An Introduction to Linear Partial Differential Equations

Qing Han

To Yansu, Raymond and Tommy

Contents

Preface vii

Chapter 1. Introduction 1
- 1.1. Notations 1
- 1.2. Well-Posed Problems 3
- 1.3. Overview 4

Chapter 2. First-Order Differential Equations 7
- 2.1. Non-Characteristic Hypersurfaces 7
- 2.2. The Method of Characteristics 11
- 2.3. A Priori Estimates 23
- Exercises 34

Chapter 3. An Overview of Second-Order PDEs 37
- 3.1. Classifications of Second Order Equations 37
- 3.2. Energy Estimates 43
- 3.3. Separation of Variables 49
- Exercises 56

Chapter 4. Laplace Equations 59
- 4.1. Dirichlet Problems 59
- 4.2. Fundamental Solutions 65
- 4.3. Mean-Value Properties 71
- 4.4. The Maximum Principle 76
- 4.5. Schauder Estimates 86
- 4.6. Weak Solutions 91
- Exercises 97

Chapter 5. Heat Equations 101
- 5.1. Initial-Value Problems 101
- 5.2. Regularity of Solutions 112
- 5.3. The Maximum Principle 117
- 5.4. Harnack Inequalities 123
- Exercises 129

Chapter 6. Wave Equations 133
- 6.1. One-Dimensional Wave Equations 133
- 6.2. Higher Dimensional Wave Equations 142
- 6.3. Energy Estimates 153
- Exercises 160

Chapter 7. First-Order Differential Systems 165
7.1. Reductions to First-Order Differential Systems 165
7.2. Cauchy-Kowalevski Theorem 171
7.3. Hyperbolic Differential Systems 180
Exercises 188

Bibliography 191

Index 193

Preface

Is it really necessary to classify partial differential equations (PDEs) and to employ different methods to discuss different types of equations? Why is it important to derive *a priori* estimates of solutions before even proving the existence of solutions? These are only a few questions any students who just start studying PDEs might ask. Students may find answers to these questions only at the end of a one-semester course in basic PDEs, sometimes after they have already lost interest in the subject. In this book, we attempt to address these issues *at the beginning.* There are several notable features in this book.

First, the importance of *a priori estimates* is addressed at the beginning and emphasized throughout this book. This is well illustrated by the first chapter on first-order PDEs. Although first-order linear PDEs can be solved by the method of characteristics, we provide a detailed analysis of *a priori* estimates of solutions in sup-norms and in integral norms. To emphasize the importance of these estimates, we demonstrate how to prove the existence of weak solutions with the help of basic results from functional analysis. The setting here is easy, since Sobolev spaces are not required. Meanwhile, all important ideas are in full display. In this book, we do attempt to derive explicit expressions for solutions whenever possible. However, these explicit expressions of solutions of special equations usually serve mostly to suggest the correct form of estimates for solutions of general equations.

The second feature is the illustration of the necessity of classifying second-order PDEs at the beginning. In the second chapter, immediately after classifying second order PDEs into elliptic, parabolic and hyperbolic type, we discuss various boundary-value problems and initial/boundary-value problems for the Laplace equation, the heat equation and the wave equation. We discuss energy methods for proving uniqueness and find solutions in the plane by the method of separation of variables. The explicit expressions of solutions demonstrate different properties of solutions of different types of PDEs. Such differences clearly indicate that there is unlikely to be a unified approach to study PDEs.

Third, we focus on simple models of PDEs and study these equations in detail. We have chapters devoted to the Laplace equation, the heat equation and the wave equation, and use several methods to study each equation. For example, for the Laplace equation, we use three different methods to study harmonic functions, the fundamental solution, the mean-value property and the maximum principle. For each method, we indicate its advantages and its shortcomings. General equations are not forgotten. We also discuss maximum principles for general elliptic and parabolic equations and energy estimates for general hyperbolic equations.

The book is designed for a one-semester course at the graduate level. As a result, many important topics have to be omitted. One topic notably missing from this book is Sobolev spaces. It is hard, if not impossible, to introduce Sobolev spaces and discuss applications in a one-semester course on basic PDEs. However, we seize every opportunity we can to address the importance of the missing topic. For example, in Chapter 3 we solve initial/boundary-value problems for the 1-dimensional heat equation and for the 1-dimensional wave equation by using separation of variables. An important role is played by the eigenvalue problem for the operator $\frac{d^2}{dx^2}$ over an interval. After successfully solving 1-dimensional problems, we point out that, in order to solve higher dimensional versions, we need to solve first the eigenvalue problem for the Laplace operator in a bounded smooth domain. For that, Sobolev spaces play an essential role. Without filling in details, we introduce Sobolev spaces and discuss how to find weak solutions of the Dirichlet problem of the Poisson equation in Section 4.6.

The choice of topics in this book, as in any books, is influenced by personal tastes of the author. Some topics in this book may not be viewed as *basic* by others. For example, differential Harnack inequality for the heat equation usually is not found in PDE textbooks at a comparable level. It is included here simply because of its importance. Attempts have been made to give a balanced coverage of different classes of partial differential equations.

Acknowledgments. This book is based on the one-semester course the author taught at University of Notre Dame in the fall of 2007. The author would like to thank Ms. Hairong Liu for a wonderful job of typing the manuscript.

Qing Han

CHAPTER 1

Introduction

This chapter serves as an introduction of entire book.

1.1. Notations

Most of the time, we denote by x points in $\mathbb{R}^n$ and write $x = (x_1, \cdots, x_n)$ in terms of its coordinates. For any $x \in \mathbb{R}^n$, we denote by $|x|$ the standard Euclidean norm, unless otherwise stated. Namely, for any $x = (x_1, \cdots, x_n)$, we have

$$|x| = \Big(\sum_{i=1}^n x_i^2\Big)^{\frac{1}{2}}.$$

Sometimes, we need to distinguish one particular direction as the time direction and write points in $\mathbb{R}^{n+1}$ by (x, t) for $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. In this case, we call $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ the space variable and $t \in \mathbb{R}$ the time variable. In $\mathbb{R}^2$, we also denote points by (x, y).

Let Ω be a domain in $\mathbb{R}^n$, an open and connected subset in $\mathbb{R}^n$. We denote by $C(\Omega)$ the collection of all continuous functions in Ω, by $C^m(\Omega)$ the collection of all functions with continuous derivatives up to order m, for any integer $m \geq 1$, and by $C^\infty(\Omega)$ the collection of all functions with continuous derivatives of arbitrary order. For any $u \in C^m(\Omega)$, we denote by $\nabla^m u$ the collection of all partial derivatives of u of order m. In particular,

$$\nabla u = (u_{x_1}, \cdots, u_{x_n}).$$

This is the gradient of u. For second derivatives, we usually write in the matrix form as follows

$$\nabla^2 u = \begin{pmatrix} u_{x_1x_1} & u_{x_1x_2} & \cdots & u_{x_1x_n} \\ u_{x_2x_1} & u_{x_2x_2} & \cdots & u_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{x_nx_1} & u_{x_nx_2} & \cdots & u_{x_nx_n} \end{pmatrix}.$$

This is a symmetric matrix, called the Hessian matrix of u. For derivatives of order higher than two, we need to use multi-indices. A multi-index $\alpha \in \mathbb{Z}_+^n$ is given by $\alpha = (\alpha_1, \cdots, \alpha_n)$ for nonnegative integers $\alpha_1, \cdots, \alpha_n$. We write

$$|\alpha| = \sum_{i=1}^n \alpha_i.$$

For any vector $\xi = (\xi_1, \cdots, \xi_n) \in \mathbb{R}^n$, we denote

$$\xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}.$$

The partial derivative $\partial^\alpha u$ is defined by

$$\partial^\alpha u = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} u,$$

and its order is $|\alpha|$. For any positive integer m, we define

$$|\nabla^m u| = \Big(\sum_{|\alpha|=m} |\partial^\alpha u|^2\Big)^{\frac{1}{2}}.$$

In particular,

$$|\nabla u| = \Big(\sum_{i=1}^n u_{x_i}^2\Big)^{\frac{1}{2}},$$

and

$$|\nabla^2 u| = \Big(\sum_{i,j=1}^n u_{x_i x_j}^2\Big)^{\frac{1}{2}}.$$

A hypersurface in $\mathbb{R}^n$ is a surface of $n-1$ dimension. Locally, a C^m-hypersurface can be expressed by $\{\varphi = 0\}$ for a C^m-function φ with $\nabla\varphi \neq 0$. Alternatively, by a rotation, we may take $\varphi(x) = x_n - \psi(x_1, \cdots, x_{n-1})$ for a C^m-function ψ of $n-1$ variables. A domain $\Omega \subset \mathbb{R}^n$ is C^m if its boundary $\partial\Omega$ is an C^m-hypersurface.

A *partial differential equation* (henceforth abbreviated as PDE) in a domain $\Omega \subset \mathbb{R}^n$ is a relation of independent variables $x \in \Omega$, an unknown function u defined in Ω and a finite number of its partial derivatives. Solving a PDE is to find this unknown function. The *order* of a PDE is the order of the highest derivative in the relation. Hence for a positive integer m, the general form of an m-th order PDE in a domain $\Omega \subset \mathbb{R}^n$ is given by

$$F\big(x, u, \nabla u(x), \nabla^2 u(x), \cdots, \nabla^m u(x)\big) = 0 \quad \text{for any } x \in \Omega.$$

Here F is a function which is continuous in all its arguments and u is a C^m-function in Ω. A C^m-solution u satisfying the above equation in the pointwise sense in Ω is often called a *classical solution*. Sometimes, we need to relax regularity requirements for solutions when classical solutions are not known to exist. Instead of going to details, we only mention that it is an important method to establish the existence of *weak solutions* first, functions with less regularity than C^m and satisfying the equation in some weak sense, and then to prove that these weak solutions actually possess required regularity to be classical solutions.

A PDE is *linear* if it is linear in the unknown functions and their derivatives, with coefficients depending on independent variables x. A general m-th order linear PDE in Ω is given by

$$\sum_{|\alpha|\le m} a_\alpha(x)\partial^\alpha u = f(x) \quad \text{for } x \in \Omega.$$

Here a_α is called the coefficient of $\partial^\alpha u$. A PDE of order m is *quasilinear* if it is linear in derivatives of solutions of order m, with coefficients depending on independent variables x and derivatives of solutions of order $< m$. In general, an m-th order quasilinear PDE in Ω is given by

$$\sum_{|\alpha|=m} a_\alpha(x, u, \cdots, \nabla^{m-1}u)\partial^\alpha u = f(x, u, \cdots, \nabla^{m-1}u) \quad \text{for } x \in \Omega.$$

Several PDEs involving one or more unknown functions and their derivatives form a differential system. We define linear and quasilinear partial differential systems accordingly.

In this book, we will focus on first-order and second-order linear PDEs and first-order linear differential systems. In a few occasions, we diverge to nonlinear PDEs.

1.2. Well-Posed Problems

What is the meaning of *solving* partial differential equations? Ideally, we obtain explicit solutions in terms of elementary functions. In practice this is only possible for very simple PDEs or very simple solutions of more general PDEs. In general, it is impossible to find explicit expressions of all solutions of all PDEs. In the absence of explicit solutions, we need to seek methods to prove existence of solutions of PDEs and discuss properties of these solutions.

A given PDE may not have solutions at all or may have many solutions. When it has many solutions, we intend to find *side conditions* to pick up the most reasonable solutions. Hadamard introduced the notion of *well-posed problems*. A PDE and side conditions are called *well-posed* if

(i) there admits a solution;

(ii) this solution is unique;

(iii) solutions depend continuously in some suitable sense on side conditions, i.e., solutions change a little if side conditions change a little.

Now the basic question can be formulated as follows: Given an equation, find side conditions to have the well-posedness. We usually refer to (i), (ii) and (iii) as the existence, uniqueness and the continuous dependence respectively.

In practice, both the uniqueness and the continuous dependence are proved by *a priori estimates*. Namely, we assume solutions already exist and then derive certain norms of solutions in terms of known functions in the equation and side conditions. It is important to note that establishing a priori estimates is in fact the first step in proving the existence of solutions. A closely related issue here is regularity of solutions such as continuity and differentiability. Solutions of a particular PDE can only be obtained if the right kind of regularity, or the right kind of norms, are employed. Two classes of norms are used often, sup-norms and L^2-norms.

Let Ω be a domain in $\mathbb{R}^n$. For any bounded function u in Ω, we define the sup-norm of u in Ω by

$$|u|_{L^\infty(\Omega)} = \sup_\Omega |u|.$$

Let m be a nonnegative integer. For any function u in Ω with bounded derivatives up to order m, we define the C^m-norm of u in Ω by

$$|u|_{C^m(\Omega)} = \sum_{|\alpha|\le m} |\partial^\alpha u|_{L^\infty(\Omega)}.$$

If Ω is a bounded C^m-domain in $\mathbb{R}^n$, then $C^m(\bar{\Omega})$, the collection of functions which are C^m in $\bar{\Omega}$, is a Banach space equipped with the C^m-norm.

Next, for any Lebesgue measurable function u in Ω, we define the L^2-norm of u in Ω by

$$\|u\|_{L^2(\Omega)} = \Big(\int_\Omega u^2\Big)^{\frac{1}{2}},$$

where integration is in the Lebesgue sense. The L^2-space in Ω is the collection of all Lebesgue measurable functions in Ω with finite L^2-norms and is denoted by

$L^2(\Omega)$. We learned from real analysis that $L^2(\Omega)$ is a Banach space equipped with the L^2-norm.

Other norms will also be used. We will introduce them when needed.

In deriving a priori estimates, we follow a common practice and use the variable constant convention. The same letter C is used to denote constants which may change from line to line, as long as it is clear from context on what quantities the constants depend upon. In most cases, we are not interested in the value of the constant, but only in its existence.

1.3. Overview

There are 7 chapters in this book.

The main topic in Chapter 2 is first-order PDEs. In Section 2.1, we introduce the basic notion of characteristic hypersurfaces. In Section 2.2, we use the method of integral curves to study initial value problems with initial values prescribed on non-characteristic hypersurfaces. In Section 2.3, we derive estimates of solutions in L^∞-norms and in L^2-norms without using explicit expressions of solutions.

Chapter 3 should be considered as an introduction to the theory of second-order linear PDEs. In Section 3.1, we introduce the Laplace equation, the heat equation and the wave equation. We also introduce their general forms, elliptic equations, parabolic equations and hyperbolic equations, which will be studied in detail in subsequent chapters. We use energy methods to discuss the uniqueness of certain boundary value problems in Section 3.2 and use separation of variables to solve these problems in the plane in Section 3.3.

In Chapter 4, we discuss the Laplace equation and the Poisson equation. The Laplace equation is probably the most important PDE with the widest range of applications. In Section 4.1, we solve Dirichlet problems for the Laplace equation in balls and derive Poisson integral formula. In the next three sections, we study harmonic functions, (i.e., solutions of the Laplace equation), by three different methods: fundamental solutions, mean-value properties and the maximum principle. These three sections are relatively independent of each other. In Section 4.5, we study the Poisson equation $\Delta u = f$ and derive Schauder estimates for its solutions. Then in Section 4.6, we briefly discuss weak solutions of the Poisson equation.

In Chapter 5, we study the heat equation and discuss properties of its solutions. This equation describes the temperature of a body conducting heat, when the density is constant, In Section 5.1, we discuss the fundamental solution of the heat equation and solve some initial-value problems. Then in Section 5.2, we use the fundamental solution to discuss regularity of arbitrary solutions. In Section 5.3, we discuss the maximum principle for the heat equation and its applications. In particular, we use the maximum principle to derive interior gradient estimates. In Section 5.4, we discuss Harnack inequalities.

In Chapter 6, we study the n-dimensional wave equation, which represents vibrations of strings or propagation of sound waves in tubes for $n = 1$, waves on the surface of shallow water for $n = 2$, and acoustic or light waves for $n = 3$. In Section 6.1, we discuss initial-value problems and various initial/boundary-value problems for the one-dimensional wave equation. In Section 6.2, we study initial value problems for the wave equation in higher dimensional spaces. Then in Section 6.3, we derive energy estimates for solutions of initial-value problems. Chapter 6 is relatively independent of Chapter 4 and Chapter 5 and can be taught after Chapter 3.

In Chapter 7, we discuss partial differential systems of first-order and focus on non-characteristic initial-value problems. In Section 7.1, we introduce non-characteristic hypersurfaces for partial differential equations and systems of arbitrary order. We also demonstrate that partial differential systems of arbitrary order can always be changed to those of first-order. In Section 7.2, we discuss the Cauchy-Kowalevski Theorem, which asserts the existence of analytic solutions of non-characteristic initial-value problems for analytic differential systems and initial values on analytic non-characteristic hypersurfaces. In Section 7.3, we discuss hyperbolic differential systems and derive energy estimates for symmetric hyperbolic differential systems.

Each chapter, except this one, ends with exercises. Level of difficulty varies considerably. Some exercises, at the most difficult level, may require efforts for days.

CHAPTER 2

First-Order Differential Equations

The main topic in this chapter is first-order PDEs. In Section 2.1, we introduce the basic notion of non-characteristic hypersurfaces. In Section 2.2, we use the method of characteristics to study initial value problems with initial values prescribed on non-characteristic hypersurfaces. In Section 2.3, we derive estimates of solutions in L^∞-norms and in L^2-norms without using explicit expressions of solutions.

2.1. Non-Characteristic Hypersurfaces

Let Ω be a domain in $\mathbb{R}^n$. A first-order PDE in Ω is given by

$$F\big(x, u, \nabla u(x)\big) = 0 \quad \text{for any } x \in \Omega, \tag{2.1}$$

where $F = F(x, u, p)$ is continuous in $(x, u, p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n$ and u is a C^1-function in Ω.

We start our study with the simplest case. Given a point in Ω, say the origin, find a reasonable condition so that (2.1) admits a solution in a neighborhood of the origin. Let Σ be a hypersurface in $\mathbb{R}^n$, a surface of dimension $n-1$, passing the origin. We will prescribe conditions on Σ to find a solution.

To illustrate, we consider first-order linear PDEs. Let Ω be a domain in $\mathbb{R}^n$ containing the origin and L be a first-order linear differential operator given by

$$Lu = \sum_{i=1}^n a_i(x) u_{x_i} + b(x) u,$$

where a_i are b are continuous functions in Ω, for any $i = 1, \cdots, n$. Here a_i and b are called the coefficients for u_{x_i} and u respectively.

For a given function f in Ω, we consider the equation

$$Lu = f(x) \quad \text{in } \Omega. \tag{2.2}$$

The function f is called the *non-homogeneous term*. If $f \equiv 0$, (2.2) is called a homogeneous equation.

Let Σ be the hyperplane $\{x_n = 0\}$. For $x \in \mathbb{R}^n$, we write $x = (x', x_n)$ for $x' = (x_1, \cdots, x_{n-1}) \in \mathbb{R}^{n-1}$. We intend to prescribe u on Σ to solve (2.2) in a neighborhood of the origin. We first check whether we can find all derivatives of u at the origin. For a given function u_0 in a neighborhood of the origin in $\mathbb{R}^{n-1}$, we set

$$u(x', 0) = u_0(x') \quad \text{for any small } x' \in \mathbb{R}^{n-1}. \tag{2.3}$$

Then we can find all x'-derivatives of u at the origin. In particular, we can

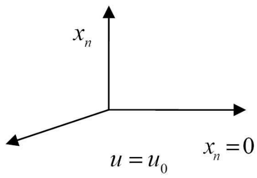

FIGURE 2.1. Initial values.

determine all first order derivatives of u at the origin except u_{x_n}. To find this, we need to use the equation. If we assume

$$a_n(0) \neq 0,$$

then we can find $u_{x_n}(0)$ from (2.2). In fact, we can compute all derivatives of u of any order at the origin by using u_0 and differentiating (2.2).

We usually call Σ an initial hypersurface and u_0 an initial value or a Cauchy value. The problem of solving (2.2) together with (2.3) is called an initial-value problem or a Cauchy problem.

More generally, consider a hypersurface Σ given by $\{\varphi = 0\}$ for a C^1-function φ in a neighborhood of the origin with $\nabla\varphi \neq 0$, with the origin on the hypersurface Σ, i.e., $\varphi(0) = 0$. We note that $\nabla\varphi$ is simply a normal vector of the hypersurface Σ. Without loss of generality, we assume $\varphi_{x_n}(0) \neq 0$. Then by the implicit function theorem, we solve $\varphi = 0$ around $x = 0$ for $x_n = \psi(x_1, \cdots, x_{n-1})$. We consider a change of variables

$$x \mapsto y = \big(x_1, \cdots, x_{n-1}, \varphi(x)\big).$$

This is a well defined transform in a neighborhood of the origin with a nonsingular Jacobian. In fact, the Jacobian J is given by

$$J = \frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)} = \begin{pmatrix} & & & 0 \\ & Id & & \vdots \\ & & & 0 \\ \varphi_{x_1} & \cdots & \varphi_{x_{n-1}} & \varphi_{x_n} \end{pmatrix},$$

and hence $\det J(0) = \varphi_{x_n}(0) \neq 0$. Now we write the operator L in new variables. Note

$$u_{x_i} = \sum_{k=1}^{n} y_{k,x_i} u_{y_k}.$$

Therefore,

$$Lu = \sum_{k=1}^{n} \Big(\sum_{i=1}^{n} a_i(x(y)) y_{k,x_i}\Big) u_{y_k} + b(x(y))u.$$

In new coordinates, the initial hypersurface Σ is given by $\{y_n = 0\}$ and the coefficient of u_{y_n} is given by

$$\sum_{i=1}^{n} a_i(x)\varphi_{x_i}.$$

We recall that $\nabla\varphi = (\varphi_{x_1}, \cdots, \varphi_{x_n})$ is simply a normal vector of $\Sigma = \{\varphi = 0\}$. When $\Sigma = \{x_n = 0\}$, or $\varphi(x) = x_n$, then $\nabla\varphi = (0, \cdots, 0, 1)$ and

$$\sum_{i=1}^{n} a_i(x)\varphi_{x_i} = a_n(x).$$

This reduces to the special case we just discussed.

Definition 2.1. For a linear operator L as in (2.2) defined in a neighborhood of $x_0 \in \mathbb{R}^n$, a C^1-hypersurface Σ is non-characteristic at $x_0 \in \Sigma$ if

$$\sum_{i=1}^{n} a_i(x_0)\xi_i \neq 0, \tag{2.4}$$

where $\xi = (\xi_1, \cdots, \xi_n)$ is a normal vector of Σ at x_0. Otherwise, it is called characteristic at x_0.

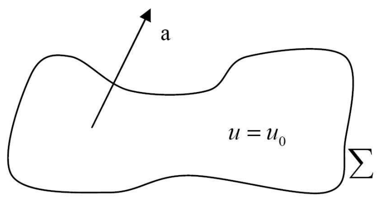

Figure 2.2. Non-characteristic hypersurfaces.

A hypersurface is non-characteristic if it is non-characteristic at every point. Strictly speaking, a hypersurface is characteristic if it is not non-characteristic, i.e., if it is characteristic at some point. In this book, we will abuse this notion. When we say a hypersurface is characteristic, we mean it is characteristic *everywhere*. This should cause few confusions. When $n = 2$, hypersurfaces are simply curves. It is more appropriate to call them characteristic curves and non-characteristic curves.

The non-characteristics condition has a simple geometric interpretation. If we view $a = (a_1, \cdots, a_n)$ as a vector in $\mathbb{R}^n$, then (2.4) holds if and only if $a(x_0)$ is not a tangent vector to Σ at x_0. We note that condition (2.4) is maintained under C^1-changes of local coordinates. This condition assures that we can compute all derivatives of solutions at x_0.

The concept of the non-characteristics can also be generalized to linear partial differential systems. Consider in a neighborhood of $x_0 \in \mathbb{R}^n$

$$\sum_{i=1}^{n} A_i(x)u_{x_i} + B(x)u = f(x),$$

where A_i and B are $N \times N$ matrices and u and f are N-column vectors. A hypersurface Σ is non-characteristic at $x_0 \in \Sigma$ if

$$\det\big(\sum_{i=1}^{n} \xi_i A_i(x_0)\big) \neq 0,$$

where $\xi = (\xi_1, \cdots, \xi_n)$ is a normal vector of Σ at the origin.

Next, we turn to general nonlinear partial differential equations

$$F(x, u, \nabla u) = 0. \tag{2.5}$$

We ask the same question as for linear equations. Given a hypersurface Σ and an initial value on it, can we compute all derivatives of solutions? First, we consider the case of hyperplane $\{x_n = 0\}$.

EXAMPLE 2.2. Consider

$$\sum_{i=1}^{n} u_{x_i}^2 = 1,$$

and

$$u(x', 0) = u_0(x').$$

It is obvious that $u = x_i$ is a solution for $u_0(x') = x_i$, $i = 1, \cdots, n-1$. However, if $|\nabla_{x'} u_0(x')|^2 > 1$, there are no solutions for such an initial value.

Hence, in the case of nonlinear PDEs, the concept of non-characteristics also depends on initial values.

Suppose an initial value is given on $\Sigma = \{x_n = 0\}$ by

$$u(x', 0) = u_0(x').$$

We ask whether (2.5) admits a solution in a neighborhood of the origin with the given initial value. To answer this question, we first assume that there is a C^1-function v in a neighborhood of the origin having the given initial value and satisfying $F = 0$ at the origin, i.e.,

$$F\big(0, v(0), \nabla v(0)\big) = 0.$$

As in the discussion of linear PDEs, we ask whether we can find u_{x_n} at the origin. By the implicit function theorem, this is possible if

$$F_{u_{x_n}}\big(0, v(0), \nabla v(0)\big) \neq 0.$$

Now we reconsider the equation in Example 2.2. We set

$$F(x, u, p) = |p|^2 - 1 \quad \text{for any } p \in \mathbb{R}^n.$$

Suppose the initial value u_0 satisfies

$$|\nabla_{x'} u_0(0)| < 1.$$

Let $v = u_0 + c x_n$ for a constant to be determined. Then $|\nabla v(0)|^2 = |\nabla_{x'} u_0(0)|^2 + c^2$. By choosing

$$c = \pm\sqrt{1 - |\nabla_{x'} u_0(0)|^2} \neq 0,$$

v satisfies the equation at $x = 0$. For such two choices of v, we have

$$F_{u_{x_n}}(0, v(0), \nabla v(0)) = 2 v_{x_n}(0) = 2c \neq 0.$$

Hence, the notion of non-characteristics depends on the choice of v.

DEFINITION 2.3. Let $F = 0$ be a first-order nonlinear PDE as in (2.5) in a neighborhood of $x_0 \in \mathbb{R}^n$ and Σ be a hypersurface. With an initial value u_0 prescribed on Σ, suppose there exists a function v such that $v = u_0$ on Σ and $F(x_0, v(x_0), \nabla v(x_0)) = 0$. Then the hypersurface Σ is non-characteristic at $x_0 \in \Sigma$ with respect to the given initial value u_0 and with respect to v if

$$\sum_{i=1}^{n} F_{u_{x_i}}\big(x_0, v(x_0), \nabla v(x_0)\big)\xi_i \neq 0,$$

where $\xi = (\xi_1, \cdots, \xi_n)$ is a normal vector of Σ at x_0.

2.2. The Method of Characteristics

In this section, we use the method of characteristics to solve first-order PDEs. We demonstrate that solutions of any first-order PDEs with initial values prescribed on non-characteristic hypersurfaces can be obtained by solving systems of Ordinary Differential Equations (ODEs). This is not true for higher-order PDEs or for systems of first-order PDEs.

Let $\Omega \subset \mathbb{R}^n$ be a domain. The general form of first-order PDEs in Ω is given by

$$F(x, u, \nabla u) = 0 \quad \text{for } x \in \Omega,$$

where F is continuous in $\Omega \times \mathbb{R} \times \mathbb{R}^n$. Let Σ be a hypersurface in $\mathbb{R}^n$ and an initial value be prescribed on Σ by

$$u = u_0 \quad \text{on } \Sigma.$$

In the following, we assume Ω is a domain containing the origin and Σ is non-characteristic at the origin.

As shown in the previous section, for nonlinear differential equations, the non-characteristics condition depends on initial values. For example, for a quasi-linear PDE of the form

$$\sum_{i=1}^{n} a_i(x, u)u_{x_i} = f(x, u),$$

the non-characteristics condition at 0 is given by

$$\sum_{i=1}^{n} a_i\big(0, u_0(0)\big)\xi_i \neq 0,$$

where $\xi = (\xi_1, \cdots, \xi_n)$ is a normal vector of Σ at 0.

We first consider a simple case where the hypersurface Σ is given by the hyperplane $\{x_n = 0\}$. Obviously, a normal vector of $\{x_n = 0\}$ is given by $(0, \cdots, 0, 1)$. Our goal is to solve the following initial value problem

$$\begin{aligned} F(x, u, \nabla u) &= 0, \\ u(x', 0) &= u_0(x'), \end{aligned}$$

for $x \in \mathbb{R}^n$ close to the origin. Here, we write $x = (x', x_n)$ for $x' \in \mathbb{R}^{n-1}$. Solving for u locally around the origin means to find values of $u(x)$ for x close to the origin.

To motivate, we consider initial-value problems of linear homogeneous equations

$$\sum_{i=1}^{n} a_i(x)u_{x_i} = 0,$$
$$u(x', 0) = u_0(x'), \tag{2.6}$$

where a_i is C^1 in a neighborhood of $0 \in \mathbb{R}^n$, $i = 1, \cdots, n$, and u_0 is at least C^1 in a neighborhood of $0 \in \mathbb{R}^{n-1}$. By introducing $a = (a_1, \cdots, a_n)$, we simply write the equation in (2.6) as

$$a(x) \cdot \nabla u = 0.$$

Here $a(x)$ is regarded as a vector field in $\mathbb{R}^n$. Then $a(x) \cdot \nabla$ is a directional derivative along $a(x)$ at x.

In the following, we assume the hyperplane $\{x_n = 0\}$ is non-characteristic at the origin, i.e.,

$$a_n(0) \neq 0.$$

Our strategy is as follows. For any $\bar{x} \in \mathbb{R}^n$ close to the origin, we construct a special curve along which u is constant. If such a curve starts from $\bar{x}$ and intersects $\mathbb{R}^{n-1} \times \{0\}$ at $(\bar{y}, 0)$ for a small $\bar{y} \in \mathbb{R}^{n-1}$, then $u(\bar{x}) = u_0(\bar{y})$.

To find such a curve $x(t)$, we consider the restriction of u on it, which gives a one-variable function $u(x(t))$. Now we calculate the t-derivative of this function and have

$$\frac{d}{dt}\big(u(x(t))\big) = \sum_{i=1}^{n} u_{x_i} \frac{dx_i}{dt}.$$

In order to have a constant value of u along this curve, we require

$$\frac{d}{dt}\big(u(x(t))\big) = 0.$$

A simple comparison with the equation in (2.6) yields the following sufficient condition

$$\frac{dx_i}{dt} = a_i(x) \quad \text{for } i = 1, \cdots, n.$$

This naturally leads to the following definition.

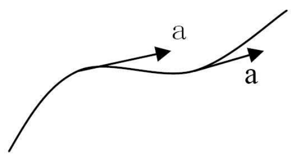

FIGURE 2.3. Integral curves.

DEFINITION 2.4. Let $x = x(t)$ be a C^1-curve in $\mathbb{R}^n$. It is an *integral curve* of (2.6) if

$$\frac{dx(t)}{dt} = a\big(x(t)\big). \tag{2.7}$$

The calculation preceding Definition 2.4 shows that solutions u of (2.6) are constant along integral curves of the coefficient vector. This yields the following method of solving (2.6). For any $\bar{x} \in \mathbb{R}^n$ near the origin, we find an integral curve of the coefficient vector through $\bar{x}$ by solving

$$\begin{aligned} \frac{dx}{dt} &= a(x), \\ x(0) &= \bar{x}. \end{aligned} \tag{2.8}$$

If it intersects the hyperplane $\{x_n = 0\}$ at some $(\bar{y}, 0)$, then we let $u(\bar{x}) = u_0(\bar{y})$.

It is easy to verify that, for any $x \in \mathbb{R}^n$ close to the origin, there is a unique integral curve, i.e., a solution of (2.8), starting from x and intersecting the hyperplane $\{x_n = 0\}$ near the origin at a unique point.

Since (2.8) is an autonomous system (i.e., the independent variable t does not appear explicitly), we consider the following system instead of (2.8)

$$\begin{aligned} \frac{dx}{dt} &= a(x), \\ x(0) &= (y, 0). \end{aligned} \tag{2.9}$$

Starting with any $y \in \mathbb{R}^{n-1}$ close to the origin, we expect the integral curve $x(y, t)$ to reach any $x \in \mathbb{R}^n$ close to the origin for small t. This is confirmed by the following result.

Lemma 2.5. *Let $k \geq 2$ be an integer and a be a C^{k-1}-vector field in a neighborhood of the origin with $a_n(0) \neq 0$. Then for any small $y \in \mathbb{R}^{n-1}$ and any small t, the solution $x = x(y, t)$ of (2.9) defines a C^k-diffeomorphism in a neighborhood of the origin in $\mathbb{R}^n$.*

Proof. This follows from the implicit function theorem easily. By standard results in ordinary differential equations, (2.9) admits a C^k-solution $x = x(y, t)$ for small $(y, t) \in \mathbb{R}^{n-1} \times \mathbb{R}$. We treat it as a map $(y, t) \mapsto x$ and calculate its Jacobian J at $(y, t) = (0, 0)$. By $x(y, 0) = (y, 0)$, we have

$$J(0) = \frac{\partial x}{\partial(y, t)}\Big|_{(y,t)=(0,0)} = \begin{pmatrix} & & & a_1(0) \\ & Id & & \vdots \\ & & & a_{n-1}(0) \\ 0 & \cdots & 0 & a_n(0) \end{pmatrix}.$$

Hence $\det J(0) = a_n(0) \neq 0$. $\square$

Hence, for any small $\bar{x}$, we can solve

$$x(y, t) = \bar{x}$$

uniquely for small $\bar{y}$ and $\bar{t}$. Then $u(\bar{x}) = u_0(\bar{y})$ yields a solution of (2.6). Note that $\bar{t}$ is not present in the expression of solutions. Hence the value of the solution $u(\bar{x})$ depends only the initial value u_0 at $(\bar{y}, 0)$ and, meanwhile, the initial value u_0 at $(\bar{y}, 0)$ influences the solution u along the integral curve starting from $(\bar{y}, 0)$. Therefore, we say the *domain of dependence* of the solution $u(\bar{x})$ on the initial value is represented by the single point $(\bar{y}, 0)$ and the *domain of influence* of the initial value at a particular point $(\bar{y}, 0)$ on solutions consists of the integral curve starting from $(\bar{y}, 0)$.

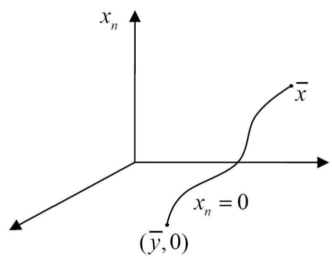

FIGURE 2.4. Solutions by integral curves.

For $n = 2$, integral curves are exactly characteristic curves. This can be seen easily by (2.7) and Definition 2.1. Hence the ODE (2.7) is often referred to as the *characteristic ODE*. This term is adopted for arbitrary dimensions. We have demonstrated how to solve homogeneous first-order linear PDEs by using characteristics ODEs. Such a method is called the *method of characteristics*. Later on, we will develop a similar method to solve general first-order PDEs.

We need to emphasize that solutions constructed by the method of characteristics are only local. In other words, they exist only in a neighborhood of the origin. A natural question here is whether there exists a global solution for globally defined a and u_0. There are several reasons that local solutions cannot be extended globally. First, $u(\bar{x})$ cannot be evaluated at $\bar{x} \in \mathbb{R}^n$ if $\bar{x}$ is not on integral curves

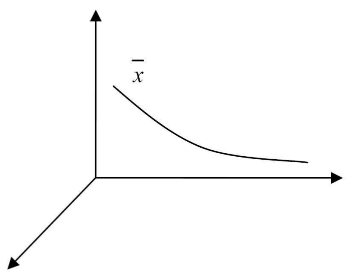

FIGURE 2.5. Undetermined $u(\bar{x})$.

from the initial hypersurface, or equivalently, integral curves from $\bar{x}$ does not intersect the initial hypersurface. Second, $u(\bar{x})$ cannot be evaluated at $\bar{x} \in \mathbb{R}^n$ if the integral curve starting from $\bar{x}$ intersects the initial hypersurface more than once. In this case, we should have a compatibility condition for initial values and we cannot prescribe initial values arbitrarily.

EXAMPLE 2.6. Consider in $\mathbb{R}^2 = \{(x, y)\}$

$$\begin{aligned} u_y + u_x &= 0, \\ u(x, 0) &= u_0(x). \end{aligned}$$

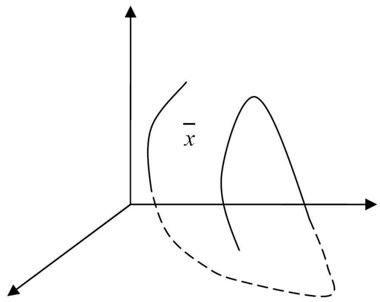

FIGURE 2.6. Undetermined $u(\bar{x})$.

Note that the vector a is given by $a = (1, 1)$ and hence $\{y = 0\}$ is non-characteristic. Characteristic ODEs are given by

$$\begin{cases} \dot{x} = 1, \\ \dot{y} = 1, \end{cases} \quad \text{and} \quad \begin{cases} x|_{t=0} = x_0, \\ y|_{t=0} = 0. \end{cases}$$

Hence

$$\begin{aligned} x =& t + x_0, \\ y =& t. \end{aligned}$$

By eliminating t, we have

$$x - y = x_0.$$

Along this curve, u is constant. Hence

$$u(x, y) = u_0(x - y).$$

We can interpret the fact that u is constant along the straight line $x - y = x_0$ in the following way. If we interpret y as time, the graph of the solution represents a wave propagating to the right with velocity 1 without changing shape. The solution u exists globally in $\mathbb{R}^2$.

Next, we discuss initial-value problems of quasi-linear PDEs. Let $\Omega \subset \mathbb{R}^n$ be a domain and consider

$$\begin{aligned} \sum_{i=1}^{n} a_i(x, u) u_{x_i} &= f(x, u) \quad \text{in } \Omega, \\ u(x', 0) &= u_0(x') \quad \text{for } x' \in \Omega', \end{aligned} \tag{2.10}$$

where a_i and f are C^1-functions in $\Omega \times \mathbb{R}$ and u_0 is C^1 in $\Omega' = \{x' \in \mathbb{R}^{n-1}; (x', 0) \in \Omega\}$. Assume the hyperplane $\{x_n = 0\}$ is non-characteristic at the origin, i.e.,

$$a_n\big(0, u_0(0)\big) \neq 0.$$

Suppose (2.10) admits a solution u. We first examine integral curves

$$\begin{aligned} \frac{dx}{dt} =& a\big(x, u\big), \\ x|_{t=0} =& (y, 0), \end{aligned}$$

where $y \in \mathbb{R}^{n-1}$. Contrary to the case of homogenous linear equations we studied earlier, we are unable to solve this ODE since u, the unknown function we intend to find, is present. However, viewing u as a function of t along these curves, we can calculate how u changes. A similar calculation as before yields

$$\begin{aligned}\frac{du}{dt} &= f(x,u),\\ u|_{t=0} &= u_0(y).\end{aligned}$$

Hence we have an ordinary differential system for x and u. This leads to the following method.

Consider the following ordinary differential system

$$\begin{aligned}\frac{dx}{dt} &= a(x,u),\\ \frac{du}{dt} &= f(x,u),\end{aligned}$$

with initial values

$$\begin{aligned}x|_{t=0} &= (y,0),\\ u|_{t=0} &= u_0(y),\end{aligned}$$

where $y \in \mathbb{R}^{n-1}$. In formulating this system, we treat x and u as functions of t only. This system consists of $n+1$ equations for $n+1$ functions and is the *characteristic ODE* of the quasilinear first-order PDE (2.10). By solving the characteristic ODE, we have a solution given by

$$x = x(y,t), \quad u = \varphi(y,t).$$

As in the proof of Lemma 2.5, we can prove that the map $(y,t) \mapsto x$ is a diffeomorphism. Hence, for any $\bar{x} \in \mathbb{R}^n$ close to the origin, there exists a unique $\bar{y} \in \mathbb{R}^{n-1}$ and $\bar{t} \in \mathbb{R}$ close to the origin such that

$$\bar{x} = x(\bar{y},\bar{t}).$$

Then the solution u at $\bar{x}$ is given by

$$u(\bar{x}) = \varphi(\bar{y},\bar{t}).$$

Similarly, we discuss initial-value problems when initial hypersurfaces are not hyperplanes. Let Σ be a hypersurface containing the origin and u_0 be a function on Σ. Instead of (2.10), we consider

$$\begin{aligned}\sum_{i=1}^n a_i(x,u)u_{x_i} &= f(x,u),\\ u &= u_0 \quad \text{on } \Sigma.\end{aligned}$$

We assume that Σ is non-characteristic at 0 with respect to u_0, i.e.,

$$\sum_{i=1}^n a_i\big(0,u_0(0)\big)\xi_i \neq 0,$$

where $\xi = (\xi_1, \cdots, \xi_n)$ is a normal vector of Σ at 0. We consider the characteristic ODE in the following form

$$\begin{aligned}\frac{dx}{dt} =& a(x,u),\\ \frac{du}{dt} =& f(x,u),\end{aligned}$$

with initial values

$$\begin{aligned}x|_{t=0} =& x_0,\\ u|_{t=0} =& u_0(x_0),\end{aligned}$$

for $x_0 \in \Sigma$. Suppose we have a solution

$$x = x(x_0, t), \quad u = \varphi(x_0, t).$$

As before, the map $(x_0, t) \mapsto x$ is a diffeomorphism. Hence for any given $\bar{x}$ close to the origin, there exists a unique $(\bar{x}_0, \bar{t}) \in \Sigma \times \mathbb{R}$ in a neighborhood of the origin such that

$$x(\bar{x}_0, \bar{t}) = \bar{x}.$$

Then define

$$u(\bar{x}) = \varphi(\bar{x}_0, \bar{t}).$$

This is the desired solution.

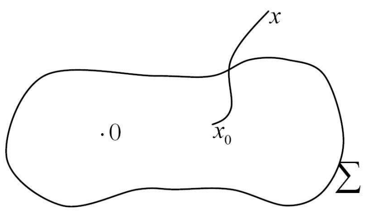

FIGURE 2.7. Solutions by the method of characteristics.

EXAMPLE 2.7. Consider in $\mathbb{R}^2 = \{(x,y)\}$

$$\begin{aligned}u_y + uu_x =& 0,\\ u(x,0) =& -x.\end{aligned}$$

Note that the vector a is given by $a = (u, 1)$ and hence $\{y = 0\}$ is non-characteristic. The characteristic ODE is given by

$$\begin{cases}\dot{x} = u,\\ \dot{y} = 1,\\ \dot{u} = 0,\end{cases} \quad \text{and} \quad \begin{cases}x|_{t=0} = x_0,\\ y|_{t=0} = 0,\\ u|_{t=0} = -x_0.\end{cases}$$

Hence

$$\begin{aligned} x =& -tx_0 + x_0, \\ y =& t, \\ u =& -x_0. \end{aligned}$$

By eliminating t and x_0, we have

$$t = y, \quad x_0 = \frac{x}{1-t} = \frac{x}{1-y}.$$

Therefore, the solution u is given by

$$u(x, y) = \frac{x}{y-1}.$$

We note that this solution is not defined at $y = 1$. In fact, the integral curve γ_{x_0} passing through the point $(x_0, 0)$ is given by the straight line

$$x = x_0 - x_0 y,$$

and the solution on this line is given by $u = -x_0$. Each γ_{x_0} passes the point $(0, 1)$.

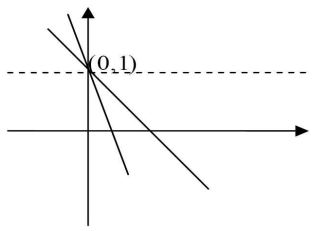

FIGURE 2.8. Integral curves in Example 2.7.

In general, smooth solutions of nonlinear PDEs are not expected to exist globally, as shown in Example 2.7. The equation discussed in Example 2.7 is called Burgers' equation.

We consider a more general equation

$$u_y + g(u)u_x = 0,$$

with an initial value is given on $\{y = 0\}$ by

$$u(x, 0) = u_0(x).$$

The characteristic ODE is given by

$$\begin{cases} \dot{x} = g(u), \\ \dot{y} = 1, \\ \dot{u} = 0, \end{cases} \qquad \text{and} \qquad \begin{cases} x|_{t=0} = x_0, \\ y|_{t=0} = 0, \\ u|_{t=0} = u_0(x_0). \end{cases}$$

First, $u = u_0(x_0)$, which is constant, along the integral curve starting from $(x_0, 0)$. Then on this integral curve,

$$\dot{x} = g\big(u_0(x_0)\big), \quad \dot{y} = 1.$$

So the integral curve is a straight line with the slope $\dfrac{1}{g(u_0(x_0))}$. Consider $x_0 < x_1$. If the slope of the straight line from x_1 is greater than that from x_0, these two straight lines will intersect. For example, for $g(u) = u$, the slope is $\frac{1}{u_0(x)}$. Therefore we have a non-global solution if u_0 is strictly decreasing. See Example 2.7.

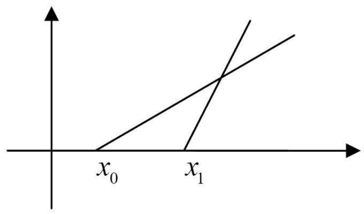

FIGURE 2.9. Integral curves.

EXAMPLE 2.8. Consider

$$u_y + uu_x = 1,$$
$$u = 0 \quad \text{on } \{y = x^2\}.$$

By setting $\varphi(x, y) = x^2 - y$, we have $\nabla\varphi(x, y) = (2x, -1)$ and hence for $a = (u, 1)$

$$a \cdot \nabla\varphi|_{\{\varphi=0\}} = (u, 1) \cdot (2x, -1)|_{\{\varphi=0\}} = -1 \neq 0.$$

Therefore, $\{y = x^2\}$ is non-characteristic. The characteristic ODE is given by

$$\begin{cases} \dot{x} = u, \\ \dot{y} = 1, \\ \dot{u} = 1, \end{cases} \quad \text{and} \quad \begin{cases} x|_{t=0} = x_0, \\ y|_{t=0} = x_0^2, \\ u|_{t=0} = 0. \end{cases}$$

Hence, we have

$$x = \frac{t^2}{2} + x_0,$$
$$y = t + x_0^2,$$
$$u = t.$$

By eliminating t and x_0, we obtain an implicit expression for u

$$y = u + (x - \frac{u^2}{2})^2.$$

So far in our discussion, initial values are prescribed on non-characteristic hypersurfaces. It becomes complicated when initial values are prescribed on characteristic hypersurfaces.

Let $\Omega \subset \mathbb{R}^n$ be a domain and consider

$$\sum_{i=1}^{n} a_i(x, u)u_{x_i} = f(x, u) \quad \text{in } \Omega,$$
$$u(x', 0) = u_0(x') \quad \text{for } x' \in \Omega',$$

where a_i and f are C^1 in $\Omega \times \mathbb{R}$ and u_0 is C^1 in $\Omega' = \{x' \in \mathbb{R}^{n-1}; (x', 0) \in \Omega\}$. For $y_0 \in \Omega'$, we assume the hyperplane $\Sigma = \{x_n = 0\}$ is characteristic at y_0, i.e.,

$$a_n\big(y_0, 0, u_0(y_0)\big) = 0.$$

Then we should have the following *compatibility condition*

$$\sum_{i=1}^{n-1} a_i\big(y_0, 0, u_0(y_0)\big) u_{0x_i}(y_0) = f\big(y_0, 0, u_0(y_0)\big). \tag{2.11}$$

Even if y_0 is the only point where $\{x_n = 0\}$ is characteristic, solutions may not exist in a neighborhood of $(y_0, 0)$ for initial values satisfying the compatibility condition (2.11) at y_0.

Now we discuss general nonlinear first-order partial differential equations. Let $\Omega \subset \mathbb{R}^n$ be a domain containing the origin and $F = F(x, u, p)$ be a C^1-function in $(x, u, p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n$. Consider

$$F(x, u, \nabla u) = 0 \quad \text{for } x \in \Omega, \tag{2.12}$$

and prescribe an initial value on $\{x_n = 0\}$

$$u(x', 0) = u_0(x') \quad \text{for } x' \text{ with } (x', 0) \in \Omega. \tag{2.13}$$

Here we write $x = (x', x_n)$. Assume there is a scalar a_0 such that

$$F\big(0, u_0(0), \nabla_{x'} u_0(0), a_0\big) = 0.$$

We seek a solution u in a neighborhood of the origin with the given initial value and $u_{x_n}(0) = a_0$. The non-characteristics condition with respect to u_0 and a_0 is given by

$$F_{p_n}\big(0, u_0(0), \nabla_{x'} u_0(0), a_0\big) \neq 0. \tag{2.14}$$

Note that by this condition, using the differential equation with the help of the implicit function theorem, we can solve for $a(x')$ for small $x' \in \mathbb{R}^{n-1}$, with $a(0) = a_0$, such that

$$F\big(x', 0, u_0(x'), \nabla_{x'} u_0(x'), a(x')\big) = 0.$$

We also require

$$u_{x_n}(x', 0) = a(x') \quad \text{for } x' \text{ small}.$$

We start with a formal consideration. Suppose we have a solution u. Set

$$p_i = u_{x_i} \quad \text{for } i = 1, \cdots, n. \tag{2.15}$$

Then

$$F(x_1, \cdots, x_n, u, p_1, \cdots, p_n) = 0. \tag{2.16}$$

Differentiating (2.16) with respect to x_i, we have

$$\sum_{j=1}^{n} F_{p_j} p_{i,x_j} = -F_{x_i} - F_u u_{x_i} \quad \text{for } i = 1, \cdots, n, \tag{2.17}$$

where we used $p_{j,x_i} = p_{i,x_j}$. We view (2.17) as a first-order quasilinear equation for p_i, for each fixed $i = 1, \cdots, n$. An important feature here is that the coefficient for p_{i,x_j} is F_{p_j}, independent of i. Characteristic ODEs associated with (2.17) are given by

$$\begin{aligned} \dot{x}_j =& F_{p_j}, \quad j = 1, \cdots, n, \\ \dot{p}_i =& -F_u p_i - F_{x_i}, \quad i = 1, \cdots, n. \end{aligned}$$

We also have

$$\dot{u} = \sum_{j=1}^{n} u_{x_j} \dot{x}_j = \sum_{j=1}^{n} p_j F_{p_j}.$$

Now we collect ordinary differential equations for x_j, u and p_i.

Consider the following ordinary differential system

$$\begin{aligned} \dot{x}_j =& F_{p_j}(x, u, p), \quad j = 1, \cdots, n, \\ \dot{p}_i =& -F_u(x, u, p) p_i - F_{x_i}(x, u, p), \quad i = 1, \cdots, n \\ \dot{u} =& \sum_{j=1}^{n} p_j F_{p_j}(x, u, p), \end{aligned} \tag{2.18}$$

with initial values at $t = 0$

$$\begin{aligned} x|_{t=0} =& (y, 0), \\ u|_{t=0} =& u_0(y), \\ p_i|_{t=0} =& u_{0x_i}(y), \quad i = 1, \cdots, n-1, \\ p_n|_{t=0} =& a(y), \end{aligned} \tag{2.19}$$

where $y \in \mathbb{R}^{n-1}$. This is the characteristic ODE for general nonlinear first-order partial differential equations. It is an ordinary differential system of $2n+1$ equations for $2n+1$ functions x, u and p. Here we view x, u and p as functions of t. Compare this with a similar ordinary differential system of $n+1$ equations for $n+1$ functions x and u for quasilinear equations. Solving (2.18) with (2.19), we have

$$x = x(y, t), \quad u = \varphi(y, t), \quad p = p(y, t).$$

We will prove that the map $(y, t) \mapsto x$ is a diffeomorphism. Hence for any given $\bar{x}$ near the origin, there exists a unique $\bar{y} \in \mathbb{R}^{n-1}$ and $\bar{t} \in \mathbb{R}$ such that

$$\bar{x} = x(\bar{y}, \bar{t}).$$

Then we define u by

$$u(\bar{x}) = \varphi(\bar{y}, \bar{t}).$$

THEOREM 2.9. *The function u defined above is a solution of (2.12)-(2.13).*

We should note that this solution u depends on the choice of the scalar a_0 and the function $a(x')$.

PROOF. The proof consists of several steps.

Step 1. The map $(y, t) \mapsto x$ is a diffeomorphism. This is proved as in the proof of Lemma 2.5. In fact, the Jacobian of the map $(y, t) \mapsto x$ at $(0, 0)$ is given by

$$J(0) = \frac{\partial x}{\partial (y, t)}|_{y=0, t=0} = \begin{pmatrix} & & & * \\ & Id & & \vdots \\ & & & * \\ 0 & \cdots & 0 & \dot{x}_n(0, 0) \end{pmatrix},$$

where

$$\dot{x}_n(0,0) = F_{p_n}(0, u_0(0), u_{0x_1}(0), \cdots, u_{0x_{n-1}}, a_0) \neq 0.$$

Hence $\det J(0) \neq 0$ by the non-characteristics condition (2.14). By the implicit function theorem, for any $\bar{x} \in \mathbb{R}^n$ close to the origin, we can solve $\bar{x} = x(\bar{y}, \bar{t})$ uniquely for $\bar{y}$ and $\bar{t}$. Then define

$$u(\bar{x}) = \varphi(y, t).$$

We will prove that this is the desired solution and

$$p_i(y,t) = u_{x_i}(\bar{x}(y,t)).$$

Step 2. We claim

$$F\big(x(y,t), \varphi(y,t), p(y,t)\big) \equiv 0.$$

Denote by $f(t)$ the function at the left hand side. Then $f(0) = 0$ since

$$f(0) = F\big(y, 0, u_0(y), \nabla_{x'} u_0(y), a(y)\big) = 0.$$

Next, we have by (2.18)

$$\begin{aligned}
\frac{df(t)}{dt} &= \frac{d}{dt} F\big(x(y,t), \varphi(y,t), p(y,t)\big) \\
&= \sum_{i=1}^{n} F_{x_i} \frac{dx_i}{dt} + F_u \frac{du}{dt} + \sum_{j=1}^{n} F_{p_j} \frac{dp_j}{dt} \\
&= \sum_{i=1}^{n} F_{x_i} F_{p_i} + F_u \sum_{j=1}^{n} p_j F_{p_j} + \sum_{j=1}^{n} F_{p_j}(-F_u p_j - F_{x_j}) = 0.
\end{aligned}$$

Hence $f(t) \equiv 0$.

Step 3. We claim

$$p_i(y,t) = u_{x_i}(x(y,t)).$$

Let

$$w_i(t) = u_{x_i}(x(y,t)) - p_i(y,t).$$

We will prove

$$w_i(t) = 0 \quad \text{for any } t \text{ and } i = 1, \cdots, n.$$

First, consider $t = 0$. For $i = 1, \cdots, n-1$, $w_i(0) = 0$ by initial values (2.19). For $i = n$, we will show $u_{x_n}(y, 0) = a(y)$. To see this, we first note by (2.18)

$$0 = \dot{u} - \sum_{j=1}^{n} p_j F_{p_j} = \sum_{j=1}^{n} \big(u_{x_j} \dot{x}_j - p_j F_{p_j}\big) = \sum_{j=1}^{n} F_{p_j}(u_{x_j} - p_j). \tag{2.20}$$

Since $w_i(0) = 0$ for $i = 1, \cdots, n-1$, we have $F_{p_n} w_n|_{t=0} = 0$. This implies $w_n(0) = 0$ since $F_{p_n}|_{t=0} \neq 0$ by the non-characteristics condition (2.14).

To prove $w_i(t) = 0$ for any t, we show $\dot{w}_i$ is a linear combination of w_l, $l = 1, \cdots, n$, i.e.,

$$\dot{w}_i = \sum_{l=1}^{n} a_{il} w_l.$$

To prove this, we differentiate (2.20) with respect to x_i and get

$$\sum_{j=1}^{n} F_{p_j} \cdot (u_{x_i x_j} - p_{j,x_i}) + \sum_{j=1}^{n} (F_{p_j})_{x_i} w_j = 0.$$

Then

$$\begin{aligned}\dot{w}_i &= \sum_{j=1}^{n} u_{x_i x_j} F_{p_j} + F_u p_i + F_{x_i} \\ &= \sum_{j=1}^{n} F_{p_j} p_{j,x_i} - \sum_{j=1}^{n} (F_{p_j})_{x_i} w_j + F_u p_i + F_{x_i}.\end{aligned}$$

By Step 2,

$$F\big(x, u(x), p_1(x), \cdots, p_n(x)\big) = 0.$$

Differentiating with respect to x_i, we have

$$F_{x_i} + F_u u_{x_i} + \sum_{j=1}^{n} F_{p_j} p_{j,x_i} = 0.$$

Hence

$$\begin{aligned}\dot{w}_i &= - F_{x_i} - F_u u_{x_i} - \sum_{j=1}^{n} (F_{p_j})_{x_i} w_j + F_u p_i + F_{x_i} \\ &= - F_u w_i - \sum_{j=1}^{n} (F_{p_j})_{x_i} w_j,\end{aligned}$$

or

$$\dot{w}_i = - \sum_{j=1}^{n} \big(F_u \delta_{ij} + (F_{p_j})_{x_i}\big) w_j.$$

This ends the proof of Step 3.

Step 2 and Step 3 imply that u is the desired solution. □

To end this section, we briefly compare methods we used to solve first-order linear or quasilinear PDEs and general first-order nonlinear PDEs. In solving a first-order quasilinear PDE, we formulate an ordinary differential system of $n+1$ equations for $n+1$ functions x and u. For a general first-order nonlinear PDE, the corresponding ordinary differential system consists of $2n+1$ equations for $2n+1$ functions x, u and ∇u. Here, we need to take into account the gradient of u by adding n more equations for ∇u. In other words, we may regard our first-order nonlinear PDE as a relation for (u, p) with a constraint $p = \nabla u$. We should emphasize this is a unique feature for single first-order PDEs. For PDEs of higher order or first-order partial differential systems, nonlinear equations are dramatically different from linear equations. In the rest of the book, we concentrate only on linear equations.

2.3. A Priori Estimates

A priori estimates play a fundamental role in partial differential equations. Usually, they are the starting point for the existence and regularity of solutions. To derive a priori estimates, we first assume solutions already exist and then control certain norms of solutions by those of known functions in equations, such as nonhomogenous terms and coefficients. Two frequently used norms are L^∞-norms and L^2-norms. The importance of L^2-norm estimates lies in the Hilbert space structure of the L^2-space. Once L^2-estimates of solutions and their derivatives are derived, standard results in Hilbert spaces such as the Riesz representation theorem can be utilized. In many cases, this is how the existence of solutions is established.

In this section, we will use first-order linear PDEs to demonstrate how to derive a priori estimates in L^∞-norms and L^2-norms. We first examine briefly first-order linear ordinary differential equations. Let β be a constant and $f = f(t)$ be a function. Consider

$$\frac{du}{dt} - \beta u = f(t).$$

A simple calculation shows

$$u(t) = e^{\beta t}u(0) + \int_0^t e^{\beta(t-s)} f(s)ds.$$

For any $T > 0$, we have

$$|u(t)| \le e^{\beta t}\big(|u(0)| + T \sup_{[0,T]} |f|\big) \quad \text{for any } t \in (0, T).$$

Here, we control the sup-norm of u in $[0, T]$ by the initial value $u(0)$ and the sup-norm of the nonhomogeneous term f in $[0, T]$.

Now we turn to PDEs. For convenience, we work in $\mathbb{R}^n \times \mathbb{R}_+$ and denote points by (x, t), with $x \in \mathbb{R}^n$ and $t \in \mathbb{R}_+$. Let a_i, b and f be functions in $\mathbb{R}^n \times \mathbb{R}_+$. We consider

$$\begin{aligned} u_t + \sum_{i=1}^n a_i(x,t)u_{x_i} + b(x,t)u &= f(x,t) \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ u(x,0) &= u_0(x) \quad \text{in } \mathbb{R}^n, \end{aligned} \tag{2.21}$$

where $a = (a_1, \cdots, a_n)$ satisfies

$$|a| \le \frac{1}{\kappa} \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \tag{2.22}$$

for a positive constant κ. Obviously, the initial hypersurface $\{t = 0\}$ is noncharacteristic. We may write the equation in (2.21) as

$$u_t + a(x,t) \cdot \nabla_x u + b(x,t)u = f(x,t).$$

We note that $a(x,t)\cdot\nabla_x + \partial_t$ is a directional derivative along the direction $(a(x,t), 1)$. With (2.22), it is easy to see that the vector $(a(x,t), 1)$ (starting from the origin) is in fact in the cone given by

$$\{(y, s); \kappa|y| \le s\} \subset \mathbb{R}^n \times \mathbb{R}.$$

This is a cone opening upward and with vertex at the origin.

For any point $P = (X, T) \in \mathbb{R}^n \times \mathbb{R}_+$, consider the cone $C_\kappa(P)$ (opening downward) with vertex at P defined by

$$C_\kappa(P) = \{(x,t); 0 < t < T,\ \kappa|x - X| < T - t\}.$$

We denote by $\partial_s C_\kappa(P)$ and $\partial_- C_\kappa(P)$ the side and bottom of the boundary respectively, i.e.,

$$\begin{aligned} \partial_s C_\kappa(P) &= \{(x,t); 0 < t \le T,\ \kappa|x - X| = T - t\}, \\ \partial_- C_\kappa(P) &= \{(x,0); \kappa|x - X| \le T\}. \end{aligned}$$

We note that $\partial_- C_\kappa(P)$ is simply a closed disc centered at X with radius T/κ in $\mathbb{R}^n \times \{0\}$. For any $(x,t) \in \partial_s C_\kappa(P)$, let $a(x,t)$ be a vector in $\mathbb{R}^n$ satisfying (2.22).

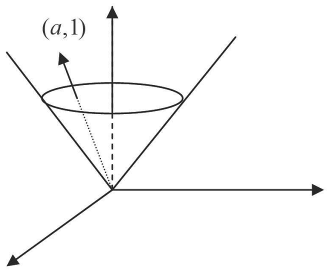

FIGURE 2.10. The cone with the vertex at the origin.

Then the vector $(a(x,t),1)$, if positioned at (x,t), points outward from the inside of the cone $C_\kappa(P)$. Hence for a function defined only in $\bar{C}_\kappa(P)$, it makes sense to calculate $u_t + a\cdot\nabla_x u$ at (x,t). This is in particular true when (x,t) is the vertex P.

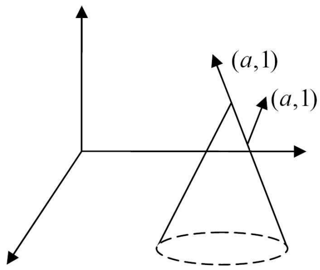

FIGURE 2.11. The cone $C_\kappa(P)$ and positions of vectors.

Now we calculate the unit outward normal vector of $\partial_s C_\kappa(P)\setminus\{P\}$. Set

$$\varphi(x,t) = \kappa|x-X| - (T-t).$$

Obviously, $\partial_s C_\kappa(P)\setminus\{P\}$ is a part of $\{\varphi = 0\}$. Then for any $(x,t)\in\partial_s C_\kappa(P)\setminus\{P\}$

$$\nabla\varphi = (\nabla_x\varphi, \varphi_t) = \big(\kappa\frac{x-X}{|x-X|}, 1\big).$$

Therefore, the unit outward normal vector γ of $\partial_s C_\kappa(P)\setminus\{P\}$ at (x,t) is given by

$$\gamma = \frac{1}{\sqrt{\kappa^2+1}}\big(\kappa\frac{x-X}{|x-X|}, 1\big).$$

For $n=1$, the cone $C_\kappa(P)$ is a triangle bounded by straight lines $\pm\kappa(x-X) = T-t$ and $t=0$. The side of the cone consists of two line segments

$$\text{the left segment:}\quad -\kappa(x-X) = T-t,\ 0<t\le T,\ \text{with a normal vector } (-\kappa,1),$$
$$\text{the right segment:}\quad \kappa(x-X) = T-t,\ 0<t\le T,\ \text{with a normal vector } (\kappa,1).$$

It is easy to see that the integral curve associated with (2.21) starting from P and going to the initial surface $\mathbb{R}^n\times\{0\}$ stays in $C_\kappa(P)$. In fact, this is true for any point $(x,t)\in C_\kappa(P)$. This suggests that solutions in $C_\kappa(P)$ should depend only on f in $C_\kappa(P)$ and the initial value u_0 on $\partial_- C_\kappa(P)$. The following result, proved using a maximum principle type argument, confirms this.

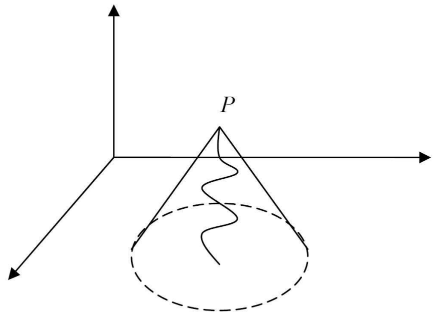

FIGURE 2.12. The domain of dependence.

