in fact the desired inequality (5) holds, we must necessarily have $1 - \frac{n}{p} + \frac{n}{q} = 0$; so that $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$, $q = \frac{np}{n-p}$. □

This observation motivates the following

DEFINITION. *If* $1 \le p < n$, *the* Sobolev conjugate *of* p *is*

$$p^* := \frac{np}{n-p}. \tag{8}$$

Note that

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \quad p^* > p. \tag{9}$$

The foregoing scaling analysis shows the estimate (5) can only possibly be true for $q = p^*$. Next we prove this inequality is in fact valid.

THEOREM 1 (Gagliardo–Nirenberg–Sobolev inequality). *Assume* $1 \le p < n$. *There exists a constant* C, *depending only on* p *and* n, *such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \le C\|Du\|_{L^p(\mathbb{R}^n)}, \tag{10}$$

for all $u \in C_c^1(\mathbb{R}^n)$.

Now we really do need u to have compact support for (10) to hold, as the example $u \equiv 1$ shows. But remarkably the constant here does not depend at all upon the size of the support of u.

Proof. 1. First assume $p = 1$.

Since u has compact support, for each $i = 1, \dots, n$ and $x \in \mathbb{R}^n$ we have

$$u(x) = \int_{-\infty}^{x_i} u_{x_i}(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)\, dy_i,$$

and so

$$|u(x)| \le \int_{-\infty}^{\infty} |Du(x_1, \dots, y_i, \dots, x_n)|\, dy_i \quad (i = 1, \dots, n).$$

Consequently

$$|u(x)|^{\frac{n}{n-1}} \le \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |Du(x_1, \dots, y_i, \dots, x_n)|\, dy_i \right)^{\frac{1}{n-1}}. \tag{11}$$

Integrate this inequality with respect to x_1:
(12)

$$\begin{aligned}
\int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx_1 &\le \int_{-\infty}^{\infty} \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |Du|\, dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&= \left(\int_{-\infty}^{\infty} |Du|\, dy_1 \right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |Du|\, dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&\le \left(\int_{-\infty}^{\infty} |Du|\, dy_1 \right)^{\frac{1}{n-1}} \left(\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du|\, dx_1 dy_i \right)^{\frac{1}{n-1}},
\end{aligned}$$

the last inequality resulting from the general Hölder inequality (§B.2).

Now integrate (12) with respect to x_2:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u|^{\frac{n}{n-1}}\,dx_1dx_2 \leq \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1dy_2\right)^{\frac{1}{n-1}}\int_{-\infty}^{\infty}\prod_{\substack{i=1\\i\neq 2}}^{n}I_i^{\frac{1}{n-1}}\,dx_2,$$

for

$$I_1 := \int_{-\infty}^{\infty}|Du|\,dy_1,\quad I_i := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1dy_i\quad (i=3,\dots,n).$$

Applying once more the extended Hölder inequality, we find

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u|^{\frac{n}{n-1}}\,dx_1dx_2 &\leq \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1dy_2\right)^{\frac{1}{n-1}}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dy_1dx_2\right)^{\frac{1}{n-1}}\\ &\quad\prod_{i=3}^{n}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1dx_2dy_i\right)^{\frac{1}{n-1}}.\end{aligned}$$

We continue by integrating with respect to $x_3,\dots,x_n$, eventually to find

$$\begin{aligned}\int_{\mathbb{R}^n}|u|^{\frac{n}{n-1}}\,dx &\leq \prod_{i=1}^{n}\left(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|Du|\,dx_1\dots dy_i\dots dx_n\right)^{\frac{1}{n-1}}\\ &= \left(\int_{\mathbb{R}^n}|Du|\,dx\right)^{\frac{n}{n-1}}.\end{aligned}\tag{13}$$

This is estimate (10) for $p=1$.

2. Consider now the case that $1<p<n$. We apply estimate (13) to $v:=|u|^\gamma$, where $\gamma>1$ is to be selected. Then

$$\begin{aligned}\left(\int_{\mathbb{R}^n}|u|^{\frac{\gamma n}{n-1}}dx\right)^{\frac{n-1}{n}} &\leq \int_{\mathbb{R}^n}|D|u|^\gamma|\,dx = \gamma\int_{\mathbb{R}^n}|u|^{\gamma-1}|Du|\,dx\\ &\leq \gamma\left(\int_{\mathbb{R}^n}|u|^{(\gamma-1)\frac{p}{p-1}}dx\right)^{\frac{p-1}{p}}\left(\int_{\mathbb{R}^n}|Du|^p dx\right)^{\frac{1}{p}}.\end{aligned}\tag{14}$$

We choose γ so that $\frac{\gamma n}{n-1} = (\gamma-1)\frac{p}{p-1}$. That is, we set

$$\gamma = \frac{p(n-1)}{n-p} > 1\ ;$$

in which case $\frac{\gamma n}{n-1} = (\gamma - 1)\frac{p}{p-1} = \frac{np}{n-p} = p^*$. Thus, in view of (5), estimate (14) becomes

$$\left(\int_{\mathbb{R}^n} |u|^{p^*} dx\right)^{\frac{1}{p^*}} \le C\left(\int_{\mathbb{R}^n} |Du|^p dx\right)^{\frac{1}{p}}.$$

□

THEOREM 2 (Estimates for $W^{1,p}, 1 \le p < n$). *Let U be a bounded, open subset of $\mathbb{R}^n$, and suppose ∂U is C^1. Assume $1 \le p < n$, and $u \in W^{1,p}(U)$. Then $u \in L^{p^*}(U)$, with the estimate*

$$\|u\|_{L^{p^*}(U)} \le C\|u\|_{W^{1,p}(U)}, \tag{15}$$

the constant C depending only on p, n, and U.

Proof. Since ∂U is C^1, there exists according to Theorem 1 in §5.4 an extension $Eu = \bar{u} \in W^{1,p}(\mathbb{R}^n)$, such that

$$\begin{cases} \bar{u} = u \text{ in } U, \ \bar{u} \text{ has compact support, and} \\ \|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(U)}. \end{cases} \tag{16}$$

Because $\bar{u}$ has compact support, we know from Theorem 1 in §5.3 that there exist functions $u_m \in C_c^\infty(\mathbb{R}^n)$ $(m = 1, 2, \dots)$ such that

$$u_m \to \bar{u} \quad \text{in } W^{1,p}(\mathbb{R}^n). \tag{17}$$

Now according to Theorem 1, $\|u_m - u_l\|_{L^{p^*}(\mathbb{R}^n)} \le C\|Du_m - Du_l\|_{L^p(\mathbb{R}^n)}$ for all $l, m \ge 1$. Thus

$$u_m \to \bar{u} \quad \text{in } L^{p^*}(\mathbb{R}^n) \tag{18}$$

as well. Since Theorem 1 also implies $\|u_m\|_{L^{p^*}(\mathbb{R}^n)} \le C\|Du_m\|_{L^p(\mathbb{R}^n)}$, assertions (17) and (18) yield the bound

$$\|\bar{u}\|_{L^{p^*}(\mathbb{R}^n)} \le C\|D\bar{u}\|_{L^p(\mathbb{R}^n)}.$$

This inequality and (16) complete the proof. □

THEOREM 3 (Estimates for $W_0^{1,p}, 1 \le p < n$). *Assume U is a bounded, open subset of $\mathbb{R}^n$. Suppose $u \in W_0^{1,p}(U)$ for some $1 \le p < n$. Then we have the estimate*

$$\|u\|_{L^q(U)} \le C\|Du\|_{L^p(U)}$$

for each $q \in [1, p^*]$, the constant C depending only on p, q, n and U.

Remark. This estimate is sometimes called *Poincaré's inequality.* The difference with Theorem 2 is that only the gradient of u appears on the righthand side of the inequality. (Other Poincaré-type inequalities will be established later, in §5.8.1.) □

Proof. Since $u \in W_0^{1,p}(U)$, there exist functions $u_m \in C_c^\infty(U)$ ($m = 1,2,\dots$) converging to u in $W^{1,p}(U)$. We extend each function u_m to be 0 on $\mathbb{R}^n - \bar{U}$ and apply Theorem 1 to discover $\|u\|_{L^{p^*}(U)} \le C\|Du\|_{L^p(U)}$. As $|U| < \infty$, we furthermore have $\|u\|_{L^q(U)} \le C\|u\|_{L^{p^*}(U)}$ if $1 \le q \le p^*$. □

Remarks. (i) In view of Theorem 3, on $W_0^{1,p}(U)$ the norm $\|Du\|_{L^p(U)}$ is equivalent to $\|u\|_{W^{1,p}(U)}$, if U is bounded.

(ii) We next consider the case

$$p = n.$$

Owing to Theorem 2 and the fact that $p^* = \frac{np}{n-p} \to +\infty$ as $p \to n$, we might expect $u \in L^\infty(U)$, provided $u \in W^{1,n}(U)$. *This is however false if* $n > 1$: for example, if $U = B^0(0,1)$ the function $u = \log\log\left(1 + \frac{1}{|x|}\right)$ belongs to $W^{1,n}(U)$, but not to $L^\infty(U)$. We will return to this borderline situation in §5.8.1 below. □

5.6.2. Morrey's inequality.

Now let us suppose

$$n < p < \infty. \tag{19}$$

We will show that if $u \in W^{1,p}(U)$, then u is in fact Hölder continuous, after possibly being redefined on a set of measure zero.

THEOREM 4 (Morrey's inequality). *Assume $n < p \le \infty$. Then there exists a constant C, depending only on p and n, such that*

$$\|u\|_{C^{0,\gamma}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{20}$$

for all $u \in C^1(\mathbb{R}^n)$, where

$$\gamma := 1 - n/p .$$

Proof. 1. First choose any ball $B(x,r) \subset \mathbb{R}^n$.

We claim there exists a constant C, depending only on n, such that

$$⨍_{B(x,r)} |u(y) - u(x)|\, dy \le C \int_{B(x,r)} \frac{|Du(y)|}{|y-x|^{n-1}}\, dy. \tag{21}$$

To prove this, fix any point $w \in \partial B(0,1)$. Then if $0 < s < r$,

$$\begin{aligned} |u(x+sw) - u(x)| &= \left| \int_0^s \frac{d}{dt} u(x+tw)\, dt \right| \\ &= \left| \int_0^s Du(x+tw) \cdot w\, dt \right| \\ &\leq \int_0^s |Du(x+tw)|\, dt. \end{aligned}$$

Hence

$$\begin{aligned} \int_{\partial B(0,1)} |u(x+sw) - u(x)|\, dS &\leq \int_0^s \int_{\partial B(0,1)} |Du(x+tw)|\, dSdt \\ &= \int_0^s \int_{\partial B(0,1)} |Du(x+tw)| \frac{t^{n-1}}{t^{n-1}}\, dSdt. \end{aligned}$$

Let $y = x + tw$, so that $t = |x-y|$. Then converting from polar coordinates, we have

$$\begin{aligned} \int_{\partial B(0,1)} |u(x+sw) - u(x)|\, dS &\leq \int_{B(x,s)} \frac{|Du(y)|}{|x-y|^{n-1}}\, dy \\ &\leq \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}}\, dy. \end{aligned}$$

Multiply by s^{n-1} and integrate from 0 to r with respect to s:

$$\int_{B(x,r)} |u(y) - u(x)|\, dy \leq \frac{r^n}{n} \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}}\, dy.$$

This is (21).

2. Now fix $x \in \mathbb{R}^n$. We apply inequality (21) as follows:

(22)

$$\begin{aligned} |u(x)| &\leq ⨍_{B(x,1)} |u(x) - u(y)|\, dy + ⨍_{B(x,1)} |u(y)|\, dy \\ &\leq C \int_{B(x,1)} \frac{|Du(y)|}{|x-y|^{n-1}}\, dy + C\|u\|_{L^p(B(x,1))} \\ &\leq C \left(\int_{\mathbb{R}^n} |Du|^p\, dy \right)^{1/p} \left(\int_{B(x,1)} \frac{dy}{|x-y|^{(n-1)\frac{p}{p-1}}} \right)^{\frac{p-1}{p}} + C\|u\|_{L^p(\mathbb{R}^n)} \\ &\leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}. \end{aligned}$$

The last estimate holds since $p > n$ implies $(n-1)\frac{p}{p-1} < n$; so that

$$\int_{B(x,1)} \frac{1}{|x-y|^{(n-1)\frac{p}{p-1}}}\, dy < \infty.$$

As $x \in \mathbb{R}^n$ is arbitrary, inequality (22) implies

$$\sup_{\mathbb{R}^n} |u| \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}. \tag{23}$$

3. Next, choose any two points $x, y \in \mathbb{R}^n$ and write $r := |x-y|$. Let $W := B(x,r) \cap B(y,r)$. Then

$$|u(x) - u(y)| \leq ⨍_W |u(x) - u(z)|\,dz + ⨍_W |u(y) - u(z)|\,dz. \tag{24}$$

But inequality (21) allows us to estimate

$$\begin{aligned} ⨍_W |u(x) - u(z)|\,dz &\leq C ⨍_{B(x,r)} |u(x) - u(z)|\,dz \\ &\leq C \left(\int_{B(x,r)} |Du|^p dz\right)^{1/p} \left(\int_{B(x,r)} \frac{dz}{|x-z|^{(n-1)\frac{p}{p-1}}}\right)^{\frac{p-1}{p}} \\ &\leq C\left(r^{n-(n-1)\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \|Du\|_{L^p(\mathbb{R}^n)} \\ &= Cr^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}. \end{aligned} \tag{25}$$

Likewise,

$$⨍_W |u(y) - u(z)|\,dz \leq Cr^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}.$$

Our substituting this estimate and (25) into (24) yields

$$|u(x) - u(y)| \leq Cr^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)} = C|x-y|^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}.$$

Thus

$$[u]_{C^{0,1-n/p}(\mathbb{R}^n)} = \sup_{x \neq y} \left\{ \frac{|u(x) - u(y)|}{|x-y|^{1-n/p}} \right\} \leq C\|Du\|_{L^p(\mathbb{R}^n)}.$$

This inequality and (23) complete the proof of (20). □

Remark. A slight variant of the proof above provides the estimate

$$|u(y) - u(x)| \leq Cr^{1-\frac{n}{p}} \left(\int_{B(x,2r)} |Du(z)|^p dz\right)^{1/p}$$

for all $u \in C^1(B(x,2r))$, $y \in B(x,r)$, $n < p < \infty$. By an approximation the same bound is valid for $u \in W^{1,p}(B(x,2r))$, $n < p < \infty$. We will use this inequality later in §5.8.2. (This estimate is in fact valid if on the right hand side we integrate over $B(x,r)$, instead of $B(x,2r)$, but the proof is a bit trickier.) □

DEFINITION. *We say u^* is a* version *of a given function u provided*

$$u = u^* \text{ a.e.}$$

THEOREM 5 (Estimates for $W^{1,p}, n < p \leq \infty$). *Let U be a bounded, open subset of $\mathbb{R}^n$, and suppose ∂U is C^1. Assume $n < p \leq \infty$, and $u \in W^{1,p}(U)$. Then u has a version $u^* \in C^{0,\gamma}(\bar{U})$, for $\gamma = 1 - \frac{n}{p}$, with the estimate*

$$\|u^*\|_{C^{0,\gamma}(\bar{U})} \leq C\|u\|_{W^{1,p}(U)}.$$

The constant C depends only on p, n and U.

Remark. In view of Theorem 5, *we will henceforth always identify a function $u \in W^{1,p}(U)$ $(p > n)$ with its continuous version.* □

Proof. Since ∂U is C^1, there exists according to Theorem 1 in §5.4 an extension $Eu = \bar{u} \in W^{1,p}(\mathbb{R}^n)$ such that

$$\text{(26)} \qquad \begin{cases} \bar{u} = u \text{ in } U, \\ \bar{u} \text{ has compact support, and} \\ \|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(U)}. \end{cases}$$

Since $\bar{u}$ has compact support, we obtain from Theorem 1 in §5.3 the existence of functions $u_m \in C_c^\infty(\mathbb{R}^n)$ such that

$$\text{(27)} \qquad u_m \to \bar{u} \quad \text{in } W^{1,p}(\mathbb{R}^n).$$

Now according to Theorem 4, $\|u_m - u_l\|_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C\|u_m - u_l\|_{W^{1,p}(\mathbb{R}^n)}$ for all $l, m \geq 1$; whence there exists a function $u^* \in C^{0,1-n/p}(\mathbb{R}^n)$ such that

$$\text{(28)} \qquad u_m \to u^* \quad \text{in } C^{0,1-n/p}(\mathbb{R}^n).$$

Owing to (27) and (28) we see that $u^* = u$ a.e. on U; so that u^* is a version of u. Since Theorem 4 also implies $\|u_m\|_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C\|u_m\|_{W^{1,p}(\mathbb{R}^n)}$, assertions (27) and (28) yield:

$$\|u^*\|_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C\|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)}.$$

This inequality and (26) complete the proof. □

5.6.3. General Sobolev inequalities.

We can now concatenate the estimates established in §§5.6.1 and 5.6.2 to obtain more complicated (and hard-to-remember) inequalities.

THEOREM 6 (General Sobolev inequalities). *Let U be a bounded open subset of $\mathbb{R}^n$, with a C^1 boundary. Assume $u \in W^{k,p}(U)$.*

(i) *If*

$$k < \frac{n}{p}, \tag{29}$$

then $u \in L^q(U)$, where

$$\frac{1}{q} = \frac{1}{p} - \frac{k}{n}.$$

We have in addition the estimate

$$\|u\|_{L^q(U)} \le C\|u\|_{W^{k,p}(U)}, \tag{30}$$

the constant C depending only on k, p, n and U.

(ii) *If*

$$k > \frac{n}{p}, \tag{31}$$

then $u \in C^{k-\left[\frac{n}{p}\right]-1,\gamma}(\bar{U})$, where

$$\gamma = \begin{cases} \left[\frac{n}{p}\right] + 1 - \frac{n}{p}, & \text{if } \frac{n}{p} \text{ is not an integer} \\ \text{any positive number} < 1, & \text{if } \frac{n}{p} \text{ is an integer.} \end{cases}$$

We have in addition the estimate

$$\|u\|_{C^{k-\left[\frac{n}{p}\right]-1,\gamma}(\bar{U})} \le C\|u\|_{W^{k,p}(U)}, \tag{32}$$

the constant C depending only on k, p, n, γ and U.

Proof. 1. Assume (29). Then since $D^\alpha u \in L^p(U)$ for all $|\alpha| = k$, the Sobolev–Nirenberg–Gagliardo inequality implies

$$\|D^\beta u\|_{L^{p^*}(U)} \le C\|u\|_{W^{k,p}(U)} \quad \text{if } |\beta| = k-1,$$

and so $u \in W^{k-1,p^*}(U)$. Similarly, we find $u \in W^{k-2,p^{**}}(U)$, where $\frac{1}{p^{**}} = \frac{1}{p^*} - \frac{1}{n} = \frac{1}{p} - \frac{2}{n}$. Continuing, we eventually discover after k steps that $u \in W^{0,q}(U) = L^q(U)$, for $\frac{1}{q} = \frac{1}{p} - \frac{k}{n}$. The stated estimate (30) follows from multiplying the relevant estimates at each stage of the above argument.

2. Assume now condition (31) holds, and $\frac{n}{p}$ is not an integer. Then as above we see

$$u \in W^{k-l,r}(U), \tag{33}$$

for

$$\frac{1}{r} = \frac{1}{p} - \frac{l}{n}, \tag{34}$$

provided $lp < n$. We choose the integer l so that

$$l < \frac{n}{p} < l + 1; \tag{35}$$

that is, we set $l = \left[\frac{n}{p}\right]$. Consequently (34) and (35) imply $r = \frac{pn}{n-pl} > n$. Hence (33) and Morrey's inequality imply that $D^\alpha u \in C^{0,1-\frac{n}{r}}(\bar{U})$ for all $|\alpha| \leq k - l - 1$. Observe also that $1 - \frac{n}{r} = 1 - \frac{n}{p} + l = \left[\frac{n}{p}\right] + 1 - \frac{n}{p}$. Thus $u \in C^{k-\left[\frac{n}{p}\right]-1,\left[\frac{n}{p}\right]+1-\frac{n}{p}}(\bar{U})$, and the stated estimate follows easily.

3. Finally, suppose (31) holds, with $\frac{n}{p}$ an integer. Set $l = \left[\frac{n}{p}\right] - 1 = \frac{n}{p} - 1$. Consequently, we have as above $u \in W^{k-l,r}(U)$ for $r = \frac{pn}{n-pl} = n$. Hence the Sobolev–Nirenberg–Gagliardo inequality shows $D^\alpha u \in L^q(U)$ for all $n \leq q < \infty$ and all $|\alpha| \leq k - l - 1 = k - \left[\frac{n}{p}\right]$. Therefore Morrey's inequality further implies $D^\alpha u \in C^{0,1-\frac{n}{q}}(\bar{U})$ for all $n < q < \infty$ and all $|\alpha| \leq k - \left[\frac{n}{p}\right] - 1$. Consequently $u \in C^{k-\left[\frac{n}{p}\right]-1,\gamma}(\bar{U})$ for each $0 < \gamma < 1$. As before, the stated estimate follows as well. □

Remark. Various general Sobolev-type inequalities can also be proved using the Fourier transform: see Problem 18. □

5.7. COMPACTNESS

We have seen in §5.6 that the Gagliardo–Nirenberg–Sobolev inequality implies the embedding of $W^{1,p}(U)$ into $L^{p^*}(U)$ for $1 \leq p < n$, $p^* = \frac{pn}{n-p}$. We will now demonstrate that $W^{1,p}(U)$ is in fact *compactly* embedded in $L^q(U)$ for $1 \leq q < p^*$. This compactness will be fundamental for our applications of linear and nonlinear functional analysis to PDE in Chapters 6–9.

DEFINITION. *Let X and Y be Banach spaces, $X \subset Y$. We say that X is* compactly embedded *in Y, written*

$$X \subset\subset Y,$$

provided

(i) $\|x\|_Y \leq C\|x\|_X$ $(x \in X)$ *for some constant C,* *and*

(ii) *each bounded sequence in X is precompact in Y.*

THEOREM 1 (Rellich-Kondrachov Compactness Theorem). *Assume U is a bounded open subset of $\mathbb{R}^n$, and ∂U is C^1. Suppose $1 \leq p < n$. Then*

$$W^{1,p}(U) \subset\subset L^q(U)$$

for each $1 \leq q < p^*$.

Proof. 1. Fix $1 \leq q < p^*$ and note that since U is bounded, Theorem 2 in §5.6.1 implies

$$W^{1,p}(U) \subset L^q(U), \ \|u\|_{L^q(U)} \leq C\|u\|_{W^{1,p}(U)}.$$

It remains therefore to show that if $\{u_m\}_{m=1}^\infty$ is a bounded sequence in $W^{1,p}(U)$, there exists a subsequence $\{u_{m_j}\}_{j=1}^\infty$ which converges in $L^q(U)$.

2. In view of the Extension Theorem from §5.4 we may with no loss of generality assume that $U = \mathbb{R}^n$ and the functions $\{u_m\}_{m=1}^\infty$ all have compact support in some bounded open set $V \subset \mathbb{R}^n$. We also may assume

$$\sup_m \|u_m\|_{W^{1,p}(V)} < \infty. \tag{1}$$

3. Let us first study the smoothed functions

$$u_m^\varepsilon := \eta_\varepsilon * u_m \quad (\varepsilon > 0, \ m = 1, 2, \dots),$$

η_ε denoting the usual mollifier. We may suppose the functions $\{u_m^\varepsilon\}_{m=1}^\infty$ all have support in V as well.

4. We first claim

$$u_m^\varepsilon \to u_m \text{ in } L^q(V) \text{ as } \varepsilon \to 0, \ \textit{uniformly} \text{ in } m. \tag{2}$$

To prove this, we first note that if u_m is smooth, then

$$\begin{aligned} u_m^\varepsilon(x) - u_m(x) &= \int_{B(0,1)} \eta(y)(u_m(x - \varepsilon y) - u_m(x))\, dy \\ &= \int_{B(0,1)} \eta(y) \int_0^1 \frac{d}{dt}(u_m(x - \varepsilon t y))\, dt dy \\ &= -\varepsilon \int_{B(0,1)} \eta(y) \int_0^1 Du_m(x - \varepsilon t y) \cdot y \, dt dy. \end{aligned}$$

Thus

$$\int_V |u_m^\varepsilon(x) - u_m(x)|\, dx \le \varepsilon \int_{B(0,1)} \eta(y) \int_0^1 \int_V |Du_m(x - \varepsilon t y)|\, dx dt dy$$
$$\le \varepsilon \int_V |Du_m(z)| dz.$$

By approximation this estimate holds if $u_m \in W^{1,p}(V)$. Hence

$$\|u_m^\varepsilon - u_m\|_{L^1(V)} \le \varepsilon \|Du_m\|_{L^1(V)} \le \varepsilon C \|Du_m\|_{L^p(V)},$$

the latter inequality holding since V is bounded. Owing to (1) we thereby discover

$$\text{(3)} \qquad u_m^\varepsilon \to u_m \quad \text{in } L^1(V), \text{ uniformly in } m.$$

But then since $1 \le q < p^*$, we see using the interpolation inequality for L^p-norms (§B.2) that

$$\|u_m^\varepsilon - u_m\|_{L^q(V)} \le \|u_m^\varepsilon - u_m\|_{L^1(V)}^\theta \|u_m^\varepsilon - u_m\|_{L^{p^*}(V)}^{1-\theta},$$

where $\frac{1}{q} = \theta + \frac{(1-\theta)}{p^*}$, $0 < \theta < 1$. Consequently (1) and the Gagliardo–Nirenberg–Sobolev inequality imply

$$\|u_m^\varepsilon - u_m\|_{L^q(V)} \le C \|u_m^\varepsilon - u_m\|_{L^1(V)}^\theta;$$

whence assertion (2) follows from (3).

5. Next we claim

$$\text{(4)} \qquad \begin{cases} \text{for each fixed } \varepsilon > 0, \text{ the sequence } \{u_m^\varepsilon\}_{m=1}^\infty \\ \text{is uniformly bounded and equicontinuous.} \end{cases}$$

Indeed, if $x \in \mathbb{R}^n$, then

$$|u_m^\varepsilon(x)| \le \int_{B(x,\varepsilon)} \eta_\varepsilon(x - y) |u_m(y)|\, dy$$
$$\le \|\eta_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \|u_m\|_{L^1(V)} \le \frac{C}{\varepsilon^n} < \infty$$

for $m = 1, 2, \dots$. Similarly

$$|Du_m^\varepsilon(x)| \le \int_{B(x,\varepsilon)} |D\eta_\varepsilon(x - y)| |u_m(y)|\, dy$$
$$\le \|D\eta_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \|u_m\|_{L^1(V)} \le \frac{C}{\varepsilon^{n+1}} < \infty,$$

for $m = 1, \dots$. Assertion (4) follows from these two estimates.

6. Now fix $\delta > 0$. We will show there exists a subsequence $\{u_{m_j}\}_{j=1}^\infty \subset \{u_m\}_{m=1}^\infty$ such that

$$\limsup_{j,k\to\infty} \|u_{m_j} - u_{m_k}\|_{L^q(V)} \le \delta. \tag{5}$$

To see this, let us first employ assertion (2) to select $\varepsilon > 0$ so small that

$$\|u_m^\varepsilon - u_m\|_{L^q(V)} \le \delta/2 \tag{6}$$

for $m = 1, 2, \dots$.

We now observe that since the functions $\{u_m\}_{m=1}^\infty$, and thus the functions $\{u_m^\varepsilon\}_{m=1}^\infty$, have support in some fixed bounded set $V \subset \mathbb{R}^n$, we may utilize (4) and the Arzela-Ascoli compactness criterion, §C.7, to obtain a subsequence $\{u_{m_j}^\varepsilon\}_{j=1}^\infty \subset \{u_m^\varepsilon\}_{m=1}^\infty$ which converges *uniformly* on V. In particular therefore

$$\limsup_{j,k\to\infty} \|u_{m_j}^\varepsilon - u_{m_k}^\varepsilon\|_{L^q(V)} = 0. \tag{7}$$

But then (6) and (7) imply

$$\limsup_{j,k\to\infty} \|u_{m_j} - u_{m_k}\|_{L^q(V)} \le \delta,$$

and so (5) is proved.

7. We next employ assertion (5) with $\delta = 1, \frac{1}{2}, \frac{1}{3}, \dots$ and use a standard diagonal argument to extract a subsequence $\{u_{m_l}\}_{l=1}^\infty \subset \{u_m\}_{m=1}^\infty$ satisfying

$$\limsup_{l,k\to\infty} \|u_{m_l} - u_{m_k}\|_{L^q(V)} = 0.$$

□

Remark. Observe that since $p^* > p$ and $p^* \to \infty$ as $p \to n$, we have in particular

$$W^{1,p}(U) \subset\subset L^p(U)$$

for all $1 \le p \le \infty$. (Observe that if $n < p \le \infty$, this follows from Morrey's inequality and the Arzela-Ascoli compactness criterion.) Note also that

$$W_0^{1,p}(U) \subset\subset L^p(U),$$

even if we do not assume ∂U to be C^1. □

5.8. ADDITIONAL TOPICS

5.8.1. Poincaré's inequalities.

We now illustrate how the compactness assertion in §5.7 can be used to generate new inequalities.

Notation. $(u)_U = ⨍_U u\,dy$ = average of u over U. □

THEOREM 1 (Poincaré's inequality). *Let U be a bounded, connected, open subset of $\mathbb{R}^n$, with a C^1 boundary ∂U. Assume $1 \le p \le \infty$. Then there exists a constant C, depending only on n, p and U, such that*

$$\|u - (u)_U\|_{L^p(U)} \le C\|Du\|_{L^p(U)} \tag{1}$$

for each function $u \in W^{1,p}(U)$.

The significance of (1) is that only the gradient of u appears on the right hand side.

Proof. We argue by contradiction. Were the stated estimate false, there would exist for each integer $k = 1, \dots$ a function $u_k \in W^{1,p}(U)$ satisfying

$$\|u_k - (u_k)_U\|_{L^p(U)} > k\|Du_k\|_{L^p(U)}. \tag{2}$$

We renormalize by defining

$$v_k := \frac{u_k - (u_k)_U}{\|u_k - (u_k)_U\|_{L^p(U)}} \quad (k = 1, \dots). \tag{3}$$

Then

$$(v_k)_U = 0, \ \|v_k\|_{L^p(U)} = 1;$$

and (2) implies

$$\|Dv_k\|_{L^p(U)} < \frac{1}{k} \quad (k = 1, 2, \dots). \tag{4}$$

In particular the functions $\{v_k\}_{k=1}^\infty$ are bounded in $W^{1,p}(U)$.

In view of the Remark after the Rellich–Kondrachov Theorem in §5.7, there exists a subsequence $\{v_{k_j}\}_{j=1}^\infty \subset \{v_k\}_{k=1}^\infty$ and a function $v \in L^p(U)$ such that

$$v_{k_j} \to v \quad \text{in } L^p(U). \tag{5}$$

From (3) it follows that

$$(v)_U = 0, \ \|v\|_{L^p(U)} = 1. \tag{6}$$

On the other hand, (4) implies for each $i = 1, \ldots, n$ and $\phi \in C_c^\infty(U)$ that

$$\int_U v\phi_{x_i}\,dx = \lim_{k_j\to\infty}\int_U v_{k_j}\phi_{x_i}\,dx = -\lim_{k_j\to\infty}\int_U v_{k_j,x_i}\phi\,dx = 0.$$

Consequently $v \in W^{1,p}(U)$, with $Dv = 0$ a.e. Thus v is constant, since U is connected (see Problem 10). However this conclusion is at variance with (6): since v is constant and $(v)_U = 0$, we must have $v \equiv 0$; in which case $\|v\|_{L^p(U)} = 0$. This contradiction establishes estimate (1). □

A particularly important special case follows.

Notation. $(u)_{x,r} = ⨍_{B(x,r)} u\,dy$ = average of u over the ball $B(x,r)$. □

THEOREM 2 (Poincaré's inequality for a ball). *Assume* $1 \le p \le \infty$. *Then there exists a constant C, depending only on n and p, such that*

$$\|u - (u)_{x,r}\|_{L^p(B(x,r))} \le Cr\|Du\|_{L^p(B(x,r))} \tag{7}$$

for each ball $B(x,r) \subset \mathbb{R}^n$ and each function $u \in W^{1,p}(B^0(x,r))$.

Proof. 1. The case $U = B^0(0,1)$ follows from Theorem 1. In general, if $u \in W^{1,p}(B^0(x,r))$ write

$$v(y) := u(x + ry) \quad (y \in B(0,1)).$$

Then $v \in W^{1,p}(B^0(0,1))$, and we have

$$\|v - (v)_{0,1}\|_{L^p(B(0,1))} \le C\|Dv\|_{L^p(B(0,1))}.$$

Changing variables, we recover estimate (7). □

Remark. Assume $u \in W^{1,n}(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, and let $B(x,r)$ be any ball. Then Theorem 2 with $p = 1$ implies

$$\begin{aligned}⨍_{B(x,r)} |u - (u)_{x,r}|\,dy &\le Cr ⨍_{B(x,r)} |Du|\,dy\\ &\le Cr\left(⨍_{B(x,r)} |Du|^n\,dy\right)^{1/n} \le C\left(\int_{\mathbb{R}^n} |Du|^n\,dy\right)^{1/n}.\end{aligned}$$

Thus $u \in BMO(\mathbb{R}^n)$, the space of functions of *bounded mean oscillation* in $\mathbb{R}^n$, with the seminorm

$$[u]_{BMO(\mathbb{R}^n)} := \sup_{B(x,r)\subset\mathbb{R}^n}\left\{⨍_{B(x,r)} |u - (u)_{x,r}|\,dy\right\}.$$

See Stein [**SE,** Chapter IV] for the theory of BMO. □

5.8.2. Difference quotients.

When we later apply Sobolev space theory to PDE, we will be forced to study difference quotient approximations to weak derivatives. Following is the relevant theory, which the reader may wish to postpone studying until the need arises, in §6.3.

a. Difference quotients and $\mathbf{W^{1,p}}$.

Assume $u : U \to \mathbb{R}$ is a locally summable function, and $V \subset\subset U$.

DEFINITIONS.

(i) *The* i^{th}-difference quotient *of size* h *is*

$$D_i^h u(x) = \frac{u(x + he_i) - u(x)}{h} \quad (i = 1, \dots, n)$$

for $x \in V$ *and* $h \in \mathbb{R}$, $0 < |h| < \operatorname{dist}(V, \partial U)$.

(ii) $D^h u := (D_1^h u, \dots, D_n^h u)$.

THEOREM 3 (Difference quotients and weak derivatives).

(i) *Suppose* $1 \le p < \infty$ *and* $u \in W^{1,p}(U)$. *Then for each* $V \subset\subset U$

$$\|D^h u\|_{L^p(V)} \le C\|Du\|_{L^p(U)} \tag{8}$$

for some constant C *and all* $0 < |h| < \frac{1}{2}\operatorname{dist}(V, \partial U)$.

(ii) *Assume* $1 < p < \infty$, $u \in L^p(V)$, *and there exists a constant* C *such that*

$$\|D^h u\|_{L^p(V)} \le C \tag{9}$$

for all $0 < |h| < \frac{1}{2}\operatorname{dist}(V, \partial U)$. *Then*

$$u \in W^{1,p}(V), \quad \textit{with} \quad \|Du\|_{L^p(V)} \le C.$$

Remark. Assertion (ii) is false for $p = 1$ (Problem 11). □

Proof. 1. Assume $1 \le p < \infty$, and temporarily suppose u is smooth. Then for each $x \in V$, $i = 1, \dots, n$, and $0 < |h| < \frac{1}{2}\operatorname{dist}(V, \partial U)$, we have

$$u(x + he_i) - u(x) = \int_0^1 u_{x_i}(x + the_i)\, dt \cdot he_i;$$

so that

$$|u(x + he_i) - u(x)| \le h \int_0^1 |Du(x + the_i)|\, dt.$$

Consequently

$$\begin{aligned}\int_V |D^h u|^p \, dx &\leq C \sum_{i=1}^n \int_V \int_0^1 |Du(x + the_i)|^p \, dtdx \\ &= C \sum_{i=1}^n \int_0^1 \int_V |Du(x + the_i)|^p \, dxdt.\end{aligned}$$

Thus

$$\int_V |D^h u|^p \, dx \leq C \int_U |Du|^p \, dx.$$

This estimate holds should u be smooth, and thus is valid by approximation for arbitrary $u \in W^{1,p}(U)$.

2. Now suppose estimate (9) holds for all $0 < |h| < \frac{1}{2}\operatorname{dist}(V, \partial U)$ and some constant C. Choose $i = 1, \dots, n$, $\phi \in C_c^\infty(V)$, and note for small enough h that

$$\int_V u(x) \left[\frac{\phi(x + he_i) - \phi(x)}{h} \right] dx = - \int_V \left[\frac{u(x) - u(x - he_i)}{h} \right] \phi(x) \, dx;$$

that is,

$$\int_V u(D_i^h \phi) \, dx = - \int_V (D_i^{-h} u) \phi \, dx. \tag{10}$$

This is the "integration-by-parts" formula for difference quotients. Estimate (9) implies

$$\sup_h \|D_i^{-h} u\|_{L^p(V)} < \infty;$$

and therefore, since $1 < p < \infty$, there exists a function $v_i \in L^p(V)$ and a subsequence $h_k \to 0$ such that

$$D_i^{-h_k} u \rightharpoonup v_i \quad \text{weakly in } L^p(V). \tag{11}$$

(See §D.4 for weak convergence.) But then

$$\begin{aligned}\int_V u\phi_{x_i} \, dx = \int_U u\phi_{x_i} \, dx &= \lim_{h_k \to 0} \int_U u D_i^{h_k} \phi \, dx \\ &= - \lim_{h_k \to 0} \int_V D_i^{-h_k} u \, \phi \, dx \\ &= - \int_V v_i \phi \, dx = - \int_U v_i \phi \, dx.\end{aligned}$$

Thus $v_i = u_{x_i}$ in the weak sense $(i = 1, \dots, n)$, and so $Du \in L^p(V)$. As $u \in L^p(V)$, we deduce therefore that $u \in W^{1,p}(V)$. □

Remark. Variants of Theorem 3 can be valid even if it is not true that $V \subset\subset U$. For example if U is the open half-ball $B^0(0,1) \cap \{x_n > 0\}$, $V = B^0(0,1/2) \cap \{x_n > 0\}$, we have the bound $\int_V |D_i^h u|^p\,dx \le \int_U |u_{x_i}|^p\,dx$ for $i = 1, \dots, n-1$. The proof is similar to that just given.

We will need this comment in Chapter 6, §6.3.2. □

b. Lipschitz functions and $\mathbf{W^{1,\infty}}$.

THEOREM 4 (Characterization of $W^{1,\infty}$). *Let U be open and bounded, with ∂U of class C^1. Then $u : U \to \mathbb{R}$ is Lipschitz continuous if and only if $u \in W^{1,\infty}(U)$.*

Proof. 1. First assume $U = \mathbb{R}^n$ and u has compact support.

Suppose $u \in W^{1,\infty}(\mathbb{R}^n)$. Then $u^\varepsilon := \eta_\varepsilon * u$, where η_ε is the usual mollifier, is smooth and satisfies

$$\begin{cases} u^\varepsilon \to u \text{ uniformly as } \varepsilon \to 0, \\ \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n)} \le \|Du\|_{L^\infty(\mathbb{R}^n)}. \end{cases}$$

Choose any two points $x, y \in \mathbb{R}^n$, $x \ne y$. We have

$$\begin{aligned} u^\varepsilon(x) - u^\varepsilon(y) &= \int_0^1 \frac{d}{dt} u^\varepsilon(tx + (1-t)y)\,dt \\ &= \int_0^1 Du^\varepsilon(tx + (1-t)y)\,dt \cdot (x - y); \end{aligned}$$

and so

$$|u^\varepsilon(x) - u^\varepsilon(y)| \le \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n)} |x - y| \le \|Du\|_{L^\infty(\mathbb{R}^n)} |x - y|.$$

We let $\varepsilon \to 0$ to discover

$$|u(x) - u(y)| \le \|Du\|_{L^\infty(\mathbb{R}^n)} |x - y|.$$

Hence u is Lipschitz continuous.

2. On the other hand assume now u is Lipschitz continuous; we must prove that u has essentially bounded weak first derivative. Since u is Lipschitz, we see

$$\|D_i^{-h} u\|_{L^\infty(\mathbb{R}^n)} \le \operatorname{Lip}(u),$$

and thus there exists a function $v_i \in L^\infty(\mathbb{R}^n)$ and a subsequence $h_k \to 0$ such that

$$D_i^{-h_k} u \rightharpoonup v_i \quad \text{weakly in } L^2_{\text{loc}}(\mathbb{R}^n). \tag{12}$$

Consequently

$$\int_{\mathbb{R}^n} u\phi_{x_i}\,dx = \lim_{h_k\to 0}\int_{\mathbb{R}^n} uD_i^{h_k}\phi\,dx$$
$$= -\lim_{h_k\to 0}\int_{\mathbb{R}^n} D_i^{-h_k}u\phi\,dx = -\int_{\mathbb{R}^n} v_i\phi\,dx$$

by (12). The above equality holds for all $\phi \in C_c^\infty(\mathbb{R}^n)$, and so $v_i = u_{x_i}$ in the weak sense $(i = 1,\dots,n)$. Consequently $u \in W^{1,\infty}(\mathbb{R}^n)$.

3. In the general case that U is bounded, with ∂U of class C^1, we as usual extend u to $Eu = \bar{u}$ and apply the above argument. □

Remark. The argument above adapts easily to prove that for any open set U, $u \in W^{1,\infty}_{\text{loc}}(U)$ if and only if u is locally Lipschitz continuous in U. There is no corresponding characterization of the spaces $W^{1,p}$ for $1 \leq p < \infty$. If $n < p < \infty$, then each function $u \in W^{1,p}$ belongs to $C^{0,1-n/p}$, but on the other hand a function Hölder continuous with exponent less than one need not belong to any Sobolev space $W^{1,p}$. □

5.8.3. Differentiability a.e.

Next we examine more closely the connections between weak partial derivatives and partial derivatives in the usual calculus sense.

DEFINITION. *A function $u : U \to \mathbb{R}$ is* differentiable *at $x \in U$ if there exists $a \in \mathbb{R}^n$ such that*

$$u(y) = u(x) + a\cdot(y-x) + o(|y-x|) \quad \textit{as } y \to x. \tag{13}$$

In other words,

$$\lim_{y\to x}\frac{|u(y)-u(x)-a\cdot(y-x)|}{|y-x|} = 0.$$

Notation. It is easy to check that a, if it exists, is unique. We henceforth write

$$Du(x)$$

for a and call Du the gradient of u. □

To be sure that this notation is consistent, we need to study the relationships between the various notions of derivatives:

THEOREM 5 (Differentiability almost everywhere). *Assume $u \in W^{1,p}_{\text{loc}}(U)$ for some $n < p \leq \infty$. Then u is differentiable* a.e. *in U, and its gradient equals its weak gradient* a.e.

Recall we always identify u with its continuous version.

Proof. 1. Assume first $n < p < \infty$. From the Remark after the proof of Theorem 4 in §5.6.2, we recall Morrey's estimate

$$|v(y) - v(x)| \leq Cr^{1-\frac{n}{p}} \left(\int_{B(x,2r)} |Dv(z)|^p dz \right)^{1/p} \quad (y \in B(x,r)), \tag{14}$$

valid for any C^1 function v and thus, by approximation, for any $v \in W^{1,p}$.

2. Choose $u \in W^{1,p}_{\text{loc}}(U)$. Now for a.e. $x \in U$, a version of Lebesgue's Differentiation Theorem (§E.4) implies

$$⨍_{B(x,r)} |Du(x) - Du(z)|^p dz \to 0 \tag{15}$$

as $r \to 0$, Du denoting as usual the weak derivative of u. Fix any such point x and set

$$v(y) := u(y) - u(x) - Du(x) \cdot (y - x)$$

in estimate (14), where

$$r = |x - y|. \tag{16}$$

We find

$$\begin{aligned} |u(y) &- u(x) - Du(x) \cdot (y - x)| \\ &\leq Cr^{1-n/p} \left(\int_{B(x,2r)} |Du(x) - Du(z)|^p dz \right)^{1/p} \\ &\leq Cr \left(⨍_{B(x,2r)} |Du(x) - Du(z)|^p dz \right)^{1/p} \\ &= o(r) \quad \text{by (15)} \\ &= o(|x - y|) \quad \text{by (16)}. \end{aligned}$$

Thus u is differentiable at x, and its gradient equals its weak gradient at x.

3. In case $p = \infty$, we note $W^{1,\infty}_{\text{loc}}(U) \subset W^{1,p}_{\text{loc}}(U)$ for all $1 \leq p < \infty$, and apply the reasoning above. □

Finally, in view of Theorem 5, we obtain

THEOREM 6 (Rademacher's Theorem). *Let u be locally Lipschitz continuous in U. Then u is differentiable almost everywhere in U.*

5.8.4. Fourier transform methods.

Next we employ the Fourier transform (§4.3) to give an alternate characterization of the spaces $H^k(\mathbb{R}^n)$. For this section all functions are complex-valued.

THEOREM 7 (Characterization of H^k by Fourier transform).
Let k be a nonnegative integer.

(i) *A function $u \in L^2(\mathbb{R}^n)$ belongs to $H^k(\mathbb{R}^n)$ if and only if*

$$(1 + |y|^k)\hat{u} \in L^2(\mathbb{R}^n). \tag{17}$$

(ii) *In addition, there exists a positive constant C such that*

$$\frac{1}{C}\|u\|_{H^k(\mathbb{R}^n)} \leq \|(1 + |y|^k)\hat{u}\|_{L^2(\mathbb{R}^n)} \leq C\|u\|_{H^k(\mathbb{R}^n)} \tag{18}$$

for each $u \in H^k(\mathbb{R}^n)$.

Proof. 1. Assume first $u \in H^k(\mathbb{R}^n)$. Then for each multiindex $|\alpha| \leq k$, we have $D^\alpha u \in L^2(\mathbb{R}^n)$. Now if $u \in C^k$ has compact support, we have

$$\widehat{D^\alpha u} = (iy)^\alpha \hat{u} \tag{19}$$

according to Theorem 2 in §4.3.1. Approximating by smooth functions we deduce formula (19) provided $u \in H^k(\mathbb{R}^n)$. Thus $(iy)^\alpha \hat{u} \in L^2(\mathbb{R}^n)$ for each $|\alpha| \leq k$. In particular choosing $\alpha = (k, 0, \ldots, 0), (0, k, \ldots, 0), \ldots, (0, \ldots, k)$, we deduce

$$\int_{\mathbb{R}^n} |y|^{2k}|\hat{u}|^2 dy \leq C \int_{\mathbb{R}^n} |D^k u|^2 dx < \infty.$$

Thus

$$\int_{\mathbb{R}^n} (1 + |y|^k)^2|\hat{u}|^2 dy \leq C\|u\|_{H^k(\mathbb{R}^n)},$$

and so $(1 + |y|^k)\hat{u} \in L^2(\mathbb{R}^n)$.

2. Suppose conversely $(1 + |y|^k)\hat{u} \in L^2(\mathbb{R}^n)$ and $|\alpha| \leq k$. Then

$$\|(iy)^\alpha \hat{u}\|_{L^2(\mathbb{R}^n)} \leq \int_{\mathbb{R}^n} |y|^{2|\alpha|}|\hat{u}|^2 dy \leq C\|(1 + |y|^k)\hat{u}\|_{L^2(\mathbb{R}^n)}. \tag{20}$$

Set

$$u_\alpha := ((iy)^\alpha \hat{u})^\vee.$$

Then for each $\phi \in C_c^\infty(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} (D^\alpha \phi)\bar{u}\, dx = \int_{\mathbb{R}^n} (\widehat{D^\alpha \phi})\bar{\hat{u}}\, dy = \int_{\mathbb{R}^n} (iy)^\alpha \hat{\phi}\bar{\hat{u}}\, dy$$
$$= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \phi \bar{u}_\alpha\, dx.$$

Thus $u_\alpha = D^\alpha u$ in the weak sense and, by (20), $D^\alpha u \in L^2(\mathbb{R}^n)$. Hence $u \in H^k(U)$, as required. □

It is sometimes useful to define also *fractional* Sobolev spaces.

DEFINITION. *Assume* $0 < s < \infty$ *and* $u \in L^2(\mathbb{R}^n)$. *Then* $u \in H^s(\mathbb{R}^n)$ *if* $(1 + |y|^s)\hat{u} \in L^2(\mathbb{R}^n)$. *For noninteger* s, *we set*

$$\|u\|_{H^s(\mathbb{R}^n)} := \|(1 + |y|^s)\hat{u}\|_{L^2(\mathbb{R}^n)}.$$

5.9. OTHER SPACES OF FUNCTIONS

5.9.1. The space $\mathbf{H}^{-1}$.

As we will see later in our systematic study in Chapters 6 and 7 of linear elliptic, parabolic and hyperbolic PDE, it is important to have an explicit characterization of the dual space of H_0^1. (See Appendix D for definitions.)

DEFINITION. *We denote by* $H^{-1}(U)$ *the* dual space *to* $H_0^1(U)$.

In other words f belongs to $H^{-1}(U)$ provided f is a bounded linear functional on $H_0^1(U)$.

Notation. We will write $\langle\ ,\ \rangle$ to denote the pairing between $H^{-1}(U)$ and $H_0^1(U)$. □

DEFINITION. *If* $f \in H^{-1}(U)$, *we define the* norm

$$\|f\|_{H^{-1}(U)} = \sup\left\{\langle f, u\rangle \mid u \in H_0^1(U), \|u\|_{H_0^1(U)} \leq 1\right\}.$$

THEOREM 1 (Characterization of H^{-1}).

(i) *Assume* $f \in H^{-1}(U)$. *Then there exist functions* $f^0, f^1, \dots, f^n$ *in* $L^2(U)$ *such that*

$$\langle f, v\rangle = \int_U f^0 v + \sum_{i=1}^n f^i v_{x_i}\, dx \quad (v \in H_0^1(U)). \tag{1}$$

(ii) *Furthermore,*

$$\|f\|_{H^{-1}(U)} = \inf\left\{\left(\int_U \sum_{i=0}^n |f^i|^2\, dx\right)^{1/2} \;\middle|\; f \text{ satisfies (1) for } f^0, \dots, f^n \in L^2(U)\right\}.$$

Notation. We write $f = f^0 - \sum_{i=1}^n f^i_{x_i}$ whenever (1) holds. □

Proof. 1. Given $u, v \in H_0^1(U)$, we define the inner product $(u,v) := \int_U Du \cdot Dv + uv\,dx$. Let $f \in H^{-1}(U)$. We apply the Riesz Representation Theorem (§D.3) to deduce the existence of a unique function $u \in H_0^1(U)$ satisfying $B[u,v] = \langle f, v\rangle$ for all $v \in H_0^1(U)$; that is,

$$\int_U Du \cdot Dv + uv\,dx = \langle f, v\rangle \tag{2}$$

for each $v \in H_0^1(U)$. This establishes (1) for

$$\begin{cases} f^0 = u \\ f^i = u_{x_i} \quad (i = 1, \dots, n). \end{cases} \tag{3}$$

2. Assume now $f \in H^{-1}(U)$,

$$\langle f, v\rangle = \int_U g^0 v + \sum_{i=1}^n g^i v_{x_i}\,dx \tag{4}$$

for $g^0, g^1, \dots, g^n \in L^2(U)$. Setting $v = u$ in (2) and using (4), we deduce

$$\int_U |Du|^2 + |u|^2 dx \leq \int_U \sum_{i=0}^n |g^i|^2 dx.$$

Thus (3) implies

$$\int_U \sum_{i=0}^n |f^i|^2 dx \leq \int_U \sum_{i=0}^n |g^i|^2 dx. \tag{5}$$

3. From (1) it follows that

$$|\langle f, v\rangle| \leq \left(\int_U \sum_{i=0}^n |f^i|^2 dx \right)^{1/2}$$

if $\|v\|_{H_0^1(U)} \leq 1$. Consequently

$$\|f\|_{H^{-1}(U)} \leq \left(\int_U \sum_{i=0}^n |f^i|^2 dx \right)^{1/2}.$$

Setting $v = \frac{u}{\|u\|_{H_0^1(U)}}$ in (2), we deduce that in fact

$$\|f\|_{H^{-1}(U)} = \left(\int_U \sum_{i=0}^n |f^i|^2 dx \right)^{1/2}. \tag{6}$$

Assertion (ii) follows now from (4)–(6). □

5.9.2. Spaces involving time.

We study next some other sorts of Sobolev spaces, these comprising functions mapping time into Banach spaces. These will prove essential in our constructions of weak solutions to linear parabolic and hyperbolic PDE in Chapter 7 and to nonlinear parabolic PDE in Chapter 9.

Let X denote a real Banach space, with norm $\| \ \|$. The reader should first of all read §E.5 about measure and integration theory for mappings taking values in X.

DEFINITION. *The space*

$$L^p(0,T;X)$$

consists of all measurable functions $\mathbf{u} : [0,T] \to X$ *with*

$$\text{(i)} \qquad \|\mathbf{u}\|_{L^p(0,T;X)} := \left(\int_0^T \|\mathbf{u}(t)\|^p dt \right)^{1/p} < \infty$$

for $1 \leq p < \infty$, *and*

$$\text{(ii)} \qquad \|\mathbf{u}\|_{L^\infty(0,T;X)} := \operatorname*{ess\,sup}_{0\leq t\leq T} \|\mathbf{u}(t)\| < \infty.$$

DEFINITION. *The space*

$$C([0,T];X)$$

comprises all continuous functions $\mathbf{u} : [0,T] \to X$ *with*

$$\|\mathbf{u}\|_{C([0,T];X)} := \max_{0\leq t\leq T} \|\mathbf{u}(t)\| < \infty.$$

DEFINITION. *Let* $\mathbf{u} \in L^1(0,T;X)$. *We say* $\mathbf{v} \in L^1(0,T;X)$ *is the* weak derivative *of* $\mathbf{u}$, *written*

$$\mathbf{u}' = \mathbf{v},$$

provided

$$\int_0^T \phi'(t)\mathbf{u}(t)\,dt = -\int_0^T \phi(t)\mathbf{v}(t)\,dt$$

for all scalar test functions $\phi \in C_c^\infty(0,T)$.

DEFINITIONS. (i) *The Sobolev space*

$$W^{1,p}(0,T;X)$$

consists of all functions $\mathbf{u} \in L^p(0,T;X)$ *such that* $\mathbf{u}'$ *exists in the weak sense and belongs to* $L^p(0,T;X)$. *Furthermore,*

$$\|\mathbf{u}\|_{W^{1,p}(0,T;X)} := \begin{cases} \left(\int_0^T \|\mathbf{u}(t)\|^p + \|\mathbf{u}'(t)\|^p \, dt\right)^{1/p} & (1 \le p < \infty) \\ \operatorname*{ess\,sup}_{0\le t\le T}(\|\mathbf{u}(t)\| + \|\mathbf{u}'(t)\|) & (p = \infty). \end{cases}$$

(ii) *We write* $H^1(0,T;X) = W^{1,2}(0,T;X)$.

THEOREM 2 (Calculus in an abstract space). *Let* $\mathbf{u} \in W^{1,p}(0,T;X)$ *for some* $1 \le p \le \infty$. *Then*

(i) $\mathbf{u} \in C([0,T];X)$ *(after possibly being redefined on a set of measure zero), and*

(ii) $\mathbf{u}(t) = \mathbf{u}(s) + \int_s^t \mathbf{u}'(\tau)\, d\tau$ *for all* $0 \le s \le t \le T$.

(iii) *Furthermore, we have the estimate*

$$\max_{0\le t\le T} \|\mathbf{u}(t)\| \le C\|\mathbf{u}\|_{W^{1,p}(0,T;X)}, \tag{7}$$

the constant C *depending only on* T.

Proof. 1. Extend $\mathbf{u}$ to be $\mathbf{0}$ on $(-\infty,0)$ and (T,∞), and then set $\mathbf{u}^\varepsilon = \eta_\varepsilon * \mathbf{u}$, η_ε denoting the usual mollifier on $\mathbb{R}^1$. We check as in the proof of Theorem 1 in §5.3.1 that $\mathbf{u}^{\varepsilon\prime} = \eta_\varepsilon * \mathbf{u}'$ on $(\varepsilon, T-\varepsilon)$.

Then as $\varepsilon \to 0$,

$$\begin{cases} \mathbf{u}^\varepsilon \to \mathbf{u} & \text{in } L^p(0,T;X), \\ (\mathbf{u}^\varepsilon)' \to \mathbf{u}' & \text{in } L^p(0,T;X). \end{cases} \tag{8}$$

Fixing $0 < s < t < T$, we compute

$$\mathbf{u}^\varepsilon(t) = \mathbf{u}^\varepsilon(s) + \int_s^t \mathbf{u}^{\varepsilon\prime}(\tau)\, d\tau.$$

Thus

$$\mathbf{u}(t) = \mathbf{u}(s) + \int_s^t \mathbf{u}'(\tau)\, d\tau \tag{9}$$

for a.e. $0 < s < t < T$, according to (8). As the mapping $t \mapsto \int_0^t \mathbf{u}'(\tau)\, d\tau$ is continuous, assertions (i), (ii) follow.

2. Estimate (7) follows easily from (9). □

The next two propositions concern what happens when $\mathbf{u}$ and $\mathbf{u}'$ lie in different spaces.

THEOREM 3 (More calculus). *Suppose* $\mathbf{u} \in L^2(0,T;H_0^1(U))$, *with* $\mathbf{u}' \in L^2(0,T;H^{-1}(U))$.

(i) *Then*

$$\mathbf{u} \in C([0,T];L^2(U))$$

(after possibly being redefined on a set of measure zero).

(ii) *The mapping*

$$t \mapsto \|\mathbf{u}(t)\|^2_{L^2(U)}$$

is absolutely continuous, with

$$\frac{d}{dt}\|\mathbf{u}(t)\|^2_{L^2(U)} = 2\langle \mathbf{u}'(t), \mathbf{u}(t)\rangle$$

for a.e. $0 \le t \le T$.

(iii) *Furthermore, we have the estimate*

$$\max_{0\le t\le T}\|\mathbf{u}(t)\|_{L^2(U)} \le C\big(\|\mathbf{u}\|_{L^2(0,T;H_0^1(U))} + \|\mathbf{u}'\|_{L^2(0,T;H^{-1}(U))}\big), \tag{10}$$

the constant C *depending only on* T.

Proof. 1. Extend $\mathbf{u}$ to the larger interval $[-\sigma, T+\sigma]$ for $\sigma > 0$, and define the regularizations $\mathbf{u}^\varepsilon = \eta_\varepsilon * \mathbf{u}$, as in the earlier proof. Then for $\varepsilon, \delta > 0$,

$$\frac{d}{dt}\|\mathbf{u}^\varepsilon(t) - \mathbf{u}^\delta(t)\|^2_{L^2(U)} = 2(\mathbf{u}^{\varepsilon'}(t) - \mathbf{u}^{\delta'}(t), \mathbf{u}^\varepsilon(t) - \mathbf{u}^\delta(t))_{L^2(U)}.$$

Thus

$$\begin{aligned}\|\mathbf{u}^\varepsilon(t) - \mathbf{u}^\delta(t)\|^2_{L^2(U)} &= \|\mathbf{u}^\varepsilon(s) - \mathbf{u}^\delta(s)\|^2_{L^2(U)} \\ &\quad + 2\int_s^t \langle \mathbf{u}^{\varepsilon'}(\tau) - \mathbf{u}^{\delta'}(\tau), \mathbf{u}^\varepsilon(\tau) - \mathbf{u}^\delta(\tau)\rangle\, d\tau\end{aligned} \tag{11}$$

for all $0 \le s, t \le T$. Fix any point $s \in (0,T)$ for which

$$\mathbf{u}^\varepsilon(s) \to \mathbf{u}(s) \quad \text{in } L^2(U).$$

Consequently (11) implies

$$\begin{aligned}\limsup_{\varepsilon,\delta\to 0}\sup_{0\le t\le T}\|\mathbf{u}^\varepsilon(t) - \mathbf{u}^\delta(t)\|^2_{L^2(U)} &\le \lim_{\varepsilon,\delta\to 0}\int_0^T \|\mathbf{u}^{\varepsilon'}(\tau) - \mathbf{u}^{\delta'}(\tau)\|^2_{H^{-1}(U)} \\ &\qquad + \|\mathbf{u}^\varepsilon(\tau) - \mathbf{u}^\delta(\tau)\|^2_{H_0^1(U)}\, d\tau \\ &= 0.\end{aligned}$$

Thus the smoothed functions $\{\mathbf{u}^\varepsilon\}_{0<\varepsilon\leq 1}$ converge in $C([0,T];L^2(U))$ to a limit $\mathbf{v} \in C([0,T];L^2(U))$. Since we also know $\mathbf{u}^\varepsilon(t) \to \mathbf{u}(t)$ for a.e. t, we deduce $\mathbf{u} = \mathbf{v}$ a.e.

2. We similarly have

$$\|\mathbf{u}^\varepsilon(t)\|_{L^2(U)}^2 = \|\mathbf{u}^\varepsilon(s)\|_{L^2(U)}^2 + 2\int_s^t \langle \mathbf{u}^{\varepsilon\prime}(\tau), \mathbf{u}^\varepsilon(\tau)\rangle \, d\tau,$$

and so, identifying $\mathbf{u}$ with $\mathbf{v}$ above,

$$\|\mathbf{u}(t)\|_{L^2(U)}^2 = \|\mathbf{u}(s)\|_{L^2(U)}^2 + 2\int_s^t \langle \mathbf{u}'(\tau), \mathbf{u}(\tau)\rangle \, d\tau \tag{12}$$

for all $0 \leq s,t \leq T$.

3. To obtain (10), we integrate (12) with respect to s, recall the inequality $|\langle \mathbf{u}', \mathbf{u}\rangle| \leq \|\mathbf{u}'\|_{H^{-1}(U)}\|\mathbf{u}\|_{H_0^1(U)}$, and make some simple estimates. □

For use later in the regularity theory for second-order parabolic and hyperbolic equations in Chapter 7, we will also need this extension of Theorem 3.

THEOREM 4 (Mappings into better spaces). *Assume that U is open, bounded, and ∂U is smooth. Take m to be a nonnegative integer.*

Suppose $\mathbf{u} \in L^2(0,T;H^{m+2}(U))$, with $\mathbf{u}' \in L^2(0,T;H^m(U))$.

(i) *Then*

$$\mathbf{u} \in C([0,T];H^{m+1}(U))$$

(after possibly being redefined on a set of measure zero).

(ii) *Furthermore, we have the estimate*

$$\max_{0\leq t\leq T}\|\mathbf{u}(t)\|_{H^{m+1}(U)} \leq C\big(\|\mathbf{u}\|_{L^2(0,T;H^{m+2}(U))} + \|\mathbf{u}'\|_{L^2(0,T;H^m(U))}\big), \tag{13}$$

the constant C depending only on T, U, and m.

Proof. 1. Suppose first that $m = 0$, in which case

$$\mathbf{u} \in L^2(0,T;H^2(U)), \ \mathbf{u}' \in L^2(0,T;L^2(U)).$$

We select a bounded open set $V \supset\supset U$, and then construct a corresponding extension $\bar{\mathbf{u}} = E\mathbf{u}$, as in §5.4. In view of estimate (10) from that section, we see

$$\bar{\mathbf{u}} \in L^2(0,T;H^2(V)),$$

and

$$\|\bar{\mathbf{u}}\|_{L^2(0,T;H^2(V))} \leq C\|\mathbf{u}\|_{L^2(0,T;H^2(U))}, \tag{14}$$

for an appropriate constant C. In addition, $\bar{\mathbf{u}}' \in L^2(0,T;L^2(V))$, with the estimate

$$\|\bar{\mathbf{u}}'\|_{L^2(0,T;L^2(V))} \leq C\|\mathbf{u}'\|_{L^2(0,T;L^2(U))}. \tag{15}$$

This follows if we consider difference quotients in the t-variable, remember the methods in §5.8.2, and observe also that E is a bounded linear operator from $L^2(U)$ into $L^2(V)$.

2. Assume for the moment that $\bar{\mathbf{u}}$ is smooth. We then compute

$$\begin{aligned}\Big|\frac{d}{dt}\Big(\int_V |D\bar{\mathbf{u}}|^2\,dx\Big)\Big| &= 2\Big|\int_V D\bar{\mathbf{u}}\cdot D\bar{\mathbf{u}}'\,dx\Big| = 2\Big|\int_V \Delta\bar{\mathbf{u}}\,\bar{\mathbf{u}}'\,dx\Big| \\ &\leq C\big(\|\bar{\mathbf{u}}\|^2_{H^2(V)} + \|\bar{\mathbf{u}}'\|^2_{L^2(V)}\big).\end{aligned}$$

There is no boundary term when we integrate by parts, since the extension $\bar{\mathbf{u}} = E\mathbf{u}$ has compact support within V. Integrating and recalling (14), (15), it follows that

$$\max_{0\leq t\leq T}\|\mathbf{u}(t)\|_{H^1(U)} \leq C\big(\|\mathbf{u}\|_{L^2(0,T;H^2(U))} + \|\mathbf{u}'\|_{L^2(0,T;L^2(U))}\big). \tag{16}$$

We obtain the same estimate if $\mathbf{u}$ is not smooth, upon approximating by $\mathbf{u}^\varepsilon := \eta_\varepsilon * \mathbf{u}$, as before. As in the previous proofs, it also follows that $\mathbf{u} \in C([0,T];H^1(U))$.

3. In the general case that $m \geq 1$, we let α be a multiindex of order $|\alpha| \leq m$, and set $\mathbf{v} := D^\alpha\mathbf{u}$. Then

$$\mathbf{v} \in L^2(0,T;H^2(U)),\ \mathbf{v}' \in L^2(0,T;L^2(U)).$$

We apply estimate (16), with $\mathbf{v}$ replacing $\mathbf{u}$, and sum over all indices $|\alpha| \leq m$, to derive (13).

□

5.10. PROBLEMS

In these exercises U always denotes an open subset of $\mathbb{R}^n$.

1. Suppose $k \in \{0,1,\dots\}$, $0 < \gamma \leq 1$. Prove $C^{k,\gamma}(\bar{U})$ is a Banach space.

2. Let U, V be open sets, with $V \subset\subset U$. Show there exists a smooth function ζ such that $\zeta \equiv 1$ on V, $\zeta = 0$ near ∂U. (Hint: Take $V \subset\subset W \subset\subset U$ and mollify χ_W.)

3. Assume $0 < \beta < \gamma \leq 1$. Prove the interpolation inequality

$$\|u\|_{C^{0,\gamma}(U)} \leq \|u\|_{C^{0,\beta}(U)}^{\frac{1-\gamma}{1-\beta}} \|u\|_{C^{0,1}(U)}^{\frac{\gamma-\beta}{1-\beta}}.$$

4. Assume U is bounded and $U \subset\subset \bigcup_{i=1}^N V_i$. Show there exist C^∞ functions ζ_i $(i = 1, \dots, N)$ such that

$$\begin{cases} 0 \leq \zeta_i \leq 1, \ \operatorname{spt} \zeta_i \subset V_i \quad (i = 1, \dots, N) \\ \sum_{i=1}^N \zeta_i = 1 \quad \text{on } U. \end{cases}$$

The functions $\{\zeta_i\}_{i=1}^N$ form a *partition of unity.*

5. Prove that if $n = 1$ and $u \in W^{1,p}(0,1)$ for some $1 \leq p < \infty$, then u is equal a.e. to an absolutely continuous function, and u' (which exists a.e.) belongs to $L^p(0,1)$.

6. Prove directly that if $u \in W^{1,p}(0,1)$ for some $1 < p < \infty$, then $|u(x) - u(y)| \leq |x-y|^{1-\frac{1}{p}} \left(\int_0^1 |u'|^p dt\right)^{1/p}$ for a.e. $x, y \in [0,1]$.

7. Denote by U the open square $\{x \in \mathbb{R}^2 \mid |x_1| < 1, |x_2| < 1\}$. Define

$$u(x) = \begin{cases} 1 - x_1 & \text{if } x_1 > 0, \ |x_2| < x_1 \\ 1 + x_1 & \text{if } x_1 < 0, \ |x_2| < -x_1 \\ 1 - x_2 & \text{if } x_2 > 0, \ |x_1| < x_2 \\ 1 + x_2 & \text{if } x_2 < 0, \ |x_1| < -x_2. \end{cases}$$

For which $1 \leq p \leq \infty$ does u belong to $W^{1,p}(U)$?

8. Integrate by parts to prove the interpolation inequality:

$$\int_U |Du|^2 dx \leq C \left(\int_U u^2 dx\right)^{1/2} \left(\int_U |D^2 u|^2 dx\right)^{1/2}$$

for all $u \in C_c^\infty(U)$. By approximation, prove this inequality if $u \in H^2(U) \cap H_0^1(U)$.

9. Integrate by parts to prove:

$$\int_U |Du|^p dx \leq C \left(\int_U |u|^p dx\right)^{1/2} \left(\int_U |D^2 u|^p dx\right)^{1/2}$$

for $2 \leq p < \infty$ and all $u \in W^{2,p}(U) \cap W_0^{1,p}(U)$. (Hint: $\int_U |Du|^p\,dx = \sum_{i=1}^n \int_U u_{x_i} u_{x_i} |Du|^{p-2} dx$.)

10. Suppose U is connected and $u \in W^{1,p}(U)$ satisfies

$$Du = 0 \quad \text{a.e. in } U.$$

Prove u is constant a.e. in U.

11. Show by example that if we have $\|D^h u\|_{L^1(V)} \leq C$ for all $0 < |h| < \frac{1}{2}\operatorname{dist}(V, \partial U)$, it does not necessarily follow that $u \in W^{1,1}(V)$.

12. Give an example of an open set $U \subset \mathbb{R}^n$ and a function $u \in W^{1,\infty}(U)$, such that u is *not* Lipschitz continuous on U. (Hint: Take U to be the open unit disk in $\mathbb{R}^2$, with a slit removed.)

13. Verify that if $n > 1$, the unbounded function $u = \log\log\left(1 + \frac{1}{|x|}\right)$ belongs to $W^{1,n}(U)$, for $U = B^0(0,1)$.

14. Let U be bounded, with a C^1 boundary. Show that a "typical" function $u \in L^p(U)$ $(1 \leq p < \infty)$ does not have a trace on ∂U. More precisely, prove there does not exist a bounded linear operator

$$T : L^p(U) \to L^p(\partial U)$$

such that $Tu = u|_{\partial U}$ whenever $u \in C(\bar{U}) \cap L^p(U)$.

15. Fix $\alpha > 0$ and let $U = B^0(0,1)$. Show there exists a constant C, depending only on n and α, such that

$$\int_U u^2 dx \leq C \int_U |Du|^2 dx\ ,$$

provided

$$|\{x \in U \mid u(x) = 0\}| \geq \alpha\ , \quad u \in H^1(U).$$

16. Assume $F : \mathbb{R} \to \mathbb{R}$ is C^1, with F' bounded. Suppose U is bounded and $u \in W^{1,p}(U)$ for some $1 < p < \infty$. Show

$$v := F(u) \in W^{1,p}(U) \quad \text{and} \quad v_{x_i} = F'(u) u_{x_i} \quad (i = 1, \dots, n).$$

17. Assume $1 < p < \infty$, and U is bounded.
(i) Prove that if $u \in W^{1,p}(U)$, then $|u| \in W^{1,p}(U)$.

(ii) Prove $u \in W^{1,p}(U)$ implies $u^+, u^- \in W^{1,p}(U)$, and

$$Du^+ = \begin{cases} Du & \text{a.e. on } \{u > 0\} \\ 0 & \text{a.e. on } \{u \leq 0\}, \end{cases}$$
$$Du^- = \begin{cases} 0 & \text{a.e. on } \{u \geq 0\} \\ -Du & \text{a.e. on } \{u < 0\}. \end{cases}$$

(Hint: $u^+ = \lim_{\varepsilon \to 0} F_\varepsilon(u)$, for

$$F_\varepsilon(z) := \begin{cases} (z^2 + \varepsilon^2)^{1/2} - \varepsilon & \text{if } z \geq 0 \\ 0 & \text{if } z < 0.) \end{cases}$$

(iii) Prove that if $u \in W^{1,p}(U)$, then

$$Du = 0 \text{ a.e. on the set } \{u = 0\}.$$

18. Use the Fourier transform to prove that if $u \in H^s(\mathbb{R}^n)$ for $s > n/2$, then $u \in L^\infty(\mathbb{R}^n)$, with the bound

$$\|u\|_{L^\infty(\mathbb{R}^n)} \leq C\|u\|_{H^s(\mathbb{R}^n)},$$

the constant C depending only on s and n.

5.11. REFERENCES

Sections 5.2-8	See Gilbarg–Trudinger [**G-T**, Chapter 7], Lieb–Loss [**L-L**], Ziemer [**Z**] and [**E-G**] for more on Sobolev spaces.
Section 5.5	W. Schlag showed me the proof of Theorem 2.
Section 5.6	J. Ralston suggested an improvement in the proof of Theorem 4.
Section 5.9	See Temam [**TE**, pp. 248–273].

Chapter 6

SECOND-ORDER ELLIPTIC EQUATIONS

This chapter investigates the solvability of uniformly elliptic, second-order partial differential equations, subject to prescribed boundary conditions. We will exploit two essentially distinct techniques, energy methods within Sobolev spaces (§§6.1–6.3) and maximum principle methods (§6.4).

6.1. DEFINITIONS

6.1.1. Elliptic equations.

We will in this chapter mostly study the boundary-value problem

$$(1)\qquad \begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases}$$

where U is an open, bounded subset of $\mathbb{R}^n$ and $u : \bar{U} \to \mathbb{R}$ is the unknown, $u = u(x)$. Here $f : U \to \mathbb{R}$ is given, and L denotes a second-order partial differential operator having either the form

(2) $$Lu = -\sum_{i,j=1}^{n} \left(a^{ij}(x)u_{x_i}\right)_{x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u$$

or else

(3) $$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)u_{x_i x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u,$$

for given coefficient functions a^{ij}, b^i, c $(i, j = 1, \dots, n)$.

We say that the PDE $Lu = f$ is in *divergence form* if L is given by (2), and is in *nondivergence form* provided L is given by (3). The requirement that $u = 0$ on ∂U in (1) is sometimes called *Dirichlet's boundary condition.*

Remark. If the highest order coefficients a^{ij} $(i, j = 1, \dots, n)$ are C^1 functions, then an operator given in divergence form can be rewritten into nondivergence structure, and vice versa. Indeed the divergence form equation (2) becomes

(2′) $$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)u_{x_i x_j} + \sum_{i=1}^{n} \tilde{b}^i(x)u_{x_i} + c(x)u$$

for $\tilde{b}^i := b^i - \sum_{j=1}^{n} a^{ij}_{x_j}$ $(i = 1, \dots, n)$, and (2′) is obviously in nondivergence form. We will see, however, there are definite advantages to considering the two different representations of L separately. The divergence form is most natural for energy methods, based upon integration by parts (§§6.1–6.3), and the nondivergence form is most appropriate for maximum principle techniques (§6.4). □

We henceforth assume as well the symmetry condition

$$a^{ij} = a^{ji} \quad (i, j = 1, \dots, n) .$$

DEFINITION. *We say the partial differential operator L is* (uniformly) elliptic *if there exists a constant $\theta > 0$ such that*

(4) $$\sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2$$

for a.e. $x \in U$ and all $\xi \in \mathbb{R}^n$.

Ellipticity thus means that for each point $x \in U$, the symmetric $n \times n$ matrix $\mathbf{A}(x) = ((a^{ij}(x)))$ is positive definite, with smallest eigenvalue greater than or equal to θ.

An obvious example is $a^{ij} \equiv \delta_{ij}$, $b^i \equiv 0$, $c \equiv 0$, in which case the operator L is $-\Delta$. Indeed we will see that solutions of the general second-order elliptic PDE $Lu = 0$ are similar in many ways to harmonic functions. However, for these partial differential equations we do not have available the various explicit formulas developed for harmonic functions in Chapter 2: we must instead work directly with the PDE. Readers should continually be alert in the following calculations for uses of the structural condition of ellipticity (4).

Physical interpretation. As just noted, second-order elliptic PDE generalize Laplace's and Poisson's equations. As in the derivation of Laplace's equation set forth in §2.2, u in applications typically represents the density of some quantity, say a chemical concentration, at equilibrium within a region U. The second-order term $\mathbf{A} : D^2u = \sum_{i,j=1} a^{ij}u_{x_i x_j}$ represents the *diffusion* of u within U, the coefficients $((a^{ij}))$ describing the anisotropic, heterogeneous nature of the medium. In particular, $\mathbf{F} := -\mathbf{A}Du$ is the diffusive flux density, and the ellipticity condition implies

$$\mathbf{F} \cdot Du \leq 0;$$

that is, the flow is from regions of higher to lower concentration. The first-order term $\mathbf{b} \cdot Du = \sum_{i=1}^n b^i u_{x_i}$ represents *transport* within U, and the zeroth-order term cu describes the local *creation* or *depletion* of the chemical (owing, say, to reactions). A careful analysis of these interpretations requires the probabilistic study of diffusion processes.

Nonlinear second-order elliptic PDE also arise naturally in the calculus of variations (as the Euler–Lagrange equations of convex energy integrands) and in differential geometry (as expressions involving curvatures). We will encounter some such nonlinear equations later, in Chapters 8 and 9. □

6.1.2. Weak solutions.

Let us consider first the boundary-value problem (1) when L has the divergence form (2). Our overall plan is first to define and then construct an appropriate weak solution u of (1), and only later to investigate the smoothness and other properties of u.

We will assume in the following exposition that

$$a^{ij}, b^i, c \in L^\infty(U) \quad (i, j = 1, \ldots, n) \tag{5}$$

and

$$f \in L^2(U). \tag{6}$$

Motivation for definition of weak solution. How should we define a weak or generalized solution? Assuming for the moment u is really a smooth solution, let us multiply the PDE $Lu = f$ by a smooth test function $v \in C_c^\infty(U)$, and integrate over U, to find

$$\int_U \sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv\, dx = \int_U fv\, dx, \tag{7}$$

where we have integrated by parts in the first term on the left hand side. There are no boundary terms since $v = 0$ on ∂U. By approximation we find the same identity holds with the smooth function v replaced by any $v \in H_0^1(U)$, and the resulting identity makes sense if only $u \in H_0^1(U)$. (We choose the space $H_0^1(U)$ to incorporate the boundary condition from (1) that "$u = 0$ on ∂U".)

DEFINITIONS. (i) *The* bilinear form $B[\ ,\ ]$ *associated with the divergence form elliptic operator* L *defined by* (2) *is*

$$B[u,v] := \int_U \sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv\, dx \tag{8}$$

for $u, v \in H_0^1(U)$.

(ii) *We say that* $u \in H_0^1(U)$ *is a* weak solution *of the boundary-value problem* (1) *if*

$$B[u,v] = (f,v) \tag{9}$$

for all $v \in H_0^1(U)$, *where* $(\ ,\ )$ *denotes the inner product in* $L^2(U)$.

Remark. The identity (9) is sometimes called the *variational formulation* of (1). This terminology will be explained later, in Example 2 of §8.1.2. □

More generally, let us consider the boundary-value problem

$$\begin{cases} Lu = f^0 - \sum_{i=1}^n f^i_{x_i} & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases} \tag{10}$$

where L is defined by (2) and $f^i \in L^2(U)$ ($i = 0, \dots, n$). In view of the theory set forth in §5.9.1 we see that the righthand term $f = f^0 - \sum_{i=1}^n f^i_{x_i}$ belongs to $H^{-1}(U)$, the dual space of $H_0^1(U)$.

DEFINITION. *We say $u \in H_0^1(U)$ is a* weak solution *of problem* (10) *provided*

$$B[u,v] = \langle f, v\rangle$$

for all $v \in H_0^1(U)$, where $\langle f,v\rangle = \int_U f^0 v + \sum_{i=1}^n f^i v_{x_i}\,dx$ and $\langle\ ,\ \rangle$ is the pairing of $H^{-1}(U)$ and $H_0^1(U)$.

Remark. We will hereafter, as above, focus attention exclusively on the case of zero boundary conditions, but in fact a problem with prescribed, nonzero boundary values can easily be transformed into this setting. We spell this out by supposing now that ∂U is C^1 and $u \in H^1(U)$ is a weak solution of

$$\begin{cases} Lu = f & \text{in } U \\ u = g & \text{on } \partial U. \end{cases}$$

This means that $u = g$ on ∂U in the trace sense, and furthermore that the bilinear form identity (9) holds for all $v \in H_0^1(U)$. That this be possible, it is necessary for g to be the trace of some H^1 function, say w. But then $\tilde{u} := u - w$ belongs to $H_0^1(U)$, and is a weak solution of the boundary-value problem

$$\begin{cases} L\tilde{u} = \tilde{f} & \text{in } U \\ \tilde{u} = 0 & \text{on } \partial U, \end{cases}$$

where $\tilde{f} := f - Lw \in H^{-1}(U)$.

See Problems 2, 3 to learn how to cast some other sorts of PDE and boundary conditions into weak formulations. □

6.2. EXISTENCE OF WEAK SOLUTIONS

6.2.1. Lax–Milgram Theorem.

We now introduce a fairly simple abstract principle from linear functional analysis, which will later in §6.2.2 provide in certain circumstances the existence and uniqueness of a weak solution to our boundary-value problem.

We assume for this section H is a real Hilbert space, with norm $\|\ \|$ and inner product $(\ ,\ )$. We let $\langle\ ,\ \rangle$ denote the pairing of H with its dual space. Readers should review as necessary the basic Hilbert space theory described in §D.2-3.

THEOREM 1 (Lax–Milgram Theorem). *Assume that*

$$B : H \times H \to \mathbb{R}$$

is a bilinear mapping, for which there exist constants $\alpha, \beta > 0$ *such that*

$$\text{(i)} \qquad |B[u,v]| \le \alpha \|u\| \, \|v\| \quad (u, v \in H)$$

and

$$\text{(ii)} \qquad \beta \|u\|^2 \le B[u,u] \quad (u \in H).$$

Finally, let $f : H \to \mathbb{R}$ *be a bounded linear functional on* H.

Then there exists a unique element $u \in H$ *such that*

$$\text{(1)} \qquad B[u,v] = \langle f, v \rangle$$

for all $v \in H$.

Proof. 1. For each fixed element $u \in H$, the mapping $v \mapsto B[u,v]$ is a bounded linear functional on H; whence the Riesz Representation Theorem (§D.3) asserts the existence of a unique element $w \in H$ satisfying

$$\text{(2)} \qquad B[u,v] = (w,v) \quad (v \in H).$$

Let us write $Au = w$ whenever (2) holds; so that

$$\text{(3)} \qquad B[u,v] = (Au, v) \quad (u, v \in H).$$

2. We first claim $A : H \to H$ is a bounded linear operator. Indeed if $\lambda_1, \lambda_2 \in \mathbb{R}$ and $u_1, u_2 \in H$, we see for each $v \in H$ that

$$\begin{aligned}
(A(\lambda_1 u_1 + \lambda_2 u_2), v) &= B[\lambda_1 u_1 + \lambda_2 u_2, v] \quad \text{by (3)} \\
&= \lambda_1 B[u_1, v] + \lambda_2 B[u_2, v] \\
&= \lambda_1 (Au_1, v) + \lambda_2 (Au_2, v) \quad \text{by (3) again} \\
&= (\lambda_1 Au_1 + \lambda_2 Au_2, v).
\end{aligned}$$

This equality obtains for each $v \in H$, and so A is linear. Furthermore

$$\|Au\|^2 = (Au, Au) = B[u, Au] \le \alpha \|u\| \, \|Au\|.$$

Consequently $\|Au\| \le \alpha \|u\|$ for all $u \in H$, and so A is bounded.

3. Next we assert

$$\text{(4)} \qquad \begin{cases} A \text{ is one-to-one, and} \\ R(A), \text{ the range of } A, \text{ is closed in } H. \end{cases}$$

To prove this, let us compute

$$\beta \|u\|^2 \le B[u,u] = (Au, u) \le \|Au\| \, \|u\|.$$

Hence $\beta\|u\| \leq \|Au\|$. This inequality easily implies (4).

4. We demonstrate now

$$R(A) = H. \tag{5}$$

For if not, then, since $R(A)$ is closed, there would exist a nonzero element $w \in H$ with $w \in R(A)^\perp$. But this fact in turn implies the contradiction $\beta\|w\|^2 \leq B[w,w] = (Aw,w) = 0$.

5. Next, we observe once more from the Riesz Representation Theorem that

$$\langle f, v\rangle = (w,v) \quad \text{for all } v \in H$$

for some element $w \in H$. We then utilize (4) and (5) to find $u \in H$ satisfying $Au = w$. Then

$$B[u,v] = (Au,v) = (w,v) = \langle f,v\rangle \quad (v \in H),$$

and this is (1).

6. Finally, we show there is at most one element $u \in H$ verifying (1). For if both $B[u,v] = \langle f,v\rangle$ and $B[\tilde{u},v] = \langle f,v\rangle$, then $B[u-\tilde{u},v] = 0$ $(v \in H)$. We set $v = u - \tilde{u}$ to find $\beta\|u-\tilde{u}\|^2 \leq B[u-\tilde{u},u-\tilde{u}] = 0$. □

Remark. If the bilinear form $B[\ ,\ ]$ is symmetric, that is, if

$$B[u,v] = B[v,u] \quad (u,v \in H),$$

we can fashion a much simpler proof by noting $((u,v)) := B[u,v]$ is a new inner product on H, to which the Riesz Representation Theorem directly applies. Consequently, the Lax–Milgram Theorem is primarily significant in that it does *not* require symmetry of $B[\ ,\ ]$. □

6.2.2. Energy estimates.

We return now to the specific bilinear form $B[\ ,\ ]$, defined in §6.1.2 by the formula

$$B[u,v] = \int_U \sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv\,dx$$

for $u,v \in H_0^1(U)$, and try to verify the hypothesis of the Lax–Milgram Theorem.

THEOREM 2 (Energy estimates). *There exist constants $\alpha, \beta > 0$ and $\gamma \geq 0$ such that*

(i)
$$|B[u,v]| \leq \alpha \|u\|_{H_0^1(U)} \|v\|_{H_0^1(U)}$$

and

(ii)
$$\beta \|u\|^2_{H_0^1(U)} \leq B[u,u] + \gamma \|u\|^2_{L^2(U)}$$

for all $u, v \in H_0^1(U)$.

Proof. 1. We readily check

$$\begin{aligned} |B[u,v]| &\leq \sum_{i,j=1}^n \|a^{ij}\|_{L^\infty} \int_U |Du|\, |Dv|\, dx \\ &\quad + \sum_{i=1}^n \|b^i\|_{L^\infty} \int_U |Du|\, |v|\, dx + \|c\|_{L^\infty} \int_U |u|\, |v|\, dx \\ &\leq \alpha \|u\|_{H_0^1(U)} \|v\|_{H_0^1(U)}, \end{aligned}$$

for some appropriate constant α.

2. Furthermore, in view of the ellipticity condition (4) from §6.1 we have

$$\begin{aligned} \theta \int_U |Du|^2\, dx &\leq \int_U \sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j}\, dx \\ &= B[u,u] - \int_U \sum_{i=1}^n b^i u_{x_i} u + cu^2 dx \qquad (6) \\ &\leq B[u,u] + \sum_{i=1}^n \|b^i\|_{L^\infty} \int_U |Du|\, |u|\, dx + \|c\|_{L^\infty} \int_U u^2 dx. \end{aligned}$$

Now from Cauchy's inequality with ε (§B.2), we observe

$$\int_U |Du|\, |u|\, dx \leq \varepsilon \int_U |Du|^2 dx + \frac{1}{4\varepsilon} \int_U u^2 dx \quad (\varepsilon > 0).$$

We insert this estimate into (6) and then choose $\varepsilon > 0$ so small that

$$\varepsilon \sum_{i=1}^n \|b^i\|_{L^\infty} < \frac{\theta}{2}.$$

Thus

$$\frac{\theta}{2} \int_U |Du|^2 dx \leq B[u,u] + C \int_U u^2 dx$$

for some appropriate constant C. In addition we recall from Poincaré's inequality in §5.6.1 that

$$\|u\|_{L^2(U)} \leq C\|Du\|_{L^2(U)}.$$

It easily follows that

$$\beta\|u\|^2_{H^1_0(U)} \leq B[u,u] + \gamma\|u\|^2_{L^2(U)}$$

for appropriate constants $\beta > 0$, $\gamma \geq 0$. □

Observe now that if $\gamma > 0$ in these energy estimates, then $B[\ ,\ ]$ does not precisely satisfy the hypotheses of the Lax–Milgram Theorem. The following existence assertion for weak solutions must confront this possibility:

THEOREM 3 (First Existence Theorem for weak solutions). *There is a number $\gamma \geq 0$ such that for each*

$$\mu \geq \gamma \tag{7}$$

and each function

$$f \in L^2(U),$$

there exists a unique weak solution $u \in H^1_0(U)$ of the boundary-value problem

$$\begin{cases} Lu + \mu u = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{8}$$

Proof. 1. Take γ from Theorem 2, let $\mu \geq \gamma$, and define then the bilinear form

$$B_\mu[u,v] := B[u,v] + \mu(u,v) \quad (u,v \in H^1_0(U)),$$

which corresponds as in §6.1 to the operator $L_\mu u := Lu + \mu u$. As before $(\ ,\ )$ means the inner product in $L^2(U)$. Then $B_\mu[\ ,\ ]$ satisfies the hypotheses of the Lax–Milgram Theorem.

2. Now fix $f \in L^2(U)$ and set $\langle f, v\rangle := (f,v)_{L^2(U)}$. This is a bounded linear functional on $L^2(U)$, and thus on $H^1_0(U)$.

We apply the Lax–Milgram Theorem to find a unique function $u \in H^1_0(U)$ satisfying

$$B_\mu[u,v] = \langle f, v\rangle$$

for all $v \in H^1_0(U)$; u is consequently the unique weak solution of (8). □

Remark. We can similarly show that for all

$$f^i \in L^2(U) \quad (i = 0, \ldots, n),$$

there exists a unique weak solution u of the PDE

$$(9) \qquad \begin{cases} Lu + \mu u = f^0 - \sum_{i=1}^n f^i_{x_i} & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Indeed, it is enough to note $\langle f, v \rangle = \int_U f^0 v + \sum_{i=1}^n f^i v_{x_i} \, dx$ is a bounded linear functional on $H_0^1(U)$, as previously discussed in §5.9.1.

In particular, we deduce that the mapping

$$L_\mu := L + \mu I : H_0^1(U) \to H^{-1}(U) \quad (\mu \geq \gamma)$$

is an isomorphism. □

Examples. In the case $Lu = -\Delta u$, so that $B[u,v] = \int_U Du \cdot Dv \; dx$, we easily check using Poincaré's inequality that Theorem 2 holds with $\gamma = 0$. A similar assertion holds for the general operator $Lu = -\sum_{i,j=1}^n \left(a^{ij} u_{x_i}\right)_{x_j} + cu$, provided $c \geq 0$ in U. □

6.2.3. Fredholm alternative.

We next employ the Fredholm theory for compact operators (discussed in §D.5) to glean more detailed information regarding the solvability of second-order elliptic PDE.

DEFINITIONS. (i) *The operator L^*, the* formal adjoint *of L, is*

$$L^* v := -\sum_{i,j=1}^n (a^{ij} v_{x_j})_{x_i} - \sum_{i=1}^n b^i v_{x_i} + \left(c - \sum_{i=1}^n b_{i,x_i}\right) v,$$

provided $b^i \in C^1(\bar{U})$ $(i = 1, \ldots, n)$.

(ii) *The* adjoint bilinear form

$$B^* : H \times H \to \mathbb{R}$$

is defined by

$$B^*[v,u] = B[u,v]$$

for all $u, v \in H_0^1(U)$.

(iii) *We say that $v \in H_0^1(U)$ is a* weak solution *of the* adjoint problem

$$\begin{cases} L^* v = f & \textit{in } U \\ v = 0 & \textit{on } \partial U, \end{cases}$$

provided

$$B^*[v,u] = (f,u)$$

for all $u \in H_0^1(U)$.

THEOREM 4 (Second Existence Theorem for weak solutions).

(i) *Precisely one of the following statements holds:*
either

$$(\alpha)\quad \begin{cases} \text{for each } f \in L^2(U) \text{ there exists a unique} \\ \text{weak solution } u \text{ of the boundary-value problem} \\ (10) \qquad \begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases} \end{cases}$$

or else

$$(\beta)\quad \begin{cases} \text{there exists a weak solution } u \not\equiv 0 \text{ of} \\ \text{the homogeneous problem} \\ (11) \qquad \begin{cases} Lu = 0 & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \end{cases}$$

(ii) *Furthermore, should assertion* (β) *hold, the dimension of the subspace* $N \subset H_0^1(U)$ *of weak solutions of* (11) *is finite and equals the dimension of the subspace* $N^* \subset H_0^1(U)$ *of weak solutions of*

$$(12) \qquad \begin{cases} L^*v = 0 & \text{in } U \\ v = 0 & \text{on } \partial U. \end{cases}$$

(iii) *Finally, the boundary-value problem* (10) *has a weak solution if and only if*

$$(f, v) = 0 \quad \text{for all } v \in N^*.$$

The dichotomy $(\alpha), (\beta)$ is the *Fredholm alternative.*

Proof. 1. Choose $\mu = \gamma$ as in Theorem 3 and define the bilinear form

$$B_\gamma[u, v] := B[u, v] + \gamma(u, v),$$

corresponding to the operator $L_\gamma u := Lu + \gamma u$. Then for each $g \in L^2(U)$ there exists a unique function $u \in H_0^1(U)$ solving

$$(13) \qquad B_\gamma[u, v] = (g, v) \quad \text{for all } v \in H_0^1(U).$$

Let us write

$$(14) \qquad u = L_\gamma^{-1} g$$

whenever (13) holds.

2. Observe next $u \in H_0^1(U)$ is a weak solution of (10) if and only if

$$(15) \qquad B_\gamma[u, v] = (\gamma u + f, v) \quad \text{for all } v \in H_0^1(U);$$

that is, if and only if

$$u = L_\gamma^{-1}(\gamma u + f). \tag{16}$$

We rewrite this equality to read

$$u - Ku = h, \tag{17}$$

for

$$Ku := \gamma L_\gamma^{-1} u \tag{18}$$

and

$$h := L_\gamma^{-1} f. \tag{19}$$

3. We now claim $K : L^2(U) \to L^2(U)$ is a bounded, linear, compact operator. Indeed, from our choice of γ and the energy estimates from §6.2.2 we note that if (13) holds, then

$$\beta \|u\|^2_{H_0^1(U)} \leq B_\gamma[u,u] = (g,u) \leq \|g\|_{L^2(U)}\|u\|_{L^2(U)} \leq \|g\|_{L^2(U)}\|u\|_{H_0^1(U)};$$

so that (18) implies

$$\|Kg\|_{H_0^1(U)} \leq C\|g\|_{L^2(U)} \quad (g \in L^2(U))$$

for some appropriate constant C. But since $H_0^1(U) \subset\subset L^2(U)$ according to the Rellich-Kondrachov compactness theorem, we deduce that K is a compact operator.

4. We may consequently apply the Fredholm alternative from §D.5: either

$$(\alpha) \quad \begin{cases} \text{for each } h \in L^2(U) \text{ the equation} \\ u - Ku = h \\ \text{has a unique solution } u \in L^2(U) \end{cases} \tag{20}$$

or else

$$(\beta) \quad \begin{cases} \text{the equation} \\ u - Ku = 0 \\ \text{has nonzero solutions in } L^2(U). \end{cases} \tag{21}$$

Should assertion (α) hold, then according to (15)–(19) there exists a unique weak solution of problem (10). On the other hand, should assertion

(β) be valid, then necessarily $\gamma \neq 0$ and we recall further from §D.5 that the dimension of the space N of the solutions of (21) is finite and equals the dimension of the space N^* of solutions of

$$v - K^* v = 0. \tag{22}$$

We readily check however that (21) holds if and only if u is a weak solution of (11); and (22) holds if and only if v is a weak solution of (12).

5. Finally, we recall (20) has a solution if and only if

$$(h, v) = 0 \tag{23}$$

for all v solving (22). But from (18), (19) and (22) we compute

$$(h, v) = \frac{1}{\gamma}(Kf, v) = \frac{1}{\gamma}(f, K^* v) = \frac{1}{\gamma}(f, v) \ .$$

Consequently the boundary-value problem (10) has a solution if and only if $(f, v) = 0$ for all weak solutions v of (12). □

THEOREM 5 (Third Existence Theorem for weak solutions).

(i) *There exists an at most countable set $\Sigma \subset \mathbb{R}$ such that the boundary-value problem*

$$\begin{cases} Lu = \lambda u + f & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases} \tag{24}$$

has a unique weak solution for each $f \in L^2(U)$ if and only if $\lambda \notin \Sigma$.

(ii) *If Σ is infinite, then $\Sigma = \{\lambda_k\}_{k=1}^{\infty}$, the values of a nondecreasing sequence with*

$$\lambda_k \to +\infty.$$

DEFINITION. *We call Σ the* (real) spectrum *of the operator L.*

Note in particular the boundary-value problem

$$\begin{cases} Lu = \lambda u & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases}$$

has a nontrivial solution $w \not\equiv 0$ if and only if $\lambda \in \Sigma$, in which case λ is called an *eigenvalue* of L, w a corresponding *eigenfunction*. The partial differential equation $Lu = \lambda u$ for $L = -\Delta$ is sometimes called *Helmholtz's equation*.

□

Proof. 1. Let γ be the constant from Theorem 2 and assume

$$\lambda > -\gamma. \tag{25}$$

Assume also with no loss of generality that $\gamma > 0$.

2. According to the Fredholm alternative, the boundary-value problem (24) has a unique weak solution for each $f \in L^2(U)$ if and only if $u \equiv 0$ is the only weak solution of the homogeneous problem

$$\begin{cases} Lu = \lambda u & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

This is in turn true if and only if $u \equiv 0$ is the only weak solution of

$$\begin{cases} Lu + \gamma u = (\gamma + \lambda) u & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{26}$$

Now (26) holds exactly when

$$u = L_\gamma^{-1}(\gamma + \lambda) u = \frac{\gamma + \lambda}{\gamma} K u, \tag{27}$$

where, as in the proof of Theorem 4, we have set $Ku = \gamma L_\gamma^{-1} u$. Recall also from that proof that $K : L^2(U) \to L^2(U)$ is a bounded, linear, compact operator.

Now if $u \equiv 0$ is the only solution of (27), we see

$$\frac{\gamma}{\gamma + \lambda} \text{ is } not \text{ an eigenvalue of } K. \tag{28}$$

Consequently we see the PDE (24) has a unique weak solution for each $f \in L^2(U)$ if and only if (28) holds.

2. According to Theorem 6 in §D.5 the collection of all eigenvalues of K comprises either a finite set or else the values of a sequence converging to zero. In the second case we see, according to (25) and (27), that the PDE (24) has a unique weak solution for all $f \in L^2(U)$, except for a sequence $\lambda_k \to +\infty$. □

Finally, we explicitly note:

THEOREM 6 (Boundedness of the inverse). *If $\lambda \notin \Sigma$, there exists a constant C such that*

$$\|u\|_{L^2(U)} \leq C \|f\|_{L^2(U)}, \tag{29}$$

whenever $f \in L^2(U)$ and $u \in H_0^1(U)$ is the unique weak solution of

$$\begin{cases} Lu = \lambda u + f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

The constant C depends only on λ, U and the coefficients of L.

This constant will blow up if λ approaches an eigenvalue.

Proof. If not, there would exist sequences $\{f_k\}_{k=1}^\infty \subset L^2(U)$ and $\{u_k\}_{k=1}^\infty \subset H_0^1(U)$ such that

$$\begin{cases} Lu_k = \lambda u_k + f_k & \text{in } U \\ u_k = 0 & \text{on } \partial U \end{cases}$$

in the weak sense, but

$$\|u_k\|_{L^2(U)} > k\|f_k\|_{L^2(U)} \quad (k = 1, \dots).$$

As we may with no loss suppose $\|u_k\|_{L^2(U)} = 1$, we see $f_k \to 0$ in $L^2(U)$. According to the usual energy estimates the sequence $\{u_k\}_{k=1}^\infty$ is bounded in $H_0^1(U)$. Thus there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$ such that

$$\begin{cases} u_{k_j} \rightharpoonup u & \text{weakly in } H_0^1(U), \\ u_{k_j} \to u & \text{in } L^2(U). \end{cases} \tag{30}$$

(See §D.4 for weak convergence.) Then u is a weak solution of

$$\begin{cases} Lu = \lambda u & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Since $\lambda \notin \Sigma$, $u \equiv 0$. However (30) implies as well that $\|u\|_{L^2(U)} = 1$, a contradiction. □

Complex solutions. The foregoing theory extends easily to include complex-valued solutions. Given complex-valued $u, v \in H^1(U)$, write

$$(u,v)_{L^2(U)} := \int_U u\bar{v}\,dx, \ (u,v)_{H^1(U)} := \int_U Du \cdot D\bar{v} + u\bar{v}\,dx,$$

and set

$$B[u,v] := \int_U \sum_{i,j=1}^n a^{ij} u_{x_i} \bar{v}_{x_j} + \sum_{i=1}^n b^i u_{x_i} \bar{v} + cu\bar{v}\,dx,$$

where $\bar{\ }$ denotes complex conjugate. We check

$$|B[u,v]| \le \alpha \|u\|_{H_0^1(U)} \|v\|_{H_0^1(U)},$$

$$\beta \|u\|_{H_0^1(U)}^2 \le \operatorname{Re} B[u,u] + \gamma \|u\|_{L^2(U)}^2 \quad (u, v \in H_0^1(U))$$

for appropriate constants $\alpha, \beta > 0$, $\gamma \ge 0$. Complex variants of the Lax–Milgram Theorem and Fredholm alternative lead to analogues of Theorems 3–6 above. □

6.3. REGULARITY

We now address the question as to whether a weak solution u of the PDE

$$Lu = f \quad \text{in } U \tag{1}$$

is in fact smooth: this is the *regularity* problem for weak solutions.

Motivation: formal derivation of estimates. To see that there is some hope that a weak solution may be better than a typical function in $H_0^1(U)$, let us consider the model problem

$$-\Delta u = f \quad \text{in } \mathbb{R}^n. \tag{2}$$

We assume for heuristic purposes that u is smooth and vanishes sufficiently rapidly as $|x| \to \infty$ to justify the following calculations. We then compute

$$\begin{aligned} \int_{\mathbb{R}^n} f^2 dx = \int_{\mathbb{R}^n} (\Delta u)^2 dx &= \sum_{i,j=1}^n \int_{\mathbb{R}^n} u_{x_i x_i} u_{x_j x_j}\, dx \\ &= -\sum_{i,j=1}^n \int_{\mathbb{R}^n} u_{x_i x_i x_j} u_{x_j}\, dx \\ &= \sum_{i,j=1}^n \int_{\mathbb{R}^n} u_{x_i x_j} u_{x_i x_j}\, dx = \int_{\mathbb{R}^n} |D^2 u|^2 dx. \end{aligned} \tag{3}$$

Thus we see the L^2-norm of the *second* derivatives of u can be estimated by (and in fact equals) the L^2-norm of f. Similarly, we can differentiate the PDE (2), to find

$$-\Delta \tilde{u} = \tilde{f},$$

for $\tilde{u} := u_{x_k}$ and $\tilde{f} := f_{x_k}$ $(k = 1, \dots, n)$. Applying the same method, we discover that the L^2-norm of the third derivatives of u can be estimated by the first derivatives of f. Continuing, we see the L^2-norm of the $(m+2)^{nd}$ derivatives of u can be controlled by the L^2-norm of the m^{th} derivatives of f, for $m = 0, 1, \dots$. □

These computations suggest that for Poisson's equation (2), we can expect a weak solution $u \in H_0^1$ to belong to H^{m+2} whenever the inhomogeneous term f belongs to H^m $(m = 1, \dots)$. Informally we say that u has "two more derivatives in L^2 than f has". This will be particularly interesting for $m = \infty$, in which case u will belong to H^m for all $m = 1, \dots,$ and thus will be in C^∞.

Observe, however, the calculations above do not really constitute a proof. We assumed u was smooth, or at least say C^3, in order to carry out the calculation (3); whereas if we start with merely a weak solution in H_0^1 we cannot immediately justify these computations. We will instead have to rely upon an analysis of certain difference quotients.

The following calculations are often technically difficult, but eventually yield extremely powerful and useful assertions concerning the smoothness of weak solutions. As always, the heart of each computation is the invocation of ellipticity: *the point is to derive analytic estimates from the structural, algebraic assumption of ellipticity.*

6.3.1. Interior regularity.

We as always assume that $U \subset \mathbb{R}^n$ is a bounded, open set. Suppose also $u \in H_0^1(U)$ is a weak solution of the PDE (1), where L has the divergence form

$$Lu = -\sum_{i,j=1}^{n} \left(a^{ij}(x)u_{x_i}\right)_{x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u. \tag{4}$$

We continue to require the uniform ellipticity condition from §6.1.1, and will as necessary make various additional assumptions about the smoothness of the coefficients a^{ij}, b^i, c.

THEOREM 1 (Interior H^2-regularity). *Assume*

$$a^{ij} \in C^1(U),\ b^i, c \in L^\infty(U) \quad (i,j = 1,\dots,n) \tag{5}$$

and

$$f \in L^2(U). \tag{6}$$

Suppose furthermore that $u \in H^1(U)$ is a weak solution of the elliptic PDE

$$Lu = f \quad \text{in } U.$$

Then

$$u \in H^2_{\text{loc}}(U)\ ; \tag{7}$$

and for each open $V \subset\subset U$ we have the estimate

$$\|u\|_{H^2(V)} \le C\left(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}\right), \tag{8}$$

the constant C depending only on V, U, and the coefficients of L.

Remarks. (i) Note carefully that we do not require $u \in H_0^1(U)$; that is, we are not necessarily assuming the boundary condition $u = 0$ on ∂U in the trace sense.

(ii) Observe additionally that since $u \in H^2_{\text{loc}}(U)$, we have

$$Lu = f \quad \text{a.e. in } U.$$

Thus u actually solves the PDE, at least for a.e. point within U. To see this, note that for each $v \in C_c^\infty(U)$, we have

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

Since $u \in H^2_{\text{loc}}(U)$, we can integrate by parts:

$$B[u, v] = (Lu, v).$$

Thus $(Lu - f, v) = 0$ for all $v \in C_c^\infty(U)$, and so $Lu = f$ a.e. □

Proof. 1. Fix any open set $V \subset\subset U$, and choose an open set W such that $V \subset\subset W \subset\subset U$. Then select a smooth function ζ satisfying

$$\begin{cases} \zeta \equiv 1 \text{ on } V, \ \zeta \equiv 0 \text{ on } \mathbb{R}^n - W, \\ 0 \leq \zeta \leq 1. \end{cases}$$

We call ζ a *cutoff* function: its purpose in the subsequent calculations will be to restrict all expressions to the subset W, which is a positive distance away from ∂U. This is necessary as we have no information concerning the behavior of u near ∂U. (As an interesting technical point, notice carefully in the following calculations why we put "ζ^2", and not just "ζ" in (11) below.)

2. Now since u is a weak solution of (1), we have $B[u, v] = (f, v)$ for all $v \in H_0^1(U)$. Consequently

$$\sum_{i,j=1}^n \int_U a^{ij} u_{x_i} v_{x_j} \, dx = \int_U \tilde{f} v \, dx, \tag{9}$$

where

$$\tilde{f} := f - \sum_{i=1}^n b^i u_{x_i} - cu. \tag{10}$$

3. Now let $|h| > 0$ be small, choose $k \in \{1, \ldots, n\}$, and then substitute

$$v := -D_k^{-h}(\zeta^2 D_k^h u) \tag{11}$$

into (9), where as in §5.8.2 $D_k^h u$ denotes the *difference quotient*

$$D_k^h u(x) = \frac{u(x+he_k) - u(x)}{h} \quad (h \in \mathbb{R},\ h \neq 0).$$

We write the resulting expression as

$$A = B, \tag{12}$$

for

$$A := \sum_{i,j=1}^n \int_U a^{ij} u_{x_i} v_{x_j}\, dx \tag{13}$$

and

$$B := \int_U \tilde{f} v\, dx. \tag{14}$$

4. *Estimate of A.* We have

$$\begin{aligned} A &= -\sum_{i,j=1}^n \int_U a^{ij} u_{x_i} \left[D_k^{-h}\left(\zeta^2 D_k^h u\right)\right]_{x_j} dx \\ &= \sum_{i,j=1}^n \int_U D_k^h\left(a^{ij} u_{x_i}\right)\left(\zeta^2 D_k^h u\right)_{x_j} dx \\ &= \sum_{i,j=1}^n \int_U a^{ij,h}\left(D_k^h u_{x_i}\right)\left(\zeta^2 D_k^h u\right)_{x_j} \\ &\qquad + \left(D_k^h a^{ij}\right) u_{x_i}\left(\zeta^2 D_k^h u\right)_{x_j} dx\ . \end{aligned} \tag{15}$$

Here we used the formulas

$$\int_U v D_k^{-h} w\, dx = -\int_U w D_k^h v\, dx \tag{16}$$

and

$$D_k^h(vw) = v^h D_k^h w + w D_k^h v, \tag{17}$$

for $v^h(x) := v(x+he_k)$.

Returning now to (15), we find

$$
\begin{aligned}
A = & \sum_{i,j=1}^{n} \int_U a^{ij,h} D_k^h u_{x_i} D_k^h u_{x_j} \zeta^2 \, dx \\
& + \sum_{i,j=1}^{n} \int_U [a^{ij,h} D_k^h u_{x_i} D_k^h u 2\zeta\zeta_{x_j} + \left(D_k^h a^{ij}\right) u_{x_i} D_k^h u_{x_j} \zeta^2 \\
& \qquad + \left(D_k^h a^{ij}\right) u_{x_i} D_k^h u 2\zeta\zeta_{x_j}] \, dx \\
=: & \, A_1 + A_2.
\end{aligned} \tag{18}
$$

The uniform ellipticity condition implies

$$
A_1 \geq \theta \int_U \zeta^2 |D_k^h Du|^2 dx. \tag{19}
$$

Furthermore we see from (5) that

$$
|A_2| \leq C \int_U \zeta |D_k^h Du| \, |D_k^h u| + \zeta |D_k^h Du| \, |Du| + \zeta |D_k^h u| \, |Du| \, dx,
$$

for some appropriate constant C. But then Cauchy's inequality with ϵ (§B.2) yields the bound

$$
|A_2| \leq \epsilon \int_U \zeta^2 |D_k^h Du|^2 dx + \frac{C}{\epsilon} \int_W |D_k^h u|^2 + |Du|^2 dx.
$$

We choose $\epsilon = \frac{\theta}{2}$ and further recall from Theorem 3,(i) in §5.8.2 the estimate

$$
\int_W |D_k^h u|^2 \, dx \leq C \int_U |Du|^2 \, dx,
$$

thereby obtaining the inequality

$$
|A_2| \leq \frac{\theta}{2} \int_U \zeta^2 |D_k^h Du|^2 dx + C \int_U |Du|^2 dx.
$$

This estimate, (19) and (18) imply finally

$$
A \geq \frac{\theta}{2} \int_U \zeta^2 |D_k^h Du|^2 dx - C \int_U |Du|^2 dx. \tag{20}
$$

5. *Estimate of B.* Recalling now (10), (11), and (14), we estimate

$$
|B| \leq C \int_U (|f| + |Du| + |u|) |v| \, dx. \tag{21}
$$

Now Theorem 3,(i) in §5.8.2 implies

$$\begin{aligned}\int_U |v|^2 dx &\leq C \int_U |D(\zeta^2 D_k^h u)|^2 dx \\ &\leq C \int_W |D_k^h u|^2 + \zeta^2 |D_k^h Du|^2 dx \\ &\leq C \int_U |Du|^2 + \zeta^2 |D_k^h Du|^2 dx.\end{aligned}$$

Thus (21) and Cauchy's inequality with ϵ imply

$$|B| \leq \epsilon \int_U \zeta^2 |D_k^h Du|^2 dx + \frac{C}{\epsilon} \int_U f^2 + u^2 dx + \frac{C}{\epsilon} \int_U |Du|^2 dx.$$

Select $\epsilon = \frac{\theta}{4}$, to obtain

$$(22) \qquad |B| \leq \frac{\theta}{4} \int_U \zeta^2 |D_k^h Du|^2 dx + C \int_U f^2 + u^2 + |Du|^2 dx.$$

6. We finally combine (12), (20) and (22), to discover

$$\int_V |D_k^h Du|^2 dx \leq \int_U \zeta^2 |D_k^h Du|^2 dx \leq C \int_U f^2 + u^2 + |Du|^2 dx$$

for $k = 1, \dots, n$ and all sufficiently small $|h| \neq 0$.