THEOREM 2.10. *Let a_i, b and f be continuous functions in $\mathbb{R}^n \times \mathbb{R}_+$ satisfying (2.22) and u be a C^1-solution of (2.21) in $\mathbb{R}^n \times \mathbb{R}_+$. Then for any $P = (X, T) \in \mathbb{R}^n \times \mathbb{R}_+$,*

$$\sup_{C_\kappa(P)} |e^{-\beta t} u| \leq \sup_{\partial_- C_\kappa(P)} |u_0| + T \sup_{C_\kappa(P)} |e^{-\beta t} f|,$$

where β is a constant such that

$$b \geq -\beta \quad \textit{in } C_\kappa(P).$$

If $b \geq 0$, we take $\beta = 0$ and have

$$\sup_{C_\kappa(P)} |u| \leq \sup_{\partial_- C_\kappa(P)} |u_0| + T \sup_{C_\kappa(P)} |f|.$$

PROOF. Take any positive number $\beta' > \beta$ and let

$$M = \sup_{\partial_- C_\kappa(P)} |u_0|, \quad F = \sup_{C_\kappa(P)} |e^{-\beta' t} f|.$$

We will prove

$$|e^{-\beta' t} u(x,t)| \leq M + tF \quad \text{for any } (x,t) \in C_\kappa(P).$$

For the upper bound, we consider

$$w(x,t) = e^{-\beta' t} u(x,t) - M - tF.$$

A simple calculation shows

$$w_t + \sum_{i=1}^n a_i(x,t) w_{x_i} + \big(b(x,t) + \beta'\big) w = -\big(b(x,t) + \beta'\big)(M + tF) + e^{-\beta' t} f - F.$$

The expression in the right hand side is nonpositive since $b + \beta' > 0$. Hence

$$w_t + a \cdot \nabla_x w + \big(b + \beta'\big) w \leq 0 \quad \text{in } C_\kappa(P).$$

Let w attain its maximum in $\overline{C_\kappa(P)}$ at $(x_0, t_0) \in \overline{C_\kappa(P)}$. We prove $w(x_0, t_0) \leq 0$. First, it is obvious if $(x_0, t_0) \in \partial_- C_\kappa(P)$, since $w(x_0, t_0) = u_0(x_0) - M \leq 0$ by the definition of M. If $(x_0, t_0) \in C_\kappa(P)$, then

$$(w_t + a \cdot \nabla_x w)|_{(x_0, t_0)} = 0.$$

If $(x_0, t_0) \in \partial_s C_\kappa(P)$, by the position of the vector $(a(x_0, t_0), 1)$ relative to the cone $C_\kappa(P)$, we have

$$(w_t + a \cdot \nabla_x w)|_{(x_0,t_0)} \geq 0.$$

Hence in both cases, we obtain

$$(b + \beta')w|_{(x_0,t_0)} \leq 0.$$

Since $b + \beta' > 0$, this implies $w(x_0, t_0) \leq 0$. (We need the positivity of $b + \beta'$ here!) Hence $w(x_0, t_0) \leq 0$ in all three cases. Therefore, $w \leq 0$ in $C_\kappa(P)$, or,

$$u(x, t) \leq e^{\beta' t}(M + tF) \quad \text{for any } (x, t) \in C_\kappa(P).$$

We simply let $\beta' \to \beta$ to get the desired upper bound. For the lower bound, we consider

$$v(x, t) = e^{-\beta' t}u(x, t) + M + tF.$$

The argument is similar and hence omitted. □

For $n = 1$, (2.21) has the form

$$u_t + a(x, t)u_x + b(x, t)u = f(x, t).$$

In this case, it is straightforward to see that

$$(w_t + aw_x)|_{(x_0,t_0)} \geq 0,$$

if w assumes its maximum at $(x_0, t_0) \in \partial_s C_\kappa(P)$. We first note that $\partial_t + \frac{1}{\kappa}\partial_x$ and $\partial_t - \frac{1}{\kappa}\partial_x$ are directional derivatives along straight lines $t - t_0 = \kappa(x - x_0)$ and $t - t_0 = -\kappa(x - x_0)$ respectively. Since w assumes its maximum at (x_0, t_0), we have

$$(w_t + \frac{1}{\kappa}w_x)|_{(x_0,t_0)} \geq 0, \quad (w_t - \frac{1}{\kappa}w_x)|_{(x_0,t_0)} \geq 0.$$

In fact, one of them is zero if $(x_0, t_0) \in \partial_s C_\kappa(P) \setminus \{P\}$. Then we obtain

$$w_t(x_0, t_0) \geq \frac{1}{k}|w_x|(x_0, t_0) \geq |aw_x|(x_0, t_0).$$

Theorem 2.10 implies the uniqueness of solutions of (2.21). Let u_1 and u_2 be two solutions of (2.21). Then $u_1 - u_2$ satisfies (2.21) with $f = 0$ in $C_\kappa(P)$ and $u_0 = 0$ on $\partial_- C_\kappa(P)$. Hence $u_1 - u_2 = 0$ in $\mathcal{C}_\kappa(P)$.

Theorem 2.10 also shows that the value $u(P)$ depends only on f in $C_\kappa(P)$ and u_0 on $\partial_- C_\kappa(P)$. Hence $C_\kappa(P)$ contains the domain of dependence of $u(P)$ on f and $\partial_- C_\kappa(P)$ contains the domain of dependence of $u(P)$ on u_0. In fact, the domain of dependence of $u(P)$ on f is the integral curve through P in $C_\kappa(P)$ and the domain of dependence of $u(P)$ on u_0 is the intercept of this integral curve with the initial hyperplane $\{t = 0\}$. We consider this from another point of view. For simplicity, we assume f is identically zero and the initial value u_0 at $t = 0$ is zero outside a bounded domain $D_0 \subset \mathbb{R}^n$. Then for any $t > 0$, $u(\cdot, t) = 0$ outside

$$D_t = \{(x, t); \kappa|x - x_0| < t \text{ for some } x_0 \in D_0\}.$$

In other words, u_0 only influences u in $\cup_{\{t>0\}} D_t$. This is the so-called finite-speed propagation.

Now we use the same idea to estimate derivatives.

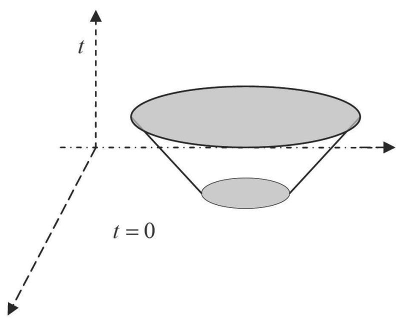

FIGURE 2.13. The domain of influence.

THEOREM 2.11. *Let a_i, b and f be C^1-functions in $\mathbb{R}^n \times \mathbb{R}_+$ satisfying (2.22) and u be a C^2-solution of (2.21) in $\mathbb{R}^n \times \mathbb{R}_+$. Then for any $P = (X, T) \in \mathbb{R}^n \times \mathbb{R}_+$,*

$$|u|_{C^1(C_\kappa(P))} \leq C\big(|u_0|_{C^1(\partial_- C_\kappa(P))} + |f|_{C^1(C_\kappa(P))}\big),$$

where C is a positive constant depending only on T and the C^1-norms of a_i and b in $C_\kappa(P)$.

PROOF. In the following, we prove

$$\sup_{C_\kappa(P)} |\nabla_x u| \leq C\big(\sup_{\partial_- C_\kappa(P)} |\nabla_x u_0| + \sup_{C_\kappa(P)} |\nabla_x f| + \sup_{C_\kappa(P)} |u|\big),$$

where C is a positive constant depending only on T and the C^1-norms of a_i and b in $C_\kappa(P)$. Then, (2.21) implies a similar estimate for u_t in $C_\kappa(P)$. Hence, we obtain the desired estimate by Theorem 2.10. The proof for for $n = 1$ is easier and we will do that first.

Proof for $n = 1$. In this case, the equation in (2.21) has the form

$$u_t + a(x,t)u_x + b(x,t)u = f(x,t).$$

By differentiating with respect to x, we have

$$(u_x)_t + a(u_x)_x + (b + a_x)u_x = f_x - b_x u.$$

We view this as an equation for u_x, which has the structure as the equation for u. Then by Theorem 2.10, we obtain

$$\sup_{C_\kappa(P)} |u_x| \leq e^{\beta_1 T}\big(\sup_{\partial_- C_\kappa(P)} |u_{0x}| + T \sup_{C_\kappa(P)} |f_x - b_x u|\big),$$

where β_1 is a nonnegative constant satisfying

$$b + a_x \geq -\beta_1 \quad \text{in } C_\kappa(P).$$

Proof for $n > 1$. Now we consider the general case. Differentiating the equation in (2.21) with respect to x_k, we obtain

$$(u_{x_k})_t + \sum_{i=1}^n a_i (u_{x_k})_{x_i} + b u_{x_k} + \sum_{i=1}^n a_{i,x_k} u_{x_i} = f_{x_k} - b_{x_k} u.$$

We note that this is not an equation for u_{x_k} as other derivatives of u are present. To proceed, we set

$$v = \frac{1}{2}\sum_{k=1}^{n} u_{x_k}^2.$$

Then

$$v_t = \sum_{k=1}^{n} u_{x_k}u_{tx_k}, \quad v_{x_i} = \sum_{k=1}^{n} u_{x_k}u_{x_ix_k}.$$

By multiplying the equation above by u_{x_k} and summing over $k = 1, \cdots, n$, we obtain

$$v_t + \sum_{i=1}^{n} a_i v_{x_i} + 2bv + \sum_{i,k=1}^{n} a_{i,x_k}u_{x_i}u_{x_k} = \sum_{k=1}^{n}(f_{x_k} - b_{x_k}u)u_{x_k}.$$

This is not an equation for v as products of two first order derivatives of u are present in the fourth term in the left hand side. However, any such product can be controlled by v. In fact, the Cauchy inequality yields

$$|u_{x_i}u_{x_k}| \le \frac{1}{2}\big(u_{x_i}^2 + u_{x_k}^2\big).$$

Hence, a summation implies

$$\sum_{i,k=1}^{n} |u_{x_i}u_{x_k}| \le n\sum_{i=1}^{n} u_{x_i}^2 = 2nv.$$

Then we have

$$v_t + \sum_{i=1}^{n} a_i v_{x_i} + 2bv - 2n\sum_{i=1}^{n} |\nabla_x a_i|_{L^\infty(C_\kappa(P))} v \le \sum_{k=1}^{n}(f_{x_k} - b_{x_k}u)u_{x_k}.$$

We apply the Cauchy inequality in a similar way to the expression in the right hand side. Therefore, we obtain

$$v_t + \sum_{i=1}^{n} a_i v_{x_i} + (2b - 2n\sum_{i=1}^{n} |\nabla_x a_i|_{L^\infty(C_\kappa(P))} - 1)v \le \frac{1}{2}\sum_{k=1}^{n}(f_{x_k} - b_{x_k}u)^2.$$

This is a first order partial differential inequality rather than an equation for v. We cannot apply Theorem 2.10 directly. However, the method in the proof still applies. We only need an estimate on the upper bound of v since v is nonnegative. Hence,

$$\sup_{C_\kappa(P)} v \le e^{\beta_1 T}\big(\sup_{\partial_- C_\kappa(P)} v + \frac{T}{2}\sup_{C_\kappa(P)}\sum_{k=1}^{n}(f_{x_k} - b_{x_k}u)^2\big),$$

where β_1 is a nonnegative constant such that

$$2b - 2n\sum_{i=1}^{n} |\nabla_x a_i|_{L^\infty(C_\kappa(P))} - 1 \ge -\beta_1.$$

With the definition of v, we get the desired estimate for $\nabla_x u$ easily. □

Next, we derive an estimate of the L^2-norm of u in terms of the L^2-norms of f and u_0. We will do this in another class of domains rather than cones. For fixed $T, \bar{t} > 0$, consider

$$D_{\kappa,T,\bar{t}} = \{(x,t); \kappa|x| < \bar{t} - t,\ 0 < t < T\}.$$

In other words,

$$D_{\kappa,T,\bar{t}} = C_\kappa(0,\bar{t}) \cap \{0 < t < T\}.$$

If $T \geq \bar{t}$, then $D_{\kappa,T,\bar{t}} = C_\kappa(0,\bar{t})$.

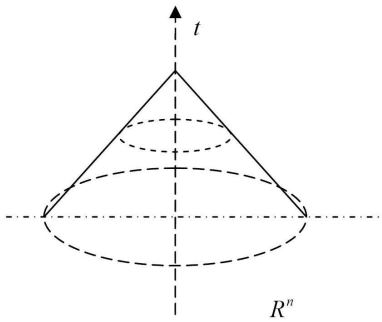

FIGURE 2.14. An integral domain.

In the following, we are only interested in the case $T < \bar{t}$. We denote by $\partial_- D_{\kappa,T,\bar{t}}$, $\partial_s D_{\kappa,T,\bar{t}}$ and $\partial_+ D_{\kappa,T,\bar{t}}$ the bottom, the side and the top of the boundary, i.e.,

$$\begin{aligned}
\partial_- D_{\kappa,T,\bar{t}} =& \{(x,0); \kappa|x| < \bar{t}\},\\
\partial_s D_{\kappa,T,\bar{t}} =& \{(x,t); \kappa|x| = \bar{t} - t,\ 0 < t < T\},\\
\partial_+ D_{\kappa,T,\bar{t}} =& \{(x,T); \kappa|x| < \bar{t} - T\}.
\end{aligned}$$

We recall the Green's formula. Let Ω be a piecewise C^1-domain in $\mathbb{R}^n$ and $\gamma = (\gamma_1, \cdots, \gamma_n)$ be the unit exterior normal vector to $\partial\Omega$. Then for any $u \in C^1(\Omega) \cap C(\bar{\Omega})$,

$$\int_\Omega u_{x_i} = \int_{\partial\Omega} u\gamma_i \quad \text{for } i = 1, \cdots, n.$$

THEOREM 2.12. *Let a_i be C^1-functions satisfying (2.22), b and f be continuous functions and u be a C^1-solution of (2.21) in $\mathbb{R}^n \times \mathbb{R}_+$. Then for any $0 < T < \bar{t}$,*

$$\int_{D_{\kappa,T,\bar{t}}} e^{-\alpha t} u^2 \leq \int_{\partial_- D_{\kappa,T,\bar{t}}} u_0^2 + \int_{D_{\kappa,T,\bar{t}}} e^{-\alpha t} f^2,$$

where α is a positive constant depending only on the C^1-norms of a_i and the sup-norm of b^- in $D_{\kappa,T,\bar{t}}$.

PROOF. For a nonnegative α, we multiply the equation in (2.21) by $2e^{-\alpha t}u$. By

$$\begin{aligned}
2e^{-\alpha t} u u_t =& (e^{-\alpha t} u^2)_t + \alpha e^{-\alpha t} u^2,\\
2a_i e^{-\alpha t} u u_{x_i} =& (e^{-\alpha t} a_i u^2)_{x_i} - e^{-\alpha t} a_{i,x_i} u^2,
\end{aligned}$$

we have

$$(e^{-\alpha t} u^2)_t + \sum_{i=1}^n (e^{-\alpha t} a_i u^2)_{x_i} + e^{-\alpha t}(\alpha + 2b - \sum_{i=1}^n a_{i,x_i}) u^2 = 2e^{-\alpha t} u f.$$

We simply write $D = D_{\kappa,T,\bar{t}}$. An integration in D yields

$$\int_{\partial_+ D} e^{-\alpha t}u^2 + \int_{\partial_s D} e^{-\alpha t}(\gamma_t + \sum_{i=1}^n a_i\gamma_i)u^2 + \int_D e^{-\alpha t}(\alpha + 2b - \sum_{i=1}^n a_{i,x_i})u^2$$
$$= \int_{\partial_- D} u_0^2 + \int_D 2e^{-\alpha t}uf,$$

where $(\gamma_1, \cdots, \gamma_n, \gamma_t)$ is the unit exterior normal vector of $\partial_s D$ given by

$$(\gamma_1, \cdots, \gamma_n, \gamma_t) = \frac{1}{\sqrt{1+\kappa^2}}\big(\kappa\frac{x}{|x|}, 1\big).$$

First, we note

$$\gamma_t + \sum_{i=1}^n a_i\gamma_i \geq 0 \quad \text{on } \partial_s D.$$

Next, we choose α such that

$$\alpha + 2b - \sum_{i=1}^n a_{i,x_i} \geq 2 \quad \text{in } D.$$

Then

$$2\int_D e^{-\alpha t}u^2 \leq \int_{\partial_- D} u_0^2 + \int_D 2e^{-\alpha t}uf.$$

Here we simply dropped integrals over $\partial_+ D$ and $\partial_s D$ since they are nonnegative. The Cauchy inequality implies

$$\int_D 2e^{-\alpha t}uf \leq \int_D e^{-\alpha t}u^2 + \int_D e^{-\alpha t}f^2.$$

We then have the desired result. □

The domain $D_{\kappa,T,\bar{t}}$ is important for us to discuss properties of solutions in $\mathbb{R}^n \times (0,T)$ for a fixed T. We note that $\bar{t}$ enters the estimate in Theorem 2.12 only through the domain $D_{\kappa,T,\bar{t}}$. Hence, we may let $\bar{t} \to \infty$ to get the following consequence.

Corollary 2.13. *Let a_i be C^1-functions satisfying (2.22), b be continuous functions and u be a C^1-solution of (2.21) in $\mathbb{R}^n \times \mathbb{R}_+$. For any $T > 0$, if $f \in L^2(\mathbb{R}^n \times (0,T))$ and $u_0 \in L^2(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n \times (0,T)} e^{-\alpha t}u^2 \leq \int_{\mathbb{R}^n} u_0^2 + \int_{\mathbb{R}^n \times (0,T)} e^{-\alpha t}f^2,$$

where α is a positive constant depending only on the C^1-norms of a_i and the sup-norm of b^- in $\mathbb{R}^n \times (0,T)$.

We point out that there are no decay assumptions on u.

A natural question for anyone who just starts studying PDEs is why we need a priori estimates. Now we use (2.21) to demonstrate that with the estimate in Corollary 2.13 we can actually prove the existence of (weak) solutions of initial value problems (2.21). We use the homogeneous initial value, i.e., $u_0 = 0$, for an illustration.

We first introduce the notion of weak solutions. In the following, we fix a $T > 0$ and consider functions in

$$D_T = \mathbb{R}^n \times (0, T).$$

Denote by $C_0^\infty(D_T)$ the collection of smooth functions in D_T with compact supports in D_T and by $\widetilde{C}_0^\infty(D_T)$ the collection of smooth functions in D_T with compact supports in x-directions. In other words, functions in $C_0^\infty(D_T)$ vanish for large x and for t close to 0 and T and functions in $\widetilde{C}_0^\infty(D_T)$ vanish only for large x.

Set

$$Lu = u_t + \sum_{i=1}^n a_i(x,t)u_{x_i} + b(x,t)u \quad \text{in } D_T. \tag{2.23}$$

For any $u, v \in \widetilde{C}_0^\infty$, we integrate vLu in D_T. To this end, we write

$$vLu = (uv)_t + \sum_{i=1}^n (a_i uv)_{x_i} - u\big(v_t + \sum_{i=1}^n (a_i v)_{x_i} - bv\big).$$

By a simple integration in D_T, we have

$$\int_{D_T} vLu = \int_{\mathbb{R}^n\times\{t=T\}} uv - \int_{\mathbb{R}^n\times\{t=0\}} uv - \int_{D_T} u\big(v_t + \sum_{i=1}^n (a_i v)_{x_i} - bv\big).$$

We note that there are no derivatives of u in the right-hand side.

Next, we define the adjoint differential operator L^* of L by

$$L^*v = -v_t - \sum_{i=1}^n (a_i v)_{x_i} + bv = -v_t - \sum_{i=1}^n a_i v_{x_i} + (b - \sum_{i=1}^n a_{i,x_i})v.$$

Then we have

$$\int_{D_T} vLu = \int_{D_T} uL^*v + \int_{\mathbb{R}^n\times\{t=T\}} uv - \int_{\mathbb{R}^n\times\{t=0\}} uv \quad \text{for any } u, v \in \widetilde{C}_0^\infty(D_T).$$

Definition 2.14. An L^2-function u is a weak solution of $Lu = f$ in D_T if

$$\int_{D_T} uL^*v = \int_{D_T} fv \quad \text{for any } v \in C_0^\infty(D_T).$$

Now we are ready to prove the existence of weak solutions of (2.21) with homogeneous initial values. The proof requires the Hahn-Banach Theorem and the Riesz Representation Theorem in functional analysis. In the following, we denote by $(\cdot,\cdot)$ the L^2-inner product in D_T.

We first note that, with L in (2.23), we can rewrite the estimate in Corollary 2.13 as

$$\|u\|_{L^2(\mathbb{R}^n\times(0,T))} \le C\big(\|u(\cdot,0)\|_{L^2(\mathbb{R}^n)} + \|Lu\|_{L^2(\mathbb{R}^n\times(0,T))}\big), \tag{2.24}$$

where C is a positive constant depending only on T, the C^1-norms of a_i and the sup-norm of b^- in $\mathbb{R}^n \times (0,T)$. This holds for any $u \in C^1(\mathbb{R}^n \times [0,T))$ with $Lu \in L^2(\mathbb{R}^n \times (0,T))$ and $u(\cdot,0) \in L^2(\mathbb{R}^n)$.

Theorem 2.15. *Let a_i be C^1-functions satisfying (2.22) and b a continuous function in D_T. Then for any $f \in L^2(D_T)$, there exists a $u \in L^2(D_T)$ such that*

$$(u, L^*v) = (f, v) \quad \textit{for any } v \in \widetilde{C}_0^\infty(D_T) \textit{ with } v = 0 \textit{ on } t = T.$$

Moreover,

$$\|u\|_{L^2(D_T)} \leq C\|f\|_{L^2(D_T)},$$

where C is a positive constant depending only on T, the C^1-norms of a_i and the sup-norm of b^- in $\mathbb{R}^n \times (0,T)$.

The function u in Theorem 2.15 is called a weak solution of (2.21) with $u_0 = 0$.

PROOF. We first note that L^* has a similar structure as L except different signs for the principal part, the part involving derivatives. Applying (2.24) to L^*, we have

$$\|v\|_{L^2(D_T)} \leq C\|L^*v\|_{L^2(D_T)} \quad \text{for any } v \in \widetilde{C}_0^\infty(D_T) \text{ with } v = 0 \text{ on } t = T,$$

where C is a positive constant depending only on T, the C^1-norms of a_i and the sup-norm of b. Here we view $t = T$ as the initial curve for the domain D_T and v has a homogeneous initial value. We denote by $\widehat{C}^\infty(D_T)$ the collection of functions $v \in \widetilde{C}_0^\infty(D_T)$ with $v = 0$ on $t = T$.

Consider the linear functional $F : L^*\widehat{C}^\infty(D_T) \to \mathbb{R}$ given by

$$F(L^*v) = (f, v) \quad \text{for any } v \in \widehat{C}^\infty(D_T).$$

Then

$$|F(L^*v)| \leq \|f\|_{L^2(D_T)}\|v\|_{L^2(D_T)} \leq C\|f\|_{L^2(D_T)}\|L^*v\|_{L^2(D_T)}.$$

Hence F is a well-defined bounded linear functional on the subspace $L^*\widehat{C}^\infty(D_T)$ of $L^2(D_T)$. Thus we apply the Hahn-Banach Theorem to obtain a bounded linear extension of F defined on $L^2(D_T)$ such that

$$\|F\| \leq C\|f\|_{L^2(D_T)}.$$

Here, $\|F\|$ is the norm of the linear functional F. By the Riesz Representation Theorem, there exists a $u \in L^2(D_T)$ such that

$$F(z) = (u, z) \quad \text{for any } z \in L^2(D_T),$$

and

$$\|u\|_{L^2(D_T)} \leq C\|f\|_{L^2(D_T)}.$$

Now we restrict z back to $L^*\widehat{C}^\infty(D_T)$ to obtain the desired result. □

Theorem 2.15 asserts the existence of weak solutions of (2.21) with $u_0 = 0$. Now we illustrate that u is indeed a classical solution if u is C^1 in D_T and continuous up to $t = 0$. Under these extra assumptions on u, we integrate by parts (u, L^*v) to get

$$\int_{D_T} vLu + \int_{\mathbb{R}^n\times\{t=0\}} uv = \int_{D_T} fv \quad \text{for any } v \in \widetilde{C}_0^\infty(D_T) \text{ with } v = 0 \text{ on } t = T.$$

There are no boundary integrals on vertical sides and on the upper side since v vanishes there. In particular,

$$\int_{D_T} vLu = \int_{D_T} fv \quad \text{for any } v \in C_0^\infty(D_T).$$

Since $C_0^\infty(D_T)$ is dense in $L^2(D_T)$, we conclude

$$Lu = f \quad \text{in } D_T.$$

Therefore,

$$\int_{\mathbb{R}^n\times\{t=0\}} uv = 0 \quad \text{for any } v \in \widetilde{C}_0^\infty(D_T) \text{ with } v = 0 \text{ on } t = T,$$

or

$$\int_{\mathbb{R}^n} u(0)\varphi = 0 \quad \text{for any } \varphi \in C_0^\infty(\mathbb{R}^n).$$

Again by the density of $C_0^\infty(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$, we conclude

$$u(\cdot, 0) = 0 \quad \text{on } \mathbb{R}^n.$$

Therefore, a crucial step in passing weak solutions to classical solutions is to improve the regularity of weak solutions.

Now we summarize the process to establish solutions by using energy estimates.

Step 1. Prove a priori estimates;

Step 2. Prove the existence of weak solutions by methods in functional analysis;

Step 3. Improve regularities of weak solutions.

In fact, the last step is also closely related to a priori estimates of derivatives of solutions.

Next, we make a crucial observation. The requirement that u possess continuous derivatives can be weakened. It suffices to assume that u has derivatives in the L^2-sense so that the integration by parts can be performed. Then we conclude $Lu = f$ almost everywhere. This clearly suggests the need to introduce a new space of functions, functions with only L^2-derivatives. This is the so-called Sobolev space, which plays a fundamental role in PDEs and will be studied in details in the future.

Exercises

(1) Find solutions of the following initial-value problems in $\mathbb{R}^2$:
   (a) $2u_y - u_x + xu = 0$ with $u(x, 0) = 2xe^{x^2/2}$.
   (b) $u_y + (1 + x^2)u_x - u = 0$ with $u(x, 0) = \arctan x$.

(2) Solve the following initial-value problems:
   (a) $u_y + u_x = u^2$ with $u(x, 0) = h(x)$.
   (b) $u_z + xu_x + yu_y = u$ with $u(x, y, 0) = h(x, y)$.

(3) Let B_1 be the unit disc in $\mathbb{R}^2$ and a and b be continuous functions in $\bar{B}_1$ with $a(x, y)x + b(x, y)y > 0$ on ∂B_1. Assume u is a C^1-solution of

$$a(x, y)u_x + b(x, y)u_y = -u \quad \text{in } \bar{B}_1.$$

Prove that u vanishes identically.

(4) Find an equation in $\mathbb{R}^2$ of the form $a(x, y)u_x + b(x, y)u_y = 0$ with smooth coefficients such that
   (a) any horizontal line $\{y = c\}$ is non-characteristic;
   (b) there does not exist a solution in entire $\mathbb{R}^2$ for any nonconstant initial value prescribed on $\{y = 0\}$.

(5) Let α be a number and $h = h(x)$ be a continuous function in $\mathbb{R}$. Consider

$$\begin{aligned} yu_x + xu_y &= \alpha u, \\ u(x, 0) &= h(x). \end{aligned}$$

(a) Find all points on $\{y = 0\}$ where $\{y = 0\}$ is characteristic. What is the compatibility condition of h at these points?
(b) Away from points in (a), find the solution of the initial-value problem. What is the domain of this solution in general?
(c) For cases $h(x) = x, \alpha = 1$ and $h(x) = x, \alpha = 3$, check whether this solution can be extended over points in (a).
(d) For each point in (a), find all characteristic curves passing it. What is the relation of these curves and the domain in (b)?

(6) Let $\alpha \in \mathbb{R}$ be a real number and $h = h(x)$ be continuous in $\mathbb{R}$ and C^1 in $\mathbb{R} \setminus \{0\}$. Consider

$$\begin{aligned} xu_x + yu_y =& \alpha u, \\ u(x, 0) =& h(x). \end{aligned}$$

(a) Check that the straight line $\{y = 0\}$ is characteristic at each point.
(b) Find all h satisfying the compatibility condition on $\{y = 0\}$. (Consider three cases, $\alpha > 0$, $\alpha = 0$ and $\alpha < 0$.)
(c) For $\alpha > 0$, find two solutions with the given initial value on $\{y = 0\}$. (This is easy to do simply by inspecting the equation, especially for $\alpha = 2$.)

(7) In the plane, solve $u_y = 4u_x^2$ near the origin with $u(x, 0) = x^2$ on $\{y = 0\}$.

(8) In the plane, find two solutions of the initial-value problem

$$\begin{aligned} xu_x + yu_y + \frac{1}{2}(u_x^2 + u_y^2) &= u, \\ u(x, 0) &= \frac{1}{2}(1 - x^2). \end{aligned}$$

(9) In the plane, find two solutions of the initial-value problem

$$\begin{aligned} \frac{1}{4}u_x^2 + uu_y &= u, \\ u(x, \frac{1}{2}x^2) &= -\frac{1}{2}x^2. \end{aligned}$$

(10) Let a_i, b and f be continuous functions satisfying (2.22) and u be a C^1-solution of (2.21) in $\mathbb{R}^n \times \mathbb{R}_+$. Prove for any $P = (X, T) \in \mathbb{R}^n \times \mathbb{R}_+$

$$\sup_{C_\kappa(P)} |e^{-\alpha t}u| \le \sup_{\partial_- C_\kappa(P)} |u_0| + \frac{1}{\alpha + \inf_{C_\kappa(P)} b} \sup_{C_\kappa(P)} |e^{-\alpha t} f|,$$

where α is a constant such that

$$\alpha + \inf_{C_\kappa(P)} b > 0.$$

(11) Consider the following first-order differential system for (u, v) in $\mathbb{R} \times \mathbb{R}_+$

$$\begin{aligned} u_t - a(x, t)v_x + b_{11}(x, t)u + b_{12}(x, t)v &= f_1(x, t), \\ v_t - a(x, t)u_x + b_{21}(x, t)u + b_{22}(x, t)v &= f_2(x, t), \end{aligned}$$

with

$$u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x),$$

where a satisfies

$$|a(x,t)| \leq \frac{1}{\kappa}.$$

Derive an L^2-estimate of (u, v) in appropriate cones.

CHAPTER 3

An Overview of Second-Order PDEs

This chapter should be considered as an introduction to second-order linear PDEs. In Section 3.1, we introduce the Laplace equation, the heat equation and the wave equation. We also introduce their general forms, elliptic equations, parabolic equations and hyperbolic equations. We use energy methods to discuss the uniqueness of boundary-value problems and initial/boundary-value problems in Section 3.2 and use separation of variables to solve these problems in the plane in Section 3.3.

3.1. Classifications of Second Order Equations

The main focus in this section is second-order linear PDEs. We proceed similarly as in Section 2.1.

Let Ω be a domain in $\mathbb{R}^n$ containing the origin and L be a second-order linear differential operator given by

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x)u_{x_ix_j} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u \quad \text{in } \Omega,$$

where a_{ij}, b_i, c are continuous in Ω. Here a_{ij}, b_i, c are called coefficients of $u_{x_ix_j}$, u_{x_i}, u respectively. We usually assume that (a_{ij}) is a symmetric matrix in Ω. For the operator L, we define its principal symbol by

$$p(x;\xi) = \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \quad \text{for any } x \in \Omega, \xi \in \mathbb{R}^n.$$

For a given function f in Ω, we consider the equation

$$Lu = f(x) \quad \text{in } \Omega. \tag{3.1}$$

The function f is called the *non-homogeneous term.*

Let Σ be the hyperplane $\{x_n = 0\}$. We now prescribe values of u and its derivatives on Σ so that we can at least find all derivatives of u at the origin. Let u_0, u_1 be functions defined in a neighborhood of the origin in $\mathbb{R}^{n-1}$. Now we prescribe

$$u(x', 0) = u_0(x'),\ u_{x_n}(x', 0) = u_1(x') \quad \text{for any small } x' \in \mathbb{R}^{n-1}. \tag{3.2}$$

Then we can find all x'-derivatives of u and all x'-derivatives of u_{x_n} at the origin. In particular, we can determine all derivatives of u of order up to 2 at the origin except $u_{x_nx_n}$. To find this, we need to use the equation. If we assume

$$a_{nn}(0) \neq 0,$$

then we can find $u_{x_nx_n}(0)$ from (3.1). In this case, we can compute all derivatives of u of any order at the origin by using values u_0, u_1 and differentiating (3.1).

We usually call Σ the *initial hypersurface* and u_0, u_1 *initial values* or *Cauchy values*. The problem of solving (3.1) together with (3.2) is called the *initial-value problem* or *the Cauchy problem*.

More generally, consider the hypersurface Σ given by $\{\varphi = 0\}$ for a C^2-function φ in a neighborhood of the origin with $\nabla\varphi \neq 0$, with the origin on the hypersurface Σ, i.e., $\varphi(0) = 0$. We note that $\nabla\varphi$ is simply a normal vector of the hypersurface Σ. Without loss of generality, we assume $\varphi_{x_n}(0) \neq 0$. Then by the implicit function theorem, we solve $\varphi = 0$ around $x = 0$ for $x_n = \psi(x_1, \cdots, x_{n-1})$. Consider a change of variables

$$x \mapsto y = (x_1, \cdots, x_{n-1}, \varphi(x)).$$

This is a well defined transform in a neighborhood of the origin with a nonsingular Jacobian. Now we write the operator L in new variables y. Note

$$u_{x_i} = \sum_{k=1}^{n} y_{k,x_i} u_{y_k},$$

and

$$u_{x_ix_j} = \sum_{k,l=1}^{n} y_{k,x_i} y_{l,x_j} u_{y_ky_l} + \sum_{k=1}^{n} y_{k,x_ix_j} u_{y_k}.$$

Hence,

$$Lu = \sum_{k,l=1}^{n} \Big(\sum_{i,j=1}^{n} a_{ij} y_{k,x_i} y_{l,x_j} \Big) u_{y_ky_l} + \sum_{k=1}^{n} \Big(b_k + \sum_{i,j=1}^{n} a_{ij} y_{k,x_ix_j} \Big) u_{y_k} + cu.$$

The initial hypersurface Σ is given by $\{y_n = 0\}$ in new coordinates. With $y_n = \varphi$, the coefficient of $u_{y_ny_n}$ is given by

$$\sum_{i,j=1}^{n} a_{ij}(x) \varphi_{x_i} \varphi_{x_j}.$$

This is the principal symbol $p(x, \xi)$ evaluated at $\xi = \nabla\varphi(x)$.

Definition 3.1. For a linear operator L as in (3.1) defined in a neighborhood of $x_0 \in \mathbb{R}^n$, a C^1-hypersurface Σ passing x_0 is non-characteristic at x_0 if

$$\sum_{i,j=1}^{n} a_{ij}(x_0) \xi_i \xi_j \neq 0, \tag{3.3}$$

where ξ is a normal vector of Σ at x_0. Otherwise, it is characteristic at x_0. A hypersurface is non-characteristic if it is non-characteristic at every point.

When the hypersurface Σ is given by $\{\varphi = 0\}$ with $\nabla\varphi \neq 0$, its normal vector is given by $\nabla\varphi = (\varphi_{x_1}, \cdots, \varphi_{x_n})$. Hence we may take $\xi = \nabla\varphi(x_0)$ in (3.3). We note that the condition (3.3) is maintained under C^2-changes of local coordinates. By this condition, we can find successively values of all derivatives of u at x_0, as far as they exist. Then, we could write *formal* power series at x_0 for solutions of initial value problems. It would be actual representations of solutions u in a neighborhood of x_0, if u were known to be analytic. This process can be carried out for analytic initial values and analytic coefficients and nonhomogeneous terms, and the result is referred to as the Cauchy-Kowalevski theorem. We will discuss it later in this book.

Now we introduce a special class of linear differential operators.

DEFINITION 3.2. The linear differential operator L in (3.1) is *elliptic* at x_0 if

$$\sum_{i,j=1}^{n} a_{ij}(x_0)\xi_i\xi_j \neq 0 \quad \text{for any } \xi \in \mathbb{R}^n \setminus \{0\}.$$

Hence, linear differential equations are elliptic if there are no characteristic hypersurfaces. In other words, every hypersurface is non-characteristic. We already assumed that (a_{ij}) is an $n\times n$ symmetric matrix. Then L is elliptic at x_0 if $(a_{ij}(x_0))$ is a definite matrix, positive definite or negative definite.

We now turn our attention to second-order linear differential equations in $\mathbb{R}^2$, where complete classifications are available. Let Ω be a domain in $\mathbb{R}^2$ and consider

$$Lu = \sum_{i,j=1}^{2} a_{ij}u_{x_ix_j} + \sum_{i=1}^{2} b_i(x)u_{x_i} + c(x)u = f(x) \quad \text{in } \Omega. \tag{3.4}$$

Here we assume (a_{ij}) is a 2×2 symmetric matrix.

DEFINITION 3.3. Let L be a differential operator defined in $\Omega \subset \mathbb{R}^2$ as in (3.4).
(1) L is elliptic at $x \in \Omega$ if $\det(a_{ij}(x)) > 0$;
(2) L is hyperbolic at $x \in \Omega$ if $\det(a_{ij}(x)) < 0$;
(3) L is degenerate at $x \in \Omega$ if $\det(a_{ij}(x)) = 0$.

It is obvious that the ellipticity defined here coincides with that in Definition 3.2 for $n = 2$.

For the operator L in (3.4), the symmetric matrix (a_{ij}) always has two eigenvalues. Then

L is elliptic if the two eigenvalues have the same sign;

L is hyperbolic if the two eigenvalues have different signs;

L is degenerate if one of the eigenvalues vanishes.

The number of characteristic curves are determined by the type of operators. For the operator L in (3.4),

there are two characteristic curves if L is hyperbolic;

there is one characteristic curve if L is degenerate;

there are no characteristic curves if L is elliptic.

Now we study several important linear differential operators in $\mathbb{R}^2$.

In $\mathbb{R}^2$, the Laplace operator Δ is give by

$$\Delta u = u_{x_1x_1} + u_{x_2x_2}.$$

Obviously, the Laplace operator is elliptic. In polar coordinates

$$x_1 = r\cos\theta, \quad x_2 = r\sin\theta,$$

the Laplace operator Δ can be expressed by

$$\Delta u = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta}.$$

The equation

$$\Delta u = 0$$

is called the *Laplace equation* and its solutions are called *harmonic functions*. By writing $x = x_1$ and $y = x_2$, we can associate with a harmonic function $u(x, y)$ a conjugate harmonic function $v(x, y)$ such that u and v satisfy the first order system of Cauchy-Riemann equations

$$u_x = v_y, \quad u_y = -v_x.$$

Any such a pair gives an analytic function

$$f(z) = u(x, y) + iv(x, y)$$

of the complex argument $z = x + iy$. Physically, $(u, -v)$ is the velocity field of an irrotational, incompressible flow. Conversely, for any analytic function f, functions $u = \mathrm{Re}f$ and $v = \mathrm{Im}f$ are harmonic. In this way, we can find many nontrivial harmonic functions in the plane. For example, for any positive integer k, $\mathrm{Re}(x + iy)^k$ and $\mathrm{Im}(x + iy)^k$ are homogeneous harmonic polynomials of degree k. Next, with $e^z = e^{x+iy} = e^x \cos y + ie^x \sin y$, we know $e^x \cos y$ and $e^x \sin y$ are harmonic functions.

Although there are no characteristic curves for the Laplace operator, initial value problems are not well-posed.

Example 3.4. Consider the Laplace equation in $\mathbb{R}^2$

$$u_{xx} + u_{yy} = 0.$$

Set

$$u_k(x, y) = \frac{1}{k} \sin(kx)e^{ky}.$$

Then u is harmonic. Moreover,

$$u_{k,x}(x, y) = \cos(kx)e^{ky}, \quad u_{k,y}(x, y) = \sin(kx)e^{ky},$$

and hence

$$u_{k,x}^2(x, y) + u_{k,y}^2(x, y) = e^{2ky}.$$

Therefore,

$$u_{k,x}^2(x, 0) + u_{k,y}^2(x, 0) = 1 \quad \text{for any } x \in \mathbb{R} \text{ and any } k,$$

and

$$u_{k,x}^2(x, y) + u_{k,y}^2(x, y) \to \infty \quad \text{as } k \to \infty, \text{ for any } x \in \mathbb{R} \text{ and } y > 0.$$

There is no continuous dependence on initial values.

In $\mathbb{R}^2$, the wave operator $\Box$ is give by

$$\Box u = u_{x_2x_2} - u_{x_1x_1}.$$

Obviously, the wave operator is hyperbolic. It is actually called the one-dimensional wave operator. This is because the wave equation $\Box u = 0$ in $\mathbb{R}^2$ represents vibrations of strings or propagation of sound waves in tubes. Because of its physical interpretation, we write u as a function of two independent variables x and t. The variable x is commonly identified with position and t with time. Then the wave equation in $\mathbb{R}^2$ has the form

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

Two families of straight lines $t = \pm x$+constant are characteristic everywhere.

In $\mathbb{R}^2$, an important class of degenerate operators is given by

$$Lu = u_{x_2} - u_{x_1x_1}.$$

This is called the heat operator. The heat equation $u_{x_2} - u_{x_1x_1} = 0$ is satisfied by the temperature distribution in a heat-conducting insulated wire. As for the wave equation, we write u as a function of two independent variables x and t. Then the heat equation in $\mathbb{R}^2$ has the form

$$u_t - u_{xx} = 0.$$

Note that $\{t = 0\}$ is characteristic everywhere. If we prescribe $u(x, 0) = u_0(x)$ in an interval of $\{t = 0\}$, then using the equation we can compute all derivatives there. However, u_0 does not determine uniquely a solution even in a neighborhood of this interval. We will see later on that we need initial values on the entire initial line $\{t = 0\}$ to compute local solutions. *Hence, the problem with Cauchy values prescribed on characteristic hypersurfaces is not well-posed.*

Definition 3.5. Let L be a differential operator defined in $\Omega \subset \mathbb{R}^2$ as in (3.4). Then L is of mixed-type if L is elliptic in a subdomain in Ω and hyperbolic in another subdomain.

Example 3.6. Consider the Tricomi equation

$$u_{x_2x_2} + x_2u_{x_1x_1} = f \quad \text{in } \mathbb{R}^2.$$

It is elliptic if $x_2 > 0$, hyperbolic if $x_2 < 0$ and degenerate if $x_2 = 0$.

Characteristic curves arise also naturally in connection with the propagation of singularities. We only consider a simple case.

Theorem 3.7. *Let Ω be a domain in $\mathbb{R}^2$ and Γ be a curve in Ω. Assume a_{ij}, b_i, c, f are continuous functions in Ω and $u \in C^1(\Omega) \cap C^2(\Omega \setminus \Gamma)$ satisfies*

$$\sum_{i,j=1}^{2} a_{ij}u_{x_ix_j} + \sum_{i=1}^{2} b_i(x)u_{x_i} + c(x)u = f(x) \quad \text{in } \Omega \setminus \Gamma.$$

If $\nabla^2 u$ has a jump across Γ, then Γ is a characteristic curve.

Proof. Without loss of generality, we assume Γ divides Ω into two parts Ω_+ and Ω_-. For any function w in Ω and any $x_0 \in \Gamma$, we set

$$w_-(x_0) = \lim_{x \to x_0,\ x \in \Omega_-} w(x), \quad w_+(x_0) = \lim_{x \to x_0,\ x \in \Omega_+} w(x),$$

and

$$[w] = w_+ - w_- \quad \text{on } \Gamma.$$

Then

$$[u] = [u_{x_1}] = [u_{x_2}] = 0 \quad \text{on } \Gamma.$$

Let (ξ_1, ξ_2) be a normal vector along Γ. Then $\xi_2\partial_{x_1} - \xi_1\partial_{x_2}$ is a directional derivative along Γ. Hence on Γ

$$(\xi_2\partial_{x_1} - \xi_1\partial_{x_2})[u_{x_1}] = \xi_2[u_{x_1x_1}] - \xi_1[u_{x_1x_2}] = 0,$$
$$(\xi_2\partial_{x_1} - \xi_1\partial_{x_2})[u_{x_2}] = \xi_2[u_{x_1x_2}] - \xi_1[u_{x_2x_2}] = 0.$$

Rigorously, we need a limiting process to justify these two identities. By the continuity of a_{ij}, b_i, c and f in Ω, we have

$$a_{11}[u_{x_1x_2}] + 2a_{12}[u_{x_1x_2}] + a_{22}[u_{x_2x_2}] = 0 \quad \text{on } \Gamma.$$

The nontrivial vector $([u_{x_1x_1}], [u_{x_1x_2}], [u_{x_2x_2}])$ satisfies a 3×3 homogeneous linear system on Γ. Hence the coefficient matrix is singular. We then have on Γ

$$\det \begin{pmatrix} \xi_2 & -\xi_1 & 0 \\ 0 & \xi_2 & -\xi_1 \\ a_{11} & 2a_{12} & a_{22} \end{pmatrix} = 0,$$

or

$$a_{11}\xi_1^2 + 2a_{12}\xi_1\xi_2 + a_{22}\xi_2^2 = 0.$$

This yields the desired result. □

The Laplace operator, the wave operator and the heat operator can be generalized to higher dimensions.

EXAMPLE 3.8. The n-dimensional Laplace operator in $\mathbb{R}^n$ is defined by

$$\Delta u = \sum_{i=1}^n u_{x_ix_i},$$

and the Laplace equation is given by $\Delta u = 0$. Solutions are called harmonic functions. In this case, (a_{ij}) is the identity matrix. Obviously, Δ is elliptic. Note that Δ is invariant under rotations. If $x = Ay$ for an orthogonal matrix A, then

$$\sum_{i=1}^n u_{x_ix_i} = \sum_{i=1}^n u_{y_iy_i}.$$

For nonzero function f, we call equation $\Delta u = f$ the *Poisson equation.*

EXAMPLE 3.9. We denote points in $\mathbb{R}^{n+1}$ by $(x_1, \cdots, x_n, t)$. The heat operator in $\mathbb{R}^{n+1}$ is given by

$$Lu = u_t - \Delta_x u.$$

It is often called the n-dimensional heat operator. Its principal symbol is given by

$$p(x, t; \xi, \tau) = -|\xi|^2 \quad \text{for any } \xi \in \mathbb{R}^n, \tau \in \mathbb{R}.$$

A hypersurface $\{\varphi(x_1, \cdots, x_n, t) = 0\}$ is non-characteristic if

$$-|\nabla_x\varphi|^2 \neq 0.$$

Likewise, a hypersurface $\{\varphi(x, t) = 0\}$ is characteristic everywhere if $\nabla_x\varphi = 0$ and $\varphi_t \neq 0$. For example, any horizontal hyperplane $\{t = t_0\}$, for a fixed $t_0 \in \mathbb{R}$, is characteristic everywhere.

EXAMPLE 3.10. We denote points in $\mathbb{R}^{n+1}$ by $(x_1, \cdots, x_n, t)$. The wave operator $\Box$ in $\mathbb{R}^{n+1}$ is given by

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

It is often called the n-dimensional wave operator. Its principal symbol is given by

$$p(x, t; \xi, \tau) = \tau^2 - |\xi|^2 \quad \text{for any } \xi \in \mathbb{R}^n, \tau \in \mathbb{R}.$$

A hypersurface $\{\varphi(x_1, \cdots, x_n, t) = 0\}$ is non-characteristic if

$$\varphi_t^2 - |\nabla_x\varphi|^2 \neq 0.$$

For any $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$, the surface

$$|x - x_0|^2 = (t - t_0)^2$$

is characteristic everywhere except at (x_0, t_0). We note that this surface, smooth except at (x_0, t_0), is the boundary of two cones.

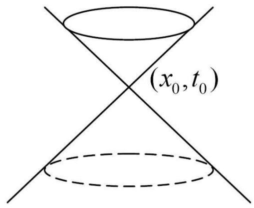

FIGURE 3.1. The characteristic surface.

The heat equation and the wave equation can be generalized to parabolic equations and hyperbolic equations in arbitrary dimensions. Again, we denote points in $\mathbb{R}^{n+1}$ by $(x_1, \cdots, x_n, t)$. Let a_{ij}, b_i, c and f be functions defined in a domain $\Omega \subset \mathbb{R}^{n+1}$, $i, j = 1, \cdots, n$. We assume $(a_{ij}(x,t))$ is an $n \times n$ positive definite matrix in Ω. A parabolic equation in Ω has the form

$$u_t - \sum_{i=1}^{n} a_{ij}(x,t) u_{x_i x_j} + \sum_{i=1}^{n} b_i(x,t) u_{x_i} + c(x,t)u = f(x,t) \quad \text{in } \Omega,$$

and a hyperbolic equation has the form

$$u_{tt} - \sum_{i=1}^{n} a_{ij}(x,t) u_{x_i x_j} + \sum_{i=1}^{n} b_i(x,t) u_{x_i} + c(x,t)u = f(x,t) \quad \text{in } \Omega.$$

3.2. Energy Estimates

In this section, we discuss the uniqueness of solutions of boundary-value problems for the Laplace equation and initial/boundary-value problems for the heat equation and the wave equation. Our main tool is the energy estimates. Specifically, we derive estimates of L^2-norms of solutions in terms of those of boundary values and/or initial values.

We first recall Green's formulas. Let Ω be a C^1-domain in $\mathbb{R}^n$ and $\gamma = (\gamma_1, \cdots, \gamma_n)$ be the unit exterior normal vector to $\partial\Omega$. Then for any $u \in C^1(\Omega) \cap C(\bar{\Omega})$,

$$\int_\Omega u_{x_i} = \int_{\partial\Omega} u\gamma_i \quad \text{for } i = 1, \cdots, n.$$

This is referred to as the divergence theorem. Now for any $w \in C^2(\Omega) \cap C^1(\bar{\Omega})$ and $v \in C^1(\Omega) \cap C(\bar{\Omega})$, substitute u above by vw_{x_i} to get

$$\int_\Omega (v w_{x_i x_i} + v_{x_i} w_{x_i}) = \int_{\partial\Omega} v w_{x_i} \gamma_i.$$

By summing up for $i = 1, \cdots, n$, we get the Green's formula of the following form

$$\int_\Omega \big(v\Delta w + \nabla v \cdot \nabla w\big) = \int_{\partial\Omega} v \frac{\partial w}{\partial n},$$

where $\frac{\partial w}{\partial n}$ is the normal derivative of w on $\partial\Omega$. If $v \equiv 1$, we get the divergence theorem

$$\int_\Omega \Delta w = \int_{\partial\Omega} \frac{\partial w}{\partial n}.$$

For any $v, w \in C^2(\Omega) \cap C^1(\bar{\Omega})$, we have another form of the Green's formula

$$\int_\Omega (v\Delta w - w\Delta v) = \int_{\partial\Omega} (v\frac{\partial w}{\partial n} - w\frac{\partial v}{\partial n}).$$

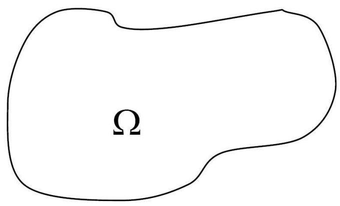

FIGURE 3.2. A smooth domain in $\mathbb{R}^n$.

Now we consider the Laplace equation. Assume $\Omega \subset \mathbb{R}^n$ is a bounded C^1-domain. The Dirichlet problem for the Laplace equation has the following form

$$\begin{aligned} \Delta u =& 0 \quad \text{in } \Omega, \\ u =& \varphi \quad \text{on } \partial\Omega, \end{aligned}$$

for a continuous function φ on $\partial\Omega$. We now prove that a C^2-solution, if exists, is unique. In fact, the difference w of any two C^2-solutions satisfies

$$\begin{aligned} \Delta w =& 0 \quad \text{in } \Omega, \\ w =& 0 \quad \text{on } \partial\Omega. \end{aligned}$$

We multiply the homogeneous Laplace equation by w and integrate in Ω. By the Green's formula, we have

$$0 = \int_\Omega w\Delta w = -\int_\Omega |\nabla w|^2 + \int_{\partial\Omega} w\frac{\partial w}{\partial n}.$$

With the homogeneous boundary value, we have

$$\int_\Omega |\nabla w|^2 = 0,$$

or $\nabla w = 0$ in Ω. Hence w is constant and this constant is zero since w is zero on the boundary. Obviously, the above argument applies to Dirichlet problems for Poisson equations $\Delta u = f$. In general, we have the following result.

LEMMA 3.11. *Assume $\Omega \subset \mathbb{R}^n$ is a bounded C^1-domain, f is a continuous function in $\bar{\Omega}$ and φ is a continuous function on $\partial\Omega$. Then*

$$\begin{aligned} \Delta u =& f \quad \text{in } \Omega, \\ u =& \varphi \quad \text{on } \partial\Omega, \end{aligned}$$

admits at most one solution in $C^2(\Omega) \cap C^1(\bar{\Omega})$.

By the maximum principle, the solution is in fact unique in $C^2(\Omega)\cap C(\bar{\Omega})$. We will discuss this in Chapter 4.

For the Neumann problem, we prescribe normal derivatives on the boundary. It has the following form

$$\begin{aligned}\Delta u =& 0 \quad \text{in } \Omega,\\ \frac{\partial u}{\partial n} =& \psi \quad \text{on } \partial\Omega,\end{aligned}$$

for a continuous function ψ on $\partial\Omega$. Similarly, we can prove that solutions are unique up to additive constants. We note that if there exists a solution of the Neumann problem, then ψ necessarily satisfies

$$\int_{\partial\Omega} \psi = 0.$$

This can be seen easily by the Green's formula.

Next, we derive an estimate of solutions of Dirichlet boundary value problems of the Poisson equation. We need the following result, which is referred to as the Poincarè Lemma.

LEMMA 3.12. *Let Ω be a bounded C^1-domain in $\mathbb{R}^n$ and u be a C^1-function in Ω with $u=0$ on $\partial\Omega$. Then*

$$\|u\|_{L^2(\Omega)} \le diam(\Omega)\|\nabla u\|_{L^2(\Omega)}.$$

Here $\mathrm{diam}(\Omega)$ denotes the diameter of Ω and is defined by

$$\mathrm{diam}(\Omega) = \sup_{x,y\in\Omega} |x-y|.$$

PROOF. For any $x_0' \in \mathbb{R}^{n-1}$, let $l_{x_0'}$ be the straight line through x_0 and normal to $\mathbb{R}^{n-1}\times\{0\}$. Consider those x_0' such that $l_{x_0'}\cap\Omega\neq\emptyset$. Let I be an open interval on $l_{x_0'}$ given by

$$I = \{(x_0', x_n); a < x_n < b\}$$

such that $I\subset\Omega$ and $(x_0',a),(x_0',b)\in\partial\Omega$. Since $u(x_0',a)=0$, then

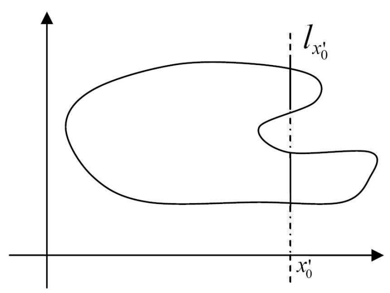

FIGURE 3.3. An integration along $l_{x_0'}$.

$$u(x_0', x_n) = \int_a^{x_n} u_{x_n}(x_0', s)ds \quad \text{for any } x\in(a,b).$$

The Hölder inequality yields

$$u^2(x_0', x_n) \le (x_n - a) \int_a^{x_n} u_{x_n}^2(x_0', s)ds \quad \text{for any } x \in (a, b).$$

By a simple integration along I, we have

$$\int_a^b u^2(x_0', x_n)dx_n \le (b-a)^2 \int_a^b u_{x_n}^2(x_0', x_n)dx_n.$$

By adding all possible intervals like I on $l_{x_0'}$, we then obtain

$$\int_{l_{x_0'} \cap \Omega} u^2(x_0', x_n)dx_n \le C_{x_0'}^2 \int_{l_{x_0'} \cap \Omega} u_{x_n}^2(x_0', x_n)dx_n,$$

where $C_{x_0'}$ is the length of $l_{x_0'}$ in Ω. Now a simple integration over x_0' yields the desired result. □

Now consider

$$\begin{aligned} \Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{3.5}$$

We note that u has a homogeneous Dirichlet boundary value on $\partial\Omega$.

Theorem 3.13. *Let $\Omega \subset \mathbb{R}^n$ be a bounded C^1-domain and $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ be a solution of (3.5). Then*

$$\|u\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)} \le C\|f\|_{L^2(\Omega)},$$

where C is a positive constant depending only on Ω.

Proof. Multiplying the equation in (3.5) by u and integrating over Ω, we obtain

$$\int_\Omega |\nabla u|^2 = -\int_\Omega uf.$$

The Hölder inequality yields

$$(\int_\Omega |\nabla u|^2)^2 \le \int_\Omega u^2 \cdot \int_\Omega f^2.$$

By Lemma 3.12, we get

$$\int_\Omega |\nabla u|^2 \le \big(\mathrm{diam}(\Omega)\big)^2 \int_\Omega f^2.$$

With Lemma 3.12 again, we have the desired estimate. □

Now we study initial/boundary-value problems for the heat equation. Suppose Ω is a bounded C^1- domain in $\mathbb{R}^n$, f is continuous in $\bar{\Omega} \times [0, \infty)$ and u_0 is continuous in $\bar{\Omega}$. Consider

$$\begin{aligned} u_t - \Delta u &= f \quad \text{in } \Omega \times (0, \infty), \\ u(\cdot, 0) &= u_0 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega \times (0, \infty). \end{aligned} \tag{3.6}$$

The geometric boundary of $\Omega \times (0, \infty)$ consists of two parts, $\Omega \times \{0\}$ and $\partial\Omega \times (0, \infty)$. We treat $\Omega \times \{0\}$ and $\partial\Omega \times (0, \infty)$ as the initial hypersurface and as the boundary respectively. We refer to values prescribed on these two parts as initial values and boundary values respectively. Problems of this type are usually called *initial/boundary value problems* or *mixed problems.* We note that u has a homogeneous Dirichlet boundary value on $\partial\Omega \times (0, \infty)$. We derive an estimate of the L^2-norms of solutions. In the following, we write $u(t) = u(\cdot, t)$ as a function in Ω for any $t \geq 0$.

THEOREM 3.14. *Let $u \in C^2(\Omega \times (0, \infty)) \cap C(\bar{\Omega} \times [0, \infty))$ be a solution of (3.6). Then*

$$\|u(t)\|_{L^2(\Omega)} \leq \|u_0\|_{L^2(\Omega)} + \int_0^t \|f(s)\|_{L^2(\Omega)} ds \quad \text{for any } t > 0.$$

Theorem 3.14 yields the uniqueness of solutions of (3.6). In fact, if $f \equiv 0$ and $u_0 \equiv 0$, then $u \equiv 0$. We also have the continuous dependence on initial values.

PROOF. Multiply the equation in (3.6) by u and integrate over Ω for each fixed $t > 0$. Note

$$uu_t - u\Delta u = \frac{1}{2}(u^2)_t - \sum_{i=1}^n (uu_{x_i})_{x_i} + |\nabla u|^2.$$

With $u(t) = 0$ on $\partial\Omega$, we have

$$\frac{1}{2}\frac{d}{dt}\int_\Omega u^2(t) + \int_\Omega |\nabla u(t)|^2 = \int_\Omega f(t)u(t).$$

Then for any $t > 0$

$$\int_\Omega |u(t)|^2 + 2\int_0^t \int_\Omega |\nabla u|^2 = \int_\Omega u_0^2 + 2\int_0^t \int_\Omega fu.$$

Set

$$E(t) = \|u(t)\|_{L^2(\Omega)}.$$

Then

$$\big(E(t)\big)^2 + 2\int_0^t \int_\Omega |\nabla u|^2 = \big(E(0)\big)^2 + 2\int_0^t \int_\Omega fu.$$

Differentiating with respect to t, we have

$$\begin{aligned} 2E(t)E'(t) \leq & 2E(t)E'(t) + 2\int_\Omega |\nabla u(t)|^2 = 2\int_\Omega f(t)u(t) \\ \leq & 2\|f(t)\|_{L^2(\Omega)} \cdot \|u(t)\|_{L^2(\Omega)}. \end{aligned}$$

Hence

$$E'(t) \leq \|f(t)\|_{L^2(\Omega)}.$$

Integrating from 0 to t, we obtain the desired estimate. □

Now we study initial/boundary-value problems for the wave equation. Again suppose Ω is a bounded C^1-domain in $\mathbb{R}^n$, f is continuous in $\Omega \times (0, \infty)$, u_0 is C^1 in Ω and u_1 is continuous in Ω. Consider

$$\begin{aligned} & u_{tt} - \Delta u = f \quad \text{in } \Omega \times (0, \infty), \\ & u(\cdot, 0) = u_0, \ u_t(\cdot,) = u_1 \quad \text{in } \Omega, \\ & u = 0 \quad \text{on } \partial\Omega \times (0, \infty). \end{aligned} \tag{3.7}$$

THEOREM 3.15. *Let $u \in C^2(\Omega\times(0,\infty))\cap C^1(\bar\Omega\times[0,\infty))$ be a solution of (3.7). Then*

$$\big(\|u_t(t)\|^2_{L^2(\Omega)}+\|\nabla_x u(t)\|^2_{L^2(\Omega)}\big)^{\frac12} \le \big(\|u_1\|^2_{L^2(\Omega)}+\|\nabla_x u_0\|^2_{L^2(\Omega)}\big)^{\frac12}+\int_0^t \|f(s)\|_{L^2(\Omega)}ds,$$

and

$$\|u(t)\|_{L^2(\Omega)} \le \|u_0\|_{L^2(\Omega)} + t\big(\|u_1\|^2_{L^2(\Omega)}+\|\nabla_x u_0\|^2_{L^2(\Omega)}\big)^{\frac12}+\int_0^t (t-s)\|f(s)\|_{L^2(\Omega)}ds.$$

Hence we also have the uniqueness and continuous dependence on initial values.

PROOF. Multiply the equation by u_t and integrate over Ω for each fixed $t>0$. Note

$$u_tu_{tt}-u_t\Delta u=\frac12(u_t^2+|\nabla u|^2)_t-\sum_{i=1}^n(u_tu_{x_i})_{x_i}.$$

Then

$$\frac12\frac{d}{dt}\int_\Omega\big(|u_t(t)|^2+|\nabla_x u(t)|^2\big)-\int_{\partial\Omega}u_t(t)\frac{\partial u}{\partial n}(t)=\int_\Omega f(t)u_t(t).$$

Note $u_t=0$ on $\partial\Omega\times(0,\infty)$ since $u=0$ on $\partial\Omega\times(0,\infty)$. Then

$$\frac12\frac{d}{dt}\int_\Omega\big(|u_t(t)|^2+|\nabla_x u(t)|^2\big)=\int_\Omega f(t)u_t(t).$$

Define the energy by

$$E(t)=\int_\Omega\big(|u_t(t)|^2+|\nabla_x u(t)|^2\big).$$

If $f\equiv 0$, then

$$\frac{d}{dt}E(t)=0.$$

Hence

$$E(t)=E(0)=\frac12\int_\Omega\big(|u_1|^2+|\nabla_x u_0|^2\big).$$

This is the conservation of energy. In general,

$$E(t)=E(0)+2\int_0^t\int_\Omega fu_t.$$

To get an estimate on the energy $E(t)$, set

$$J(t)=\big(E(t)\big)^{\frac12}.$$

Then

$$\big(J(t)\big)^2=\big(J(0)\big)^2+2\int_0^t\int_\Omega fu_t.$$

Differentiating with respect to t, we get

$$2J(t)J'(t)=2\int_\Omega f(t)u_t(t)\le 2\|f(t)\|_{L^2(\Omega)}\|u_t(t)\|_{L^2(\Omega)}\le 2J(t)\|f(t)\|_{L^2(\Omega)}.$$

Hence

$$J'(t)\le\|f(t)\|_{L^2(\Omega)}.$$

Integrating from 0 to t, we obtain

$$J(t)\le J(0)+\int_0^t\|f(s)\|_{L^2(\Omega)}ds.$$

This is the desired estimate on the energy. Next, to estimate the L^2-norm of u, we set

$$F(t) = \|u(t)\|_{L^2(\Omega)},$$

i.e.,

$$\big(F(t)\big)^2 = \int_\Omega |u(t)|^2.$$

A simple differentiation yields

$$2F(t)F'(t) = 2\int_\Omega u(t)u_t(t) \le 2\|u(t)\|_{L^2(\Omega)}\|u_t(t)\|_{L^2(\Omega)}.$$

Hence

$$F'(t) \le \|u_t\|_{L^2(\Omega)} \le J(0) + \int_0^t \|f(s)\|_{L^2(\Omega)}ds.$$

Integrating from 0 to t, we get

$$\|u(t)\|_{L^2(\Omega)} \le \|u_0\|_{L^2(\Omega)} + tJ(0) + \int_0^t\int_0^{t'} \|f(s)\|_{L^2(\Omega)}dsdt'.$$

This implies the desired estimate on u itself easily. $\square$

There are other forms of estimates on energies. By squaring the first estimate in Theorem 3.15, we obtain

$$\int_\Omega \big(|u_t(t)|^2 + |\nabla_x u(t)|^2\big) \le 2\int_\Omega \big(|u_1|^2 + |\nabla_x u_0|^2\big) + 2t\int_0^t\int_\Omega |f|^2.$$

Integrating from 0 to t, we get

$$\int_0^t\int_\Omega \big(|u_t|^2 + |\nabla_x u|^2\big) \le 2t\int_\Omega \big(|u_1|^2 + |\nabla_x u_0|^2\big) + t^2\int_0^t\int_\Omega |f|^2.$$

To end this section, we briefly review methods used in deriving estimates in Theorems 3.13-3.15. In the proof of Theorems 3.13-3.14, we multiply the Laplace equation and the heat equation by u and integrate over Ω. While in the proof of Theorem 3.15, we multiply the wave equation by u_t and integrate over Ω. These strategies also work for general elliptic equations, parabolic equations and hyperbolic equations.

3.3. Separation of Variables

In this section, we use the separation of variables to solve initial/boundary-value problems of the heat equation and the wave equation in $\mathbb{R}\times\mathbb{R}_+$.

We first discuss initial/boundary-value problems for the 1-dimensional heat equation of the following form

$$\begin{aligned} &u_t - u_{xx} = 0 \quad \text{for any } x\in(0,\pi) \text{ and any } t>0,\\ &u(x,0) = u_0(x) \quad \text{for any } x\in(0,\pi),\\ &u(0,t) = u(\pi,t) = 0 \quad \text{for any } t>0. \end{aligned} \tag{3.8}$$

Physically, u represents the temperature in an insulated rod with ends kept at 0 temperature.

We now use separation of variables to find solutions of

$$\begin{aligned} &u_t - u_{xx} = 0 \quad \text{for any } x \in (0,\pi) \text{ and any } t > 0, \\ &u(0,t) = u(\pi,t) = 0 \quad \text{for any } t > 0. \end{aligned}$$

Set

$$u(x,t) = a(t)w(x).$$

Then

$$a'(t)w(x) - a(t)w''(x) = 0,$$

or

$$\frac{a'(t)}{a(t)} - \frac{w''(x)}{w(x)} = 0.$$

The first term is a function of t and the second term is a function of x. Since their difference is zero, they should be equal to a constant. We denote this common constant by $-\lambda$. Then

$$a'(t) + \lambda a(t) = 0 \quad \text{for any } t > 0,$$

and

$$\begin{aligned} &w''(x) + \lambda w(x) = 0 \quad \text{for any } x \in (0,\pi), \\ &w(0) = w(\pi) = 0. \end{aligned}$$

The latter is the homogeneous eigenvalue problem of $-\dfrac{d^2}{dx^2}$ in $(0,\pi)$. It is well known that $\lambda_k = k^2$ and $w_k(x) = \sin kx$ are a sequence of solutions and $\{w_k\}$ forms an orthogonal basis for Fourier series in $L^2(0,\pi)$. For any $v \in L^2(0,\pi)$, we have

$$v(x) = \sum_{k=1}^{\infty} a_k \sin kx,$$

where

$$a_k = \frac{2}{\pi}\int_0^{\pi} v(x)\sin kx dx.$$

The convergence for the series of v is in the L^2-sense. By the completeness of $\{\sin kx\}$ in $L^2(0,\pi)$, we have

$$\|v\|_{L^2(0,\pi)} = \Big(\frac{\pi}{2}\sum_{k=1}^{\infty} a_k^2\Big)^{\frac{1}{2}}.$$

Back to $a(t)$, we have for each $\lambda = k^2$

$$a(t) = ae^{-k^2 t},$$

where a is a constant. Now we set for each integer $k \geq 1$

$$u_k(x,t) = e^{-k^2 t}\sin kx \quad \text{for any } (x,t) \in (0,\pi)\times(0,\infty).$$

Then u_k satisfies the heat equation and the boundary value in (3.8). In order to get a solution satisfying the equation, the boundary value and the initial value in (3.8), we consider an infinite linear combination of u_k and choose coefficients appropriately.

Now we are ready to solve the initial/boundary-value problem (3.8). We treat t as a parameter and expand $u(x,t)$ as a function of $x \in (0,\pi)$ in its Fourier series with respect to $\{\sin kx\}$. Formally, we write

$$u(x,t) = \sum_{k=1}^{\infty} a_k(t) \sin kx.$$

Then

$$u_t - u_{xx} = \sum_{k=1}^{\infty} \big(a_k'(t) + k^2 a_k(t)\big) \sin kx.$$

With

$$a_k'(t) + k^2 a_k(t) = 0 \quad \text{for any } k = 1, 2, \cdots,$$

we have

$$a_k(t) = a_k(0) e^{-k^2 t} \quad \text{for any } k = 1, 2, \cdots,$$

and hence

$$u(x,t) = \sum_{k=1}^{\infty} a_k(0) e^{-k^2 t} \sin kx.$$

Let

$$u_0(x) = \sum_{k=1}^{\infty} u_{0k} \sin kx,$$

where

$$u_{0k} = \frac{2}{\pi} \int_0^{\pi} u_0(x) \sin kx dx.$$

Then $a_k(0) = u_{0k}$ and the solution u is formally given by

$$u(x,t) = \sum_{k=1}^{\infty} u_{0k} e^{-k^2 t} \sin kx. \tag{3.9}$$

Next we prove that u in (3.9) indeed solves (3.8).

THEOREM 3.16. *If $u_0 \in L^2(0,\pi)$, then u given by (3.9) is smooth in $[0,\pi] \times (0,\infty)$ and satisfies*

$$\begin{aligned} &u_t - u_{xx} = 0 \quad \text{for any } x \in (0,\pi) \text{ and any } t > 0, \\ &u(0,t) = u(\pi,t) = 0 \quad \text{for any } t > 0, \end{aligned}$$

and

$$\lim_{t \to 0+} \|u(\cdot,t) - u_0\|_{L^2(0,\pi)} = 0.$$

PROOF. By $u_0 \in L^2(0,\pi)$, we have

$$\|u_0\|_{L^2(0,\pi)}^2 = \frac{\pi}{2} \sum_{k=1}^{\infty} |u_{0k}|^2 < \infty.$$

For any nonnegative integers i and j, any $x \in [0,\pi]$ and $t \in (0,\infty)$, we have

$$|\partial_x^i \partial_t^j u(x,t)| \le \sum_{k=1}^{\infty} |u_{0k}| \frac{k^{i+2j}}{e^{k^2 t}}.$$

Hence the Cauchy inequality implies

$$|\partial_x^i \partial_t^j u(x,t)| \le \sum_{k=1}^{\infty} \big(|u_{0k}|^2\big)^{\frac{1}{2}} \big(\sum_{k=1}^{\infty} \frac{k^{2i+4j}}{e^{2k^2 t}}\big)^{\frac{1}{2}} \le C_{i,j,t} \|u_0\|_{L^2(0,\pi)},$$

where $C_{i,j,t}$ is a positive constant depending only on i, j and t. This shows the absolute convergence of series for $\partial_x^i \partial_t^j u(x,t)$, for any $x \in [0,\pi], t > 0$ and any nonnegative integers i and j. Hence u is smooth in $[0,\pi] \times (0,\infty)$. The proof for the convergence as $t \to 0+$ is left as an exercise. $\square$

By examining the proof, we have the following estimate. For any integer $m \ge 0$ and any $t_0 > 0$

$$\|u\|_{C^m([0,\pi]\times(t_0,\infty))} \le C \|u_0\|_{L^2(0,\pi)},$$

where C is a positive constant depending only on m and t_0. This estimate controls the C^m-norm of u in $(0,\pi) \times (t_0,\infty)$ for any $t_0 > 0$ in terms of the L^2-norm of u_0 on $(0,\pi)$. It is referred to as an interior estimate (with respect to t). We note that u becomes smooth immediately after $t = 0$ even if the initial value u_0 is only L^2.

Naturally, we ask whether u in Theorem 3.16 is continuous up to $t = 0$, or generally whether u is smooth up to $t = 0$. In order to have continuity of u up to $t = 0$, first we should require $u_0 \in C[0,\pi]$. Next, u_0 should satisfy a compatibility condition. By comparing the initial value with the homogeneous boundary value at the corner, we have

$$u_0(0) = 0, \quad u_0(\pi) = 0.$$

In general, in order to have smoothness of u up to $t = 0$, we assume $u_0 \in C^\infty[0,\pi]$. Then from the heat equation, we have $u_{xx}(0,0) = 0$ and $u_{xx}(\pi,0) = 0$. Hence

$$u_0''(0) = 0, \quad u_0''(\pi) = 0.$$

Continuing this process, we have a necessary condition

$$u_0^{(2\ell)}(0) = 0, \quad u_0^{(2\ell)}(\pi) = 0 \quad \text{for any } \ell = 0, 1, \cdots. \tag{3.10}$$

Now, we prove this is also a sufficient condition.

THEOREM 3.17. *If $u_0 \in C^\infty[0,\pi]$ and (3.10) holds, then u given by (3.9) is smooth in $[0,\pi] \times [0,\infty)$ with $u(\cdot,0) = u_0$.*

PROOF. In order to prove the smoothness up to $t = 0$, we need to bound $\partial_x^i \partial_t^j u(x,t)$ independent of t for t small, for any nonnegative integers i and j. To this end, we first improve estimates of u_{0k}. With (3.10), we have by simple integrations by parts

$$u_{0k} = \frac{2}{\pi} \int_0^\pi u_0(x) \sin kx dx = \frac{2}{\pi} \int_0^\pi u_0'(x) \frac{\cos kx}{k} dx = -\frac{2}{\pi} \int_0^\pi u_0''(x) \frac{\sin kx}{k^2} dx.$$

We note that values at end points are not present since $u_0(0) = u_0(\pi) = 0$ in the first integration by parts and $\sin kx = 0$ at $x = 0$ and $x = \pi$ in the second integration by parts. Hence for any $m \ge 1$, we continue this process with the help of (3.10) for $\ell = 0, \cdots, [(m-1)/2]$ and obtain

$$u_{0k} = (-1)^{\frac{m-1}{2}} \frac{2}{\pi} \int_0^\pi u_0^{(m)}(x) \frac{\cos kx}{k^m} dx \quad \text{if } m \text{ is odd},$$

and

$$u_{0k} = (-1)^{\frac{m}{2}} \frac{2}{\pi} \int_0^\pi u_0^{(m)}(x) \frac{\sin kx}{k^m} dx \quad \text{if } m \text{ is even.}$$

By the Cauchy inequality, we get for any $m \geq 0$

$$|u_{0k}| \leq \sqrt{\frac{2}{\pi}} \|u_0^{(m)}\|_{L^2(0,\pi)} \frac{1}{k^m} \quad \text{for any } k = 1, 2, \cdots,$$

where we used

$$\int_0^\pi \sin^2 kx dx = \int_0^\pi \cos^2 kx dx = \frac{\pi}{2}.$$

Now for any $x \in [0, \pi]$ and $t > 0$

$$|\partial_x^i \partial_t^j (x,t)| \leq \sum_{j=1}^\infty |u_{0k}| k^{i+2j} \leq \sqrt{\frac{2}{\pi}} \|u_0^{(m)}\|_{L^2(0,\pi)} \sum_{k=1}^\infty \frac{k^{i+2j}}{k^m} < \infty,$$

which is independent of t, if we take $m = i + 2j + 2$. This implies that $\partial_x^i \partial_t^j (x,t)$ is continuous in $[0, \pi] \times [0, \infty)$. □

If we are only interested in the continuity of u up to $t = 0$, we have the following result.

COROLLARY 3.18. *If $u_0 \in C^1[0, \pi]$ with $u_0'' \in L^2(0, \pi)$ and $u_0(0) = u_0(\pi) = 0$, then u given by (3.9) is smooth in $[0, \pi] \times (0, \infty)$ and continuous in $[0, \pi] \times [0, \infty)$ and satisfies (3.8).*

PROOF. It follows from Theorem 3.16 that u is smooth in $[0, \pi] \times (0, \infty)$ and satisfies the heat equation and the homogeneous boundary condition in (3.8). The continuity of u up to $t = 0$ follows from the proof of Theorem 3.17 with $i = j = 0$ and $m = 2$. □

The regularity assumption on u_0 in Corollary 3.18 does not seem to be optimal. It is natural to ask whether it suffices to assume $u_0 \in C[0, \pi]$. To answer this question, we should analyze whether the series in (3.9) converges uniformly in t since we need to take limit $t \to 0$. We will not pursue this issue here.

Now we provide another expression of u in (3.9). With explicit expressions of u_{0k} in terms of u_0, we can write

$$u(x,t) = \int_0^\pi G(x, y; t) u_0(y) dy, \tag{3.11}$$

where

$$G(x, y; t) = \frac{2}{\pi} \sum_{k=1}^\infty e^{-k^2 t} \sin kx \sin ky \quad \text{for any } x, y \in [0, \pi] \text{ and } t > 0.$$

The function G is called the Green's function associated with the initial-boundary value problem (3.8). For each fixed t, the series for G is uniformly convergent for $x, y \in [0, \pi]$. In fact, this uniform convergence justifies the exchange of summation and integration in obtaining (3.11). The Green's function G satisfies the following properties:

(G1) Symmetry: $G(x, y; t) = G(y, x; t)$.

(G2) Smoothness: $G(x, y; t)$ is smooth in $x, y \in [0, \pi]$ and $t > 0$.

(G3) Solution of the heat equation: $G_t - G_{xx} = 0$.

(G4) Homogeneous boundary values: $G(0, y; t) = G(\pi, y; t) = 0$.

These properties follow easily from the explicit expression for G. They imply that u in (3.11) is a smooth function for $t > 0$ and satisfies the heat equation with homogeneous boundary values. We can also prove with help of the explicit expression of G that u in (3.11) is continuous up to $t = 0$ and satisfies $u(\cdot, 0) = u_0$ under appropriate assumptions on u_0. We point out that G can also be expressed in terms of the fundamental solution of the heat equation. See Chapter 5 for discussions of the fundamental solution.

Next, we discuss initial/boundary-value problems for the 1-dimensional wave equation of the following form

$$\begin{aligned} &u_{tt} - u_{xx} = 0 \quad \text{for any } x \in (0, \pi) \text{ and any } t > 0, \\ &u(x, 0) = u_0(x),\ u_t(x, 0) = u_1(x) \quad \text{for any } x \in (0, \pi), \\ &u(0, t) = u(\pi, t) = 0 \quad \text{for any } t > 0. \end{aligned} \tag{3.12}$$

We proceed as for the heat equation. Set

$$u(x, t) = \sum_{k=1}^{\infty} T_k(t) \sin kx.$$

Then

$$u_{tt} - u_{xx} = \sum_{k=1}^{\infty} \big(T_k''(t) + k^2 T_k(t)\big) \sin kx.$$

With

$$T_k''(t) + k^2 T_k(t) = 0 \quad \text{for any } k = 1, 2, \cdots,$$

we have

$$T_k(t) = a_k \sin kt + b_k \cos kt \quad \text{for any } k = 1, 2, \cdots.$$

Therefore,

$$u(x, t) = \sum_{k=1}^{\infty} (a_k \sin kt + b_k \cos kt) \sin kx. \tag{3.13}$$

If we write

$$u_0(x) = \sum_{k=1}^{\infty} u_{0k} \sin kx, \quad u_1(x) = \sum_{k=1}^{\infty} u_{1k} \sin kx,$$

we can determine a_k, b_k from u_0 and u_1 by

$$\begin{aligned} a_k &= \frac{u_{1k}}{k} = \frac{2}{k\pi} \int_0^{\pi} u_1(x) \sin kx dx, \\ b_k &= u_{0k} = \frac{2}{\pi} \int_0^{\pi} u_0(x) \sin kx dx. \end{aligned}$$

Unlike the case in the heat equation, in order to get regularity now, we need to impose similar regularity conditions on initial values. Preceding as for the heat equation, we note that if u is a C^2-solution, then

$$\begin{aligned} &u_0(0) = 0, \quad u_1(0) = 0, \quad u_0''(0) = 0, \\ &u_0(\pi) = 0, \quad u_1(\pi) = 0, \quad u_0''(\pi) = 0. \end{aligned} \tag{3.14}$$

THEOREM 3.19. *If* $u_0 \in C^2[0,\pi], u_0''' \in L^2(0,\pi), u_1 \in C^1[0,\pi], u_1'' \in L^2(0,\pi)$ *and* u_0, u_1 *satisfy (3.14), then* u *given by (3.13) is* C^2 *in* $[0,\pi]\times[0,\infty)$ *and is a solution of (3.12).*

PROOF. We first improve estimates for a_k and b_k. By (3.14) and integration by parts, we have

$$a_k = \frac{2}{k\pi}\int_0^\pi u_1(x)\sin kx dx = -\frac{2}{\pi k^3}\int_0^\pi u_1''(x)\sin kx dx,$$
$$b_k = \frac{2}{\pi}\int_0^\pi u_0(x)\sin kx dx = -\frac{2}{\pi k^3}\int_0^\pi u_0'''(x)\cos kx dx.$$

In other words, $\{k^3 a_k\}$ is the Fourier coefficients of $-u_1''(x)$ with respect to $\{\sin kx\}$ and $\{k^3 b_k\}$ is the Fourier coefficients of $-u_0'''(x)$ with respect to $\{1, \cos kx\}$. Hence

$$\sum_{k=1}^\infty (k^6 a_k^2 + k^6 b_k^2) < \infty.$$

Note for any integers i, j with $0 \le i+j \le 2$ and any $x \in [0,\pi]$ and $t > 0$

$$|\partial_x^i \partial_t^j u(x,t)| \le \sum_{k=1}^\infty k^{i+j}(|a_k| + |b_k|),$$

and hence

$$|\partial_x^i \partial_t^j u(x,t)| \le \big(\sum_{k=1}^\infty k^{2(i+j+1)}(|a_k|+|b_k|)^2\big)^{\frac12}\cdot\big(\sum_{k=1}^\infty \frac{1}{k^2}\big)^{\frac12} < \infty.$$

Therefore, u is C^2 in $[0,\pi]\times[0,\infty)$. Then u satisfies the wave equation and the initial condition by a simple differentiation term by term. □

By examining the proof, we have

$$\|u\|_{C^2([0,\pi]\times(0,\infty))} \le C\big(\sum_{i=0}^3 \|u_0^{(i)}\|_{L^2(0,\pi)} + \sum_{i=0}^2 \|u_1^{(i)}\|_{L^2(0,\pi)}\big),$$

where C is a positive constant independent of u.

In fact, in order to get a C^2-solution of (3.12), it suffices to assume $u_0 \in C^2[0,\pi]$ and $u_1 \in C^1[0,\pi]$. We will prove this for a more general initial/boundary-value problem of the wave equation in Section 6.1. See Theorem 6.3.

Now, we compare the regularity results in Theorems 3.16, 3.17 and 3.19. For initial/boundary-value problems of the heat equation in Theorem 3.16, solutions become smooth immediately after $t = 0$, even for L^2-initial values. This is the so-called *interior smoothness* (with respect to time). We also proved in Theorem 3.17 that solutions are smooth up to $t = 0$ if initial values are also smooth with a compatibility condition. This property is called the *global smoothness.* However, solutions of initial/boundary-value problems of the wave equation exhibit a different property. We proved in Theorem 3.19 that solutions have a similar degree of regularity as initial values. In fact, this is a general result. Solutions of the wave equation do not have better regularity than initial values, and in some cases they are less regular than initial values. We will see in Chapter 6 how solutions of the wave equation depend on initial values.

To end this section, we point out that methods employed in this section to solve initial/boundary-value problems for the one-dimensional heat equation and wave equation can actually be generalized to higher dimensions. We use the heat equation as an illustration. Let Ω be a bounded smooth domain in $\mathbb{R}^n$. We consider

$$\begin{aligned} u_t - \Delta u &= 0 \quad \text{in } \Omega \times (0,\infty), \\ u(\cdot, 0) &= u_0 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega \times (0,\infty), \end{aligned} \tag{3.15}$$

for a continuous function u_0 in $\bar{\Omega}$. To solve (3.15) by separation of variables, we need solutions of the eigenvalue problem of $-\Delta$ in Ω with homogeneous boundary values, i.e.,

$$\begin{aligned} \Delta\varphi + \lambda\varphi &= 0 \quad \text{in } \Omega, \\ \varphi &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{3.16}$$

Of course, this is much harder to solve than its one-dimensional counterpart. Nevertheless, a similar result still holds. In fact, solutions of (3.16) are given by a sequence (λ_k, φ_k) where λ_k is a sequence of nondecreasing positive numbers such that $\lambda_k \to \infty$ as $k \to \infty$ and φ_k is a sequence of smooth functions in $\bar{\Omega}$ which forms a basis in $L^2(\Omega)$. Then we can use a similar method to find a solution of (3.15) of the form

$$u(x,t) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t} \varphi_k(x) \quad \text{for any } (x,t) \in \Omega \times (0,\infty).$$

We should remark that solving (3.16) is a complicated process. We need to work in Sobolev spaces, spaces of functions with L^2-integrable derivatives. The process is similar as that indicated at the end of Chapter 2. First, we prove the existence of weak solutions of (3.16) and then improve regularity of solutions. A brief discussion can be found in Section 4.6.

Exercises

(1) Classify the following second-order PDEs:

(a) $\displaystyle\sum_{i=1}^{n} u_{x_i x_i} + \sum_{1 \le i < j \le n} u_{x_i x_j} = 0.$

(b) $\displaystyle\sum_{1 \le i < j \le n} u_{x_i x_j} = 0.$

(2) Discuss the uniqueness of the following problems by energy methods:

(a) $\begin{cases} \Delta u - u^3 = f & \text{in } \Omega, \\ u = \varphi & \text{on } \partial\Omega. \end{cases}$

(b) $\begin{cases} \Delta u - u \int_\Omega u^2(y)dy = f & \text{in } \Omega, \\ u = \varphi & \text{on } \partial\Omega. \end{cases}$

(3) Let Ω be a bounded C^1-domain in $\mathbb{R}^n$ and u be a C^2-function in $\bar{\Omega}\times[0,T]$ satisfying

$$\begin{aligned} u_t-\Delta u=&f \quad \text{in } \Omega\times(0,\infty),\\ u(\cdot,0)=&u_0 \quad \text{in } \Omega,\\ u=&0 \quad \text{on } \partial\Omega\times(0,\infty). \end{aligned}$$

Prove

$$\sup_{0\le t\le T}\int_\Omega |\nabla u(\cdot,t)|^2+\int_0^T\int_\Omega u_t^2\le C\Big(\int_\Omega|\nabla u_0|^2+\int_0^T\int_\Omega f^2\Big),$$

where C is a positive constant independent of u, u_0 and f.

(4) Prove the convergence part in Theorem 3.16.

(5) For $u_0\in L^2(0,\pi)$, let u be given in (3.9). Then for any nonnegative integers i and j

$$\sup_{[0,\pi]}|\partial_x^i\partial_t^j u(\cdot,t)|\to 0 \quad \text{as } t\to\infty.$$

(6) Let G be defined in (3.11). Prove

$$|G(x,y;t)|\le\frac{1}{\sqrt{\pi t}} \quad \text{for any } x,y\in[0,\pi] \text{ and } t>0.$$

(7) Solve the following problem by separation of variables

$$\begin{aligned} &u_t-u_{xx}=0 \quad \text{for any } 0<x<\pi, t>0,\\ &u(x,0)=u_0(x) \quad \text{for any } 0<x<\pi,\\ &u_x(0,t)=u_x(\pi,t)=0 \quad \text{for any } t>0. \end{aligned}$$

(8) For any $u_0\in L^2(0,\pi)$ and $f\in L^2((0,\pi)\times(0,\infty))$, find a formal explicit expression of a solution of the following problem

$$\begin{aligned} &u_t-u_{xx}=f \quad \text{for any } x\in(0,\pi) \text{ and any } t>0,\\ &u(x,0)=u_0(x) \quad \text{for any } x\in(0,\pi),\\ &u(0,t)=u(\pi,t)=0 \quad \text{for any } t>0. \end{aligned}$$

(9) For any $u_0,u_1\in L^2(0,\pi)$ and $f\in L^2((0,\pi)\times(0,\infty))$, find a formal explicit expression of a solution of the following problem

$$\begin{aligned} &u_{tt}-u_{xx}=f \quad \text{for any } x\in(0,\pi) \text{ and any } t>0,\\ &u(x,0)=u_0(x),\ u_t(x,0)=u_1(x) \quad \text{for any } x\in(0,\pi),\\ &u(0,t)=u(\pi,t)=0 \quad \text{for any } t>0. \end{aligned}$$

(10) Let Ω be a bounded C^1-domain in $\mathbb{R}^n$ and u be C^2 in x and C^1 in t in $\bar{\Omega}\times[0,T]$ for a constant $T>0$. Suppose u satisfies

$$\begin{aligned} u_t-\Delta u&=0 \quad \text{in } \Omega\times(0,T),\\ u(\cdot,T)&=0 \quad \text{in } \Omega,\\ u&=0 \quad \text{on } \partial\Omega\times(0,T). \end{aligned}$$

Then $u = 0$ in $\Omega \times (0, T)$.
Hint: The function $J(t) = \log \int_\Omega u^2(\cdot, t)$ is a decreasing convex function.

CHAPTER 4

Laplace Equations

The Laplace equation $\Delta u = 0$ is probably the most important PDE with the widest range of applications. In Section 4.1, we solve Dirichlet problems for the Laplace equation in balls and derive Poisson integral formula. In the next three sections, we study harmonic functions, solutions of the Laplace equation, by three different methods: fundamental solutions, mean-value properties and the maximum principle. These three sections are relatively independent of each other. In Section 4.5, we study the Poisson equation $\Delta u = f$ and derive Schauder estimates for its solutions. Then in Section 4.6, we briefly discuss weak solutions of the Poisson equation.

4.1. Dirichlet Problems

The Laplace operator Δ is defined on a C^2-function u in a domain in $\mathbb{R}^n$ by

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

The equation $\Delta u = 0$ is called the Laplace equation and its C^2-solutions are called harmonic functions.

One of important properties of the Laplace equation is its spherical symmetry. The equation is preserved by rotations about some point $a \in \mathbb{R}^n$. Hence, it is plausible that there exist special solutions that are invariant under rotations about a. We begin this section by seeking a harmonic function u in $\mathbb{R}^n$ which depends only on $r = |x - a|$ for some fixed $a \in \mathbb{R}^n$. By setting $v(r) = u(x)$, we have

$$\Delta u = v'' + \frac{n-1}{r} v' = 0,$$

and hence

$$v(r) = \begin{cases} c_1 + c_2 \log r & \text{for } n = 2, \\ c_3 + c_4 r^{2-n} & \text{for } n \geq 3, \end{cases}$$

where c_i are constants for $i = 1, 2, 3, 4$. We are interested in solutions with a singularity such that

$$\int_{\partial B_r} \frac{\partial v}{\partial r} d\sigma = 1 \quad \text{for any } r > 0.$$

Hence we set for any fixed $a \in \mathbb{R}^n$

$$\Gamma(a, x) = \begin{cases} \frac{1}{2\pi} \log |a - x| & \text{for } n = 2, \\ \frac{1}{\omega_n (2-n)} |a - x|^{2-n} & \text{for } n \geq 3, \end{cases}$$

where ω_n is the surface area of the unit sphere in $\mathbb{R}^n$. In summary, for any fixed $a \in \mathbb{R}^n$, $\Gamma(a, \cdot)$ is harmonic in $\mathbb{R}^n \setminus \{a\}$, i.e.,

$$\Delta_x \Gamma(a, x) = 0 \quad \text{for any } x \neq a,$$

and has a singularity at $x = a$. Moreover,

$$\int_{\partial B_r(a)} \frac{\partial \Gamma}{\partial n_x}(a, x)\, dS_x = 1 \qquad \text{for any } r > 0.$$

The function Γ is called the fundamental solution of the Laplace operator.

Now we prove the Green's identity, which plays an important role in discussions of harmonic functions. In the next result, we replace a, x by x, y and write $\Gamma(x, y)$ instead of $\Gamma(a, x)$.

THEOREM 4.1. *Suppose Ω is a bounded C^1-domain in $\mathbb{R}^n$ and that $u \in C^1(\bar{\Omega}) \cap C^2(\Omega)$. Then for any $x \in \Omega$*

$$u(x) = \int_{\Omega} \Gamma(x, y) \Delta u(y) dy - \int_{\partial\Omega} \big(\Gamma(x, y) \frac{\partial u}{\partial n_y}(y) - u(y) \frac{\partial \Gamma}{\partial n_y}(x, y)\big)\, dS_y.$$

REMARK 4.2. (i) For any $x \in \Omega$, $\Gamma(x, \cdot)$ is integrable in Ω although it has a singularity.

(ii) For $x \notin \bar{\Omega}$, the expression in the right hand side is zero. This follows from the Green's formula in Ω and $\Delta_y \Gamma(x, y) = 0$ if $x \notin \bar{\Omega}$.

(iii) By letting $u = 1$, we have

$$\int_{\partial\Omega} \frac{\partial \Gamma}{\partial n_y}(x, y) dS_y = 1 \quad \text{for any } x \in \Omega.$$

PROOF. By applying the Green's formula to u and $\Gamma(x, \cdot)$ in $\Omega \setminus B_r(x)$ for small $r > 0$, we get

$$\int_{\Omega \setminus B_r(x)} (\Gamma \Delta u - u \Delta \Gamma) dy = \int_{\partial\Omega} (\Gamma \frac{\partial u}{\partial n} - u \frac{\partial \Gamma}{\partial n}) dS_y - \int_{\partial B_r(x)} (\Gamma \frac{\partial u}{\partial n} - u \frac{\partial \Gamma}{\partial n}) dS_y.$$

Noting $\Delta \Gamma = 0$ in $\Omega \setminus B_r(x)$, we have

$$\int_{\Omega} \Gamma \Delta u dy = \int_{\partial\Omega} (\Gamma \frac{\partial u}{\partial n} - u \frac{\partial \Gamma}{\partial n}) dS_y - \lim_{r \to 0} \int_{\partial B_r(x)} (\Gamma \frac{\partial u}{\partial n} - u \frac{\partial \Gamma}{\partial n}) dS_y.$$

For $n \geq 3$, we get by the definition of Γ

$$\begin{aligned} \Big| \int_{\partial B_r(x)} \Gamma \frac{\partial u}{\partial n} dS_y \Big| &= \Big| \frac{1}{(2 - n)\omega_n} r^{2-n} \int_{\partial B_r(x)} \frac{\partial u}{\partial n} dS_y \Big| \\ &\leq \frac{r}{n - 2} \sup_{\partial B_r(x)} |\nabla u| \to 0 \quad \text{as } r \to 0, \end{aligned}$$

and

$$\int_{\partial B_r(x)} u \frac{\partial \Gamma}{\partial n}\, dS_y = \frac{1}{\omega_n r^{n-1}} \int_{\partial B_r(x)} u dS_y \to u(x) \quad \text{as } r \to 0.$$

For $n = 2$, we get the same conclusion similarly. $\square$

Now we use Theorem 4.1 to discuss the Dirichlet boundary-value problem

$$\begin{aligned} \Delta u &= f \quad \text{in } \Omega, \\ u &= \varphi \quad \text{on } \partial\Omega, \end{aligned} \tag{4.1}$$

for some $f \in C(\bar{\Omega})$ and $\varphi \in C(\partial\Omega)$. Lemma 3.11 asserts the uniqueness of a solution in $C^2(\Omega) \cap C^1(\bar{\Omega})$. Another method to obtain this result is by the maximum principle to be discussed later in this chapter. If u solves (4.1), then by Theorem 4.1 u can be expressed in terms of f and φ, with one *unknown term* $\frac{\partial u}{\partial n}$ on $\partial\Omega$. We intend to eliminate this term by adjusting Γ. We need to emphasize that we cannot prescribe $\frac{\partial u}{\partial n}$ on $\partial\Omega$ together with u on $\partial\Omega$.

For any fixed $x \in \Omega$, we consider a function $\Phi(x, \cdot) \in C^2(\Omega) \cap C^1(\bar{\Omega})$ with $\Delta_y \Phi(x, y) = 0$ in Ω. Then by the Green's formula,

$$0 = \int_\Omega \Phi(x,y)\Delta u(y)dy - \int_{\partial\Omega} \big(\Phi(x,y)\frac{\partial u}{\partial n_y}(y) - u(y)\frac{\partial \Phi}{\partial n_y}(x,y)\big)dS_y.$$

By adding this to the Green's identity in Theorem 4.1, we obtain for any $x \in \Omega$

$$u(x) = \int_\Omega \gamma(x,y)\Delta u(y)dy - \int_{\partial\Omega} \big(\gamma(x,y)\frac{\partial u}{\partial n_y}(y) - u(y)\frac{\partial \gamma}{\partial n_y}(x,y)\big)dS_y,$$

where

$$\gamma(x,y) = \Gamma(x,y) + \Phi(x,y).$$

Now we choose Φ appropriately so that $\gamma(x, \cdot) = 0$ on $\partial\Omega$. This leads to the important concept of Green's functions.

For each fixed $x \in \Omega$, we consider $\Phi(x, \cdot) \in C^1(\bar{\Omega}) \cap C^2(\Omega)$ such that

$$\begin{aligned} \Delta_y \Phi(x,y) &= 0 \quad \text{for} \quad y \in \Omega \\ \Phi(x,y) &= -\Gamma(x,y) \quad \text{for} \quad y \in \partial\Omega. \end{aligned}$$

By Corollary 4.7, $\Phi(x, \cdot)$ is smooth in Ω for each fixed x if it exists. Denote by $G(x, y)$ the resulting $\gamma(x, y)$, which is called the Green's function. If such a G exists, then the solution u of the Dirichlet problem (4.1) can be expressed by

$$u(x) = \int_\Omega G(x,y)f(y)dy + \int_{\partial\Omega} \varphi(y)\frac{\partial G}{\partial n_y}(x,y)dS_y. \tag{4.2}$$

Note that the Green's function $G(x, y)$ is defined as a function of $y \in \bar{\Omega}$ for each fixed $x \in \Omega$. Now we discuss properties of G as a function of x and y. We should point out that the Green's function is unique. This follows from Lemma 3.11 or Corollary 4.16 to be established, since a difference of any two Green's functions is harmonic with a zero boundary value.

Lemma 4.3. *The Green's function $G(x, y)$ is symmetric in $\Omega \times \Omega$, i.e., $G(x, y) = G(y, x)$ for $x \neq y \in \Omega$.*

Proof. For any $x_1, x_2 \in \Omega$ with $x_1 \neq x_2$, take $r > 0$ small such that $B_r(x_1) \cap B_r(x_2) = \phi$. Set $G_i(y) = G(x_i, y)$ and $\Gamma_i(y) = \Gamma(x_i, y)$. We apply the Green's formula in $\Omega \setminus B_r(x_1) \cup B_r(x_2)$ and get

$$\int_{\Omega\setminus B_r(x_1)\cup B_r(x_2)} \big(G_1\Delta G_2 - G_2\Delta G_1\big) = \int_{\partial\Omega} \big(G_1\frac{\partial G_2}{\partial n} - G_2\frac{\partial G_1}{\partial n}\big)$$
$$- \int_{\partial B_r(x_1)} \big(G_1\frac{\partial G_2}{\partial n} - G_2\frac{\partial G_1}{\partial n}\big) - \int_{\partial B_r(x_2)} \big(G_1\frac{\partial G_2}{\partial n} - G_2\frac{\partial G_1}{\partial n}\big).$$

Since G_i is harmonic for $y \neq x_i$, $i = 1, 2$, and vanishes on $\partial\Omega$, we have

$$\int_{\partial B_r(x_1)} \big(G_1\frac{\partial G_2}{\partial n} - G_2\frac{\partial G_1}{\partial n}\big) + \int_{\partial B_r(x_2)} \big(G_1\frac{\partial G_2}{\partial n} - G_2\frac{\partial G_1}{\partial n}\big) = 0.$$

Now we replace G_1 in the first integral by Γ_1 and replace G_2 in the second integral by Γ_2. Since $G_1 - \Gamma_1$ is harmonic in Ω and G_2 is harmonic in $\Omega \setminus B_r(x_2)$, we have

$$\int_{\partial B_r(x_1)} \big((G_1 - \Gamma_1)\frac{\partial G_2}{\partial n} - G_2\frac{\partial(G_1 - \Gamma_1)}{\partial n}\big) \to 0 \quad \text{as } r \to 0.$$

Similarly,

$$\int_{\partial B_r(x_2)} \big(G_1\frac{\partial(G_2 - \Gamma_2)}{\partial n} - (G_2 - \Gamma_2)\frac{\partial G_1}{\partial n}\big) \to 0 \quad \text{as } r \to 0.$$

Therefore, we obtain

$$\int_{\partial B_r(x_1)} \big(\Gamma_1\frac{\partial G_2}{\partial n} - G_2\frac{\partial \Gamma_1}{\partial n}\big) + \int_{\partial B_r(x_2)} \big(G_1\frac{\partial \Gamma_2}{\partial n} - \Gamma_2\frac{\partial G_1}{\partial n}\big) \to 0,$$

as $r \to 0$. On the other hand, we have

$$\int_{\partial B_r(x_1)} \Gamma_1\frac{\partial G_2}{\partial n} \to 0, \quad \int_{\partial B_r(x_2)} \Gamma_2\frac{\partial G_1}{\partial n} \to 0 \quad \text{as} \quad r \to 0,$$

and

$$\int_{\partial B_r(x_1)} G_2\frac{\partial \Gamma_1}{\partial n} \to G_2(x_1), \quad \int_{\partial B_r(x_2)} G_1\frac{\partial \Gamma_2}{\partial n} \to G_1(x_2) \quad \text{as} \quad r \to 0.$$

The two limits above can be proved similarly as the last limit in the proof of Theorem 4.1. This implies $G_2(x_1) - G_1(x_2) = 0$, or $G(x_2, x_1) = G(x_1, x_2)$. □

Finding Green's functions involves solving a Dirichlet problem for the Laplace equation. Meanwhile, Green's functions are introduced to solve Dirichlet problems. We need to break this cycle. In fact, we can construct Green's functions explicitly for special domains. The next result yields an expression for Green's functions in balls.

THEOREM 4.4. *The Green's function in the ball $B_R \subset \mathbb{R}^n$ is given by*

$$G(x,y) = \frac{1}{2\pi}\left(\log|x-y| - \log\left|\frac{R}{|x|}x - \frac{|x|}{R}y\right|\right) \quad \textit{for } n = 2,$$

and

$$G(x,y) = \frac{1}{(2-n)\omega_n}\left(|x-y|^{2-n} - \left|\frac{R}{|x|}x - \frac{|x|}{R}y\right|^{2-n}\right) \quad \textit{for } n \geq 3.$$

For $x = 0$, we have

$$G(0, y) = \frac{1}{(2-n)\omega_n}\big(|y|^{2-n} - R^{2-n}\big) \quad \text{for } n \geq 3$$

and

$$G(0, y) = \frac{1}{2\pi}\big(\log|y| - \log|R|\big) \quad \text{for } n = 2.$$

Proof. We need to adjust the fundamental solution by adding a harmonic function so that the sum vanishes on the boundary. Fix an $x \neq 0$ with $|x| < R$, and consider $X \in \mathbb{R}^n \setminus \bar{B}_R$ given by $X = \dfrac{R^2}{|x|^2}x$. In other words, X and x are reflexive of each other with respect to the sphere ∂B_R. Note that the map $x \mapsto X$ is conformal, i.e., preserves angles. For $|y| = R$, we have by the similarity of triangles

$$\frac{|x|}{R} = \frac{R}{|X|} = \frac{|y-x|}{|y-X|},$$

which implies

$$|y - x| = \frac{|x|}{R}|y - X| = \left|\frac{|x|}{R}y - \frac{R}{|x|}x\right| \quad \text{for any } y \in \partial B_R. \tag{4.3}$$

Therefore, in order to have a zero boundary value, we take for $n \geq 3$

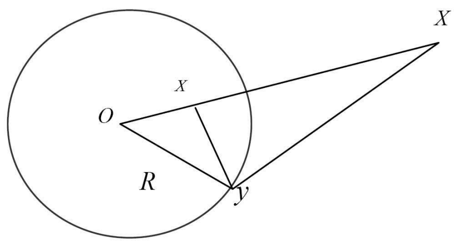

Figure 4.1. The reflection about the sphere.

$$G(x, y) = \frac{1}{(2-n)\omega_n}\left(\frac{1}{|x-y|^{n-2}} - \left(\frac{R}{|x|}\right)^{n-2}\frac{1}{|y-X|^{n-2}}\right).$$

We point out that the second term is smooth for $y \in B_R$, since $X \notin \bar{B}_R$. The case $n = 2$ is similar. $\square$

Next, we calculate normal derivatives of the Green's function on spheres.

Corollary 4.5. *Suppose G is the Green's function in B_R as in Theorem 4.4. Then*

$$\frac{\partial G}{\partial n_y}(x, y) = \frac{R^2 - |x|^2}{\omega_n R|x-y|^n} \quad \textit{for any } x \in B_R \textit{ and } y \in \partial B_R.$$

Proof. We only consider the case $n \geq 3$. With $X = R^2x/|x|^2$, we have

$$G(x, y) = \frac{1}{(2-n)\omega_n}\left(|y-x|^{2-n} - \big(\frac{R}{|x|}\big)^{n-2}|y-X|^{2-n}\right),$$

for any $x \in B_R$ and $y \in \partial B_R$. Hence we get for such x and y

$$G_{y_i}(x,y) = \frac{1}{\omega_n}\left(\frac{y_i - x_i}{|y-x|^n} - \left(\frac{R}{|x|}\right)^{n-2} \cdot \frac{y_i - X_i}{|y-X|^n}\right) = \frac{y_i}{\omega_n R^2}\frac{R^2 - |x|^2}{|x-y|^n},$$

by (4.3) in the proof of Theorem 4.4. With $n_i = \dfrac{y_i}{R}$ for $|y| = R$, we obtain

$$\frac{\partial G}{\partial n_y}(x,y) = \sum_{i=1}^n n_i G_{y_i}(x,y) = \frac{1}{\omega_n R} \cdot \frac{R^2 - |x|^2}{|x-y|^n}.$$

This yields the desired result. □

Denote by $K(x,y)$ the function in Corollary 4.5, i.e.,

$$K(x,y) = \frac{R^2 - |x|^2}{\omega_n R|x-y|^n} \quad \text{for any } x \in B_R,\ y \in \partial B_R.$$

It is called the Poisson kernel and has the following properties:

(K1) $K(x,y)$ is smooth for any $x \in B_R$ and $y \in \partial B_R$.

(K2) $K(x,y) > 0$ for any $x \in B_R$ and $y \in \partial B_R$.

(K3) For any fixed $x_0 \in \partial B_R$, $\lim_{x \to x_0, |x|<R} K(x,y) = 0$ uniformly in y for $|y - x_0| > \delta > 0$.

(K4) $\Delta_x K(x,y) = 0$ for any $x \in B_R$ and $y \in \partial B_R$.

(K5) $\int_{\partial B_R} K(x,y) dS_y = 1$ for any $x \in B_R$.

Here (K1), (K2) and (K3) follow easily from the explicit expression for K and (K4) follows easily from the definition $K(x,y) = \dfrac{\partial}{\partial n_y} G(x,y)$ and the fact that $G(x,y)$ is harmonic in x. An easy derivation of (K5) is based on (4.2). By taking a $C^2(\bar{B}_R)$ harmonic function u in (4.2), we conclude

$$u(x) = \int_{\partial B_R} K(x,y)u(y)dS_y \quad \text{for any } |x| < R.$$

This is called the Poisson integral formula. Then we have (K5) easily by taking $u \equiv 1$.

Now we are ready to solve the Laplace equation in balls with prescribed Dirichlet boundary values.

THEOREM 4.6. *For any $\varphi \in C(\partial B_R)$, the function u defined by*

$$u(x) = \int_{\partial B_R} \frac{R^2 - |x|^2}{\omega_n R|x-y|^n}\varphi(y)dS_y \quad \text{for any } x \in B_R \tag{4.4}$$

is smooth in B_R and continuous up to ∂B_R and satisfies

$$\Delta u = 0 \quad \text{in } B_R,$$
$$u = \varphi \quad \text{on } \partial B_R.$$

PROOF. By (K1) and (K4), we conclude easily that u defined by (4.4) is smooth and harmonic in B_R. We only need to prove the continuity of u up to the boundary ∂B_R. Let $x_0 \in \partial B_R$ and $x \in B_R$. By (K5), we have

$$u(x) - \varphi(x_0) = \int_{\partial B_R} K(x,y)\big(\varphi(y) - \varphi(x_0)\big)dS_y = I_1 + I_2,$$

where

$$I_1 = \int_{|y-x_0|<\delta, y\in\partial B_R} \cdots, \quad I_2 = \int_{|y-x_0|>\delta, y\in\partial B_R} \cdots,$$

for a positive constant δ to be determined. For any given $\varepsilon > 0$, we choose $\delta = \delta(\varepsilon) > 0$ so small that

$$|\varphi(y) - \varphi(x_0)| < \varepsilon \quad \text{for any } y \in \partial B_R \text{ with } |y - x_0| < \delta,$$

by the continuity of φ. Then $|I_1| \le \varepsilon$ by (K2) and (K5). Let $M = \sup_{\partial B_R} |\varphi|$. By (K3), we find a δ' such that

$$K(x, y) \le \frac{\varepsilon}{2M\omega_n R^{n-1}} \quad \text{for any } |x - x_0| < \delta', \ |y - x_0| > \delta,$$

where δ' depends on ε and $\delta = \delta(\varepsilon)$, and hence only on ε. Then $|I_2| < \varepsilon$. Hence

$$|u(x) - \varphi(x_0)| < 2\varepsilon \quad \text{for any } x \in B_R \text{ with } |x - x_0| < \delta'.$$

This shows the continuity of u at $x_0 \in \partial B_R$. $\square$

For $n = 2$, with $x = (r\cos\theta, r\sin\theta)$ and $y = (R\cos\eta, R\sin\eta)$ in (4.4), we have

$$u(r\cos\theta, r\sin\theta) = \frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2Rr\cos(\theta - \eta) + r^2}\varphi(R\cos\eta, R\sin\eta)d\eta.$$

Now we discuss some properties of the solution (4.4). First, $u(x)$ in (4.4) is smooth for $|x| < R$, even if the boundary value φ is simply continuous. In fact, this is a general fact. Any harmonic function is smooth. We will prove this result in the next section.

In (4.4), by letting $x = 0$, we have

$$u(0) = \frac{1}{\omega_n R^{n-1}}\int_{\partial B_R} u(y)dS_y.$$

This is the so-called mean-value property. The value of harmonic functions at center of spheres is equal to their average on spheres.

Moreover, by (K2) and (K5), u in (4.4) satisfies

$$\inf_{\partial B_R} \varphi \le u \le \sup_{\partial B_R} \varphi \quad \text{in } B_R.$$

In other words, harmonic functions in balls are bounded from above by their maximum on boundary and bounded from below by their minimum on boundary. Such a result is referred to as the *maximum principle*. Again, this is a general fact and will be proved for any harmonic functions in any bounded domains.

The mean-value property and the maximum principle are main topics in Section 4.3 and Section 4.4.

4.2. Fundamental Solutions

In this section, we use fundamental solutions of the Laplace equation to discuss regularity of harmonic functions.

First, as an application of Theorem 4.1, we prove that harmonic functions are smooth.

LEMMA 4.7. *Let Ω be a domain in $\mathbb{R}^n$ and $u \in C^2(\Omega)$ be a harmonic function in Ω. Then u is smooth in Ω.*

We note that, in its definition, a harmonic function is required only to be C^2.

PROOF. For any $x_0 \in \Omega$, we take a bounded C^1-domain Ω' in Ω such that $x \in \Omega'$ and $\bar{\Omega}'$ is a compact subset of Ω. Obviously, u is C^1 in $\bar{\Omega}'$ and $\Delta u = 0$ in Ω'. Then by Theorem 4.1, we have

$$u(x) = -\int_{\partial\Omega'} \big(\Gamma(x,y)\frac{\partial u}{\partial n_y}(y) - u(y)\frac{\partial \Gamma}{\partial n_y}(x,y)\big)dS_y \quad \text{for any } x \in \Omega'.$$

There is no singularity in the integrand since $x \in \Omega'$ and $y \in \partial\Omega'$. This implies easily that u is smooth in Ω'. $\square$

In the following, we will prove that harmonic functions are analytic. As the first step, we estimate derivatives of harmonic functions. For convenience, we consider harmonic functions in balls.

The following result is referred to as interior gradient estimates. It asserts that derivatives of a harmonic function at any point is controlled by its maximal value in a ball centered around the point.

LEMMA 4.8. *Suppose $u \in C\big(\bar{B}_R(x_0)\big)$ is harmonic in $B_R(x_0) \subset \mathbb{R}^n$. Then*

$$|\nabla u(x_0)| \leq \frac{C}{R}\max_{\bar{B}_R(x_0)} |u|,$$

where C is a positive constant depending only on n.

PROOF. Without loss of generality, we assume $x_0 = 0$.

We first consider $R = 1$ and employ the local version of the Green's identity. Choose a cut-off function $\varphi \in C_0^\infty(B_{3/4})$ such that $\varphi = 1$ in $B_{1/2}$ and $0 \leq \varphi \leq 1$. For $y \neq x$, $\Gamma(x,y)$ is harmonic in y and hence

$$\Delta_y\big(\varphi(y)\Gamma(x,y)\big) = \Delta_y\varphi(y)\Gamma(x,y) + 2\nabla_y\varphi(y)\cdot\nabla_y\Gamma(x,y).$$

This is zero for $|y| < 1/2$ and $3/4 < |y| < 1$ since φ is constant there. Apply the Green's formula to u and $\varphi\Gamma(x,\cdot)$ in $B_1 \setminus B_\rho(x)$ for $x \in B_{1/4}$ and ρ small enough. We proceed as in the proof of Theorem 4.1 to obtain

$$u(x) = -\int_{\frac{1}{2}<|y|<\frac{3}{4}} u(y)\big(\Delta_y\varphi(y)\Gamma(x,y) + 2\nabla_y\varphi(y)\cdot\nabla_y\Gamma(x,y)\big)dy,$$

for any $x \in B_{1/4}$. Then

$$\nabla u(x) = -\int_{\frac{1}{2}<|y|<\frac{3}{4}} u(y)\big(\Delta_y\varphi(y)\nabla_x\Gamma(x,y) + 2\nabla_y\varphi(y)\cdot\nabla_x\nabla_y\Gamma(x,y)\big)dy,$$

for any $x \in B_{1/4}$. There is no singularity in the integrand for $|x| < 1/4$ and $1/2 < |y| < 3/4$. Hence, we obtain

$$\sup_{B_{\frac{1}{4}}} |Du| \leq C\sup_{B_1}|u|,$$

where C is a constant depending only on n.

The general case follows from a simple dilation. Define

$$\tilde{u}(x) = u(Rx) \quad \text{for any } x \in B_1.$$

Then $\tilde{u}$ is a harmonic function in B_1. By applying the result we just proved to $\tilde{u}$, we obtain

$$\sup_{B_{\frac{1}{4}R}} |\nabla u| \leq \frac{C}{R} \sup_{B_R} |u|.$$

This implies the desired result. □

Next, we estimate derivatives of harmonic functions of arbitrary order.

LEMMA 4.9. *Suppose $u \in C\big(\bar{B}_R(x_0)\big)$ is harmonic in $B_R(x_0) \subset \mathbb{R}^n$. Then for any multi-index α with $|\alpha| = m$*

$$|\partial^\alpha u(x_0)| \leq \frac{C^m e^{m-1} m!}{R^m} \sup_{\bar{B}_R(x_0)} |u|,$$

where C is a positive constant depending only on n.

PROOF. We prove by an induction on $m \geq 1$. It holds for $m = 1$ by Lemma 4.17. We assume it holds for m and consider $m+1$. Let v be an arbitrary derivative of u of order m. Obviously, it is harmonic in $B_R(x_0)$. For $0 < \theta < 1$, we apply Lemma 4.17 to v in $B_{(1-\theta)R}(x_0)$ and get

$$|\nabla v(x_0)| \leq \frac{C}{(1-\theta)R} \max_{\bar{B}_{(1-\theta)R}(x_0)} |v|.$$

For any $x \in B_{(1-\theta)R}(x_0)$, we have $B_{\theta R}(x) \subset B_R(x_0)$. By the induction assumption, we obtain

$$|v(x)| \leq \frac{C^m e^{m-1} m!}{(\theta R)^m} \max_{\bar{B}_{\theta R}(x)} |u| \quad \text{for any } x \in B_{(1-\theta)R}(x_0),$$

and hence

$$\max_{\bar{B}_{(1-\theta)R}(x_0)} |v| \leq \frac{C^m e^{m-1} m!}{(\theta R)^m} \max_{\bar{B}_R(x_0)} |u|.$$

Therefore,

$$|\nabla v(x_0)| \leq \frac{C^{m+1} e^{m-1} m!}{(1-\theta)\theta^m R^{m+1}} \max_{\bar{B}_R(x_0)} |u|.$$

By taking $\theta = \dfrac{m}{m+1}$, we have

$$\frac{1}{(1-\theta)\theta^m} = \left(1 + \frac{1}{m}\right)^m (m+1) < e(m+1).$$

Hence the desired result is established for any derivatives of order $m+1$. □

Harmonic functions are not only smooth but also analytic in their domains. Real analytic functions will be studied in Section 7.2. Now we simply introduce the notion. A function $u : \mathbb{R}^n \to \mathbb{R}$ is called *analytic* near 0 if its Taylor series about 0 is convergent to u in a neighborhood of $0 \in \mathbb{R}^n$, i.e.,

$$u(x) = \sum_\alpha \frac{1}{\alpha!} \partial^\alpha u(0) x^\alpha \quad \text{for any } |x| < r.$$

THEOREM 4.10. *Harmonic functions are analytic.*

PROOF. Suppose u is a harmonic function in $\Omega \subset \mathbb{R}^n$. For any fixed $x \in \Omega$, we prove that u is equal to its Taylor series in a neighborhood of x. To do this, we take $B_{2R}(x) \subset \Omega$ and $h \in \mathbb{R}^n$ with $|h| \le R$. For any integer $m \ge 1$, we have by the Taylor expansion

$$u(x+h) = u(x) + \sum_{i=1}^{m-1} \frac{1}{i!} \left[(h_1\partial_{x_1} + \cdots + h_n\partial_{x_n})^i u \right](x) + R_m(h),$$

where

$$R_m(h) = \frac{1}{m!} \left[(h_1\partial_{x_1} + \cdots + h_n\partial_{x_n})^m u \right](x_1 + \theta h_1, \ldots, x_n + \theta h_n),$$

for some $\theta \in (0,1)$. Note that $x + h \in B_R(x)$ for $|h| < R$. Hence by Lemma 4.9, we obtain

$$|R_m(h)| \le \frac{1}{m!}|h|^m \cdot n^m \cdot \frac{C^m e^{m-1} m!}{R^m} \max_{\bar{B}_{2R}(x)} |u| \le \left(\frac{Cne|h|}{R} \right)^m \max_{\bar{B}_{2R}(x)} |u|.$$

Then for any h with $Cne|h| < R/2$, $R_m(h) \to 0$ as $m \to \infty$. □

The proof of Theorem 4.10 shows that the radius of convergence is uniform for harmonic functions, depending only on n. In other words, any harmonic functions defined in a fixed ball are equal to their Taylor expansions at least in balls with a fixed radius and such a radius depends only on the dimension n and is independent of harmonic functions. It is easy to see that each homogenous polynomial of fixed degree in Taylor expansions of harmonic functions is still harmonic, which is called a homogenous harmonic polynomial. Homogeneous harmonic polynomials constitute an important subject in the study of harmonic functions. Because of the homogeneity, homogeneous harmonic polynomials in $\mathbb{R}^n$ can be identified with corresponding spherical harmonics, their restrictions on the unit sphere $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$.

To end this section, we classify homogeneous harmonic polynomials in $\mathbb{R}^2$ and $\mathbb{R}^3$ and study their zero sets. For any integer $m \ge 1$, we denote by $\mathcal{H}_m(\mathbb{R}^n)$ the collection of all homogeneous harmonic polynomials of degree m in $\mathbb{R}^n$. Obviously, $\mathcal{H}_m(\mathbb{R}^n)$ is a linear space.

We start with $n = 2$ and denote points in $\mathbb{R}^2$ by (x, y). A general homogeneous polynomial P_m of degree m in $\mathbb{R}^2$ contains $m + 1$ coefficients. Then ΔP_m is a homogeneous polynomial of degree $m-2$ and therefore contains $m-1$ terms. These terms have to vanish if P_m is a harmonic function. This gives $m-1$ relations among the $m + 1$ constants if P_m is a harmonic function; so that these constants can be expressed linearly in terms of $(m+1)-(m-1)$ or 2 of them. Therefore, $\mathcal{H}_m(\mathbb{R}^2)$ is a linear space of dimension 2. We note that $\mathrm{Re}(x+iy)^m$ and $\mathrm{Im}(x+iy)^m$ are linearly independent homogeneous harmonic polynomials and hence form a basis in $\mathcal{H}_m(\mathbb{R}^2)$. In polar coordinates (r, θ), these two homogeneous harmonic polynomials are given by $r^m \cos m\theta$ and $r^m \sin m\theta$. Hence, any homogeneous harmonic polynomial P_m of degree m can be expressed by

$$P_m = c_1 r^m \sin m\theta + c_2 r^m \cos m\theta = a r^m \cos(m\theta + \theta_0),$$

for some constants c_1, c_2, a and $\theta_0 \in [0, 2\pi)$. We note that the nodal set, or the zero set, $P_m^{-1}(0)$ of P_m consists of m straight lines passing the origin and forming equal angles by any two consecutive lines.

Now we turn to $n = 3$. A general homogeneous polynomial P_m of degree m in $\mathbb{R}^3$ contains $\frac{1}{2}(m+1)(m+2)$ coefficients. Then ΔP_m is a homogeneous polynomial of degree $m-2$ and therefore contains $\frac{1}{2}m(m-1)$ terms. These terms have to vanish if P_m is a harmonic function. This gives $\frac{1}{2}m(m-1)$ relations among the $\frac{1}{2}(m+1)(m+2)$ constants if P_m is a harmonic function; so that these constants can be expressed linearly in terms of $\frac{1}{2}\{(m+1)(m+2) - m(m-1)\}$ or $2m+1$ of them. Therefore, $\mathcal{H}_m(\mathbb{R}^3)$ is a linear space of dimension $2m+1$. A basis of this linear space is given in terms of Legendre functions. We now discuss how to construct such a basis.

Let (x, y, z) and (r, θ, ϕ) be rectangular coordinates and corresponding polar coordinates in $\mathbb{R}^3$, i.e.,

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta.$$

Let $P_m = r^m Q_m(\theta, \varphi)$ be a homogeneous harmonic polynomial of degree m in $\mathbb{R}^3$. Then

$$m(m+1)Q_m + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\big(\sin\theta\frac{\partial Q_m}{\partial\theta}\big) + \frac{1}{\sin^2\theta}\frac{\partial^2 Q_m}{\partial\varphi^2} = 0.$$

By writing $\mu = \cos\theta$, we have

$$m(m+1)Q_m + \frac{\partial}{\partial\mu}\big((1-\mu^2)\frac{\partial Q_m}{\partial\mu}\big) + \frac{1}{1-\mu^2}\frac{\partial^2 Q_m}{\partial\varphi^2} = 0.$$

By setting $Q_m(\theta, \varphi) = f(\theta)g(\varphi)$, we obtain

$$m(m+1) + \frac{1}{f}\frac{\partial}{\partial\mu}\big((1-\mu^2)\frac{\partial f}{\partial\mu}\big) + \frac{1}{1-\mu^2}\frac{1}{g}\frac{\partial^2 g}{\partial\varphi^2} = 0.$$

The first two terms in this equation are independent of g, and therefore so is the last. Hence, the value of $\frac{1}{g}\frac{\partial^2 g}{\partial\varphi^2}$ must be a constant. Since g is 2π-periodic in φ, we take this constant to be $-k^2$ for an integer k. Thus

$$\frac{\partial^2 g}{\partial\varphi^2} = -k^2 g,$$

and hence

$$g(\varphi) = A\cos k\varphi + B\sin k\varphi,$$

where A and B are constants. Then f satisfies

$$\frac{\partial}{\partial\mu}\big((1-\mu^2)\frac{\partial f}{\partial\mu}\big) + \big(m(m+1) - \frac{k^2}{1-\mu^2}\big)f = 0,$$

which is known as *Legendre's associated equation.* If $f_{m,k}(\mu)$ is a solution, then

$$r^m\big(A\cos k\varphi + B\sin k\varphi\big)f_{m,k}(\cos\theta) \tag{4.5}$$

is a harmonic function wherever it is defined in $\mathbb{R}^3$. We are interested in only those $f_{m,k}$ such that (4.5) gives a homogeneous polynomial of degree m in $\mathbb{R}^3$. A lengthy calculation then yields for $k = 0, 1, \cdots, m$

$$f_{m,k}(\mu) = (1-\mu^2)^{\frac{k}{2}}\frac{d^{m+k}}{d\mu^{m+k}}(1-\mu^2)^m. \tag{4.6}$$

We omit details here. For $k = 0$, we have

$$f_{m,0}(\mu) = \frac{d^m}{d\mu^m}(1-\mu^2)^m.$$

This is the *Legendre function*. For each fixed positive integer m, the collection in (4.5) with $f_{m,k}$ given by (4.6) for $0 \leq k \leq m$ consists of $2m+1$ linearly independent homogeneous harmonic polynomials of degree m and hence forms a basis of $\mathcal{H}_m(\mathbb{R}^3)$. (We note that there is only one function for $k = 0$ in (4.5).)

Now, we demonstrate that these homogenous harmonic polynomials in (4.5) exhibit different properties for different k by plotting their nodal sets. Because of the homogeneity, it is convenient to consider restrictions of nodal sets to the unit sphere $\mathbb{S}^2$, which is called *the nodal curves*.

If $k = 0$, the corresponding harmonic polynomial is a constant multiple of the Legendre function $f_{m,0}(\cos\theta)$. A simple calculus argument shows that $f_{m,0}(\mu)$ has m distinct zeros between -1 and 1, arranged symmetrically about $\mu = 0$. Hence $f_{m,0}(\cos\theta)$ has m distinct zeros between 0 and π, arranged symmetrically about $\theta = \pi/2$. Accordingly on $\mathbb{S}^2$, the function $f_{m,0}(\cos\theta)$ vanishes on m latitude circles. These circles are symmetrically situated with respect to the equator $\theta = \pi/2$, and, if m is an odd number, the equator itself is one of these circles. Similarly, level sets

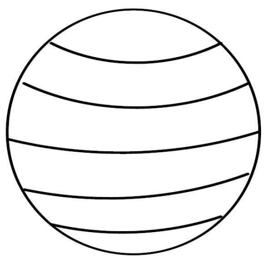

FIGURE 4.2. Nodal curves of zonal harmonics.

of this function consist of latitude circles. Because of this division of the sphere into zones by sets of latitude circles, the function $f_{m,0}(\cos\theta)$ and its constant multiples are called *zonal harmonics*.

If $0 < k < m$, the spherical harmonic is of the form

$$\big(A\cos k\varphi + B\sin k\varphi\big)\sin^k\theta\frac{d^{m+k}}{d\mu^{m+k}}(\mu^2-1)^m|_{\mu=\cos\theta}.$$

The first factor vanishes when $A\cos k\varphi + B\sin k\varphi = 0$, i.e., when $\tan k\varphi = -A/B$, and on $\mathbb{S}^2$ this corresponds to k great circles through the pole $\theta = 0$, the angle between the planes of any two consecutive great circles being π/k. The second factor vanishes at the points $\theta = 0$ and $\theta = \pi$, and the third on $m-k$ latitude circles, arranged like the corresponding circles in the case of zonal harmonics. Since the two sets of circles intersect orthogonally, these harmonics are called *tesseral harmonics*.

Finally, if $k = m$, the spherical harmonic is of the form

$$\big(A\cos m\varphi + B\sin m\varphi\big)\sin^m\theta,$$

which vanishes when $\theta = 0$ or π, or when $\tan m\varphi = -A/B$. This corresponds on $\mathbb{S}^2$ to the points $\theta = 0$ and $\theta = \pi$, and to m great circles through these points, the angle between the planes of any two consecutive great circles being π/m. As $\mathbb{S}^2$ is thus divided up into $2m$ sectors, these functions called *sectorial harmonics*. We

FIGURE 4.3. Nodal curves of tesseral harmonics.

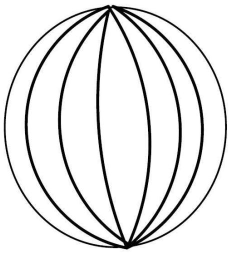

FIGURE 4.4. Nodal curves of sectorial harmonics.

note that sectorial harmonics are simply linear combinations of $\mathrm{Re}(x+iy)^m$ and $\mathrm{Im}(x+iy)^m$.

In summary, zonal harmonics, tesseral harmonics and sectorial harmonics exhibit different nodal curves. Generally, we do not expect a simple pattern for nodal sets of arbitrary spherical harmonics.

Finally, we point out that it is beyond the reach of this book to classify homogeneous harmonic polynomials in arbitrary dimensions.

4.3. Mean-Value Properties

As a simple consequence of the Poisson integral formula, harmonic functions are equal to their mean values over arbitrary spheres. This is the so-called mean-value property. In this section, we will use the mean-value property to discuss harmonic functions. The fundamental solution is not used throughout this section.

We first prove by a simple calculation that harmonic functions satisfy the mean-value property.

THEOREM 4.11. *Let Ω be a domain in $\mathbb{R}^n$ and $u \in C^2(\Omega)$ be harmonic in Ω. Then*

$$u(x) = \frac{1}{\omega_n r^{n-1}} \int_{\partial B_r(x)} u(y)\,dS_y \quad \textit{for any } B_r(x) \subset \Omega,$$

where ω_n is the area of the unit sphere ∂B_1 in $\mathbb{R}^n$.

PROOF. Take any ball $B_r(x) \subset \Omega$. For any $\rho \in (0, r)$, we apply the Green's formula in $B_\rho(x)$ and get

$$\begin{aligned} \int_{B_\rho(x)} \triangle u &= \int_{\partial B_\rho(x)} \frac{\partial u}{\partial n}\, dS = \rho^{n-1} \int_{\partial B_1} \frac{\partial u}{\partial \rho}(x+\rho w) dS_w \\ &= \rho^{n-1} \frac{\partial}{\partial \rho} \int_{\partial B_1} u(x+\rho w) dS_w. \end{aligned} \tag{4.7}$$

Hence for the harmonic function u, we have for any $\rho \in (0, r)$

$$\frac{\partial}{\partial \rho} \int_{\partial B_1} u(x+\rho w)\; dS_w = 0.$$

Integrating from 0 to r, we obtain

$$\int_{\partial B_1} u(x+rw)\; dS_w = \int_{\partial B_1} u(x)\; dS_w = u(x)\omega_n,$$

or

$$u(x) = \frac{1}{\omega_n} \int_{\partial B_1} u(x+rw)\; dS_w.$$

A simple change of variables yields the desired result. □

Theorem 4.11 asserts that harmonic functions equal to their mean values on spheres. There are two versions of mean-value properties, mean values on spheres and mean values on balls.

DEFINITION 4.12. We assume that Ω is a domain in $\mathbb{R}^n$. For $u \in C(\Omega)$, (i) u satisfies the first mean-value property if

$$u(x) = \frac{1}{\omega_n r^{n-1}} \int\limits_{\partial B_r(x)} u(y) dS_y \quad \text{for any } B_r(x) \subset \Omega;$$

(ii) u satisfies the second mean-value property if

$$u(x) = \frac{n}{\omega_n r^n} \int\limits_{B_r(x)} u(y) dy \quad \text{for any } B_r(x) \subset \Omega,$$

where ω_n denotes the surface area of the unit sphere in $\mathbb{R}^n$.

We note that $\omega_n r^{r-1}$ is the surface area of the sphere $\partial B_r(x)$ and $\omega_n r^n/n$ is the volume of the ball $B_r(x)$.

These two definitions are equivalent. In fact, if we write (i) as

$$u(x) r^{n-1} = \frac{1}{\omega_n} \int\limits_{\partial B_r(x)} u(y) dS_y,$$

we integrate with respect to r to get (ii). If we write (ii) as

$$u(x) r^n = \frac{n}{\omega_n} \int\limits_{B_r(x)} u(y) dy,$$

we differentiate to get (i).

By a change of variables, we also write mean-value properties in the following equivalent forms:

(i) u satisfies the first mean-value property if

$$u(x) = \frac{1}{\omega_n} \int\limits_{\partial B_1} u(x+ry)dS_y \quad \text{for any } B_r(x) \subset \Omega;$$

(ii) u satisfies the second-mean value property if

$$u(x) = \frac{n}{\omega_n} \int_{B_1} u(x+ry)\ dy \quad \text{for any } B_r(x) \subset \Omega.$$

For a function u satisfying mean-value properties, u is required only to be continuous. However, a harmonic function is required to be C^2. We now prove these two are equivalent. We already proved that harmonic functions satisfy the mean-value property in Theorem 4.11. We now prove that any function with the mean-value property is harmonic.

THEOREM 4.13. *If $u \in C(\Omega)$ has the mean-value property in Ω, then u is smooth and harmonic in Ω.*

PROOF. For the smoothness, we prove that u is equal to the convolution of itself with some smooth function. To this end, we choose a function $\varphi \in C_0^\infty(B_1)$ with $\int_{B_1} \varphi = 1$ and $\varphi(x) = \psi(|x|)$, i.e.,

$$\omega_n \int_0^1 r^{n-1}\psi(r)\ dr = 1.$$

The existence of such a function can be verified easily. We define $\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi\big(\frac{x}{\varepsilon}\big)$ for any $\varepsilon > 0$. Obviously, $\text{supp}\varphi_\varepsilon \subset B_\varepsilon$. We claim

$$u(x) = \int_\Omega u(y)\varphi_\varepsilon(y-x)dy \quad \text{for any } x \in \Omega \text{ with } d(x, \partial\Omega) > \varepsilon.$$

Then it follows that u is smooth. Moreover, by (4.7) in the proof of Theorem 4.11 and the mean-value property, we have for any $B_r(x) \subset \Omega$

$$\int_{B_r(x)} \Delta u = r^{n-1}\frac{\partial}{\partial r}\int_{\partial B_1} u(x+rw)dS_w = r^{n-1}\frac{\partial}{\partial r}\big(\omega_n u(x)\big) = 0.$$

This implies $\Delta u = 0$ in Ω.

Now we prove the claim. For any $x \in \Omega$ and $\varepsilon <$ dist $(x, \partial\Omega)$, we have

$$\begin{aligned}
\int_\Omega u(y)\varphi_\varepsilon(y-x)dy &= \int_{B_\varepsilon} u(x+y)\varphi_\varepsilon(y)dy = \frac{1}{\varepsilon^n}\int_{B_\varepsilon} u(x+y)\varphi\left(\frac{y}{\varepsilon}\right)dy \\
&= \int_{B_1} u(x+\varepsilon y)\varphi(y)dy \\
&= \int_0^1 r^{n-1}dr \int_{\partial B_1} u(x+\varepsilon rw)\varphi(rw)dS_w \\
&= \int_0^1 \psi(r)r^{n-1}dr \int_{\partial B_1} u(x+\varepsilon rw)dS_w \\
&= u(x)\omega_n \int_0^1 \psi(r)r^{n-1}dr = u(x),
\end{aligned}$$

where in the last equality we used the mean-value property. □

By combining Theorems 4.11 and 4.13, we conclude the following result.