In view of Theorem 3,(ii) in §5.8.2, we deduce $Du \in H^1_{\text{loc}}(U)$, and thus $u \in H^2_{\text{loc}}(U)$, with the estimate

$$(23) \qquad \|u\|_{H^2(V)} \leq C \left(\|f\|_{L^2(U)} + \|u\|_{H^1(U)} \right).$$

7. We now refine estimate (23) by noting that if $V \subset\subset W \subset\subset U$, then the same argument shows

$$(24) \qquad \|u\|_{H^2(V)} \leq C \left(\|f\|_{L^2(W)} + \|u\|_{H^1(W)} \right),$$

for an appropriate constant C depending on V, W, etc. Choose a new cutoff function ζ satisfying

$$\begin{cases} \zeta \equiv 1 \text{ on } W, \ \operatorname{spt} \zeta \subset U, \\ 0 \leq \zeta \leq 1. \end{cases}$$

Now set $v = \zeta^2 u$ in identity (9) and perform elementary calculations, to discover

$$\int_U \zeta^2 |Du|^2 dx \leq C \int_U f^2 + u^2 dx.$$

Thus

$$\|u\|_{H^1(W)} \le C\left(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}\right).$$

This inequality and (24) yield (8). □

Our intention next is to iterate the argument above, thereby deducing our weak solution lies in various higher Sobolev spaces (provided the coefficients are smooth enough and the righthand side lies in sufficiently good spaces).

THEOREM 2 (Higher interior regularity). *Let m be a nonnegative integer, and assume*

(25) $$a^{ij}, b^i, c \in C^{m+1}(U) \quad (i,j = 1,\dots,n)$$

and

(26) $$f \in H^m(U).$$

Suppose $u \in H^1(U)$ is a weak solution of the elliptic PDE

$$Lu = f \quad \text{in } U.$$

Then

(27) $$u \in H^{m+2}_{\mathrm{loc}}(U);$$

and for each $V \subset\subset U$ we have the estimate

(28) $$\|u\|_{H^{m+2}(V)} \le C(\|f\|_{H^m(U)} + \|u\|_{L^2(U)}),$$

the constant C depending only on m, U, V and the coefficients of L.

Proof. 1. We will establish (27), (28) by induction on m, the case $m = 0$ being Theorem 1 above.

2. Assume now assertions (27) and (28) are valid for some nonnegative integer m and all open sets U, coefficients a^{ij}, b^i, c, etc., as above. Suppose then

(29) $$a^{ij}, b^i, c \in C^{m+2}(U),$$

(30) $$f \in H^{m+1}(U),$$

and $u \in H^1(U)$ is a weak solution of $Lu = f$ in U. By the induction hypotheses, we have

$$u \in H^{m+2}_{\text{loc}}(U), \tag{31}$$

with the estimate

$$\|u\|_{H^{m+2}(W)} \leq C(\|f\|_{H^m(U)} + \|u\|_{L^2(U)}), \tag{32}$$

for each $W \subset\subset U$ and an appropriate constant C, depending only on W, the coefficients of L, etc. Fix $V \subset\subset W \subset\subset U$.

3. Now let α be any multiindex with

$$|\alpha| = m+1, \tag{33}$$

and choose any test function $\tilde{v} \in C_c^\infty(W)$. Insert

$$v := (-1)^{|\alpha|} D^\alpha \tilde{v}$$

into the identity $B[u,v] = (f,v)_{L^2(U)}$, and perform some integrations by parts, eventually to discover

$$B[\tilde{u}, \tilde{v}] = (\tilde{f}, \tilde{v}) \tag{34}$$

for

$$\tilde{u} := D^\alpha u \in H^1(W) \tag{35}$$

and

$$\begin{aligned} \tilde{f} := D^\alpha f - \sum_{\substack{\beta \leq \alpha \\ \beta \neq \alpha}} \binom{\alpha}{\beta} \Big[- \sum_{i,j=1}^n (D^{\alpha-\beta} a^{ij} D^\beta u_{x_i})_{x_j} \\ + \sum_{i=1}^n D^{\alpha-\beta} b^i D^\beta u_{x_i} + D^{\alpha-\beta} c D^\beta u \Big]. \end{aligned} \tag{36}$$

Since the identity (34) holds for each $\tilde{v} \in C_c^\infty(W)$, we see that $\tilde{u}$ is a weak solution of

$$L\tilde{u} = \tilde{f} \quad \text{in } W.$$

In view of (29)–(32) and (36), we have $\tilde{f} \in L^2(W)$, with

$$\|\tilde{f}\|_{L^2(W)} \leq C(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}). \tag{37}$$

4. In light of Theorem 1 then, we see $\tilde{u} \in H^2(V)$, with the estimate

$$\begin{aligned}\|\tilde{u}\|_{H^2(V)} &\le C(\|\tilde{f}\|_{L^2(W)} + \|\tilde{u}\|_{L^2(W)}) \\ &\le C(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}).\end{aligned}$$

This inequality holds for each multiindex α with $|\alpha| = m+1$, and $\tilde{u} = D^\alpha u$ as above. Consequently $u \in H^{m+3}(V)$, and

$$\|u\|_{H^{m+3}(V)} \le C(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}).$$

□

We can now repeatedly apply Theorem 2 for $m = 0, 1, 2, \dots$ to deduce the infinite differentiability of u.

THEOREM 3 (Infinite differentiability in the interior). *Assume*

$$a^{ij}, b^i, c \in C^\infty(U) \quad (i,j = 1, \dots, n)$$

and

$$f \in C^\infty(U).$$

Suppose $u \in H^1(U)$ is a weak solution of the elliptic PDE

$$Lu = f \quad \text{in } U.$$

Then

$$u \in C^\infty(U).$$

We are again making no assumptions here about the behavior of u on ∂U. Therefore, in particular, we are asserting that any possible singularities of u on the boundary do not "propagate" into the interior.

Proof. According to Theorem 2, we have $u \in H^m_{\mathrm{loc}}(U)$ for each integer $m = 1, 2, \dots$. Hence Theorem 6 in §5.6.3 implies $u \in C^k(U)$ for each $k = 1, 2, \dots$. □

6.3.2. Boundary regularity.

Now we extend the estimates from §6.3.1 to study the smoothness of weak solutions up to the boundary.

THEOREM 4 (Boundary H^2-regularity). *Assume*

$$a^{ij} \in C^1(\bar{U}), \; b^i, c \in L^\infty(U) \quad (i,j = 1, \dots, n) \tag{38}$$

and

$$f \in L^2(U). \tag{39}$$

Suppose that $u \in H_0^1(U)$ is a weak solution of the elliptic boundary-value problem

$$\begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{40}$$

Assume finally

$$\partial U \text{ is } C^2. \tag{41}$$

Then

$$u \in H^2(U),$$

and we have the estimate

$$\|u\|_{H^2(U)} \leq C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}), \tag{42}$$

the constant C depending only on U and the coefficients of L.

Remarks. (i) If $u \in H_0^1(U)$ is the *unique* weak solution of (40), estimate (42) simplifies to read

$$\|u\|_{H^2(U)} \leq C\|f\|_{L^2(U)}.$$

This follows from Theorem 6 in §6.2.

(ii) Observe also that in contrast to Theorem 1 in §6.3.1, we are now assuming $u = 0$ along ∂U (in the trace sense). □

Proof. 1. We first investigate the special case that U is a half-ball:

$$U = B^0(0,1) \cap \mathbb{R}^n_+. \tag{43}$$

Set $V := B^0(0, \frac{1}{2}) \cap \mathbb{R}^n_+$. Then select a smooth cutoff function ζ satisfying

$$\begin{cases} \zeta \equiv 1 \text{ on } B(0, \frac{1}{2}), \; \zeta \equiv 0 \text{ on } \mathbb{R}^n - B(0,1), \\ 0 \leq \zeta \leq 1. \end{cases}$$

So $\zeta \equiv 1$ on V and ζ vanishes near the *curved* part of ∂U.

2. Since u is a weak solution of (3), we have $B[u,v] = (f,v)$ for all $v \in H_0^1(U)$; consequently

$$\sum_{i,j=1}^{n} \int_U a^{ij} u_{x_i} v_{x_j}\, dx = \int_U \tilde{f} v\, dx, \tag{44}$$

for

$$\tilde{f} := f - \sum_{i=1}^{n} b^i u_{x_i} - cu. \tag{45}$$

3. Now let $h > 0$ be small, choose $k \in \{1, \dots, n-1\}$, and write

$$v := -D_k^{-h}(\zeta^2 D_k^h u).$$

Let us note carefully

$$\begin{aligned} v(x) &= -\frac{1}{h} D_k^{-h}(\zeta^2(x)[u(x+he_k) - u(x)]) \\ &= -\frac{1}{h^2}(\zeta^2(x-he_i)[u(x) - u(x-he_k)] \\ &\qquad - \zeta^2(x)[u(x+he_k) - u(x)]) \end{aligned}$$

if $x \in U$. Now since $u = 0$ along $\{x_n = 0\}$ in the trace sense and $\zeta \equiv 0$ near the curved portion of ∂U, we see $v \in H_0^1(U)$.

We may therefore substitute v into the identity (44), and write the resulting expression as

$$A = B, \tag{46}$$

for

$$A := \sum_{i,j=1}^{n} \int_U a^{ij} u_{x_i} v_{x_j}\, dx \tag{47}$$

and

$$B := \int_U \tilde{f} v\, dx. \tag{48}$$

4. We can now estimate the terms A and B in almost exactly the same way that we estimated their counterparts in the proof of Theorem 1. After some calculations we find

$$A \geq \frac{\theta}{2} \int_U \zeta^2 |D_k^h Du|^2 dx - C \int_U |Du|^2 dx \tag{49}$$

and

$$|B| \le \frac{\theta}{4} \int_U \zeta^2 |D_k^h Du|^2 dx + C \int_U f^2 + u^2 + |Du|^2 dx, \tag{50}$$

for appropriate constants C. We then combine (46), (49), and (50) to discover

$$\int_V |D_k^h Du|^2 dx \le C \int_U f^2 + u^2 + |Du|^2 dx$$

for $k = 1, \dots, n-1$. Thus recalling the Remark after Theorem 3 in §5.8.2, we deduce

$$u_{x_k} \in H^1(V) \quad (k = 1, \dots, n-1),$$

with the estimate

$$\sum_{\substack{k,l=1 \\ k+l<2n}}^{n} \|u_{x_k x_l}\|_{L^2(V)} \le C(\|f\|_{L^2(U)} + \|u\|_{H^1(U)}). \tag{51}$$

5. We must now augment (51) with an estimate of the L^2-norm of $u_{x_n x_n}$ over V. For this we recall from the Remarks after Theorem 1 that $Lu = f$ a.e. in U. Remembering the definition of L, we can rewrite this equality into nondivergence form, as

$$-\sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} \tilde{b}^i u_{x_i} + cu = f, \tag{52}$$

for $\tilde{b}^i := b^i - \sum_{j=1}^n a^{ij}_{x_j}$ $(i = 1, \dots, n)$. So we discover

$$a^{nn} u_{x_n x_n} = -\sum_{\substack{i,j=1 \\ i+j<2n}}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} \tilde{b}^i u_{x_i} + cu - f. \tag{53}$$

Now according to the uniform ellipticity condition, $\sum_{i,j=1}^n a^{ij}(x)\xi_i\xi_j \ge \theta|\xi|^2$ for all $x \in U$, $\xi \in \mathbb{R}^n$. We set $\xi = e_n = (0, \dots, 0, 1)$, to conclude

$$a^{nn}(x) \ge \theta > 0 \tag{54}$$

for all $x \in U$. But then (38), (53) and (54) imply

$$|u_{x_n x_n}| \le C\Big(\sum_{\substack{i,j=1 \\ i+j<2n}}^{n} |u_{x_i x_j}| + |Du| + |u| + |f| \Big) \tag{55}$$

in U. Utilizing this estimate in inequality (51), we conclude $u \in H^2(V)$, and

$$\|u\|_{H^2(V)} \leq C(\|f\|_{L^2(U)} + \|u\|_{H^1(U)}) \tag{56}$$

for some appropriate constant C.

6. We now drop the assumption that U is a half-ball and so has the special form (43). In the general case we choose any point $x^0 \in \partial U$ and note that since ∂U is C^2, we may assume—upon relabeling the coordinate axes if needs be—that

$$U \cap B(x^0, r) = \{x \in B(x^0, r) \mid x_n > \gamma(x_1, \dots, x_{n-1})\}$$

for some $r > 0$ and some C^2 function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$. As usual, we change variables utilizing §C.1 and write

$$y = \mathbf{\Phi}(x), \ x = \mathbf{\Psi}(y). \tag{57}$$

7. Choose $s > 0$ so small that the half-ball $U' := B^0(0, s) \cap \{y_n > 0\}$ lies in $\mathbf{\Phi}(U \cap U(x^0, r))$. Set $V' := B^0(0, s/2) \cap \{y_n > 0\}$. Finally define

$$u'(y) := u(\mathbf{\Psi}(y)) \quad (y \in U'). \tag{58}$$

It is straightforward to check

$$u' \in H^1(U') \tag{59}$$

and

$$u' = 0 \quad \text{on } \partial U' \cap \{y_n = 0\} \tag{60}$$

in the trace sense.

8. We now claim u' is a weak solution of the PDE

$$L'u' = f' \quad \text{in } U', \tag{61}$$

for

$$f'(y) := f(\mathbf{\Psi}(y)) \tag{62}$$

and

$$L'u' := -\sum_{k,l=1}^{n} (a'^{kl} u'_{y_k})_{y_l} + \sum_{k=1}^{n} b'^{k} u'_{y_k} + c'u', \tag{63}$$

where

$$(64)\qquad a'^{kl}(y) := \sum_{r,s=1}^{n} a^{rs}(\Psi(y))\Phi^k_{x_r}(\Psi(y))\Phi^l_{x_s}(\Psi(y)) \quad (k,l=1,\dots,n),$$

$$(65)\qquad b'^{k}(y) := \sum_{r=1}^{n} b^{r}(\Psi(y))\Phi^k_{x_r}(\Psi(y)) \quad (k=1,\dots,n),$$

and

$$(66)\qquad c'(y) := c(\Psi(y))$$

for $y \in U'$, $k,l = 1,\dots,n$.

If $v' \in H^1_0(U')$ and $B'[\ ,\ ]$ denotes the bilinear form associated with the operator L', we have

$$(67)\qquad B'[u',v'] = \int_{U'} \sum_{k,l=1}^{n} a'^{kl} u'_{y_k} v'_{y_l} + \sum_{k=1}^{n} b'^{k} u'_{y_k} v' + c'u'v'\,dy.$$

Now define

$$v(x) := v'(\Phi(x)).$$

Then from (67) we calculate

$$(68)\qquad \begin{aligned} B'[u',v'] &= \sum_{i,j=1}^{n} \sum_{k,l=1}^{n} \int_U a'^{kl} u_{x_i} \Psi^i_{y_k} v_{x_j} \Psi^j_{y_l}\,dy \\ &\quad + \sum_{i=1}^{n}\sum_{k=1}^{n} \int_U b'^{k} u_{x_i} \Psi^i_{y_k} v\,dy + \int_U c'uv\,dy. \end{aligned}$$

Now according to (64), we find for each $i,j = 1,\dots,n$ that

$$\sum_{k,l=1}^{n} a'^{kl} \Psi^i_{y_k} \Psi^j_{y_l} = \sum_{r,s=1}^{n} \sum_{k,l=1}^{n} a^{rs} \Phi^k_{x_r} \Phi^l_{x_s} \Psi^i_{y_k} \Psi^j_{y_l} = a^{ij},$$

since $D\Phi = (D\Psi)^{-1}$. Similarly for $i = 1,\dots,n$, we have

$$\sum_{k=1}^{n} b'^{k} \Psi^i_{y_k} = \sum_{k=1}^{n}\sum_{r=1}^{n} b^{r} \Phi^k_{x_r} \Psi^i_{y_k} = b^i.$$

Substituting these calculations into (68) and changing variables yields, since $|\det D\mathbf{\Phi}| = 1$,

$$
\begin{aligned}
B'[u',v'] &= \int_U \sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv\,dx \\
&= B[u,v] = (f,v)_{L^2(U)} = (f',v')_{L^2(U')}.
\end{aligned}
$$

This establishes (61).

9. We now check that *the operator L' is uniformly elliptic in U'*. Indeed if $y \in U'$ and $\xi \in \mathbb{R}^n$, we note that

$$
\begin{aligned}
\sum_{k,l=1}^n a'^{kl}(y)\xi_k\xi_l &= \sum_{r,s=1}^n \sum_{k,l=1}^n a^{rs}(\mathbf{\Psi}(y))\Phi^k_{x_r}\Phi^l_{x_s}\xi_k\xi_l \\
&= \sum_{r,s=1}^n a^{rs}(\mathbf{\Psi}(y))\eta_r\eta_s \ \geq\ \theta|\eta|^2,
\end{aligned} \tag{69}
$$

where $\eta = \xi D\mathbf{\Phi}$; that is, $\eta_r = \sum_{k=1}^n \Phi^k_{x_r}\xi_k$ $(r = 1,\dots,n)$. But then, since $D\mathbf{\Phi}D\mathbf{\Psi} = I$, we have $\xi = \eta D\mathbf{\Psi}$; and so $|\xi| \leq C|\eta|$ for some constant C. This inequality and (69) imply

$$
\sum_{k,l=1}^n a'^{kl}(y)\xi_k\xi_l \geq \theta'|\xi|^2 \tag{70}
$$

for some $\theta' > 0$ and all $y \in U'$, $\xi \in \mathbb{R}^n$.

Observe also from (64) that the coefficients a'^{kl} are C^1, since $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are C^2.

10. In view of (61) and (70), we may apply the results from steps 1–5 in the proof above to ascertain that $u' \in H^2(V')$, with the bound

$$
\|u'\|_{H^2(V')} \leq C(\|f'\|_{L^2(U')} + \|u'\|_{L^2(U')}).
$$

Consequently

$$
\|u\|_{H^2(V)} \leq C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}) \tag{71}
$$

for $V := \mathbf{\Psi}(V')$.

Since ∂U is compact, we can as usual cover ∂U with finitely many sets $V_1,\dots,V_N$ as above. We sum the resulting estimates, along with the interior estimate, to find $u \in H^2(U)$, with the inequality (42). □

Now we derive higher regularity for our weak solutions, all the way up to ∂U.

THEOREM 5 (Higher boundary regularity). *Let m be a nonnegative integer, and assume*

(72) $$a^{ij}, b^i, c \in C^{m+1}(\bar{U}) \quad (i,j = 1,\dots,n)$$

and

(73) $$f \in H^m(U).$$

Suppose that $u \in H_0^1(U)$ is a weak solution of the boundary-value problem

(74) $$\begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Assume finally

(75) $$\partial U \text{ is } C^{m+2}.$$

Then

(76) $$u \in H^{m+2}(U),$$

and we have the estimate

(77) $$\|u\|_{H^{m+2}(U)} \le C\left(\|f\|_{H^m(U)} + \|u\|_{L^2(U)}\right),$$

the constant C depending only on m, U and the coefficients of L.

Remark. If u is the *unique* solution of (74), then estimate (77) simplifies to read

$$\|u\|_{H^{m+2}(U)} \le C\|f\|_{H^m(U)}.$$

□

Proof. 1. We first investigate the special case

(78) $$U := B^0(0,s) \cap \mathbb{R}^n_+$$

for some $s > 0$. Fix $0 < t < s$ and set $V := B^0(0,t) \cap \mathbb{R}^n_+$.

2. We intend to prove by induction on m that whenever $u = 0$ along $\{x_n = 0\}$ in the trace sense, (72) and (73) imply

(79) $$u \in H^{m+2}(V),$$

with the estimate

(80) $$\|u\|_{H^{m+2}(V)} \le C\left(\|f\|_{H^m(U)} + \|u\|_{L^2(U)}\right),$$

for a constant C depending only on U, V and the coefficients of L. The case $m = 0$ follows as in the proof of Theorem 4 above.

Suppose then

$$a^{ij}, b^i, c \in C^{m+2}(\bar{U}), \tag{81}$$

$$f \in H^{m+1}(U), \tag{82}$$

and u is a weak solution of $Lu = f$ in U, which vanishes in the trace sense along $\{x_n = 0\}$. Fix any $0 < t < r < s$, and write $W := B^0(0,r) \cap \mathbb{R}^n_+$. By the induction assumption we have

$$u \in H^{m+2}(W), \tag{83}$$

with the estimate

$$\|u\|_{H^{m+2}(W)} \le C\left(\|f\|_{H^m(U)} + \|u\|_{L^2(U)}\right). \tag{84}$$

Furthermore according to the interior regularity Theorem 2, $u \in H^{m+3}_{\text{loc}}(U)$.

3. Next, let α be any multiindex with

$$|\alpha| = m+1 \tag{85}$$

and

$$\alpha_n = 0. \tag{86}$$

Then

$$\tilde{u} := D^\alpha u \tag{87}$$

belongs to $H^1(U)$, and vanishes along the plane $\{x_n = 0\}$ in the trace sense. Furthermore, as in the proof of Theorem 2, $\tilde{u}$ is a weak solution of $L\tilde{u} = \tilde{f}$ in U, for

$$\tilde{f} := D^\alpha f - \sum_{\substack{\beta \le \alpha \\ \beta \ne \alpha}} \binom{\alpha}{\beta} \left[\sum_{i,j=1}^n -\left(D^{\alpha-\beta} a^{ij} D^\beta u_{x_i}\right)_{x_j} + \sum_{i=1}^n D^{\alpha-\beta} b^i D^\beta u_{x_i} + D^{\alpha-\beta} c D^\beta u \right].$$

In view of (72), (73), (82) and (84), we see $\tilde{f} \in L^2(W)$, with

(88) $$\|\tilde{f}\|_{L^2(W)} \leq C\left(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}\right).$$

Consequently the proof of Theorem 4 shows $\tilde{u} \in H^2(V)$, with the estimate

$$\begin{aligned}\|\tilde{u}\|_{H^2(V)} &\leq C\left(\|\tilde{f}\|_{L^2(W)} + \|\tilde{u}\|_{L^2(W)}\right)\\ &\leq C\left(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}\right).\end{aligned}$$

In light of (85)–(88), we thus deduce

(89) $$\|D^\beta u\|_{L^2(V)} \leq C\left(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}\right)$$

for any multiindex β with $|\beta| = m+3$ and

(90) $$\beta_n = 0, 1, \text{ or } 2.$$

4. We must extend estimate (89) to remove the restriction (90). For this, let us suppose by induction

(91) $$\|D^\beta u\|_{L^2(V)} \leq C\left(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}\right)$$

for any multiindex β with $|\beta| = m+3$ and

(92) $$\beta_n = 0, 1, \dots, j,$$

for some $j \in \{2, \dots, m+2\}$. Assume then $|\beta| = m+3$,

(93) $$\beta_n = j+1.$$

Let us write $\beta = \gamma + \delta$, for $\delta = (0, \dots, 2)$ and $|\gamma| = m+1$. Since $u \in H^{m+3}_{\text{loc}}(U)$ and $Lu = f$ in U, we have $D^\gamma Lu = D^\gamma f$ a.e. in U. Now

$$\begin{aligned}D^\gamma Lu = a^{nn} D^\beta u + \{ &\text{ sum of terms involving at most } j\\ &\text{derivatives of } u \text{ with respect to } x_n, \text{ and}\\ &\text{at most } m+3 \text{ derivatives in all } \}.\end{aligned}$$

Since $a^{nn} \geq \theta > 0$, we thus find by utilizing (91), (92) that

(94) $$\|D^\beta u\|_{L^2(V)} \leq C\left(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}\right)$$

provided $|\beta| = m+3$ and $\beta_n = j+1$. By induction on j then, we have

$$\|u\|_{H^{m+3}(U)} \leq C\left(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)}\right).$$

This estimate in turn completes the induction on m, begun in step 2.

5. We have now shown that (72) and (73) imply (79) and (80), provided U has the form (78). The general case follows once we straighten out the boundary, using the ideas explained in the proof of Theorem 4. □

We finally iterate the foregoing estimates to obtain

THEOREM 6 (Infinite differentiability up to the boundary). *Assume*

$$a^{ij}, b^i, c \in C^\infty(\bar{U}) \quad (i,j = 1, \dots, n)$$

and

$$f \in C^\infty(\bar{U}).$$

Suppose $u \in H_0^1(U)$ *is a weak solution of the boundary-value problem*

$$\begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Assume also that ∂U *is* C^∞. *Then*

$$u \in C^\infty(\bar{U}).$$

Proof. According to Theorem 5 we have $u \in H^m(U)$ for each integer $m = 1, 2, \dots$. Thus Theorem 6 in §5.6.3 implies $u \in C^k(\bar{U})$ for each $k = 1, 2, \dots$. □

The computations in this section have basically been repeated applications of "energy" methods to higher and higher partial derivatives. The basic tool of integration by parts has eventually taken us from weak solutions (belonging merely to $H_0^1(U)$) to smooth, classical solutions.

6.4. MAXIMUM PRINCIPLES

This section develops the *maximum principle* for second-order elliptic partial differential equations.

Maximum principle methods are based upon the observation that if a C^2 function u attains its maximum over an open set U at a point $x_0 \in U$, then

$$(1) \qquad Du(x_0) = 0, \quad D^2(x_0) \le 0,$$

the latter inequality meaning that the symmetric matrix $D^2u = ((u_{x_i x_j}))$ is nonpositive definite at x_0. Deductions based upon (1) are consequently *pointwise* in character, and are thus utterly different from the integral-based energy methods set forth in §§6.1–6.3.

In particular we will need to require that our solutions u are at least C^2, so that it makes sense to consider the pointwise values of Du, D^2u. (In view of the regularity theory from §6.3 we know however that a weak solution is this smooth, at least provided the coefficients, etc. are sufficiently regular.) As we will shortly learn, it is also most appropriate now to consider elliptic operators L having the *nondivergence form*

(2) $$Lu = -\sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} + cu,$$

where the coefficients a^{ij}, b^i, c are continuous and—as always— the uniform ellipticity condition (4) in §6.1 holds. We continue also to assume, without loss of generality, the symmetry condition $a^{ij} = a^{ji}$ $(i, j = 1, \dots, n)$.

6.4.1. Weak maximum principle.

First, we identify circumstances under which a function must attain its maximum (or minimum) on the boundary. We always assume $U \subset \mathbb{R}^n$ is open, bounded.

THEOREM 1 (Weak maximum principle). *Assume $u \in C^2(U) \cap C(\bar{U})$ and*

$$c \equiv 0 \quad \text{in } U.$$

(i) *If*

(3) $$Lu \leq 0 \quad \text{in } U,$$

then

$$\max_{\bar{U}} u = \max_{\partial U} u.$$

(ii) *If*

(4) $$Lu \geq 0 \quad \text{in } U,$$

then

$$\min_{\bar{U}} u = \min_{\partial U} u.$$

Remark. A function satisfying (3) is called a *subsolution*. We are thus asserting *a subsolution attains its maximum on ∂U*. Similarly, if (4) holds, u is a *supersolution* and attains its minimum on ∂U. □

Proof. 1. Let us first suppose we have the strict inequality

(5) $$Lu < 0 \quad \text{in } U,$$

and yet there exists a point $x_0 \in U$ with

(6) $$u(x_0) = \max_{\bar{U}} u.$$

Now at this maximum point x_0, we have

$$Du(x_0) = 0, \tag{7}$$

and

$$D^2u(x_0) \leq 0. \tag{8}$$

2. Since the matrix $A = ((a^{ij}(x_0)))$ is symmetric and positive definite, there exists an orthogonal matrix $O = ((o_{ij}))$ so that

$$OAO^T = \operatorname{diag}(d_1, \dots, d_n), \quad OO^T = I, \tag{9}$$

with $d_k > 0$ $(k = 1, \dots, n)$. Write $y = x_0 + O(x - x_0)$. Then $x - x_0 = O^T(y - x_0)$, and so

$$u_{x_i} = \sum_{k=1}^n u_{y_k} o_{ik}, \; u_{x_i x_j} = \sum_{k,l=1}^n u_{y_k y_l} o_{ik} o_{jl} \quad (i, j = 1, \dots, n).$$

Hence at the point x_0,

$$\begin{aligned} \sum_{i,j=1}^n a^{ij} u_{x_i x_j} &= \sum_{k,l=1}^n \sum_{i,j=1}^n a^{ij} u_{y_k y_l} o_{ik} o_{jl} \\ &= \sum_{k=1}^n d_k u_{y_k y_k} \quad \text{by (9)} \\ &\leq 0, \end{aligned} \tag{10}$$

since $d_k > 0$ and $u_{y_k y_k}(x_0) \leq 0$ $(k = 1, \dots, n)$, according to (8).

3. Thus at x_0

$$Lu = -\sum_{i,j=1}^n a^{ij} u_{x_i x_j} + \sum_{i=1}^n b^i u_{x_i} \geq 0,$$

in light of (7) and (10). So (5) and (6) are incompatible, and we have a contradiction.

4. In the general case that (3) holds, write

$$u^\epsilon(x) := u(x) + \epsilon e^{\lambda x_1} \quad (x \in U),$$

where $\lambda > 0$ will be selected below and $\epsilon > 0$. Recall (as in the proof of Theorem 4 in §6.3.2) that the uniform ellipticity condition implies $a^{ii}(x) \geq \theta$ $(i = 1, \dots, n, \; x \in U)$. Therefore

$$\begin{aligned} Lu^\epsilon &= Lu + \epsilon L(e^{\lambda x_1}) \\ &\leq \epsilon e^{\lambda x_1} \left[-\lambda^2 a^{11} + \lambda b^1 \right] \\ &\leq \epsilon e^{\lambda x_1} [-\lambda^2 \theta + \|\mathbf{b}\|_{L^\infty} \lambda] \\ &< 0 \quad \text{in } U, \end{aligned}$$

provided we choose $\lambda > 0$ sufficiently large. Then according to steps 1 and 2 above $\max_{\bar{U}} u^\epsilon = \max_{\partial U} u^\epsilon$. Let $\epsilon \to 0$ to find $\max_{\bar{U}} u = \max_{\partial U} u$. This proves (i).

5. Since $-u$ is a subsolution whenever u is a supersolution, assertion (ii) follows. □

We next modify the maximum principle to allow for a *nonnegative* zeroth-order coefficient c. Remember from §A.3 that $u^+ = \max(u, 0)$, $u^- = -\min(u, 0)$.

THEOREM 2 (Weak maximum principle for $c \geq 0$). *Assume $u \in C^2(U) \cap C(\bar{U})$ and*

$$c \geq 0 \quad \textit{in } U.$$

(i) *If*

$$Lu \leq 0 \quad \textit{in } U,$$

then

$$\max_{\bar{U}} u \leq \max_{\partial U} u^+. \tag{11}$$

(ii) *Likewise, if*

$$Lu \geq 0 \quad \textit{in } U,$$

then

$$\min_{\bar{U}} u \geq -\max_{\partial U} u^-. \tag{12}$$

Remark. So in particular, if $Lu = 0$ in U, then

$$\max_{\bar{U}} |u| = \max_{\partial U} |u|. \tag{13}$$

□

Proof. 1. Let u be a subsolution and set $V := \{x \in U \mid u(x) > 0\}$. Then

$$\begin{aligned} Ku &:= Lu - cu \\ &\leq -cu \leq 0 \quad \text{in } V. \end{aligned}$$

The operator K has no zeroth-order term and consequently Theorem 1 implies $\max_{\bar{V}} u = \max_{\partial V} u = \max_{\partial U} u^+$. This gives (11) in the case that $V \neq \emptyset$. Otherwise $u \leq 0$ everywhere in U, and (11) likewise follows.

2. Assertion (ii) follows from (i) applied to $-u$, once we observe that $(-u)^+ = u^-$. □

6.4.2. Strong maximum principle.

We next substantially strengthen the foregoing assertions, by demonstrating that a subsolution u cannot attain its maximum at an interior point of a connected region at all, unless u is constant. This statement is the *strong maximum principle*, which depends on the following subtle analysis of the outer normal derivative $\frac{\partial u}{\partial \nu}$ at a boundary maximum point.

LEMMA (Hopf's Lemma). *Assume $u \in C^2(U) \cap C^1(\bar{U})$ and*

$$c \equiv 0 \quad \text{in } U.$$

Suppose further

$$Lu \leq 0 \quad \text{in } U,$$

and there exists a point $x^0 \in \partial U$ such that

$$u(x^0) > u(x) \quad \text{for all } x \in U. \tag{14}$$

Assume finally that U satisfies the interior ball condition at x^0; that is, there exists an open ball $B \subset U$ with $x^0 \in \partial B$.

(i) *Then*

$$\frac{\partial u}{\partial \nu}(x^0) > 0,$$

where ν is the outer unit normal to B at x^0.

(ii) *If*

$$c \geq 0 \quad \text{in } U,$$

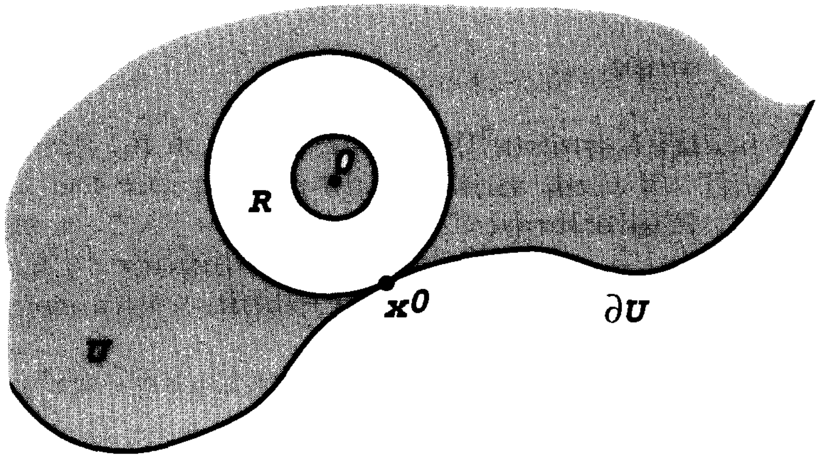
the same conclusion holds provided

$$u(x^0) \geq 0.$$

Remark. The importance of (i) is the *strict* inequality: that $\frac{\partial u}{\partial \nu}(x^0) \geq 0$ is obvious. Note that the interior ball condition automatically holds if ∂U is C^2. □

Proof. 1. Assume $c \geq 0$ and $u(x^0) \geq 0$. We may as well further assume $B = B^0(0,r)$ for some radius $r > 0$. Define

$$v(x) := e^{-\lambda|x|^2} - e^{-\lambda r^2} \quad (x \in B(0,r))$$



for $\lambda > 0$ as selected below. Then using the uniform ellipticity condition, we compute:

$$\begin{aligned} Lv &= -\sum_{i,j=1}^{n} a^{ij} v_{x_i x_j} + \sum_{i=1}^{n} b^i v_{x_i} + cv \\ &= e^{-\lambda|x|^2} \sum_{i,j=1}^{n} a^{ij} \left(-4\lambda^2 x_i x_j + 2\lambda\delta^{ij}\right) \\ &\quad - e^{-\lambda|x|^2} \sum_{i=1}^{n} b^i 2\lambda x_i + c(e^{-\lambda|x|^2} - e^{-\lambda r^2}) \\ &\leq e^{-\lambda|x|^2}(-4\theta\lambda^2|x|^2 + 2\lambda \operatorname{tr} \mathbf{A} + 2\lambda|\mathbf{b}||x| + c), \end{aligned}$$

for $\mathbf{A} = ((a^{ij}))$, $\mathbf{b} = (b^1, \dots, b^n)$. Consider next the open annular region $R := B^0(0,r) - B(0,r/2)$. We have

$$Lv \leq e^{-\lambda|x|^2}(-\theta\lambda^2 r^2 + 2\lambda \operatorname{tr} \mathbf{A} + 2\lambda|\mathbf{b}|r + c) \leq 0 \tag{15}$$

in R, provided $\lambda > 0$ is fixed large enough.

2. In view of (14) there exists a constant $\epsilon > 0$ so small that

$$u(x^0) \geq u(x) + \epsilon v(x) \quad (x \in \partial B(0, r/2)). \tag{16}$$

In addition note

$$u(x^0) \geq u(x) + \epsilon v(x) \quad (x \in \partial B(0, r)), \tag{17}$$

since $v \equiv 0$ on $\partial B(0,r)$.

3. From (15) we see

$$L(u + \epsilon v - u(x^0)) \leq -cu(x^0) \leq 0 \quad \text{in } R,$$

and from (16), (17) we observe

$$u + \epsilon v - u(x^0) \leq 0 \quad \text{on } \partial R.$$

In view of the weak maximum principle, Theorem 1, $u+\epsilon v-u(x^0) \leq 0$ in R. But $u(x^0) + \epsilon v(x^0) - u(x^0) = 0$, and so

$$\frac{\partial u}{\partial \nu}(x^0) + \epsilon \frac{\partial v}{\partial \nu}(x^0) \geq 0.$$

Consequently

$$\frac{\partial u}{\partial \nu}(x^0) \geq -\epsilon \frac{\partial v}{\partial \nu}(x^0) = -\frac{\epsilon}{r} Dv(x^0) \cdot x^0 = 2\lambda \epsilon r e^{-\lambda r^2} > 0,$$

as required. □

Hopf's Lemma is the primary technical tool in the next proof:

THEOREM 3 (Strong maximum principle). *Assume $u \in C^2(U) \cap C(\bar{U})$ and*

$$c \equiv 0 \quad \textit{in } U.$$

Suppose also U is connected, open and bounded.

(i) *If*

$$Lu \leq 0 \quad \textit{in } U$$

and u attains its maximum over $\bar{U}$ at an interior point, then

$$u \textit{ is constant within } U.$$

(ii) *Similarly, if*

$$Lu \geq 0 \quad \textit{in } U$$

and u attains its minimum over $\bar{U}$ at an interior point, then

$$u \textit{ is constant within } U.$$

Proof. Write $M := \max_{\bar{U}} u$ and $C := \{x \in U \mid u(x) = M\}$. Then if $u \not\equiv M$, set

$$V := \{x \in U \mid u(x) < M\}.$$

Choose a point $y \in V$ satisfying $\operatorname{dist}(y, C) < \operatorname{dist}(y, \partial U)$, and let B denote the largest ball with center y whose interior lies in V. Then there exists some point $x^0 \in C$, with $x^0 \in \partial B$. Clearly V satisfies the interior ball condition at x^0; whence Hopf's Lemma, (i), implies $\frac{\partial u}{\partial \nu}(x^0) > 0$. But this is a contradiction: since u attains its maximum at $x^0 \in U$, we have $Du(x^0) = 0$. □

If the zeroth-order term c is *nonnegative*, we have this version of the strong maximum principle:

THEOREM 4 (Strong maximum principle with $c \geq 0$). *Assume $u \in C^2(U) \cap C(\bar{U})$ and*

$$c \geq 0 \quad \textit{in } U.$$

Suppose also U is connected.

(i) *If*

$$Lu \leq 0 \quad \textit{in } U$$

and u attains a nonnegative maximum over $\bar{U}$ at an interior point, then

$$u \textit{ is constant within } U.$$

(ii) *Similarly, if*

$$Lu \geq 0 \quad \textit{in } U$$

and U attains a nonpositive minimum over $\bar{U}$ at an interior point, then

$$u \textit{ is constant within } U.$$

The proof is like that above, except that we use statement (ii) in Hopf's Lemma.

6.4.3. Harnack's inequality.

Harnack's inequality states the values of a nonnegative solution are comparable, at least in any subregion away from the boundary. We assume as usual that

$$Lu = -\sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} + cu.$$

THEOREM 5 (Harnack's inequality). *Assume $u \geq 0$ is a C^2 solution of*

$$Lu = 0 \quad \textit{in } U,$$

and suppose $V \subset\subset U$ is connected. Then there exists a constant C such that

$$\sup_V u \leq C \inf_V u. \tag{18}$$

The constant C depends only on V and the coefficients of L.

If the coefficients are smooth, the proof is a (much easier) special case of the calculations to be presented later for the parabolic Harnack inequality, in §7.1.4. For the general situation of merely bounded and measurable coefficients, see Gilbarg–Trudinger [**G-T**].

6.5. EIGENVALUES AND EIGENFUNCTIONS

We consider in this section the boundary-value problem

$$\begin{cases} Lw = \lambda w & \text{in } U \\ w = 0 & \text{on } \partial U, \end{cases} \tag{1}$$

where U is open, bounded, and recall that λ is an eigenvalue of L provided there exists a nontrivial solution w of (1). From the theory developed in §6.2 we recall that the set Σ of eigenvalues of L is at most countable.

The theorems in §6.5.1 below are analogues for elliptic PDE of the standard linear algebra assertion that a real symmetric matrix possesses real eigenvalues and an orthonormal basis of eigenvectors. Similarly, the results in §6.5.2 are PDE versions of the Perron–Frobenius theorem that a matrix with positive entries has a real, positive eigenvalue and a corresponding eigenvector with positive entries (cf. Gantmacher [**GA**]).

6.5.1. Eigenvalues of symmetric elliptic operators.

For simplicity, we consider now an elliptic operator having the divergence form

$$Lu = -\sum_{i,j=1}^{n} \left(a^{ij} u_{x_i}\right)_{x_j}, \tag{2}$$

where $a^{ij} \in C^\infty(\bar{U})$ $(i, j = 1, \dots, n)$. We suppose the usual uniform ellipticity condition to hold, and as usual suppose

$$a^{ij} = a^{ji} \quad (i, j = 1, \dots, n). \tag{3}$$

The operator L is thus formally symmetric, and in particular the associated bilinear form $B[\ ,\ ]$ satisfies $B[u, v] = B[v, u]$ $(u, v \in H_0^1(U))$. Assume also U is connected.

THEOREM 1 (Eigenvalues of symmetric elliptic operators).

(i) *Each eigenvalue of L is real.*

(ii) *Furthermore, if we repeat each eigenvalue according to its (finite) multiplicity, we have*

$$\Sigma = \{\lambda_k\}_{k=1}^{\infty},$$

where

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots,$$

and

$$\lambda_k \to \infty \qquad \text{as } k \to \infty.$$

(iii) *Finally, there exists an orthonormal basis $\{w_k\}_{k=1}^{\infty}$ of $L^2(U)$, where $w_k \in H_0^1(U)$ is an eigenfunction corresponding to λ_k:*

$$\begin{cases} Lw_k = \lambda_k w_k & \text{in } U \\ w_k = 0 & \text{on } \partial U, \end{cases} \tag{4}$$

for $k = 1, 2, \ldots$.

Remark. Owing to the regularity theory in §6.3, $w_k \in C^\infty(U)$ (and $w_k \in C^\infty(\bar{U})$ if ∂U is smooth), for $k = 1, 2, \ldots$. □

Proof. 1. As in §6.2,

$$S := L^{-1}$$

is a bounded, linear, compact operator mapping $L^2(U)$ into itself.

2. We claim further that S is symmetric. To see this, select $f, g \in L^2(U)$. Then $Sf = u$ means $u \in H_0^1(U)$ is the weak solution of

$$\begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases}$$

and likewise $Sg = v$ means $v \in H_0^1(U)$ solves

$$\begin{cases} Lv = g & \text{in } U \\ v = 0 & \text{on } \partial U \end{cases}$$

in the weak sense. Thus

$$(Sf, g) = (u, g) = B[v, u]$$

and

$$(f, Sg) = (f, v) = B[u, v].$$

Since $B[u,v] = B[v,u]$, we see $(Sf, g) = (f, Sg)$ for all $f, g \in L^2(U)$. Therefore S is symmetric.

3. Notice also

$$(Sf, f) = (u, f) = B[u, u] \geq 0 \quad (f \in L^2(U)).$$

Consequently the theory of compact, symmetric operators from §D.6 implies that all the eigenvalues of S are real, positive, and there are corresponding eigenfunctions which make up an orthonormal basis of $L^2(U)$. But observe as well that for $\eta \neq 0$, we have $Sw = \eta w$ if and only if $Lw = \lambda w$ for $\lambda = \frac{1}{\eta}$. The theorem follows. □

We next scrutinize more carefully the first eigenvalue of L.

DEFINITION. *We call* $\lambda_1 > 0$ *the* principal eigenvalue *of* L.

THEOREM 2 (Variational principle for the principal eigenvalue).
(i) *We have*

$$\lambda_1 = \min\{B[u,u] \mid u \in H_0^1(U), \|u\|_{L^2} = 1\}. \tag{5}$$

(ii) *Furthermore, the above minimum is attained for a function* w_1, *positive within* U, *which solves*

$$\begin{cases} Lw_1 = \lambda_1 w_1 & \text{in } U \\ w_1 = 0 & \text{on } \partial U. \end{cases}$$

(iii) *Finally, if* $u \in H_0^1(U)$ *is any weak solution of*

$$\begin{cases} Lu = \lambda_1 u & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases}$$

then u *is a multiple of* w_1.

Remarks. (i) Assertion (iii) says *the principal eigenvalue* λ_1 *is simple.* In particular

$$0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots.$$

(ii) Expression (5) is *Rayleigh's formula*, and is equivalent to the statement

$$\lambda_1 = \min_{\substack{u \in H_0^1(U) \\ u \not\equiv 0}} \frac{B[u,u]}{\|u\|_{L^2(U)}^2}.$$

□

Proof. 1. In view of (4) we see

$$B[w_k, w_k] = \lambda_k \|w_k\|^2_{L^2(U)} = \lambda_k, \tag{6}$$

and

$$B[w_k, w_l] = \lambda_k (w_k, w_l) = 0 \tag{7}$$

for $k, l = 1, 2, \dots, \ k \neq l$.

2. As $\{w_k\}_{k=1}^\infty$ is an orthonormal basis of $L^2(U)$, if $u \in H_0^1(U)$ and $\|u\|_{L^2(U)} = 1$, we can write

$$u = \sum_{k=1}^\infty d_k w_k \tag{8}$$

for $d_k = (u, w_k)_{L^2(U)}$, the series converging in $L^2(U)$: see §D.2. In addition

$$\sum_{k=1}^\infty d_k^2 = \|u\|^2_{L^2(U)} = 1. \tag{9}$$

3. Furthermore from (6) and (7) we see that $\left\{ \frac{w_k}{\lambda_k^{1/2}} \right\}_{k=1}^\infty$ is an orthonormal subset of $H_0^1(U)$, endowed with the new inner product $B[\ ,\ ]$.

We claim further that $\left\{ \frac{w_k}{\lambda_k^{1/2}} \right\}_{k=1}^\infty$ is in fact an orthonormal basis of $H_0^1(U)$, with this new inner product. To see this, it suffices to verify that

$$B[w_k, u] = 0 \quad (k = 1, 2, \dots)$$

implies $u \equiv 0$. But this assertion is clearly true, since the identities

$$B[w_k, u] = \lambda_k (w_k, u) = 0 \quad (k = 1, \dots)$$

force $u \equiv 0$, as $\{w_k\}_{k=1}^\infty$ is a basis of $L^2(U)$. Consequently

$$u = \sum_{k=1}^\infty \mu_k \frac{w_k}{\lambda_k^{1/2}}$$

for $\mu_k = B\left[u, \frac{w_k}{\lambda_k^{1/2}} \right]$, the series converging in $H_0^1(U)$. But then according to (8), $\mu_k = d_k \lambda_k^{1/2}$; and so the series (8) in fact converges also in $H_0^1(U)$.

4. Thus (6) and (8) imply

$$B[u,u] = \sum_{k=1}^{\infty} d_k^2 \lambda_k \geq \lambda_1 \quad \text{by (9)}.$$

As equality holds for $u = w_1$, we obtain formula (5).

5. We next *claim* that if $u \in H_0^1(U)$ and $\|u\|_{L^2(U)} = 1$, then u is a weak solution of

$$\text{(10)} \qquad \begin{cases} Lu = \lambda_1 u & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases}$$

if only only if

$$\text{(11)} \qquad B[u,u] = \lambda_1.$$

Obviously (10) implies (11). On the other hand, suppose (11) is valid. Then, writing $d_k = (u, w_k)$ as above, we have

$$\text{(12)} \qquad \sum_{k=1}^{\infty} d_k^2 \lambda_1 = \lambda_1 = B[u,u] = \sum_{k=1}^{\infty} d_k^2 \lambda_k.$$

Hence

$$\text{(13)} \qquad \sum_{k=1}^{\infty} (\lambda_k - \lambda_1) d_k^2 = 0.$$

Consequently

$$d_k = (u, w_k) = 0 \quad \text{if} \quad \lambda_k > \lambda_1.$$

Since λ_1 has finite multiplicity, it follows that

$$\text{(14)} \qquad u = \sum_{k=1}^{m} (u, w_k) w_k$$

for some m, where $Lw_k = \lambda_1 w_k$. Therefore

$$\text{(15)} \qquad Lu = \sum_{k=1}^{m} (u, w_k) Lw_k = \lambda_1 u.$$

This proves (10).

6. Next we will show that if $u \in H_0^1(U)$ is a weak solution of (10), $u \not\equiv 0$, then either

$$\text{(16)} \qquad u > 0 \quad \text{in } U$$

or else

$$u < 0 \quad \text{in } U. \tag{17}$$

To see this, let us assume without loss that $\|u\|_{L^2} = 1$, and note

$$\alpha + \beta = 1 \tag{18}$$

for

$$\alpha := \int_U (u^+)^2 dx, \quad \beta := \int_U (u^-)^2 dx.$$

Furthermore since $u^\pm \in H_0^1(U)$, with

$$Du^+ = \begin{cases} Du & \text{a.e. on } \{u \geq 0\} \\ 0 & \text{a.e. on } \{u \leq 0\}, \end{cases}$$
$$Du^- = \begin{cases} 0 & \text{a.e. on } \{u \geq 0\} \\ -Du & \text{a.e. on } \{u \leq 0\} \end{cases}$$

(cf. Problem 17 in Chapter 5), we have $B[u^+, u^-] = 0$. Accordingly

$$\begin{aligned} \lambda_1 = B[u,u] &= B[u^+,u^+] + B[u^-,u^-] \\ &\geq \lambda_1 \|u^+\|^2_{L^2(U)} + \lambda_1 \|u^-\|^2_{L^2(U)} \quad \text{by (5)} \\ &= (\alpha + \beta)\lambda_1 = \lambda_1. \end{aligned}$$

But then we see that the inequality above must in fact be an equality, and so

$$B[u^+,u^+] = \lambda_1 \|u^+\|^2_{L^2(U)}, \; B[u^-,u^-] = \lambda_1 \|u^-\|^2_{L^2(U)}.$$

Therefore the claim proved in step 5 asserts

$$\begin{cases} Lu^+ = \lambda_1 u^+ & \text{in } U \\ u^+ = 0 & \text{on } \partial U \end{cases} \tag{19}$$

and

$$\begin{cases} Lu^- = \lambda_1 u^- & \text{in } U \\ u^- = 0 & \text{on } \partial U \end{cases} \tag{20}$$

in the weak sense.

7. Next, since the coefficients a^{ij} are smooth, we deduce from (19) that $u^+ \in C^\infty(U)$ and

$$Lu^+ = \lambda_1 u^+ \geq 0 \;\text{ in } U.$$

The function u^+ is therefore a supersolution. Thus the strong maximum principle implies either $u^+ > 0$ in U or else $u^+ \equiv 0$ in U. Similar arguments apply to u^-, and so (16) and (17) hold.

8. Finally assume that u and $\tilde{u}$ are two nontrivial weak solutions of (10). In view of steps 6 and 7 above

$$\int_U \tilde{u}\,dx \neq 0,$$

and so there exists a real constant χ such that

$$\int_U u - \chi\tilde{u}\,dx = 0. \tag{21}$$

But since $u - \chi\tilde{u}$ is also a weak solution of (10), steps 6, 7, and the equality (21) imply $u \equiv \chi\tilde{u}$ in U. Hence the eigenvalue λ_1 is simple. □

6.5.2. Eigenvalues of nonsymmetric elliptic operators.

We will now consider a uniformly elliptic operator L in the nondivergence form:

$$Lu = -\sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} + cu.$$

Let us for simplicity assume $a^{ij}, b^i, c \in C^\infty(\bar{U})$, U is open, bounded, connected, and ∂U is smooth. We suppose also $a^{ij} = a^{ji}$ $(i,j = 1,\dots,n)$, and

$$c \geq 0 \quad \text{in } U. \tag{22}$$

Notice however that in general the operator L will not equal its formal adjoint. We therefore *cannot* invoke as above the abstract theory from §D.6. And in fact L will in general have complex eigenvalues and eigenfunctions.

Remarkably, however, the principal eigenvalue of L is real, and the corresponding eigenfunction is of one sign within U.

THEOREM 3 (Principal eigenvalue for nonsymmetric elliptic operators).

(i) *There exists a real eigenvalue λ_1 for the operator L, taken with zero boundary conditions. Furthermore if $\lambda \in \mathbb{C}$ is any other eigenvalue, we have*

$$\operatorname{Re}(\lambda) \leq \lambda_1.$$

(ii) *There exists a corresponding eigenfunction w_1, which is positive within U.*

(iii) *The eigenvalue λ_1 is simple; that is, if u is any solution of* (1), *then u is a multiple of w_1.*

Proof*. 1. Choose $m = [\frac{n}{2}] + 2$ and consider the Banach space $X = H^m(U) \cap H_0^1(U)$. According to Theorem 3 in §5.3.2 $X \subset C^2(\bar{U})$. We define the linear, compact operator $A : X \to X$ by setting $Af = u$, where u is the unique solution of

(23)
$$\begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Next define the cone

$$C = \{u \in X \mid u \geq 0 \text{ in } U\}.$$

According to the maximum principle, $A : C \to C$.

2. Hereafter fix any function $w \in C$, $w \not\equiv 0$. Employing the strong maximum principle and Hopf's Lemma, we deduce

(24)
$$v > 0 \quad \text{in } U, \quad \frac{\partial v}{\partial \nu} < 0 \quad \text{on } \partial U$$

for $v = A(w)$.

Remember that $w = 0$ on ∂U. So in view of (24) there exists a constant $\mu > 0$ so that

(25)
$$\mu v \geq w \qquad \text{in } U.$$

3. Fix $\epsilon > 0$, $\eta > 0$, and consider then the equation

(26)
$$u = \eta A[u + \epsilon w]$$

for the unknown $u \in C$. We claim that

(27) if (26) has a solution u, then $\eta \leq \mu$.

To verify this assertion, suppose in fact $u \in C$ solves (26). We compute

$$u \geq \eta A[\epsilon w] = \eta \epsilon v \geq \frac{\eta}{\mu}\, \epsilon w,$$

according to (25). Hence

$$u \geq \eta A u \geq \frac{\eta^2 \epsilon}{\mu} A w = \frac{\eta^2 \epsilon}{\mu}\, v \geq \left(\frac{\eta}{\mu}\right)^2 \epsilon w.$$

*Omit on first reading.

Continuing, we deduce

$$u \geq \left(\frac{\eta}{\mu}\right)^k \epsilon w \quad (k = 1, \dots),$$

a contradiction unless $\eta \leq \mu$. This observation confirms the assertion (27).

4. Define

$$S_\epsilon := \{u \in C \mid \text{ there exists } 0 \leq \eta \leq 2\mu \text{ such that } u = \eta A[u + \epsilon w]\}.$$

We next assert

$$S_\epsilon \text{ is unbounded in } X. \tag{28}$$

For otherwise we could apply Schaefer's fixed point theorem (to be proved later, as Theorem 4 in §9.2.2), to deduce that the equation

$$u = 2\mu A[u + \epsilon w]$$

has a solution, in contradiction to (27).

5. Owing to (28), there exist

$$0 \leq \eta_\epsilon \leq 2\mu \tag{29}$$

and $v_\epsilon \in C$, with $\|v_\epsilon\|_X \geq 1$, satisfying

$$v_\epsilon = \eta_\epsilon A[v_\epsilon + \epsilon w]. \tag{30}$$

Renormalize by setting

$$u_\epsilon := \frac{v_\epsilon}{\|v_\epsilon\|_X}. \tag{31}$$

Using (29)–(31) and the compactness of the operator A, we obtain a subsequence $\epsilon_k \to 0$ so that

$$\eta_{\epsilon_k} \to \eta \text{ and } u_{\epsilon_k} \to u \text{ in } X.$$

Then (31) implies

$$\|u\|_X = 1, \; u \in C. \tag{32}$$

Since $u_\epsilon = \eta_\epsilon A\left[u_\epsilon + \frac{\epsilon w}{\|v_\epsilon\|_X}\right]$, we deduce in the limit that $u = \eta A u$. In view of (32), $\eta > 0$. We may consequently rewrite the above to read

$$\begin{cases} L w_1 = \lambda_1 w_1 & \text{in } U \\ w_1 = 0 & \text{on } \partial U, \end{cases}$$

for $\lambda_1 = \eta$, $u = w_1$. Thus λ_1 is a real eigenvalue for the operator L, taken with zero boundary conditions, and $w_1 \geq 0$ is a corresponding eigenfunction. In view of the strong maximum principle and Hopf's Lemma, we have

$$w_1 > 0 \text{ in } U, \quad \frac{\partial w_1}{\partial \nu} < 0 \text{ on } \partial U. \tag{33}$$

Additionally, we know w_1 is smooth, owing to the regularity theory in §6.3.

6. All expressions occurring in steps 1–5 above are real. Suppose now $\lambda \in \mathbb{C}$ and u is a complex-valued solution of

$$\begin{cases} Lu = \lambda u & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{34}$$

Now choose any smooth function $w : U \to \mathbb{R}$, with $w > 0$ in U, and set $v := \frac{u}{w}$. We compute

$$\begin{aligned} \lambda v &= \frac{1}{w} L(vw) \quad \text{by (34)} \\ &= Lv - cv - \frac{2}{w} \sum_{i,j=1}^{n} a^{ij} w_{x_j} v_{x_i} + \frac{v}{w} Lw. \end{aligned} \tag{35}$$

Writing

$$Kv := -\sum_{i,j=1}^{n} a^{ij} v_{x_i x_j} + \sum_{i=1}^{n} b_i' v_{x_i}$$

for $b_i' := b^i - \frac{2}{w} \sum_{j=1}^{n} a^{ij} w_{x_j}$ $(1 \leq i \leq n)$, we deduce from (35) that

$$Kv + \left(\frac{Lw}{w} - \lambda \right) v = 0 \quad \text{in } U. \tag{36}$$

Take complex conjugates:

$$K\bar{v} + \left(\frac{Lw}{w} - \bar{\lambda} \right) \bar{v} = 0 \quad \text{in } U. \tag{37}$$

Next we compute

$$K(|v|^2) = K(v\bar{v}) = \bar{v}Kv + vK\bar{v} - 2 \sum_{i,j=1}^{n} a^{ij} v_{x_i} \bar{v}_{x_j} \leq \bar{v}Kv + vK\bar{v}, \tag{38}$$

since

$$\sum_{i,j=1}^{n} a^{ij} \xi_i \bar{\xi}_j = \sum_{i,j=1}^{n} a^{ij} (\mathrm{Re}(\xi_i) \mathrm{Re}(\xi_j) + \mathrm{Im}(\xi_i) \mathrm{Im}(\xi_j)) \geq 0$$

for $\xi \in \mathbb{C}^n$. Combining (36)–(38), we discover

$$K(|v|^2) \leq 2\left(\operatorname{Re}\lambda - \frac{Lw}{w}\right)|v|^2.$$

Now choose

$$w := w_1^{1-\epsilon} \tag{39}$$

for $0 < \epsilon < 1$. Then

$$Lw = \frac{(1-\epsilon)}{w_1^\epsilon} Lw_1 + \frac{\epsilon(1-\epsilon)}{w_1^{1+\epsilon}} \sum_{i,j=1}^n a^{ij} w_{1,x_i} w_{1,x_j} + \epsilon c w_1^{1-\epsilon} \geq (1-\epsilon)\lambda_1 w.$$

Consequently

$$K(|v|^2) \leq 2(\operatorname{Re}\lambda - (1-\epsilon)\lambda_1)|v|^2 \quad \text{in } U.$$

Thus if $\operatorname{Re}(\lambda) \leq (1-\epsilon)\lambda_1$, then $K(|v|^2) \leq 0$ in U. As $v = 0$ on ∂U, according to (33) and (39), we deduce from the maximum principle that $v = \frac{u}{w} = 0$ in U. Thus $u \equiv 0$ in U and so λ cannot be an eigenvalue. This conclusion obtains for each $\epsilon > 0$, and so $\operatorname{Re}\lambda \geq \lambda_1$ if λ is any complex eigenvalue.

7. Finally, let u be any (possibly complex-valued) solution of

$$\begin{cases} Lu = \lambda_1 u & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{40}$$

Since $\operatorname{Re}(u)$ and $\operatorname{Im}(u)$ also solve (40), we may as well suppose from the outset u is real-valued. Replacing u by $-u$ if needs be, we may also suppose $u > 0$ somewhere in U. Now set

$$\chi := \sup\{\mu > 0 \mid w_1 - \mu u \geq 0 \text{ in } U\}. \tag{41}$$

Then $0 < \chi < \infty$. Write $v = w_1 - \chi u$; so that $v \geq 0$ in U and

$$\begin{cases} Lv = \lambda_1 v \geq 0 & \text{in } U \\ v = 0 & \text{on } \partial U. \end{cases}$$

Now if v is not identically zero, the strong maximum principle and Hopf's Lemma imply

$$v > 0 \text{ in } U, \quad \frac{\partial v}{\partial \nu} < 0 \text{ on } \partial U.$$

Thus

$$v - \epsilon u \geq 0 \quad \text{in } U \text{ for some } \epsilon > 0,$$

and so

$$w_1 - (\chi + \epsilon)u \geq 0 \quad \text{in } U,$$

a contradiction to (41). Hence $v \equiv 0$ in U, and so u is a multiple of w_1. □

6.6. PROBLEMS

In the following exercises we assume the coefficients of the various PDE are smooth and satisfy the uniform ellipticity condition. Also $U \subset \mathbb{R}^n$ is always an open, bounded set, with smooth boundary.

1. Let
$$Lu = -\sum_{i,j=1}^{n} \left(a^{ij} u_{x_i}\right)_{x_j} + cu.$$
Prove that there exists a constant $\mu > 0$ such that the corresponding bilinear form $B[\ ,\ ]$ satisfies the hypotheses of the Lax–Milgram Theorem, provided
$$c(x) \geq -\mu \quad (x \in U).$$

2. A function $u \in H_0^2(U)$ is a weak solution of this boundary-value problem for the *biharmonic equation*
$$(*) \qquad \begin{cases} \Delta^2 u = f & \text{in } U \\ u = \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial U \end{cases}$$
provided
$$\int_U \Delta u \Delta v \, dx = \int_U f v \, dx$$
for all $v \in H_0^2(U)$. Given $f \in L^2(U)$, prove that there exists a unique weak solution of $(*)$.

3. Assume U is connected. A function $u \in H^1(U)$ is a weak solution of *Neumann's problem*
$$(*) \qquad \begin{cases} -\Delta u = f & \text{in } U \\ \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial U \end{cases}$$
if
$$\int_U Du \cdot Dv \, dx = \int_U f v \, dx$$
for all $v \in H^1(U)$. Suppose $f \in L^2(U)$. Prove $(*)$ has a weak solution if and only if
$$\int_U f \, dx = 0.$$

4. Let $u \in H^1(\mathbb{R}^n)$ have compact support and be a weak solution of the semilinear PDE
$$-\Delta u + c(u) = f \quad \text{in } \mathbb{R}^n,$$

where $f \in L^2(\mathbb{R}^n)$ and $c : \mathbb{R} \to \mathbb{R}$ is smooth, with $c(0) = 0$ and $c' \geq 0$. Prove $u \in H^2(\mathbb{R}^n)$.

(Hint: Mimic the proof of Theorem 1 in §6.3.1, but without the cutoff function ζ.)

5. Let u be a smooth solution of $Lu = -\sum_{i,j=1}^n a^{ij} u_{x_i x_j} = 0$ in U. Set $v := |Du|^2 + \lambda u^2$. Show that

$$Lv \leq 0 \quad \text{in } U, \text{ if } \lambda \text{ is large enough.}$$

Deduce

$$\|Du\|_{L^\infty(U)} \leq C(\|Du\|_{L^\infty(\partial U)} + \|u\|_{L^\infty(\partial U)}).$$

6. Assume u is a smooth solution of $Lu = -\sum_{i,j=1}^n a^{ij} u_{x_i x_j} = f$ in U, $u = 0$ on ∂U, where f is bounded. Fix $x^0 \in \partial U$. A *barrier* at x^0 is a C^2 function w such that

$$Lw \geq 1 \text{ in } U, \quad w(x^0) = 0, \quad w \geq 0 \text{ on } \partial U.$$

Show that if w is a barrier at x^0, there exists a constant C such that

$$|Du(x^0)| \leq C \left| \frac{\partial w}{\partial \nu}(x^0) \right|.$$

7. Assume U is connected. Use (a) energy methods and (b) the maximum principle to show that the only smooth solutions of the Neumann boundary-value problem

$$\begin{cases} -\Delta u = 0 & \text{in } U \\ \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial U \end{cases}$$

are $u \equiv$ constant.

8. Assume $u \in H^1(U)$ is a bounded weak solution of

$$-\sum_{i,j=1}^n \left(a^{ij} u_{x_i}\right)_{x_j} = 0 \quad \text{in } U.$$

Let $\phi : \mathbb{R} \to \mathbb{R}$ be convex and smooth, and set $w = \phi(u)$. Show w is a weak subsolution; that is,

$$B[w, v] \leq 0$$

for all $v \in H_0^1(U)$, $v \geq 0$.

9. *(Courant minimax principle).* Let $L = -\sum_{i,j=1}^{n}(a^{ij}u_{x_i})_{x_j}$, where $((a^{ij}))$ is symmetric. Assume the operator L, with zero boundary conditions, has eigenvalues $0 < \lambda_1 < \lambda_2 \leq \cdots$. Show

$$\lambda_k = \max_{S \in \Sigma_{k-1}} \min_{\substack{u \in S^\perp \\ \|u\|_{L^2}=1}} B[u,u] \quad (k = 1, 2, \dots).$$

Here Σ_{k-1} denotes the collection of $(k-1)$-dimensional subspaces of $H_0^1(U)$.

10. Let λ_1 be the principal eigenvalue of the uniformly elliptic, nonsymmetric operator

$$Lu = -\sum_{i,j=1}^{n} a^{ij}u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} + cu,$$

taken with zero boundary conditions. Prove the "max-min" representation formula:

$$\lambda_1 = \sup_u \inf_x \frac{Lu(x)}{u(x)},$$

the "sup" taken over functions $u \in C^\infty(\bar{U})$ with $u > 0$ in U, $u = 0$ on ∂U, and the "inf" taken over points $x \in U$. (Hint: Consider the eigenfunction w_1^* corresponding to λ_1 for the adjoint operator L^*.)

6.7. REFERENCES

Section 6.1 See Gilbarg–Trudinger [**G-T,** Chapters 5, 8].

Section 6.3 Consult Gilbarg–Trudinger [**G-T**], Krylov [**KR**] and Ladyzhenskaya–Uraltseva [**L-U**] for more on regularity theory for elliptic PDE.

Section 6.4 Gilbarg–Trudinger [**G-T,** Chapter 3]. Protter and Weinberger [**P-W**] is a good reference for further maximum principle methods.

Section 6.5 See Smoller [**S,** pp. 122–125]. The last part of the proof of Theorem 3 is modified from Protter–Weinberger [**P-W,** §2.8]. O. Hald simplified my proof of Theorem 2.

Chapter 7

LINEAR EVOLUTION EQUATIONS

This long chapter studies various linear partial differential equations that involve time. We often call such PDE *evolution equations*, the idea being that the solution evolves in time from a given initial configuration. We will study by energy methods general second-order parabolic and hyperbolic equations, and also certain first-order hyperbolic systems. The Fourier transform, utilized in §7.3.3, and the semigroup technique, discussed in §7.4, provide alternative approaches.

7.1. SECOND-ORDER PARABOLIC EQUATIONS

Second-order parabolic PDE are natural generalizations of the heat equation (§2.3). We will study in this section the existence and uniqueness of appropriately defined weak solutions, their smoothness and other properties.

7.1.1. Definitions.

a. Parabolic equations.

For this chapter we assume U to be an open, bounded subset of $\mathbb{R}^n$, and as before set $U_T = U \times (0, T]$ for some fixed time $T > 0$.

We will first study the initial/boundary-value problem

$$\text{(1)} \qquad \begin{cases} u_t + Lu = f & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g & \text{on } U \times \{t = 0\}, \end{cases}$$

where $f : U_T \to \mathbb{R}$ and $g : U \to \mathbb{R}$ are given, and $u : \bar{U}_T \to \mathbb{R}$ is the unknown, $u = u(x,t)$. The letter L denotes for each time t a second-order partial differential operator, having either the divergence form

$$\text{(2)} \qquad Lu = -\sum_{i,j=1}^{n} (a^{ij}(x,t)u_{x_i})_{x_j} + \sum_{i=1}^{n} b^i(x,t)u_{x_i} + c(x,t)u$$

or else the nondivergence form

$$\text{(3)} \qquad Lu = -\sum_{i,j=1}^{n} a^{ij}(x,t)u_{x_i x_j} + \sum_{i=1}^{n} b^i(x,t)u_{x_i} + c(x,t)u,$$

for given coefficients a^{ij}, b^i, c $(i, j = 1, \dots, n)$.

DEFINITION. *We say that the partial differential operator* $\frac{\partial}{\partial t} + L$ *is* (uniformly) parabolic *if there exists a constant* $\theta > 0$ *such that*

$$\text{(4)} \qquad \sum_{i,j=1}^{n} a^{ij}(x,t)\xi_i\xi_j \geq \theta|\xi|^2$$

for all $(x,t) \in U_T$, $\xi \in \mathbb{R}^n$.

Remark. Note in particular that for each fixed time $0 \leq t \leq T$ the operator L is a uniformly elliptic operator in the spatial variable x. □

An obvious example is $a^{ij} \equiv \delta_{ij}, b^i \equiv c \equiv f \equiv 0$; in which case $L = -\Delta$ and the PDE $\frac{\partial u}{\partial t} + Lu$ becomes the heat equation. We will see in fact that solutions of the general second-order parabolic PDE are similar in many ways to solutions of the heat equation.

General second-order parabolic equations describe in physical applications the time-evolution of the density of some quantity u, say a chemical concentration, within the region U. As noted for the equilibrium setting (i.e., second-order elliptic PDE, in §6.1.1), the second-order term $\sum_{i,j=1}^{n} a^{ij}u_{x_i x_j}$ describes diffusion, the first-order term $\sum_{i=1}^{n} b^i u_{x_i}$ describes transport, and the zeroth-order term cu describes creation or depletion.

The Fokker–Planck and Kolmogorov equations from the probabilistic study of diffusion processes are also second-order parabolic equations.

b. Weak solutions.

Mimicking the developments in §6.1.2 for elliptic equations, we consider first the case that L has the divergence form (2) and try to find an appropriate notion of weak solution for the initial/boundary-value problem (1). We assume for now that

$$a^{ij}, b^i, c \in L^\infty(U_T) \quad (i,j = 1,\dots,n), \tag{5}$$

$$f \in L^2(U_T), \tag{6}$$

$$g \in L^2(U). \tag{7}$$

We will also always suppose $a^{ij} = a^{ji}$ $(i,j = 1,\dots,n)$.

Let us now define, by analogy with the notation introduced in Chapter 6, the time-dependent bilinear form

$$B[u,v;t] := \int_U \sum_{i,j=1}^{n} a^{ij}(\cdot,t)u_{x_i}v_{x_j} + \sum_{i=1}^{n} b^i(\cdot,t)u_{x_i}v + c(\cdot,t)uv\,dx \tag{8}$$

for $u, v \in H_0^1(U)$ and a.e. $0 \le t \le T$.

Motivation for definition of weak solution. To make plausible the following definition of weak solution, let us first temporarily suppose that $u = u(x,t)$ is in fact a smooth solution of our parabolic problem (1). We now switch our viewpoint, by associating with u a mapping

$$\mathbf{u} : [0,T] \to H_0^1(U)$$

defined by

$$[\mathbf{u}(t)](x) := u(x,t) \quad (x \in U,\ 0 \le t \le T).$$

In other words, we are going to consider u not as a function of x and t together, but rather as a mapping $\mathbf{u}$ of t into the space $H_0^1(U)$ of functions of x. This point of view will greatly clarify the following presentation.

Returning to problem (1), let us similarly define

$$\mathbf{f} : [0,T] \to L^2(U)$$

by

$$[\mathbf{f}(t)](x) := f(x,t) \quad (x \in U,\ 0 \le t \le T).$$

Then if we fix a function $v \in H_0^1(U)$, we can multiply the PDE $\frac{\partial u}{\partial t} + Lu = f$ by v and integrate by parts, to find

$$(9) \qquad (\mathbf{u}', v) + B[\mathbf{u}, v; t] = (\mathbf{f}, v) \quad \left({}' = \frac{d}{dt}\right)$$

for each $0 \le t \le T$, the pairing $(\ ,\ )$ denoting inner product in $L^2(U)$.

Next, observe that

$$(10) \qquad u_t = g^0 + \sum_{j=1}^n g^j_{x_j} \quad \text{in } U_T$$

for $g^0 := f - \sum_{i=1}^n b^i u_{x_i} - cu$ and $g^j := \sum_{i=1}^n a^{ij} u_{x_i}$ $(j = 1, \dots, n)$. Consequently (10) and the definitions from §5.9.1 imply the right hand side of (10) lies in the Sobolev space $H^{-1}(U)$, with

$$\|u_t\|_{H^{-1}(U)} \le \left(\sum_{j=0}^n \|g^j\|^2_{L^2(U)} \right)^{1/2} \le C\left(\|u\|_{H_0^1(U)} + \|f\|_{L^2(U)} \right).$$

This estimate suggests it may be reasonable to look for a weak solution with $\mathbf{u}' \in H^{-1}(U)$ for a.e. time $0 \le t \le T$; in which case the first term in (9) can be reexpressed as $\langle \mathbf{u}', v\rangle$, $\langle\ ,\ \rangle$ being the pairing of $H^{-1}(U)$ and $H_0^1(U)$. □

All these considerations motivate the following

DEFINITION. *We say a function*

$$\mathbf{u} \in L^2(0,T; H_0^1(U)), \ \textit{with } \mathbf{u}' \in L^2(0,T; H^{-1}(U)),$$

is a weak solution *of the parabolic initial/boundary-value problem* (1) *provided*

(i) $\langle \mathbf{u}', v\rangle + B[\mathbf{u}, v; t] = (\mathbf{f}, v)$

for each $v \in H_0^1(U)$ and a.e. time $0 \le t \le T$, and

(ii) $\mathbf{u}(0) = g$.

Remark. In view of Theorem 3 in §5.9.2, we see $\mathbf{u} \in C([0,T]; L^2(U))$, and thus the equality (ii) makes sense. □

7.1.2. Existence of weak solutions.

a. Galerkin approximations.

We intend to build a weak solution of the parabolic problem

(11) $$\begin{cases} u_t + Lu = f & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g & \text{on } U \times \{t=0\} \end{cases}$$

by first constructing solutions of certain finite-dimensional approximations to (11) and then passing to limits. This is called *Galerkin's method.*

More precisely, assume the functions $w_k = w_k(x)$ $(k = 1, \dots)$ are smooth,

(12) $$\{w_k\}_{k=1}^\infty \text{ is an orthogonal basis of } H_0^1(U),$$

and

(13) $$\{w_k\}_{k=1}^\infty \text{ is an orthonormal basis of } L^2(U).$$

(For instance, we could take $\{w_k\}_{k=1}^\infty$ to be the complete set of appropriately normalized eigenfunctions for $L = -\Delta$ in $H_0^1(U)$: see §6.5.1.)

Fix now a positive integer m. We will look for a function $\mathbf{u}_m : [0,T] \to H_0^1(U)$ of the form

(14) $$\mathbf{u}_m(t) := \sum_{k=1}^m d_m^k(t) w_k,$$

where we hope to select the coefficients $d_m^k(t)$ $(0 \le t \le T,\ k = 1, \dots, m)$ so that

(15) $$d_m^k(0) = (g, w_k) \quad (k = 1, \dots, m)$$

and

(16) $$(\mathbf{u}_m', w_k) + B[\mathbf{u}_m, w_k; t] = (\mathbf{f}, w_k) \quad (0 \le t \le T,\ k = 1, \dots, m).$$

(Here, as before, (,) denotes the inner product in $L^2(U)$.)

Thus we seek a function $\mathbf{u}_m$ of the form (14) that satisfies the "projection" (16) of problem (11) onto the finite dimensional subspace spanned by $\{w_k\}_{k=1}^m$.

THEOREM 1 (Construction of approximate solutions). *For each integer $m = 1, 2, \ldots$ there exists a unique function $\mathbf{u}_m$ of the form* (14) *satisfying* (15), (16).

Proof. Assuming $\mathbf{u}_m$ has the structure (14), we first note from (13) that

$$(\mathbf{u}_m'(t), w_k) = d_m^{k\,\prime}(t). \tag{17}$$

Furthermore

$$B[\mathbf{u}_m, w_k; t] = \sum_{l=1}^{m} e^{kl}(t) d_m^l(t), \tag{18}$$

for $e^{kl}(t) := B[w_l, w_k; t]$ $(k, l = 1, \ldots, m)$. Let us further write $f^k(t) := (\mathbf{f}(t), w_k)$ $(k = 1, \ldots, m)$. Then (16) becomes the linear system of ODE

$$d_m^{k\,\prime}(t) + \sum_{l=1}^{m} e^{kl}(t) d_m^l(t) = f^k(t) \quad (k = 1, \ldots, m), \tag{19}$$

subject to the initial conditions (15). According to standard existence theory for ordinary differential equations, there exists a unique absolutely continuous function $\mathbf{d}_m(t) = (d_m^1(t), \ldots, d_m^m(t))$ satisfying (15) and (19) for a.e. $0 \le t \le T$. And then $\mathbf{u}_m$ defined by (14) solves (16) for a.e. $0 \le t \le T$. □

b. Energy estimates.

We propose now to send m to infinity and to show a subsequence of our solutions $\mathbf{u}_m$ of the approximate problems (15), (16) converges to a weak solution of (11). For this we will need some uniform estimates.

THEOREM 2 (Energy estimates). *There exists a constant C, depending only on U, T and the coefficients of L, such that*

$$\max_{0\le t\le T} \|\mathbf{u}_m(t)\|_{L^2(U)} + \|\mathbf{u}_m\|_{L^2(0,T;H_0^1(U))} + \|\mathbf{u}_m'\|_{L^2(0,T;H^{-1}(U))} \le C(\|\mathbf{f}\|_{L^2(0,T;L^2(U))} + \|g\|_{L^2(U)}) \tag{20}$$

for $m = 1, 2, \ldots$.

Proof. 1. Multiply equation (16) by $d_m^k(t)$, sum for $k = 1, \ldots, m$, and then recall (14) to find

$$(\mathbf{u}_m', \mathbf{u}_m) + B[\mathbf{u}_m, \mathbf{u}_m; t] = (\mathbf{f}, \mathbf{u}_m) \tag{21}$$

for a.e. $0 \le t \le T$. We proved in §6.2.2 that there exist constants $\beta > 0$, $\gamma \ge 0$ such that

$$\beta \|\mathbf{u}_m\|^2_{H^1_0(U)} \le B[\mathbf{u}_m, \mathbf{u}_m; t] + \gamma \|\mathbf{u}_m\|^2_{L^2(U)} \tag{22}$$

for all $0 \le t \le T$, $m = 1, \dots$. Furthermore $|(\mathbf{f}, \mathbf{u}_m)| \le \frac{1}{2}\|\mathbf{f}\|^2_{L^2(U)} + \frac{1}{2}\|\mathbf{u}_m\|^2_{L^2(U)}$, and $(\mathbf{u}_m', \mathbf{u}_m) = \frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}_m\|^2_{L^2(U)}\right)$ for a.e. $0 \le t \le T$. Consequently (21) yields the inequality

$$\frac{d}{dt}\left(\|\mathbf{u}_m\|^2_{L^2(U)}\right) + 2\beta \|\mathbf{u}_m\|^2_{H^1_0(U)} \le C_1 \|\mathbf{u}_m\|^2_{L^2(U)} + C_2 \|\mathbf{f}\|^2_{L^2(U)} \tag{23}$$

for a.e. $0 \le t \le T$, and appropriate constants C_1 and C_2.

2. Now write

$$\eta(t) := \|\mathbf{u}_m(t)\|^2_{L^2(U)} \tag{24}$$

and

$$\xi(t) := \|\mathbf{f}(t)\|^2_{L^2(U)}. \tag{25}$$

Then (23) implies

$$\eta'(t) \le C_1 \eta(t) + C_2 \xi(t)$$

for a.e. $0 \le t \le T$. Thus the differential form of Gronwall's inequality (§B.2) yields the estimate

$$\eta(t) \le e^{C_1 t}\left(\eta(0) + C_2 \int_0^t \xi(s)\, ds\right) \quad (0 \le t \le T). \tag{26}$$

Since $\eta(0) = \|\mathbf{u}_m(0)\|^2_{L^2(U)} \le \|g\|^2_{L^2(U)}$ by (15), we obtain from (24)–(26) the estimate

$$\max_{0 \le t \le T} \|\mathbf{u}_m(t)\|^2_{L^2(U)} \le C\left(\|g\|^2_{L^2(U)} + \|\mathbf{f}\|^2_{L^2(0,T;L^2(U))}\right).$$

3. Returning once more to inequality (23), we integrate from 0 to T and employ the inequality above to find

$$\begin{aligned} \|\mathbf{u}_m\|^2_{L^2(0,T;H^1_0(U))} &= \int_0^T \|\mathbf{u}_m\|^2_{H^1_0(U)}\, dt \\ &\le C\left(\|g\|^2_{L^2(U)} + \|\mathbf{f}\|^2_{L^2(0,T;L^2(U))}\right). \end{aligned}$$

4. Fix any $v \in H^1_0(U)$, with $\|v\|_{H^1_0(U)} \le 1$, and write $v = v^1 + v^2$, where $v^1 \in \operatorname{span}\{w_k\}_{k=1}^m$ and $(v^2, w_k) = 0$ $(k = 1, \dots, m)$. Since the functions

$\{w_k\}_{k=0}^{\infty}$ are orthogonal in $H_0^1(U)$, $\|v^1\|_{H_0^1(U)} \leq \|v\|_{H_0^1(U)} \leq 1$. Utilizing (16), we deduce for a.e. $0 \leq t \leq T$ that

$$(\mathbf{u}_m', v^1) + B[\mathbf{u}_m, v^1; t] = (\mathbf{f}, v^1).$$

Then (14) implies

$$\langle \mathbf{u}_m', v \rangle = (\mathbf{u}_m', v) = (\mathbf{u}_m', v^1) = (\mathbf{f}, v^1) - B[\mathbf{u}_m, v^1; t].$$

Consequently

$$|\langle \mathbf{u}_m', v \rangle| \leq C\big(\|\mathbf{f}\|_{L^2(U)} + \|\mathbf{u}_m\|_{H_0^1(U)}\big),$$

since $\|v^1\|_{H_0^1(U)} \leq 1$. Thus

$$\|\mathbf{u}_m'\|_{H^{-1}(U)} \leq C\big(\|\mathbf{f}\|_{L^2(U)} + \|\mathbf{u}_m\|_{H_0^1(U)}\big),$$

and therefore

$$\begin{aligned}\int_0^T \|\mathbf{u}_m'\|_{H^{-1}(U)}^2 \, dt &\leq C \int_0^T \|\mathbf{f}\|_{L^2(U)}^2 + \|\mathbf{u}_m\|_{H_0^1(U)}^2 \, dt \\ &\leq C(\|g\|_{L^2(U)}^2 + \|\mathbf{f}\|_{L^2(0,T;L^2(U))}^2).\end{aligned}$$

□

c. Existence and uniqueness.

Next we pass to limits as $m \to \infty$, to build a weak solution of our initial/boundary-value problem (11).

THEOREM 3 (Existence of weak solution). *There exists a weak solution of* (11).

Proof. 1. According to the energy estimates (20), we see that the sequence $\{\mathbf{u}_m\}_{m=1}^{\infty}$ is bounded in $L^2(0,T;H_0^1(U))$, and $\{\mathbf{u}_m'\}_{m=1}^{\infty}$ is bounded in $L^2(0,T;H^{-1}(U))$.

Consequently there exists a subsequence $\{\mathbf{u}_{m_l}\}_{l=1}^{\infty} \subset \{\mathbf{u}_m\}_{m=1}^{\infty}$ and a function $\mathbf{u} \in L^2(0,T;H_0^1(U))$, with $\mathbf{u}' \in L^2(0,T;H^{-1}(U))$, such that

$$(27) \qquad \begin{cases} \mathbf{u}_{m_l} \rightharpoonup \mathbf{u} & \text{weakly in } L^2(0,T;H_0^1(U)) \\ \mathbf{u}_{m_l}' \rightharpoonup \mathbf{u}' & \text{weakly in } L^2(0,T;H^{-1}(U)). \end{cases}$$

(See §D.4 and Problem 4.)

2. Next fix an integer N and choose a function $\mathbf{v} \in C^1([0,T]; H_0^1(U))$ having the form

(28)
$$\mathbf{v}(t) = \sum_{k=1}^{N} d^k(t) w_k,$$

where $\{d^k\}_{k=1}^N$ are given smooth functions. We choose $m \geq N$, multiply (16) by $d^k(t)$, sum $k = 1, \dots, N$, and then integrate with respect to t to find

(29)
$$\int_0^T \langle \mathbf{u}_m', \mathbf{v} \rangle + B[\mathbf{u}_m, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt.$$

We set $m = m_l$ and recall (27), to find upon passing to weak limits that

(30)
$$\int_0^T \langle \mathbf{u}', \mathbf{v} \rangle + B[\mathbf{u}, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt.$$

This equality then holds for all functions $\mathbf{v} \in L^2(0,T; H_0^1(U))$, as functions of the form (28) are dense in this space. Hence in particular

(31)
$$\langle \mathbf{u}', v \rangle + B[\mathbf{u}, v; t] = (\mathbf{f}, v)$$

for each $v \in H_0^1(U)$ and a.e. $0 \leq t \leq T$. From Theorem 3 in §5.9.2 we see that furthermore $\mathbf{u} \in C([0,T]; L^2(U))$.

3. In order to prove $\mathbf{u}(0) = g$, we first note from (30) that

(32)
$$\int_0^T -\langle \mathbf{v}', \mathbf{u} \rangle + B[\mathbf{u}, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt + (\mathbf{u}(0), \mathbf{v}(0))$$

for each $\mathbf{v} \in C^1([0,T]; H_0^1(U))$ with $\mathbf{v}(T) = 0$. Similarly, from (29) we deduce

(33)
$$\int_0^T -\langle \mathbf{v}', \mathbf{u}_m \rangle + B[\mathbf{u}_m, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt + (\mathbf{u}_m(0), \mathbf{v}(0)).$$

We set $m = m_l$ and once again employ (27) to find

(34)
$$\int_0^T -\langle \mathbf{v}', \mathbf{u} \rangle + B[\mathbf{u}, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt + (g, \mathbf{v}(0)),$$

since $\mathbf{u}_{m_l}(0) \to g$ in $L^2(U)$. As $\mathbf{v}(0)$ is arbitrary, comparing (32) and (34), we conclude $\mathbf{u}(0) = g$. □

THEOREM 4 (Uniqueness of weak solutions). *A weak solution of* (11) *is unique.*

Proof. It suffices to check that the only weak solution of (11) with $\mathbf{f} \equiv g \equiv 0$ is

$$\mathbf{u} \equiv 0. \tag{35}$$

To prove this, observe that by setting $\mathbf{v} = \mathbf{u}$ in identity (31) (for $\mathbf{f} \equiv 0$) we learn, using Theorem 3 in §5.9.2, that

$$\frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}\|^2_{L^2(U)}\right) + B[\mathbf{u},\mathbf{u};t] = \langle \mathbf{u}', \mathbf{u}\rangle + B[\mathbf{u},\mathbf{u};t] = 0. \tag{36}$$

Since

$$B[\mathbf{u},\mathbf{u};t] \geq \beta\|\mathbf{u}\|^2_{H^1_0(U)} - \gamma\|\mathbf{u}\|^2_{L^2(U)} \geq -\gamma\|\mathbf{u}\|^2_{L^2(U)},$$

Gronwall's inequality and (36) imply (35). □

7.1.3. Regularity.

In this section we discuss the regularity of our weak solutions $\mathbf{u}$ to the initial/boundary-value problem for second-order parabolic equations. Our eventual goal is to prove that $\mathbf{u}$ is smooth, provided the coefficients of the PDE, the boundary of the domain, etc. are smooth. The following presentation mirrors that from §6.3.

Motivation: formal derivation of estimates. (i) To gain some insight as to what regularity assertions could possibly be valid, let us temporarily suppose $u = u(x,t)$ is a smooth solution of this initial-value problem for the heat equation:

$$\begin{cases} u_t - \Delta u = f & \text{in } \mathbb{R}^n \times (0,T] \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases} \tag{37}$$

and assume also u goes to zero as $|x| \to \infty$ sufficiently rapidly to justify the following computations. We then calculate for $0 \leq t \leq T$:

$$\begin{aligned} \int_{\mathbb{R}^n} f^2\,dx &= \int_{\mathbb{R}^n} (u_t - \Delta u)^2\,dx \\ &= \int_{\mathbb{R}^n} u_t^2 - 2\Delta u u_t + (\Delta u)^2\,dx \\ &= \int_{\mathbb{R}^n} u_t^2 + 2Du\cdot Du_t + (\Delta u)^2\,dx. \end{aligned} \tag{38}$$

Now $2Du \cdot Du_t = \frac{d}{dt}(|Du|^2)$, and consequently

$$\int_0^t \int_{\mathbb{R}^n} 2Du \cdot Du_t \, dxds = \int_{\mathbb{R}^n} |Du|^2 \, dx \Big|_{s=0}^{s=t}.$$

Furthermore, as demonstrated in §6.3,

$$\int_{\mathbb{R}^n} (\Delta u)^2 \, dx = \int_{\mathbb{R}^n} |D^2 u|^2 \, dx.$$

We utilize the two equalities above in (38) and integrate in time, to obtain

$$\begin{aligned} \sup_{0 \le t \le T} \int_{\mathbb{R}^n} |Du|^2 \, dx + \int_0^T \int_{\mathbb{R}^n} u_t^2 + |D^2 u|^2 \, dxdt \\ \le C \left(\int_0^T \int_{\mathbb{R}^n} f^2 \, dxdt + \int_{\mathbb{R}^n} |Dg|^2 \, dx \right). \end{aligned} \tag{39}$$

We therefore see that we can estimate the L^2-norms of u_t and $D^2 u$ within $\mathbb{R}^n \times (0,T)$, in terms of the L^2-norm of f on $\mathbb{R}^n \times (0,T)$ and the L^2-norm of Dg on $\mathbb{R}^n$.

(ii) Next differentiate the PDE with respect to t and set $\tilde{u} := u_t$. Then

$$\begin{cases} \tilde{u}_t - \Delta \tilde{u} = \tilde{f} & \text{in } \mathbb{R}^n \times (0,T] \\ \tilde{u} = \tilde{g} & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{40}$$

for $\tilde{f} := f_t$, $\tilde{g} := u_t(\cdot, 0) = f(\cdot, 0) + \Delta g$. Multiplying by $\tilde{u}$, integrating by parts and invoking Gronwall's inequality, we deduce:

$$\begin{aligned} \sup_{0 \le t \le T} \int_{\mathbb{R}^n} |u_t|^2 \, dx + \int_0^T \int_{\mathbb{R}^n} |Du_t|^2 \, dxdt \\ \le C \left(\int_0^T \int_{\mathbb{R}^n} f_t^2 \, dxdt + \int_{\mathbb{R}^n} |D^2 g|^2 + f(\cdot, 0)^2 \, dx \right). \end{aligned} \tag{41}$$

But

$$\max_{0 \le t \le T} \|f(\cdot, t)\|_{L^2(\mathbb{R}^n)} \le C(\|f\|_{L^2(\mathbb{R}^n \times (0,T))} + \|f_t\|_{L^2(\mathbb{R}^n \times (0,T))}), \tag{42}$$

according to Theorem 2,(iii) of §5.9.2. Furthermore, writing $-\Delta u = f - u_t$, we find as in §6.3 that

$$\int_{\mathbb{R}^n} |D^2 u|^2 \, dx \le C \int_{\mathbb{R}^n} f^2 + u_t^2 \, dx. \tag{43}$$

Combining (41)–(43) leads us to the estimate

$$\sup_{0\le t\le T}\int_{\mathbb{R}^n}|u_t|^2+|D^2u|^2\,dx+\int_0^T\int_{\mathbb{R}^n}|Du_t|^2\,dxdt \le C\left(\int_0^T\int_{\mathbb{R}^n}f_t^2+f^2\,dxdt+\int_{\mathbb{R}^n}|D^2g|^2\,dx\right), \tag{44}$$

for some constant C. □

The foregoing formal computations suggest that we have estimates corresponding to (39) and (44) for our weak solution to a general second-order parabolic PDE. These calculations do not constitute a proof however, since our weak solution of (11), constructed in §7.1.2, is not smooth enough to justify the foregoing computations.

We will instead calculate using the Galerkin approximations. To streamline the presentation, we hereafter assume that $\{w_k\}_{k=1}^\infty$ is the complete collection of eigenfunctions for $-\Delta$ on $H_0^1(U)$, and that U is bounded, open, with ∂U smooth. We furthermore suppose that

$$\begin{cases}\text{the coefficients } a^{ij}, b^i, c\ (i,j=1,\dots,n) \text{ are smooth on}\\ \bar{U} \text{ and do not depend on } t.\end{cases} \tag{45}$$

THEOREM 5 (Improved regularity).

(i) *Assume*

$$g\in H_0^1(U),\ \ \mathbf{f}\in L^2(0,T;L^2(U)).$$

Suppose also $\mathbf{u}\in L^2(0,T;H_0^1(U))$, *with* $\mathbf{u}'\in L^2(0,T;H^{-1}(U))$, *is the weak solution of*

$$\begin{cases} u_t+Lu=f & \text{in } U_T\\ u=0 & \text{on } \partial U\times[0,T]\\ u=g & \text{on } U\times\{t=0\}.\end{cases}$$

Then in fact

$$\mathbf{u}\in L^2(0,T;H^2(U))\cap L^\infty(0,T;H_0^1(U)),\ \ \mathbf{u}'\in L^2(0,T;L^2(U)),$$

and we have the estimate

$$\operatorname*{ess\,sup}_{0\le t\le T}\|\mathbf{u}(t)\|_{H_0^1(U)}+\|\mathbf{u}\|_{L^2(0,T;H^2(U))}+\|\mathbf{u}'\|_{L^2(0,T;L^2(U))} \le C\left(\|\mathbf{f}\|_{L^2(0,T;L^2(U))}+\|g\|_{H_0^1(U)}\right), \tag{46}$$

the constant C depending only on U,T and the coefficients of L.

(ii) *If, in addition,*

$$g \in H^2(U), \ \mathbf{f}' \in L^2(0,T;L^2(U)),$$

then

$$\mathbf{u} \in L^\infty(0,T;H^2(U)), \ \mathbf{u}' \in L^\infty(0,T;L^2(U)) \cap L^2(0,T;H_0^1(U)),$$
$$\mathbf{u}'' \in L^2(0,T;H^{-1}(U)),$$

with the estimate

$$\begin{aligned} &\operatorname*{ess\,sup}_{0\leq t\leq T}(\|\mathbf{u}(t)\|_{H^2(U)} + \|\mathbf{u}'(t)\|_{L^2(U)}) + \|\mathbf{u}'\|_{L^2(0,T;H_0^1(U))} \\ &+\|\mathbf{u}''\|_{L^2(0,T;H^{-1}(U))} \leq C\big(\|\mathbf{f}\|_{H^1(0,T;L^2(U))} + \|g\|_{H^2(U)}\big). \end{aligned} \tag{47}$$

Remark. Assertions (i), (ii) of Theorem 5 are precise versions of the formal estimates (39), (44) (for the heat equation on $U = \mathbb{R}^n$). □

Proof. 1. Fixing $m \geq 1$, we multiply equation (16) in §7.1.2 by ${d_m^k}'(t)$ and sum $k = 1, \ldots, m$, to discover

$$(\mathbf{u}_m', \mathbf{u}_m') + B[\mathbf{u}_m, \mathbf{u}_m'] = (\mathbf{f}, \mathbf{u}_m')$$

for a.e. $0 \leq t \leq T$. Now

$$\begin{aligned} B[\mathbf{u}_m, \mathbf{u}_m'] &= \int_U \sum_{i,j=1}^n a^{ij} \mathbf{u}_{m,x_i} \mathbf{u}_{m,x_j}' \, dx \\ &\quad + \int_U \sum_{i=1}^n b^i \mathbf{u}_{m,x_i} \mathbf{u}_m' + c\mathbf{u}_m \mathbf{u}_m' \, dx \\ &=: A + B. \end{aligned}$$

Since $a^{ij} = a^{ji}$ $(i,j = 1, \ldots, n)$ and these coefficients do not depend on t, we see $A = \frac{d}{dt}\big(\frac{1}{2}A[\mathbf{u}_m, \mathbf{u}_m]\big)$, for the symmetric bilinear form

$$A[u,v] := \int_U \sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} \, dx \quad (u, v \in H_0^1(U)).$$

Furthermore,

$$|B| \leq \frac{C}{\epsilon}\|\mathbf{u}_m\|_{H_0^1(U)}^2 + \epsilon\|\mathbf{u}_m'\|_{L^2(U)}^2, \ |(\mathbf{f}, \mathbf{u}_m')| \leq \frac{C}{\epsilon}\|\mathbf{f}\|_{L^2(U)}^2 + \epsilon\|\mathbf{u}_m'\|_{L^2(U)}^2$$

for each $\epsilon > 0$.

2. Combining the above inequalities, we deduce

$$\|\mathbf{u}_m'\|^2_{L^2(U)} + \frac{d}{dt}\left(\frac{1}{2}A[\mathbf{u}_m, \mathbf{u}_m]\right)$$
$$\leq \frac{C}{\epsilon}(\|\mathbf{u}_m\|^2_{H^1_0(U)} + \|\mathbf{f}\|^2_{L^2(U)}) + 2\epsilon\|\mathbf{u}_m'\|^2_{L^2(U)}.$$

Choosing $\epsilon = \frac{1}{4}$ and integrating, we find

$$\int_0^T \|\mathbf{u}_m'\|^2_{L^2(U)}dt + \sup_{0\leq t\leq T} A[\mathbf{u}_m(t), \mathbf{u}_m(t)]$$
$$\leq C\left(A[\mathbf{u}_m(0), \mathbf{u}_m(0)] + \int_0^T \|\mathbf{u}_m\|^2_{H^1_0(U)} + \|\mathbf{f}\|^2_{L^2(U)}dt\right)$$
$$\leq C(\|g\|^2_{H^1_0(U)} + \|\mathbf{f}\|^2_{L^2(0,T;L^2(U))}),$$

according to Theorem 2 in §7.1.2, where we estimated $\|\mathbf{u}_m(0)\|_{H^1_0(U)} \leq \|g\|_{H^1_0(U)}$. As $A[u,u] \geq \theta \int_U |Du|^2 dx$ for each $u \in H^1_0(U)$, we find that

$$\sup_{0\leq t\leq T} \|\mathbf{u}_m(t)\|^2_{H^1_0(U)} \leq C(\|g\|^2_{H^1_0(U)} + \|\mathbf{f}\|^2_{L^2(0,T;L^2(U))}). \tag{48}$$

Passing to limits as $m = m_l \to \infty$, we deduce $\mathbf{u} \in L^\infty(0,T;H^1_0(U))$, $\mathbf{u}' \in L^2(0,T;L^2(U))$, with the stated bounds; cf. Problem 5.

3. In particular, for a.e. t we have the identity

$$(\mathbf{u}', v) + B[\mathbf{u}, v] = (\mathbf{f}, v)$$

for each $v \in H^1_0(U)$. This equality we rewrite to read

$$B[\mathbf{u}, v] = (\mathbf{h}, v)$$

for $\mathbf{h} := \mathbf{f} - \mathbf{u}'$. Since $\mathbf{h}(t) \in L^2(U)$ for a.e. $0 \leq t \leq T$, we deduce from the elliptic regularity Theorem 4 in §6.3.2 that $\mathbf{u}(t) \in H^2(U)$ for a.e. $0 \leq t \leq T$, with the estimate

$$\|\mathbf{u}\|^2_{H^2(U)} \leq C(\|\mathbf{h}\|^2_{L^2(U)} + \|\mathbf{u}\|^2_{L^2(U)})$$
$$\leq C(\|\mathbf{f}\|^2_{L^2(U)} + \|\mathbf{u}'\|^2_{L^2(U)} + \|\mathbf{u}\|^2_{L^2(U)}). \tag{49}$$

Integrating and utilizing the estimates from step 2, we complete the proof of (i).

4. The goal next is to establish higher regularity for our weak solution. So now suppose $g \in H^2(U) \cap H^1_0(U)$, $\mathbf{f} \in H^1(0,T;L^2(U))$. Fix $m \geq 1$ and

differentiate equation (16) in §7.1.2 with respect to t. Owing to (45), we find

$$(\tilde{\mathbf{u}}_m', w_k) + B[\tilde{\mathbf{u}}_m, w_k] = (\mathbf{f}', w_k) \quad (k = 1, \dots, m), \tag{50}$$

where $\tilde{\mathbf{u}}_m := \mathbf{u}_m'$. Multiply (50) by $d_m^{k'}(t)$ and sum $k = 1, \dots, m$:

$$(\tilde{\mathbf{u}}_m', \tilde{\mathbf{u}}_m) + B[\tilde{\mathbf{u}}_m, \tilde{\mathbf{u}}_m] = (\mathbf{f}', \tilde{\mathbf{u}}_m).$$

Employing Gronwall's inequality, we deduce

$$\begin{aligned} \sup_{0 \le t \le T} \|\mathbf{u}_m'(t)\|_{L^2(U)}^2 + \int_0^T \|\mathbf{u}_m'\|_{H_0^1(U)}^2 \, dt & \\ \le C(\|\mathbf{u}_m'(0)\|_{L^2(U)}^2 &+ \|\mathbf{f}'\|_{L^2(0,T;L^2(U))}^2) \\ \le C(\|\mathbf{f}\|_{H^1(0,T;L^2(U))}^2 &+ \|\mathbf{u}_m(0)\|_{H^2(U)}^2). \end{aligned} \tag{51}$$

We employed (16) in the last inequality.

5. We must estimate the last term in (51). Remember that we have taken $\{w_k\}_{k=1}^\infty$ to be the complete collection of (smooth) eigenfunctions for $-\Delta$ on $H_0^1(U)$. In particular, $\Delta \mathbf{u}_m = 0$ on ∂U. Thus

$$\|\mathbf{u}_m(0)\|_{H^2(U)}^2 \le C\|\Delta \mathbf{u}_m(0)\|_{L^2(U)}^2 = C(\mathbf{u}_m(0), \Delta^2 \mathbf{u}_m(0)).$$

Since $\Delta^2 \mathbf{u}_m(0) \in \operatorname{span}\{w_k\}_{k=1}^m$ and $(\mathbf{u}_m(0), w_k) = (g, w_k)$ for $k = 1, \dots, m$, we have

$$\begin{aligned} \|\mathbf{u}_m(0)\|_{H^2(U)}^2 &\le C(g, \Delta^2 \mathbf{u}_m(0)) = C(\Delta g, \Delta \mathbf{u}_m(0)) \\ &\le \frac{1}{2}\|\mathbf{u}_m(0)\|_{H^2(U)}^2 + C\|g\|_{H^2(U)}^2. \end{aligned}$$

Hence $\|\mathbf{u}_m(0)\|_{H^2(U)} \le C\|g\|_{H^2(U)}$. Therefore (51) implies

$$\begin{aligned} \sup_{0 \le t \le T} \|\mathbf{u}_m'(t)\|_{L^2(U)}^2 + \int_0^T \|\mathbf{u}_m'\|_{H_0^1(U)}^2 \, dt & \\ \le C(\|\mathbf{f}\|_{H^1(0,T;L^2(U))}^2 &+ \|g\|_{H^2(U)}^2). \end{aligned} \tag{52}$$

6. Now

$$B[\mathbf{u}_m, w_k] = (\mathbf{f} - \mathbf{u}_m', w_k) \quad (k = 1, \dots, m).$$

Let λ_k denote the k^{th} eigenvalue of $-\Delta$ on $H_0^1(U)$. Multiplying the above identity by $\lambda_k d_m^k(t)$ and summing $k = 1, \dots, m$, we deduce for $0 \le t \le T$ that

$$B[\mathbf{u}_m, -\Delta \mathbf{u}_m] = (\mathbf{f} - \mathbf{u}_m', -\Delta \mathbf{u}_m). \tag{53}$$

Since $\Delta \mathbf{u}_m = 0$ on ∂U, we see $B[\mathbf{u}_m, -\Delta \mathbf{u}_m] = (L\mathbf{u}_m, -\Delta \mathbf{u}_m)$. Next we invoke the inequality

$$\beta \|u\|^2_{H^2(U)} \leq (Lu, -\Delta u) + \gamma \|u\|^2_{L^2(U)} \quad (u \in H^2(U) \cap H^1_0(U)) \tag{54}$$

for constants $\beta > 0$, $\gamma \geq 0$. See Problem 8 and also the Remark following the proof.

We thereupon conclude from (53) that

$$\|\mathbf{u}_m\|_{H^2(U)} \leq C(\|\mathbf{f}\|_{L^2(U)} + \|\mathbf{u}'_m\|_{L^2(U)} + \|\mathbf{u}_m\|_{L^2(U)}).$$

This inequality, (52), (48) and Theorem 3 in §5.9.2 imply

$$\sup_{0\leq t\leq T} (\|\mathbf{u}'_m(t)\|^2_{L^2(U)} + \|\mathbf{u}_m(t)\|^2_{H^2(U)}) + \int_0^T \|\mathbf{u}'_m\|^2_{H^1_0(U)}\, dt \\ \leq C(\|\mathbf{f}\|^2_{H^1(0,T;L^2(U))} + \|g\|^2_{H^2(U)}).$$

Passing to limits as $m = m_l \to \infty$, we deduce the same bound for $\mathbf{u}$.

7. It remains to show $\mathbf{u}'' \in L^2(0,T;H^{-1}(U))$. To do so, take $v \in H^1_0(U)$, with $\|v\|_{H^1_0(U)} \leq 1$, and write $v = v^1 + v^2$, as in the proof of Theorem 2 in §7.1.2. Then for a.e. time $0 \leq t \leq T$:

$$\langle \mathbf{u}''_m, v\rangle = (\mathbf{u}''_m, v) = (\mathbf{u}''_m, v^1) = (\mathbf{f}', v^1) - B[\mathbf{u}'_m, v^1; t]$$

according to (50), since $\mathbf{u}''_m = \tilde{\mathbf{u}}'_m$. Consequently

$$|\langle \mathbf{u}''_m, v\rangle| \leq C\big(\|\mathbf{f}'\|_{L^2(U)} + \|\mathbf{u}'_m\|_{H^1_0(U)}\big),$$

since $\|v^1\|_{H^1_0(U)} \leq 1$. Thus

$$\|\mathbf{u}''_m\|_{H^{-1}(U)} \leq C\big(\|\mathbf{f}'\|_{L^2(U)} + \|\mathbf{u}'_m\|_{H^1_0(U)}\big),$$

and so $\mathbf{u}''_m$ is bounded in $L^2(0,T;H^{-1}(U))$. Passing to limits, we deduce that $\mathbf{u}'' \in L^2(0,T;H^{-1}(U))$, with the stated estimate. □

Remark. If L were symmetric, we could alternatively have taken $\{w_k\}_{k=1}^\infty$ to be a basis of eigenfunctions of L on $H^1_0(U)$, and so avoided the need for inequality (54). □

Let us now build upon the previous regularity assertion:

THEOREM 6 (Higher regularity). *Assume*

$$g \in H^{2m+1}(U), \ \frac{d^k \mathbf{f}}{dt^k} \in L^2(0,T; H^{2m-2k}(U)) \quad (k = 0, \dots, m).$$

Suppose also the following m^{th}*-order* compatibility conditions *hold:*

$$\begin{cases} g_0 := g \in H_0^1(U), \ g_1 := \mathbf{f}(0) - Lg_0 \in H_0^1(U), \\ \dots, \ g_m := \frac{d^{m-1}\mathbf{f}}{dt^{m-1}}(0) - Lg_{m-1} \in H_0^1(U). \end{cases}$$

Then

$$\frac{d^k \mathbf{u}}{dt^k} \in L^2(0,T; H^{2m+2-2k}(U)) \quad (k = 0, \dots, m+1);$$

and we have the estimate

$$\begin{aligned} \sum_{k=0}^{m+1} \left\| \frac{d^k \mathbf{u}}{dt^k} \right\|_{L^2(0,T;H^{2m+2-2k}(U))} \\ \le C \left(\sum_{k=0}^{m} \left\| \frac{d^k \mathbf{f}}{dt^k} \right\|_{L^2(0,T;H^{2m-2k}(U))} + \|g\|_{H^{2m+1}(U)} \right), \end{aligned} \tag{55}$$

the constant C *depending only on* m*,* U*,* T *and the coefficients of* L*.*

Remark. Taking into account Theorem 4 in §5.9.2, we see that

$$\mathbf{f}(0) \in H^{2m-1}(U), \ \mathbf{f}'(0) \in H^{2m-3}(U), \ \dots, \ \mathbf{f}^{(m-1)}(0) \in H^1(U),$$

and consequently

$$g_0 \in H^{2m+1}(U), \ g_1 \in H^{2m-1}(U), \ \dots, \ g_m \in H^1(U).$$

The compatibility conditions are consequently the requirements that, in addition, each of these functions equals 0 on ∂U, in the trace sense. □

Proof. 1. The proof is an induction on m, the case $m = 0$ being Theorem 5,(i) above.

2. Assume now the theorem is valid for some nonnegative integer m, and suppose then

$$g \in H^{2m+3}(U), \ \frac{d^k \mathbf{f}}{dt^k} \in L^2(0,T; H^{2m+2-2k}(U)) \quad (k = 0, \dots, m+1), \tag{56}$$

and the $(m+1)^{th}$-order compatibility conditions hold. Now set $\tilde{\mathbf{u}} := \mathbf{u}'$. Differentiating the PDE with respect to t, we check that $\tilde{\mathbf{u}}$ is the unique weak solution of

$$\begin{cases} \tilde{u}_t + L\tilde{u} = \tilde{f} & \text{in } U_T \\ \tilde{u} = 0 & \text{on } \partial U \times [0,T] \\ \tilde{u} = \tilde{g} & \text{on } U \times \{t = 0\}, \end{cases} \tag{57}$$

for $\tilde{f} := f_t$, $\tilde{g} := f(\cdot,0) - Lg$. In particular, for $m = 0$ we rely upon Theorem 5,(ii) to be sure that $\tilde{\mathbf{u}} \in L^2(0,T;H_0^1(U))$, $\tilde{\mathbf{u}}' \in L^2(0,T;H^{-1}(U))$.

Since f and g satisfy the $(m+1)^{th}$-order compatibility conditions, it follows that $\tilde{f}$ and $\tilde{g}$ satisfy the m^{th}-order compatibility condition. Thus applying the induction assumption, we deduce

$$\frac{d^k\tilde{\mathbf{u}}}{dt^k} \in L^2(0,T;H^{2m+2-2k}(U)) \quad (k = 0,\dots,m+1)$$

and

$$\sum_{k=0}^{m+1} \left\| \frac{d^k\tilde{\mathbf{u}}}{dt^k} \right\|_{L^2(0,T;H^{2m+2-2k}(U))} \leq C\left(\sum_{k=0}^{m} \left\| \frac{d^k\tilde{\mathbf{f}}}{dt^k} \right\|_{L^2(0,T;H^{2m-2k}(U))} + \|\tilde{g}\|_{H^{2m+1}(U)} \right)$$

for $\tilde{\mathbf{f}} := \mathbf{f}'$. Since $\tilde{\mathbf{u}} = \mathbf{u}'$, we can rewrite the foregoing:

$$\frac{d^k\mathbf{u}}{dt^k} \in L^2(0,T;H^{2m+4-2k}(U)) \quad (k = 1,\dots,m+2),$$

$$\begin{aligned} (58) \qquad & \sum_{k=1}^{m+2} \left\| \frac{d^k\mathbf{u}}{dt^k} \right\|_{L^2(0,T;H^{2m+4-2k}(U))} \\ & \leq C\left(\sum_{k=1}^{m+1} \left\| \frac{d^k\mathbf{f}}{dt^k} \right\|_{L^2(0,T;H^{2m+2-2k}(U))} + \|\mathbf{f}(0)\|_{H^{2m+1}(U)} + \|Lg\|_{H^{2m+1}(U)} \right) \\ & \leq C\left(\sum_{k=0}^{m+1} \left\| \frac{d^k\mathbf{f}}{dt^k} \right\|_{L^2(0,T;H^{2m+2-2k}(U))} + \|g\|_{H^{2m+3}(U)} \right). \end{aligned}$$

Here we used the estimate

$$(59) \qquad \|\mathbf{f}(0)\|_{H^{2m+1}(U)} \leq C(\|\mathbf{f}\|_{L^2(0,T;H^{2m+2}(U))} + \|\mathbf{f}'\|_{L^2(0,T;H^{2m}(U))}),$$

which follows from Theorem 4 in §5.9.2.

3. Now write for a.e. $0 \leq t \leq T$: $L\mathbf{u} = \mathbf{f} - \mathbf{u}' =: \mathbf{h}$. According to Theorem 5 in §6.3.2, we have

$$\begin{aligned} \|\mathbf{u}\|_{H^{2m+4}(U)} & \leq C(\|\mathbf{h}\|_{H^{2m+2}(U)} + \|\mathbf{u}\|_{L^2(U)}) \\ & \leq C(\|\mathbf{f}\|_{H^{2m+2}(U)} + \|\mathbf{u}'\|_{H^{2m+2}(U)} + \|\mathbf{u}\|_{L^2(U)}). \end{aligned}$$

Integrating with respect to t from 0 to T and adding the resulting expression to (58), we deduce

$$
\begin{aligned}
&\sum_{k=0}^{m+2}\left\|\frac{d^k\mathbf{u}}{dt^k}\right\|_{L^2(0,T;H^{2m+4-2k}(U))} \\
(60)\qquad &\leq C\left(\sum_{k=0}^{m+1}\left\|\frac{d^k\mathbf{f}}{dt^k}\right\|_{L^2(0,T;H^{2m+2-2k}(U))} + \|g\|_{H^{2m+3}(U)} + \|\mathbf{u}\|^2_{L^2(0,T;L^2(U))}\right).
\end{aligned}
$$

Since

$$\|\mathbf{u}\|_{L^2(0,T;L^2(U))} \leq C(\|\mathbf{f}\|_{L^2(0,T;L^2(U))} + \|g\|_{L^2(U)}),$$

we thereby obtain the assertion of the theorem for $m+1$. □

THEOREM 7 (Infinite differentiability). *Assume*

$$g \in C^\infty(\bar{U}),\ f \in C^\infty(\bar{U}_T),$$

and the m^{th}-order compatibility conditions hold for $m = 0, 1, \dots$. Then the parabolic initial/boundary-value problem (11) *has a unique solution*

$$u \in C^\infty(\bar{U}_T).$$

Proof. Apply Theorem 6 for $m = 0, 1, \dots$. □

As we did for elliptic operators in Chapter 6, we have succeeded in repeatedly applying fairly straightforward "energy" estimates to produce a smooth solution of our parabolic initial/boundary-value problem (1). This assertion requires the compatibility conditions (53) hold for all m, and it is easy to see that these conditions are necessary for the existence of a solution smooth on all of $\bar{U}_T$.

Remark. Interior estimates, analogous to those developed for elliptic PDE in §6.3.1, can also be derived, and these in particular do not require the compatibility conditions. □

7.1.4. Maximum principles.

This section develops the maximum principle and Harnack's inequality for second-order parabolic operators.

a. Weak maximum principle.

We will from now on assume that the operator L has the nondivergence form

(61) $$Lu = -\sum_{i,j=1}^{n} a^{ij}u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} + cu,$$

where the coefficients a^{ij}, b^i, c are continuous. We will always suppose the uniform parabolicity condition from §7.1.1, and also that $a^{ij} = a^{ji}$ $(i,j = 1,\dots,n)$. Recall also that the parabolic boundary of U_T is $\Gamma_T = \bar{U}_T - U_T$.

THEOREM 8 (Weak maximum principle). *Assume* $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ *and*

(62) $$c \equiv 0 \quad \text{in } U_T.$$

(i) *If*

(63) $$u_t + Lu \leq 0 \quad \text{in } U_T,$$

then

$$\max_{\bar{U}_T} u = \max_{\Gamma_T} u.$$

(ii) *Likewise, if*

(64) $$u_t + Lu \geq 0 \quad \text{in } U_T,$$

then

$$\min_{\bar{U}_T} u = \min_{\Gamma_T} u.$$

Remark. A function u satisfying the inequality (63) is called a *subsolution*, and so we are asserting that a subsolution attains its maximum on the parabolic boundary Γ_T. Similarly, u is a *supersolution* provided (64) holds, in which case u attains its minimum on Γ_T. □

Proof. 1. Let us first suppose we have the strict inequality

(65) $$u_t + Lu < 0 \quad \text{in } U_T,$$

but there exists a point $(x_0, t_0) \in U_T$ with $u(x_0, t_0) = \max_{\bar{U}_T} u$.

2. If $0 < t_0 < T$, then (x_0, t_0) belongs to the interior of U_T and consequently

(66) $$u_t = 0 \quad \text{at } (x_0, t_0),$$

since u attains its maximum at this point. On the other hand $Lu \geq 0$ at (x_0, t_0), as explained in the proof of Theorem 1 in §6.4. Thus $u_t + Lu \geq 0$ at (x_0, t_0), a contradiction to (65).

3. Now suppose $t_0 = T$. Then since u attains its maximum over $\bar{U}_T$ at (x_0, t_0), we see

$$u_t \geq 0 \quad \text{at } (x_0, t_0).$$

Since we still have the inequality $Lu \geq 0$ at (x_0, t_0), we once more deduce the contradiction

$$u_t + Lu \geq 0 \quad \text{at } (x_0, t_0).$$

4. In the general case that (63) holds, write $u^\epsilon(x,t) := u(x,t) - \epsilon t$ where $\epsilon > 0$. Then

$$u^\epsilon_t + Lu^\epsilon = u_t + Lu - \epsilon < 0 \quad \text{in } U_T,$$

and so $\max_{\bar{U}_T} u^\epsilon = \max_{\Gamma_T} u^\epsilon$. Let $\epsilon \to 0$ to find $\max_{\bar{U}_T} u = \max_{\Gamma_T} u$. This proves assertion (i).

5. As $-u$ is a subsolution whenever u is a supersolution, assertion (ii) follows at once. □

Next we allow for zeroth-order terms.

THEOREM 9 (Weak maximum principle for $c \geq 0$). *Assume $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ and*

$$c \geq 0 \quad \textit{in } U_T.$$

(i) *If*

$$u_t + Lu \leq 0 \quad \textit{in } U_T,$$

then

$$\max_{\bar{U}_T} u \leq \max_{\Gamma_T} u^+.$$

(ii) *If*

$$u_t + Lu \geq 0 \quad \textit{in } U_T,$$

then

$$\min_{\bar{U}_T} u \geq -\max_{\Gamma_T} u^-.$$

Remark. In particular, if $u_t + Lu = 0$ within U_T, then

$$\max_{\bar{U}_T} |u| = \max_{\Gamma_T} |u|.$$

□

Proof. 1. Assume u satisfies

$$u_t + Lu < 0 \quad \text{in } U_T \tag{67}$$

and attains a *positive* maximum at a point $(x_0, t_0) \in U_T$. Since $u(x_0, t_0) > 0$ and $c \geq 0$, we as above derive the contradiction

$$u_t + Lu \geq 0 \quad \text{at } (x_0, t_0).$$

2. If instead of (67), we have only

$$u_t + Lu \leq 0 \quad \text{in } U_T,$$

then as before $u^\epsilon(x,t) := u(x,t) - \epsilon t$ satisfies

$$u^\epsilon_t + Lu^\epsilon < 0 \quad \text{in } U_T.$$

Furthermore if u attains a positive maximum at some point in U_T, then u^ϵ also attains a positive maximum at some point belonging to U_T, provided $\epsilon > 0$ is small enough. But then, as in the previous proof, we obtain a contradiction.

3. Assertion (ii) follows similarly. □

Remark. Unlike the situation for elliptic PDE, various versions of the maximum principle obtain for parabolic PDE, *even if the zeroth-order coefficient is negative*: see Problem 7. □

b. Harnack's inequality.

Harnack's inequality states that if u is a nonnegative solution of our parabolic PDE, then the maximum of u in some interior region at a positive time can be estimated by the minimum of u in the same region *at a later time.*

THEOREM 10 (Parabolic Harnack inequality). *Assume $u \in C^2_1(U_T)$ solves*

$$u_t + Lu = 0 \quad \text{in } U_T, \tag{68}$$

and

$$u \geq 0 \quad \text{in } U_T.$$

Suppose $V \subset\subset U$ is connected. Then for each $0 < t_1 < t_2 \leq T$, there exists a constant C such that

$$\sup_V u(\cdot, t_1) \leq C \inf_V u(\cdot, t_2). \tag{69}$$

The constant C depends only on V, t_1, t_2, and the coefficients of L.

This is true if the coefficients are continuous, or even merely bounded and measurable; see Lieberman [**LB**]. We will however provide a proof only for the special case that $b^i \equiv c \equiv 0$ and the a^{ij} are smooth $(i, j = 1, \dots, n)$. The following computations are elementary, but fairly tricky.

Proof*. 1. We may assume $u > 0$ in U_T, for otherwise we could apply the result to $u + \epsilon$ and then let $\epsilon \to 0^+$.

Set

$$v := \log u \quad \text{in } U_T. \tag{70}$$

Using (68), we compute

$$v_t = \sum_{i,j=1}^n a^{ij} v_{x_i x_j} + a^{ij} v_{x_i} v_{x_j} \quad \text{in } U_T. \tag{71}$$

Define

$$w := \sum_{i,j=1}^n a^{ij} v_{x_i x_j}, \ \tilde{w} := \sum_{i,j=1}^n a^{ij} v_{x_i} v_{x_j}; \tag{72}$$

so that (71) reads

$$v_t = w + \tilde{w}. \tag{73}$$

2. We calculate using (72), (73) that for $k, l = 1, \dots, n$:

$$v_{x_k x_l t} = w_{x_k x_l} + \sum_{i,j=1}^n 2a^{ij} v_{x_i x_k x_l} v_{x_j} + 2a^{ij} v_{x_i x_k} v_{x_j x_l} + R,$$

where the remainder term R satisfies an estimate of the form

$$|R| \leq \epsilon |D^2 v|^2 + C(\epsilon)|Dv|^2 + C \tag{74}$$

*Omit on first reading.

for each $\epsilon > 0$. Thus

$$\begin{aligned} w_t &= \sum_{k,l=1}^n a^{kl} v_{x_k x_l t} + a_t^{kl} v_{x_k x_l} \\ &= \sum_{k,l=1}^n a^{kl} w_{x_k x_l} + 2\sum_{i,j=1}^n a^{ij} v_{x_j} w_{x_i} \\ &\quad + 2 \sum_{i,j,k,l=1}^n a^{ij} a^{kl} v_{x_i x_k} v_{x_j x_l} + R, \end{aligned}$$

where R now denotes another remainder term satisfying estimate (74). Therefore choosing $\epsilon > 0$ small enough in (74) and remembering the uniform parabolicity condition, we discover

$$w_t - \sum_{k,l=1}^n a^{kl} w_{x_k x_l} + \sum_{k=1}^n b^k w_{x_k} \geq \theta^2 |D^2 v|^2 - C|Dv|^2 - C, \tag{75}$$

where

$$b^k := -2\sum_{l=1}^n a^{kl} v_{x_l} \quad (k = 1, \dots, n). \tag{76}$$

3. Estimate (75) is a differential inequality for w, and our task next is to obtain a similar inequality for $\tilde{w}$. Indeed, using (72) and (71) we compute

$$\begin{aligned} \tilde{w}_t - \sum_{k,l=1}^n a^{kl} \tilde{w}_{x_k x_l} &= 2 \sum_{i,j=1}^n a^{ij} v_{x_i} \left(v_{t x_j} - \sum_{k,l=1}^n a^{kl} v_{x_k x_l x_j} \right) \\ &\quad - 2 \sum_{i,j,k,l=1}^n a^{ij} a^{kl} v_{x_i x_k} v_{x_j x_l} + R, \end{aligned}$$

R yet another remainder term satisfying (74) for each $\epsilon > 0$. Recalling (71) and (76), we simplify to discover:

$$\begin{aligned} &\tilde{w}_t - \sum_{k,l=1}^n a^{kl} \tilde{w}_{x_k x_l} + \sum_{k=1}^n b^k \tilde{w}_{x_k} \\ &\qquad \geq -C|D^2 v|^2 - C|Dv|^2 - C \quad \text{in } U_T. \end{aligned} \tag{77}$$

4. Next set

$$\hat{w} := w + \kappa \tilde{w}, \tag{78}$$

$\kappa > 0$ to be selected below. Combining (75) and (77), we deduce

$$(79)\qquad \hat{w}_t - \sum_{k,l=1}^{n} a^{kl}\hat{w}_{x_k x_l} + \sum_{k=1}^{n} b^k \hat{w}_{x_k} \geq \frac{\theta^2}{2}|D^2 v|^2 - C|Dv|^2 - C$$

provided $0 < \kappa \leq \frac{1}{2}$ is now fixed to be sufficiently small.

5. Suppose next $V \subset\subset U$ is an open ball and $0 < t_1 < t_2 \leq T$. Choose a cutoff function $\zeta \in C^\infty(U_T)$ such that

$$(80)\qquad \begin{cases} 0 \leq \zeta \leq 1, \ \zeta = 0 \text{ on } \Gamma_T, \\ \zeta \equiv 1 \text{ on } V \times [t_1, t_2]. \end{cases}$$

Note that ζ vanishes along $\{t = 0\}$.

Let μ be a positive constant (to be adjusted below), and assume then

$$(81)\qquad \begin{cases} \zeta^4 \hat{w} + \mu t \text{ attains a } \textit{negative} \text{ minimum} \\ \text{at some point } (x_0, t_0) \in U \times (0, T]. \end{cases}$$

Consequently

$$(82)\qquad \zeta \hat{w}_{x_k} + 4\zeta_{x_k}\hat{w} = 0 \quad \text{at } (x_0, t_0) \quad (k = 1, \ldots, n).$$

In addition

$$0 \geq (\zeta^4 \hat{w} + \mu t)_t - \sum_{k,l=1}^{n} a^{kl}(\zeta^4 \hat{w} + \mu t)_{x_k x_l} \quad \text{at } (x_0, t_0).$$

Hence at (x_0, t_0):

$$(83)\qquad 0 \geq \mu + \zeta^4 \Big(\hat{w}_t - \sum_{k,l=1}^{n} a^{kl}\hat{w}_{x_k x_l} \Big) - 2 \sum_{k,l=1}^{n} a^{kl}(\zeta^4)_{x_l}\hat{w}_{x_k} + \hat{R},$$

where

$$(84)\qquad |\hat{R}| \leq C\zeta^2 |\hat{w}|.$$

Recall now (79) and (82):

$$0 \geq \mu + \zeta^4 \left(\frac{\theta^2}{2}|D^2 v|^2 - C|Dv|^2 - C - \sum_{k=1}^{n} b^k \hat{w}_{x_k} \right) + \hat{R},$$

$\hat{R}$ another remainder term satisfying estimate (84). Utilizing (82) and (76), we deduce

$$(85)\qquad 0 \geq \mu + \zeta^4 \left(\frac{\theta^2}{2}|D^2 v|^2 - C|Dv|^2 - C \right) + \tilde{R},$$

where now

(86) $$|\tilde{R}| \leq C\zeta^2|\hat{w}| + C\zeta^3|Dv|\,|\hat{w}|.$$

Remember that (85), (86) are valid at the point (x_0, t_0) where $\zeta^4\hat{w} + \mu t$ attains a negative minimum. In particular, at this point $\hat{w} = w + \kappa\tilde{w} < 0$. Recalling the definition (72) of $w, \tilde{w}$, we deduce

(87) $$|Dv|^2 \leq C|D^2v|,$$

and so

$$|\hat{w}| \leq C|D^2v| \quad \text{at } (x_0, t_0).$$

Consequently (86) implies the estimate

(88) $$|\tilde{R}| \leq C\zeta^2|D^2v| + C\zeta^3|D^2v|^{3/2} \leq \epsilon\zeta^4|D^2v|^2 + C(\epsilon),$$

where we employed Young's inequality with ϵ from §B.2. Making use of (85), (87), (88) we at last discover a contradiction to (81), provided μ is large enough.

6. Therefore $\zeta^4\hat{w} + \mu t \geq 0$ in U_T, and so in particular

$$\hat{w} + \mu t \geq 0 \quad \text{in } V \times [t_1, t_2].$$

Using (73), we deduce then that

(89) $$v_t \geq \alpha|Dv|^2 - \beta \quad \text{in } V \times [t_1, t_2]$$

for appropriate constants $\alpha, \beta > 0$.

7. The differential inequality (89) for $v = \log u$ now leads us to the Harnack inequality, as follows. Fix $x_1, x_2 \in V$, $t_2 > t_1$. Then

$$\begin{aligned} v(x_2,t_2) - v(x_1,t_1) &= \int_0^1 \frac{d}{ds} v(sx_2 + (1-s)x_1, st_2 + (1-s)t_1)\,ds \\ &= \int_0^1 Dv \cdot (x_2 - x_1) + v_t(t_2 - t_1)\,ds \\ &\geq \int_0^1 -|Dv|\,|x_2 - x_1| + (t_2 - t_1)[\alpha|Dv|^2 - \beta]\,ds \text{ by (89)} \\ &\geq -\gamma, \end{aligned}$$

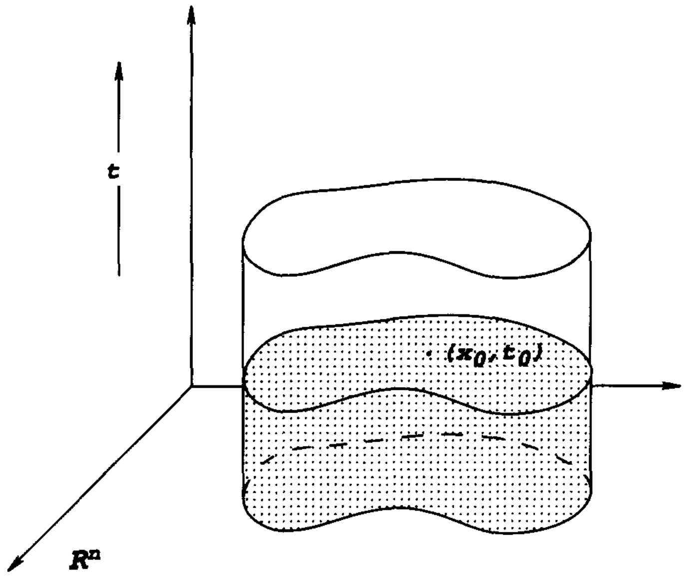
where γ depends only upon $\alpha, \beta, |x_1 - x_2|, |t_1 - t_2|$. Thus (70) implies

$$\log u(x_2, t_2) \geq \log u(x_1, t_1) - \gamma,$$

and so

$$u(x_2, t_2) \geq e^{-\gamma} u(x_1, t_1).$$

This inequality obtains for each $x_1, x_2 \in V$, and so (69) is valid in the case V is a ball. In the general case we cover $V \subset\subset U$ with balls and repeatedly apply the estimate above. □



Parabolic strong maximum principle

c. Strong maximum principle.

Now we employ Harnack's inequality to establish

THEOREM 11 (Strong maximum principle). *Assume $u \in C_1^2(U_T) \cap C(\bar{U}_T)$ and*

$$c \equiv 0 \quad \text{in } U_T.$$

Suppose also U is connected.

(i) *If*

$$u_t + Lu \leq 0 \quad \text{in } U_T$$

and u attains its maximum over $\bar{U}_T$ at a point $(x_0, t_0) \in U_T$, then

$$u \text{ is constant on } U_{t_0}.$$

(ii) *Likewise, if*

$$u_t + Lu \geq 0 \quad \text{in } U_T$$

and u attains its minimum over $\bar{U}_T$ at a point $(x_0, t_0) \in U_T$, then

$$u \text{ is constant on } U_{t_0}.$$

Remark. Thus our uniformly parabolic partial differential equations support "infinite propagation speed of disturbances". □

We will for the following proofs assume that u and the coefficients of L are in fact smooth.

Proof. 1. Assume $u_t + Lu \leq 0$ in U_T, and u attains its maximum at some point $(x_0, t_0) \in U_T$.

Select a smooth, open set $W \subset\subset U$, with $x_0 \in W$. Let v solve

$$\begin{cases} v_t + Lv = 0 & \text{in } W_T \\ v = u & \text{on } \Delta_T, \end{cases}$$

where Δ_T denotes the parabolic boundary of W_T.

Then by the weak maximum principle $u \leq v$. Since

$$u \leq v \leq M,$$

for $M := \max_{\bar{U}_T} u$, we deduce that $v = M$ at (x_0, t_0).

2. Now write $\tilde{v} := M - v$. Then, since $c \equiv 0$, we have

$$\tilde{v}_t + L\tilde{v} = 0,\ \tilde{v} \geq 0 \quad \text{in } W_T. \tag{90}$$

Choose any $V \subset\subset W$ with $x_0 \in V$, V connected. Let $0 < t < t_0$. Then owing to the Harnack inequality,

$$\max_V \tilde{v}(\cdot, t) \leq C \inf_V \tilde{v}(\cdot, t_0). \tag{91}$$

But $\inf_V \tilde{v}(\cdot, t_0) \leq \tilde{v}(x_0, t_0) = 0$. As $\tilde{v} \geq 0$, (91) therefore implies $\tilde{v} \equiv 0$ on $V \times \{t\}$, for each $0 < t < t_0$. This deduction holds for each set V as above, and so $\tilde{v} \equiv 0$ in W_{t_0}. But therefore $v \equiv M$ in W_{t_0}. As $v = u$ on Δ_T, we conclude $u \equiv M$ on $\partial W \times [0, t_0]$.

This conclusion holds for all sets W as above, and therefore $u \equiv M$ on U_{t_0}. □

THEOREM 12 (Strong maximum principle for $c \geq 0$). *Assume $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ and*

$$c \geq 0 \quad \textit{in } U_T.$$

Suppose also U is connected.

(i) *If*

$$u_t + Lu \leq 0 \quad \textit{in } U_T$$

and u attains a nonnegative maximum over $\bar{U}_T$ at a point $(x_0, t_0) \in U_T$, then

$$u \textit{ is constant on } U_{t_0}.$$

(ii) *Similarly, if*

$$u_t + Lu \geq 0 \quad \textit{in } U_T$$

and u attains a nonpositive minimum over $\bar{U}_T$ at a point $(x_0, t_0) \in U_T$, then

$$u \textit{ is constant on } U_{t_0}.$$

Proof. 1. As above, set $M := \max_{\bar{U}_T} u$. Assume $M \geq 0$, $u_t + Lu \leq 0$ in U_T, and u attains this maximum of M at some point $(x_0, t_0) \in U_T$.

If $M = 0$, the foregoing proof directly applies, as then

$$\tilde{v}_t + L\tilde{v} = 0, \ \tilde{v} \geq 0 \quad \text{in } W_T.$$

2. Assume instead that $M > 0$. Select as in the previous proof a smooth, open set $W \subset\subset U$, with $x_0 \in W$. Now let v solve

$$\begin{cases} v_t + Kv = 0 & \text{in } W_T \\ v = u^+ & \text{on } \Delta_T, \end{cases}$$

where

$$Kv := Lv - cv.$$

Note $0 \leq v \leq M$. Since $u_t + Ku = -cu \leq 0$ on $\{u \geq 0\}$, we deduce from the weak maximum principle that $u \leq v$. As before it follows that $v = (x_0, t_0)$.

2. Now write $\tilde{v} := M - v$. Then, since the operator K has no zeroth-order term, we have

$$\tilde{v}_t + K\tilde{v} = 0, \ \tilde{v} \geq 0 \quad \text{in } W_T.$$

Select any $V \subset\subset W$ with $x_0 \in V$, V connected. Let $0 < t < t_0$. Then the Harnack inequality implies as above that $v \equiv u^+ \equiv M$ on $\partial W \times [0, t_0]$. Since $M > 0$, we conclude that $u \equiv M$ on $\partial W \times [0, t_0]$.

This deduction is valid for all sets W as above, and therefore $u \equiv M$ on U_{t_0}. □

7.2. SECOND-ORDER HYPERBOLIC EQUATIONS

Second-order hyperbolic equations are natural generalizations of the wave equation (§2.4). We will build in this section appropriately defined weak solutions, and study their uniqueness, smoothness and other properties. It is interesting, given the utterly different physical character of second-order parabolic and hyperbolic PDE, that we can provide rather similar functional analytic constructions of weak solutions.

7.2.1. Definitions.

a. Hyperbolic equations.

As in §7.1 we write $U_T = U \times (0, T]$, where $T > 0$ and $U \subset \mathbb{R}^n$ is an open, bounded set.

We will next study the initial/boundary-value problem

$$\text{(1)} \qquad \begin{cases} u_{tt} + Lu = f & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g, u_t = h & \text{on } U \times \{t = 0\}, \end{cases}$$

where $f : U_T \to \mathbb{R}$, $g, h : U \to \mathbb{R}$ are given, and $u : \bar{U}_T \to \mathbb{R}$ is the unknown, $u = u(x,t)$. The symbol L denotes for each time t a second-order partial differential operator, having either the divergence form

$$\text{(2)} \qquad Lu = -\sum_{i,j=1}^{n} \left(a^{ij}(x,t)u_{x_i}\right)_{x_j} + \sum_{i=1}^{n} b^i(x,t)u_{x_i} + c(x,t)u$$

or else the nondivergence form

$$\text{(3)} \qquad Lu = -\sum_{i,j=1}^{n} a^{ij}(x,t)u_{x_i x_j} + \sum_{i=1}^{n} b^i(x,t)u_{x_i} + c(x,t)u$$

for given coefficients a^{ij}, b^i, c $(i,j = 1,\dots,n)$.

DEFINITION. *We say the partial differential operator* $\frac{\partial^2}{\partial t^2} + L$ *is* (uniformly) hyperbolic *if there exists a constant* $\theta > 0$ *such that*

$$\text{(4)} \qquad \sum_{i,j=1}^{n} a^{ij}(x,t)\xi_i\xi_j \geq \theta|\xi|^2$$

for all $(x,t) \in U_T$, $\xi \in \mathbb{R}^n$.

If $a^{ij} \equiv \delta_{ij}$, $b^i \equiv c \equiv f \equiv 0$, then $L = -\Delta$ and the partial differential equation becomes the wave equation, already studied in Chapter 2. General second-order hyperbolic PDE model wave transmission in heterogeneous, nonisotropic media.

b. Weak solutions.

As before, in §6.1.2 and §7.1.1, we first assume L has the divergence form (2) and look for an appropriate notion of weak solution for problem (1). We will suppose initially that

$$\text{(5)} \qquad a^{ij}, b^i, c \in C^1(\bar{U}_T) \quad (i,j = 1,\dots,n),$$

$$\text{(6)} \qquad f \in L^2(U_T),$$

$$\text{(7)} \qquad g \in H_0^1(U),\ h \in L^2(U),$$

and always assume $a^{ij} = a^{ji}$ $(i,j = 1,\dots,n)$.

As in §7.1.1, let us also introduce the time-dependent bilinear form

$$\text{(8)} \qquad B[u,v;t] := \int_U \sum_{i,j=1}^{n} a^{ij}(\cdot,t)u_{x_i}v_{x_j} + \sum_{i=1}^{n} b^i(\cdot,t)u_{x_i}v + c(\cdot,t)uv\,dx$$

for $u, v \in H_0^1(U)$ and $0 \leq t \leq T$.

Motivation for definition of weak solution. We begin by supposing $u = u(x,t)$ to be a smooth solution of (1) and defining the associated mapping

$$\mathbf{u} : [0,T] \to H_0^1(U),$$

by

$$[\mathbf{u}(t)](x) := u(x,t) \quad (x \in U,\ 0 \le t \le T).$$

We similarly introduce the function

$$\mathbf{f} : [0,T] \to L^2(U)$$

defined by

$$[\mathbf{f}(t)](x) := f(x,t) \quad (x \in U,\ 0 \le t \le T).$$

Now fix any function $v \in H_0^1(U)$, multiply the PDE $u_{tt} + Lu = f$ by v, and integrate by parts, to obtain the identity

$$(9) \qquad (\mathbf{u}'', v) + B[\mathbf{u}, v; t] = (\mathbf{f}, v)$$

for $0 \le t \le T$, where $(\ ,\ )$ denotes the inner product in $L^2(U)$. Almost exactly as in the parallel discussion for parabolic PDE in §7.1.1, we see from the PDE $u_{tt} + Lu = f$ that

$$u_{tt} = g^0 + \sum_{j=1}^n g^j_{x_j}$$

for $g^0 := f - \sum_{i=1}^n b^i u_{x_i} - cu$ and $g^j := \sum_{i=1}^n a^{ij} u_{x_i}$ $(j = 1, \dots, n)$. This suggests that we should look for a weak solution $\mathbf{u}$ with $\mathbf{u}'' \in H^{-1}(U)$ for a.e. $0 \le t \le T$, and then reinterpret the first term of (9) as $\langle \mathbf{u}'', v \rangle$, $\langle\ ,\ \rangle$ denoting as usual the pairing between $H^{-1}(U)$ and $H_0^1(U)$. □

DEFINITION. *We say a function*

$$\mathbf{u} \in L^2(0,T; H_0^1(U)), \ \textit{with}\ \mathbf{u}' \in L^2(0,T; L^2(U)),\ \mathbf{u}'' \in L^2(0,T; H^{-1}(U)),$$

is a weak solution *of the hyperbolic initial/boundary-value problem* (1) *provided*

(i) $\langle \mathbf{u}'', v \rangle + B[\mathbf{u}, v; t] = (\mathbf{f}, v)$

for each $v \in H_0^1(U)$ *and a.e. time* $0 \le t \le T$, *and*

(ii) $\mathbf{u}(0) = g$, $\mathbf{u}'(0) = h$.

Remark. In view of Theorem 2 in §5.9.2, we know $\mathbf{u} \in C([0,T];L^2(U))$ and $\mathbf{u}' \in C([0,T];H^{-1}(U))$. Consequently the equalities (ii) above make sense. □

7.2.2. Existence of weak solutions.

a. Galerkin approximations.

By analogy with the approach taken in §7.1.2 we will construct our weak solution of the hyperbolic initial/boundary-value problem

$$(10) \qquad \begin{cases} u_{tt} + Lu = f & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g, u_t = h & \text{on } U \times \{t=0\} \end{cases}$$

by first solving a finite dimensional approximation. We thus once more employ *Galerkin's method* by selecting smooth functions $w_k = w_k(x)$ $(k = 1,\dots)$ such that

$$(11) \qquad \{w_k\}_{k=1}^{\infty} \text{ is an orthogonal basis of } H_0^1(U)$$

and

$$(12) \qquad \{w_k\}_{k=1}^{\infty} \text{ is an orthonormal basis of } L^2(U).$$

Fix a positive integer m, and write

$$(13) \qquad \mathbf{u}_m(t) := \sum_{k=1}^{m} d_m^k(t) w_k,$$

where we intend to select the coefficients $d_m^k(t)$ $(0 \le t \le T,\ k = 1,\dots,m)$ to satisfy

$$(14) \qquad d_m^k(0) = (g, w_k) \quad (k = 1,\dots,m),$$

$$(15) \qquad {d_m^k}'(0) = (h, w_k) \quad (k = 1,\dots,m),$$

and

$$(16) \qquad (\mathbf{u}_m'', w_k) + B[\mathbf{u}_m, w_k; t] = (\mathbf{f}, w_k) \quad (0 \le t \le T,\ k = 1,\dots,m).$$

THEOREM 1 (Construction of approximate solutions). *For each integer $m = 1,2,\dots,$ there exists a unique function $\mathbf{u}_m$ of the form* (13) *satisfying* (14)–(16).

Proof. Assuming $\mathbf{u}_m$ to be given by (13), we observe using (12)

$$(\mathbf{u}_m''(t), w_k) = d_m^{k\,''}(t). \tag{17}$$

Furthermore, exactly as in the proof of Theorem 1 in §7.1.2, we have

$$B[\mathbf{u}_m, w_k; t] = \sum_{l=1}^m e^{kl}(t) d_m^l(t)$$

for $e^{kl}(t) := B[w_l, w_k; t]$ $(k, l = 1, \dots, m)$. We also write $f^k(t) := (\mathbf{f}(t), w_k)$ $(k = 1, \dots, m)$. Consequently (16) becomes the linear system of ODE

$$d_m^{k\,''}(t) + \sum_{l=1}^m e^{kl}(t) d_m^l(t) = f^k(t) \quad (0 \le t \le T,\ k = 1, \dots, m), \tag{18}$$

subject to the initial conditions (14), (15). According to standard theory for ordinary differential equations, there exists a unique C^2 function $\mathbf{d}_m(t) = (d_m^1(t), \dots, d_m^m(t))$, satisfying (14), (15), and solving (18) for $0 \le t \le T$.

□

b. Energy estimates.

Our plan is hereafter to send $m \to \infty$, and so we will need some estimates, uniform in m.

THEOREM 2 (Energy estimates). *There exists a constant C, depending only on U, T and the coefficients of L, such that*

$$\begin{aligned} \max_{0\le t\le T} \left(\|\mathbf{u}_m(t)\|_{H_0^1(U)} + \|\mathbf{u}_m'(t)\|_{L^2(U)} \right) &+ \|\mathbf{u}_m''\|_{L^2(0,T;H^{-1}(U))} \\ &\le C \left(\|\mathbf{f}\|_{L^2(0,T;L^2(U))} + \|g\|_{H_0^1(U)} + \|h\|_{L^2(U)} \right), \end{aligned} \tag{19}$$

for $m = 1, 2, \dots$.

Proof. 1. Multiply equality (16) by $d_m^{k\prime}(t)$, sum $k = 1, \dots, m$, and recall (13) to discover

$$(\mathbf{u}_m'', \mathbf{u}_m') + B[\mathbf{u}_m, \mathbf{u}_m'; t] = (\mathbf{f}, \mathbf{u}_m') \tag{20}$$

for a.e. $0 \le t \le T$. Observe next $(\mathbf{u}_m'', \mathbf{u}_m') = \frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}_m'\|_{L^2(U)}^2\right)$. Furthermore, we can write

$$\begin{aligned} B[\mathbf{u}_m, \mathbf{u}_m'; t] &= \int_U \sum_{i,j=1}^n a^{ij} \mathbf{u}_{m,x_i} \mathbf{u}_{m,x_j}' \, dx \\ &\quad + \int_U \sum_{i=1}^n b^i \mathbf{u}_{m,x_i} \mathbf{u}_m' + c \mathbf{u}_m \mathbf{u}_m' \, dx \\ &=: B_1 + B_2. \end{aligned} \tag{21}$$

Since $a^{ij} = a^{ji}$ $(i, j = 1, \dots, n)$, we see

$$B_1 = \frac{d}{dt}\left(\frac{1}{2}A[\mathbf{u}_m, \mathbf{u}_m; t]\right) - \frac{1}{2}\int_U \sum_{i,j=1}^{n} a_t^{ij}\mathbf{u}_{m,x_i}\mathbf{u}_{m,x_j}\,dx, \tag{22}$$

for the symmetric bilinear form

$$A[u, v; t] := \int_U \sum_{i,j=1}^{n} a^{ij} u_{x_i} v_{x_j}\,dx \quad (u, v \in H_0^1(U)).$$

The equality (22) implies

$$B_1 \geq \frac{d}{dt}\left(\frac{1}{2}A[\mathbf{u}_m, \mathbf{u}_m; t]\right) - C\|\mathbf{u}_m\|^2_{H_0^1(U)}, \tag{23}$$

and we note also

$$|B_2| \leq C\left(\|\mathbf{u}_m\|^2_{H_0^1(U)} + \|\mathbf{u}_m'\|^2_{L^2(U)}\right). \tag{24}$$

Combining estimates (20)–(24), we discover

$$\begin{aligned}\frac{d}{dt}\left(\|\mathbf{u}_m'\|^2_{L^2(U)} + A[\mathbf{u}_m, \mathbf{u}_m; t]\right) &\leq C\left(\|\mathbf{u}_m'\|^2_{L^2(U)} + \|\mathbf{u}_m\|^2_{H_0^1(U)} + \|\mathbf{f}\|^2_{L^2(U)}\right)\\ &\leq C\left(\|\mathbf{u}_m'\|^2_{L^2(U)} + A[\mathbf{u}_m, \mathbf{u}_m; t] + \|\mathbf{f}\|^2_{L^2(U)}\right),\end{aligned} \tag{25}$$

where we used the inequality

$$\theta\int_U |Du|^2 dx \leq A[u, u; t] \quad (u \in H_0^1(U)), \tag{26}$$

which follows from the uniform hyperbolicity condition.

2. Now write

$$\eta(t) := \|\mathbf{u}_m'(t)\|^2_{L^2(U)} + A[\mathbf{u}_m(t), \mathbf{u}_m(t); t] \tag{27}$$

and

$$\xi(t) := \|\mathbf{f}(t)\|^2_{L^2(U)}. \tag{28}$$

Then inequality (25) reads

$$\eta'(t) \leq C_1\eta(t) + C_2\xi(t)$$

for $0 \le t \le T$ and appropriate constants C_1, C_2. Thus Gronwall's inequality (§B.2) yields the estimate

$$\eta(t) \le e^{C_1 t}\left(\eta(0) + C_2 \int_0^t \xi(s)\,ds\right) \quad (0 \le t \le T). \tag{29}$$

However

$$\begin{aligned}\eta(0) &= \|\mathbf{u}_m'(0)\|_{L^2(U)}^2 + A[\mathbf{u}_m(0), \mathbf{u}_m(0); 0] \\ &\le C\left(\|h\|_{L^2(U)}^2 + \|g\|_{H_0^1(U)}^2\right),\end{aligned}$$

according to (14) and (15) and the estimate $\|\mathbf{u}_m(0)\|_{H_0^1(U)} \le \|g\|_{H_0^1(U)}$. Thus formulas (27)–(29) provide the bound

$$\begin{aligned}&\|\mathbf{u}_m'(t)\|_{L^2(U)}^2 + A[\mathbf{u}_m(t), \mathbf{u}_m(t); t] \\ &\qquad \le C\left(\|g\|_{H_0^1(U)}^2 + \|h\|_{L^2(U)}^2 + \|\mathbf{f}\|_{L^2(0,T;L^2(U))}^2\right).\end{aligned}$$

Since $0 \le t \le T$ was arbitrary, we see from this estimate and (26) that

$$\begin{aligned}&\max_{0\le t\le T}\left(\|\mathbf{u}_m(t)\|_{H_0^1(U)}^2 + \|\mathbf{u}_m'(t)\|_{L^2(U)}^2\right) \\ &\qquad \le C\big(\|g\|_{H_0^1(U)}^2 + \|h\|_{L^2(U)}^2 + \|\mathbf{f}\|_{L^2(0,T;L^2(U))}^2\big).\end{aligned}$$

3. Fix any $v \in H_0^1(U)$, $\|v\|_{H_0^1(U)} \le 1$, and write $v = v^1 + v^2$, where $v^1 \in \operatorname{span}\{w_k\}_{k=1}^m$ and $(v^2, w_k) = 0$ $(k = 1, \dots, m)$. Note $\|v^1\|_{H_0^1(U)} \le 1$. Then (13) and (16) imply

$$\langle \mathbf{u}_m'', v\rangle = (\mathbf{u}_m'', v) = (\mathbf{u}_m'', v^1) = (\mathbf{f}, v^1) - B[\mathbf{u}_m, v^1; t].$$

Thus

$$|\langle \mathbf{u}_m'', v\rangle| \le C(\|\mathbf{f}\|_{L^2(U)} + \|\mathbf{u}_m\|_{H_0^1(U)}),$$

since $\|v^1\|_{H_0^1(U)} \le 1$. Consequently

$$\begin{aligned}\int_0^T \|\mathbf{u}_m''\|_{H^{-1}(U)}^2 dt &\le C\int_0^T \|\mathbf{f}\|_{L^2(U)}^2 + \|\mathbf{u}_m\|_{H_0^1(U)}^2 dt \\ &\le C\big(\|g\|_{H_0^1(U)}^2 + \|h\|_{L^2(U)}^2 + \|\mathbf{f}\|_{L^2(0,T;L^2(U))}^2\big).\end{aligned}$$

□

c. Existence and uniqueness.

Now we pass to limits in our Galerkin approximations.

THEOREM 3 (Existence of weak solution). *There exists a weak solution of* (1).

Proof. 1. According to the energy estimates (19), we see that the sequence $\{\mathbf{u}_m\}_{m=1}^\infty$ is bounded in $L^2(0,T;H_0^1(U))$, $\{\mathbf{u}_m'\}_{m=1}^\infty$ is bounded in $L^2(0,T;L^2(U))$ and $\{\mathbf{u}_m''\}_{m=1}^\infty$ is bounded in $L^2(0,T;H^{-1}(U))$.

As a consequence there exists a subsequence $\{\mathbf{u}_{m_l}\}_{l=1}^\infty \subset \{\mathbf{u}_m\}_{m=1}^\infty$ and $\mathbf{u} \in L^2(0,T;H_0^1(U))$, with $\mathbf{u}' \in L^2(0,T;L^2(U))$, $\mathbf{u}'' \in L^2(0,T;H^{-1}(U))$, such that

$$
(30) \qquad \begin{cases} \mathbf{u}_{m_l} \rightharpoonup \mathbf{u} & \text{weakly in } L^2(0,T;H_0^1(U)) \\ \mathbf{u}_{m_l}' \rightharpoonup \mathbf{u}' & \text{weakly in } L^2(0,T;L^2(U)) \\ \mathbf{u}_{m_l}'' \rightharpoonup \mathbf{u}'' & \text{weakly in } L^2(0,T;H^{-1}(U)). \end{cases}
$$

2. Next fix an integer N and choose a function $\mathbf{v} \in C^1([0,T];H_0^1(U))$ of the form

$$
(31) \qquad \mathbf{v}(t) = \sum_{k=1}^N d^k(t) w_k,
$$

where $\{d^k\}_{k=1}^N$ are smooth functions. We select $m \geq N$, multiply (16) by $d^k(t)$, sum $k = 1, \dots, N$, and then integrate with respect to t, to discover

$$
(32) \qquad \int_0^T \langle \mathbf{u}_m'', \mathbf{v} \rangle + B[\mathbf{u}_m, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt.
$$

We set $m = m_l$ and recall (30), to find in the limit that

$$
(33) \qquad \int_0^T \langle \mathbf{u}'', \mathbf{v} \rangle + B[\mathbf{u}, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt.
$$

This equality then holds for all functions $\mathbf{v} \in L^2(0,T;H_0^1(U))$, since functions of the form (31) are dense in this space. From (33) it follows further that

$$
\langle \mathbf{u}'', v \rangle + B[\mathbf{u}, v; t] = (\mathbf{f}, v)
$$

for all $v \in H_0^1(U)$ and a.e. $0 \leq t \leq T$. Furthermore, $\mathbf{u} \in C([0,T];L^2(U))$ and $\mathbf{u}' \in C([0,T];H^{-1}(U))$.

3. We must now verify

$$
(34) \qquad \mathbf{u}(0) = g,
$$

$$\mathbf{u}'(0) = h. \tag{35}$$

For this, choose any function $\mathbf{v} \in C^2([0,T]; H_0^1(U))$, with $\mathbf{v}(T) = \mathbf{v}'(T) = 0$. Then integrating by parts twice with respect to t in (33), we find

$$\begin{aligned}\int_0^T (\mathbf{v}'', \mathbf{u}) + B[\mathbf{u}, \mathbf{v}; t]\, dt = \int_0^T &(\mathbf{f}, \mathbf{v})\, dt \\ &- (\mathbf{u}(0), \mathbf{v}'(0)) + \langle \mathbf{u}'(0), \mathbf{v}(0)\rangle.\end{aligned} \tag{36}$$

Similarly from (32) we deduce

$$\begin{aligned}\int_0^T (\mathbf{v}'', \mathbf{u}_m) + B[\mathbf{u}_m, \mathbf{v}; t]\, dt = \int_0^T &(\mathbf{f}, \mathbf{v})\, dt \\ &- \left(\mathbf{u}_m(0), \mathbf{v}'(0)\right) + \left(\mathbf{u}_m'(0), \mathbf{v}(0)\right).\end{aligned}$$

We set $m = m_l$ and recall (14), (15) and (30), to deduce

$$\int_0^T (\mathbf{v}'', \mathbf{u}) + B[\mathbf{u}, \mathbf{v}; t]\, dt = \int_0^T (\mathbf{f}, \mathbf{v})\, dt - (g, \mathbf{v}'(0)) + (h, \mathbf{v}(0)). \tag{37}$$

Comparing identities (36) and (37), we conclude (34), (35), since $\mathbf{v}(0)$, $\mathbf{v}'(0)$ are arbitrary. Hence $\mathbf{u}$ is a weak solution of (1). □

Remark. Recalling the energy estimates from Theorem 2, we observe that in fact $\mathbf{u} \in L^\infty(0,T; H_0^1(U))$, $\mathbf{u}' \in L^\infty(0,T; L^2(U))$, $\mathbf{u}'' \in L^2(0,T; H^{-1}(U))$: see Theorem 5 below. □

THEOREM 4 (Uniqueness of weak solution). *A weak solution of* (1) *is unique.*

Remark. The following tricky demonstration would be greatly simplified if we knew $\mathbf{u}'(t)$ itself were smooth enough to insert in place of v in the definition of weak solution. This is not so, however. □

Proof. 1. It suffices to show that the only weak solution of (1) with $\mathbf{f} \equiv g \equiv h \equiv 0$ is

$$\mathbf{u} \equiv \mathbf{0}. \tag{38}$$

To verify this, fix $0 \le s \le T$ and set

$$\mathbf{v}(t) := \begin{cases} \int_t^s \mathbf{u}(\tau)\, d\tau & \text{if } 0 \le t \le s \\ \mathbf{0} & \text{if } s \le t \le T. \end{cases}$$

Then $\mathbf{v}(t) \in H_0^1(U)$ for each $0 \leq t \leq T$, and so

$$\int_0^s \langle \mathbf{u}'', \mathbf{v} \rangle + B[\mathbf{u}, \mathbf{v}; t] \, dt = 0.$$

Since $\mathbf{u}'(0) = \mathbf{v}(s) = \mathbf{0}$, we obtain after integrating by parts in the first term above:

$$\int_0^s -(\mathbf{u}', \mathbf{v}') + B[\mathbf{u}, \mathbf{v}; t] \, dt = 0. \tag{39}$$

Now $\mathbf{v}' = -\mathbf{u}$ $(0 \leq t \leq s)$, and so

$$\int_0^s \langle \mathbf{u}', \mathbf{u} \rangle - B[\mathbf{v}', \mathbf{v}; t] \, dt = 0.$$

Thus

$$\int_0^s \frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}\|_{L^2(U)}^2 - \frac{1}{2}B[\mathbf{v}, \mathbf{v}; t]\right) dt = -\int_0^s C[\mathbf{u}, \mathbf{v}; t] + D[\mathbf{v}, \mathbf{v}; t] \, dt,$$

where

$$C[u, v; t] := \int_U \sum_{i=1}^n b^i v_{x_i} u + \frac{1}{2} b_{i,x_i} uv \, dx$$

and

$$D[u, v; t] := \frac{1}{2} \int_U \sum_{i,j=1}^n a_{ij,t} u_{x_i} v_{x_j} + \sum_{i=1}^n b_{i,t} u_{x_i} v + c_t uv \, dx,$$

for $u, v \in H_0^1(U)$. Hence

$$\frac{1}{2}\|\mathbf{u}(s)\|_{L^2(U)}^2 + \frac{1}{2}B[\mathbf{v}(0), \mathbf{v}(0); t] = -\int_0^s C[\mathbf{u}, \mathbf{v}; t] + D[\mathbf{v}, \mathbf{v}; t] \, dt,$$

and consequently

$$\begin{aligned} &\|\mathbf{u}(s)\|_{L^2(U)}^2 + \|\mathbf{v}(0)\|_{H_0^1(U)}^2 \\ &\qquad \leq C\left(\int_0^s \|\mathbf{v}\|_{H_0^1(U)}^2 + \|\mathbf{u}\|_{L^2(U)}^2 dt + \|\mathbf{v}(0)\|_{L^2(U)}^2\right). \end{aligned} \tag{40}$$

2. Now let us write

$$\mathbf{w}(t) := \int_0^t \mathbf{u}(\tau) \, d\tau \quad (0 \leq t \leq T);$$

whereupon (40) becomes

$$\|\mathbf{u}(s)\|^2_{L^2(U)} + \|\mathbf{w}(s)\|^2_{H^1_0(U)}$$
$$\text{(41)} \qquad \leq C\left(\int_0^s \|\mathbf{w}(t) - \mathbf{w}(s)\|^2_{H^1_0(U)} + \|\mathbf{u}(t)\|^2_{L^2(U)}\, dt + \|\mathbf{w}(s)\|^2_{L^2(U)}\right).$$

But $\|\mathbf{w}(t)-\mathbf{w}(s)\|^2_{H^1_0(U)} \leq 2\|\mathbf{w}(t)\|^2_{H^1_0(U)}+2\|\mathbf{w}(s)\|^2_{H^1_0(U)}$, and $\|\mathbf{w}(s)\|_{L^2(U)} \leq \int_0^s \|\mathbf{u}(t)\|_{L^2(U)}dt$. Therefore (41) implies

$$\|\mathbf{u}(s)\|^2_{L^2(U)} + (1 - 2sC_1)\|\mathbf{w}(s)\|^2_{H^1_0(U)} \leq C_1 \int_0^s \|\mathbf{w}\|^2_{H^1_0(U)} + \|\mathbf{u}\|^2_{L^2(U)}dt.$$

Choose T_1 so small that

$$1 - 2T_1C_1 \geq \frac{1}{2}.$$

Then if $0 \leq s \leq T_1$, we have

$$\|\mathbf{u}(s)\|^2_{L^2(U)} + \|\mathbf{w}(s)\|^2_{H^1_0(U)} \leq C \int_0^s \|\mathbf{u}\|^2_{L^2(U)} + \|\mathbf{w}\|^2_{H^1_0(U)}dt.$$

Consequently the integral form of Gronwall's inequality (§B.2) implies $\mathbf{u} \equiv \mathbf{0}$ on $[0, T_1]$.

3. We apply the same argument on the intervals $[T_1, 2T_1]$, $[2T_1, 3T_1]$, etc., eventually to deduce (38). □

7.2.3. Regularity.

As in our earlier treatments of second-order elliptic and parabolic PDE, the next task is to study the smoothness of our weak solutions.

Motivation: formal derivation of estimates. (i) Suppose for the moment $u = u(x, t)$ is a smooth solution of this initial-value problem for the wave equation:

$$\begin{cases} u_{tt} - \Delta u = f & \text{in } \mathbb{R}^n \times (0, T] \\ u = g, u_t = h & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases}$$

and assume also u goes to zero as $|x| \to \infty$ sufficiently rapidly to justify the following calculations. Then as in §2.4.3, we compute

$$\begin{aligned} \frac{d}{dt}\left(\int_{\mathbb{R}^n} |Du|^2 + u_t^2\, dx\right) &= 2\int_{\mathbb{R}^n} Du \cdot Du_t + u_t u_{tt}\, dx \\ &= 2\int_{\mathbb{R}^n} u_t(u_{tt} - \Delta u)\, dx = 2\int_{\mathbb{R}^n} u_t f\, dx \\ &\leq \int_{\mathbb{R}^n} u_t^2\, dx + \int_{\mathbb{R}^n} f^2\, dx. \end{aligned}$$

Applying Gronwall's inequality, we deduce

$$\sup_{0\le t\le T}\int_{\mathbb{R}^n}|Du|^2+u_t^2\,dx\le C\Big(\int_0^T\int_{\mathbb{R}^n}f^2\,dxdt+\int_{\mathbb{R}^n}|Dg|^2+h^2\,dx\Big), \tag{42}$$

with the constant C depending only on T.

(ii) Next differentiate the PDE with respect to t and set $\tilde u := u_t$. Then

$$\begin{cases}\tilde u_{tt}-\Delta\tilde u=\tilde f & \text{in } \mathbb{R}^n\times(0,T]\\ \tilde u=\tilde g,\ \tilde u_t=\tilde h & \text{on } \mathbb{R}^n\times\{t=0\},\end{cases} \tag{43}$$

for $\tilde f := f_t$, $\tilde g := h$, $\tilde h := u_{tt}(\cdot,0)=f(\cdot,0)+\Delta g$. Applying estimate (42) to $\tilde u$, we discover

$$\begin{aligned}&\sup_{0\le t\le T}\int_{\mathbb{R}^n}|Du_t|^2+u_{tt}^2\,dx\\ &\quad\le C\Big(\int_0^T\int_{\mathbb{R}^n}f_t^2\,dxdt+\int_{\mathbb{R}^n}|D^2g|^2+|Dh|^2+f(\cdot,0)^2\,dx\Big).\end{aligned} \tag{44}$$

Now

$$\max_{0\le t\le T}\|f(\cdot,t)\|_{L^2(\mathbb{R}^n)}\le C(\|f\|_{L^2(\mathbb{R}^n\times(0,T))}+\|f_t\|_{L^2(\mathbb{R}^n\times(0,T))}), \tag{45}$$

according to Theorem 2 in §5.9.2. Furthermore, writing $-\Delta u=f-u_{tt}$, we deduce as in §6.3 that

$$\int_{\mathbb{R}^n}|D^2u|^2\,dx\le C\int_{\mathbb{R}^n}f^2+u_{tt}^2\,dx \tag{46}$$

for each $0\le t\le T$. Combining (44)–(46), we conclude

$$\begin{aligned}&\sup_{0\le t\le T}\int_{\mathbb{R}^n}|D^2u|^2+|Du_t|^2+u_{tt}^2\,dx\\ &\quad\le C\Big(\int_0^T\int_{\mathbb{R}^n}f_t^2+f^2\,dxdt+\int_{\mathbb{R}^n}|D^2g|^2+|Dh|^2\,dx\Big),\end{aligned} \tag{47}$$

the constant C depending only on T. □

This estimate suggests that bounds similar to (42) and (47) should be valid for our weak solution of a general second-order hyperbolic PDE.

We will calculate using the Galerkin approximations. To simplify the presentation, we hereafter assume that $\{w_k\}_{k=1}^\infty$ is the complete collection of eigenfunctions for $-\Delta$ on $H_0^1(U)$, and also that U is bounded, open, with ∂U smooth. In addition we suppose

$$\begin{cases}\text{the coefficients } a^{ij},b^i,c\ (i,j=1,\dots,n) \text{ are smooth on}\\ \bar U \text{ and do not depend on } t.\end{cases} \tag{48}$$

THEOREM 5 (Improved regularity).

(i) *Assume*

$$g \in H_0^1(U),\ h \in L^2(U),\ \mathbf{f} \in L^2(0,T;L^2(U)),$$

and suppose also $\mathbf{u} \in L^2(0,T;H_0^1(U))$, *with* $\mathbf{u}' \in L^2(0,T;L^2(U))$, $\mathbf{u}'' \in L^2(0,T;H^{-1}(U))$, *is the weak solution of the problem*

$$\text{(49)} \qquad \begin{cases} u_{tt} + Lu = f & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g, u_t = h & \text{on } U \times \{t = 0\}. \end{cases}$$

Then in fact

$$\mathbf{u} \in L^\infty(0,T;H_0^1(U)),\ \mathbf{u}' \in L^\infty(0,T;L^2(U)),$$

and we have the estimate

$$\text{(50)} \qquad \begin{aligned} &\operatorname*{ess\,sup}_{0\le t\le T} \left(\|\mathbf{u}(t)\|_{H_0^1(U)} + \|\mathbf{u}'(t)\|_{L^2(U)}\right) \\ &\qquad\le C(\|\mathbf{f}\|_{L^2(0,T;L^2(U))} + \|g\|_{H_0^1(U)} + \|h\|_{L^2(U)}). \end{aligned}$$

(ii) *If, in addition,*

$$g \in H^2(U),\ h \in H_0^1(U),\ \mathbf{f}' \in L^2(0,T;L^2(U)),$$

then

$$\begin{gathered} \mathbf{u} \in L^\infty(0,T;H^2(U)),\ \mathbf{u}' \in L^\infty(0,T;H_0^1(U)), \\ \mathbf{u}'' \in L^\infty(0,T;L^2(U)),\ \mathbf{u}''' \in L^2(0,T;H^{-1}(U)), \end{gathered}$$

with the estimate:

$$\text{(51)} \qquad \begin{aligned} &\operatorname*{ess\,sup}_{0\le t\le T} \left(\|\mathbf{u}(t)\|_{H^2(U)} + \|\mathbf{u}'(t)\|_{H_0^1(U)} + \|\mathbf{u}''(t)\|_{L^2(U)}\right) \\ &+ \|\mathbf{u}'''\|_{L^2(0,T;H^{-1}(U))} \le C(\|\mathbf{f}\|_{H^1(0,T;L^2(U))} + \|g\|_{H^2(U)} + \|h\|_{H^1(U)}). \end{aligned}$$

Remark. Assertions (i), (ii) of this theorem are precise versions of the formal estimates (42), (47) (for the wave equation in $U = \mathbb{R}^n$). □

Proof. 1. In the proof of Theorem 2, we have already derived the bounds

$$\sup_{0\le t\le T}(\|\mathbf{u}_m(t)\|_{H_0^1(U)}+\|\mathbf{u}_m'(t)\|_{L^2(U)})\\ \le C(\|\mathbf{f}\|_{L^2(0,T;L^2(U))}+\|g\|_{H_0^1(U)}+\|h\|_{L^2(U)}). \tag{52}$$

Passing to limits as $m=m_l\to\infty$, we deduce (50).

2. Assume now the hypotheses of assertion (ii). Fix a positive integer m, and next differentiate the identity (16) with respect to t. Writing $\tilde{\mathbf{u}}_m:=\mathbf{u}_m'$ we obtain

$$(\tilde{\mathbf{u}}_m'',w_k)+B[\tilde{\mathbf{u}}_m,w_k]=(\mathbf{f}',w_k)\quad(k=1,\dots,m). \tag{53}$$

Multiplying by $d_m^{k\,\prime\prime}(t)$ and adding for $k=1,\dots,m$, we discover

$$(\tilde{\mathbf{u}}_m'',\tilde{\mathbf{u}}_m')+B[\tilde{\mathbf{u}}_m,\tilde{\mathbf{u}}_m']=(\mathbf{f}',\tilde{\mathbf{u}}_m'). \tag{54}$$

Arguing as in the proof of the energy estimates, we observe

$$\frac{d}{dt}(\|\tilde{\mathbf{u}}_m'\|^2_{L^2(U)}+A[\tilde{\mathbf{u}}_m,\tilde{\mathbf{u}}_m])\\ \le C(\|\tilde{\mathbf{u}}_m'\|^2_{L^2(U)}+A[\tilde{\mathbf{u}}_m,\tilde{\mathbf{u}}_m]+\|\mathbf{f}'\|^2_{L^2(U)}), \tag{55}$$

the bilinear form $A[\ ,\ ]$ defined as before.

3. Now

$$B[\mathbf{u}_m,w_k]=(\mathbf{f}-\mathbf{u}_m'',w_k)\quad(k=1,\dots,m). \tag{56}$$

Recall we are taking $\{w_k\}_{k=1}^\infty$ to be the complete collection of eigenfunctions for $-\Delta$ on $H_0^1(U)$. Multiplying (56) by $\lambda_k d_m^k(t)$ and summing $k=1,\dots,m$, we deduce

$$B[\mathbf{u}_m,-\Delta\mathbf{u}_m]=(\mathbf{f}-\mathbf{u}_m'',-\Delta\mathbf{u}_m). \tag{57}$$

Since $\Delta\mathbf{u}_m=0$ on ∂U, we have

$$B[\mathbf{u}_m,-\Delta\mathbf{u}_m]=(L\mathbf{u}_m,-\Delta\mathbf{u}_m). \tag{58}$$

Next we employ the inequality

$$\beta\|u\|^2_{H^2(U)}\le(Lu,-\Delta u)+\gamma\|u\|^2_{L^2(U)}\quad(u\in H^2(U)\cap H_0^1(U)); \tag{59}$$

see Problem 8. We deduce from (56)–(59) that

$$\|\mathbf{u}_m\|^2_{H^2(U)}\le C(\|\mathbf{f}\|^2_{L^2(U)}+\|\mathbf{u}_m''\|^2_{L^2(U)}+\|\mathbf{u}_m\|^2_{L^2(U)}). \tag{60}$$

Using this estimate in (55), recalling $\tilde{\mathbf{u}}_m = \mathbf{u}_m'$, and applying Gronwall's inequality, we deduce

$$(61)\qquad \begin{aligned}\sup_{0\le t\le T}&\left(\|\mathbf{u}_m(t)\|^2_{H^2(U)} + \|\mathbf{u}_m'(t)\|^2_{H^1(U)} + \|\mathbf{u}_m''(t)\|^2_{L^2(U)}\right)\\ &\le C\left(\|\mathbf{f}\|^2_{H^1(0,T;L^2(U))} + \|g\|^2_{H^2(U)} + \|h\|^2_{H^1(U)}\right).\end{aligned}$$

Here we estimated $\|\mathbf{u}_m(0)\|_{H^2(U)} \le C\|g\|_{H^2(U)}$, as in the proof of Theorem 5 in §7.1.3.

Passing to limits as $m = m_l \to \infty$, we derive the same bound for $\mathbf{u}$.

4. As in the earlier proof of Theorem 5 in §7.1.3, we likewise deduce $\mathbf{u}''' \in L^2(0,T;H^{-1}(U))$, with the stated estimate. □

Remark. If L were symmetric, we could alternatively have taken $\{w_k\}_{k=1}^\infty$ to be a basis of eigenfunctions of L on $H_0^1(U)$, and so avoided the need for inequality (59). □

Now let m be a nonnegative integer.

THEOREM 6 (Higher regularity). *Assume*

$$\begin{cases} g \in H^{m+1}(U),\ h \in H^m(U), \\ \frac{d^k\mathbf{f}}{dt^k} \in L^2(0,T;H^{m-k}(U)) \quad (k = 0,\dots,m). \end{cases}$$

Suppose also the following m^{th}-order compatibility conditions *hold:*

$$(62)\qquad \begin{cases} g_0 := g \in H_0^1(U),\ h_1 := h \in H_0^1(U),\dots, \\ g_{2l} := \frac{d^{2l-2}\mathbf{f}}{dt^{2l-2}}(\cdot,0) - Lg_{2l-2} \in H_0^1(U) \qquad (\textit{if}\ \ m = 2l) \\ h_{2l+1} := \frac{d^{2l-1}\mathbf{f}}{dt^{2l-1}}(\cdot,0) - Lh_{2l-1} \in H_0^1(U) \quad (\textit{if}\ \ m = 2l+1). \end{cases}$$

Then

$$(63)\qquad \frac{d^k\mathbf{u}}{dt^k} \in L^\infty(0,T;H^{m+1-k}(U)) \quad (k = 0,\dots,m+1),$$

and we have the estimate

$$(64)\qquad \begin{aligned}&\operatorname*{ess\,sup}_{0\le t\le T} \sum_{k=0}^{m+1} \left\|\frac{d^k\mathbf{u}}{dt^k}\right\|_{H^{m+1-k}(U)} \\ &\le C\left(\sum_{k=0}^{m} \left\|\frac{d^k\mathbf{f}}{dt^k}\right\|_{L^2(0,T;H^{m-k}(U))} + \|g\|_{H^{m+1}(U)} + \|h\|_{H^m(U)}\right).\end{aligned}$$

Remark. In view of Theorem 2 in §5.9.2, we see that

(65) $$\mathbf{f}(0) \in H^{m-1}(U),\ \mathbf{f}'(0) \in H^{m-2}(U), \dots,\ \mathbf{f}^{(m-2)}(0) \in H^1(U),$$

and consequently

(66) $$\begin{aligned} &g_0 \in H^{m+1}(U),\ h_1 \in H^m(U),\ g_2 \in H^{m-1}(U),\ h_3 \in H^{m-2}(U),\\ &\dots, g_{2l} \in H^1(U) \text{ (if } m = 2l),\ h_{2l+1} \in H^1(U) \text{ (if } m = 2l+1). \end{aligned}$$

The compatibility conditions are consequently the requirements that, in addition, each of these functions equals 0 on ∂U, in the trace sense. □

Proof. 1. The proof is by an induction, the case $m = 0$ following from Theorem 5,(i) above.

2. Assume next the theorem is valid for some nonnegative integer m, and suppose

(67) $$\begin{cases} g \in H^{m+2}(U),\ h \in H^{m+1}(U), \\ \frac{d^k \mathbf{f}}{dt^k} \in L^2(0,T; H^{m+1-k}(U)) \quad (k = 0, \dots, m+1). \end{cases}$$

Suppose also the $(m+1)^{th}$-order compatibility conditions obtain. Now set $\tilde{\mathbf{u}} := \mathbf{u}'$. Differentiating the PDE with respect to t, we check that $\tilde{\mathbf{u}}$ is the unique, weak solution of

(68) $$\begin{cases} \tilde{u}_{tt} + L\tilde{u} = \tilde{f} & \text{in } U_T \\ \tilde{u} = 0 & \text{on } \partial U \times [0,T] \\ \tilde{u} = \tilde{g},\ \tilde{u}_t = \tilde{h} & \text{on } U \times \{t = 0\}, \end{cases}$$

for

(69) $$\tilde{f} := f_t,\ \tilde{g} := h,\ \tilde{h} := f(\cdot, 0) - Lg.$$

In particular, for $m = 0$ we rely upon Theorem 5,(ii) to be sure that $\tilde{\mathbf{u}} \in L^2(0,T; H_0^1(U))$, $\tilde{\mathbf{u}}' \in L^2(0,T; L^2(U))$, $\tilde{\mathbf{u}}'' \in L^2(0,T; H^{-1}(U))$.

Since f, g and h satisfy the $(m+1)^{th}$-order compatibility conditions, $\tilde{f}, \tilde{g}$ and $\tilde{h}$ satisfy the m^{th}-order compatibility conditions. Thus applying the induction assumption, we see

$$\frac{d^k \tilde{\mathbf{u}}}{dt^k} \in L^\infty(0,T; H^{m+1-k}(U)) \quad (k = 0, \dots, m+1),$$

with the estimate

$$\begin{aligned} &\operatorname*{ess\,sup}_{0 \le t \le T} \sum_{k=0}^{m+1} \left\| \frac{d^k \tilde{\mathbf{u}}}{dt^k} \right\|_{H^{m+1-k}(U)} \\ &\quad \le C \left(\sum_{k=0}^{m} \left\| \frac{d^k \tilde{\mathbf{f}}}{dt^k} \right\|_{L^2(0,T; H^{m-k}(U))} + \|\tilde{g}\|_{H^{m+1}(U)} + \|\tilde{h}\|_{H^m(U)} \right). \end{aligned}$$

Since $\tilde{\mathbf{u}} = \mathbf{u}'$, we can rewrite:

$$\text{(70)} \quad \operatorname*{ess\,sup}_{0\le t\le T} \sum_{k=1}^{m+2} \left\| \frac{d^k\mathbf{u}}{dt^k} \right\|_{H^{m+2-k}(U)}$$
$$\le C\left(\sum_{k=1}^{m+1} \left\| \frac{d^k\mathbf{f}}{dt^k} \right\|_{L^2(0,T;H^{m+1-k}(U))} + \|h\|_{H^{m+1}(U)} + \|Lg\|_{H^m(U)} + \|\mathbf{f}(0)\|_{H^m(U)} \right)$$
$$\le C\left(\sum_{k=0}^{m+1} \left\| \frac{d^k\mathbf{f}}{dt^k} \right\|_{L^2(0,T;H^{m+1-k}(U))} + \|g\|_{H^{m+2}(U)} + \|h\|_{H^{m+1}(U)} \right).$$

Here we used the inequality

$$\|\mathbf{f}\|_{C([0,T];H^m(U))} \le C(\|\mathbf{f}\|_{L^2(0,T;H^m(U))} + \|\mathbf{f}'\|_{L^2(0,T;H^m(U))}),$$

which follows from Theorem 2 in §5.9.2.

3. Now write for a.e. $0 \le t \le T$: $L\mathbf{u} = \mathbf{f} - \mathbf{u}'' =: \mathbf{h}$. We have

$$\begin{aligned} \|\mathbf{u}\|_{H^{m+2}(U)} &\le C(\|\mathbf{h}\|_{H^m(U)} + \|\mathbf{u}\|_{L^2(U)}) \\ &\le C(\|\mathbf{f}\|_{H^m(U)} + \|\mathbf{u}''\|_{H^m(U)} + \|\mathbf{u}\|_{L^2(U)}). \end{aligned}$$

Taking the essential supremum with respect to t, adding this inequality to (70) and making standard estimates, we deduce:

$$\operatorname*{ess\,sup}_{0\le t\le T} \sum_{k=0}^{m+2} \left\| \frac{d^k\mathbf{u}}{dt^k} \right\|_{H^{m+2-k}(U)}$$
$$\le C\left(\sum_{k=0}^{m+1} \left\| \frac{d^k\mathbf{f}}{dt^k} \right\|_{L^2(0,T;H^{m+1-k}(U))} + \|g\|_{H^{m+2}(U)} + \|h\|_{H^{m+1}(U)} \right).$$

This is the assertion of the theorem for $m+1$. □

THEOREM 7 (Infinite differentiability). *Assume*

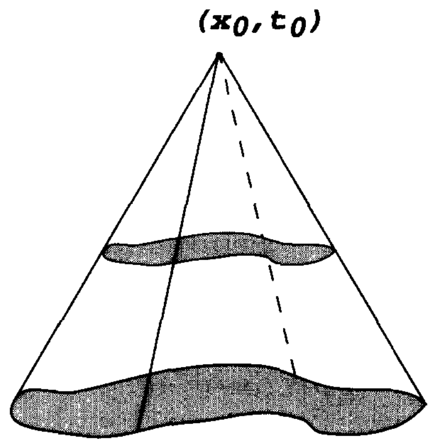
$$g, h \in C^\infty(\bar{U}), \ f \in C^\infty(\bar{U}_T),$$

and the m^{th}-order compatibility conditions hold for $m = 0, 1, \ldots$.

Then the hyperbolic initial/boundary-value problem (1) *has a unique solution*

$$u \in C^\infty(\bar{U}_T).$$

Proof. Apply Theorem 6 for $m = 0, 1, \ldots$. □



Domain of dependence

7.2.4. Propagation of disturbances.

Our study of second-order hyperbolic equations has thus far pretty much paralleled our treatment of second-order parabolic PDE, in §7.1. In the corresponding earlier section §7.1.4, we discussed maximum principles for second-order parabolic equations, and noted in particular that the strong maximum principle implies an "infinite propagation speed" of initial disturbances for such PDE. Now strong maximum principles are false for second-order hyperbolic partial differential equations, and we will instead address here the opposite phenomenon, namely the "finite propagation speed" of initial disturbances. This study extends some ideas already introduced in §2.4.3.

For simplicity we will consider in this section the case $U = \mathbb{R}^n$ and L has the simple nondivergence form

$$Lu = -\sum_{i,j=1}^{n} a^{ij} u_{x_i x_j}, \tag{71}$$

where the coefficients are smooth, independent of time, and there are no lower-order terms. We as usual require the uniform hyperbolicity condition.

Let us assume now u is a smooth solution of the PDE

$$u_{tt} + Lu = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty). \tag{72}$$

We wish to prove a uniqueness/finite propagation speed assertion analogous to that obtained for the wave equation in §2.4.3. For this, we fix a point $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$, and then try to find some sort of a curved "cone-like" region C, with vertex (x_0, t_0), such that $u \equiv 0$ within C if $u \equiv u_t \equiv 0$ on $C_0 = C \cap \{t = 0\}$.

Motivated by the geometric optics computation in Example 3 of §4.5.3, let us guess that the boundary of such a region C is given as a level set $\{p = 0\}$, where p solves the Hamilton–Jacobi PDE

$$p_t - \left(\sum_{i,j=1}^{n} a^{ij}(x)p_{x_i}p_{x_j}\right)^{1/2} = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty). \tag{73}$$

We will simplify (73) by separating variables, to write

$$p(x,t) = q(x) + t - t_0 \quad (x \in \mathbb{R}^n,\ 0 \le t \le t_0), \tag{74}$$

where q solves

$$\begin{cases} \sum_{i,j=1}^{n} a^{ij} q_{x_i} q_{x_j} = 1,\ q > 0 & \text{in } \mathbb{R}^n - \{x_0\} \\ q(x_0) = 0. \end{cases} \tag{75}$$

We henceforth assume that q is a smooth solution of (75) on $\mathbb{R}^n - \{x_0\}$. (In fact $q(x)$ = distance of x to x_0, in the Riemannian metric determined by $((a^{ij}))$.) We write

$$C := \{(x,t) \mid p(x,t) < 0\} = \{(x,t) \mid q(x) < t_0 - t\}.$$

For each $t > 0$, we further define

$$C_t := \{x \mid q(x) < t_0 - t\} = \text{ cross section of } C \text{ at time } t. \tag{76}$$

Since (75) implies $Dq \neq 0$ in $\mathbb{R}^n - \{x_0\}$, ∂C_t is a smooth, $(n-1)$-dimensional surface for $0 \le t < t_0$.

THEOREM 8 (Finite propagation speed). *Assume u is a smooth solution of the hyperbolic equation (72). If $u \equiv u_t \equiv 0$ on C_0, then $u \equiv 0$ within the cone C.*

Remark. We see in particular that if u is a solution of (72) with the initial conditions

$$u = g,\ u_t = h \quad \text{on } \mathbb{R}^n \times \{t = 0\}, \tag{77}$$

then $u(x_0, t_0)$ depends only upon the values of g and h within C_0. □

Proof. 1. We modify a proof from §2.4.3, and so define the energy

$$e(t) := \frac{1}{2}\int_{C_t} u_t^2 + \sum_{i,j=1}^{n} a^{ij} u_{x_i} u_{x_j}\, dx \quad (0 \le t \le t_0).$$

2. In order to compute $\dot{e}(t)$, we first note that if f is a continuous function of x, then

$$\frac{d}{dt}\left(\int_{C_t} f\,dx\right) = -\int_{\partial C_t} \frac{f}{|Dq|}\,dS$$

according to the coarea formula from §C.3. Thus

$$\begin{aligned}\dot{e}(t) &= \int_{C_t} u_t u_{tt} + \sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j t}\,dx \\ &\quad - \frac{1}{2}\int_{\partial C_t}\left(u_t^2 + \sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j}\right)\frac{1}{|Dq|}\,dS \\ &=: A - B.\end{aligned} \tag{78}$$

Integrating by parts, we calculate

$$\begin{aligned}A &= \int_{C_t} u_t\left(u_{tt} - \sum_{i,j=1}^n (a^{ij} u_{x_i})_{x_j}\right)dx + \int_{\partial C_t} \sum_{i,j=1}^n a^{ij} u_{x_i} \nu^j u_t\,dS \\ &= -\int_{C_t} u_t \sum_{i,j=1}^n a^{ij}_{x_j} u_{x_i}\,dx + \int_{\partial C_t} \sum_{i,j=1}^n a^{ij} u_{x_i} \nu^j u_t\,dS,\end{aligned} \tag{79}$$

with $\boldsymbol{\nu} = (\nu^1, \dots, \nu^n)$ being as usual the outer unit normal to ∂C_t. But according to the generalized Cauchy-Schwarz inequality (§B.2)

$$\left|\sum_{i,j=1}^n a^{ij} u_{x_i} \nu^j\right| \leq \left(\sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j}\right)^{1/2}\left(\sum_{i,j=1}^n a^{ij} \nu^i \nu^j\right)^{1/2}. \tag{80}$$

In addition, since $q = t_0 - t$ on ∂C_t, we have $\boldsymbol{\nu} = \frac{Dq}{|Dq|}$ on ∂C_t. Hence

$$\sum_{i,j=1}^n a^{ij} \nu^i \nu^j = \sum_{i,j=1}^n \frac{a^{ij} q_{x_i} q_{x_j}}{|Dq|^2} = \frac{1}{|Dq|^2}$$

by (75). Consequently inequality (80) reads

$$\left|\sum_{i,j=1}^n a^{ij} u_{x_i} \nu^j\right| \leq \left(\sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j}\right)^{1/2} \frac{1}{|Dq|}. \tag{81}$$

Then returning to (79), we estimate using (81) and Cauchy's inequality:

$$\begin{aligned}|A| &\leq Ce(t) + \int_{\partial C_t}\left(\sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j}\right)^{1/2} |u_t|\,\frac{1}{|Dq|}\,dS \\ &\leq Ce(t) + \frac{1}{2}\int_{\partial C_t}\left(u_t^2 + \sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j}\right)\frac{1}{|Dq|}\,dS \\ &= Ce(t) + B.\end{aligned}$$

3. Therefore inequality (78) gives

$$\dot{e}(t) \leq Ce(t).$$

Since hypothesis (76) implies $e(0) = 0$, we deduce using Gronwall's inequality that

$$e(t) = 0 \quad \text{for all } 0 \leq t \leq t_0.$$

Hence $u_t \equiv Du \equiv 0$ in C, and consequently $u \equiv 0$ in C.

□

7.2.5. Equations in two variables.

In this section we briefly consider second-order hyperbolic partial differential equations involving only two variables, and demonstrate that in this setting rather more precise information can be obtained. The very rough idea is that since a function of two variables has "only" three second partial derivatives, algebraic and analytic simplifications in the structure of the PDE may be possible, which are unavailable for more than two variables.

We begin by considering a general linear second-order PDE in two variables

$$\sum_{i,j=1}^{2} a^{ij} u_{x_i x_j} + \sum_{i=1}^{2} b^i u_{x_i} + cu = 0, \tag{82}$$

where the coefficients a^{ij}, b^i, c $(i,j = 1,2)$, with $a^{ij} = a^{ji}$, and the unknown u are functions of the two variables x_1 and x_2 in some region $U \subset \mathbb{R}^2$. Note that for the moment, and in contrast to the theory developed above, we do *not* identify either x_1 or x_2 with the variable t denoting time.

We now pose the following basic question: *Is it possible to simplify the structure of the* PDE (82) *by introducing new independent variables*? In other words, can we expect to convert the PDE into some "nicer" form by rewriting in terms of new variables $y = \mathbf{\Phi}(x)$?

More precisely, let us set

$$\begin{cases} y_1 = \Phi^1(x_1, x_2) \\ y_2 = \Phi^2(x_1, x_2) \end{cases} \tag{83}$$

for some appropriate function $\mathbf{\Phi} = (\Phi^1, \Phi^2)$. To investigate this possibility let us now write

$$u(x) = v(\mathbf{\Phi}(x)). \tag{84}$$

That is, we define $v(y) := u(\mathbf{\Psi}(y))$, where $\mathbf{\Psi} = \mathbf{\Phi}^{-1}$.

From (84), we compute

$$\begin{cases} u_{x_i} = \sum_{k=1}^2 v_{y_k} \Phi^k_{x_i} \\ u_{x_i x_j} = \sum_{k,l=1}^2 v_{y_k y_l} \Phi^k_{x_i} \Phi^l_{x_j} + \sum_{k=1}^2 v_{y_k} \Phi^k_{x_i x_j} \end{cases}$$

for $i, j = 1, 2$. Substituting into (82), we discover v solves the PDE

$$\sum_{k,l=1}^{n} \tilde{a}^{kl} v_{y_k y_l} + \cdots = 0, \tag{85}$$

for

$$\tilde{a}^{kl} := \sum_{i,j=1}^{2} a^{ij} \Phi^k_{x_i} \Phi^l_{x_j} \quad (k, l = 1, 2), \tag{86}$$

where the dots in (85) represent terms of lower order.

We examine the first term in the PDE (85) in the hope we can perhaps choose the transformation $\mathbf{\Phi} = (\Phi^1, \Phi^2)$ so this expression is particularly simple. Let us try to achieve

$$\tilde{a}^{11} \equiv \tilde{a}^{22} \equiv 0. \tag{87}$$

In view of formula (86) this will be possible provided we can choose both Φ^1 and Φ^2 to solve the nonlinear first-order PDE

$$a^{11}(v_{x_1})^2 + 2a^{12} v_{x_1} v_{x_2} + a^{22}(v_{x_2})^2 = 0 \quad \text{in } U. \tag{88}$$

Observe this is the *characteristic equation* associated with the partial differential equation (82), as discussed in §4.6.2.

To proceed further, let us suppose

$$\det \mathbf{A} = a^{11}a^{22} - (a^{12})^2 < 0 \quad \text{in } U; \tag{89}$$

in which case we say the PDE (82) is *hyperbolic.*

Utilizing condition (89), we can then factor equation (88) as follows:

$$\begin{aligned} &\left(a^{11} v_{x_1} + \left[a^{12} + \left((a^{12})^2 - a^{11}a^{22}\right)^{1/2}\right] v_{x_2}\right) \\ &\quad \cdot \left(a^{11} v_{x_1} + \left[a^{12} - \left((a^{12})^2 - a^{11}a^{22}\right)^{1/2}\right] v_{x_2}\right) \\ &\quad = a^{11}\left(a^{11}(v_{x_1})^2 + 2a^{12} v_{x_1} v_{x_2} + a^{22}(v_{x_2})^2\right) = 0. \end{aligned} \tag{90}$$

Now the left hand side of (90) is the product of two linear first-order PDE:

$$(91_1) \qquad a^{11}v_{x_1} + \left[a^{12} + \left((a^{12})^2 - a^{11}a^{22}\right)^{1/2}\right] v_{x_2} = 0 \quad \text{in } U$$

and

$$(91_2) \qquad a^{11}v_{x_1} + \left[a^{12} - \left((a^{12})^2 - a^{11}a^{22}\right)^{1/2}\right] v_{x_2} = 0 \quad \text{in } U.$$

We now assume that we can choose Φ^1 to be a smooth solution of the PDE (91_1), with $D\Phi^1 \neq 0$ in U. Then Φ^1 is constant along trajectories $\mathbf{x} = (x^1, x^2)$ of the ODE

$$(92) \qquad \begin{cases} \dot{x}^1 = a^{11} \\ \dot{x}^2 = \left[a^{12} + \left((a^{12})^2 - a^{11}a^{22}\right)^{1/2}\right]. \end{cases}$$

Similarly, suppose Φ^2 is a smooth solution of (91_2), with $D\Phi^2 \neq 0$ in U; then Φ^2 is constant along trajectories $\mathbf{x} = (x^1, x^2)$ of the ODE

$$(93) \qquad \begin{cases} \dot{x}^1 = a^{11} \\ \dot{x}^2 = \left[a^{12} - \left((a^{12})^2 - a^{11}a^{22}\right)^{1/2}\right]. \end{cases}$$

Curves which are trajectories of either the ODE (92) or (93) are called *characteristics* of the original partial differential equation (82). Returning now to (83) we see that trajectories of solutions of the characteristic ODE (92) and (93) provide our new coordinate lines.

Additionally we can verify using (89) that

$$(94) \qquad \tilde{a}^{12} = \sum_{i,j=1}^{2} a^{ij}\Phi^1_{x_i}\Phi^2_{x_j} \neq 0 \quad \text{in } U.$$

Combining then (85), (86), (87) and (94), we see that our PDE (82) becomes in the y coordinates

$$(95) \qquad v_{y_1y_2} + \cdots = 0,$$

the dots as before denoting terms of lower order. Let us call equation (95) the *first canonical form* for the hyperbolic PDE (82).

If we change variables again by setting $z_1 = y_1 + y_2$, $z_2 = y_1 - y_2$, then (95) becomes

$$(96) \qquad w_{z_1z_1} - w_{z_2z_2} + \cdots = 0.$$

If we then further rename the variables $t = z_1$, $x = z_2$, then (96) reads

$$w_{tt} - w_{xx} + \cdots = 0, \tag{97}$$

the second-order term of which is the one-dimensional wave operator. Equation (97) is the *second canonical form.*

Hence any hyperbolic PDE in two variables of the form (82) can be converted by a change of variables into the wave equation plus a lower order term, assuming we can find the mapping $\mathbf{\Phi}$ as above.

7.3. HYPERBOLIC SYSTEMS OF FIRST-ORDER EQUATIONS

We next broaden our study of hyperbolic PDE (which we may informally interpret as equations supporting "wave-like" solutions) to the case of first-order systems. We continue in the manner of §§7.1 and 7.2 by first employing energy bounds to construct weak solutions for symmetric hyperbolic systems. For nonsymmetric, constant coefficient hyperbolic systems, however, we will instead employ Fourier transform methods.

7.3.1. Definitions.

We investigate in this section *systems* of linear first-order partial differential equations having the form

$$\mathbf{u}_t + \sum_{j=1}^{n} \mathbf{B}_j \mathbf{u}_{x_j} = \mathbf{f} \quad \text{in } \mathbb{R}^n \times (0, \infty), \tag{1}$$

subject to the initial condition

$$\mathbf{u} = \mathbf{g} \quad \text{on } \mathbb{R}^n \times \{t = 0\}. \tag{2}$$

The unknown is $\mathbf{u} : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \dots, u^m)$, and the functions $\mathbf{B}_j : \mathbb{R}^n \times [0, \infty) \to \mathbb{M}^{m \times m}$ $(j = 1, \dots, n)$, $\mathbf{f} : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^m$, $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ are given.

Notation. For each $y \in \mathbb{R}^n$, set

$$\mathbf{B}(x, t; y) := \sum_{j=1}^{n} y_j \mathbf{B}_j(x, t) \quad (x \in \mathbb{R}^n,\ t \geq 0).$$

□

DEFINITION. *The system of* PDE (1) *is called* hyperbolic *if the* $m \times m$ *matrix* $\mathbf{B}(x,t;y)$ *is diagonalizable for each* $x, y \in \mathbb{R}^n$, $t \geq 0$.

In other words, (1) is hyperbolic provided for each x, y, t, the matrix $\mathbf{B}(x,t,y)$ has m real eigenvalues

$$\lambda_1(x,t;y) \leq \lambda_2(x,t;y) \leq \cdots \leq \lambda_m(x,t;y)$$

and corresponding eigenvectors $\{\mathbf{r}_k(x,t;y)\}_{k=1}^m$ that form a basis of $\mathbb{R}^m$.

There are two important special cases:

DEFINITIONS. (i) *We say* (1) *is a* symmetric hyperbolic *system if* $\mathbf{B}_j(x,t)$ *is a symmetric* $m \times m$ *matrix for each* $x \in \mathbb{R}^n$, $t \geq 0$ $(j = 1, \ldots, m)$.

(ii) *The system* (1) *is* strictly hyperbolic *if for each* $x, y \in \mathbb{R}^n$, $y \neq 0$, *and each* $t \geq 0$, *the matrix* $\mathbf{B}(x,t;y)$ *has* m *distinct real eigenvalues:*

$$\lambda_1(x,t;y) < \lambda_2(x,t;y) < \cdots < \lambda_m(x,t;y).$$

Motivation for the definition of hyperbolicity. We justify the hyperbolicity condition as follows. Assume $\mathbf{f} \equiv 0$ and, further, the matrices B_j are constant $(j = 1, \ldots, n)$. Thus

$$\sum_{j=1}^n y_j B_j = \mathbf{B}(y) \tag{3}$$

depends only on $y \in \mathbb{R}^n$.

As in §4.2 let us look for a *plane wave* solution of (1), (2). That is, we seek a solution $\mathbf{u}$ having the form

$$\mathbf{u}(x,t) = \mathbf{v}(y \cdot x - \sigma t) \quad (x \in \mathbb{R}^n,\ t \geq 0) \tag{4}$$

for some direction $y \in \mathbb{R}^n$, velocity $\frac{\sigma}{|y|}$ $(\sigma \in \mathbb{R})$, and profile $\mathbf{v} : \mathbb{R} \to \mathbb{R}^m$. Plugging (4) into (1), we compute:

$$\left(-\sigma I + \sum_{j=1}^n y_j B_j\right) \mathbf{v}' = 0.$$

This equality asserts $\mathbf{v}'$ is an eigenvector of the matrix $\mathbf{B}(y)$ corresponding to the eigenvalue σ.

The hyperbolicity condition requires that there are m distinct plane wave solutions of (1) for each direction y. These are

$$(y \cdot x - \lambda_k(y)t)\mathbf{r}_k(y) \quad (k = 1, \ldots, m),$$

where

$$\lambda_1(y) \leq \lambda_2(y) \leq \cdots \leq \lambda_m(y)$$

are the eigenvalues of $\mathbf{B}(y)$ and $\{\mathbf{r}_k(y)\}_{k=1}^m$ the corresponding eigenvectors. The eigenvalues for $|y| = 1$ are the *wave speeds*. □

7.3.2. Symmetric hyperbolic systems.

In this section we apply energy methods and the vanishing viscosity technique to build a solution to the hyperbolic initial-value problem

$$\begin{cases} \mathbf{u}_t + \sum_{j=1}^n \mathbf{B}_j \mathbf{u}_{x_j} = \mathbf{f} & \text{in } \mathbb{R}^n \times (0,T] \\ \mathbf{u} = \mathbf{g} & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases} \tag{5}$$

where $T > 0$, under the fundamental assumption that

$$\text{the matrices } \mathbf{B}_j(x,t) \text{ are symmetric} \quad (j = 1, \dots, n), \tag{6}$$

for $x \in \mathbb{R}^n$, $0 \leq t \leq T$. We will further assume $\mathbf{B}_j \in C^2(\mathbb{R}^n \times [0,T]; \mathbb{M}^{m \times m})$, with

$$\sup_{\mathbb{R}^n \times [0,T]} (|\mathbf{B}_j|, |D_{x,t}\mathbf{B}_j|, |D^2_{x,t}\mathbf{B}_j|) < \infty \quad (j = 1, \dots, n), \tag{7}$$

and

$$\mathbf{g} \in H^1(\mathbb{R}^n; \mathbb{R}^m), \ \mathbf{f} \in H^1(\mathbb{R}^n \times (0,T); \mathbb{R}^m). \tag{8}$$

Remark. More general systems having the form

$$\mathbf{B}_0 \mathbf{u}_t + \sum_{j=1}^n \mathbf{B}_j \mathbf{u}_{x_j} = \mathbf{f} \tag{9}$$

are also called *symmetric*, provided the matrices $\mathbf{B}_j$ are symmetric for $j = 0, \dots, n$. The theory set forth below easily extends to such systems, provided $\mathbf{B}_0$ is positive definite.

Symmetric hyperbolic systems of the type (9) generalize the second-order hyperbolic PDE studied in §7.2. For suppose v is a smooth solution of the scalar equation

$$v_{tt} - \sum_{i,j=1}^n a^{ij} v_{x_i x_j} = 0, \tag{10}$$

where without loss we may take $a^{ij} = a^{ji}$ $(i,j = 1, \dots, n)$. Writing

$$\mathbf{u} = (u^1, \dots, u^{n+1}) := (v_{x_1}, \dots, v_{x_n}, v_t),$$

we discover $\mathbf{u}$ solves a system of the form (9), for $m = n+1$, $\mathbf{f} \equiv 0$,

$$\mathbf{B}_j = \begin{pmatrix} 0 & \dots & 0 & -a^{1j} \\ & \ddots & & \vdots \\ 0 & \dots & 0 & -a^{nj} \\ -a^{1j} & \dots & -a^{nj} & 0 \end{pmatrix}_{(n+1)\times(n+1)} \quad (j = 1, \dots, n),$$

$$\mathbf{B}_0 = \begin{pmatrix} a^{11} & \dots & a^{1n} & 0 \\ & \ddots & & \vdots \\ a^{1n} & \dots & a^{nn} & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}_{(n+1)\times(n+1)}.$$

Observe that the uniform hyperbolicity condition for (10) implies that the matrix $\mathbf{B}_0$ is positive definite. □

a. Weak solutions.

To ease notation, let us define the bilinear form

$$B[\mathbf{u},\mathbf{v};t] := \int_{\mathbb{R}^n} \sum_{j=1}^{n} (\mathbf{B}_j(\cdot,t)\mathbf{u}_{x_j}) \cdot \mathbf{v}\,dx$$

for $0 \leq t \leq T$, $\mathbf{u},\mathbf{v} \in H^1(\mathbb{R}^n;\mathbb{R}^m)$.

DEFINITION. *We say*

$$\mathbf{u} \in L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m)), \text{ with } \mathbf{u}' \in L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m)),$$

is a weak solution *of the initial-value problem* (5) *for the symmetric hyperbolic system provided*

(i) $(\mathbf{u}',\mathbf{v}) + B[\mathbf{u},\mathbf{v};t] = (\mathbf{f},\mathbf{v})$

for each $\mathbf{v} \in H^1(\mathbb{R}^n;\mathbb{R}^m)$ *and a.e.* $0 \leq t \leq T$, *and*

(ii) $\mathbf{u}(0) = \mathbf{g}$.

Here and afterwards $(\ ,\ )$ denotes the inner product in $L^2(\mathbb{R}^n;\mathbb{R}^m)$.

Remark. According to Theorem 2 in §5.9.2, $\mathbf{u} \in C([0,T];L^2(\mathbb{R}^n;\mathbb{R}^m))$ and so the initial condition (ii) makes sense. □

b. Vanishing viscosity method.

We will approximate problem (5) by the parabolic initial-value problem

$$(11) \qquad \begin{cases} \mathbf{u}^\epsilon_t - \epsilon\Delta\mathbf{u}^\epsilon + \sum_{j=1}^n \mathbf{B}_j\mathbf{u}^\epsilon_{x_j} = \mathbf{f} & \text{in } \mathbb{R}^n \times (0,T] \\ \mathbf{u}^\epsilon = \mathbf{g}^\epsilon & \text{on } \mathbb{R}^n \times \{t=0\} \end{cases}$$

for $0 < \epsilon \leq 1$, $\mathbf{g}^\epsilon := \eta_\epsilon * \mathbf{g}$. The idea is that for each $\epsilon > 0$, problem (11) has a unique smooth solution $\mathbf{u}^\epsilon$, which converges to zero as $|x| \to \infty$. The plan is to show that as $\epsilon \to 0$, the $\mathbf{u}^\epsilon$ converge to a limit function $\mathbf{u}$, which is a weak solution of (5).

THEOREM 1 (Existence of approximate solutions). *For each* $\epsilon > 0$, *there exists a unique solution* $\mathbf{u}^\epsilon$ *of* (11), *with*

$$(12) \qquad \mathbf{u}^\epsilon \in L^2(0,T;H^3(\mathbb{R}^n;\mathbb{R}^m)),\ \mathbf{u}^{\epsilon\prime} \in L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m)).$$

Proof. 1. Set $X = L^\infty((0,T); H^1(\mathbb{R}^n;\mathbb{R}^m))$. For each $\mathbf{v} \in X$, consider the linear system

$$\begin{cases} \mathbf{u}_t - \epsilon\Delta\mathbf{u} = \mathbf{f} - \sum_{j=1}^n \mathbf{B}_j \mathbf{v}_{x_j} & \text{in } \mathbb{R}^n \times (0,T] \\ \mathbf{u} = \mathbf{g}^\varepsilon & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

As the right hand side is bounded in L^2, there exists a unique solution $\mathbf{u} \in L^2(0,T; H^2(\mathbb{R}^n;\mathbb{R}^m))$, $\mathbf{u}' \in L^2(0,T; L^2(\mathbb{R}^n;\mathbb{R}^m))$. Indeed, we can utilize the fundamental solution Φ of the heat equation (§2.3.1) to represent $\mathbf{u}^\varepsilon$ in terms of $\mathbf{g}^\varepsilon$ and $\mathbf{f} - \sum_{j=1}^n \mathbf{B}_j \mathbf{v}_{x_j}$.

Similarly, take $\tilde{\mathbf{v}} \in X$ and let $\tilde{\mathbf{u}}$ solve

$$\begin{cases} \tilde{\mathbf{u}}_t - \epsilon\Delta\tilde{\mathbf{u}} = \mathbf{f} - \sum_{j=1}^n \mathbf{B}_j \tilde{\mathbf{v}}_{x_j} & \text{in } \mathbb{R}^n \times (0,T] \\ \tilde{\mathbf{u}} = \mathbf{g}^\varepsilon & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

2. Subtracting, we find $\hat{\mathbf{u}} := \mathbf{u} - \tilde{\mathbf{u}}$ satisfies

$$\text{(13)} \qquad \begin{cases} \hat{\mathbf{u}}_t - \epsilon\Delta\hat{\mathbf{u}} = -\sum_{j=1}^n \mathbf{B}_j \hat{\mathbf{v}}_{x_j} & \text{in } \mathbb{R}^n \times (0,T] \\ \hat{\mathbf{u}} = \mathbf{0} & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases}$$

for $\hat{\mathbf{v}} := \mathbf{v} - \tilde{\mathbf{v}}$. From the representation formula of $\hat{\mathbf{u}}$ in terms of the fundamental solution Φ and $\sum_{j=1}^n \mathbf{B}_j \hat{\mathbf{v}}_{x_j}$, we obtain the estimate:

$$\text{(14)} \qquad \begin{aligned} \operatorname*{ess\,sup}_{0\le t\le T} \|\hat{\mathbf{u}}(t)\|_{H^1(\mathbb{R}^n;\mathbb{R}^m)} &\le C(\epsilon)\Big\|\sum_{j=1}^n \mathbf{B}_j \hat{\mathbf{v}}_{x_j}\Big\|_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))} \\ &\le C(\epsilon)\|\hat{\mathbf{v}}\|_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))} \\ &\le C(\epsilon)T^{1/2} \operatorname*{ess\,sup}_{0\le t\le T} \|\hat{\mathbf{v}}(t)\|_{H^1(\mathbb{R}^n;\mathbb{R}^m)}. \end{aligned}$$

Thus

$$\text{(15)} \qquad \|\hat{\mathbf{u}}\| \le C(\epsilon)T^{1/2}\|\hat{\mathbf{v}}\|.$$

3. If T is so small that

$$\text{(16)} \qquad C(\epsilon)T^{1/2} \le 1/2,$$

then (15) reads $\|\mathbf{u} - \tilde{\mathbf{u}}\| \le \frac{1}{2}\|\mathbf{v} - \tilde{\mathbf{v}}\|$. According to Banach's fixed point theorem (§9.2.1) the mapping $\mathbf{v} \mapsto \mathbf{u}$ has a unique fixed point. Then $\mathbf{u} = \mathbf{u}^\epsilon$ solves (11), provided (16) holds.

If (16) fails, we choose $0 < T_1 < T$ so that $CT_1 = 1/2$ and repeat the above argument on the time intervals $[0, T_1], [T_1, 2T_1]$, etc.

Assertion (12) follows from parabolic regularity theory (cf. §7.2.3). □

c. Energy estimates.

We intend to send $\varepsilon \to 0$ in (11), and for this as usual need some uniform estimates.

THEOREM 2 (Energy estimates). *There exists a constant C, depending only on n and the coefficients, such that*

$$\begin{aligned}(17)\quad &\max_{0\le t\le T}\|\mathbf{u}^\epsilon(t)\|_{H^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{u}^{\epsilon'}\|_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}\\ &\le C(\|g\|_{H^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))} + \|\mathbf{f}'\|_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))})\end{aligned}$$

for each $0 < \epsilon \le 1$.

Proof. 1. We compute

$$(18)\qquad \frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) = (\mathbf{u}^\epsilon, \mathbf{u}^{\epsilon'}) = \left(\mathbf{u}^\epsilon, \mathbf{f} - \sum_{j=1}^n B_j \mathbf{u}^\epsilon_{x_j} + \epsilon\Delta\mathbf{u}^\epsilon\right).$$

Now

$$(19)\qquad |(\mathbf{u}^\epsilon, \mathbf{f})| \le \|\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}$$

and

$$(20)\qquad (\mathbf{u}^\epsilon, \epsilon\Delta\mathbf{u}^\epsilon) = -\epsilon\|D\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} \le 0.$$

2. Suppose $\mathbf{v} \in C^\infty_c(\mathbb{R}^n;\mathbb{R}^m)$. Then

$$\begin{aligned}\left(\mathbf{v}, \sum_{j=1}^n \mathbf{B}_j\mathbf{v}_{x_j}\right) &= \sum_{j=1}^n \int_{\mathbb{R}^n} (\mathbf{B}_j\mathbf{v}_{x_j})\cdot\mathbf{v}\,dx\\ &= \frac{1}{2}\sum_{j=1}^n \int_{\mathbb{R}^n} ((\mathbf{B}_j\mathbf{v})\cdot\mathbf{v})_{x_j}\,dx - \frac{1}{2}\sum_{j=1}^n\int_U (\mathbf{B}_{j,x_j}\mathbf{v})\cdot\mathbf{v}\,dx,\end{aligned}$$

the last equality following from the symmetry assumption (6). As $\mathbf{v}$ has compact support, we deduce using (7) that

$$\left|(\mathbf{v}, \sum_{j=1}^n \mathbf{B}_j\mathbf{v}_{x_j})\right| \le \frac{1}{2}\sum_{j=1}^n\left|\int_{\mathbb{R}^n}(\mathbf{B}_{j,x_j}\mathbf{v})\cdot\mathbf{v}\,dx\right| \le C\|\mathbf{v}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}.$$

By approximation therefore:

$$\left|(\mathbf{u}^\epsilon, \sum_{j=1}^n \mathbf{B}_j\mathbf{u}^\epsilon_{x_j})\right| \le C\|\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}.$$

Utilizing this bound, (19) and (20) in (18), we obtain the estimate

$$\frac{d}{dt}\left(\|\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) \le C\left(\|\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right).$$

We next apply Gronwall's inequality, to deduce

$$\text{(21)} \quad \max_{0\le t\le T} \|\mathbf{u}^\epsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le C\left(\|\mathbf{g}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|^2_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}\right),$$

since $\|\mathbf{g}^\epsilon\|_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le \|\mathbf{g}\|_{L^2(\mathbb{R}^n;\mathbb{R}^m)}$.

3. Fix $k \in \{1,\dots,n\}$ and write $\mathbf{v}^k := \mathbf{u}^\epsilon_{x_k}$. Differentiating (11) with respect to x_k, we find

$$\begin{cases} \mathbf{v}^k_t - \epsilon\Delta\mathbf{v}^k + \sum_{j=1}^n \mathbf{B}_j\mathbf{v}^k_{x_j} = \mathbf{f}_{x_k} - \sum_{j=1}^n \mathbf{B}_{j,x_k}\mathbf{u}^\epsilon_{x_j} & \text{in } \mathbb{R}^n\times(0,T] \\ \mathbf{v}^k = \mathbf{g}^\varepsilon_{x_k} & \text{on } \mathbb{R}^n\times\{t=0\}. \end{cases}$$

Reasoning as above, we find

$$\text{(22)} \quad \frac{d}{dt}\left(\|\mathbf{v}^k\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) \le C\left(\|D\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} + \|D\mathbf{f}\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})}\right).$$

Sum the previous inequalities for $k = 1,\dots,n$, to deduce

$$\frac{d}{dt}\left(\|D\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})}\right) \le C\left(\|D\mathbf{u}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} + \|D\mathbf{f}\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})}\right).$$

Gronwall's inequality now provides the bound

$$\text{(23)} \quad \begin{aligned} &\max_{0\le t\le T} \|D\mathbf{u}^\epsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} \\ &\quad \le C(\|D\mathbf{g}\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} + \|\mathbf{f}\|^2_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))}), \end{aligned}$$

since $\|D\mathbf{g}^\epsilon\|_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} \le \|D\mathbf{g}\|_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})}$.

4. Next set $\mathbf{v} := \mathbf{u}^{\epsilon'}$ and differentiate (11) with respect to t, to discover

$$\text{(24)} \quad \begin{cases} \mathbf{v}_t - \epsilon\Delta\mathbf{v} + \sum_{j=1}^n \mathbf{B}_j\mathbf{v}_{x_j} = \mathbf{f}' - \sum_{j=1}^n \mathbf{B}_{j,t}\mathbf{u}^\epsilon_{x_i} & \text{in } \mathbb{R}^n\times(0,T] \\ \mathbf{v} = \mathbf{f} - \sum_{j=1}^n \mathbf{B}_j\mathbf{g}^\epsilon_{x_j} + \epsilon\Delta\mathbf{g}^\epsilon & \text{on } \mathbb{R}^n\times\{t=0\}. \end{cases}$$

Reasoning as before, we compute

$$\text{(25)} \quad \begin{aligned} \max_{0\le t\le T} \|\mathbf{u}^{\epsilon'}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} &\le C\big(\|D\mathbf{g}\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})} + \epsilon^2\|\Delta\mathbf{g}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \\ &+ \|\mathbf{f}(0)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))} + \|\mathbf{f}'\|^2_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}\big). \end{aligned}$$

Now

$$\text{(26)} \quad \|\Delta\mathbf{g}^\epsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le \frac{C}{\epsilon^2}\|D\mathbf{g}\|^2_{L^2(\mathbb{R}^n;\mathbb{M}^{m\times n})},$$

since $\mathbf{g}^\epsilon = \eta_\epsilon * \mathbf{g}$. Furthermore

$$\text{(27)} \quad \|\mathbf{f}(0)\|_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le C(\|\mathbf{f}\|^2_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))} + \|\mathbf{f}'\|^2_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}).$$

This bound, together with (21) and (23), completes the proof. □

d. Existence and uniqueness.

THEOREM 3 (Existence of weak solution). *There exists a weak solution of the initial value problem* (5).

Proof. 1. According to the energy estimates (17) there exists a subsequence $\epsilon_k \to 0$ and a function $\mathbf{u} \in L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))$, such that $\mathbf{u}' \in L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))$, with

$$(28)\qquad \begin{cases} \mathbf{u}^{\epsilon_k} \rightharpoonup \mathbf{u} & \text{weakly in } L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m)) \\ \mathbf{u}^{\epsilon_k'} \rightharpoonup \mathbf{u}' & \text{weakly in } L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m)). \end{cases}$$

2. Choose a function $\mathbf{v} \in C^1([0,T];H^1(\mathbb{R}^n;\mathbb{R}^m))$. Then from (11) we deduce:

$$(29)\qquad \int_0^T (\mathbf{u}^{\epsilon'},\mathbf{v}) + \epsilon D\mathbf{u}^\epsilon : D\mathbf{v} + B[\mathbf{u}^\epsilon,\mathbf{v};t]\,dt = \int_0^T (\mathbf{f},\mathbf{v})\,dt.$$

Let $\epsilon = \epsilon_k \to 0$:

$$(30)\qquad \int_0^T (\mathbf{u}',\mathbf{v}) + B[\mathbf{u},\mathbf{v};t]\,dt = \int_0^T (\mathbf{f},\mathbf{v})\,dt.$$

This identity is valid for all $\mathbf{v} \in C([0,T];H^1(\mathbb{R}^n;\mathbb{R}^m))$, and so

$$(\mathbf{u}',\mathbf{v}) + B[\mathbf{u},\mathbf{v};t] = (\mathbf{f},\mathbf{v})$$

for a.e. t and each $\mathbf{v} \in H^1(\mathbb{R}^n;\mathbb{R}^m)$.

3. Assume now $\mathbf{v}(T) = 0$. Then (29) implies

$$\int_0^T -(\mathbf{u}^\epsilon,\mathbf{v}') + \epsilon D\mathbf{u}^\epsilon : D\mathbf{v} + B[\mathbf{u}^\epsilon,\mathbf{v};t]\,dt = \int_0^T (\mathbf{f},\mathbf{v})\,dt + (\mathbf{g}^\epsilon,\mathbf{v}(0)).$$

Upon sending $\epsilon = \epsilon_k \to 0$ we obtain:

$$\int_0^T -(\mathbf{u},\mathbf{v}) + B[\mathbf{u},\mathbf{v};t]\,dt = \int_0^T (\mathbf{f},\mathbf{v})\,dt + (\mathbf{g},\mathbf{v}(0)).$$

Integrating by parts in (30) gives us the identity

$$\int_0^T -(\mathbf{u},\mathbf{v}) + B[\mathbf{u},\mathbf{v};t]\,dt = \int_0^T (\mathbf{f},\mathbf{v})\,dt + (\mathbf{u}(0),\mathbf{v}(0)).$$

Consequently $\mathbf{u}(0) = \mathbf{g}$, as $\mathbf{v}(0)$ is arbitrary. □

THEOREM 4 (Uniqueness of weak solution). *A weak solution of* (5) *is unique.*

Proof. 1. It suffices to show the only weak solution of (5) with $\mathbf{f} \equiv \mathbf{g} \equiv 0$ is $\mathbf{u} \equiv 0$.

To verify this, note

$$(\mathbf{u}', \mathbf{u}) + B[\mathbf{u}, \mathbf{u}; t] = 0 \quad \text{for a.e. } 0 \le t \le T. \tag{31}$$

Since $|B[\mathbf{u}, \mathbf{u}; t]| \le C\|\mathbf{u}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}$, we as usual compute from (31) that

$$\frac{d}{dt}\left(\|\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) \le C\|\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)};$$

whence Gronwall's inequality forces $\|\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} = 0$ $(0 \le t \le T)$, since $\mathbf{u}(0) = 0$. □

7.3.3. Systems with constant coefficients.

In this section we apply the Fourier transform (§4.3) to solve the *constant coefficient* system

$$\mathbf{u}_t + \sum_{j=1}^n B_j \mathbf{u}_{x_j} = \mathbf{0} \quad \text{in } \mathbb{R}^n \times (0, \infty), \tag{32}$$

with the initial condition

$$\mathbf{u} = \mathbf{g} \quad \text{on } \mathbb{R}^n \times \{t = 0\}. \tag{33}$$

We assume that the $\{B_j\}_{j=1}^n$ are constant $m \times m$ matrices and that the $m \times m$ matrix

$$\mathbf{B}(y) := \sum_{j=1}^n y_j B_j \tag{34}$$

has for each $y \in \mathbb{R}^n$ m *real* eigenvalues

$$\lambda_1(y) \le \lambda_2(y) \le \cdots \le \lambda_m(y). \tag{35}$$

There is no hypothesis concerning the eigenvectors, and so we are supposing only a very weak sort of hyperbolicity here. We also make no assumption of symmetry for the matrices $\{B_j\}_{j=1}^n$. Consequently the foregoing energy estimates do not apply. We need a new tool, which we discover in the Fourier transform.

THEOREM 5 (Existence of solution). *Assume*

$$\mathbf{g} \in H^s(\mathbb{R}^n; \mathbb{R}^m) \quad \left(s > \frac{n}{2} + m\right).$$

Then there is a unique solution $\mathbf{u} \in C^1([0,\infty); \mathbb{R}^m)$ *of the initial-value problem* (32), (33).

See §5.8.4 for the definition of the fractional Sobolev spaces H^s.

Proof. 1. We apply the Fourier transform (§4.3.1), as follows. First, temporarily assume $\mathbf{u} = (u^1, \dots, u^m)$ is a smooth solution. Then set

$$\hat{\mathbf{u}} = (\hat{u}^1, \dots, \hat{u}^m),$$

where ˆ denotes the Fourier transform in the variable x: we do not transform with respect to the time variable t. Equation (32) becomes

$$\hat{\mathbf{u}}_t + i \sum_{j=1}^n y_j B_j \hat{\mathbf{u}} = \mathbf{0};$$

that is,

$$\hat{\mathbf{u}}_t + i\mathbf{B}(y)\hat{\mathbf{u}} = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty). \tag{36}$$

In addition

$$\hat{\mathbf{u}} = \hat{\mathbf{g}} \quad \text{on } \mathbb{R}^n \times \{t = 0\}. \tag{37}$$

For each fixed $y \in \mathbb{R}^n$ we solve (36), (37) by integrating in time, to find

$$\hat{\mathbf{u}}(y,t) = e^{-it\mathbf{B}(y)}\hat{\mathbf{g}}(y) \quad (y \in \mathbb{R}^n, t \geq 0). \tag{38}$$

Consequently $\mathbf{u} = (e^{-it\mathbf{B}(y)}\hat{\mathbf{g}})^\vee$; so that

$$\mathbf{u}(x,t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot y} e^{-it\mathbf{B}(y)}\hat{\mathbf{g}}(y)\, dy \quad (x \in \mathbb{R}^n,\ t \geq 0). \tag{39}$$

2. We have derived formula (39) assuming $\mathbf{u}$ to be a smooth solution of (32), (33). We now verify that the function $\mathbf{u}$ defined by (39) is in truth a solution, and so must first check that the integral in (39) converges.

Since $\mathbf{g} \in H^s(\mathbb{R}^n; \mathbb{R}^m)$, we know according to §5.8.4 that there exists $\mathbf{f} \in L^2(\mathbb{R}^n; \mathbb{R}^m)$ such that

$$|\hat{\mathbf{g}}(y)| \leq C(1 + |y|^s)^{-1}|\mathbf{f}(y)| \quad (y \in \mathbb{R}^n). \tag{40}$$

So in order to investigate the convergence of the integral (39), we must estimate $\|e^{-itB(y)}\|$.

3. For a fixed y, let Γ denote the path $\partial B(0,r)$ in the complex plane, traversed counterclockwise, the radius r selected so large that the eigenvalues $\lambda_1(y), \dots, \lambda_m(y)$ lie within Γ.

We have the formula:

$$e^{-it\mathbf{B}(y)} = \frac{1}{2\pi i}\int_\Gamma e^{-itz}(zI - \mathbf{B}(y))^{-1}dz. \tag{41}$$

To verify this, let $\mathbf{A}(t,y)$ denote the right hand side of (41) and fix $x \in \mathbb{R}^m$. Then

$$\begin{aligned}\mathbf{B}(y)\mathbf{A}(t,y)x &= \frac{1}{2\pi i}\int_\Gamma e^{-itz}\mathbf{B}(y)(zI - B(y))^{-1}x\,dz\\ &= \frac{1}{2\pi i}\int_\Gamma e^{-itz}(z(zI - \mathbf{B}(y))^{-1}x - x)\,dz\\ &= -\frac{1}{i}\frac{d}{dt}\mathbf{A}(t,y)x,\end{aligned}$$

since $\int_\Gamma e^{-itz}dz = 0$. Consequently

$$\left(\frac{d}{dt} + i\mathbf{B}(y)\right)\mathbf{A}(t,y) = 0. \tag{42}$$

In addition

$$\begin{aligned}\mathbf{A}(0,y)x &= \frac{1}{2\pi i}\int_\Gamma (zI - \mathbf{B}(y))^{-1}x\,dz\\ &= \frac{1}{2\pi i}\int_\Gamma \frac{x + \mathbf{B}(y)(zI - \mathbf{B}(y))^{-1}x}{z}\,dz\\ &= x + \frac{1}{2\pi i}\int_\Gamma \frac{\mathbf{B}(y)(zI - \mathbf{B}(y))^{-1}x}{z}\,dz.\end{aligned} \tag{43}$$

Now set

$$w := (zI - \mathbf{B}(y))^{-1}x\,; \tag{44}$$

so that $zw - \mathbf{B}(y)w = x$. Taking the product with $\bar{w}$, we deduce $|w| \leq \frac{C}{|z|}$ for some constant C. Using this estimate and letting r go to infinity, we conclude from (43) that $\mathbf{A}(0,y)x = x$. This equality and (42) verify the representation formula (41).