COROLLARY 4.14. *Harmonic functions are smooth and satisfy the mean-value property.*

Now we prove the maximum principle for harmonic functions.

THEOREM 4.15. *Let $u \in C^2(\Omega) \cap C(\bar{\Omega})$ be harmonic in Ω. Then u assumes its maximum and minimum only on $\partial\Omega$ unless u is constant.*

PROOF. We prove only for the maximum. Set

$$D = \left\{ x \in \Omega;\ u(x) = M \equiv \max_{\bar{\Omega}} u \right\} \subset \Omega.$$

It is obvious that D is relatively closed; namely, for any sequence $\{x_m\} \subset D$, if $x_m \to x \in \Omega$, then $x \in D$. This follows easily from the continuity of u. Next we show that D is open. For any $x_0 \in D$, take $\bar{B}_r(x_0) \subset \Omega$ for some $r > 0$. By the mean-value property, we have

$$M = u(x_0) = \frac{n}{\omega_n r^n} \int_{B_r(x_0)} u(y)dy \leq M \frac{n}{\omega_n r^n} \int_{B_r(x_0)} dy = M.$$

This implies $u = M$ in $B_r(x_0)$. Hence D is both relatively closed and open in Ω. Therefore either $D = \phi$ or $D = \Omega$. A similar argument can be used to prove for the minimum. $\square$

A consequence is the uniqueness of solutions of the Dirichlet problem in a bounded domain.

COROLLARY 4.16. *For any $f \in C(\Omega)$ and $\varphi \in C(\partial\Omega)$, there exists at most one solution $u \in C^2(\Omega) \cap C(\bar{\Omega})$ of the following problem*

$$\begin{aligned} \Delta u &= f \ \text{ in } \Omega, \\ u &= \varphi \ \text{ on } \partial\Omega. \end{aligned}$$

PROOF. Let w be the difference of any two solutions. Then $\Delta w = 0$ in Ω and $w = 0$ on $\partial\Omega$. Theorem 4.15 implies $w = 0$ in Ω. $\square$

Compare Corollary 4.16 with Lemma 3.11, where the uniqueness was proved by energy estimates for solutions $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$.

In general, the uniqueness does not hold for unbounded domains. Consider the Dirichlet problem in an unbounded domain Ω

$$\begin{aligned} \Delta u &= 0 \text{ in } \Omega \\ u &= 0 \text{ on } \partial\Omega. \end{aligned}$$

First, we consider the case $\Omega = \{x \in \mathbb{R}^n; |x| > 1\}$. We have a nontrivial solution u given by

$$u(x) = \begin{cases} \log|x| & \text{for } n = 2; \\ |x|^{2-n} - 1 & \text{for } n \geq 3. \end{cases}$$

Note that $u \to \infty$ as $|x| \to \infty$ for $n = 2$ and u is bounded for $n \geq 3$. Next, we consider the upper half space $\Omega = \{x \in \mathbb{R}^n;\ x_n > 0\}$. Then $u(x) = x_n$ is a nontrivial solution, which is unbounded.

Next, we employ the mean-value property to prove interior gradient estimates.

LEMMA 4.17. *Suppose $u \in C\big(\bar{B}_R(x_0)\big)$ is harmonic in $B_R(x_0)$. Then*

$$|\nabla u(x_0)| \le \frac{n}{R} \max_{\bar{B}_R(x_0)} |u|.$$

We note that Lemma 4.17 improves Lemma 4.8.

PROOF. For simplicity, we assume $u \in C^1(\bar{B}_R(x_0))$. Since u is smooth, then $\Delta(u_{x_i}) = 0$, i.e., u_{x_i} is also harmonic in $B_R(x_0)$. Hence u_{x_i} satisfies the mean-value property. By the divergence theorem, we have

$$u_{x_i}(x_0) = \frac{n}{\omega_n R^n} \int_{B_R(x_0)} u_{x_i}(y)dy = \frac{n}{\omega_n R^n} \int_{\partial B_R(x_0)} u(y)\ \nu_i dS_y,$$

and hence

$$|u_{x_i}(x_0)| \le \frac{n}{\omega_n R^n} \max_{\partial B_R(x_0)} |u| \cdot \omega_n R^{n-1} \le \frac{n}{R} \max_{\bar{B}_R(x_0)} |u|.$$

This yields the desired result. $\square$

When u is nonnegative, we can improve Lemma 4.17.

LEMMA 4.18. *Suppose $u \in C\big(\bar{B}_R(x_0)\big)$ is a nonnegative harmonic function in $B_R(x_0)$. Then*

$$|\nabla u(x_0)| \le \frac{n}{R} u(x_0).$$

This result is often referred to as the differential Harnack inequality. It has many important consequences.

PROOF. As in the proof of Lemma 4.17, by the divergence theorem and the nonnegativeness of u, we have

$$|u_{x_i}(x_0)| \le \frac{n}{\omega_n R^n} \int_{\partial B_R(x_0)} u(y)\ dS_y = \frac{n}{R} u(x_0),$$

where in the last equality we used the mean-value property. $\square$

As an application, we prove the following result which is referred to as the Liouville Theorem.

COROLLARY 4.19. *A harmonic function in $\mathbb{R}^n$ bounded from above or below is constant.*

PROOF. Suppose u is a harmonic function in $\mathbb{R}^n$ with $u \ge 0$. Then for any $x \in \mathbb{R}^n$, we apply Lemma 4.18 to u in $B_R(x)$ and then let $R \to \infty$. We conclude $\nabla u(x) = 0$ for any $x \in \mathbb{R}^n$ and hence u is constant. $\square$

As another application, we prove the Harnack inequality, which asserts that nonnegative harmonic functions have comparable values in compact subsets.

LEMMA 4.20. *Suppose u is a nonnegative harmonic function in $B_R(x_0)$. Then*

$$u(x) \le Cu(y) \quad \text{for any } x, y \in B_{\frac{R}{2}}(x_0),$$

where C is a positive constant depending only on n.

PROOF. By considering $u+\varepsilon$ for any $\varepsilon>0$, we assume $u>0$ in $B_R(x_0)$. For any $x\in B_{R/2}(x_0)$, we apply Lemma 4.18 to u in $B_{R/2}(x)$ and get

$$|\nabla u(x)|\le \frac{2n}{R}u(x),$$

or

$$|\nabla \log u(x)|\le \frac{2n}{R}.$$

For any $x,y\in B_{R/2}(x_0)$,

$$\log\frac{u(x)}{u(y)}=\int_0^1\frac{d}{dt}\log u(ty+(1-t)x)dt.$$

With $ty+(1-t)x\in B_{R/2}(x_0)$ for any $t\in[0,1]$ and $|x-y|\le R$, a simple integration yields

$$\log\frac{u(x)}{u(y)}\le|x-y|\int_0^1|\nabla\log u(ty+(1-t)x)|dt\le\frac{2n}{R}|x-y|\le 2n.$$

Therefore

$$u(x)\le e^{2n}u(y).$$

This yields the desired result. □

In fact, Lemma 4.20 can be proved directly by the mean-value property.

ANOTHER PROOF OF LEMMA 4.20. Consider any $B_{4r}(\bar{x})\subset B_R(x_0)$. We first claim

$$u(x)\le 2^n u(y)\quad\text{for any } x,y\in B_r(\bar{x}).$$

To see this, we note that $B_r(x)\subset B_{2r}(y)\subset B_{4r}(\bar{x})$ for any $x,y\in B_r(\bar{x})$. Then the mean-value properties imply

$$u(x)=\frac{n}{\omega_n r^n}\int_{B_r(x)}u\le\frac{n}{\omega_n r^n}\int_{B_{2r}(y)}u\le 2^n u(y).$$

With $r=R/8$, we choose finitely many $\bar{x}_1,\cdots,\bar{x}_N\in B_{R/2}(x_0)$ such that $\{B_r(\bar{x}_i)\}$ covers $B_{R/2}(x_0)$. Obviously, $B_{4r}(\bar{x}_i)\subset B_R(x_0)$ and N is a constant depending only on n. Then we have the desired result by choosing $C=2^{nN}$. □

4.4. The Maximum Principle

One of the important tools in studying harmonic functions is the maximum principle. In this section, we discuss the maximum principle for a class of elliptic differential equations slightly more general than the Laplace equation and derive a priori estimates for boundary-value problems, interior gradient estimates and the Harnack inequality.

Throughout this section, we assume Ω is a bounded domain in $\mathbb{R}^n$ and c is a continuous function in $\bar{\Omega}$. Sometimes, we assume $c\le 0$ in Ω. For $u\in C^2(\Omega)\cap C(\bar{\Omega})$ and $f\in C(\Omega)$, we consider

$$\Delta u+c(x)u=f(x)\quad\text{in } \Omega. \tag{4.8}$$

Obviously, u is harmonic if $c=f=0$. A C^2-function u is called a subsolution (or supersolution) of (4.8) if $\Delta u+cu\ge f$ (or $\Delta u+cu\le f$). If $c=0$ and $f=0$, subsolutions (or supersolutions) are called subharmonic (or superharmonic) functions. In other words, a C^2-function u is subharmonic (or superharmonic) if $\Delta u \geq 0$ (or $\Delta u \leq 0$).

Now we start to prove the maximum principle without using mean-value properties. We prove a simple result first.

LEMMA 4.21. *Suppose $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfies $\Delta u + cu > 0$ in Ω with $c(x) \leq 0$ in Ω. If u has a nonnegative maximum in $\bar{\Omega}$, then u cannot attain this maximum in Ω.*

PROOF. Suppose u attains its nonnegative maximum of $\bar{\Omega}$ at $x_0 \in \Omega$. Then $\nabla u(x_0) = 0$ and the Hessian matrix $\big(\nabla^2 u(x_0)\big)$ is nonpositive definite. By taking a trace, we have

$$\Delta u(x_0) = \mathrm{tr}(\nabla^2 u(x_0)) \leq 0,$$

and hence

$$(\Delta u + cu)(x_0) \leq 0,$$

which is a contradiction. □

REMARK 4.22. If $c(x) \equiv 0$, the requirement for the nonnegativeness on u can be removed. This remark holds for many results in the rest of this section.

REMARK 4.23. Lemma 4.21 in fact holds for more general elliptic differential equations. For $u \in C^2(\Omega) \cap C(\bar{\Omega})$ and $f \in C(\Omega)$, consider

$$Lu \equiv \sum_{i,j=1}^{n} a_{ij}(x) u_{x_i x_j} + \sum_{i=1}^{n} b_i(x) u_{x_i} + c(x) u = f(x) \quad \text{in } \Omega.$$

We assume that a_{ij}, b_i and c are continuous and hence bounded in $\bar{\Omega}$ and that L is uniformly elliptic in Ω in the following sense

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2 \quad \text{for any } x \in \Omega \text{ and any } \xi \in \mathbb{R}^n,$$

for some positive constant λ. The uniform ellipticity means a uniform positive lower bound of all eigenvalues of (a_{ij}) in Ω. Many results in this section hold for $Lu = f$.

We now prove the maximum principle for subsolutions.

THEOREM 4.24. *Suppose $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfies $\Delta u + cu \geq 0$ in Ω with $c(x) \leq 0$ in Ω. Then u attains on $\partial\Omega$ its nonnegative maximum in $\bar{\Omega}$, i.e.,*

$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u^+.$$

A continuous function in $\bar{\Omega}$ always attains its maximum in $\bar{\Omega}$. Theorem 4.24 asserts that any subsolution continuous up to the boundary attains its nonnegative maximum on the boundary $\partial\Omega$, possibly also in Ω. Theorem 4.24 is often called the weak maximum principle. A stronger version asserts that subsolutions attain their maximum only on the boundary. We will prove the strong maximum principle later.

PROOF. Without loss of generality, we assume Ω is contained in a ball of radius R and centered at the origin, i.e., $\Omega \subset B_R$. For any $\varepsilon > 0$, consider

$$u_\varepsilon(x) = u(x) - \varepsilon(R^2 - |x|^2).$$

By $c \le 0$ and $|x| < R$ in Ω, we have

$$\Delta u_\varepsilon + cu_\varepsilon = \Delta u + cu + 2n\varepsilon - \varepsilon c(R^2 - |x|^2) \ge 2n\varepsilon > 0 \quad \text{in } \Omega.$$

By Lemma 4.21, u_ε attains its nonnegative maximum only on $\partial\Omega$, i.e.,

$$\sup_\Omega u_\varepsilon \le \sup_{\partial\Omega} u_\varepsilon^+.$$

Then we obtain

$$\sup_\Omega u \le \sup_\Omega u_\varepsilon + \varepsilon R^2 \le \sup_{\partial\Omega} u_\varepsilon^+ + \varepsilon R^2 \le \sup_{\partial\Omega} u^+ + \varepsilon R^2.$$

We finish the proof by letting $\varepsilon \to 0$. □

The following result is often referred to as the comparison principle.

Corollary 4.25. *Suppose $c \le 0$ in Ω and $u, v \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfy*

$$\begin{aligned} \Delta u + cu &\ge \Delta v + cv \quad \text{in } \Omega \\ u &\le v \quad \text{on } \partial\Omega. \end{aligned}$$

Then $u \le v$ in Ω.

Proof. The difference $w = u - v$ satisfies $\Delta w + cw \ge 0$ in Ω and $w \le 0$ on $\partial\Omega$. Theorem 4.24 implies $w \le 0$ in Ω. □

The comparison principle provides a reason that functions u satisfying $\Delta u + cu \ge f$ are called subsolutions. They are less than solutions v of $\Delta v + cv = f$ with the same boundary values.

A consequence of the comparison principle is the uniqueness of solutions of Dirichlet problems.

Corollary 4.26. *Let Ω be a bounded domain in $\mathbb{R}^n$ and $u \in C^2(\Omega) \cap C(\bar{\Omega})$ be a solution of*

$$\begin{aligned} \Delta u + cu &= f \quad \text{in } \Omega \\ u &= \varphi \quad \text{on } \partial\Omega, \end{aligned}$$

for some $f \in C(\bar{\Omega})$ and $\varphi \in C(\partial\Omega)$. If $c(x) \le 0$ in Ω, then u is the unique solution.

The boundedness of domains Ω is essential, since it guarantees the existence of maximum and minimum of u in $\bar{\Omega}$. The uniqueness does not hold if domains are unbounded. Examples are given in Section 4.3. Equally important is the nonpositiveness of the coefficient c. For example, $u = \sin x$ is a nontrivial solution of the problem

$$\begin{aligned} u'' + u &= 0 \quad \text{in } (0, \pi) \\ u(0) = u(\pi) &= 0. \end{aligned}$$

The weak maximum principle asserts subsolutions of elliptic differential equations with nonpositive zero order coefficients attain on boundary their nonnegative maximum. In fact, they can attain their nonnegative maximum only on boundary, unless they are constant. This is the strong maximum principle. To prove this, we need the following Hopf lemma. Let u be a subsolution in Ω which attains its maximum on $\partial\Omega$, say at $x_0 \in \partial\Omega$. Then $\frac{\partial u}{\partial n}(x_0) \ge 0$. The Hopf lemma asserts that it is in fact positive.

THEOREM 4.27. *Let B be an open ball in $\mathbb{R}^n$ with $x_0 \in \partial B$. Suppose $u \in C^2(B) \cap C^1(B \cup \{x_0\})$ satisfies $\Delta u + cu \geq 0$ in B with $c(x) \leq 0$ in B. Assume in addition that*

$$u(x) < u(x_0) \quad \text{for any } x \in B \text{ and } u(x_0) \geq 0.$$

Then

$$\frac{\partial u}{\partial n}(x_0) > 0.$$

PROOF. We assume that B is centered at the origin with radius r. We assume further that $u \in C(\bar{B})$ and $u(x) < u(x_0)$ for any $x \in \bar{B} \backslash \{x_0\}$, since we can construct a ball $B_* \subset B$ tangent to B at x_0.

Set $D = B \cap B_{r/2}(x_0)$ and for some α to be determined

$$h(x) = e^{-\alpha|x|^2} - e^{-\alpha r^2}.$$

A direct calculation yields

$$\begin{aligned} \Delta h + ch =& e^{-\alpha|x|^2}\big(4\alpha^2|x|^2 - 2n\alpha + c\big) - ce^{-\alpha r^2} \\ \geq& e^{-\alpha|x|^2}\big(4\alpha^2|x|^2 - 2n\alpha + c\big), \end{aligned}$$

where we used $c \leq 0$ in B. By $|x| \geq r/2$ in D and choosing α sufficiently large, we conclude

$$\Delta h + ch > 0 \quad \text{in } D.$$

Consider $v(x) = u(x) + \varepsilon h(x)$. Then,

$$\Delta v + cv = \Delta u + cu + \varepsilon(\Delta h + ch) > 0 \quad \text{in } D,$$

for any $\varepsilon > 0$. By Lemma 4.21, v cannot attain its nonnegative maximum in D.

Next, we prove, for some small $\varepsilon > 0$, v attains at x_0 its nonnegative maximum. Consider v on the boundary ∂D.

(i) For $x \in \partial D \cap B$, since $u(x) < u(x_0)$, so $u(x) < u(x_0) - \delta$ for some $\delta > 0$. Take ε small such that $\varepsilon h < \delta$ on $\partial D \cap B$. Hence, for such an ε, we have $v(x) < u(x_0)$ for any $x \in \partial D \cap B$.

(ii) On $\partial D \cap \partial B$, $h(x) = 0$ and $u(x) < u(x_0)$ for $x \neq x_0$. Hence $v(x) < u(x_0)$ on $\partial D \cap \partial B \setminus \{x_0\}$ and $v(x_0) = u(x_0)$.

Therefore, we conclude

$$\frac{\partial v}{\partial n}(x_0) \geq 0,$$

or

$$\frac{\partial u}{\partial n}(x_0) \geq -\varepsilon \frac{\partial h}{\partial n}(x_0) = 2\varepsilon\alpha r e^{-\alpha r^2} > 0.$$

This yields the desired result. □

Now, we are ready to prove the strong maximum principle.

THEOREM 4.28. *Let $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfy $\Delta u + cu \geq 0$ with $c(x) \leq 0$ in Ω. Then the nonnegative maximum of u in $\bar{\Omega}$ can be attained only on $\partial\Omega$ unless u is a constant.*

Proof. Let M be the nonnegative maximum of u in $\bar{\Omega}$. Set $D = \{x \in \Omega; u(x) = M\}$. It is relatively closed in Ω. We prove $D = \Omega$ by contradiction. If D is a proper subset of Ω, we can find an open ball $B \subset \Omega \setminus D$ with a point on its boundary belonging to D. In fact, we may choose a point $p \in \Omega \setminus D$ such that $d(p, D) < d(p, \partial\Omega)$ first, construct a small ball in $\Omega \setminus D$ of center p and then extend this ball. It hits D before hitting $\partial\Omega$. Suppose $x_0 \in \partial B \cap D$. Obviously, we have $\Delta u + cu \geq 0$ in B and

$$u(x) < u(x_0) \quad \text{for any } x \in B \text{ and } u(x_0) = M \geq 0.$$

Theorem 4.27 implies $\dfrac{\partial u}{\partial n}(x_0) > 0$, where $\mathbf{n}$ is the outward normal vector at x_0 to the ball B. While x_0 is an interior maximal point of Ω, we have $\nabla u(x_0) = 0$. This leads to a contradiction. □

The following result improves Corollary 4.25.

Corollary 4.29. *Suppose $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfies $\Delta u + cu \geq 0$ in Ω with $c(x) \leq 0$ in Ω. If $u \leq 0$ on $\partial\Omega$, then either $u < 0$ in Ω or $u \equiv 0$ in Ω.*

In order to discuss boundary-value problems with general boundary conditions, we need the following result, which is a corollary of Theorem 4.27 and Theorem 4.28.

Corollary 4.30. *Suppose Ω is a bounded C^1-domain in $\mathbb{R}^n$ and that $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ satisfies $\Delta u + cu \geq 0$ in Ω with $c(x) \leq 0$. Assume u attains its nonnegative maximum at $x_0 \in \bar{\Omega}$. Then $x_0 \in \partial\Omega$ and*

$$\frac{\partial u}{\partial n}(x_0) > 0,$$

unless u is a constant in $\bar{\Omega}$.

Now we discuss the uniqueness of solutions of a class of boundary-value problems with general boundary conditions.

Corollary 4.31. *Suppose Ω is a bounded C^1-domain in $\mathbb{R}^n$, c is a continuous function in Ω and α is a continuous function on $\partial\Omega$. Let $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ be a solution of the boundary-value problem*

$$\begin{aligned} \Delta u + cu &= f \quad \text{in } \Omega \\ \frac{\partial u}{\partial n} + \alpha u &= \varphi \quad \text{on } \partial\Omega, \end{aligned}$$

for some $f \in C(\bar{\Omega})$ and $\varphi \in C(\partial\Omega)$. Assume in addition that $c \leq 0$ in Ω and $\alpha \geq 0$ on $\partial\Omega$. Then u is the unique solution if $c \not\equiv 0$ or $\alpha \not\equiv 0$. If $c \equiv 0$ and $\alpha \equiv 0$, u is unique up to additive constants.

Proof. Suppose u is a solution of the homogeneous problem

$$\begin{aligned} \Delta u + cu &= 0 \quad \text{in } \Omega \\ \frac{\partial u}{\partial n} + \alpha u &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Case 1. $c \not\equiv 0$ or $\alpha \not\equiv 0$. Suppose that u has a positive maximum at $x_0 \in \bar{\Omega}$. If u is a positive constant, there leads to a contradiction to the condition $c \not\equiv 0$ in

Ω or $\alpha \not\equiv 0$ on $\partial\Omega$. Otherwise $x_0 \in \partial\Omega$ and $\frac{\partial u}{\partial n}(x_0) > 0$ by Corollary 4.30, which contradicts the boundary value. Therefore $u \equiv 0$.

Case 2. $c \equiv 0$ and $\alpha \equiv 0$. Suppose u is a nonconstant solution. Then its maximum in $\bar{\Omega}$ is assumed only on $\partial\Omega$ by Theorem 4.28, say at $x_0 \in \partial\Omega$. Again Corollary 4.30 implies $\frac{\partial u}{\partial n}(x_0) > 0$. This contradiction shows that u is a constant. □

Proving the uniqueness of solutions of boundary-value problems is an important application of maximum principles. Equally important or more important is to derive a priori estimates. In derivations of a priori estimates, essential steps consist of constructions of auxiliary functions. Next, we derive a priori estimates for solutions of Dirichlet problems.

THEOREM 4.32. *Suppose $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfies*

$$\begin{aligned} \Delta u + cu &= f \quad \text{in } \Omega, \\ u &= \varphi \quad \text{on } \partial\Omega, \end{aligned}$$

for some $f \in C(\bar{\Omega})$ and $\varphi \in C(\partial\Omega)$. If $c(x) \le 0$, then

$$\max_{\Omega} |u| \le \max_{\partial\Omega} |\varphi| + C \max_{\Omega} |f|,$$

where C is a positive constant depending only on n and diam(Ω).

If $\Omega = B_R(x_0)$, then

$$\max_{B_R(x_0)} |u| \le \max_{\partial B_R(x_0)} |\varphi| + \frac{R^2}{2n} \max_{B_R(x_0)} |f|.$$

This follows from the proof below easily.

PROOF. We construct an auxiliary function w in Ω such that
(i) $(\Delta + c)(w \pm u) = (\Delta + c)w \pm f \le 0$, or $(\Delta + c)w \le \mp f$ in Ω;
(ii) $w \pm u = w \pm \varphi \ge 0$, or $w \ge \mp\varphi$ on $\partial\Omega$.
Set

$$F = \max_{\Omega} |f|, \quad \Phi = \max_{\partial\Omega} |\varphi|.$$

We need

$$\begin{aligned} \Delta w + cw &\le -F \quad \text{in } \Omega, \\ w &\ge \Phi \quad \text{on } \partial\Omega. \end{aligned}$$

Without loss of generality, we assume Ω is contained in a ball of radius R and centered at the origin, i.e., $\Omega \subset B_R$. Set

$$w = \Phi + \frac{F}{2n}(R^2 - |x|^2).$$

Then by $c \le 0$ and $|x| \le R$ in Ω, we have

$$\Delta w + cw = -F + c\Phi + \frac{cF}{2n}(R^2 - |x|^2) \le -F.$$

We also have $w \le \Phi$ on $\partial\Omega$. Hence w satisfies (i) and (ii). By Corollary 4.25, the comparison principle, we conclude $-w \le u \le w$ in Ω, and in particular,

$$|u(x)| \le \Phi + \frac{1}{2n}(R^2 - |x|^2)F \quad \text{for any } x \in \Omega.$$

This yields the desired result. $\square$

Now we consider a class of general boundary-value problems.

THEOREM 4.33. *Suppose $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ satisfies*

$$\begin{aligned} \Delta u + cu &= f \quad \textit{in } \Omega, \\ \frac{\partial u}{\partial n} + \alpha u &= \varphi \quad \textit{on } \partial\Omega, \end{aligned}$$

for some $f \in C(\bar{\Omega})$ and $\varphi \in C(\partial\Omega)$. If $c(x) \le 0$ in Ω and $\alpha(x) \ge \alpha_0$ on $\partial\Omega$ for a positive constant α_0, then

$$\max_{\Omega} |u| \le C\big(\max_{\partial\Omega} |\varphi| + \max_{\Omega} |f|\big),$$

where C is a positive constant depending only on α_0 and diam(Ω).

If $c \equiv 0$ in Ω and $\alpha \equiv 0$ on $\partial\Omega$, the homogeneous problem in Theorem 4.33 (with $f \equiv 0$ in Ω and $\varphi \equiv 0$ on $\partial\Omega$) admits a nontrivial solution $u \equiv 1$ in Ω. Hence there does not hold a sup-norm estimate in this case.

PROOF. By assuming $\Omega \subset B_R$ for some $R > 0$, we prove

$$\sup_{\Omega} |u| \le \frac{1}{\alpha_0} \max_{\partial\Omega} |\varphi| + \frac{1}{2n}\big(\frac{1+R^2}{\alpha_0} + R^2\big) \max_{\Omega} |f|.$$

Set

$$\Phi = \max_{\partial\Omega} |\varphi|, \quad F = \max_{\Omega} |f|,$$

and

$$v(x) = \frac{1}{\alpha_0}\Phi + \frac{1}{2n}\big(\frac{1+R^2}{\alpha_0} + R^2 - |x|^2\big)F.$$

Then

$$\Delta v + cv = -F + cv \le -F \quad \text{in } \Omega,$$

and

$$\begin{aligned} \frac{\partial v}{\partial n} + \alpha v =& \frac{\alpha}{\alpha_0}\Phi + \frac{F}{2n}\big(-2x\cdot \mathbf{n} + \alpha\big(\frac{1+R^2}{\alpha_0} + R^2 - |x|^2\big)\big) \\ \ge& \Phi + \frac{F}{2n}(-2x\cdot\mathbf{n} + 1 + R^2) \ge \Phi \quad \text{on } \partial\Omega, \end{aligned}$$

where $\mathbf{n}$ is the unit exterior normal vector of $\partial\Omega$. With

$$w = v \pm u,$$

we have

$$\begin{aligned} \Delta w + cw &\le 0 \quad \text{in } \Omega, \\ \frac{\partial w}{\partial n} + \alpha w &\ge 0 \quad \text{on } \partial\Omega. \end{aligned}$$

It is easy to prove $w \ge 0$ in Ω, or

$$|u| \le v \quad \text{in } \Omega.$$

This yields the desired result. $\square$

In the following, we derive gradient estimates, estimates of derivatives. The basic method to derive gradient estimates, the so-called the Bernstein method, involves a differentiation of the equation with respect to $x_k, k = 1, \dots, n$, followed by a multiplication by u_{x_k} and summation over k. The maximum principle is then applied to the resulting equation in the function $v = |\nabla u|^2$, possibly with some modifications. There are two kinds of gradient estimates, global gradient estimates and interior gradient estimates. We use the Laplace equation to demonstrate ideas of proving interior gradient estimates. Compare with Lemma 4.17.

THEOREM 4.34. *Suppose $u \in C(\bar{B}_1)$ is harmonic in B_1. Then*
$$\sup_{B_{\frac{1}{2}}} |\nabla u| \leq C \sup_{\partial B_1} |u|,$$
where C is a positive constant depending only on n.

PROOF. A direct calculation shows
$$\triangle(|\nabla u|^2) = 2\sum_{i,j=1}^n u_{x_i x_j}^2 + 2\sum_{i=1}^n u_{x_i}(\triangle u)_{x_i} = 2\sum_{i,j=1}^n u_{x_i x_j}^2,$$
where we used $\triangle u = 0$ in B_1. Hence $|\nabla u|^2$ is a subharmonic function. To get interior estimates, we need to introduce a cut-off function. For any nonnegative function $\varphi \in C_0^1(B_1)$, we have
$$\triangle(\varphi|\nabla u|^2) = (\triangle\varphi)|\nabla u|^2 + 4\sum_{i,j=1}^n \varphi_{x_i} u_{x_j} u_{x_i x_j} + 2\varphi \sum_{i,j=1}^n u_{x_i x_j}^2.$$
By the Cauchy inequality, we have
$$4|\varphi_{x_i} u_{x_j} u_{x_i x_j}| \leq 2\varphi u_{x_i x_j}^2 + \frac{2}{\varphi}\varphi_{x_i}^2 u_{x_j}^2.$$
Then
$$\triangle(\varphi|\nabla u|^2) \geq \big(\Delta\varphi - \frac{2|\nabla\varphi|^2}{\varphi}\big)|\nabla u|^2.$$
To get a bounded $|\nabla\varphi|^2/\varphi$ in B_1, we take $\varphi = \eta^2$ for some $\eta \in C_0^1(B_1)$. Moreover, we require $\eta \equiv 1$ in $B_{1/2}$. Then
$$\triangle(\eta^2|\nabla u|^2) \geq \big(2\eta\triangle\eta - 6|\nabla\eta|^2\big)|\nabla u|^2 \geq -C|\nabla u|^2,$$
where C is a positive constant depending only on η and n. Note
$$\triangle(u^2) = 2|\nabla u|^2 + 2u\triangle u = 2|\nabla u|^2,$$
since u is harmonic. By taking a constant α large enough, we get
$$\triangle(\eta^2|\nabla u|^2 + \alpha u^2) \geq (2\alpha - C)|\nabla u|^2 \geq 0.$$
We apply Theorem 4.24, the maximum principle, to get
$$\sup_{B_1}(\eta^2|\nabla u|^2 + \alpha u^2) \leq \sup_{\partial B_1}(\eta^2|\nabla u|^2 + \alpha u^2).$$
This implies the desired result easily since $\eta = 0$ on ∂B_1 and $\eta = 1$ in $B_{1/2}$. $\square$

Next we derive the differential Harnack inequality for harmonic functions. Compare this with Lemma 4.18.

THEOREM 4.35. *Suppose u is a positive harmonic function in B_1. Then*

$$\sup_{B_{\frac{1}{2}}} |\nabla \log u| \leq C,$$

where C is a positive constant depending only on n.

PROOF. Set $v = \log u$. A direct calculation shows

$$\triangle v = -|\nabla v|^2.$$

Next, we prove an interior gradient estimate for v. By setting $w = |\nabla v|^2$, we get

$$\triangle w + 2\sum_{i=1}^{n} v_{x_i} w_{x_i} = 2\sum_{i,j=1}^{n} v_{x_i x_j}^2.$$

As before, we need to introduce a cut-off function. First note

$$\sum_{i,j=1}^{n} v_{x_i x_j}^2 \geq \sum_{i=1}^{n} v_{x_i x_i}^2 \geq \frac{1}{n}(\triangle v)^2 = \frac{|\nabla v|^4}{n} = \frac{w^2}{n}. \tag{4.9}$$

Take a nonnegative function $\varphi \in C_0^1(B_1)$. As in the proof of Theorem 4.34, we have

$$\begin{aligned}
&\triangle(\varphi w) + 2\sum_{i=1}^{n} v_{x_i}(\varphi w)_{x_i} \\
=&2\varphi \sum_{i,j=1}^{n} v_{x_i x_j}^2 + 4\sum_{i,j=1}^{n} \varphi_{x_i} v_{x_j} v_{x_i x_j} + 2w\sum_{i=1}^{n} \varphi_{x_i} v_{x_i} + (\triangle \varphi) w \\
\geq&\varphi \sum_{i,j=1}^{n} v_{x_i x_j}^2 - 2|\nabla \varphi||\nabla v|^3 + \left(\triangle \varphi - \frac{4|\nabla \varphi|^2}{\varphi}\right)|\nabla v|^2.
\end{aligned}$$

Here we keep one $v_{x_i x_j}^2$ instead of dropping it entirely in the proof of Theorem 4.34. To get a bounded $|\nabla \varphi|^2/\varphi$ in B_1, we take $\varphi = \eta^4$ for some $\eta \in C_0^1(B_1)$. We obtain by (4.9)

$$\begin{aligned}
&\triangle(\eta^4 w) + 2\sum_{i=1}^{n} v_{x_i}(\eta^4 w)_{x_i} \\
\geq&\frac{1}{n}\eta^4 |\nabla v|^4 - 8\eta^3 |\nabla \eta||\nabla v|^3 + 4\eta^2(\eta \triangle \eta - 13|\nabla \eta|^2)|\nabla v|^2.
\end{aligned}$$

Note

$$\frac{1}{2n}t^4 - 8|\nabla \eta| t^3 + 4(\eta \triangle \eta - 13|\nabla \eta|^2)t^2 \geq -C \quad \text{for any } t \in \mathbb{R},$$

where C is a positive constant depending only on n and η. Hence with $t = \eta|\nabla v|$, we get

$$\triangle(\eta^4 w) + 2\sum_{i=1}^{n} v_{x_i}(\eta^4 w)_{x_i} \geq \frac{1}{2n}\eta^4 w^2 - C.$$

where C is a positive constant depending only on n and η.

Suppose $\eta^4 w$ attains its maximum at $x_0 \in B_1$. Then $\nabla(\eta^4 w) = 0$ and $\triangle(\eta^4 w) \leq 0$ at x_0. Hence

$$\eta^4 w^2(x_0) \leq C.$$

If $w(x_0) \geq 1$, then $\eta^4 w(x_0) \leq C$. Otherwise $\eta^4 w(x_0) \leq \eta^4(x_0)$. In both cases we conclude

$$\eta^4 w \leq C_* \quad \text{in } B_1,$$

where C_* is a positive constant depending only on n and η. With the help of the definition of w and $\eta = 1$ in $B_{1/2}$, this implies the desired result. □

The following result is called the Harnack inequality. Compare it with Lemma 4.20.

Corollary 4.36. *Suppose u is a nonnegative harmonic function in B_1. Then*

$$u(x_1) \leq Cu(x_2) \quad \text{for any } x_1, x_2 \in B_{\frac{1}{2}},$$

where C is a positive constant depending only on n.

The proof is identical to the first proof of Lemma 4.20 and is omitted.

To end this section, we discuss isolated singularities of harmonic functions. We note that the fundamental solution of the Laplace operator has an isolated singularity and is harmonic elsewhere. The next result asserts that an isolated singularity of harmonic functions, if "better" than that of the fundamental solution, can be removed.

Theorem 4.37. *Suppose u is harmonic in $B_R \setminus \{0\} \subset \mathbb{R}^n$ and satisfies*

$$u(x) = \begin{cases} o(\log|x|), & n = 2 \\ o(|x|^{2-n}), & n \geq 3 \end{cases} \quad \text{as } |x| \to 0.$$

Then u can be defined at 0 so that it is C^2 and harmonic in B_R.

Proof. Assume u is continuous in $0 < |x| \leq R$. Let v solve

$$\begin{aligned} \Delta v &= 0 \quad \text{in } B_R, \\ v &= u \quad \text{on } \partial B_R. \end{aligned}$$

The existence of v is guaranteed by Theorem 4.6. By the maximum principle,

$$|v| \leq M \quad \text{in } B_R,$$

where $M = \max_{\partial B_R} |u|$. We prove $u = v$ in $B_R \setminus \{0\}$. Set $w = v - u$ in $B_R \setminus \{0\}$ and $M_r = \max_{\partial B_r} |w|$ for any $r < R$. We only consider the case $n \geq 3$. First, we have

$$|w(x)| \leq M_r \cdot \frac{r^{n-2}}{|x|^{n-2}} \quad \text{for any } x \in \partial B_r.$$

Note that w and $\dfrac{1}{|x|^{n-2}}$ are harmonic in $B_R \setminus B_r$ with $w = 0$ on ∂B_R. Hence the maximum principle implies

$$|w(x)| \leq M_r \cdot \frac{r^{n-2}}{|x|^{n-2}} \text{ for any } x \in B_R \setminus B_r,$$

where

$$M_r = \max_{\partial B_r} |v - u| \leq \max_{\partial B_r} |v| + \max_{\partial B_r} |u| \leq M + \max_{\partial B_r} |u|.$$

Then for each fixed $x \neq 0$,

$$|w(x)| \leq \frac{r^{n-2}}{|x|^{n-2}} M + \frac{1}{|x|^{n-2}} r^{n-2} \max_{\partial B_r} |u| \to 0 \text{ as } \ r \to 0.$$

This implies $w = 0$ in $B_R \setminus \{0\}$. □

4.5. Schauder Estimates

In this section, we discuss regularity of solutions of the Poisson equation. Let Ω be a domain in $\mathbb{R}^n$ and f be a continuous function in Ω. Then the Poisson equation has the form

$$\Delta u = f. \tag{4.10}$$

The Laplace operator Δ acts on C^2-functions and Δu is continuous for any C^2-function u. Conversely, we ask whether u is C^2 if f is continuous. The following example provides a negative answer.

EXAMPLE 4.38. For any $R < 1$, consider in $B_R \subset \mathbb{R}^2$

$$\Delta u = \frac{x_2^2 - x_1^2}{2|x|^2}\left\{\frac{4}{(-\log|x|)^{1/2}} + \frac{1}{2(-\log|x|)^{3/2}}\right\}. \tag{4.11}$$

By letting $f(x)$ be the function in the right hand side of (4.11), we note that f is continuous in $\bar{B}_R$ if we set it equal to zero at the origin. The function

$$u(x) = (x_1^2 - x_2^2)(-\log|x|)^{1/2} \in C(\bar{B}_R) \cap C^\infty(\bar{B}_R \setminus \{0\})$$

satisfies (4.11) in $B_R \setminus \{0\}$. Note that u is not in $C^2(B_R)$ since $\lim_{x\to 0} u_{x_1x_1} = \infty$. In fact, (4.11) has no C^2-solutions. Assume on the contrary that there exists a C^2-solution v in B_R. Then the function $w = u - v$ is harmonic and bounded in $B_R \setminus \{0\}$. By Theorem 4.37, w may be redefined at the origin so that $\triangle w = 0$ in B_R and therefore belongs to $C^2(B_R)$. In particular, the (finite) limit $\lim_{x\to 0} u_{x_1x_1}$ exists, which is a contradiction.

It turns out that u fails to be C^2 because there is no control on the module of continuity of f. If there is a better assumption than the continuity of f, we can control the module of continuity of $\nabla^2 u$. Here, controlling the module of continuity means imposing conditions on how fast functions approach constants. The simplest way to do this is known by Hölder norms.

In the following, we discuss the Schauder estimates for Poisson equations. Schauder estimates are among the most important results in the theory of elliptic partial differential equations. They give Hölder regularity estimates for solutions with Hölder continuous data and form the basis of general existence theorems. There are two classes of Schauder estimates, interior estimates and boundary estimates. In this section, we only discuss interior Schauder estimates for Poisson equations. In its simplest form, the interior Schauder estimate asserts that the C^α-norm of any second-order derivatives of u in a ball is estimated by the sum of the C^α-norm of Δu and the sup-norm of u in a larger ball. It is an optimal result in the sense that *u is as regular as the data allow*. Even though Δu is just one particular combination of second-order derivatives of u, the Hölder continuity of Δu implies the same regularity for all second-order derivatives.

We first define pointwise Hölder continuity.

DEFINITION 4.39. Let k be a positive integer and $\alpha \in (0, 1)$ be a constant. Suppose u is a function defined in B_1.

(1) A function u is C^α at 0 if

$$|u(x) - u(0)| \leq c|x|^\alpha \quad \text{for any } x \in B_1,$$

for a positive constant c. The Hölder semi-norm of u at 0 is defined by

$$[u]_{C^\alpha}(0) \equiv \sup_{|x|\leq 1} \frac{|u(x) - u(0)|}{|x|^\alpha}.$$

(2) A function u is $C^{k,\alpha}$ at 0 if there exists a polynomial $p(x)$ of degree k such that

$$|u(x) - p(x)| \leq c|x|^{k+\alpha} \quad \text{for any } x \in B_1,$$

for a positive constant c.

We note that the definition of Hölder continuity here gives a quantitative speed of how fast functions approach constant or a polynomial. We are only interested in asymptotic behavior of u as $x \to 0$. The ball B_1 can be replaced by any B_r for small r. This affects only the definition of Hölder semi-norms. The asymptotic behavior of u as $x \to 0$ remains the same.

We also note that the polynomial p in (2) is unique if it exists. We point out that u in Definition 4.39(2) is not necessarily C^k in a neighborhood of 0. If u is, then p is evidently the Taylor expansion of order k of u at 0. In particular, for $k = 2$,

$$\begin{aligned} p(x) =& u(0) + \nabla u(0) \cdot x + \frac{1}{2} x^T \nabla^2 u(0) x \\ =& u(0) + \sum_{i=1}^n u_{x_i}(0) x_i + \frac{1}{2} \sum_{i,j=1}^n u_{x_i x_j}(0) x_i x_j. \end{aligned}$$

We first state a global estimate on solutions in terms of nonhomogeneous terms and boundary values.

Lemma 4.40. *Suppose $u \in C^2(B_1) \cap C(\bar{B}_1)$ satisfies*

$$\Delta u = f \text{ in } B_1,$$

for some $f \in C(B_1)$. Then

$$\sup_{B_1} |u| \leq \sup_{\partial B_1} |u| + \frac{1}{2n} \sup_{B_1} |f|.$$

This is a special case of Theorem 4.32.

Now we begin to estimate second derivatives of solutions.

Lemma 4.41. *Suppose $u \in C^2(B_1)$ satisfies*

$$\Delta u = f \text{ in } B_1,$$

for some $f \in C(B_1)$. Then for any $\alpha \in (0,1)$, there exist constants $c_0 > 0, \mu \in (0,1)$ and $\varepsilon_0 > 0$, depending only on n and α, such that, if $|u| \leq 1$ and $|f| \leq \varepsilon_0$ in B_1, then there exists a second order harmonic polynomial

$$p(x) = \frac{1}{2} x^T A x + B \cdot x + C,$$

satisfying

$$|u(x) - p(x)| \leq \mu^{2+\alpha} \quad \text{for any } |x| \leq \mu,$$

and

$$|A| + |B| + |C| \le c_0.$$

Here, A is an $n \times n$ matrix, B is a vector and C is a constant.

PROOF. Without loss of generality, we assume $u \in C(\bar{B}_1)$. Suppose v is the harmonic function satisfying

$$\begin{aligned} \Delta v &= 0 \text{ in } B_1, \\ v &= u \text{ on } \partial B_1. \end{aligned}$$

The existence of v is guaranteed by Theorem 4.6. Clearly $|v| \le 1$ in B_1, by the maximum principle since $|u| \le 1$ on ∂B_1. Then Lemma 4.9 implies

$$\sum_{k=0}^{3} |\nabla^k v|_{L^\infty(B_{\frac{1}{2}})} \le c(n),$$

where $c(n)$ is a positive constant depending only on n, which is often called a universal constant. Hence the second order Taylor polynomial of v at 0

$$p(x) = \frac{1}{2} x^T A x + B \cdot x + C$$

has universal bounded coefficients and is harmonic. Now the function $u - v$ satisfies

$$\begin{aligned} \Delta(u - v) &= f \quad \text{in} \quad B_1, \\ u - v &= 0 \quad \text{on } \partial B_1. \end{aligned}$$

By Lemma 4.40, we have

$$|u(x) - v(x)| \le \frac{1}{2n} \sup_{B_1} |f| \quad \text{for any } x \in B_1.$$

By the mean value theorem, we have for any $x \in B_{1/2}$

$$\begin{aligned} |u(x) - p(x)| &\le |v(x) - p(x)| + \frac{1}{2n} \sup_{B_1} |f| \le c|x|^3 \sup_{B_{\frac{1}{2}}} |\nabla^3 v| + \frac{1}{2n} \sup_{B_1} |f| \\ &\le c(n)|x|^3 + \frac{1}{2n} \sup_{B_1} |f|. \end{aligned}$$

Now take μ small enough such that

$$c(n)\mu^{1-\alpha} \le \frac{1}{2},$$

and then take ε_0 such that

$$\varepsilon_0 \le \mu^{2+\alpha}.$$

We then have

$$|u(x) - p(x)| \le \frac{1}{2}\mu^{2+\alpha} + \frac{1}{2}\varepsilon_0 \le \mu^{2+\alpha} \quad \text{for any } x \in B_\mu.$$

This is the desired estimate. □

Now we prove the pointwise Schauder estimate.

THEOREM 4.42. *Suppose $u \in C^2(B_1)$ satisfies*

$$\Delta u = f \text{ in } B_1,$$

where f is Hölder continuous at 0. Then, u is $C^{2,\alpha}$ at 0. Moreover, there exists a second order polynomial

$$p(x) = \frac{1}{2}x^T Ax + Bx + C$$

satisfying

$$|u(x) - p(x)| \le c_0 |x|^{2+\alpha}\big(|u|_{L^\infty} + |f(0)| + [f]_{C^\alpha}(0)\big) \text{ for any } x \in B_1,$$

and

$$|A| + |B| + |C| \le c_0\big(|u|_{L^\infty} + |f(0)| + [f]_{C^\alpha}(0)\big),$$

where c_0 is a positive constant depending only on n and α.

Obviously, for $u \in C^2(B_1)$, we have

$$p(x) = u(0) + \nabla u(0) \cdot x + \frac{1}{2}x^T \nabla^2 u(0)x.$$

PROOF. Without loss of generality, we assume $f(0) = 0$. In general, we set $v = u - \frac{1}{2n}f(0)|x|^2$. Then $\Delta v = f - f(0)$ in B_1. Furthermore, we assume $|u| \le 1$ and $[f]_{C^\alpha}(0) \le \varepsilon_0$ for small $\varepsilon_0 > 0$. The general case can be recovered by considering

$$\frac{u}{|u|_{L^\infty} + \frac{1}{\varepsilon_0}[f]_{C^\alpha}(0)}.$$

First we claim that there exist harmonic polynomials for any $k = 1, 2, \cdots$,

$$P_k(x) = \frac{1}{2}x^T A_k x + B_k x + C_k,$$

satisfying

$$|u(x) - P_k(x)| \le \mu^{(2+\alpha)k} \quad \text{for any } |x| \le \mu^k,$$

and

$$|A_k - A_{k+1}| \le c\mu^{\alpha k},$$
$$|B_k - B_{k+1}| \le c\mu^{(\alpha+1)k},$$
$$|C_k - C_{k+1}| \le c\mu^{(\alpha+2)k},$$

where $\mu \in (0, 1)$ and c are positive constants depending only on n and α.

Note that the case $k = 1$ corresponds to Lemma 4.41. Assume it holds for some $k \ge 1$. Set

$$w(y) = \frac{(u - P_k)(\mu^k y)}{\mu^{(2+\alpha)k}} \quad \text{for any } y \in B_1.$$

Then

$$\Delta w(y) = \frac{f(\mu^k y)}{\mu^{\alpha k}} \quad \text{for any } y \in B_1,$$

with

$$\sup_{y \in B_1} \frac{|f(\mu^k y)|}{\mu^{\alpha k}} \le [f]_{C^\alpha}(0) \le \varepsilon_0.$$

By Lemma 4.41, there is a harmonic polynomial p_0 with bounded coefficients such that

$$|w(y) - p_0(y)| \le \mu^{2+\alpha} \quad \text{for } |y| \le \mu.$$

Now we scale back to get

$$|u(x) - P_k(x) - \mu^{(2+\alpha)k} p_0 \left(\frac{x}{\mu^k} \right)| \le \mu^{(k+1)(2+\alpha)} \quad \text{for } |x| \le \mu^{k+1}.$$

Clearly we proved the $(k+1)$-th step by letting

$$P_{k+1}(x) = P_k(x) + \mu^{(2+\alpha)k} p_0 \left(\frac{x}{\mu^k} \right).$$

It is easy to see that A_k, B_k and C_k converge and the limiting polynomial

$$p(x) = \frac{1}{2} x^T A_\infty x + B_\infty x + C_\infty$$

satisfies

$$|P_k(x) - p(x)| \le c\big(|x|^2 \mu^{\alpha k} + |x| \mu^{(\alpha+1)k} + \mu^{(\alpha+2)k}\big) \le c\mu^{(2+\alpha)k},$$

for any $|x| \le \mu^k$. Hence

$$|u(x) - p(x)| \le |u(x) - P_k(x)| + |P_k(x) - p(x)| \le c\mu^{(\alpha+2)k} \quad \text{for any } |x| \le \mu^k.$$

For any $x \in B_1$, there exists a nonnegative integer k such that $\mu^{k+1} \le |x| \le \mu^k$. This implies

$$|u(x) - p(x)| \le c|x|^{2+\alpha} \quad \text{for any } x \in B_1,$$

which yields the desired result. □

Next, we put Theorem 4.42 in the classical form of Schauder estimates. We first introduce Höder continuous functions in domains.

Definition 4.43. Let k be a nonnegative integer, $\alpha \in (0,1)$ be a constant and Ω be a domain in $\mathbb{R}^n$.

(1) A function u is C^α in Ω, denoted by $u \in C^\alpha(\Omega)$, if

$$|u(x) - u(y)| \le c|x-y|^\alpha \quad \text{for any } x, y \in \Omega,$$

for some constant c. The Hölder semi-norm of u in Ω is defined by

$$[u]_{C^\alpha(\Omega)} \equiv \sup_{x,y \in \Omega} \frac{|u(x) - u(y)|}{|x-y|^\alpha}.$$

(2) A function u is $C^{k,\alpha}$ in Ω, denoted by $u \in C^{k,\alpha}(\Omega)$, if $u \in C^k(\Omega)$ and $\nabla^k u \in C^\alpha(\Omega)$.

(3) The $C^{k,\alpha}$-norm in Ω is defined by

$$|u|_{C^{k,\alpha}(\Omega)} = |u|_{C^k(\Omega)} + [\nabla^k u]_{C^\alpha(\Omega)}.$$

In Definition 4.43, we may replace Ω by $\bar{\Omega}$ and define $C^{k,\alpha}(\bar{\Omega})$ similarly. It is easy to check that $C^{k,\alpha}(\bar{\Omega})$ is a Banach space equipped with the $C^{k,\alpha}$-norm.

For $u \in C^\alpha(\Omega)$, it is easy to see that u is C^α at x for any $x \in \Omega$ as in Definition 4.39. In fact, the converse is also true. If u is C^α at x for any $x \in \Omega$ as in Definition 4.39, then $u \in C^\alpha(\Omega)$. In other words, the (global) Hölder continuity is equivalent to the everywhere pointwise Hölder continuity. We will leave verification as an exercise.

The following result is often referred to as the interior Schauder regularity or the interior Schauder estimate.

THEOREM 4.44. *Let $\alpha \in (0,1)$ be a constant and $u \in C^2(B_1)$ satisfy*

$$\Delta u = f \text{ in } B_1.$$

If $f \in C^\alpha(B_1)$, then $u \in C^{2,\alpha}(B_1)$. Moreover,

$$|u|_{C^{2,\alpha}(B_{\frac{1}{2}})} \le c\big(|u|_{L^\infty(B_1)} + |f|_{C^\alpha(B_1)}\big),$$

where c is a positive constant depending only on n and α.

PROOF. We only sketch the proof. For any $x \in B_{1/2}$, we apply Theorem 4.42 in $B_{(1-|x|)}(x)$ to obtain a quadratic polynomial p_x such that

the absolute values of coefficients of p_x are bounded by $c(|u|_{L^\infty(B_1)} + |f|_{C^\alpha(B_1)})$

and

$$|u(y) - p_x(y-x)| \le c(|u|_{L^\infty(B_1)} + |f|_{C^\alpha(B_1)})|y-x|^{2+\alpha} \quad \text{for any } y \in B_{\frac{1-|x|}{2}}(x),$$

where c is a positive constant depending only on n and α. Then we have the desired result easily by the equivalence of the (global) Hölder continuity and the everywhere pointwise Hölder continuity. □

To end this section, we mention briefly another approach for Schauder estimates. Let Ω be a domain in $\mathbb{R}^n$ and f be a continuous function in Ω. We consider the Poisson equation (4.10) in Ω.

If $u \in C_0^\infty(\Omega)$ is a solution of (4.10), then f evidently has a compact support in Ω and hence by Theorem 4.1

$$u(x) = \int_\Omega \Gamma(x,y) f(y)\,dy,$$

where $\Gamma(x,y)$ is the fundamental solution of Δ in $\mathbb{R}^n$.

Now let Ω be a bounded domain in $\mathbb{R}^n$ and f be a bounded function in Ω. We define

$$u_f(x) = \int_\Omega \Gamma(x,y) f(y)\,dy. \tag{4.12}$$

This is called the potential of f in Ω. If u_f is C^2 and satisfies (4.10) in Ω, i.e., $\Delta u_f = f$ in Ω, then any solution of (4.10) differs from u_f by an addition of a harmonic function. Since harmonic functions are analytic, regularity of arbitrary solutions of (4.10) is determined by those of u_f defined in (4.12). The classical approach for the Schauder theory is to prove $u_f \in C^{2,\alpha}(\Omega)$ and $\Delta u_f = f$ in Ω if $f \in C^\alpha(\Omega)$ for some $\alpha \in (0,1)$. We will not provide proofs here. Example 4.38 shows that u_f is not necessarily C^2 if f is only continuous.

4.6. Weak Solutions

In this section, we discuss briefly how to extend the notion of classical solutions of the Poisson equation to less regularized solutions, solutions with derivatives only in integral sense, the so-called weak solutions. The same process can be applied to general linear elliptic equations of second order, even nonlinear elliptic equations, of divergence form. This section serves only as an introduction to this important advanced topic in the theory of elliptic differential equations. A complete presentation will constitute a book much thicker than this one.

To introduce weak solutions, we make use of divergence structure or variation structure of the Laplace operator. Namely, we write the Laplace operator as

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

Then the Laplace equation can be considered as an Euler equation of the Dirichlet energy.

For any domain $\Omega \subset \mathbb{R}^n$, we define the Dirichlet energy of u in Ω by

$$E(u) = \frac{1}{2}\int_\Omega |\nabla u|^2.$$

Suppose u is a C^1-minimizer of $E(u)$ among all C^1-functions in Ω with the same boundary value. In particular, a small perturbation of u in any subregion of Ω increases the energy. Hence, for any C^1-function φ with a compact support in Ω,

$$\frac{1}{2}\int_\Omega |\nabla(u+\varepsilon\varphi)|^2 \geq \frac{1}{2}\int_\Omega |\nabla u|^2 \quad \text{for any } \varepsilon.$$

In other words,

$$F(\varepsilon) \equiv \frac{1}{2}\int_\Omega |\nabla(u+\varepsilon\varphi)|^2$$

has a minimum at $\varepsilon = 0$. This implies $F'(0) = 0$, or

$$\int_\Omega \nabla u \cdot \nabla\varphi = 0 \quad \text{for any } \varphi \in C_0^1(\Omega).$$

In general, for any $f \in L^2(\Omega)$, we consider

$$J(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 + \int_\Omega uf. \tag{4.13}$$

If u is a C^1-minimizer, we obtain similarly

$$-\int_\Omega \nabla u \cdot \nabla\varphi = \int_\Omega f\varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

This suggests the following definition.

Definition 4.45. Let $f \in L^2(\Omega)$. Then $u \in C^1(\Omega)$ is a weak solution of $\Delta u = f$ in Ω if

$$-\int_\Omega \nabla u \cdot \nabla\varphi = \int_\Omega f\varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

In Definition 4.45, φ is called a test function. If u is a weak C^2-solution, we obtain by a simple integration by parts

$$\int_\Omega \varphi\Delta u = \int_\Omega f\varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

This implies easily that

$$\Delta u = f \quad \text{a.e. in } \Omega.$$

If f is continuous in Ω, then $\Delta u = f$ in Ω. In other words, a weak C^2-solution of the Poisson equation is a classical solution.

Next, we illustrate how to find a weak solution of the Poisson equation with a homogeneous Dirichlet boundary value, i.e.,

$$\begin{aligned} \Delta u =& f \quad \text{in } \Omega, \\ u =& 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.14}$$

We intend to minimize $J(u)$ given in (4.13) in the space

$$\mathcal{C} = \{u \in C^1(\Omega) \cap C(\bar{\Omega});\ u = 0 \text{ on } \partial\Omega\}.$$

We first demonstrate that J has a lower bound in $\mathcal{C}$. First, by the Cauchy inequality, we have

$$\int_\Omega |uf| \leq \big(\int_\Omega u^2\big)^{\frac{1}{2}}\big(\int_\Omega f^2\big)^{\frac{1}{2}}.$$

By Lemma 3.12, the Poincarè inequality, we have for any $u \in \mathcal{C}$

$$\int_\Omega u^2 \leq C\int_\Omega |\nabla u|^2,$$

where C is a positive constant depending only on Ω. Then

$$\int_\Omega |uf| \leq \sqrt{C}\big(\int_\Omega |\nabla u|^2\big)^{\frac{1}{2}}\big(\int_\Omega f^2\big)^{\frac{1}{2}} \leq \frac{1}{4}\int_\Omega |\nabla u|^2 + C\int_\Omega f^2.$$

Hence for any $u \in \mathcal{C}$

$$J(u) \geq \frac{1}{4}\int_\Omega |\nabla u|^2 - C\int_\Omega f^2, \tag{4.15}$$

and in particular

$$J(u) \geq -C\int_\Omega f^2.$$

Therefore, J has a lower bound in $\mathcal{C}$. We set

$$J_0 = \inf\{J(u);\ u \in \mathcal{C}\}.$$

If J_0 is realized by some $u \in \mathcal{C}$, then u is a minimizer of J and hence a weak solution of $\Delta u = f$.

To discuss whether J realizes J_0 in $\mathcal{C}$, we consider a minimizing sequence $\{u_k\} \subset \mathcal{C}$ with $J(u_k) \to J_0$ as $k \to \infty$ and ask whether there exists a function $u \in \mathcal{C}$ such that $J(u_k) \to J(u)$. If such a u exists, then J_0 is realized by u by the convergence of $J(u_k)$. Here we do not require that $\{u_k\}$ converge to u in the C^1-norm. However, in order to construct u, $\{u_k\}$ has to converge in some sense. So what is a reasonable norm to provide such a convergence? To answer this question, we examine the expression of $J(u)$ in (4.13). First, we have by (4.15)

$$\int_\Omega |\nabla u_k|^2 \leq 4J(u_k) + 4C\int_\Omega f^2.$$

Since $J(u_k)$ is bounded, there is a uniform bound on the L^2-norms of gradients of u_k. Hence a convergence of u_k, if exists, should be related to the L^2-norm. For this, we need to introduce Sobolev norms.

For any $u \in C^1_0(\Omega)$, we define the H^1_0-norm of u by

$$\|u\|_{H^1_0} = \big(\int_\Omega |\nabla u|^2\big)^{\frac{1}{2}}.$$

Here in H_0^1, the super-index 1 indicates the order of differentiation and the sub-index 0 refers to functions with compact support in Ω. We then complete the space $C_0^1(\Omega)$ with the H_0^1-norm. The resulting space is called the Sobolev space $H_0^1(\Omega)$ and its elements are called H_0^1-functions in Ω. By the Poincarè inequality, we can prove $H_0^1(\Omega) \subset L^2(\Omega)$. In general, functions in $H_0^1(\Omega)$ may not have classical derivatives. However, it has weak derivatives in an integral sense. For any $u \in H_0^1(\Omega)$ and any $i = 1, \cdots, n$, there exists a function $v_i \in L^2(\Omega)$ such that

$$\int_\Omega u\varphi_{x_i} = -\int_\Omega v_i\varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

Here v_i is called a weak x_i-derivative of u and is often denoted by u_{x_i}, the same notion of classical derivatives. Weak derivatives are unique by the density of continuous functions in the L^2-space. Hence, $\|u\|_{H_0^1}$ can be computed for any $u \in H_0^1(\Omega)$ by using weak derivatives of u. By a simple integration by parts, classical derivatives of C^1-functions are weak derivatives. We also point out that $H_0^1(\Omega)$ is in fact a Hilbert space with the inner product

$$(u, v) = \int_\Omega \nabla u \cdot \nabla v \quad \text{for any } u, v \in H_0^1(\Omega),$$

where ∇u and ∇v are weak gradients of u and v.

The following compactness result, called Rellich's Theorem, plays an important role in the theory of Sobolev spaces. *Any sequence in $H_0^1(\Omega)$ with bounded $H_0^1(\Omega)$-norms has a subsequence convergent in the $L^2(\Omega)$-norm to some function in $H_0^1(\Omega)$.* We emphasize that the limit function is in $H_0^1(\Omega)$ although the convergence is only in L^2. In general, we do not have a convergence in the H_0^1-norm.

It is easy to see that the functional $J(u)$ in (4.13) can be extended to functions $u \in H_0^1(\Omega)$, with classical gradients replaced by weak gradients. Instead of minimizing J in $C_0^1(\Omega)$, we minimize J in $H_0^1(\Omega)$. One advantage of $H_0^1(\Omega)$ over $C_0^1(\Omega)$ is that $H_0^1(\Omega)$ is complete under the $H_0^1(\Omega)$-norm, which is more naturally associated with the functional $J(u)$ in (4.13). Another important aspect of the Sobolev space $H_0^1(\Omega)$ in studies of elliptic differential equations lies in the following simple fact. *The concept of weak solutions in Definition 4.45 can be extended to H_0^1-functions.* We also point out that the Poincarè inequaltiy holds for $H_0^1(\Omega)$-functions by a simple approximation of $H_0^1(\Omega)$-functions by $C_0^1(\Omega)$-functions. Hence, (4.15) also holds for $H_0^1(\Omega)$-functions.

We return to our minimizing problem. First, it is easy to see

$$J_0 = \inf_{u \in H_0^1(\Omega)} J(u).$$

Let $\{u_k\}$ be a minimizing sequence of $J(u)$ in $H_0^1(\Omega)$. Then (4.15) shows that $\|u_k\|_{H_0^1}$ is uniformly bounded. By Rellich's theorem, there exists a subsequence $\{u_{k'}\}$ and a $u \in H_0^1(\Omega)$ such that

$$u_{k'} \to u \quad \text{in the } L^2\text{-norm as } k' \to \infty.$$

Next, with the help of the Hilbert space structure of $H_0^1(\Omega)$ it is not difficult to prove

$$J(u) \leq \liminf_{k' \to \infty} J(u_k').$$

This implies $J(u) = J_0$. In fact, we can also prove $u_{k'} \to u$ in the $H_0^1(\Omega)$-norm as $k' \to \infty$. Now we try to interpret in what sense u solves (4.14). First, u is a minimizer of J in $H_0^1(\Omega)$, i.e.,

$$J(u+\varphi) \geq J(u) \quad \text{for any } \varphi \in C_0^1(\Omega).$$

Then by the same analysis preceding Definition 4.45, we have

$$-\int_\Omega \nabla u \cdot \nabla \varphi = \int_\Omega f\varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

Here ∇u is the weak gradient of u. Hence, u is a weak solution of $\Delta u = f$. Concerning the boundary value, we point out that u is not defined in the pointwise sense. We cannot conclude $u = 0$ at each point on $\partial\Omega$. Here the boundary condition $u = 0$ on $\partial\Omega$ is interpreted precisely by the fact $u \in H_0^1(\Omega)$, i.e., u is the limit of a sequence of $C_0^1(\Omega)$-functions in the $H_0^1(\Omega)$-norm. One consequence is that $u|_{\partial\Omega}$ is a well-defined zero function in $L^2(\partial\Omega)$. This finishes the existence part of the variational approach. We have found a weak solution of (4.14), the Poisson equation with a homogeneous Dirichlet boundary value.

Now we ask whether u possesses a better regularity. The answer is yes. With $f \in L^2(\Omega)$, the solution u has weak derivatives of second order in the following sense. For any $1 \leq i, j \leq n$, there exists a function $v_{ij} \in L^2(\Omega)$ such that

$$\int_\Omega u\varphi_{x_i x_j} = \int_\Omega v_{ij}\varphi \quad \text{for any } \varphi \in C_0^2(\Omega).$$

We also denote v_{ij} by $u_{x_i x_j}$. The solution u is a so-called H^2-function in Ω. With weak derivatives $u_{x_i x_j}$ defined as L^2-functions, we also conclude

$$\Delta u = f \quad \text{a.e. in } \Omega.$$

In fact, if f is an H^k-function for any $k \geq 1$, a function with weak derivatives of order up to k as L^2-functions, then u is an H^{k+2}-function. In particular, if f is smooth, then u is smooth. We omit details of discussions.

With similar methods, we can solve the eigenvalue problem of the Laplace operator in a smooth bounded domain Ω in $\mathbb{R}^n$. Consider

$$\begin{aligned} -\Delta u =& \lambda u \quad \text{in } \Omega, \\ u =& 0 \quad \text{on } \partial\Omega. \end{aligned}$$

First, set

$$\lambda_1 = \inf\Big\{ \frac{\int_\Omega |\nabla u|^2}{\int_\Omega u^2}; u \in H_0^1(\Omega) \setminus \{0\} \Big\}.$$

With a minimizing sequence, we can prove that λ_1 is achieved in $H_0^1(\Omega)$, say by $u_1 \in H_0^1(\Omega)$. We further assume $\int_\Omega u_1^2 = 1$. Then for any $\varphi \in C_0^1(\Omega)$,

$$F(\varepsilon) = \frac{\int_\Omega |\nabla u_1 + \varepsilon \nabla \varphi|^2}{\int_\Omega (u_1 + \varepsilon\varphi)^2}$$

has a minimum at $\varepsilon = 0$. Hence $F'(0) = 0$, or

$$\int_\Omega \nabla u_1 \cdot \nabla \varphi = \lambda_1 \int_\Omega u_1 \varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

It is easy to check that $\lambda_1 > 0$. Here λ_1 is the first eigenvalue of the Laplace operator and u_1 is a corresponding eigenfunction in the weak sense. In fact, we can prove u_1 is a smooth function in $\bar{\Omega}$ and satisfies

$$\begin{aligned} -\Delta u_1 =& \lambda_1 u_1 \quad \text{in } \Omega, \\ u_1 =& 0 \quad \text{on } \partial\Omega \end{aligned}$$

in the classical sense. The definition of λ_1 yields

$$\frac{\int_\Omega |\nabla u|^2}{\int_\Omega u^2} \geq \lambda_1 \quad \text{for any } u \in H_0^1(\Omega),$$

or

$$\int_\Omega u^2 \leq \frac{1}{\lambda_1} \int_\Omega |\nabla u|^2 \quad \text{for any } u \in H_0^1(\Omega).$$

Therefore, $1/\lambda_1$ is the best constant in the Poincarè inequality. See Lemma 3.12.

To find the next eigenvalue, we set

$$\lambda_2 = \inf \Big\{ \frac{\int_\Omega |\nabla u|^2}{\int_\Omega u^2}; u \in H_0^1(\Omega) \setminus \{0\}, \ \int_\Omega u u_1 = 0 \ \Big\}.$$

Obviously, $\lambda_2 \geq \lambda_1$. With a similar minimizing process, we can prove that λ_2 is achieved in $H_0^1(\Omega)$, say by $u_2 \in H_0^1(\Omega)$ with $\int_\Omega u_1 u_2 = 0$. We further assume $\int_\Omega u_2^2 = 1$. We can check similarly

$$\int_\Omega \nabla u_2 \cdot \nabla \varphi = \lambda_2 \int_\Omega u_2 \varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

Next, we set

$$\lambda_3 = \inf \Big\{ \frac{\int_\Omega |\nabla u|^2}{\int_\Omega u^2}; u \in H_0^1(\Omega) \setminus \{0\}, \ \int_\Omega u u_i = 0, i = 1, 2 \ \Big\}.$$

By continuing this process, we obtain an increasing sequence $\{\lambda_i\}$ and a sequence $\{u_i\} \subset H_0^1(\Omega)$ with $\int_\Omega u_i u_j = \delta_{ij}$ such that

$$\int_\Omega \nabla u_i \cdot \nabla \varphi = \lambda_i \int_\Omega u_i \varphi \quad \text{for any } \varphi \in C_0^1(\Omega).$$

We can prove that u_i in fact is a smooth function in $\bar{\Omega}$ for any $i = 1, 2, \cdots$, and satisfies

$$\begin{aligned} -\Delta u_i =& \lambda_i u_i \quad \text{in } \Omega, \\ u_i =& 0 \quad \text{on } \partial\Omega \end{aligned}$$

in the classical sense. Moreover, $\{u_i\}$ forms a complete sequence in $L^2(\Omega)$. In other words, for any $v \in L^2(\Omega)$

$$v = \sum_{i=1}^{\infty} (v, u_i)_{L^2(\Omega)} u_i,$$

where $(\cdot, \cdot)_{L^2(\Omega)}$ is the L^2-inner product in Ω and the convergence of this series is in the L^2-norm in Ω.

With such a sequence of eigenvalues and eigenfunctions in a bounded smooth domain, we can solve initial/boundary-value problems of the heat equation and the wave equation in arbitrary dimension, as described at the end of Section 3.3.

Exercises

(1) Suppose $u(x)$ is harmonic in some domain in $\mathbb{R}^n$. Then

$$v(x) = |x|^{2-n} u\big(\frac{x}{|x|^2}\big)$$

is also harmonic in a suitable domain.

(2) For $n = 2$, find the Green's function for the Laplace operator on the first quadrant.

(3) Find the Green's function for the Laplace operator in the upper half space $\{x_n > 0\}$ and then derive a formal integral representation for a solution of the Dirichlet problem

$$\begin{aligned} \Delta u = 0 \quad & \text{in } \{x_n > 0\}, \\ u = \varphi \quad & \text{on } \{x_n = 0\}. \end{aligned}$$

(4) Let λ be a positive constant. Find the fundamental solution for $\Delta_3 u + \lambda u = 0$ in $\mathbb{R}^3$.

(5) (a) Prove by the Poisson formula the following Harnack inequality: Suppose u is harmonic in $B_R(x_0)$ and $u \geq 0$. Then

$$\left(\frac{R}{R+r}\right)^{n-2} \frac{R-r}{R+r} u(x_0) \leq u(x) \leq \left(\frac{R}{R-r}\right)^{n-2} \frac{R+r}{R-r} u(x_0),$$

where $r = |x - x_0| < R$.

(b) Prove by (a) the Liouville Theorem: If u is a harmonic function in $\mathbb{R}^n$ and bounded above or below, then $u \equiv$ const.

(6) Let u be a harmonic function in $\mathbb{R}^n$ with $\int_{\mathbb{R}^n} |u|^p < \infty$ for some $p \in (1, \infty)$. Then $u \equiv 0$.

(7) Let u be a harmonic function in $\mathbb{R}^n$ and $u(x) = O(|x|^m)$ as $|x| \to \infty$ for a positive integer m. Then u is a polynomial of degree m.

(8) Suppose $u \in C(\bar{B}_1^+)$ is harmonic in $B_1^+ = \{x \in B_1; x_n > 0\}$ with $u = 0$ on $\{x_n = 0\} \cap B_1$. Then the odd extension of u in B_1 is harmonic in B_1.

(9) Suppose u is harmonic in the open upper half space $\{x_n > 0\}$ and continuous in its closure. If $u = 0$ on $\{x_n = 0\}$ and $u(x) = O(|x|)$ as $|x| \to \infty$, then $u(x) = cx_n$ for some constant c.

(10) Let Ω be a domain in $\mathbb{R}^n$ and $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfies $\Delta u + \lambda u = 0$ in Ω for a constant $\lambda > 0$. Then u is analytic in Ω.

(11) Suppose Ω is a domain in $\mathbb{R}^n$ and $\{u_m\}$ is a sequence of uniformly bounded harmonic functions in Ω. Prove that $\{u_m\}$ has a subsequence convergent to a function u uniformly on any compact subsets of Ω and that u is harmonic in Ω.

(12) Prove that the square of a nonnegative subharmonic function is subharmonic. Give an example to show that the condition *nonnegative* cannot be omitted.

(13) Suppose u is harmonic in the open upper half space $\{x_n > 0\}$ and continuous and bounded in its closure. Then

$$\sup_{\{x_n \geq 0\}} u = \sup_{\{x_n = 0\}} u.$$

(14) Let u be a C^2-solution of

$$\begin{aligned} \Delta u &= 0 \quad \text{in } \mathbb{R}^n \setminus B_R, \\ u &= 0 \quad \text{on } \partial B_R. \end{aligned}$$

Prove that $u \equiv 0$ if

$$\begin{aligned} \lim_{|x|\to\infty} \frac{u(x)}{\ln |x|} &= 0 \quad \text{for } n = 2, \\ \lim_{|x|\to\infty} u(x) &= 0 \quad \text{for } n \geq 3. \end{aligned}$$

(15) Suppose $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ satisfies

$$\begin{aligned} -\Delta u + u^3 &= 0 \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} + \alpha u &= \varphi \quad \text{on } \partial\Omega, \end{aligned}$$

for continuous functions φ and α on $\partial\Omega$ with $\alpha(x) \geq \alpha_0 > 0$. Prove

$$\max_{\Omega} |u| \leq \frac{1}{\alpha_0} \max_{\partial\Omega} |\varphi|.$$

(16) Let Ω be a smooth bounded domain in $\mathbb{R}^n$ with $\partial\Omega = \Gamma_1 \cup \Gamma_2$, c be a continuous function in Ω with $c \leq 0$ and α be a continuous function on $\partial\Omega$ with $\alpha \geq 0$. Discuss the uniqueness of the following problem

$$\begin{aligned} \Delta u + cu &= f \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} + \alpha u &= \varphi \quad \text{on } \Gamma_1, \\ u &= \psi \quad \text{on } \Gamma_2. \end{aligned}$$

(17) Let k and $\lambda > 0$ be constants and B_1^+ be the upper half disc in $\mathbb{R}^2$. Suppose $u \in C^2(B_1^+) \cap C(\bar{B}_1^+)$ is a solution of the following problem

$$\begin{aligned} \Delta u + \frac{k}{y} u_y - \lambda u &= f \quad \text{in } B_1, \\ u &= 0 \quad \text{on } \partial B_1^+. \end{aligned}$$

Then

$$\max_{\bar{B}_1^+} |u| \leq \frac{1}{\lambda} \max_{\bar{B}_1^+} |f|.$$

(18) Let u be a nonzero harmonic function in $B_1 \subset \mathbb{R}^n$ and set

$$N(r) = \frac{r \int_{B_r} |\nabla u|^2}{\int_{\partial B_r} u^2} \quad \text{for any } r \in (0,1).$$

(a) Prove $N(r)$ is a nondecreasing function in $r \in (0,1)$ and identify

$$\lim_{r \to 0+} N(r).$$

(b) Prove for any $0 < r < R < 1$

$$\frac{1}{R^{n-1}} \int_{\partial B_R} u^2 \leq \left(\frac{R}{r}\right)^{2N(R)} \frac{1}{r^{n-1}} \int_{\partial B_r} u^2.$$

Remark: The quantity $N(r)$ is called the *frequency*. The estimate in (b) is referred to as the *doubling condition* for $R = 2r$.

(19) Let k be a nonnegative integer, $\alpha \in (0, 1)$ be a constant and u be a C^k-function in $[0, 1]$. Suppose for any $x \in [0, 1]$

$$|u(y) - T_k(y; x)| \leq C|y - x|^{k+\alpha} \quad \text{for any } y \in [0, 1],$$

where $T_k(\cdot; x)$ is the k-th Taylor expansion of u at x. Then $u \in C^{k,\alpha}[0, 1]$ and $[\nabla^k u]_{C^\alpha[0,1]} \leq C$.

CHAPTER 5

Heat Equations

The n-dimensional heat equation is given by $u_t - \Delta u = 0$ for functions $u = u(x,t)$, with $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Here, x is called the space variable and t the time variable. The heat equation models the temperature of a body conducting heat, when the density is constant. In this chapter, we discuss properties of solutions of the heat equation. In Section 5.1, we discuss the fundamental solution of the heat equation and solve initial-value problems. Then in Section 5.2, we use the fundamental solution to discuss regularity of arbitrary solutions. In Section 5.3, we discuss the maximum principle for the heat equation and its applications. In particular, we use the maximum principle to derive interior gradient estimates. In Section 5.4, we discuss Harnack inequalities.

5.1. Initial-Value Problems

In this section, we discuss initial-value problems of the heat equation and derive an explicit expression for their solutions.

We first examine the equation itself. The n-dimensional heat equation is given by

$$u_t - \Delta u = 0, \tag{5.1}$$

for $u = u(x,t)$ with $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. We note that (5.1) is not preserved by the change $t \mapsto -t$. This indicates that the heat equation describes an irreversible process and distinguishes between past and future. This fact will be well illustrated by Harnack inequalities which we will derive in Section 5.4. Next, (5.1) is preserved under linear transforms $x' = \lambda x$ and $t' = \lambda^2 t$ for any nonzero constant λ, which leave the quotient $|x|^2/t$ invariant. Because of this, the expression $|x|^2/t$ appears frequently in connection with the heat equation (5.1). In fact, the fundamental solution we will derive has such an expression.

If u is a solution of (5.1) in a domain in $\mathbb{R}^n \times \mathbb{R}$, then for any (x_0, t_0) in this domain and any $r > 0$

$$\tilde{u}(x,t) = u(x_0 + rx, t_0 + r^2 t)$$

is a solution of (5.1) in an appropriate domain in $\mathbb{R}^n \times \mathbb{R}$.

Let u be a polynomial in $\mathbb{R}^n \times \mathbb{R}$. It is of p-homogeneous degree d if

$$u(\lambda x, \lambda^2 t) = \lambda^d u(x,t) \quad \text{for any } (x,t) \in \mathbb{R}^n \times \mathbb{R} \text{ and } \lambda > 0.$$

Now we let P be a homogeneous polynomial of degree d in $\mathbb{R}^n$ and proceed to find a p-homogeneous polynomial of degree d such that

$$\begin{aligned} u_t - \Delta u &= 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}, \\ u(\cdot, 0) &= P \quad \text{in } \mathbb{R}^n. \end{aligned}$$

To do this, we simply expand u as a power function of t with coefficients as functions of x, i.e.,

$$u(x,t)=\sum_{k=0}^{\infty}a_k(x)t^k.$$

Then a simple calculation shows that

$$a_0=P,\quad a_k=\frac{1}{k}\Delta a_{k-1}\quad\text{for any } k\geq 1.$$

Since P is a polynomial of degree d, then $\Delta^{[d/2]+1}P=0$, where $[d/2]$ is the integral part of $d/2$, i.e., $[d/2]=d/2$ if d is an even integer and $[d/2]=(d-1)/2$ if d is an odd integer. Hence

$$u(x,t)=\sum_{k=0}^{[\frac{d}{2}]}\Delta^k P(x)t^k.$$

For $n=1$, let u_d be a p-homogeneous polynomial of degree d in $\mathbb{R}\times\mathbb{R}$ satisfying the heat equation and $u_d(x,0)=x^d$. The first five such polynomials are given by

$$\begin{aligned}
u_1(x,t)=&x,\\
u_2(x,t)=&x^2+2t,\\
u_3(x,t)=&x^3+6xt,\\
u_4(x,t)=&x^4+12x^2t+12t^2,\\
u_5(x,t)=&x^5+20x^3t+60xt^2.
\end{aligned}$$

In the following, we discuss initial-value problem of the heat equation and derive an explicit expression for its solution.

We first introduce briefly Fourier transforms. Fourier transforms are an important subject and has a close connection with many fields in mathematics, especially with partial differential equations.

By allowing functions to be complex valued, we define the Schwartz class $\mathcal{S}$ as the collection of all functions $u\in C^\infty(\mathbb{R}^n;\mathbb{C})$ such that

$$(1+|x|^2)^{\frac{m}{2}}\partial_x^\alpha u(x)\text{ is bounded in }\mathbb{R}^n\text{ for any }\alpha\in\mathbb{Z}_+^n\text{ and }m\in\mathbb{Z}_+.$$

In other words, the Schwartz class consists of smooth functions whose arbitrary derivatives decay faster than any polynomials. For example, $u(x)=e^{-|x|^2}$ is in the Schwartz class.

For any $u\in\mathcal{S}$, define a new function $\hat{u}$ in $\mathbb{R}^n$ by

$$\hat{u}(\xi)=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-ix\cdot\xi}u(x)dx\quad\text{for any }\xi\in\mathbb{R}^n.$$

The operator $u\mapsto\hat{u}$ is called the Fourier transform operator and $\hat{u}$ is called the Fourier transform of u.

We now state some properties of Fourier transforms.

(F1) Linearity: The Fourier transform operator is linear, i.e., for any $u_1,u_2\in\mathcal{S}$ and $c_1,c_2\in\mathbb{C}$

$$(cu_1+c_2u_2)^\wedge=c_1\hat{u}_1+c_2\hat{u}_2.$$

(F2) Differentiation: For any $u \in \mathcal{S}$ and any $j = 1, \cdots, n$, we have by integration by parts

$$\widehat{\partial_j u}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \partial_{x_j} u(x) dx = i\xi_j \hat{u}(\xi),$$

and hence for any multi-index $\alpha \in \mathbb{Z}_+^n$

$$\widehat{\partial_x^\alpha u}(\xi) = (i\xi)^\alpha \hat{u}(\xi).$$

For any polynomial $P(\xi)$ on $\mathbb{R}^n$ given by

$$P(\xi) = \sum_{|\alpha| \le m} a_\alpha \xi^\alpha,$$

we define a differential operator $P(\partial)$ by

$$P(\partial)u = \sum_{|\alpha| \le \alpha} a_\alpha \partial^\alpha u.$$

Then

$$\widehat{P(\partial)u}(\xi) = P(i\xi)\hat{u}(\xi).$$

(F3) Multiplication by polynomials: It is easy to see that $\hat{u} \in C^\infty(\mathbb{R}^n)$ for any $u \in \mathcal{S}$. Then for any $j = 1, \cdots, n$

$$\partial_{\xi_j} \hat{u}(\xi) = -\frac{i}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} x_j u(x) dx = -i\widehat{x_j u}(\xi),$$

and hence for any multi-index $\beta \in \mathbb{Z}_+^n$

$$\partial_\xi^\beta \hat{u}(\xi) = (-i)^{|\beta|} \widehat{x^\beta u}(\xi).$$

As a consequence, we have $\hat{u} \in \mathcal{S}$ for $u \in \mathcal{S}$, i.e., $|\xi^\alpha \partial_\xi^\beta \hat{u}(\xi)|$ is bounded in $\mathbb{R}^n$ for any $\alpha, \beta \in \mathbb{Z}_+^n$. To see this, we note by (F2) and (F3)

$$\begin{aligned}
\xi^\alpha \partial_\xi^\beta \hat{u}(\xi) =& (-i)^{|\beta|} \xi^\alpha \widehat{x^\beta u}(\xi) = (-i)^{|\alpha|+|\beta|} (i\xi)^\alpha \widehat{x^\beta u}(\xi) = (-i)^{|\alpha|+|\beta|} \widehat{\partial_x^\alpha (x^\beta u)}(\xi) \\
=& \frac{1}{(2\pi)^{\frac{n}{2}}} (-i)^{|\alpha|+|\beta|} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \partial_x^\alpha \big(x^\beta u(x)\big) dx.
\end{aligned}$$

Hence

$$\sup_{\xi \in \mathbb{R}^n} |\xi^\alpha \partial_\xi^\beta \hat{u}(\xi)| \le \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |\partial_x^\alpha \big(x^\beta u(x)\big)| dx < \infty,$$

since each term in the integrand decays faster than any polynomials.

(F4) Translation and dilation: For any $u \in \mathcal{S}$, $a \in \mathbb{R}^n$ and $k \in \mathbb{R}$,

$$\widehat{u(\cdot - a)}(\xi) = e^{-i\xi\cdot a} \hat{u}(\xi),$$

and

$$\widehat{u(k\cdot)}(\xi) = \frac{1}{|k|^n} \hat{u}(\frac{\xi}{k}).$$

(F5) Convolution: For any $u, v \in \mathcal{S}$, it is easy to check $u * v \in \mathcal{S}$, where $u * v$ is the convolution of u and v defined by

$$(u * v)(x) = \int_{\mathbb{R}^n} u(x - y) v(y) dy.$$

We then have

$$\widehat{u * v}(\xi) = (2\pi)^{\frac{n}{2}} \hat{u}(\xi) \hat{v}(\xi).$$

We leave verification of Properties (F1)-(F5) as an exercise.

The following example will be useful in discussions of the heat equation. We first note

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

EXAMPLE 5.1. Consider $u(x) = e^{-|x|^2}$ in $\mathbb{R}^n$. Then $u \in \mathcal{S}$ and a direct calculation yields

$$\hat{u}(\xi) = \frac{1}{2^{\frac{n}{2}}} e^{-\frac{1}{4}|\xi|^2}.$$

Therefore by (F4), for $f(x) = e^{-A|x|^2}$ with $A > 0$,

$$\hat{f}(\xi) = \frac{1}{(2A)^{\frac{n}{2}}} e^{-\frac{|\xi|^2}{4A}}.$$

The Fourier transform is an important subject in mathematics. Now we attempt to use it to solve certain classes of partial differential equations. Let P be a polynomial in $\mathbb{R}^n$ of degree m. Consider the following linear partial differential equation of degree m

$$P(\partial)u = f \quad \text{in } \mathbb{R}^n.$$

By applying Fourier transforms and (F2), we have

$$P(i\xi)\hat{u}(\xi) = \hat{f}(\xi),$$

and then

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{P(i\xi)}.$$

Hence a solution u is a function whose Fourier transform is $\hat{f}(\xi)/P(i\xi)$. This naturally leads to inverse Fourier transforms.

THEOREM 5.2. *Suppose $u \in \mathcal{S}$. Then*

$$u(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \hat{u}(\xi) d\xi.$$

The right hand side is simply the Fourier transform of $\hat{u}$ evaluated at $-x$. Hence, $u(x) = \hat{\hat{u}}(-x)$.

PROOF. First, by the definition of the Fourier transform, we have

$$R.H.S = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{ix\cdot\xi} e^{-iy\cdot\xi} u(y) dy \right) d\xi.$$

We cannot simply change the order of integrations since there is no absolute convergence. Instead, we introduce a limiting process to have the absolute convergence. Then

$$\begin{aligned} R.H.S. &= \frac{1}{(2\pi)^n} \lim_{\lambda\to\infty} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi} e^{-\frac{|\xi|^2}{\lambda}} u(y) dy d\xi \\ &= \frac{1}{(2\pi)^n} \lim_{\lambda\to\infty} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi - \frac{|\xi|^2}{\lambda}} u(y) d\xi dy, \end{aligned}$$

where we changed the order of integrations in the last step. A simple calculation of the exponential factor yields

$$R.H.S. = \frac{1}{(2\pi)^n} \lim_{\lambda\to\infty} \int_{\mathbb{R}^n} u(y) e^{-\frac{\lambda}{4}|x-y|^2} \int_{\mathbb{R}^n} e^{-\sum_{j=1}^n \left(i\frac{(x_j-y_j)\sqrt{\lambda}}{2} + \frac{\xi_j}{\sqrt{\lambda}}\right)^2} d\xi dy.$$

For any $a \in \mathbb{R}$, we first note

$$\int_{\mathbb{R}} e^{-\left(i\frac{a\sqrt{\lambda}}{2} + \frac{\eta}{\sqrt{\lambda}}\right)^2} d\eta = \int_{\mathbb{R}} e^{-\frac{\eta^2}{\lambda}} d\eta = \sqrt{\lambda} \int_{\mathbb{R}} e^{-\eta^2} d\eta = \sqrt{\lambda\pi}.$$

This integral is independent of a. Hence

$$\int_{\mathbb{R}^n} e^{-\sum_{j=1}^n \left(i\frac{(x_j-y_j)\sqrt{\lambda}}{2} + \frac{\xi_j}{\sqrt{\lambda}}\right)^2} d\xi = \prod_{j=1}^n \int_{\mathbb{R}} e^{-\left(i\frac{(x_j-y_j)\sqrt{\lambda}}{2} + \frac{\xi_j}{\sqrt{\lambda}}\right)^2} d\xi_j = (\lambda\pi)^{\frac{n}{2}}.$$

Then

$$R.H.S. = \frac{1}{(4\pi)^{\frac{n}{2}}} \lim_{\lambda\to\infty} \int_{\mathbb{R}^n} \lambda^{\frac{n}{2}} e^{-\frac{\lambda|x-y|^2}{4}} u(y) dy.$$

Let

$$K_\lambda(z) = \frac{1}{(4\pi)^{\frac{n}{2}}} \lambda^{\frac{n}{2}} e^{-\frac{|z|^2\lambda}{4}}.$$

Then $K_\lambda > 0$ in $\mathbb{R}^n$ and $\int_{\mathbb{R}^n} K_\lambda = 1$. It is straightforward to prove

$$R.H.S. = \lim_{\lambda\to\infty} \int_{\mathbb{R}^n} K_\lambda(x-y)u(y)dy = u(x).$$

We will provide detail in the proof of Theorem 5.5. □

Next, we set

$$(u, v)_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} u\bar{v} \quad \text{for any } u, v \in L^2(\mathbb{R}^n).$$

Here u and v may be complex-valued. The next result is referred to as the Parseval's identity.

THEOREM 5.3. *For any $u, v \in \mathcal{S}$,*

$$(u, v)_{L^2(\mathbb{R}^n)} = (\hat{u}, \hat{v})_{L^2(\mathbb{R}^n)}.$$

In particular, for any $u \in \mathcal{S}$

$$\|u\|_{L^2(\mathbb{R}^n)} = \|\hat{u}\|_{L^2(\mathbb{R}^n)}.$$

PROOF. We note

$$\begin{aligned}(\hat{u}, \hat{v})_{L^2(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \hat{u}(\xi)\overline{\hat{v}(\xi)} d\xi = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \overline{\hat{v}(\xi)} e^{-ix\cdot\xi} u(x) dx d\xi \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \overline{\hat{v}(\xi) e^{ix\cdot\xi}} u(x) d\xi dx = \int_{\mathbb{R}^n} u(x)\bar{v}(x) dx = (u, v)_{L^2(\mathbb{R}^n)},\end{aligned}$$

where we applied Theorem 5.2 to v. This yields the desired result. □

We now extend Fourier transforms to a larger class of functions than $\mathcal{S}$. Note that the Fourier transform is a linear operator from $\mathcal{S}$ to $\mathcal{S}$. Since $\mathcal{S}$ is dense in $L^2(\mathbb{R}^n)$, we can complete the space $\mathcal{S}$ under $\|\cdot\|_{L^2(\mathbb{R}^n)}$ to get $L^2(\mathbb{R}^n)$. Hence, the Fourier transform can be extended to a bounded linear operator in $L^2(\mathbb{R}^n)$. Since it is an isometry in $\mathcal{S}$ with respect to the L^2-norm, it is also an isometry in $L^2(\mathbb{R}^n)$.

We note that the Fourier transform also extends to functions in $L^1(\mathbb{R}^n)$. For any $u \in \mathcal{S}$,

$$\hat{u}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x)dx.$$

Hence

$$|\hat{u}(\xi)| \leq \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |u(x)|dx,$$

or

$$\sup_{\mathbb{R}^n} |\hat{u}| \leq \frac{1}{(2\pi)^{\frac{n}{2}}} \|u\|_{L^1(\mathbb{R}^n)}.$$

Note that $\mathcal{S}$ is dense in $L^1(\mathbb{R}^n)$. By completing $\mathcal{S}$ under the $L^1(\mathbb{R}^n)$, we conclude that the Fourier transform operator is a bounded linear operator from $L^1(\mathbb{R}^n)$ to the set of bounded continuous functions in $\mathbb{R}^n$.

Now we are ready to derive an expression for solutions of general initial-value problems of the heat equation. We consider

$$\begin{aligned} u_t - \Delta u &= 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \\ u(\cdot, 0) &= u_0 \quad \text{in } \mathbb{R}^n. \end{aligned} \tag{5.2}$$

Here we require $u \in C^2(\mathbb{R}^n \times (0,\infty)) \cap C(\mathbb{R}^n \times [0,\infty))$. Although called an initial-value problem, (5.2) is not the type of initial-value problems we discussed in Section 3.1. The heat equation is of the second order, while only one condition is prescribed on the initial hypersurface $\{t = 0\}$, which is characteristic.

We first use Fourier transforms to derive a formal solution. In the following, we employ the Fourier transform of u with respect to $x \in \mathbb{R}^n$ and write

$$\hat{u}(\xi, t) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x,t)dx.$$

Then by (F2)

$$\begin{aligned} \hat{u}_t + |\xi|^2 \hat{u} &= 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \\ \hat{u}(\cdot, 0) &= \hat{u}_0 \quad \text{in } \mathbb{R}^n. \end{aligned}$$

This is an initial-value problem for an ODE with ξ as a parameter and its solution is given by

$$\hat{u}(\xi, t) = \hat{u}_0(\xi) e^{-|\xi|^2 t}.$$

Now we treat t as a parameter. Let $K(x,t)$ satisfy

$$\hat{K}(\xi, t) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-|\xi|^2 t} \quad \text{for } t > 0.$$

Then

$$\hat{u}(\xi, t) = (2\pi)^{\frac{n}{2}} \hat{K}(\xi, t) \hat{u}_0(\xi).$$

Therefore Theorem 5.2 and (F5) imply

$$u(x,t) = \int_{\mathbb{R}^n} K(x-y, t) u_0(y) dy.$$

This is called the Poisson integral formula. By Theorem 5.2 and Example 5.1, we have

$$K(x,t) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} e^{-|\xi|^2 t} d\xi,$$

or

$$K(x,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}} \quad \text{for any } x \in \mathbb{R}^n,\ t > 0, \tag{5.3}$$

which is called the fundamental solution of the heat equation.

Now we verify directly that the Poisson integral formula defines a solution under suitable conditions on u_0. The proof follows the same line as that of the Poisson integral formula for the Laplace equation in Theorem 4.6.

We first note that the fundamental solution K satisfies the following properties:

(K1) $K(x,t)$ is smooth for any $x \in \mathbb{R}^n$ and $t > 0$.

(K2) $K(x,t) > 0$ for any $x \in \mathbb{R}^n$ and $t > 0$.

(K3) $(\partial_t - \Delta_x)K(x,t) = 0$ for any $x \in \mathbb{R}^n$ and $t > 0$.

(K4) $\int_{\mathbb{R}^n} K(\cdot,t) = 1$ for any $t > 0$.

(K5) For any $\delta > 0$,

$$\lim_{t\to 0+} \int_{|x|>\delta} K(x,t)dx = 0.$$

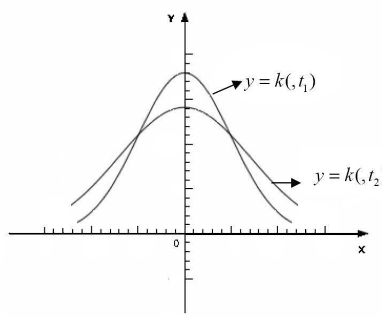

FIGURE 5.1. Graphs of fundamental solutions for $t_2 > t_1 > 0$.

Here (K1) and (K2) follow from the explicit expression for K in (5.3). We may also get (K3) from (5.3) by a straightforward calculation. However, an easy way to get (K3) is to use the integral expression of K preceding (5.3). For (K4) and (K5), we simply note

$$\int_{|x|>\delta} K(x,t)dx = \frac{1}{\pi^{\frac{n}{2}}} \int_{|\eta|>\frac{\delta}{2\sqrt{t}}} e^{-|\eta|^2} d\eta.$$

This implies (K4) for $\delta = 0$ and (K5) for $\delta > 0$.

Now we are ready to solve initial-value problems of the heat equation.

THEOREM 5.4. *For any bounded $u_0 \in C(\mathbb{R}^n)$, the function u defined by*

$$u(x,t) = \int_{\mathbb{R}^n} K(x-y,t)u_0(y)dy \tag{5.4}$$

is smooth in $\mathbb{R}^n \times (0,\infty)$ and continuous up to $t = 0$ and satisfies

$$\begin{aligned} u_t - \Delta u &= 0 \quad \text{in } \mathbb{R}^n \times (0,\infty),\\ u(\cdot,0) &= u_0 \quad \text{in } \mathbb{R}^n. \end{aligned}$$

PROOF. *Step 1.* We first prove that $u(x,t)$ is smooth for any $x \in \mathbb{R}^n$ and $t > 0$. It suffices to check

$$\int_{\mathbb{R}^n} |x-y|^m e^{-\frac{|x-y|^2}{4t}} |u_0(y)| dy < \infty,$$

for any $m \geq 0$. This follows from the exponential decay of the integrand. Then u is a solution of the heat equation by a simple differentiation. We point out for future references that we only use the boundedness of u_0.

Step 2. For any fixed $x_0 \in \mathbb{R}^n$, we claim

$$\lim_{(x,t)\to(x_0,0)} u(x,t) = u_0(x_0).$$

For simplicity, we assume $x_0 = 0$. By (K4), we have

$$u(x,t) - u_0(0) = \int_{\mathbb{R}^n} K(x-y,t)\big(u_0(y) - u_0(0)\big) dy.$$

Then

$$|u(x,t) - u_0(0)| \leq \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} |u_0(y) - u_0(0)| dy.$$

For any $\varepsilon > 0$, by the continuity of u_0, there exists a positive constant $\delta = \delta(\varepsilon)$ such that

$$|u_0(y) - u_0(0)| < \varepsilon \quad \text{for any} \quad |y| \leq \delta.$$

Next, we assume $|u_0| \leq M$ by the boundedness of u_0. Then by (K2) and (K4)

$$\begin{aligned}
&|u(x,t) - u_0(0)| \\
\leq &\int_{B_\delta} K(x-y,t)|u_0(y) - u_0(0)| dy + \int_{\mathbb{R}^n \setminus B_\delta} K(x-y,t)|u_0(y) - u_0(0)| dy \\
\leq &\varepsilon + 2M \int_{\mathbb{R}^n \setminus B_\delta} K(x-y,t) dy.
\end{aligned}$$

We note $|x-y| \geq \delta/2$ for any $y \in \mathbb{R}^n \setminus B_\delta$ and $x \in B_{\delta/2}$. Hence

$$|u(x,t) - u_0(0)| \leq \varepsilon + 2M \int_{\mathbb{R}^n \setminus B_{\frac{\delta}{2}}} K(z,t) dz \quad \text{for any } x \in B_{\frac{\delta}{2}}.$$

By (K5), there exists a $t_0 > 0$ depending only on ε, M and δ such that

$$|u(x,t) - u_0(0)| \leq 2\varepsilon \quad \text{for any } x \in B_{\frac{\delta}{2}},\ t \in (0, t_0).$$

We then have the desired result. □

Now we discuss a result more general than Theorem 5.4. The boundedness assumption on u_0 in Theorem 5.4 can be relaxed. To seek a reasonable assumption on initial values, we examine the expression of the fundamental solution K. We note that K in (5.3) has an exponential decay in space variables with a large decay rate for small time. This suggests that we can allow an exponential growth for initial values. In the convolution (5.4), a fixed exponential growth from initial values can be offset by the fast exponential decay in the fundamental solution at least for a short time period. To see this clearly, we consider an example. For any $\alpha > 0$, set

$$G(x,t) = \frac{1}{(1-4\alpha t)^{\frac{n}{2}}} e^{\frac{\alpha}{1-4\alpha t}|x|^2} \quad \text{for any } x \in \mathbb{R}^n,\ t < \frac{1}{4\alpha}.$$

It is straightforward to check

$$G_t - \Delta G = 0 \quad \text{for any } x \in \mathbb{R}^n,\ t < \frac{1}{4\alpha}.$$

Note

$$G(x,0) = e^{\alpha|x|^2} \quad \text{for any } x \in \mathbb{R}^n.$$

Hence, G has an exponential growth initially for $t = 0$, and in fact for any $t < 1/4\alpha$. The growth rate becomes arbitrarily large as t approaches $1/4\alpha$ and G does not exist beyond $t = 1/4\alpha$.

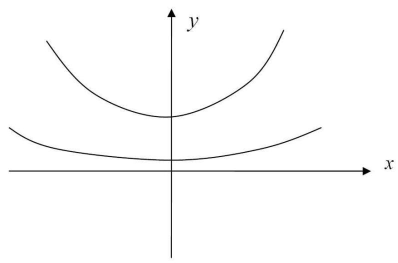

FIGURE 5.2. Graphs of G for different time.

Now we formulate a general result. If u_0 is continuous and has an exponential growth, then (5.4) still defines a solution of the initial-value problem in a short period of time.

THEOREM 5.5. *Suppose* $u_0 \in C(\mathbb{R}^n)$ *satisfy*

$$|u_0(x)| \le Me^{A|x|^2} \quad \textit{for any } x \in \mathbb{R}^n,$$

for some constants $M, A \ge 0$. *Then* u *defined in (5.4) is a* $C^\infty(\mathbb{R}^n \times (0,T)) \times C(\mathbb{R}^n \times [0,T))$*-solution of*

$$u_t - \Delta u = 0 \quad \textit{in } \mathbb{R}^n \times (0,T],$$
$$u(\cdot,0) = u_0 \quad \textit{in } \mathbb{R}^n,$$

for any $T < \frac{1}{4A}$. *Moreover,*

$$|u(x,t)| \le M_1 e^{A_1|x|^2} \quad \textit{for any } x \in \mathbb{R}^n, 0 < t \le T,$$

where $A_1 \ge 0$ *and* $M_1 > 0$ *are constants depending only on* A, M *and* T.

We first prove the following lemma.

LEMMA 5.6. *Suppose* $f, g \in C(\mathbb{R}^n)$ *satisfy*

$$|f(x)| \le Me^{A|x|^2}, \quad |g(x)| \le \bar{M}e^{\bar{A}|x|^2},$$

for some constants $M, \bar{M} > 0$ *and some* $A, \bar{A}$. *If* $A + \bar{A} < 0$, *then*

$$h(x) = \int_{\mathbb{R}^n} g(x-y)f(y)dy$$

is continuous in $\mathbb{R}^n$ *and*

$$|h(x)| \le \tilde{M}e^{\tilde{A}|x|^2},$$

where $\tilde{M}$ *and* $\tilde{A}$ *are constants depending only on* $M, \bar{M}, A$ *and* $\bar{A}$.

PROOF. With $A + \bar{A} < 0$, we have

$$\begin{aligned}|h(x)| &\leq M\bar{M}\int_{\mathbb{R}^n} e^{A|y|^2+\bar{A}|x-y|^2}dy\\ &= M\bar{M}e^{\frac{A\bar{A}}{A+\bar{A}}|x|^2}\int_{\mathbb{R}^n} e^{(A+\bar{A})|y-\frac{\bar{A}}{A+\bar{A}}x|^2}dy\\ &= M\bar{M}\big(\frac{-\pi}{A+\bar{A}}\big)^{\frac{n}{2}}e^{\frac{A\bar{A}}{A+\bar{A}}|x|^2}.\end{aligned}$$

This yields the desired estimate. □

Now we prove Theorem 5.5. The proof follows the same line as that of Theorem 5.4.

PROOF OF THEOREM 5.5. *Step 1.* The case $A = 0$ is covered by Theorem 5.4. We only consider $A > 0$ and take T such that $-\frac{1}{4T} + A < 0$ or $T < \frac{1}{4A}$. Then Lemma 5.6 implies u is continuous in $\mathbb{R}^n \times (0, T]$. To show that u has continuous derivatives of arbitrary order in $\mathbb{R}^n \times (0, T]$, we only need to verify

$$\int_{\mathbb{R}^n} |x-y|^m e^{-\frac{|x-y|^2}{4t}}|u_0(y)|dy < \infty,$$

for any $m \geq 0$. This can be proved similarly as the proof of Lemma 5.6.

Step 2. For any fixed $x_0 \in \mathbb{R}^n$, we claim

$$\lim_{(x,t)\to(x_0,0)} u(x,t) = u_0(x_0).$$

For simplicity, we assume $x_0 = 0$. By (K4), we have

$$u(x,t) - u_0(0) = \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}\big(u_0(y) - u_0(0)\big)dy.$$

Set $v(y) = u_0(y) - u_0(0)$. Then

$$|u(x,t) - u_0(0)| \leq \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}|v(y)|dy.$$

For any $\varepsilon > 0$, by the continuity of v, there exists a $\delta = \delta(\varepsilon)$ such that

$$|v(y)| < \varepsilon \quad \text{for any} \quad |y| \leq \delta.$$

Note $|v(y)| \leq \tilde{M}e^{\tilde{A}|y|^2}$ for some constants $\tilde{M}$ and $\tilde{A}$. Then there exists a constant $B > 0$ such that

$$|v(y)| \leq \varepsilon e^{B|y|^2} \quad \text{for any } |y| \geq \delta.$$

In fact, we can choose B such that $\varepsilon e^{(B-\tilde{A})\delta^2} \geq \tilde{M}$. Then by $e^{B|y|^2} \geq 1$, we have

$$|v(y)| \leq \varepsilon e^{B|y|^2} \quad \text{for any } y \in \mathbb{R}^n.$$

Therefore,

$$|u(x,t) - u_0(0)| \leq \frac{\varepsilon}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}e^{B|y|^2}dy.$$

A direct calculation then yields

$$|u(x,t) - u_0(0)| \leq \frac{\varepsilon}{(1-4Bt)^{\frac{n}{2}}}e^{\frac{B}{1-4Bt}|x|^2} \quad \text{for any } t < \frac{1}{4B}.$$

We have the desired result by letting $(x,t) \to (0,0)$. □

Now we discuss some properties of the solution (5.4) of the initial-value problem (5.2). First from (5.4), $u(x,t)$ for $t>0$ depends on values of u_0 at all points. Equivalently, values of u_0 near a point $x_0 \in \mathbb{R}^n$ affect the value of $u(x,t)$ at all x as long as $t>0$. Thus values here travel at infinite speed.

Next, the function $u(x,t)$ in (5.4) becomes smooth for $t>0$, even if the initial value u_0 is simply bounded. This is well illustrated in Step 1 in the proof of Theorem 5.4. We did not use any regularity assumption on u_0 there. Compare with Theorem 3.16. Later on, we will prove a general result that any solutions of the heat equation in a domain in $\mathbb{R}^n \times (0,\infty)$ are smooth away from the boundary. Moreover, for each fixed time, any solutions of the heat equation are analytic in space variables away from the boundary.

We need to point out that (5.4) represents only one of infinitely many solutions. Solutions are not unique without further conditions on u, like the boundedness assumption or the exponential growth assumption. In fact, there exists a nontrivial solution $u \in C^\infty(\mathbb{R}^n \times \mathbb{R})$ of $u_t - \Delta u = 0$, with $u \equiv 0$ for $t \le 0$. In the following, we construct such a solution of the 1-dimensional heat equation due to Tychonov.

PROPOSITION 5.7. *There exists a nonzero smooth function* $u \in C^\infty(\mathbb{R}\times[0,\infty))$ *satisfying*

$$\begin{aligned} u_t - u_{xx} =&0 \quad \textit{in } \mathbb{R}\times[0,\infty),\\ u(\cdot,0) =&0 \quad \textit{in } \mathbb{R}. \end{aligned}$$

PROOF. We construct a smooth function in $\mathbb{R}\times\mathbb{R}$ such that $u_t - u_{xx} = 0$ in $\mathbb{R}\times\mathbb{R}$ and $u \equiv 0$ for $t<0$. We treat $x=0$ as the initial curve and attempt to find a smooth solution of the following initial-value problem

$$\begin{aligned} &u_t - u_{xx} = 0 \quad \text{in } \mathbb{R}\times\mathbb{R},\\ &u(0,t) = a(t),\ u_x(0,t) = 0 \quad \text{for any } t \in \mathbb{R}, \end{aligned}$$

for an appropriate function a in $\mathbb{R}$. We write u as a power series in x as

$$u(x,t) = \sum_{k=1}^{\infty} a_k(t) x^k.$$

By a simple substitution in $u_t = u_{xx}$ and a comparison of coefficients of power of x, we have

$$a_0 = a,\ a_1 = 0,\ a_k' = (k+1)(k+2)a_{k+2} \text{ for any } k \ge 0.$$

Therefore, we have a formal solution

$$u(x,t) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} a^{(k)}(t) x^{2k}.$$

We choose $a(t)$ appropriately so that $u(x,t)$ defined above is a smooth function and is identically zero for $t<0$. To this end, we define

$$a(t) = \begin{cases} e^{-\frac{1}{t^2}} & \text{for } t>0,\\ 0 & \text{for } t \le 0. \end{cases}$$

Then it is straightforward to verify that the series defining u is absolutely convergent in $\mathbb{R}\times\mathbb{R}$. This implies that u is continuous. Similarly, we can prove that u is smooth. We skip details and leave the rest of the proof as an exercise. $\square$

To end this section, we discuss briefly terminal-value problems. For a fixed constant $T > 0$, we consider

$$u_t - u_{xx} = 0 \quad \text{in } \mathbb{R} \times (0, T),$$
$$u(\cdot, T) = \varphi \quad \text{in } \mathbb{R}.$$

Here the initial value φ is prescribed at the terminal time T. We are interested in the well-posedness of this problem. First, we ask whether there exists a function a in $\mathbb{R}$ such that the solution u of the initial-value problem

$$u_t - u_{xx} = 0 \quad \text{in } \mathbb{R} \times (0, T),$$
$$u(\cdot, 0) = a \quad \text{in } \mathbb{R}$$

satisfies

$$u(\cdot, T) = \varphi \quad \text{in } \mathbb{R}.$$

Obviously, a necessary condition is given by $\varphi \in C^\infty(\mathbb{R})$. In fact, φ has to be analytic, as will be shown by Corollary 5.11.

Now we consider an example. For any positive integer m, we set

$$\varphi_m(x) = u_m(x, T) = \varepsilon \sin(mx).$$

Then, a solution is given by

$$u_m(x, t) = \varepsilon e^{m^2(T-t)} \sin(mx),$$

with

$$a_m(x) = v_m(x, 0) = \varepsilon e^{m^2 T} \sin(mx).$$

We note

$$\max_{\mathbb{R}} |\varphi_m| = \varepsilon,$$

and

$$\max_{\mathbb{R}} |a_m| = \varepsilon e^{m^2 T} \to \infty \text{ as } m \to \infty.$$

There is no continuous dependence on initial values (prescribed at the terminal time T).

5.2. Regularity of Solutions

In this section, we discuss regularity of solutions of the heat equation with the help of the fundamental solution.

We often discuss the heat equation $u_t - \Delta u = 0$ in cylinders of the form $\Omega \times (t_0, t_1]$, where Ω is a domain in $\mathbb{R}^n$ and $t_0 < t_1$ are two scalars. The *parabolic boundary* $\partial_p(\Omega \times (t_0, t_1])$ is defined by

$$\partial_p\big(\Omega \times (t_0, t_1]\big) = \big(\Omega \times \{t_0\}\big) \cup \big(\partial\Omega \times [t_0, t_1)\big).$$

In other words, parabolic boundary consists of the bottom and the side of the boundary. For any $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and any $r > 0$, we define

$$Q_r(x_0, t_0) = B_r(x_0) \times (t_0 - r^2, t_0].$$

We note that $Q_r(x_0, t_0)$ for the heat equation play the same role as balls for the Laplace equation. If u is a solution of the heat equation $u_t - \Delta u = 0$ in $Q_r(0)$, then

$$u_r(x, t) = u(rx, r^2 t)$$

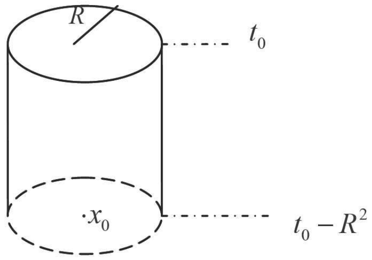

FIGURE 5.3. The region $Q_R(x_0, t_0)$.

is a solution of the heat equation in $Q_1(0)$. For any (x, t) and (y, s), the parabolic distance is defined by

$$d_P\big((x,t),(y,s)\big) = \big(|x-y|^2 + |t-s|\big)^{\frac{1}{2}}.$$

In the following, we denote by $C^{2,1}$ the collection of functions which are C^2 in x and C^1 in t. We first have the following regularity result for solutions of the heat equation.

THEOREM 5.8. *Let u be a $C^{2,1}$-solution of $u_t - \Delta u = 0$ in $Q_R(x_0, t_0)$ for some $R > 0$. Then u is smooth in $Q_R(x_0, t_0)$.*

PROOF. We claim for any $(x, t) \in Q_R(x_0, t_0)$

$$\begin{aligned} u(x,t) = &\int_{B_R(x_0)} K\big(x-y, t-(t_0-R^2)\big)u(y, t_0 - R^2)dy \\ &+ \int_{t_0-R^2}^{t}\int_{\partial B_R(x_0)} \big[K(x-y,t-s)\frac{\partial u}{\partial n_y}(y,s) - u(y,s)\frac{\partial K}{\partial n_y}(x-y,t-s)\big]dS_y ds. \end{aligned}$$

This implies the smoothness of u in $Q_R(x_0, t_0)$ easily since the boundary integrals in the right-hand side are only on the parabolic boundary of $B_R(x_0) \times (t_0 - R^2, t]$ and there is no singularity for integrands for $(x, t) \in Q_R(x_0, t_0)$.

We denote by (y, s) points in $Q_R(x_0, t_0)$. Set

$$K = K(x-y, t-s) = \frac{1}{\big(4\pi(t-s)\big)^{\frac{n}{2}}}e^{-\frac{|x-y|^2}{4(t-s)}} \quad \text{for } s < t.$$

Then

$$K_s + \Delta_y K = 0,$$

and hence

$$\begin{aligned} 0 =& K(u_s - \Delta_y u) = (uK)_s + \sum_{i=1}^n (uK_{y_i} - Ku_{y_i})_{y_i} - u(K_s + \Delta_y K) \\ =& (uK)_s + \sum_{i=1}^n (uK_{y_i} - Ku_{y_i})_{y_i}. \end{aligned}$$

For any fixed $(x,t) \in Q_R(x_0,t_0)$ and any $\varepsilon > 0$, we integrate with respect to (y,s) in $B_R(x_0) \times (t_0 - R^2, t-\varepsilon)$. Then

$$\int_{B_R(x_0)} K(x-y,\varepsilon)u(y,t-\varepsilon)dy = \int_{B_R(x_0)} K\big(x-y,t-(t_0-R^2)\big)u(y,t_0-R^2)dy$$
$$+\int_{t_0-R^2}^{t-\varepsilon}\int_{\partial B_R(x_0)} \big[K(x-y,t-s)\frac{\partial u}{\partial n_y}(y,s) - u(y,s)\frac{\partial K}{\partial n_y}(x-y,t-s)\big]dS_y ds.$$

Now it suffices to prove

$$\lim_{\varepsilon\to 0+}\int_{B_R(x_0)} K(x-y,\varepsilon)u(y,t-\varepsilon) = u(x,t).$$

The proof is almost identical to Step 2 in the proof of Theorem 5.4. The integral over a finite domain here introduces few changes. We omit details. □

Now we prove the interior gradient estimates by a similar method.

THEOREM 5.9. *Let u be a $C^{2,1}$-solution of $u_t - \Delta u = 0$ in $Q_R(x_0,t_0)$ for some $R > 0$. Then*

$$|\nabla_x u(x_0,t_0)| \le \frac{c}{R}\sup_{Q_R(x_0,t_0)} |u|,$$

where c is a positive constant depending only on n.

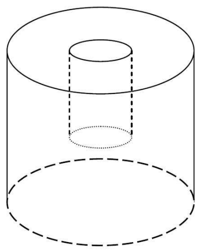

FIGURE 5.4. Domains for interior gradient estimates.

PROOF. We first modify the proof of Theorem 5.8 to express u using the fundamental solution and cutoff functions. We denote by (y,s) points in $Q_R(x_0,t_0)$. For any smooth function v is $B_R(x_0) \times (t_0-R^2, t_0]$, we have

$$0 = v(u_s - \Delta_y u) = (uv)_s + \sum_{i=1}^n (uv_{y_i} - vu_{y_i})_{y_i} - u(v_s + \Delta_y v).$$

We first choose a cutoff function $\varphi \in C^\infty(Q_R(x_0,t_0))$ such that

$$\text{supp}\varphi \subset Q_{\frac{3}{4}R}(x_0,t_0), \quad \varphi \equiv 1 \text{ in } Q_{\frac{1}{2}R}(x_0,t_0).$$

Let K be the fundamental solution of the heat equation as in (5.3) and set for any fixed $(x,t) \in Q_{R/4}(x_0,t_0)$

$$v = \varphi K = \varphi(y,s)K(x-y,t-s) \quad \text{for } s < t.$$

For any $\varepsilon > 0$, we integrate with respect to (y,s) in $B_R(x_0) \times (t_0 - R^2, t-\varepsilon)$. We first note that there is no boundary integral over the parabolic boundary of

$B_R(x_0) \times (t_0 - R^2, t-\varepsilon)$ since the support of φ has an empty intersection with this part of the boundary. Hence

$$\int_{B_R(x_0)} (\varphi u)(y, t-\varepsilon) K(x-y, \varepsilon) dy = \int_{B_R(x_0)\times(t_0-R^2, t-\varepsilon)} u(\partial_s + \Delta_y)(\varphi K) dyds.$$

Then similarly as in the proof of Theorem 5.4, with $\varepsilon \to 0$ we have

$$\varphi(x,t)u(x,t) = \int_{B_R(x_0)\times(t_0-R^2,t)} u(\partial_s + \Delta_y)(\varphi K) dyds.$$

With

$$K = \frac{1}{\big(4\pi(t-s)\big)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{4(t-s)}} \quad \text{for } s < t,$$

we note

$$K_s + \Delta_y K = 0.$$

Hence for any $(x,t) \in Q_{R/4}(x_0, t_0)$

$$u(x,t) = \int_{B_R(x_0)\times(t_0-R^2,t)} u(\varphi_s K + \Delta_y \varphi K + 2\nabla_y \varphi \cdot \nabla_y K) dyds.$$

We note that each term in the integrand involves a derivative of φ, which is zero in $Q_{R/2}(x_0, t_0)$ since $\varphi \equiv 1$ there. Then the integral domain D is given by

$$D = B_{\frac{3}{4}R}(x_0) \times (t_0 - (\frac{3}{4}R)^2, t] \setminus B_{\frac{1}{2}R}(x_0) \times (t_0 - (\frac{1}{2}R)^2, t].$$

Hence the distance between any $(y,s) \in D$ and any $(x,t) \in Q_{R/4}(x_0,t_0)$ has a positive lower bound. Therefore, the integrand has no singularity in the integral domain. (This gives an alternate proof of the smoothness of u in $Q_{R/4}(x_0,t_0)$.)

Next, we have for any $(x,t) \in Q_{R/4}(x_0, t_0)$

$$\nabla_x u(x,t) = \int_D u\big((\varphi_s + \Delta_y \varphi)\nabla_x K + 2\nabla_y \varphi \cdot \nabla_x \nabla_y K\big) dyds.$$

For the cutoff function φ in $Q_R(x_0,t_0)$, we require further

$$|\nabla_y \varphi| \le \frac{c}{R}, \quad |\varphi_s| + |\nabla_y^2 \varphi| \le \frac{c}{R^2},$$

where c is a positive constant depending only on n. With the explicit expression of K, we have

$$|\nabla_x K| \le c\frac{|x-y|}{(t-s)^{\frac{n}{2}+1}} e^{-\frac{|x-y|^2}{4(t-s)}},$$

and

$$|\nabla_x^2 K| \le c\frac{|x-y|^2 + (t-s)}{(t-s)^{\frac{n}{2}+2}} e^{-\frac{|x-y|^2}{4(t-s)}}.$$

Obviously, for any $(x,t) \in Q_{R/4}(x_0,t_0)$ and any $(y,s) \in D$,

$$|x-y| \le R, \quad 0 < t-s \le R^2.$$

Therefore, for any $(x,t) \in Q_{R/4}(x_0,t_0)$,

$$|\nabla_x u(x,t)| \le c\int_D \Big(\frac{1}{R(t-s)^{\frac{n}{2}+1}} e^{-\frac{|x-y|^2}{4(t-s)}} + \frac{R}{(t-s)^{\frac{n}{2}+2}} e^{-\frac{|x-y|^2}{4(t-s)}}\Big)|u(y,s)| dyds.$$

Now we claim

$$\int_D \frac{1}{(t-s)^{\frac{n}{2}+i}} e^{-\frac{|x-y|^2}{4(t-s)}} dyds \leq \frac{c}{R^{2i-2}} \quad \text{for } i = 1, 2.$$

Then we obtain easily

$$|\nabla_x u(x,t)| \leq \frac{c}{R} \sup_{Q_R(x_0,t_0)} |u| \quad \text{for any } (x,t) \in Q_{\frac{R}{4}}(x_0, t_0).$$

Next, we prove the claim. We decompose D into two parts

$$\begin{aligned} D_1 =& B_{\frac{R}{2}}(x_0) \times \big(t_0 - (\frac{3}{4}R)^2, t_0 - (\frac{1}{2}R)^2\big), \\ D_2 =& \big(B_{\frac{3}{4}R}(x_0) \setminus B_{\frac{1}{2}R}(x_0)\big) \times \big(t_0 - (\frac{3}{4}R)^2, t\big). \end{aligned}$$

For any $(x,t) \in Q_{R/4}(x_0, t_0)$ and $(y,s) \in D_1$, we have

$$t - s \geq \frac{1}{4}R^2.$$

Hence for any $(x,t) \in Q_{R/4}(x_0, t_0)$

$$\int_{D_1} \frac{1}{(t-s)^{\frac{n}{2}+i}} e^{-\frac{|x-y|^2}{4(t-s)}} dyds \leq \int_{D_1} \frac{2^{n+2i}}{R^{n+2i}} dyds \leq \frac{c}{R^{2i-2}},$$

where we used the fact that the volume of D_1 has an order of R^{n+2}. Now we consider D_2. For any $(x,t) \in Q_{R/4}(x_0, t_0)$ and $(y,s) \in D_2$, we have

$$|y - x| \geq \frac{1}{4}R,$$

and hence

$$\frac{1}{(t-s)^{\frac{n}{2}+i}} e^{-\frac{|x-y|^2}{4(t-s)}} \leq \frac{1}{(t-s)^{\frac{n}{2}+i}} e^{-\frac{R^2}{4^3(t-s)}}.$$

Then for any $(x,t) \in Q_{R/4}(x_0, t_0)$ and $(y,s) \in D_2$,

$$I \equiv \int_{D_2} \frac{1}{(t-s)^{\frac{n}{2}+i}} e^{-\frac{|x-y|^2}{4(t-s)}} dyds \leq cR^n \int_{t_0 - (\frac{3}{4}R)^2}^{t} \frac{1}{(t-s)^{\frac{n}{2}+i}} e^{-\frac{R^2}{4^3(t-s)}} ds.$$

With $t - s = R^2\tau$, we obtain

$$I \leq \frac{c}{R^{2i-2}} \int_0^{(\frac{3}{4})^2} \frac{1}{\tau^{\frac{n}{2}+i}} e^{-\frac{1}{4^3\tau}} d\tau \leq \frac{c}{R^{2i-2}}.$$

This finishes the proof of the claim. □

Next, we estimate derivatives of arbitrary order.

COROLLARY 5.10. *Let u be a $C^{2,1}$-solution of $u_t - \Delta u = 0$ in $Q_R(x_0, t_0)$ for some $R > 0$. Then*

$$|\partial_t^k \nabla_x^m u(x_0, t_0)| \leq \frac{c^{m+2k}}{R^{m+2k}} n^k e^{m+2k-1} (m+2k)! \sup_{Q_R(x_0,t_0)} |u|,$$

where c is a positive constant depending only on n.

PROOF. For x-derivatives, we proceed as in the proof of Theorem 4.9 and obtain for any multi-index α with $|\alpha| = m$

$$|\partial_x^\alpha u(x_0, t_0)| \leq \frac{c^m e^{m-1} m!}{R^m} \sup_{Q_R(x_0,t_0)} |u|.$$

For t-derivatives, we have $u_t = \Delta u$ and hence $\partial_t^k u = \Delta^k u$ for any nonnegative integer k. We note that there are n^k terms of x-derivatives of u of order $2k$ in $\Delta^k u$. Hence

$$|\partial_t^k \nabla_x^m u(x_0, t_0)| \leq n^k \max_{|\beta|=m+2k} |\partial_x^\beta u(x_0, t_0)|.$$

This implies the desired result easily. □

Now we prove the analyticity of solutions of the heat equation on any time slice.

COROLLARY 5.11. *Let u be a $C^{2,1}$-solution of $u_t - \Delta u = 0$ in $Q_R(x_0, t_0)$ for some $R > 0$. Then $u(\cdot, t)$ is analytic in $B_R(x_0)$ for any $t \in (t_0 - R^2, t_0]$. Moreover, for each nonnegative integer k, $\partial_t^k u(\cdot, t)$ is analytic in $B_R(x_0)$ for any $t \in (t_0 - R^2, t_0]$.*

The proof is identical to that of Theorem 4.10 and is omitted. As shown in Proposition 5.7, solutions of $u_t - \Delta u = 0$ are not analytic in t in general.

5.3. The Maximum Principle

In this section, we discuss the maximum principle for the heat equation and its applications. Suppose $\Omega \subset \mathbb{R}^n$ is a bounded smooth domain. For any $T > 0$, set

$$Q_T = \Omega \times (0, T] = \{(x, t); x \in \Omega, 0 < t \leq T\},$$

and

$$\partial_p Q_T = (\Omega \times \{t = 0\}) \cup (\partial\Omega \times [0, T]).$$

As in the previous section, $\partial_p Q_T$ is called the *parabolic boundary* of Q_T. We denote by $C^{2,1}(Q_T)$ the collection of functions which are C^2 in x and C^1 in t in Q_T. Without confusion, we simply write Q instead Q_T.

We first prove the maximum principle, which asserts that any subsolutions of the heat equation attains their maximum on the parabolic boundary.

THEOREM 5.12. *For $u \in C^{2,1}(Q) \cap C(\bar{Q})$, if $u_t - \Delta u \leq 0$, then the maximum of u in $\bar{Q}$ is attained on $\partial_p Q$, i.e.,*

$$\max_{\bar{Q}} u = \max_{\partial_p Q} u.$$

PROOF. 1) We first consider a special case. If $u_t - \Delta u < 0$, we prove by contradiction that the maximum of u in $\bar{Q}$ cannot be attained in Q. Suppose there exists a point $P_0 \in Q$ such that

$$u(P_0) = \max_{\bar{Q}} u.$$

Then $u_{x_i}(P_0) = 0$ and $\{u_{x_i x_j}(P_0)\}$ is a nonpositive definite matrix. Moreover, $u_t(P_0) = 0$ for $t \in (0, T)$ and $u_t(P_0) \geq 0$ for $t = T$. Hence $u_t - \Delta u \geq 0$ at P_0, which leads to a contradiction.

2) For any $\varepsilon > 0$, we consider an auxiliary function

$$v(x,t) = u(x,t) - \varepsilon t.$$

Then

$$v_t - \Delta v = u_t - \Delta u - \epsilon < 0.$$

By 1), v cannot attain its maximum inside Q. Hence

$$\max_{\bar{Q}} v = \max_{\partial_p Q} v.$$

Then

$$\begin{aligned}\max_{\bar{Q}} u(x,t) &= \max_{\bar{Q}}(v(x,t) + \varepsilon t)) \le \max_{\bar{Q}} v(x,t) + \varepsilon T\\ &= \max_{\partial_p Q} v(x,t) + \varepsilon T \le \max_{\partial_p Q} u(x,t) + \varepsilon T.\end{aligned}$$

Letting $\varepsilon \to 0$, we get the desired result. □

The following result is referred to as the comparison principle.

COROLLARY 5.13. *Suppose $u, v \in C^{2,1}(Q) \cap C(\bar{Q})$ satisfy $u_t - \Delta u \le v_t - \Delta v$ in Q and $u \le v$ on $\partial_p Q$. Then $u \le v$ in $\bar{Q}$.*

Next, we consider a slightly more general operator given by

$$Lu \equiv u_t - \Delta u + cu = f \quad \text{in } Q.$$

We prove the following weak maximum principle for L.

THEOREM 5.14. *Suppose $u \in C^{2,1}(Q) \cap C(\bar{Q})$ satisfies $Lu \le 0$ with $c(x,t) \ge 0$. Then the nonnegative maximum of u must be attained on $\partial_p Q$, i.e.,*

$$\max_{\bar{Q}} u \le \max_{\partial_p Q}\{0, u\}.$$

The proof is a simple modification of that of Theorem 5.12 and is omitted.

Now, we consider a more general case.

THEOREM 5.15. *Suppose $c(x,t) \ge -c_0$ in Q for a nonnegative constant c_0 and $u \in C^{2,1}(Q) \cap C(\bar{Q})$ satisfies $Lu \le 0$. If $u \le 0$ on $\partial_p Q$, then $u \le 0$ in Q.*

Continuous functions in $\bar{Q}$ always have global minimum. Therefore, $c \ge -c_0$ always holds in Q for some nonnegative constant c_0 if c is continuous in $\bar{Q}$. Such a condition is introduced to emphasize the role of the minimum of c.

PROOF. Let $v(x,t) = e^{-c_0 t} u(x,t)$. Then

$$v_t - \Delta v + (c(x,t) + c_0) v = e^{-c_0 t} Lu \le 0.$$

Since $c(x,t) + c_0 \ge 0$, by Theorem 5.14, we get

$$\max_{Q} v(x,t) \le \max_{\partial_p Q}\{0, v(x,t)\} \le \max_{\partial_p Q}\{0, e^{-c_0 t} u(x,t)\} \le 0.$$

Hence $u(x,t) \le 0$ in Q. □

The following result is referred to as the comparison principle, which generalizes Corollary 5.13.

COROLLARY 5.16. *Suppose $c(x,t) \geq -c_0$ in Q for a nonnegative constant c_0 and $u, v \in C^{2,1}(Q) \cap C(\bar{Q})$ satisfy $Lu \leq Lv$ in Q and $u \leq v$ on $\partial_p Q$. Then $u \leq v$ in $\bar{Q}$.*

There is also a strong maximum principle for parabolic differential equations. We will not discuss it here. In the following, we discuss applications of maximum principles.

We first derive an estimate in sup-norms of solutions of initial/boundary-value problems with Dirichlet boundary values. Compare this with the estimates in integral norms in Theorem 3.14.

THEOREM 5.17. *Suppose $u \in C^{2,1}(Q) \cap C(\bar{Q})$ is a solution of*

$$\begin{aligned} Lu \equiv u_t - \Delta u + cu &= f \quad \text{in } Q_T, \\ u(\cdot, 0) &= u_0 \quad \text{in } \Omega, \\ u &= \varphi \quad \text{on } \partial\Omega \times (0,T), \end{aligned}$$

for some $f \in C(\bar{Q})$, $u_0 \in C(\bar{\Omega})$ and $\varphi \in C(\partial\Omega \times [0,T])$. If $c(x,t) \geq -c_0$ in Q for a nonnegative constant c_0, then

$$\max_{\bar{Q}} |u| \leq e^{c_0 T}\big(\max\{\sup_{\Omega} |u_0|, \sup_{\partial\Omega \times (0,T)} |\varphi|\} + T \sup_{Q} |f|\big).$$

PROOF. Set

$$B = \max\{\sup_{\Omega} |u_0|, \sup_{\partial\Omega \times (0,T)} |\varphi|\}, \quad F = \sup_{Q} |f|,$$

and

$$v(x,t) = e^{c_0 t}(Ft + B).$$

Then

$$Lv = (c_0 + c)e^{c_0 t}(Ft + B) + e^{c_0 t} F \geq F \geq \pm Lu \quad \text{in } Q,$$

and

$$v \geq B \geq \pm u \quad \text{on } \partial_p Q.$$

Here we used $c + c_0 \geq 0$ and $e^{c_0 t} \geq 1$ in Q. By Corollary 5.16, we obtain

$$v \geq \pm u \quad \text{in } Q,$$

or

$$|u(x,t)| \leq e^{c_0 t}(Ft + B) \quad \text{for any } (x,t) \in Q_T.$$

This implies the desired estimate. □

In the following, we study a priori estimates for initial-value problems.

THEOREM 5.18. *Suppose u is a bounded C^2-solution of*

$$\begin{aligned} Lu \equiv u_t - \Delta u + cu &= f \quad \text{in } \mathbb{R}^n \times (0,T], \\ u(\cdot, 0) &= u_0 \quad \text{in } \mathbb{R}^n, \end{aligned}$$

for some bounded $f \in C(\mathbb{R}^n \times (0,T])$ and $u_0 \in C(\mathbb{R}^n)$. If $c \geq -c_0$ for a nonnegative constant c_0, then

$$\sup_{\mathbb{R}^n \times (0,T)} |u| \leq e^{c_0 T}\big(\sup_{\mathbb{R}^n \times (0,T)} |u_0| + T \sup_{\mathbb{R}^n \times (0,T)} |f|\big).$$

PROOF. Set

$$H = \mathbb{R}^n \times (0, T] = \{(x, t); x \in \mathbb{R}^n, t \in (0, T]\},$$

and

$$F = \sup_H |f|, \quad B = \sup_{\mathbb{R}^n} |u_0|.$$

We assume $\sup_H |u| \le M$ for a positive constant M since u is bounded. For any $R > 0$, set

$$Q_R = B_R \times (0, T].$$

Consider an auxiliary function

$$w(x, t) = e^{c_0 t}(Ft + B) + v_R(x, t) \pm u(x, t) \quad \text{in } Q_R,$$

where v_R is a function to be chosen. Then

$$Lw = (c + c_0)e^{c_0 t}(Ft + B) + e^{c_0 t}F \pm Lu + Lv_R \ge Lv_R \quad \text{in } Q_R.$$

Here we used $c + c_0 \ge 0$ and $e^{c_0 t} \ge 1$. Moreover,

$$w(\cdot, 0) = B + v_R(\cdot, 0) \pm u_0 \quad \text{in } B_R,$$

and

$$w \ge v_R \pm u \quad \text{on } \partial B_R \times (0, T].$$

We require

$$\begin{aligned} Lv_R \ge& 0 \quad \text{in } Q_R, \\ v_R(\cdot, 0) \ge& 0 \quad \text{in } B_R, \\ v_R \pm u \ge& 0 \quad \text{on } \partial B_R \times (0, T]. \end{aligned}$$

To construct such a v_R, we consider

$$v_R(x, t) = \frac{M}{R^2} e^{c_0 t}(2nt + |x|^2).$$

Obviously, $v_R \ge 0$ for $t = 0$ and $v_R \ge M$ on $|x| = R$. Next,

$$Lv_R = \frac{M}{R^2} e^{c_0 t}(c + c_0)(2nt + |x|^2) \ge 0 \quad \text{in } Q_R.$$

With such a v_R, we have

$$\begin{aligned} Lw \ge 0 \quad & \text{in } Q_R, \\ w \ge 0 \quad & \text{on } \partial_p Q_R. \end{aligned}$$

Hence Corollary 5.16 yields $w \ge 0$ in Q_R. Therefore for any fixed (x, t), by taking R large such that $(x, t) \in Q_R$, we have

$$|u(x, t)| \le e^{c_0 t}(Ft + B) + \frac{M}{R^2} e^{c_0 t}(2nt + |x|^2).$$

By letting $R \to +\infty$, we have

$$|u(x, t)| \le e^{c_0 t}(Ft + B) \quad \text{for any } (x, t) \in H.$$

This yields the desired estimate. $\square$

Next, we prove the uniqueness of solutions of the heat equation under an assumption of exponential growth.