4. Define a new path Δ in the complex plane as follows. For fixed y, draw circles $B_k = B(\lambda_k(y), 1)$ of radius 1, centered at $\lambda_k(y)$ $(k = 1, \dots, m)$. Then take Δ to be the boundary of $\bigcup_{k=1}^m B_k$, traversed counterclockwise.

Deforming the path Γ into Δ, we deduce from (41) that

$$e^{-it\mathbf{B}(y)} = \frac{1}{2\pi i}\int_{\Delta} e^{-itz}(zI - \mathbf{B}(y))^{-1}dz. \tag{45}$$

Now

$$|e^{-itz}| \leq e^{t} \quad (z \in \Delta). \tag{46}$$

Furthermore

$$\det(zI - \mathbf{B}(y)) = \prod_{k=1}^{m}(z - \lambda_k(y));$$

whence

$$|\det(zI - \mathbf{B}(y))| \geq 1 \quad \text{if } z \in \Delta. \tag{47}$$

Now

$$(zI - \mathbf{B}(y))^{-1} = \frac{\operatorname{cof}(zI - \mathbf{B}(y))^{T}}{\det(zI - \mathbf{B}(y))},$$

where "cof" denotes the cofactor matrix (see §8.1.4). We deduce

$$\begin{aligned} \|(zI - \mathbf{B}(y))^{-1}\| &\leq \|\operatorname{cof}(zI - \mathbf{B}(y))\| \\ &\leq C(1 + |z|^{m-1} + \|\mathbf{B}(y)\|^{m-1}) \\ &\leq C(1 + |y|^{m-1}) \quad \text{if } z \in \Delta. \end{aligned} \tag{48}$$

We have utilized in this calculation the elementary inequality

$$|\lambda_k(y)| \leq C|y| \quad (k = 1, \dots, m).$$

Combining (45)–(48), we derive the estimate

$$\|e^{-it\mathbf{B}(y)}\| \leq Ce^{t}(1 + |y|^{m-1}) \quad (y \in \mathbb{R}^n). \tag{49}$$

5. Return now to (37). We deduce using (40), (49) that

$$\begin{aligned} \int_{\mathbb{R}^n} |e^{ix\cdot y}e^{-it\mathbf{B}(y)}\hat{\mathbf{g}}(y)|\,dy &\leq C\int_{\mathbb{R}^n} \|e^{-it\mathbf{B}(y)}\|(1 + |y|^s)^{-1}|\mathbf{f}(y)|\,dy \\ &\leq Ce^{t}\int_{\mathbb{R}^n} |\mathbf{f}(y)|(1 + |y|^{m-1})(1 + |y|^s)^{-1}dy \\ &\leq C\left(\int_{\mathbb{R}^n} |\mathbf{f}|^2 dy\right)^{1/2}\left(\int_{\mathbb{R}^n} \frac{dy}{1 + |y|^{2(s-m+1)}}\right)^{1/2} \\ &< \infty, \end{aligned}$$

since $s > \frac{n}{2} + m - 1$. Hence the integral in (39) converges, and it follows easily that the function

$$\mathbf{u}(x,t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot y} e^{-it\mathbf{B}(y)} \hat{\mathbf{g}}(y)\, dy$$

is continuous on $\mathbb{R}^n \times [0,\infty)$.

6. To show $\mathbf{u}$ is C^1, observe for $0 < |h| \leq 1$ that

$$\frac{\mathbf{u}(x,t+h) - \mathbf{u}(x,t)}{h} = \frac{1}{(2\pi)^{n/2} h} \int_{\mathbb{R}^n} e^{ix\cdot y} (e^{-i(t+h)\mathbf{B}(y)} - e^{-it\mathbf{B}(y)}) \hat{\mathbf{g}}(y)\, dy.$$

Since

$$e^{-i(t+h)\mathbf{B}(y)} - e^{-it\mathbf{B}(y)} = -i \int_t^{t+h} \mathbf{B}(y) e^{-is\mathbf{B}(y)}\, ds,$$

we can estimate as above that

$$\left| \frac{1}{h} (e^{-i(t+h)\mathbf{B}(y)} - e^{-it\mathbf{B}(y)}) \right| \leq Ce^{t+1}(1 + |y|^m).$$

Therefore

$$\left| \frac{\mathbf{u}(x,t+h) - \mathbf{u}(x,t)}{h} \right| \leq Ce^{t+1} \int_{\mathbb{R}^n} |\mathbf{f}(y)|(1 + |y|^m)(1 + |y|^s)^{-1} dy,$$

and the integrand is summable since $s > \frac{n}{2} + m$. Thus $\mathbf{u}_t$ exists and is continuous on $\mathbb{R}^n \times [0,\infty)$. A similar argument shows $\mathbf{u}_{x_i}$ exists and is continuous ($i = 1, \dots, n$). According to the Dominated Convergence Theorem, we can furthermore differentiate under the integral sign in (39), to confirm that $\mathbf{u}$ solves the system $\mathbf{u}_t + \sum_{j=1}^n B_j \mathbf{u}_{x_j} = \mathbf{0}$. □

In Chapter 11 we will encounter *nonlinear* first-order systems of hyperbolic equations.

7.4. SEMIGROUP THEORY

Semigroup theory is the abstract study of first-order ordinary differential equations with values in Banach spaces, driven by linear, but possibly unbounded, operators. In this section we outline the basics of the theory, and present as well two applications to linear PDE. This approach provides an elegant alternative to some of the existence theory for evolution equations set forth in §§7.1–7.3.

7.4.1. Definitions, elementary properties.

We begin in an abstract setting. Let X denote a real Banach space, and consider then the ordinary differential equation

$$\text{(1)} \qquad \begin{cases} \mathbf{u}'(t) = A\mathbf{u}(t) & (t \geq 0) \\ \mathbf{u}(0) = u, \end{cases}$$

where $' = \frac{d}{dt}$, $u \in X$ is given, and A is a linear operator. More precisely, suppose $D(A)$, the *domain* of A, is a linear subspace of X and we are given a possibly unbounded linear operator

$$\text{(2)} \qquad A : D(A) \to X.$$

We investigate the existence and uniqueness of a solution

$$\mathbf{u} : [0, \infty) \to X$$

of the ODE (1). The key problem is to ascertain reasonable conditions on the operator A so that (a) the ODE has a unique solution $\mathbf{u}$ for each initial point $u \in X$, and (b) many interesting PDE can be cast into the abstract form (1). (We have in mind the situation that X is an L^p space of functions and A is a linear partial differential operator involving variables other than t. In this case A is necessarily an unbounded operator.)

a. Semigroups.

Let us for the moment informally assume $\mathbf{u} : [0, \infty) \to X$ is a solution of the differential equation (1), and that (1) in fact has a unique solution for each initial point $u \in X$.

Notation. We will write

$$\text{(3)} \qquad \mathbf{u}(t) := S(t)u \quad (t \geq 0)$$

to display explicitly the dependence of $\mathbf{u}(t)$ on the initial value $u \in X$. For each time $t \geq 0$ we may therefore regard $S(t)$ as a mapping from X into X. □

What properties does the family of operators $\{S(t)\}_{t \geq 0}$ satisfy? Clearly $S(t) : X \to X$ is linear. Furthermore

$$\text{(4)} \qquad S(0)u = u \quad (u \in X)$$

and

$$\text{(5)} \qquad S(t+s)u = S(t)S(s)u = S(s)S(t)u \quad (t, s \geq 0,\ u \in X).$$

Condition (5) is simply our assumption that the ODE (1) has a unique solution for each initial point. Finally, it seems reasonable to suppose that for each $u \in X$

(6) the mapping $t \mapsto S(t)u$ is continuous from $[0, \infty)$ into X.

DEFINITIONS. (i) *A family* $\{S(t)\}_{t\geq 0}$ *of bounded linear operators mapping* X *into* X *is called a* semigroup *if conditions (4)–(6) are satisfied.*

(ii) *We say* $\{S(t)\}_{t\geq 0}$ *is a* contraction semigroup *if in addition*

$$\|S(t)\| \leq 1 \quad (t \geq 0), \tag{7}$$

$\|\ \|$ *here denoting the operator norm. Thus*

$$\|S(t)u\| \leq \|u\| \quad (t \geq 0,\ u \in X).$$

The notion of contraction semigroup captures many properties of a nice flow on X generated by the ODE (1).

b. Elementary properties, generators.

The real problem now is to determine which operators A generate contraction semigroups. We will answer this in §7.4.2, after recording in this section some further general facts.

Henceforth assume $\{S(t)\}_{t\geq 0}$ is a contraction semigroup on X.

DEFINITIONS. *Write*

$$D(A) := \left\{ u \in X \mid \lim_{t\to 0+} \frac{S(t)u - u}{t} \text{ exists in } X \right\} \tag{8}$$

and

$$Au := \lim_{t\to 0+} \frac{S(t)u - u}{t} \quad (u \in D(A)). \tag{9}$$

We call $A : D(A) \to X$ *the* (infinitesimal) generator *of the semigroup* $\{S(t)\}_{t\geq 0}$; $D(A)$ *is the* domain *of* A.

THEOREM 1 (Differential properties of semigroups). *Assume* $u \in D(A)$. *Then*

(i) $S(t)u \in D(A)$ *for each* $t \geq 0$,

(ii) $AS(t)u = S(t)Au$ *for each* $t > 0$,

(iii) *the mapping* $t \mapsto S(t)u$ *is differentiable for each* $t > 0$,

and

(iv) $\frac{d}{dt}S(t)u = AS(t)u \quad (t > 0)$.

Proof. 1. Let $u \in D(A)$. Then

$$\begin{aligned}
&\lim_{s\to 0+} \frac{S(s)S(t)u - S(t)u}{s} \\
&= \lim_{s\to 0+} \frac{S(t)S(s)u - S(t)u}{s} \quad \text{by the semigroup property (5)} \\
&= S(t) \lim_{s\to 0+} \frac{S(s)u - u}{s} = S(t)Au.
\end{aligned}$$

Thus $S(t)u \in D(A)$ and $AS(t)u = S(t)Au$. Assertions (i) and (ii) are proved.

2. Let $u \in D(A)$, $h > 0$. Then if $t > 0$,

$$\begin{aligned}
&\lim_{h\to 0+} \left\{ \frac{S(t)u - S(t-h)u}{h} - S(t)Au \right\} \\
&= \lim_{h\to 0+} \left\{ S(t-h)\left(\frac{S(h)u - u}{h}\right) - S(t)Au \right\} \\
&= \lim_{h\to 0+} \left\{ S(t-h)\left(\frac{S(h)u - u}{h} - Au\right) + (S(t-h) - S(t))Au \right\} = 0,
\end{aligned}$$

since $\frac{S(h)u-u}{h} \to Au$ and $\|S(t-h)\| \leq 1$. Consequently

$$\lim_{h\to 0+} \frac{S(t)u - S(t-h)u}{h} = S(t)Au.$$

Similarly

$$\lim_{h\to 0+} \frac{S(t+h)u - S(t)u}{h} = S(t) \lim_{h\to 0+} \frac{S(h)u - u}{h} = S(t)Au.$$

Thus $\frac{d}{dt}S(t)u$ exists for each time $t > 0$, and equals $S(t)Au = AS(t)u$. □

Remark. Since $t \mapsto AS(t)u = S(t)Au$ is continuous, the mapping $t \mapsto S(t)u$ is C^1 in $(0, \infty)$, if $u \in D(A)$. □

THEOREM 2 (Properties of generators).

(i) *The domain $D(A)$ is dense in X,*

and

(ii) *A is a closed operator.*

Remark. To say A is *closed* means that whenever $u_k \in D(A)$ $(k = 1, \ldots)$ and $u_k \to u$, $Au_k \to v$ as $k \to \infty$, then

$$u \in D(A), \quad v = Au.$$

□

Proof. 1. Fix any $u \in X$ and define then $u^t := \int_0^t S(s)u\,ds$. In view of (6), $\frac{u^t}{t} \to u$ in X, as $t \to 0+$.

2. We claim

$$u^t \in D(A) \quad (t > 0). \tag{10}$$

Indeed if $r > 0$, we have

$$\begin{aligned}\frac{S(r)u^t - u^t}{r} &= \frac{1}{r}\left[S(r)\left(\int_0^t S(s)u\,ds\right) - \left(\int_0^t S(s)u\,ds\right)\right]\\ &= \frac{1}{r}\int_0^t S(r+s)u - S(s)u\,ds,\end{aligned}$$

where we used the semigroup property (5). Thus

$$\begin{aligned}\frac{S(r)u^t - u^t}{r} &= \frac{1}{r}\int_t^{t+r} S(s)u\,ds - \frac{1}{r}\int_0^r S(s)u\,ds\\ &\to S(t)u - u, \quad \text{as } r \to 0+\,.\end{aligned}$$

Hence $u^t \in D(A)$, with $Au^t = S(t)u - u$. This proves (10) and completes the proof of assertion (i).

3. To prove A is closed, let $u_k \in D(A)$ $(k = 1, \dots)$ and suppose

$$u_k \to u,\ Au_k \to v \quad \text{in } X. \tag{11}$$

We must prove $u \in D(A)$, $v = Au$. According to Theorem 1

$$S(t)u_k - u_k = \int_0^t S(s)Au_k\,ds.$$

Let $k \to \infty$ and recall (11):

$$S(t)u - u = \int_0^t S(s)v\,ds.$$

Hence we have

$$\lim_{t\to 0+}\frac{S(t)u - u}{t} = \lim_{t\to 0+}\frac{1}{t}\int_0^t S(s)v\,ds = v.$$

But then by definition $u \in D(A)$, $v = Au$. □

c. Resolvents.

DEFINITIONS. (i) *We say a real number* λ *belongs to* $\rho(A)$, *the* resolvent set *of* A, *provided the operator*

$$\lambda I - A : D(A) \to X$$

is one-to-one and onto.

(ii) *If* $\lambda \in \rho(A)$, *the* resolvent operator $R_\lambda : X \to X$ *is defined by*

$$R_\lambda u := (\lambda I - A)^{-1} u.$$

According to the Closed Graph Theorem (§D.3), $R_\lambda : X \to D(A) \subseteq X$ is a bounded linear operator. Furthermore,

$$AR_\lambda u = R_\lambda A u \quad \text{if } u \in D(A).$$

THEOREM 3 (Properties of resolvent operators).

(i) *If* $\lambda, \mu \in \rho(A)$, *we have*

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu \quad \textit{(resolvent identity)} \tag{12}$$

and

$$R_\lambda R_\mu = R_\mu R_\lambda. \tag{13}$$

(ii) *If* $\lambda > 0$, *then* $\lambda \in \rho(A)$,

$$R_\lambda u = \int_0^\infty e^{-\lambda t} S(t) u \, dt \quad (u \in X), \tag{14}$$

and so $\|R_\lambda\| \leq \frac{1}{\lambda}$.

Remark. Thus *the resolvent operator is the Laplace transform of the semigroup* (cf. Example 4 in §4.3.2). □

Proof. 1. Verification of the identities (12), (13) is left to the reader (Problem 11).

2. Note first that since $\lambda > 0$ and $\|S(t)\| \leq 1$, the integral on the right hand side of (14) is defined. Let $\tilde{R}_\lambda u$ denote this integral. Then for $h > 0$

and $u \in X$,

$$\begin{aligned}\frac{S(h)\tilde{R}_\lambda u - \tilde{R}_\lambda u}{h} &= \frac{1}{h}\left\{\int_0^\infty e^{-\lambda t}[S(t+h)u - S(t)u]\,dt\right\}\\ &= -\frac{1}{h}\int_0^h e^{-\lambda(t-h)}S(t)u\,dt\\ &\quad + \frac{1}{h}\int_0^\infty (e^{-\lambda(t-h)} - e^{-\lambda t})S(t)u\,dt\\ &= -e^{\lambda h}\frac{1}{h}\int_0^h e^{-\lambda t}S(t)u\,dt\\ &\quad + \left(\frac{e^{\lambda h}-1}{h}\right)\int_0^\infty e^{-\lambda t}S(t)u\,dt.\end{aligned}$$

Hence

$$\lim_{h\to 0+}\frac{S(h)\tilde{R}_\lambda u - \tilde{R}_\lambda u}{h} = -u + \lambda\tilde{R}_\lambda u.$$

Thus $A\tilde{R}_\lambda u = -u + \lambda\tilde{R}_\lambda u$; that is,

$$(\lambda I - A)\tilde{R}_\lambda u = u \quad (u \in X). \tag{15}$$

On the other hand if $u \in D(A)$,

$$\begin{aligned}A\tilde{R}_\lambda u = A\int_0^\infty e^{-\lambda t}S(t)u\,dt &= \int_0^\infty e^{-\lambda t}AS(t)u\,dt\\ &= \int_0^\infty e^{-\lambda t}S(t)Au\,dt = \tilde{R}_\lambda Au.\end{aligned} \tag{16}$$

Our passing A under the integral sign is justified since A is a closed operator: see Problem 12. Thus

$$\tilde{R}_\lambda(\lambda I - A)u = u \quad (u \in D(A)).$$

In view of (15) and the formula above $\lambda I - A$ is one-to-one and onto. Consequently $\lambda \in \rho(A)$, $\tilde{R}_\lambda = (\lambda I - A)^{-1} = R_\lambda$. □

7.4.2. Generating contraction semigroups.

We now characterize the generators of contraction semigroups:

THEOREM 4 (Hille–Yosida Theorem). *Let A be a closed, densely-defined linear operator on X. Then A is the generator of a contraction semigroup $\{S(t)\}_{t\geq 0}$ if and only if*

$$(0,\infty) \subset \rho(A) \quad \text{and} \quad \|R_\lambda\| \leq \frac{1}{\lambda} \quad \text{for } \lambda > 0. \tag{17}$$

Proof. 1. If A is a generator, then from Theorem 3,(iii) we immediately deduce (17).

2. Conversely, suppose A is closed, densely-defined, and satisfies (17). We must build a contraction semigroup with A as its generator. For this, fix $\lambda > 0$ and define

$$A_\lambda := -\lambda I + \lambda^2 R_\lambda = \lambda A R_\lambda. \tag{18}$$

The operator A_λ is a kind of regularized approximation to A.

3. We first claim

$$A_\lambda u \to Au \quad \text{as } \lambda \to \infty \quad (u \in D(A)). \tag{19}$$

Indeed, since $\lambda R_\lambda u - u = AR_\lambda u = R_\lambda Au$, $\|\lambda R_\lambda u - u\| \leq \|R_\lambda\| \, \|Au\| \leq \frac{1}{\lambda}\|Au\| \to 0$. Thus $\lambda R_\lambda u \to u$ as $\lambda \to \infty$ if $u \in D(A)$. But since $\|\lambda R_\lambda\| \leq 1$ and $D(A)$ is dense, we deduce then as well

$$\lambda R_\lambda u \to u \quad \text{as } \lambda \to \infty, \text{ for all } u \in X. \tag{20}$$

Now if $u \in D(A)$, then

$$A_\lambda u = \lambda A R_\lambda u = \lambda R_\lambda A u.$$

In view of (20), our claim (19) is proved.

4. Next, define

$$S_\lambda(t) := e^{tA_\lambda} = e^{-\lambda t} e^{\lambda^2 t R_\lambda} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k}{k!} R_\lambda^k.$$

Observe that since $\|R_\lambda\| \leq \lambda^{-1}$,

$$\|S_\lambda(t)\| \leq e^{-\lambda t} \sum_{k=0}^{\infty} \frac{\lambda^{2k} t^k}{k!} \|R_\lambda\|^k \leq e^{-\lambda t} \sum_{k=0}^{\infty} \frac{\lambda^k t^k}{k!} = 1.$$

Consequently $\{S_\lambda(t)\}_{t \geq 0}$ is a contraction semigroup, and it is easy to check its generator is A_λ, with $D(A_\lambda) = X$.

5. Let $\lambda, \mu > 0$. Since $R_\lambda R_\mu = R_\mu R_\lambda$, we see $A_\lambda A_\mu = A_\mu A_\lambda$, and so

$$A_\mu S_\lambda(t) = S_\lambda(t) A_\mu \quad \text{for each } t > 0.$$

Thus if $u \in D(A)$, we can compute

$$\begin{aligned} S_\lambda(t)u - S_\mu(t)u &= \int_0^t \frac{d}{ds}[S_\mu(t-s)S_\lambda(s)u]\, ds \\ &= \int_0^t S_\mu(t-s)S_\lambda(s)(A_\lambda u - A_\mu u)\, ds, \end{aligned}$$

because $\frac{d}{dt}S_\lambda(t)u = A_\lambda S_\lambda(t)u = S_\lambda(t)A_\lambda u$. Consequently $\|S_\lambda(t)u - S_\mu(t)u\| \le t\|A_\lambda u - A_\mu u\| \to 0$ as $\lambda, \mu \to \infty$. Hence

$$\text{(21)} \qquad S(t)u := \lim_{\lambda\to\infty} S_\lambda(t)u \quad \text{exists for each } t \ge 0,\ u \in D(A).$$

As $\|S_\lambda(t)\| \le 1$, the limit (21) in fact exists for all $u \in X$, uniformly for t in compact subsets of $[0, \infty)$. It is now straightforward to verify $\{S(t)\}_{t\ge0}$ is a contraction semigroup on X.

6. It remains to show A is the generator of $\{S(t)\}_{t\ge0}$. Write B to denote this generator. Now

$$\text{(22)} \qquad S_\lambda(t)u - u = \int_0^t S_\lambda(s)A_\lambda u\, ds.$$

In addition

$$\|S_\lambda(s)A_\lambda u - S(s)Au\| \le \|S_\lambda(s)\|\,\|A_\lambda u - Au\| + \|(S_\lambda(s) - S(s))Au\| \to 0$$

as $\lambda \to \infty$, if $u \in D(A)$. Passing therefore to limits in (22), we deduce

$$S(t)u - u = \int_0^t S(s)Au\, ds$$

if $u \in D(A)$. Thus $D(A) \subseteq D(B)$ and

$$Bu = \lim_{t\to0+} \frac{S(t)u - u}{t} = Au \quad (u \in D(A)).$$

Now if $\lambda > 0$, $\lambda \in \rho(A)\cap\rho(B)$. Also $(\lambda I - B)(D(A)) = (\lambda I - A)(D(A)) = X$, according to (17). Hence $(\lambda I - B)|_{D(A)}$ is one-to-one and onto; whence $D(A) = D(B)$. Therefore $A = B$, and so A is indeed the generator of $\{S(t)\}_{t\ge0}$. □

Remark. Let $\omega \in \mathbb{R}$. A semigroup $\{S(t)\}_{t\ge0}$ is called *ω-contractive* if $\|S(t)\| \le e^{\omega t}$ $(t \ge 0)$. An easy variant of Theorem 4 asserts that a closed, densely-defined linear operator A generates an ω-contractive semigroup if and only if

$$\text{(23)} \qquad (\omega, \infty) \subset \rho(A) \quad \text{and} \quad \|R_\lambda\| \le \frac{1}{\lambda - \omega} \text{ for all } \lambda > \omega.$$

This version of the Hille–Yosida Theorem will be required for our first example below. □

7.4.3. Applications.

We demonstrate in this section that certain second-order parabolic and hyperbolic PDE can be realized within the semigroup framework.

a. Second-order parabolic PDE.

We consider first the parabolic initial/boundary-value problem

$$(24)\qquad \begin{cases} u_t + Lu = 0 & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g & \text{on } U \times \{t = 0\}, \end{cases}$$

a special case (corresponding to $f \equiv 0$) of (1) in §7.1.1. We assume L has the divergence structure (2) from §7.1.1, satisfies the usual strong ellipticity condition, and has smooth coefficients, *which do not depend on t.* We additionally suppose that the bounded open set U has a smooth boundary.

We propose to reinterpret (24) as the flow determined by a semigroup on $X = L^2(U)$. For this, we set

$$(25)\qquad D(A) := H_0^1(U) \cap H^2(U),$$

and define

$$(26)\qquad Au := -Lu \quad \text{if } u \in D(A).$$

Clearly then A is an unbounded linear operator on X. Recall from §6.2.2 the energy estimate

$$(27)\qquad \beta\|u\|^2_{H_0^1(U)} \leq B[u,u] + \gamma\|u\|^2_{L^2(U)},$$

for constants $\beta > 0$, $\gamma \geq 0$, where $B[\ ,\ ]$ is the bilinear form associated with L.

THEOREM 5 (Second-order parabolic PDE as semigroups). *The operator A generates a γ-contraction semigroup $\{S(t)\}_{t\geq 0}$ on $L^2(U)$.*

Proof. 1. We must verify the hypotheses of the variant of the Hille–Yosida Theorem mentioned in the concluding Remark in §7.4.2, with γ replacing ω.

First, $D(A)$ given by (25) is clearly dense in $L^2(U)$.

2. We claim now that the operator A is closed. Indeed, let $\{u_k\}_{k=1}^\infty \subset D(A)$ with

$$(28)\qquad u_k \to u, \ Au_k \to f \quad \text{in } L^2(U).$$

According to the regularity Theorem 4 in §6.3.2,

$$\|u_k - u_l\|_{H^2(U)} \leq C(\|Au_k - Au_l\|_{L^2(U)} + \|u_k - u_l\|_{L^2(U)})$$

for all k and l. Thus (28) implies $\{u_k\}_{k=1}^\infty$ is a Cauchy sequence in $H^2(U)$ and so

$$u_k \to u \quad \text{in } H^2(U). \tag{29}$$

Therefore $u \in D(A)$. Furthermore (29) implies $Au_k \to Au$ in $L^2(U)$, and consequently $f = Au$.

3. Next we check the resolvent conditions (23), with γ replacing ω. According to Theorem 3 in §6.2.2, for each $\lambda \geq \gamma$ the boundary-value problem

$$\begin{cases} Lu + \lambda u = f & \text{in } U \\ u = 0 & \text{in } \partial U \end{cases} \tag{30}$$

has a unique weak solution $u \in H_0^1(U)$ for each $f \in L^2(U)$. Owing to elliptic regularity theory, in fact $u \in H^2(U) \cap H_0^1(U)$. Thus $u \in D(A)$. We may now rewrite (30), using (26), and find

$$\lambda u - Au = f. \tag{31}$$

Thus $(\lambda I - A) : D(A) \to X$ is one-to-one and onto, provided $\lambda \geq \gamma$. Hence $\rho(A) \supset [\gamma, \infty)$.

4. Consider the weak form of the boundary-value problem (30):

$$B[u, v] + \lambda(u, v) = (f, v)$$

for each $v \in H_0^1(U)$, where $(\, , \,)$ is the inner product in $L^2(U)$. Set $v = u$ and recall (27) to compute for $\lambda > \gamma$:

$$(\lambda - \gamma)\|u\|_{L^2(U)}^2 \leq \|f\|_{L^2(U)}\|u\|_{L^2(U)}.$$

Hence, since $u = R_\lambda f$, we have the estimate

$$\|R_\lambda f\|_{L^2(U)} \leq \frac{1}{\lambda - \gamma}\|f\|_{L^2(U)}.$$

This bound is valid for all $f \in L^2(U)$ and so

$$\|R_\lambda\| \leq \frac{1}{\lambda - \gamma} \quad (\lambda > \gamma), \tag{32}$$

as required. □

Remark. Semigroup theory provides an elegant method for constructing a solution to the initial/boundary-value problem (24). It is worth noting however that this technique requires that the coefficients a^{ij}, b^i, c $(i, j = 1, \dots, n)$ of L be independent of t. The Galerkin method in §7.1 works without this restriction. On the other hand, semigroup theory constructs at the outset a more regular solution than that produced by the Galerkin technique, at least until we develop regularity theory. □

b. Second-order hyperbolic PDE.

We turn our attention next to the hyperbolic initial/boundary-value problem

(33) $$\begin{cases} u_{tt} + Lu = 0 & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g,\ u_t = h & \text{on } U \times \{t=0\}, \end{cases}$$

for the operator L and open set U as above. We recast (33) as a first-order system by setting $v := u_t$. Then the foregoing reads

$$\begin{cases} u_t = v,\ v_t + Lu = 0 & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g,\ v = h & \text{on } U \times \{t=0\}. \end{cases}$$

We will further assume L has the symmetric form:

$$Lu = -\sum_{i,j=1}^{\infty} (a^{ij} u_{x_i})_{x_j} + cu,$$

where

(34) $$c \geq 0,\ a^{ij} = a^{ji} \quad (i,j = 1,\dots,n).$$

Thus for some constant $\beta > 0$:

(35) $$\beta \|u\|^2_{H^1_0(U)} \leq B[u,u] \quad \text{for all } u \in H^1_0(U).$$

Now take

$$X = H^1_0(U) \times L^2(U),$$

with the norm

(36) $$\|(u,v)\| := (B[u,u] + \|v\|^2_{L^2(U)})^{1/2}.$$

Define

$$D(A) := [H^2(U) \cap H^1_0(U)] \times H^1_0(U)$$

and set

(37) $$A(u,v) := (v, -Lu) \quad \text{for } (u,v) \in D(A).$$

We will show A verifies the hypothesis of the Hille–Yosida Theorem.

THEOREM 6 (Second-order hyperbolic PDE as semigroups). *The operator A generates a contraction semigroup $\{S(t)\}_{t\geq 0}$ on $H_0^1(U) \times L^2(U)$.*

Proof. 1. The domain of A is clearly dense in $H_0^1(U) \times L^2(U)$.

2. To see A is closed, let $\{(u_k, v_k)\}_{k=1}^{\infty} \subset D(A)$, with

$$(u_k, v_k) \to (u, v), \ A(u_k, v_k) \to (f, g) \quad \text{in } H_0^1(U) \times L^2(U).$$

Since $A(u_k, v_k) = (v_k, -Lu_k)$, we conclude $f = v$ and $Lu_k \to -g$ in $L^2(U)$. As in the previous proof, it follows that $u_k \to u$ in $H^2(U)$ and $g = -Lu$. Thus $(u, v) \in D(A)$, $A(u, v) = (v, -Lu) = (f, g)$.

3. Now let $\lambda > 0$, $(f, g) \in X := H_0^1(U) \times L^2(U)$, and consider the operator equation

$$\lambda(u, v) - A(u, v) = (f, g). \tag{38}$$

This is equivalent to the two scalar equations

$$\begin{cases} \lambda u - v = f & (u \in H^2(U) \cap H_0^1(U)), \\ \lambda v + Lu = g & (v \in H_0^1(U)). \end{cases} \tag{39}$$

But (39) implies

$$\lambda^2 u + Lu = \lambda f + g \quad (u \in H^2(U) \cap H_0^1(U)). \tag{40}$$

Since $\lambda^2 > 0$, estimate (35) and the regularity theory imply there exists a unique solution u of (40). Defining then $v := \lambda u - f \in H_0^1(U)$, we have shown that (38) has a unique solution (u, v). Consequently $\rho(A) \supset (0, \infty)$.

4. Whenever (39) holds, we write $(u, v) = R_\lambda(f, g)$. Now from the second equation in (39), we deduce

$$\lambda \|v\|_{L^2}^2 + B[u, v] = (g, v)_{L^2}.$$

Substituting $v = \lambda u - f$, we obtain

$$\begin{aligned} \lambda \left(\|v\|_{L^2}^2 + B[u, u] \right) &= (g, v)_{L^2} + B[u, f] \\ &\leq (\|g\|_{L^2}^2 + B[f, f])^{1/2} (\|v\|_{L^2}^2 + B[u, u])^{1/2}. \end{aligned}$$

Here we used the generalized Cauchy–Schwarz inequality (§B.2), which holds owing to the symmetry condition (34). In light of our definition (36),

$$\|(u, v)\| \leq \frac{1}{\lambda} \|(f, g)\|;$$

and so

$$\|R_\lambda\| \leq \frac{1}{\lambda} \quad (\lambda > 0),$$

as required. □

See Friedman [**FR1**] or Yosida [**Y**] for the theory of *analytic semigroups*. Aspects of *nonlinear* semigroup theory will be developed later, in §9.6.

7.5. PROBLEMS

In the following exercises we assume $U \subset \mathbb{R}^n$ is an open, bounded set, with smooth boundary, and $T > 0$.

1. Prove there is at most one smooth solution of this initial/boundary-value problem for the heat equation with Neumann boundary conditions:
$$\begin{cases} u_t - \Delta u = f & \text{in } U_T \\ \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial U \times [0,T] \\ u = g & \text{on } U \times \{t = 0\}. \end{cases}$$

2. Assume u is a smooth solution of
$$\begin{cases} u_t - \Delta u = 0 & \text{in } U \times (0,\infty) \\ u = 0 & \text{on } \partial U \times [0,\infty) \\ u = g & \text{on } U \times \{t = 0\}. \end{cases}$$
Prove the exponential decay estimate:
$$\|u(\cdot,t)\|_{L^2(U)} \le e^{-\lambda_1 t}\|g\|_{L^2(U)} \quad (t \ge 0),$$
where $\lambda_1 > 0$ is the principal eigenvalue of $-\Delta$ (with zero boundary conditions) on U.

3. (Galerkin's method for Poisson's equation.) Suppose $f \in L^2(U)$ and assume that $u_m = \sum_{k=1}^m d_m^k w_k$ solves
$$\int_U Du_m \cdot Dw_k \, dx = \int_U f \cdot w_k \, dx$$
for $k = 1, \ldots, m$. Show that a subsequence of $\{u_m\}_{m=1}^\infty$ converges weakly in $H_0^1(U)$ to the weak solution u of
$$\begin{cases} -\Delta u = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

4. Assume
$$\begin{cases} \mathbf{u}_k \rightharpoonup \mathbf{u} & \text{in } L^2(0,T; H_0^1(U)) \\ \mathbf{u}_k' \rightharpoonup \mathbf{v} & \text{in } L^2(0,T; H^{-1}(U)). \end{cases}$$
Prove that $\mathbf{v} = \mathbf{u}'$. (Hint: Let $\phi \in C_c^1(0,T)$, $w \in H_0^1(U)$. Then
$$\int_0^T \langle \mathbf{u}_k', \phi w\rangle \, dt = -\int_0^T \langle \mathbf{u}_k, \phi' w\rangle \, dt.)$$

5. Suppose H is a Hilbert space and
$$\mathbf{u}_k \rightharpoonup \mathbf{u} \quad \text{in } L^2(0,T; H).$$

Suppose further we have the uniform bounds

$$\operatorname*{ess\,sup}_{0\le t\le T} \|\mathbf{u}_k(t)\| \le C \quad (k = 1, \dots),$$

for some constant C. Prove

$$\operatorname*{ess\,sup}_{0\le t\le T} \|\mathbf{u}(t)\| \le C.$$

(Hint: If $v \in H$ and $0 \le a \le b \le T$, we have

$$\int_a^b (v, \mathbf{u}_k(t))\, dt \le C\|v\||b-a|.)$$

6. Suppose u is a smooth solution of

$$\begin{cases} u_t - \Delta u + cu = 0 & \text{in } U \times (0,\infty) \\ u = 0 & \text{on } \partial U \times [0,\infty) \\ u = g & \text{on } U \times \{t = 0\} \end{cases}$$

and the function c satisfies

$$c \ge \gamma > 0.$$

Prove

$$|u(x,t)| \le Ce^{-\gamma t} \quad ((x,t) \in U_T).$$

7. Suppose u is a smooth solution of the PDE from Problem 6, that $g \ge 0$, and c is bounded (but *not* necessarily nonnegative). Show that $u \ge 0$. (Hint: What PDE does $v := e^{\lambda t}u$ solve?)
8. Prove inequality (54) in §7.1.3, (59) in §7.2.3. (Hints: Assume u is smooth, $u = 0$ on ∂U. Transform the term $(Lu, -\Delta u)$ by integrating by parts twice. Simplify and then estimate the boundary terms using trace inequalities.)
9. Show there exists at most one smooth solution of this initial/boundary-value problem for the *telegraph equation*

$$\begin{cases} u_{tt} + du_t - u_{xx} = f & \text{in } (0,1) \times (0,T) \\ u = 0 & \text{on } (\{0\} \times [0,T]) \cup (\{1\} \times [0,T]) \\ u = g,\, u_t = h & \text{on } (0,1) \times \{t = 0\}. \end{cases}$$

Here d is a constant.

10. Prove there exists at most one smooth solution u of this problem for the *beam equation*

$$\begin{cases} u_{tt} + u_{xxxx} = 0 & \text{in } (0,1) \times (0,T) \\ u = u_x = 0 & \text{on } (\{0\} \times [0,T]) \cup (\{1\} \times [0,T]) \\ u = g, u_t = h & \text{on } [0,1] \times \{t = 0\}. \end{cases}$$

11. Prove the resolvent identities (12) and (13) in §7.4.1.

12. Justify the equality

$$A \int_0^\infty e^{-\lambda t} S(t)x\, dt = \int_0^\infty e^{-\lambda t} AS(t)x\, dt$$

used in (16) of §7.4.1. (Hint: Approximate the integral by a Riemann sum and recall A is a closed operator.)

13. Define for $t > 0$

$$[S(t)g](x) = \int_{\mathbb{R}^n} \Phi(x-y,t)g(y)dy \quad (x \in \mathbb{R}^n),$$

where $g : \mathbb{R}^n \to \mathbb{R}$ and Φ is the fundamental solution of the heat equation. Also set $S(0)g = g$.

(i) Prove $\{S(t)\}_{t \geq 0}$ is a contraction semigroup on $L^2(\mathbb{R}^n)$.

(ii) Show $\{S(t)\}_{t \geq 0}$ is *not* a contraction semigroup on $L^\infty(\mathbb{R}^n)$.

14. Let $\{S(t)\}_{t \geq 0}$ be a contraction semigroup on X, with generator A. Inductively define $D(A^k) := \{x \in D(A^{k-1}) \mid A^{k-1}x \in D(A)\}$ $(k = 2, \dots)$. Show that if $x \in D(A^k)$ for some k, then $S(t)x \in D(A^k)$ for each $t \geq 0$.

15. Use Problem 14 to prove that if u is the semigroup solution in $X = L^2(U)$ of

$$\begin{cases} u_t - \Delta u = 0 & \text{in } U_T \\ u = 0 & \text{on } \partial U \times [0,T] \\ u = g & \text{on } U \times \{t = 0\}, \end{cases}$$

with $g \in C_c^\infty(U)$, then $u(\cdot, t) \in C^\infty(U)$ for each $0 \leq t \leq T$.

7.6. REFERENCES

Section 7.1 See Ladyzhenskaya [**L**, Chapter 3], J.-L. Lions [**L1**], Lions–Magenes [**L-M**], Wloka [**WL**]. W. Schlag helped me with the proof of Theorems 5, 6. The proof of the parabolic Harnack inequality is similar to calculations found in Davies [**DA**].

The books of Krylov [**KR**], Ladyzhenskaya–Solonnikov–Uraltseva [**L-S-U**], and Lieberman [**LB**] have much more on regularity theory.

Section 7.2 See Ladyzhenskaya [**L,** Chapter 4], J.-L. Lions [**L1**], Lions–Magenes [**L-M**] and Wloka [**WL**].

Section 7.3 The Fourier transform argument is taken from John [**J**]; cf. also Treves [**T,** §15]. D. Serre has shown me a much more precise version of Theorem 5, under the further assumption of strict hyperbolicity.

Section 7.4 See Friedman [**FR1,** Part 2, §1] and Yosida [**Y,** Chapter IX].

Section 7.5 Problem 8: see [**B-E**] for a proof, and see also Ladyzhenskaya–Uraltseva [**L-U,** p. 182].

Chapter 8

THE CALCULUS OF VARIATIONS

8.1. INTRODUCTION

8.1.1. Basic ideas.

We introduce some new ideas by supposing first of all that we wish to solve some partial differential equation, which for simplicity we write in the abstract form

$$A[u] = 0. \tag{1}$$

In this formula $A[\cdot]$ denotes a given, possibly nonlinear partial differential operator and u is the unknown. There is, of course, no general theory for solving all such PDE.

The *calculus of variations* identifies an important class of such nonlinear problems that can be solved using relatively simple techniques from nonlinear functional analysis. This is the class of *variational problems*, that is, PDE of the form (1), where the nonlinear operator $A[\cdot]$ is the "derivative" of an appropriate "energy" functional $I[\cdot]$. Symbolically we write

$$A[\cdot] = I'[\cdot]. \tag{2}$$

Then problem (1) reads

$$I'[u] = 0. \tag{3}$$

The advantage of this new formulation is that we now can recognize solutions of (1) as being critical points of $I[\cdot]$. These in certain circumstances may be relatively easy to find: if, for instance, the functional $I[\cdot]$ has a minimum at u, then presumably (3) is valid and thus u is a solution of the original PDE (1). *The point is that whereas it is usually extremely difficult to solve* (1) *directly, it may be much easier to discover minimum (or maximum, or other critical) points of the functional* $I[\cdot]$.

In addition of course, many of the laws of physics and other scientific disciplines arise directly as variational principles.

8.1.2. First variation, Euler–Lagrange equation.

Suppose now $U \subset \mathbb{R}^n$ is a bounded, open set with smooth boundary ∂U, and we are given a smooth function

$$L : \mathbb{R}^n \times \mathbb{R} \times \bar{U} \to \mathbb{R}.$$

We call L the *Lagrangian.*

Notation. We will write

$$L = L(p, z, x) = L(p_1, \ldots, p_n, z, x_1, \ldots, x_n)$$

for $p \in \mathbb{R}^n$, $z \in \mathbb{R}$, and $x \in U$. Thus "p" is the name of the variable for which we substitute $Dw(x)$ below, and "z" is the variable for which we substitute $w(x)$. We also set

$$\begin{cases} D_pL = (L_{p_1}, \ldots, L_{p_n}) \\ D_zL = L_z \\ D_xL = (L_{x_1}, \ldots, L_{x_n}) \,. \end{cases}$$

This notation will clarify the theory to follow. □

We now make the vague ideas in §8.1.1 more precise by now assuming $I[\cdot]$ to have the explicit form

$$I[w] = \int_U L(Dw(x), w(x), x)\, dx, \tag{4}$$

for smooth functions $w : \bar{U} \to \mathbb{R}$ satisfying, say, the boundary condition

$$w = g \quad \text{on } \partial U. \tag{5}$$

Let us now additionally suppose some particular smooth function u, satisfying the requisite boundary condition $u = g$ on ∂U, happens to be a minimizer of $I[\cdot]$ among all functions w satisfying (5). *We will demonstrate that u is then automatically a solution of a certain nonlinear partial differential equation.*

To confirm this, first choose any smooth function $v \in C_c^\infty(U)$ and consider the real-valued function

$$i(\tau) := I[u + \tau v] \quad (\tau \in \mathbb{R}). \tag{6}$$

Since u is a minimizer of $I[\cdot]$ and $u + \tau v = u = g$ on ∂U, we observe that $i(\cdot)$ has a minimum at $\tau = 0$. Therefore

$$i'(0) = 0. \tag{7}$$

We explicitly compute this derivative (called the *first variation*) by writing out

$$i(\tau) = \int_U L(Du + \tau Dv, u + \tau v, x)\, dx. \tag{8}$$

Thus

$$i'(\tau) = \int_U \sum_{i=1}^n L_{p_i}(Du + \tau Dv, u + \tau v, x) v_{x_i} + L_z(Du + \tau Dv, u + \tau v, x) v\, dx.$$

Let $\tau = 0$, to deduce from (7) that

$$0 = i'(0) = \int_U \sum_{i=1}^n L_{p_i}(Du, u, x) v_{x_i} + L_z(Du, u, x) v\, dx.$$

Finally, since v has compact support, we can integrate by parts and obtain

$$0 = \int_U \left[-\sum_{i=1}^n (L_{p_i}(Du, u, x))_{x_i} + L_z(Du, u, x) \right] v\, dx.$$

As this equality holds for all test functions v, we conclude u solves the nonlinear PDE

$$(9) \qquad -\sum_{i=1}^{n}\left(L_{p_i}(Du,u,x)\right)_{x_i} + L_z(Du,u,x) = 0 \quad \text{in } U.$$

This is the *Euler–Lagrange equation* associated with the energy functional $I[\cdot]$ defined by (4). Observe that (9) is a quasilinear, second-order PDE in divergence form.

In summary, any smooth minimizer of $I[\cdot]$ is a solution of the Euler–Lagrange partial differential equation (9), and thus—conversely—we can try to find a solution of (9) by searching for minimizers of (4).

Example 1 (Dirichlet's principle). Take

$$L(p,z,x) = \tfrac{1}{2}|p|^2.$$

Then $L_{p_i} = p_i$ $(i = 1,\dots,n)$, $L_z = 0$; and so the Euler–Lagrange equation associated with the functional

$$I[w] := \tfrac{1}{2}\int_U |Dw|^2\,dx$$

is

$$\Delta u = 0 \quad \text{in } U.$$

This fact is *Dirichlet's principle*, previously introduced in §2.2.5. □

Example 2 (Generalized Dirichlet's principle). Write

$$L(p,z,x) = \tfrac{1}{2}\sum_{i,j=1}^{n} a^{ij}(x)p_ip_j - zf(x),$$

where $a^{ij} = a^{ji}$ $(i,j = 1,\dots,n)$. Then $L_{p_i} = \sum_{j=1}^{n} a^{ij}(x)p_j$ $(i = 1,\dots,n)$, $L_z = -f(x)$. Hence the Euler–Lagrange equation associated with the functional

$$I[w] := \int_U \tfrac{1}{2}\sum_{i,j=1}^{n} a^{ij}w_{x_i}w_{x_j} - wf\,dx$$

is the divergence structure linear equation

$$-\sum_{i,j=1}^{n}\left(a^{ij}u_{x_j}\right)_{x_i} = f \quad \text{in } U.$$

We will see later (in §8.1.3 and §8.2) that the uniform ellipticity condition on the a^{ij} $(i,j=1,\dots,n)$ is a natural further assumption, required to prove the existence of a minimizer. Consequently from the nonlinear viewpoint of the calculus of variations, the divergence structure form of a linear second-order elliptic PDE is completely natural. This observation provides some much belated motivation for the bilinear form techniques utilized in Chapter 6. □

Example 3 (Nonlinear Poisson equation). Given a smooth function $f: \mathbb{R} \to \mathbb{R}$, define its antiderivative $F(z) = \int_0^z f(y)\,dy$. Then the Euler–Lagrange equation associated with the functional

$$I[w] := \int_U \frac{1}{2}|Dw|^2 - F(w)\,dx$$

is the nonlinear Poisson equation

$$-\Delta u = f(u) \quad \text{in } U.$$

□

Example 4 (Minimal surfaces). Let

$$L(p,z,x) = (1+|p|^2)^{1/2}\ ;$$

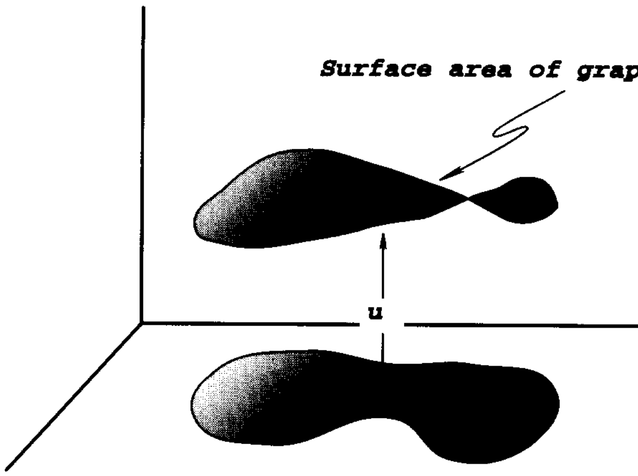
so that

$$I[w] = \int_U (1+|Dw|^2)^{1/2}\,dx$$

is the area of the graph of the function $w: U \to \mathbb{R}$. The associated Euler–Lagrange equation is

$$\sum_{i=1}^n \left(\frac{u_{x_i}}{(1+|Du|^2)^{1/2}} \right)_{x_i} = 0 \quad \text{in } U. \tag{10}$$

This partial differential equation is the *minimal surface equation*. The expression $\operatorname{div}\left(\frac{Du}{(1+|Du|^2)^{1/2}}\right)$ on the left side of (10) is n times the *mean curvature* of the graph of u. Thus *a minimal surface has zero mean curvature.* □



A minimal surface

8.1.3. Second variation.

We continue in the spirit of the calculations from §8.1.2 by computing now the *second variation* of $I[\cdot]$ at the function u. This we find by observing that since u gives a minimum for $I[\cdot]$, we must have

$$i''(0) \geq 0,$$

$i(\cdot)$ defined as above by (6). In view of (8) we can calculate

$$\begin{aligned} i''(\tau) = \int_U \sum_{i,j=1}^n & L_{p_i p_j}(Du + \tau Dv, u + \tau v, x) v_{x_i} v_{x_j} \\ & + 2 \sum_{i=1}^n L_{p_i z}(Du + \tau Dv, u + \tau v, x) v_{x_i} v \\ & + L_{zz}(Du + \tau Dv, u + \tau v, x) v^2 \, dx. \end{aligned}$$

Setting $\tau = 0$, we derive the inequality

$$\begin{aligned} 0 \leq i''(0) = \int_U \sum_{i,j=1}^n & L_{p_i p_j}(Du, u, x) v_{x_i} v_{x_j} \\ & + 2 \sum_{i=1}^n L_{p_i z}(Du, u, x) v_{x_i} v + L_{zz}(Du, u, x) v^2 \, dx, \end{aligned} \tag{11}$$

holding for all test functions $v \in C_c^\infty(U)$.

We can extract useful information from inequality (11), as follows. First, note after a routine approximation argument that estimate (11) is valid for any Lipschitz continuous function v vanishing on ∂U. We then fix $\xi \in \mathbb{R}^n$ and define

$$v(x) := \epsilon\rho\left(\frac{x\cdot\xi}{\epsilon}\right)\zeta(x) \quad (x \in U), \tag{12}$$

where $\zeta \in C_c^\infty(U)$ and $\rho : \mathbb{R} \to \mathbb{R}$ is the periodic "zig-zag" function defined by

$$\rho(x) = \begin{cases} x & \text{if } 0 \le x \le \frac{1}{2} \\ 1 - x & \text{if } \frac{1}{2} \le x \le 1 \end{cases}, \quad \rho(x+1) = \rho(x) \quad (x \in \mathbb{R}).$$

Thus

$$|\rho'| = 1 \quad \text{a.e.} \tag{13}$$

Observe further $v_{x_i}(x) = \rho'\left(\frac{x\cdot\xi}{\epsilon}\right)\xi_i\zeta + O(\epsilon)$ as $\epsilon \to 0$, and so our substituting (12) into (11) yields

$$0 \le \int_U \sum_{i,j=1}^n L_{p_ip_j}(Du, u, x)(\rho')^2\xi_i\xi_j\zeta^2\,dx + O(\epsilon).$$

We recall (13) and then send $\epsilon \to 0$, thereby obtaining the inequality

$$0 \le \int_U \sum_{i,j=1}^n L_{p_ip_j}(Du, u, x)\xi_i\xi_j\zeta^2\,dx.$$

Since this estimate holds for all $\zeta \in C_c^\infty(U)$, we deduce

$$\sum_{i,j=1}^n L_{p_ip_j}(Du, u, x)\xi_i\xi_j \ge 0 \quad (\xi \in \mathbb{R}^n,\ x \in U). \tag{14}$$

We will see later in §8.2 that this necessary condition contains a clue as to the basic convexity assumption on the Lagrangian L required for the existence theory.

8.1.4. Systems.

a. Euler–Lagrange equations.

The foregoing considerations generalize quite easily to the case of systems, the only new complications being largely notational. Recall from §A.1 that $\mathbb{M}^{m\times n}$ is the space of real $m \times n$ matrices, and assume the smooth Lagrangian function

$$L : \mathbb{M}^{m\times n} \times \mathbb{R}^m \times \bar{U} \to \mathbb{R}$$

is given.

Notation. We will write

$$L = L(P, z, x) = L(p_1^1, \dots, p_n^m, z^1, \dots, z^m, x_1, \dots, x_n)$$

for $P \in \mathbb{M}^{m\times n}$, $z \in \mathbb{R}^m$, and $x \in U$, where

$$P = \begin{pmatrix} p_1^1 & \dots & p_n^1 \\ & \ddots & \\ p_1^m & \dots & p_n^m \end{pmatrix}_{m\times n} .$$

(We are now employing superscripts to denote rows: this notational convention simplifies the following formulas.) □

As in §8.1.2 we associate with L the functional

$$(15) \qquad I[\mathbf{w}] := \int_U L(D\mathbf{w}(x), \mathbf{w}(x), x)\, dx,$$

defined for smooth functions $\mathbf{w} : \bar{U} \to \mathbb{R}^m$, $\mathbf{w} = (w^1, \dots, w^m)$, satisfying the boundary conditions $\mathbf{w} = \mathbf{g}$ on ∂U, $\mathbf{g} : \partial U \to \mathbb{R}^m$ being given. Here

$$D\mathbf{w}(x) = \begin{pmatrix} w_{x_1}^1 & \dots & x_{x_n}^1 \\ & \ddots & \\ w_{x_1}^m & \dots & w_{x_n}^m \end{pmatrix}_{m\times n}$$

is the gradient matrix of $\mathbf{w}$ at x.

Let us now show that any smooth minimizer $\mathbf{u} = (u^1, \dots, u^m)$ of $I[\cdot]$, taken among functions equal to $\mathbf{g}$ on ∂U, must solve a certain *system* of nonlinear partial differential equations. We therefore fix $\mathbf{v} = (v^1, \dots, v^m) \in C_c^\infty(U; \mathbb{R}^m)$ and write

$$i(\tau) := I[\mathbf{u} + \tau \mathbf{v}].$$

As before

$$i'(0) = 0.$$

From this we readily deduce as above the equality

$$0 = i'(0) = \int_U \sum_{i=1}^n \sum_{k=1}^m L_{p_i^k}(D\mathbf{u}, \mathbf{u}, x) v_{x_i}^k + \sum_{k=1}^m L_{z^k}(D\mathbf{u}, \mathbf{u}, x) v^k \, dx .$$

As this identity is valid for all choices of $v^1, \dots, v^m$, we conclude after integrating by parts:

$$(16) \quad -\sum_{i=1}^n \left(L_{p_i^k}(D\mathbf{u}, \mathbf{u}, x) \right)_{x_i} + L_{z^k}(D\mathbf{u}, \mathbf{u}, x) = 0 \quad \text{in } U \ (k = 1, \dots, m).$$

This coupled, quasilinear *system* of PDE comprise the *Euler–Lagrange equations* for the functional $I[\cdot]$ defined by (15).

b. Null Lagrangians.

Surprisingly, it turns out to be interesting to study certain systems of nonlinear partial differential equations, for which *every* smooth function is a solution.

DEFINITION. *The function L is called a* null Lagrangian *if the system of Euler–Lagrange equations*

$$(17)\quad -\sum_{i=1}^{n}\left(L_{p_i^k}(D\mathbf{u},\mathbf{u},x)\right)_{x_i} + L_{z^k}(D\mathbf{u},\mathbf{u},x) = 0 \quad \text{in } U\ (k=1,\dots,m)$$

is automatically solved by all smooth functions $\mathbf{u}: U \to \mathbb{R}^m$.

The importance of null Lagrangians is that the corresponding energy

$$I[\mathbf{w}] = \int_U L(D\mathbf{w},\mathbf{w},x)\,dx$$

depends only on the boundary conditions:

THEOREM 1 (Null Lagrangians and boundary conditions). *Let L be a null Lagrangian. Assume* $\mathbf{u},\tilde{\mathbf{u}}$ *are two functions in* $C^2(\bar{U},\mathbb{R}^m)$ *such that*

$$(18)\quad \mathbf{u} \equiv \tilde{\mathbf{u}} \quad \text{on } \partial U\ .$$

Then

$$(19)\quad I[\mathbf{u}] = I[\tilde{\mathbf{u}}]\ .$$

Proof. Define

$$i(\tau) := I[\tau\mathbf{u} + (1-\tau)\tilde{\mathbf{u}}] \quad (0 \le \tau \le 1)\ .$$

Then

$$\begin{aligned}
i'(\tau) = &\int_U \sum_{i=1}^{n}\sum_{k=1}^{m} L_{p_i^k}(\tau D\mathbf{u} + (1-\tau)D\tilde{\mathbf{u}}, \tau\mathbf{u} + (1-\tau)\tilde{\mathbf{u}}, x)(u^k_{x_i} - \tilde{u}^k_{x_i}) \\
&+ \sum_{k=1}^{m} L_{z^k}(\tau D\mathbf{u} + (1-\tau)D\tilde{\mathbf{u}}, \tau\mathbf{u} + (1-\tau)\tilde{\mathbf{u}}, x)(u^k - \tilde{u}^k)\,dx \\
= &\sum_{k=1}^{m}\int_U \Bigg[-\sum_{i=1}^{n}\left(L_{p_i^k}(\tau D\mathbf{u} + (1-\tau)D\tilde{\mathbf{u}}, \tau\mathbf{u} + (1-\tau)\tilde{\mathbf{u}}, x)\right)_{x_i} \\
&+ L_{z^k}(\tau D\mathbf{u} + (1-\tau)D\tilde{\mathbf{u}}, \tau\mathbf{u} + (1-\tau)\tilde{\mathbf{u}}, x)\Bigg](u^k - \tilde{u}^k)\,dx \\
= &\ 0,
\end{aligned}$$

the last equality holding since the Euler–Lagrange system of PDE is satisfied by $\tau\mathbf{u} + (1-\tau)\tilde{\mathbf{u}}$. The identity (19) follows. □

In the scalar case that $m = 1$ the only null Lagrangians are the boring examples where L is linear in the variable p. For the case of systems ($m > 1$), however, there are certain nontrivial examples, which will turn out to be important for us later.

Notation. If A is an $n \times n$ matrix, we denote by

$$\operatorname{cof} A$$

the *cofactor* matrix, whose $(k,i)^{th}$ entry is $(\operatorname{cof} A)_i^k = (-1)^{i+k} d(A)_i^k$, where $d(A)_i^k$ = determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the k^{th}-row and i^{th}-column from A. □

LEMMA (Divergence-free rows). *Let* $\mathbf{u} : \mathbb{R}^n \to \mathbb{R}^n$ *be a smooth function. Then*

$$\sum_{i=1}^{n} (\operatorname{cof} D\mathbf{u})^k_{i,x_i} = 0 \quad (k = 1, \dots, n). \tag{20}$$

Proof. 1. From linear algebra we recall the identity

$$(\det P)I = P^T(\operatorname{cof} P) \quad (P \in \mathbb{M}^{n\times n}); \tag{21}$$

that is,

$$(\det P)\delta_{ij} = \sum_{k=1}^{n} p_i^k (\operatorname{cof} P)_j^k \quad (i,j = 1, \dots, n). \tag{22}$$

Thus in particular

$$\frac{\partial \det P}{\partial p_m^k} = (\operatorname{cof} P)_m^k \quad (k,m = 1, \dots, n). \tag{23}$$

2. Now set $P = Du$ in (22), differentiate with respect to x_j, and sum $j = 1, \dots, n$, to find

$$\sum_{j,k,m=1}^{n} \delta_{ij} (\operatorname{cof} D\mathbf{u})_m^k u^k_{x_m x_j} = \sum_{k,j=1}^{n} u^k_{x_i x_j} (\operatorname{cof} D\mathbf{u})_j^k + u^k_{x_i} (\operatorname{cof} D\mathbf{u})^k_{j,x_j}$$

for $i = 1, \dots, n$. This identity simplifies to read

$$\sum_{k=1}^{n} u^k_{x_i} \left(\sum_{j=1}^{n} (\operatorname{cof} D\mathbf{u})^k_{j,x_j} \right) = 0 \quad (i = 1, \dots, n). \tag{24}$$

3. Now if $\det D\mathbf{u}(x_0) \neq 0$, we deduce from (24) that

$$\sum_{j=1}^{n} (\operatorname{cof} D\mathbf{u})^k_{j,x_j} = 0 \quad (k = 1, \dots, n)$$

at x_0. If instead $\det D\mathbf{u}(x_0) = 0$, we choose a number $\epsilon > 0$ so small that $\det(D\mathbf{u}(x_0) + \epsilon I) \neq 0$, apply steps 1–3 to $\tilde{\mathbf{u}} := \mathbf{u} + \varepsilon x$, and send $\epsilon \to 0$. □

THEOREM 2 (Determinants as null Lagrangians). *The determinant function*

$$L(P) = \det P \quad (P \in \mathbb{M}^{n \times n})$$

is a null Lagrangian.

Proof. We must show that for any smooth function $\mathbf{u} : U \to \mathbb{R}^n$,

$$\sum_{i=1}^{n} \left(L_{p_i^k}(D\mathbf{u}) \right)_{x_i} = 0 \quad (k = 1, \dots, n).$$

According to (23) we have $L_{p_i^k} = (\operatorname{cof} P)^k_i$ $(i, k = 1, \dots, n)$. But then employing the notation and conclusion of the lemma, we see

$$\sum_{i=1}^{n} \left(L_{p_i^k}(D\mathbf{u}) \right)_{x_i} = \sum_{i=1}^{n} (\operatorname{cof} D\mathbf{u})^k_{i,x_i} = 0 \quad (k = 1, \dots, n).$$

□

Some other interesting null Lagrangians are introduced in the exercises.

c. Application.

A nice application is a quick analytic proof of a topological fixed point theorem.

THEOREM 3 (Brouwer's Fixed Point Theorem). *Assume*

$$\mathbf{u} : B(0,1) \to B(0,1)$$

is continuous, where $B(0,1)$ denotes the closed unit ball in $\mathbb{R}^n$. Then $\mathbf{u}$ has a fixed point; that is, there exists a point $x \in B(0,1)$ with

$$\mathbf{u}(x) = x.$$

Proof. 1. Write $B = B(0,1)$. We first of all claim that there does not exist a smooth function

$$\mathbf{w} : B \to \partial B \tag{25}$$

such that

$$\mathbf{w}(x) = x \quad \text{for all } x \in \partial B. \tag{26}$$

2. Suppose to the contrary that such a function $\mathbf{w}$ exists. Let us temporarily write $\tilde{\mathbf{w}}$ for the identity function, so that $\tilde{\mathbf{w}}(x) = x$ for all $x \in B$. In view of (26), $\mathbf{w} \equiv \tilde{\mathbf{w}}$ on ∂B. Since the determinant is a null Lagrangian, Theorem 1 implies

$$\int_B \det D\mathbf{w}\, dx = \int_B \det D\tilde{\mathbf{w}}\, dx = |B| \neq 0. \tag{27}$$

On the other hand, (25) implies $|\mathbf{w}|^2 \equiv 1$; and so differentiating, we find

$$(D\mathbf{w})^T \mathbf{w} = \mathbf{0}. \tag{28}$$

Since $|\mathbf{w}| = 1$, (28) says 0 is an eigenvalue of $D\mathbf{w}^T$ for each $x \in B$. Therefore $\det D\mathbf{w} \equiv 0$ in B. This contradicts (27), and thereby proves no smooth function $\mathbf{w}$ satisfying (25), (26) can exist.

2. Next we show there does not exist any continuous function $\mathbf{w}$ verifying (25), (26). Indeed if $\mathbf{w}$ were such a function, we continuously extend $\mathbf{w}$ by setting $\mathbf{w}(x) = x$ if $x \in \mathbb{R}^n - B$. Observe that $\mathbf{w}(x) \neq 0$ ($x \in \mathbb{R}^n$). Fix $\epsilon > 0$ so small that $\mathbf{w}_1 := \eta_\epsilon * \mathbf{w}$ satisfies $\mathbf{w}_1(x) \neq 0$ ($x \in \mathbb{R}^n$). Note also that since η_ϵ is radial, we have $\mathbf{w}_1(x) = x$ if $x \in \mathbb{R}^n - B(0,2)$, for $\epsilon > 0$ sufficiently small. Then

$$\mathbf{w}_2 := \frac{2\mathbf{w}_1}{|\mathbf{w}_1|}$$

would be a smooth mapping satisfying (25), (26) (with the ball $B(0,2)$ replacing $B = B(0,1)$), in contradiction to step 1.

3. Finally suppose $\mathbf{u} : B \to B$ is continuous, but has no fixed point. Define now the mapping $\mathbf{w} : B \to \partial B$ by setting $\mathbf{w}(x)$ to be the point on ∂B hit by the ray emanating from $\mathbf{u}(x)$ and passing through x. This mapping is well defined since $\mathbf{u}(x) \neq x$ for all $x \in B$. In addition $\mathbf{w}$ is continuous and satisfies (25), (26).

But this in turn is a contradiction to step 2. □

We will employ Brouwer's fixed point theorem several times in Chapter 9.

8.2. EXISTENCE OF MINIMIZERS

In this section we will identify some conditions on the Lagrangian L which ensure that the functional $I[\cdot]$ does indeed have a minimizer, at least within an appropriate Sobolev space.

8.2.1. Coercivity, lower semicontinuity.

Let us start with some largely heuristic insights as to when the functional

$$I[w] := \int_U L(Dw(x), w(x), x)\, dx, \tag{1}$$

defined for appropriate functions $w : U \to \mathbb{R}$ satisfying

$$w = g \quad \text{on } \partial U, \tag{2}$$

should have a minimizer.

a. Coercivity.

We first of all note that even a smooth function f mapping $\mathbb{R}$ to $\mathbb{R}$ and bounded below need not attain its infimum. Consider, for instance, $f = e^x$ or $(1 + x^2)^{-1}$. These examples suggest that we in general will need some hypothesis controlling $I[w]$ for "large" functions w. Certainly the most effective way to ensure this would be to hypothesize that $I[w]$ "grows rapidly as $|w| \to \infty$".

More specifically, let us assume

$$1 < q < \infty \tag{3}$$

is fixed. We will then suppose

$$\begin{cases} \text{there exist constants } \alpha > 0,\ \beta \geq 0 \text{ such that} \\ \qquad L(p, z, x) \geq \alpha |p|^q - \beta \\ \text{for all } p \in \mathbb{R}^n,\ z \in \mathbb{R},\ x \in U. \end{cases} \tag{4}$$

Therefore

$$I[w] \geq \alpha \|Dw\|^q_{L^q(U)} - \gamma \tag{5}$$

for $\gamma := \beta |U|$. Thus $I[w] \to \infty$ as $\|Dw\|_{L^q} \to \infty$. It is customary to call (5) a *coercivity condition* on $I[\cdot]$.

Turning once more to our basic task of finding minimizers for the functional $I[\cdot]$, we observe from inequality (5) that it seems reasonable to define

$I[w]$ not only for smooth functions w, but also for functions w in the Sobolev space $W^{1,q}(U)$ that satisfy the boundary condition (2) in the trace sense. After all, the wider the class of functions w for which $I[w]$ is defined, the more candidates we will have for a minimizer.

We will henceforth write

$$\mathcal{A} := \{w \in W^{1,q}(U) \mid w = g \text{ on } \partial U \text{ in the trace sense}\}$$

to denote this class of *admissible* functions w. Note in view of (4) that $I[w]$ is defined (but may equal $+\infty$) for each $w \in \mathcal{A}$.

b. Lower semicontinuity.

Next, let us observe that although a continuous function $f : \mathbb{R} \to \mathbb{R}$ satisfying a coercivity condition does indeed attain its infimum, our integral functional $I[\cdot]$ in general will not. To understand the problem, set

$$m := \inf_{w \in \mathcal{A}} I[w] \tag{6}$$

and choose functions $u_k \in \mathcal{A}$ $(k = 1, \dots)$ so that

$$I[u_k] \to m \quad \text{as } k \to \infty. \tag{7}$$

We call $\{u_k\}_{k=1}^{\infty}$ a *minimizing sequence.*

We would now like to show that some subsequence of $\{u_k\}_{k=1}^{\infty}$ converges to an actual minimizer. For this, however, we need some kind of compactness, and this is definitely a problem since the space $W^{1,q}(U)$ is infinite dimensional. Indeed, if we utilize the coercivity inequality (5), it turns out (cf. §8.2.2) that we can only conclude that the minimizing sequence lies in a bounded subset of $W^{1,q}(U)$. But this does *not* imply that there exists any subsequence which converges in $W^{1,q}(U)$.

We therefore turn our attention to the *weak topology* (cf. §D.4). Since we are assuming $1 < q < \infty$, so that $L^q(U)$ is reflexive, we conclude that there exists a subsequence $\{u_{k_j}\}_{j=1}^{\infty} \subset \{u_k\}_{k=1}^{\infty}$ and a function $u \in W^{1,q}(U)$ so that

$$\begin{cases} u_{k_j} \rightharpoonup u \text{ weakly in } L^q(U) \\ Du_{k_j} \rightharpoonup Du \text{ weakly in } L^q(U;\mathbb{R}^n). \end{cases} \tag{8}$$

We will hereafter abbreviate (8) by saying

$$u_{k_j} \rightharpoonup u \text{ weakly in } W^{1,q}(U). \tag{9}$$

Furthermore, it will be true that $u = g$ on ∂U in the trace sense, and so $u \in \mathcal{A}$.

Consequently by shifting to the weak topology we have recovered enough compactness from the coercivity inequality (5) to deduce (9) for an appropriate subsequence. But now another difficulty arises, for in essentially all cases of interest the *functional* $I[\cdot]$ *is not continuous with respect to weak convergence.* In other words, we *cannot* deduce from (7) and (9) that

$$I[u] = \lim_{j\to\infty} I[u_{k_j}], \tag{10}$$

and thus u is a minimizer. The problem is that $Du_{k_j} \rightharpoonup Du$ does *not* imply $Du_{k_j} \to Du$ a.e.: it is quite possible for instance that the gradients Du_{k_j}, although bounded in L^q, are oscillating more and more wildly as $k_j \to \infty$.

What saves us is the final, key observation that we do not really need the full strength of (10). It would suffice instead to know only

$$I[u] \leq \liminf_{j\to\infty} I[u_{k_j}]. \tag{11}$$

Then from (7) we could deduce $I[u] \leq m$. But owing to (6), $m \leq I[u]$. Consequently u is indeed a minimizer.

DEFINITION. *We say that a function* $I[\cdot]$ *is* (sequentially) weakly lower semicontinuous *on* $W^{1,q}(U)$, *provided*

$$I[u] \leq \liminf_{k\to\infty} I[u_k]$$

whenever

$$u_k \rightharpoonup u \quad \textit{weakly in } W^{1,q}(U).$$

Our goal therefore is now to identify reasonable conditions on the nonlinearity L that ensure $I[\cdot]$ is weakly lower semicontinuous.

8.2.2. Convexity.

We next look back to our second variation analysis in §8.1.3 and recall we derived there the inequality

$$\sum_{i,j=1}^{n} L_{p_i p_j}(Du(x), u(x), x)\xi_i\xi_j \geq 0 \quad (\xi \in \mathbb{R}^n, x \in U)$$

holding as a necessary condition, whenever u is a smooth minimizer. This inequality strongly suggests that it is reasonable to assume that L is convex in its first argument.

THEOREM 1 (Weak lower semicontinuity). *Assume that L is bounded below, and in addition*

$$\text{the mapping } p \mapsto L(p,z,x) \text{ is convex,}$$

for each $z \in \mathbb{R}$, $x \in U$. Then

$$I[\cdot] \text{ is weakly lower semicontinuous on } W^{1,q}(U).$$

Proof. 1. Choose any sequence $\{u_k\}_{k=1}^{\infty}$ with

$$u_k \rightharpoonup u \quad \text{weakly in } W^{1,q}(U), \tag{12}$$

and set $l := \liminf_{k\to\infty} I[u_k]$. We must show

$$I[u] \leq l. \tag{13}$$

2. Note first from (12) and §D.4 that

$$\sup_k \|u_k\|_{W^{1,q}(U)} < \infty. \tag{14}$$

Upon passing to a subsequence if necessary, we may as well also suppose

$$l = \lim_{k\to\infty} I[u_k]. \tag{15}$$

Furthermore we see from the compactness theorem in §5.7 that $u_k \to u$ strongly in $L^q(U)$; and thus, passing if necessary to yet another subsequence, we have

$$u_k \to u \quad \text{a.e. in } U. \tag{16}$$

3. Fix $\epsilon > 0$. Then (16) and Egoroff's Theorem (§E.2) assert

$$u_k \to u \quad \text{uniformly on } E_\epsilon, \tag{17}$$

where E_ϵ is a measurable set with

$$|U - E_\epsilon| \leq \epsilon. \tag{18}$$

Now write

$$F_\epsilon := \left\{ x \in U \,\middle|\, |u(x)| + |Du(x)| \leq \frac{1}{\epsilon} \right\}. \tag{19}$$

Then

$$|U - F_\epsilon| \to 0 \quad \text{as } \epsilon \to 0. \tag{20}$$

We finally set

$$G_\epsilon := E_\epsilon \cap F_\epsilon, \tag{21}$$

and observe from (18), (20): $|U - G_\epsilon| \to 0$ as $\epsilon \to 0$.

4. Now let us observe since L is bounded below, we may as well assume

$$L \geq 0 \tag{22}$$

(for otherwise we could apply the following arguments to $\tilde{L} = L + \beta \geq 0$ for some appropriate constant β). Consequently

$$\begin{aligned} I[u_k] &= \int_U L(Du_k, u_k, x)\,dx \geq \int_{G_\epsilon} L(Du_k, u_k, x)\,dx \\ &\geq \int_{G_\epsilon} L(Du, u_k, x)\,dx + \int_{G_\epsilon} D_pL(Du, u_k, x) \cdot (Du_k - Du)\,dx, \end{aligned} \tag{23}$$

the last inequality following from the convexity of L in its first argument; see §B.1. Now in view of (17), (19) and (21):

$$\lim_{k\to\infty} \int_{G_\epsilon} L(Du, u_k, x)\,dx = \int_{G_\epsilon} L(Du, u, x)\,dx. \tag{24}$$

In addition, since $D_pL(Du, u_k, x) \to D_pL(Du, u, x)$ uniformly on G_ϵ and $Du_k \rightharpoonup Du$ weakly in $L^q(U; \mathbb{R}^n)$, we have

$$\lim_{k\to\infty} \int_{G_\epsilon} D_pL(Du, u_k, x) \cdot (Du_k - Du)\,dx = 0. \tag{25}$$

Owing now to (24), (25), we deduce from (23) that

$$l = \lim_{k\to\infty} I[u_k] \geq \int_{G_\epsilon} L(Du, u, x)\,dx.$$

This inequality holds for each $\epsilon > 0$. We now let ϵ tend to zero, and recall (22) and the Monotone Convergence Theorem (§E.3) to conclude

$$l \geq \int_U L(Du, u, x)\,dx = I[u],$$

as required. □

Remark. It is very important to understand how the foregoing proof deals with the weak convergence $Du_k \rightharpoonup Du$. The key is the convexity inequality (23), on the right hand side of which Du_k appears linearly. Weak convergence is, by its very definition, compatible with linear expressions, and so the limit (25) holds. Remember that it is not in general true that $Du_k \to Du$ a.e., even if we pass to a subsequence.

The convergence of u_k to u in L^q is much stronger, and so we do not need any convexity assumption concerning $z \mapsto L(p, z, x)$. □

We can at last establish that $I[\cdot]$ has a minimizer among the functions in $\mathcal{A}$.

THEOREM 2 (Existence of minimizer). *Assume that L satisfies the coercivity inequality* (4) *and is convex in the variable p. Suppose also the admissible set $\mathcal{A}$ is nonempty.*

Then there exists at least one function $u \in \mathcal{A}$ solving

$$I[u] = \min_{w \in \mathcal{A}} I[w] .$$

Proof. 1. Set $m := \inf_{w \in \mathcal{A}} I[w]$. If $m = +\infty$ we are done, and so we henceforth assume m is finite. Select a minimizing sequence $\{u_k\}_{k=1}^{\infty}$. Then

$$I[u_k] \to m. \tag{26}$$

2. We may as well take $\beta = 0$ in inequality (4), since we could otherwise just as well consider $\tilde{L} := L + \beta$. Thus $L \geq \alpha |p|^q$, and so

$$I[w] \geq \alpha \int_U |Dw|^q \, dx. \tag{27}$$

Since m is finite, we conclude from (26) and (27) that

$$\sup_k \|Du_k\|_{L^q(U)} < \infty. \tag{28}$$

3. Now fix any function $w \in \mathcal{A}$. Since u_k and w both equal g on ∂U in the trace sense, we have $u_k - w \in W_0^{1,q}(U)$. Therefore Poincaré's inequality implies

$$\begin{aligned} \|u_k\|_{L^q(U)} &\leq \|u_k - w\|_{L^q(U)} + \|w\|_{L^q(U)} \\ &\leq C\|Du_k - Dw\|_{L^q(U)} + C \leq C, \end{aligned}$$

by (28). Hence $\sup_k \|u_k\|_{L^q(U)} < \infty$. This estimate and (28) imply $\{u_k\}_{k=1}^\infty$ is bounded in $W^{1,q}(U)$.

4. Consequently there exist a subsequence $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$ and a function $u \in W^{1,q}(U)$ such that

$$u_{k_j} \rightharpoonup u \quad \text{weakly in } W^{1,q}(U).$$

We assert next that $u \in \mathcal{A}$. To see this, note that for $w \in \mathcal{A}$ as above, $u_k - w \in W_0^{1,q}(U)$. Now $W_0^{1,q}(U)$ is a closed, linear subspace of $W^{1,q}(U)$, and so, by Mazur's Theorem (§D.4), is weakly closed. Hence $u - w \in W_0^{1,q}(U)$. Consequently the trace of u on ∂U is g.

In view of Theorem 1 then, $I[u] \leq \liminf_{j\to\infty} I[u_{k_j}] = m$. But since $u \in \mathcal{A}$, it follows that

$$I[u] = m = \min_{w\in\mathcal{A}} I[w].$$

□

We turn next to the problem of uniqueness. In general there can be many minimizers, and so to ensure uniqueness we require further assumptions. Suppose for instance

$$L = L(p, x) \text{ does not depend on } z, \tag{29}$$

and

$$\begin{cases} \text{there exists } \theta > 0 \text{ such that} \\ \sum_{i,j=1}^n L_{p_i p_j}(p, x)\xi_i\xi_j \geq \theta|\xi|^2 \quad (p, \xi \in \mathbb{R}^n;\ x \in U). \end{cases} \tag{30}$$

Condition (30) says the mapping $p \mapsto L(p, x)$ is uniformly convex for each x.

THEOREM 3 (Uniqueness of minimizer). *Suppose* (29), (30) *hold. Then a minimizer $u \in \mathcal{A}$ of $I[\cdot]$ is unique.*

Proof. 1. Assume $u, \tilde{u} \in \mathcal{A}$ are both minimizers of $I[\cdot]$ over $\mathcal{A}$. Then $v := \frac{u+\tilde{u}}{2} \in \mathcal{A}$. We claim

$$I[v] \leq \frac{I[u] + I[\tilde{u}]}{2}, \tag{31}$$

with a strict inequality, unless $u = \tilde{u}$ a.e.

2. To see this, note from the uniform convexity assumption that we have

$$L(p, x) \geq L(q, x) + D_pL(q, x)\cdot(p - q) + \frac{\theta}{2}|p - q|^2 \quad (x \in U,\ p, q \in \mathbb{R}^n). \tag{32}$$

Set $q = \frac{Du+D\tilde{u}}{2}$, $p = Du$, and integrate over U:

$$I[v] + \int_U D_pL\left(\frac{Du + D\tilde{u}}{2}, x\right) \cdot \left(\frac{Du - D\tilde{u}}{2}\right) dx + \frac{\theta}{8}\int_U |Du - D\tilde{u}|^2\,dx \le I[u]. \tag{33}$$

Similarly, set $q = \frac{Du+D\tilde{u}}{2}$, $p = D\tilde{u}$ in (32) and integrate:

$$I[v] + \int_U D_pL\left(\frac{Du + D\tilde{u}}{2}, x\right) \cdot \left(\frac{D\tilde{u} - Du}{2}\right) dx + \frac{\theta}{8}\int_U |Du - D\tilde{u}|^2\,dx \le I[\tilde{u}]. \tag{34}$$

Add and divide by 2, to deduce

$$I[v] + \frac{\theta}{8}\int_U |Du - D\tilde{u}|^2\,dx \le \frac{I[u] + I[\tilde{u}]}{2}.$$

This proves (31).

3. As $I[u] = I[\tilde{u}] = \min_{w\in\mathcal{A}} I[w] \le I[v]$, we deduce $Du = D\tilde{u}$ a.e. in U. Since $u = \tilde{u} = g$ on ∂U in the trace sense, it follows that $u = \tilde{u}$ a.e. □

8.2.3. Weak solutions of Euler–Lagrange equation.

We wish next to demonstrate that any minimizer $u \in \mathcal{A}$ of $I[\cdot]$ solves the Euler–Lagrange equation in some suitable sense. This does *not* follow from the calculations in §8.1 since we do not know u is smooth, only $u \in W^{1,q}(U)$. And in fact we will need some growth conditions on L and its derivatives. Let us hereafter suppose

$$|L(p, z, x)| \le C(|p|^q + |z|^q + 1), \tag{35}$$

and also

$$\begin{cases} |D_pL(p, z, x)| \le C(|p|^{q-1} + |z|^{q-1} + 1) \\ |D_zL(p, z, x)| \le C(|p|^{q-1} + |z|^{q-1} + 1) \end{cases} \tag{36}$$

for some constant C and all $p \in \mathbb{R}^n$, $z \in \mathbb{R}$, $x \in U$.

Motivation for definition of weak solution. We now turn our attention to the boundary-value problem for the Euler–Lagrange PDE associated with our functional L, which for a smooth minimizer u reads

$$\begin{cases} -\sum_{i=1}^n (L_{p_i}(Du, u, x))_{x_i} + L_z(Du, u, x) = 0 & \text{in } U \\ u = g & \text{on } \partial U. \end{cases} \tag{37}$$

If we multiply (37) by a test function $v \in C_c^\infty(U)$ and integrate by parts, we arrive at the equality

$$\int_U \sum_{i=1}^n L_{p_i}(Du,u,x)v_{x_i} + L_z(Du,u,x)v\,dx = 0. \tag{38}$$

Of course this is the identity we first obtained in our derivation of (37) in §8.1.2.

Now assume $u \in W^{1,q}(U)$. Then using (36) we see

$$|D_pL(Du,u,x)| \le C(|Du|^{q-1} + |u|^{q-1} + 1) \in L^{q'}(U),$$

where $q' = \frac{q}{q-1}$, $\frac{1}{q} + \frac{1}{q'} = 1$. Similarly

$$|D_zL(Du,u,x)| \le C(|Du|^{q-1} + |u|^{q-1} + 1) \in L^{q'}(U). \tag{39}$$

Consequently we see using a standard approximation argument that the equality (38) is valid for any $v \in W_0^{1,q}(U)$. This motivates the following

DEFINITION. *We say $u \in \mathcal{A}$ is a* weak solution *of the boundary-value problem* (37) *for the Euler–Lagrange equation provided*

$$\int_U \sum_{i=1}^n L_{p_i}(Du,u,x)v_{x_i} + L_z(Du,u,x)v\,dx = 0$$

for all $v \in W_0^{1,q}(U)$.

THEOREM 4 (Solution of Euler–Lagrange equation). *Assume L verifies the growth conditions* (35), (36), *and $u \in \mathcal{A}$ satisfies*

$$I[u] = \min_{w\in\mathcal{A}} I[w].$$

Then u is a weak solution of (37).

Proof. We proceed as in §8.1.2, taking care about differentiating under the integrals. Fix any $v \in W_0^{1,q}(U)$ and set

$$i(\tau) := I[u + \tau v] \quad (\tau \in \mathbb{R}).$$

In view of (35) we see that $i(\tau)$ is finite for all τ.

Let $\tau \neq 0$ and write the difference quotient

$$\begin{aligned} \frac{i(\tau)-i(0)}{\tau} &= \int_U \frac{L(Du+\tau Dv, u+\tau v, x) - L(Du,u,x)}{\tau}\,dx \\ &= \int_U L^\tau(x)\,dx, \end{aligned} \tag{40}$$

where

$$L^\tau(x) := \frac{1}{\tau}[L(Du(x) + \tau Dv(x), u(x) + \tau v(x), x) - L(Du(x), u(x), x)]$$

for a.e. $x \in U$. Clearly

$$L^\tau(x) \to \sum_{i=1}^n L_{p_i}(Du, u, x)v_{x_i} + L_z(Du, u, x)v \quad \text{a.e.} \tag{41}$$

as $\tau \to 0$. Furthermore

$$\begin{aligned} L^\tau(x) &= \frac{1}{\tau}\int_0^\tau \frac{d}{ds}L(Du + sDv, u + sv, x)\,ds \\ &= \frac{1}{\tau}\int_0^\tau \sum_{i=1}^n L_{p_i}(Du + sDv, u + sv, x)v_{x_i} \\ &\qquad\qquad + L_z(Du + sDv, u + sv, x)v\,ds. \end{aligned}$$

Next recall from §B.2 Young's inequality: $ab \le \frac{a^q}{q} + \frac{b^{q'}}{q'}$, where $\frac{1}{q} + \frac{1}{q'} = 1$. Then since $u, v \in W^{1,q}(U)$, inequalities (36) and Young's inequality imply after some elementary calculations that

$$|L^\tau(x)| \le C(|Du|^q + |u|^q + |Dv|^q + |v|^q + 1) \in L^1(U)$$

for each $\tau \neq 0$. Consequently we may invoke the Dominated Convergence Theorem to conclude from (40), (41) that $i'(0)$ exists and equals

$$\int_U \sum_{i=1}^n L_{p_i}(Du, u, x)v_{x_i} + L_z(Du, u, x)v\,dx.$$

But then since $i(\cdot)$ has a minimum for $\tau = 0$, we know $i'(0) = 0$; and thus u is a weak solution. □

Remark. In general, the Euler–Lagrange equation (37) will have other solutions which do not correspond to minima of $I[\cdot]$; see §8.5. However, *in the special case that the joint mapping* $(p, z) \mapsto L(p, z, x)$ *is convex for each* x, *then each weak solution is in fact a minimizer.*

To see this, suppose $u \in \mathcal{A}$ solves

$$\begin{cases} -\sum_{i=1}^n (L_{p_i}(Du, u, x))_{x_i} + L_z(Du, u, x) = 0 & \text{in } U \\ u = g & \text{on } \partial U \end{cases} \tag{42}$$

in the weak sense and select any $w \in \mathcal{A}$. Utilizing the convexity of the mapping $(p, z) \mapsto L(p, z, x)$, we have

$$L(p,z,x) + D_pL(p,z,x)\cdot(q-p) + D_zL(p,z,x)\cdot(w-z) \le L(q,w,x).$$

Let $p = Du(x)$, $q = Dw(x)$, $z = u(x)$, $w = w(x)$ and integrate over U:

$$I[u] + \int_U D_pL(Du,u,x)\cdot(Dw - Du) + D_zL(Du,u,x)(w-u)\,dx \le I[w].$$

In view of (42) the second term on the right is zero, and therefore $I[u] \le I[w]$ for each $w \in \mathcal{A}$. □

8.2.4. Systems.

a. Convexity.

We now adopt again the notation for systems set forth in §8.1.4, and consider the existence question for minimizers of the functional

$$I[\mathbf{w}] := \int_U L(D\mathbf{w}(x), \mathbf{w}(x), x)\,dx,$$

defined for appropriate functions $\mathbf{w} : U \to \mathbb{R}^m$, where now $L : \mathbb{M}^{m\times n} \times \mathbb{R}^m \times \bar{U} \to \mathbb{R}$ is given.

It turns out the theory developed in §8.2.2 extends with no difficulty to the case at hand. Let us therefore assume the coercivity inequality

$$L(P,z,x) \ge \alpha|P|^q - \beta \quad (P \in \mathbb{M}^{m\times n}, z \in \mathbb{R}^m, x \in U) \tag{43}$$

for constants $\alpha > 0, \beta \ge 0$, and set also

$$\mathcal{A} = \{\mathbf{w} \in W^{1,q}(U;\mathbb{R}^m) \mid \mathbf{w} = \mathbf{g} \text{ on } \partial U \text{ in the trace sense}\},$$

where $\mathbf{g} : \partial U \to \mathbb{R}^m$ is given.

THEOREM 5 (Existence of minimizer). *Assume that L satisfies the coercivity inequality* (43) *and is convex in the variable P. Suppose also the admissible set $\mathcal{A}$ is nonempty.*

Then there exists $\mathbf{u} \in \mathcal{A}$ solving

$$I[\mathbf{u}] = \min_{\mathbf{w}\in\mathcal{A}} I[\mathbf{w}].$$

The proof follows almost exactly the proofs of Theorems 1 and 2 in §8.2.2. Similarly to Theorem 3 above we have:

THEOREM 6 (Uniqueness of minimizer). *Assume L does not depend on z and the mapping $P \mapsto L(P,x)$ is uniformly convex. Then a minimizer $\mathbf{u} \in \mathcal{A}$ of $I[\cdot]$ is unique.*

Now suppose additionally

$$
(44) \qquad \begin{cases} |L(P,z,x)| \leq C(|P|^q + |z|^q + 1) \\ |D_P L(P,z,x)| \leq C(|P|^{q-1} + |z|^{q-1} + 1) \\ |D_z L(P,z,x)| \leq C(|P|^{q-1} + |z|^{q-1} + 1) \end{cases}
$$

for some constant C and all $P \in \mathbb{M}^{m \times n}$, $z \in \mathbb{R}^m$, $x \in U$.

We consider now the system of Euler–Lagrange equations

$$
(45) \qquad \begin{cases} -\sum_{i=1}^n (L_{p_i^k}(D\mathbf{u},\mathbf{u},x))_{x_i} + L_{z^k}(D\mathbf{u},\mathbf{u},x) = 0 & \text{in } U \\ u^k = g^k & \text{on } \partial U \end{cases}
$$

for $k = 1, \dots, m$, and define $u \in \mathcal{A}$ to be a *weak solution* provided

$$
\sum_{k=1}^m \int_U \sum_{i=1}^n L_{p_i^k}(D\mathbf{u},\mathbf{u},x) w_{x_i}^k + L_{z^k}(D\mathbf{u},\mathbf{u},x) w^k \, dx = 0
$$

for all $\mathbf{w} \in W_0^{1,q}(U;\mathbb{R}^m)$, $\mathbf{w} = (w^1, \dots, w^m)$.

THEOREM 7 (Solution of Euler–Lagrange system). *Assume L verifies the growth conditions* (44) *and $\mathbf{u} \in \mathcal{A}$ satisfies*

$$
I[\mathbf{u}] = \min_{\mathbf{w} \in \mathcal{A}} I[\mathbf{w}].
$$

Then $\mathbf{u}$ is a weak solution of (45).

The proof is almost precisely like that of Theorem 4.

b. Polyconvexity.

It is rather surprising that there are some mathematically and physically interesting systems which are not covered by Theorem 5 above, but which can still be studied using the calculus of variations. These include certain problems where the Lagrangian L is *not* convex in P, but $I[\cdot]$ is nonetheless weakly lower semicontinuous.

LEMMA (Weak continuity of determinants). *Assume $n < q < \infty$ and*

$$
\mathbf{u}_k \rightharpoonup \mathbf{u} \quad \text{weakly in } W^{1,q}(U;\mathbb{R}^n).
$$

Then

$$
\det D\mathbf{u}_k \rightharpoonup \det D\mathbf{u} \quad \text{weakly in } L^{q/n}(U).
$$

Proof. 1. First we recall the matrix identity $(\det P)I = P(\operatorname{cof} P)^T$; consequently

$$\det P = \sum_{j=1}^{n} p_j^i (\operatorname{cof} P)_j^i \quad (i = 1, \dots, n).$$

2. Now let $\mathbf{w} \in C^\infty(U; \mathbb{R}^n)$, $\mathbf{w} = (w^1, \dots, w^n)$. Then

$$\det D\mathbf{w} = \sum_{j=1}^{n} w_{x_j}^i (\operatorname{cof} D\mathbf{w})_j^i \quad (i = 1, \dots, n). \tag{46}$$

But the lemma in §8.1.4 asserts $\sum_{j=1}^n (\operatorname{cof} D\mathbf{w})_{j,x_j}^i = 0$. Thus formula (46) says

$$\det D\mathbf{w} = \sum_{j=1}^{n} (w^i (\operatorname{cof} D\mathbf{w})_j^i)_{x_j}.$$

Consequently *the determinant of the gradient matrix can be written as a divergence.* Therefore if $v \in C_c^\infty(U)$, we have

$$\int_U v \det D\mathbf{w}\, dx = -\sum_{j=1}^{n} \int_U v_{x_j} w^i (\operatorname{cof} D\mathbf{w})_j^i \, dx \quad (i = 1, \dots, n). \tag{47}$$

3. We have established the identity (47) for a smooth function w; and so a standard approximation argument yields

$$\int_U v \det D\mathbf{u}_k \, dx = -\sum_{j=1}^{n} \int_U v_{x_j} u_k^i (\operatorname{cof} D\mathbf{u}_k)_j^i \, dx \tag{48}$$

for $k = 1, 2, \dots$. Now since $n < q < \infty$ and $\mathbf{u}_k \rightharpoonup \mathbf{u}$ in $W^{1,q}(U; \mathbb{R}^n)$, we know from Morrey's inequality that $\{\mathbf{u}_k\}_{k=1}^\infty$ is bounded in $C^{0,1-n/q}(U; \mathbb{R}^n)$. Thus using the Arzela–Ascoli compactness criterion, §C.7, we deduce $\mathbf{u}_k \to \mathbf{u}$ uniformly in U. Returning then to identity (48) we see that we could conclude

$$\lim_{k\to\infty} \int_U v \det D\mathbf{u}_k \, dx = -\sum_{j=1}^{n} \int_U v_{x_j} u^i (\operatorname{cof} D\mathbf{u})_j^i \, dx = \int_U v \det D\mathbf{u} \, dx, \tag{49}$$

if we knew

$$\lim_{k\to\infty} \int_U \psi (\operatorname{cof} D\mathbf{u}_k)_j^i \, dx = \int_U \psi (\operatorname{cof} D\mathbf{u})_j^i \, dx \tag{50}$$

for $i, j = 1, \dots, n$ and each $\psi \in C_c^\infty(U)$. However $(\operatorname{cof} D\mathbf{u}_k)_j^i$ is the determinant of an $(n-1) \times (n-1)$ matrix, which can be analyzed as above by being written as a sum of determinants of appropriate $(n-2)\times(n-2)$ submatrices, times uniformly convergent factors. We continue and eventually must show only the obvious fact that the entries of the matrices $D\mathbf{u}_k$ converge weakly to the corresponding entries of $D\mathbf{u}$. In this way we verify (50), and thus (49).

4. Finally, since $\{\mathbf{u}_k\}_{k=1}^\infty$ is bounded in $W^{1,q}(U;\mathbb{R}^n)$ and $|\det D\mathbf{u}_k| \leq C|D\mathbf{u}_k|^n$, we see that $\{\det D\mathbf{u}_k\}_{k=1}^\infty$ is bounded in $L^{q/n}(U)$. Hence any subsequence has a weakly convergent subsequence in $L^{q/n}(U)$, which—owing to (49)—can only converge to $\det D\mathbf{u}$. □

We next utilize this lemma to establish a weak lower semicontinuity assertion analogous to Theorem 1, except that we will not assume that the Lagrangian L is necessarily convex in P. Instead let us suppose that $m = n$ and L has the form

$$L(P,z,x) = F(P,\det P,z,x) \quad (P \in \mathbb{M}^{n\times n}, z \in \mathbb{R}^n, x \in U) \tag{51}$$

where $F : \mathbb{M}^{n\times n} \times \mathbb{R} \times \mathbb{R}^n \times \bar{U} \to \mathbb{R}$ is smooth. We additionally hypothesize that

$$\begin{cases} \text{for each fixed } z \in \mathbb{R}^m,\ x \in \mathbb{R}, \text{ the joint mapping} \\ (P,r) \mapsto F(P,r,z,x) \text{ is convex.} \end{cases} \tag{52}$$

A Lagrangian L of the form (51) is called *polyconvex* provided (52) holds.

THEOREM 8 (Lower semicontinuity of polyconvex functionals). *Suppose $n < q < \infty$. Assume also L is bounded below and is polyconvex. Then*

$$I[\cdot] \text{ is weakly lower semicontinuous on } W^{1,q}(U;\mathbb{R}^n).$$

Proof. Choose any sequence $\{\mathbf{u}_k\}_{k=1}^\infty$ with

$$\mathbf{u}_k \rightharpoonup \mathbf{u} \quad \text{weakly in } W^{1,q}(U;\mathbb{R}^n). \tag{53}$$

According to the lemma,

$$\det D\mathbf{u}_k \rightharpoonup \det D\mathbf{u} \quad \text{weakly in } L^{q/n}(U). \tag{54}$$

We can now argue almost exactly as in the proof of Theorem 1. Indeed, using the notation from that proof, we have

$$
\begin{aligned}
I[\mathbf{u}_k] &= \int_U L(D\mathbf{u}_k, \mathbf{u}_k, x)\,dx \geq \int_{G_\epsilon} L(D\mathbf{u}_k, \mathbf{u}_k, x)\,dx \\
&= \int_{G_\epsilon} F(D\mathbf{u}_k, \det D\mathbf{u}_k, \mathbf{u}_k, x)\,dx \\
&\geq \int_{G_\epsilon} F(D\mathbf{u}, \det D\mathbf{u}, \mathbf{u}_k, x)\,dx \\
&\quad + \int_{G_\epsilon} F_p(D\mathbf{u}, \det D\mathbf{u}, \mathbf{u}_k, x) \cdot (D\mathbf{u}_k - D\mathbf{u}) \\
&\qquad + F_r(D\mathbf{u}, \det D\mathbf{u}, \mathbf{u}_k, x)(\det D\mathbf{u}_k - \det D\mathbf{u})\,dx,
\end{aligned}
$$

in view of (52). Reasoning as in the proof of Theorem 1 we deduce from (53), (54) that the limit of the last term is zero as $k \to \infty$. □

As before, we immediately deduce

THEOREM 9 (Existence of minimizers, polyconvex functionals). *Assume that $n < q < \infty$, and that L satisfies the coercivity inequality* (43) *and is polyconvex. Suppose also the admissible set $\mathcal{A}$ is nonempty.*

Then there exists $\mathbf{u} \in \mathcal{A}$ solving

$$I[\mathbf{u}] = \min_{w \in \mathcal{A}} I[\mathbf{w}].$$

Example. We explain the interest in the hypothesis of polyconvexity with an application from nonlinear elasticity theory, where $n = 3$. We consider an elastic body, which initially has the reference configuration U. We then displace each point $x \in \partial U$ to a new position $\mathbf{g}(x)$ and wish to determine new displacement $\mathbf{u}(x)$ of each internal point $x \in U$.

If the material is *hyperelastic*, there exists by definition an associated energy density L such that the physical displacement $\mathbf{u}$ minimizes the internal energy functional

$$I[\mathbf{w}] := \int_U L(D\mathbf{w}, x)\,dx$$

over all admissible displacements $\mathbf{w} \in \mathcal{A}$. Now it seems reasonable physically that L, which represents the internal energy density from stretching and compression, may explicitly depend on the local change in volume, that is, on $\det D\mathbf{w}$. In other words, it is physically appropriate to suppose that L has the form $L(P, x) = F(P, \det P, x)$. Then F describes in its first argument changes in internal energy due to changes in line elements, and in its second argument changes in internal energy due to changes in volume elements. See Ball [**B**] for more explanation. □

8.3. REGULARITY

We discuss in this section the smoothness of minimizers to our energy functionals. This is generally a quite difficult topic, and so we will make a number of strong simplifying assumptions, most notably that L depends only on p. Thus we henceforth assume our functional $I[\cdot]$ to have the form

$$I[w] := \int_U L(Dw) - wf\,dx, \tag{1}$$

for $f \in L^2(U)$. We will also take $q = 2$, and suppose as well the growth condition

$$|D_pL(p)| \le C(|p|+1) \quad (p \in \mathbb{R}^n). \tag{2}$$

Then any minimizer $u \in \mathcal{A}$ is a weak solution of the Euler–Lagrange PDE

$$-\sum_{i=1}^n (L_{p_i}(Du))_{x_i} = f \quad \text{in } U; \tag{3}$$

that is,

$$\int_U \sum_{i=1}^n L_{p_i}(Du)v_{x_i}\,dx = \int_U fv\,dx \tag{4}$$

for all $v \in H_0^1(U)$.

8.3.1. Second derivative estimates.

We now intend to show if $u \in H^1(U)$ is a weak solution of the nonlinear PDE (3), then in fact $u \in H^2_{\text{loc}}(U)$. But to establish this we will need to strengthen our growth conditions on L. Let us first of all suppose

$$|D^2L(p)| \le C \quad (p \in \mathbb{R}^n). \tag{5}$$

In addition let us assume that L is uniformly convex, and so there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^n L_{p_ip_j}(p)\xi_i\xi_j \ge \theta|\xi|^2 \quad (p, \xi \in \mathbb{R}^n). \tag{6}$$

Clearly this is some sort of nonlinear analogue of our uniform ellipticity condition for linear PDE in Chapter 6. The idea will therefore be to try to utilize, or at least mimic, some of the calculations from that chapter.

THEOREM 1 (Second derivatives for minimizers).

(i) *Let $u \in H^1(U)$ be a weak solution of the nonlinear partial differential equation* (3), *where L satisfies* (5), (6). *Then*

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

(ii) *If in addition $u \in H^1_0(U)$ and ∂U is C^2, then*

$$u \in H^2(U),$$

with the estimate

$$\|u\|_{H^2(U)} \leq C\|f\|_{L^2(U)}.$$

Proof. 1. We will largely follow the proof of Theorem 1 in §6.3.1, the corresponding assertion of local H^2 regularity for solutions of linear second-order elliptic PDE.

Fix any open set $V \subset\subset U$ and choose then an open set W so that $V \subset\subset W \subset\subset U$. Select a smooth cutoff function ζ satisfying

$$\begin{cases} \zeta \equiv 1 \text{ on } V, \quad \zeta \equiv 0 \text{ in } \mathbb{R}^n - W \\ 0 \leq \zeta \leq 1. \end{cases}$$

Let $|h| > 0$ be small, choose $k \in \{1, \dots, n\}$, and substitute

$$v := -D_k^{-h}(\zeta^2 D_k^h u)$$

into (4). We are employing here the notation from §5.8.2:

$$D_k^h u(x) = \frac{u(x + he_k) - u(x)}{h} \quad (x \in W).$$

Using the identity $\int_U u D_k^{-h} v\,dx = -\int_U v D_k^h u\,dx$, we deduce

$$\sum_{i=1}^n \int_U D_k^h(L_{p_i}(Du))(\zeta^2 D_k^h u)_{x_i}\,dx = -\int_U f D_k^{-h}(\zeta^2 D_k^h u)\,dx. \tag{7}$$

Now

$$\begin{aligned} D_k^h L_{p_i}(Du(x)) &= \frac{L_{p_i}(Du(x + he_k)) - L_{p_i}(Du(x))}{h} \\ &= \frac{1}{h}\int_0^1 \frac{d}{ds} L_{p_i}(sDu(x + he_k) + (1 - s)Du(x))\,ds \\ &= \frac{1}{h}\int_0^1 \sum_{j=1}^n L_{p_i p_j}(sDu(x + he_k) + (1-s)Du(x))\,ds \\ &\qquad\qquad (u_{x_j}(x + he_k) - u_{x_j}(x)) \\ &= \sum_{j=1}^n a_{ij}^h(x) D_k^h u_{x_j}(x), \end{aligned} \tag{8}$$

for

$$(9)\quad a_{ij}^h(x) := \int_0^1 L_{p_ip_j}(sDu(x+he_k)+(1-s)Du(x))\,ds \quad (i,j=1,\dots,n).$$

We substitute (8) into (7) and perform simple calculations, to arrive at the identity:

$$(10)\quad \begin{aligned} A_1 + A_2 &:= \sum_{i,j=1}^n \int_U \zeta^2 a_{ij}^h D_k^h u_{x_j} D_k^h u_{x_i}\,dx \\ &\qquad + \sum_{i,j=1}^n \int_U a_{ij}^h D_k^h u_{x_j} D_k^h u 2\zeta\zeta_{x_i}\,dx \\ &= -\int_U f D_k^{-h}(\zeta^2 D_k^h u)\,dx =: B. \end{aligned}$$

Now the uniform convexity condition (6) implies

$$(11)\quad A_1 \geq \theta \int_U \zeta^2 |D_k^h Du|^2\,dx.$$

Furthermore we see from (5) that

$$(12)\quad \begin{aligned} |A_2| &\leq C\int_W \zeta |D_k^h Du||D_k^h u|\,dx \\ &\leq \epsilon \int_W \zeta^2 |D_k^h Du|^2\,dx + \frac{C}{\epsilon}\int_W |D_k^h u|^2\,dx. \end{aligned}$$

Furthermore, as in the proof of Theorem 1 in §6.3.1, we have

$$|B| \leq \epsilon \int_U \zeta^2 |D_k^h Du|^2 dx + \frac{C}{\epsilon}\int_U f^2 + |Du|^2 dx.$$

We select $\epsilon = \frac{\theta}{4}$, to deduce from the foregoing bounds on A_1, A_2, B, the estimate

$$\int_U \zeta^2 |D_k^h Du|^2\,dx \leq C\int_W f^2 + |D_k^h u|^2\,dx \leq C\int_U f^2 + |Du|^2\,dx,$$

the last inequality valid according to Theorem 3,(i) in §5.8.2.

2. Since $\zeta \equiv 1$ on V, we find

$$\int_V |D_k^h Du|^2\,dx \leq C\int_U f^2 + |Du|^2\,dx$$

for $k = 1, \ldots, n$ and all sufficiently small $|h| > 0$. Consequently Theorem 3, (iii) in §5.8.2 implies $Du \in H^1(V)$, and so $u \in H^2(V)$. This is true for each $V \subset\subset U$; thus $u \in H^2_{\text{loc}}(U)$.

3. If $u \in H^1_0(U)$ is a weak solution of (3) and ∂U is C^2, we can then mimic the proof of the boundary regularity Theorem 4 in §6.3.2 to prove $u \in H^2(U)$, with estimate

$$||u||_{H^2(U)} \leq C(||f||_{L^2(U)} + ||u||_{H^1(U)});$$

details are left to the reader. Now from (6) follows the inequality

$$(DL(p) - DL(0)) \cdot p \geq \theta |p|^2 \quad (p \in \mathbb{R}^n).$$

If we then put $v = u$ in (4), we can employ this estimate to derive the bound

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

and so finish the proof. □

8.3.2. Remarks on higher regularity.

We would next like to show that if L is infinitely differentiable, then so is u. By analogy with the regularity theory developed for second-order linear elliptic PDE in §6.3, it may seem natural to try to extend the H^2_{loc} estimate from the previous section to obtain further estimates in the higher Sobolev spaces $H^k_{\text{loc}}(U)$ for $k = 3, 4, \ldots$.

This method will *not* work for the nonlinear partial differential equation (3) however. The reason is this. For linear equations we could, roughly speaking, differentiate the equation many times and still obtain a linear PDE of the same general form as that we began with. See for instance the proof of Theorem 2 in §6.3.1 Whereas if we differentiate a nonlinear differential equation many times, the resulting increasingly complicated expressions quickly become impossible to handle. Much deeper ideas are called for, the full development of which is beyond the scope of this book. We will nevertheless at least outline the basic plan.

To start with, choose a test function $w \in C^\infty_c(U)$, select $k \in \{1, \ldots, n\}$, and set $v = -w_{x_k}$ in the identity (4), where for simplicity we now take $f \equiv 0$. Since we now know $u \in H^2_{\text{loc}}(U)$, we can integrate by parts to find

$$\text{(13)} \qquad \int_U \sum_{i,j=1}^n L_{p_i p_j}(Du) u_{x_k x_j} w_{x_i} \, dx = 0.$$

Next write

$$\tilde{u} := u_{x_k} \tag{14}$$

and

$$a^{ij} := L_{p_i p_j}(Du) \quad (i,j = 1,\dots,n). \tag{15}$$

Fix also any $V \subset\subset U$. Then after an approximation we find from (13)–(15) that

$$\int_U \sum_{i,j=1}^n a^{ij}(x)\tilde{u}_{x_j} w_{x_i}\, dx = 0 \tag{16}$$

for all $w \in H_0^1(V)$. This is to say that $\tilde{u} \in H^1(V)$ is a weak solution of the linear, second order elliptic PDE

$$-\sum_{i,j=1}^n \left(a^{ij}\tilde{u}_{x_j}\right)_{x_i} = 0 \quad \text{in } V. \tag{17}$$

But we *cannot* just apply our regularity theory from §6.3 to conclude from (17) that $\tilde{u}$ is smooth, the reason being that we can deduce from (5) and (15) only that

$$a^{ij} \in L^\infty(V) \quad (i,j = 1,\dots,n).$$

However a deep theorem, due independently to DeGiorgi and to Nash, asserts that any weak solution of (17) must in fact be locally Hölder continuous for some exponent $\gamma > 0$. (See Gilbarg–Trudinger [**G-T,** Chapter 8].) Thus if $W \subset\subset V$ we have $\tilde{u} \in C^{0,\gamma}(W)$, and so

$$u \in C^{1,\gamma}_{\text{loc}}(U).$$

Return to the definition (15). If L is smooth, we now know $a^{ij} \in C^{0,\gamma}_{\text{loc}}(U)$ $(i,j = 1,\dots,n)$. Then (3) and an older theorem of Schauder [**G-T,** Chapter 4] assert that in fact

$$u \in C^{2,\gamma}_{\text{loc}}(U).$$

But then $a^{ij} \in C^{1,\gamma}_{\text{loc}}(U)$; and so another version of Schauder's estimate implies

$$u \in C^{3,\gamma}_{\text{loc}}(U).$$

We can continue this so-called "bootstrap" argument, eventually to deduce u is $C^{k,\gamma}_{\text{loc}}(U)$ for $k = 1,\dots,$ and so $u \in C^\infty(U)$.

See Giaquinta [**GI**] for much more about regularity theory in the calculus of variations.

8.4. CONSTRAINTS

In this section we consider applications of the calculus of variations to certain constrained minimization problems, and, in particular, discuss the role of *Lagrange multipliers* in the corresponding Euler–Lagrange PDE.

8.4.1. Nonlinear eigenvalue problems.

We investigate first problems with *integral constraints.* To be specific, let us consider the problem of minimizing the energy functional

$$I[w] := \frac{1}{2}\int_U |Dw|^2\, dx \tag{1}$$

over all functions w with, say, $w = 0$ on ∂U, but subject now also to the side condition that

$$J[w] := \int_U G(w)\, dx = 0, \tag{2}$$

where $G : \mathbb{R} \to \mathbb{R}$ is a given, smooth function.

We will henceforth write $g = G'$. Assume now

$$|g(z)| \leq C(|z| + 1), \tag{3}$$

and so

$$|G(z)| \leq C(|z|^2 + 1) \quad (z \in \mathbb{R}) \tag{4}$$

for some constant C.

Let us introduce as well the appropriate admissible class

$$\mathcal{A} := \{w \in H^1_0(U) \mid J[w] = 0\}.$$

We suppose also that the open set U is bounded, connected and has a smooth boundary.

THEOREM 1 (Existence of constrained minimizer). *Assume the admissible set $\mathcal{A}$ is nonempty. Then there exists $u \in \mathcal{A}$ satisfying*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Proof. Choose, as usual, a minimizing sequence $\{u_k\}_{k=1}^\infty \subset \mathcal{A}$ with

$$I[u_k] \to m = \inf_{w \in \mathcal{A}} I[w].$$

Then as above we can extract a subsequence

$$u_{k_j} \rightharpoonup u \quad \text{weakly in } H_0^1(U), \tag{5}$$

with $I[u] \leq m$. We will be done once we show

$$J[u] = 0, \tag{6}$$

so that $u \in \mathcal{A}$. Utilizing the compactness theory from §5.6, we deduce from (5) that

$$u_{k_j} \to u \quad \text{in } L^2(U). \tag{7}$$

Consequently

$$\begin{aligned} |J(u)| = |J(u) - J(u_k)| &\leq \int_U |G(u) - G(u_k)|\,dx \\ &\leq C\int_U |u - u_k|(1 + |u| + |u_k|)\,dx \quad \text{by (3)} \\ &\to 0 \quad \text{as } k \to \infty\,. \end{aligned} \tag{8}$$

□

Far more interesting than the mere existence of constrained minimizers is an examination of the corresponding Euler–Lagrange equation.

THEOREM 2 (Lagrange multiplier). *Let $u \in \mathcal{A}$ satisfy*

$$I[u] = \min_{w \in \mathcal{A}} I[w]. \tag{9}$$

Then there exists a real number λ such that

$$\int_U Du \cdot Dv\,dx = \lambda \int_U g(u)v\,dx \tag{10}$$

for all $v \in H_0^1(U)$.