THEOREM 5.19. *Let* $u \in C^{2,1}(\mathbb{R}^n \times (0,T)) \cap C(\mathbb{R}^n \times [0,T])$ *satisfy*

$$\begin{aligned} u_t - \Delta u &= 0 \quad \text{in } \mathbb{R}^n \times (0,T), \\ u(\cdot, 0) &= 0 \quad \text{in } \mathbb{R}^n. \end{aligned}$$

If

$$|u(x,t)| \le Me^{A|x|^2} \quad \text{for any } (x,t) \in \mathbb{R}^n \times (0,T),$$

for some positive constants M *and* A, *then* $u \equiv 0$ *in* $\mathbb{R}^n \times (0,T)$.

PROOF. For any constant $\alpha > A$, we prove

$$u(x,t) = 0 \quad \text{for any } (x,t) \in \mathbb{R}^n \times (0, \frac{1}{4\alpha}).$$

Our strategy is as follows. We construct an appropriate auxiliary function v_R in $B_R \times (0, 1/4\alpha)$ so that

$$|u(x,t)| \le v_R(x,t) \quad \text{for any } (x,t) \in B_R \times \big(0, \frac{1}{4\alpha}\big),$$

and for any fixed $(x,t) \in B_R \times (0, 1/4\alpha)$

$$v_R(x,t) \to 0 \quad \text{as } R \to \infty.$$

The function v_R is constructed with the help of the fundamental solution. Now we start our proof.

For any constant $R > 0$, consider

$$v_R(x,t) = \frac{Me^{(A-\alpha)R^2}}{(1-4\alpha t)^{\frac{n}{2}}} e^{\frac{\alpha |x|^2}{1-4\alpha t}} \quad \text{for } (x,t) \in B_R \times \big(0, \frac{1}{4\alpha}\big).$$

We note that v_R is simply the example we discussed preceding Theorem 5.5. Then

$$\partial_t v_R - \Delta v_R = 0 \quad \text{in } B_R \times \big(0, \frac{1}{4\alpha}\big).$$

Obviously,

$$v_R(\cdot, 0) \ge 0 = \pm u(\cdot, 0) \quad \text{in } B_R.$$

Next,

$$\begin{aligned} v_R(x,t) \ge & Me^{(A-\alpha)R^2} e^{\alpha R^2} = Me^{AR^2} \ge \pm u(x,t) \\ & \text{for any } (x,t) \in \partial B_R \times \big(0, \frac{1}{4\alpha}\big). \end{aligned}$$

By Corollary 5.13, the comparison principle, we have

$$\pm u(x,t) \le v_R(x,t) \quad \text{for any } (x,t) \in B_R \times \big(0, \frac{1}{4\alpha}\big),$$

or

$$|u(x,t)| \le v_R(x,t) \quad \text{for any } (x,t) \in B_R \times \big(0, \frac{1}{4\alpha}\big).$$

For any fixed $(x,t) \in \mathbb{R}^n \times (0, 1/4\alpha)$, we choose $R > |x|$. We note $v_R(x,t) \to 0$ as $R \to \infty$, by $\alpha > A$. This implies easily $u(x,t) = 0$. $\Box$

As the final application of the maximum principle, we provide another proof of interior gradient estimates by the maximum principle. We do this only for solutions of the heat equation. Recall for any r

$$Q_r = B_r \times (-r^2, 0].$$

THEOREM 5.20. *Suppose $u \in C^{2,1}(Q_1) \cap C(\bar{Q}_1)$ satisfies $u_t - \Delta u = 0$ in Q_1. Then*

$$\sup_{Q_{\frac{1}{2}}} |\nabla_x u| \leq C \sup_{\partial_p Q_1} |u|,$$

where C is a positive constant depending only on n.

PROOF. Set $v = |\nabla_x u|^2$. Then

$$v_t - \Delta v = -2 \sum_{i,j=1}^{n} u_{x_i x_j}^2 + 2 \sum_{i=1}^{n} u_{x_i} (u_t - \triangle u)_{x_i} = -2 \sum_{i,j=1}^{n} u_{x_i x_j}^2.$$

To get interior estimates, we intorduce a cut-off function. For any smooth function φ in $C^\infty(Q_1)$ with supp$\varphi \subset Q_{3/4}$, we have

$$(\varphi v)_t - \Delta(\varphi v) = (\varphi_t - \triangle\varphi)|\nabla_x u|^2 - 4 \sum_{i,j=1}^{n} \varphi_{x_i} u_{x_j} u_{x_i x_j} - 2\varphi \sum_{i,j=1}^{n} u_{x_i x_j}^2.$$

By taking $\varphi = \eta^2$ for some $\eta \in C^\infty(Q_1)$ with $\eta \equiv 1$ in $Q_{1/2}$ and supp$\eta \subset Q_{3/4}$, we have

$$\begin{aligned} &(\eta^2|\nabla_x u|^2)_t - \triangle(\eta^2|\nabla_x u|^2) \\ =&(2\eta\eta_t - 2\eta\triangle\eta - 2|\nabla_x \eta|^2)|\nabla_x u|^2 - 8\eta \sum_{i,j=1}^{n} \eta_{x_i} u_{x_j} u_{x_i x_j} - 2\eta^2 \sum_{i,j=1}^{n} u_{x_i x_j}^2. \end{aligned}$$

By the Cauchy inequality, we obtain

$$8|\eta\eta_{x_i} u_{x_j} u_{x_i x_j}| \leq 4\eta_{x_i}^2 u_{x_j}^2 + 2\eta^2 u_{x_i x_j}^2,$$

and hence

$$(\eta^2|\nabla_x u|^2)_t - \triangle(\eta^2|\nabla_x u|^2) \leq \Big(2\eta\eta_t - 2\eta\triangle\eta + 6|\nabla_x \eta|^2\Big)|\nabla_x u|^2 \leq C|\nabla_x u|^2,$$

where C is a positive constant depending only on η and n. Note

$$(u^2)_t - \triangle(u^2) = -2|\nabla_x u|^2 + 2u(u_t - \triangle u) = -2|\nabla_x u|^2.$$

By taking a constant α large enough, we get

$$(\partial_t - \Delta)(\eta^2|\nabla_x u|^2 + \alpha u^2) \leq (C - 2\alpha)|\nabla_x u|^2 \leq 0.$$

We apply Corollary 5.13 to get

$$\sup_{Q_1}(\eta^2|\nabla_x u|^2 + \alpha u^2) \leq \sup_{\partial_p Q_1} (\eta^2|\nabla_x u|^2 + \alpha u^2).$$

This implies the desired result since $\eta = 0$ on $\partial_p Q_1$ and $\eta = 1$ in $Q_{1/2}$. □

To end our discussions, we compare maximum principles for elliptic and parabolic equations of the following forms

$$L_e u = -\Delta u + c(x)u \quad \text{in } \Omega,$$

and

$$L_p u = u_t - \Delta u + c(x,t)u \quad \text{in } Q_T \equiv \Omega \times (0,T).$$

If $c \geq 0$, then

$$L_e u \leq 0 \text{ implies } u \text{ attains its nonnegative maximum on } \partial\Omega,$$
$$L_p u \leq 0 \text{ implies } u \text{ attains its nonnegative maximum on } \partial_p Q.$$

If $c \equiv 0$, the nonnegativity condition can be removed. For $c \geq 0$, comparison principles can be stated as follows:

$$L_e u \leq L_e w \text{ in } \Omega,\ u \leq v \text{ on } \partial\Omega \ \Rightarrow\ u \leq v \text{ in } \Omega,$$
$$L_p u \leq L_p w \text{ in } Q_T,\ u \leq v \text{ on } \partial_p Q_T \ \Rightarrow\ u \leq v \text{ in } Q_T.$$

We should note that comparison principles for parabolic equations hold for $c \geq -c_0$, for a nonnegative constant c_0.

In practice, we need to construct auxiliary functions for comparisons. Usually, we take $|x|^2$ or $e^{\pm\alpha|x|^2}$ for elliptic equations and $Kt + |x|^2$ for parabolic equations. Sometimes, auxiliary functions are constructed with the help of the fundamental solutions for the Laplace equation and the heat equation.

5.4. Harnack Inequalities

For positive harmonic functions, the Harnack inequality asserts that values in compact subsets are comparable. In this section, we study the Harnack inequality for positive solutions of the heat equation. In order to investigate a proper form of the Harnack inequality for solutions of the heat equation, we begin our discussion with the fundamental solution of the heat equation.

Consider for any $\xi \in \mathbb{R}^n$

$$u(x,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x-\xi|^2}{4t}} \quad \text{for any } x \in \mathbb{R}^n,\ t > 0.$$

Then u satisfies the heat equation $u_t - \Delta u = 0$ in $\mathbb{R}^n \times (0, \infty)$. Hence, for any (x_1, t_1) and $(x_2, t_2) \in \mathbb{R}^n \times (0, \infty)$

$$\frac{u(x_1,t_1)}{u(x_2,t_2)} = \left(\frac{t_2}{t_1}\right)^{\frac{n}{2}} e^{\frac{|x_2-\xi|^2}{4t_2} - \frac{|x_1-\xi|^2}{4t_1}}.$$

Note

$$\frac{(p+q)^2}{a+b} \leq \frac{p^2}{a} + \frac{q^2}{b} \quad \text{for any } a, b > 0,$$

where equality holds if and only if $bp = aq$. This implies for any $t_2 > t_1 > 0$

$$\frac{|x_2-\xi|^2}{t_2} \leq \frac{|x_2-x_1|^2}{t_2-t_1} + \frac{|x_1-\xi|^2}{t_1},$$

where equality holds if and only if

$$\xi = \frac{t_2 x_1 - t_1 x_2}{t_2 - t_1}.$$

Therefore,

$$u(x_1,t_1) \leq \left(\frac{t_2}{t_1}\right)^{\frac{n}{2}} e^{\frac{|x_2-x_1|^2}{4(t_2-t_1)}} u(x_2,t_2) \quad \text{for any } t_2 > t_1 > 0,$$

where equality holds if ξ is chosen as above. This simple calculation suggests that the Harnack inequality for the heat equation has an "evolution" feature: the value of a positive solution at a certain time is controlled from above by the value at

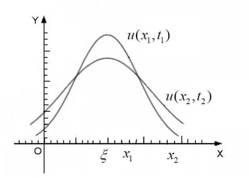

FIGURE 5.5. Values of the fundamental solution at different time.

a later time. In general, we cannot estimate the value of a positive solution at a certain time from above by values at a previous time. Hence if we attempt to establish the following estimate

$$u(x_1, t_1) \leq Cu(x_2, t_2),$$

the constant C should depend on t_2/t_1, $|x_2 - x_1|$ and most importantly $(t_2 - t_1)^{-1}(> 0)$. Usually, we choose following regions for points $(x_1, t_1) \in D_1$ and $(x_2, t_2) \in D_2$.

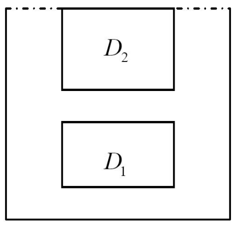

FIGURE 5.6. Domains for Harnack inequalities.

Suppose u is a positive function and set $v = \log u$. In order to derive an estimate for the quotient

$$\frac{u(x_1, t_1)}{u(x_2, t_2)},$$

it suffices to get an estimate for the difference

$$v(x_1, t_1) - v(x_2, t_2).$$

To this end, we need an estimate of v_t and $|\nabla v|$. For a hint of proper forms, we again turn our attention to the fundamental solution of the heat equation.

Consider

$$u(x, t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}} \quad \text{for any } x \in \mathbb{R}^n, t > 0.$$

Then

$$v(x, t) = \log u(x, t) = -\frac{n}{2} \log(4\pi t) - \frac{|x|^2}{4t},$$

and hence

$$v_t = -\frac{n}{2t} + |\nabla v|^2.$$

We have the following result for arbitrary positive solutions of the heat equation.

THEOREM 5.21. *Suppose $u \in C^{2,1}(\mathbb{R}^n \times (0,T])$ satisfies*

$$u_t = \Delta u, \quad u > 0 \quad \text{in } \mathbb{R}^n \times (0,T].$$

Then $v = \log u$ satisfies

$$v_t + \frac{n}{2t} \geq |\nabla v|^2 \quad \text{in } \mathbb{R}^n \times (0,T]. \tag{5.5}$$

The inequality (5.5) is referred to as *the differential Harnack inequality.* We first prove a corollary.

COROLLARY 5.22. *Suppose $u \in C^{2,1}(\mathbb{R}^n \times (0,T])$ satisfies*

$$u_t = \Delta u, \quad u > 0 \quad \text{in } \mathbb{R}^n \times (0,T].$$

Then for any $(x_1,t_1),(x_2,t_2) \in \mathbb{R}^n \times (0,T]$ with $t_2 > t_1 > 0$

$$\frac{u(x_1,t_1)}{u(x_2,t_2)} \leq \left(\frac{t_2}{t_1}\right)^{\frac{n}{2}} \exp\{\frac{|x_2-x_1|^2}{4(t_2-t_1)}\}. \tag{5.6}$$

PROOF. Along any path $\big(t, x(t)\big)$ with $x(t_i) = x_i$, $i = 1, 2$, we have by Theorem 5.21

$$\frac{d}{dt}v(x(t),t) = v_t + \nabla v \cdot \frac{dx}{dt} \geq |\nabla v|^2 + \nabla v \cdot \frac{dx}{dt} - \frac{n}{2t}.$$

By completing square, we obtain

$$\frac{d}{dt}v(x(t),t) \geq -\frac{1}{4}\left|\frac{dx}{dt}\right|^2 - \frac{n}{2t},$$

and hence

$$v(x_2,t_2) \geq v(x_1,t_1) - \frac{n}{2}\log\frac{t_2}{t_1} - \frac{1}{4}\int_{t_1}^{t_2}\left|\frac{dx}{dt}\right|^2 dt.$$

To seek an optimal path which makes the last integral minimal, we require

$$\frac{d^2x}{dt^2} = 0$$

along the path. Hence we set

$$x = at + b \quad \text{for some } a, b \in \mathbb{R}^n.$$

By $x_i = at_i + b$, $i = 1, 2$, we take

$$a = \frac{x_2 - x_1}{t_2 - t_1} \quad \text{and} \quad b = \frac{t_2x_1 - t_1x_2}{t_2 - t_1}.$$

Then,

$$\int_{t_1}^{t_2}\left|\frac{dx}{dt}\right|^2 dt = \frac{|x_2-x_1|^2}{t_2-t_1}.$$

Therefore, we obtain

$$v(x_2,t_2) \geq v(x_1,t_1) - \frac{n}{2}\log\frac{t_2}{t_1} - \frac{1}{4}\frac{|x_2-x_1|^2}{t_2-t_1},$$

or

$$\frac{u(x_2,t_2)}{u(x_1,t_1)} \geq \left(\frac{t_1}{t_2}\right)^{\frac{n}{2}} \exp\{-\frac{|x_2-x_1|^2}{4(t_2-t_1)}\}.$$

This is the desired estimate. □

Now we start to prove the differential Harnack inequality. The basic idea is to apply the maximum principle to an appropriate combination of derivatives of v. In our case, we consider $|\nabla v|^2 - v_t$ and intend to derive an upper bound. First, we need to derive a parabolic equation satisfied by $|\nabla v|^2 - v_t$. A careful analysis shows there are uncontrollable terms in this equation. Therefore, we should introduce a parameter $\alpha \in (0,1)$ and consider $\alpha|\nabla v|^2 - v_t$ instead. After we apply the maximum principle, we let $\alpha \to 1-$. Since the proof below is quite involved, we divide it into several steps.

Proof of Theorem 5.21. *Step 1.* We first derive some equations involving derivatives of $v = \log u$. A simple calculation shows

$$v_t = \Delta v + |\nabla v|^2.$$

Note that the differential Harnack inequality asserts

$$\Delta v + \frac{n}{2t} \geq 0,$$

which implies Δv is almost concave.

Consider $w = \Delta v$. Then we have

$$w_t = \Delta v_t = \Delta(\Delta v + |\nabla v|^2) = \Delta w + \Delta|\nabla v|^2.$$

Since

$$\Delta|\nabla v|^2 = 2|\nabla^2 v|^2 + 2\nabla v \cdot \nabla(\Delta v) = 2|\nabla^2 v|^2 + 2\nabla v \cdot \nabla w,$$

then

$$w_t - \Delta w - 2\nabla v \cdot \nabla w = 2|\nabla^2 v|^2. \tag{5.7}$$

Note that the uncontrolled expression ∇v appears as a coefficient of the equation in (5.7). It is convenient to derive an equation for it. Set $\tilde{w} = |\nabla v|^2$. Then,

$$\begin{aligned}
\tilde{w}_t &= 2\nabla v \cdot \nabla v_t = 2\nabla v \cdot \nabla(\Delta v + |\nabla v|^2) \\
&= 2\nabla v \cdot \nabla(\Delta v) + 2\nabla v \cdot \nabla \tilde{w} \\
&= \Delta|\nabla v|^2 - 2|\nabla^2 v|^2 + 2\nabla v \cdot \nabla \tilde{w} \\
&= \Delta \tilde{w} + 2\nabla v \cdot \nabla \tilde{w} - 2|\nabla^2 v|^2.
\end{aligned}$$

Therefore, $\tilde{w}$ satisfies

$$\tilde{w}_t - \Delta \tilde{w} - 2\nabla v \cdot \nabla \tilde{w} = -2|\nabla^2 v|^2. \tag{5.8}$$

Note by the Cauchy inequality

$$|\nabla^2 v|^2 = \sum_{i,j=1}^n v_{x_i x_j}^2 \geq \sum_{i=1}^n v_{x_i x_i}^2 \geq \frac{1}{n}(\sum_{i=1}^n v_{x_i x_i})^2 = \frac{1}{n}(\Delta v)^2.$$

So (5.7) implies

$$w_t - \Delta w - 2\nabla v \cdot \nabla w \geq \frac{2}{n} w^2.$$

Step 2. Set for a constant $\alpha \in (0,1]$

$$F = \alpha|\nabla v|^2 - v_t.$$

Then

$$F = \alpha|\nabla v|^2 - \Delta v - |\nabla v|^2 = -\Delta v - (1-\alpha)|\nabla v|^2 = -w - (1-\alpha)\tilde{w},$$

and hence by (5.7) and (5.8)

$$F_t - \Delta F - 2\nabla v \cdot \nabla F = -2\alpha|\nabla^2 v|^2.$$

Next, we estimate $|\nabla^2 v|^2$ by F. Note

$$\begin{aligned}
|\nabla^2 v|^2 &\geq \frac{1}{n}(\Delta v)^2 = \frac{1}{n}\big(|\nabla v|^2 - v_t\big)^2 = \frac{1}{n}((1-\alpha)|\nabla v|^2 + F)^2 \\
&= \frac{1}{n}\big(F^2 + (1-\alpha)^2|\nabla v|^4 + 2(1-\alpha)|\nabla v|^2 F\big) \\
&\geq \frac{1}{n}\big(F^2 + 2(1-\alpha)|\nabla v|^2 F\big).
\end{aligned}$$

We obtain

$$F_t - \Delta F - 2\nabla v \cdot \nabla F \leq -\frac{2\alpha}{n}\big(F^2 + 2(1-\alpha)|\nabla v|^2 F\big). \tag{5.9}$$

We should point out that the term $-|\nabla v|^2$ at the right-hand side plays an important role later on.

Step 3. Now we introduce a cut-off function. Set

$$G = tF\varphi,$$

where $\varphi \geq 0$ and $\varphi \in C_0^\infty(\mathbb{R}^n)$. We derive a differential inequality for G. Note

$$\begin{aligned}
G_t &= F\varphi + tF_t\varphi, \\
\nabla G &= t\nabla F\varphi + tF\nabla\varphi, \\
\Delta G &= t\Delta F\varphi + 2t\nabla F \cdot \nabla\varphi + tF\Delta\varphi.
\end{aligned}$$

Then,

$$\begin{aligned}
tF_t\varphi &= G_t - \frac{G}{t}, \\
t\nabla F\varphi &= \nabla G - \frac{\nabla\varphi}{\varphi}G, \\
t\Delta F\varphi &= \Delta G - \frac{2}{\varphi}\nabla\varphi \cdot \left(\nabla G - \frac{\nabla\varphi}{\varphi}G\right) - \frac{\Delta\varphi}{\varphi}G \\
&= \Delta G - 2\frac{\nabla\varphi}{\varphi} \cdot \nabla G + 2\frac{|\nabla\varphi|^2}{\varphi^2}G - \frac{\Delta\varphi}{\varphi}G.
\end{aligned}$$

Multiplying (5.9) by $t^2\varphi^2$ and substituting above inequalities, we obtain

$$\begin{aligned}
&t\varphi(G_t - \Delta G) + 2t(\nabla\varphi - \varphi\nabla v) \cdot \nabla G \\
&\leq G\left\{\varphi - \frac{2\alpha}{n}G + t\left(2\frac{|\nabla\varphi|^2}{\varphi} - \Delta\varphi - \frac{4\alpha(1-\alpha)}{n}\varphi|\nabla v|^2 - 2\nabla v \cdot \nabla\varphi\right)\right\}.
\end{aligned}$$

To eliminate $|\nabla v|$ from the right-hand side, we complete square for the last two terms. Here we need $\alpha < 1$! Otherwise, we cannot control the expression $-2\nabla v \cdot \nabla\varphi$ in the right-hand side. Then we have

$$\begin{aligned}
&t\varphi(G_t - \Delta G) + 2t(\nabla\varphi - \varphi\nabla v) \cdot \nabla G \\
&\leq G\left\{\varphi - \frac{2\alpha}{n}G + t\left(2\frac{|\nabla\varphi|^2}{\varphi} - \Delta\varphi - \frac{n}{4\alpha(1-\alpha)}\frac{|\nabla\varphi|^2}{\varphi}\right)\right\}.
\end{aligned}$$

We point out that there are no unknown expressions in the right-hand side except G. By choosing $\varphi = \eta^2$ for some $\eta \geq 0$ and $\eta \in C_0^\infty(\mathbb{R}^n)$, we get

$$\begin{aligned} &t\eta^2(G_t - \Delta G) + 2t(2\eta\nabla\eta - \eta^2\nabla v)\cdot\nabla G \\ &\leq G\left\{\eta^2 - \frac{2\alpha}{n}G + t\left(6|\nabla\eta|^2 - 2\eta\Delta\eta - \frac{n}{\alpha(1-\alpha)}|\nabla\eta|^2\right)\right\}. \end{aligned}$$

Now we fix a cut-off function $\eta_0 \in C_0^\infty(B_1)$, with $0 < \eta_0 \leq 1$ in B_1 and $\eta_0 = 1$ in $B_{1/2}$. For an arbitrary $R \geq 1$, we consider $\eta(x) = \eta_0(x/R)$. Then we obtain in $B_R \times (0, T)$

$$t\eta^2(G_t - \Delta G) + 2t(2\eta\nabla\eta - \eta^2\nabla v)\cdot\nabla G \leq G\left(1 - \frac{2\alpha}{n}G + \frac{C_\alpha t}{R^2}\right),$$

where C_α is a positive constant depending only on α and η_0.

Step 4. We claim

$$1 - \frac{2\alpha}{n}G + \frac{C_\alpha t}{R^2} \geq 0 \quad \text{in } B_R \times (0,T). \tag{5.10}$$

Note that G vanishes on the parabolic boundary of $B_R \times (0,T)$ since $G = tF\varphi$. Suppose (5.10) does not hold. Then $1 - \frac{2\alpha}{n}G + \frac{C_\alpha t}{R^2}$ has a negative minimum at $(x_0, t_0) \in B_R \times (0, T]$. Moreover, we have

$$G(x_0, t_0) > 0,$$

and

$$G_t \geq 0, \quad \nabla G = 0, \quad \Delta G \leq 0 \quad \text{at } (x_0, t_0).$$

Then at (x_0, t_0), we get

$$\begin{aligned} 0 &\leq t\eta^2(G_t - \Delta G) + 2t(2\eta\nabla\eta - \eta^2\nabla v)\cdot\nabla G \\ &\leq G\left(1 - \frac{2\alpha}{n}G + \frac{C_\alpha t}{R^2}\right) < 0. \end{aligned}$$

This is a contradiction. Hence (5.10) holds in $B_R \times (0,T)$.

Therefore, we obtain

$$1 - \frac{2\alpha}{n}t\eta^2(\alpha|\nabla v|^2 - v_t) + \frac{C_\alpha t}{R^2} \geq 0 \quad \text{in } B_R \times (0,T). \tag{5.11}$$

For any fixed $(x,t) \in \mathbb{R}^n \times (0,T)$, choose $R > |x|$. Recall $\eta(x) = \eta_0(x/R)$ and $\eta_0 = 1$ in $B_{1/2}$. Letting $R \to +\infty$, we obtain

$$1 - \frac{2\alpha}{n}t(\alpha|\nabla v|^2 - v_t) \geq 0.$$

We then let $\alpha \to 1-$ and get the desired estimate. □

By a simple modification of the proof above, we obtain the following differential Harnack inequality for positive solutions in a finite region.

Theorem 5.23. *Suppose u satisfies*

$$u_t - \Delta u = 0, \; u > 0 \quad \text{in } B_1 \times (0,1).$$

Then for any $\alpha \in (0,1)$, $v = \log u$ satisfies

$$v_t - \alpha|\nabla v|^2 + \frac{n}{2\alpha t} + C \geq 0 \quad \text{in } B_{\frac{1}{2}} \times (0,1),$$

where C is a positive constant depending only on n and α.

PROOF. We simply take $R = 1$ in (5.11). $\square$

Now we state the Harnack inequality in a finite region.

COROLLARY 5.24. *Suppose u satisfies*

$$u_t - \Delta u = 0, \ u > 0 \quad \text{in } B_1 \times (0, 1).$$

Then for any $(x_1, t_1), (x_2, t_2) \in B_{1/2} \times (0, 1)$ with $t_2 > t_1$,

$$u(x_1, t_1) \leq Cu(x_2, t_2),$$

where C is a positive constant depending only on t_2/t_1 and $(t_2 - t_1)^{-1}$.

The proof is left as an exercise.

Exercises

(1) Prove the following statements by straightforward calculations.
  (a) $K(x, t) = t^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}$ satisfies the heat equation for $t > 0$.
  (b) For any $\alpha > 0$, $G(x, t) = (1-4\alpha t)^{-\frac{n}{2}} e^{\frac{\alpha |x|^2}{1-4\alpha t}}$ satisfies the heat equation for $t < 1/4\alpha$.

(2) For $u_0 \in C_0(\mathbb{R}^n)$ and u defined in (5.4), prove $\lim_{t\to\infty} u(x, t) = 0$ uniformly in x.

(3) For $u_0 \in C_0(\mathbb{R}^n)$ and u defined in (5.4), find an appropriate condition on u_0 so that u is C^k up to $t = 0$ and give a proof.

(4) Let u_0 be a bounded and continuous function in $[0, \infty)$ with $u_0(0) = 0$. Find an integral representation for the solution of the following problem

$$\begin{aligned} u_t - u_{xx} &= 0 \quad \text{for } x > 0, t > 0, \\ u(x, 0) &= u_0(x) \quad \text{for } x > 0, \\ u(0, t) &= 0 \quad \text{for } t > 0. \end{aligned}$$

(5) Prove that u constructed in the proof of Proposition 5.7 is smooth in $\mathbb{R} \times \mathbb{R}$.

(6) Suppose the initial temperature u_0 of an infinite rod is given by

$$u_0 = \begin{cases} 1 & x \in [0, 1], \\ 0 & \text{otherwise.} \end{cases}$$

At what time does a point with coordinate x attain the highest temperature? In particular, consider points $x = 1/2$ and $x = 3/2$.

(7) Suppose $u \in C^2(Q_T) \cap C(\bar{Q}_T)$ satisfies

$$\begin{aligned} u_t - \Delta u + c(x, t)u &= -u^2 \quad \text{in } Q_T = \Omega \times (0, T), \\ u|_{t=0} &= u_0, \\ u|_{\partial\Omega\times(0,T)} &= 0, \end{aligned}$$

where $c(x,t) \in C(\bar{Q}_T)$ with $c \geq -c_0$ and $u_0 \in C(\bar{\Omega})$ with $u_0 \geq 0$. Prove

$$0 \leq u \leq e^{c_0 t} \sup_{\Omega} |u_0| \quad \text{in } \Omega.$$

(8) Suppose $u \in C^2(Q_T) \cap C(\bar{Q}_T)$ satisfies

$$\begin{aligned} u_t - \Delta u =& e^{-u} - f(x) \quad \text{in } Q_T, \\ u|_{t=0} =& u_0 \quad \text{in } \Omega, \\ u|_{\partial\Omega \times (0,T)} =& \varphi. \end{aligned}$$

Prove

$$-M \leq u \leq e^{MT} + M \quad \text{in } \Omega,$$

where

$$M = T \sup_{\Omega} |f| + \sup\{\sup_{\Omega} |u_0|, \sup_{\partial\Omega \times (0,T)} |\varphi|\}.$$

(9) Let $Q = (0,l) \times (0,\infty)$ and $u \in C^3(Q) \cap C^1(\bar{Q})$ satisfy

$$\begin{aligned} u_t - u_{xx} =& 0 \quad \text{in } Q, \\ u|_{t=0} =& \varphi \quad \text{in } (0,l) \\ u|_{x=0} = 0,\ u|_{x=l} =& 0 \quad \text{for any } t > 0, \end{aligned}$$

for a function $\varphi \in C^1[0,l]$ with $\varphi(0) = \varphi(l)$. Prove

$$\sup_{Q} |u_x| \leq \sup_{[0,l]} |\varphi'|.$$

(10) Let $h(t)$ be a continuous increasing function with $h(0) \geq 0$ and

$$Q_T = \{(x,t); 0 \leq x \leq h(t), 0 \leq t \leq T\}.$$

Let g be an increasing nonnegative function in $[0,T]$ and φ be a decreasing C^1-function in $[0,h(0)]$ with where $\varphi(0) = g(0)$ and $\varphi(h(0)) = 0$. Suppose $u \in C^{2,1}(Q)$ is a solution of

$$\begin{aligned} u_t - u_{xx} =& 0 \quad \text{in } Q, \\ u(x,0) =& \varphi(x) \quad \text{for } x \in [0,h(0)], \\ u(0,t) =& g(t), \\ u(h(t),t) =& 0, \end{aligned}$$

Prove $\partial_x u \leq 0$ in Q_T.

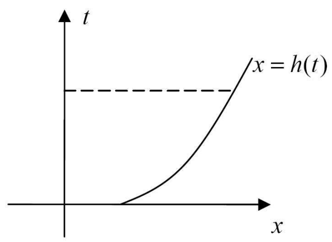

FIGURE 5.7. A domain.

(11) Let $Q = (0, l) \times (0, \infty)$ and $u_h = u_h(x, t)$ be a solution of

$$\begin{aligned} u_t - u_{xx} =& 0 \quad \text{in } Q, \\ u(x, 0) =& 0 \quad \text{in } (0, l), \\ u|_{x=l} =& 0, \\ \big(\partial_x u + h(u_0 - u)\big)|_{x=0} =& 0, \end{aligned}$$

for positive constants u_0 and h. Prove

(a) $0 \le u \le u_0$ in $(0, l) \times (0, \infty)$;

(b) u_h is monotone increasing for h.

(12) Let $u_1, \cdots, u_m$ be $C^{2,1}$-functions in $Q = \Omega \times (0, T] \subset \mathbb{R}^n \times \mathbb{R}_+$ satisfying

$$\partial_t u_i = \Delta u_i \quad \text{in } Q, \text{ for } i = 1, \cdots, m.$$

Suppose f is a convex function in $\mathbb{R}^m$. Then

$$\sup_Q f(u_1, \cdots, u_m) \le \sup_{\partial_p Q} f(u_1, \cdots, u_m).$$

(13) Let $u \in C^{2,1}(\mathbb{R}^n \times (0, T])$ satisfy

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

Suppose u and ∇u are bounded in $\mathbb{R}^n \times (0, T]$. Prove

$$\sup_{\mathbb{R}^n \times (0,T]} |\nabla u| \le \frac{1}{\sqrt{2t}} \sup_{\mathbb{R}^n} |u(\cdot, 0)|.$$

Hint: With $|u| \le M$ at $t = 0$, consider

$$w = u^2 + 2t|\nabla u|^2 - M^2.$$

(14) Prove Corollary 5.24.

CHAPTER 6

Wave Equations

In this chapter, we study the wave equation in $\mathbb{R}^n \times \mathbb{R}_+$. The wave equation represents vibrations of strings or propagation of sound waves in tubes for $n = 1$, waves on the surface of shallow water for $n = 2$, and acoustic or light waves for $n = 3$. In Section 6.1, we discuss initial-value problems and initial/boundary-value problems for the one-dimensional wave equation. In Section 6.2, we study initial-value problems for the wave equation in higher dimensional spaces. Then in Section 6.3, we derive energy estimates for solutions of initial-value problems.

6.1. One-Dimensional Wave Equations

In this section, we discuss initial-value problems and various initial/boundary-value problems for the one-dimensional wave equation. We first study initial-value problems.

For $f \in C(\mathbb{R} \times \mathbb{R}_+)$, $\varphi \in C^2(\mathbb{R})$ and $\psi \in C^1(\mathbb{R})$, we seek a solution $u \in C^2(\mathbb{R} \times \mathbb{R}_+)$ of the following problem

$$\begin{aligned} u_{tt} - u_{xx} &= f \quad \text{in } \mathbb{R} \times \mathbb{R}_+, \\ u(\cdot, 0) = \varphi,\ u_t(\cdot, 0) &= \psi \quad \text{on } \mathbb{R}. \end{aligned} \tag{6.1}$$

We will use different methods to derive expressions of solutions in various special cases.

As discussed in Section 3.1, characteristic curves for the one-dimensional wave equation are given by straight lines $t = \pm x + c$. In particular, for any $(\bar{x}, \bar{t}) \in \mathbb{R} \times \mathbb{R}_+$, there are two characteristic curves through $(\bar{x}, \bar{t})$ given by

$$t - x = \bar{t} - \bar{x} \quad \text{and} \quad t + x = \bar{t} + \bar{x}.$$

We first consider the case $f \equiv 0$. By setting

$$\xi = x - t,\ \eta = x + t,$$

we have

$$u_{\xi\eta} = 0,$$

and hence

$$u(\xi, \eta) = g(\xi) + h(\eta),$$

for some functions g and h in $\mathbb{R}$. Therefore,

$$u(x, t) = g(x - t) + h(x + t). \tag{6.2}$$

With initial values, we have

$$\varphi(x) = g(x) + h(x), \quad \psi(x) = -g'(x) + h'(x).$$

Then

$$g(x) = \frac{1}{2}\varphi(x) - \frac{1}{2}\int_0^x \psi(s)ds + c$$
$$h(x) = \frac{1}{2}\varphi(x) + \frac{1}{2}\int_0^x \psi(s)ds - c,$$

for a constant c. Hence,

$$u(x,t) = \frac{1}{2}\big(\varphi(x-t) + \varphi(x+t)\big) + \frac{1}{2}\int_{x-t}^{x+t} \psi(s)ds. \tag{6.3}$$

This is called the *d'Alembert's formula*. Such a formula clearly shows that regularity of $u(\cdot, t)$ for any $t > 0$ is the same as that of the initial value $u(\cdot, 0)$. There is no improvement of regularity.

We see from (6.3) that $u(x, t)$ is determined uniquely by initial values in the interval $[x-t, x+t]$ of the x-axis whose end points are cut out by the characteristic curves through the point (x, t). This interval represents the *domain of dependence* for the solution at the point (x, t). Conversely, the initial values at a point $(x_0, 0)$ of the x-axis influence $u(x, t)$ at points (x, t) in the wedge-shaped region bounded by characteristic curves through $(x_0, 0)$, i.e., for $x_0 - t < x < x_0 + t$.

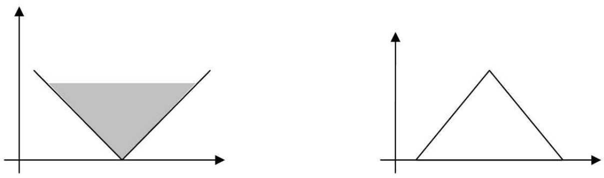

FIGURE 6.1. The domain of influence and the domain of dependence.

Now we derive an important formula for the solution of the wave equation. Let u be a C^2-solution of

$$u_{tt} - u_{xx} = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}_+,$$

and A, B, C, D be four vertices of a parallelogram bounded by four characteristic curves in $\mathbb{R} \times \mathbb{R}^+$ as follows. (This parallelogram is in fact a rectangle.) Then

$$u(A) + u(D) = u(B) + u(C). \tag{6.4}$$

In other words, the sums of the values of u in opposite vertices are equal. This follows easily from (6.2).

Next, we use the method of characteristics to solve (6.1) if $f \equiv 0$ and $\varphi \equiv 0$. By writing

$$u_{tt} - u_{xx} = (\partial_t + \partial_x)(\partial_t - \partial_x)u,$$

we decompose (6.1) to two initial-value problems for first order PDEs as follows

$$\begin{aligned} u_t - u_x &= v \quad \text{in } \mathbb{R} \times \mathbb{R}_+, \\ u(\cdot, 0) &= 0 \quad \text{on } \mathbb{R}, \end{aligned} \tag{6.5}$$

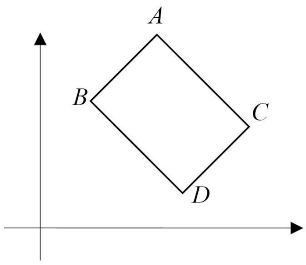

FIGURE 6.2. A rectangle.

and

$$(6.6)\qquad \begin{aligned} v_t + v_x &= 0 \quad \text{in } \mathbb{R} \times \mathbb{R}_+, \\ v(\cdot, 0) &= \psi \quad \text{on } \mathbb{R}. \end{aligned}$$

We note that integral curves are exactly characteristic curves for first order linear PDEs in the plane. First, we discuss (6.6). Its associated integral curves are given by $x = t + c$. Along these curves, the equation in (6.6) becomes

$$\frac{dv(x(t), t)}{dt} = 0.$$

Hence

$$v(x(t), t) = v(x(0), 0) = \psi(x(0)).$$

For any $(\bar{x}, \bar{t})$, we have $x(t) = t + (\bar{x} - \bar{t})$ and hence

$$v(\bar{x}, \bar{t}) = v(x(\bar{t}), \bar{t}) = \psi(x(0)) = \psi(\bar{x} - \bar{t}).$$

Therefore, (6.6) has a solution

$$v(x, t) = \psi(x - t).$$

For (6.5), its integral curves are given by $x = -t + c$ and hence $x(t) = -t + \bar{x} + \bar{t}$ for any $(\bar{x}, \bar{t})$. Along these curves, the equation in (6.5) becomes

$$\frac{du(x(t), t)}{dt} = v(x(t), t) = \psi(x(t) - t).$$

Then

$$u\big(x(t), t\big) = \int_0^t \psi\big(x(\tau) - \tau\big) d\tau,$$

and hence

$$u(\bar{x}, \bar{t}) = \int_0^{\bar{t}} \psi\big(x(\tau) - \tau\big) d\tau = \int_0^{\bar{t}} \psi(\bar{x} + \bar{t} - 2\tau) d\tau = \frac{1}{2} \int_{\bar{x} - \bar{t}}^{\bar{x} + \bar{t}} \psi(s) ds.$$

Therefore,

$$u(x, t) = \frac{1}{2} \int_{x-t}^{x+t} \psi(s) ds.$$

This is simply a special case of the D'Alemberts formula.

Now we derive an expression of solutions in the general case. For any $(x, t) \in \mathbb{R} \times \mathbb{R}_+$, consider the triangle

$$C_1(x, t) = \{(y, \tau); |y - x| < t - \tau, \tau > 0\}.$$

This is exactly the cone we introduced in Section 2.3 for $n = 1$. The boundary of $C_1(x,t)$ consists of three parts

$$L_+ = \{(y,\tau); \tau = -y + x + t,\ 0 < \tau < t\},$$
$$L_- = \{(y,\tau); \tau = y - x + t,\ 0 < \tau < t\},$$

and

$$L_0 = \{(y,0); x - t < y < x + t\}.$$

We note that L_+ and L_- are parts of the characteristic curves through (x,t). Let $\gamma = (\gamma_1, \gamma_2)$ be the unit exterior normal vector of $\partial C_1(x,t)$. Then

$$\gamma = \begin{cases} (1,1)/\sqrt{2} & \text{on } L_+, \\ (1,-1)/\sqrt{2} & \text{on } L_-, \\ (0,-1) & \text{on } L_0. \end{cases}$$

Hence by the Green's formula, we have

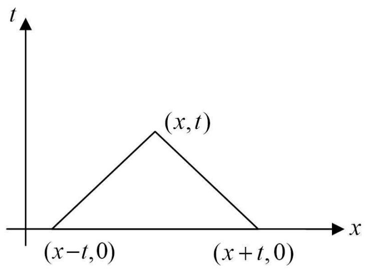

FIGURE 6.3. Integral curves.

$$\int_{C_1(x,t)} f = \int_{C_1(x,t)} (u_{tt} - u_{xx}) = \int_{\partial C_1(x,t)} (u_t\gamma_2 - u_x\gamma_1)$$
$$= -\int_{x-t}^{x+t} u_t(s,0)ds + \int_{L_+} \frac{1}{\sqrt{2}}(u_t - u_x) + \int_{L_-} \frac{1}{\sqrt{2}}(u_t + u_x),$$

where the orientation of integrals on L_+ and L_- is counterclockwise. Note that $(\partial_t - \partial_x)/\sqrt{2}$ is a directional derivative along L_+ with unit length and with the direction matching the orientation of the integral on L_+. Hence

$$\int_{L_+} \frac{1}{\sqrt{2}}(u_t - u_x) = u(x,t) - u(x+t,0).$$

On the other hand, $(\partial_t + \partial_x)/\sqrt{2}$ is a directional derivative along L_- with unit length and with the opposite direction matching the orientation of the integral on L_-. Hence

$$\int_{L_-} \frac{1}{\sqrt{2}}(u_t + u_x) = -\big(u(x-t,0) - u(x,t)\big).$$

Therefore, a simple substitution yields

$$u(x,t) = \frac{1}{2}\big(\varphi(x+t)+\varphi(x-t)\big)+\frac{1}{2}\int_{x-t}^{x+t} \psi(s)ds+\frac{1}{2}\int_0^t d\tau \int_{x-(t-\tau)}^{x+(t-\tau)} f(y,\tau)dy. \tag{6.7}$$

Theorem 6.1. *Suppose* $\varphi \in C^2(\mathbb{R}), \psi \in C^1(\mathbb{R})$ *and* $f \in C^1(\mathbb{R} \times [0, \infty))$. *Then* u *in (6.7) is a* C^2*-solution of (6.1).*

The proof is a straightforward calculation and is omitted. Obviously, C^2-solutions of (6.1) are unique.

The formula (6.7) illustrates that the value $u(x, t)$ is determined by f in the triangle $C_1(x, t)$ and by φ and ψ on the interval $[x - t, x + t] \times \{0\}$. This has the same feature as solutions of initial value problems of first-order linear differential equations discussed in Section 2.3.

Without using the explicit expression of solutions in (6.7), we can also derive energy estimates, the estimates of the L^2-norms of solutions of (6.1) and their derivatives. For any constants $0 < T < \bar{t}$, we use the following domain for energy estimates

$$\{(x, t); |x| < \bar{t} - t, 0 < t < T\}.$$

We postpone derivation until the final section of this chapter.

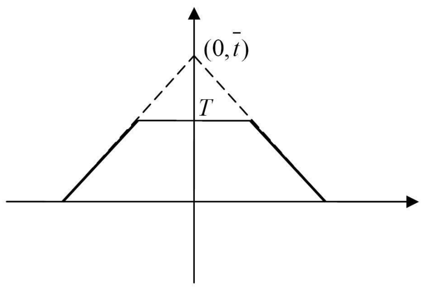

Figure 6.4. A domain of integration.

In the rest of this section, we study initial/boundary-value problems. For simplicity, we only discuss homogeneous wave equations.

First, we study half-space problems. Assume $\varphi \in C^2[0, \infty)$, $\psi \in C^1[0, \infty)$ and $\alpha \in C^2[0, \infty)$, and consider

$$\begin{aligned} &u_{tt} - u_{xx} = 0 \quad \text{in } [0, \infty) \times (0, \infty), \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = \psi \quad \text{on } [0, \infty), \\ &u(0, t) = \alpha(t) \quad \text{for } t > 0. \end{aligned} \tag{6.8}$$

We construct a C^2-solution under appropriate compatibility conditions. If (6.8) admits a solution which is C^2 in $[0, \infty) \times [0, \infty)$, a simple calculation shows

$$\varphi(0) = \alpha(0), \ \psi(0) = \alpha'(0), \ \varphi''(0) = \alpha''(0). \tag{6.9}$$

We first consider the case $\alpha \equiv 0$ and use the method of extension to solve (6.8). The compatibility condition (6.9) has the following form

$$\varphi(0) = 0, \quad \psi(0) = 0, \quad \varphi''(0) = 0.$$

Now we assume this holds and proceed to construct a C^2-solution of (6.8). We extend φ and ψ as odd functions in $\mathbb{R}$. The extended φ and ψ are C^2 and C^1 in $\mathbb{R}$

respectively. Let u be the unique C^2-solution of

$$\begin{aligned} u_{tt} - u_{xx} &= 0 \quad \text{in } \mathbb{R} \times \mathbb{R}_+, \\ u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) &= \psi \quad \text{in } \mathbb{R}. \end{aligned}$$

In fact, u is given by the D'Alemberts formula (6.3). We now prove that $u(x, t)$ is a solution of (6.8) when we restrict $x \in [0, \infty)$. We only need to prove

$$u(0, t) = 0 \quad \text{for any } t > 0.$$

In fact, for $v(x, t) = -u(-x, t)$, a simple calculation yields

$$\begin{aligned} v_{tt} - v_{xx} &= 0 \quad \text{in } \mathbb{R} \times \mathbb{R}_+, \\ v(\cdot, 0) = \varphi, \ v_t(\cdot, 0) &= \psi \quad \text{in } \mathbb{R}. \end{aligned}$$

By the uniqueness, $u(x, t) = v(x, t) = -u(-x, t)$ and hence $u(0, t) = 0$.

Now we consider the general case of (6.8) and use the method of reflection to construct a solution in $[0, \infty) \times [0, \infty)$. To do this, we divide $[0, \infty) \times [0, \infty)$ into two parts by the straight line $t = x$ and construct u in each region. First, we consider

$$\Omega_1 = \{(x, t); x > t > 0\}.$$

By (6.3), the solution u_1 in Ω_1 is given by

$$u_1(x, t) = \frac{1}{2}\big(\varphi(x + t) + \varphi(x - t)\big) + \frac{1}{2}\int_{x-t}^{x+t} \psi(s)ds \quad \text{for any } (x, t) \in \Omega_1.$$

Set

$$\gamma(x) = u_1(x, x) = \frac{1}{2}\big(\varphi(2x) + \varphi(0)\big) + \frac{1}{2}\int_0^{2x} \psi(s)ds.$$

Next, we set

$$\Omega_2 = \{(x, t); t > x > 0\},$$

and consider

$$\begin{aligned} u_{tt} - u_{xx} &= 0 \quad \text{in } \Omega_2 \\ u(0, t) = \alpha(t), \ u(x, x) &= \gamma(x). \end{aligned}$$

We denote its solution by u_2. For any $(x, t) \in \Omega_2$, consider a rectangle formed by (x, t), $(0, t - x)$, $(\frac{t-x}{2}, \frac{t-x}{2})$ and $(\frac{t+x}{2}, \frac{t+x}{2})$. By (6.4), we have

$$\begin{aligned} u_2(x, t) &= \alpha(-x + t) - \gamma(\frac{-x + t}{2}) + \gamma(\frac{x + t}{2}) \\ &= \alpha(-x + t) + \frac{1}{2}\big(\varphi(x + t) - \varphi(-x + t)\big) + \frac{1}{2}\int_{-x+t}^{x+t} \psi(s)ds. \end{aligned}$$

Set $u = u_1$ in Ω_1 and $u = u_2$ in Ω_2. Now we check $u, u_t, u_x, u_{tt}, u_{xx}, u_{tx}$ are continuous along $t = x$. By a direct calculation, we have

$$\begin{aligned} u_1(x, t)|_{t=x} - u_2(x, t)|_{t=x} &= \gamma(0) - \alpha(0) = \varphi(0) - \alpha(0), \\ \partial_x u_1(x, t)|_{t=x} - \partial_x u_2(x, t)|_{t=x} &= -\psi(0) + \alpha'(0), \\ \partial_x^2 u_1(x, t)|_{t=x} - \partial_x^2 u_2(x, t)|_{t=x} &= -\varphi''(0) + \alpha''(0). \end{aligned}$$

Then (6.9) implies

$$u_1 = u_2, \quad \partial_x u_1 = \partial_x u_2, \quad \partial_x^2 u_1 = \partial_x^2 u_2 \quad \text{on } \{t = x\}.$$

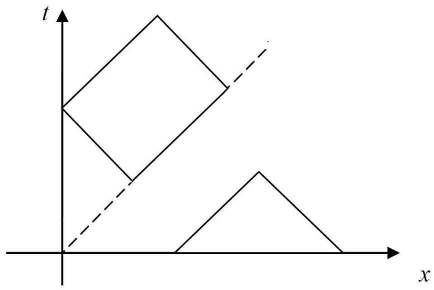

FIGURE 6.5. The method of reflection.

It is easy to get $\partial_t u_1 = \partial_t u_2$ on $\{t = x\}$ by $u_1 = u_2$ and $\partial_x u_1 = \partial_x u_2$ on $\{t = x\}$. Similarly, we get $\partial_{xt} u_1 = \partial_{xt} u_2$ and $\partial_{tt} u_1 = \partial_{tt} u_2$ on $\{t = x\}$. Therefore, u is C^2 across $t = x$. Hence, we obtain the following result.

THEOREM 6.2. *Suppose* $\varphi \in C^2[0, \infty), \psi \in C^1[0, \infty), \alpha \in C^2[0, \infty)$ *and satisfy the compatibility condition (6.9). Then there exists a* C^2*-solution of (6.8) in* $[0, \infty) \times [0, \infty)$.

As for solutions of initial-value problems, we can also derive a priori energy estimates for solutions of (6.8). For any constants $0 < T < \bar{t}$, we use the following domain for energy estimates

$$\{(x, t); 0 < x < \bar{t} - t, 0 < t < T\}.$$

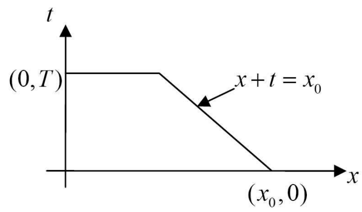

FIGURE 6.6. A domain of integration.

Now we consider the initial/boundary-value problem discussed in Section 3.3. For a positive $l > 0$, assume $\varphi \in C^2[0, l]$, $\psi \in C^1[0, l]$ and $\alpha, \beta \in C^2[0, \infty)$. Consider

$$\begin{aligned} &u_{tt} - u_{xx} = 0 \quad \text{in } [0, l] \times (0, \infty), \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = \psi \quad \text{on } [0, l], \\ &u(0, t) = \alpha(t), \ u(l, t) = \beta(t) \quad \text{for } t > 0. \end{aligned} \tag{6.10}$$

The compatibility condition is given by

$$\begin{aligned} &\varphi(0) = \alpha(0), \ \psi(0) = \alpha'(0), \ \varphi''(0) = \alpha''(0), \\ &\varphi(l) = \beta(0), \ \psi(l) = \beta'(0), \ \varphi''(l) = \beta''(0). \end{aligned} \tag{6.11}$$

For $\alpha = \beta \equiv 0$, we can use the method of extension to construct solutions. We first extend $\varphi(x) = -\varphi(-x)$ for any $x \in [0, l]$ and then extend $\varphi(x + 2l) = \varphi(x)$ for any x. So φ is odd and $2l$-periodic in $\mathbb{R}$. We extend ψ similarly. The extended functions φ and ψ are C^2 and C^1 on $\mathbb{R}$ respectively. Let u be the unique solution of

$$\begin{aligned} &u_{tt} - u_{xx} = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}_+, \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = \psi \quad \text{in } \mathbb{R}. \end{aligned}$$

We now prove that $u(x, t)$ is a solution of (6.10) when we restrict $x \in [0, l]$, i.e.,

$$u(0, t) = 0, \ u(l, t) = 0 \quad \text{for any } t > 0.$$

The proof is similar to that of Theorem 6.2. We use $v(x, t) = -u(-x, t)$ to prove $u(0, t) = 0$ and $w(x, t) = -u(2l - x, t)$ to prove $u(l, t) = 0$.

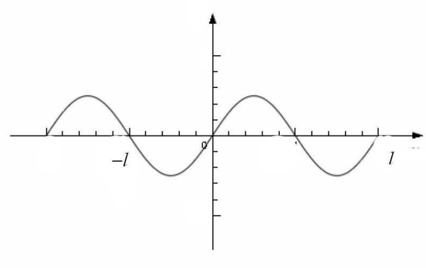

FIGURE 6.7. Extensions of initial values.

Next we use the method of reflection to construct a solution of (6.10) in the general case. We divide $[0, l] \times [0, l]$ into four regions by $t = x$ and $t = -x + l$ and construct u in each region. The process is similar and hence is omitted. Therefore,

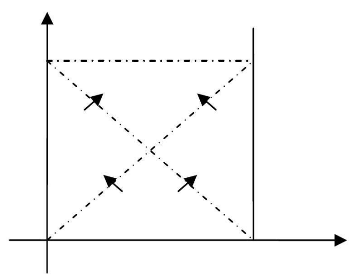

FIGURE 6.8. A method of reflection.

we obtain the following result.

THEOREM 6.3. *Suppose $\varphi \in C^2[0, \infty), \psi \in C^1[0, \infty), \alpha, \beta \in C^2[0, \infty)$ and satisfy the compatibility condition (6.11). Then there exists a C^2-solution of (6.10) in $[0, \infty) \times [0, \infty)$.*

Theorem 6.3 includes Theorem 3.19 in Chapter 2 as a special case.

Now we summarize various problems discussed in this section. We emphasize that characteristic curves play an important role in studies of the 1-dimensional wave equation.

First, presentation of problems depends on characteristic curves. Let Ω be a piecewise smooth domain in $\mathbb{R}^2$ whose boundary is not characteristic. We intend to prescribe appropriate boundary values to ensure the well-posedness for the wave equation. To do this, we take an arbitrary point on the boundary and examine characteristic curves through this point. We then count how many characteristic curves enter the domain Ω in the positive t-direction. In this section, we discussed cases where Ω is given by the upper half space $\mathbb{R} \times \mathbb{R}_+$, the first quadrant $\mathbb{R}_+ \times \mathbb{R}_+$ and $I \times \mathbb{R}_+$ for a finite integral I. We note that the number of boundary values is the same as the number of characteristic curves entering the domain in the positive t-direction. In summary, we have

$u|_{t=0} = \varphi,\ u_t|_{t=0} = \psi$ for initial-value problems;

$u|_{t=0} = \varphi,\ u_t|_{t=0} = \psi,\ u|_{x=0} = \alpha$ for half-space problems;

$u|_{t=0} = \varphi,\ u_t|_{t=0} = \psi,\ u|_{x=0} = \alpha,\ u|_{x=l} = \beta$ for initial/boundary-value problems.

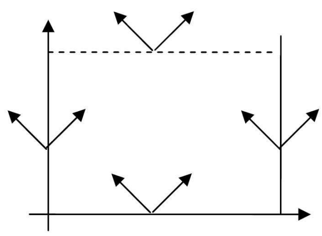

FIGURE 6.9. Characteristic directions.

Second, characteristic curves determine the domain of influence and the domain of dependence. In fact, as illustrated by (6.7), initial values propagate along characteristic curves.

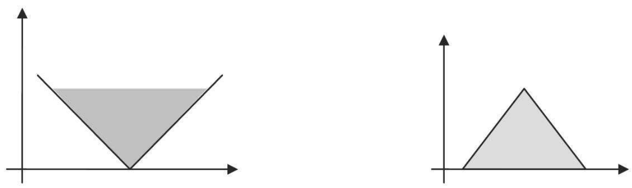

FIGURE 6.10. The domain of influence and the domain of dependence.

Last, characteristic curves also determine domains for energy estimates. We will explore this in detail in the final section of this chapter.

6.2. Higher Dimensional Wave Equations

In this section, we discuss initial-value problems for the wave equation in higher dimensions. Let φ and ψ be C^2 and C^1 functions in $\mathbb{R}^n$ respectively and f be a continuous function in $\mathbb{R}^n \times \mathbb{R}_+$. Consider

$$\begin{aligned} &u_{tt} - \Delta u = f \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = \psi \quad \text{in } \mathbb{R}^n. \end{aligned} \tag{6.12}$$

We intend to derive an expression for C^2-solution u in $\mathbb{R}^n \times [0, \infty)$. To do this, we decompose (6.12) to three problems

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = 0 \quad \text{in } \mathbb{R}^n, \end{aligned} \tag{6.13}$$

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot, 0) = 0, \ u_t(\cdot, 0) = \psi \quad \text{in } \mathbb{R}^n, \end{aligned} \tag{6.14}$$

and

$$\begin{aligned} &u_{tt} - \Delta u = f \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot, 0) = 0, \ u_t(\cdot, 0) = 0 \quad \text{in } \mathbb{R}^n. \end{aligned} \tag{6.15}$$

Obviously, a sum of solutions of (6.13)-(6.15) yields a solution of (6.12).

The next result is referred to as the Duhamel's Principle.

THEOREM 6.4. *Suppose $u_2 = M_\psi(x, t)$ is a solution of (6.14). Then*

$$\begin{aligned} u_1(x, t) &= \partial_t M_\varphi(x, t), \\ u_3(x, t) &= \int_0^t M_{f_\tau}(x, t - \tau) d\tau \quad \textit{where } f_\tau = f(\cdot, \tau) \end{aligned}$$

are solutions of (6.13) and (6.15) respectively.

The proof is based on straightforward calculations.

PROOF. First, $M_\varphi(x, t)$ satisfies

$$\begin{aligned} &\partial_{tt} M_\varphi - \Delta M_\varphi = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &M_\varphi|_{t=0} = 0, \ \partial_t M_\varphi|_{t=0} = \varphi \quad \text{in } \mathbb{R}^n. \end{aligned}$$

Then

$$\begin{aligned} &\partial_{tt} u_1 - \Delta u_1 = (\partial_{tt} - \Delta)\partial_t M_\varphi = \partial_t(\partial_{tt} M_\varphi - \Delta M_\varphi) = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u_1|_{t=0} = \partial_t M_\varphi(x, t)|_{t=0} = \varphi \quad \text{in } \mathbb{R}^n, \\ &\partial_t u_1|_{t=0} = (\partial_{tt} M_\varphi)|_{t=0} = (\Delta M_\varphi)|_{t=0} = 0 \quad \text{in } \mathbb{R}^n. \end{aligned}$$

Next, $w(x, t) = M_{f_\tau}(x, t - \tau)$ satisfies

$$\begin{aligned} &w_{tt} - \Delta w = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &w|_{t=\tau} = 0, \ \partial_t w|_{t=\tau} = f(\cdot, \tau) \quad \text{in } \mathbb{R}^n. \end{aligned}$$

Then

$$\partial_t u_3 = M_{f_\tau}(x, t - \tau)|_{\tau=t} + \int_0^t \partial_t M_{f_\tau}(x, t - \tau) d\tau = \int_0^t \partial_t M_{f_\tau}(x, t - \tau) d\tau,$$

and

$$\begin{aligned}\partial_{tt}u_3 &= \partial_t M_{f_\tau}(x,t-\tau)|_{\tau=t} + \int_0^t \partial_{tt}M_{f_\tau}(x,t-\tau)d\tau\\ &= f(x,t) + \int_0^t \Delta M_{f_\tau}(x,t-\tau)d\tau = f(x,t) + \Delta\int_0^t M_{f_\tau}(x,t-\tau)d\tau\\ &= f(x,t) + \Delta u_3.\end{aligned}$$

Hence $\partial_{tt}u_3 - \Delta u_3 = f$ in $\mathbb{R}^n \times \mathbb{R}_+$ and $u_3|_{t=0} = 0$, $\partial_t u_3|_{t=0} = 0$ in $\mathbb{R}^n$. □

In the following, we concentrate on (6.14). Suppose u is a C^2-solution of (6.14) in $\mathbb{R}^n \times [0,\infty)$, i.e.,

$$\begin{aligned}&u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+,\\ &u(\cdot,0) = 0,\ u_t(\cdot,0) = \psi \quad \text{in } \mathbb{R}^n.\end{aligned}$$

We solve this initial-value problem by the method of spherical average due to Poisson. Set for any $t > 0$ and $r > 0$

$$I(r,t) = \frac{1}{\omega_n r^{n-1}}\int_{|y-x|=r} u(y,t)dS_y,$$

where ω_n is the surface area of the unit sphere in $\mathbb{R}^n$. Obviously, $I(r,t)$ is the average of $u(\cdot,t)$ over the sphere $\partial B_r(x)$ and also depends on x. First, $I(r,t)$ is defined for $r > 0$. By changing it to the form

$$I(r,t) = \frac{1}{\omega_n}\int_{|\omega|=1} u(x + r\omega, t)dS_\omega,$$

we note that $I(r,t)$ is defined for all $r \in (-\infty,\infty)$ and is even for r, since replacing r by $-r$ does not change the value of the integral. It is clear that u can be recovered from I, since

$$I(0,r) = \lim_{r\to 0} I(r,t) = \frac{1}{\omega_n}\int_{|\omega|=1} u(x,t)dS_\omega = u(x,t).$$

Next, we change the equation of u into an equation of $I(r,t)$. We claim that $I(r,t)$ satisfies the following Euler-Poisson-Darboux equation

$$I_{tt} = I_{rr} + \frac{n-1}{r}I_r \quad \text{for } r > 0, t > 0,$$

with initial values

$$I(r,0) = 0,\ I_t(r,0) = \Psi(r) \quad \text{for } r > 0,$$

where

$$\Psi(r) = \frac{1}{\omega_n}\int_{|\omega|=1} \psi(x + r\omega)dS_\omega \quad \text{for } r > 0.$$

To verify the equation satisfied by I, we first have by differentiating under the integral sign

$$\begin{aligned}I_r(r,t) =& \frac{1}{\omega_n}\int_{|\omega|=1} \frac{\partial u}{\partial n}(x + r\omega, t)dS_\omega = \frac{1}{\omega_n r^{n-1}}\int_{|y-x|=r} \frac{\partial u}{\partial n}(y,t)dS_y\\ =& \frac{1}{\omega_n r^{n-1}}\int_{|y-x|\le r} \Delta u(y,t)dy.\end{aligned}$$

Then

$$r^{n-1}I_r = \frac{1}{\omega_n}\int_{|y-x|\le r}\Delta u(y,t)dy = \frac{1}{\omega_n}\int_{|y-x|\le r} u_{tt}(y,t)dy.$$

Hence

$$(r^{n-1}I_r)_r = \frac{1}{\omega_n}\int_{|y-x|=r} u_{tt}(y,t)dS_y = \frac{1}{\omega_n}\partial_{tt}\int_{|y-x|=r} u(y,t)dS_y = r^{n-1}I_{tt}.$$

For initial values, we simply note for any $r > 0$

$$I(r,0) = \frac{1}{\omega_n}\int_{|\omega|=1} u(x+r\omega,0)dS_\omega = 0,$$

$$I_t(r,0) = \frac{1}{\omega_n}\int_{|\omega|=1}\psi(x+r\omega)dS_\omega.$$

In general, it is a tedious process to solve initial-value problems for the Euler-Poisson-Darboux equation for general dimension. However, this process is relatively easy for $n = 3$. If $n = 3$, we have

$$I_{tt} = I_{rr} + \frac{2}{r}I_r,$$

or

$$(rI)_{tt} = (rI)_{rr} \quad \text{for } r > 0, t > 0,$$

with

$$(rI)(r,0) = 0, \ (rI)_t(r,0) = r\Psi(r) \quad \text{for } r > 0.$$

This is a half-space problem for $rI(r,t)$ studied in the previous section for the *one-dimensional* wave equation. We note $(r\Psi)|_{r=0} = 0$ and $(rI)|_{r=0} = 0$. Hence the compatibility condition in Theorem 6.2 is satisfied. In order to find a solution rI, we extend $r\Psi(r)$ to $r \in (-\infty,\infty)$ as an odd function. (In fact, $\Psi(r)$ as the average of ψ over $\partial B_r(x)$ can be extended as an even function for $r \in (-\infty,\infty)$ as before.) Then

$$(rI)(r,t) = \frac{1}{2}\int_{r-t}^{r+t} s\Psi(s)ds = \frac{1}{2}\int_{t-r}^{r+t} s\Psi(s)ds,$$

where we changed $r-t$ to $t-r$ for the lower integral limit by the oddness of the integrand. Hence,

$$I(r,t) = \frac{1}{2r}\int_{t-r}^{t+r} s\Psi(s)ds.$$

Note that the area of the unit sphere in $\mathbb{R}^3$ is 4π. Then we have

$$\lim_{r\to 0} I(r,t) = t\Psi(t) = \frac{t}{4\pi}\int_{|\omega|=1}\psi(x+t\omega)dS_\omega = \frac{1}{4\pi t}\int_{|y-x|=t}\psi(y)dS_y.$$

Therefore, we obtain formally an expression of a solution of (6.14) in $\mathbb{R}^3 \times \mathbb{R}_+$ given by

$$u(x,t) = \frac{1}{4\pi t}\int_{|y-x|=t}\psi(y)dS_y. \tag{6.16}$$

We note that $u(x,t)$ depends on the initial value in the sphere $\partial B_t(x)$.

Any C^2-solution u of the initial-value problem (6.14) in $\mathbb{R}^3 \times \mathbb{R}_+$ is given by the formula (6.16), hence is unique. Now we prove it yields the classical solution.

THEOREM 6.5. *For any $\psi \in C^2(\mathbb{R}^3)$, u in (6.16) is a C^2-solution of (6.14) in* $\mathbb{R}^3 \times \mathbb{R}_+$.