Remark. Thus u is a weak solution of the nonlinear boundary-value problem

$$\begin{cases} -\Delta u = \lambda g(u) & \text{in } U \\ \quad u = 0 & \text{on } \partial U, \end{cases} \tag{11}$$

where λ is the *Lagrange multiplier* corresponding to the *integral constraint*

$$J[u] = 0. \tag{12}$$

A problem of the form (11) for the unknowns (u, λ), with $u \not\equiv 0$, is a *nonlinear eigenvalue problem.* □

Proof. 1. Fix any function $v \in H_0^1(U)$. Assume first

$$g(u) \text{ is not equal to zero a.e. within } U. \tag{13}$$

Choose then any function $w \in H_0^1(U)$ with

$$\int_U g(u)w\,dx \neq 0; \tag{14}$$

this is possible because of (13). Now write

$$\begin{aligned} j(\tau,\sigma) &:= J[u+\tau v+\sigma w] \\ &= \int_U G(u+\tau v+\sigma w)\,dx \quad (\tau,\sigma \in \mathbb{R}). \end{aligned} \tag{15}$$

Clearly

$$j(0,0) = \int_U G(u)\,dx = 0. \tag{16}$$

In addition, j is C^1 and

$$\frac{\partial j}{\partial \tau}(\tau,\sigma) = \int_U g(u+\tau v+\sigma w)v\,dx, \tag{17}$$

$$\frac{\partial j}{\partial \sigma}(\tau,\sigma) = \int_U g(u+\tau v+\sigma w)w\,dx. \tag{18}$$

Consequently (14) implies

$$\frac{\partial j}{\partial \sigma}(0,0) \neq 0. \tag{19}$$

According to the Implicit Function Theorem (§C.6), there exists a C^1 function $\phi : \mathbb{R} \to \mathbb{R}$ such that

$$\phi(0) = 0 \tag{20}$$

and

$$j(\tau,\phi(\tau)) = 0 \tag{21}$$

for all sufficiently small τ, say $|\tau| \leq \tau_0$. Differentiating, we discover

$$\frac{\partial j}{\partial \tau}(\tau,\phi(\tau)) + \frac{\partial j}{\partial \sigma}(\tau,\phi(\tau))\phi'(\tau) = 0;$$

whence (17) and (18) yield

$$\phi'(0) = -\frac{\int_U g(u)v\,dx}{\int_U g(u)w\,dx}. \tag{22}$$

2. Now set

$$w(\tau) := \tau v + \phi(\tau)w \quad (|\tau| \le \tau_0)$$

and write

$$i(\tau) := I[u + w(\tau)].$$

Since (21) implies $J[u + w(\tau)] = 0$, we see that $u + w(\tau) \in \mathcal{A}$. So the C^1 function $i(\cdot)$ has a minimum at 0. Thus

$$\begin{aligned} 0 = i'(0) &= \int_U (Du + \tau Dv + \phi(\tau)Dw) \cdot (Dv + \phi'(\tau)Dw)\,dx|_{\tau=0} \\ &= \int_U Du \cdot (Dv + \phi'(0)Dw)\,dx. \end{aligned} \tag{23}$$

Recall now (22) and *define*

$$\lambda := \frac{\int_U Du \cdot Dw\,dx}{\int_U g(u)w\,dx},$$

to deduce from (23) the desired equality

$$\int_U Du \cdot Dv\,dx = \lambda \int_U g(u)v\,dx$$

for all $v \in H_0^1(U)$.

3. Suppose now instead of (13) that

$$g(u) = 0 \quad \text{a.e. in } U.$$

Approximating g by bounded functions, we deduce $DG(u) = g(u)Du = 0$ a.e. Hence, since U is connected, $G(u)$ is constant a.e. It follows that $G(u) = 0$ a.e., because $J[u] = \int_U G(u)\,dx = 0$. As $u = 0$ on ∂U in the trace sense, it follows that $G(0) = 0$.

But then $u = 0$ a.e., as otherwise $I[u] > I[0] = 0$. Since $g = 0$ a.e., the identity (10) is trivially valid in this case, for any λ. □

8.4.2. Unilateral constraints, variational inequalities.

We study now calculus of variation problems with certain pointwise, *one-sided constraints* on the values of $u(x)$ for each $x \in U$. For definiteness let us consider the problem of minimizing, say, the energy functional

$$I[w] := \int_U \tfrac{1}{2}|Dw|^2 - fw\,dx, \tag{24}$$

among all functions w belonging to the set

$$\mathcal{A} := \{w \in H_0^1(U) \mid w \geq h \text{ a.e. in } U\}, \tag{25}$$

where $h : \bar{U} \to \mathbb{R}$ is a given smooth function, called the *obstacle*. The convex admissible set $\mathcal{A}$ thus comprises those functions $w \in H_0^1(U)$ satisfying the one-sided or *unilateral* constraint that $w \geq h$. We suppose as well that f is a given, smooth function.

THEOREM 3 (Existence of minimizer). *Assume the admissible set $\mathcal{A}$ is nonempty. Then there exists a unique function $u \in \mathcal{A}$ satisfying*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Proof. 1. The existence of a minimizer follows very easily from the general ideas discussed before. We need only note explicitly that if $\{u_{k_j}\}_{j=1}^{\infty} \subset \mathcal{A}$ is a minimizing sequence with $u_{k_j} \rightharpoonup u$ weakly in $H_0^1(U)$, then by compactness we have $u_{k_j} \to u$ strongly in $L^2(U)$. Since $u_{k_j} \geq h$ a.e., it follows that $u \geq h$ a.e. Therefore $u \in \mathcal{A}$.

2. We now prove uniqueness. Assume u and $\tilde{u} \in \mathcal{A}$ are two minimizers, with $u \neq \tilde{u}$. Then $w := \frac{u+\tilde{u}}{2} \in \mathcal{A}$, and

$$\begin{aligned} I[w] &= \int_U \tfrac{1}{2}\left|\left(\tfrac{Du+D\tilde{u}}{2}\right)\right|^2 - f\left(\tfrac{u+\tilde{u}}{2}\right)dx \\ &= \int \tfrac{1}{8}(|Du|^2 + 2Du\cdot D\tilde{u} + |D\tilde{u}|^2) - f\left(\tfrac{u+\tilde{u}}{2}\right)dx. \end{aligned}$$

Now $2a\cdot b = |a|^2 + |b|^2 - |a-b|^2$. Thus

$$\begin{aligned} I[w] &= \int_U \tfrac{1}{8}(2|Du|^2 + 2|D\tilde{u}|^2 - |Du - D\tilde{u}|^2) - f\left(\tfrac{u+\tilde{u}}{2}\right)dx \\ &< \tfrac{1}{2}\int_U \tfrac{1}{2}|Du|^2 - fu\,dx + \tfrac{1}{2}\int_U \tfrac{1}{2}|D\tilde{u}|^2 - f\tilde{u}\,dx \\ &= \tfrac{1}{2}I[u] + \tfrac{1}{2}I[\tilde{u}], \end{aligned}$$

the strict inequality holding since $u \not\equiv \tilde{u}$. This is a contradiction, since u and $\tilde{u}$ are minimizers. □

We next compute the analogue of the Euler–Lagrange equation, which for the case at hand turns out to be an inequality.

THEOREM 4 (Variational characterization of minimizer). *Let $u \in \mathcal{A}$ be the unique solution of*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Then

$$\int_U Du \cdot D(w-u)\,dx \geq \int_U f(w-u)\,dx \quad \text{for all} \quad w \in \mathcal{A}. \tag{26}$$

We call (26) a *variational inequality.*

Proof. 1. Fix any element $w \in \mathcal{A}$. Then for each $0 \leq \tau \leq 1$,

$$u + \tau(w-u) = (1-\tau)u + \tau w \in \mathcal{A},$$

since $\mathcal{A}$ is convex. Thus if we set

$$i(\tau) := I[u + \tau(w-u)],$$

we see that $i(0) \leq i(\tau)$ for all $0 \leq \tau \leq 1$. Hence

$$i'(0) \geq 0. \tag{27}$$

2. Now if $0 < \tau \leq 1$,

$$\begin{aligned}\frac{i(\tau)-i(0)}{\tau} &= \frac{1}{\tau}\int_U \frac{|Du + \tau D(w-u)|^2 - |Du|^2}{2} - f(u+\tau(w-u)-u)\,dx \\ &= \int_U Du \cdot D(w-u) + \frac{\tau|D(w-u)|^2}{2} - f(w-u)\,dx.\end{aligned}$$

Thus (27) implies

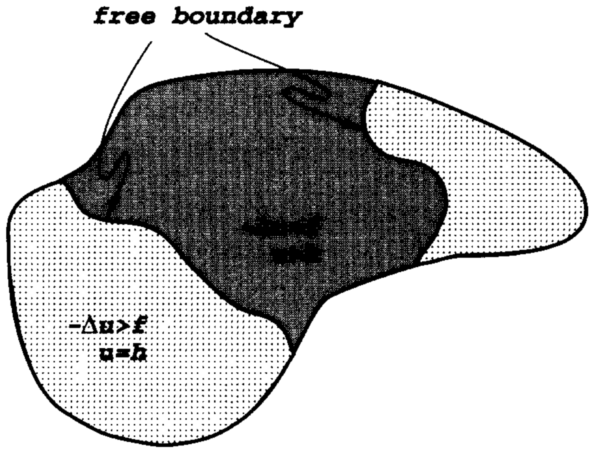
$$0 \leq i'(0) = \int_U Du \cdot D(w-u) - f(w-u)\,dx.$$

□

Notice that we obtain the inequality (27), since we can in effect take only "one-sided" variations, away from the constraint.

Interpretation of the variational inequality. To gain some insight into the variational inequality (26), let us quote without proof a regularity assertion (see Kinderlehrer–Stampacchia [K-S]), which states $u \in W^{2,\infty}(U)$, provided ∂U is smooth. Hence the set

$$O := \{x \in U \mid u(x) > h(x)\}$$



The free boundary for the obstacle problem

is open, and

$$C := \{x \in U \mid u(x) = h(x)\}$$

is (relatively) closed.

We claim that in fact $u \in C^\infty(O)$ and

$$-\Delta u = f \quad \text{in } O. \tag{28}$$

To see this, fix any test function $v \in C_c^\infty(O)$. Then if $|\tau|$ is sufficiently small, $w := u + \tau v \geq h$, and so $w \in \mathcal{A}$. Thus (26) implies $\tau \int_O Du \cdot Dv - fv\,dx \geq 0$. This inequality is valid for all sufficiently small τ, both positive and negative, and so in fact

$$\int_O Du \cdot Dv - fv\,dx = 0$$

for all $v \in C_c^\infty(O)$. Hence u is a weak solution of (28); whence linear regularity theory (§6.3) shows $u \in C^\infty(O)$.

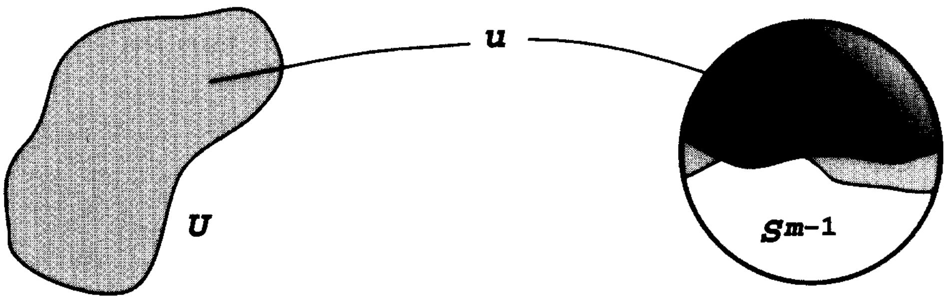
Now if $v \in C_c^\infty(U)$ satisfies $v \geq 0$ and if $0 < \tau \leq 1$, then $w := u + \tau v \in \mathcal{A}$, whence $\int_U Du \cdot Dv - fv\,dx \geq 0$. But since $u \in W^{2,\infty}(U)$, we can integrate by parts to deduce $\int_U (-\Delta u - f)v\,dx \geq 0$ for all nonnegative functions $v \in C_c^\infty(U)$. Thus

$$-\Delta u \geq f \quad \text{a.e. in } U. \tag{29}$$

We summarize our conclusions by observing from (28), (29) that

$$\begin{cases} u \geq h, \ -\Delta u \geq f & \text{a.e. in } U \\ \qquad\quad -\Delta u = f & \text{on } U \cap \{u > h\}. \end{cases} \tag{30}$$

□



A harmonic map into a sphere

Remark. The set

$$F := \partial O \cap U$$

is called the *free boundary.* Many interesting problems in applied mathematics involve partial differential equations with free boundaries. Such of these problems as can be recast as variational inequalities become relatively easy to study, especially in that there is no explicit mention of the free boundary in the inequalities (30). Applications arise in stopping time optimal control problems for Brownian motion, in groundwater hydrology, in plasticity theory, etc. See Kinderlehrer–Stampacchia [**K-S**]. □

8.4.3. Harmonic maps.

We consider next the case of *pointwise* constraints as exemplified by *harmonic maps into spheres.* We are interested now in the problem of minimizing the energy

$$I[\mathbf{w}] := \frac{1}{2}\int_U |D\mathbf{w}|^2\,dx \tag{31}$$

over all functions belonging to the admissible class

$$\mathcal{A} := \{\mathbf{w} \in H^1(U;\mathbb{R}^m) \mid \mathbf{w} = \mathbf{g} \text{ on } \partial U, |\mathbf{w}| = 1 \text{ a.e.}\}. \tag{32}$$

The idea is that we are trying to minimize the energy over all appropriate maps from $U \subset \mathbb{R}^n$ into the unit sphere $S^{m-1} = \partial B(0,1) \subset \mathbb{R}^m$. This problem and its variants arises for instance as a greatly simplified model for the behavior of liquid crystals.

It is straightforward to verify that there exists at least one minimizer in $\mathcal{A}$, provided $\mathcal{A} \neq \emptyset$.

THEOREM 5 (Euler–Lagrange equation for harmonic maps). *Let* $\mathbf{u} \in \mathcal{A}$ *satisfy*

$$I[\mathbf{u}] = \min_{\mathbf{w}\in\mathcal{A}} I[\mathbf{w}].$$

Then

$$\int_U D\mathbf{u} : D\mathbf{v}\,dx = \int_U |D\mathbf{u}|^2 \mathbf{u}\cdot\mathbf{v}\,dx \tag{33}$$

for each $\mathbf{v} \in H_0^1(U;\mathbb{R}^m) \cap L^\infty(U;\mathbb{R}^m)$.

Remark. We interpret (33) as saying that $\mathbf{u} = (u^1, \dots, u^m)$ is a weak solution of the boundary-value problem

$$\begin{cases} -\Delta \mathbf{u} = |D\mathbf{u}|^2 \mathbf{u} & \text{in } U \\ \mathbf{u} = \mathbf{g} & \text{on } \partial U. \end{cases} \tag{34}$$

The *function* $\lambda = |D\mathbf{u}|^2$ is the Lagrange multiplier corresponding to the *pointwise* constraint $|\mathbf{u}| = 1$. Note carefully that for a single, integral constant (§8.4.1) the Lagrange multiplier is a number, but for a pointwise constraint it is a function. □

Proof. 1. Fix $\mathbf{v} \in H_0^1(U;\mathbb{R}^m) \cap L^\infty(U;\mathbb{R}^m)$. Then since $|\mathbf{u}| = 1$ a.e., we have

$$|\mathbf{u} + \tau\mathbf{v}| \neq 0 \text{ a.e.}$$

for each sufficiently small τ. Consequently

$$\mathbf{v}(\tau) := \frac{\mathbf{u}+\tau\mathbf{v}}{|\mathbf{u}+\tau\mathbf{v}|} \in \mathcal{A}. \tag{35}$$

Thus

$$i(\tau) := I[\mathbf{v}(\tau)]$$

has a minimum at $\tau = 0$, and so, as usual,

$$i'(0) = 0. \tag{36}$$

2. Now

$$i'(0) = \int_U D\mathbf{u} : D\mathbf{v}'(0)\,dx \quad \left({}' = \frac{d}{d\tau}\right). \tag{37}$$

But we compute directly from (35) that

$$\mathbf{v}'(\tau) = \frac{\mathbf{v}}{|\mathbf{u}+\tau\mathbf{v}|} - \frac{[(\mathbf{u}+\tau\mathbf{v})\cdot\mathbf{v}](\mathbf{u}+\tau\mathbf{v})}{|\mathbf{u}+\tau\mathbf{v}|^3};$$

whence $\mathbf{v}'(0) = \mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{u}$. Inserting this equality into (36), (37), we find

$$0 = \int_U D\mathbf{u} : D\mathbf{v} - D\mathbf{u} : D((\mathbf{u}\cdot\mathbf{v})\mathbf{u})\,dx. \tag{38}$$

However since $|\mathbf{u}|^2 \equiv 1$, we have

$$(D\mathbf{u})^T\mathbf{u} = \mathbf{0}.$$

Using this fact, we then verify

$$D\mathbf{u} : D((\mathbf{u}\cdot\mathbf{v})\mathbf{u}) = |D\mathbf{u}|^2(\mathbf{u}\cdot\mathbf{v}) \quad \text{a.e. in } U.$$

This identity employed in (38) gives (33). □

8.4.4. Incompressibility.

a. Stokes' problem.

Suppose $U \subset \mathbb{R}^3$ is open, bounded, simply connected, and set

$$I[\mathbf{w}] := \int_U \tfrac{1}{2}|D\mathbf{w}|^2 - \mathbf{f} \cdot \mathbf{w}\, dx,$$

for $\mathbf{w}$ belonging to

$$\mathcal{A} = \{\mathbf{w} \in H_0^1(U; \mathbb{R}^3) \mid \operatorname{div} \mathbf{w} = 0 \text{ in } U\}.$$

Here $\mathbf{f} \in L^2(U; \mathbb{R}^3)$ is given.

There is no problem in showing by customary methods that there exists a unique minimizer $\mathbf{u} \in \mathcal{A}$. We interpret $\mathbf{u}$ as representing the velocity field of a steady fluid flow within the region U, subject to the external force $\mathbf{f}$. The constraint that $\operatorname{div} \mathbf{u} = 0$ ensures that the flow is incompressible: see the Remark at the end of this section.

How does the constraint manifest itself in the Euler–Lagrange equation?

THEOREM 6 (Pressure as Lagrange multiplier). *There exists a scalar function $p \in L^2_{\text{loc}}(U)$ such that*

$$\int_U D\mathbf{u} : D\mathbf{v}\, dx = \int_U p \operatorname{div} \mathbf{v} + \mathbf{f} \cdot \mathbf{v}\, dx \tag{39}$$

for all $\mathbf{v} \in H^1(U; \mathbb{R}^3)$ with compact support within U.

Remark. We interpret (39) as saying that $(\mathbf{u}, p)$ form a weak solution of *Stokes' problem*

$$\begin{cases} -\Delta \mathbf{u} = \mathbf{f} - Dp & \text{in } U \\ \operatorname{div} \mathbf{u} = \mathbf{0} & \text{in } U \\ \mathbf{u} = \mathbf{0} & \text{on } \partial U. \end{cases} \tag{40}$$

The function p is the *pressure* and arises as a Lagrange multiplier corresponding to the *incompressibility condition* $\operatorname{div} \mathbf{u} = 0$. □

Proof. 1. Assume first $\mathbf{v} \in \mathcal{A}$. Then for each $\tau \in \mathbb{R}$, $\mathbf{u} + \tau\mathbf{v} \in \mathcal{A}$. Thus

$$0 = i'(0) = \int_U D\mathbf{u} : D\mathbf{v} - \mathbf{f} \cdot \mathbf{v}\, dx. \tag{41}$$

2. Fix now $V \subset\subset U$, V smooth and simply connected, and select $\mathbf{w} \in H_0^1(V;\mathbb{R}^3)$ with $\operatorname{div}\mathbf{w} = 0$. Choose $0 < \epsilon < \operatorname{dist}(V, \partial U)$ and set $\mathbf{v} = \mathbf{v}^\epsilon := \eta_\epsilon * \mathbf{w}$ in (41), η_ϵ denoting the usual mollifier and $\mathbf{w}$ defined to be zero in $U - V$. Then

$$(42) \qquad 0 = \int_U D\mathbf{u} : D\mathbf{v}^\epsilon - \mathbf{f} \cdot \mathbf{v}^\epsilon \, dx = \int_U D\mathbf{u}^\epsilon : D\mathbf{w} - \mathbf{f}^\epsilon \cdot \mathbf{w} \, dx$$

for

$$(43) \qquad \mathbf{u}^\epsilon := \eta_\epsilon * \mathbf{u}, \quad \mathbf{f}^\epsilon := \eta_\epsilon * \mathbf{f}.$$

As $\mathbf{u}^\epsilon$ is smooth, (42) implies

$$(44) \qquad \int_V (-\Delta \mathbf{u}^\epsilon - \mathbf{f}^\epsilon) \cdot \mathbf{w} \, dx = 0$$

for each $\mathbf{w} \in H_0^1(V;\mathbb{R}^3)$ with $\operatorname{div}\mathbf{w} = 0$.

3. Fix any smooth vector field $\boldsymbol{\zeta} \in C_c^\infty(V;\mathbb{R}^3)$ and put $\mathbf{w} = \operatorname{curl}\boldsymbol{\zeta}$ in (44). This is legitimate since $\operatorname{div}\mathbf{w} = \operatorname{div}(\operatorname{curl}\boldsymbol{\zeta}) = 0$. Then, temporarily writing $\mathbf{h} := \Delta\mathbf{u}^\epsilon + \mathbf{f}^\epsilon$, we find

$$0 = \int_V \mathbf{h} \cdot \operatorname{curl}\boldsymbol{\zeta} \, dx = \int_V h^1(\zeta^3_{x_2} - \zeta^2_{x_3}) + h^2(\zeta^1_{x_3} - \zeta^3_{x_1}) + h^3(\zeta^2_{x_1} - \zeta^1_{x_2}) \, dx,$$

for $\mathbf{h} = (h^1, h^2, h^3)$, $\boldsymbol{\zeta} = (\zeta^1, \zeta^2, \zeta^3)$. An integration by parts reveals

$$0 = \int_V \zeta^1(h^3_{x_2} - h^2_{x_3}) + \zeta^2(h^1_{x_3} - h^3_{x_1}) + \zeta^3(h^2_{x_1} - h^1_{x_2}) \, dx.$$

As $\zeta^1, \zeta^2, \zeta^3 \in C_c^\infty(V)$ are arbitrary, we deduce $\operatorname{curl}\mathbf{h} = 0$ in V. Since V is simply connected, there consequently exists a smooth function p^ϵ in V such that

$$(45) \qquad Dp^\epsilon = \mathbf{h} = \Delta\mathbf{u}^\epsilon + \mathbf{f}^\epsilon \quad \text{in } V.$$

4. If necessary we can add a constant to p^ϵ to ensure $\int_V p^\epsilon \, dx = 0$.

In view of this normalization, there exists a smooth vector field $\mathbf{v}^\epsilon : V \to \mathbb{R}^3$ solving

$$(46) \qquad \begin{cases} \operatorname{div}\mathbf{v}^\epsilon = p^\epsilon & \text{in } V \\ \mathbf{v}^\epsilon = 0 & \text{on } \partial V. \end{cases}$$

In addition we have the estimate

$$(47) \qquad \|\mathbf{v}^\epsilon\|_{H^1(V;\mathbb{R}^3)} \leq C\|p^\epsilon\|_{L^2(V)},$$

the constant C depending only on V. (We omit the proof of the existence of the vector field $\mathbf{v}^\epsilon$: the construction both is intricate and requires knowledge of certain estimates for Laplace's equation with Neumann-type boundary conditions beyond the scope of this book. See Dacorogna–Moser [**D-M**] for details.)

Now compute

$$\begin{aligned}
\int_V (p^\epsilon)^2\,dx &= \int_V p^\epsilon \operatorname{div}\mathbf{v}^\epsilon\,dx \quad \text{by (46)} \\
&= -\int_V Dp^\epsilon\cdot\mathbf{v}^\epsilon\,dx \\
&= \int_V (-\Delta\mathbf{u}^\epsilon - \mathbf{f}^\epsilon)\cdot\mathbf{v}^\epsilon\,dx \quad \text{by (45)} \\
&= \int_V D\mathbf{u}^\epsilon : D\mathbf{v}^\epsilon - \mathbf{f}^\epsilon\cdot\mathbf{v}^\epsilon\,dx \\
&\le \|\mathbf{v}^\epsilon\|_{H^1(V;\mathbb{R}^3)}(\|\mathbf{u}^\epsilon\|_{H^1(V)} + \|\mathbf{f}^\epsilon\|_{L^2(V)}) \\
&\le C\|p^\epsilon\|_{L^2(U)}(\|\mathbf{u}\|_{H_0^1(U)} + \|\mathbf{f}\|_{L^2(U)}) \quad \text{by (47)}.
\end{aligned}$$

Thus

$$\|p^\epsilon\|_{L^2(V)} \le C\left(\|\mathbf{u}\|_{H^1(V)} + \|\mathbf{f}\|_{L^2(V)}\right). \tag{48}$$

5. In view of estimate (48) there exists a subsequence $\epsilon_j \to 0$ so that

$$p^{\epsilon_j} \rightharpoonup p \quad \text{weakly in } L^2(V) \tag{49}$$

for some $p \in L^2(V)$. Now (45) implies

$$\int_V D\mathbf{u}^\epsilon : D\mathbf{v}\,dx = \int_V p^\epsilon \operatorname{div}\mathbf{v} + \mathbf{f}^\epsilon\cdot\mathbf{v}\,dx$$

for all $\mathbf{v} \in H_0^1(V;\mathbb{R}^3)$. Sending $\epsilon = \epsilon_j \to 0$ we find

$$\int_V D\mathbf{u} : D\mathbf{v}\,dx = \int_V p\operatorname{div}\mathbf{v} + \mathbf{f}\cdot\mathbf{v}\,dx \tag{50}$$

as well.

6. Finally choose a sequence of sets $V_k \subset\subset U$ $(k = 1,\dots)$ as above, with $V_1 \subset V_2 \subset V_3 \subset \cdots$ and $U = \bigcup_{k=1}^\infty V_k$. Utilizing steps 2–5 we find $p_k \in L^2(V_k)$ $(k = 1,\dots)$ so that

$$\int_{V_k} D\mathbf{u} : D\mathbf{v}\,dx = \int_{V_k} p_k \operatorname{div}\mathbf{v} + \mathbf{f}\cdot\mathbf{v}\,dx \tag{51}$$

for each $\mathbf{v} \in H_0^1(V_k;\mathbb{R}^3)$. Adding constants as necessary to each p_k, we deduce from (51) that if $1 \le l \le k$, then $p_k = p_l$ on V_l. We finally define $p = p_k$ on V_k $(k = 1,\dots)$. □

b. Incompressible nonlinear elasticity.

We return now to the model of nonlinear elasticity discussed before in §8.2.4. Suppose that $\mathbf{u}$ represents the displacement of an elastic body which has the rest configuration U. Let us suppose now that the elastic body is incompressible, which now means

$$\det D\mathbf{u} = 1.$$

We therefore suppose the energy density function $L : \mathbb{M}^{3\times 3} \times U \to \mathbb{R}$ is given, and consider the problem of minimizing the elastic energy

$$I[\mathbf{w}] := \int_U L(D\mathbf{w}, x)\, dx$$

over all $\mathbf{w}$ in the admissible set

$$\mathcal{A} := \{\mathbf{w} \in W^{1,q}(U; \mathbb{R}^3) \mid \mathbf{w} = \mathbf{g} \text{ on } \partial U, \det D\mathbf{w} = 1 \text{ a.e.}\}$$

for some $q > 3$.

THEOREM 7 (Minimizers with determinant constraint). *Assume the mapping*

$$P \mapsto L(P, x)$$

is convex, and L satisfies the coercivity condition

$$L(P, x) \geq \alpha |P|^q - \beta \quad (P \in \mathbb{M}^{3\times 3}, x \in U)$$

for some $\alpha > 0$, $\beta \geq 0$. Suppose finally $\mathcal{A} \neq \emptyset$.

Then there exists $\mathbf{u} \in \mathcal{A}$ satisfying

$$I[\mathbf{u}] = \min_{w \in \mathcal{A}} I[\mathbf{w}].$$

Proof. We as usual select a minimizing sequence, with

$$\mathbf{u}_{k_j} \rightharpoonup \mathbf{u} \quad \text{weakly in } W^{1,q}(U; \mathbb{R}^3).$$

Since

$$I[\mathbf{u}] \leq \liminf_{j\to\infty} I\left[\mathbf{u}_{k_j}\right],$$

we must only show that $u \in \mathcal{A}$. However, since in view of the lemma in §8.2.4 we have $\det D\mathbf{u}_{k_j} \rightharpoonup \det D\mathbf{u}$ weakly in $L^{q/n}(U)$, we see that $\det D\mathbf{u} = 1$ a.e., as required. □

Remark. It may seem odd that the incompressibility condition in example (a) is

$$\operatorname{div} \mathbf{u} = 0 \tag{52}$$

and in example (b) is

$$\det D\mathbf{u} = 1. \tag{53}$$

The explanation is that $\mathbf{u}$ represents a velocity in (52) and a displacement in (53). More generally if $\mathbf{w}$ is a velocity field, say of a fluid, we compute the motion of a particle initially at a point x by solving the ODE

$$\begin{cases} \dot{\mathbf{y}}(t) = \mathbf{w}(\mathbf{y}(t), t) & (t \in \mathbb{R}) \\ \mathbf{y}(0) = x. \end{cases}$$

Write $\mathbf{y}(t) = \mathbf{y}(t, x)$ to display the dependence on the initial position x. Then for each $t > 0$, the mapping $x \mapsto \mathbf{y}(t, x)$ is volume-preserving if

$$J(x, t) = \det D_x \mathbf{y}(t, x) = 1 \quad \text{for all } x.$$

Clearly $J(0, x) = 1$, and a calculation verifies *Euler's formula:*

$$J_t(x, t) = (\operatorname{div} \mathbf{w})(\mathbf{y}(t, x), t) J(x, t),$$

the divergence taken with respect to the spatial variables. Hence if $\operatorname{div} \mathbf{w} \equiv 0$, the flow is volume preserving. □

8.5. CRITICAL POINTS

Thus far we have studied the problem of locating minimizers of various energy functionals, subject perhaps to constraints, and of discovering the appropriate nonlinear Euler–Lagrange equations they satisfy. For this section we turn our attention to the problem of finding additional solutions of the Euler–Lagrange PDE, by looking for other *critical points.* These critical points will not in general be minimizers, but rather "saddle points" of $I[\cdot]$.

8.5.1. Mountain Pass Theorem.

We develop here some machinery which ensures that an abstract functional $I[\cdot]$ has a critical point.

a. Critical points, deformations.

Hereafter H denotes a real Hilbert space, with norm $\| \quad \|$ and inner product $(\ , \ )$. Let $I : H \to \mathbb{R}$ be a nonlinear functional on H.

DEFINITION. *We say I is* differentiable *at $u \in H$ if there exists $v \in H$ such that*

$$I[w] = I[u] + (v, w-u) + o(\|w-u\|) \quad (w \in H). \tag{1}$$

The element v, if it exists, is unique. We then write $I'[u] = v$.

DEFINITION. *We say I belongs to $C^1(H;\mathbb{R})$ if $I'[u]$ exists for each $u \in H$, and the mapping $I' : H \to H$ is continuous.*

Remark. The theory we will develop below holds if $I \in C^1(H;\mathbb{R})$, but the proofs will be greatly streamlined provided we additionally assume

$$I' : H \to H \text{ is Lipschitz continuous on bounded subsets of } H. \tag{2}$$

□

Notation. (i) We denote by $\mathcal{C}$ the collection of functions $I \in C^1(H;\mathbb{R})$ satisfying (2).

(ii) If $c \in \mathbb{R}$, we write

$$A_c := \{u \in H \mid I[u] \le c\}, \ K_c := \{u \in H \mid I[u] = c, \ I'[u] = 0\}.$$

□

DEFINITIONS. (i) *We say $u \in H$ is a* critical point *if $I'[u] = 0$.*

(ii) *The real number c is a* critical value *if $K_c \neq \emptyset$.*

We now want to prove that if c is *not* a critical level, we can nicely deform the set $A_{c+\epsilon}$ into $A_{c-\epsilon}$ for some $\epsilon > 0$. The idea will be to solve an appropriate ODE in H and to follow the resulting flow "downhill". As H is generally infinite dimensional, we will need some kind of compactness condition.

DEFINITION. *A functional $I \in C^1(H;\mathbb{R})$ satisfies the* Palais-Smale compactness condition *if each sequence $\{u_k\}_{k=1}^\infty \subset H$ such that*

(i) *$\{I[u_k]\}_{k=1}^\infty$ is bounded*

and

(ii) *$I'[u_k] \to 0$ in H,*

is precompact in H.

THEOREM 1 (Deformation Theorem). *Assume $I \in \mathcal{C}$ satisfies the Palais-Smale condition. Suppose also*

$$K_c = \emptyset. \tag{3}$$

Then for each sufficiently small $\epsilon > 0$, there exists a constant $0 < \delta < \epsilon$ and a function

$$\eta \in C([0,1] \times H; H)$$

such that the mappings

$$\eta_t(u) = \eta(t,u) \quad (0 \le t \le 1, u \in H)$$

satisfy

(i) $\eta_0(u) = u \quad (u \in H)$,
(ii) $\eta_1(u) = u \quad (u \notin I^{-1}[c-\epsilon, c+\epsilon])$,
(iii) $I[\eta_t(u)] \le I[u] \quad (u \in H, 0 \le t \le 1)$,
(iv) $\eta_1(A_{c+\delta}) \subset A_{c-\delta}$.

Proof. 1. We first claim that there exist constants $0 < \sigma$, $\epsilon < 1$ such that

$$\|I'[u]\| \ge \sigma \text{ for each } u \in A_{c+\epsilon} - A_{c-\epsilon}. \tag{4}$$

The proof is by contradiction. Were (4) false for all constants $\sigma, \epsilon > 0$, there would exist sequences $\sigma_k \to 0$, $\epsilon_k \to 0$ and elements

$$u_k \in A_{c+\epsilon_k} - A_{c-\epsilon_k} \tag{5}$$

with

$$\|I'[u_k]\| \le \sigma_k. \tag{6}$$

According to the Palais-Smale condition, there is a subsequence $\{u_{k_j}\}_{j=1}^\infty$ and an element $u \in H$ with $u_{k_j} \to u$ in H. But then since $I \in C^1(H;\mathbb{R})$, (5) and (6) imply $I[u] = c$, $I'[u] = 0$. Consequently $K_c \neq \emptyset$, a contradiction to our hypothesis (3).

2. Now fix δ to satisfy

$$0 < \delta < \epsilon, \quad 0 < \delta < \frac{\sigma^2}{2}. \tag{7}$$

Write

$$A := \{u \in H \mid I[u] \le c - \epsilon \text{ or } I[u] \ge c + \epsilon\},$$
$$B := \{u \in H \mid c - \delta \le I[u] \le c + \delta\}.$$

Since I' is bounded on bounded sets, we verify that the mapping $u \mapsto \operatorname{dist}(u, A) + \operatorname{dist}(u, B)$ is bounded below by a positive constant on each bounded subset of H. Consequently, the function

$$g(u) := \frac{\operatorname{dist}(u, A)}{\operatorname{dist}(u, A) + \operatorname{dist}(u, B)} \quad (u \in H)$$

satisfies

$$0 \leq g \leq 1, \ g = 0 \text{ on } A, \ g = 1 \text{ on } B. \tag{8}$$

Set

$$h(t) := \begin{cases} 1, & 0 \leq t \leq 1 \\ 1/t, & t \geq 1. \end{cases} \tag{9}$$

Finally define the mapping $V : H \to H$ by

$$V(u) := -g(u)h(\|I'[u]\|)I'[u] \quad (u \in H). \tag{10}$$

Observe that V is bounded.

3. Consider now for each $u \in H$ the ODE

$$\begin{cases} \frac{d\eta}{dt}(t) = V(\eta(t)) & (t > 0) \\ \eta(0) = u. \end{cases} \tag{11}$$

As V is bounded and Lipschitz continuous on bounded sets there is a unique solution, existing for all times $t \geq 0$. We write $\eta = \eta(t, u) = \eta_t(u)$ $(t \geq 0, u \in H)$ to display the dependence of the solution on both the time t and the initial position $u \in H$. Restricting ourselves to times $0 \leq t \leq 1$, we see that the mapping $\eta \in C([0,1] \times H; H)$ so defined satisfies assertions (i) and (ii).

4. We now compute

$$\begin{aligned} \frac{d}{dt} I[\eta_t(u)] &= \left(I'[\eta_t(u)], \frac{d}{dt}\eta_t(u) \right) \\ &= (I'[\eta_t(u)], V(\eta_t(u))) \\ &= -g(\eta_t(u))h(\|I'[\eta_t(u)]\|) \ \|I'[\eta_t(u)]\|^2. \end{aligned} \tag{12}$$

In particular

$$\frac{d}{dt} I[\eta_t(u)] \leq 0 \quad (u \in H, 0 \leq t \leq 1),$$

and so assertion (iii) is valid.

5. Now fix any point

$$u \in A_{c+\delta}. \tag{13}$$

We want to prove

$$\eta_1(u) \in A_{c-\delta} \tag{14}$$

and thereby verify assertion (iv). If $\eta_t(u) \notin B$ for some $0 \leq t \leq 1$, we are done; and so we may as well suppose instead $\eta_t(u) \in B$ $(0 \leq t \leq 1)$. Then $g(\eta_t(u)) = 1$ $(0 \leq t \leq 1)$. Consequently, calculation (12) yields

$$\frac{d}{dt} I[\eta_t(u)] = -h(\|I'[\eta_t(u)]\|)\|I'[\eta_t(u)]\|^2. \tag{15}$$

Now if $\|I'[\eta_t(u)]\| \geq 1$, then (9) and (4) imply

$$\frac{d}{dt} I[\eta_t(u)] = -\|I'[\eta_t(u)]\| \leq -\sigma.$$

On the other hand, if $\|I'[\eta_t(u)]\| \leq 1$, (9) and (4) yield

$$\frac{d}{dt} I[\eta_t(u)] \leq -\sigma^2.$$

Both these inequalities, and (15), then imply

$$I[\eta_t(u)] \leq I[u] - \sigma^2 \leq c + \delta - \sigma^2 \leq c - \delta \text{ by (7).}$$

This estimate establishes (14) and completes the proof. □

b. Mountain Pass Theorem.

Next we employ an interesting "min-max" technique, using the deformation η built above to deduce the existence of a critical point.

THEOREM 2 (Mountain Pass Theorem). *Assume $I \in C$ satisfies the Palais-Smale condition. Suppose also*

(i) $I[0] = 0$,

(ii) *there exist constants $r, a > 0$ such that*

$$I[u] \geq a \quad \textit{if} \quad \|u\| = r,$$

and

(iii) *there exists an element $v \in H$ with*

$$\|v\| > r, \quad I[v] \leq 0.$$

Define

$$\Gamma := \{\mathbf{g} \in C([0,1]; H) \mid \mathbf{g}(0) = 0,\ \mathbf{g}(1) = v\}.$$

Then

$$c = \inf_{\mathbf{g}\in\Gamma} \max_{0\leq t\leq 1} I[\mathbf{g}(t)]$$

is a critical value of I.

Think of the graph of $I[\cdot]$ as a landscape with a low spot at 0, surrounded by a ring of mountains. Beyond these mountains lies another low spot at v. The idea is to look for a path $\mathbf{g}$ connecting 0 to v, which passes through a mountain pass, that is, a saddle point for $I[\cdot]$. But note carefully: we are only asserting the existence of a critical point at the "energy level" c, which may not necessarily correspond to a true saddle point.

Proof. 1. Clearly

$$c \geq a. \tag{16}$$

2. Assume that c is not a critical value of I, so that

$$K_c = \emptyset. \tag{17}$$

Choose then any sufficiently small number

$$0 < \epsilon < \frac{a}{2}. \tag{18}$$

According to the Deformation Theorem 1, there exist a constant $0 < \delta < \epsilon$ and a homeomorphism $\boldsymbol{\eta} : H \to H$ with

$$\boldsymbol{\eta}(A_{c+\delta}) \subset A_{c-\delta} \tag{19}$$

and

$$\boldsymbol{\eta}(u) = u \quad \text{if} \quad u \notin I^{-1}[c-\delta, c+\epsilon]. \tag{20}$$

3. Now select $\mathbf{g} \in \Gamma$ satisfying

$$\max_{0\leq t\leq 1} I[\mathbf{g}(t)] \leq c + \delta. \tag{21}$$

Then $\hat{\mathbf{g}} := \boldsymbol{\eta} \circ \mathbf{g}$ also belongs to Γ, since $\boldsymbol{\eta}(\mathbf{g}(0)) = \boldsymbol{\eta}(0) = 0$ and $\boldsymbol{\eta}(\mathbf{g}(1)) = \boldsymbol{\eta}(v) = v$, according to (20). But then (21) implies $\max_{0\leq t\leq 1} I[\hat{\mathbf{g}}(t)] \leq c - \delta$; whence $c = \inf_{\mathbf{g}\in\Gamma} \max_{0\leq t\leq 1} I[\mathbf{g}(t)] \leq c - \delta$, a contradiction. □

8.5.2. Application to semilinear elliptic PDE.

To illustrate the utility of the Mountain Pass Theorem, let us investigate now the semilinear boundary-value problem:

$$\begin{cases} -\Delta u = f(u) & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{22}$$

We assume f is smooth, and for some

$$1 < p < \frac{n+2}{n-2}$$

we have

$$|f(z)| \leq C(1+|z|^p), \quad |f'(z)| \leq C(1+|z|^{p-1}) \quad (z \in \mathbb{R}), \tag{23}$$

where C is a constant. We will suppose also

$$0 \leq F(z) \leq \gamma f(z)z \text{ for some constant } \gamma < \tfrac{1}{2}, \tag{24}$$

where $F(z) := \int_0^z f(s)ds$ and $z \in \mathbb{R}$. We hypothesize finally for constants $0 < a \leq A$ that

$$a|z|^{p+1} \leq |F(z)| \leq A|z|^{p+1} \quad (z \in \mathbb{R}). \tag{25}$$

Now (25) implies $f(0) = 0$, and so obviously $u \equiv 0$ is a trivial solution of (22). We want to find another.

Remark. Observe that the PDE

$$-\Delta u = |u|^{p-1}u$$

falls under the hypotheses above. We will return to this particular nonlinearity again in §9.4.2. □

THEOREM 3 (Existence). *The boundary-value problem* (22) *has at least one weak solution* $u \not\equiv 0$.

Proof. 1. Define

$$I[u] := \int_U \tfrac{1}{2}|Du|^2 - F(u)\,dx \tag{26}$$

for $u \in H_0^1(U)$. We intend to apply the Mountain Pass Theorem to $I[\cdot]$.

We set $H = H_0^1(U)$, with the norm $\|u\| = \left(\int_U |Du|^2\,dx\right)^{1/2}$ and inner product $(u,v) = \int_U Du \cdot Dv\,dx$. Then

$$I[u] = \tfrac{1}{2}\|u\|^2 - \int_U F(u)\,dx =: I_1[u] - I_2[u].$$

2. We first claim

$$\text{(27)} \qquad I \text{ belongs to the class } \mathcal{C}.$$

To see this, note first that for each $u, w \in H$,

$$I_1[w] = \tfrac{1}{2}\|w\|^2 = \tfrac{1}{2}\|u + w - u\|^2 = \tfrac{1}{2}\|u\|^2 + (u, w-u) + \tfrac{1}{2}\|w-u\|^2.$$

Hence I_1 is differentiable at u, with $I_1'[u] = u$. Consequently, $I_1 \in \mathcal{C}$.

3. We must next examine the term I_2. Recall from the Lax–Milgram Theorem (§6.2.1) that for each element $v^* \in H^{-1}(U)$, the problem

$$\begin{cases} -\Delta v = v^* & \text{in } U \\ \quad v = 0 & \text{on } \partial U \end{cases}$$

has a unique solution $v \in H_0^1(V)$. We will write $v = Kv^*$; so that

$$\text{(28)} \qquad K : H^{-1}(U) \to H_0^1(U) \text{ is an isometry.}$$

Note in particular that if $w \in L^{2n/n+2}(U)$, then the linear functional w^* defined by

$$\langle w^*, u\rangle := \int_U wu\,dx \quad (u \in H_0^1(U))$$

belongs to $H^{-1}(U)$. (We will misuse notation and say "$w \in H^{-1}(U)$".) Observe next that $p\left(\frac{2n}{n+2}\right) < \frac{n+2}{n-2} \cdot \frac{2n}{n+2} = 2^*$, and so $f(u) \in L^{2n/n+2}(U)$ $\subset H^{-1}(U)$ if $u \in H_0^1(U)$.

We now demonstrate that if $u \in H_0^1(U)$, then

$$\text{(29)} \qquad I_2'[u] = K[f(u)].$$

To see this, note first that

$$F(a+b) = F(a) + f(a)b + \int_0^1 (1-s)f'(a+sb)\,ds\; b^2.$$

Thus for each $w \in H_0^1(U)$,

$$\begin{aligned} I_2[w] &= \int_U F(w)\,dx = \int_U F(u+w-u)\,dx \\ \text{(30)} \qquad &= \int_U F(u) + f(u)(w-u)\,dx + R \\ &= I_2(u) + \int_U DK[f(u)] \cdot D(w-u)\,dx + R, \end{aligned}$$

where the remainder term R satisfies, according to (23),

$$\begin{aligned}|R| &\le C\int_U (1+|u|^{p-1}+|w-u|^{p-1})|w-u|^2\,dx\\ &\le C\left(\int_U |w-u|^2+|w-u|^{p+1}\,dx\right)+C\left(\int_U |u|^{p+1}\,dx\right)^{\frac{p-1}{p+1}}\\ &\qquad\qquad\left(\int_U |w-u|^{p+1}\,dx\right)^{\frac{2}{p+1}}.\end{aligned}$$

Since $p+1<2^*$, the Sobolev inequalities show $R=o(\|w-u\|)$. Thus we see from (28) that

$$I_2[w]=I_2[u]+(K[f(u)],w)+o(\|w-u\|),$$

as required.

Finally we note that if $u,\bar{u}\in H_0^1(U)$ with $\|u\|,\|\bar{u}\|\le L$, then

$$\begin{aligned}\|I_2'[u]-I_2'[\bar{u}]\| &= \|K[f(u)]-K[f(\bar{u})]\|_{H_0^1(U)}\\ &= \|f(u)-f(\bar{u})\|_{H^{-1}(U)}\\ &\le \|f(u)-f(\bar{u})\|_{L^{\frac{2n}{n+2}}(U)}.\end{aligned}$$

But

$$\begin{aligned}&\|f(u)-f(\bar{u})\|_{L^{\frac{2n}{n+2}}(U)}\\ &\quad\le C\left(\int_U ((1+|u|^{p-1}+|\bar{u}|^{p-1})|u-\bar{u}|)^{\frac{2n}{n+2}}\,dx\right)^{\frac{n+2}{2n}}\\ &\quad\le C\left(\int_U (1+|u|^{p-1}+|\bar{u}|^{p-1})^{\frac{2n}{n+2}\frac{n+2}{4}}dx\right)^{\frac{2}{n}}\|u-\bar{u}\|_{L^{2^*}(U)}\\ &\quad\le C(L)\|u-\bar{u}\|_{L^{2^*}(U)}\le C(L)\|u-\bar{u}\|,\end{aligned}$$

where we used (23). Thus $I_2':H_0^1(U)\to H_0^1(U)$ is Lipschitz continuous on bounded sets. Consequently $I_2\in\mathcal{C}$, and we have established assertion (27).

4. Now we verify the Palais-Smale condition. For this suppose $\{u_k\}_{k=1}^\infty\subset H_0^1(U)$, with

$$\{I[u_k]\}_{k=1}^\infty \text{ bounded} \tag{31}$$

and

$$I'[u_k]\to 0 \quad \text{in } H_0^1(U). \tag{32}$$

According to the foregoing

$$(33) \qquad u_k - K(f(u_k)) \to 0 \quad \text{in } H_0^1(U).$$

Thus for each $\epsilon > 0$ we have

$$|(I'[u_k], v)| = \left| \int_U Du_k \cdot Dv - f(u_k) v \, dx \right| \leq \epsilon \|v\| \quad (v \in H_0^1(U))$$

for k large enough. Let $v = u_k$ above to find

$$\left| \int_U |Du_k|^2 - f(u_k) u_k \, dx \right| \leq \epsilon \|u_k\|$$

for each $\epsilon > 0$ and for all k sufficiently large. For $\epsilon = 1$ in particular, we see that

$$(34) \qquad \int_U f(u_k) u_k \, dx \leq \|u_k\|^2 + \|u_k\|$$

for all k sufficiently large. But since (31) says

$$\left(\tfrac{1}{2} \|u_k\|^2 - \int_U F(u_k) \, dx \right) \leq C < \infty$$

for all k and some constant C, we deduce

$$\begin{aligned} \|u_k\|^2 &\leq C + 2 \int_U F(u_k) \, dx \\ &\leq C + 2\gamma \left(\|u_k\|^2 + \|u_k\| \right) \quad \text{by (34), (24).} \end{aligned}$$

Since $2\gamma < 1$, we discover that $\{u_k\}_{k=1}^{\infty}$ is bounded in $H_0^1(U)$. Hence there exists a subsequence $\{u_{k_j}\}_{j=1}^{\infty}$ and $u \in H_0^1(U)$, with $u_{k_j} \rightharpoonup u$ weakly in $H_0^1(U)$ and $u_{k_j} \to u$ in $L^{p+1}(U)$, the latter assertion holding since $p+1 < 2^*$. But then $f(u_k) \to f(u)$ in $H^{-1}(U)$, whence $K[f(u_k)] \to K[f(u)]$ in $H_0^1(U)$. Consequently (33) implies

$$(35) \qquad u_{k_j} \to u \quad \text{in } H_0^1(U).$$

5. We finally verify the remaining hypotheses of the Mountain Pass Theorem. Clearly $I[0] = 0$. Suppose now that $u \in H_0^1(U)$, with $\|u\| = r$, for $r > 0$ to be selected below. Then

$$(36) \qquad I[u] = I_1[u] - I_2[u] = \frac{r^2}{2} - I_2[u].$$

Now hypothesis (25) implies, since $p+1<2^*$, that

$$\begin{aligned}|I_2[u]| &\leq C\int_U |u|^{p+1}\,dx \leq C\left(\int_U |u|^{2^*}dx\right)^{\frac{p+1}{2^*}}\\ &\leq C\|u\|^{p+1} \leq Cr^{p+1}.\end{aligned}$$

In view of (36), then

$$I[u] \geq \frac{r^2}{2} - Cr^{p+1} \geq \frac{r^2}{4} = a > 0,$$

provided $r>0$ is small enough, since $p+1>2$. Now fix some element $u \in H$, $u \not\equiv 0$. Write $v := tu$ for $t>0$ to be selected. Then

$$\begin{aligned}I[v] &= I_1[tu] - I_2[tu]\\ &= t^2 I_1[u] - \int_U F(tu)\,dx\\ &\leq t^2 I_1[u] - at^{p+1}\int_U |u|^{p+1}dx \quad \text{by (25)}\\ &< 0\end{aligned}$$

for $t>0$ large enough.

6. We have at last checked all the hypotheses of the Mountain Pass Theorem. There must consequently exist a function $u \in H_0^1(U)$, $u \not\equiv 0$, with

$$I'[u] = u - K[f(u)] = 0.$$

In particular for each $v \in H_0^1(U)$, we have

$$\int_U Du\cdot Dv\,dx = \int_U f(u)v\,dx,$$

and so u is a weak solution of (22). □

See §9.4.2 for further discussion about nonlinear Poisson equations, and in particular the significance of the *critical exponent* $\frac{n+2}{n-2}$ in hypothesis (23).

8.6. PROBLEMS

In the exercises U always denotes a bounded, open subset of $\mathbb{R}^n$, with smooth boundary.

1. This problem illustrates that a weakly convergent sequence can in fact be rather badly behaved.

   (i) Prove $u_k(x) = \sin(kx) \rightharpoonup 0$ as $k \to \infty$ in $L^2(0,1)$.

   (ii) Fix $a, b \in \mathbb{R}$, $0 < \lambda < 1$. Define

$$u_k(x) = \begin{cases} a & \text{if } j/k \le x < \lambda(j+1)/k \\ b & \text{if } \lambda(j+1)/k \le x < (j+1)/k \end{cases} \quad (j = 0, \dots, k-1).$$

   Prove $u_k \rightharpoonup \lambda a + (1-\lambda)b$ in $L^2(0,1)$.

2. Find $L = L(p, z, x)$ so that the PDE

$$-\Delta u + D\phi \cdot Du = f \quad \text{in } U$$

is the Euler–Lagrange equation corresponding to the functional $I[w] := \int_U L(Dw, w, x)\,dx$. Here $\phi, f : \bar{U} \to \mathbb{R}$ are given smooth functions.

3. The *elliptic regularization* of the heat equation is the PDE

$$(*) \qquad u_t - \Delta u - \epsilon u_{tt} = 0 \quad \text{in } U_T,$$

where $\epsilon > 0$ and $U_T = U \times (0, T]$. Show that $(*)$ is the Euler–Lagrange equation corresponding to an energy functional $I_\varepsilon[w] := \int\int_{U_T} L_\varepsilon(Dw, w_t, w, x, t)\,dxdt$.

4. Assume $\eta : \mathbb{R}^n \to \mathbb{R}$ is C^1.

   (i) Show $L(P, z, x) = \eta(z) \det P$ $(P \in \mathbb{M}^{n\times n}, z \in \mathbb{R}^n)$ is a null Lagrangian.

   (ii) Deduce that if $\mathbf{u} : \mathbb{R}^n \to \mathbb{R}^n$ is C^1, then

$$\int_U \eta(\mathbf{u}) \det D\mathbf{u}\,dx$$

   depends only on $\mathbf{u}|_{\partial U}$.

5. (Continuation). Fix $x_0 \notin \mathbf{u}(\partial U)$, and choose η as above so that $\int_{\mathbb{R}^n} \eta\,dz = 1$, $\operatorname{spt}\eta \subset B(x_0, r)$, r chosen so small that $B(x_0, r) \cap \mathbf{u}(\partial U) = \emptyset$. Define

$$\deg(\mathbf{u}, x_0) = \int_U \eta(\mathbf{u}) \det D\mathbf{u}\,dx,$$

the *degree* of $\mathbf{u}$ relative to x_0. Prove the degree is an integer.

6. Let $\Sigma \subset \mathbb{R}^3$ denote the graph of the smooth function $u : U \to \mathbb{R}$, $U \subset \mathbb{R}^2$. Then

$$(*) \qquad \int_U (1 + |Du|^2)^{-1} \det D^2 u \; dx$$

represents the integral of the Gauss curvature over Σ. Prove this expression depends only upon Du restricted to ∂U. (The Gauss–Bonnet Theorem in differential geometry computes $(*)$ in terms of the geodesic curvature of $\partial \Gamma$.)

7. Let $m = n$. Prove

$$L(P) = \operatorname{tr}(P^2) - \operatorname{tr}(P)^2 \quad (P \in \mathbb{M}^{n \times n})$$

is a null Lagrangian.

8. Explain why the methods in §8.2 will not work to prove the existence of a minimizer of the functional

$$I[w] := \int_U (1 + |Dw|^2)^{1/2} \, dx$$

over $\mathcal{A} = \{w \in W^{1,q}(U) \mid w = g \text{ on } \partial U\}$, for any $1 \leq q < \infty$.

9. (Second variation for systems). Assume $\mathbf{u} : U \to \mathbb{R}^m$ is a smooth minimizer of the functional

$$I[\mathbf{w}] := \int_U L(D\mathbf{w}, \mathbf{w}, x) \, dx.$$

(i) Show

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{m} \frac{\partial^2 L}{\partial p_i^k \partial p_j^l}(D\mathbf{u}, \mathbf{u}, x) \eta_k \eta_l \xi_i \xi_j \geq 0$$

for all $x \in U$, $\xi \in \mathbb{R}^n$, $\eta \in \mathbb{R}^m$.

(ii) Give an example of a nonconvex function $L : \mathbb{M}^{m \times n} \to \mathbb{R}$ satisfying

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{m} \frac{\partial^2 L(P)}{\partial p_i^k \partial p_j^l} \eta_k \eta_l \xi_i \xi_j \geq 0$$

for all $P \in \mathbb{M}^{m \times n}$, $\xi \in \mathbb{R}^n$, $\eta \in \mathbb{R}^m$.

10. Use the methods of §8.4.1 to show the existence of a nontrivial weak solution $u \in H_0^1(U)$, $u \not\equiv 0$, of

$$\begin{cases} -\Delta u = |u|^{q-1} u & \text{in } U \\ \quad u = 0 & \text{on } \partial U \end{cases}$$

for $1 < q < \frac{n+2}{n-2}$, $n > 2$.

11. Assume $\beta : \mathbb{R} \to \mathbb{R}$ is smooth, with

$$0 < a \leq \beta'(z) \leq b \quad (z \in \mathbb{R})$$

for constants a, b. Let $f \in L^2(U)$. Formulate what it means for $u \in H^1(U)$ to be a weak solution of the nonlinear boundary-value problem

$$\begin{cases} -\Delta u = f & \text{in } U \\ \frac{\partial u}{\partial \nu} + \beta(u) = 0 & \text{on } \partial U. \end{cases}$$

Prove there exists a unique weak solution.

12. Assume u is a smooth minimizer of the area integral

$$I[w] = \int_U (1 + |Dw|^2)^{1/2}\, dx,$$

subject to given boundary conditions $w = g$ on ∂U and the constraint

$$J[w] = \int_U w\, dx = 1.$$

Prove the graph of u is a surface of constant mean curvature. (Hint: Recall Example 4 in §8.1.2.)

13.

(i) Show there exists a unique minimizer $u \in \mathcal{A}$ of

$$I[w] = \int_U \tfrac{1}{2}|Dw|^2 - fw\, dx,$$

where $f \in L^2(U)$ and

$$\mathcal{A} = \{w \in H_0^1(U) \mid |Dw| \leq 1 \text{ a.e.}\}.$$

(ii) Prove

$$\int_U Du \cdot D(w - u)\, dx \geq \int_U f(w - u)\, dx$$

for all $w \in \mathcal{A}$.

8.7. REFERENCES

Section 8.1 See Giaquinta [**GI**] for more about the calculus of variations. Another good source is Giaquinta–Hildebrandt [**G-H**], and

Zeidler [**ZD**, Vol. 3] is a general reference for variational methods.

Section 8.2 The proof of Theorem 1 in §8.2.2 is from Ladyzhenskaya-Uraltseva [**L-U**].

Section 8.3 See Giaquinta [**GI**] for additional information about regularity (and partial regularity) of minimizers.

Section 8.4 The book by Kinderlehrer–Stampacchia [**K-S**] explains much more about variational inequalities.

Section 8.5 The Mountain Pass Theorem is due to Ambrosetti and Rabinowitz. Consult Rabinowitz [**RA**] (the source of §8.5) for further results on saddle point methods.

Chapter 9

NONVARIATIONAL TECHNIQUES

We gather in this chapter various techniques for proving the existence, nonexistence, uniqueness, and various other properties of solutions for nonlinear elliptic and parabolic partial differential equations that are *not* of variational form.

9.1. MONOTONICITY METHODS

Let us look first at this boundary-value problem for a divergence structure quasilinear PDE:

$$(1) \qquad \begin{cases} -\operatorname{div} \mathbf{a}(Du) = f & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases}$$

where $f \in L^2(U)$ is given, as is the smooth vector field $\mathbf{a} : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{a} = (a^1, \dots, a^n)$. As usual the unknown is $u : \bar{U} \to \mathbb{R}$, $u = u(x)$, where U is a bounded, open subset of $\mathbb{R}^n$ with smooth boundary. Now *if* there exists a function $F : \mathbb{R}^n \to \mathbb{R}$ such that $\mathbf{a}$ is the gradient of F,

$$\mathbf{a}(p) = DF(p) \quad (p \in \mathbb{R}^n), \tag{2}$$

then (1) is the Euler-Lagrange equation corresponding to the Lagrangian $L(p, z, x) = F(p) - f(x)z$. However if there exists no such potential F, the variational methods from Chapter 8 simply do not apply to the problem (1).

We inquire instead if there is rather some direct method of constructing a solution of (1), and in particular ask what are reasonable conditions to place upon the nonlinearity. For motivation let us note that *if* (2) were valid and *if* F were convex (the natural assumption for the variational theory, as we have seen in §8.2.2), then for each $p, q \in \mathbb{R}^n$:

$$\begin{aligned}(\mathbf{a}(p) - \mathbf{a}(q)) \cdot (p - q) &= \sum_{i=1}^{n} (F_{p_i}(p) - F_{p_i}(q))(p_i - q_i) \\ &= \int_0^1 \sum_{i,j=1}^{n} F_{p_i p_j}(p + t(q - p))(p_j - q_j)(p_i - q_i)\, dt \geq 0,\end{aligned}$$

the last inequality following from the convexity of F.

This calculation suggests the following

DEFINITION. *A vector field* $\mathbf{a} : \mathbb{R}^n \to \mathbb{R}^n$ *is called* monotone *provided*

$$(\mathbf{a}(p) - \mathbf{a}(q)) \cdot (p - q) \geq 0 \tag{3}$$

for all $p, q \in \mathbb{R}^n$.

We will show below that the quasilinear PDE does indeed possess a weak solution, under the primary structural assumption that the nonlinearity be monotone. Later we will realize that this condition in effect says that $-\operatorname{div} \mathbf{a}(Du) = f$ is a nonlinear elliptic partial differential equation. So let us henceforth assume that the smooth vector field $\mathbf{a}$ is monotone, and that

$$|\mathbf{a}(p)| \leq C(1 + |p|), \tag{4}$$

$$\mathbf{a}(p) \cdot p \geq \alpha |p|^2 - \beta \tag{5}$$

for all $p \in \mathbb{R}^n$ and appropriate constants $C, \alpha > 0$, $\beta \geq 0$. We will see momentarily that (5) amounts to a coercivity condition on the nonlinearity. We intend now to build a solution of the boundary-value problem (1) as the limit of certain finite dimensional approximations, thereby extending Galerkin's method from Chapter 7 to a new class of nonlinear problems.

More precisely, assume that the functions $w_k = w_k(x)$ $(k = 1, \dots)$ are smooth and

$$\{w_k\}_{k=1}^{\infty} \text{ is an orthonormal basis of } H_0^1(U),$$

taken with the inner product $(u, v) = \int_U Du \cdot Dv \, dx$. (We could for instance take $\{w_k\}_{k=1}^{\infty}$ to be the set of appropriately normalized eigenfunctions for $-\Delta$ in $H_0^1(U)$.)

We will look for a function $u_m \in H_0^1(U)$ of the form

$$(6) \qquad u_m = \sum_{k=1}^{m} d_m^k w_k,$$

where we hope to select the coefficients d_m^k so that

$$(7) \qquad \int_U \mathbf{a}(Du_m) \cdot Dw_k \, dx = \int_U f w_k \, dx \quad (k = 1, \dots, m).$$

This amounts to our requiring that u_m solves the "projection" of the problem (1) onto the finite dimensional subspace spanned by $\{w_k\}_{k=1}^{m}$.

We start with a technical assertion.

LEMMA (Zeros of a vector field). *Assume the continuous function* $\mathbf{v} : \mathbb{R}^n \to \mathbb{R}^n$ *satisfies*

$$(8) \qquad \mathbf{v}(x) \cdot x \geq 0 \quad \text{if } |x| = r,$$

for some $r > 0$. *Then there exists a point* $x \in B(0, r)$ *such that*

$$\mathbf{v}(x) = 0.$$

Proof. Suppose the assertion were false; then $\mathbf{v}(x) \neq 0$ for all $x \in B(0, r)$. Define the continuous mapping $\mathbf{w} : B(0, r) \to \partial B(0, r)$ by setting

$$\mathbf{w}(x) := -\frac{r}{|\mathbf{v}(x)|}\mathbf{v}(x) \quad (x \in B(0, r)).$$

According to Brouwer's fixed point theorem (§8.1.4), there exists a point $z \in B(0, r)$ with

$$(9) \qquad \mathbf{w}(z) = z.$$

But then $z \in \partial B(0, r)$, and so (8) and (9) imply the contradiction

$$r^2 = z \cdot z = \mathbf{w}(z) \cdot z = -\frac{r}{|\mathbf{v}(z)|}\mathbf{v}(z) \cdot z \leq 0.$$

□

THEOREM 1 (Construction of approximate solutions). *For each integer* $m = 1, \dots,$ *there exists a function* u_m *of the form* (6) *satisfying the identities* (7).

Proof. 1. Define the continuous function $\mathbf{v} : \mathbb{R}^m \to \mathbb{R}^m$, $\mathbf{v} = (v^1, \dots, v^m)$, by setting

$$v^k(d) := \int_U \mathbf{a}\Big(\sum_{j=1}^m d_j Dw_j\Big) \cdot Dw_k - f w_k \, dx \quad (k = 1, \dots, m) \tag{10}$$

for each point $d = (d_1, \dots, d_m) \in \mathbb{R}^m$. Now

$$\begin{aligned}
\mathbf{v}(d) \cdot d &= \int_U \mathbf{a}\Big(\sum_{j=1}^m d_j Dw_j\Big) \cdot \Big(\sum_{j=1}^m d_j Dw_j\Big) - f\Big(\sum_{j=1}^m d_j w_j\Big) \, dx \\
&\geq \int_U \alpha\Big|\sum_{j=1}^m d_j Dw_j\Big|^2 - \beta - f\Big(\sum_{j=1}^m d_j w_j\Big) \, dx \quad \text{by (5)} \\
&= \alpha|d|^2 - \beta|U| - \sum_{j=1}^m d_j \int_U f w_j \, dx \\
&\geq \frac{\alpha}{2}|d|^2 - \beta|U| - C\sum_{j=1}^m (f, w_j)^2_{L^2(U)}.
\end{aligned}$$

Now let $u \in H_0^1(U)$ solve the PDE $-\Delta u = f$. Then

$$\int_U Du \cdot Dw_j \, dx = \int_U f w_j \, dx \quad (j = 1, \dots),$$

and consequently

$$\sum_{j=1}^m (f, w_j)^2_{L^2(U)} = \sum_{j=1}^m (u, w_j)^2 \leq \|u\|^2_{H_0^1(U)} \leq C\|f\|^2_{L^2(U)}.$$

Hence $\mathbf{v}(d) \cdot d \geq \frac{\alpha}{2}|d|^2 - C$, for some constant C, and so $\mathbf{v}(d) \cdot d \geq 0$ if $|d| = r$, provided we select $r > 0$ sufficiently large.

2. We apply the lemma, to conclude that $\mathbf{v}(d) = 0$ for some point $d \in \mathbb{R}^m$. Then (10) implies u_m defined by (6) satisfies (7). □

We want to take the limit as $m \to \infty$, and for this will require some uniform estimates.

THEOREM 2 (Energy estimates). *There exists a constant C, depending only on U and $\mathbf{a}$, such that*

$$\|u_m\|_{H_0^1(U)} \leq C(1 + \|f\|_{L^2(U)}) \tag{11}$$

for $m = 1, 2, \dots$.

Proof. Multiply equality (7) by d_m^k and sum for $k = 1, \dots, m$:

$$\int_U \mathbf{a}(Du_m) \cdot Du_m \, dx = \int_U f u_m \, dx.$$

In view of the coercivity inequality (5), we find

$$\alpha \int_U |Du_m|^2 \, dx \leq C + \int_U f u_m \, dx \leq C + \epsilon \int_U u_m^2 \, dx + \frac{1}{4\epsilon} \int_U f^2 dx.$$

We recall Poincaré's inequality, and then choose $\epsilon > 0$ small enough to deduce (11). □

We wish now to employ the L^2 inequalities (11) to pass to limits as $m \to \infty$, obtaining thereby a weak solution of problem (1), which is to say, a function $u \in H_0^1(U)$ satisfying the identity

$$\int_U \mathbf{a}(Du) \cdot Dv \, dx = \int_U f v \, dx \quad \text{for all } v \in H_0^1(U). \tag{12}$$

Employing estimate (11) we can extract a subsequence $\{u_{m_j}\}_{j=1}^\infty$ that converges weakly in $H_0^1(U)$ to a limit u, which we hope to show verifies (12). However, we encounter a major problem here: we cannot directly conclude that

$$\mathbf{a}(Du_{m_j}) \to \mathbf{a}(Du)$$

in any sense whatsoever. Take note: *nonlinearities are (usually) not continuous with respect to weak convergence.* (See Problem 2.)

What saves us is the monotonicity assumption on vector field $\mathbf{a}$.

THEOREM 3 (Existence of weak solution). *There exists a weak solution of the nonlinear boundary-value problem* (1).

Proof. 1. As noted in the foregoing discussion, we can extract a subsequence $\{u_{m_j}\}_{j=1}^\infty \subset \{u_m\}_{m=1}^\infty$ and a function $u \in H_0^1(U)$ such that

$$u_{m_j} \rightharpoonup u \quad \text{weakly in } H_0^1(U) \tag{13}$$

and

$$u_{m_j} \to u \quad \text{in } L^2(U). \tag{14}$$

We must show u satisfies (12).

2. In view of the growth condition (4), $\{\mathbf{a}(Du_m)\}_{m=1}^{\infty}$ is bounded in $L^2(U;\mathbb{R}^n)$; and so we may as well suppose—upon passing to a further subsequence if necessary—that

$$\mathbf{a}(Du_{m_j}) \rightharpoonup \boldsymbol{\xi} \quad \text{weakly in } L^2(U;\mathbb{R}^n), \tag{15}$$

for some $\boldsymbol{\xi} \in L^2(U;\mathbb{R}^n)$. Using identity (7), we deduce

$$\int_U \boldsymbol{\xi} \cdot Dw_k \, dx = \int_U f w_k \, dx$$

for each $k = 1, \dots$. And so

$$\int_U \boldsymbol{\xi} \cdot Dv \, dx = \int_U fv \, dx \quad \text{for each } v \in H_0^1(U). \tag{16}$$

3. To proceed further, let us note from the monotonicity condition (3) that

$$\int_U (\mathbf{a}(Du_m) - \mathbf{a}(Dw)) \cdot (Du_m - Dw) \, dx \geq 0 \tag{17}$$

for $m = 1, \dots$ and all $w \in H_0^1(U)$. But as observed before, equation (7) yields the identity

$$\int_U \mathbf{a}(Du_m) \cdot Du_m \, dx = \int_U f u_m \, dx.$$

Substitute into (17), to find:

$$\int_U f u_m - \mathbf{a}(Du_m) \cdot Dw - \mathbf{a}(Dw) \cdot (Du_m - Dw) \, dx \geq 0.$$

Let $m = m_j \to \infty$ and recall (13)–(15), to deduce

$$\int_U fu - \boldsymbol{\xi} \cdot Dw - \mathbf{a}(Dw) \cdot (Du - Dw) \, dx \geq 0.$$

We simplify using identity (16) with $v = u$, and discover

$$\int_U (\boldsymbol{\xi} - \mathbf{a}(Dw)) \cdot D(u - w) \, dx \geq 0 \quad \text{for all } w \in H_0^1(U). \tag{18}$$

4. Fix any $v \in H_0^1(U)$ and set $w := u - \lambda v$ $(\lambda > 0)$ in (18). We obtain then

$$\int_U (\boldsymbol{\xi} - \mathbf{a}(Du - \lambda Dv)) \cdot Dv \, dx \geq 0.$$

Send $\lambda \to 0$:

$$\int_U (\boldsymbol{\xi} - \mathbf{a}(Du)) \cdot Dv \, dx \geq 0 \quad \text{for all } v \in H_0^1(U). \tag{19}$$

Replacing v by $-v$, we deduce that in fact equality holds above. Then (16) and (19) taken together yield

$$\int_U \mathbf{a}(Du) \cdot Dv \, dx = \int_U fv \, dx \quad \text{for all } v \in H_0^1(U).$$

Hence u is indeed a weak solution of (1). □

Remark. This use of monotonicity is the *method of Browder and Minty*, a remarkable technique which somehow employs the inequality condition of monotonicity to justify passing to weak limits within a nonlinearity. □

Let us assume now the vector field $\mathbf{a}$ satisfies the condition of *strict monotonicity*; that is,

$$(\mathbf{a}(p) - \mathbf{a}(q)) \cdot (p - q) \geq \theta |p - q|^2 \tag{20}$$

for all $p, q \in \mathbb{R}^n$ and some constant $\theta > 0$.

THEOREM 4 (Uniqueness of weak solution). *Assume the strict monotonicity property* (20) *holds. Then there exists only one weak solution of* (1).

Proof. Assume that u and $\tilde{u}$ are two weak solutions. Consequently

$$\int_U \mathbf{a}(Du) \cdot Dv \, dx = \int_U \mathbf{a}(D\tilde{u}) \cdot Dv \, dx = \int_U fv \, dx,$$

and so

$$\int_U [\mathbf{a}(Du) - \mathbf{a}(D\tilde{u})] \cdot Dv \, dx = 0$$

for each $v \in H_0^1(U)$. We set $v := u - \tilde{u}$, and use (20) to deduce

$$\int_U |Du - D\tilde{u}|^2 \, dx = 0.$$

Thus $u = \tilde{u}$ a.e. in U. □

Remark. Under the strengthened monotonicity assumption (20) our weak solution u in fact belongs to $H^2(U)$, and so satisfies

$$-\operatorname{div} \mathbf{a}(Du) = f \quad \text{a.e. in } U.$$

To demonstrate this, we select $q, \xi \in \mathbb{R}^n$ and set $p = q + h\xi$, $h \neq 0$, in (20). We obtain, after dividing by h^2, the inequality

$$\sum_{i=1}^{n} \frac{[a^i(q+h\xi) - a^i(q)]}{h} \xi_i \geq \theta |\xi|^2.$$

Now send $h \to 0$:

$$\sum_{i,j=1}^{n} a^i_{p_j}(q)\xi_i\xi_j \geq \theta|\xi|^2 \quad (q, \xi \in \mathbb{R}^n). \tag{21}$$

We can thus interpret the nonlinear PDE $-\operatorname{div}\mathbf{a}(Du) = f$ as being uniformly elliptic. The proof of H^2 regularity of the weak solution now follows almost precisely as in the proof of Theorem 1 in §6.3.1. □

9.2. FIXED POINT METHODS

We study next the applicability of topological *fixed point theorems* to nonlinear partial differential equations. There are at least three distinct classes of such abstract theorems that are useful. These are:

(a) fixed point theorems for *strict contractions*,

(b) fixed point theorems for *compact mappings*,

and

(c) fixed point theorems for *order-preserving operators*.

We present below applications of types (a) and (b). The utility of order-preserving properties for nonlinear PDE will be explained later, in §9.3.

9.2.1. Banach's Fixed Point Theorem.

Hereafter X denotes a Banach space. The simplest fixed point theorem of all is:

THEOREM 1 (Banach's Fixed Point Theorem). *Assume*

$$A : X \to X$$

is a nonlinear mapping, and suppose that

$$\|A[u] - A[\tilde{u}]\| \leq \gamma \|u - \tilde{u}\| \quad (u, \tilde{u} \in X) \tag{1}$$

for some constant $\gamma < 1$. Then A has a unique fixed point.

DEFINITION. *We say that A is a* strict contraction *if* (1) *holds.*

Proof. Fix any point $u_0 \in X$ and thereafter iteratively define $u_{k+1} = A[u_k]$ for $k = 0, 1, \dots$. Then

$$\|A[u_{k+1}] - A[u_k]\| \leq \gamma \|u_{k+1} - u_k\| = \gamma \|A[u_k] - A[u_{k-1}]\|,$$

and so

$$\|A[u_{k+1}] - A[u_k]\| \leq \gamma^k \|A[u_0] - u_0\|$$

for $k = 1, 2, \dots$. Consequently if $k \geq l$,

$$\begin{aligned} \|u_k - u_l\| = \|A[u_{k-1}] - A[u_{l-1}]\| &\leq \sum_{j=l-1}^{k-2} \|A[u_{j+1}] - A[u_j]\| \\ &\leq \|A[u_0] - u_0\| \sum_{j=l-1}^{k-2} \gamma^j. \end{aligned}$$

Hence $\{u_k\}_{k=1}^{\infty}$ is a Cauchy sequence in X, and so there exists a point $u \in X$ with $u_k \to u$ in X. Clearly then $A[u] = u$. Hence u is a fixed point for A, and hypothesis (1) ensures uniqueness. □

Applications of Banach's fixed point theorem to nonlinear PDE usually involve perturbation arguments of various sorts: given a well-behaved linear elliptic partial differential equation, it is often straightforward to cast a small nonlinear modification as a contraction mapping. The hallmark of such proofs is the occurrence of a parameter which must be taken small enough to ensure the strict contraction property.

Sometimes however we can eliminate such a smallness hypothesis by an iteration, as now illustrated.

Example 1 (Reaction-diffusion equations). Let us investigate the solvability of the initial/boundary-value problem for the *reaction-diffusion system*

$$(2) \qquad \begin{cases} \mathbf{u}_t - \Delta \mathbf{u} = \mathbf{f}(\mathbf{u}) & \text{in } U_T \\ \mathbf{u} = \mathbf{0} & \text{on } \partial U \times [0, T] \\ \mathbf{u} = \mathbf{g} & \text{on } U \times \{t = 0\}. \end{cases}$$

Here $\mathbf{u} = (u^1, \dots, u^m)$, $\mathbf{g} = (g^1, \dots, g^m)$, and as usual $U_T = U \times (0, T]$, where $U \in \mathbb{R}^n$ is open and bounded, with smooth boundary. The time $T > 0$ is fixed. We assume that the initial function $\mathbf{g}$ belongs to $H_0^1(U; \mathbb{R}^m)$. Concerning the nonlinearity, let us suppose

$$(3) \qquad \mathbf{f} : \mathbb{R}^m \to \mathbb{R}^m \text{ is Lipschitz continuous.}$$

This hypothesis in particular implies

$$|\mathbf{f}(z)| \leq C(1+|z|) \tag{4}$$

for each $z \in \mathbb{R}^m$ and some constant C.

Adapting the terminology from §7.1, we say that a function

$$\mathbf{u} \in L^2(0,T;H_0^1(U;\mathbb{R}^m)), \text{ with } \mathbf{u}' \in L^2\ (0,T;H^{-1}(U;\mathbb{R}^m)), \tag{5}$$

is a *weak solution* of (2) provided

$$\langle \mathbf{u}',\mathbf{v}\rangle + B[\mathbf{u},\mathbf{v}] = (\mathbf{f}(\mathbf{u}),\mathbf{v}) \quad \text{a.e. } 0 \leq t \leq T \tag{6}$$

for each $\mathbf{v} \in H_0^1(U;\mathbb{R}^m)$, and

$$\mathbf{u}(0) = \mathbf{g}. \tag{7}$$

In (6) $\langle\ ,\ \rangle$ denotes the pairing of $H^{-1}(U;\mathbb{R}^m)$ and $H_0^1(U;\mathbb{R}^m)$, $B[\ ,\ ]$ is the bilinear form associated with $-\Delta$ in $H_0^1(U;\mathbb{R}^m)$, and $(\ ,\ )$ denotes the inner product in $L^2(U;\mathbb{R}^m)$. The norm in $H_0^1(U;\mathbb{R}^m)$ is taken to be

$$||\mathbf{u}||_{H_0^1(U;\mathbb{R}^m)} = \left(\int_U |D\mathbf{u}|^2\,dx\right)^{\frac{1}{2}}.$$

Recall from §5.9.2 that (5) implies $\mathbf{u} \in C([0,T];L^2(U;\mathbb{R}^m))$, after possible redefinition of $\mathbf{u}$ on a set of measure zero.

THEOREM 2 (Existence). *There exists a unique weak solution of* (2).

Proof. 1. We will apply Banach's theorem in the space

$$X = C([0,T];L^2(U;\mathbb{R}^m)),$$

with the norm

$$||\mathbf{v}|| = \max_{0\leq t\leq T} ||\mathbf{v}(t)||_{L^2(U;\mathbb{R}^m)}.$$

Let the operator A be defined as follows. Given a function $\mathbf{u} \in X$, set $\mathbf{h}(t) := \mathbf{f}(\mathbf{u}(t))$ $(0 \leq t \leq T)$. In light of the growth estimate (4), we see $\mathbf{h} \in L^2(0,T;L^2(U;\mathbb{R}^m))$. Consequently the theory set forth in §7.1 ensures that the linear parabolic PDE

$$\begin{cases} \mathbf{w}_t - \Delta\mathbf{w} = \mathbf{h} & \text{in } U_T \\ \mathbf{w} = \mathbf{0} & \text{on } \partial U \times [0,T] \\ \mathbf{w} = \mathbf{g} & \text{on } U \times \{t=0\} \end{cases} \tag{8}$$

has a unique weak solution

$$\mathbf{w} \in L^2(0,T;H_0^1(U;\mathbb{R}^m)), \text{ with } \mathbf{w}' \in L^2(0,T;H^{-1}(U;\mathbb{R}^m)). \tag{9}$$

Thus $\mathbf{w} \in X$ satisfies

$$\langle \mathbf{w}', \mathbf{v}\rangle + B[\mathbf{w}, \mathbf{v}] = (\mathbf{h}, \mathbf{v}) \quad \text{a.e. } 0 \le t \le T \tag{10}$$

for each $\mathbf{v} \in H_0^1(U;\mathbb{R}^m)$, and $\mathbf{w}(0) = \mathbf{g}$.

Define $A : X \to X$ by setting $A[\mathbf{u}] = \mathbf{w}$.

2. We now claim that

$$\begin{cases} \text{if } T > 0 \text{ is small enough, then} \\ A \text{ is a strict contraction.} \end{cases} \tag{11}$$

To prove this, choose $\mathbf{u}, \tilde{\mathbf{u}} \in X$, and define $\mathbf{w} = A[\mathbf{u}]$, $\tilde{\mathbf{w}} = A[\tilde{\mathbf{u}}]$ as above. Consequently $\mathbf{w}$ verifies (10) for $\mathbf{h} = \mathbf{f}(\mathbf{u})$, and $\tilde{\mathbf{w}}$ satisfies a similar identity for $\tilde{\mathbf{h}} := \mathbf{f}(\tilde{\mathbf{u}})$.

We calculate as in §7.1

$$\begin{aligned} &\frac{d}{dt}\|\mathbf{w} - \tilde{\mathbf{w}}\|^2_{L^2(U;\mathbb{R}^m)} + 2\|\mathbf{w} - \tilde{\mathbf{w}}\|^2_{H_0^1(U;\mathbb{R}^m)} \\ &\qquad = 2(\mathbf{w} - \mathbf{w}, \mathbf{h} - \tilde{\mathbf{h}}) \\ &\qquad \le \epsilon\|\mathbf{w} - \tilde{\mathbf{w}}\|^2_{L^2(U;\mathbb{R}^m)} + \frac{1}{\epsilon}\|\mathbf{f}(\mathbf{u}) - \mathbf{f}(\tilde{\mathbf{u}})\|^2_{L^2(U;\mathbb{R}^m)} \\ &\qquad \le \epsilon C\|\mathbf{w} - \tilde{\mathbf{w}}\|^2_{H_0^1(U;\mathbb{R}^m)} + \frac{1}{\epsilon}\|\mathbf{f}(\mathbf{u}) - \mathbf{f}(\tilde{\mathbf{u}})\|^2_{L^2(U;\mathbb{R}^m)}, \end{aligned} \tag{12}$$

by Poincaré's inequality. Selecting $\epsilon > 0$ sufficiently small, we deduce

$$\frac{d}{dt}\|\mathbf{w} - \tilde{\mathbf{w}}\|^2_{L^2(U;\mathbb{R}^m)} \le C\|\mathbf{f}(\mathbf{u}) - \mathbf{f}(\tilde{\mathbf{u}})\|^2_{L^2(U;\mathbb{R}^m)} \le C\|\mathbf{u} - \tilde{\mathbf{u}}\|^2_{L^2(U;\mathbb{R}^m)},$$

since $\mathbf{f}$ is Lipschitz. Consequently

$$\begin{aligned} \|\mathbf{w}(s) - \tilde{\mathbf{w}}(s)\|^2_{L^2(U;\mathbb{R}^m)} &\le C\int_0^s \|\mathbf{u}(t) - \tilde{\mathbf{u}}(t)\|^2_{L^2(U;\mathbb{R}^m)}\,dt \\ &\le CT\|\mathbf{u} - \tilde{\mathbf{u}}\|^2 \end{aligned} \tag{13}$$

for each $0 \le s \le T$. Maximizing the left hand side with respect to s, we discover

$$\|\mathbf{w} - \tilde{\mathbf{w}}\|^2 \le CT\|\mathbf{u} - \tilde{\mathbf{u}}\|^2.$$

Hence

$$\|A[\mathbf{u}] - A[\tilde{\mathbf{u}}]\| \le (CT)^{1/2}\|\mathbf{u} - \tilde{\mathbf{u}}\|, \tag{14}$$

and thus A is a strict contraction, provided $T > 0$ is so small that $(CT)^{1/2} = \gamma < 1$.

3. Given any $T > 0$ we select $T_1 > 0$ so small that $(CT_1)^{1/2} < 1$. We can then apply Banach's fixed point theorem to find a weak solution $\mathbf{u}$ of the problem (2) existing on the time interval $[0, T_1]$. Since $\mathbf{u}(t) \in H_0^1(U; \mathbb{R}^m)$ for a.e. $0 \leq t \leq T_1$, we can upon redefining T_1 if necessary assume $\mathbf{u}(T_1) \in H_0^1(U; \mathbb{R}^m)$. We can then repeat the argument above to extend our solution to the time interval $[T_1, 2T_1]$. Continuing, after finitely many steps we construct a weak solution existing on the full interval $[0, T]$.

4. To demonstrate uniqueness, suppose both $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are two weak solutions of (2). Then we have $\mathbf{w} = \mathbf{u}$, $\tilde{\mathbf{w}} = \tilde{\mathbf{u}}$ in inequality (13); whence

$$\|\mathbf{u}(s) - \tilde{\mathbf{u}}(s)\|^2_{L^2(U;\mathbb{R}^m)} \leq C \int_0^s \|\mathbf{u}(t) - \tilde{\mathbf{u}}(t)\|^2_{L^2(U;\mathbb{R}^m)}\, dt$$

for $0 \leq s \leq T$. According to Gronwall's inequality, $\mathbf{u} \equiv \tilde{\mathbf{u}}$. □

Remark. In common applications problem (2) records the evolution of the densities $u^1, \dots, u^m$ of various chemicals, which both diffuse within a medium and interact with each other. The diffusion term is $\Delta \mathbf{u}$ (or more generally $(a_1 \Delta u^1, \dots, a_m \Delta u^m)$ where the constants $a_k > 0$ characterize the diffusion of the k^{th} chemical). The reaction term $\mathbf{f}(\mathbf{u})$ models the chemistry. In the foregoing example we made the unreasonable assumption that $\mathbf{f}$ is globally Lipschitz. In more realistic models $\mathbf{f}$ is often a polynomial in $\mathbf{u}$ and there are interesting problems as to the global existence or blow-up of a solution. (A simple such problem is treated in §9.4.1.) □

9.2.2. Schauder's, Schaefer's Fixed Point Theorems.

Next we extend Brouwer's fixed point theorem (§8.1.4) to Banach spaces. The key assumption is now compactness. Throughout this section X continues to denote a real Banach space.

THEOREM 3 (Schauder's Fixed Point Theorem). *Suppose $K \subset X$ is compact and convex, and assume also*

$$A : K \to K$$

is continuous. Then A has a fixed point in K.

Proof. 1. Fix $\epsilon > 0$ and choose finitely many points $u_1, \dots, u_{N_\epsilon} \in K$, so that the open balls $\{B^0(u_i, \epsilon)\}_{i=1}^{N_\epsilon}$ cover K:

$$K \subset \bigcup_{i=1}^{N_\epsilon} B^0(u_i, \epsilon). \tag{15}$$

This is possible since K is compact. Let K_ϵ denote the closed convex hull of the points $\{u_1, \dots, u_{N_\epsilon}\}$:

$$K_\epsilon := \left\{ \sum_{i=1}^{N_\epsilon} \lambda_i u_i \mid 0 \leq \lambda_i \leq 1, \sum_{i=1}^{N_\epsilon} \lambda_i = 1 \right\}.$$

Then $K_\epsilon \subset K$, since K is convex. Now define $P_\epsilon : K \to K_\epsilon$ by writing

$$P_\epsilon[u] := \frac{\sum_{i=1}^{N_\epsilon} \operatorname{dist}(u, K - B^0(u_i, \epsilon)) u_i}{\sum_{i=1}^{N_\epsilon} \operatorname{dist}(u, K - B^0(u_i, \epsilon))} \quad (u \in K).$$

The denominator is never zero, because of (15). Now clearly P_ϵ is continuous, and furthermore for each $u \in K$, we have

$$\|P_\epsilon[u] - u\| \leq \frac{\sum_{i=1}^{N_\epsilon} \operatorname{dist}(u, K - B^0(u_i, \epsilon)) \|u_i - u\|}{\sum_{i=1}^{N_\epsilon} \operatorname{dist}(u, K - B^0(u_i, \epsilon))} \leq \epsilon. \tag{16}$$

2. Consider next the operator $A_\epsilon : K_\epsilon \to K_\epsilon$ defined by

$$A_\epsilon[u] := P_\epsilon[A[u]\,] \quad (u \in K_\epsilon).$$

Now K_ϵ is homeomorphic to the closed unit ball in $\mathbb{R}^{M_\epsilon}$ for some $M_\epsilon \leq N_\epsilon$. Brouwer's fixed point theorem (§8.1.4) therefore ensures the existence of a point $u_\epsilon \in K_\epsilon$ with

$$A_\epsilon[u_\epsilon] = u_\epsilon.$$

3. As K is compact, there exists a subsequence $\epsilon_j \to 0$ and a point $u \in K$, such that $u_{\epsilon_j} \to u$ in X. We claim u is a fixed point of A. Indeed, using estimate (16) we deduce

$$\|u_{\epsilon_j} - A[u_{\epsilon_j}]\| = \|A_{\epsilon_j}[u_{\epsilon_j}] - A[u_{\epsilon_j}]\| = \|P_{\epsilon_j}[A[u_{\epsilon_j}]\,] - A[u_{\epsilon_j}]\| \leq \epsilon_j.$$

Consequently, since A is continuous, we conclude $u = A[u]$. □

We next transform Schauder's fixed point theorem into an alternative form which is often more useful for applications to nonlinear partial differential equations.

DEFINITION. *A nonlinear mapping $A : X \to X$ is called* compact *provided for each bounded sequence $\{u_k\}_{k=1}^\infty$ the sequence $\{A[u_k]\}_{k=1}^\infty$ is precompact; that is, there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty$ such that $\{A[u_{k_j}]\}_{j=1}^\infty$ converges in X.*

THEOREM 4 (Schaefer's Fixed Point Theorem). *Suppose*

$$A : X \to X$$

is a continuous and compact mapping. Assume further that the set

$$\{u \in X \mid u = \lambda A[u] \text{ for some } 0 \leq \lambda \leq 1\}$$

is bounded. Then A has a fixed point.

Remark. The assertion is that if we have a bound on any possible fixed points of any of the operators λA for $0 \leq \lambda \leq 1$, then we have the existence of a fixed point for A. This is in accordance with the remarkable informal principle that "if we can prove appropriate estimates for solutions of a nonlinear PDE, under the assumption that such solutions exist, then in fact these solutions do exist". This is the method of *a priori** estimates. □

The advantage of Schaefer's theorem over Schauder's for applications is that we do not have to identify an explicit convex, compact set.

Proof. 1. Choose a constant M so large that

$$\|u\| < M \quad \text{if } u = \lambda A[u] \text{ for some } 0 \leq \lambda \leq 1. \tag{17}$$

Define then

$$\tilde{A}[u] := \begin{cases} A[u] & \text{if } \|A[u]\| \leq M \\ \frac{MA[u]}{\|A[u]\|} & \text{if } \|A[u]\| \geq M. \end{cases} \tag{18}$$

Observe $\tilde{A} : B(0,M) \to B(0,M)$. Now set K = closed convex hull of $\tilde{A}(B(0,M))$. Then since A and thus $\tilde{A}$ are compact mappings, K is a compact, convex subset of X. Furthermore $\tilde{A} : K \to K$.

2. Invoking Schauder's fixed point theorem, we infer the existence of a point $u \in K$ with

$$\tilde{A}[u] = u. \tag{19}$$

We now claim additionally that u is a fixed point of A. For if not, then according to (18) and (19) we would have

$$\|A[u]\| > M$$

* *a priori* = from before (Latin).

and

$$u = \lambda A[u] \quad \text{for } \lambda = \frac{M}{\|A[u]\|} < 1. \tag{20}$$

But $\|u\| = \|\tilde{A}[u]\| = M$, a contradiction to (17) and (20). □

Applications of Schauder's and Schaefer's fixed point theorems to PDE depend upon quite different considerations than applications of Banach's theorem. The crucial assumption is now not that some parameter be small, but rather that we have some sort of compactness. As the inverses of linear elliptic operators are typically smoothing, compactness is indeed available for certain nonlinear elliptic equations. Following is a quick, albeit fairly crude, example:

Example 2 (A quasilinear elliptic PDE). We present now a simple application of Schaefer's theorem by solving the semilinear boundary-value problem

$$\begin{cases} -\Delta u + b(Du) + \mu u = 0 & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases} \tag{21}$$

where U is bounded and ∂U is smooth. We assume $b : \mathbb{R}^n \to \mathbb{R}$ is smooth, Lipschitz continuous, and thus satisfies the growth condition

$$|b(p)| \leq C(|p| + 1) \tag{22}$$

for some constant C and all $p \in \mathbb{R}^n$.

THEOREM 5 (Existence). *If $\mu > 0$ is sufficiently large, there exists a function $u \in H^2(U) \cap H_0^1(U)$ solving the boundary-value problem* (21).

Proof. 1. Given $u \in H_0^1(U)$, set

$$f := -b(Du). \tag{23}$$

Owing to estimate (22), we see that $f \in L^2(U)$. Now let $w \in H_0^1(U)$ be the unique weak solution of the linear problem

$$\begin{cases} -\Delta w + \mu w = f & \text{in } U \\ w = 0 & \text{on } \partial U. \end{cases} \tag{24}$$

By the regularity theory proved in §6.3, we know additionally that $w \in H^2(U)$, with the estimate

$$\|w\|_{H^2(U)} \leq C\|f\|_{L^2(U)} \tag{25}$$

for some constant C.

Let us henceforth write $A[u] = w$ whenever w is derived from u via (23), (24). In light of (22) and (25), we have the estimate

$$\|A[u]\|_{H^2(U)} \leq C(\|u\|_{H_0^1(U)} + 1). \tag{26}$$

2. We now assert that $A : H_0^1(U) \to H_0^1(U)$ is continuous and compact. Indeed, if

$$u_k \to u \quad \text{in } H_0^1(U), \tag{27}$$

then in view of estimate (26) we have

$$\sup_k \|w_k\|_{H^2(U)} < \infty, \tag{28}$$

for $w_k = A[u_k]$ $(k = 1, \dots)$. Thus there is a subsequence $\{w_{k_j}\}_{j=1}^{\infty}$ and a function $w \in H_0^1(U)$ with

$$w_{k_j} \to w \quad \text{in } H_0^1(U). \tag{29}$$

Now

$$\int_U Dw_{k_j} \cdot Dv + \mu w_{k_j} v \, dx = -\int_U b(Du_{k_j}) v \, dx$$

for each $v \in H_0^1(U)$. Consequently using (22), (27) and (29), we see

$$\int_U Dw \cdot Dv + \mu w v \, dx = -\int_U b(Du) v \, dx$$

for each $v \in H_0^1(U)$. Thus $w = A[u]$.

Hence (27) implies $A[u_k] \to A[u]$ in $H_0^1(U)$, and so A is continuous. A similar argument shows that A is compact, since if $\{u_k\}_{k=1}^{\infty}$ is bounded in $H_0^1(U)$, estimate (22) asserts that $\{A[u_k]\}_{k=1}^{\infty}$ is bounded in $H^2(U)$, and so possesses a strongly convergent subsequence in $H_0^1(U)$.

3. Finally, we must show that if μ is large enough, the set

$$\{u \in H_0^1(U) \mid u = \lambda A[u] \text{ for some } 0 \leq \lambda \leq 1\}$$

is bounded in $H_0^1(U)$. So assume $u \in H_0^1(U)$,

$$u = \lambda A[u] \quad \text{for some } 0 < \lambda \leq 1.$$

Then $\frac{u}{\lambda} = A[u]$; or, in other words, $u \in H^2(U) \cap H_0^1(U)$ and

$$-\Delta u + \mu u = \lambda b(Du) \quad \text{a.e. in } U.$$

Multiply this identity by u and integrate over U, to compute

$$\begin{aligned}\int_U |Du|^2 + \mu|u|^2 dx &= -\int_U \lambda b(Du)u\,dx \le \int_U C(|Du|+1)|u|\,dx \\ &\le \frac{1}{2}\int_U |Du|^2 dx + C\int_U |u|^2 + 1\,dx.\end{aligned}$$

Thus if $\mu > 0$ is sufficiently large, we have $\|u\|_{H_0^1(U)} \le C$, for some constant C that does not depend on $0 \le \lambda \le 1$.

4. Applying Schaefer's fixed point theorem in the space $X = H_0^1(U)$, we conclude that A has a fixed point $u \in H_0^1(U) \cap H^2(U)$, which in turn solves our semilinear PDE (21). □

Remark and warning. A plausible plan for constructively solving (21) would be to select some u^0 and then iteratively solve the linear boundary-value problems

$$\begin{cases} -\Delta u^{k+1} + \mu u^{k+1} = -b(Du^k) & \text{in } U \\ \qquad u^{k+1} = 0 & \text{on } \partial U \end{cases} \quad (k = 0, 1, \dots).$$

However, we *cannot* assert that $\{u^k\}_{k=0}^\infty$ then converges to a solution of (21). Schauder's and Schaefer's fixed point theorems do not say that any sequence converges to a fixed point. (But see the proof in §9.3 following.) □

See Gilbarg-Trudinger [**G-T**] for much more sophisticated applications of fixed point theorems to nonlinear elliptic PDE.

9.3. METHOD OF SUBSOLUTIONS AND SUPERSOLUTIONS

Our application of Schaefer's theorem above in §9.2.2 depends upon the regularity estimates for solutions of elliptic equations. We turn attention now to another basic property of elliptic PDE, namely the maximum principle, and demonstrate how various resulting comparison arguments can be used to solve certain semilinear problems. The idea is to exploit *ordering properties* for solutions. More precisely, we will show that if we can find a subsolution $\underline{u}$ and a supersolution $\bar{u}$ of a particular boundary-value problem, and if furthermore $\underline{u} \le \bar{u}$, then there in fact exists a solution satisfying

$$\underline{u} \le u \le \bar{u}.$$

We will investigate this boundary-value problem for the nonlinear Poisson equation:

$$\begin{cases} -\Delta u = f(u) & \text{in } U \\ \quad u = 0 & \text{on } \partial U, \end{cases} \tag{1}$$

where $f : \mathbb{R} \to \mathbb{R}$ is smooth, with

$$|f'| \leq C \quad (z \in \mathbb{R}) \tag{2}$$

for some constant C.

DEFINITIONS. (i) *We say that* $\bar{u} \in H^1(U)$ *is a* weak supersolution *of problem* (1) *if*

$$\int_U D\bar{u} \cdot Dv \, dx \geq \int_U f(\bar{u})v \, dx \tag{3}$$

for each $v \in H_0^1(U)$, $v \geq 0$ *a.e.*

(ii) *Similarly,* $\underline{u} \in H^1(U)$ *is a* weak subsolution *provided*

$$\int_U D\underline{u} \cdot Dv \, dx \leq \int_U f(\underline{u})v \, dx \tag{4}$$

for each $v \in H_0^1(U)$, $v \geq 0$ *a.e.*

(iii) *We say* $u \in H_0^1(U)$ *is a* weak solution *of* (1) *if*

$$\int_U Du \cdot Dv \, dx = \int_U f(u)v \, dx$$

for each $v \in H_0^1(U)$.

Remark. If $\bar{u}, \underline{u} \in C^2(U)$, then from (3) and (4) it follows that

$$-\Delta\bar{u} \geq f(\bar{u}), \; -\Delta\underline{u} \leq f(\underline{u}) \quad \text{in } U.$$

□

THEOREM 1 (Existence of a solution between sub- and supersolutions). *Assume there exist a weak supersolution* $\bar{u}$ *and a weak subsolution* $\underline{u}$ *of* (1), *satisfying*

$$\underline{u} \leq 0, \;\; \bar{u} \geq 0 \text{ on } \partial U \text{ in the trace sense}, \;\; \underline{u} \leq \bar{u} \text{ a.e. in } U. \tag{5}$$

Then there exists a weak solution u *of* (1), *such that*

$$\underline{u} \leq u \leq \bar{u} \quad \text{a.e. in } U.$$

Proof. 1. Fix a number $\lambda > 0$ so large that

$$\text{the mapping } z \mapsto f(z) + \lambda z \text{ is nondecreasing;} \tag{6}$$

this is possible as a consequence of hypothesis (2).

Now write $u_0 = \underline{u}$, and then given u_k $(k = 0, 1, 2, \dots)$ inductively define $u_{k+1} \in H_0^1(U)$ to be the unique weak solution of the linear boundary-value problem

$$\begin{cases} -\Delta u_{k+1} + \lambda u_{k+1} = f(u_k) + \lambda u_k & \text{in } U \\ u_{k+1} = 0 & \text{on } \partial U. \end{cases} \tag{7}$$

2. We claim

$$\underline{u} = u_0 \le u_1 \le \cdots \le u_k \le \cdots \quad \text{a.e. in } U. \tag{8}$$

To confirm this, first note from (7) for $k = 0$ that

$$\int_U Du_1 \cdot Dv + \lambda u_1 v \, dx = \int_U (f(u_0) + \lambda u_0) v \, dx \tag{9}$$

for each $v \in H_0^1(U)$. Subtract (9) from (4), recall $u_0 = \underline{u}$, and set

$$v := (u_0 - u_1)^+ \in H_0^1(U), \ v \ge 0 \ \text{ a.e.}$$

We find

$$\int_U D(u_0 - u_1) \cdot D(u_0 - u_1)^+ + \lambda (u_0 - u_1)(u_0 - u_1)^+ \, dx \le 0. \tag{10}$$

But

$$D(u_0 - u_1)^+ = \begin{cases} D(u_0 - u_1) & \text{a.e. on } \{u_0 \ge u_1\} \\ 0 & \text{a.e. on } \{u_0 \le u_1\}. \end{cases}$$

(See Problem 17 in Chapter 5.) Consequently,

$$\int_{\{u_0 \ge u_1\}} |D(u_0 - u_1)|^2 + \lambda (u_0 - u_1)^2 \, dx \le 0;$$

so that $u_0 \le u_1$ a.e. in U.

Now assume inductively

$$u_{k-1} \le u_k \quad \text{a.e. in } U. \tag{11}$$

From (7) we find

$$\int_U Du_{k+1} \cdot Dv + \lambda u_{k+1} v \, dx = \int_U (f(u_k) + \lambda u_k) v \, dx \tag{12}$$

and

$$\int_U Du_k \cdot Dv + \lambda u_k v \, dx = \int_U (f(u_{k-1}) + \lambda u_{k-1}) v \, dx$$

for each $v \in H_0^1(U)$. Subtract and set $v := (u_k - u_{k+1})^+$. We deduce

$$\begin{aligned} &\int_{\{u_k \geq u_{k+1}\}} |D(u_k - u_{k+1})|^2 + \lambda (u_k - u_{k+1})^2 \, dx \\ &= \int_U [(f(u_{k-1}) + \lambda u_{k-1}) - (f(u_k) + \lambda u_k)](u_k - u_{k+1})^+ \, dx \leq 0, \end{aligned}$$

the last inequality holding in view of (11) and (6). Therefore $u_k \leq u_{k+1}$ a.e. in U, as asserted.

3. Next we show

(13) $$u_k \leq \bar{u} \quad \text{a.e. in } U \ (k = 0, 1, \dots).$$

Statement (13) is valid for $k = 0$ by hypothesis (5). Assume now for induction

(14) $$u_k \leq \bar{u} \quad \text{a.e. in } U.$$

Then subtracting (3) from (12) and taking $v := (u_{k+1} - \bar{u})^+$, we find

$$\begin{aligned} &\int_{\{u_{k+1} \geq \bar{u}\}} |D(u_{k+1} - \bar{u})|^2 + \lambda (u_{k+1} - \bar{u})^2 dx \\ &\leq \int_U [(f(u_k) + \lambda u_k) - (f(\bar{u}) + \lambda \bar{u})](u_{k+1} - \bar{u})^+ \, dx \leq 0, \end{aligned}$$

by (14) and (6). Thus $u_{k+1} \leq \bar{u}$ a.e. in U.

4. In light of (8) and (13), we have

(15) $$\underline{u} \leq \cdots \leq u_k \leq u_{k+1} \leq \cdots \leq \bar{u} \quad \text{a.e. in } U.$$

Therefore

$$u(x) := \lim_{k \to \infty} u_k(x)$$

exists for a.e. x. Furthermore we have

(16) $$u_k \to u \quad \text{in } L^2(U),$$

as guaranteed by the Dominated Convergence Theorem and (15). Finally, since we have $\|f(u_k)\|_{L^2(U)} \leq C(\|u_k\|_{L^2(U)} + 1)$, we deduce from (7) that

$\sup_k \|u_k\|_{H_0^1(U)} < \infty$. Hence there is a subsequence $\{u_{k_j}\}_{j=1}^\infty$ which converges weakly in $H_0^1(U)$ to $u \in H_0^1(U)$.

5. We at last verify that u is a weak solution of problem (1). For this, fix $v \in H_0^1(U)$. Then from (7) we find

$$\int_U Du_{k+1} \cdot Dv + \lambda u_{k+1} v \, dx = \int_U (f(u_k) + \lambda u_k) v \, dx.$$

Let $k \to \infty$:

$$\int_U Du \cdot Dv + \lambda uv \, dx = \int_U (f(u) + \lambda u) v \, dx.$$

Canceling the term involving λ, we at last confirm that

$$\int_U Du \cdot Dv \, dx = \int_U f(u) v \, dx,$$

as desired. □

This proof illustrates the use of integration-by-parts, rather than the maximum principle, to establish comparisons between sub- and supersolutions.

9.4. NONEXISTENCE

We now complement the theory in §§9.1–9.3 with some *nonexistence* assertions for solutions of various nonlinear partial differential equations. The overall procedure will be to assume there exists a solution, and then to obtain certain inequalities, which in turn force a contradiction.

9.4.1. Blow-up.

We begin by considering an initial/boundary-value problem for a parabolic equation with a simple quadratic nonlinearity:

$$\text{(1)} \qquad \begin{cases} u_t - \Delta u = u^2 & \text{in } U_T \\ u = 0 & \text{on } \partial U \times (0, T) \\ u = g & \text{on } U \times \{t = 0\}. \end{cases}$$

We will show that if $T > 0$ and $g \geq 0$ are large enough in an appropriate sense, then there does *not* exist a smooth solution u of (1). We can regard the nonlinear heat equation in (1) as a simple reaction-diffusion equation (cf. Example 1 in §9.2.1). The nonlinear term alone corresponds to the ODE

$$\text{(2)} \qquad \dot{u} = u^2 \quad \left(\cdot = \frac{d}{dt} \right),$$

which certainly blows up in finite time, provided $u(0) > 0$. The purely diffusive effects on the other hand yield the heat equation, which tends to smooth out irregularities. The following analysis must therefore untangle the competing effects of blow-up from the u^2 term and smoothing from the Δu term.

We proceed by choosing w_1 to be an eigenfunction corresponding to the principal eigenvalue $\lambda_1 > 0$ of $-\Delta$ in $H_0^1(U)$. Then owing to the theory in §6.5.1, w_1 is smooth,

$$\begin{cases} -\Delta w_1 = \lambda_1 w & \text{in } U \\ w_1 = 0 & \text{on } \partial U, \end{cases}$$

and we may furthermore assume

(3) $$w_1 > 0 \text{ in } U, \int_U w_1 \, dx = 1.$$

Suppose u is a smooth solution of (1), with $g \geq 0, g \not\equiv 0$; so that $u > 0$ within U_T by the strong maximum principle. Define

$$\eta(t) := \int_U u(x,t) w_1(x) \, dx \quad (0 \leq t \leq T).$$

Then

(4) $$\begin{aligned} \dot{\eta}(t) &= \int_U u_t w_1 \, dx = \int_U (\Delta u + u^2) w_1 \, dx \\ &= \int_U u \Delta w_1 + u^2 w_1 \, dx = -\lambda_1 \eta(t) + \int_U u^2 w_1 \, dx. \end{aligned}$$

Furthermore

(5) $$\begin{aligned} \eta(t) &= \int_U u w_1 \, dx = \int_U u w_1^{1/2} w_1^{1/2} \, dx \\ &\leq \left(\int_U u^2 w_1 \, dx \right)^{1/2} \left(\int_U w_1 \, dx \right)^{1/2} \\ &= \left(\int_U u^2 w_1 \, dx \right)^{1/2} \quad \text{by (3).} \end{aligned}$$

Consequently

$$\eta(t)^2 \leq \int_U u^2 w_1 \, dx.$$

Employing this inequality in (4), we find

(6) $$\dot{\eta}(t) \geq -\lambda_1 \eta(t) + \eta(t)^2 \quad (0 \leq t \leq T).$$

Writing $\xi(t) := e^{\lambda_1 t}\eta(t)$ gives

$$\dot{\xi}(t) = e^{\lambda_1 t}\dot{\eta}(t) + \lambda_1 e^{\lambda_1 t}\eta(t) \geq e^{\lambda_1 t}\eta(t)^2 = e^{-\lambda_1 t}\xi(t)^2$$

for $0 \leq t \leq T$. Thus

$$\left(\frac{-1}{\xi(t)}\right)^{\cdot} = \frac{\dot{\xi}(t)}{\xi(t)^2} \geq e^{-\lambda_1 t};$$

so that integrating we find

$$\frac{-1}{\xi(t)} \geq \frac{-1}{\xi(0)} + \frac{1 - e^{-\lambda_1 t}}{\lambda_1}.$$

Rearranging, we deduce

$$\xi(t) \geq \frac{\xi(0)\lambda_1}{\lambda_1 - \xi(0)(1 - e^{-\lambda_1 t})}, \tag{7}$$

provided the denominator is not zero.

But now suppose

$$\eta(0) = \xi(0) > \lambda_1. \tag{8}$$

Then

$$\xi(t) \to +\infty \quad \text{as } t \to t^*, \tag{9}$$

where

$$t^* := \frac{-1}{\lambda_1} \log\left(\frac{\eta(0) - \lambda_1}{\eta(0)}\right). \tag{10}$$

The conclusion is that *if*

$$\eta(0) = \int_U g w_1 \, dx > \lambda_1,$$

then there cannot exist a smooth solution u of (1). Furthermore, either the solution is not smooth enough to justify the calculation above, or else

$$\lim_{t \to t_*} \int_U u(x,t) w_1(x) \, dx = \infty$$

for some $0 < t_* \leq t^*$. In this case we say u "blows up" at time t_*.

9.4.2. Derrick–Pohozaev identity.

We investigate next a nonlinear elliptic PDE to which different differential inequality methods apply, namely the nonlinear boundary-value problem

$$(11)\qquad \begin{cases} -\Delta u = |u|^{p-1}u & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Now the theory in §8.5.2 applies to (11) provided

$$(12)\qquad 1 < p < \frac{n+2}{n-2},$$

and proves the existence of a nontrivial solution $u \not\equiv 0$. Let us now instead suppose

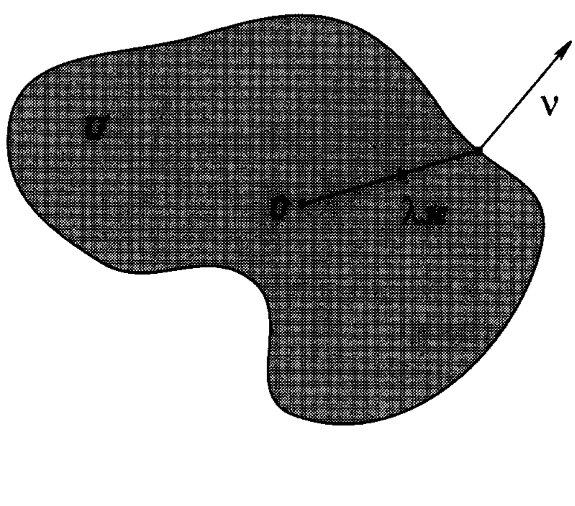
$$(13)\qquad \frac{n+2}{n-2} < p < \infty.$$

Our goal is to demonstrate under a certain geometric condition on U that (13) implies $u \equiv 0$ is the only smooth solution of (11). We see therefore that the restriction to condition (12) in §8.5.2 was in some sense natural, and consequently say $p = \frac{n+2}{n-2}$ is a *critical exponent.*

DEFINITION. *An open set U is called* star-shaped *with respect to* 0 *provided for each $x \in \bar{U}$, the line segment*

$$\{\lambda x \mid 0 \leq \lambda \leq 1\}$$

lies in $\bar{U}$.



A star-shaped domain

Clearly if U is convex and $0 \in U$, then U is star-shaped with respect to 0. But a general star-shaped region need not be convex.

LEMMA (Normals to a star-shaped region). *Assume ∂U is C^1 and U is star-shaped with respect to 0. Then*

$$x \cdot \boldsymbol{\nu}(x) \geq 0 \quad \textit{for all } x \in \partial U,$$

where $\boldsymbol{\nu}$ denotes the unit outward normal.

Proof. 1. Since ∂U is C^1, if $x \in \partial U$ then for each $\epsilon > 0$ there exists $\delta > 0$ such that $|y - x| < \delta$ and $y \in \bar{U}$ imply $\boldsymbol{\nu}(x) \cdot \frac{(y-x)}{|y-x|} \leq \epsilon$. In particular

$$\limsup_{\substack{y \to x \\ y \in \bar{U}}} \boldsymbol{\nu}(x) \cdot \frac{(y-x)}{|y-x|} \leq 0.$$

Let $y = \lambda x$ for $0 < \lambda < 1$. Then $y \in \bar{U}$, since U is star-shaped. Thus

$$\boldsymbol{\nu}(x) \cdot \frac{x}{|x|} = - \lim_{\lambda \to 1^-} \boldsymbol{\nu}(x) \cdot \frac{(\lambda x - x)}{|\lambda x - x|} \geq 0.$$

□

We next prove that there can exist no nontrivial solution to problem (11) for supercritical growth, provided U is star-shaped. The proof is a remarkable calculation initiated by multiplying the PDE $-\Delta u = |u|^{p-1}u$ by $x \cdot Du$ and continually integrating by parts.

THEOREM 1 (Nonexistence of nontrivial solution). *Assume $u \in C^2(\bar{U})$ is a solution of problem* (11), *and the exponent p satisfies inequality* (13). *Suppose further U is star-shaped with respect to 0, and ∂U is C^1. Then*

$$u \equiv 0 \quad \textit{within } U.$$

Proof. 1. We multiply the PDE by $x \cdot Du$ and integrate over U, to find

$$\int_U (-\Delta u)(x \cdot Du)\,dx = \int_U |u|^{p-1}u(x \cdot Du)\,dx. \tag{14}$$

We rewrite this expression as

$$A = B.$$

2. The term on the left is

$$
\begin{aligned}
A &:= -\sum_{i,j=1}^{n} \int_U u_{x_i x_i} x_j u_{x_j}\, dx \\
&= \sum_{i,j=1}^{n} \int_U u_{x_i} (x_j u_{x_j})_{x_i}\, dx - \sum_{i,j=1}^{n} \int_{\partial U} u_{x_i} \nu^i x_j u_{x_j}\, dS \qquad (15) \\
&=: A_1 + A_2.
\end{aligned}
$$

3. Now

$$
\begin{aligned}
A_1 &= \sum_{i,j=1}^{n} \int_U u_{x_i} \delta_{ij} u_{x_j} + u_{x_i} x_j u_{x_j x_i}\, dx \\
&= \int_U |Du|^2 + \sum_{j=1}^{n} \left(\frac{|Du|^2}{2} \right)_{x_j} x_j\, dx \qquad (16) \\
&= \left(1 - \frac{n}{2}\right) \int_U |Du|^2 dx + \int_{\partial U} \frac{|Du|^2}{2} (\boldsymbol{\nu} \cdot x)\, dS.
\end{aligned}
$$

On the other hand, since $u = 0$ on ∂U, $Du(x)$ is parallel to the normal $\boldsymbol{\nu}(x)$ at each point $x \in \partial U$. Thus $Du(x) = \pm |Du(x)| \boldsymbol{\nu}(x)$. Using this equality we calculate

$$
A_2 = -\int_{\partial U} |Du|^2 (\boldsymbol{\nu} \cdot x)\, dS. \qquad (17)
$$

Combine (15)–(17), to deduce

$$
A = \frac{2-n}{2} \int_U |Du|^2 dx - \frac{1}{2} \int_{\partial U} |Du|^2 (\boldsymbol{\nu} \cdot x)\, dS.
$$

4. Returning to (14), we compute

$$
\begin{aligned}
B &:= \sum_{j=1}^{n} \int_U |u|^{p-1} u x_j u_{x_j}\, dx \\
&= \sum_{j=1}^{n} \int_U \left(\frac{|u|^{p+1}}{p+1} \right)_{x_j} x_j\, dx = -\frac{n}{p+1} \int_U |u|^{p+1} dx.
\end{aligned}
$$

5. This calculation and (14) yield

$$
\left(\frac{n-2}{2} \right) \int_U |Du|^2 dx + \frac{1}{2} \int_{\partial U} |Du|^2 (\boldsymbol{\nu} \cdot x)\, dS = \frac{n}{p+1} \int_U |u|^{p+1} dx. \qquad (18)
$$

In view of the lemma above, we then obtain the inequality

$$\left(\frac{n-2}{2}\right)\int_U |Du|^2\,dx \le \frac{n}{p+1}\int_U |u|^{p+1}\,dx. \tag{19}$$

But once we multiply the PDE $-\Delta u = |u|^{p-1}u$ by u and integrate by parts, we produce the equality

$$\int_U |Du|^2 dx = \int_U |u|^{p+1} dx.$$

Substituting into (19), we thus conclude

$$\left(\frac{n-2}{2} - \frac{n}{p+1}\right)\int_U |u|^{p+1} dx \le 0.$$

Hence if $u \not\equiv 0$, it follows that $\frac{n-2}{2} - \frac{n}{p+1} \le 0$; that is, $p \le \frac{n+2}{n-2}$. □

Remark. Equality (18) is sometimes called the *Derrick–Pohozaev identity*. □

9.5. GEOMETRIC PROPERTIES OF SOLUTIONS

9.5.1. Star-shaped level sets.

We explain in this section a simple method that is occasionally useful for studying the geometric properties of the level sets of solutions to various PDE. The easiest such case occurs when we look at harmonic functions in an open set U having the form

$$U = W - \bar{V},$$

where $V \subset\subset W$, for open sets V, W, each of which is star-shaped with respect to 0. Write

$$\Gamma_0 = \partial W, \quad \Gamma_1 = \partial V.$$

We consider the problem

$$\begin{cases} -\Delta u = 0 & \text{in } U \\ u = 1 & \text{on } \Gamma_1 \\ u = 0 & \text{on } \Gamma_0. \end{cases} \tag{1}$$

Physically u corresponds to the electrostatic potential generated within the region U, once we fix the potential value to be one on Γ_1 and zero on Γ_0. According to the strong maximum principle $0 < u < 1$ within U.

THEOREM 1 (Star-shaped level sets). *For each* $0 < \lambda < 1$ *the level set*

$$\Gamma_\lambda := \{x \in U \mid u(x) = \lambda\}$$

is a smooth surface and is the boundary of a set star-shaped with respect to 0.

Proof. 1. For each $\mu > 0$, the function $x \mapsto u(\mu x)$ is harmonic, and thus so is

$$v(x) := \frac{d}{d\mu}(u(\mu x))|_{\mu=1} = Du(x) \cdot x \quad (x \in U).$$

Now since $u = 0$ on Γ_0, $Du(x)$ points in the direction of $-\boldsymbol{\nu}(x)$ at each point $x \in \partial W$. Additionally, we have $x \cdot \boldsymbol{\nu}(x) \geq 0$ on Γ_0, since W is star-shaped with respect to 0. Consequently $v = Du \cdot x \leq 0$ on Γ_0. Similarly $v \leq 0$ on Γ_1. According then to the strong maximum principle for harmonic functions, $v < 0$ in U. In particular, $Du \neq 0$ within U. Consequently the Implicit Function Theorem implies that Γ_λ is a smooth surface for $0 < \lambda < 1$.

2. Extend u to equal 1 on all of V and write

$$U_\lambda := \{x \in W \mid u > \lambda\}.$$

Then U_λ is an open subset of W and $\partial U_\lambda = \Gamma_\lambda$. By the strong maximum principle, U_λ is connected.

Now let $x \in \Gamma_\lambda$ and let $\boldsymbol{\nu}(x)$ denote the outer unit normal to Γ_λ at x. Then $Du(x)$ points in the direction of $-\boldsymbol{\nu}(x)$. Since $v(x) < 0$, we have $x \cdot \boldsymbol{\nu}(x) > 0$. This inequality holds for each $x \in \Gamma_\lambda$.

3. It follows that Γ_λ is the boundary of a set star-shaped with respect to 0. To see this, return to the proof of the lemma in §9.4.2, and notice that if Γ_λ were not star-shaped with respect to 0, we could then find a point $x \in \Gamma_\lambda$ for which $y = \mu x \notin \Gamma_\lambda$ if μ is close to 1, $\mu < 1$. But then we can derive the contradiction

$$\boldsymbol{\nu}(x) \cdot \frac{x}{|x|} = -\lim_{\mu \to 1^-} \boldsymbol{\nu}(x) \cdot \frac{(\mu x - x)}{|\mu x - x|} \leq 0.$$

□

9.5.2. Radial symmetry.

In this section we take $U = B^0(0,1)$ to be the open unit ball in $\mathbb{R}^n$ and investigate this boundary-value problem for a semilinear Poisson PDE:

$$\text{(2)} \qquad \begin{cases} -\Delta u = f(u) & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

We are interested in *positive solutions*:

$$u > 0 \quad \text{in } U, \tag{3}$$

and will assume $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, but is otherwise arbitrary. Our intention is to prove that u is necessarily radial, that is, $u(x)$ depends only on $r = |x|$. This is quite an unexpectedly strong conclusion, since we are making essentially no assumption on the nonlinearity.

a. Maximum principles.

Our proofs will depend upon an extension of the maximum principle for second-order elliptic PDE.

LEMMA 1 (A refinement of Hopf's Lemma). *Suppose $V \subset \mathbb{R}^n$ is open, $v \in C^2(\bar{V})$, and $c \in L^\infty(V)$. Assume*

$$\begin{cases} -\Delta v + cv \geq 0 & \text{in } V \\ \quad\quad\quad v \geq 0 & \text{in } V. \end{cases} \tag{4}$$

Suppose also $v \not\equiv 0$.

(i) *If $x^0 \in \partial V$, $v(x^0) = 0$, and V satisfies the interior ball condition at x_0, then*

$$\frac{\partial v}{\partial \nu}(x^0) < 0. \tag{5}$$

(ii) *Furthermore,*

$$v > 0 \quad \text{in } V. \tag{6}$$

Observe that we are here making no hypothesis concerning the sign of the zeroth-order coefficient c.

Proof. Let $w := e^{-\lambda x_1} v$, where $\lambda > 0$ will be selected below. Then $v = e^{\lambda x_1} w$, and so

$$cv \geq \Delta v = \Delta(e^{\lambda x_1} w) = \lambda^2 v + 2\lambda e^{\lambda x_1} w_{x_1} + e^{\lambda x_1} \Delta w.$$

Thus

$$-\Delta w - 2\lambda w_{x_1} \geq (\lambda^2 - c) w \geq 0 \quad \text{in } V,$$

if $\lambda = \|c\|_{L^\infty}^{1/2}$.

Consequently w is a supersolution for the elliptic operator $Kw := -\Delta w - 2\lambda w_{x_1}$, which has no zeroth-order term. The strong maximum principle implies $w > 0$ in V. According to Hopf's Lemma (§6.4.2) therefore, $\frac{\partial w}{\partial \nu}(x^0) < 0$. But

$$\frac{\partial w}{\partial \nu}(x^0) = Dw(x^0) \cdot \boldsymbol{\nu}(x^0) = e^{-\lambda x_1^0} \frac{\partial v}{\partial \nu}(x^0)$$

since $v(x^0) = 0$. Assertion (i) therefore holds, and assertion (ii) follows since $w > 0$ in V. □

LEMMA 2 (Boundary estimates). *Let $u \in C^2(\bar{U})$ satisfy (2), (3). Then for each point $x^0 \in \partial U \cap \{x_n > 0\}$, either*

$$u_{x_n}(x^0) < 0 \tag{7}$$

or else

$$u_{x_n}(x^0) = 0, \ u_{x_n x_n}(x^0) > 0. \tag{8}$$

In either case, u is strictly decreasing as a function of x_n near x^0.

Proof. 1. Fix any point $x^0 \in \partial U \cap \{x_n > 0\}$ and let $\nu = \boldsymbol{\nu}(x^0) = (\nu_1, \dots, \nu_n)$ denote the outer unit normal to ∂U at x^0. Note $\nu_n > 0$.

2. We first claim

$$u_{x_n}(x^0) < 0,$$

provided

$$f(0) \geq 0. \tag{9}$$

Indeed

$$\begin{aligned} 0 = -\Delta u - f(u) &= -\Delta u - f(u) + f(0) - f(0) \\ &\leq -\Delta u + cu, \end{aligned}$$

for $c(x) := -\int_0^1 f'(su(x))\,ds$. According to Lemma 1, $\frac{\partial u}{\partial \nu}(x^0) < 0$. Since Du is parallel to $\boldsymbol{\nu}$ on ∂U and $\nu_n > 0$, we conclude $u_{x_n}(x^0) < 0$.

3. Now suppose

$$f(0) < 0. \tag{10}$$

If $u_{x_n}(x^0) < 0$, we are done. Otherwise, since Du is parallel to $\boldsymbol{\nu}$,

$$Du(x^0) = 0. \tag{11}$$

As (2) is invariant under a rotation of coordinate axes, we may as well suppose $x^0 = (0, \dots, 1)$, $\nu = (0, \dots, 1)$.

4. We assert

$$u_{x_i x_j}(x_0) = -f(0)\nu_i \nu_j \quad \text{for each } i, j = 1, \dots, n. \tag{12}$$

Since $u = 0$ on ∂U, we have $u(x', \gamma(x')) = 0$ for all $x' \in \mathbb{R}^{n-1}$, $|x'| \leq 1$, where $\gamma(x') = (1 - |x'|^2)^{1/2}$. Differentiating with respect to x_i and x_j $(i, j = 1, \dots, n-1)$ and using (11), we conclude

$$u_{x_i x_j}(x^0) = 0 \quad (i, j = 1, \dots, n-1). \tag{13}$$

Since $u_{x_n} \leq 0$ on $\partial U \cap \{x_n > 0\}$ and $u_{x_n}(x^0) = 0$, the mapping $x' \mapsto u_{x_n}(x', \gamma(x'))$ has a maximum at $x' = 0$. Thus

$$u_{x_n x_i}(x^0) = 0 \quad (i = 1, \dots, n-1). \tag{14}$$

Finally, (13), (14) and the PDE (2) force $u_{x_n x_n}(x^0) = -f(0)$. This equality is (12) for $\nu = (0, \dots, 1)$. Returning to the original coordinate axes, we obtain (12).

5. Setting $i = j = n$ in (12), we find using (10) that

$$u_{x_n x_n}(x^0) = -f(0)\nu_n \nu_n > 0.$$

□

b. Moving planes.

We introduce next a "moving plane" P_λ, across which we will reflect our partial differential equation.

Notation. (i) If $0 \leq \lambda \leq 1$, define the plane

$$P_\lambda := \{x \in \mathbb{R}^n \mid x_n = \lambda\}.$$

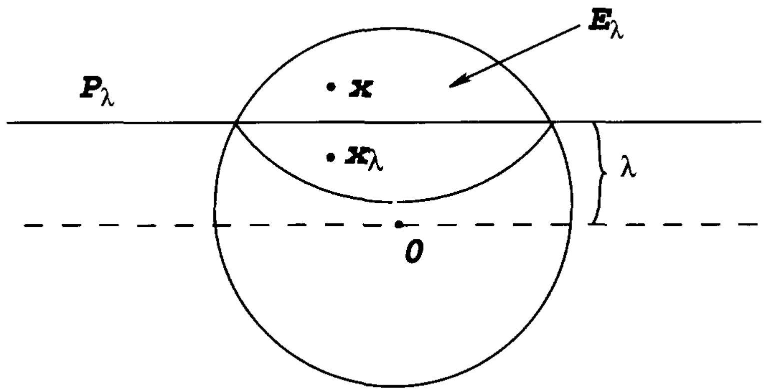
(ii) Write $x_\lambda := (x_1, \dots, x_{n-1}, 2\lambda - x_n)$ to denote the *reflection* of x in P_λ.

(iii) $E_\lambda := \{x \in U \mid \lambda < x_n < 1\}$. □

THEOREM 2 (Radial symmetry). *Let $u \in C^2(\bar{U})$ solve (2), (3). Then u is radial; that is,*

$$u(x) = v(r) \quad (r = |x|)$$

for some strictly decreasing function $v : [0, 1] \to [0, \infty)$.



Reflection through a plane

Proof. 1. We consider for each $0 \leq \lambda < 1$ the statement

$$(15_\lambda) \qquad u(x) < u(x_\lambda) \text{ for each point } x \in E_\lambda.$$

2. According to Lemma 2, (15_λ) is valid for each $\lambda < 1$, λ sufficiently close to 1. Set

$$(16) \qquad \lambda_0 := \inf\{0 \leq \lambda < 1 \mid (15_\mu) \text{ holds for each } \lambda \leq \mu < 1\}.$$

We will prove:

$$(17) \qquad \lambda_0 = 0.$$

Assume instead $\lambda_0 > 0$. Write $w(x) := u(x_{\lambda_0}) - u(x)$ $(x \in E_{\lambda_0})$. Then

$$-\Delta w = f(u(x_{\lambda_0})) - f(u(x)) = -cw \quad \text{in } E_{\lambda_0},$$

for $c(x) := -\int_0^1 f'(su(x_{\lambda_0}) + (1-s)u(x))\,ds$. As $w \geq 0$ in E_{λ_0}, we deduce from Lemma 1 (applied to $V = E_{\lambda_0}$) that $w > 0$ in E_{λ_0}, $w_{x_n} > 0$ on $P_{\lambda_0} \cap U$. Thus

$$(18) \qquad u(x) < u(x_{\lambda_0}) \quad \text{in } E_{\lambda_0},$$

and

$$(19) \qquad u_{x_n} < 0 \quad \text{on } P_{\lambda_0} \cap U.$$

Using (18), (19) and Lemma 2, we conclude

$$(20) \qquad u(x) < u(x_{\lambda_0 - \varepsilon}) \quad \text{in } E_{\lambda_0 - \varepsilon} \text{ for all } 0 \leq \varepsilon \leq \varepsilon_0,$$

if ε_0 is small enough. Assertion (20) contradicts our choice (16) of λ_0, if $\lambda_0 > 0$.

3. Since $\lambda_0 = 0$, we see $u(x_1, \dots, x_{n-1}, -x_n) \geq u(x_1, \dots, x_n)$ for all $x \in U \cap \{x_n > 0\}$. A similar argument in $U \cap \{x_n < 0\}$ shows $u(x_1, \dots, x_{n-1}, -x_n) \leq u(x_1, \dots, x_n)$ for all $x \in U \cap \{x_n > 0\}$. Thus u is symmetric in the plane P_0 and $u_{x_n} = 0$ on P_0.

This argument applies as well after any rotation of coordinate axes, and so the theorem follows. □

9.6. GRADIENT FLOWS

In this section we augment our discussion of abstract semigroup theory for linear operators (§7.4) by introducing certain nonlinear semigroups, generated by convex functions. Applications include various nonlinear second-order parabolic partial differential equations.

9.6.1. Convex functions on Hilbert spaces.

Convexity has been an essential ingredient in much of our analysis of nonlinear PDE thus far. We now broaden our view by considering convex functions defined on (possibly infinite dimensional) Hilbert spaces.

Hereafter H will denote a real Hilbert space, with inner product $(\ ,\ )$ and norm $\|\ \|$.

DEFINITION. *A function*

$$I : H \to (-\infty, \infty]$$

is convex *provided*

$$I[\tau u + (1-\tau)v] \le \tau I[u] + (1-\tau)I[v]$$

for all $u, v \in H$ *and each* $0 \le \tau \le 1$.

Note carefully that we allow I to take on the value $+\infty$ (but not $-\infty$). The function I is called *proper* if I is not identically equal to $+\infty$. The *domain* of I is

$$D(I) := \{u \in H \mid I[u] < +\infty\}.$$

DEFINITION. *We say* $I : H \to (-\infty, +\infty]$ *is* lower semicontinuous *if*

$$\begin{cases} u_k \to u \quad \text{in } H \text{ implies} \\ I[u] \le \liminf_{k\to\infty} I[u_k]. \end{cases}$$

As in the finite dimensional case (cf. §B.1), it is important to understand when the graph of I has a supporting hyperplane.

DEFINITIONS. *Let* $I : H \to (-\infty, +\infty]$ *be convex and proper.*

(i) *For each* $u \in H$, *we write*

$$\partial I[u] := \{v \in H \mid I[w] \ge I[u] + (v, w-u) \text{ for all } w \in H\}. \tag{1}$$

The mapping $\partial I : H \to 2^H$ *is the* subdifferential *of* I.

(ii) *We say* $u \in D(\partial I)$, *the* domain of ∂I, *provided* $\partial I[u] \ne \emptyset$.

The geometric interpretation of (1) is that $v \in \partial I[u]$ if and only if v is the "slope" of an affine functional touching the graph of I from below at the point u. Since this graph may have a "corner" at u, $\partial I[u]$ could be multivalued.

THEOREM 1 (Properties of subdifferentials). *Let $I : H \to (-\infty, +\infty]$ be convex, proper and lower semicontinuous. Then*

(i) $D(\partial I) \subseteq D(I)$.

(ii) *If $v \in \partial I[u]$ and $\tilde{v} \in \partial I[\tilde{u}]$, then*

$$(v - \tilde{v}, u - \tilde{u}) \geq 0 \quad \textit{(monotonicity)}.$$

(iii) $I[u] = \min_{w \in H} I[w]$ *if and only if* $0 \in \partial I[u]$.

(iv) *For each $w \in H$ and $\lambda > 0$, the problem*

$$u + \lambda \partial I[u] \ni w$$

has a unique solution $u \in D(\partial I)$.

Assertion (iv) means that there exists $u \in D(\partial I)$ and $v \in \partial I[u]$ such that

$$u + \lambda v = w.$$

Proof. 1. Let $u \in D(\partial I)$, $v \in \partial I[u]$. Then $I[w] \geq I[u] + (v, w - u)$ for all $w \in H$. Since I is proper, there exists a point u_0 with $I[u_0] < +\infty$. Thus $I[u] \leq I[u_0] + (v, u - u_0) < \infty$ and so $u \in D(I)$. This proves (i).

2. Given $v \in \partial I[u]$, $\tilde{v} \in \partial I[\tilde{u}]$, we know

$$I[\tilde{u}] \geq I[u] + (v, \tilde{u} - u), \ \ I[u] \geq I[\tilde{u}] + (\tilde{v}, u - \tilde{u}).$$

As (i) implies $I[u]$, $I[\tilde{u}] < +\infty$, we may add the foregoing inequalities and rearrange to obtain (ii).

3. If $I[u] = \min I$, then

$$I[w] \geq I[u] + (0, w - u) \text{ for all } w \in H. \tag{2}$$

Hence $0 \in \partial I[u]$. If conversely $0 \in \partial I[u]$, then (2) holds, and so $I[u] = \min I$.

4. Given $w \in H$ and $\lambda > 0$, define

$$J[u] := \tfrac{1}{2}\|u\|^2 + \lambda I[u] - (u, w) \quad (u \in H). \tag{3}$$

We intend to show that J attains its minimum over H.

Let us first claim that

$$\begin{cases} u_k \rightharpoonup u \ \ \textit{weakly} \text{ in } H \text{ implies} \\ I[u] \leq \liminf\limits_{k \to \infty} I[u_k]. \end{cases} \tag{4}$$

In other words, we are asserting for a convex function I that lower semicontinuity with respect to strong convergence of sequences implies lower semicontinuity with respect to weak convergence. To see this, suppose $u_k \rightharpoonup u$ in H and

$$\liminf_{k\to\infty} I[u_k] = \lim_{j\to\infty} I[u_{kj}] = l < \infty$$

for some subsequence $[u_{kj}]_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$. For each $\varepsilon > 0$ the set $K_\varepsilon = \{w \in H \mid I[w] \leq l+\varepsilon\}$ is closed and convex, and is thus weakly closed according to Mazur's Theorem (§D.4). Since all but finitely many of the points $\{u_{kj}\}_{j=1}^\infty$ lie in K_ε, u lies in K_ε, and consequently

$$I[u] \leq l + \varepsilon = \liminf_{k\to\infty} I[u_k] + \varepsilon.$$

This is true for each $\varepsilon > 0$ and thus (4) follows.

5. Next we assert that

$$I[u] \geq -C - C\|u\| \quad (u \in H) \tag{5}$$

for some constant C. To verify this claim we suppose to the contrary that for each $k = 1, 2, \dots$ there exists a point $u_k \in H$ such that

$$I[u_k] \leq -k - k\|u_k\|. \tag{6}$$

If the sequence $\{u_k\}_{k=1}^\infty$ is bounded in H, there exists according to §D.4 a weakly convergent subsequence: $u_{k_j} \rightharpoonup u$. But then (4) and (6) imply the contradiction $I[u] = -\infty$. Thus we may as well assume, passing if necessary to a subsequence, that $\|u_k\| \to \infty$. Select $u_0 \in H$ so that $I[u_0] < \infty$. Set

$$z_k := \frac{u_k}{\|u_k\|} + \left(1 - \frac{1}{\|u_k\|}\right) u_0 \quad (k = 1, 2, \dots) .$$

Then convexity implies

$$I[z_k] \leq \frac{1}{\|u_k\|} I[u_k] + \left(1 - \frac{1}{\|u_k\|}\right) I[u_0] \leq -k + |I[u_0]|.$$

As $\{z_k\}_{k=1}^\infty$ is bounded, we can extract a weakly convergent subsequence: $z_{k_j} \rightharpoonup z$, and again derive the contradiction $I[z] = -\infty$. We thereby establish the claim (5).

6. Return now to the function J defined by (3). Choose a minimizing sequence $\{u_k\}_{k=1}^\infty \subset H$ so that

$$J[u_k] \to \inf_{w\in H} J[w] = m.$$

Owing to (3), (5) m is a finite number. Thus we deduce from (3), (5) that the sequence $\{u_k\}_{k=1}^{\infty}$ is bounded. We may then extract a weakly convergent subsequence: $u_{k_j} \rightharpoonup u$. As the mapping $u \mapsto \|u\|^2$ is weakly lower semicontinuous, J has a minimum at u. Then assertion (iii) says $0 \in \partial J[u]$. A computation verifies that $\partial J[u] = u - w + \lambda \partial I[u]$, and so

$$u + \lambda \partial I[u] \ni w.$$

7. To confirm uniqueness, suppose as well

$$\tilde{u} + \lambda \partial I[\tilde{u}] \ni w.$$

Then $u + \lambda v = w$, $\tilde{u} + \lambda \tilde{v} = w$ for $v \in \partial I[u]$, $\tilde{v} \in \partial I[\tilde{u}]$. Owing to the monotonicity assertion (ii),

$$0 \leq (u - \tilde{u}, v - \tilde{v}) = \left(u - \tilde{u}, -\frac{u}{\lambda} + \frac{\tilde{u}}{\lambda}\right) = -\frac{1}{\lambda}\|u - \tilde{u}\|^2.$$

Since $\lambda > 0$, $u = \tilde{u}$. □

We introduce next nonlinear analogues of the operators R_λ, A_λ introduced in §7.4.

DEFINITIONS. (1) *For each* $\lambda > 0$ *define the* nonlinear resolvent $J_\lambda : H \to D(\partial I)$ *by setting*

$$J_\lambda[w] := u,$$

where u *is the unique solution of*

$$u + \lambda \partial I[u] \ni w.$$

(2) *For each* $\lambda > 0$ *define the* Yosida approximation $A_\lambda : H \to H$ *by*

$$A_\lambda[w] := \frac{w - J_\lambda[w]}{\lambda} \quad (w \in H). \tag{7}$$

Think of A_λ as a sort of regularization or smoothing of the operator $A = \partial I$.

THEOREM 2 (Properties of J_λ, A_λ). *For each* $\lambda > 0$ *and* $w, \hat{w} \in H$, *the following statements hold:*

(i) $\|J_\lambda[w] - J_\lambda[\hat{w}]\| \leq \|w - \hat{w}\|$,

(ii) $\|A_\lambda[w] - A_\lambda[\hat{w}]\| \leq \frac{2}{\lambda}\|w - \hat{w}\|$,

(iii) $0 \leq (w - \hat{w}, A_\lambda[w] - A_\lambda[\hat{w}])$,

(iv) $A_\lambda[w] \in \partial I[J_\lambda[w]]$.

(v) *If* $w \in D(\partial I)$, *then*

$$\sup_{\lambda>0} \|A_\lambda[w]\| \leq |A^0[w]|,$$

where $|A^0[w]| := \min_{z\in\partial I[w]} \|z\|$.

(vi) *For each* $w \in \overline{D(\partial I)}$,

$$\lim_{\lambda\to 0} J_\lambda[w] = w.$$

Proof. 1. Let $u = J_\lambda[w]$, $\tilde{u} = J_\lambda[\tilde{w}]$. Then $u + \lambda v = w$, $\tilde{u} + \lambda\tilde{v} = \tilde{w}$ for some $v \in \partial I[u]$, $\tilde{v} \in \partial I[\tilde{u}]$. Therefore

$$\begin{aligned}\|w - \tilde{w}\|^2 &= \|u - \tilde{u} + \lambda(v - \tilde{v})\|^2 \\ &= \|u - \tilde{u}\|^2 + 2\lambda(u - \tilde{u}, v - \tilde{v}) + \lambda^2\|v - \tilde{v}\|^2 \\ &\geq \|u - \tilde{u}\|^2,\end{aligned}$$

according to Theorem 1,(ii). This proves assertion (i), and assertion (ii) follows at once from the definition (7) of the Yosida approximation A_λ.

2. We verify (iii) by using (7) to compute:

$$\begin{aligned}(w - \hat{w}, A_\lambda[w] - A_\lambda[\hat{w}]) &= \frac{1}{\lambda}(\|w - \hat{w}\|^2 - (w - \hat{w}, J_\lambda[w] - J_\lambda[\hat{w}])) \\ &\geq \frac{1}{\lambda}(\|w - \hat{w}\|^2 - \|w - \hat{w}\| \, \|J_\lambda[w] - J_\lambda[\hat{w}]\|) \geq 0,\end{aligned}$$

according to (i).

3. To prove (iv), note $u = J_\lambda[w]$ if and only if $u + \lambda v = w$ for some $v \in \partial I[u] = \partial I[J_\lambda[w]]$. But

$$v = \frac{w - u}{\lambda} = \frac{w - J_\lambda[w]}{\lambda} = A_\lambda[w].$$

4. Assume next $w \in D(\partial I)$, $z \in \partial I[w]$. Let $u = J_\lambda[w]$; so that $u + \lambda v = w$, where $v \in \partial I[u]$. By monotonicity

$$0 \leq (w - u, z - v) = \left(w - J_\lambda[w], z - \frac{w - J_\lambda[w]}{\lambda}\right) = (\lambda A_\lambda[w], z - A_\lambda[w]).$$

Consequently

$$\lambda\|A_\lambda[w]\|^2 \leq (\lambda A_\lambda[w], z) \leq \lambda\|A_\lambda[w]\| \, \|z\|,$$

and so

$$\|A_\lambda[w]\| \leq \|z\|.$$

This estimate is valid for all $\lambda > 0$, $z \in \partial I[w]$. Assertion (v) follows.

5. If $w \in D(\partial I)$, then

$$\|J_\lambda[w] - w\| = \lambda\|A_\lambda[w]\| \leq \lambda|A^0[w]|\ ,$$

and hence $J_\lambda[w] \to w$ as $\lambda \to 0$. Now let $w \in \overline{D(\partial I)} - D(\partial I)$. There exists for each $\varepsilon \to 0$ a point $\hat{w} \in D(\partial I)$ with $\|w - \hat{w}\| \leq \varepsilon$. Then

$$\begin{aligned}\|J_\lambda[w] - w\| &\leq \|J_\lambda[w] - J_\lambda[\hat{w}]\| + \|J_\lambda[\hat{w}] - \hat{w}\| + \|w - \hat{w}\| \\ &\leq 2\|w - \hat{w}\| + \|J_\lambda[\hat{w}] - \hat{w}\| \\ &\leq 2\varepsilon + \|J_\lambda[\hat{w}] - \hat{w}\|.\end{aligned}$$

Since $\hat{w} \in D(\partial I)$, $J_\lambda[\hat{w}] \to \hat{w}$ as $\lambda \to 0$. Thus

$$\limsup_{\lambda \to 0} \|J_\lambda[w] - w\| \leq 2\varepsilon$$

for each $\varepsilon > 0$. □

9.6.2. Subdifferentials and nonlinear semigroups.

As above, let H be a real Hilbert space, and take $I : H \to (-\infty, +\infty]$ to be convex, proper, lower-semicontinuous. Let us for simplicity assume as well

$$\text{(8)} \qquad \partial I \text{ is densely defined,} \quad \text{i.e. } \overline{D(\partial I)} = H.$$

By analogy with the theory of linear semigroups set forth in §7.4, we propose now to study the differential equation

$$\text{(9)} \qquad \begin{cases} \mathbf{u}'(t) + A[\mathbf{u}(t)] \ni 0 & (t \geq 0) \\ \mathbf{u}(0) = u, \end{cases}$$

where $u \in H$ is given and $A = \partial I$ is a nonlinear, discontinuous operator, which is perhaps multivalued. Assuming for the moment (9) has a unique solution for each initial point u, we write

$$\text{(10)} \qquad \mathbf{u}(t) = S(t)u \quad (t > 0)$$

and regard $S(t)$ so defined as a mapping from H into H for each time $t \geq 0$.

Remark. We will employ the notation (10) to emphasize similarities with linear semigroup theory, previously introduced in §7.4. But carefully note here and afterwards that the mapping $u \mapsto S(t)u$ is in general *nonlinear.* □

As in §7.4 it is reasonable to expect further that

$$S(0)u = u \quad (u \in H), \tag{11}$$

$$S(t+s)u = S(t)S(s)u \quad (t, s \geq 0,\ u \in H), \tag{12}$$

and for each $u \in H$

(13) the mapping $t \mapsto S(t)u$ is continuous from $[0, \infty)$ into H.

DEFINITIONS. (i) *A family* $\{S(t)\}_{t\geq 0}$ *of nonlinear operators mapping* H *into* H *is called a* nonlinear semigroup *if conditions* (11)–(13) *are satisfied.*

(ii) *We say* $\{S(t)\}_{t\geq 0}$ *is a* contraction semigroup *if in addition*

$$\|S(t)u - S(t)\hat{u}\| \leq \|u - \hat{u}\| \quad (t \geq 0,\ u, \hat{u} \in H). \tag{14}$$

Our intention is to show that the operator $A = \partial I$ generates a nonlinear semigroup of contractions on H. In particular we will prove that the ODE

$$\begin{cases} \mathbf{u}'(t) \in -\partial I[\mathbf{u}(t)] \quad (t \geq 0) \\ \mathbf{u}(0) = u, \end{cases} \tag{15}$$

for a given initial point $u \in H$, is well-posed. This is a kind of infinite dimensional "gradient flow" governed by ∂I. Later in §9.6.3 we will see that certain quasilinear parabolic PDE can be cast into the abstract from (15).

THEOREM 3 (Solution of gradient flow). *For each* $u \in D(\partial I)$ *there exists a unique function*

$$\mathbf{u} \in C([0,\infty); H), \ \textit{with } \mathbf{u}' \in L^\infty(0,\infty; H), \tag{16}$$

such that

(i) $\mathbf{u}(0) = u$,

(ii) $\mathbf{u}(t) \in D(\partial I)$ *for each* $t > 0$,

and

(iii) $\mathbf{u}'(t) \in -\partial I[\mathbf{u}(t)]$ *for a.e.* $t \geq 0$.

Proof. 1. We first build approximate solutions by solving for each $\lambda > 0$ the ODE

$$\text{(17)} \qquad \begin{cases} \mathbf{u}_\lambda'(t) + A_\lambda[\mathbf{u}_\lambda(t)] = 0 & (t \geq 0) \\ \mathbf{u}_\lambda(0) = u. \end{cases}$$

According to Theorem 2,(ii) the Yosida approximation $A_\lambda : H \to H$ is an everywhere defined, Lipschitz continuous mapping, and thus (17) has a unique solution $\mathbf{u}_\lambda \in C^1([0,\infty); H)$.

Our plan is to show that as $\lambda \to 0^+$, the functions $\mathbf{u}_\lambda$ converge to a solution of (15). This is subtle, however, as the operator $A = \partial I$ is in general nonlinear, multivalued, and not everywhere defined.

2. First, let us take another point $v \in H$ and consider as well the ODE

$$\text{(18)} \qquad \begin{cases} \mathbf{v}_\lambda'(t) + A_\lambda[\mathbf{v}_\lambda(t)] = 0 & (t \geq 0) \\ \mathbf{v}_\lambda(0) = v. \end{cases}$$

Then

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\mathbf{u}_\lambda - \mathbf{v}_\lambda\|^2 &= (\mathbf{u}_\lambda' - \mathbf{v}_\lambda', \mathbf{u}_\lambda - \mathbf{v}_\lambda) \\ &= (-A_\lambda[\mathbf{u}_\lambda] + A_\lambda[\mathbf{v}_\lambda], \mathbf{u}_\lambda - \mathbf{v}_\lambda) \leq 0, \end{aligned}$$

owing to Theorem 2,(iii). Hence

$$\text{(19)} \qquad \|\mathbf{u}_\lambda(t) - \mathbf{v}_\lambda(t)\| \leq \|u - v\| \quad (t \geq 0).$$

In particular, if $h > 0$ and $v = \mathbf{u}_\lambda(h)$, then by uniqueness $\mathbf{v}_\lambda(t) = \mathbf{u}_\lambda(t+h)$. Consequently (19) implies

$$\|\mathbf{u}_\lambda(t+h) - \mathbf{u}_\lambda(t)\| \leq \|\mathbf{u}(h) - u\|.$$

Divide by h and send $h \to 0$:

$$\text{(20)} \qquad \|\mathbf{u}_\lambda'(t)\| \leq \|\mathbf{u}_\lambda'(0)\| = \|A_\lambda[u]\| \leq |A^0[u]|,$$

the last inequality resulting from Theorem 2,(v).

3. We next take $\lambda, \mu > 0$ and compute:

$$\text{(21)} \qquad \begin{aligned} \frac{1}{2}\frac{d}{dt}\|\mathbf{u}_\lambda - \mathbf{u}_\mu\|^2 &= (\mathbf{u}_\lambda' - \mathbf{u}_\mu', \mathbf{u}_\lambda - \mathbf{u}_\mu) \\ &= (-A_\lambda[\mathbf{u}_\lambda] + A_\mu[\mathbf{u}_\mu], \mathbf{u}_\lambda - \mathbf{u}_\mu). \end{aligned}$$

Now

$$\begin{aligned} \mathbf{u}_\lambda - \mathbf{u}_\mu &= (\mathbf{u}_\lambda - J_\lambda[\mathbf{u}_\lambda]) + (J_\lambda[\mathbf{u}_\lambda] - J_\mu[\mathbf{u}_\mu]) + (J_\mu[\mathbf{u}_\mu] - \mathbf{u}_\mu) \\ &= \lambda A_\lambda[\mathbf{u}_\lambda] + J_\lambda[\mathbf{u}_\lambda] - J_\mu[\mathbf{u}_\mu] - \mu A_\mu[\mathbf{u}_\mu]. \end{aligned}$$

Consequently

$$(22)\qquad \begin{aligned}(A_\lambda[\mathbf{u}_\lambda] - A_\mu[\mathbf{u}_\mu], \mathbf{u}_\lambda - \mathbf{u}_\mu) &= (A_\lambda[\mathbf{u}_\lambda] - A_\mu[\mathbf{u}_\mu], J_\lambda[\mathbf{u}_\lambda] - J_\mu[\mathbf{u}_\mu]) \\ &\quad + (A_\lambda[\mathbf{u}_\lambda] - A_\mu[\mathbf{u}_\mu], \lambda A_\lambda[\mathbf{u}_\lambda] - \mu A_\mu[\mathbf{u}_\mu]).\end{aligned}$$

Since $A_\lambda[\mathbf{u}_\lambda] \in \partial I[J_\lambda[\mathbf{u}_\lambda]]$ and $A_\mu[\mathbf{u}_\mu] \in \partial I[J_\mu[\mathbf{u}_\mu]]$, the monotonicity property implies that the first term on the right hand side of (22) is nonnegative. Thus

$$\begin{aligned}(A_\lambda[\mathbf{u}_\lambda] - A_\mu[\mathbf{u}_\mu], \mathbf{u}_\lambda - \mathbf{u}_\mu) \geq \lambda\|A_\lambda[\mathbf{u}_\lambda]\|^2 &+ \mu\|A_\mu[\mathbf{u}_\mu]\|^2 \\ &- (\lambda+\mu)\|A_\lambda[\mathbf{u}_\lambda]\|\ \|A_\mu[\mathbf{u}_\mu]\|.\end{aligned}$$

Since

$$\begin{aligned}(\lambda+\mu)\|A_\lambda[\mathbf{u}_\lambda]\|\ \|A_\mu[\mathbf{u}_\mu]\| &\leq \lambda\left(\|A_\lambda[\mathbf{u}_\lambda]\|^2 + \tfrac{1}{4}\|A_\mu[\mathbf{u}_\mu]\|^2\right) \\ &\quad + \mu\left(\|A_\mu[\mathbf{u}_\mu]\|^2 + \tfrac{1}{4}\|A_\lambda[\mathbf{u}_\lambda]\|^2\right),\end{aligned}$$

we deduce

$$(A_\lambda[\mathbf{u}_\lambda] - A_\mu[\mathbf{u}_\mu], \mathbf{u}_\lambda - \mathbf{u}_\mu) \geq -\frac{\lambda}{4}\|A_\mu[\mathbf{u}_\mu]\|^2 - \frac{\mu}{4}\|A_\lambda[\mathbf{u}_\lambda]\|^2.$$

But $\|A_\lambda[\mathbf{u}_\lambda]\| = \|\mathbf{u}'_\lambda\| \leq |A^0[u]|$ according to (20); whence

$$(A_\lambda[\mathbf{u}_\lambda] - A_\mu[\mathbf{u}_\mu], \mathbf{u}_\lambda - \mathbf{u}_\mu) \geq -\frac{\lambda+\mu}{4}\,|A^0[u]|^2.$$

Recalling (21), (22), we obtain the inequality

$$\frac{d}{dt}\|\mathbf{u}_\lambda - \mathbf{u}_\mu\|^2 \leq \frac{(\lambda+\mu)}{2}|A^0[u]|^2 \quad (t \geq 0);$$

and hence

$$(23)\qquad \|\mathbf{u}_\lambda(t) - \mathbf{u}_\mu(t)\|^2 \leq \frac{(\lambda+\mu)}{2}t|A^0u|^2 \quad (t \geq 0).$$

In view of estimate (23) there exists a function $\mathbf{u} \in C([0,\infty); H)$ such that

$$\mathbf{u}_\lambda \to \mathbf{u} \quad \text{uniformly in } C([0,T], H)$$

as $\lambda \to 0$, for each time $T > 0$. Furthermore estimate (20) implies

$$(24)\qquad \mathbf{u}'_\lambda \rightharpoonup \mathbf{u}' \quad \text{weakly in } L^2(0,T;H)$$

for each $T > 0$, and

$$(25)\qquad \|\mathbf{u}'(t)\| \leq |A^0[u]| \quad \text{for a.e. } t.$$

4. We must show $\mathbf{u}(t) \in D(\partial I)$ for each $t \geq 0$ and

$$\mathbf{u}'(t) + \partial I[\mathbf{u}(t)] \ni 0 \quad \text{for a.e. } t \geq 0.$$

Now

$$\|J_\lambda[\mathbf{u}_\lambda](t) - \mathbf{u}_\lambda(t)\| = \lambda\|A_\lambda[\mathbf{u}_\lambda](t)\| = \lambda\|\mathbf{u}'_\lambda(t)\| \leq \lambda|A^0[u]|$$

by (20). Hence

$$J_\lambda[\mathbf{u}_\lambda] \to \mathbf{u} \quad \text{uniformly in } C([0,T];H) \tag{26}$$

for each $T > 0$.

For each time $t \geq 0$,

$$-\mathbf{u}'_\lambda(t) = A_\lambda[\mathbf{u}_\lambda(t)] \in \partial I[J_\lambda[\mathbf{u}_\lambda(t)]].$$

Thus given $w \in H$, we have:

$$I[w] \geq I[J_\lambda[\mathbf{u}_\lambda(t)]] - (\mathbf{u}'_\lambda(t), w - J_\lambda[\mathbf{u}_\lambda(t)]).$$

Consequently if $0 \leq s \leq t$,

$$(t-s)I[w] \geq \int_s^t I[J_\lambda[\mathbf{u}_\lambda(r)]]\,dr - \int_s^t (\mathbf{u}'_\lambda(r), w - J_\lambda[\mathbf{u}_\lambda(r)])\,dr.$$

In view of (26), the lower semicontinuity of I, and Fatou's Lemma (§E.3), we conclude upon sending $\lambda \to 0$ that

$$(t-s)I[w] \geq \int_s^t I[\mathbf{u}(r)]\,dr - \int_s^t (\mathbf{u}'(r), w - \mathbf{u}(r))\,dr$$

for each $0 \leq s \leq t$. Therefore

$$I[w] \geq I[\mathbf{u}(t)] + (-\mathbf{u}'(t), w - \mathbf{u}(t))$$

if t is a Lebesgue point of $\mathbf{u}'$, $I[\mathbf{u}]$. Hence for a.e. $t \geq 0$,

$$I[w] \geq I[\mathbf{u}(t)] + (-\mathbf{u}'(t), w - \mathbf{u}(t))$$

for all $w \in H$. Thus $\mathbf{u}(t) \in D(\partial I)$, with

$$-\mathbf{u}'(t) \in \partial I[\mathbf{u}(t)]$$

for a.e. $t \geq 0$.

5. Finally we prove $\mathbf{u}(t) \in D(\partial I)$ for each $t \geq 0$. To see this, fix $t \geq 0$ and choose $t_k \to t$ such that $\mathbf{u}(t_k) \in D(\partial I)$, $-\mathbf{u}'(t_k) \in \partial I[\mathbf{u}(t_k)]$. In view of (25) we may assume, upon passing if necessary to a subsequence, that

$$\mathbf{u}'(t_k) \rightharpoonup \mathbf{v} \quad \text{weakly in } H.$$

Fix $w \in H$. Then

$$I[w] \geq I[\mathbf{u}(t_k)] + (-\mathbf{u}'(t_k), w - \mathbf{u}(t_k)).$$

Let $t_k \to t$ and recall that $\mathbf{u} \in C([0,\infty]; H)$ and I is lower semicontinuous. We obtain the inequality

$$I[w] \geq I[\mathbf{u}(t)] + (-\mathbf{v}, w - \mathbf{u}(t)).$$

Hence $\mathbf{u}(t) \in D(\partial I)$ and $-\mathbf{v} \in \partial I[\mathbf{u}(t)]$.

6. We have shown $\mathbf{u}$ satisfies assertions (i)–(iii). To prove uniqueness, assume $\tilde{\mathbf{u}}$ is another solution and compute

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{u} - \tilde{\mathbf{u}}\|^2 = (\mathbf{u}' - \tilde{\mathbf{u}}', \mathbf{u} - \tilde{\mathbf{u}}) \leq 0 \quad \text{for a.e. } t \geq 0,$$

since $-\mathbf{u}' \in \partial I[\mathbf{u}]$, $-\tilde{\mathbf{u}}' \in \partial I[\tilde{\mathbf{u}}]$. □

Remarks. (i) The operator $A = \partial I$ in fact generates a nonlinear contraction semigroup on all of H. If $u, v \in D(\partial I)$, we write as above

$$\lim_{\lambda \to 0} \mathbf{u}_\lambda(t) = \mathbf{u}(t) = S(t)u$$

and

$$\lim_{\lambda \to 0} \mathbf{v}_\lambda(t) = \mathbf{v}(t) = S(t)v.$$

Owing to (19) we see

$$\|S(t)u - S(t)v\| \leq \|u - v\| \quad (t \geq 0)$$

if $u, v \in D(\partial I)$. Using this inequality we uniquely extend the semigroup of nonlinear operators $\{S(t)\}_{t\geq 0}$ to $H = \overline{D(\partial I)}$.

(ii) We have assumed that $D(\partial I)$ is dense in H purely to streamline the exposition: in general $\{S(t)\}_{t\geq 0}$ is a semigroup of contractions on $\overline{D(\partial I)} \subseteq H$. □

9.6.3. Applications.

We now illustrate how some of the abstract theory set forth in §§9.6.1–2 applies to certain nonlinear parabolic partial differential equations.

Let us hereafter suppose U is a bounded, open subset of $\mathbb{R}^n$, with smooth boundary ∂U. We choose $H = L^2(U)$, and set

$$I[u] := \begin{cases} \int_U L(Du)\,dx & \text{if } u \in H_0^1(U) \\ +\infty & \text{otherwise,} \end{cases} \tag{27}$$

where $L : \mathbb{R}^n \to \mathbb{R}$ is smooth, convex, and satisfies

$$|D^2 L(p)| \leq C \quad (p \in \mathbb{R}^n), \tag{28}$$

$$\sum_{i,j=1}^{n} L_{p_i p_j}(p)\xi_i \xi_j \geq \theta |\xi|^2 \quad (p, \xi \in \mathbb{R}^n) \tag{29}$$

for constants $C, \theta > 0$.

THEOREM 4 (Characterization of ∂I).

(i) $I : H \to (-\infty, +\infty]$ *is convex, proper and lower semicontinuous.*

(ii) $D(\partial I) = H^2(U) \cap H_0^1(U)$.

(iii) *If* $u \in D(\partial I)$, *then* ∂I *is single-valued and*

$$\partial I[u] = -\sum_{i=1}^{n} (L_{p_i}(Du))_{x_i} \quad a.e.$$

Proof. 1. I is clearly proper and convex. Furthermore, since I is weakly sequentially lower semicontinuous (cf. Theorem 1 in §8.2.2), I is lower semicontinuous.

2. Define the nonlinear operator A by setting

$$\begin{cases} D(A) & := H^2(U) \cap H_0^1(U), \\ A[u] & := -\sum_{i=1}^{n} (L_{p_i}(Du))_{x_i} \quad (u \in D(A)). \end{cases}$$

We must prove $A = \partial I$.

3. First let $u \in D(A)$, $v = A[u]$, $w \in L^2(U)$. If $w \notin H_0^1(U)$, then $I[w] = +\infty$ and so clearly

$$I[w] \geq I[u] + (v, w - u). \tag{30}$$

Assume next $w \in H_0^1(U)$. Thus

$$(v, w-u) = -\int_U \sum_{i=1}^n (L_{p_i}(Du))_{x_i}(w-u)\,dx$$
$$= \int_U \sum_{i=1}^n L_{p_i}(Du)(Dw - Du)\,dx.$$

Since L is convex,

$$L(Dw) \geq L(Du) + D_pL(Du)\cdot(Dw - Du) \quad \text{a.e. in } U.$$

Integrating over U gives (30).

4. We have thus far shown that $A \subseteq \partial I$, that is, $D(A) \subseteq D(\partial I)$ and $Au \in \partial I[u]$ for $u \in D(A)$. To conclude, we must prove that $A \supseteq \partial I$.

Select any function $f \in L^2(U)$. If we minimize the functional

$$J[w] := \int_U L(Dw) + \frac{w^2}{2} - fw\,dx$$

over the admissible class $\mathcal{A} = H_0^1(U)$, we will find $u \in H_0^1(U)$, which is a weak solution of

$$(31) \qquad u - \sum_{i=1}^n (L_{p_i}(Du))_{x_i} = f \quad \text{in } U.$$

According to calculations similar to those for the proof of Theorem 1,(ii) in §8.3, we see that in fact $u \in H^2(U)$, with the estimate

$$(32) \qquad \|u\|_{H^2(U)} \leq C\|f\|_{L^2(U)}.$$

Thus $u \in D(A)$, and $u + A[u] = f$. Consequently the range of $I + A$ is all of H. But this implies $A = \partial I$. For if $v \in D(\partial I)$, $w \in \partial I[v]$, then there exists $u \in D(A)$ such that $u + A[u] = v + w$. Since $A[u] \in \partial I[u]$, $w \in \partial I[v]$, the uniqueness assertion of Theorem 1,(iv) implies $u = v$, $w = A[u]$. Thus $A = \partial I$. □

We can now look at the initial/boundary-value problem:

$$(33) \qquad \begin{cases} u_t - \sum_{i=1}^n (L_{p_i}(Du))_{x_i} = 0 & \text{in } U \times (0,\infty) \\ u = 0 & \text{on } \partial U \times (0,\infty) \\ u = g & \text{on } U \times \{t=0\}, \end{cases}$$

where $g \in L^2(U)$. In accordance with Theorem 4, we can recast this problem into the abstract form

$$(34) \qquad \begin{cases} \mathbf{u}'(t) = -\partial I[\mathbf{u}(t)] & (t \geq 0) \\ \mathbf{u}(0) = g. \end{cases}$$

We apply Theorem 3. If $g \in H^2(U) \cap H_0^1(U)$, there exists a unique function

$$\mathbf{u} \in C([0,\infty); L^2(U)), \text{ with } \mathbf{u}' \in L^\infty((0;\infty); L^2(U)),$$

that is, a weak solution of (33). In view of the estimate

$$\|\mathbf{u}(t)\|_{H^2(U)} \leq C\|\mathbf{u}'(t)\|_{L^2(U)},$$

we see $\mathbf{u} \in L^\infty((0,\infty), H^2(U) \cap H_0^1(U))$ as well.

9.7. PROBLEMS

In these problems U always denotes a bounded, open subset of $\mathbb{R}^n$, with smooth boundary.

1. Assume the vector field $\mathbf{v}$ is smooth. Give another proof of the lemma in §9.1 for this case by solving the ODE

$$\begin{cases} \dot{\mathbf{x}}(t) = -\mathbf{v}(\mathbf{x}(t)) & (t \geq 0) \\ \mathbf{x}(0) = y. \end{cases}$$

Let us write the solution as $\mathbf{x}(t,y)$ to display the dependence on the initial point y. For each fixed time $t > 0$, the map $y \mapsto \mathbf{x}(t,y)$ is continuous and so has a fixed point. Conclude $\mathbf{v}$ has a zero in the closed ball $B(0,r)$.

2. Assume $a : \mathbb{R} \to \mathbb{R}$ is continuous and $a(f_n) \rightharpoonup a(f)$ weakly in $L^2(0,1)$ whenever $f_n \rightharpoonup f$ weakly in $L^2(0,1)$. Show a is an affine function; that is, a has the form

$$a(z) = \alpha z + \beta \quad (z \in \mathbb{R})$$

for constants α, β.

3. Assume

$$\begin{cases} u_t - \Delta u = f & \text{in } U \times (0,\infty) \\ u = 0 & \text{on } \partial U \times (0,\infty) \\ u = g & \text{on } U \times \{t = 0\}, \end{cases}$$

where $g \in L^2(U)$, $f \in L^\infty(U_T)$ for each $T > 0$. Suppose $\tau > 0$ and f is τ-periodic in t; that is, $f(x,t) = f(x,t+\tau)$ $(x \in U, t \geq 0)$. Prove there exists a unique function $g \in L^2(U)$ for which the corresponding solution u is τ-periodic.

4. Consider the nonlinear boundary-value problem

$$\begin{cases} -\Delta u + b(Du) = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

Use Banach's fixed point theorem to show there exists a unique weak solution $u \in H^2(U) \cap H_0^1(U)$ provided $b : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous, with $\mathrm{Lip}(b)$ small enough.

5. Assume $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, bounded, with $f(0) = 0$ and $f'(0) > \lambda_1$, λ_1 denoting the principal eigenvalue for $-\Delta$ on $H_0^1(U)$. Use the method of sub- and supersolutions to show there exists a weak solution u of

$$\begin{cases} -\Delta u = f(u) & \text{in } U \\ u = 0 & \text{on } \partial U \\ u > 0 & \text{in } U. \end{cases}$$

6. Assume that $\underline{u}, \bar{u}$ are smooth sub- and supersolutions of the boundary-value problem (1) in §9.3. Use the maximum principle to verify directly

$$\underline{u} = u_0 \leq u_1 \leq \cdots \leq u_k \leq \cdots \leq \bar{u},$$

where the $\{u_k\}_{k=0}^{\infty}$ are defined as in §9.3.

7. Let $\epsilon > 0$. Define

$$\beta_\epsilon(z) := \begin{cases} 0 & \text{if } z \geq 0 \\ \frac{z}{\epsilon} & \text{if } z \leq 0, \end{cases}$$

and suppose $u_\epsilon \in H_0^1(U)$ is the weak solution of

$$(*) \qquad \begin{cases} -\Delta u_\epsilon + \beta_\epsilon(u_\epsilon) = f & \text{in } U \\ u_\epsilon = 0 & \text{on } \partial U, \end{cases}$$

where $f \in L^2(U)$. Prove that as $\epsilon \to 0$, $u^\epsilon \rightharpoonup u$ weakly in $H_0^1(U)$, u being the unique solution of the variational inequality

$$\int_U Du \cdot D(w-u)\,dx \geq \int_U f(w-u)\,dx$$

for all $w \in H_0^1(U)$ with $w \geq 0$ a.e.

Approximating the variational inequality by $(*)$ is the *penalty method.*

8. Let $K \subset \mathbb{R}^n$ be a closed, convex, nonempty set. Define

$$I[x] := \begin{cases} 0 & \text{if } x \in K \\ +\infty & \text{if } x \notin K. \end{cases}$$

Explicitly determine $A = \partial I$, $J_\lambda = (I + \lambda A)^{-1}$, $A_\lambda = \frac{I - J_\lambda}{\lambda}$ $(\lambda > 0)$ in terms of the geometry of K.

9. Give a simple example showing the flow

$$(*) \qquad \mathbf{u}' \in -\partial I[\mathbf{u}] \quad (t > 0)$$

may be irreversible. (That is, find a Hilbert space H and a convex, proper, lower semicontinuous function $I : H \to (-\infty, +\infty]$ such that the semigroup solution of $(*)$ satisfies

$$S(t)u = S(t)\hat{u}$$

for some $t > 0$ and $u \neq \hat{u}$.)

9.8. REFERENCES

Section 9.1 See J.-L. Lions [**L2**] and Zeidler [**ZD,** Vol. 2]. The booklet [**E**] is a survey of methods for confronting weak convergence problems for nonlinear PDE.

Section 9.2 Zeidler [**ZD,** Vol. 1] has more on fixed point techniques. Consult also Gilbarg–Trudinger [**G-T,** Chapter 11].

Section 9.3 See Smoller [**S,** Chapter 10].

Section 9.4 Section 9.4.1 is based upon Payne [**PA,** §9].

Section 9.5 Section 9.5.2 is due to Gidas–Ni–Nirenberg[**G-N-N**].

Section 9.6 This material is from Brezis [**BR2**], which contains much more on nonlinear semigroups on Hilbert spaces. See Barbu [**BA**] or Zeidler [**ZD,** Vol. 2] for nonlinear semigroup theory in Banach spaces.

HAMILTON–JACOBI EQUATIONS

10.1. INTRODUCTION, VISCOSITY SOLUTIONS

This chapter investigates the existence, uniqueness and other properties of appropriately defined weak solutions of the initial-value problem for the *Hamilton–Jacobi equation*:

$$(1) \qquad \begin{cases} u_t + H(Du, x) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Here the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is given, as is the initial function $g : \mathbb{R}^n \to \mathbb{R}$. The unknown is $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$, $u = u(x,t)$, and $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$. We will write $H = H(p,x)$, so that "p" is the name of the variable for which we substitute the gradient Du in the PDE.

We recall from our study of characteristics in §3.2 that in general there can be no smooth solution of (1) lasting for all times $t \geq 0$. We recall further that if H depends only on p and is convex, then the Hopf–Lax formula (expression (21) in §3.3.2) provides us with a type of generalized solution.

In this chapter we consider the general case that H depends also on x and, more importantly, is no longer necessarily convex in the variable p. We will discover in these new circumstances a different way to define a weak solution of (1).

Our approach is to consider first this approximate problem:

$$\text{(2)}\qquad \begin{cases} u^\epsilon_t + H(Du^\epsilon, x) - \epsilon\Delta u^\epsilon = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ \qquad\qquad\qquad\quad u^\epsilon = g & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases}$$

for $\epsilon > 0$. The idea is that whereas (1) involves a fully nonlinear first-order PDE, (2) is an initial-value problem for a quasilinear parabolic PDE, which turns out to have a smooth solution. The term $\epsilon\Delta$ in (2) in effect regularizes the Hamilton–Jacobi equation. Then of course we hope that as $\epsilon \to 0$ the solutions u^ϵ of (2) will converge to some sort of weak solution of (1). This technique is the method of *vanishing viscosity.*

However, as $\epsilon \to 0$ we can expect to lose control over the various estimates of the function u^ϵ and its derivatives: these estimates depend strongly on the regularizing effect of $\epsilon\Delta$ and blow up as $\epsilon \to 0$. However, it turns out that we can often in practice at least be sure that the family $\{u^\epsilon\}_{\epsilon>0}$ is bounded and equicontinuous on compact subsets of $\mathbb{R}^n \times [0,\infty)$. Consequently the Arzela-Ascoli compactness criterion, §C.7, ensures that

$$\text{(3)}\qquad u^{\epsilon_j} \to u \quad \text{locally uniformly in } \mathbb{R}^n \times [0,\infty),$$

for some subsequence $\{u^{\epsilon_j}\}_{j=1}^\infty$ and some limit function

$$\text{(4)}\qquad u \in C(\mathbb{R}^n \times [0,\infty)).$$

Now we can surely expect that u is some kind of solution of our initial-value problem (1), but as we only know u is continuous, and have absolutely no information as to whether Du and u_t exist in any sense, such an interpretation is difficult.

Similar problems have arisen before in Chapters 8 and 9, where we had to deal with the weak convergence of various would-be approximate solutions to other nonlinear partial differential equations. Remember in particular that in §9.1 we solved a divergence structure quasilinear elliptic PDE by passing to limits using the method of Browder and Minty. Roughly speaking, we there integrated by parts to throw "hard-to-control" derivatives onto a fixed test function, and only then tried to go to limits to discover a solution. We will for the Hamilton–Jacobi equation (1) attempt something similar. We will fix a smooth test function v and will pass from (2) to (1) as $\epsilon \to 0$ by first "putting the derivatives onto v".

But since (1) is fully nonlinear, and in particular is not of divergence structure, we cannot just integrate by parts, as we did in §9.1, to switch to differentiations on v. Instead we will exploit the maximum principle to accomplish this transition, at least at certain points.

We will call the solution we build a *viscosity solution*, in honor of the vanishing viscosity technique. Our main goal will then be to discover an intrinsic characterization of such generalized solutions of (1).

10.1.1. Definitions.

Motivation for definition of viscosity solution. We henceforth assume that H, g are continuous and will as necessary later add further hypotheses.

The technique alluded to above works as follows. Fix any smooth test function $v \in C^\infty(\mathbb{R}^n \times (0,\infty))$ and suppose

$$\begin{cases} u - v \text{ has a } \textit{strict} \text{ local maximum at some point} \\ (x_0, t_0) \in \mathbb{R}^n \times (0,\infty). \end{cases} \tag{5}$$

This means

$$(u-v)(x_0,t_0) > (u-v)(x,t)$$

for all points (x,t) sufficiently close to (x_0,t_0), with $(x,t) \neq (x_0,t_0)$.

Now recall (3). We claim for each sufficiently small $\epsilon_j > 0$, there exists a point $(x_{\epsilon_j}, t_{\epsilon_j})$ such that

$$u^{\epsilon_j} - v \text{ has a local maximum at } (x_{\epsilon_j}, t_{\epsilon_j}) \tag{6}$$

and

$$(x_{\epsilon_j}, t_{\epsilon_j}) \to (x_0, t_0) \quad \text{as } j \to \infty. \tag{7}$$

To confirm this, note that for each sufficiently small $r > 0$, (5) implies $\max_{\partial B}(u-v) < (u-v)(x_0,t_0)$, B denoting the closed ball in $\mathbb{R}^{n+1}$ with center (x_0,t_0) and radius r. In view of (3), $u^{\epsilon_j} \to u$ uniformly on B, and so $\max_{\partial B}(u^{\epsilon_j} - v) < (u^{\epsilon_j} - v)(x_0,t_0)$ provided ϵ_j is small enough. Consequently $u^\epsilon - v$ attains a local maximum at some point in the interior of B. We can next replace r by a sequence of radii tending to zero to obtain (6), (7).

Now owing to (6) we see that the equations

$$Du^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) = Dv(x_{\epsilon_j}, t_{\epsilon_j}), \tag{8}$$

$$u_t^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) = v_t(x_{\epsilon_j}, t_{\epsilon_j}), \tag{9}$$

and the inequality

$$-\Delta u^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) \geq -\Delta v(x_{\epsilon_j}, t_{\epsilon_j}) \tag{10}$$

hold. We consequently can calculate

$$
(11)\quad \begin{aligned} & v_t(x_{\epsilon_j}, t_{\epsilon_j}) + H(Dv(x_{\epsilon_j}, t_{\epsilon_j}), x_{\epsilon_j}) \\ & \qquad = u_t^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) + H(Du^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}), x_{\epsilon_j}) \quad \text{by (8),(9)} \\ & \qquad = \epsilon_j \Delta u^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) \quad \text{by (2)} \\ & \qquad \le \epsilon_j \Delta v(x_{\epsilon_j}, t_{\epsilon_j}) \quad \text{by (10).} \end{aligned}
$$

Now let $\epsilon_j \to 0$ and remember (7). Since v is smooth and H is continuous, we deduce

$$
(12)\qquad v_t(x_0, t_0) + H(Dv(x_0, t_0), x_0) \le 0.
$$

We have established this inequality assuming (5). Suppose now instead that

$$
(13)\qquad u - v \text{ has a local maximum at } (x_0, t_0),
$$

but that this maximum is not necessarily strict. Then $u - \tilde{v}$ has a strict local maximum at (x_0, t_0), for $\tilde{v}(x,t) := v(x,t) + \delta(|x - x_0|^2 + (t - t_0)^2)$ $(\delta > 0)$. We thus conclude as above that $\tilde{v}_t(x_0, t_0) + H(D\tilde{v}(x_0, t_0), x_0) \le 0$; whereupon (12) again follows.

Consequently (13) implies inequality (12). Similarly, we deduce the reverse inequality

$$
(14)\qquad v_t(x_0, t_0) + H(Dv(x_0, t_0), x_0) \ge 0,
$$

provided

$$
(15)\qquad u - v \text{ has a local minimum at } (x_0, t_0).
$$

The proof is exactly like that above, except that the inequalities in (10), and thus in (11), are reversed.

In summary, we have discovered for any smooth function v that inequality (12) follows from (13), and (14) from (15). We have in effect put the derivatives onto v, at the expense of certain inequalities holding. □

Our intention now is to *define* a weak solution of (1) in terms of (12), (13) and (14), (15).

DEFINITION. *A bounded, uniformly continuous function u is called a* viscosity solution *of the initial-value problem* (1) *for the Hamilton–Jacobi equation provided:*

(i) $u = g$ *on* $\mathbb{R}^n \times \{t = 0\}$, *and*

(ii) *for each* $v \in C^\infty(\mathbb{R}^n \times (0,\infty))$,

$$(16)\quad \begin{cases} \textit{if } u - v \textit{ has a local maximum at a point } (x_0,t_0) \in \mathbb{R}^n \times (0,\infty), \\ \textit{then} \\ \qquad v_t(x_0,t_0) + H(Dv(x_0,t_0),x_0) \leq 0, \end{cases}$$

and

$$(17)\quad \begin{cases} \textit{if } u - v \textit{ has a local minimum at a point } (x_0,t_0) \in \mathbb{R}^n \times (0,\infty), \\ \textit{then} \\ \qquad v_t(x_0,t_0) + H(Dv(x_0,t_0),x_0) \geq 0. \end{cases}$$

Remark. Note carefully that by definition a viscosity solution satisfies (16), (17), and so all subsequent deductions must be based on these inequalities. The previous discussion was purely motivational.

For emphasis, we repeat the same point, which has caused some confusion among students. To verify that a given function u is a viscosity solution of the Hamilton–Jacobi equation $u_t + H(Du,x) = 0$, we must confirm that (16), (17) hold for all smooth functions v. Now the argument above shows that *if* u is constructed using the vanishing viscosity method, it is indeed a viscosity solution. But we will also see later in §10.3 that viscosity solutions can be built in entirely different ways, which have nothing whatsoever to do with vanishing viscosity.

The point is that the inequalities (16), (17) provide an intrinsic characterization, and indeed the very definition, of our generalized solutions.

□

We devote the rest of this chapter to demonstrating that viscosity solutions provide an appropriate and useful notion of weak solutions for our Hamilton–Jacobi PDE.

10.1.2. Consistency.

Let us begin by checking that the notion of viscosity solution is consistent with that of a classical solution. First of all, note that if $u \in C^1(\mathbb{R}^n \times [0,\infty))$ solves (1) and if u is bounded and uniformly continuous, then u is a viscosity solution. That is, we assert that any *classical solution of* $u_t + H(Du,x) = 0$ *is also a viscosity solution.* The proof is easy. If v is smooth and $u - v$ obtains a local maximum at (x_0,t_0), then

$$\begin{cases} Du(x_0,t_0) = Dv(x_0,t_0) \\ u_t(x_0,t_0) = v_t(x_0,t_0). \end{cases}$$

Consequently

$$\begin{aligned} v_t(x_0,t_0) &+ H(Dv(x_0,t_0),x_0) \\ &= u_t(x_0,t_0) + H(Du(x_0,t_0),x_0) = 0, \end{aligned}$$

since u solves (1). A similar equality holds at any point (x_0, t_0) where $u - v$ has a local minimum.

Next we assert that *any sufficiently smooth viscosity solution is a classical solution*, and, even more, that if a viscosity solution is differentiable at some point, then it solves the Hamilton–Jacobi PDE there. We will need the following calculus fact:

LEMMA (Touching by a C^1 function). *Assume $u : \mathbb{R}^n \to \mathbb{R}$ is continuous and is also differentiable at some point x_0. Then there exists a function $v \in C^1(\mathbb{R}^n)$ such that*

$$u(x_0) = v(x_0) \tag{18}$$

and

$$u - v \textit{ has a strict local maximum at } x_0. \tag{19}$$

Proof. 1. We may as well assume

$$x_0 = 0,\ u(0) = Du(0) = 0; \tag{20}$$

for otherwise we could consider $\tilde{u}(x) := u(x + x_0) - u(x_0) - Du(x_0) \cdot x$ in place of u.

2. In view of (20) and our hypothesis, we have

$$u(x) = |x|\rho_1(x), \tag{21}$$

where

$$\rho_1 : \mathbb{R}^n \to \mathbb{R} \text{ is continuous, } \ \rho_1(0) = 0. \tag{22}$$

Set

$$\rho_2(r) := \max_{x \in B(r)} \{|\rho_1(x)|\} \quad (r \geq 0). \tag{23}$$

Then

$$\rho_2 : [0, \infty) \to [0, \infty) \text{ is continuous, } \ \rho_2(0) = 0, \tag{24}$$

and

(25) ρ_2 is nondecreasing.

3. Now write

$$v(x) := \int_{|x|}^{2|x|} \rho_2(r)\,dr + |x|^2 \quad (x \in \mathbb{R}^n).$$

Since $|v(x)| \le |x|\rho_2(2|x|) + |x|^2$, we observe

(26) $$v(0) = Dv(0) = 0.$$

Furthermore if $x \neq 0$, we have

$$Dv(x) = \frac{2x}{|x|}\rho_2(2|x|) - \frac{x}{|x|}\rho_2(|x|) + 2x,$$

and so $v \in C^1(\mathbb{R}^n)$.

4. Finally note that if $x \neq 0$,

$$\begin{aligned} u(x) - v(x) &= |x|\rho_1(x) - \int_{|x|}^{2|x|} \rho_2(r)\,dr - |x|^2 \\ &\le |x|\rho_2(|x|) - \int_{|x|}^{2|x|} \rho_2(r)\,dr - |x|^2 \\ &\le -|x|^2 \quad \text{by (25)} \\ &< 0 = u(0) - v(0). \end{aligned}$$

Thus $u - v$ has a strict local maximum at 0, as required. □

THEOREM 1 (Consistency of viscosity solutions). *Let u be a viscosity solution of* (1), *and suppose u is differentiable at some point $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$. Then*

$$u_t(x_0, t_0) + H(Du(x_0, t_0), x_0) = 0.$$

Proof. 1. Applying the lemma above to u, with $\mathbb{R}^{n+1}$ replacing $\mathbb{R}^n$ and (x_0, t_0) replacing x_0, we deduce there exists a C^1 function v such that

(27) $u - v$ has a strict maximum at (x_0, t_0).

2. Now set $v^\epsilon := \eta_\epsilon * v$, η_ϵ denoting the usual mollifier in the $n+1$ variables (x, t). Then

(28) $$\begin{cases} v^\epsilon \to v \\ Dv^\epsilon \to Dv \quad \text{uniformly near } (x_0, t_0) \\ v_t^\epsilon \to v_t; \end{cases}$$

and so (27) implies

(29) $$u - v^\epsilon \text{ has a maximum at some point } (x_\epsilon, t_\epsilon),$$

with

(30) $$(x_\epsilon, t_\epsilon) \to (x_0, t_0) \quad \text{as } \epsilon \to 0.$$

Applying then the definition of viscosity solution, we see

$$v_t^\epsilon(x_\epsilon, t_\epsilon) + H(Dv^\epsilon(x_\epsilon, t_\epsilon), x_\epsilon) \leq 0.$$

Let $\epsilon \to 0$ and use (28), (30) to deduce

(31) $$v_t(x_0, t_0) + H(Dv(x_0, t_0), x_0) \leq 0.$$

But in view of (27), we see that since u is differentiable at (x_0, t_0),

$$Du(x_0, t_0) = Dv(x_0, t_0), \;\; u_t(x_0, t_0) = v_t(x_0, t_0).$$

Substitute above, to conclude from (31) that

(32) $$u_t(x_0, t_0) + H(Du(x_0, t_0), x_0) \leq 0.$$

3. Now apply the lemma above to $-u$ in $\mathbb{R}^{n+1}$, to find a C^1 function v such that $u - v$ has a strict minimum at (x_0, t_0). Then, arguing as above, we likewise deduce

$$u_t(x_0, t_0) + H(Du(x_0, t_0), x_0) \geq 0.$$

This inequality and (32) complete the proof. □

10.2. UNIQUENESS

Our goal now is to establish the uniqueness of a viscosity solution of our initial-value problem for Hamilton–Jacobi PDE. To be slightly more general, let us fix a time $T > 0$ and consider the problem

(1) $$\begin{cases} u_t + H(Du, x) = 0 & \text{in } \mathbb{R}^n \times (0, T] \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

We say that a bounded, uniformly continuous function u is a viscosity solution of (1) provided $u = g$ on $\mathbb{R}^n \times \{t = 0\}$, and the inequalities in (16) (or (17)) from §10.1.1 hold if $u - v$ has a local maximum (or minimum) at a point $(x_0, t_0) \in \mathbb{R}^n \times (0, T)$.

LEMMA (Extrema at a terminal time). *Assume u is a viscosity solution of (1) and $u - v$ has a local maximum (minimum) at a point $(x_0, t_0) \in \mathbb{R}^n \times (0, T]$. Then*

(2) $$v_t(x_0, t_0) + H(Dv(x_0, t_0), x_0) \leq 0 \quad (\geq 0).$$

The point is that we are now allowing for $t_0 = T$.

Proof. Assume $u - v$ has a local maximum at the point (x_0, T); as before we may assume that this is a strict local maximum. Write

$$\tilde{v}(x,t) := v(x,t) + \frac{\epsilon}{T-t} \quad (x \in \mathbb{R}^n, \ 0 < t < T).$$

Then for $\epsilon > 0$ small enough, $u - \tilde{v}$ has a local maximum at a point (x_ϵ, t_ϵ), where $0 < t_\epsilon < T$ and $(x_\epsilon, t_\epsilon) \to (x_0, T)$. Consequently

$$\tilde{v}_t(x_\epsilon, t_\epsilon) + H(D\tilde{v}(x_\epsilon, t_\epsilon), x_\epsilon) \leq 0,$$

and so

$$v_t(x_\epsilon, t_\epsilon) + \frac{\epsilon}{(T-t_\epsilon)^2} + H(Dv(x_\epsilon, t_\epsilon), x_\epsilon) \leq 0.$$

Letting $\epsilon \to 0$, we find

$$v_t(x_0, T) + H(Dv(x_0, T), x_0) \leq 0.$$

This proves (2) if $u - v$ has a maximum at (x_0, T). A similar proof gives the reverse inequality should $u - v$ have a minimum at (x_0, T). □

To go further, let us hereafter suppose the Hamiltonian H to satisfy these conditions of Lipschitz continuity:

$$\text{(3)} \qquad \begin{cases} |H(p,x) - H(q,x)| \leq C|p-q| \\ |H(p,x) - H(p,y)| \leq C|x-y|(1+|p|) \end{cases}$$

for $x, y, p, q \in \mathbb{R}^n$ and some constant $C \geq 0$.

We come next to the central fact concerning viscosity solutions of the initial-value problem (1), namely uniqueness. This important assertion justifies our taking the inequalities (16) and (17) from §10.1.1 as the foundation of our theory.

THEOREM 1 (Uniqueness of viscosity solution). *Under assumption* (3) *there exists at most one viscosity solution of* (1).

Remark. The following proof is based upon an unusual idea of "doubling the number of variables". See the proof of Theorem 3 in §11.4.3 for a related technique. □

Proof*. 1. Assume u and $\tilde{u}$ are both viscosity solutions with the same initial conditions, but

$$\sup_{\mathbb{R}^n\times[0,T]}(u-\tilde{u}) =: \sigma > 0. \tag{4}$$

Choose $0<\epsilon,\ \lambda<1$ and set

$$\begin{aligned}\Phi(x,y,t,s) :=& u(x,t)-\tilde{u}(y,s)-\lambda(t+s)\\ &-\frac{1}{\epsilon^2}(|x-y|^2+(t-s)^2)-\epsilon(|x|^2+|y|^2),\end{aligned} \tag{5}$$

for $x,y\in\mathbb{R}^n,\ t,s\geq 0$. Then there exists a point $(x_0,y_0,t_0,s_0)\in\mathbb{R}^{2n}\times[0,T]^2$ such that

$$\Phi(x_0,y_0,t_0,s_0)=\max_{\mathbb{R}^{2n}\times[0,T]^2}\Phi(x,y,t,s). \tag{6}$$

2. We may fix $0<\epsilon,\ \lambda<1$ so small that (4) implies

$$\Phi(x_0,y_0,t_0,s_0)\geq\sup_{\mathbb{R}^n\times[0,T]}\Phi(x,x,t,t)\geq\frac{\sigma}{2}. \tag{7}$$

In addition, $\Phi(x_0,y_0,t_0,s_0)\geq\Phi(0,0,0,0)$; and therefore

$$\begin{aligned}\lambda(t_0+s_0)+\frac{1}{\epsilon^2}(|x_0-y_0|^2+(t_0-s_0)^2)+\epsilon(|x_0|^2+|y_0|^2)\\ \leq u(x_0,t_0)-\tilde{u}(y_0,s_0)-u(0,0)+\tilde{u}(0,0).\end{aligned} \tag{8}$$

Since u and $\tilde{u}$ are bounded, we deduce

$$|x_0-y_0|,|t_0-s_0|=O(\epsilon)\quad\text{as }\epsilon\to 0. \tag{9}$$

Furthermore (8) implies $\epsilon(|x_0|^2+|y_0|^2)=O(1)$, and consequently

$$\begin{aligned}\epsilon(|x_0|+|y_0|)&=\epsilon^{1/4}\epsilon^{3/4}(|x_0|+|y_0|)\\ &\leq\epsilon^{1/2}+C\epsilon^{3/2}(|x_0|^2+|y_0|^2)\\ &\leq C\epsilon^{1/2}.\end{aligned}$$

Thus

$$\epsilon(|x_0|+|y_0|)=O(\epsilon^{1/2}). \tag{10}$$

*Omit on first reading.

3. Since $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(x_0, x_0, t_0, t_0)$, we also have

$$\begin{aligned} u(x_0,t_0) - \tilde{u}(y_0,s_0) - \lambda(t_0+s_0) - \frac{1}{\epsilon^2}(|x_0-y_0|^2 + (t_0-s_0)^2) \\ - \epsilon(|x_0|^2+|y_0|^2) \geq u(x_0,t_0) - \tilde{u}(x_0,t_0) - 2\lambda t_0 - 2\epsilon|x_0|^2. \end{aligned}$$

Hence

$$\begin{aligned} \frac{1}{\epsilon^2}(|x_0-y_0|^2 + (t_0-s_0)^2) \leq \tilde{u}(x_0,t_0) - \tilde{u}(y_0,s_0) + \lambda(t_0-s_0) \\ + \epsilon(x_0+y_0)\cdot(x_0-y_0). \end{aligned}$$

In view of (9), (10) and the uniform continuity of $\tilde{u}$, we deduce

$$|x_0-y_0|, |t_0-s_0| = o(\epsilon). \tag{11}$$

4. Now write $\omega(\cdot)$ to denote the modulus of continuity of u; that is,

$$|u(x,t) - u(y,s)| \leq \omega(|x-y| + |t-s|)$$

for all $x, y \in \mathbb{R}^n$, $0 \leq t,\ s \leq T$, and $\omega(r) \to 0$ as $r \to 0$. Similarly, $\tilde{\omega}(\cdot)$ will denote the modulus of continuity of $\tilde{u}$.

Then (7) implies

$$\begin{aligned} \frac{\sigma}{2} \leq u(x_0,t_0) - \tilde{u}(y_0,s_0) &= u(x_0,t_0) - u(x_0,0) + u(x_0,0) - \tilde{u}(x_0,0) \\ &\quad + \tilde{u}(x_0,0) - \tilde{u}(x_0,t_0) + \tilde{u}(x_0,t_0) - \tilde{u}(y_0,s_0) \\ &\leq \omega(t_0) + \tilde{\omega}(t_0) + \tilde{\omega}(o(\epsilon)), \end{aligned}$$

by (9), (11) and the initial condition. We can now take $\epsilon > 0$ to be so small that the foregoing implies $\frac{\sigma}{4} \leq \omega(t_0) + \tilde{\omega}(t_0)$; and this in turn implies $t_0 \geq \mu > 0$ for some constant $\mu > 0$. Similarly we have $s_0 \geq \mu > 0$.

5. Now observe in light of (6) that the mapping $(x,t) \mapsto \Phi(x, y_0, t, s_0)$ has a maximum at the point (x_0, t_0). In view of (5) then,

$$u - v \text{ has a maximum at } (x_0, t_0)$$

for

$$v(x,t) := \tilde{u}(y_0,s_0) + \lambda(t+s_0) + \frac{1}{\epsilon^2}(|x-y_0|^2 + (t-s_0)^2) + \epsilon(|x|^2 + |y_0|^2).$$

Since u is a viscosity solution of (1) we conclude, using the lemma if necessary, that

$$v_t(x_0,t_0) + H(D_x v(x_0,t_0), x_0) \leq 0.$$

Therefore

$$\lambda + \frac{2(t_0 - s_0)}{\epsilon^2} + H\left(\frac{2}{\epsilon^2}(x_0 - y_0) + 2\epsilon x_0, x_0\right) \leq 0. \tag{12}$$

We further observe that since the mapping $(y, s) \mapsto -\Phi(x_0, y, t_0, s)$ has a minimum at the point (y_0, s_0),

$$\tilde{u} - \tilde{v} \text{ has a minimum at } (y_0, s_0)$$

for

$$\tilde{v}(y, s) := u(x_0, t_0) - \lambda(t_0 + s) - \frac{1}{\epsilon^2}(|x - y_0|^2 + (t_0 - s)^2) - \epsilon(|x_0|^2 + |y|^2).$$

As $\tilde{u}$ is a viscosity solution of (1), we know then that

$$\tilde{v}_s(y_0, s_0) + H(D_y\tilde{v}(y_0, s_0), y_0) \geq 0.$$

Consequently

$$-\lambda + \frac{2(t_0 - s_0)}{\epsilon^2} + H\left(\frac{2}{\epsilon^2}(x_0 - y_0) - 2\epsilon y_0, y_0\right) \geq 0. \tag{13}$$

6. Next, subtract (13) from (12):

$$2\lambda \leq H\left(\frac{2}{\epsilon^2}(x_0 - y_0) - 2\epsilon y_0, y_0\right) - H\left(\frac{2}{\epsilon^2}(x_0 - y_0) + 2\epsilon x_0, x_0\right). \tag{14}$$

In view of hypothesis (3) therefore,

$$\lambda \leq C\epsilon(|x_0| + |y_0|) + C|x_0 - y_0|\left(1 + \frac{|x_0 - y_0|}{\epsilon^2} + \epsilon(|x_0| + |y_0|)\right). \tag{15}$$

We employ estimates (10), (11) in (15), and then let $\epsilon \to 0$, to discover $0 < \lambda \leq 0$. This contradiction completes the proof. □

10.3. CONTROL THEORY, DYNAMIC PROGRAMMING

It remains for us to establish the existence of a viscosity solution to our initial-value problem for the Hamilton–Jacobi partial differential equation. One method would be now to prove the existence of a smooth solution u^ϵ of the regularized equation (2) in §10.1 and then to make good enough uniform estimates. This technique in fact works, but requires knowledge of certain bounds for the heat equation beyond the scope of this book.

In this section we provide an alternative approach of independent interest, which is suitable for Hamiltonians which are convex in p.

We will first of all introduce some of the basic issues concerning *control theory* for ordinary differential equations and the connection with Hamilton–Jacobi PDE afforded by the method of *dynamic programming.* This discussion will make clearer the connections of the theory developed above in §§10.1–2 with that set forth earlier in §3.3.1. The remarkable fact is that *the defining viscosity solution inequalities* (16), (17) *in* §10.1.1 *are a consequence of the optimality conditions of control theory.*

10.3.1. Introduction to control theory.

We will now study the possibility of optimally controlling the solution $\mathbf{x}(\cdot)$ of the ordinary differential equation

$$
(1) \qquad \begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) & (t < s < T) \\ \mathbf{x}(t) = x. \end{cases}
$$

Here $\dot{} = \frac{d}{ds}$, $T > 0$ is a fixed terminal time, and $x \in \mathbb{R}^n$ is a given initial point, taken on by our solution $\mathbf{x}(\cdot)$ at the starting time $t \geq 0$. At later times $t < s < T$, $\mathbf{x}(\cdot)$ evolves according to the ODE, where

$$
\mathbf{f} : \mathbb{R}^n \times A \to \mathbb{R}^n
$$

is a given bounded, Lipschitz continuous function, and A is some given compact subset of, say, $\mathbb{R}^m$. The function $\boldsymbol{\alpha}(\cdot)$ appearing in (1) is a *control,* that is, some appropriate scheme for adjusting parameters from the set A as time evolves, thereby affecting the dynamics of the system modeled by (1).

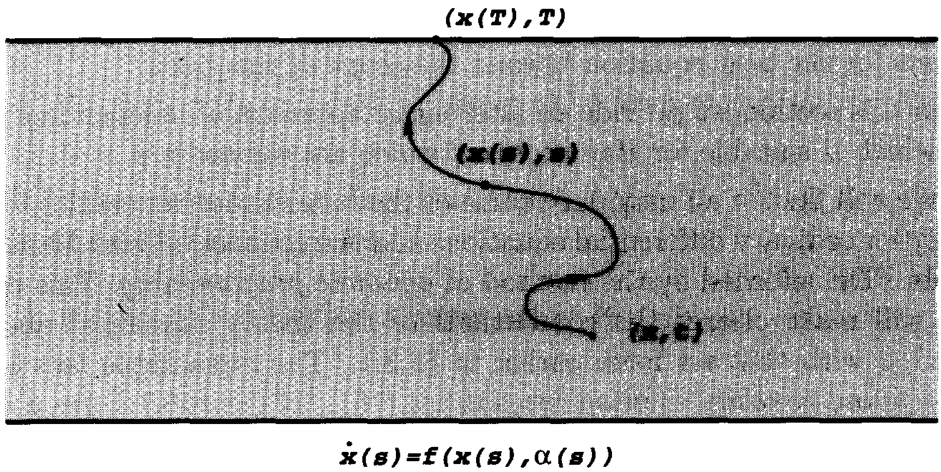
Let us write

$$
(2) \qquad \mathcal{A} = \{\boldsymbol{\alpha} : [0, T] \to A \mid \boldsymbol{\alpha}(\cdot) \text{ is measurable}\}
$$

to denote the set of *admissible controls.* Then since

$$
(3) \qquad |\mathbf{f}(x, a)| \leq C, \ |\mathbf{f}(x, a) - \mathbf{f}(y, a)| \leq C|x - y| \quad (x, y \in \mathbb{R}^n, \ a \in A)
$$

for some constant C, we see that for each control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$, the ODE (1) has a unique, Lipschitz continuous solution $\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$, existing on the time interval $[t, T]$ and solving the ODE for a.e. time $t < s < T$. We call $\mathbf{x}(\cdot)$ the *response* of the system to the control $\boldsymbol{\alpha}(\cdot)$, and $\mathbf{x}(s)$ the *state* of the system at time s.



Response of system to the control $\boldsymbol{\alpha}(\cdot)$

Our goal is to find a control $\boldsymbol{\alpha}^*(\cdot)$ which optimally steers the system. In order to define what "optimal" means however, we must first introduce a *cost criterion.* Given $x \in \mathbb{R}^n$ and $0 \le t \le T$, let us define for each admissible control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ the corresponding *cost*

$$C_{x,t}[\boldsymbol{\alpha}(\cdot)] := \int_t^T h(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)), \tag{4}$$

where $\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$ solves the ODE (1) and

$$h : \mathbb{R}^n \times A \to \mathbb{R}, \quad g : \mathbb{R}^n \to \mathbb{R}$$

are given functions. We call h the *running cost per unit time* and g the *terminal cost*, and will henceforth assume

$$\begin{cases} |h(x,a)|,\ |g(x)| \le C \\ |h(x,a) - h(y,a)|,\ |g(x) - g(y)| \le C|x-y| \end{cases} \quad (x, y \in \mathbb{R}^n, a \in A) \tag{5}$$

for some constant C.

Given now $x \in \mathbb{R}^n$ and $0 \le t \le T$, we would like to find if possible a control $\boldsymbol{\alpha}^*(\cdot)$ which minimizes the cost functional (4) among all other admissible controls.

10.3.2. Dynamic programming.

The method of *dynamic programming* investigates the above problem by turning attention to the *value function*

$$u(x,t) := \inf_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} C_{x,t}[\boldsymbol{\alpha}(\cdot)] \quad (x \in \mathbb{R}^n,\ 0 \le t \le T). \tag{6}$$

The plan is this: having defined $u(x,t)$ as the least cost given we start at the position x at time t, we want to study u as a function of x and t. We are therefore embedding our given control problem (1), (4) into the larger class of all such problems, as x and t vary. The idea then is to show that u solves a certain Hamilton–Jacobi type PDE, and finally to show conversely that a solution of this PDE helps us to synthesize an optimal feedback control.

Hereafter, we fix $x \in \mathbb{R}^n$, $0 \leq t < T$.

THEOREM 1 (Optimality conditions). *For each $h > 0$ so small that $t + h \leq T$, we have*

$$u(x,t) = \inf_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} \left\{ \int_t^{t+h} h(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + u(\mathbf{x}(t+h), t+h) \right\}, \tag{7}$$

where $\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$ solves the ODE (1) *for the control $\boldsymbol{\alpha}(\cdot)$.*

Proof. 1. Choose any control $\boldsymbol{\alpha}_1(\cdot) \in \mathcal{A}$ and solve the ODE

$$\begin{cases} \dot{\mathbf{x}}_1(s) = \mathbf{f}(\mathbf{x}_1(s), \boldsymbol{\alpha}_1(s)) & (t < s < t+h) \\ \mathbf{x}_1(t) = x. \end{cases} \tag{8}$$

Fix $\epsilon > 0$ and choose then $\boldsymbol{\alpha}_2(\cdot) \in \mathcal{A}$ so that

$$u(\mathbf{x}_1(t+h), t+h) + \epsilon \geq \int_{t+h}^T h(\mathbf{x}_2(s), \boldsymbol{\alpha}_2(s))\, ds + g(\mathbf{x}_2(T)), \tag{9}$$

where

$$\begin{cases} \dot{\mathbf{x}}_2(s) = \mathbf{f}(\mathbf{x}_2(s), \boldsymbol{\alpha}_2(s)) & (t+h < s < T) \\ \mathbf{x}_2(t+h) = \mathbf{x}_1(t+h). \end{cases} \tag{10}$$

Now define the control

$$\boldsymbol{\alpha}_3(s) := \begin{cases} \boldsymbol{\alpha}_1(s) & \text{if } t \leq s < t+h \\ \boldsymbol{\alpha}_2(s) & \text{if } t+h \leq s \leq T, \end{cases} \tag{11}$$

and let

$$\begin{cases} \dot{\mathbf{x}}_3(s) = \mathbf{f}(\mathbf{x}_3(s), \boldsymbol{\alpha}_3(s)) & (t < s < T) \\ \mathbf{x}_3(t) = x. \end{cases} \tag{12}$$

By uniqueness of solutions to the differential equation (1), we have

$$\mathbf{x}_3(s) = \begin{cases} \mathbf{x}_1(s) & \text{if } t \leq s \leq t+h \\ \mathbf{x}_2(s) & \text{if } t+h \leq s \leq T. \end{cases} \tag{13}$$

Thus the definition (6) implies

$$\begin{aligned} u(x,t) &\le C_{x,t}[\alpha_3(\cdot)] \\ &= \int_t^T h(\mathbf{x}_3(s),\alpha_3(s))\,ds + g(\mathbf{x}_3(T)) \\ &= \int_t^{t+h} h(\mathbf{x}_1(s),\alpha_1(s))\,ds + \int_{t+h}^T h(\mathbf{x}_2(s),\alpha_2(s))\,ds + g(\mathbf{x}_2(T)) \\ &\le \int_t^{t+h} h(\mathbf{x}_1(s),\alpha_1(s))\,ds + u(\mathbf{x}_1(t+h),t+h) + \epsilon, \end{aligned}$$

the last inequality resulting from (9). As $\alpha_1(\cdot) \in \mathcal{A}$ was arbitrary, we conclude

$$(14) \qquad u(x,t) \le \inf_{\alpha(\cdot)\in\mathcal{A}} \left\{ \int_t^{t+h} h(\mathbf{x}(s),\alpha(s))\,ds + u(\mathbf{x}(t+h),t+h) \right\} + \epsilon,$$

$\mathbf{x}(\cdot) = \mathbf{x}^{\alpha(\cdot)}(\cdot)$ solving (1).

2. Fixing again $\epsilon > 0$, select now $\alpha_4(\cdot) \in \mathcal{A}$ so that

$$(15) \qquad u(x,t) + \epsilon \ge \int_t^T h(\mathbf{x}_4(s),\alpha_4(s))ds + g(\mathbf{x}_4(T)),$$

where

$$\begin{cases} \dot{\mathbf{x}}_4(s) = \mathbf{f}(\mathbf{x}_4(s),\alpha_4(s)) \quad (t < s < T) \\ \mathbf{x}_4(t) = x. \end{cases}$$

Observe then from (6) that

$$(16) \qquad u(\mathbf{x}_4(t+h),t+h) \le \int_{t+h}^T h(\mathbf{x}_4(s),\alpha_4(s))ds + g(\mathbf{x}_4(T)).$$

Therefore

$$u(x,t) + \epsilon \ge \inf_{\alpha(\cdot)\in\mathcal{A}} \left\{ \int_t^{t+h} h(\mathbf{x}(s),\alpha(s))ds + u(\mathbf{x}(t+h),t+h) \right\},$$

$\mathbf{x}(\cdot) = \mathbf{x}^{\alpha(\cdot)}(\cdot)$ solving (1). This inequality and (14) complete the proof of (7). □

10.3.3. Hamilton–Jacobi–Bellman equation.

Our eventual goal is writing down as a PDE an "infinitesimal version" of the optimality conditions (7). But first we must check that the value function u is bounded and Lipschitz continuous.

LEMMA (Estimates for value function). *There exists a constant C such that*

$$|u(x,t)| \le C,$$
$$|u(x,t) - u(\hat{x},\hat{t})| \le C(|x-\hat{x}| + |t-\hat{t}|)$$

for all $x, \hat{x} \in \mathbb{R}^n$, $0 \le t, \hat{t} \le T$.

Proof. 1. Clearly hypothesis (5) implies u is bounded on $\mathbb{R}^n \times [0,T]$.

2. Fix $x, \hat{x} \in \mathbb{R}^n$, $0 \le t < T$. Let $\epsilon > 0$ and then choose $\hat{\boldsymbol{\alpha}}(\cdot) \in \mathcal{A}$ so that

$$u(\hat{x},t) + \epsilon \ge \int_t^T h(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))\, ds + g(\hat{\mathbf{x}}(T)), \tag{17}$$

where $\hat{\mathbf{x}}(\cdot)$ solves the ODE

$$\begin{cases} \dot{\hat{\mathbf{x}}}(s) = \mathbf{f}(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s)) & (t < s < T) \\ \hat{\mathbf{x}}(t) = \hat{x}. \end{cases} \tag{18}$$

Then

$$\begin{aligned} u(x,t) - u(\hat{x},t) \le & \int_t^T h(\mathbf{x}(s), \hat{\boldsymbol{\alpha}}(s))\, ds + g(\mathbf{x}(T)) \\ & - \int_t^T h(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))\, ds - g(\hat{\mathbf{x}}(T)) + \epsilon, \end{aligned} \tag{19}$$

where $\mathbf{x}(\cdot)$ solves

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \hat{\boldsymbol{\alpha}}(s)) & (t < s < T) \\ \mathbf{x}(t) = x. \end{cases} \tag{20}$$

Since $\mathbf{f}$ is Lipschitz continuous, (18), (20) and Gronwall's inequality (§B.2) imply $|\mathbf{x}(s) - \hat{\mathbf{x}}(s)| \le C|x - \hat{x}|$ $(t \le s \le T)$. Hence we deduce from (5) and (19) that $u(x,t) - u(\hat{x},t) \le C|x-\hat{x}| + \epsilon$. The same argument with the roles of x and $\hat{x}$ reversed implies

$$|u(x,t) - u(\hat{x},t)| \le C|x - \hat{x}| \quad (x, \hat{x} \in \mathbb{R}^n,\ 0 \le t \le T).$$

3. Now let $x \in \mathbb{R}^n$, $0 \le t < \hat{t} \le T$. Take $\epsilon > 0$ and choose $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ so that

$$u(x,t) + \epsilon \ge \int_t^T h(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)),$$

$\mathbf{x}(\cdot)$ solving the ODE (1). Define

$$\hat{\boldsymbol{\alpha}}(s) := \boldsymbol{\alpha}(s + t - \hat{t}) \ \text{ for } \hat{t} \le s \le T$$

and let $\hat{\mathbf{x}}(\cdot)$ solve

$$\begin{cases} \dot{\hat{\mathbf{x}}}(s) = \mathbf{f}(\hat{\mathbf{x}}(s), \boldsymbol{\alpha}(s)) & (\hat{t} < s < T) \\ \hat{\mathbf{x}}(\hat{t}) = x. \end{cases}$$

Then $\hat{\mathbf{x}}(s) = \mathbf{x}(s + t - \hat{t})$. Hence

$$\begin{aligned} u(x,\hat{t}) - u(x,t) &\leq \int_{\hat{t}}^{T} h(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))ds + g(\hat{\mathbf{x}}(T)) \\ &\quad - \int_{t}^{T} h(\mathbf{x}(s), \boldsymbol{\alpha}(s))ds - g(\mathbf{x}(T)) + \epsilon \qquad (21) \\ &= - \int_{T+t-\hat{t}}^{T} h(\mathbf{x}(s), \boldsymbol{\alpha}(s))ds + g(\mathbf{x}(T+t-\hat{t})) - g(\mathbf{x}(T)) + \epsilon \\ &\leq C|t - \hat{t}| + \epsilon. \end{aligned}$$

Next pick $\hat{\boldsymbol{\alpha}}(\cdot)$ so that

$$u(x,\hat{t}) + \epsilon \geq \int_{\hat{t}}^{T} h(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))ds + g(\hat{\mathbf{x}}(T)),$$

where

$$\begin{cases} \dot{\hat{\mathbf{x}}}(s) = \mathbf{f}(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s)) & (\hat{t} < s < T) \\ \hat{\mathbf{x}}(\hat{t}) = x. \end{cases}$$

Define

$$\boldsymbol{\alpha}(s) := \begin{cases} \hat{\boldsymbol{\alpha}}(s + \hat{t} - t) & \text{if } t \leq s \leq T + t - \hat{t} \\ \hat{\boldsymbol{\alpha}}(T) & \text{if } T + t - \hat{t} \leq s \leq T, \end{cases}$$

and let $\mathbf{x}(\cdot)$ solve (1). Then $\boldsymbol{\alpha}(s) = \hat{\boldsymbol{\alpha}}(s + \hat{t} - t)$, $\mathbf{x}(s) = \hat{\mathbf{x}}(s + \hat{t} - t)$ for $t \leq s \leq T + t - \hat{t}$. Consequently

$$\begin{aligned} u(x,t) - u(x,\hat{t}) &\leq \int_{t}^{T} h(\mathbf{x}(s), \boldsymbol{\alpha}(s))ds + g(\mathbf{x}(T)) \\ &\quad - \int_{\hat{t}}^{T} h(\hat{\mathbf{x}}(s), \boldsymbol{\alpha}(s))ds - g(\hat{\mathbf{x}}(T)) + \epsilon \\ &= \int_{T+t-\hat{t}}^{T} h(\mathbf{x}(s), \boldsymbol{\alpha}(s))ds + g(\mathbf{x}(T)) - g(\mathbf{x}(T+t-\hat{t})) + \epsilon \\ &\leq C|t - \hat{t}| + \epsilon. \end{aligned}$$

This inequality and (21) prove

$$|u(x,t) - u(x,\hat{t})| \leq C|t - \hat{t}| \quad (0 \leq t \leq \hat{t} \leq T,\ x \in \mathbb{R}^n).$$

□

We prove next that the value function solves a Hamilton–Jacobi type partial differential equation.

THEOREM 2 (A PDE for the value function). *The value function u is the unique viscosity solution of this terminal-value problem for the* Hamilton–Jacobi–Bellman equation*:*

$$
(22) \qquad \begin{cases} u_t + \min_{a \in A} \{\mathbf{f}(x,a) \cdot Du + h(x,a)\} = 0 & \textit{in } \mathbb{R}^n \times (0,T) \\ u = g & \textit{on } \mathbb{R}^n \times \{t = T\}. \end{cases}
$$

Remarks. (i) The Hamilton–Jacobi–Bellman PDE has the form

$$
u_t + H(Du, x) = 0 \quad \text{in } \mathbb{R}^n \times (0,T),
$$

for the Hamiltonian

$$
(23) \qquad H(p,x) := \min_{a \in A} \{\mathbf{f}(x,a) \cdot p + h(x,a)\} \quad (p, x \in \mathbb{R}^n).
$$

From the inequalities (5), we deduce that H satisfies the estimates (3) in §10.2.

(ii) Since (22) is a *terminal-value problem*, we must specify what we mean by a solution. Let us say that a bounded, uniformly continuous function u is a *viscosity solution* of (22) provided:

(i) $u = g$ on $\mathbb{R}^n \times \{t = T\}$,

and

(ii) for each $v \in C^\infty(\mathbb{R}^n \times (0,T))$

$$
(24) \qquad \begin{cases} \text{if } u - v \text{ has a local maximum at a point } (x_0, t_0) \in \mathbb{R}^n \times (0,T), \\ \text{then} \\ \qquad v_t(x_0,t_0) + H(Dv(x_0,t_0), x_0) \geq 0, \end{cases}
$$

and

$$
(25) \qquad \begin{cases} \text{if } u - v \text{ has a local minimum at a point } (x_0, t_0) \in \mathbb{R}^n \times (0,T), \\ \text{then} \\ \qquad v_t(x_0,t_0) + H(Dv(x_0,t_0), x_0) \leq 0. \end{cases}
$$

Observe that for our terminal-value problem (22) we *reverse* the sense of the inequalities from those for the initial-value problem.

(iii) The reader should check that if u is the viscosity solution of (22), then $w(x,t) := u(x, T-t)$ $(x \in \mathbb{R}^n, 0 \leq t \leq T)$ is the viscosity solution of the initial-value problem

$$
\begin{cases} w_t - H(Dw, x) = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ w = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}
$$

□

Proof. 1. In view of the lemma, u is bounded and Lipschitz continuous. In addition, we see directly from (4) and (6) that

$$u(x,T) = \inf_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} C_{x,T}[\boldsymbol{\alpha}(\cdot)] = g(x) \quad (x \in \mathbb{R}^n).$$

2. Now let $v \in C^\infty(\mathbb{R}^n \times (0,T))$, and assume

$$u - v \text{ has a local maximum at a point } (x_0, t_0) \in \mathbb{R}^n \times (0,T).$$

We must prove

$$v_t(x_0,t_0) + \min_{a\in A}\{\mathbf{f}(x_0,a)\cdot Dv(x_0,t_0) + h(x_0,a)\} \geq 0. \tag{26}$$

Suppose not. Then there exist $a \in A$ and $\theta > 0$ such that

$$v_t(x,t) + \mathbf{f}(x,a)\cdot Dv(x,t) + h(x,a) \leq -\theta < 0 \tag{27}$$

for all points (x,t) sufficiently close to (x_0,t_0), say

$$|x - x_0| + |t - t_0| < \delta. \tag{28}$$

Since $u - v$ has a local maximum at (x_0,t_0), we may as well also suppose

$$\begin{cases} (u-v)(x,t) \leq (u-v)(x_0,t_0) \\ \text{for all } (x,t) \text{ satisfying (28).} \end{cases} \tag{29}$$

Consider now the constant control $\boldsymbol{\alpha}(s) \equiv a$ $(t_0 \leq s \leq T)$ and the corresponding dynamics

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s),a) & (t_0 < s < T) \\ \mathbf{x}(t_0) = x. \end{cases} \tag{30}$$

Choose $0 < h < \delta$ so small that $|\mathbf{x}(s) - x_0| < \delta$ for $t_0 \leq s \leq t_0 + h$. Then

$$v_t(\mathbf{x}(s),s) + \mathbf{f}(\mathbf{x}(s),a)\cdot Dv(\mathbf{x}(s),s) + h(\mathbf{x}(s),a) \leq -\theta \quad (t_0 \leq s \leq t_0+h), \tag{31}$$

according to (27), (28). But utilizing (29) we find

$$\begin{aligned} &u(\mathbf{x}(t_0+h),t_0+h) - u(x_0,t_0) \leq v(\mathbf{x}(t_0+h),t_0+h) - v(x_0,t_0) \\ &= \int_{t_0}^{t_0+h} \frac{d}{ds}v(\mathbf{x}(s),s)\,ds = \int_{t_0}^{t_0+h} v_t(\mathbf{x}(s),s) + Dv(\mathbf{x}(s),s)\cdot\dot{\mathbf{x}}(s)\,ds \\ &= \int_{t_0}^{t_0+h} v_t(\mathbf{x}(s),s) + \mathbf{f}(\mathbf{x}(s),a)\cdot Dv(\mathbf{x}(s),s)\,ds. \end{aligned} \tag{32}$$

In addition, the optimality condition (7) provides us with the inequality

$$u(x_0,t_0) \le \int_{t_0}^{t_0+h} h(\mathbf{x}(s),a)\,ds + u(\mathbf{x}(t_0+h),t_0+h). \tag{33}$$

Combining (32) and (33), we discover

$$0 \le \int_{t_0}^{t_0+h} v_t(\mathbf{x}(s),s) + \mathbf{f}(\mathbf{x}(s),a)\cdot Dv(\mathbf{x}(s),s) + h(\mathbf{x}(s),a)\,ds \le -\theta h,$$

according to (31). This contradiction establishes (26).

3. Now suppose

$$u - v \text{ has a local minimum at a point } (x_0,t_0) \in \mathbb{R}^n \times (0,T);$$

we must prove

$$v_t(x_0,t_0) + \min_{a\in A}\{\mathbf{f}(x_0,a)\cdot Dv(x_0,t_0) + h(x_0,a)\} \le 0. \tag{34}$$

Suppose not. Then there exists $\theta > 0$ such that

$$v_t(x,t) + \mathbf{f}(x,a)\cdot Dv(x,t) + h(x,a) \ge \theta > 0 \tag{35}$$

for all $a \in A$ and all (x,t) sufficiently close to (x_0,t_0), say

$$|x-x_0| + |t-t_0| < \delta. \tag{36}$$

Since $u-v$ has a local minimum at (x_0,t_0), we may as well also suppose

$$\begin{cases} (u-v)(x,t) \ge (u-v)(x_0,t_0) \\ \text{for all } (x,t) \text{ satisfying (36).} \end{cases} \tag{37}$$

Choose $0 < h < \delta$ so small that $|\mathbf{x}(s) - x_0| < \delta$ for $t_0 \le s \le t_0 + h$, where $\mathbf{x}(\cdot)$ solves

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s),\boldsymbol{\alpha}(s)) & (t_0 < s < T) \\ \mathbf{x}(t_0) = x_0 \end{cases} \tag{38}$$

for some control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$. This is possible owing to hypothesis (3).

Then utilizing (37), we find for any control $\boldsymbol{\alpha}(\cdot)$ that

$$\begin{aligned} u(\mathbf{x}(t_0+h),\, t_0+h) &- u(x_0,t_0) \\ &\ge v(\mathbf{x}(t_0+h),t_0+h) - v(x_0,t_0) \\ &= \int_{t_0}^{t_0+h} \frac{d}{ds} v(\mathbf{x}(s),s)\,ds \\ &= \int_{t_0}^{t_0+h} v_t(\mathbf{x}(s),s) + \mathbf{f}(\mathbf{x}(s),\boldsymbol{\alpha}(s))\cdot Dv(\mathbf{x}(s),s)\,ds, \end{aligned} \tag{39}$$

by (38). On the other hand, according to the optimality condition (7) we can select a control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ so that

$$u(x_0, t_0) \geq \int_{t_0}^{t_0+h} h(\mathbf{x}(s), \boldsymbol{\alpha}(s))\,ds + u(\mathbf{x}(t_0+h), t_0+h) - \frac{\theta h}{2}. \tag{40}$$

Combining (39) and (40), we discover

$$\begin{aligned}\frac{\theta h}{2} \geq \int_{t_0}^{t_0+h} & v_t(\mathbf{x}(s), s) + \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) \cdot Dv(\mathbf{x}(s), s) \\ & + h(\mathbf{x}(s), \boldsymbol{\alpha}(s))\,ds \geq \theta h,\end{aligned}$$

according to (35). This contradiction proves (34). □

Remark (Design of optimal controls). We have now shown that the value function u, defined by (6), is the unique viscosity solution of the terminal-value problem (22) for the Hamilton–Jacobi–Bellman equation. How does this PDE help us solve the problem of synthesizing an optimal control? In informal terms, the method is this. Given an initial time $0 < t \leq T$ and an initial state $x \in \mathbb{R}^n$, we consider the optimal ODE

$$\begin{cases} \dot{\mathbf{x}}^*(s) = \mathbf{f}(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s)) & (t < s < T) \\ \mathbf{x}^*(t) = x, \end{cases} \tag{41}$$

where at each time s, $\boldsymbol{\alpha}^*(s) \in A$ is selected so that

$$\begin{aligned}\mathbf{f}(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s)) \cdot Du(\mathbf{x}^*(s), s) &+ h(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s)) \\ &= H(Du(\mathbf{x}^*(s), s), \mathbf{x}^*(s)).\end{aligned} \tag{42}$$

In other words, given the system is at the point $\mathbf{x}^*(s)$ at time s, we adjust the optimal control value $\boldsymbol{\alpha}^*(s)$ so as to attain the minimum in the definition (23) of the Hamiltonian H. We call $\boldsymbol{\alpha}^*(\cdot)$ so defined a *feedback control.*

It is fairly easy to check that this prescription does in fact generate a minimum cost trajectory, at least in regions where u and $\boldsymbol{\alpha}^*(\cdot)$ are smooth (so that (42) makes sense). There are however problems in interpreting (42) at points where the gradient Du does not exist. □

10.3.4. Hopf–Lax formula revisited.

Remember that earlier in §3.3 we investigated this initial-value problem for the Hamilton–Jacobi equation:

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0,T] \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{43}$$

under the assumptions that

$$p \mapsto H(p) \text{ is convex, } \lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty,$$

and

$$g : \mathbb{R}^n \to \mathbb{R} \text{ is Lipschitz continuous.}$$

Notice that we are now taking $0 \leq t \leq T$, to be consistent with §10.2. We introduced as well the Hopf–Lax formula for a solution:

$$u(x,t) = \min_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \quad (x \in \mathbb{R}^n, t > 0), \tag{44}$$

where L is the Legendre transform of H:

$$L(q) = \sup_{p\in\mathbb{R}^n} \{p \cdot q - H(p)\} \quad (q \in \mathbb{R}^n). \tag{45}$$

In order to tie together the theory set forth here and in §3.3, let us now check that the Hopf–Lax formula gives the correct viscosity solution, as defined in §10.1.1. (The proof is really just a special case of that for Theorem 2.)