PROOF. By setting

$$\Psi(x,t) = \frac{1}{4\pi}\int_{|\omega|=1} \psi(x+t\omega)dS_\omega,$$

we have by (6.16)

$$u(x,t) = t\Psi(x,t).$$

Then

$$u_t = \Psi + t\Psi_t, \quad u_{tt} = 2\Psi_t + t\Psi_{tt}.$$

We calculate derivatives of Ψ as follows

$$\begin{aligned}
\Psi_t =& \frac{1}{4\pi}\int_{|\omega|=1} \frac{\partial\psi}{\partial n}(x+t\omega)dS_\omega = \frac{1}{4\pi t^2}\int_{|y-x|=t}\frac{\partial\psi}{\partial n}(y)dS_y \\
=& \frac{1}{4\pi t^2}\int_{|y-x|\le t}\Delta\psi(y)dy = \frac{1}{4\pi t^2}\int_0^t d\rho\int_{|y-x|=\rho}\Delta\psi(y)dS_y, \\
\Psi_{tt}(t) =& -\frac{1}{2\pi t^3}\int_{|y-x|\le t}\Delta\psi(y)dy + \frac{1}{4\pi t^2}\int_{|y-x|=t}\Delta\psi(y)dS_y \\
=& -\frac{2}{t}\Psi_t(t) + \frac{1}{4\pi t^2}\int_{|y-x|=t}\Delta_y\psi(y)dS_y.
\end{aligned}$$

By $y = x + \omega t$, we have

$$u_{tt} = \frac{t}{4\pi t^2}\int_{|y-x|=t}\Delta_y\psi(y)dS_y = \frac{t}{4\pi}\int_{|\omega|=1}\Delta_x\psi(x+t\omega)dS_\omega = \Delta u,$$

where we used $\Delta_x\psi(x+t\omega) = \Delta_y\psi(y)$. It is easy to see $u(\cdot,0) = 0$ and $u_t(\cdot,0) = \psi$ in $\mathbb{R}^3$. □

Now we consider the initial-value problem (6.12) of the wave equation in $\mathbb{R}^3$ as follows

$$\begin{aligned}
&u_{tt} - \Delta u = f \quad \text{in } \mathbb{R}^3 \times \mathbb{R}_+, \\
&u(\cdot,0) = \varphi,\ u_t(\cdot,0) = \psi \quad \text{in } \mathbb{R}^3.
\end{aligned}$$

By Theorem 6.4, we have

$$\begin{aligned}
u(x,t) =& \partial_t\left(\frac{1}{4\pi t}\int_{|y-x|=t}\varphi(y)dS_y\right) + \frac{1}{4\pi t}\int_{|y-x|=t}\psi(y)dS_y \\
&+ \frac{1}{4\pi}\int_{|y-x|\le t}\frac{f(y,t-|y-x|)}{|y-x|}dy.
\end{aligned} \tag{6.17}$$

In fact, we only need to verify the final term in (6.17). We note

$$\begin{aligned}
&\frac{1}{4\pi}\int_0^t (t-\tau)\int_{|\omega|=1} f(x+\omega(t-\tau),\tau)dS_\omega d\tau \quad (\text{with } t-\tau=\tau')\\
=&\frac{1}{4\pi}\int_0^t \tau\int_{|\omega|=1} f(x+\tau\omega,t-\tau)dS_\omega d\tau\\
=&\frac{1}{4\pi}\int_0^t \int_{|y-x|=\tau} \frac{f(y,t-\tau)}{\tau}dS_y d\tau\\
=&\frac{1}{4\pi}\int_{|y-x|\le t} \frac{f(y,t-|y-x|)}{|y-x|}dy.
\end{aligned}$$

Hence we obtain the following result.

THEOREM 6.6. *For any $\varphi \in C^3(\mathbb{R}^3)$, $\psi \in C^2(\mathbb{R}^3)$ and $f \in C^2(\mathbb{R}^3 \times \mathbb{R}_+)$, the function u given by (6.17) is a C^2-solution of (6.12) in $\mathbb{R}^3 \times \mathbb{R}_+$.*

We can also express the final term in (6.17) as

$$\frac{1}{4\pi}\int_0^t \frac{1}{t-\tau}\int_{|y-x|=t-\tau} f(y,\tau)dS_y d\tau.$$

Hence the value of the solution u at (x,t) depends only on the values of f in points (y,τ) with

$$|y-x| = t-\tau, \quad 0<\tau<t.$$

This is exactly the backward characteristic cone with vertex (x,t).

We now use the method of descent to solve initial-value problems of the wave equation in $\mathbb{R}^2 \times \mathbb{R}_+$. As for three dimensional case, we first consider (6.14). Let ψ be C^2 in $\mathbb{R}^2$ and $u(x,t)$ be C^2 in $\mathbb{R}^2 \times [0,\infty)$ and satisfy (6.14), i.e.,

$$\begin{aligned}
&u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^2 \times \mathbb{R}_+,\\
&u(\cdot,0) = 0,\ u_t(\cdot,0) = \psi \quad \text{in } \mathbb{R}^2.
\end{aligned}$$

Any solutions in $\mathbb{R}^2$ can be viewed as solutions of the same problem in $\mathbb{R}^3$ which are independent of the third space variable. Then for any $(x,t) \in \mathbb{R}^2 \times \mathbb{R}_+$ (with $x_3 = 0$), we have by (6.16)

$$u(x,t) = \frac{1}{4\pi t}\int_{|\bar{y}-x|=t} \psi(y)dS_{\bar{y}},$$

where $\bar{y} = (y_1,y_2,y_3) = (y,y_3)$. The integral here is a surface integral in $\mathbb{R}^3$. Now we evaluate it as a 2-dimensional integral by eliminating y_3. The sphere $|\bar{y}-x| = t$ in $\mathbb{R}^3$ has two pieces given by

$$y_3 = \pm\sqrt{t^2 - |y-x|^2},$$

and its surface area element is

$$dS_{\bar{y}} = \big(1 + (\partial_{y_1}y_3)^2 + (\partial_{y_2}y_3)^2\big)^{\frac{1}{2}}dy_1dy_2 = \frac{t}{\sqrt{t^2-|y-x|^2}}dy_1dy_2.$$

Therefore, we obtain

$$u(x,t) = \frac{1}{2\pi}\int_{|y-x|\le t} \frac{\psi(y)}{\sqrt{t^2-|y-x|^2}}dy. \tag{6.18}$$

We note that $u(x,t)$ depends on initial values in the solid disc $B_t(x)$.

THEOREM 6.7. *For any* $\psi \in C^2(\mathbb{R}^2)$, *$u$ in (6.18) is a C^2-solution of (6.14) in* $\mathbb{R}^2 \times \mathbb{R}_+$.

Next, we discuss briefly how to obtain explicit expressions for solutions of the wave equation in arbitrary dimension. For odd dimension, we seek an appropriate combination of $I(r,t)$ and its derivatives to satisfy the 1-dimensional wave equation and then proceed as for $n = 3$. For even dimension, we again use the method of descent.

Let $n \geq 3$ be an odd integer. Then the spherical average $I(r,t)$ satisfies

$$I_{tt} = I_{rr} + \frac{n-1}{r} I_r \quad \text{for } r > 0, t > 0. \tag{6.19}$$

First, we write (6.19) as

$$I_{tt} = \frac{1}{r}\big(rI_{rr} + (n-1)I_r\big).$$

By

$$(rI)_{rr} = rI_{rr} + 2I_r,$$

we obtain

$$I_{tt} = \frac{1}{r}\big((rI)_{rr} + (n-3)I_r\big),$$

or

$$(rI)_{tt} = (rI)_{rr} + (n-3)I_r. \tag{6.20}$$

If $n = 3$, then rI satisfies the 1-dimensional wave equation. This is how we solved the wave equation in 3-dimension. By differentiating (6.19) with respect to r, we have

$$I_{rtt} = I_{rrr} + \frac{n-1}{r} I_{rr} - \frac{n-1}{r^2} I_r = \frac{1}{r^2}\big(r^2 I_{rrr} + (n-1)rI_{rr} - (n-1)I_r\big).$$

By

$$(r^2 I_r)_{rr} = r^2 I_{rrr} + 4rI_{rr} + 2I_r,$$

we obtain

$$I_{rtt} = \frac{1}{r^2}\big((r^2 I)_{rr} + (n-5)rI_{rr} - (n+1)I_r\big),$$

or

$$(r^2 I_r)_{tt} = (r^2 I_r)_{rr} + (n-5)rI_{rr} - (n+1)I_r. \tag{6.21}$$

The second term in the right-hand side of (6.21) has a coefficient $n - 5$, which is 2 less than the coefficient of the second term in the right-hand side of (6.20). Also the third term involving I_r in the right-hand side of (6.21) has a similar expression as the second term in the right-hand side of (6.20). Therefore an appropriate combination of (6.20) and (6.21) eliminates terms involving I_r. In particular, for $n = 5$, we have

$$(r^2 I_r + 3rI)_{tt} = (r^2 I_r + 3rI)_{rr}.$$

In other words, $r^2 I_r + 3rI$ satisfies the 1-dimensional wave equation. We can continue this process to obtain appropriate forms for all odd dimensions. Next, we note

$$r^2 I_r + 3rI = \frac{1}{r}(r^3 I)_r.$$

It turns out that the correct combination of I and its derivatives for arbitrary odd dimension n is given by

$$J(r,t) = \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{\frac{n-3}{2}} \left(r^{n-2}I(r,t)\right).$$

We leave derivation as an exercise.

We need to point out that the method of spherical average is by no means the only way to derive explicit expressions of solutions of the wave equation. Other methods are available. Refer to an exercise for an alternative approach to solving three-dimensional wave equation.

Now we compare several formulas we obtained in the previous section and in this section. Let u be a C^2-solution of the initial-value problem

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot,0) = \varphi,\ u_t(\cdot,0) = \psi \quad \text{in } \mathbb{R}^n. \end{aligned} \tag{6.22}$$

We write u_n for dimension n. Then

$$\begin{aligned} u_1(x,t) &= \frac{1}{2}\big(\varphi(x+t) + \varphi(x-t)\big) + \frac{1}{2}\int_{x-t}^{x+t} \psi(y)dy, \\ u_2(x,t) &= \partial_t \left(\frac{1}{2\pi}\int_{|y-x|\le t} \frac{\varphi(y)}{\sqrt{t^2 - |y-x|^2}}dy\right) + \frac{1}{2\pi}\int_{|y-x|\le t} \frac{\psi(y)}{\sqrt{t^2 - |y-x|^2}}dy, \\ u_3(x,t) &= \partial_t \left(\frac{1}{4\pi t}\int_{|y-x|=t} \varphi(y)dS_y\right) + \frac{1}{4\pi t}\int_{|y-x|=t} \psi(y)dS_y. \end{aligned}$$

These formulas display many important properties of solutions u.

According to these expressions, the value u at (x,t) depends on the values of φ and ψ on the interval $[x-t, x+t]$ for $n=1$, on the solid disc $B_t(x)$ of center x and radius t for $n=2$, and on the sphere $\partial B_t(x)$ of center x and radius t for $n=3$. These regions are the *domain of dependence* of solutions on initial values. Conversely, initial values φ and ψ at a point x_0 on the initial hypersurface $t=0$

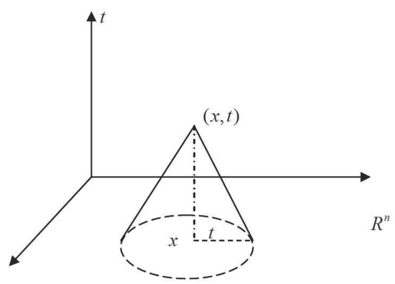

FIGURE 6.11. The domain of dependence.

influence u in points (x,t) in the solid cone $|x-x_0| \le t$ for $n=2$ and only the cone $|x-x_0| = t$ for $n=3$ at a later time t.

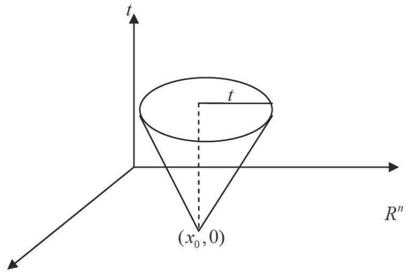

FIGURE 6.12. The domain of influence.

The central issue here is that the solution at a given point is determined by initial values in a proper subset of the initial hypersurface. An important consequence is that the process of solving initial-value problems for the wave equation can be localized in space. Specifically, changing initial values outside domain of dependence of a point does not change values of solutions there. This is a unique property of the wave equation which distinguishes it from the heat equation.

Before exploring the difference between $n = 2$ and $n = 3$, we first note that it takes time (literally) for initial values to make influences. Suppose the initial values φ, ψ have their support contained in a ball $B_r(x_0)$. Then at a later time t, the support of $u(\cdot, t)$ is contained in the union of all balls $B_t(\bar{x})$ for $\bar{x} \in B_r(x_0)$. It is easy to see that such a union is in fact the ball with center x_0 and radius $r + t$. The support of u spreads at a finite speed. To put it in another perspective, we fix an $x \notin B_r(x_0)$. Then $u(x,t) = 0$ for $t < |x - x_0| - r$. This is the so called *finite speed propagation*.

For $n = 2$, if the support of φ and ψ takes the entire disc $B_r(x_0)$, then the support of $u(\cdot, t)$ will take the entire disc $B_{r+t}(x_0)$ in general. The influence from initial values never disappears in a finite time at any particular point, like the surface waves arising from a stone dropped into water.

For $n = 3$, the behavior of solutions is different. Again, we assume the support of φ and ψ are contained in a ball $B_r(x_0)$. Then at a later time t, the support of $u(\cdot, t)$ is in fact contained in the union of all spheres $\partial B_t(\bar{x})$ for $\bar{x} \in B_r(x_0)$. Such a union is a ball $B_{t+r}(x_0)$ for $t \le r$, as in the two-dimensional case, and an annular region with center x_0 and outer and inner radius $t + r$ and $t - r$ respectively for $t > r$. Such an annular region has a thickness $2r$ and spreads at a finite speed. In other words, $u(x,t)$ is not zero only if

$$t - r < |x - x_0| < t + r,$$

or

$$|x - x_0| - r < t < |x - x_0| + r.$$

Therefore, for a fixed $x \in \mathbb{R}^3$, $u(x,t) = 0$ for $t < |x - x_0| - r$ (corresponding to the finite speed propagation) and for $t > |x - x_0| + r$. So, influence from initial values lasts only for an interval of length $2r$ in time. This phenomenon is called *Huygens' principle* for the wave equation.

In fact, Huygens' principle holds for the wave equation in every odd space dimension n except $n = 1$ and does not hold in even space dimension.

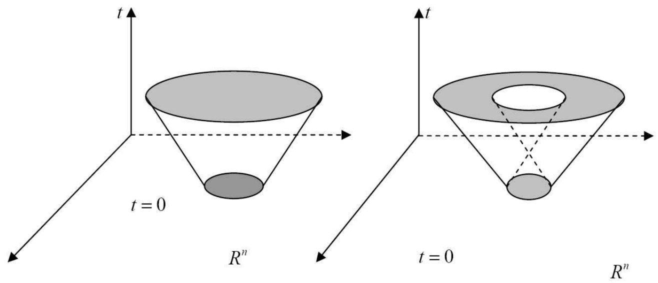

FIGURE 6.13. A comparison between $n = 2$ and $n = 3$.

Next, we discuss whether solutions decay as $t \to \infty$. In this aspect, there is a sharp difference between 1-dimension and higher dimensions. By the d'Alembert's formula, it is obvious that solutions to the 1-dimensional homogeneous wave equation do not decay as $t \to \infty$. However, solutions in higher dimensions have a different behavior.

THEOREM 6.8. *Let ψ be a smooth function with compact support in $\mathbb{R}^n$ and u be a solution of*

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \textit{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot, 0) = 0, \ u_t(\cdot, 0) = \psi \quad \textit{on } \mathbb{R}^n. \end{aligned}$$

Then for any $t > 1$

$$|u(\cdot, t)|_{L^\infty(\mathbb{R}^n)} \le \frac{C}{\sqrt{t}} \big(\|\psi\|_{L^1(\mathbb{R}^n)} + \|\nabla\psi\|_{L^1(\mathbb{R}^n)} \big) \quad \textit{for } n = 2,$$

and

$$|u(\cdot, t)|_{L^\infty(\mathbb{R}^n)} \le \frac{C}{t} \|\nabla\psi\|_{L^1(\mathbb{R}^n)} \quad \textit{for } n = 3,$$

where C is a positive constant independent of u and ψ.

We note that decay rates vary according to dimensions. In fact, solutions to the n-dimensional wave equation have a decay rate of $t^{-\frac{n-1}{2}}$. These decay estimates play an important role in study of global solutions of nonlinear wave equations.

PROOF. We first consider $n = 3$. By (6.16), the solution u is given by

$$u(x, t) = \frac{t}{4\pi} \int_{|\omega|=1} \psi(x + t\omega) dS_\omega.$$

Since ψ has compact support, we have

$$\psi(x + t\omega) = -\int_t^\infty \frac{\partial}{\partial s} \psi(x + s\omega) ds.$$

Then

$$u(x, t) = -\frac{t}{4\pi} \int_t^\infty \int_{|\omega|=1} \frac{\partial}{\partial s} \psi(x + s\omega) dS_\omega ds.$$

For $s \ge t$, we write $t \le s^2/t$ and hence

$$|u(x, t)| \le \frac{1}{4\pi t} \int_t^\infty s^2 \int_{|\omega|=1} |\nabla\psi(x + s\omega)| dS_\omega ds \le \frac{1}{4\pi t} \|\nabla\psi\|_{L^1(\mathbb{R}^3)}.$$

This implies the desired result for $n = 3$. In fact, it holds for any $t > 0$.

Now we consider $n = 2$. By (6.18) and a change of variables, we have

$$u(x,t) = \frac{1}{2\pi}\int_0^t \frac{r}{\sqrt{t^2-r^2}}\int_{|\omega|=1}\psi(x+r\omega)dS_\omega dr.$$

For some $\varepsilon > 0$ to be determined, we write u as

$$u(x,t) = \frac{1}{2\pi}\Big(\int_0^{t-\varepsilon} + \int_{t-\varepsilon}^t\Big) = \frac{1}{2\pi}\big(I_1 + I_2\big).$$

For I_1, we have

$$\begin{aligned}|I_1| =& \Big|\int_0^{t-\varepsilon}\frac{r}{\sqrt{t^2-r^2}}\int_{|\omega|=1}\psi(x+r\omega)dS_\omega dr\Big|\\ \leq& \frac{1}{\sqrt{t^2-(t-\varepsilon)^2}}\int_0^{t-\varepsilon} r\int_{|\omega|=1}|\psi(x+r\omega)|dS_\omega dr\\ \leq& \frac{1}{\sqrt{2\varepsilon t-\varepsilon^2}}\|\psi\|_{L^1(\mathbb{R}^2)}.\end{aligned}$$

Next, for $r \in (t-\varepsilon, t)$, as in the proof for $n = 3$, we have

$$\int_{|\omega|=1}\psi(x+r\omega)dS_\omega = -\int_r^\infty\int_{|\omega|=1}\frac{\partial}{\partial s}\psi(x+s\omega)dS_\omega ds,$$

and hence

$$\Big|r\int_{|\omega|=1}\psi(x+r\omega)dS_\omega\Big| \leq \int_r^\infty s\int_{|\omega|=1}|\nabla\psi(x+s\omega)|dS_\omega ds \leq \|\nabla\psi\|_{L^1(\mathbb{R}^2)}.$$

Then

$$\begin{aligned}|I_2| =& \Big|\int_{t-\varepsilon}^t\frac{r}{\sqrt{t^2-r^2}}\int_{|\omega|=1}\psi(x+r\omega)dS_\omega dr\Big|\\ \leq& \int_{t-\varepsilon}^t\frac{1}{\sqrt{t^2-r^2}}dr\cdot\sup_{r\in[t-\varepsilon,t]}\left(r\int_{|\omega|=1}\psi(x+r\omega)dS_\omega\right)\\ \leq& \int_{t-\varepsilon}^t\frac{1}{\sqrt{t^2-r^2}}dr\cdot\|\nabla\psi\|_{L^1(\mathbb{R}^2)}.\end{aligned}$$

A simple calculation yields

$$\int_{t-\varepsilon}^t\frac{1}{\sqrt{t^2-r^2}}dr = \int_{t-\varepsilon}^t\frac{1}{\sqrt{(t+r)(t-r)}}dr \leq \frac{1}{\sqrt{t}}\int_{t-\varepsilon}^t\frac{1}{\sqrt{t-r}}dr = \frac{2\sqrt{\varepsilon}}{\sqrt{t}}.$$

Therefore, we obtain

$$|u(x,t)| \leq \frac{1}{2\pi}\left(\frac{1}{\sqrt{2\varepsilon t-\varepsilon^2}}\|\psi\|_{L^1(\mathbb{R}^2)} + \frac{2\sqrt{\varepsilon}}{\sqrt{t}}\|\nabla\psi\|_{L^1(\mathbb{R}^2)}\right).$$

For any $t > 1$, we take $\varepsilon = 1/2$ and obtain the desired result. $\square$

Decay estimates in Theorem 6.8 are optimal for large t. We have the following result for arbitrary t.

THEOREM 6.9. *Let ψ be a smooth function with compact support in $\mathbb{R}^n$ and u be a solution of*

$$u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+,$$
$$u(\cdot, 0) = 0, \ u_t(\cdot, 0) = \psi \quad \text{on } \mathbb{R}^n.$$

Then for any $t > 0$

$$|u(\cdot, t)|_{L^\infty(\mathbb{R}^n)} \leq C\|\nabla\psi\|_{L^1(\mathbb{R}^n)} \quad \text{for } n = 2,$$

and

$$|u(\cdot, t)|_{L^\infty(\mathbb{R}^n)} \leq C\|\nabla^2\psi\|_{L^1(\mathbb{R}^n)} \quad \text{for } n = 3,$$

where C is a positive constant independent of u and ψ.

The proof is left as an exercise.

Now we compare regularity of solutions for $n = 1$ and $n = 3$. For $n = 1$, the regularity of u is clearly the same as $u(\cdot, 0)$ and one order better than $u_t(\cdot, 0)$. In other words, $u \in C^m$ and $u_t \in C^{m-1}$ initially at $t = 0$ guarantee $u \in C^m$ at a later time. However, such a result does not hold for $n = 3$. By writing spherical averages as integrals over the unit sphere $|\omega| = 1$ and differentiating under integral sign, we have

$$u_3(x, t) = \frac{1}{4\pi t^2}\int_{|y-x|=t} \big(\varphi(y) + \sum_{i=1}^{3} \varphi_{y_i}(y)(y_i - x_i) + t\psi(y)\big) dS_y.$$

This formula indicates that u can be less regular than initial values. There is a possible loss of one order of differentiability. Namely, $u \in C^m$ and $u_t \in C^{m-1}$ initially at $t = 0$ only guarantee $u \in C^{m-1}$ at a later time.

Next, we construct a solution of the 3-dimensional homogeneous wave equation with a loss of differentiation described above. Let $u = u(r, t)$ be a spherically symmetric solution of the wave equation in $\mathbb{R}^3 \times \mathbb{R}_+$. Then

$$u_{tt} - u_{rr} - \frac{2}{r}u_r = 0,$$

or

$$(ru)_{tt} - (ru)_{rr} = 0.$$

In other words, $ru(r, t)$ is a solution of the 1-dimensional wave equation. Hence

$$u(r, t) = \frac{1}{r}\left(f(t + r) + g(t - r)\right) \quad \text{for any } r \neq 0,$$

for some functions f and g on $\mathbb{R}$. In the following, we consider $f = g$. For any positive integer m, we define

$$f(r) = \begin{cases} 0 & \text{for } r \leq 1, \\ (r-1)^{2m+1} & \text{for } 1 < r < 2. \end{cases}$$

We extend f smoothly for $r > 1$. We note that f is C^{2m} in $\mathbb{R}$ and not C^{2m+1} at $r = 1$. Then for $t = 0$,

$$u(r, 0) = \frac{1}{r}\left(f(r) + f(-r)\right), \quad u_t(r, 0) = \frac{1}{r}\left(f'(r) + f'(-r)\right).$$

Hence $u(\cdot, 0) \in C^{2m}(\mathbb{R})$, $u_t(\cdot, 0) \in C^{2m-1}(\mathbb{R})$, and $u(r, 0)$ is not C^{2m+1} and $u_t(r, 0)$ is not C^{2m} at $r = 1$. We now claim that $u(r, t)$ is not C^{2m} at $(r, t) = (0, 1)$. The intuitive idea is that the singularity of initial values at $r = 1$ propagates along a characteristic hypersurface and focuses at $t = 1$. We note that $(r, t) = (0, 1)$ is the vertex of the characteristic hypersurface $\{(x, t); t < |x|\}$ which intersects $t = 0$ on $r = 1$. First, $u(r, 0) = u_t(r, 0) = 0$ for $r \leq 1$. Then $u(x, t) = 0$ for $0 < t \leq |x|$. In particular,

$$\lim_{t \to 1-} \partial_t^{2m} u(0, t) = 0.$$

If u were C^{2m} at $(r, t) = (0, 1)$, then $\Delta^m u = \partial_t^{2m} u = 0$ at $(0, 1)$. Next for $r > 0$ sufficiently small, we have $1 + r > 1$ and $1 - r < 1$, and hence

$$u(r, 1) = \frac{1}{r} f(1 + r) = r^{2m} = |x|^{2m}.$$

Therefore for $r > 0$ sufficiently small,

$$\Delta_x^m u(r, 1) = \left(\partial_{rr} + \frac{2}{r} \partial_r \right)^m r^{2m} = (2m + 1)!.$$

We conclude that u is not C^{2m} at $(r, t) = (0, 1)$.

This example demonstrates that solutions of the higher dimensional wave equation do not have good point-wise behavior. However, the differentiability is preserved in the L^2-sense. Before we move on to energy estimates in the next section, we show by a simple case what is involved.

Suppose u is a C^2-solution of (6.22) for general n. We assume that φ and ψ have compact support. By finite speed propagation, $u(\cdot, t)$ also has compact support for any $t > 0$. We multiply the equation in (6.22) by u_t and integrate in $B_R \times (0, t)$. Here we choose R sufficiently large such that the support of $u(\cdot, s)$ is contained in B_R for any $s \in (0, t)$. Note

$$u_t u_{tt} - u_t \Delta u = \frac{1}{2}(u_t^2 + |\nabla_x u|^2)_t - \sum_{i=1}^{n} (u_t u_{x_i})_{x_i}.$$

Then a simple integration in $B_R \times (0, t)$ yields

$$\frac{1}{2} \int_{\mathbb{R}^n \times \{t\}} (u_t^2 + |\nabla_x u|^2) = \frac{1}{2} \int_{\mathbb{R}^n \times \{0\}} (u_t^2 + |\nabla_x u|^2).$$

This is the conservation of energy: the L^2-norm of derivatives at each time slice is constant, independent of time.

6.3. Energy Estimates

In this section, we derive energy estimates of solutions of initial-value problems for a slightly more general hyperbolic equations instead of just wave equations. Fix a positive number T. Let a, c and f be continuous functions in $\mathbb{R}^n \times [0, T]$ and φ and ψ be continuous functions in $\mathbb{R}^n$. We consider the following initial-value problem

$$\begin{aligned} &u_{tt} - a\Delta u + cu = f \quad \text{in } \mathbb{R}^n \times [0, T], \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = \psi \quad \text{in } \mathbb{R}^n. \end{aligned} \tag{6.23}$$

We assume a is a positive function satisfying

$$\lambda \leq a(x, t) \leq \Lambda \quad \text{for any } (x, t) \in \mathbb{R}^n \times [0, T] \text{ and } \xi \in \mathbb{R}^n, \tag{6.24}$$

for some positive constants λ and Λ. If the principal part of the equation is given by the wave operator $u_{tt} - \Delta u$, then $a = 1$ and we can choose $\lambda = \Lambda = 1$ in (6.24).

With $\kappa = 1/\sqrt{\Lambda}$, we introduce for fixed $\bar{t} > T$

$$D_{\kappa,T,\bar{t}} = \{(x,t); \kappa|x| < \bar{t} - t,\ 0 < t < T\}.$$

We denote by $\partial_- D_{\kappa,T,\bar{t}}$, $\partial_s D_{\kappa,T,\bar{t}}$ and $\partial_+ D_{\kappa,T,\bar{t}}$ the bottom, the side and the top of the boundary. See discussions in Section 2.3 for details.

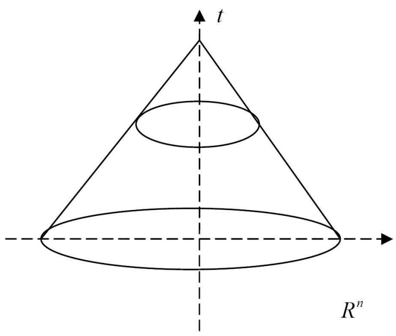

FIGURE 6.14. An integral domain.

THEOREM 6.10. *Assume a is C^1, c and f are continuous in $\mathbb{R}^n \times [0,T]$, φ is C^1 in $\mathbb{R}^n$ and ψ is continuous in $\mathbb{R}^n$. Let (6.24) be assumed and u be a C^2-solution of (6.23). Then for any $\eta \geq \eta_0$*

$$\begin{aligned}\int_{\partial_+ D_{\kappa,T,\bar{t}}} & e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2) + (\eta - \eta_0)\int_{D_{\kappa,T,\bar{t}}} e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2) \\ &\leq \int_{\partial_- D_{\kappa,T,\bar{t}}} (\varphi^2 + \psi^2 + a|\nabla\varphi|^2) + \int_{D_{\kappa,T,\bar{t}}} e^{-\eta t} f^2,\end{aligned}$$

where η_0 is a positive constant depending only on λ, the C^1-norm of a and the L^∞-norm of c in $D_{\kappa,T,\bar{t}}$.

PROOF. In the following, we simply set $D = D_{\kappa,T,\bar{t}}$. We multiply the equation in (6.23) by $2e^{-\eta t}u_t$ and integrate in D, for a scalar η to be determined. First, we note

$$2e^{-\eta t}u_t u_{tt} = e^{-\eta t}(u_t^2)_t + \eta e^{-\eta t}u_t^2,$$

and

$$\begin{aligned}-2e^{-\eta t}au_t\Delta u &= \sum_{i=1}^n \big(-2(e^{-\eta t}au_t u_{x_i})_{x_i} + 2e^{-\eta t}au_{x_i}u_{tx_i} + 2e^{-\eta t}a_{x_i}u_t u_{x_i}\big) \\ &= \sum_{i=1}^n \big(-2(e^{-\eta t}au_t u_{x_i})_{x_i} + (e^{-\eta t}au_{x_i}^2)_t + 2e^{-\eta t}a_{x_i}u_t u_{x_i} \\ &\qquad + \eta e^{-\eta t}au_{x_i}^2 - e^{-\eta t}a_t u_{x_i}^2\big),\end{aligned}$$

where we used $2u_{x_i}u_{tx_i} = (u_{x_i}^2)_t$. Therefore, we obtain

$$(e^{-\eta t}u_t^2 + e^{-\eta t}a|\nabla u|^2)_t - 2\sum_{i=1}^n (e^{-\eta t}u_t u_{x_i})_{x_i} + \eta e^{-\eta t}(u_t^2 + a|\nabla u|^2)$$
$$+ \sum_{i=1}^n 2e^{-\eta t}a_{x_i}u_t u_{x_i} - e^{-\eta t}a_t|\nabla u|^2 + 2e^{-\eta t}cuu_t = 2e^{-\eta t}u_t f.$$

To control the final term in the left hand side involving uu_t, we note

$$(e^{-\eta t}u^2)_t + \eta e^{-\eta t}u^2 - 2e^{-\eta t}uu_t = 0.$$

Then a simple addition yields

$$\big(e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2)\big)_t - \sum_{i=1}^n 2(e^{-\eta t}au_t u_{x_i})_{x_i} + \eta e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2)$$
$$= -2e^{-\eta t}\sum_{i=1}^n a_{x_i}u_t u_{x_i} + e^{-\eta t}a_t|\nabla u|^2 - 2e^{-\eta t}(c-1)uu_t + 2e^{-\eta t}u_t f.$$

The first three terms in the right hand side is quadratic in u_t, u_{x_i} and u. Then by (6.24) and the Cauchy inequality applied to each term in the right hand side, we have

$$\big(e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2)\big)_t - 2\sum_{i=1}^n (e^{-\eta t}au_t u_{x_i})_{x_i}$$
$$+ (\eta - \eta_0)e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2) \le e^{-\eta t}f^2,$$

where η_0 is a positive constant depending only on λ, the C^1-norm of a and the L^∞-norm of c. By integrating over D, we obtain

$$\int_{\partial_+ D} e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2) + (\eta - \eta_0)\int_D e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2)$$
$$+ \int_{\partial_s D} e^{-\eta t}\big((u^2 + u_t^2 + a|\nabla u|^2)\gamma_t - 2\sum_{i=1}^n au_t u_{x_i}\gamma_i\big)$$
$$\le \int_{\partial_- D}(u^2 + u_t^2 + a|\nabla u|^2) + \int_D e^{-\eta t}f^2,$$

where $\gamma = (\gamma_1, \cdots, \gamma_n, \gamma_t)$ is the unit exterior normal vector along $\partial_s D$. We only need to prove that the integrand for $\partial_s D$ is nonnegative. We claim

$$(u_t^2 + a|\nabla u|^2)\gamma_t - 2\sum_{i=1}^n au_t u_{x_i}\gamma_i \ge 0 \quad \text{on } \partial_s D.$$

To prove this, we first note by the Cauchy inequality

$$|\sum_{i=1}^n u_{x_i}\gamma_i| \le \big(\sum_{i=1}^n u_{x_i}^2\big)^{\frac{1}{2}} \cdot \big(\sum_{i=1}^n \gamma_i^2\big)^{\frac{1}{2}} = |\nabla u|\sqrt{1-\gamma_t^2}.$$

By the explicit expression of γ (see Section 2.3), we have

$$\gamma_t = \frac{1}{\sqrt{1+\kappa^2}},$$

and hence with (6.24) and $\kappa = 1/\sqrt{\Lambda}$

$$\begin{aligned}(u_t^2 + a|\nabla u|^2)\gamma_t - 2\sum_{i=1}^n a u_t u_{x_i}\gamma_i \geq& \frac{1}{\sqrt{1+\kappa^2}}\big(u_t^2 + a|\nabla u|^2 - 2\kappa a|u_t|\cdot|\nabla u|\big)\\ \geq& \frac{1}{\sqrt{1+\kappa^2}}\big(u_t^2 + a|\nabla u|^2 - 2\sqrt{a}|u_t|\cdot|\nabla u|\big),\end{aligned}$$

which is nonnegative. Therefore, the boundary integral on $\partial_s D$ is nonnegative. $\square$

A consequence of Theorem 6.10 is the uniqueness of solutions of (6.23). We can also discuss domain of dependence and domain of influence as in the previous section.

Next, we consider (6.23) in a general domain

$$D = \{(x,t); h_-(x) < t < h_+(x), x \in \Omega\},$$

where Ω is a bounded domain in $\mathbb{R}^n$ and h_- and h_+ are two piecewise C^1-functions in Ω with $h_- = h_+$ on $\partial\Omega$. We now perform a similar integration in D as in the

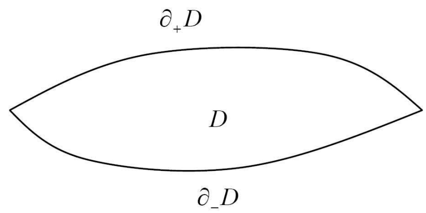

FIGURE 6.15. A general domain.

above proof to get

$$\begin{aligned}&\int_{\partial_+ D} e^{-\eta t}\big((u^2 + u_t^2 + a|\nabla u|^2)\gamma_{+t} - 2\sum_{i=1}^n a u_t u_{x_i}\gamma_{+i}\big)\\ &\qquad + (\eta - \eta_0)\int_D e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2)\\ &\leq \int_{\partial_- D} e^{-\eta t}\big((u^2 + u_t^2 + a|\nabla u|^2)\gamma_{-t} - 2\sum_{i=1}^n a u_t u_{x_i}\gamma_{-i}\big) + \int_D e^{-\eta t} f^2,\end{aligned}$$

where $\gamma_\pm = (\gamma_{\pm 1}, \cdots, \gamma_{\pm n}, \gamma_{\pm t})$ are unit normal vector pointing to positive t-direction along $\partial_\pm D$. We are interested in whether the integrand for $\partial_+ D$ is nonnegative. As in the proof above, we have by the Cauchy inequality

$$|\sum_{i=1}^n u_{x_i}\gamma_{+i}| \leq \big(\sum_{i=1}^n u_{x_i}^2\big)^{\frac{1}{2}} \cdot \big(\sum_{i=1}^n \gamma_{+i}^2\big)^{\frac{1}{2}} = |\nabla u|\sqrt{1-\gamma_{+t}^2}.$$

Then it is easy to see

$$\begin{aligned}&(u_t^2 + a|\nabla u|^2)\gamma_{+t} - 2\sum_{i=1}^n a u_t u_{x_i}\gamma_{+i}\\ \geq&(u_t^2 + a|\nabla u|^2)\gamma_{+t} - 2\sqrt{a(1-\gamma_{+t}^2)}\cdot\sqrt{a}|\nabla u| \geq 0 \quad \text{on } \partial_+ D\end{aligned}$$

if

$$\gamma_{+t} \geq \sqrt{a(1 - \gamma_{+t}^2)},$$

or

$$\gamma_{+t} \geq \sqrt{\frac{a}{1+a}} \quad \text{on } \partial_+ D. \tag{6.25}$$

With (6.25), we obtain

$$\int_{\partial_+ D} e^{-\eta t} u^2 \gamma_{+t} + (\eta - \eta_0) \int_D e^{-\eta t} (u^2 + u_t^2 + a|\nabla u|^2)$$
$$\leq \int_{\partial_- D} e^{-\eta t} \big((u^2 + u_t^2 + a|\nabla u|^2)\gamma_{-t} - 2 \sum_{i=1}^n a u_t u_{x_i} \gamma_{-i} \big) + \int_D e^{-\eta t} f^2.$$

In particular, if $u = u_t = 0$ on $\partial_- D$ and $f = 0$ in D, then $u = 0$ in D.

Now we introduce the notion of space-like and time-like surfaces.

Definition 6.11. Let Σ be a C^1-hypersurface in $\mathbb{R}^n \times \mathbb{R}_+$ and $\gamma = (\gamma_x, \gamma_t)$ be a unit normal vector with $\gamma_t \geq 0$. Then Σ is *space-like* at (x, t) if

$$\gamma_t(x,t) > \sqrt{\frac{a(x,t)}{1 + a(x,t)}};$$

Σ is *time-like* at (x, t) if

$$\gamma_t(x,t) < \sqrt{\frac{a(x,t)}{1 + a(x,t)}}.$$

If the hypersurface Σ is given by $t = t(x)$, it is easy to check that Σ is space-like at $(x, t(x))$ if

$$|\nabla t(x)| < \frac{1}{\sqrt{a(x, t(x))}}.$$

For the wave equation

$$u_{tt} - \Delta u = f, \tag{6.26}$$

we have $a = 1$. Then the hypersurface Σ is space-like at (x, t) if $\gamma_t(x, t) > 1/\sqrt{2}$. If Σ is given by $t = t(x)$, then Σ is space-like at $(x, t(x))$ if

$$|\nabla t(x)| < 1.$$

In the following, we use the wave equation to discuss the importance of space-like surfaces.

Let Σ be a space-like surface. Then for any $(x_0, t_0) \in \Sigma$, the domain of influence of (x_0, t_0) is given by the cone $\{(x,t); t - t_0 > |x - x_0|\}$ and hence is always above Σ. It follows that prescribing initial values on space-like surfaces yields a well-posed problem. In fact, integral domains for energy estimates can be constructed accordingly. However, we cannot prescribe initial values on time-like hypersurfaces in general. This can be explained as follows. Let Σ be a time-like hypersurface. For some point p on Σ, the domain of influence of p includes other points on Σ. Values of solutions at those points, say q, depend on those at p. Hence, initial values cannot be prescribed arbitrarily at points like q.

Of course, the above illustration is valid only for $n > 1$. The one-dimensional wave equation $u_{tt} - u_{xx} = 0$ has a symmetric form with respect to x and t, and

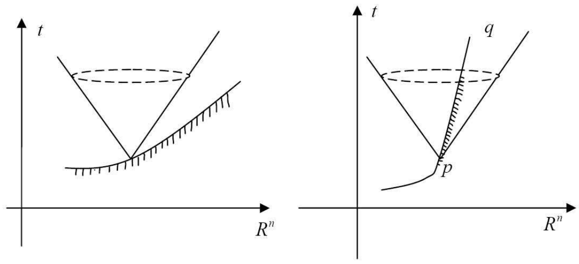

FIGURE 6.16. Space-like and time-like surfaces.

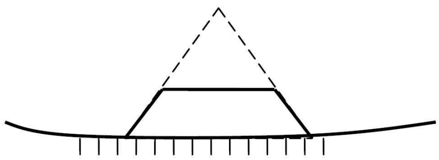

FIGURE 6.17. An integral domain for space-like initial hypersurfaces.

we can always turn around the influence domain if we are allowed to exchange the role of space and time. We emphasize that this works only for initial-value prob-

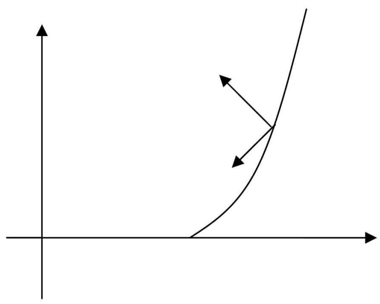

FIGURE 6.18. An illustration for $n = 1$.

lems. For initial/boundary-value problems for the one-dimensional wave equation we discussed in Section 6.1, we do not have freedom to exchange the role of space and time. The space and time variables are already fixed. The initial curve is not necessarily given by $t = 0$. It can be any space-like curve, i.e., any curve $t = t(x)$ with $|t'(x)| < 1$. The two vertical lines are time-like. We point out that we did not prescribe initial values there. (Recall that initial values consist of two conditions for a second-order PDE.) Instead, we prescribed Dirichlet boundary values $u = 0$ there.

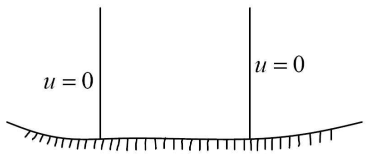

FIGURE 6.19. Initial/boundary-value problems.

To end our discussion of space-like and time-like surfaces, we consider an example of initial-value problems with initial values prescribed on a time-like hypersurface. Consider

$$\begin{aligned} &u_{tt} = u_{xx} + u_{yy} \quad \text{for } x > 0, y, t \in \mathbb{R}, \\ &u = \frac{1}{m^2}\sin my,\ \frac{\partial u}{\partial n} = \frac{1}{m}\sin my \quad \text{on } x = 0. \end{aligned}$$

Here we treat $\{x = 0\}$ as the initial hypersurface. We note that $\{x = 0\}$ is time-like. A solution is given by

$$u_m(x, y) = \frac{1}{m^2} e^{mx} \sin my.$$

Note

$$u_m \to 0 \text{ on } x = 0, \quad \frac{\partial u_m}{\partial n} \to 0 \text{ on } x = 0 \text{ as } m \to \infty.$$

Meanwhile,

$$\sup_{\mathbb{R}^2} |u_m(x, \cdot)| = \frac{1}{m^2} e^{mx} \to \infty \text{ as } m \to \infty \text{ for any } x > 0.$$

Therefore, there is no continuous dependence on initial values.

Now we return to Theorem 6.10. As in Theorem 2.12 and Corollary 2.13, we note that $\bar{t}$ enters the estimate only through the domain $D_{\kappa,T,\bar{t}}$. Hence, we let $\bar{t} \to \infty$ and obtain the following result.

THEOREM 6.12. *Assume a is C^1, c and f are continuous in $\mathbb{R}^n \times [0, T]$, φ is C^1 in $\mathbb{R}^n$ and ψ is continuous in $\mathbb{R}^n$. Let (6.24) be assumed and u be a C^2-solution of (6.23). If $f \in L^2(\mathbb{R}^n \times (0, T))$ and $\varphi, \nabla_x\varphi, \psi \in L^2(\mathbb{R}^n)$, then for any $\eta \geq \eta_0$*

$$\begin{aligned} \int_{\mathbb{R}^n \times \{T\}} e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2) + (\eta - \eta_0) \int_{\mathbb{R}^n \times (0,T)} e^{-\eta t}(u^2 + u_t^2 + a|\nabla u|^2) \\ \leq \int_{\mathbb{R}^n \times \{0\}} (\varphi^2 + \psi^2 + a|\nabla\varphi|^2) + \int_{\mathbb{R}^n \times (0,T)} e^{-\eta t} f^2, \end{aligned}$$

where η_0 is a positive constant depending only on λ, the C^1-norm of a and the L^∞-norm of c in $\mathbb{R}^n \times [0, T]$.

With Theorem 6.12, we can prove the existence of weak solutions of (6.23) by a similar process used in the proof of Theorem 2.15. However, there is a significant difference. The weak solutions in Theorem 2.15 are in L^2 because estimates of the L^2-norms of solutions are established in Corollary 2.13. In the present situation, Theorem 6.12 establishes an estimate of the L^2-norms of solutions and *their* *derivatives*. This naturally leads to a new norm defined by

$$\|u\|_{H^1(\mathbb{R}^n\times(0,T))} = \left(\int_{\mathbb{R}^n\times(0,T)} (u^2+u_t^2+|\nabla_x u|^2)\right)^{\frac{1}{2}}.$$

The superscript 1 in H^1 indicates the order of derivatives. With such a norm, we can define the Sobolev space $H^1(\mathbb{R}^n\times(0,T))$ as the completion of smooth functions of finite H^1-norms with respect to the H^1-norm. Obviously, $H^1(\mathbb{R}^n\times(0,T))$ defined in this way is complete. In fact, it is a Hilbert space, since the H^1-norm is naturally induced by an H^1-inner product given by

$$(u,v)_{H^1(\mathbb{R}^n\times(0,T))} = \int_{\mathbb{R}^n\times(0,T)} (uv+u_tv_t+\nabla_x u\cdot\nabla_x v).$$

Then proceeding as in the proof of Theorem 2.15, we can prove (6.23) admits a weak H^1-solution in $\mathbb{R}^n\times(0,T)$ if $\varphi=\psi=0$. We will not provide details here. In fact, we did not define the notion of weak H^1-solutions. The purpose of the short discussion here, similar as that at the end of Section 3.3, is to demonstrate the importance of Sobolev spaces in PDEs. We will study Sobolev spaces in detail in the future.

Exercises

(1) Let l be a positive number and consider

$$\begin{aligned}&u_{tt}-u_{xx}=0 \quad \text{in } (0,l)\times(0,\infty),\\&u(\cdot,0)=\varphi,\ u_t(\cdot,0)=\psi \quad \text{in } [0,l],\\&u(0,t)=0,\ u_x(l,t)=0 \quad \text{for } t>0.\end{aligned}$$

Find a compatibility condition and prove the existence of a C^2-solution under such a condition.

(2) Let φ_1 and φ_2 be functions defined in $\{x<0\}$ and $\{x>0\}$ respectively. Consider the characteristic initial-value problem

$$\begin{aligned}&u_{tt}-u_{xx}=0 \quad \text{for } t>|x|,\\&u(x,-x)=\varphi_1(x) \quad \text{for } x<0,\\&u(x,x)=\varphi_2(x) \quad \text{for } x>0.\end{aligned}$$

Solve and find the domain of dependence for any point (x,t) with $t>|x|$.

(3) Let φ_1 and φ_2 be functions defined in $\{x>0\}$. Consider the Goursat problem

$$\begin{aligned}&u_{tt}-u_{xx}=0 \quad \text{for } 0<t<x,\\&u(x,0)=\varphi_1(x), u(x,x)=\varphi_2(x) \quad \text{for } x>0.\end{aligned}$$

Solve and find the domain of dependence for any point (x,t) with $0<t<x$.

(4) Let α be a constant and φ and ψ be C^2-functions on $\mathbb{R}_+$ which vanish near $x = 0$. Consider

$$u_{tt} - u_{xx} = 0 \quad \text{for } x > 0, t > 0,$$
$$u(x,0) = \varphi(x), \; u_t(x,0) = \psi(x) \quad \text{for } x > 0,$$
$$u_t(0,t) = \alpha u_x(0,t) \quad \text{for } t > 0.$$

Find a solution for $\alpha \neq -1$ and prove that there exist no solutions in general for $\alpha = -1$.

(5) Let λ be a positive constant. Use the method of descent to solve the following initial-value problems for $n = 2$

$$u_{tt} = \Delta u + \lambda^2 u \quad \text{in } \mathbb{R}^2 \times \mathbb{R}_+,$$
$$u(\cdot,0) = 0, \; u_t(\cdot,0) = \psi \quad \text{in } \mathbb{R}^2,$$

and

$$u_{tt} = \Delta u - \lambda^2 u \quad \text{in } \mathbb{R}^2 \times \mathbb{R}_+,$$
$$u(\cdot,0) = 0, \; u_t(\cdot,0) = \psi \quad \text{in } \mathbb{R}^2.$$

Hint: You need to use complex functions temporarily to solve the second problem.

(6) Let u be a solution of the following initial-value problem

$$u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+,$$
$$u(\cdot,0) = 0, \; u_t(\cdot,0) = \psi \quad \text{on } \mathbb{R}^n.$$

(a) Let $n \geq 3$ be odd and $I(r,t)$ be the spherical average of u at (x,t). Set

$$J(r,t) = \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{\frac{n-3}{2}} \big(r^{n-2} I(r,t)\big).$$

Prove $J(r,t)$ satisfies the 1-dimensional wave equation, i.e., $J_{tt} - J_{rr} = 0$.

(b) Let $n \geq 3$ be odd. Derive a formula for $u(x,t)$.

(c) Let $n \geq 2$ be even. Derive a formula for $u(x,t)$ by the method of descent.

(7) Let ψ be a smooth function with compact support in $\mathbb{R}^n$ and u be a solution of

$$u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+,$$
$$u(\cdot,0) = 0, \; u_t(\cdot,0) = \psi \quad \text{on } \mathbb{R}^n.$$

Prove for any $t > 1$

$$|u(\cdot,t)|_{L^\infty(\mathbb{R}^n)} \leq Ct^{-\frac{n-1}{2}} \sum_{i=0}^{\frac{n-2}{2}} \|\nabla^i \psi\|_{L^1(\mathbb{R}^n)} \quad \text{for even } n \geq 2,$$

and

$$|u(\cdot,t)|_{L^\infty(\mathbb{R}^n)} \leq Ct^{-\frac{n-1}{2}} \|\nabla^{\frac{n-1}{2}} \psi\|_{L^1(\mathbb{R}^n)} \quad \text{for odd } n \geq 3,$$

where C is a positive constant depending only on n.

(8) Let ψ be a smooth function with compact support in $\mathbb{R}^n$ and u be a solution of

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ &u(\cdot, 0) = 0, \ u_t(\cdot, 0) = \psi \quad \text{on } \mathbb{R}^n. \end{aligned}$$

Prove for any $n \geq 2$ and for any $t > 0$

$$|u(\cdot, t)|_{L^\infty(\mathbb{R}^n)} \leq C\|\nabla^{n-1}\psi\|_{L^1(\mathbb{R}^n)},$$

where C is a positive constant depending only on n.

(9) Let u be a solution of the following initial-value problem

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \text{in } \mathbb{R}^3 \times \mathbb{R}_+, \\ &u(\cdot, 0) = \varphi, \ u_t(\cdot, 0) = \psi \quad \text{on } \mathbb{R}^3. \end{aligned}$$

(a) For any fixed $(x_0, t_0) \in \mathbb{R}^3 \times \mathbb{R}_+$, set for any $x \in B_{t_0}(x_0) \setminus \{x_0\}$

$$\mathbf{v}(x) = \left(\frac{\nabla_x u(x, t)}{|x - x_0|} + \frac{x - x_0}{|x - x_0|^3} u(x, t) + \frac{x - x_0}{|x - x_0|^2} u_t(x, t) \right) \Bigg|_{t = t_0 - |x - x_0|}.$$

Prove $div\mathbf{v} = 0$.

(b) Derive an expression of $u(x_0, t_0)$ in terms of φ and ψ by integrating $div\mathbf{v}$ in $B_{t_0}(x_0) \setminus B_\varepsilon(x_0)$ and then letting $\varepsilon \to 0$.

Remark: This exercise gives an alternative approach to solving the 3-dimensional wave equation.

(10) Let a be a positive constant and u be a C^2-solution of the following characteristic initial-value problem

$$\begin{aligned} &u_{tt} - \Delta u = 0 \quad \text{in } \{(x, t) \in \mathbb{R}^3 \times \mathbb{R}_+; t > |x| > a\}, \\ &u(x, |x|) = 0 \quad \text{for any } |x| > a. \end{aligned}$$

(a) For any fixed $(x_0, t_0) \in \mathbb{R}^3 \times \mathbb{R}_+$ with $t_0 > |x_0| > a$, integrate $div\mathbf{v}$ (introduced in the previous problem) in a region bounded by $|x - x_0| + |x| = t_0$, $|x| = a$ and $|x - x_0| = \varepsilon$. By letting $\varepsilon \to 0$, express $u(x_0, t_0)$ by an integral over ∂B_a.

(b) For any $\omega \in \mathbb{S}^2$ and $\tau > 0$, prove the limit

$$\lim_{r \to \infty} \big(ru(r\omega, r + \tau) \big)$$

exists and the convergence is uniform for $\omega \in \mathbb{S}^2$ and $\tau \in (0, \tau_0]$ for any fixed $\tau_0 > 0$.

Remark: The limit in (b) is called the *radiation field*.

(11) Set $Q_T = \{(x, t); 0 < x < l, 0 < t < T\}$. Consider the equation

$$Lu \equiv 2u_{tt} + 3u_{tx} + u_{xx} = 0.$$

(a) Give a correct presentation of the boundary-value problem in Q_T;

(b) Find an explicit expression of a solution with prescribed boundary values;

(c) Derive an estimate of the integral of $u_x^2 + u_t^2$ in Q_T.

Hint: For (b), divide Q_T into three regions separated by characteristic curves from $(0,0)$. For (c), integrate an appropriate linear combination of $u_t Lu$ and $u_x Lu$ to make integrands on $[0,l]\times\{t\}$ and $\{l\}\times[0,t]$ positive definite.

(12) For some constant $a>0$, consider the following characteristic initial-value problem for the wave equation

$$\begin{aligned} u_{tt}-\Delta u =& f(x,t) \quad \text{in } a<|x|<t+a,\\ u =& \varphi(x,t) \quad \text{on } |x|>a,\ t=|x|-a,\\ u =& \psi(x,t) \quad \text{on } |x|=a,\ t>0, \end{aligned}$$

where f is a C^1-function in $a<|x|<t+a$, φ is a C^1-function on $r_0<|x|=t-a$ and ψ is a C^1-function on $|x|=a$ and $t>0$. Derive an energy estimate in an appropriate domain in $a<|x|<t+a$.

CHAPTER 7

First-Order Differential Systems

In this chapter, we discuss partial differential systems of first-order and focus on non-characteristic initial-value problems. In Section 7.1, we introduce non-characteristic hypersurfaces for partial differential equations and systems of arbitrary order. We also demonstrate that partial differential systems of arbitrary order can always be changed to those of first-order. In Section 7.2, we discuss the Cauchy-Kowalevski Theorem, which asserts the existence of analytic solutions of non-characteristic initial-value problems for analytic differential systems and initial values on analytic non-characteristic hypersurfaces. In Section 7.3, we discuss hyperbolic differential systems and derive energy estimates for symmetric hyperbolic differential systems.

7.1. Reductions to First-Order Differential Systems

The main focus in this section is linear partial differential systems. We start with linear PDEs of arbitrary order and proceed similarly as in Section 2.1 and Section 3.1.

Let Ω be a domain in $\mathbb{R}^n$ containing the origin and L be an m-th order linear differential operator given by

$$Lu = \sum_{|\alpha|\le m} a_\alpha(x)\partial^\alpha u \quad \text{in } \Omega,$$

where a_α is continuous in Ω, for any $\alpha \in \mathbb{Z}^n_+$. Here a_α is called the coefficient for $\partial^\alpha u$.

Definition 7.1. The principal part L_0 and the principal symbol p of L are defined by

$$L_0 u = \sum_{|\alpha|=m} a_\alpha(x)\partial^\alpha u \quad \text{in } \Omega,$$

and

$$p(x;\xi) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha \quad \text{for any } x\in\Omega \text{ and } \xi\in\mathbb{R}^n.$$

The principal symbol p plays an important role in prescribing appropriate conditions associated with L. We usually write first-order and second-order linear differential operators in following forms respectively

$$Lu = \sum_{i=1}^n a_i(x)u_{x_i} + b(x)u,$$

and

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x)u_{x_ix_j} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u.$$

Their principal symbols are given by

$$p(x;\xi) = \sum_{i=1}^{n} a_i(x)\xi_i,$$

and

$$p(x;\xi) = \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j.$$

For second-order differential operators, we usually assume that (a_{ij}) is a symmetric matrix in Ω.

For a given function f in Ω, we consider the equation

$$Lu = f(x) \quad \text{in } \Omega. \tag{7.1}$$

The function f is called the *non-homogeneous term.*

Let Σ be the hyperplane $\{x_n = 0\}$. We now prescribe values of u and its derivatives on Σ so that we can at least find all derivatives of u at the origin. Let $u_0, u_1, \cdots, u_{m-1}$ be functions defined in a neighborhood of the origin in $\mathbb{R}^{n-1}$. First, we prescribe

$$u(x', 0) = u_0(x') \quad \text{for any small } x' \in \mathbb{R}^{n-1}.$$

Then we can find all x'-derivatives of u at the origin. Similarly, for any $j = 1, \cdots, m-1$, we prescribe

$$\partial_{x_n}^j u(x', 0) = u_j(x') \quad \text{for any small } x' \in \mathbb{R}^{n-1}.$$

Then we can find all x'-derivatives of $\partial_{x_n}^j u$ at the origin for $j = 1, \cdots, m-1$. Hence with the help of $u_0(x'), u_1(x'), \cdots, u_{m-1}(x')$ on Σ, we can determine all derivatives of u of order up to m at the origin except $\partial_{x_n}^m u$. To find this, we need to use the equation. If we assume

$$a_{(0,\cdots,0,m)}(0) \neq 0,$$

then we can find $\partial_{x_n}^m u(0)$ from (7.1). In this case, we can compute all derivatives of u of any order at the origin by using values $u_0, u_1, \cdots, u_{m-1}$ and differentiating (7.1).

Now we summarize these conditions on $\Sigma = \{x_n = 0\}$ by writing

$$\partial_{x_n}^j u(x', 0) = u_j(x') \quad \text{for any small } x' \in \mathbb{R}^{n-1} \text{ and } j = 0, 1, \cdots, m-1. \tag{7.2}$$

We usually call Σ the initial hypersurface and $u_0, \cdots, u_{m-1}$ initial values or Cauchy values. The problem of solving (7.1) together with (7.2) is called the initial value problem or the Cauchy problem.

More generally, consider the hypersurface Σ given by $\{\varphi = 0\}$ for a C^m-function φ in a neighborhood of the origin with $\nabla\varphi \neq 0$, with the origin on the hypersurface Σ, i.e., $\varphi(0) = 0$. We note that $\nabla\varphi$ is simply a normal vector of the hypersurface Σ. Without loss of generality, we assume $\varphi_{x_n}(0) \neq 0$. Then by the implicit function theorem, we solve $\varphi = 0$ around $x = 0$ for $x_n = \psi(x_1, \cdots, x_{n-1})$. Consider a change of variables

$$x \mapsto y = (x_1, \cdots, x_{n-1}, \varphi(x)).$$

This is a well defined transform in a neighborhood of the origin with a nonsingular Jacobian. Now we write the operator L in new variables y as

$$Lu = \sum_{|\alpha|\le m} \tilde{a}_\alpha(y)\partial_y^\alpha u.$$

The initial hypersurface Σ is given by $\{y_n = 0\}$ in new coordinates. We are interested in the coefficient $\tilde{a}_{(0,\cdots,0,m)}$ of $\partial_{y_n}^m u$. Note

$$u_{x_i} = \sum_{k=1}^n y_{k,x_i} u_{y_k} = \varphi_{x_i} u_{y_n} + \text{ terms not involving } u_{y_n},$$

and hence

$$\partial_x^\alpha u = \partial_{x_1}^{\alpha_1}\cdots\partial_{x_n}^{\alpha_n} u = \varphi_{x_1}^{\alpha_1}\cdots\varphi_{x_n}^{\alpha_n}\partial_{y_n}^m u + \text{terms not involving } \partial_{y_n}^m u.$$

Therefore,

$$\begin{aligned} Lu = \sum_{|\alpha|\le m} a_\alpha(x)\partial_x^\alpha u &= \sum_{|\alpha|=m} a_\alpha\big(x(y)\big)\varphi_{x_1}^{\alpha_1}\cdots\varphi_{x_n}^{\alpha_n}\partial_{y_n}^m u \\ &\quad + \text{ terms not involving } \partial_{y_n}^m u. \end{aligned}$$

The coefficient of $\partial_{y_n}^m u$ is given by

$$\sum_{|\alpha|=m} a_\alpha\big(x(y)\big)\varphi_{x_1}^{\alpha_1}\cdots\varphi_{x_n}^{\alpha_n}.$$

Back to x-variables, we note that this is simply

$$p\big(x; \nabla\varphi(x)\big),$$

where p is the principal symbol of the operator L.

Definition 7.2. For a linear operator L as in (7.1) defined in a neighborhood of $x_0 \in \mathbb{R}^n$, a C^1-hypersurface Σ passing x_0 is non-characteristic at x_0 if

$$p(x_0;\xi) = \sum_{|\alpha|=m} a_\alpha(x_0)\xi^\alpha \ne 0, \tag{7.3}$$

where ξ is a normal vector of Σ at x_0. Otherwise, it is characteristic at x_0. A hypersurface is non-characteristic if it is non-characteristic at every point.

When the hypersurface Σ is given by $\{\varphi = 0\}$ with $\nabla\varphi \ne 0$, its normal vector is given by $\nabla\varphi = (\varphi_{x_1}, \cdots, \varphi_{x_n})$. Hence we may take $\xi = \nabla\varphi(x_0)$ in (7.3). We note that the condition (7.3) is maintained under C^m-changes of local coordinates. By this condition, we can find successively values of all derivatives of u at x_0, as far as they exist. Then, we could write *formal* power series at x_0 for solutions of initial-value problems. It would be actual representations of solutions u in a neighborhood of x_0, if u were known to be analytic. This process can be carried out for analytic initial values and analytic coefficients and nonhomogeneous terms and the result is referred to as the Cauchy-Kowalevski theorem. We will discuss it in Section 7.2.

Now we introduce a special class of linear differential operators.

Definition 7.3. The linear differential operator L in (7.1) is *elliptic* at x_0 if $p(x_0;\xi) \ne 0$ for any $\xi \in \mathbb{R}^n \setminus \{0\}$, where p is the principal symbol of L.

Hence, linear differential equations are elliptic if there are no characteristic hypersurfaces. In other words, every hypersurface is non-characteristic.

For example, consider a first-order linear differential operator of the form

$$Lu = \sum_{i=1}^{n} a_i(x)u_{x_i} + b(x)u \quad \text{in } \Omega \subset \mathbb{R}^n.$$

Its principal symbol is given by

$$p(x;\xi) = \sum_{i=1}^{n} a_i(x)\xi_i \quad \text{for any } x \in \Omega \text{ and any } \xi \in \mathbb{R}^n \setminus \{0\}.$$

Hence first-order linear differential equations with real coefficients are never elliptic. Complex coefficients may yield elliptic equations. For example, we take $a_1 = 1/2$ and $a_2 = 1/2i$ in $\mathbb{R}^2$. Then $\partial_z = (\partial_x + i\partial_y)/2$ is elliptic.

For second-order linear differential operators, we write

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x)u_{x_ix_j} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u \quad \text{in } \Omega \subset \mathbb{R}^n.$$

Its principal symbol is given by

$$p(x;\xi) = \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \quad \text{for any } x \in \Omega \text{ and any } \xi \in \mathbb{R}^n \setminus \{0\}.$$

Then L is elliptic at $x \in \Omega$ if

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \neq 0 \quad \text{for any } \xi \in \mathbb{R}^n \setminus \{0\}.$$

If $(a_{ij}(x))$ is a real-valued $n \times n$ symmetric matrix, L is elliptic at x if $(a_{ij}(x))$ is a definite matrix at x, positive definite or negative definite.

The concept of non-characteristics can also be generalized to linear partial differential systems. Let $\Omega \subset \mathbb{R}^n$ be a domain. Consider a linear partial differential systems of N equations for N unknowns as follows

$$Lu \equiv \sum_{|\alpha| \le m} A_\alpha(x)\partial^\alpha u = f \quad \text{in } \Omega, \tag{7.4}$$

where A_α are $N \times N$ matrices and u and f are N-column vectors.

We define principle parts, principle symbols and non-characteristic hypersurfaces similarly as for single differential equations.

Definition 7.4. The principal part L_0 and the principal symbol p of L are defined by

$$L_0 u = \sum_{|\alpha| = m} A_\alpha(x)\partial^\alpha u \quad \text{in } \Omega,$$

and

$$p(x;\xi) = \det\Big(\sum_{|\alpha| = m} A_\alpha(x)\xi^\alpha\Big) \quad \text{for any } x \in \Omega \text{ and } \xi \in \mathbb{R}^n.$$

DEFINITION 7.5. For a linear operator L as in (7.4) defined in a neighborhood of $x_0 \in \mathbb{R}^n$, a C^1-hypersurface Σ passing x_0 is non-characteristic at x_0 if

$$p(x_0;\xi) = \det\big(\sum_{|\alpha|=m} A_\alpha(x_0)\xi^\alpha\big) \neq 0,$$

where ξ is a normal vector of Σ at x_0. Otherwise, it is characteristic at x_0. A hypersurface is non-characteristic if it is non-characteristic at every point.

Let Σ be a non-characteristic hypersurface. We prescribe initial values on Σ as follows. Let ν be a normal vector on Σ. We prescribe

$$\frac{\partial^j u}{\partial \nu^j} = u_j \quad \text{on } \Sigma \text{ for } j = 0, 1, \cdots, m-1, \tag{7.5}$$

where $u_0, u_1, \cdots, u_{m-1}$ are N-column vectors on Σ.

We now demonstrate that we can always reduce the order of differential systems to 1 by increasing the number of equations and the number of components of solution vectors.

PROPOSITION 7.6. *Let Σ be a non-characteristic hypersurface at $x_0 \in \Sigma$. Then the initial-value problem (7.4)-(7.5) in a neighborhood of x_0 is equivalent to an initial-value problem of a first-order differential system with initial values prescribed on Σ.*

PROOF. We assume x_0 is the origin. In the following, we write $x = (x', x_n)$ and $\alpha = (\alpha', \alpha_n)$.

Step 1. Straightening initial hypersurfaces. Assume Σ is given by $\{\varphi = 0\}$ for a C^m-function φ in a neighborhood of the origin with $\varphi_{x_n} \neq 0$. Then we introduce a change of coordinates $x = (x', x_n) \mapsto (x', \varphi(x))$. In the new coordinates, still denoted by x, the hypersurface Σ is given by $\{x_n = 0\}$ and the initial condition (7.5) is given by

$$\partial_{x_n}^j u(x', 0) = u_j(x') \quad \text{for } j = 0, 1, \cdots, m-1.$$

Step 2. Reductions to canonical forms and zero initial values. In the new coordinates, $\{x_n = 0\}$ is non-characteristic at 0. Then, the coefficient matrix $A_{(0,\cdots,0,m)}(x)$ is a nonsingular matrix at $x = 0$ and hence also in a neighborhood of the origin. Multiplying the partial differential system by the inverse of this matrix, we assume that $A_{(0,\cdots,0,m)}$ is the identity matrix in a neighborhood of the origin. Next, we may assume

$$u_j(x') = 0 \quad \text{for } j = 0, 1, \cdots, m-1.$$

To this end, we introduce a function v such that

$$u = v + \sum_{j=0}^{m-1} \frac{1}{j!} u_j(x') x_n^j.$$

Then the differential system for v is the same as that for u with f replaced by

$$f(x) - \sum_{j=0}^{m-1} \sum_{|\alpha| \le m} A_\alpha(x) \partial^\alpha\big(\frac{1}{j!} u_j(x') x_n^j\big).$$

With Step 1 and Step 2, we assume (7.4)-(7.5) have the following form

$$\partial_{x_n}^m u + \sum_{\alpha_n=0}^{m-1} \sum_{|\alpha'| \le m-\alpha_n} A_\alpha \partial^\alpha u = f,$$

with

$$\partial_{x_n}^j u(x',0) = 0 \quad \text{for } j = 0, 1, \cdots, m-1.$$

Step 3. Lowering the order. We now replace the above system by an equivalent one of order $m-1$. Introduce new functions

$$U_0 = u, \quad U_i = u_{x_i}, \ i = 1, \cdots, n,$$

and

$$U = (U_0, U_1, \cdots, U_n)^T.$$

Then

$$U_{0,x_n} = U_n, \quad U_{i,x_n} = U_{n,x_i}, \ i = 1, \cdots, n-1.$$

Hence

$$\begin{aligned}
&\partial_{x_n}^{m-1} U_0 - \partial_{x_n}^{m-2} U_n = 0, \\
&\partial_{x_n}^{m-1} U_i - \partial_{x_i} \partial_{x_n}^{m-2} U_n = 0, \quad i = 1, \cdots, n-1.
\end{aligned}$$

To get an $(m-1)$-th order differential equation for U_n, we write the equation for u as

$$\partial_{x_n}^m u + \sum_{\alpha_n=1}^{m-1} \sum_{|\alpha'| \le m-\alpha_n} A_\alpha \partial^\alpha u + \sum_{|\alpha'| \le m} A_{(\alpha',0)} \partial^{(\alpha',0)} u = f.$$

We substitute $U_n = u_{x_n}$ in the first two terms in the right-hand side to get

$$\partial_{x_n}^{m-1} U_n + \sum_{\alpha_n=0}^{m-2} \sum_{|\alpha'| \le m-\alpha_n-1} A_\alpha \partial^\alpha U_n + \sum_{|\alpha'| \le m} A_{(\alpha',0)} \partial^{(\alpha',0)} u = f.$$

In the last summation in the left hand side, any m-th order term of u can be changed to an $(m-1)$-th order term for U_i for some $i = 1, \cdots, n-1$, since no derivatives with respect to x_n are involved. Now we can write a differential system for U in the following form

$$\partial_{x_n}^{m-1} U + \sum_{\alpha_n=0}^{m-2} \sum_{|\alpha'| \le m-\alpha_n-1} A_\alpha^{(1)} \partial^\alpha U = F^{(1)}.$$

The initial value for U is given by

$$\partial_{x_n}^j U(x',0) = 0 \quad \text{for } j = 0, 1, \cdots, m-2.$$

Now we prove that the initial-value problem for the new $(m-1)$-th order differential system is equivalent to that for the original m-th order differential system. In particular, if U is a solution of the new $(m-1)$-th order differential system, then U_0 is a solution of the original m-th order differential system. To this end, we prove $U_i = U_{0,x_i}$, for $i = 1, \cdots, n$. First, by the first equation in the system for U and initial conditions for U, we have

$$\partial_{x_n}^{m-2}(U_n - U_{0,x_n}) = 0,$$

and on $t = 0$

$$\partial_{x_n}^j (U_n - U_{0,x_n}) = 0 \quad \text{for } j = 0, \cdots, m-3.$$

This implies easily $U_n = U_{0,x_n}$. Next, for $i = 1, \cdots, n-1$,

$$\partial_{x_n}^{m-1} U_i - \partial_{x_i} \partial_{x_n}^{m-2} U_n = \partial_{x_n}^{m-1} U_i - \partial_{x_i} \partial_{x_n}^{m-1} U_0 = \partial_{x_n}^{m-1} (U_i - \partial_{x_i} U_0).$$

By the second equation in the system for U and initial conditions, we have

$$\partial_{x_n}^{m-1} (U_i - \partial_{x_i} U_0) = 0,$$

and on $t = 0$

$$\partial_{x_n}^j (U_i - \partial_{x_i} U_0) = 0 \quad \text{for } j = 0, \cdots, m-2.$$

Hence, $U_i = U_{0,x_i}$, for $i = 1, \cdots, n-1$. Substituting $U_i = U_{0,x_i}$ in the final equation in the system for U, we conclude that U_0 is a solution for the original m-th order differential system.

Now, we can repeat the procedure to reduce m to 1. □

We point that straightening initial hypersurfaces and reducing initial values to zero are frequently used techniques in discussions of initial-value problems.

7.2. Cauchy-Kowalevski Theorem

For any given first-order linear partial differential system in a neighborhood of $x_0 \in \mathbb{R}^n$ and an initial value u_0 prescribed on a hypersurface Σ passing x_0, we first intend to find a solution u formally. To this end, we need to determine all derivatives of u at x_0, in terms of the initial value u_0 and all known functions in the equation. Obviously, all tangential derivatives (with respect to Σ) of u are given by derivatives of u_0. In order to find derivatives of u involving the normal direction, we need help of the equation. It has been established that, if Σ is non-characteristic at x_0, the initial-value problem leads to evaluations of all derivatives of u at x_0. This is clearly a necessary first step to the determination of a solution of the initial-value problem. If the coefficient matrices and initial values are analytic, a Taylor series solution could be developed for u. The Cauchy-Kowalevski Theorem asserts the convergence of this Taylor series in a neighborhood of x_0.

To motivate our discussion, we study an example of first-order partial differential systems which may not have solutions in any neighborhood of the origin, unless initial values prescribed on analytic non-characteristic hypersurfaces are analytic.

Example 7.7. Let $g = g(x)$ be a real valued function in $\mathbb{R}$. Consider the following system in $\mathbb{R}^2_+ = \{(x, y); y > 0\}$

$$\begin{aligned} u_x - v_y &= 0, \\ u_y + v_x &= 0, \end{aligned} \tag{7.6}$$

with prescribed initial values given by

$$u = g(x), \ v = 0 \quad \text{on } y = 0.$$

Note that (7.6) is simply the Cauchy-Riemann equation in $\mathbb{C} = \mathbb{R}^2$. It can be written in the matrix form as follows

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_y + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Note that $\{y = 0\}$ is non-characteristic. In fact, there are no characteristic curves. To see this, we take any $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ and then have

$$\xi_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \xi_2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \xi_1 & \xi_2 \\ -\xi_2 & \xi_1 \end{pmatrix}.$$

Its determinant is $\xi_1^2 + \xi_2^2$, which is not zero for any $\xi \neq 0$. Therefore there are no characteristic curves. We now write (7.6) in complex form. Suppose we have a solution (u, v) for (7.6) with the given initial values and let $w = u + iv$. Then

$$w_x + iw_y = 0 \text{ for } y > 0 \text{ with } w|_{y=0} = g(x).$$

Therefore, w is (complex) analytic in the upper half plane and its imaginary part is zero on the x-axis. By the Schwartz reflection principle, w can be extended across $y = 0$ to an analytic function in $\mathbb{C} = \mathbb{R}^2$. This implies in particular that g is (real) analytic since $w|_{y=0} = g$. We conclude that (7.6) does not admit any solutions with the given initial value on $\{y = 0\}$ unless g is real analytic.

Example 7.7 naturally leads to discussions of analytic solutions. We introduced real analyticity in Section 4.3. We now discuss this subject in detail.

To introduce (real) analytic functions, we need to study convergence of multiple infinite series of the form

$$\sum_\alpha c_\alpha,$$

where c_α are real numbers defined for all multi-indices $\alpha \in \mathbb{Z}_+^n$. Throughout this section, the term *convergence* always refers to *absolute convergence*. Hence, a series $\sum_\alpha c_\alpha$ is convergent if and only if $\sum_\alpha |c_\alpha| < \infty$. Here, the summation is over all multi-indices $\alpha \in \mathbb{Z}_+^n$.

Definition 7.8. A function $u : \mathbb{R}^n \to \mathbb{R}$ is called analytic near 0 if there exists an $r > 0$ and constants $\{u_\alpha\}$ such that

$$u(x) = \sum_\alpha u_\alpha x^\alpha \quad \text{for any } |x| < r.$$

If u is analytic near 0, then u is smooth near 0. Moreover, the constants u_α are given by

$$u_\alpha = \frac{1}{\alpha!}\partial^\alpha u(0) \quad \text{for any } \alpha \in \mathbb{Z}_+^n.$$

Thus u equals its Taylor series about 0, i.e.,

$$u(x) = \sum_\alpha \frac{1}{\alpha!}\partial^\alpha u(0) x^\alpha \quad \text{for any } |x| < r.$$

Now we give an important analytic function.

Example 7.9. For $r > 0$, set

$$u(x) = \frac{r}{r - (x_1 + \cdots + x_n)} \quad \text{for any } |x| < \frac{r}{\sqrt{n}}.$$

Then

$$u(x) = \left(1 - \frac{x_1 + \cdots + x_n}{r}\right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{x_1 + \cdots + x_n}{r}\right)^k$$

$$= \sum_{k=0}^{\infty} \frac{1}{r^k} \sum_{|\alpha|=k} \binom{k}{\alpha} x^\alpha = \sum_{\alpha} \frac{|\alpha|!}{r^{|\alpha|}\alpha!} x^\alpha.$$

This power series is absolutely convergent for $|x| < r/\sqrt{n}$ since

$$\sum_{\alpha} \frac{|\alpha|!}{r^{|\alpha|}\alpha!} |x^\alpha| = \sum_{k=0}^{\infty} \left(\frac{|x_1| + \cdots + |x_n|}{r}\right)^k < \infty,$$

for $|x_1| + \cdots + |x_n| \leq |x|\sqrt{n} < r$. We also note

$$\partial^\alpha u(0) = \frac{|\alpha|!}{r^{|\alpha|}} \quad \text{for any } \alpha \in \mathbb{Z}_+^n.$$

We point out that all derivatives of u at 0 are positive.

An effective method to prove analyticity of functions is to control their derivatives by derivatives of functions known to be analytic. For this, we first introduce the following terminology.

Definition 7.10. Let u and v be two C^∞-functions defined in a neighborhood of the origin in $\mathbb{R}^n$. Then v majorizes u, denoted by $v \gg u$ or $u \ll v$, if

$$\partial^\alpha v(0) \geq |\partial^\alpha u(0)| \quad \text{for any } \alpha \in \mathbb{Z}_+^n.$$

We also call v a majorant of u.

We have the following simple method to verify analyticity.

Lemma 7.11. *If $v \gg u$ and v is analytic near 0, then u is analytic near 0.*

Proof. We prove that the Taylor series of u about 0 converges for $|x| < r$ if the Taylor series of v about 0 converges for $|x| < r$. We simply note

$$\sum_{\alpha} \frac{1}{\alpha!} |\partial^\alpha u(0) x^\alpha| \leq \sum_{\alpha} \frac{1}{\alpha!} \partial^\alpha v(0) |x^\alpha| < \infty \quad \text{for any } |x| < r.$$

We hence have the desired convergence for u. □

Next, we prove that every analytic function has a majorant.

Lemma 7.12. *If the Taylor series of u is convergent for $|x| < r$ and $0 < s\sqrt{n} < r$, then u has an analytic majorant for $|x| \leq s/\sqrt{n}$.*

Proof. By setting $y = s(1, \cdots, 1)$, we have $|y| = s\sqrt{n} < r$ and hence

$$\sum_{\alpha} \frac{1}{\alpha!} \partial^\alpha u(0) y^\alpha$$

is a convergent series. Then there exists a constant C such that

$$\left|\frac{1}{\alpha!} \partial^\alpha u(0) y^\alpha\right| \leq C \quad \text{for any } \alpha \in \mathbb{Z}_+^n,$$

and in particular,

$$\left|\frac{1}{\alpha!} \partial^\alpha u(0)\right| \leq \frac{C}{y_1^{\alpha_1} \cdots y_n^{\alpha_n}} \leq C \frac{|\alpha|!}{s^{|\alpha|}\alpha!}.$$

Now set

$$v(x) \equiv \frac{Cs}{s-(x_1+\cdots+x_n)} = C\sum_\alpha \frac{|\alpha|!}{s^{|\alpha|}\alpha!}.$$

Then v majorizes u for $|x| \le s/\sqrt{n}$. □

So far, our discussions are limited to scalar-valued functions. All results can be generalized to vector-valued functions easily. For example, a vector-valued function $u = (u_1, \cdots, u_N)$ is analytic if each of its component is analytic.

We have following results for compositions of functions.

LEMMA 7.13. *Let u, v be C^∞-functions of a neighborhood of $0 \in \mathbb{R}^n$ into $\mathbb{R}^m$ and f, g be C^∞-functions of a neighborhood of $0 \in \mathbb{R}^m$ into $\mathbb{R}^N$, with $u(0) = 0$, $f(0) = 0$, $u \ll v$ and $f \ll g$. Then $f \circ u \ll g \circ v$.*

LEMMA 7.14. *Let u be an analytic function of a neighborhood of $0 \in \mathbb{R}^n$ into $\mathbb{R}^m$ and f be an analytic function of a neighborhood of $u(0) \in \mathbb{R}^m$ into $\mathbb{R}^N$. Then $f \circ u$ is an analytic function in a neighborhood of $0 \in \mathbb{R}^n$.*

We leave proof as an exercise.

Now we are ready to discuss real analytic solutions of non-characteristic initial-value problems for real analytic equations and initial values. We study first-order quasilinear partial differential systems of N equations for N unknowns with initial values prescribed on non-characteristic hyperplanes.

For convenience, we study initial-value problems in $\mathbb{R}^{n+1} = \{(x,t)\}$ with initial values prescribed on $\{t = 0\}$. Consider

$$u_t = \sum_{j=1}^n A_j(x,t,u)u_{x_j} + F(x,t,u), \tag{7.7}$$

with

$$u = u_0 \quad \text{on } \{t = 0\}, \tag{7.8}$$

where $A_1, \cdots, A_n$ are $N \times N$ matrices in (x,t) and F are N-column vectors. We assume all functions are analytic in their arguments.

The next result is referred to as the Cauchy-Kowalevski theorem.

THEOREM 7.15. *Let $A_1, \cdots, A_n$ be analytic $N \times N$ matrices and F be an analytic N-column vector near $(0, 0, u_0(0)) \in \mathbb{R}^{n+1+N}$ and u_0 be an analytic function near $0 \in \mathbb{R}^n$. Then (7.7)-(7.8) admits an analytic solution u near $0 \in \mathbb{R}^{n+1}$.*

PROOF. Without loss of generality, we assume $u_0(x) = 0$. To this end, we introduce an analytic function v by $v(x,t) = u(x,t) - u_0(x)$. Then the differential system for v is similar as that for u. Next, we add t as an additional component of u by introducing u_{N+1} such that $u_{N+1,t} = 1$ and $u_{N+1}|_{t=0} = 0$. This increases the number of equations and the number of components of solution vectors in (7.7) by 1 and meanwhile deletes t from $A_1, \cdots, A_n$ and F. We still denote by N the number of equations and the number of components of solution vectors.

In the following, we study

$$u_t = \sum_{j=1}^n A_j(x,u)u_{x_j} + F(x,u), \tag{7.9}$$

with

$$u = 0 \quad \text{on } \{t = 0\}, \tag{7.10}$$

where $A_1, \cdots, A_n$ are analytic $N \times N$ matrices and F is an analytic N-column vector in a neighborhood of the origin in $\mathbb{R}^{n+N}$. We seek an analytic solution u in a neighborhood of the origin in $\mathbb{R}^{n+1}$. To this end, we will compute $\partial^\alpha u(0)$ in terms of derivatives of $A_1, \cdots, A_n$ and F at $(0, 0) \in \mathbb{R}^{n+N}$ and then prove that the Taylor series of u at 0 converges in a neighborhood of $0 \in \mathbb{R}^{n+1}$. We note that t does not appear explicitly in the right hand side of (7.9).

Since $u = 0$ on $\{t = 0\}$, we have

$$\partial_x^\alpha u(0) = 0 \quad \text{for any } \alpha \in \mathbb{Z}_+^n.$$

Fix $i = 1, \cdots, n$ and differentiate (7.9) with respect to x_i to get

$$u_{x_i t} = \sum_{j=1}^n (A_j u_{x_i x_j} + A_{j,x_i} u_{x_j} + A_{j,u} u_{x_i} u_{x_j}) + F_u u_{x_i} + F_{x_i}.$$

In view of (7.10), we have

$$u_{x_i t}(0) = F_{x_i}(0, 0).$$

More generally, we obtain by induction

$$\partial_x^\alpha \partial_t u(0) = \partial_x^\alpha F(0, 0) \quad \text{for any } \alpha \in \mathbb{Z}_+^n.$$

Next, for any $\alpha \in \mathbb{Z}_+^n$, we have

$$\begin{aligned} \partial_x^\alpha \partial_t^2 u =& \partial_x^\alpha (u_t)_t = \partial_x^\alpha \big(\sum_{j=1}^n A_j u_{x_j} + F \big)_t \\ =& \partial_x^\alpha \big(\sum_{j=1}^n (A_j u_{x_j t} + A_{j,u} u_t u_{x_j}) + F_u u_t \big). \end{aligned}$$

Here we used the fact that A_j and F are independent of t. Thus,

$$\partial_x^\alpha \partial_t^2 u(0) = \partial_x^\alpha \big(\sum_{j=1}^n (A_j u_{x_j t} + A_{j,u} u_t u_{x_j}) + F_u u_t \big)\big|_{x=0,t=0,u=0}.$$

The expression on the right hand side can be worked out to be a polynomial with nonnegative coefficients involving various derivatives of $A_1, \cdots, A_n$ and F and derivatives $\partial_x^\beta \partial_t^l u$ with $|\beta| + l \le |\alpha| + 2$ and $l \le 1$.

More generally, for any $\alpha \in \mathbb{Z}_+^n$ and $k \ge 0$, we have

$$\partial_x^\alpha \partial_t^k u(0) = p_{\alpha,k}(\partial_x^\eta \partial_u^\gamma A_1, \ \cdots, \ \partial_x^\eta \partial_u^\gamma A_n, \ \partial_x^\eta \partial_u^\gamma F, \ \partial_x^\beta \partial_t^l u)\big|_{x=0,t=0,u=0}, \tag{7.11}$$

where $p_{\alpha,k}$ is a polynomial with nonnegative coefficients and indices η, γ, β range over $\eta, \beta \in \mathbb{Z}_+^n$ and $\gamma \in \mathbb{Z}_+^N$ with $|\eta| + |\gamma| \le |\alpha| + k - 1$, $|\beta| + l \le |\alpha| + k$ and $l \le k - 1$.

We need to point out that $p_{\alpha,k}(\partial_x^\eta \partial_u^\gamma A_1, \cdots)$ is considered as a polynomial of components of $\partial_x^\eta \partial_u^\gamma A_1, \cdots$. We denote by $p_{\alpha,k}(|\partial_x^\eta \partial_u^\gamma A_1|, \cdots)$ the value of $p_{\alpha,k}$ when all components of $\partial_x^\eta \partial_u^\gamma A_1, \cdots$ are replaced by their absolute values. Since $p_{\alpha,k}$ has nonnegative coefficients, we conclude

$$\begin{aligned} &\big|p_{\alpha,k}(\partial_x^\eta \partial_u^\gamma A_1, \ \cdots, \ \partial_x^\eta \partial_u^\gamma A_n, \ \partial_x^\eta \partial_u^\gamma F, \ \partial_x^\beta \partial_t^l u)\big|_{x=0,t=0,u=0}\big| \\ \le& p_{\alpha,k}(|\partial_x^\eta \partial_u^\gamma A_1|, \ \cdots, \ |\partial_x^\eta \partial_u^\gamma A_n|, \ |\partial_x^\eta \partial_u^\gamma F|, \ |\partial_x^\beta \partial_t^l u|)\big|_{x=0,t=0,u=0}. \end{aligned} \tag{7.12}$$

We now consider a new differential system

$$
\begin{aligned}
v_t =& \sum_{j=1}^{n} B_j(x,v) v_{x_i} + G(x,v) \\
v =& 0 \quad \text{on } \{t=0\},
\end{aligned}
\tag{7.13}
$$

where $B_1, \cdots, B_N$ are analytic $N \times N$ matrices and G is an analytic N-column vector in a neighborhood of the origin in $\mathbb{R}^{n+N}$. We will choose $B_1, \cdots, B_n$ and G such that

$$
B_j \gg A_j \text{ for } j = 1, \cdots, n \quad \text{and} \quad G \gg F. \tag{7.14}
$$

Hence, for any $(\eta, \gamma) \in \mathbb{Z}_+^{n+N}$,

$$
\partial_x^\eta \partial_u^\gamma B_j(0) \geq |\partial_x^\eta \partial_u^\gamma A_j(0)| \text{ for } j = 1, \cdots, n \quad \text{and} \quad \partial_x^\eta \partial_u^\gamma G(0) \geq |\partial_x^\eta \partial_u^\gamma F(0)|.
$$

The above inequalities should be understood as for each components.

Let v be a solution of (7.13). We now claim

$$
|\partial_x^\alpha \partial_t^k u(0)| \leq \partial_x^\alpha \partial_t^k v(0) \quad \text{for any } (\alpha, k) \in \mathbb{Z}_+^{n+1}.
$$

The proof is by induction on the order of t-derivatives. The general step follows since

$$
\begin{aligned}
|\partial_x^\alpha \partial_t^k u(0)| =& |p_{\alpha,k}(\partial_x^\eta \partial_u^\gamma A_1, \cdots, \partial_x^\eta \partial_u^\gamma A_n, \partial_x^\eta \partial_u^\gamma F, \partial_x^\beta \partial_t^l u)\big|_{x=0,t=0,u=0}| \\
\leq& p_{\alpha,k}(|\partial_x^\eta \partial_u^\gamma A_1|, \cdots, |\partial_x^\eta \partial_u^\gamma A_n|, |\partial_x^\eta \partial_u^\gamma F|, |\partial_x^\beta \partial_t^l u|)\big|_{x=0,t=0,u=0} \\
\leq& p_{\alpha,k}(\partial_x^\eta \partial_u^\gamma B_1, \cdots, \partial_x^\eta \partial_u^\gamma B_n, \partial_x^\eta \partial_u^\gamma G, \partial_x^\beta \partial_t^l v)\big|_{x=0,t=0,v=0} \\
=& \partial_x^\alpha \partial_t^k v(0),
\end{aligned}
$$

where we used (7.11), (7.12) and the fact that $p_{\alpha,k}$ has positive coefficients. Thus

$$
v \gg u. \tag{7.15}
$$

It remains to prove the Taylor series of v at 0 converges in a neighborhood of $0 \in \mathbb{R}^{n+1}$.

To this end, we consider

$$
B_1 = \cdots = B_n = \frac{Cr}{r - (x_1 + \cdots + x_n + u_1 + \cdots + u_N)} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix},
$$

and

$$
G = \frac{Cr}{r - (x_1 + \cdots + x_n + u_1 + \cdots + u_N)} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.
$$

for positive constants C and r with $|x| + |u| < r/\sqrt{n+N}$. As demonstrated in the proof of Lemma 7.12, we may choose C sufficiently large and r sufficiently small such that (7.14) holds.

Set

$$
v = w \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},
$$

for some scalar-valued function w in a neighborhood of $0 \in \mathbb{R}^{n+1}$. Then (7.13) is reduced to

$$\begin{aligned} w_t =& \frac{Cr}{r-(x_1+\cdots+x_n+Nw)}(N\sum_{i=1}^n w_{x_i}+1), \\ w =& 0 \quad \text{on } \{t=0\}. \end{aligned}$$

This is a (single) first-order quasilinear partial differential equation and has a solution of the form

$$w(x_1, \cdots, x_n, t) = \tilde{w}(x_1+\cdots+x_n, t).$$

Then $\tilde{w} = \tilde{w}(z,t)$ satisfies

$$\begin{aligned} \tilde{w}_t =& \frac{Cr}{r-z-Nw}(nN\tilde{w}_z+1), \\ \tilde{w} =& 0 \quad \text{on } \{t=0\}. \end{aligned}$$

By using the method of characteristics as in Section 2.2, we have an explicit solution

$$\tilde{w}(z,t) = \frac{1}{(n+1)N}\big\{r-z-[(r-z)^2-2Cr(n+1)Nt]^{\frac{1}{2}}\big\}.$$

and hence

$$w(x,t) = \frac{1}{(n+1)N}\big\{r-\sum_{i=1}^n x_i-[(r-\sum_{i=1}^n x_i)^2-2Cr(n+1)Nt]^{\frac{1}{2}}\big\}.$$

This expression is analytic for $|(x,t)| < s$, for sufficiently small $s > 0$. Hence, the corresponding solution v of (7.13) is analytic for $|(x,t)| < s$. By Lemma 7.11 and (7.15), u is analytic for $|(x,t)| < s$. Since the Taylor series of the analytic functions u_t and $\sum_{j=1}^n A_j(x,u)u_{x_j} + F(x,u)$ have the same coefficients at the origin, they agree throughout the region $|(x,t)| < s$. □

At the beginning of the proof, we introduced an extra component for the solution vector to get rid of t in coefficient matrices of the differential system. Had we chosen to preserve t, we would have to solve the initial-value problem

$$\begin{aligned} \tilde{w}_t =& \frac{Cr}{r-z-t-Nw}(nN\tilde{w}_z+1), \\ \tilde{w} =& 0 \quad \text{on } \{t=0\}. \end{aligned}$$

It is difficult, if not impossible, to find an explicit expression of the solution $\tilde{w}$.

The solution given in Theorem 7.15 is the only analytic solution since all derivatives of the solution are computed at the origin and they uniquely determine the analytic solution. A natural question is whether non-analytic solution is also unique. For this, we discuss initial value problems of linear differential systems as follows

$$\begin{aligned} A_0(x,t)u_t + \sum_{j=1}^n A_j(x,t)u_{x_j} + B(x,t)u =& F(x,t), \\ u(x,0) =& 0, \end{aligned} \tag{7.16}$$

where $A_0, A_1, \cdots, A_n, B$ are analytic $N\times N$ matrices and F is an analytic N-column vector in a neighborhood of the origin in $\mathbb{R}^{n+1}$.

The next result is referred to as the local Holmgren uniqueness theorem.

THEOREM 7.16. *Let $A_0, A_1, \cdots, A_n, B$ be analytic $N \times N$ matrices and F be an analytic N-column vector near the origin in $\mathbb{R}^{n+1}$ and u_0 be an analytic function near $0 \in \mathbb{R}^n$. If $\{t = 0\}$ is non-characteristic at the origin, then any C^1-solution of (7.16) is analytic in a sufficiently small neighborhood of the origin in $\mathbb{R}^{n+1}$.*

In the proof, we need a notion of adjoint operators. Let L be a differential operator defined by

$$Lu = A_0(x,t)u_t + \sum_{i=1}^n A_i(x,t)u_{x_i} + B(x,t)u.$$

For any vector-valued functions u and v, we write

$$v^T Lu = (v^T A_0 u)_t + \sum_{i=1}^n (v^T A_i u)_{x_i} - \big(A_0^T v_t + \sum_{i=1}^n (A_i^T v)_{x_i} - B^T v\big)^T u.$$

We define the adjoint operator L^* of L by

$$\begin{aligned} L^* v = &- (A_0^T v)_t - \sum_{i=1}^n (A_i^T v)_{x_i} + B^T v \\ = &- A_0^T v_t - \sum_{i=1}^n A_i^T v_{x_i} + (B^T - A_{0,t}^T - \sum_{i=1}^n A_{i,x_i}^T)v. \end{aligned}$$

Then

$$v^T Lu = (v^T A_0 u)_t + \sum_{i=1}^n (v^T A_i u)_{x_i} + \big(L^* v\big)^T u.$$

PROOF. We introduce an analytic change of coordinates so that the initial hypersurface $\Sigma = \{t = 0\}$ becomes a paraboloid

$$t = \sum_{i=1}^n x_i^2.$$

We will prove that any C^1-solution u of $Lu = 0$ with a zero initial value on Σ is in fact zero. For any $\varepsilon > 0$, we set

$$\Omega_\varepsilon = \{(x,t); \sum_{i=1}^n x_i^2 < t < \varepsilon\}.$$

We will prove $u = 0$ in Ω_ε for a sufficiently small ε. In the following, we denote by $\partial^+\Omega_\varepsilon$ and $\partial^-\Omega_\varepsilon$ the upper and lower boundary of Ω_ε respectively, i.e.,

$$\begin{aligned} \partial^+\Omega_\varepsilon = &\{(x,t); \sum_{i=1}^n x_i^2 < t = \varepsilon\}, \\ \partial^-\Omega_\varepsilon = &\{(x,t); \sum_{i=1}^n x_i^2 = t < \varepsilon\}. \end{aligned}$$

We note that $\det(A_0(0)) \neq 0$ since Σ is non-characteristic at the origin. Hence A_0 is nonsingular in a neighborhood of the origin. By multiplying the equation in (7.16) by A_0^{-1}, we assume $A_0 = I$.

For any vector-valued function v in a neighborhood of the origin containing Ω_ε, we have

$$0 = \int_{\Omega_\varepsilon} v^T Lu = \int_{\Omega_\varepsilon} u^T L^* v + \int_{\partial^+\Omega_\varepsilon} uv^T.$$

There is no boundary integral on $\partial^-\Omega_\varepsilon$ since $u = 0$ there. Let $P_k = P_k(x)$ be an arbitrary polynomial in $\mathbb{R}^n$, $k = 1, \cdots, n$, and form $P = (P_1, \cdots, P_N)$. We consider the initial value problem

$$\begin{aligned} L^* v =& 0 \quad \text{in } B_r, \\ v =& P \quad \text{on } B_r \cap \{t = \varepsilon\}, \end{aligned}$$

where B_r is the ball with center at the origin and radius r in $\mathbb{R}^{n+1}$. The principal part of L^* is the same as that of L, except a different sign and a transpose. We fix r so that $\{t = \varepsilon\} \cap B_r$ is non-characteristic for L^*, for each small ε. By Theorem 7.15, an analytic solution v exists in B_r for any ε small. We need to point out that the domain of convergence of v is independent of P, whose components are polynomials. We choose ε small such that $\Omega_\varepsilon \subset B_r$. Then we have

$$\int_{\partial^+\Omega_\varepsilon} uP^T = 0.$$

By Weierstrass approximation theorem, any continuous function in a compact domain can be approximated in the L^∞-norm by a sequence of polynomials. Therefore, $u = 0$ on $\partial^+\Omega_\varepsilon$ for any small ε and hence in Ω_ε. □

Theorem 7.15 guarantees the existence of solutions of initial-value problems in the analytic setting. As the next example shows, we do not expect any estimates of solutions in terms of initial values.

Example 7.17. In $\mathbb{R}^2$, consider the first-order homogeneous linear differential system (7.6)

$$\begin{aligned} u_x - v_y =& 0, \\ u_y + v_x =& 0. \end{aligned}$$

Note that all coefficients are constant. As shown in Example 7.7, there are no characteristic curves. Consider for any integer $k \geq 1$

$$u_k(x, y) = \sin(kx)e^{ky}, \quad v_k(x, y) = \cos(kx)e^{ky}.$$

Then u_k and v_k satisfy (7.6) and on $\{y = 0\}$

$$u_k(x, 0) = \sin(kx), \quad v_k(x, 0) = \cos(kx).$$

Obviously,

$$u^2(x, 0) + v^2(x, 0) = 1,$$

and

$$u_k^2(x, y) + v_k^2(x, y) = e^{2ky} \to \infty \text{ as } k \to \infty \text{ for any } y > 0.$$

Therefore, there is no continuous dependence on initial values. This illustrates that initial-value problems are not well posed for (7.6).

7.3. Hyperbolic Differential Systems

In this section, we study the non-characteristic initial-value problem for partial differential systems without analyticity and introduce an important class of first-order partial differential systems – hyperbolic differential systems. The basic idea underlying the hyperbolicity of a differential system is that the initial-value problem should be well-posed. Specifically, initial values on non-characteristic hypersurface Σ are sufficient to determine unique solutions that depend continuously on values specified at points on Σ.

We consider a first-order linear differential system in $\mathbb{R}^n \times \mathbb{R} = \{(x,t)\}$ in the following form

$$A_0(x,t)u_t + \sum_{k=1}^{n} A_k(x,t)u_{x_k} + B(x,t)u = f(x,t), \tag{7.17}$$

where u and f are column vectors with N elements and $A_0, A_1, \cdots, A_n$ and B are $N \times N$ matrices. We prescribe an initial value on the hyperplane $\{t = 0\}$ by

$$u(x,0) = u_0(x). \tag{7.18}$$

In the following, we always assume that $A_0(x,t)$ is nonsingular for any (x,t), i.e.,

$$\det(A_0(x,t)) \neq 0.$$

Hence, the hypersurface $\{t = 0\}$ is non-characteristic.

For (7.17), principal symbols defined in Definition 7.4 have the following form

$$p(x,t;\xi,\lambda) = \det\big(\sum_{k=1}^{n} \xi_k A_k(x,t) + \lambda A_0(x,t)\big), \tag{7.19}$$

for any $(x,t) \in \mathbb{R}^{n+1}$ and $(\xi,\lambda) \in \mathbb{R}^{n+1}$. For any $\xi \in \mathbb{R}^n$, $p(x,t;\xi,\lambda)$ is a polynomial of λ of degree N. If $A_0 = I$, $p(x,t;\xi,\cdot)$ is simply the characteristic polynomial of the matrix

$$-\sum_{k=1}^{n} \xi_k A_k(x,t).$$

Now, we start to investigate solvability of initial value problems associated with (7.17). As discussed in Section 7.2, we can find all derivatives of u at $t = 0$, in terms of $A_0, A_1, \cdots, A_n, B, f$ and u_0. This determines a formal solution for small t. If the coefficient matrices and initial values are analytic, a Taylor series solution could be developed for u. The Cauchy-Kowaleski Theorem asserts the convergence of this Taylor series at least in a neighborhood of any point $(x, 0)$.

In this section, we discuss the non-analytic case. Extra assumptions are needed for the well-posedness of initial-value problems.

As a motivation, we study a constant coefficient system

$$A_0 u_t + \sum_{k=1}^{n} A_k u_{x_k} = 0. \tag{7.20}$$

For some constant $\lambda \in \mathbb{C}$ and constant vectors $\xi \in \mathbb{R}^n$ and $v \in \mathbb{C}^N$, we seek a solution u of the form

$$u(x,t) = e^{i(\xi\cdot x + \lambda t)}v.$$

An easy calculation shows that u is a solution of (7.20) if and only if

$$\Big(\sum_{k=1}^{n}\xi_k A_k + \lambda A_0\Big)v = 0. \tag{7.21}$$

Suppose (7.21) holds for some $\lambda \in \mathbb{C}\setminus\mathbb{R}$, $\xi \in \mathbb{R}^n \setminus \{0\}$ and $v \in \mathbb{C}^N$, with $\mathrm{Im}\lambda < 0$. We consider a family of solutions

$$u_r(x,t) = \frac{1}{\sqrt{r}}\mathrm{Re}\{e^{ir(\xi\cdot x+\lambda t)}v\} = \frac{1}{\sqrt{r}}e^{-t\mathrm{Im}\lambda}\mathrm{Re}\{e^{ir(\xi\cdot x+t\mathrm{Re}\lambda)}v\},$$

for $r > 0$. Then,

$$u_r(x,0) \to 0 \quad \text{as } r \to \infty,$$

and

$$u_r(x,t) \to \infty \quad \text{as } r \to \infty \text{ for any } t > 0.$$

This simple example illustrates that the initial-value problem for (7.20) is not well-posed if $\mathrm{Im}\lambda \neq 0$.

In summary, a necessary condition for initial-value problems (7.17) to be well-posed is that the principal symbol $p(x,t;\xi,\cdot)$ admits only real roots for any $\xi \in \mathbb{R}^n \setminus \{0\}$.

We are led naturally to the following definition of hyperbolicity.

Definition 7.18. The differential system (7.17) is *hyperbolic* at (x,t) if for any unit vector $\xi \in \mathbb{R}^n$ the principal symbol $p(x,t;\xi,\cdot)$ has N real zeros $\lambda_1,\cdots,\lambda_N$ and there exist N linearly independent $v_1,\cdots,v_N$ in $\mathbb{R}^N$ satisfying

$$\Big(\sum_{k=1}^{n}\xi_k A_k(x,t) + \lambda_i A_0(x,t)\Big)v_i = 0, \quad i = 1,\cdots,N.$$

The differential system (7.17) is *strictly hyperbolic* at (x,t) if $p(x,t;\xi,\cdot)$ has N distinct real zeros $\lambda_1,\cdots,\lambda_N$.

The notion of hyperbolicity is invariant under the change of coordinates of the form

$$x = x(y,s), t = s.$$

An important class of hyperbolic differential system is symmetric hyperbolic differential systems.

Definition 7.19. The differential system (7.17) is *symmetric hyperbolic* at (x,t) if $A_0(x,t)$, $A_1(x,t)$, $\cdots$, $A_n(x,t)$ are symmetric and $A_0(x,t)$ is positive definite.

Obviously, symmetric hyperbolic differential systems are indeed hyperbolic.

The canonical form of (7.17) is given when $A_0 = I$. In general, we can obtain canonical forms easily. Since A_0 is assumed to be nonsingular, we may multiply (7.17) by A_0^{-1} to get an equivalent system

$$u_t + \sum_{k=1}^{n}\tilde{A}_k(x,t)u_{x_k} + \tilde{B}(x,t)u = \tilde{f}(x,t).$$

Its principal symbol is simply given by

$$p(x,t;\xi,\lambda) = \det\big(\sum_{k=1}^{n} \xi_k \tilde{A}_k(x,t) + \lambda I\big).$$

If the differential system (7.17) is symmetric hyperbolic, we wish to preserve the symmetry. Since A_0 is positive definite, we may write

$$A_0 = M^t M,$$

where $M(x,t)$ is an $N \times N$ nonsingular matrix. Introduce a new dependent variable vector v by

$$v = Mu.$$

Then v satisfies

$$v_t + \sum_{k=1}^{n} \tilde{A}_k(x,t) v_{x_k} + \tilde{B}(x,t)v = \tilde{f}(x,t),$$

where $\tilde{A}_k = M^{-t} A_k M^{-1}$ is still symmetric.

Now we study solvability of the initial-value problem (7.17)-(7.18). If coefficient matrices and nonhomogeneous terms in (7.17) and initial values in (7.18) are analytic, then analytic solutions exist locally in any neighborhood of $(x,0)$. This is the Cauchy-Kawaleski Theorem. No hyperbolicity condition is required. However, we do not have any estimates for solutions in general, as Example 7.17 suggests.

If $N = 1$, the system (7.17) is reduced to a single equation for a scalar-valued function u. The hyperbolicity condition is redundant. The initial-value problem (7.17)-(7.18) can be solved by integrating along integral curves. An energy estimate was derived in Section 2.3.

If $N > 1$, the symmetry plays an essential role in solving the initial-value problem (7.17)-(7.18). Symmetric hyperbolic differential systems in general dimensions behave like single differential equations of a similar form. An energy estimate can be derived similarly and hence the problem (7.17)-(7.18) can be solved in weak sense. Now, we derive such an estimate.

We consider the initial-value problem for a first-order linear differential system in $\mathbb{R}^n \times \mathbb{R} = \{(x,t)\}$ in the following form

$$\begin{aligned} A_0(x,t)u_t + \sum_{k=1}^{n} A_k(x,t) u_{x_k} + B(x,t)u &= f(x,t), \\ u(x,0) &= u_0(x), \end{aligned} \tag{7.22}$$

where u and f are column vectors with N elements and $A_1, \cdots, A_n$ and B are $N \times N$ matrices. In the following, we always assume

$$A_0, A_1, \cdots, A_n \text{ are symmetric in } \mathbb{R}^n \times (0,T),$$

and

$$A_0 \text{ is positive definite.}$$

We take positive constants a and κ such that

$$A_0 \geq aI \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \tag{7.23}$$

and

$$A_0 + \kappa \sum_{i=1}^{n} \xi_i A_i \geq 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}_+, \tag{7.24}$$

for any $\xi \in \mathbb{R}^n$ with $|\xi| \leq 1$.

We recall the domain $D_{\kappa,T,\bar{t}}$ defined in Section 2.3

$$D_{\kappa,T,\bar{t}} = \{(x,t); \kappa|x| < \bar{t} - t,\ 0 < t < T\},$$

for fixed $T, \bar{t} > 0$. Its boundary consists of several pieces given by

$$\begin{aligned}\partial_- D_{\kappa,T,\bar{t}} =& \{(x,0); \kappa|x| < \bar{t}\},\\ \partial_s D_{\kappa,T,\bar{t}} =& \{(x,t); \kappa|x| = \bar{t} - t,\ 0 < t < T\},\\ \partial_+ D_{\kappa,T,\bar{t}} =& \{(x,T); \kappa|x| < \bar{t} - T\}.\end{aligned}$$

THEOREM 7.20. *Let $A_0, A_1, \cdots, A_n$ be C^1 symmetric $N \times N$ matrices satisfying (7.23)-(7.24), B a continuous $N \times N$ matrix, f a continuous function and u be a C^1-solution of (7.22) in $\mathbb{R}^n \times \mathbb{R}_+$. Then for any $0 < T < \bar{t}$,*

$$\int_{D_{\kappa,T,\bar{t}}} e^{-\alpha t}|u|^2 \leq C \left\{ \int_{\partial_- D_{\kappa,T,\bar{t}}} |u_0|^2 + \int_{D_{\kappa,T,\bar{t}}} e^{-\alpha t}|f|^2 \right\},$$

where α and C are positive constants depending only on a in (7.23), the C^1-norms of A_i and the sup-norm of B in $D_{\kappa,T,\bar{t}}$.

The estimate in Theorem 7.20 implies the uniqueness of solutions and the property of finite speed propagation.

Before proving Theorem 7.20, we point out the role of symmetry. The proof proceeds similarly as that for scalar equations Section 2.3. We take an inner product of the equation (7.22) with $2u$ and hence need to analyze terms such as $2u^T A_0 u_t$ and similar terms for x-derivatives. If A_0 is symmetric, then

$$2u^T A_0 u_t = u_t^T A_0 u + u^T A_0 u_t = (u^T A_0 u)_t - u^T A_{0,t} u.$$

It is not necessary to assume symmetry for B since no derivatives of u are involved with B. We simply note

$$2u^T B u = u^T (B + B^T) u,$$

and $B^T + B$ is always symmetric.

PROOF. For a nonnegative α, we multiply the equation in (7.22) by $2e^{-\alpha t}u^T$. By

$$\begin{aligned}2e^{-\alpha t}u^T A_0 u_t =& (e^{-\alpha t}u^T A_0 u)_t + \alpha e^{-\alpha t}|u|^2 - e^{-\alpha t}u^T A_{0,t} u,\\ 2e^{-\alpha t}u^T A_i u_{x_i} =& (e^{-\alpha t}u^T A_i u)_{x_i} - e^{-\alpha t}u^T A_{i,x_i} u,\end{aligned}$$

we have

$$\begin{aligned}&(e^{-\alpha t}u^T A_0 u)_t + \sum_{i=1}^{n} (e^{-\alpha t}u^T A_i u)_{x_i}\\ &\qquad + e^{-\alpha t}u^T (\alpha A_0 - \sum_{i=1}^{n} A_{i,x_i} + B + B^T) u = 2e^{-\alpha t}u^T f.\end{aligned}$$

We write $D = D_{\kappa,T,\bar{t}}$. An integration in D yields

$$\int_{\partial_+ D} e^{-\alpha t} u^T A_0 u + \int_{\partial_s D} e^{-\alpha t} u^T (\gamma_t + \sum_{i=1}^n A_i \gamma_i) u$$
$$+ \int_D e^{-\alpha t} u^T (\alpha A_0 - \sum_{i=1}^n A_{i,x_i} + B + B^T) u$$
$$= \int_{\partial_- D} u^T A_0 u + \int_D 2e^{-\alpha t} u^T f,$$

where $(\gamma_1, \cdots, \gamma_n, \gamma_t)$ is the unit exterior normal vector of $\partial_s D$ given by

$$(\gamma_1, \cdots, \gamma_n, \gamma_t) = \frac{1}{\sqrt{1+\kappa^2}} \big(\kappa \frac{x}{|x|}, 1\big).$$

First, we note by (7.24)

$$\gamma_t + \sum_{i=1}^n a_i \gamma_i = \frac{1}{\sqrt{1+\kappa^2}} (A_0 + \kappa \sum_{i=1}^n \frac{x_i}{|x|} A_i) \geq 0 \quad \text{on } \partial_s D.$$

Next, we choose α such that

$$\alpha A_0 - \sum_{i=1}^n A_{i,x_i} + B + B^T \geq 2aI \quad \text{in } D.$$

Then

$$2a \int_D e^{-\alpha t} |u|^2 \leq \sup |A_0| \int_{\partial_- D} |u_0|^2 + \int_D 2e^{-\alpha t} u^T f.$$

Here we simply dropped integrals over $\partial_+ D$ and $\partial_s D$ since they are nonnegative. The Cauchy inequality implies

$$\int_D 2e^{-\alpha t} u^T f \leq a \int_D e^{-\alpha t} |u|^2 + \frac{1}{a} \int_D e^{-\alpha t} |f|^2.$$

We then have the desired result. □

By letting $\bar{t} \to \infty$, we have the following result.

THEOREM 7.21. *Let $A_0, A_1, \cdots, A_n$ be C^1 symmetric $N \times N$ matrices satisfying (7.23)-(7.24), B a continuous $N \times N$ matrix, f a continuous functions and u be a C^1-solution of (7.22) in $\mathbb{R}^n \times \mathbb{R}_+$. For any $T > 0$, if $f \in L^2(\mathbb{R}^n \times (0,T))$ and $u_0 \in L^2(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n \times (0,T)} e^{-\alpha t} |u|^2 \leq C \left\{ \int_{\mathbb{R}^n} |u_0|^2 + \int_{\mathbb{R}^n \times (0,T)} e^{-\alpha t} |f|^2 \right\},$$

where α and C are positive constants depending only on a, the C^1-norms of A_i and the sup-norm of B in $\mathbb{R}^n \times (0,T)$.

Proceeding as in Section 2.3, we can discuss weak solutions of (7.22). For a fixed $T > 0$, we consider functions in

$$D_T = \mathbb{R}^n \times (0,T).$$

Denote by $C_0^\infty(D_T)$ the collection of smooth functions in D_T with compact supports in D_T and by $\widetilde{C}_0^\infty(D_T)$ the collection of smooth functions in D_T with compact supports in x-directions.

Set

$$Lu = A_0(x,t)u_t + \sum_{i=1}^{n} A_i(x,t)u_{x_i} + B(x,t)u \quad \text{in } D_T. \tag{7.25}$$

We introduce the adjoint operator L^* of L by

$$\begin{aligned} L^*v &= -(A_0 v)_t - \sum_{i=1}^{n}(A_i v)_{x_i} + B^T v \\ &= -A_0 v_t - \sum_{i=1}^{n} A_i v_{x_i} + (B^T - A_{0,t} - \sum_{i=1}^{n} A_{i,x_i})v. \end{aligned}$$

For any $u, v \in \widetilde{C}_0^\infty$, we have

$$v^T Lu = (v^T A_0 u)_t + \sum_{i=1}^{n}(v^T A_i u)_{x_i} - \big(L^* v\big)^T u.$$

By a simple integration in D_T, we have

$$\int_{D_T} v^T Lu = \int_{D_T} u^T L^* v + \int_{\mathbb{R}^n\times\{t=T\}} v^T A_0 u - \int_{\mathbb{R}^n\times\{t=0\}} v^T A_0 u$$
$$\text{for any } u, v \in \widetilde{C}_0^\infty(D_T).$$

We note that there are no derivatives of u in the right-hand side.

Definition 7.22. An L^2-function u is a weak solution of $Lu = f$ in D_T if

$$\int_{D_T} u^T L^* v = \int_{D_T} f^T v \quad \text{for any } v \in C_0^\infty(D_T).$$

Now we can prove the existence of weak solutions of (7.22) with homogeneous initial values by the Hahn-Banach Theorem and the Riesz Representation Theorem as in Section 2.3.

Theorem 7.23. *Let $A_0, A_1, \cdots, A_n$ be C^1 symmetric $N\times N$ matrices satisfying (7.23)-(7.24) and B a continuous $N \times N$ matrix. Then for any $f \in L^2(D_T)$, there exists a $u \in L^2(D_T)$ such that*

$$\int_{D_T} u^T L^* v = \int_{D_T} f^T v \quad \textit{for any } v \in \widetilde{C}_0^\infty(D_T) \textit{ with } v = 0 \textit{ on } t = T.$$

Moreover,

$$\|u\|_{L^2(D_T)} \le C\|f\|_{L^2(D_T)},$$

where C is a positive constant depending only on a, the C^1-norms of A_i and the sup-norm of B in $\mathbb{R}^n \times (0,T)$.

The symmetry plays an essential role in Theorem 7.20, Theorem 7.21 and Theorem 7.23. The remaining question is whether we can change an arbitrary hyperbolic differential system into a symmetric hyperbolic differential system. The discussion of this question in general case is beyond the scope of this book. Now we study a special case.

When the space dimension is one, we can change strictly hyperbolic differential systems to symmetric hyperbolic differential systems easily. For $n = 1$, we consider

$$u_t + A(x,t)u_x + B(x,t)u = f(x,t). \tag{7.26}$$

The principal symbol is given by

$$p(x,t;\xi,\lambda) = \det\big(\xi A(x,t) + \lambda I\big).$$

The hyperbolicity condition means that all eigenvalues of A are real and eigenvectors span $\mathbb{R}^N$. If the system (7.26) is a constant coefficient homogeneous system of the form

$$u_t + Au_x = 0,$$

then solution can be written explicitly. Suppose A has N real eigenvalues $\lambda_1, \cdots, \lambda_N$ and corresponding eigenvectors $v_1, \cdots, v_N$, which span $\mathbb{R}^N$. Then solution u is given by

$$u(x,t) = \sum_{k=1}^{N} \varphi_k(x - \lambda_k t) v_k,$$

where functions $\varphi_1, \cdots, \varphi_N$ are determined by the initial condition

$$u(x,0) = \sum_{k=1}^{N} \varphi_k(x) v_k.$$

Back to the general system (7.26), we introduce a new solution vector v by

$$u = Mv,$$

where M is an $N \times N$ nonsingular matrix. Then a straightforward calculation yields

$$v_t + M^{-1}AMv_x + (M^{-1}BM + M^{-1}AM_x + M^{-1}M_t)v = M^{-1}f.$$

This is a symmetric hyperbolic differential system if $M^{-1}AM$ is symmetric. Now we take M whose columns consist of eigenvectors of A. Then $M^{-1}AM$ is a diagonal matrix with eigenvalues of A as diagonal entries, which is obviously symmetric. Therefore, v satisfies a symmetric hyperbolic differential system. If M is C^1, an energy estimate can be derived. In fact, we can obtain a solution v by integrating along characteristic curves. However, the matrix M consisting of eigenvectors of A may not be C^1 when eigenvalues of high multiplicity are present. One condition to ensure regularity of eigenvectors is that all eigenvalues are simple, i.e., the differential system (7.26) is strictly hyperbolic.

To end this section, we demonstrate that second-order hyperbolic differential equations can always be changed to symmetric hyperbolic differential systems. We consider

$$w_{tt} - \sum_{i,j=1}^{n} a_{ij}(x,t) w_{x_i x_j} - \sum_{i=1}^{n} b_i(x,t) w_{x_i} - c(x,t)w = f, \tag{7.27}$$

where (a_{ij}) is positive definite in $\mathbb{R}^n \times [0,T]$. Here we assume a_{ij}, b_i and c are smooth, with

$$\lambda I \le (a_{ij}) \le \Lambda I \quad \text{in } \mathbb{R}^n \times [0,T], \tag{7.28}$$

for some positive constants $\lambda \le \Lambda$. For convenience, we write

$$A = (a_{ij}),$$

and

$$\begin{aligned}\mathbf{a_i} =&(a_{i1}, \cdots, a_{in}) \quad \text{for any } i = 1, \cdots, n,\\ \mathbf{b} =&(b_1, \cdots, b_n).\end{aligned}$$

Set

$$U \equiv \begin{pmatrix} u_{-1} \\ u_0 \\ u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} w \\ w_t \\ w_{x_1} \\ \vdots \\ w_{x_n} \end{pmatrix}. \tag{7.29}$$

This is an $(n+2)$-vector. We write the equation (7.27) as

$$\begin{aligned} u_{-1,t} &= u_0, \\ u_{0,t} &= \sum_{i,j=1}^{n} a_{ij} u_{j,x_i} + \sum_{i=1}^{n} b_i u_i + c u_{-1} + f, \\ \sum_{j=1}^{n} a_{ij} u_{j,t} &= \sum_{j=1}^{n} a_{ij} u_{0,x_j}, \quad i = 1, \cdots, n. \end{aligned}$$

Then we put it in a matrix form

$$A_0 U_t + \sum_{i=1}^{n} A_i U_{x_i} + BU = F, \tag{7.30}$$

where

$$\begin{aligned} A_0 &= \begin{pmatrix} I_{2\times 2} & \\ & A \end{pmatrix}, \\ A_i &= -\begin{pmatrix} 0 & 0 & \mathbf{0} \\ 0 & 0 & \mathbf{a_i} \\ \mathbf{0} & \mathbf{a_i}^T & 0_{n\times n} \end{pmatrix}, \quad i = 1, \cdots, n, \\ B &= -\begin{pmatrix} 0 & 1 & \mathbf{0} \\ c & 0 & \mathbf{b} \\ \mathbf{0} & \mathbf{0} & 0_{n\times n} \end{pmatrix}, \end{aligned}$$

and

$$F^T = \begin{pmatrix} 0 & f & 0 \cdots & 0 \end{pmatrix}.$$

Obviously, (7.30) is symmetric hyperbolic since $A_0, A_1, \cdots, A_n$ are symmetric and A_0 is positive definite.

We now consider the initial-value problem

$$\begin{aligned} w_{tt} - \sum_{i,j=1}^{n} a_{ij} w_{x_i x_j} - \sum_{i=1}^{n} b_i w_{x_i} - cw =& f \quad \text{in } \mathbb{R}^n \times (0, T), \\ w(\cdot, 0) = w_0, \ w_t(\cdot, 0) =& w_1 \quad \text{on } \mathbb{R}^n. \end{aligned} \tag{7.31}$$

Then U defined in (7.29) satisfies (7.30) and initial conditions

$$U = (w_0, w_1, w_{0,x_1}, \cdots, w_{0,x_n})^T \quad \text{on } \mathbb{R}^n. \tag{7.32}$$

Now we verify that the initial-value problem (7.31) is equivalent to the initial-value problem (7.30) and (7.32). We only need to prove that a solution $U = (u_{-1}, u_0, u_1, \cdots, u_n)^T$ of (7.30) and (7.32) yields a solution $w = u_{-1}$ of (7.31). First, by the first equation in the system (7.30) and the corresponding initial condition, we have $u_0 = u_{-1,t}$. Hence we have $u_{-1,t} = w_1$ on $t = 0$. Next, by the last n equations in (7.30), we have $u_{j,t} = u_{0,x_j} = u_{-1,x_jt}$, or $(u_j - u_{-1,x_j})_t = 0$. Then we get $u_j = u_{-1,x_j}$ since they match on $t = 0$. Hence $w = u_{-1}$ is a solution of (7.31).

We then have the following result as a consequence of Theorem 7.21.

THEOREM 7.24. *Let a_{ij} be C^1 and b_i, c be continuous in $\mathbb{R}^n \times (0,T)$, (a_{ij}) satisfy (7.28) and w be a C^2-solution of (7.31) in $\mathbb{R}^n \times (0,T)$. If $f \in L^2(\mathbb{R}^n \times (0,T))$ and $w_0, \nabla w_0, w_1 \in L^2(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n \times (0,T)} e^{-\alpha t}(w^2 + |\nabla w|^2 + w_t^2) \le C\Big\{ \int_{\mathbb{R}^n} (|w_0|^2 + |\nabla w_0|^2 + w_1^2) + \int_{\mathbb{R}^n \times (0,T)} e^{-\alpha t}|f|^2 \Big\},$$

where α and C are positive constants depending only on λ, Λ, the C^1-norms of a_{ij} and the sup-norm of b_i, c in $\mathbb{R}^n \times (0,T)$.

Theorem 6.12 in Section 6.3 is a special case of Theorem 7.24.

Exercises

(1) Classify the following 4-th order equation in $\mathbb{R}^3$

$$2\partial_x^4 u + 2\partial_x^2\partial_y^2 u + \partial_y^4 u - 2\partial_x^2\partial_z^2 u + \partial_z^4 u = f.$$

(2) Prove Lemma 7.13 and Lemma 7.14.

(3) Consider the initial-value problem

$$u_{tt} - u_{xx} - u = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}_+,$$
$$u(x,0) = x, \;\; u_t(x,0) = -x.$$

Find a solution in a power series expansion about the origin and identify this solution.

(4) Consider the initial-value problem for $u : \mathbb{R} \times (0,T) \to \mathbb{R}^N$ as follows

$$u_t + A(x,t)u_x = f(x,t,u) \quad \text{in } \mathbb{R} \times (0,T),$$

with

$$u(\cdot, 0) = 0 \quad \text{on } \mathbb{R},$$

where A is an $N \times N$ *diagonal* C^1-matrix on $\mathbb{R} \times (0,T)$ and $f : \mathbb{R} \times (0,T) \times \mathbb{R}^N \to \mathbb{R}^N$ is a C^2-map. Under appropriate conditions on f, prove the above initial-value problem admits a C^1-solution by using the contraction mapping principle.
Hint: It may be helpful to write in a system of equations instead of in a matrix form.

(5) Consider in $D = \{(x,t); x > 0, t > 0\} \subset \mathbb{R}^2$

$$\begin{aligned} u_t + au_x + b_{11}u + b_{12}v &= f \\ v_x + b_{12}u + b_{22}v &= g, \end{aligned}$$

with the condition

$$u(x,0) = \varphi(x) \text{ for } x > 0 \text{ and } v(0,t) = \psi(t) \text{ for } t > 0.$$

Assume a is C^1 and b_{ij} are continuous in D.

(a) Assume $a \le 0$ in D. Derive an energy estimate for (u, v) in an appropriate domain in D.

(b) Assume $a \le 0$ in D. For sufficiently small T, derive an estimate for $\sup_{[0,T]} |u(0, \cdot)|$ in terms of sup-norms of f, g, φ and ψ.

(c) Discuss whether similar estimates can be derived if a is positive somewhere along $\{(0,t); t > 0\}$.

(6) Consider in a neighborhood of the origin in $\mathbb{R}^2 = \{(x,t)\}$

$$\begin{aligned} u_t + au_x + b_{11}u + b_{12}v &= f \\ v_x + b_{12}u + b_{22}v &= g, \end{aligned}$$

with the condition

$$u(x,0) = \varphi(x) \text{ and } v(0,t) = \psi(t).$$

Assume a, b_{ij} are analytic in a neighborhood of $0 \in \mathbb{R}^2$ and φ, ψ are analytic in a neighborhood of $0 \in \mathbb{R}$.

(a) Prove all derivatives of u and v at 0 can be expressed in terms of derivatives of a, b_{ij}, φ and ψ at 0.

(b) Prove there exists an analytic solution (u, v) in a neighborhood of $0 \in \mathbb{R}^2$.

Bibliography

[1] Arnolds, V. I., *Lectures on Partial Differential Equations*, Universitext, Springer, 2004.

[2] Caffarelli, L., Cabré, X., *Fully Nonlinear Elliptic Equations*, AMS Colloquium Publications, Vol. 43, American Math. Society, Providence, RI, 1993.

[3] Evans, L., *Partial Differential Equations*, Graduate Studies in Mathematics, Vol. 19, American Math. Society, Providence, RI, 1998.

[4] Friedman, A., *Partial Differential Equations of Parabolic Type,* Prentice-Hall, Englewood Cliffs, 1964.

[5] Gilbarg, D., Trudinger, N., *Elliptic Partial Differential Equations of Second Order*, (2nd ed.), Grundlehren der Mathematischen Wisenschaften, Vol. 224, Spinger-Verlag, Berlin-New York, 1983.

[6] Han, Q., Lin, F.-H., *Elliptic Partial Differential Equations*, Courant Institute Lecture Notes, Volume 1, American Math. Society, Providence, RI, 2000.

[7] John, F., *Partial Differential Equations*, (4th ed.), Applied Math. Sciences, Vol. 1, Springer-Verlag, New York, 1991.

[8] MacRobert, T. M., *Spherical Harmonics, An Elementary Treatise on Harmonic Functions with Applications*, Pergamon Press, Oxford-New York-Toronto, 1967.

[9] Taylor, M., *Partial Differential Equations I: Basic Theory*, Applied Math. Sciences, Vol. 115, Springer-Verlag, New York, 1996.

Index

a priori estimates, 3
adjoint differential operators, 32, 178
analytic functions, 67, 172
auxiliary functions, 81, 120

Bernstein method, 83
Burgers' equation, 18

canonical forms
  first-order systems, 181
Cauchy problems, 8, 38, 166
Cauchy values, 8, 38, 166, 169
Cauchy-Kowalevski theorem, 174
characteristic curves, 9
characteristic hypersurfaces, 9, 38
  non-characteristic hypersurfaces, 167, 169
characteristic ODEs, 14, 16, 21
compact supports, 32
comparison principles, 78, 80, 118
compatibility conditions, 20, 52, 54, 137, 139
conservation of energies, 48, 153
convergence of series, 67, 172
  absolute convergence, 172
convolutions, 103

d'Alembert's formula, 134
decay estimates, 150
degenerate differential equations, 39
diameters, 45
differential Harnack inequalities, 75
Dirichlet energy, 92
Dirichlet problems, 44, 61, 74
  Green's function, 61
divergence theorem, 43
domain of influence, 13
domains, 1
domains of dependence, 13, 27, 134, 148
domains of influence, 27, 134, 148
doubling condition, 99
Duhamel's principle, 142

eigenvalue problems, 50, 56
elliptic differential equations, 39, 168
energy estimates
  first-order PDEs, 31
  second-order hyperbolic equations, 159
  symmetric hyperbolic differential systems, 184
  wave equations, 47
Euclidean norms, 1
Euler-Poisson-Darboux equation, 143

finite speed propagation, 149
finite-speed propagation, 27
first-order linear differential systems, 180
  canonical forms, 181
  initial-value problems, 180
first-order linear PDEs, 7
  initial-value problems, 24
Fourier series, 50
Fourier transforms, 102
  inverse Fourier transforms, 104
frequency, 99
fundamental solutions, 107
  Laplace equations, 60

gradient estimates
  interior gradient estimates, 66, 75, 83, 114, 121
gradients, 1
Green's formula, 30, 43
Green's function, 53, 61
  Green's function in balls, 62
Green's identity, 60

Hölder continuity, 90
  pointwise Hölder continuity, 86
half-space problems, 137
harmonic functions, 40, 59
  conjugate harmonic functions, 40
  converegence of Taylor series, 67
  differential Harnack inequalities, 75, 83
  doubling condition, 99
  frequency, 99
  Harnack inequalities, 75, 85

interior gradient estimates, 66, 75, 83
Liouville Theorem, 75
mean-value properties, 72
nodal sets, 68
removable singularity, 85
superharmonic functions, 77
Harnack inequalities, 75, 85, 129
differential Harnack inequalities, 75, 83, 125, 128
heat equation
fundamental solutions, 107
heat equations
1 dimension, 41
analyticity of solutions, 117
differential Harnack inequalities, 125, 128
Harnack inequalities, 129
initial/boundary-value problems, 46, 49
interior gradient estimates, 114, 121
n dimensions, 42
Poisson integral formula, 106
Hessian matrices, 1
Holmgren uniqueness theorem, 177
Hopf lemma, 78
Huygens' principle, 149
hyperbolic differential equations, 39, 43
hyperbolic differential systems, 181
strictly hyperbolic systems, 181
symmetric hyperbolic systems, 181
hypersurfaces, 2

initial hypersurfaces, 8, 38, 166
initial values, 8, 38, 166, 169
initial-value problems, 166
first-order PDEs, 8, 11
first-order systems, 180
second-order PDEs, 38
wave equations, 133, 142
initial/boundary-value problems
heat equations, 46, 49
wave equations, 47, 54, 139
integral curves, 13

Laplace equations, 39, 42
fundamental solutions, 60
Green's indentity, 60
Poisson integral formula, 64
Poisson kernel, 64
Legendre equations, 69
Legendre functions, 70
linear differential systems
m-th order, 168
first-order, 180
linear PDEs, 2
first-order, 7
m-th order, 165
second-order, 37

Liouville Theorem, 75
loss of differentiations, 152

majorants, 173
maximum principles, 74
strong maximum principles, 79
weak maximum principles, 77, 117, 118
mean-value properties, 72
method of characteristics, 14
method of descent, 146
method of extensions, 137, 140
method of reflections, 138, 140
method of spherical averages, 143
mixed problems, 47
multi-indices, 1

Neumann problems, 45
nodal curves, 70
nodal sets, 68
non-characteristic curves, 9
non-characteristic hypersurfaces, 9, 11, 38
non-homogeneous terms, 7, 37, 166

parabolic boundaries, 112, 117
parabolic differential equations, 43
parabolic distance, 113
Parseval's identity, 105
partial differential equations (PDEs), 2
linear PDEs, 2
quasilinear PDEs, 2
PDEs of mixed type, 41
Poincarè Lemma, 45
Poisson equations, 42, 86
weak solutions, 92
Poisson integral formula, 64, 106
Poisson kernel, 64
principal parts, 165, 168
principal symbols, 37, 165, 168, 180
propagation of singularities, 41

quasilinear PDEs, 2

radiation field, 162
Rellich's theorem, 94
removable singularity, 85

Schauder estimates, 90
pointwise Schauder estimates, 88
Schwartz class, 102
second-order linear PDEs, 37
in the plane, 39
separation of variables, 49
space variables, 1
space-like surfaces, 157
spherical harmonics, 68
sectorial harmonics, 70
tesseral harmonics, 70

zonal harmonics, 70
strictly hyperbolic systems, 181
subsolutions, 77
subharmonic functions, 77
supersolutions, 77
superharmonic functions, 77
symmetric hyperbolic systems, 181

Taylor series, 67, 172
terminal value problems, 112
time variables, 1
time-time surfaces, 157
Tricomi equation, 41

uniform ellipticity, 77

wave equations
1 dimension, 40, 133
2 dimensions, 146
3 dimensions, 144
Huygens' principle, 149
decay estimates, 150
energy estimates, 153
half-space problems, 137
initial-value problems, 133, 142
initial/boundary-value problems, 47, 54, 139
n dimensions, 42, 142
radiation field, 162
weak solutions, 32, 92, 159, 185
Weierstrass approximation theorem, 179
well-posed problems, 3