THEOREM 3 (Hopf–Lax formula as viscosity solution). *Assume in addition that g is bounded. Then the unique viscosity solution of the initial-value problem* (43) *is given by the formula* (44).

Proof. 1. As shown in §3.3 the function u defined by (44) is Lipschitz continuous and takes on the initial function g at time $t = 0$. It is easy to verify as well that u is also bounded on $\mathbb{R}^n \times (0,T]$, since g is bounded.

2. Now let $v \in C^\infty(\mathbb{R}^n \times (0,\infty))$ and assume $u - v$ has a local maximum at $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$. According to Lemma 1 in §3.3.2,

$$u(x_0,t_0) = \min_{x\in\mathbb{R}^n} \left\{ (t_0 - t)L\left(\frac{x_0 - x}{t_0 - t}\right) + u(x,t) \right\} \tag{46}$$

for each $0 \leq t < t_0$. Thus for each $0 \leq t < t_0$, $x \in \mathbb{R}^n$

$$u(x_0,t_0) \leq (t_0 - t)L\left(\frac{x_0 - x}{t_0 - t}\right) + u(x,t). \tag{47}$$

But since $u - v$ has a local maximum at (x_0, t_0),

$$u(x_0,t_0) - v(x_0,t_0) \geq u(x,t) - v(x,t)$$

for (x,t) close to (x_0,t_0). Combining this estimate with (47), we find

$$v(x_0,t_0) - v(x,t) \leq (t_0 - t)L\left(\frac{x_0 - x}{t_0 - t}\right) \tag{48}$$

for $t < t_0$, (x,t) close to (x_0,t_0). Now write $h = t_0 - t$ and set $x = x_0 - hq$, where $q \in \mathbb{R}^n$ is given. Inequality (48) becomes

$$v(x_0,t_0) - v(x_0 - hq, t_0 - h) \leq hL(q).$$

Divide by $h > 0$ and send $h \to 0$:

$$v_t(x_0,t_0) + Dv(x_0,t_0)\cdot q - L(q) \leq 0.$$

This is true for all $q \in \mathbb{R}^n$ and so

$$v_t(x_0,t_0) + H(Dv(x_0,t_0)) \leq 0, \tag{49}$$

since

$$H(p) = \sup_{q\in\mathbb{R}^n}\{p\cdot q - L(q)\}, \tag{50}$$

by the convex duality of H and L. We have, as desired, established the inequality (49) whenever $u - v$ has a local maximum at (x_0,t_0).

3. Now suppose instead $u - v$ has a local minimum at a point $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. We must prove

$$v_t(x_0,t_0) + H(Dv(x_0,t_0)) \geq 0. \tag{51}$$

Suppose to the contrary estimate (51) fails; in which case

$$v_t(x,t) + H(Dv(x,t)) \leq -\theta < 0$$

for some $\theta > 0$ and all points (x,t) close enough to (x_0,t_0). In view of (50)

$$v_t(x,t) + Dv(x,t)\cdot q - L(q) \leq -\theta \tag{52}$$

for all (x,t) near (x_0,t_0) and all $q \in \mathbb{R}^n$.

Now from (46) we see that if $h > 0$ is small enough

$$u(x_0,t_0) = hL\left(\frac{x_0 - x_1}{h}\right) + u(x_1, t_0 - h) \tag{53}$$

for some point x_1 close to x_0. We then compute

$$\begin{aligned} v(x_0,t_0) - v(x_1,t_0-h) &= \int_0^1 \frac{d}{ds} v(sx_0+(1-s)x_1,\ t_0+(s-1)h)\,ds \\ &= \int_0^1 Dv(sx_0+(1-s)x_1,\ t_0+(s-1)h)\cdot(x_0-x_1) \\ &\quad + v_t(sx_0+(1-s)x_1,\ t_0+(s-1)h)h\,ds \\ &= h\left(\int_0^1 Dv(\cdots)\cdot q + v_t(\cdots)\,ds\right), \end{aligned}$$

where $q = \frac{x_0-x_1}{h}$. Now if $h > 0$ is sufficiently small, we may apply (52), to find

$$v(x_0,t_0) - v(x_1,t_0-h) \le hL\left(\frac{x_0-x_1}{h}\right) - \theta h.$$

But then (53) forces

$$v(x_0,t_0) - v(x_1,t_0-h) \le u(x_0,t_0) - u(x_1,t_0-h) - \theta h,$$

a contradiction, since $u - v$ has a local minimum at (x_0,t_0). Consequently the desired inequality (51) is indeed valid. □

10.4. PROBLEMS

1. Let $\{u^k\}_{k=1}^\infty$ be viscosity solutions of the Hamilton–Jacobi equations

$$u_t^k + H(Du^k, x) = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty)$$

$(k = 1,\dots)$, and suppose $u^k \to u$ uniformly. Assume as well that H is continuous. Show u is a viscosity solution of

$$u_t + H(Du, x) = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty).$$

Hence the uniform limits of viscosity solutions is a viscosity solution.

2. Let u^i $(i = 1,2)$ be viscosity solutions of

$$\begin{cases} u_t^i + H(Du^i, x) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ \qquad\qquad\quad u^i = g^i & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Assume H satisfies condition (3) in §10.2. Prove the *contraction* property

$$\sup_{\mathbb{R}^n} |u^1(\cdot,t) - u^2(\cdot,t)| \le \sup_{\mathbb{R}^n} |g^1 - g^2| \quad (t \ge 0).$$

3. Suppose for each $\epsilon > 0$ that u^ϵ is a smooth solution of the parabolic equation

$$u_t^\epsilon + H(Du^\epsilon, x) - \epsilon \sum_{i,j=1}^n a^{ij} u^\epsilon_{x_i x_j} = 0$$

in $\mathbb{R}^n \times (0,\infty)$, where the smooth coefficients a^{ij} $(i,j = 1,\dots,n)$ satisfy the uniform ellipticity condition from Chapter 6. Suppose also that H is continuous, and that $u^\epsilon \to u$ uniformly as $\epsilon \to 0$.

Prove that u is a viscosity solution of $u_t + H(Du,x) = 0$. (This exercise shows that viscosity solutions do not depend upon the precise structure of the parabolic smoothing.)

4. (i) Show that $u(x) := 1 - |x|$ is a viscosity solution of

$$(*) \qquad \begin{cases} |u'| = 1 & \text{in } (-1,1) \\ u(-1) = u(1) = 0. \end{cases}$$

This means that for each $v \in C^\infty(-1,1)$, if $u - v$ has a maximum (minimum) at a point $x_0 \in (-1,1)$, then $|v'(x_0)| \leq 1$ (≥ 1).

(ii) Show that $\tilde{u}(x) := |x| - 1$ is *not* a viscosity solution of $(*)$.

(iii) Show that $\tilde{u}$ is a viscosity solution of

$$(**) \qquad \begin{cases} -|\tilde{u}'| = -1 & \text{in } (-1,1) \\ \tilde{u}(-1) = \tilde{u}(1) = 0. \end{cases}$$

(Hint: What is the meaning of a viscosity solution of $(**)$?)

(iv) Why do problems $(*), (**)$ have different viscosity solutions?

5. Let $U \subset \mathbb{R}^n$ be open, bounded. Set $u(x) := \operatorname{dist}(x, \partial U)$ $(x \in U)$. Prove u is Lipschitz continuous and is a viscosity solution of the eikonal equation

$$|Du| = 1 \quad \text{in } U.$$

This means that for each $v \in C^\infty(U)$, if $u - v$ has a maximum (minimum) at a point $x_0 \in U$, then $|Dv(x_0)| \leq 1$ (≥ 1).

10.5. REFERENCES

Section 10.1 The definition of viscosity solutions presented here is due to **[C-E-L]**, who recast an earlier definition set forth in the basic paper **[C-L]** of Crandall–Lions.

Section 10.2 The uniqueness proof follows **[C-E-L]**, with some recent improvements I learned from M. Crandall.

Section 10.3 P.-L. Lions [**LI**] observed the connection between the definition of viscosity solution and the optimality conditions of control theory. The books of Fleming–Soner [**F-S**] and Bardi–Capuzzo Dolcetta [**B-CD**] provide much more information about viscosity solutions and the connections with deterministic and stochastic optimal control.

Chapter 11

SYSTEMS OF CONSERVATION LAWS

11.1. INTRODUCTION

In this final chapter we study *systems* of nonlinear, divergence structure first-order hyperbolic PDE, which arise as models of *conservation laws.*

Physical interpretation. In the most general circumstance we would like to investigate a vector function

$$\mathbf{u} = \mathbf{u}(x,t) = (u^1(x,t), \dots, u^m(x,t)) \quad (x \in \mathbb{R}^n,\ t \geq 0),$$

the components of which are the densities of various conserved quantities in some physical system under investigation. Given then any smooth, bounded region $U \subset \mathbb{R}^n$, we note that the integral

$$(1) \qquad \int_U \mathbf{u}(x,t)\,dx$$

represents the total amount of these quantities within U at time t. Now conservation laws typically assert that the rate of change within U is governed by a *flux* function $\mathbf{F} : \mathbb{R}^m \to \mathbb{M}^{m\times n}$, which controls the rate of loss or increase of $\mathbf{u}$ through ∂U. Otherwise stated, it is appropriate to assume for each time t

$$\frac{d}{dt}\int_U \mathbf{u}\,dx = -\int_{\partial U} \mathbf{F}(\mathbf{u})\nu\,dS, \tag{2}$$

ν denoting as usual the outward unit normal along U. Rewriting (2), we deduce

$$\int_U \mathbf{u}_t\,dx = -\int_{\partial U} \mathbf{F}(\mathbf{u})\nu\,dS = -\int_U \operatorname{div}\mathbf{F}(\mathbf{u})\,dx. \tag{3}$$

As the region $U \subset \mathbb{R}^n$ was arbitrary, we derive from (3) this initial-value problem for a *general system of conservation laws*:

$$\begin{cases} \mathbf{u}_t + \operatorname{div}\mathbf{F}(\mathbf{u}) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ \mathbf{u} = \mathbf{g} & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases} \tag{4}$$

the given function $\mathbf{g} = (g^1,\dots,g^m)$ describing the initial distribution of $\mathbf{u} = (u^1,\dots,u^m)$. □

At present a good mathematical understanding of problem (4) is largely unavailable (but see Zheng [**ZH**]). For this reason we shall henceforth consider instead the initial-value problem for a *system of conservation laws in one space dimension*:

$$\begin{cases} \mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ \mathbf{u} = \mathbf{g} & \text{on } \mathbb{R} \times \{t=0\}, \end{cases} \tag{5}$$

where $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m$ and $\mathbf{g} : \mathbb{R} \to \mathbb{R}^m$ are given and $\mathbf{u} : \mathbb{R} \times [0,\infty) \to \mathbb{R}^m$ is the unknown, $\mathbf{u} = \mathbf{u}(x,t)$. We call $\mathbb{R}^m$ the *state space*, and write

$$\mathbf{F} = \mathbf{F}(z) = (F^1(z),\dots,F^m(z)) \quad (z \in \mathbb{R}^m)$$

for the smooth flux function.

We intend to study the solvability of problem (5), properties of its solutions, etc.

Example 1. The *p-system* is this collection of two conservation laws:

$$\begin{cases} u^1_t - u^2_x = 0 & \text{(compatibility condition)} \\ u^2_t - p(u^1)_x = 0 & \text{(Newton's law)}, \end{cases} \tag{6}$$

in $\mathbb{R} \times (0, \infty)$, where $p : \mathbb{R} \to \mathbb{R}$ is given. Here

$$\mathbf{F}(z) = (-z_2, -p(z_1)) \tag{7}$$

for $z = (z_1, z_2)$. The p-system arises as a rewritten form of the scalar *nonlinear wave equation*

$$u_{tt} - (p(u_x))_x = 0 \quad \text{in } \mathbb{R} \times (0, \infty). \tag{8}$$

Taking $u^1 := u_x$, $u^2 := u_t$, we obtain the system (6), with the stated interpretations. □

Example 2. *Euler's equations* for compressible gas flow in one dimension are

$$\begin{cases} \rho_t + (\rho v)_x = 0 & \text{(conservation of mass)} \\ (\rho v)_t + (\rho v^2 + p)_x = 0 & \text{(conservation of momentum)} \\ (\rho E)_t + (\rho E v + p v)_x = 0 & \text{(conservation of energy)} \end{cases} \tag{9}$$

in $\mathbb{R} \times (0, \infty)$. Here ρ is the *mass density*, v the *velocity*, and E the *energy density* per unit mass. We assume

$$E = e + \frac{v^2}{2},$$

where e is the *internal energy* per unit mass and the term $\frac{v^2}{2}$ corresponds to the *kinetic energy* per unit mass. The letter p in (9) denotes the *pressure*. We assume p is a known function

$$p = p(\rho, e) \tag{10}$$

of ρ and e; formula (10) is a *constitutive relation*. Writing $\mathbf{u} = (u^1, u^2, u^3) = (\rho, \rho v, \rho E)$, we check (9) is a system of conservation laws of the requisite form

$$\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0 \quad \text{in } \mathbb{R} \times (0, \infty)$$

for $\mathbf{F} = (F^1, F^2, F^3)$,

$$\begin{cases} F^1(z) = z_2 \\ F^2(z) = \frac{(z_2)^2}{z_1} + p(z_1, \frac{z_3}{z_1} - \frac{1}{2}(\frac{z_2}{z_1})^2) \\ F^3(z) = \frac{z_2 z_3}{z_1} + p(z_1, \frac{z_3}{z_1} - \frac{1}{2}(\frac{z_2}{z_1})^2)\frac{z_2}{z_1}, \end{cases} \tag{11}$$

where $z = (z_1, z_2, z_3)$, $z_1 > 0$. □

Example 3. The one-dimensional *shallow water equations* are

$$\begin{cases} \phi_t + (v\phi)_x = 0 & \text{(conservation of mass)} \\ v_t + (v^2/2 + \phi)_x = 0 & \text{(conservation of momentum)} \end{cases}$$

in $\mathbb{R} \times [0, \infty)$. In these equations v is the horizontal velocity, ϕ denotes gh, where h is the height and g the gravitational constant, and

$$\mathbf{F}(z) = \left(z_1 z_2, \frac{(z_2)^2}{2} + z_1 \right).$$

□

11.1.1. Integral solutions.

The great difficulty in this subject is discovering a proper notion of weak solution for the initial-value problem (5). We have already encountered similar issues in §3.4 in our study of the much simpler case of a single or *scalar* conservation law (i.e., $m = 1$ above).

Following then the development in §3.4.1, let us suppose

$$\text{(12)} \qquad \begin{cases} \mathbf{v} : \mathbb{R} \times [0, \infty) \to \mathbb{R}^m \text{ is smooth,} \\ \text{with compact support, } \mathbf{v} = (v^1, \dots, v^m). \end{cases}$$

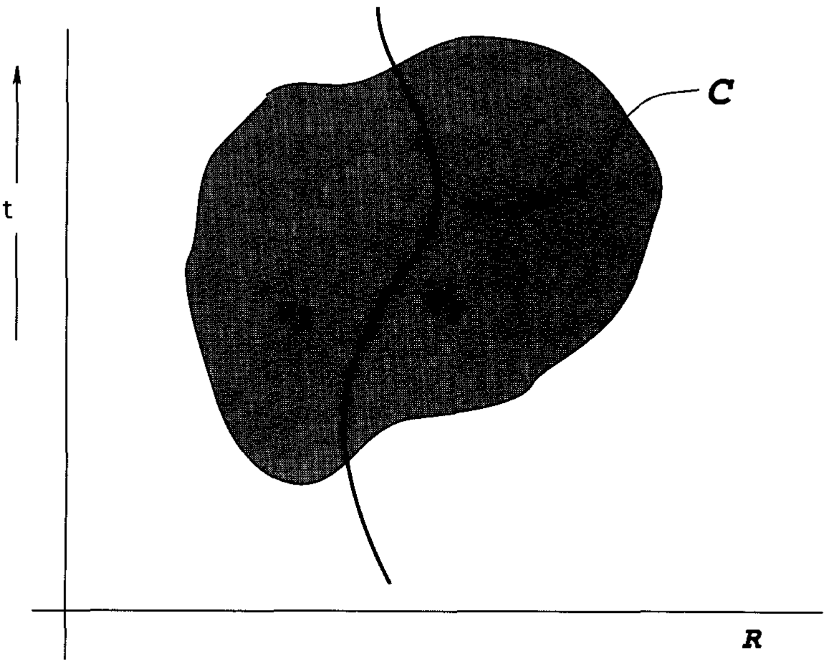
We temporarily assume $\mathbf{u}$ is a smooth solution of our problem (5), take the dot product of the PDE $\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0$ with the test function $\mathbf{v}$, and integrate by parts, to obtain the equality

$$\text{(13)} \qquad \int_0^\infty \int_{-\infty}^\infty \mathbf{u} \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}) \cdot \mathbf{v}_x \, dxdt + \int_{-\infty}^\infty \mathbf{g} \cdot \mathbf{v} \, dx|_{t=0} = 0.$$

This identity, which we derived supposing $\mathbf{u}$ to be a smooth solution, makes sense if $\mathbf{u}$ is merely bounded.

DEFINITION. *We say that* $\mathbf{u} \in L^\infty(\mathbb{R}^n \times (0, \infty); \mathbb{R}^m)$ *is an* integral solution *of the initial-value problem* (5) *provided the equality* (13) *holds for all test functions* $\mathbf{v}$ *satisfying* (12).

Continuing now to parallel the development in §3.4.1 for a single conservation law, let us now consider the situation that we have an integral solution $\mathbf{u}$ of (5) which is smooth on either side of a curve C, along which $\mathbf{u}$ has simple jump discontinuities. More precisely, let us assume that $V \subset \mathbb{R} \times (0, \infty)$ is some region cut by a smooth curve C into a left hand part V_l and a right hand part V_r.



Rankine-Hugoniot condition

Assuming that $\mathbf{u}$ is smooth in V_l, we select the test function $\mathbf{v}$ with compact support in V_l and deduce from (13) that

$$\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0 \quad \text{in } V_l. \tag{14}$$

Similarly, we have

$$\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0 \quad \text{in } V_r, \tag{15}$$

provided $\mathbf{u}$ is smooth in V_r. Now choose a test function $\mathbf{v}$ with compact support in V, but which does not necessarily vanish along the curve C. Then utilizing the identity (13) we find

$$\begin{aligned} 0 &= \int_0^\infty \int_{-\infty}^\infty \mathbf{u} \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}) \cdot \mathbf{v}_x \, dxdt \\ &= \iint_{V_l} \mathbf{u} \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}) \cdot \mathbf{v}_x \, dxdt + \iint_{V_r} \mathbf{u} \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}) \cdot \mathbf{v}_x \, dxdt. \end{aligned} \tag{16}$$

As $\mathbf{v}$ has compact support within V, we deduce

$$\begin{aligned} \iint_{V_l} \mathbf{u} \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}) \cdot \mathbf{v}_x \, dxdt &= -\iint_{V_l} [\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x] \cdot \mathbf{v} \, dxdt \\ &\quad + \int_C (\mathbf{u}_l \nu^2 + \mathbf{F}(\mathbf{u}_l)\nu^1) \cdot \mathbf{v} \, dl \\ &= \int_C (\mathbf{u}_l \nu^2 + \mathbf{F}(\mathbf{u}_l)\nu^1) \cdot \mathbf{v} \, dl, \end{aligned} \tag{17}$$

owing to (14). Here $\boldsymbol{\nu} = (\nu^1, \nu^2)$ is the unit normal to the curve C pointing from V_l into V_r, and the subscript "l" denotes the limit from the left. Likewise (16) implies

$$\iint_{V_r} \mathbf{u} \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}) \cdot \mathbf{v}_x \, dxdt = -\int_C (\mathbf{u}_r \nu^2 + \mathbf{F}(\mathbf{u}_r)\nu^1) \cdot \mathbf{v} \, dl,$$

"r" denoting the limit from the right. Adding this identity to (17) and remembering (16), we deduce

$$\int_C [(\mathbf{F}(\mathbf{u}_l) - \mathbf{F}(\mathbf{u}_r))\nu^1 + (\mathbf{u}_l - \mathbf{u}_r)\nu^2] \cdot \mathbf{v} \, dl = 0.$$

This identity obtains for all smooth functions $\mathbf{v}$ as above; whence

$$(\mathbf{F}(\mathbf{u}_l) - \mathbf{F}(\mathbf{u}_r))\nu^1 + (\mathbf{u}_l - \mathbf{u}_r)\nu^2 = 0 \text{ along } C. \tag{18}$$

Suppose now the curve C is represented parametrically as $\{(x, t) \mid x = s(t)\}$ for some smooth function $s(\cdot) : [0, \infty) \to \mathbb{R}$. Then $\boldsymbol{\nu} = (\nu^1, \nu^2) = (1 + \dot{s}^2)^{-1/2}(1, -\dot{s})$. Consequently (18) reads

$$\mathbf{F}(\mathbf{u}_l) - \mathbf{F}(\mathbf{u}_r) = \dot{s}(\mathbf{u}_l - \mathbf{u}_r) \tag{19}$$

in V along the curve C.

Notation.

$$\begin{cases} [[\mathbf{u}]] = \mathbf{u}_l - \mathbf{u}_r = \textit{jump} \text{ in } \mathbf{u} \text{ across the curve } C \\ [[\mathbf{F}(\mathbf{u})]] = \mathbf{F}(\mathbf{u}_l) - \mathbf{F}(\mathbf{u}_r) = \textit{jump} \text{ in } \mathbf{F}(\mathbf{u}) \\ \sigma = \dot{s} = \textit{speed} \text{ of the curve } C. \end{cases}$$

□

Let us then rewrite (19) as the identity

$$[[\mathbf{F}(\mathbf{u})]] = \sigma[[\mathbf{u}]] \tag{20}$$

along the discontinuity curve, the *Rankine–Hugoniot jump condition.* Note carefully this is a vector equality.

11.1.2. Traveling waves, hyperbolic systems.

We have seen in §3.4 that the notion of an integral solution for conservation laws is not adequate: such solutions need not be unique. We are therefore intent upon discovering some additional requirements for a good definition of a generalized solution. This will presumably entail as in §3.4

an entropy criterion based upon an analysis of shock waves. This expectation, now as carried over to systems, is largely correct, but first of all we must study more carefully the nonlinearity $\mathbf{F}$ in the hopes of discovering mathematically appropriate and physically correct structural conditions to impose.

Let us start by first considering the wider class of *semilinear systems* having the nondivergence form:

$$\mathbf{u}_t + \mathbf{B}(\mathbf{u})\mathbf{u}_x = 0 \quad \text{in } \mathbb{R} \times (0,\infty), \tag{21}$$

where $\mathbf{B} : \mathbb{R}^m \to \mathbb{M}^{m\times m}$. This system is for smooth functions equivalent to the conservation law in (5), provided

$$\mathbf{B} = D\mathbf{F} = \begin{pmatrix} F^1_{z_1} & \dots & F^1_{z_m} \\ & \ddots & \\ F^m_{z_1} & \dots & F^m_{z_m} \end{pmatrix}_{m\times m}.$$

We consider now the possibility of finding particular solutions which have the form of a traveling wave:

$$\mathbf{u}(x,t) = \mathbf{v}(x-\sigma t) \quad (x \in \mathbb{R}^n, t > 0), \tag{22}$$

where the profile $\mathbf{v} : \mathbb{R} \to \mathbb{R}^m$ and the velocity $\sigma \in \mathbb{R}$ are to be found. We substitute expression (22) into the PDE (21), and thereby obtain the equality

$$-\sigma\mathbf{v}'(x-\sigma t) + \mathbf{B}(\mathbf{v}(x-\sigma t))\mathbf{v}'(x-\sigma t) = 0. \tag{23}$$

Observe (23) says that σ is an eigenvalue of the matrix $\mathbf{B}(\mathbf{v})$ corresponding to the eigenvector $\mathbf{v}'$.

This conclusion suggests (exactly as for the linear theory in §7.3) that if we wish to find traveling waves or, more generally, wavelike solutions of our system of PDE, we should make some sort of hyperbolicity hypothesis concerning the eigenvalues of $\mathbf{B}$.

DEFINITION. *If for each $z \in \mathbb{R}^m$ the eigenvalues of $\mathbf{B}(z)$ are real and distinct, we call the system* (21) strictly hyperbolic.

We henceforth assume the system of partial differential equations (21) (and the special case $\mathbf{B} = D\mathbf{F}$ of conservation laws) to be strictly hyperbolic.

Notation. (i) We will write

$$\lambda_1(z) < \lambda_2(z) < \cdots < \lambda_m(z) \quad (z \in \mathbb{R}^m) \tag{24}$$

to denote the real and distinct eigenvalues of $\mathbf{B}(z)$, in increasing order.

(ii) Then for each $k = 1, \ldots, m$, we let

$$\mathbf{r}_k(z)$$

denote a corresponding nonzero eigenvector, so that

$$\mathbf{B}(z)\mathbf{r}_k(z) = \lambda_k(z)\mathbf{r}_k(z) \quad (k = 1, \ldots, m,\ z \in \mathbb{R}^m). \tag{25}$$

Since we are always assuming the strict hyperbolicity condition, the vectors $\{\mathbf{r}_k(z)\}_{k=1}^m$ span $\mathbb{R}^m$ for each $z \in \mathbb{R}^m$.

(iii) Next, since a matrix and its transpose have the same spectrum, we can introduce for each $k = 1, \ldots, m$ a nonzero eigenvector

$$\mathbf{l}_k(z)$$

for the matrix $\mathbf{B}(z)^T$, corresponding to the eigenvalue $\lambda_k(z)$. Thus

$$\mathbf{B}(z)^T\mathbf{l}_k(z) = \lambda_k(z)\mathbf{l}_k(z) \quad (k = 1, \ldots, m,\ z \in \mathbb{R}^m). \tag{26}$$

This equality is usually written

$$\mathbf{l}_k(z)\mathbf{B}(z) = \lambda(z)\mathbf{l}_k(z) \quad (k = 1, \ldots, m,\ z \in \mathbb{R}^m). \tag{27}$$

Thus $\{\mathbf{l}_k(z)\}_{k=1}^m$ can be regarded as *left eigenvectors* of $\mathbf{B}(z)$, and $\{\mathbf{r}_k(z)\}_{k=1}^\infty$ are *right eigenvectors*. □

Remark. Additionally, we observe

$$\mathbf{l}_l(z) \cdot \mathbf{r}_k(z) = 0 \ \text{ if } k \neq l \quad (z \in \mathbb{R}^m). \tag{28}$$

To confirm this, we compute using (25) and (26) that

$$\begin{aligned}\lambda_k(z)(\mathbf{l}_l(z) \cdot \mathbf{r}_k(z)) &= \mathbf{l}_l(z) \cdot (\mathbf{B}(z)\mathbf{r}_k(z)) = (\mathbf{B}(z)^T\mathbf{l}_l(z)) \cdot \mathbf{r}_k(z)\\ &= \lambda_l(z)(\mathbf{l}_l(z) \cdot \mathbf{r}_k(z));\end{aligned}$$

whence (28) follows since $\lambda_k(z) \neq \lambda_l(z)$ if $k \neq l$. □

Let us first show that the notion of strict hyperbolicity is independent of coordinates.

THEOREM 1 (Invariance of hyperbolicity under change of coordinates). *Let* $\mathbf{u}$ *be a smooth solution of the strictly hyperbolic system* (21). *Assume also* $\mathbf{\Phi} : \mathbb{R}^m \to \mathbb{R}^m$ *is a smooth diffeomorphism, with inverse* $\mathbf{\Psi}$. *Then*

$$\tilde{\mathbf{u}} := \mathbf{\Phi}(\mathbf{u}) \tag{29}$$

solves the strictly hyperbolic system

$$\tilde{\mathbf{u}}_t + \tilde{\mathbf{B}}(\tilde{\mathbf{u}})\tilde{\mathbf{u}}_x = 0 \quad \text{in } \mathbb{R} \times (0, \infty), \tag{30}$$

for

$$\tilde{\mathbf{B}}(\tilde{z}) := D\mathbf{\Phi}(\mathbf{\Psi}(\tilde{z}))\mathbf{B}(\mathbf{\Psi}(\tilde{z}))D\mathbf{\Psi}(\tilde{z}) \quad (\tilde{z} \in \mathbb{R}^m). \tag{31}$$

Proof. 1. We compute $\tilde{\mathbf{u}}_t = D\mathbf{\Phi}(\mathbf{u})\mathbf{u}_t$, $\tilde{\mathbf{u}}_x = D\mathbf{\Phi}(\mathbf{u})\mathbf{u}_x$, and so equation (30) is valid for $\mathbf{B}(\tilde{z}) = D\mathbf{\Phi}(z)\mathbf{B}(z)D\mathbf{\Phi}^{-1}(z)$, where $\tilde{z} = \mathbf{\Phi}(z)$. Substituting $z = \mathbf{\Psi}(\tilde{z})$, we obtain (31).

2. We must prove the system (30) is strictly hyperbolic. If $\lambda_k(z)$ is an eigenvalue of $\mathbf{B}(z)$, with corresponding right eigenvector $\mathbf{r}_k(z)$, we have

$$\mathbf{B}(z)\mathbf{r}_k(z) = \lambda_k(z)\mathbf{r}_k(z).$$

Setting

$$\tilde{\mathbf{r}}_k(\tilde{z}) := D\mathbf{\Phi}(\mathbf{\Psi}(\tilde{z}))\mathbf{r}_k(\mathbf{\Psi}(\tilde{z})), \tag{32}$$

$$\tilde{\lambda}_k(\tilde{z}) := \lambda_k(\mathbf{\Psi}(\tilde{z})), \tag{33}$$

we compute

$$\tilde{\mathbf{B}}(\tilde{z})\tilde{\mathbf{r}}_k(\tilde{z}) = \tilde{\lambda}_k(\tilde{z})\tilde{\mathbf{r}}_k(\tilde{z}). \tag{34}$$

Similarly if $\mathbf{l}_k(z)$ is a left eigenvector, we write

$$\tilde{\mathbf{l}}_k(\tilde{z}) := \mathbf{l}_k(\mathbf{\Psi}(\tilde{z}))D\mathbf{\Psi}(\tilde{z}), \tag{35}$$

and calculate

$$\tilde{\mathbf{l}}_k(\tilde{z})\tilde{\mathbf{B}}(\tilde{z}) = \tilde{\lambda}_k(\tilde{z})\tilde{\mathbf{l}}_k(\tilde{z}). \tag{36}$$

In view of (32)–(36), we conclude that the system (30) is strictly hyperbolic. □

Next we study how $\lambda_k(z)$, $\mathbf{r}_k(z)$ and $\mathbf{l}_k(z)$ change as z varies:

THEOREM 2 (Dependence of eigenvalues and eigenvectors on parameters). *Assume the matrix function* $\mathbf{B}$ *is smooth, strictly hyperbolic. Then*

(i) *the eigenvalues* $\lambda_k(z)$ *depend smoothly on* $z \in \mathbb{R}^m$ $(k = 1, \dots, m)$.

(ii) *Furthermore, we can select the right eigenvectors* $\mathbf{r}_k(z)$ *and left eigenvectors* $\mathbf{l}_k(z)$ *to depend smoothly on* $z \in \mathbb{R}^m$ *and satisfy the normalization*

$$|\mathbf{r}_k(z)|,\ |\mathbf{l}_k(z)| = 1 \quad (k = 1, \dots, m).$$

Proof. 1. Since $\mathbf{B}(z)$ is strictly hyperbolic, for each $z_0 \in \mathbb{R}^m$ we have

$$\lambda_1(z_0) < \lambda_2(z_0) < \cdots < \lambda_m(z_0). \tag{37}$$

Fix $k \in \{1, \dots, m\}$ and any point $z_0 \in \mathbb{R}^m$, and let $\mathbf{r}_k(z_0)$ satisfy

$$\begin{cases} B(z_0)\mathbf{r}_k(z_0) = \lambda_k(z_0)\mathbf{r}_k(z_0) \\ \quad |\mathbf{r}_k(z_0)| = 1. \end{cases}$$

Upon rotating coordinates if necessary, we may assume

$$\mathbf{r}_k(z_0) = e_m = (0, \dots, 1). \tag{38}$$

We first show that near z_0, there exist smooth functions $\lambda_k(z), \mathbf{r}_k(z)$ such that

$$\begin{cases} \mathbf{B}(z)\mathbf{r}_k(z) = \lambda_k(z)\mathbf{r}_k(z) \\ \quad |\mathbf{r}_k(z)| = 1. \end{cases}$$

2. We will apply the Implicit Function Theorem to the smooth function $\Phi : \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^{m+1}$ defined by

$$\Phi(r, \lambda, z) = (\mathbf{B}(z)r - \lambda r, |r|^2) \quad (r, z \in \mathbb{R}^m,\ \lambda \in \mathbb{R}).$$

Now

$$\frac{\partial \Phi(r, \lambda, z)}{\partial(r, \lambda)} = \begin{pmatrix} \mathbf{B}(z) - \lambda I & \begin{matrix} -r_1 \\ \vdots \\ -r_m \end{matrix} \\ 2r_1 \dots 2r_m & 0 \end{pmatrix}_{(m+1)\times(m+1)},$$

and so, according to (38), it suffices to check that

$$\det \begin{pmatrix} \mathbf{B}(z_0) - \lambda_k(z_0) I & \begin{matrix} 0 \\ \vdots \\ -1 \end{matrix} \\ 0 \dots\dots 2 & 0 \end{pmatrix} \neq 0. \tag{39}$$

3. Note that for $\epsilon > 0$ sufficiently small, the matrix

$$\mathbf{B}_\epsilon = \mathbf{B}(z_0) - (\lambda_k(z_0) + \epsilon)I \tag{40}$$

is invertible. In light of (38),

$$\mathbf{B}_\epsilon e_m = -\epsilon e_m.$$

Therefore

$$\begin{pmatrix} & & 0 \\ \mathbf{B}_\epsilon & & \vdots \\ & & -1 \\ 0\dots 2 & & 0 \end{pmatrix} \begin{pmatrix} & & 0 \\ I & & \vdots \\ & & (-\epsilon)^{-1} \\ 0\dots 0 & & 1 \end{pmatrix} = \begin{pmatrix} & & 0 \\ \mathbf{B}_\epsilon & & \vdots \\ & & 0 \\ 0\dots 2 & & 2(-\epsilon)^{-1} \end{pmatrix}.$$

Consequently, since the determinant of the second matrix before the equals sign is one, we have

$$\begin{aligned} \det \begin{pmatrix} & & 0 \\ \mathbf{B}_\epsilon & & \vdots \\ & & -1 \\ 0\dots 2 & & 0 \end{pmatrix} &= 2(\det \mathbf{B}_\epsilon)(-\epsilon)^{-1} \\ &= 2\prod_{j\neq k}(\lambda_j(z_0) - (\lambda_k(z_0) + \epsilon))(-\epsilon)(-\epsilon)^{-1} \\ &\to 2\prod_{j\neq k}(\lambda_j(z_0) - \lambda_k(z_0)) \quad \text{as } \epsilon \to 0. \end{aligned}$$

As $B(z_0)$ is strictly hyperbolic, the last expression is nonzero. Condition (39) is verified. We may thus invoke the Implicit Function Theorem to find near z_0 smooth functions $\lambda_k(z)$ and $\mathbf{r}_k(z)$, satisfying the conclusion of the theorem.

4. It remains to show that we can define $\lambda_k(z)$ and $\mathbf{r}_k(z)$ for all $z \in \mathbb{R}^m$, and not just near any particular point z_0. To do so, let us write

$$R := \sup\{r > 0 \mid \lambda_k(z), \mathbf{r}_k(z) \text{ as above exist and are smooth on } B(0,r)\}.$$

If $R = \infty$, we are done. Otherwise, we cover $\partial B(0,R)$ with finitely many open balls into which we can smoothly extend $\lambda_k(\cdot)$ and $\mathbf{r}_k(\cdot)$, using steps 1–3 above. This yields a contradiction to the definition of R.

A similar proof works for the left eigenvectors. □

Remark. Observe that we are not only globally and smoothly defining the eigenvalues and eigenspaces of $\mathbf{B}$, but are also globally providing the eigenspaces with an orientation.

This proof depends fundamentally upon the one-dimensionality of the eigenspaces. See Problem 2 for an example of what could go wrong in the hyperbolic, but not strictly hyperbolic, setting. □

Example 1 (continued). For the p-system (6), we have

$$\mathbf{u}_t + \mathbf{B}(\mathbf{u})\mathbf{u}_x = 0,$$

for

$$\mathbf{B}(z) = D\mathbf{F}(z) = \begin{pmatrix} 0 & -1 \\ -p'(z_1) & 0 \end{pmatrix}.$$

The eigenvalues are $\lambda_1 = -\sigma$, $\lambda_2 = \sigma$, for $\sigma := p'(z_1)^{1/2}$. These are real and distinct provided we hereafter suppose the strict hyperbolicity condition

$$p' > 0. \tag{41}$$

For the nonlinear wave equation (8) this is the physical assumption that the *stress* $p(u_x)$ is a strictly increasing function of the *strain* u_x. □

Example 2 (continued). Euler's equations (9) comprise a strictly hyperbolic system provided we assume $p > 0$ and

$$\frac{\partial p}{\partial \rho} > 0, \ \frac{\partial p}{\partial e} > 0, \tag{42}$$

where $p = p(\rho, e)$ is the constitutive relation between the mass density, the internal energy density and the pressure. This assertion is however difficult to verify directly, as the flux function $\mathbf{F}$ defined by (11) is complicated.

Let us rather change variables and regard the density ρ, velocity v and internal energy e as the unknowns. We can then rewrite Euler's equations (9) in terms of these quantities, and, so doing, obtain after some calculations the system

$$\begin{cases} \rho_t + v\rho_x + \rho v_x = 0 \\ v_t + vv_x + \frac{1}{\rho}p_x = 0 \\ e_t + ve_x + \frac{p}{\rho}v_x = 0, \end{cases} \tag{43}$$

provided $\rho > 0$. These equations are not in conservation form. Setting now $\mathbf{u} = (u^1, u^2, u^3) = (\rho, v, e)$, we rewrite (43) as

$$\mathbf{u}_t + \mathbf{B}(\mathbf{u})\mathbf{u}_x = 0 \quad \text{in } \mathbb{R} \times (0, \infty), \tag{44}$$

for

$$\mathbf{B}(z) = z_2 I + \hat{\mathbf{B}}(z), \tag{45}$$

where

$$\hat{\mathbf{B}}(z) := \begin{pmatrix} 0 & z_1 & 0 \\ \frac{1}{z_1}\frac{\partial p}{\partial \rho}(z_1, z_3) & 0 & \frac{1}{z_1}\frac{\partial p}{\partial e}(z_1, z_3) \\ 0 & \frac{1}{z_1}p(z_1, z_3) & 0 \end{pmatrix}.$$

The characteristic polynomial of $\hat{\mathbf{B}}$ is $-\lambda(\lambda^2 - \sigma^2)$, for $\sigma^2 = \frac{p}{z_1^2}\frac{\partial p}{\partial e} + \frac{\partial p}{\partial \rho}$. Recalling (45) and reverting to physical notation, we see that the eigenvalues of $\mathbf{B}$ are

$$\lambda_1 = v - \sigma,\ \lambda_2 = v,\ \lambda_3 = v + \sigma, \tag{46}$$

where

$$\sigma := \left(\frac{p}{\rho^2}\frac{\partial p}{\partial e} + \frac{\partial p}{\partial \rho}\right)^{1/2} > 0$$

is the local *sound speed.* We therefore see that the system (44) is strictly hyperbolic, provided assumption (42) is valid. Remembering now Theorem 1, we deduce that Euler's equations (9) are also strictly hyperbolic, with eigenvalues given by (46). □

11.2. RIEMANN'S PROBLEM

In this section we investigate in detail the system of conservation laws

$$\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0 \quad \text{in } \mathbb{R} \times (0, \infty), \tag{1}$$

with the piecewise-constant initial data

$$\mathbf{g} = \begin{cases} u_l & \text{if } x < 0 \\ u_r & \text{if } x > 0. \end{cases} \tag{2}$$

This is *Riemann's problem.* We call the given vectors u_l and u_r the left and right *initial states.*

11.2.1. Simple waves.

We commence our study of (1), (2) in very much the same spirit as in §11.1.2, in that we look for solutions of (1) having a special form. Before we searched for traveling waves, that is, solutions of the type $\mathbf{u}(x,t) = \mathbf{v}(x - \sigma t)$. We now seek *simple waves.* These are solutions of (1) having the structure

$$\mathbf{u}(x,t) = \mathbf{v}(w(x,t)) \quad (x \in \mathbb{R}, t > 0), \tag{3}$$

where $\mathbf{v} : \mathbb{R} \to \mathbb{R}^m$, $\mathbf{v} = (v^1, \dots, v^m)$, and $w : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ are to be found. To discover the requisite properties of $\mathbf{v}$ and w, let us substitute (3) into (1) and obtain the equality

$$\dot{\mathbf{v}}(w) + D\mathbf{F}(\mathbf{v}(w))\dot{\mathbf{v}}(w)w_x = 0. \tag{4}$$

Now in view of equation (25) from §11.1.2, with $\mathbf{B} = D\mathbf{F}$, we see (4) will be valid if for some $k \in \{1, \dots, m\}$ w solves the PDE

$$w_t + \lambda_k(\mathbf{v}(w))w_x = 0 \tag{5}$$

and $\mathbf{v}$ solves the ODE

$$\dot{\mathbf{v}}(s) = \mathbf{r}_k(\mathbf{v}(s)) \quad \left(\cdot = \frac{d}{ds} \right). \tag{6}$$

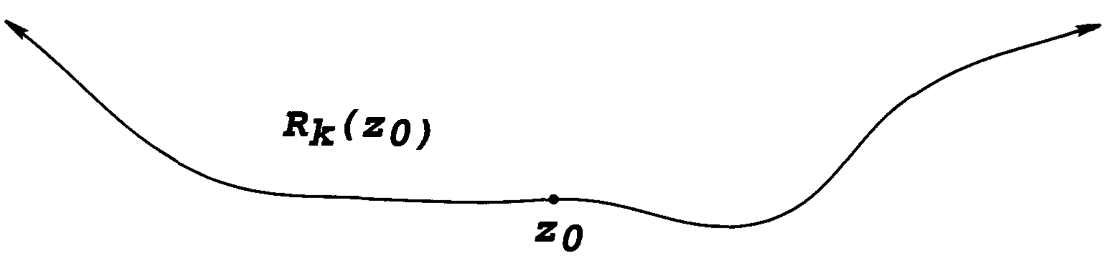
If (5) and (6) hold, we call the function $\mathbf{u}$ defined by (3) a *k-simple wave.* The point of all this is that we can regard (6) as an ODE for the vector function $\mathbf{v}$, and then—once $\mathbf{v}$ has been found by solving (6)—can interpret equation (5) as a *scalar* conservation law for w.

Let us next identify circumstances under which we can employ the construction (3)–(6) to build a continuous solution $\mathbf{u}$ of (1). We must examine first the ODE (6).

DEFINITION. *Given a fixed state $z_0 \in \mathbb{R}^m$, we define the k^{th}*-rarefaction curve

$$R_k(z_0)$$

to be the path in $\mathbb{R}^m$ of the solution of the ODE (6) *which passes through z_0.*



Rarefaction curve

Given then the solution $\mathbf{v}$ of (6), we turn to the PDE (5), which we rewrite as the scalar conservation law

$$w_t + F_k(w)_x = 0 \tag{7}$$

for

$$F_k(s) := \int_0^s \lambda_k(\mathbf{v}(t))\,dt \quad (s \in \mathbb{R}). \tag{8}$$

The PDE (7) will fall under the general theory developed in §3.4 provided F_k is strictly convex (or else strictly concave). Let us therefore compute

$$F'_k(s) = \lambda_k(\mathbf{v}(s)), \tag{9}$$

$$F''_k(s) = D\lambda_k(\mathbf{v}(s)) \cdot \dot{\mathbf{v}}(s) = D\lambda_k(\mathbf{v}(s)) \cdot \mathbf{r}_k(\mathbf{v}(s)). \tag{10}$$

Owing to (10), the function F_k will be convex if

$$D\lambda_k(z) \cdot \mathbf{r}_k(z) > 0 \quad (z \in \mathbb{R}^m),$$

and concave if

$$D\lambda_k(z) \cdot \mathbf{r}_k(z) < 0 \quad (z \in \mathbb{R}^m).$$

The function F_k is linear provided

$$D\lambda_k(z) \cdot \mathbf{r}_k(z) \equiv 0 \quad (z \in \mathbb{R}^m).$$

These possibilities motivate the following

DEFINITIONS. (i) *The pair* $(\lambda_k(z), \mathbf{r}_k(z))$ *is called* genuinely nonlinear *provided*

$$D\lambda_k(z) \cdot \mathbf{r}_k(z) \neq 0 \quad \textit{for all } z \in \mathbb{R}^m. \tag{11}$$

(ii) *We say* $(\lambda_k(z), \mathbf{r}_k(z))$ *is* linearly degenerate *if*

$$D\lambda_k(z) \cdot \mathbf{r}_k(z) = 0 \quad \textit{for all } z \in \mathbb{R}^m. \tag{12}$$

Notation. If the pair $(\lambda_k, \mathbf{r}_k)$ is genuinely nonlinear, write

$$R_k^+(z_0) := \{z \in R_k(z_0) \mid \lambda_k(z) > \lambda_k(z_0)\}$$

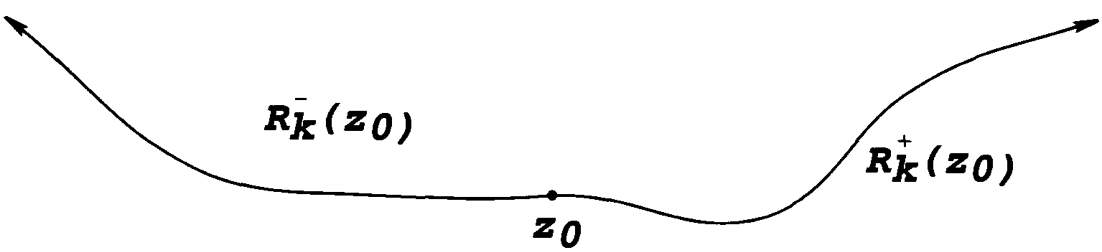
and

$$R_k^-(z_0) := \{z \in R_k(z_0) \mid \lambda_k(z) < \lambda_k(z_0)\}.$$

Then

$$R_k(z_0) = R_k^+(z_0) \cup \{z_0\} \cup R_k^-(z_0).$$

□



Two parts of the rarefaction curve

11.2.2. Rarefaction waves.

We turn our attention again to Riemann's problem (1), (2).

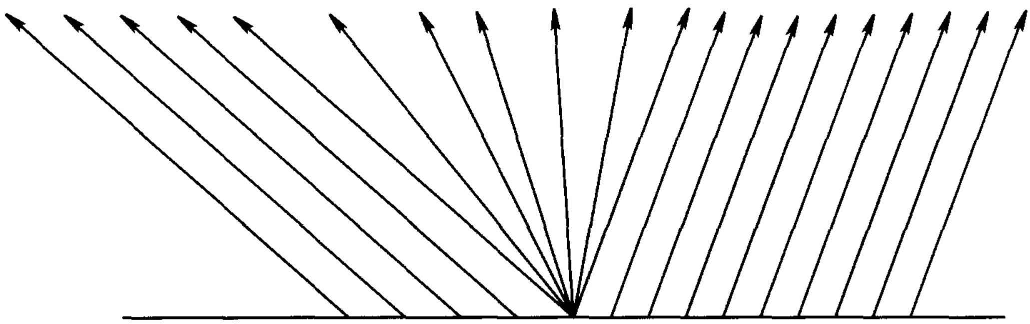
THEOREM 1 (Existence of k-rarefaction waves). *Suppose that for some* $k \in \{1, \dots, m\}$,

(i) *the pair* $(\lambda_k, \mathbf{r}_k)$ *is genuinely nonlinear, and*

(ii) $u_r \in R_k^+(u_l)$.

Then there exists a continuous integral solution $\mathbf{u}$ *of Riemann's problem* (1), (2), *which is a k-simple wave constant along lines through the origin.*

Remark. We call $\mathbf{u}$ a (centered) *k-rarefaction wave.*



k-rarefaction wave

Proof. 1. We first choose w_l and $w_r \in \mathbb{R}$ so that $u_l = \mathbf{v}(w_l)$, $u_r = \mathbf{v}(w_r)$. Suppose for the moment

$$w_l < w_r. \tag{13}$$

2. Consider then the scalar Riemann problem consisting of the PDE (7) together with the initial condition

$$g = \begin{cases} w_l & \text{if } x < 0 \\ w_r & \text{if } x > 0. \end{cases} \tag{14}$$

Now in view of hypothesis (ii) we have $\lambda_k(u_r) > \lambda_k(u_l)$; that is, according to (9), $F_k'(w_r) > F_k'(w_l)$. But then it follows from (i) that the function F_k defined by (8) is strictly convex. Accordingly we can apply Theorem 4 in §3.4.4 to the scalar Riemann problem (7), (14), whose unique weak solution is a continuous rarefaction wave connecting the states w_l and w_r. More specifically,

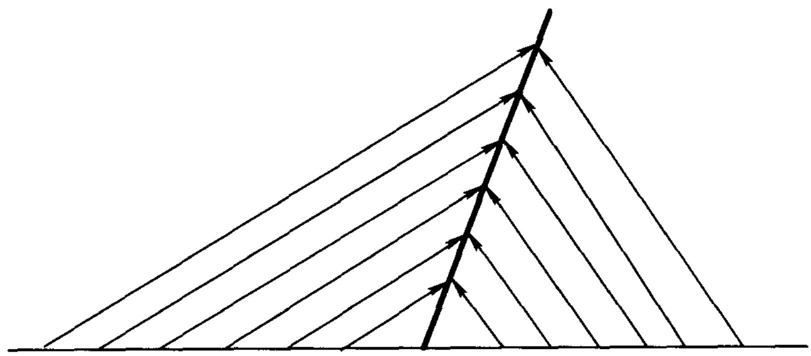
$$w(x,t) = \begin{cases} w_l & \text{if } \frac{x}{t} < F_k'(w_l) \\ G_k\left(\frac{x}{t}\right) & \text{if } F_k'(w_l) < \frac{x}{t} < F_k'(w_r) \quad (x \in \mathbb{R},\ t > 0) \\ w_r & \text{if } \frac{x}{t} > F_k'(w_r), \end{cases}$$

where $G_k = (F_k')^{-1}$. Thus $\mathbf{u}(x,t) = \mathbf{v}(w(x,t))$, where $\mathbf{v}$ solves the ODE (6) and passes through u_l, is a continuous integral solution of (1), (2).

3. The case $w_l > w_r$ is treated similarly, since F_k is then concave. □

11.2.3. Shock waves, contact discontinuities.

We consider next the possibility that the states u_l and u_r may be joined not by a rarefaction wave as above, but rather by a shock.



k-**shock wave**

a. The shock set.

Recalling the Rankine–Hugoniot condition from §11.1.1, we see that necessarily we must have the equality $\mathbf{F}(u_l) - \mathbf{F}(u_r) = \sigma(u_l - u_r)$, where $\sigma \in \mathbb{R}$, for such a shock wave to exist. This observation motivates the following

DEFINITION. *Given a fixed state $z_0 \in \mathbb{R}^m$, we define the* shock set

$$S(z_0) := \{z \in \mathbb{R}^m \mid \mathbf{F}(z) - \mathbf{F}(z_0) = \sigma(z - z_0) \textit{ for a constant } \sigma = \sigma(z, z_0)\}.$$

THEOREM 2 (Structure of the shock set). *Fix $z_0 \in \mathbb{R}^m$. In some neighborhood of z_0, $S(z_0)$ consists of the union of m smooth curves $S_k(z_0)$ $(k = 1, \dots, m)$, with the following properties:*

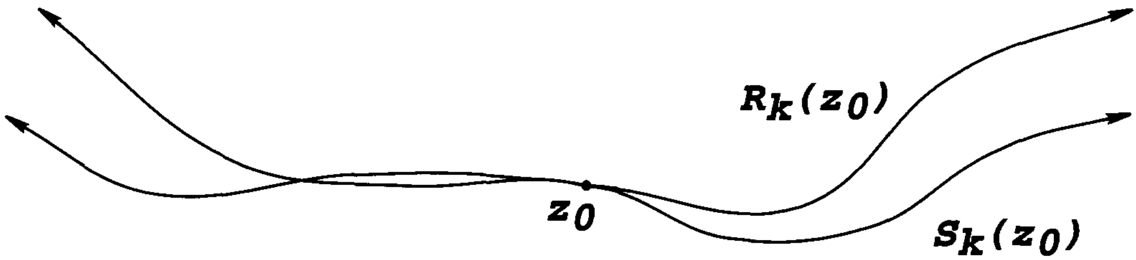
(i) *The curve $S_k(z_0)$ passes through z_0, with tangent $\mathbf{r}_k(z_0)$.*

(ii) $\lim_{\substack{z \to z_0 \\ z \in S_k(z_0)}} \sigma(z, z_0) = \lambda_k(z_0)$.

(iii) $\sigma(z, z_0) = \frac{\lambda_k(z) + \lambda_k(z_0)}{2} + O(|z - z_0|^2)$, *as $z \to z_0$ with $z \in S_k(z_0)$.*

Proof. 1. Define

$$\mathbf{B}(z) := \int_0^1 D\mathbf{F}(z_0 + t(z - z_0))\, dt \quad (z \in \mathbb{R}^m).$$



Contact between R_k and S_k

Then

(15) $$\mathbf{B}(z)(z - z_0) = \mathbf{F}(z) - \mathbf{F}(z_0).$$

In particular $z \in S(z_0)$ if and only if

(16) $$(\mathbf{B}(z) - \sigma I)(z - z_0) = 0$$

for some scalar $\sigma = \sigma(z, z_0)$.

2. We study equation (16) by first of all noting

(17) $$\mathbf{B}(z_0) = D\mathbf{F}(z_0).$$

Now in view of the strict hyperbolicity, the characteristic polynomial $\lambda \mapsto \det(\lambda I - \mathbf{B}(z_0))$ has m distinct, real roots, and hence the polynomial $\lambda \mapsto \det(\lambda I - \mathbf{B}(z))$ likewise has m distinct roots if z is close to z_0. Recalling Theorem 2 in §11.1.2 we see that near z_0 there exist smooth functions $\hat{\lambda}_1(z) < \cdots < \hat{\lambda}_m(z)$ and unit vectors $\{\hat{\mathbf{r}}_k(z), \hat{\mathbf{l}}_k(z)\}_{k=1}^m$ satisfying

$$\hat{\lambda}_k(z_0) = \lambda_k(z_0),\ \hat{\mathbf{r}}_k(z_0) = \mathbf{r}_k(z_0),\ \hat{\mathbf{l}}_k(z_0) = \mathbf{l}_k(z_0) \quad (k = 1, \dots, m),$$

and

(18) $$\begin{cases} \mathbf{B}(z)\hat{\mathbf{r}}_k(z) = \hat{\lambda}_k(z)\hat{\mathbf{r}}_k(z) \\ \hat{\mathbf{l}}_k(z)\mathbf{B}(z) = \hat{\lambda}_k(z)\hat{\mathbf{l}}_k(z) \end{cases} \quad (k = 1, \dots, m).$$

Note that $\{\hat{\mathbf{r}}_k(z)\}$, $\{\hat{\mathbf{l}}_k(z)\}_{k=1}^m$ are bases of $\mathbb{R}^m$, and also

(19) $$\hat{\mathbf{l}}_l(z) \cdot \hat{\mathbf{r}}_k(z) = 0 \quad (k \neq l).$$

3. Equation (16) will hold provided $\sigma = \hat{\lambda}_k(z)$ for some $k \in \{1, \dots, m\}$, and $(z - z_0)$ is parallel to $\hat{\mathbf{r}}_k(z)$. In light of (19), these conditions are equivalent to asking

(20) $$\hat{\mathbf{l}}_l(z) \cdot (z - z_0) = 0 \quad (l \neq k).$$

These equalities amount to $m-1$ equations for the m unknown components of z, which we intend to solve using the Implicit Function Theorem. So define $\mathbf{\Phi}_k : \mathbb{R}^m \to \mathbb{R}^{m-1}$ by setting

$$\mathbf{\Phi}_k(z) := (\dots, \hat{\mathbf{l}}_{k-1}(z) \cdot (z - z_0), \hat{\mathbf{l}}_{k+1}(z) \cdot (z - z_0), \dots).$$

Now $\mathbf{\Phi}_k(z_0) = 0$ and

$$D\mathbf{\Phi}_k(z_0) = \begin{pmatrix} \mathbf{l}_1(z_0) \\ \vdots \\ \mathbf{l}_{k-1}(z_0) \\ \mathbf{l}_{k+1}(z_0) \\ \vdots \\ \mathbf{l}_m(z_0) \end{pmatrix}_{(m-1)\times m},$$

the entries of this matrix being regarded as row vectors. Since the vectors $\{\mathbf{l}_k(z_0)\}_{k=1}^m$ form a basis of $\mathbb{R}^m$, we see

$$\text{rank } D\mathbf{\Phi}_k(z_0) = m - 1.$$

Accordingly, there exists a smooth curve $\phi_k : \mathbb{R} \to \mathbb{R}^m$ such that

$$\phi_k(0) = z_0 \tag{21}$$

and

$$\mathbf{\Phi}_k(\phi_k(t)) = 0 \quad \text{for all } t \text{ close to } 0. \tag{22}$$

The path of the curve $\phi_k(\cdot)$ for t near zero defines $S_k(z_0)$. We may reparameterize as necessary to ensure

$$|\dot{\phi}_k(t)| = 1 \quad \left(\cdot = \frac{d}{dt} \right). \tag{23}$$

4. Now (20)–(22) imply

$$\phi_k(t) = z_0 + \mu(t)\hat{\mathbf{r}}_k(\phi_k(t)) \tag{24}$$

for all t near zero, where $\mu : \mathbb{R} \to \mathbb{R}$ is a smooth function satisfying $\mu(0) = 0$, $\dot{\mu}(0) = 1$. Differentiating (24) with respect to t and setting $t = 0$, we thus find

$$\dot{\phi}_k(0) = \mathbf{r}_k(z_0). \tag{25}$$

Hence the curve $S_k(z_0)$ has tangent $\mathbf{r}_k(z_0)$ at z_0. Assertion (i) is proved.

5. In light of the foregoing analysis, there exists a smooth function $\sigma : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ such that

$$\mathbf{F}(\phi_k(t)) - \mathbf{F}(z_0) = \sigma(\phi_k(t), z_0)(\phi_k(t) - z_0) \tag{26}$$

for all t close to zero. Differentiating with respect to t and setting $t = 0$, we deduce from (21) that

$$D\mathbf{F}(z_0)\dot{\phi}_k(0) = \sigma(z_0, z_0)\dot{\phi}_k(0).$$

In light of (25), we see that $\sigma(z_0, z_0) = \lambda_k(z_0)$. This establishes assertion (ii).

6. Now write $\sigma(t) := \sigma(\phi_k(t), z_0)$; so that (26) reads

$$\mathbf{F}(\phi_k(t)) - \mathbf{F}(z_0) = \sigma(t)(\phi_k(t) - z_0).$$

Differentiate twice with respect to t:

$$\begin{aligned}(D^2\mathbf{F}(\phi_k(t))\dot{\phi}_k(t))\dot{\phi}_k(t) + D\mathbf{F}(\phi_k(t))\ddot{\phi}_k(t) \\ = \ddot{\sigma}(t)(\phi_k(t) - z_0) + 2\dot{\sigma}(t)\dot{\phi}_k(t) + \sigma(t)\ddot{\phi}_k(t).\end{aligned}$$

Evaluate this expression at $t = 0$ and recall $\sigma(0) = \lambda_k(z_0)$, $\phi_k(0) = z_0$, $\dot{\phi}_k(0) = \mathbf{r}_k(z_0)$:

$$(2\dot{\sigma}(0)I - D^2\mathbf{F}(z_0)\mathbf{r}_k(z_0))\mathbf{r}_k(z_0) = (D\mathbf{F}(z_0) - \lambda_k(z_0)I)\ddot{\phi}_k(0). \tag{27}$$

7. Let $\psi_k(t) = \mathbf{v}(t)$ be a unit speed parameterization of the rarefaction curve $R_k(z_0)$ near z_0 (as in (6) above). Then

$$\psi_k(0) = z_0, \ \dot{\psi}_k(t) = \mathbf{r}_k(\psi_k(t)). \tag{28}$$

Thus

$$D\mathbf{F}(\psi_k(t))\mathbf{r}_k(t) = \lambda_k(t)\mathbf{r}_k(t),$$

for

$$\lambda_k(t) := \lambda_k(\psi_k(t)), \ \mathbf{r}_k(t) := \mathbf{r}_k(\psi_k(t)).$$

Next differentiate with respect to t and set $t = 0$:

$$(D^2\mathbf{F}(z_0)\mathbf{r}_k(z_0) - \dot{\lambda}_k(0)I)\mathbf{r}_k(z_0) = -(D\mathbf{F}(z_0) - \lambda_k(z_0)I)\dot{r}_k(0). \tag{29}$$

Add (27) and (29), to obtain

$$(2\dot{\sigma}(0) - \dot{\lambda}_k(0))\mathbf{r}_k(z_0) = (D\mathbf{F}(z_0) - \lambda_k(z_0)I)(\ddot{\phi}_k(0) - \dot{r}_k(0)). \tag{30}$$

Take the dot product with $\mathbf{l}_k(z_0)$ and observe $\mathbf{l}_k \cdot \mathbf{r}_k \neq 0$, to conclude

$$2\dot{\sigma}(0) = \dot{\lambda}_k(0). \tag{31}$$

We deduce from (31) that

$$2\sigma(t) - \lambda_k(z_0) - \lambda_k(t) = O(t^2) \quad \text{as } t \to 0.$$

Assertion (iii) follows. □

We see from Theorem 2,(iii) that the curves $R_k(z_0)$ and $S_k(z_0)$ agree at least to first order at z_0. Next is the assertion that in the linearly degenerate case these curves in fact coincide.

THEOREM 3 (Linear degeneracy). *Suppose for some $k \in \{1, \dots, m\}$ that the pair $(\lambda_k, \mathbf{r}_k)$ is linearly degenerate. Then for each $z_0 \in \mathbb{R}^m$,*

(i) $R_k(z_0) = S_k(z_0)$

and

(ii) $\sigma(z, z_0) = \lambda_k(z) = \lambda_k(z_0)$ *for all $z \in S_k(z_0)$.*

Proof. Let $\mathbf{v} = \mathbf{v}(s)$ solve the ODE

$$\begin{cases} \dot{\mathbf{v}}(s) = \mathbf{r}_k(\mathbf{v}(s)) \quad (s \in \mathbb{R}) \\ \mathbf{v}(0) = z_0. \end{cases}$$

Then the mapping $s \mapsto \lambda_k(\mathbf{v}(s))$ is constant, and so

$$\begin{aligned} \mathbf{F}(\mathbf{v}(s)) - \mathbf{F}(z_0) &= \int_0^s D\mathbf{F}(\mathbf{v}(t))\dot{\mathbf{v}}(t)\,dt = \int_0^s D\mathbf{F}(\mathbf{v}(t))\mathbf{r}_k(\mathbf{v}(t))\,dt \\ &= \int_0^s \lambda_k(\mathbf{v}(t))\mathbf{r}_k(\mathbf{v}(t))\,dt = \lambda_k(z_0)\int_0^s \dot{\mathbf{v}}(t)\,dt \\ &= \lambda_k(z_0)(\mathbf{v}(s) - z_0). \end{aligned}$$

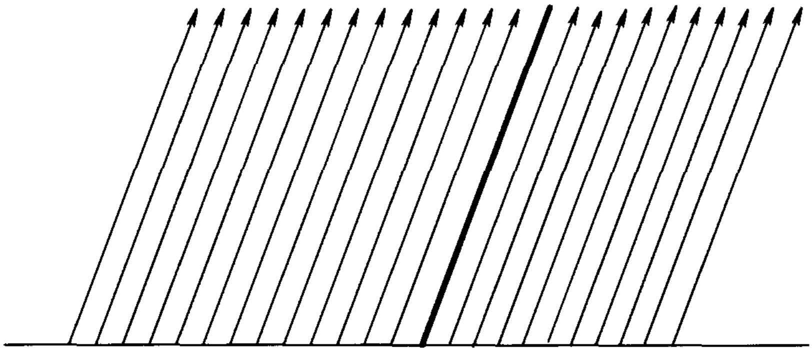
□

b. Contact discontinuities, shock waves.

We next undertake to analyze in light of Theorems 2 and 3 the possibility of solving Riemann's problem by joining two given states u_l and u_r by some kind of shock wave.

Contact discontinuities. Suppose first that $(\lambda_k, \mathbf{r}_k)$ is *linearly degenerate* and

$$u_r \in S_k(u_l). \tag{32}$$



k-contact discontinuity

We then define an integral solution of our system of conservation laws by setting

(33) $$\mathbf{u}(x,t) = \begin{cases} u_l & \text{if } x < \sigma t \\ u_r & \text{if } x > \sigma t, \end{cases}$$

for

(34) $$\sigma = \sigma(u_r, u_l) = \lambda_k(u_l) = \lambda_k(u_r).$$

Now observe from our analysis in §3.2.1 that since $\lambda_k(u_l) = \lambda_k(u_r) = \sigma$, the projected characteristics to the left and right are *parallel* to the line of discontinuity. We interpret this situation physically by saying that fluid particles do not cross the discontinuity. The line $x = \sigma t$ is called a *k-contact discontinuity.*

Shock waves. We next turn our attention to the case $(\lambda_k, \mathbf{r}_k)$ is *genuinely nonlinear*, and

(35) $$u_r \in S_k(u_l)$$

as before. If we consider the integral solution

(36) $$\mathbf{u}(x,t) = \begin{cases} u_l & \text{if } x < \sigma t \\ u_r & \text{if } x > \sigma t, \end{cases}$$

for

(37) $$\sigma = \sigma(u_r, u_l),$$

we see that there are two essentially different cases according as to whether

(38) $$\lambda_k(u_r) < \lambda_k(u_l)$$

or else

$$\lambda_k(u_l) < \lambda_k(u_r). \tag{39}$$

Now in view of assertion (iii) from Theorem 2, we have then either

$$\lambda_k(u_r) < \sigma(u_r, u_l) < \lambda_k(u_l) \tag{40}$$

or

$$\lambda_k(u_l) < \sigma(u_r, u_l) < \lambda_k(u_r), \tag{41}$$

provided u_r is close enough to u_l.

This dichotomy is reminiscent of a corresponding situation in §3.4, for a scalar conservation law. By analogy with the entropy conditions introduced there, let us hereafter agree to reject the inequalities (41) as allowing for "nonphysical shocks" from which characteristics emanate as we move forward in time. We rather take (40) as being physically correct. The informal viewpoint is that then the characteristics from the left and right run into the line of discontinuity, whereupon "information is lost" and so "entropy increases". This interpretation was largely justified mathematically in §3.4 with our uniqueness theorem for weak solutions that satisfied this sort of entropy condition.

Refocusing our attention again to systems, we therefore agree to regard (40) as the correct inequalities to be satisfied:

DEFINITION. *Assume the pair $(\lambda_k, \mathbf{r}_k)$ is genuinely nonlinear at u_l. We say that the pair (u_l, u_r) is* admissible *provided*

$$u_r \in S_k(u_l)$$

and

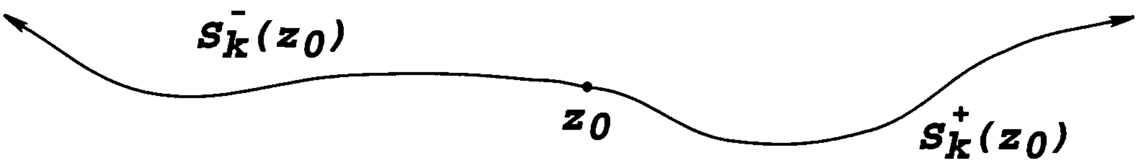
$$\lambda_k(u_r) < \sigma(u_r, u_l) < \lambda_k(u_l). \tag{42}$$

We refer to (42) as the *Lax entropy condition.* If (u_l, u_r) is admissible, we call our solution $\mathbf{u}$ defined by (36), (37) a *k-shock wave.*

By analogy with our decomposition of $R_k(z_0)$ into $R_k^{\pm}(z_0)$, let us introduce this

DEFINITION. *If the pair $(\lambda_k, \mathbf{r}_k)$ is genuinely nonlinear, we write*

$$S_k^+(z_0) := \{z \in S_k(z_0) \mid \lambda_k(z_0) < \sigma(z, z_0) < \lambda_k(z)\}$$



Shock curve

and

$$S_k^-(z_0) := \{z \in S_k(z_0) \mid \lambda_k(z) < \sigma(z, z_0) < \lambda_k(z_0)\}.$$

Then

$$S_k(z_0) = S_k^+(z_0) \cup \{z_0\} \cup S_k^-(z_0)$$

near z_0. Note then that the pair (u_l, u_r) is admissible if and only if

$$u_r \in S_k^-(u_l).$$

11.2.4. Local solution of Riemann's problem.

Next we glue together the physically relevant parts of the rarefaction and shock curves.

DEFINITIONS. (i) *If the pair $(\lambda_k, \mathbf{r}_k)$ is genuinely nonlinear, write*

$$T_k(z_0) := R_k^+(z_0) \cup \{z_0\} \cup S_k^-(z_0).$$

(ii) *If the pair $(\lambda_k, \mathbf{r}_k)$ is linearly degenerate, we set*

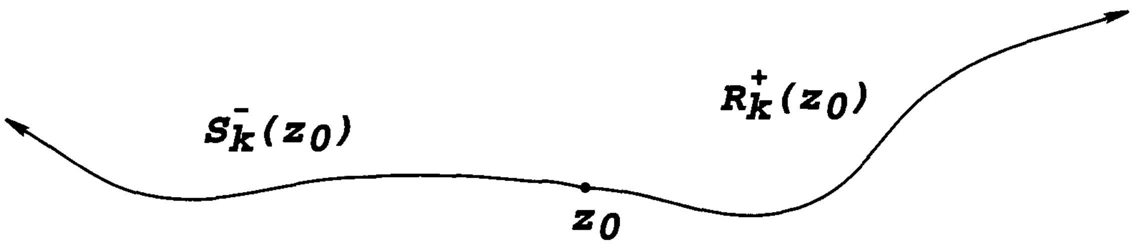
$$T_k(z_0) := R_k(z_0) = S_k(z_0).$$

Owing to Theorem 2,(ii) the curve $T_k(z_0)$ is C^1. Employing this notation we see that nearby states u_l and u_r can be joined by a k-rarefaction wave, a shock wave or a contact discontinuity provided

$$u_l \in T_k(u_r). \tag{43}$$

We now at last ask if we can find a solution to Riemann's problem provided only that u_r is close to u_l (but (43) may fail for each $k = 1, \ldots, m$). The hope is that by moving along various paths T_k for different values of k, we may be able to connect u_l to u_r, utilizing a sequence of rarefaction waves, shock waves, and/or contact discontinuities.

THEOREM 4 (Local solution of Riemann's problem). *Assume for each $k = 1, \ldots, m$ that the pair $(\lambda_k, \mathbf{r}_k)$ is either genuinely nonlinear or else linearly degenerate. Suppose further the left state u_l is given. Then for each right state u_r sufficiently close to u_l there exists an integral solution $\mathbf{u}$ of Riemann's problem, which is constant on lines through the origin.*



Structure of the T-curve

Proof. 1. We intend to apply the Inverse Function Theorem to a mapping $\Phi : \mathbb{R}^m \to \mathbb{R}^m$, defined near 0 as follows.

First, for each family of curves T_k ($k = 1, \dots, m$) choose the nonsingular parameter τ_k to measure arc length; that is, if $z, \tilde{z} \in \mathbb{R}^m$ with $\tilde{z} \in T_k(z)$, then

$$\tau_k(\tilde{z}) - \tau_k(z) = \text{ (signed) distance from } z \text{ to } \tilde{z} \text{ along the curve } T_k(z).$$

We take the plus sign for $\tau_k(\tilde{z})$ if $\tilde{z} \in R_k^+(z)$, the minus sign if $\tilde{z} \in S_k^-(z)$.

2. Given then $t = (t_1, \dots, t_m) \in \mathbb{R}^m$, with $|t|$ small, we define $\Phi(t) = z$ as follows. First, temporarily write

$$u_l = z_0. \tag{44}$$

Then choose states $z_1, \dots, z_m$ to satisfy

$$\begin{cases} z_1 \in T_1(z_0), \ \ \tau_1(z_1) - \tau_1(z_0) = t_1, \\ z_2 \in T_2(z_1), \ \ \tau_2(z_2) - \tau_2(z_1) = t_2, \\ \quad \vdots \\ z_m \in T_m(z_{m-1}), \ \ \tau_m(z_m) - \tau_m(z_{m-1}) = t_m. \end{cases} \tag{45}$$

Now write

$$z = z_m, \tag{46}$$

and define

$$\Phi(t) = z. \tag{47}$$

Note Φ is C^1, and

$$\Phi(0) = z_0. \tag{48}$$

3. We claim

$$D\Phi(0) \text{ is nonsingular.} \tag{49}$$

To see this, observe

$$\Phi(0,\dots,t_k,\dots,0)-\Phi(0,\dots,0)=t_k\mathbf{r}_k(z_0)+o(t_k)\quad\text{as } t_k\to 0.$$

Thus

$$\frac{\partial\Phi}{\partial t_k}(0)=\mathbf{r}_k(z_0)\quad(k=1,\dots,m),$$

and so

$$D\Phi(0)=(\mathbf{r}_1(z_0),\dots,\mathbf{r}_m(z_0))_{m\times m},$$

the entries regarded as column vectors. This matrix is nonsingular, since $\{\mathbf{r}_k(z_0)\}_{k=1}^m$ is a basis.

4. In light of (49), the Inverse Function Theorem applies: for each state u_r sufficiently close to u_l there exists a unique parameter $t=(t_1,\dots,t_m)$ close to zero such that $\Phi(t)=u_r$.

Recall next that if z_{k-1} and z_k are joined by a k-rarefaction wave, this wave is

$$\begin{cases} z_{k-1} & \text{if } \frac{x}{t}<\lambda_k(z_{k-1}) \\ G_k(\frac{x}{t}) & \text{if } \lambda_k(z_{k-1})<\frac{x}{t}<\lambda_k(z_k), \quad \text{for } G_k=(F'_k)^{-1} \\ z_k & \text{if } \lambda_k(z_k)<\frac{x}{t}. \end{cases}$$

Moreover if z_{k-1}, z_k are joined by a k-shock, it has the form

$$\begin{cases} z_{k-1} & \text{if } \frac{x}{t}<\sigma(z_k,z_{k-1}) \\ z_k & \text{if } \sigma(z_k,z_{k-1})<\frac{x}{t}, \end{cases}$$

where $\lambda_k(z_k)<\sigma(z_k,z_{k-1})<\lambda_k(z_{k-1})$. In both cases the waves are constant outside the regions $\lambda_k(z_0)-\epsilon<\frac{x}{t}<\lambda_k(z_0)+\epsilon$, for small $\epsilon>0$, provided z_k, z_{k-1} are close enough to z_0. This is true for $k=1,\dots,m$.

Since $\lambda_1(z_0)<\cdots<\lambda_m(z_0)$, we see then that the rarefactions, shock waves and/or contact discontinuities connecting $u_l=z_0$ to z_1, z_1 to z_2, z_2 to z_3, $\dots$, z_{m-1} to $z_m=u_r$ do not intersect. □

11.3. SYSTEMS OF TWO CONSERVATION LAWS

In this section we more deeply analyze the initial-value problem for $m=2$, which is to say, for a pair of conservation laws:

$$(1)\qquad \begin{cases} u_t^1+F^1(u^1,u^2)_x=0 & \text{in } \mathbb{R}\times(0,\infty) \\ u_t^2+F^2(u^1,u^2)_x=0 & \\ u^1=g^1,\ u^2=g^2 & \text{on } \mathbb{R}\times\{t=0\}. \end{cases}$$

Here $\mathbf{F} = (F^1, F^2), \mathbf{g} = (g^1, g^2), \mathbf{u} = (u^1, u^2)$.

11.3.1. Riemann invariants.

Our intention is first to demonstrate that we can transform (1) into a much simpler form by performing an appropriate nonlinear change of dependent variables. The idea is to find two functions $w^1, w^2 : \mathbb{R}^2 \to \mathbb{R}$ with nice properties along the rarefaction curves R_1, R_2:

DEFINITION. *We say*

$$w^i : \mathbb{R}^2 \to \mathbb{R}$$

is an i^{th}-Riemann invariant *provided*

(2) $$Dw^i(z) \text{ is parallel to } \mathbf{l}_j(z) \quad (z \in \mathbb{R}^2, i \neq j).$$

We will see momentarily how useful condition (2) is, but let us first pause to ask whether Riemann invariants exist. It turns out that since we are now taking $m = 2$, this is easy. Indeed, because $\mathbf{l}_j(z) \cdot \mathbf{r}_i(z) = 0$ $(i \neq j)$, we see (2) is equivalent in $\mathbb{R}^2$ to the statement

(2′) $$Dw^i(z) \cdot \mathbf{r}_i(z) = 0 \quad (i = 1, 2, \; z \in \mathbb{R}^2);$$

which is to say

(3) $$w^i \text{ is constant along the rarefaction curve } R_i \quad (i = 1, 2).$$

In particular, any smooth function w^i satisfying (3) satisfies also (2′), (2) and so is an i^{th}-Riemann invariant.

Remark. In the case that $m > 2$, Riemann invariants do not in general exist. □

Now we can regard $w = (w_1, w_2) = (w^1(z_1, z_2), w^2(z_1, z_2))$ as being new coordinates on the state space $\mathbb{R}^2$, replacing $z = (z_1, z_2)$. More precisely, we define $\mathbf{w} : \mathbb{R}^2 \to \mathbb{R}^2$ by setting

(4) $$\mathbf{w}(z) = \mathbf{w}(z_1, z_2) = (w^1(z_1, z_2), w^2(z_1, z_2)).$$

The inverse mapping is $\mathbf{z}(w) = \mathbf{z}(w_1, w_2) = (z^1(w_1, w_2), z^2(w_1, w_2))$.

Let us now utilize the transformation (4) to simplify our system of two conservation laws (1). For this, let us suppose henceforth $\mathbf{u} = (u^1, u^2)$ is a smooth solution of (1). We now change dependent variables by setting

(5) $$\mathbf{v}(x, t) := \mathbf{w}(\mathbf{u}(x, t)) \quad (x \in \mathbb{R}, t > 0).$$

What system of PDE does $\mathbf{v} = (v^1, v^2)$ satisfy?

THEOREM 1 (Conservation laws and Riemann invariants). *The functions v^1, v^2 solve the system*

$$\begin{cases} v^1_t + \lambda_2(\mathbf{u})v^1_x = 0 \\ v^2_t + \lambda_1(\mathbf{u})v^2_x = 0 \end{cases} \quad \text{in } \mathbb{R} \times (0,\infty). \tag{6}$$

The point is that the system (6), although not in conservation law form, is in many ways rather simpler than (1). Note in particular that whereas the PDE for u^1 involves the term u^2_x, the PDE for v^1 does *not* entail v^2_x. Similarly, the PDE for v^2 does not involve v^1_x.

Proof. According to (5), we see that for $i = 1, 2$, $i \neq j$,

$$\begin{aligned} v^i_t + \lambda_j(\mathbf{u})v^i_x &= Dw^i(\mathbf{u}) \cdot \mathbf{u}_t + \lambda_j(\mathbf{u})Dw^i(\mathbf{u}) \cdot \mathbf{u}_x \\ &= Dw^i(\mathbf{u}) \cdot (-\mathbf{F}(\mathbf{u})_x + \lambda_j(\mathbf{u})\mathbf{u}_x) \\ &= Dw^i(\mathbf{u}) \cdot (-D\mathbf{F}(\mathbf{u}) + \lambda_j(\mathbf{u})I)\mathbf{u}_x = 0, \end{aligned}$$

since, by definition, Dw^i is parallel to $\mathbf{l}_j$. □

Remarks. (i) We can interpret the system of PDE (6) by introducing the ODE

$$\dot{x}_i(s) = \lambda_j(\mathbf{u}(x_i(s), s)) \quad (s \geq 0) \tag{7}$$

for $i = 1, 2$, $j \neq i$. Then we see from (6) that

(8) $\qquad v^i$ is constant along the curve $(x_i(s), s) \quad (s \geq 0)$

for $i = 1, 2$.

(ii) Recall from §11.2 our condition of genuine nonlinearity reads

$$D\lambda_i(z) \cdot \mathbf{r}_i(z) \neq 0 \quad (z \in \mathbb{R}^2,\ i = 1, 2). \tag{9}$$

Since we can also regard λ_i as a function of $w = (w_1, w_2)$, we can rewrite (9) to read

$$\frac{\partial \lambda_i}{\partial w_j} \neq 0 \quad (w \in \mathbb{R}^2,\ i \neq j). \tag{10}$$

To see (9) is equivalent to (10), observe that if (10) fails, then

$$0 = \frac{\partial \lambda_i}{\partial w_j} = \sum_{k=1}^{2} \frac{\partial \lambda_i}{\partial z_k}\frac{\partial z^k}{\partial w_j}. \tag{11}$$

But since $\sum_{k=1}^2 \frac{\partial w^i}{\partial z_k}\frac{\partial z^k}{\partial w_j} = \delta_{ij} = 0$ for $i \neq j$, we see that (11) asserts that $D\lambda_i$ is parallel to Dw^i. However, Dw^i is perpendicular to $\mathbf{r}_i$, and so we obtain a contradiction to (9). Hence (9) implies (10), and the converse implication is established in the same way. □

Example (Barotropic compressible gas dynamics). We illustrate the foregoing ideas by examining in detail Euler's equations for barotropic compressible gas dynamics, a special case of the general Euler equations (Example 2 in §11.1) when the internal energy e is constant. The relevant PDE are

$$\text{(12)} \qquad \begin{cases} \rho_t + (\rho v)_x = 0 & \text{(conservation of mass)} \\ (\rho v)_t + (\rho v^2 + p)_x = 0 & \text{(conservation of momentum),} \end{cases}$$

where we now assume

$$\text{(13)} \qquad p = p(\rho)$$

for some smooth function $p : \mathbb{R} \to \mathbb{R}$. Formula (13) is called a *barotropic equation of state.* We assume the strict hyperbolicity condition

$$\text{(14)} \qquad p' > 0.$$

Setting $\mathbf{u} = (u^1, u^2) = (\rho, \rho v)$, we can rewrite (12), (13) to read

$$\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0,$$

for

$$\mathbf{F} = (F^1, F^2) = (z_2, (z_2)^2/z_1 + p(z_1))$$

and $z = (z_1, z_2)$, provided $z_1 > 0$. Then

$$D\mathbf{F} = \begin{pmatrix} 0 & 1 \\ -(\frac{z_2}{z_1})^2 + p'(z_1) & \frac{2z_2}{z_1} \end{pmatrix}.$$

Consequently

$$\text{(15)} \qquad \lambda_1 = \frac{z_2}{z_1} - p'(z_1)^{1/2}, \ \lambda_2 = \frac{z_2}{z_1} + p'(z_1)^{1/2}.$$

In physical notation:

$$\text{(16)} \qquad \lambda_1 = v - \sigma, \ \lambda_2 = v + \sigma,$$

for the sound speed

$$\text{(17)} \qquad \sigma := p'(\rho)^{1/2}.$$

Remembering (7), we consider next the ODE

$$\text{(18)} \qquad \dot{x}_1(t) = v(x_1(t), t) + \sigma(x_1(t), t),$$

(19) $$\dot{x}_2(t) = v(x_2(t), t) - \sigma(x_2(t), t),$$

where $\sigma(x,t) := p'(\rho(x,t))^{\frac{1}{2}}$, $t \geq 0$. We know from (8) that the Riemann invariant $v^1 = w^1(\mathbf{u})$ is constant along the trajectories of (18) and $v^2 = w^2(\mathbf{u})$ is constant along trajectories of (19).

To compute w^1 and w^2 directly, let us carry out some computations on the system (12). First we transform (12) in nondivergence form:

(20) $$\rho_t + \rho v_x + \rho_x v = 0,$$

(21) $$\rho_t v + \rho v_t + \rho_x v^2 + 2\rho v v_x + p_x = 0.$$

Multiplying (20) by $\sigma^2 = p'(\rho)$ and recalling (13) gives us

(22) $$p_t + v p_x + \sigma^2 \rho v_x = 0.$$

In addition (20), (21) combine to yield

(23) $$\rho v_t + \rho v v_x + p_x = 0.$$

We now manipulate (22), (23) so that the directions $\lambda_1, \lambda_2 = v \mp \sigma$ appear explicitly. To accomplish this, we multiply (23) by σ and then add to and subtract from (22):

(24) $$\begin{cases} p_t + (v+\sigma)p_x + \rho\sigma(v_t + (v+\sigma)v_x) = 0 \\ p_t + (v-\sigma)p_x - \rho\sigma(v_t + (v-\sigma)v_x) = 0. \end{cases}$$

We then deduce from (24) that

(25) $$\begin{cases} \frac{d}{dt}[p(x_1(t),t)] + \rho(x_1(t),t)\sigma(x_1(t),t)\frac{d}{dt}[v(x_1(t),t)] = 0 \\ \frac{d}{dt}[p(x_2(t),t)] - \rho(x_2(t),t)\sigma(x_2(t),t)\frac{d}{dt}[v(x_2(t),t)] = 0. \end{cases}$$

As $\frac{dp}{dt} = \sigma^2 \frac{d\rho}{dt}$, we see

(26) $$\frac{\sigma}{\rho}\frac{d\rho}{dt} \pm \frac{dv}{dt} = 0 \text{ along the trajectories of (18), (19),}$$

provided $\rho > 0$.

Think now of the Riemann invariants as functions of ρ and v. Then since $v^1 = w^1(\rho, v)$ is constant along the curve determined by $x_1(\cdot)$, we have

$$\begin{aligned} 0 &= \frac{d}{dt}[w^1(\rho(x_1(t),t), v(x_1(t),t))] \\ &= \frac{\partial w^1}{\partial \rho}\frac{d}{dt}[\rho(x_1(t),t)] + \frac{\partial w^1}{\partial v}\frac{d}{dt}[v(x_1(t),t)]. \end{aligned}$$

This is consistent with (26) if

$$\frac{\partial w^1}{\partial \rho} = \frac{\sigma(\rho)}{\rho}, \ \frac{\partial w^1}{\partial v} = 1.$$

We similarly deduce

$$\frac{\partial w^2}{\partial \rho} = \frac{\sigma(\rho)}{\rho}, \ \frac{\partial w^2}{\partial v} = -1.$$

Integrating, we conclude that the Riemann invariants are, up to additive constants,

$$w^1 = \int_1^\rho \frac{\sigma(s)}{s}\, ds + v, \quad w^2 = \int_1^\rho \frac{\sigma(s)}{s}\, ds - v.$$

We leave it as an exercise to check that w^1, w^2, taken now as functions of $z = (z_1, z_2)$, satisfy the definition of Riemann invariants. □

11.3.2. Nonexistence of smooth solutions.

Illustrating now the usefulness of Riemann invariants, we establish the following criterion for the nonexistence of a smooth solution:

THEOREM 2 (Riemann invariants and blow-up). *Assume* $\mathbf{g}$ *is smooth, with compact support. Suppose also the genuine nonlinearity condition*

$$\frac{\partial \lambda_i}{\partial w_j} > 0 \ \ \text{in } \mathbb{R}^2 \quad (i = 1, 2, \ i \neq j) \tag{27}$$

holds. Then the initial-value problem (1) *cannot have a smooth solution* $\mathbf{u}$ *existing for all times* $t \geq 0$ *if*

$$\text{either} \ \ v_x^1 < 0 \ \ \text{or} \ \ v_x^2 < 0 \ \ \text{somewhere on } \mathbb{R} \times \{t = 0\}. \tag{28}$$

Proof. 1. Assume for the time being that $\mathbf{u}$ is a smooth solution of (1). Write

$$a := v_x^1, \ b := v_x^2, \tag{29}$$

where $\mathbf{v} = \mathbf{w}(\mathbf{u})$, $\mathbf{v} = (v^1, v^2)$, solves the system of PDE (6). We differentiate the first equation of (6) with respect to x, to compute:

$$a_t + \lambda_2 a_x + \frac{\partial \lambda_2}{\partial w_1} a^2 + \frac{\partial \lambda_2}{\partial w_2} ab = 0. \tag{30}$$

We employ then the second equation of (6), which we rewrite as

$$v_t^2 + \lambda_2 v_x^2 = (\lambda_2 - \lambda_1)b.$$

Substituting this expression into (30) gives

$$a_t + \lambda_2 a_x + \frac{\partial \lambda_2}{\partial w_1} a^2 + \left[\frac{1}{\lambda_2 - \lambda_1}\frac{\partial \lambda_2}{\partial w_2}(v_t^2 + \lambda_2 v_x^2)\right] a = 0. \tag{31}$$

2. To integrate (31), fix $x_0 \in \mathbb{R}$ and set

$$\xi(t) := \exp\left(\int_0^t \frac{1}{\lambda_2 - \lambda_1}\frac{\partial \lambda_2}{\partial w_2}(v_t^2 + \lambda_2 v_x^2)(x_1(s), s)\, ds\right), \tag{32}$$

where

$$\begin{cases} \dot{x}_1(s) = \lambda_2(\mathbf{u}(x_1(s), s)) \quad (s \geq 0) \\ x_1(0) = x_0. \end{cases} \tag{33}$$

Next is the key observation from (8) that v^1 is constant along the curve $(x_1(s), s)$. So write

$$v^1(x_1(s), s) = v_0^1 = v^1(x_0, 0) \quad (s \geq 0).$$

Thus we see that the expression $\left(\frac{1}{\lambda_2 - \lambda_1}\frac{\partial \lambda_2}{\partial w_2}\right)$, considered as a function of $\mathbf{v} = \mathbf{w}(\mathbf{u})$, depends only on v^2. Let us set

$$\gamma(\mu) := \int_0^\mu \left(\frac{1}{\lambda_2 - \lambda_1}\frac{\partial \lambda_2}{\partial w_2}\right)(v_0^1, v)\, dv.$$

Then (32), (33) imply

$$\begin{aligned} \xi(t) &= \exp\left(\int_0^t \frac{d}{ds}[\gamma(v^2(x_1(s), s))]\, ds\right) \\ &= \exp(\gamma(v^2(x_1(t), t)) - \gamma(v^2(x_0, 0))). \end{aligned} \tag{34}$$

3. We now transform (31) to read

$$\frac{d}{dt}((\xi\alpha)^{-1}) = \frac{-1}{(\xi\alpha)^2}\frac{d}{dt}(\xi\alpha) = \frac{\partial \lambda_2}{\partial w_1}\xi^{-1}$$

where $\alpha(t) := a(x_1(t), t)$ and we assume $\alpha \neq 0$. Consequently

$$(\xi(t)\alpha(t))^{-1} = (\alpha(0))^{-1} + \int_0^t \frac{\partial \lambda_2}{\partial w_1}\xi^{-1}(s)\, ds.$$

This equality in turn rearranges to become

$$\alpha(t) = \alpha(0)\xi^{-1}(t)\left(1+\alpha(0)\int_0^t \frac{\partial\lambda_2}{\partial w_1}\xi^{-1}(s)\,ds\right)^{-1}. \tag{35}$$

4. Now in view of the system of PDE (6), $\mathbf{v}$ is bounded. Thus we deduce from (34) that $0 < \theta \leq \xi(t) \leq \Theta$ for all times $t > 0$, for appropriate constants θ, Θ. Therefore it follows from (27) and (35) that α is bounded for all $t > 0$ if and only if $\alpha(0) \geq 0$, that is, if

$$v_x^1(x_0, 0) \geq 0.$$

A similar calculation holds with v^2 replacing v^1. We conclude that if either $v_x^1 < 0$ or else $v_x^2 < 0$ somewhere on $\mathbb{R} \times \{t = 0\}$, there cannot then exist a smooth solution of (1), lasting for all times $t \geq 0$. □

11.4. ENTROPY CRITERIA

In our study of Riemann's problem in §11.2 we have taken Lax's entropy condition

$$\lambda_k(u_r) < \sigma(u_r, u_l) < \lambda_k(u_l) \tag{1}$$

for some $k \in \{1, \dots, m\}$ as the selection criteria for admissible shock waves.

There is great ongoing interest in discovering other mathematically correct and physically appropriate entropy conditions of various sorts, with the aim of applying these to more complicated integral solutions of our system of conservation laws, so as to obtain uniqueness criteria, more information concerning allowable discontinuities, etc.

One general principle, instances of which we have already seen for scalar conservation laws in §4.5.1 and for Hamilton–Jacobi equations in §10.1, is that physically and mathematically correct solutions should arise as the limit of solutions to the regularized system

$$\mathbf{u}_t^\varepsilon + \mathbf{F}(\mathbf{u}^\varepsilon)_x - \varepsilon\mathbf{u}_{xx}^\varepsilon = 0 \quad \text{in } \mathbb{R} \times (0,\infty) \tag{2}$$

as $\varepsilon \to 0$. The idea is to interpret the term "$\varepsilon\mathbf{u}_{xx}^\varepsilon$" as providing a small viscosity effect, which will presumably "smear out" sharp shocks. The hope is to study various aspects of the problem (2) in the limit $\varepsilon \to 0$, and thereby to discover more general entropy criteria, to augment Lax's condition (1).

The next sections discuss aspects of this general program.

11.4.1. Vanishing viscosity, traveling waves.

We begin our investigation of the parabolic system (2) by first seeking a traveling wave solution, having the form

$$\mathbf{u}^\varepsilon(x,t) = \mathbf{v}\left(\frac{x-\sigma t}{\varepsilon}\right) \quad (x \in \mathbb{R},\ t \geq 0), \tag{3}$$

where as usual the speed σ and profile $\mathbf{v}$ must be found. Substituting (3) into (2), we find $\mathbf{v} : \mathbb{R} \to \mathbb{R}^m$, $\mathbf{v} = \mathbf{v}(s)$, must solve the ODE

$$\ddot{\mathbf{v}} = -\sigma\dot{\mathbf{v}} + D\mathbf{F}(\mathbf{v})\dot{\mathbf{v}} \quad \left(\cdot = \frac{d}{ds}\right). \tag{4}$$

Assume now $u_l, u_r \in \mathbb{R}^m$ are given, and furthermore

$$\lim_{s\to-\infty} \mathbf{v} = u_l, \quad \lim_{s\to+\infty} \mathbf{v} = u_r, \quad \lim_{s\to\pm\infty} \dot{\mathbf{v}} = 0. \tag{5}$$

Then from (3) we deduce

$$\lim_{\varepsilon\to 0} \mathbf{u}^\varepsilon(x,t) = \begin{cases} u_l & \text{if } x < \sigma t \\ u_r & \text{if } x > \sigma t. \end{cases} \tag{6}$$

Hence the limit as $\varepsilon \to 0$ of our solution to (2) gives us a shock wave connecting the states u_l, u_r. The plan now is to study carefully the form of σ and $\mathbf{v}$, and thereby glean more detailed information about the structure of the shock determined by (6).

The first and primary question is whether there in fact exist σ and $\mathbf{v}$ solving (4), (5). Integrating (4), we deduce

$$\dot{\mathbf{v}} = \mathbf{F}(\mathbf{v}) - \sigma\mathbf{v} + c \tag{7}$$

for some constant $c \in \mathbb{R}^m$. We conclude from (5) that

$$\mathbf{F}(u_l) - \sigma u_l + c = \mathbf{F}(u_r) - \sigma u_r + c. \tag{8}$$

Hence

$$\mathbf{F}(u_l) - \mathbf{F}(u_r) = \sigma(u_l - u_r). \tag{9}$$

In view of (5), (8) and (9), our ODE (7) becomes

$$\dot{\mathbf{v}} = \mathbf{F}(\mathbf{v}) - \mathbf{F}(u_l) - \sigma(\mathbf{v} - u_l). \tag{10}$$

Now think of the left state u_l as being given, and suppose we are trying to build a traveling wave connecting u_l to a nearby state u_r. From (9) we see that necessarily $u_r \in S_k(u_l)$ for some $k \in \{1,\dots,m\}$, and

$$\sigma = \sigma(u_r, u_l). \tag{11}$$

We refine this observation as follows:

THEOREM 1 (Existence of traveling waves for genuinely nonlinear systems). *Assume the pair* $(\lambda_k, \mathbf{r}_k)$ *is genuinely nonlinear for* $k = 1, \dots, m$. *Let* u_r *be selected sufficiently close to* u_l. *Then there exists a traveling wave solution of* (2) *connecting* u_l *to* u_r *if and only if*

$$u_r \in S_k^-(u_l) \tag{12}$$

for some $k \in \{1, \dots, m\}$.

Proof. 1. Assume first σ and $\mathbf{v}$ solve (4), (5). Then, as noted above, necessarily $u_r \in S_k(u_l)$ for some $k \in \{1, \dots, m\}$, $\sigma = \sigma(u_r, u_l)$. Now set

$$\mathbf{G}(z) := \mathbf{F}(z) - \mathbf{F}(u_l) - \sigma(z - u_l). \tag{13}$$

Our ODE (10) then reads

$$\dot{\mathbf{v}} = \mathbf{G}(\mathbf{v}), \tag{14}$$

and we have

$$\mathbf{G}(u_l) = \mathbf{G}(u_r) = 0, \tag{15}$$

according to (9). We compute

$$D\mathbf{G}(u_l) = D\mathbf{F}(u_l) - \sigma I;$$

and so the eigenvalues of $D\mathbf{G}$ at u_l are $\{\lambda_k(u_l) - \sigma\}_{k=1}^m$, with corresponding right and left eigenvectors $\{\mathbf{r}_k, \mathbf{l}_k\}_{k=1}^m$, $\mathbf{r}_k = \mathbf{r}_k(u_l)$, $\mathbf{l}_k = \mathbf{l}_k(u_l)$.

2. Now since $u_r \in S_k(u_l)$ and $|u_r - u_l|$ is small, we know from Theorem 2,(iii) in §11.2.3 that

$$\sigma = \frac{\lambda_k(u_r) + \lambda_k(u_l)}{2} + o(|u_r - u_l|).$$

Thus

$$\lambda_k(u_l) - \sigma = \frac{\lambda_k(u_l) - \lambda_k(u_r)}{2} + o(|u_r - u_l|).$$

In order that there be an orbit of the ODE (14) connecting u_l at $s = -\infty$ to $u_r \in S_k(u_r)$ at $s = +\infty$, it must be that $\lambda_k(u_l) - \sigma > 0$; for otherwise the trajectory would not converge to u_l as $s \to -\infty$. Thus if $|u_r - u_l|$ is small enough, $\lambda_k(u_r) < \lambda_k(u_l)$, which is to say $u_r \in S_k^-(u_l)$.

3. We omit proof of the sufficiency of condition (12): see Majda–Pego [**M-P**]. □

The preceding result employs the genuine nonlinearity assumption, but the assertion holds in general, provided we introduce an appropriate variant of Lax's entropy condition (1). So let us suppose now $u_r \in S_k(u_l)$ for some $k \in \{1, \dots, m\}$ and furthermore

$$\text{(16)} \qquad \begin{cases} \sigma(z, u_l) > \sigma(u_r, u_l) \text{ for each } z \text{ lying} \\ \text{on the curve } S_k(u_l) \text{ between } u_r \text{ and } u_l. \end{cases}$$

Condition (16) is *Liu's entropy criterion.* (Observe this condition is automatic provided $(\lambda_k, \mathbf{r}_k)$ is genuinely nonlinear, $u_r \in S_k^-(u_l)$, and u_r is sufficiently close to u_l.)

We can motivate (16) by again seeking traveling wave solutions of system (2). So assume u_l is given. Then provided $|u_r - u_l|$ is small enough, it turns out that there exists a traveling wave solution $\mathbf{u}^\varepsilon(x,t) = \mathbf{v}(\frac{x-\sigma t}{\varepsilon})$, $\mathbf{v}$ solving (4), (5), if and only if the entropy condition (16) is satisfied. See Conlon [**CO**] for a proof.

To make this all a bit clearer, we next present in detail a specific application.

Example (Traveling waves for the p-system). Let us consider again the p-system

$$\text{(17)} \qquad \begin{cases} u_t^1 - u_x^2 = 0 & \text{(compatibility condition)} \\ u_t^2 - p(u^1)_x = 0 & \text{(Newton's law),} \end{cases}$$

under the usual strict hyperbolicity condition

$$\text{(18)} \qquad p' > 0.$$

We investigate the existence of traveling wave solutions to the regularized system

$$\text{(19)} \qquad \begin{cases} u_t^{\varepsilon,1} - u_x^{\varepsilon,2} = 0 \\ u_t^{\varepsilon,2} - p(u^{\varepsilon,1})_x = \varepsilon u_{xx}^{\varepsilon,2}. \end{cases}$$

Notice we have added the viscosity term only to the second equation. This makes sense physically, as the first line of (17) is only a mathematical compatibility condition.

Assume now $\mathbf{u}^\varepsilon = \mathbf{v}(\frac{x-\sigma t}{\varepsilon})$ is a traveling wave solution of (19), with

$$\text{(20)} \qquad \lim_{s\to-\infty} \mathbf{v} = u_l, \ \lim_{s\to\infty} \mathbf{v} = u_r, \ \lim_{s\to\pm\infty} \dot{\mathbf{v}} = 0.$$

Writing $\mathbf{v} = (v^1, v^2)$, we compute from (19) that

$$\text{(21)} \qquad \begin{cases} -\sigma\dot{v}^1 - \dot{v}^2 = 0 \quad (\dot{} = \frac{d}{ds}) \\ -\sigma\dot{v}^2 - p(v^1)^{\cdot} = \ddot{v}^2. \end{cases}$$

An integration using (20) gives

$$\text{(22)} \qquad \begin{cases} \sigma v^1 + v^2 = \sigma v_l^1 + v_l^2 = \sigma v_r^1 + v_r^2 \\ \dot{v}^2 = \sigma(v_l^2 - v^2) + p(v_l^1) - p(v^1) = \sigma(v_r^2 - v^2) + p(v_r^1) - p(v^1), \end{cases}$$

for $u_l = (v_l^1, v_l^2)$, $u_r = (v_r^1, v_r^2)$. In particular,

$$\begin{cases} \sigma v_l^1 + v_l^2 = \sigma v_r^1 + v_r^2 \\ \sigma v_l^2 + p(v_l^1) = \sigma v_r^2 + p(v_r^1). \end{cases}$$

Solving these equations for σ, we obtain

$$\text{(23)} \qquad \sigma^2 = \frac{p(v_r^1) - p(v_l^1)}{v_r^1 - v_l^1}.$$

Suppose hereafter $v_r^1 > v_l^1$. In view of (18) we can take $\sigma > 0$. In this situation the Liu entropy criterion reads

$$\text{(24)} \qquad \frac{p(z_1) - p(v_l^1)}{z_1 - v_l^1} > \frac{p(v_r^1) - p(v_l^1)}{v_r^1 - v_l^1}$$

for all z on the curve $S_k(u_l)$ between u_l and u_r, $z = (z_1, z_2)$.

We now claim the system of ODE (22), with asymptotic boundary conditions (20), has a solution if and only if the entropy condition (24) holds. To confirm this, combine the two equations in (22) to eliminate v^2:

$$\dot{v}^1 = \frac{p(v^1) - p(v_l^1)}{\sigma} - \sigma(v^1 - v_l^1) =: g(v^1).$$

Now $g(v_l^1) = 0$ and $g(v_r^1) = 0$, according to (23). Thus in order that the ODE (24) have a solution, with $\lim_{s \to -\infty} v^1 = v_l^1$, $\lim_{s \to \infty} v^1 = v_r^1$, we require

$$g(z_1) > 0 \quad \text{for } v_l^1 < z_1 < v_r^1.$$

But this is precisely the entropy criterion (24). A similar calculation works if $v_r^1 < v_l^1$. □

11.4.2. Entropy/entropy-flux pairs.

Both Lax's and Liu's entropy criteria provide restrictions on possible left and right hand states joined by a shock wave (or a traveling wave for the viscous approximation). It is however of considerable interest to widen still further the entropy criteria, so as to apply to more general integral solutions of our conservation laws.

One idea is to require an integral solution satisfy certain "entropy-type" inequalities.

DEFINITION. *Two smooth functions* $\Phi, \Psi : \mathbb{R}^m \to \mathbb{R}$ *comprise an* entropy/entropy-flux pair *for the conservation law* $\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0$ *provided*

$$\Phi \text{ is convex} \tag{25}$$

and

$$D\Phi(z)D\mathbf{F}(z) = D\Psi(z) \quad (z \in \mathbb{R}^m). \tag{26}$$

To motivate condition (26), suppose for the moment $\mathbf{u}$ is a smooth solution of the system of PDE $\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0$. We then compute

$$\begin{aligned} \Phi(\mathbf{u})_t + \Psi(\mathbf{u})_x &= D\Phi(\mathbf{u}) \cdot \mathbf{u}_t + D\Psi(\mathbf{u}) \cdot \mathbf{u}_x \\ &= (-D\Phi(\mathbf{u})D\mathbf{F}(\mathbf{u}) + D\Psi(\mathbf{u})) \cdot \mathbf{u}_x = 0 \end{aligned} \tag{27}$$

by (26). This computation says the quantity $\Phi(\mathbf{u})$ satisfies a scalar conservation law, with flux $\Psi(\mathbf{u})$.

Now in general integral solutions of (1) will not be smooth enough, owing to shocks and other irregularities, to justify the foregoing computation. The new idea is instead to replace (27) with an *inequality:*

$$\Phi(\mathbf{u})_t + \Psi(\mathbf{u})_x \leq 0 \quad \text{in } \mathbb{R} \times (0, \infty). \tag{28}$$

In applications $\Phi(\mathbf{u})$ will sometimes be the negative of physical entropy and $\Psi(\mathbf{u})$ the entropy flux. The inequality (28) therefore asserts entropy evolves according to its flux, but may also undergo sharp increases, for instance along shocks.

Let us hereafter rigorously understand (28) to mean

$$\begin{cases} \int_0^\infty \int_{-\infty}^\infty \Phi(\mathbf{u})v_t + \Psi(\mathbf{u})v_x \, dxdt \geq 0 \\ \text{for each } v \in C_c^\infty(\mathbb{R} \times (0, \infty)), \ v \geq 0. \end{cases} \tag{29}$$

We consider once more the initial-value problem

$$\begin{cases} \mathbf{u}_t + \mathbf{F}(\mathbf{u})_x = 0 & \text{in } \mathbb{R} \times (0, \infty) \\ \mathbf{u} = \mathbf{g} & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{30}$$

DEFINITION. *We call* $\mathbf{u}$ *an* entropy solution *of* (30) *provided* $\mathbf{u}$ *is an integral solution and* $\mathbf{u}$ *satisfies the inequalities* (29) *for each entropy/entropy-flux pair* (Φ, Ψ).

Let us now attempt to build for general initial data $\mathbf{g}$ an entropy solution. As in §11.4.1 we expect such a "physically correct" solution $\mathbf{u}$ to be a limit of solutions $\mathbf{u}^\varepsilon$ of approximating viscous problems

$$\text{(31)} \qquad \begin{cases} \mathbf{u}^\varepsilon_t + \mathbf{F}(\mathbf{u}^\varepsilon)_x - \varepsilon \mathbf{u}^\varepsilon_{xx} = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ \mathbf{u}^\varepsilon = \mathbf{g} & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases}$$

We assume $\mathbf{u}^\varepsilon$ is a smooth solution of (31), converging to 0 as $|x| \to \infty$ sufficiently rapidly to justify the calculations below. Let us further suppose $\{\mathbf{u}^\varepsilon\}_{0<\varepsilon\leq 1}$ is uniformly bounded in L^∞ and furthermore

$$\text{(32)} \qquad \mathbf{u}^\varepsilon \to \mathbf{u} \quad \text{a.e. as } \varepsilon \to 0$$

for some limit function $\mathbf{u}$. (In practice it is extremely difficult to verify this a.e. convergence.)

THEOREM 2 (Entropy and vanishing viscosity). *The function* $\mathbf{u}$ *is an entropy solution of the conservation law* (30).

Proof. 1. Choose any smooth entropy/entropy-flux pair (Φ, Ψ). Left multiplying (30) by $D\Phi(\mathbf{u}^\varepsilon)$ and recalling (26), we compute

$$\text{(33)} \qquad \begin{aligned} \Phi(\mathbf{u}^\varepsilon)_t + \Psi(\mathbf{u}^\varepsilon)_x &= \varepsilon D\Phi(\mathbf{u}^\varepsilon)\mathbf{u}^\varepsilon_{xx} \\ &= \varepsilon \Phi(\mathbf{u}^\varepsilon)_{xx} - \varepsilon (D^2\Phi(\mathbf{u}^\varepsilon)\mathbf{u}^\varepsilon_x) \cdot \mathbf{u}^\varepsilon_x. \end{aligned}$$

As Φ is convex,

$$\text{(34)} \qquad (D^2\Phi(\mathbf{u}^\varepsilon)\mathbf{u}^\varepsilon_x) \cdot \mathbf{u}^\varepsilon_x \geq 0.$$

2. Multiply (33) by $v \in C^\infty_c(\mathbb{R} \times (0,\infty))$, $v \geq 0$. We integrate by parts and discover:

$$\begin{aligned} &\int_0^\infty \int_{-\infty}^\infty \Phi(\mathbf{u}^\varepsilon) v_t + \Psi(\mathbf{u}^\varepsilon) v_x \, dxdt \\ &\quad = \int_0^\infty \int_{-\infty}^\infty \varepsilon (D^2\Phi(\mathbf{u}^\varepsilon)\mathbf{u}^\varepsilon_x) \cdot \mathbf{u}^\varepsilon_x v - \varepsilon \Phi(\mathbf{u}^\varepsilon) v_{xx} \, dxdt \\ &\quad \geq - \int_0^\infty \int_{-\infty}^\infty \varepsilon \Phi(\mathbf{u}^\varepsilon) v_{xx} \, dxdt, \end{aligned}$$

the last inequality holding in view of (34) and the nonnegativity of v.

Now let $\varepsilon \to 0$. Recalling (32) and the Dominated Convergence Theorem, we obtain

$$\int_0^\infty \int_{-\infty}^\infty \Phi(\mathbf{u})v_t + \Psi(\mathbf{u})v_x \, dxdt \geq 0.$$

Thus $\mathbf{u}$ verifies the entropy/entropy-flux inequalities (29). If Φ and Ψ are not smooth, we obtain the same conclusion after an approximation.

3. Finally fix $\mathbf{v} \in C_c^\infty(\mathbb{R} \times [0,\infty); \mathbb{R}^m)$ and take the dot product of the PDE in (31) with $\mathbf{v}$. After integrating by parts we obtain

$$\int_0^\infty \int_{-\infty}^\infty \mathbf{u}^\varepsilon \cdot \mathbf{v}_t + \mathbf{F}(\mathbf{u}^\varepsilon)\mathbf{v}_x + \varepsilon \mathbf{u}^\varepsilon \cdot \mathbf{v}_{xx} \, dxdt + \int_{-\infty}^\infty \mathbf{g} \cdot \mathbf{v} \, dx|_{t=0} = 0.$$

We send $\varepsilon \to 0$, to deduce $\mathbf{u}$ is an integral solution of (30). □

Example 1. In the case of a scalar conservation law (i.e. $m = 1$), for any convex Φ we can find a corresponding flux function Ψ, namely

$$\Psi(z) = \int_{z_0}^z \Phi'(w)F'(w) \, dw \quad (z \in \mathbb{R}).$$

See §11.4.3 following for an application. □

Example 2. For the p-system we have $m = 2$. To verify (25), (26) we must find Φ, Ψ, with Φ convex and

$$(\Phi_{z_1}, \Phi_{z_2}) \begin{pmatrix} 0 & -1 \\ -p'(z_1) & 0 \end{pmatrix} = \begin{pmatrix} \Psi_{z_1} \\ \Psi_{z_2} \end{pmatrix}.$$

A solution is

$$\Phi(z) = \frac{z_2^2}{2} + \int_0^{z_1} p(w) \, dw, \quad \Psi(z) = -p(z_1)z_2 \quad (z \in \mathbb{R}^2).$$

Note Φ is convex, since $p' > 0$. □

See Problems 4, 5 for other examples.

11.4.3. Uniqueness for a scalar conservation law.

As a further illustration of the ideas in §11.4.2, let us now consider again the initial-value problem for a *scalar* conservation law

$$(35) \qquad \begin{cases} u_t + F(u)_x = 0 & \text{in } \mathbb{R} \times (0,\infty) \\ u = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases}$$

Hence the unknown $u = u(x,t)$ is real-valued and $F : \mathbb{R} \to \mathbb{R}$ is a given smooth flux function.

In §3.4 we carefully studied the problem (35), making use of the primary assumption that F be strictly convex, to derive the Lax–Oleinik formula (see §3.4.2). Let us now drop the assumption that F be convex and devise an appropriate notion of weak solution. As above, we introduce entropies:

DEFINITION. *Two smooth functions* $\Phi, \Psi : \mathbb{R} \to \mathbb{R}$ *comprise an* entropy/entropy-flux pair *for the conservation law* $u_t + F(u)_x = 0$ *provided*

$$\Phi \textit{ is convex} \tag{36}$$

and

$$\Phi'(z)F'(z) = \Psi'(z) \quad (z \in \mathbb{R}). \tag{37}$$

As noted in Example 1 above, for each convex Φ there exists a corresponding flux Ψ.

The entropy condition for u reads

$$\Phi(u)_t + \Psi(u)_x \leq 0 \quad \text{ on } \mathbb{R} \times (0, \infty)$$

for each entropy/entropy-flux pair Φ, Ψ. This means

$$\begin{cases} \int_0^\infty \int_{-\infty}^\infty \Phi(u)v_t + \Psi(u)v_x \, dxdt \geq 0 \\ \text{for each } v \in C_c^\infty(\mathbb{R} \times (0,\infty)), \ v \geq 0. \end{cases} \tag{38}$$

DEFINITION. *We call* $u \in C([0,\infty), L^1(\mathbb{R})) \cap L^\infty(\mathbb{R} \times (0,\infty))$ *an* entropy solution *of* (35) *provided* u *satisfies the inequalities* (38) *for each entropy/entropy-flux pair* (Φ, Ψ), *and* $u(\cdot, t) \to g$ *in* L^1 *as* $t \to 0$.

Remarks. (i) This definition supersedes our earlier definition of "entropy solution" in §3.4.3.

(ii) Taking $\Phi(z) = \pm z$, $\Psi(z) = \pm F(z)$ in (38), we deduce

$$\int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \, dxdt = 0$$

for all $v \geq 0$ and thus for all $v \in C_c^1(\mathbb{R} \times (0,\infty))$. It is an exercise to prove then that

$$\int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \, dxdt + \int_{-\infty}^\infty gv \, dx|_{t=0} = 0$$

for all $v \in C_c^1(\mathbb{R} \times [0,\infty))$, since $u(\cdot, t) \to g$ in L^1. Thus *an entropy solution is an integral solution.* □

We discussed in §11.4.2 the construction of an entropy solution, and we now prove uniqueness.

THEOREM 3 (Uniqueness of entropy solutions for a single conservation law). *There exists—up to a set of measure zero—at most one entropy solution of* (35).

As in the proof of Theorem 1 in §10.2, the basic idea will be to "double the variables" in the problem.

Proof*. 1. Let u be an entropy solution of (35). Then

$$\int_0^\infty \int_{-\infty}^\infty \Phi(u)v_t + \Psi(u)v_x \, dxdt \geq 0 \tag{39}$$

for all $v \in C_c^\infty(\mathbb{R} \times (0,\infty)), v \geq 0$, where Φ is smooth, convex and

$$\Psi(z) = \int_{z_0}^z \Phi'(w)F'(w)\,dw$$

for any z_0. Fix $\alpha \in \mathbb{R}$ and take

$$\Phi_k(z) := \beta_k(z-\alpha) \quad (z \in \mathbb{R}), \tag{40}$$

where for each $k = 1, \ldots,$ $\beta_k : \mathbb{R} \to \mathbb{R}$ is smooth, convex and

$$\begin{cases} \beta_k(z) \to |z| & \text{uniformly} \\ \beta_k'(z) \to \operatorname{sgn}(z) & \text{boundedly, a.e.} \end{cases}$$

Thus $\Phi_k(z) \to |z-\alpha|$ uniformly for $z \in \mathbb{R}$. A flux corresponding to (40) is

$$\Psi_k(z) = \int_\alpha^z \beta_k'(w-\alpha)F'(w)\,dw.$$

Consequently for each z

$$\Psi_k(z) \to \int_\alpha^z \operatorname{sgn}(w-\alpha)F'(w)\,dw = \operatorname{sgn}(z-\alpha)(F(z)-F(\alpha)).$$

Putting Φ_k, Ψ_k into (39) and sending $k \to \infty$, we deduce

$$\int_0^\infty \int_{-\infty}^\infty |u-\alpha|v_t + \operatorname{sgn}(u-\alpha)(F(u)-F(\alpha))v_x \, dxdt \geq 0 \tag{41}$$

for each $\alpha \in \mathbb{R}$ and v as above.

2. Next let $\tilde{u}$ be another entropy solution. Then

$$\int_0^\infty \int_{-\infty}^\infty |\tilde{u}-\tilde{\alpha}|\tilde{v}_s + \operatorname{sgn}(\tilde{u}-\tilde{\alpha})(F(\tilde{u})-F(\tilde{\alpha}))\tilde{v}_y \, dyds \geq 0 \tag{42}$$

*Omit on first reading.

where $\tilde{\alpha} \in \mathbb{R}$ and $\tilde{v} \in C_c^\infty(\mathbb{R} \times (0,\infty))$, $\tilde{v} \geq 0$.

Now let $w \in C_c^\infty(\mathbb{R} \times \mathbb{R} \times (0,\infty) \times (0,\infty))$, $w \geq 0$, $w = w(x,y,t,s)$. Fixing $(y,s) \in \mathbb{R} \times (0,\infty)$, we take $\alpha = \tilde{u}(y,s)$, $v(x,t) = w(x,y,t,s)$ in (41). Integrating with respect to y, s, we produce the inequality

$$\begin{aligned} (43)\quad \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty &|u(x,t) - \tilde{u}(y,s)| w_t \\ &+ \operatorname{sgn}(u(x,t) - \tilde{u}(y,s))(F(u(x,t)) - F(\tilde{u}(y,s))) w_x \, dx dy dt ds \geq 0. \end{aligned}$$

Likewise, for each fixed $(x,t) \in \mathbb{R} \times (0,\infty)$ we take $\tilde{\alpha} = u(x,t)$, $\tilde{v}(y,s) = w(x,y,t,s)$ in (42). Integrating with respect to x, t gives:

$$\begin{aligned} (44)\quad \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty &|\tilde{u}(y,s) - u(x,t)| w_s \\ &+ \operatorname{sgn}(\tilde{u}(y,s) - u(x,t))\,(F(\tilde{u}(y,s)) - F(u(x,t))) w_y \, dx dy dt ds \geq 0. \end{aligned}$$

Add (43), (44):

$$\begin{aligned} (45)\quad \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty &|u(x,t) - \tilde{u}(y,s)|(w_t + w_s) \\ &+ \operatorname{sgn}(u(x,t) - \tilde{u}(y,s))\,(F(u(x,t)) - F(\tilde{u}(y,s))) \\ &(w_x + w_y)\, dx dy dt ds \geq 0. \end{aligned}$$

3. We design as follows a clever choice for w in (45). Select η to be a standard mollifier as in §C.4 (with $n = 1$) and, as usual, write $\eta_\varepsilon(x) = \frac{1}{\varepsilon}\eta\left(\frac{x}{\varepsilon}\right)$. Take

$$w(x,y,t,s) := \eta_\varepsilon\left(\frac{x-y}{2}\right) \eta_\varepsilon\left(\frac{t-s}{2}\right) \phi\left(\frac{x+y}{2}, \frac{t+s}{2}\right),$$

where $\phi \in C_c^\infty(\mathbb{R} \times (0,\infty))$, $\phi \geq 0$. We insert this choice of w into (45) and thereby obtain:

$$\begin{aligned} (46)\quad \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty &\left\{ |u(x,t) - \tilde{u}(y,s)| \phi_t\left(\frac{x+y}{2}, \frac{t+s}{2}\right) \right. \\ &\left. + \operatorname{sgn}(u(x,t) - \tilde{u}(y,s))\,(F(u(x,t)) - F(\tilde{u}(y,s))) \phi_x\left(\frac{x+y}{2}, \frac{t+s}{2}\right) \right\} \\ &\eta_\varepsilon\left(\frac{x-y}{2}\right) \eta_\varepsilon\left(\frac{t-s}{2}\right) dx dy dt ds \geq 0. \end{aligned}$$

Change variables by writing

$$\begin{cases} \bar{x} = \frac{x+y}{2}, \bar{t} = \frac{t+s}{2} \\ \bar{y} = \frac{x-y}{2}, \bar{s} = \frac{t-s}{2}. \end{cases}$$

Then (46) implies

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(\bar{y},\bar{s})\eta_\varepsilon(\bar{y})\eta_\varepsilon(\bar{s})\,d\bar{y}d\bar{s} \geq 0, \tag{47}$$

where

$$\begin{aligned} G(\bar{y},\bar{s}) := \int_0^{\infty}\int_{-\infty}^{\infty} & |u(\bar{x}+\bar{y},\bar{t}+\bar{s}) - \tilde{u}(\bar{x}-\bar{y},\bar{t}-\bar{s})|\phi_t(\bar{x},\bar{t}) \\ & + \operatorname{sgn}(u(\bar{x}+\bar{y},\bar{t}+\bar{s}) - \tilde{u}(\bar{x}-\bar{y},\bar{t}-\bar{s})) \\ & (F(u(\bar{x}+\bar{y},\bar{t}+\bar{s})) - F(\tilde{u}(\bar{x}-\bar{y},\bar{t}-\bar{s})))\phi_x(\bar{x},\bar{t})\,d\bar{x}d\bar{t}. \end{aligned} \tag{48}$$

Now $u(\bar{x}+\bar{y},\bar{t}+\bar{s}) \to u(\bar{x},\bar{t})$, $\tilde{u}(\bar{x}-\bar{y},\bar{t}-\bar{s}) \to \tilde{u}(\bar{x},\bar{t})$ in L^1_{loc} as $\bar{y},\bar{s} \to 0$. Since the mappings $(a,b) \mapsto |a-b|$, $\operatorname{sgn}(a-b)(F(a)-F(b))$ are Lipschitz continuous, we deduce upon letting $\varepsilon \to 0$ in (47) that

$$\begin{aligned} \int_0^{\infty}\int_{-\infty}^{\infty} & |u(\bar{x},\bar{t}) - \tilde{u}(\bar{x},\bar{t})|\,\phi_t(\bar{x},\bar{t}) \\ & + \operatorname{sgn}(u(\bar{x},\bar{t}) - \tilde{u}(\bar{x},\bar{t}))\,(F(u(\bar{x},\bar{t})) - F(\tilde{u}(\bar{x},\bar{t})))\,\phi_x(\bar{x},\bar{t})\,d\bar{x}d\bar{t} \geq 0. \end{aligned}$$

Rewriting $x = \bar{x}$, $t = \bar{t}$, we have therefore

$$\int_0^{\infty}\int_{-\infty}^{\infty} a(x,t)\phi_t(x,t) + b(x,t)\phi_x(x,t)\,dxdt \geq 0, \tag{49}$$

for

$$\begin{cases} a(x,t) := |u(x,t) - \tilde{u}(x,t)| \\ b(x,t) := \operatorname{sgn}(u(x,t) - \tilde{u}(x,t))(F(u(x,t)) - F(\tilde{u}(x,t))). \end{cases}$$

4. We now employ the inequality (49) to establish the L^1-contraction inequalities

$$\begin{cases} \int_{-\infty}^{\infty} |u(x,t) - \tilde{u}(x,t)|\,dx \leq \int_{-\infty}^{\infty} |u(x,s) - \tilde{u}(x,s)|\,dx \\ \text{for a.e. } 0 \leq s \leq t. \end{cases} \tag{50}$$

To prove this assertion, we take $0 < s < t$, $r > 0$, and let $\phi(x,t) = \alpha(x)\beta(t)$ in (49), where

$$\begin{cases} \alpha : \mathbb{R} \to \mathbb{R} \text{ is smooth,} \\ \alpha(x) = 1 \text{ if } |x| \leq r,\ \alpha(x) = 0 \text{ if } |x| \geq r+1, \\ |\alpha'(x)| \leq 2 \end{cases}$$

and

$$\begin{cases} \beta : \mathbb{R} \to \mathbb{R} \text{ is Lipschitz,} \\ \beta(\tau) = 0 \text{ if } \ 0 \leq \tau \leq s \text{ or } \tau \geq t+\delta, \\ \beta(\tau) = 1 \text{ if } \ s+\delta \leq \tau \leq t, \\ \beta \text{ is linear on } [s,s+\delta] \text{ and } [t,t+\delta], \end{cases}$$

for $0 < \delta < t - s$. We deduce

$$\frac{1}{\delta}\int_s^{s+\delta}\int_{-\infty}^{\infty} a(x,\tau)\alpha(x)\,dxd\tau \geq \frac{1}{\delta}\int_t^{t+\delta}\int_{-\infty}^{\infty} a(x,\tau)\alpha(x)\,dxd\tau$$
$$- \int_s^{t+\delta}\int_{\{r\leq|x|\leq r+1\}} b(x,\tau)\alpha'(x)b(\tau)\,dxd\tau.$$

Let $r \to \infty$:

$$\frac{1}{\delta}\int_t^{t+\delta}\int_{-\infty}^{\infty} a(x,\tau)\,dxd\tau \leq \frac{1}{\delta}\int_s^{s+\delta}\int_{-\infty}^{\infty} a(x,\tau)\,dxd\tau.$$

Next let $\delta \to 0$ to deduce (50) for a.e. $0 \leq s \leq t$.

5. In light of (50) and the fact $u(\cdot,t), \tilde{u}(\cdot,t) \to g$ in L^1 as $t \to 0$, we at last conclude $u = \tilde{u}$ a.e. □

11.5. PROBLEMS

1. Verify that the shallow water equations (Example 3 in §11.1) form a strictly hyperbolic system, provided $\phi > 0$.
2. Define for $z \in \mathbb{R}$, $z \neq 0$, the matrix function
$$\mathbf{B}(z) := e^{-\frac{1}{z^2}}\begin{pmatrix}\cos(\frac{2}{z}) & \sin(\frac{2}{z})\\ \sin(\frac{2}{z}) & -\cos(\frac{2}{z})\end{pmatrix},$$
and set $\mathbf{B}(0) = 0$. Show that $\mathbf{B}$ is C^∞ and has real eigenvalues, but we cannot find unit-length right eigenvectors $\{\mathbf{r}_1(z), \mathbf{r}_2(z)\}$ depending continuously on z near 0. What happens to the eigenspaces as $z \to 0$?
3. Confirm that the functions w^1, w^2 computed for the barotropic gas dynamics in §11.3.1 are indeed Riemann invariants.
4. Suppose that Φ is an entropy for the shallow water equations. Prove
$$\frac{\partial^2\Phi}{\partial v^2} = \phi\frac{\partial^2\Phi}{\partial\phi^2}.$$
5. Show that $\Phi = \rho v^2/2 + P(\rho)$ is an entropy for the barotropic Euler equations (from §11.3.1), provided $P''(\rho) = p'(\rho)/\rho$, $\rho > 0$. Confirm that Φ is convex in the proper variables. What is the corresponding entropy flux Ψ?
6. Assume that u is an entropy solution of the scalar conservation law $u_t + F(u)_x = 0$, and that, as in §3.4.1, u is smooth on either side of a curve $\{x = s(t)\}$.

(i) Prove that along this curve the left and right hand limits of u satisfy the relations:

$$F(z) \geq \frac{F(u_r) - F(u_l)}{u_r - u_l}(z - u_r) + F(u_r) \quad \text{if} \quad u_l \leq z \leq u_r,$$

and

$$F(z) \leq \frac{F(u_r) - F(u_l)}{u_r - u_l}(z - u_r) + F(u_r) \quad \text{if} \quad u_r \leq z \leq u_l.$$

This is the *condition E.*

(ii) What does condition E imply if F is uniformly convex?

11.6. REFERENCES

T.-P. Liu wrote the initial draft of this entire chapter.

Section 11.1 The shallow water equations are discussed in LeVeque [**LV**]. P. Colella helped me with the calculations in §11.1.2. The proof of Theorem 2 follows suggestions of W. Han. See Courant–Friedrichs [**C-F**], Xiao–Zhang [**X-Z**], Zheng [**ZH**] for more.

Section 11.3 I followed Logan [**LO**] for the example.

Section 11.4 See Majda–Pego [**M-P**], Conlon [**CO**], Smoller [**S**] for more about viscous traveling waves. The uniqueness theorem in §11.4.3 is due to Kružkov.

Section 11.5 Problem 2 is taken from Kato [**K**, p. 111].

APPENDICES

APPENDIX A: NOTATION

A.1. Notation for matrices.

(i) We write $A = ((a_{ij}))$ to mean A is an $m \times n$ matrix with $(i,j)^{th}$ entry a_{ij}. (Sometimes, as in §8.1.4, it will be convenient to use superscripts to denote rows.)

A diagonal matrix is denoted $\operatorname{diag}(d_1, \dots, d_n)$.

(ii) $\mathbb{M}^{m\times n}$ = space of real $m \times n$ matrices.

$\mathbb{S}^{n\times n}$ = space of real symmetric $n \times n$ matrices.

(iii) $\operatorname{tr} A$ = trace of the matrix A.

(iv) $\det A$ = determinant of the matrix A.

(v) $\operatorname{cof} A$ = cofactor matrix of A (See §8.1.4).

(vi) A^T = transpose of the matrix A.

(vii) If $A = ((a_{ij}))$ and $B = ((b_{ij}))$ are $m \times n$ matrices, then

$$A : B = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij},$$

and

$$|A| = (A : A)^{\frac{1}{2}} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}.$$

(viii) If $A \in \mathbb{S}^{n \times n}$ and, as below, $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, the corresponding quadratic form is $x \cdot Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j$.

(ix) If $A \in \mathbb{S}^{n \times n}$, we write

$$A \geq \theta I$$

if $x \cdot Ax \geq \theta |x|^2$ for all $x \in \mathbb{R}^n$.

(x) We sometimes will write yA to mean $A^T y$, for $A \in \mathbb{M}^{m \times n}$ and $y \in \mathbb{R}^m$.

A.2. Geometric notation.

(i) $\mathbb{R}^n = n$-dimensional real Euclidean space, $\mathbb{R} = \mathbb{R}^1$.

(ii) $e_i = (0, \dots, 0, 1, \dots, 0) = i^{th}$ standard coordinate vector.

(iii) A typical point in $\mathbb{R}^n$ is $x = (x_1, \dots, x_n)$.
We will also, depending upon the context, regard x as a row or column vector.

(iv) $\mathbb{R}^n_+ = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_n > 0\}$ = open upper half-space.
$\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$.

(v) A typical point in $\mathbb{R}^{n+1}$ will often be denoted as $(x, t) = (x_1, \dots, x_n, t)$, and we usually interpret $t = x_{n+1} =$ *time*.
A point $x \in \mathbb{R}^n$ will sometimes be written $x = (x', x_n)$ for $x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$.

(vi) U, V, and W usually denote open subsets of $\mathbb{R}^n$. We write

$$V \subset\subset U$$

if $V \subset \bar{V} \subset U$ and $\bar{V}$ is compact, and say V is *compactly contained* in U.

(vii) ∂U = boundary of U, $\bar{U} = U \cup \partial U$ = closure of U.

(viii) $U_T = U \times (0, T]$.

(ix) $\Gamma_T = \bar{U}_T - U_T$ = parabolic boundary of U_T.

(x) $B^0(x, r) = \{y \in \mathbb{R}^n \mid |x - y| < r\}$ = *open* ball in $\mathbb{R}^n$ with center x and radius $r > 0$.

(xi) $B(x,r)$ = *closed* ball with center x, radius $r > 0$.

(xii) $C(x,t,r) = \{y \in \mathbb{R}^n, s \in \mathbb{R} \mid |x-y| \leq r, t-r^2 \leq s \leq t\}$ = closed cylinder with top center (x,t), radius $r > 0$, height r^2.

(xiii) $\alpha(n)$ = volume of unit ball $B(0,1)$ in $\mathbb{R}^n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$.
$n\alpha(n)$ = surface area of unit sphere $\partial B(0,1)$ in $\mathbb{R}^n$.

(xiv) If $a = (a_1, \dots, a_n)$ and $b = (b_1, \dots, b_n)$ belong to $\mathbb{R}^n$,

$$a \cdot b = \sum_{i=1}^{n} a_i b_i, \ |a| = \left(\sum_{i=1}^{n} a_i^2\right)^{\frac{1}{2}}.$$

(xv) $\mathbb{C}^n$ = n-dimensional complex space, $\mathbb{C}$ = complex plane.
If $z \in \mathbb{C}$, we write $\mathrm{Re}(z)$ for the real part of z, and $\mathrm{Im}(z)$ for the imaginary part.

A.3. Notation for functions.

(i) If $u : U \to \mathbb{R}$, we write

$$u(x) = u(x_1, \dots, x_n) \quad (x \in U).$$

We say u is *smooth* provided u is infinitely differentiable.

(ii) If u, v are two functions, we write

$$u \equiv v$$

to mean that u is identically equal to v; that is, the functions u, v agree for all values of their arguments.

Let us also set

$$u := v$$

to define u as equaling v.

The support of a function u is denoted

$$\operatorname{spt} u.$$

(iii) $u^+ = \max(u,0)$, $u^- = -\min(u,0)$, $u = u^+ - u^-$, $|u| = u^+ + u^-$.
The *sign* function is

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

(iv) If $\mathbf{u} : U \to \mathbb{R}^m$, we write

$$\mathbf{u}(x) = (u^1(x), \ldots, u^m(x)) \quad (x \in U).$$

The function u^k is the k^{th} *component* of $\mathbf{u}$, $k = 1, \ldots, m$.

(v) If Σ is a smooth $(n-1)$-dimensional surface in $\mathbb{R}^n$, we write

$$\int_\Sigma f \, dS$$

for the integral of f over Σ, with respect to $(n-1)$-dimensional surface measure. If C is a curve in $\mathbb{R}^n$, we denote by

$$\int_C f \, dl$$

the integral of f over C with respect to arclength.

(vi) *Averages*:

$$⨍_{B(x,r)} f dy = \frac{1}{\alpha(n) r^n} \int_{B(x,r)} f dy = \text{ average of } f \text{ over the ball } B(x,r)$$

and

$$⨍_{\partial B(x,r)} f dS = \frac{1}{n\alpha(n) r^{n-1}} \int_{\partial B(x,r)} f dS$$
$$= \text{ average of } f \text{ over the sphere } \partial B(x,r).$$

(vii) $\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E. \end{cases}$ χ_E is the *indicator function* of E.

(viii) A function $u : U \to \mathbb{R}$ is called *Lipschitz continuous* if

$$|u(x) - u(y)| \le C|x - y|$$

for some constant C and all $x, y \in U$. We write

$$\operatorname{Lip}[u] := \sup_{\substack{x,y \in U \\ x \neq y}} \frac{|u(x) - u(y)|}{|x - y|}.$$

(ix) The convolution of the functions f, g is denoted

$$f * g.$$

Notation for derivatives. Assume $u : U \to \mathbb{R}$, $x \in U$.

(i) $\frac{\partial u}{\partial x_i}(x) = \lim_{h\to 0} \frac{u(x+he_i)-u(x)}{h}$, provided this limit exists.

(ii) We usually write u_{x_i} for $\frac{\partial u}{\partial x_i}$.

(iii) Similarly, $\frac{\partial^2 u}{\partial x_i \partial x_j} = u_{x_i x_j}$, $\frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} = u_{x_i x_j x_k}$, etc.

(iv) *Multiindex Notation*:

(a) A vector of the form $\alpha = (\alpha_1, \dots, \alpha_n)$, where each component α_i is a nonnegative integer, is called a *multiindex* of *order*

$$|\alpha| = \alpha_1 + \cdots + \alpha_n.$$

(b) Given a multiindex α, define

$$D^\alpha u(x) := \frac{\partial^{|\alpha|} u(x)}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} u.$$

(c) If k is a nonnegative integer,

$$D^k u(x) := \{D^\alpha u(x) \mid |\alpha| = k\},$$

the set of all partial derivatives of order k. Assigning some ordering to the various partial derivatives, we can also regard $D^k u(x)$ as a point in $\mathbb{R}^{n^k}$.

(d) $$|D^k u| = \left(\sum_{|\alpha|=k} |D^\alpha u|^2\right)^{1/2}.$$

(e) *Special Cases*: If $k = 1$, we regard the elements of Du as being arranged in a vector:

$$Du = (u_{x_1}, \dots, u_{x_n}) = \textit{ gradient vector.}$$

If $k = 2$, we regard the elements of $D^2 u$ as being arranged in a matrix:

$$D^2 u = \begin{pmatrix} \frac{\partial^2 u}{\partial x_1^2} & \cdots & \frac{\partial^2 u}{\partial x_1 \partial x_n} \\ & \ddots & \\ \frac{\partial^2 u}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 u}{\partial x_n^2} \end{pmatrix}_{n\times n} = \textit{ Hessian matrix.}$$

(v) $$\Delta u = \sum_{i=1}^n u_{x_i x_i} = \operatorname{tr}(D^2 u) = \textit{ Laplacian} \text{ of } u.$$

(vi) We sometimes employ a subscript attached to the symbols D, D^2, etc. to denote the variables being differentiated. For example if $u = u(x,y)$ ($x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$), then $D_x u = (u_{x_1}, \dots, u_{x_n})$, $D_y u = (u_{y_1}, \dots, u_{y_m})$.

Function spaces.

(i) $C^k(U) = \{u : U \to \mathbb{R} \mid u \text{ continuous}\}$
$C(\bar{U}) = \{u \in C(U) \mid u \text{ uniformly continuous}\}$
$C^k(U) = \{u : U \to \mathbb{R} \mid u \text{ is } k\text{-times continuously differentiable}\}$
$C^k(\bar{U}) = \{u \in C^k(U) \mid D^\alpha u \text{ is uniformly continuous for all } |\alpha| \leq k\}$. Thus if $u \in C^k(\bar{U})$, then $D^\alpha u$ continuously extends to $\bar{U}$ for each multiindex α, $|\alpha| \leq k$.

(ii) $C^\infty(U) = \{u : U \to \mathbb{R} \mid u \text{ is infinitely differentiable}\} = \bigcap_{k=0}^\infty C^k(U)$
$C^\infty(\bar{U}) = \bigcap_{k=0}^\infty C^k(\bar{U})$.

(iii) $C_c(U)$, $C_c^k(U)$, etc. denote these functions in $C(U)$, $C^k(U)$, etc. with *compact support.*

(iv) $L^p(U) = \{u : U \to \mathbb{R} \mid u \text{ is Lebesgue measurable}, \|u\|_{L^p(U)} < \infty\}$, where

$$\|u\|_{L^p(U)} = \left(\int_U |f|^p dx\right)^{\frac{1}{p}} \quad (1 \leq p < \infty).$$

$L^\infty(U) = \{u : U \to \mathbb{R} \mid u \text{ is Lebesgue measurable}, \|u\|_{L^\infty(U)} < \infty\}$, where

$$\|u\|_{L^\infty(U)} = \text{ess sup}_U |u|.$$

$L^p_{\text{loc}}(U) = \{u : U \to \mathbb{R} \mid v \in L^p(V) \text{ for each } V \subset\subset U\}$.
(See also §D.1.)

(v) $\|Du\|_{L^p(U)} = \||Du|\|_{L^p(U)}$
$\|D^2u\|_{L^p(U)} = \||D^2u|\|_{L^p(U)}$.

(vi) $W^{k,p}(U)$, $H^k(U)$, etc. $(k = 0, 1, 2, \ldots,\ 1 \leq p \leq \infty)$ denote Sobolev spaces: see Chapter 5.

(vii) $C^{k,\beta}(\bar{U})$, $C^{k,\beta}(\bar{U})$ $(k = 0, \ldots, 0 < \beta \leq 1)$ denote Hölder spaces: see Chapter 5.

(viii) *Functions of x and t.* It is occasionally useful to introduce spaces of functions with differing smoothness in the x- and t-variables, although there is no standard notation for such spaces. We will for this book write

$$C_1^2(U_T) = \{u : U_T \to \mathbb{R} \mid u, D_x u, D_x^2 u, u_t \in C(U_T)\}.$$

In particular, if $u \in C_1^2(U_T)$, then $u, D_x u$, etc. are continuous up to the top $U \times \{t = T\}$.

A.4. Vector-valued functions.

(i) If now $m > 1$ and $\mathbf{u} : U \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \dots, u^m)$, we define

$$D^\alpha \mathbf{u} = (D^\alpha u^1, \dots, D^\alpha u^m) \quad \text{for each multiindex } \alpha.$$

Then

$$D^k \mathbf{u} = \{D^\alpha \mathbf{u} \mid |\alpha| = k\}$$

and

$$|D^k \mathbf{u}| = \left(\sum_{|\alpha|=k} |D^\alpha \mathbf{u}|^2 \right)^{1/2},$$

as before.

(ii) In the special case $k = 1$, we write

$$D\mathbf{u} = \begin{pmatrix} \frac{\partial u^1}{\partial x_1} & \dots & \frac{\partial u^1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial u^m}{\partial x_1} & \dots & \frac{\partial u^m}{\partial x_n} \end{pmatrix}_{m \times n} = \textit{gradient matrix.}$$

(iii) If $m = n$, we have

$$\operatorname{div} \mathbf{u} = \operatorname{tr}(D\mathbf{u}) = \sum_{i=1}^n u^i_{x_i} = \text{divergence of } \mathbf{u}.$$

(iv) The spaces $C(U; \mathbb{R}^m)$, $L^p(U; \mathbb{R}^m)$, etc. consist of those functions $\mathbf{u} : U \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \dots, u^m)$, with $u^i \in C(U)$, $L^p(U)$, etc. $(i = 1, \dots, m)$.

Remark on sub- and superscripts. As illustrated above, we will adhere to the convention of setting in boldface *mappings* which take values in $\mathbb{R}^m$ for $m > 1$ (or else in Banach or Hilbert spaces). The component functions of such mappings will be given *super*scripts. On the other hand, a typical *point* $x \in \mathbb{R}^n$ is not boldface and has components with *sub*scripts, $x = (x_1, \dots, x_n)$.

Matrix-valued mappings will also be set in boldface, and their component functions written with either superscripts or a mixture of sub- and superscripts, depending on the context. □

A.5. Notation for estimates.

Constants. We employ the letter C to denote any constant that can be explicitly computed in terms of known quantities. The exact value denoted by C may therefore change from line to line in a given computation. The big advantage is that our calculations will be simpler looking, since we continually absorb "extraneous" factors into the term C.

DEFINITIONS. (i) (Big-oh notation.) *We write*

$$f = O(g) \quad \text{as } x \to x_0,$$

provided there exists a constant C such that

$$|f(x)| \le C|g(x)|$$

for all x sufficiently close to x_0.

(ii) (Little-oh notation.) *We write*

$$f = o(g) \quad \text{as } x \to x_0,$$

provided

$$\lim_{x \to x_0} \frac{|f(x)|}{|g(x)|} = 0.$$

Remark. The expression "$O(g)$" (or "$o(g)$") is not by itself defined. There must always be an accompanying limit, for example "as $x \to x_0$" above, although this limit is often implicit. □

A.6. Some comments about notation.

The foregoing notation is largely standard within the PDE literature, with a few significant exceptions:

(i) We employ the symbol "Du", and not "∇u", to denote the gradient of the function u. The reason is that "D^2u" then naturally denotes the Hessian matrix of u, whereas "$\nabla^2 u$" would be confused with the Laplacian. The multiindex notation also looks better with the letter D.

(ii) Most books and papers on partial differential equations denote by "Ω" the open subset of $\mathbb{R}^n$ within which a given PDE holds.

As indicated above, we will instead mostly use the symbol "U" for such a region. The advantages are several. First of all, since a typical solution is denoted u, it makes sense to denote its domain by U, and not to switch to a Greek letter. Furthermore, once we call a given open set U, the letters V and W are then available for subregions.

Lastly, it is important to save Ω as the standard symbol for a probability space. Many important partial differential equations have probabilistic representation formulas (cf. Freidlin [**FD**]), and although such are beyond the scope of this book, it seems wise to avoid the possibility of future notational confusion.

APPENDIX B: INEQUALITIES

B.1. Convex functions.

Definition. *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is called* convex *provided*

$$f(\tau x + (1-\tau)y) \leq \tau f(x) + (1-\tau)f(y) \tag{1}$$

for all $x, y \in \mathbb{R}^n$ *and each* $0 \leq \tau \leq 1$.

THEOREM 1 (Supporting hyperplanes). *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *is convex. Then for each* $x \in \mathbb{R}^n$ *there exists* $r \in \mathbb{R}^n$ *such that the inequality*

$$f(y) \geq f(x) + r \cdot (y - x) \tag{2}$$

holds for all $y \in \mathbb{R}^n$.

Remarks. (i) The mapping $y \mapsto f(x) + r \cdot (y-x)$ determines the *supporting hyperplane* to f at x. Inequality (2) says the graph of f lies above each supporting hyperplane. If f is differentiable at x, $r = Df(x)$.

(ii) If f is C^2, then f is convex if and only if $D^2 f \geq 0$. The C^2 function f is *uniformly convex* if $D^2 f \geq \theta I$ for some constant $\theta > 0$: this means

$$\sum_{i,j=1}^{n} f_{x_i x_j}(x)\xi_i \xi_j \geq \theta |\xi|^2 \quad (x, \xi \in \mathbb{R}^n).$$

□

THEOREM 2 (Jensen's inequality). *Assume* $f : \mathbb{R} \to \mathbb{R}$ *is convex, and* $U \subset \mathbb{R}^n$ *is open, bounded. Let* $u : U \to \mathbb{R}$ *be summable. Then*

$$f\left(⨍_U u\,dx \right) \leq ⨍_U f(u)\,dx. \tag{3}$$

Remember from §A.3 the notation $⨍_U u\,dx = \frac{1}{|U|}\int_U u\,dx$ = *average* of u over U.

Proof. Since f is convex, for each $p \in \mathbb{R}$ there exists $r \in \mathbb{R}$ such that

$$f(q) \geq f(p) + r(q-p) \quad \text{for all } q \in \mathbb{R}.$$

Let $p = ⨍_U u\,dx$, $q = u(x)$:

$$f(u(x)) \geq f\left(⨍_U u\,dx \right) + r\left(u(x) - ⨍_U u\,dx \right).$$

Integrate with respect to x over U.

□

Convex functions are discussed more fully in §3.3.2 and §9.6.1.

B.2. Elementary inequalities.

Following is a collection of elementary, but fundamental, inequalities. These estimates are continually employed throughout the text and should be memorized.

a. Cauchy's inequality.

$$ab \leq \frac{a^2}{2} + \frac{b^2}{2} \quad (a, b \in \mathbb{R}). \tag{4}$$

Proof. $0 \leq (a-b)^2 = a^2 - 2ab + b^2$. □

b. Cauchy's inequality with ϵ.

$$ab \leq \epsilon a^2 + \frac{b^2}{4\epsilon} \quad (a, b > 0,\ \epsilon > 0). \tag{5}$$

Proof. Write

$$ab = ((2\epsilon)^{1/2} a)\left(\frac{b}{(2\epsilon)^{1/2}}\right)$$

and apply Cauchy's inequality. □

c. Young's inequality. Let $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad (a, b > 0). \tag{6}$$

Proof. The mapping $x \mapsto e^x$ is convex, and consequently

$$ab = e^{\log a + \log b} = e^{\frac{1}{p}\log a^p + \frac{1}{q}\log b^q} \leq \frac{1}{p} e^{\log a^p} + \frac{1}{q} e^{\log b^q} = \frac{a^p}{p} + \frac{b^q}{q}.$$

□

d. Young's inequality with ϵ.

$$ab \leq \epsilon a^p + C(\epsilon) b^q \quad (a, b > 0,\ \epsilon > 0) \tag{7}$$

for $C(\epsilon) = (\epsilon p)^{-q/p} q^{-1}$.

Proof. Write $ab = \left((\epsilon p)^{1/p} a\right)\left(\frac{b}{(\epsilon p)^{1/p}}\right)$ and apply Young's inequality. □

e. Hölder's inequality. Assume $1 \leq p, q \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then if $u \in L^p(U)$, $v \in L^q(U)$, we have

$$\int_U |uv|\,dx \leq \|u\|_{L^p(U)} \|v\|_{L^q(U)}. \tag{8}$$

Proof. By homogeneity, we may assume $\|u\|_{L^p} = \|v\|_{L^q} = 1$. Then Young's inequality implies for $1 < p, q < \infty$ that

$$\int_U |uv| dx \le \frac{1}{q} \int_U |u|^p dx + \frac{1}{q} \int_U |v|^q dx = 1 = \|u\|_{L^p} \|v\|_{L^q}.$$

□

f. Minkowski's inequality. Assume $1 \le p \le \infty$ and $u, v \in L^p(U)$. Then

$$\|u + v\|_{L^p(U)} \le \|u\|_{L^p(U)} + \|v\|_{L^p(U)}. \tag{9}$$

Proof.

$$\begin{aligned} \|u + v\|^p_{L^p(U)} &= \int_U |u + v|^p dx \le \int_U |u + v|^{p-1} (|u| + |v|)\, dx \\ &\le \left(\int_U |u + v|^p dx \right)^{\frac{p-1}{p}} \left(\left(\int |u|^p dx \right)^{1/p} + \left(\int_U |v|^p dx \right)^{1/p} \right) \\ &= \|u + v\|^{p-1}_{L^p(U)} (\|u\|_{L^p(U)} + \|v\|_{L^p(U)}). \end{aligned}$$

□

Remark. Similar proofs establish the discrete versions of Holder's and Minkowski's inequalities:

$$\begin{cases} \left|\sum_{k=1}^n a_k b_k\right| \le \left(\sum_{k=1}^n |a_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^n |b_k|^q\right)^{\frac{1}{q}}, \\ \left(\sum_{k=1}^n |a_k + b_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^n |a_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |b_k|^p\right)^{\frac{1}{p}}, \end{cases} \tag{10}$$

for $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n) \in \mathbb{R}^n$ and $1 \le p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. □

g. General Hölder inequality. Let $1 \le p_1, \dots, p_m \le \infty$, with $\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_m} = 1$, and assume $u_k \in L^{p_k}(U)$ for $k = 1, \dots, m$. Then

$$\int_U |u_1 \cdots u_m|\, dx \le \prod_{k=1}^m \|u_i\|_{L^{p_k}(U)}. \tag{11}$$

Proof. Induction, using Hölder's inequality. □

h. Interpolation inequality for L^p-norms. Assume $1 \le s \le r \le t \le \infty$ and

$$\frac{1}{r} = \frac{\theta}{s} + \frac{(1 - \theta)}{t}.$$

Suppose also $u \in L^s(U) \cap L^t(U)$. Then $u \in L^r(U)$, and

$$\|u\|_{L^r(U)} \le \|u\|^\theta_{L^s(U)} \|u\|^{1-\theta}_{L^t(U)}. \tag{12}$$

Proof. We compute

$$\int_U |u|^r dx = \int_U |u|^{\theta r}|u|^{(1-\theta)r} dx$$
$$\leq \left(\int_U |u|^{\theta r \frac{s}{\theta r}} dx\right)^{\frac{\theta r}{s}} \left(\int_U |u|^{(1-\theta)r\frac{t}{(1-\theta)r}} dx\right)^{\frac{(1-\theta)r}{t}}.$$

We have invoked Hölder's inequality, which applies since $\frac{\theta r}{s} + \frac{(1-\theta)r}{t} = 1$. □

i. Cauchy–Schwarz inequality.

$$(13) \qquad |x \cdot y| \leq |x||y| \quad (x, y \in \mathbb{R}^n).$$

Proof. Let $\epsilon > 0$ and note

$$0 \leq |x \pm \epsilon y|^2 = |x|^2 \pm 2\epsilon x \cdot y + \epsilon^2 |y|^2.$$

Consequently

$$\pm x \cdot y \leq \frac{1}{2\epsilon}|x|^2 + \frac{\epsilon}{2}|y|^2.$$

Minimize the right hand side by setting $\epsilon = \frac{|x|}{|y|}$, provided $y \neq 0$. □

Remark. Likewise, if A is a symmetric, nonnegative $n \times n$ matrix,

$$(14) \qquad \left|\sum_{i,j=1}^n a_{ij}x_i y_j\right| \leq \left(\sum_{i,j=1}^n a_{ij}x_i x_j\right)^{1/2} \left(\sum_{i,j=1}^n a_{ij}y_i y_j\right)^{1/2} \quad (x, y \in \mathbb{R}^n).$$

□

j. Gronwall's inequality (differential form).

(i) Let $\eta(\cdot)$ be a nonnegative, absolutely continuous function on $[0, T]$, which satisfies for a.e. t the differential inequality

$$(15) \qquad \eta'(t) \leq \phi(t)\eta(t) + \psi(t),$$

where $\phi(t)$ and $\psi(t)$ are nonnegative, summable functions on $[0, T]$. Then

$$(16) \qquad \eta(t) \leq e^{\int_0^t \phi(s)\,ds}\left[\eta(0) + \int_0^t \psi(s)\,ds\right]$$

for all $0 \leq t \leq T$.

(ii) In particular, if

$$\eta' \leq \phi\eta \text{ on } [0, T] \text{ and } \eta(0) = 0,$$

then

$$\eta \equiv 0 \text{ on } [0, T].$$

Proof. From (15) we see

$$\frac{d}{ds}\left(\eta(s)e^{-\int_0^s \phi(r)\,dr}\right) = e^{-\int_0^s \phi(r)\,dr}(\eta'(s) - \phi(s)\eta(s)) \leq e^{-\int_0^s \phi(r)\,dr}\psi(s)$$

for a.e. $0 \leq s \leq T$. Consequently for each $0 \leq t \leq T$, we have

$$\eta(t)e^{-\int_0^t \phi(r)\,dr} \leq \eta(0) + \int_0^t e^{-\int_0^s \phi(r)\,dr}\psi(s)\,ds \leq \eta(0) + \int_0^t \psi(s)\,ds.$$

This implies inequality (16). □

k. Gronwall's inequality (integral form).

(i) Let $\xi(t)$ be a nonnegative, summable function on $[0,T]$ which satisfies for a.e. t the integral inequality

$$\xi(t) \leq C_1 \int_0^t \xi(s)\,ds + C_2 \tag{17}$$

for constants $C_1, C_2 \geq 0$. Then

$$\xi(t) \leq C_2(1 + C_1 t e^{C_1 t}) \tag{18}$$

for a.e. $0 \leq t \leq T$.

(ii) In particular, if

$$\xi(t) \leq C_1 \int_0^t \xi(s)\,ds$$

for a.e. $0 \leq t \leq T$, then

$$\xi(t) = 0 \quad \text{a.e.}$$

Proof. Let $\eta(t) := \int_0^t \xi(s)\,ds$; then $\eta \leq C_1\eta + C_2$ a.e. in $[0,T]$. According to the differential form of Gronwall's inequality above:

$$\eta(t) \leq e^{C_1 t}(\eta(0) + C_2 t) = C_2 t e^{C_1 t}.$$

Then (17) implies

$$\xi(t) \leq C_1\eta(t) + C_2 \leq C_2(1 + C_1 t e^{C_1 t}).$$

□

APPENDIX C: CALCULUS FACTS

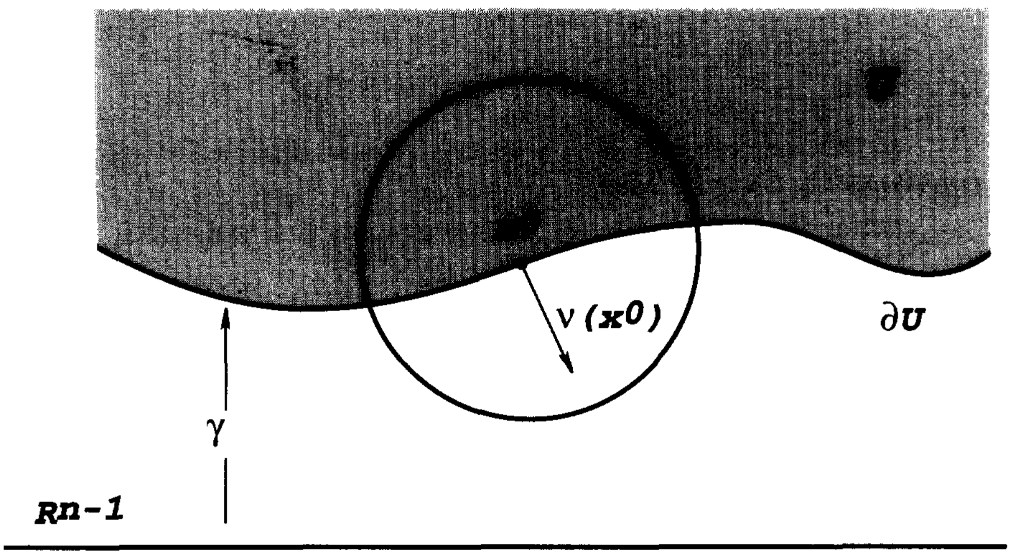
C.1. Boundaries.

Let $U \subset \mathbb{R}^n$ be open and bounded, $k \in \{1, 2, \dots\}$.

DEFINITION. *We say ∂U is C^k if for each point $x^0 \in \partial U$ there exist $r > 0$ and a C^k function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that—upon relabeling and reorienting the coordinates axes if necessary—we have*

$$U \cap B(x^0, r) = \{x \in B(x^0, r) \mid x_n > \gamma(x_1, \dots, x_{n-1})\}.$$

Likewise, ∂U is C^∞ if ∂U is C^k for $k = 1, 2, \dots$, and ∂U is analytic *if the mapping γ is analytic.*



The boundary of U

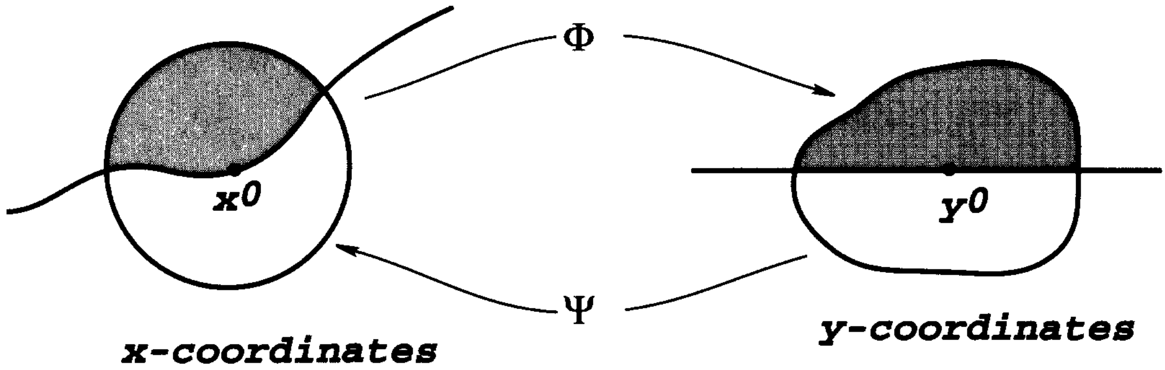
DEFINITIONS. (i) *If ∂U is C^1, then along ∂U is defined the outward pointing* unit normal vector field

$$\boldsymbol{\nu} = (\nu^1, \dots, \nu^n).$$

The unit normal *at any point $x^0 \in \partial U$ is $\boldsymbol{\nu}(x^0) = \nu = (\nu_1, \dots, \nu_n)$.*

(ii) *Let $u \in C^1(\bar{U})$. We call*

$$\frac{\partial u}{\partial \nu} := \boldsymbol{\nu} \cdot Du$$



Straightening out the boundary

the (outward) normal derivative *of* u.

We will frequently need to change coordinates near a point of ∂U so as to "flatten out" the boundary. To be more specific, fix $x^0 \in \partial U$, and choose r, γ, etc. as above. Define then

$$\begin{cases} y_i = x_i =: \Phi^i(x) & (i = 1, \dots, n-1) \\ y_n = x_n - \gamma(x_1, \dots, x_{n-1}) =: \Phi^n(x), \end{cases}$$

and write

$$y = \mathbf{\Phi}(x).$$

Similarly, we set

$$\begin{cases} x_i = y_i =: \Psi^i(y) & (i = 1, \dots, n-1) \\ x_n = y_n + \gamma(y_1, \dots, y_n) =: \Psi^n(y), \end{cases}$$

and write

$$x = \mathbf{\Psi}(y).$$

Then $\mathbf{\Phi} = \mathbf{\Psi}^{-1}$, and the mapping $x \mapsto \mathbf{\Phi}(x) = y$ "straightens out ∂U" near x^0. Observe also that $\det \mathbf{\Phi} = \det \mathbf{\Psi} = 1$.

C.2. Gauss-Green Theorem.

In this section we assume U is a bounded, open subset of $\mathbb{R}^n$, and ∂U is C^1.

THEOREM 1 (Gauss-Green Theorem). *Suppose* $u \in C^1(\bar{U})$*. Then*

(1) $$\int_U u_{x_i}\,dx = \int_{\partial U} u\nu^i\,dS \quad (i = 1, \dots, n).$$

THEOREM 2 (Integration-by-parts formula). *Let $u, v \in C^1(\bar{U})$. Then*

$$\int_U u_{x_i} v \, dx = -\int_U u v_{x_i} \, dx + \int_{\partial U} u v \nu^i \, dS \quad (i = 1, \dots, n). \tag{2}$$

Proof. Apply Theorem 1 to uv. □

THEOREM 3 (Green's formulas). *Let $u, v \in C^2(\bar{U})$. Then*

(i) $\int_U \Delta u \, dx = \int_{\partial U} \frac{\partial u}{\partial \nu} \, dS$,

(ii) $\int_U Dv \cdot Du \, dx = -\int_U u \Delta v \, dx + \int_{\partial U} \frac{\partial v}{\partial \nu} u \, dS$,

(iii) $\int_U u \Delta v - v \Delta u \, dx = \int_{\partial U} u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \, dS$.

Proof. Using (2), with u_{x_i} in place of u and $v \equiv 1$, we see

$$\int_U u_{x_i x_i} \, dx = \int_{\partial U} u_{x_i} \nu^i \, dS.$$

Sum $i = 1, \dots, n$ to establish (i).

To derive (ii), we employ (2) with $v = u_{x_i}$. Write (ii) with u and v interchanged and then subtract, to obtain (iii). □

C.3. Polar coordinates, coarea formula.

Next we convert n-dimensional integrals into integrals over spheres.

THEOREM 4 (Polar coordinates).

(i) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous and summable. Then*

$$\int_{\mathbb{R}^n} f \, dx = \int_0^\infty \left(\int_{\partial B(x_0, r)} f \, dS \right) dr$$

for each point $x_0 \in \mathbb{R}^n$.

(ii) *In particular*

$$\frac{d}{dr} \left(\int_{B(x_0, r)} f \, dx \right) = \int_{\partial B(x_0, r)} f \, dS \tag{ii}$$

for each $r > 0$.

Theorem 4 is a special case of

THEOREM 5 (Coarea formula). *Let $u : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz continuous and assume that for a.e. $r \in \mathbb{R}$ the level set*

$$\{x \in \mathbb{R}^n \mid u(x) = r\}$$

is a smooth, $(n-1)$-dimensional hypersurface in $\mathbb{R}^n$. Suppose also $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and summable. Then

$$\int_{\mathbb{R}^n} f|Du|dx = \int_{-\infty}^{\infty} \left(\int_{\{u=r\}} f dS \right) dr.$$

Theorem 4 follows from Theorem 5 by taking $u(x) = |x - x_0|$. See [**E-G,** Chapter 3] for more on the coarea formula. The word "coarea" is pronounced, and sometimes spelled "co-area".

C.4. Convolution and smoothing.

We next introduce tools that will allow us to build smooth approximations to given functions.

Notation. If $U \subset \mathbb{R}^n$ is open, $\epsilon > 0$, write $U_\epsilon := \{x \in U \mid \operatorname{dist}(x, \partial U) > \epsilon\}$. □

DEFINITIONS. (i) *Define $\eta \in C^\infty(\mathbb{R}^n)$ by*

$$\eta(x) := \begin{cases} C \exp\left(\frac{1}{|x|^2-1}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

the constant $C > 0$ selected so that $\int_{\mathbb{R}^n} \eta\, dx = 1$.

(ii) *For each $\epsilon > 0$, set*

$$\eta_\epsilon(x) := \frac{1}{\epsilon^n} \eta\left(\frac{x}{\epsilon}\right).$$

We call η the *standard mollifier.* The functions η_ϵ are C^∞ and satisfy

$$\int_{\mathbb{R}^n} \eta_\epsilon\, dx = 1, \ \operatorname{spt}(\eta_\epsilon) \subset B(0, \epsilon).$$

DEFINITION. *If $f : U \to \mathbb{R}$ is locally integrable, define its* mollification

$$f^\epsilon := \eta_\epsilon * f \quad \text{in } U_\epsilon.$$

That is,

$$f^\epsilon(x) = \int_U \eta_\epsilon(x - y) f(y)\, dy = \int_{B(0,\epsilon)} \eta_\epsilon(y) f(x - y)\, dy$$

for $x \in U_\epsilon$.

THEOREM 6 (Properties of mollifiers).

(i) $f^\epsilon \in C^\infty(U_\epsilon)$.

(ii) $f^\epsilon \to f$ *a.e. as* $\epsilon \to 0$.

(iii) *If* $f \in C(U)$, *then* $f^\epsilon \to f$ *uniformly on compact subsets of* U.

(iv) *If* $1 \leq p < \infty$ *and* $f \in L^p_{\text{loc}}(U)$, *then* $f^\epsilon \to f$ *in* $L^p_{\text{loc}}(U)$.

Proof. 1. Fix $x \in U_\epsilon$, $i \in \{1, \dots, n\}$, and h so small that $x + he_i \in U_\epsilon$. Then

$$\begin{aligned} \frac{f^\epsilon(x+he_i) - f^\epsilon(x)}{h} &= \frac{1}{\epsilon^n} \int_U \frac{1}{h} \left[\eta\left(\frac{x+he_i-y}{\epsilon}\right) - \eta\left(\frac{x-y}{\epsilon}\right) \right] f(y)\, dy \\ &= \frac{1}{\epsilon^n} \int_V \frac{1}{h} \left[\eta\left(\frac{x+he_i-y}{\epsilon}\right) - \eta\left(\frac{x-y}{\epsilon}\right) \right] f(y)\, dy \end{aligned}$$

for some open set $V \subset\subset U$. As

$$\frac{1}{h} \left[\eta\left(\frac{x+he_i-y}{\epsilon}\right) - \eta\left(\frac{x-y}{\epsilon}\right) \right] \to \frac{1}{\epsilon} \frac{\partial \eta}{\partial x_i}\left(\frac{x-y}{\epsilon}\right)$$

uniformly on V, $\frac{\partial f^\epsilon}{\partial x_i}(x)$ exists and equals

$$\int_U \frac{\partial \eta_\epsilon}{\partial x_i}(x-y) f(y)\, dy.$$

A similar argument shows that $D^\alpha f^\epsilon(x)$ exists, and

$$D^\alpha f^\epsilon(x) = \int_U D^\alpha \eta_\epsilon(x-y) f(y)\, dy \quad (x \in U_\epsilon),$$

for each multiindex α. This proves (i).

2. According to Lebesgue's Differentiation Theorem (§E.4),

$$\lim_{r\to 0} ⨍_{B(x,r)} |f(y) - f(x)|\, dy = 0 \tag{3}$$

for a.e. $x \in U$. Fix such a point x. Then

$$\begin{aligned} |f^\epsilon(x) - f(x)| &= \left| \int_{B(x,\epsilon)} \eta^\epsilon(x-y)[f(y) - f(x)]\, dy \right| \\ &\leq \frac{1}{\epsilon^n} \int_{B(x,\epsilon)} \eta\left(\frac{x-y}{\epsilon}\right) |f(y) - f(x)|\, dy \\ &\leq C ⨍_{B(x,\epsilon)} |f(y) - f(x)|\, dy \to 0 \quad \text{as } \epsilon \to 0, \end{aligned}$$

by (3). Assertion (ii) follows.

3. Assume now $f \in C(U)$. Given $V \subset\subset U$ we choose $V \subset\subset W \subset\subset U$ and note that f is uniformly continuous on W. Thus the limit (3) holds uniformly for $x \in V$. Consequently the calculation above implies $f^\epsilon \to f$ uniformly on V.

4. Next, assume $1 \leq p < \infty$ and $f \in L^p_{\text{loc}}(U)$. Choose an open set $V \subset\subset U$ and, as above, an open set W so that $V \subset\subset W \subset\subset U$. We claim that for sufficiently small $\epsilon > 0$

$$\|f^\epsilon\|_{L^p(V)} \leq \|f\|_{L^p(W)}. \tag{4}$$

To see this, we note that if $1 < p < \infty$ and $x \in V$,

$$\begin{aligned}
|f^\epsilon(x)| &= \left| \int_{B(x,\epsilon)} \eta_\epsilon(x-y) f(y)\, dy \right| \\
&\leq \int_{B(x,\epsilon)} \eta_\epsilon^{1-1/p}(x-y)\eta_\epsilon^{1/p}(x-y)|f(y)|\, dy \\
&\leq \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)\, dy \right)^{1-1/p} \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|^p dy \right)^{1/p}.
\end{aligned}$$

Since $\int_{B(x,\epsilon)} \eta_\epsilon(x-y)\, dy = 1$, this inequality implies

$$\begin{aligned}
\int_V |f^\epsilon(x)|^p dx &\leq \int_V \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|^p dy \right) dx \\
&\leq \int_W |f(y)|^p \left(\int_{B(y,\epsilon)} \eta_\epsilon(x-y)\, dx \right) dy = \int_W |f(y)|^p dy,
\end{aligned}$$

provided $\epsilon > 0$ is sufficiently small. This is inequality (4).

5. Now fix $V \subset\subset W \subset\subset U$, $\delta > 0$, and choose $g \in C(W)$ so that

$$\|f - g\|_{L^p(W)} < \delta.$$

Then

$$\begin{aligned}
\|f^\epsilon - f\|_{L^p(V)} &\leq \|f^\epsilon - g^\epsilon\|_{L^p(V)} + \|g^\epsilon - g\|_{L^p(V)} + \|g - f\|_{L^p(V)} \\
&\leq 2\|f - g\|_{L^p(W)} + \|g^\epsilon - g\|_{L^p(V)} \quad \text{by (4)} \\
&\leq 2\delta + \|g^\epsilon - g\|_{L^p(V)}.
\end{aligned}$$

Since $g^\epsilon \to g$ uniformly on V, we have $\limsup_{\epsilon\to 0} \|f^\epsilon - f\|_{L^p(V)} \leq 2\delta$. □

C.5. Inverse Function Theorem.

Let $U \subset \mathbb{R}^n$ be an open set and suppose $\mathbf{f} : U \to \mathbb{R}^n$ is C^1, $\mathbf{f} = (f^1, \dots, f^n)$. Assume $x_0 \in U$, $z_0 = \mathbf{f}(x_0)$.

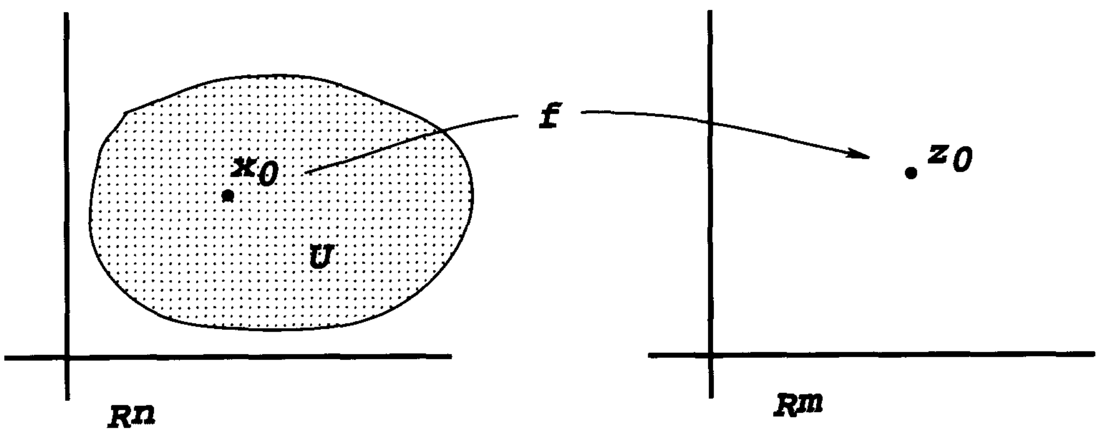
Notation. Remember from §A.4 that we write

$$D\mathbf{f} = \begin{pmatrix} f^1_{x_1} & \dots & f^1_{x_n} \\ & \ddots & \\ f^n_{x_1} & \dots & f^n_{x_n} \end{pmatrix}_{n \times n} = \textit{gradient matrix} \text{ of } \mathbf{f}.$$

□

DEFINITION.

$$J\mathbf{f} = \textit{ Jacobian of } \mathbf{f} = |\det D\mathbf{f}| = \left| \frac{\partial(f^1, \dots, f^n)}{\partial(x_1, \dots, x_n)} \right|.$$



THEOREM 7 (Inverse Function Theorem). *Assume* $\mathbf{f} \in C^1(U; \mathbb{R}^n)$ *and*

$$J\mathbf{f}(x_0) \neq 0.$$

Then there exist an open set $V \subset U$, with $x_0 \in V$, and an open set $W \subset \mathbb{R}^n$, with $z_0 \in W$, such that

(i) *the mapping*

$$\mathbf{f} : V \to W$$

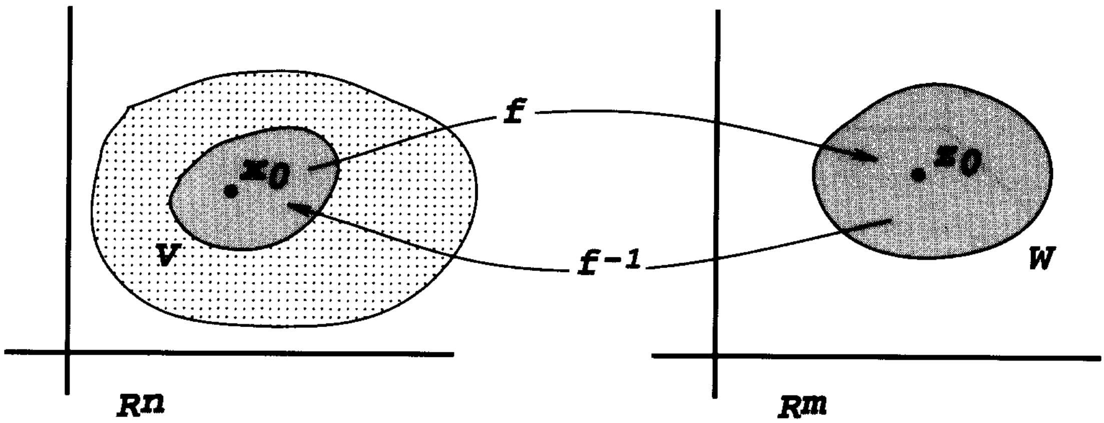
is one-to-one and onto, and

(ii) *the inverse function*

$$\mathbf{f}^{-1} : W \to V$$

is C^1.

(iii) *If* $\mathbf{f} \in C^k$, *then* $\mathbf{f}^{-1} \in C^k$ $(k = 2, \dots)$.



C.6. Implicit Function Theorem.

Let n, m be positive integers.

Notation. We write a typical point in $\mathbb{R}^{n+m}$ as

$$(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m)$$

for $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. □

Let $U \subset \mathbb{R}^{n+m}$ be an open set and suppose $\mathbf{f} : U \to \mathbb{R}^m$ is C^1, $\mathbf{f} = (f^1, \dots, f^m)$. Assume $(x_0, y_0) \in U$, $z_0 = \mathbf{f}(x_0, y_0)$.

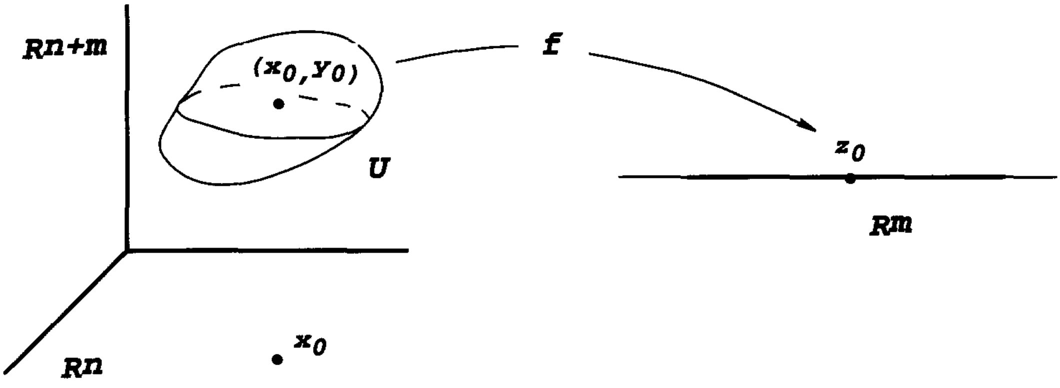
Notation.

$$D\mathbf{f} = \begin{pmatrix} f^1_{x_1} & \dots & f^1_{x_n} & f^1_{y_1} & \dots & f^1_{y_m} \\ & \ddots & & & \ddots & \\ f^m_{x_1} & \dots & f^m_{x_n} & f^m_{y_1} & \dots & f^m_{y_m} \end{pmatrix}_{m \times (n+m)}$$
$$= (D_x\mathbf{f}, D_y\mathbf{f}) = \textit{ gradient matrix of } \mathbf{f}.$$

□

DEFINITION.

$$J_y\mathbf{f} = |\det D_y\mathbf{f}| = \left| \frac{\partial(f^1, \dots, f^m)}{\partial(y_1, \dots, y_m)} \right|.$$



THEOREM 8 (Implicit Function Theorem). *Assume* $\mathbf{f} \in C^1(U;\mathbb{R}^m)$ *and*

$$J_y\mathbf{f}(x_0, y_0) \neq 0.$$

Then there exists an open set $V \subset U$, *with* $(x_0, y_0) \in V$, *an open set* $W \subset \mathbb{R}^n$, *with* $x_0 \in W$, *and a* C^1 *mapping* $\mathbf{g} : W \to \mathbb{R}^m$ *such that*

(i) $\mathbf{g}(x_0) = y_0$,

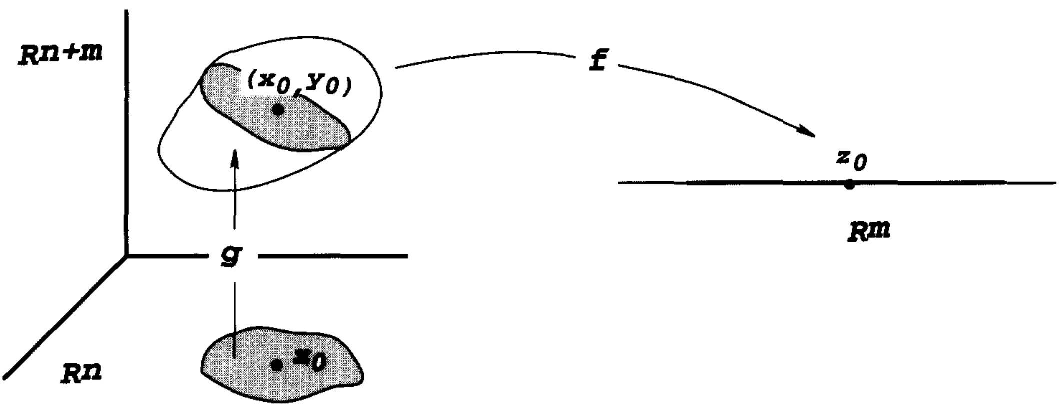
(ii) $\mathbf{f}(x, \mathbf{g}(x)) = z_0 \quad (x \in W)$,

and

(iii) *if* $(x, y) \in V$ *and* $\mathbf{f}(x, y) = z_0$, *then* $y = \mathbf{g}(x)$.

(iv) *If* $\mathbf{f} \in C^k$, *then* $\mathbf{g} \in C^k$ $(k = 2, \dots)$.

The function $\mathbf{g}$ is *implicitly defined* near x_0 by the equation $\mathbf{f}(x, y) = z_0$.



C.7. Uniform convergence.

We record here the *Arzela-Ascoli compactness criterion* for uniform convergence:

Suppose that $\{f_k\}_{k=1}^{\infty}$ is a sequence of real-valued functions defined on $\mathbb{R}^n$, such that

$$|f_k(x)| \leq M \quad (k = 1, \ldots, \ x \in \mathbb{R}^n)$$

for some constant M, and the $\{f_k\}_{k=1}^{\infty}$ are *uniformly equicontinuous.* Then there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty} \subseteq \{f_k\}_{k=1}^{\infty}$ and a continuous function f, such that

$$f_{k_j} \to f \quad \text{uniformly on compact subsets of } \mathbb{R}^n.$$

To say the $\{f_k\}_{k=1}^{\infty}$ are uniformly equicontinuous means that for each $\varepsilon > 0$, there exists $\delta > 0$, such that $|x - y| < \delta$ implies $|f_k(x) - f_k(y)| < \varepsilon$, for $x, y \in \mathbb{R}^n, \ k = 1, \ldots$.

APPENDIX D: LINEAR FUNCTIONAL ANALYSIS

D.1. Banach spaces.

Let X denote a real linear space.

DEFINITION. *A mapping* $\| \quad \| : X \to [0, \infty)$ *is called a* norm *if*

(i) $\|u + v\| \leq \|u\| + \|v\|$ *for all* $u, v \in X$.
(ii) $\|\lambda u\| = |\lambda| \|u\|$ *for all* $u \in X$, $\lambda \in \mathbb{R}$.
(iii) $\|u\| = 0$ *if and only if* $u = 0$.

Inequality (i) is the *triangle inequality.*

Hereafter we assume X is a normed linear space.

DEFINITION. *We say a sequence* $\{u_k\}_{k=1}^{\infty} \subset X$ converges *to* $u \in X$, *written*

$$u_k \to u,$$

if

$$\lim_{k \to \infty} \|u_k - u\| = 0.$$

DEFINITIONS. (i) *A sequence* $\{u_k\}_{k=1}^{\infty} \subset X$ *is called a* Cauchy sequence *provided for each* $\epsilon > 0$ *there exists* $N > 0$ *such that*

$$\|u_k - u_l\| < \epsilon \quad \textit{for all } k, l \geq N.$$

(ii) X *is* complete *if each Cauchy sequence in* X *converges; that is, whenever* $\{u_k\}_{k=1}^{\infty}$ *is a Cauchy sequence, there exists* $u \in X$ *such that* $\{u_k\}_{k=1}^{\infty}$ *converges to* u.

(iii) *A* Banach space X *is a complete, normed linear space.*

DEFINITION. *We say X is* separable *if X contains a countable dense subset.*

Examples. (i) L^p spaces. Assume U is an open subset of $\mathbb{R}^n$, and $1 \leq p \leq \infty$. If $f : U \to \mathbb{R}$ is measurable, we define

$$\|f\|_{L^p(U)} := \begin{cases} \left(\int_U |f|^p dx\right)^{1/p} & \text{if } 1 \leq p < \infty \\ \operatorname{ess\,sup}_U |f| & \text{if } p = \infty. \end{cases}$$

We define $L^p(U)$ to be the linear space of all measurable functions $f : U \to \mathbb{R}$ for which $\|f\|_{L^p(U)} < \infty$. Then $L^p(U)$ is a Banach space, provided we identify two functions which agree a.e.

(ii) Hölder spaces. See §5.1.

(iii) Sobolev spaces. See §5.2. □

D.2. Hilbert spaces.

Let H be a real linear space.

DEFINITION. *A mapping $(\, , \,) : H \times H \to \mathbb{R}$ is called an* inner product *if*

(i) $(u, v) = (v, u)$ *for all* $u, v \in H$,
(ii) *the mapping* $u \mapsto (u, v)$ *is linear for each* $v \in H$,
(iii) $(u, u) \geq 0$ *for all* $u \in H$,
(iv) $(u, u) = 0$ *if and only if* $u = 0$.

Notation. If $(\, , \,)$ is an inner product, the associated *norm* is

(1) $$\|u\| := (u, u)^{1/2} \quad (u \in H).$$

□

The *Cauchy-Schwarz inequality* states

(2) $$|(u, v)| \leq \|u\| \|v\| \quad (u, v \in H).$$

This inequality is proved as in §B.2. Using (2), we easily verify (1) defines a norm on H.

DEFINITION. *A* Hilbert space *H is a Banach space endowed with an inner product which generates the norm.*

Examples. a. The space $L^2(U)$ is a Hilbert space, with

$$(f, g) = \int_U fg \, dx.$$

b. The Sobolev space $H^1(U)$ is a Hilbert space, with

$$(f,g) = \int_U fg + Df \cdot Dg \, dx.$$

DEFINITIONS. (i) *Two elements* $u, v \in H$ *are* orthogonal *if* $(u,v) = 0$.

(ii) *A countable basis* $\{w_k\}_{k=1}^\infty \subset H$ *is called* orthonormal *if*

$$\begin{cases} (w_k, w_l) = 0 & (k, l = 1, \dots; k \neq l) \\ \|w_k\| = 1 & (k = 1, \dots). \end{cases}$$

If $u \in H$ and $\{w_k\}_{k=1}^\infty \subset H$ is an orthonormal basis, we can write

$$u = \sum_{k=1}^\infty (u, w_k) w_k,$$

the series converging in H. In addition

$$\|u\|^2 = \sum_{k=1}^\infty (u, w_k)^2.$$

DEFINITION. *If* S *is a subspace of* H, $S^\perp = \{u \in H \mid (u,v) = 0$ *for all* $v \in S\}$ *is the* subspace orthogonal to S.

D.3. Bounded linear operators.

a. Linear operators on Banach spaces.

Let X and Y be real Banach spaces.

DEFINITIONS. (i) *A mapping* $A : X \to Y$ *is a* linear operator *provided*

$$A[\lambda u + \mu v] = \lambda A u + \mu A v$$

for all $u, v \in X$, $\lambda, \mu \in \mathbb{R}$.

(ii) *The* range *of* A *is* $R(A) := \{v \in Y \mid v = Au$ *for some* $u \in X\}$ *and the* null space *of* A *is* $N(A) := \{u \in X \mid Au = 0\}$. □

DEFINITION. *A linear operator* $A : X \to Y$ *is* bounded *if*

$$\|A\| := \sup\{\|Au\|_Y \mid \|u\|_X \leq 1\} < \infty.$$

It is easy to check that a bounded linear operator $A : X \to Y$ is continuous.

DEFINITION. *A linear operator $A : X \to Y$ is called* closed *if whenever $u_k \to u$ in X and $Au_k \to v$ in Y, then*

$$Au = v.$$

THEOREM 1 (Closed Graph Theorem). *Let $A : X \to Y$ be a closed, linear operator. Then A is bounded.*

DEFINITIONS. *Let $A : X \to X$ be a bounded linear operator.*

(i) *The* resolvent set *of A is*

$$\rho(A) = \{\eta \in \mathbb{R} \mid (A - \eta I) \text{ is one-to-one and onto}\}.$$

(ii) *The* spectrum *of A is*

$$\sigma(A) = \mathbb{R} - \rho(A).$$

If $\eta \in \rho(A)$, the Closed Graph Theorem then implies that the inverse $(A - \eta I)^{-1} : X \to X$ is a bounded linear operator.

DEFINITIONS. (i) *We say $\eta \in \sigma(A)$ is an* eigenvalue *of A provided*

$$N(A - \eta I) \neq \{0\}.$$

We write $\sigma_p(A)$ to denote the collection of eigenvalues of A; $\sigma_p(A)$ is the point spectrum.

(ii) *If η is an eigenvalue and $w \neq 0$ satisfies*

$$Aw = \eta w,$$

we say w is an associated eigenvector.

DEFINITIONS. (i) *A bounded linear operator $u^* : X \to \mathbb{R}$ is called a* bounded linear functional *on X.*

(ii) *We write X^* to denote the collection of all bounded linear functionals on X; X^* is the* dual space *of X.*

DEFINITIONS. (i) *If $u \in X$, $u^* \in X^*$ we write*

$$\langle u^*, u \rangle$$

to denote the real number $u^*(u)$. The symbol $\langle\ ,\ \rangle$ denotes the pairing *of X^* and X.*

(ii) *We define*

$$\|u^*\| := \sup\{\langle u^*, u\rangle \mid \|u\| \leq 1\}.$$

(iii) *A Banach space is* reflexive *if $(X^*)^* = X$. More precisely, this means that for each $u^{**} \in (X^*)^*$, there exists $u \in X$ such that*

$$\langle u^{**}, u^* \rangle = \langle u^*, u \rangle \quad \text{for all } u^* \in X^*.$$

b. Linear operators on Hilbert spaces.

Now let H be a real Hilbert space, with inner product $(\ ,\ )$.

THEOREM 2 (Riesz Representation Theorem). *H^* can be canonically identified with H; more precisely, for each $u^* \in H^*$ there exists a unique element $u \in H$ such that*

$$\langle u^*, v\rangle = (u, v) \quad \textit{for all } v \in H.$$

The mapping $u^* \mapsto u$ is a linear isomorphism of H^* onto H.

DEFINITIONS. (i) *If $A : H \to H$ is a bounded, linear operator, its* adjoint *$A^* : H \to H$ satisfies*

$$(Au, v) = (u, A^*v)$$

for all $u, v \in H$.

(ii) *A is* symmetric *if $A^* = A$.*

D.4. Weak convergence.

Let X denote a real Banach space.

DEFINITION. *We say a sequence $\{u_k\}_{k=1}^\infty \subset X$* converges weakly *to $u \in X$, written*

$$u_k \rightharpoonup u,$$

if

$$\langle u^*, u_k\rangle \to \langle u^*, u\rangle$$

for each bounded linear functional $u^* \in X^*$.

It is easy to check that if $u_k \to u$, then $u_k \rightharpoonup u$. It is also true that *any weakly convergent sequence is bounded.* In addition, if $u_k \rightharpoonup u$, then

$$\|u\| \leq \liminf_{k\to\infty} \|u_k\|.$$

THEOREM 3 (Weak compactness). *Let X be a reflexive Banach space and suppose the sequence $\{u_k\}_{k=1}^\infty \subset X$ is bounded. Then there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$ and $u \in X$ such that*

$$u_{k_j} \rightharpoonup u.$$

In other words, bounded sequences in a reflexive Banach space are weakly precompact. In particular, *a bounded sequence in a Hilbert space contains a weakly convergent subsequence.*

Mazur's Theorem asserts that a convex, closed subset of X is weakly closed.

IMPORTANT EXAMPLE. We will most often employ weak convergence ideas in the following context. Take $U \subset \mathbb{R}^n$ to be open, and assume $1 \leq p < \infty$. Then

$$\text{the dual space of } X = L^p(U) \text{ is } X^* = L^q(U),$$

where $\frac{1}{p} + \frac{1}{q} = 1$, $1 < q \leq \infty$. More precisely, each bounded linear functional on $L^p(U)$ can be represented as $f \mapsto \int_U gf\,dx$ for some $g \in L^q(U)$. Therefore

$$f_k \rightharpoonup f \quad \text{weakly in } L^p(U)$$

means

$$\text{(3)} \qquad \int_U g f_k \, dx \to \int_U g f \, dx \quad \text{as } k \to \infty, \text{ for all } g \in L^q(U).$$

Now the identification of $L^q(U)$ as the dual of $L^p(U)$ shows that

$$L^p(U) \text{ is reflexive if } 1 < p < \infty.$$

In particular Theorem 3 then assures us that from a bounded sequence in $L^p(U)$ $(1 < p < \infty)$ we can extract a *weakly* convergent subsequence, that is, a sequence satisfying (3). This is an important compactness assertion, but note very carefully: *the convergence (3) does not imply that* $f_k \to f$ *pointwise or almost everywhere.* It may very well be, for example, that the functions $\{f_k\}_{k=1}^\infty$ oscillate more and more rapidly as $k \to \infty$. See Problem 1 in Chapter 8. □

D.5. Compact operators, Fredholm theory.

Let X and Y be real Banach spaces.

DEFINITION. *A bounded linear operator*

$$K : X \to Y$$

is called compact *provided for each bounded sequence* $\{u_k\}_{k=1}^\infty \subset X$, *the sequence* $\{Ku_k\}_{k=1}^\infty$ *is precompact in* Y; *that is, there exists a subsequence* $\{u_{k_j}\}_{j=1}^\infty$ *such that* $\{Ku_{k_j}\}_{j=1}^\infty$ *converges in* Y.

Now let H denote a real Hilbert space, with inner product $(\ ,\ )$. It is easy to see that if a linear operator $K : H \to H$ is compact and $u_k \rightharpoonup u$, then $Ku_k \to Ku$.

THEOREM 4 (Compactness of adjoints). *If* $K : H \to H$ *is compact, so is* $K^* : H \to H$.

Proof. Let $\{u_k\}_{k=1}^{\infty}$ be a bounded sequence in H and extract a weakly convergent subsequence $u_{k_j} \rightharpoonup u$ in H. We will prove $K^*u_{k_j} \to K^*u$. Indeed,

$$\begin{aligned}\|K^*u_{k_j} - K^*u\|^2 &= (K^*u_{k_j} - K^*u, K^*[u_{k_j} - u]) \\ &= (KK^*u_{k_j} - KK^*u, u_{k_j} - u).\end{aligned}$$

Now since K^* is linear, $K^*u_{k_j} \rightharpoonup K^*u$, and so $KK^*u_{k_j} \to KK^*x$. Thus $K^*u_{k_j} \to K^*u$. □

THEOREM 5 (Fredholm alternative). *Let $K : H \to H$ be a compact linear operator. Then*

(i) $N(I - K)$ *is finite dimensional,*

(ii) $R(I - K)$ *is closed,*

(iii) $R(I - K) = N(I - K^*)^{\perp}$,

(iv) $N(I - K) = \{0\}$ *if and only if* $R(I - K) = H$,

and

(v) $\dim N(I - K) = \dim N(I - K^*)$.

Proof. 1. If $\dim N(I - K) = +\infty$, we can find an infinite orthonormal set $\{u_k\}_{k=1}^{\infty} \subset N(I - K)$. Then

$$Ku_k = u_k \quad (k = 1, \dots).$$

Now $\|u_k - u_l\|^2 = \|u_k\|^2 - 2(u_k, u_l) + \|u_l\|^2 = 2$ if $k \neq l$, and so $\|Ku_k - Ku_l\| = \sqrt{2}$ for $k \neq l$. This however contradicts the compactness of K, as $\{Ku_k\}_{k=1}^{\infty}$ would then contain no convergent subsequence. Assertion (i) is proved.

2. We next claim there exists a constant $\gamma > 0$ such that

$$\|u - Ku\| \geq \gamma \|u\| \quad \text{for all } u \in N(I - K)^{\perp}. \tag{4}$$

Indeed, if not, there would exist for $k = 1, \dots$ elements $u_k \in N(I - K)^{\perp}$ with $\|u_k\| = 1$ and $\|u_k - Ku_k\| < \frac{1}{k}$. Consequently

$$u_k - Ku_k \to 0. \tag{5}$$

But since $\{u_k\}_{k=1}^{\infty}$ is bounded, there exists a weakly convergent subsequence $u_{k_j} \rightharpoonup u$. By compactness $Ku_{k_j} \to Ku$, and then (5) implies $u_{k_j} \to u$. We therefore have $u \in N(I - K)$ and so

$$(u_{k_j}, u) = 0 \quad (j = 1, \dots).$$

Let $k_j \to \infty$ to derive a contradiction to (4).

3. Next let $\{v_k\}_{k=1}^{\infty} \subset R(I-K)$, $v_k \to v$. We can find $u_k \in N(I-K)^{\perp}$ solving $u_k - Ku_k = v_k$. Using (4) we deduce

$$\|v_k - v_l\| \geq \gamma \|u_k - u_l\|.$$

Thus $u_k \to u$ and $u - Ku = v$. This proves (ii).

4. Assertion (iii) is now a consequence of (ii) and the general fact

$$\overline{R(A)} = N(A^*)^{\perp} \text{ for each bounded linear operator } A : H \to H.$$

5. To verify (iv), let us suppose to start with that $N(I-K) = \{0\}$, but $H_1 = (I-K)(H) \subsetneq H$. According to (ii) H_1 is a closed subspace of H. Furthermore $H_2 \equiv (I-K)(H_1) \subsetneq H_1$, since $I-K$ is one-to-one. Similarly if we write $H_k \equiv (I-K)^k(H)$ $(k = 1, \dots)$, we see that H_k is a closed subspace of H, $H_{k+1} \subsetneq H_k$ $(k = 1, \dots)$.

Choose $u_k \in H_k$ with $\|u_k\| = 1$, $u_k \in H_{k+1}^{\perp}$. Then $Ku_k - Ku_l = -(u_k - Ku_k) + (u_l - Ku_l) + (u_k - u_l)$. Now if $k > l$, $H_{k+1} \subsetneq H_k \subseteq H_{l+1} \subsetneq H_l$. Thus $u_k - Ku_k$, $u_l - Ku_l$, $u_k \in H_{l+1}$. Since $u_l \in H_{l+1}^{\perp}$, $\|u_l\| = 1$, we deduce $\|Ku_k - Ku_l\| \geq 1$ $(k, l = 1, \dots)$. But this is impossible since K is compact.

6. Now conversely assume $R(I-K) = H$. Then owing to (iii) we see that $N(I-K^*) = \{0\}$. Since K^* is compact, we may utilize step 5 to conclude $R(I-K^*) = H$. But then $N(I-K) = R(I-K^*)^{\perp} = \{0\}$. This conclusion and step 5 complete the proof of assertion (iv).

7. Next we assert

$$\dim N(I-K) \geq \dim R(I-K)^{\perp}.$$

To prove this, suppose instead $\dim N(I-K) < \dim R(I-K)^{\perp}$. Then there exists a bounded linear mapping $A : N(I-K) \to R(I-K)^{\perp}$ which is one-to-one, but *not* onto. Extend A to a linear mapping $A : H \to R(I-K)^{\perp}$ by setting $Au = 0$ for $u \in N(I-K)^{\perp}$. Now A has a finite dimensional range and so A, and thus $K + A$, are compact. Furthermore $N(I - (K+A)) = \{0\}$. Indeed, if $Ku + Au = u$, then $u - Ku = Au \in R(I-K)^{\perp}$; whence $u - Ku = Au = 0$. Thus $u \in N(I-K)$ and so in fact $u = 0$, since A is one-to-one on $N(I-K)$. Now apply assertion (iv) to $\tilde{K} = K + A$. We conclude $R(I - (K+A)) = H$. But this is impossible: if $v \in R(I-K)^{\perp}$, but $v \notin R(A)$, the equation

$$u - (Ku + Au) = v$$

has no solution.

8. Since $R(I-K^*)^\perp = N(I-K)$, we deduce from step 7

$$\begin{aligned}\dim N(I-K^*) &\geq \dim R(I-K^*)^\perp \\ &= \dim N(I-K).\end{aligned}$$

The opposite inequality comes from interchanging the roles of K and K^*. This establishes (v). □

Remark. Theorem 5 asserts in particular either

$$(\alpha) \qquad \begin{cases} \text{for each } f \in H, \text{ the equation } u - Ku = f \\ \text{has a unique solution} \end{cases}$$

or else

$$(\beta) \qquad \begin{cases} \text{the homogeneous equation } u - Ku = 0 \\ \text{has solutions } u \neq 0. \end{cases}$$

This dichotomy is the *Fredholm alternative.* In addition, should (β) obtain, the space of solutions of the homogeneous problem is finite dimensional, and the nonhomogeneous equation

$$(\gamma) \qquad u - Ku = f$$

has a solution if and only if $f \in N(I-K^*)^\perp$. □

Now we investigate the spectrum of a compact linear operator.

THEOREM 6 (Spectrum of a compact operator). *Assume* $\dim H = \infty$ *and* $K : H \to H$ *is compact. Then*

(i) $0 \in \sigma(K)$,

(ii) $\sigma(K) - \{0\} = \sigma_p(K) - \{0\}$,

and

$$\text{(iii)} \quad \begin{cases} \sigma(K) - \{0\} \textit{ is finite, or else} \\ \sigma(K) - \{0\} \textit{ is a sequence tending to } 0. \end{cases}$$

Proof. 1. Assume $0 \notin \sigma(K)$. Then $K : H \to H$ is bijective and so $I = K \circ K^{-1}$, being the composition of a compact and a bounded linear operator, is compact. This is impossible, since $\dim H = \infty$.

2. Assume $\eta \in \sigma(K)$, $\eta \neq 0$. Then if $N(K - \eta I) = \{0\}$, the Fredholm alternative would imply $R(K - \eta I) = H$. But then $\eta \in \rho(K)$, a contradiction.

3. Suppose now $\{\eta_k\}_{k=1}^\infty$ is a sequence of *distinct* elements of $\sigma(K) - \{0\}$, and $\eta_k \to \eta$. We will show $\eta = 0$.

Indeed, since $\eta_k \in \sigma_p(K)$ there exists $w_k \neq 0$ such that $Kw_k = \eta_k w_k$. Let H_k denote the subspace of H spanned by $\{w_1, \dots, w_k\}$. Then $H_k \subsetneq H_{k+1}$ for each $k = 1, 2, \dots$, since the $\{w_k\}_{k=1}^{\infty}$ are linearly independent.

Observe also $(K - \eta_k I)H_k \subseteq H_{k-1}$ $(k = 2, \dots)$. Choose now for $k = 1, \dots$ an element $u_k \in H_k$, with $u_k \in H_{k-1}^{\perp}$ and $\|u_k\| = 1$. Now if $k > l$, $H_{l-1} \subsetneq H_l \subseteq H_{k-1} \subsetneq H_k$. Thus

$$\left\| \frac{Ku_k}{\eta_k} - \frac{Ku_l}{\eta_l} \right\| = \left\| \frac{(Ku_k - \eta_k u_k)}{\eta_k} - \frac{(Ku_l - \eta_l u_l)}{\eta_l} + u_k - u_l \right\| \geq 1,$$

since $Ku_k - \eta_k u_k$, $Ku_l - \eta_l u_l$, $u_l \in H_{k-1}$. If $\eta_k \to \eta \neq 0$, we obtain a contradiction to the compactness of K. □

D.6. Symmetric operators.

Now let $S : H \to H$ be bounded, symmetric, and write

$$m := \inf_{\substack{u \in H \\ \|u\|=1}} (Su, u), \quad M := \sup_{\substack{u \in H \\ \|u\|=1}} (Su, u).$$

LEMMA (Bounds on spectrum). *We have*
(i) $\sigma(S) \subset [m, M]$, *and*
(ii) $m, M \in \sigma(S)$.

Proof. 1. Let $\eta > M$. Then

$$(\eta u - Su, u) \geq (\eta - M)\|u\|^2 \quad (u \in H).$$

Hence the Lax–Milgram Theorem (§6.2.1) asserts $\eta I - S$ is one-to-one and onto, and thus $\eta \in \rho(S)$. Similarly $\eta \in \rho(S)$ if $\rho < m$. This proves (i).

2. We will prove $M \in \sigma(S)$. Since the pairing $[u, v] := (Mu - Su, v)$ is symmetric, with $[u, u] \geq 0$ for all $u \in H$, the Cauchy–Schwarz inequality implies

$$|(Mu - Su, v)| \leq (Mu - Su, u)^{1/2}(Mv - Sv, v)^{1/2}$$

for all $u, v \in H$. In particular

$$\|Mu - Su\| \leq C(Mu - Su, u)^{1/2} \quad (u \in H) \tag{6}$$

for some constant C.

Now let $\{u_k\}_{k=1}^{\infty} \subset H$ satisfy $\|u_k\| = 1$ $(k = 1, \dots)$ and $(Su_k, u_k) \to M$. Then (6) implies $\|Mu_k - Su_k\| \to 0$. Now if $M \in \rho(S)$, then

$$u_k = (MI - S)^{-1}(Mu_k - Su_k) \to 0,$$

a contradiction. Thus $M \in \sigma(S)$, and likewise $m \in \sigma(S)$. □

THEOREM 7 (Eigenvectors of a compact, symmetric operator). *Let H be a separable Hilbert space, and suppose $S : H \to H$ is a compact and symmetric operator. Then there exists a countable orthonormal basis of H consisting of eigenvectors of S.*

Proof. 1. Let $\{\eta_k\}$ comprise the sequence of distinct eigenvalues of S, excepting 0. Set $\eta_0 = 0$. Write $H_0 = N(S)$, $H_k = N(S - \eta_k I)$ $(k = 1, \dots)$. Then $0 \leq \dim H_0 \leq \infty$, and $0 < \dim H_k < \infty$, according to the Fredholm alternative.

2. Let $u \in H_k$, $v \in H_l$ for $k \neq l$. Then $Su = \eta_k u$, $Sv = \eta_l v$ and so

$$\eta_k(u, v) = (Su, v) = (u, Sv) = \eta_l(u, v).$$

As $\eta_k \neq \eta_l$, we deduce $(u, v) = 0$. Consequently we see the subspaces H_k and H_l are orthogonal.

3. Now let $\tilde{H}$ be the smallest subspace of H containing $H_0, H_1, \dots$. Thus $\tilde{H} = \{\sum_{k=0}^{m} a_k u_k \mid m \in \{0, \dots\}, u_k \in H_k, a_k \in \mathbb{R}\}$. We next demonstrate $\tilde{H}$ is dense in H. Clearly $S(\tilde{H}) \subseteq \tilde{H}$. Furthermore $S(\tilde{H}^\perp) \subseteq \tilde{H}^\perp$: indeed if $u \in \tilde{H}^\perp$ and $v \in \tilde{H}$, then $(Su, v) = (u, Sv) = 0$.

Now the operator $\tilde{S} \equiv S|_{\tilde{H}^\perp}$ is compact and symmetric. In addition $\sigma(\tilde{S}) = \{0\}$, since any nonzero eigenvalue of $\tilde{S}$ would be an eigenvalue of S as well. According to the lemma then, $(\tilde{S}u, u) = 0$ for all $u \in \tilde{H}^\perp$. But if $u, v \in \tilde{H}^\perp$,

$$2(\tilde{S}u, v) = (\tilde{S}(u + v), u + v) - (\tilde{S}u, u) - (\tilde{S}v, v) = 0.$$

Hence $\tilde{S} = 0$. Consequently $\tilde{H}^\perp \subset N(S) \subset \tilde{H}$, and so $\tilde{H}^\perp = \{0\}$. Thus $\tilde{H}$ is dense in H.

4. Choose an orthonormal basis for each subspace H_k $(k = 0, \dots)$, noting that since H is separable, H_0 has a countable orthonormal basis. We obtain thereby an orthonormal basis of eigenvectors. □

Most of these proofs are from Brezis [**BR1**]. See also Gilbarg–Trudinger [**G-T,** Chapter 5] and Yosida [**Y**].

APPENDIX E: MEASURE THEORY

This appendix provides a quick outline of some fundamentals of measure theory.

E.1. Lebesgue measure.

Lebesgue measure provides a way of describing the "size" or "volume" of certain subsets of $\mathbb{R}^n$.

DEFINITION. *A collection $\mathcal{M}$ of subsets of $\mathbb{R}^n$ is called a σ-*algebra *if*

(i) $\emptyset, \mathbb{R}^n \in \mathcal{M}$,

(ii) $A \in \mathcal{M}$ *implies* $\mathbb{R}^n - A \in \mathcal{M}$,

and

(iii) *if* $\{A_k\}_{k=1}^\infty \subset \mathcal{M}$, *then* $\bigcup_{k=1}^\infty A_k, \bigcap_{k=1}^\infty A_k \in \mathcal{M}$.

THEOREM 1 (Existence of Lebesgue measure and Lebesgue measurable sets). *There exist a σ-algebra $\mathcal{M}$ of subsets of $\mathbb{R}^n$ and a mapping*

$$| \; | : \mathcal{M} \to [0, +\infty]$$

with the following properties:

(i) *Every open subset of $\mathbb{R}^n$, and thus every closed subset of $\mathbb{R}^n$, belong to $\mathcal{M}$.*

(ii) *If B is any ball in $\mathbb{R}^n$, then $|B|$ equals the n-dimensional volume of B.*

(iii) *If $\{A_k\}_{k=1}^\infty \subset \mathcal{M}$ and the sets $\{A_k\}_{k=1}^\infty$ are pairwise disjoint, then*

$$(1) \qquad \left| \bigcup_{k=1}^\infty A_k \right| = \sum_{k=1}^\infty |A_k| \quad (\text{“countable additivity”}).$$

(iv) *If $A \subseteq B$, where $B \in \mathcal{M}$ and $|B| = 0$, then $A \in \mathcal{M}$ and $|A| = 0$.*

Notation. The sets in $\mathcal{M}$ are called *Lebesgue measurable sets* and $| \cdot |$ is n-dimensional *Lebesgue measure.* □

Remarks. (i) From (ii) and (iii), we see that $|A|$ equals the volume of any set A with piecewise smooth boundary.

(ii) We deduce from (1) that

$$(2) \qquad |\emptyset| = 0,$$

and

$$(3) \qquad \left| \bigcup_{k=1}^\infty A_k \right| \leq \sum_{k=1}^\infty |A_k| \quad (\text{“countable subadditivity”})$$

for any countable collection of measurable sets $\{A_k\}_{k=1}^\infty$.

Notation. If some property holds everywhere on $\mathbb{R}^n$, except for a measurable set with Lebesgue measure zero, we say the property holds *almost everywhere*, abbreviated “a.e.”. □

E.2. Measurable functions and integration.

DEFINITION. *Let $f : \mathbb{R}^n \to \mathbb{R}$. We say f is a* measurable function *if*

$$f^{-1}(U) \in \mathcal{M}$$

for each open subset $U \subset \mathbb{R}$.

Note in particular that if f is continuous, then f is measurable. The sum and product of two measurable functions are measurable. In addition if $\{f_k\}_{k=0}^{\infty}$ are measurable functions, then so are limsup f_k and liminf f_k.

THEOREM 2 (Egoroff's Theorem). *Let $\{f_k\}_{k=1}^{\infty}, f$ be measurable functions, and*

$$f_k \to f \quad \text{a.e. on } A,$$

where $A \subset \mathbb{R}^n$ is measurable, $|A| < \infty$. Then for each $\varepsilon > 0$ there exists a measurable subset $E \subset A$ such that

(i) $|A - E| \leq \varepsilon$

and

(ii) $f_k \to f$ *uniformly on E.*

Now if f is a nonnegative, measurable function, it is possible, by an approximation of f with simple functions, to define the Lebesgue integral

$$\int_{\mathbb{R}^n} f \, dx.$$

Cf. §E.5 below. This agrees with the usual integral if f is continuous or Riemann integrable. If f is measurable, but is not necessarily nonnegative, we define

$$\int_{\mathbb{R}^n} f \, dx = \int_{\mathbb{R}^n} f^+ \, dx - \int_{\mathbb{R}^n} f^- \, dx,$$

provided at least one of the terms on the right hand side is finite. In this case we say f is *integrable.*

DEFINITION. *A measurable function f is* summable *if*

$$\int_{\mathbb{R}^n} |f| \, dx < \infty.$$

Remark. Note carefully our terminology: a measurable function is *integrable* if it has an integral (which may equal $+\infty$ or $-\infty$) and is *summable* if this integral is finite. □

Notation. If the real-valued function f is measurable, we define the *essential supremum* of f to be

$$\text{ess sup } f := \inf\{\mu \in \mathbb{R} \mid |\{f > \mu\}| = 0\}.$$

□

E.3. Convergence theorems for integrals.

The Lebesgue theory of integration is especially useful since it provides the following powerful convergence theorems.

THEOREM 3 (Fatou's Lemma). *Assume the functions $\{f_k\}_{k=1}^\infty, f$ are nonnegative and summable. Then*

$$\int_{\mathbb{R}^n} \liminf_{k\to\infty} f_k \, dx \leq \liminf_{k\to\infty} \int_{\mathbb{R}^n} f_k \, dx.$$

THEOREM 4 (Monotone Convergence Theorem). *Assume the functions $\{f_k\}_{k=1}^\infty$ are measurable, with*

$$f_1 \leq f_2 \leq \cdots \leq f_k \leq f_{k+1} \leq \cdots.$$

Then

$$\int_{\mathbb{R}^n} \lim_{k\to\infty} f_k \, dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k \, dx.$$

THEOREM 5 (Dominated Convergence Theorem). *Assume the functions $\{f_k\}_{k=1}^\infty$ are integrable and*

$$f_k \to f \ \text{a.e.}$$

Suppose also

$$|f_k| \leq g \ \text{a.e.},$$

for some summable function g. Then

$$\int_{\mathbb{R}^n} f_k \, dx \to \int_{\mathbb{R}^n} f \, dx.$$

E.4. Differentiation.

An important fact is that a summable function is "approximately continuous" at almost every point.

THEOREM 6 (Lebesgue's Differentiation Theorem). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally summable.*

(i) *Then for a.e. point $x_0 \in \mathbb{R}^n$,*

$$\fint_{B(x_0,r)} f\,dx \to f(x_0) \quad \text{as } r \to 0.$$

(ii) *In fact, for a.e. point $x_0 \in \mathbb{R}^n$,*

$$\fint_{B(x_0,r)} |f(x) - f(x_0)|\,dx \to 0 \quad \text{as } r \to 0. \tag{4}$$

A point x_0 at which (4) holds is called a *Lebesgue point* of f.

Remark. More generally, if $f \in L^p_{loc}(\mathbb{R}^n)$ for some $1 \le p < \infty$, then for a.e. point $x_0 \in \mathbb{R}^n$ we have

$$\fint_{B(x_0,r)} |f(x) - f(x_0)|^p\,dx \to 0 \quad \text{as } r \to 0.$$

□

E.5. Banach space-valued functions.

We extend the notions of measurability, integrability, etc. to mappings

$$\mathbf{f} : [0,T] \to X$$

where $T > 0$ and X is a real Banach space, with norm $\|\ \|$.

DEFINITIONS. (i) *A function $\mathbf{s} : [0,T] \to X$ is called* simple *if it has the form*

$$\mathbf{s}(t) = \sum_{i=1}^{m} \chi_{E_i}(t) u_i \quad (0 \le t \le T), \tag{5}$$

where each E_i is a Lebesgue measurable subset of $[0,T]$ and $u_i \in X$ ($i = 1, \dots, m$).

(ii) *A function $\mathbf{f} : [0,T] \to X$ is* strongly measurable *if there exist simple functions $\mathbf{s}_k : [0,T] \to X$ such that*

$$\mathbf{s}_k(t) \to \mathbf{f}(t) \quad \text{for a.e. } 0 \le t \le T.$$

(iii) *A function $\mathbf{f} : [0,T] \to X$ is* weakly measurable *if for each $u^* \in X^*$, the mapping $t \mapsto \langle u^*, \mathbf{f}(t)\rangle$ is Lebesgue measurable.*

DEFINITION. *We say* $\mathbf{f} : [0,T] \to X$ *is* almost separably valued *if there exists a subset* $N \subset [0,T]$, *with* $|N| = 0$, *such that the set* $\{\mathbf{f}(t) \mid t \in [0,T] - N\}$ *is separable.*

THEOREM 7 (Pettis). *The mapping* $\mathbf{f} : [0,T] \to X$ *is strongly measurable if and only if* $\mathbf{f}$ *is weakly measurable and is almost separably valued.*

DEFINITIONS. (i) *If* $\mathbf{s}(t) = \sum_{i=1}^{m} \chi_{E_i}(t) u_i$ *is simple, we define*

$$\int_0^T \mathbf{s}(t)\,dt := \sum_{i=1}^{m} |E_i| u_i. \tag{6}$$

(ii) *We say* $\mathbf{f} : [0,T] \to X$ *is* summable *if there exists a sequence* $\{\mathbf{s}_k\}_{k=1}^{\infty}$ *of simple functions such that*

$$\int_0^T \|\mathbf{s}_k(t) - \mathbf{f}(t)\|\,dt \to 0 \text{ as } k \to \infty. \tag{7}$$

(iii) *If* f *is summable, we define*

$$\int_0^T \mathbf{f}(t)\,dt = \lim_{k\to\infty} \int_0^T \mathbf{s}_k(t)\,dt. \tag{8}$$

THEOREM 8 (Bochner). *A strongly measurable function* $\mathbf{f} : [0,T] \to X$ *is summable if and only if* $t \mapsto \|\mathbf{f}(t)\|$ *is summable. In this case*

$$\left\| \int_0^T \mathbf{f}(t)\,dt \right\| \le \int_0^T \|\mathbf{f}(t)\|\,dt,$$

and

$$\left\langle u^*, \int_0^T \mathbf{f}(t)\,dt \right\rangle = \int_0^T \langle u^*, \mathbf{f}(t) \rangle\,dt$$

for each $u^* \in X^*$.

A good book for measure theory is Folland [**F2**]. See Yosida [**Y,** Chapter V, Sections 4–5] for proofs of Theorems 7, 8.

BIBLIOGRAPHY

[A] S. S. Antman, The equations for large vibrations of strings, American Math. Society Monthly **87** (1980), 359–370.

[AR] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, 1986.

[B] J. M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, Archive for Rational Mech. and Analysis **63** (1977), 337–403.

[BA] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Noordhoff, 1976.

[B-CD] M. Bardi and I. Capuzzo Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*, forthcoming.

[B-L-P] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North Holland, 1978.

[BE] S. Benton, *The Hamilton-Jacobi Equation: A Global Approach*, Academic Press, 1977.

[BR1] H. Brezis, *Analyse Fonctionnelle*, Masson, 1983.

[BR2] H. Brezis, *Operateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, North Holland, 1973.

[B-E] H. Brezis and L. C. Evans, A variational inequality approach to the Bellman-Dirichlet equation for two elliptic operators, Archive for Rational Mech. and Analysis **71** (1979), 1–13.

[C] C. Carathéodory, *The Calculus of Variations and Partial Differential Equations of First Order*, Chelsea, 1982.

[CH] C. Chester, *Techniques in Partial Differential Equations*, McGraw–Hill, 1971.

[CO] J. Conlon, A theorem in ordinary differential equations with an application to hyperbolic conservation laws, Adv. in Math. **35** (1980), 1–18.

[C-F] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves*, Springer, 1976.

[C-H] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, 1962.

[C-E-L] M. G. Crandall, L. C. Evans, and P.-L. Lions, Some properties of viscosity solutions of Hamilton–Jacobi equations, Trans. American Math. Society **282** (1984), 487–502.

[C-L] M. G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton–Jacobi equations, Trans. American Math. Society **277** (1983), 1–42.

[D-M] B. Dacorogna and J. Moser, On a partial differential equation involving a Jacobian determinant, Ann. Inst. H. Poincaré **7** (1990), 1–26.

[DA] E. B. Davies, *Heat Kernels and Spectral Theory*, Cambridge University Press, 1989.

[DB] E. DiBenedetto, *Partial Differential Equations*, Birkhäuser, 1995.

[DP] R. DiPerna, Decay and asymptotic behavior of solutions to nonlinear hyperbolic systems of conservation laws, Indiana University Math. J. **24** (1975), 1047–1071.

[D] P. G. Drazin, *Solitons*, Cambridge University Press, 1983.

[D-Z] P. Duchateau and D. W. Zachman, *Partial Differential Equations*, Schaum's Outline Series in Math., McGraw-Hill, 1986.

[E] L. C. Evans, *Weak Convergence Methods for Nonlinear Partial Differential Equations*, CBMS # 74, American Math. Society, 1990.

[E-G] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992 (An errata sheet for the first printing of this book is available through the math.berkeley.edu website).

[F-L-S] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures in Physics*, Vol. II, Addison–Wesley, 1966.

[F-S] W. Fleming and M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer, 1993.

[F1] G. Folland, *Introduction to Partial Differential Equations*, Princeton University Press, 1976.

[F2] G. Folland, *Real Analysis: Modern Techniques and their Applications*, Wiley-Interscience, 1984.

[FD] M. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton University Press, 1985.

[FR1] A. Friedman, *Partial Differential Equations*, Holt, Rinehart, Winston, 1969.

[FR2] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice–Hall, 1964.

[GA] F. Gantmacher, *The Theory of Matrices*, Chelsea, 1977.

[G] P. Garabedian, *Partial Differential Equations*, Wiley, 1964.

[GI] M. Giaquinta, *Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems*, Princeton University Press, 1983.

[G-H] M. Giaquinta and S. Hildebrandt, *Calculus of Variations*, Vol. 1-2, Springer, 1996.

[G-N-N] B. Gidas, W. Ni, and L. Nirenberg, Symmetry and related properties via the maximum principle, Comm. Math. Physics **68** (1979), 209–243.

[G-T] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, (2nd ed.), Springer, 1983.

[H1] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, Vol. 1-4, Springer, 1983-85.

[H2] L. Hörmander, Fourier integral operators I, Acta Mathematica **127** (1971), 79-183.

[J] F. John, *Partial Differential Equations*, (4th ed.), Springer.

[K] T. Kato, *Perturbation Theory for Linear Operators*, (2nd ed.), Springer, 1980.

[K-S] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, 1980.

[KR] N. V. Krylov, *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, American Math. Society, 1996.

[L] O. A. Ladyzhenskaya, *The Boundary Value Problems of Mathematical Physics*, Springer, 1985.

[L-S-U] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uraltseva, *Linear and Quasilinear Equations of Parabolic Type*, American Math. Society, 1968.

[L-U] O. A. Ladyzhenskaya and N. N. Uraltseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, 1968.

[LA] P. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, SIAM, 1973.

[LV] R. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser, 1992.

[L-L] E. Lieb and M. Loss, *Analysis*, American Math. Society, 1997.

[LB] G. M. Lieberman, *Second Order Parabolic Partial Differential Equations*, World Scientific, 1996.

[L1] J.-L. Lions, *Equations Differentielles Operationnelles et Problèmes aux Limits*, Springer, 1961.

[L2] J.-L. Lions, *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*, Dunod, 1969.

[L-M] J.-L. Lions and E. Magenes, *Nonhomogeneous boundary value problems and applications*, Vol. I-III, Springer, 1972.

[LI] P.-L. Lions, *Generalized Solutions of Hamilton-Jacobi Equations*, Research Notes in Mathematics 69, Pitman, 1982.

[LO] J. D. Logan, *Applied Mathematics*, Wiley, 1987.

[M-P] A. Majda and R. Pego, Stable viscosity matrices for systems of conservation laws, J. Differential Equations **56** (1985), 229–262.

[M] V. P. Mikhailov, *Partial Differential Equations*, Mir, 1978.

[O] P. Olver, *Applications of Lie Groups to Differential Equations*, Springer, 1986.

[PA] L. E. Payne, *Improperly Posed Problems in Partial Differential Equations*, SIAM, 1975.

[PY] I. G. Petrovsky, *Lectures on Partial Differential Equations*, Wiley-Interscience, 1954.

[P] M. Pinsky, *Partial Differential Equations and Boundary-Value Problems, with Applications*, (2nd ed.), McGraw–Hill, 1991.

[P-T] M. Pinsky and M. Taylor, Pointwise Fourier inversion: a wave equation approach, J. Fourier Analysis (1997).

[P-W] M. Protter and H. Weinberger, *Maximum Principles in Differential Equations*, Prentice–Hall, 1967.

[RA] P. Rabinowitz, *Minimax Methods in Critical Point Theory with Applications to Differential Equations*, CBMS # 65, American Math. Society, 1986.

[R] J. Rauch, *Partial Differential Equations*, Springer, 1992.

[R-R] M. Renardy and R. Rogers, *A First Graduate Course in Partial Differential Equations*, Springer, 1993.

[RD] W. Rudin, *Principles of Mathematical Analysis*, (3rd ed.), McGraw–Hill, 1976.

[RU] H. Rund, *The Hamilton–Jacobi Theory in the Calculus of Variations*, Krieger, 1973.

[S] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer, 1983.

[SN] I. Sneddon, *Elements of Partial Differential Equations*, McGraw–Hill, 1957.

[SE] E. Stein, *Harmonic Analysis*, Princeton University Press, 1993.

[S-W] E. Stein and G. Weiss, *Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[ST] W. Strauss, *Partial Differential Equations: An Introduction*, Wiley, 1992.

[TA] M. Taylor, *Partial Differential Equations*, Vol. I-III, Springer, 1996.

[TE] R. Temam, *Navier–Stokes Equations*, North Holland, 1977.

[T-Z] D. W. Thoe and E. C. Zachmanoglou, *Introduction to Partial Differential Equations with Applications*, Dover, 1986.

[T] F. Treves, *Basic Linear Partial Differential Equations*, Academic Press, 1975.

[V] J. L. Vazquez, An introduction to the mathematical theory of the porous medium equation, in *Shape Optimization and Free Boundaries*, (ed. by Delfour), Kluwer Acad. Publ., 1992.

[W] N. A. Watson, A theory of subtemperatures in several variables, Proc. London Math. Society **26** (1973), 385–417.

[WE] H. Weinberger, *A First Course in Partial Differential Equations*, Blaisdell, 1965.

[WH] G. B. Whitham, *Linear and Nonlinear Waves*, Wiley-Interscience, 1974.

[WL] J. Wloka, *Partial Differential Equations*, Cambridge University Press, 1987.

[X-Z] L. Xiao and T. Zhang, *The Riemann Problem and Interaction of Waves in Gas Dynamics*, Pitman, 1989.

[Y] K. Yosida, *Functional Analysis*, (6th ed.), Springer, 1980.

[ZD] E. Zeidler, *Nonlinear Functional Analysis and its Applications,* Vol. 1-4, Springer.

[ZH] Y. Zheng, *Two-Dimensional Riemann Problems for Systems of Conservation Laws*, forthcoming.

[Z] W. Ziemer, *Weakly Differentiable Functions*, Springer, 1989.

[ZW] D. Zwillinger, *Handbook of Differential Equations*, Academic Press, 1984.

INDEX