SECOND ORDER PARABOLIC DIFFERENTIAL EQUATIONS

二阶抛物微分方程

Gary M. Lieberman

World Scientific
世界图书出版公司

SECOND ORDER PARABOLIC DIFFERENTIAL EQUATIONS

Gary M. Lieberman

Iowa State University

World Scientific

世界图书出版公司

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

First published 1996
First reprint 1998

SECOND ORDER PARABOLIC DIFFERENTIAL EQUATIONS

ISBN 981-02-2883-X

PREFACE

My goal in writing this book was to create a companion volume to "Elliptic Partial Differential Equations of Second Order" by David Gilbarg and Neil S. Trudinger. Like that book, this one is an essentially self-contained exposition of the theory of second order parabolic (instead of elliptic) partial differential equations of second order, with emphasis on the theory of certain initial-boundary value problems in bounded space-time domains. In addition to the Cauchy-Dirichlet problem, which is the parabolic analog of the Dirichlet problem, I also study oblique derivative problems. Preparatory material on such topics as functional analysis and harmonic analysis is included to make the book accessible to a large audience. (Dave and Neil, I hope I succeeded.)

This book would not be possible withour the help of many people. First and foremost are David Gilbarg, my Ph.D. advisor, who started me on the road to *a priori* estimates; Neil Trudinger, whose continued encouragment, collaboration, and invitations to the Centre for Mathematics and its Applications (in fact, several chapters of this book were written at the C. M. A.) kept me going; and Howard Levine, who showed me the glory of parabolic equations. Next, but just as important, are my family: my parents, Alvin and Tillie Lieberman, without whom I would not have been possible; and my wife, Linda Lewis Lieberman, and my step-children, Ben and Jenny Lewis, who joined me during the writing of this book. Without Linda, Ben, and Jenny, the project may have been completely more quickly, but definitely not more pleasantly. Without Olga Ladyzhenskaya and Nina Ural'tseva, the whole field of *a priori* estimates would be much smaller; I thank them for their pioneering work, continued advances, and constant inspiration. I also thank Cliff Bergman, Jan Nyhus, Ruth deBoer, and Mike Fletcher for their invaluable assistance with and education on $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX and LATEX. Russell Brown and Wei Hu were my collaborators on the material in Section 6.1, and it was their prodding that led to a useful version of the results therein.

I also thank my editors at World Scientific over the years: J. G. Xu, Lam Poh Fong, and Chow Mun Zing.

There were many others who provided encouragement and advice. In particular, Suncica Canic, Emmanuele DiBenedetto, Nicola Garofalo, Nina Ivochkina, Nikolai Krylov, Paul Sacks, Mikhail Safonov, and Michael Smiley deserve special mention.

My final thanks are to all those people who gave me advice through the years and who shared their work with me.

July, 1996
Ames, Iowa USA
Gary M. Lieberman

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1. Outline of this book

In Chapter 2, we prove maximum principles for general, linear parabolic operators:

$$Pu = -u_t + a^{ij} D_{ij} u + b^i D_i u + cu,$$

where we use the summation convention and notation discussed in Section 3 of this chapter. Under weak hypotheses on the coefficients of this operator (specifically, that the matrix (a^{ij}) be positive semidefinite and that c be bounded from above), the inequalities $Pu \geq 0$ in Ω and $u \leq 0$ on $\mathcal{P}\Omega$ (the parabolic boundary of Ω) imply that $u \leq 0$ in Ω. Nirenberg's strong maximum principle implies that either $u < 0$ or $u \equiv 0$ on a suitable subset of Ω provided the conditions on the coefficients are strengthened only slightly. Also, maximum estimates for solutions of various initial-boundary value problems are examined.

Chapter 3 introduces the notion of weak solution for the problem

$$-u_t + D_i(a^{ij} D_j u) = f \text{ in } \Omega, u = \varphi \text{ on } \mathcal{P}\Omega$$

when a^{ij} is a constant, positive definite matrix and f and φ are sufficiently smooth. Unlike most discussions of this problem, we deal with noncylindrical domains directly. In particular, we show that this problem is solvable in sufficiently small space-time balls. General properties of weak solutions and weak derivatives are studied also. From the theory in balls, we derive an existence theorem in a wide class of domains via a version of the Perron process.

The key to our study of linear equations is the Schauder-type estimates in Chapter 4. These estimates relate the norms of solutions of initial-boundary value problems for these equations to the norms of the known quantities in the problems. Specifically, the norms are parabolic Hölder norms, which can be considered as bounds on fractional derivatives of the functions. The main estimates were first proved by Barrar [12,13] and Friedman [64,65] using a representation formula for solutions of the inhomogeneous heat equation. Our proof is based on Campanato's approach [33] and does not use any representation formulae, but it does need the existence result of Chapter 3. The estimates of Chapter 4 are used in Chapter 5 to prove the existence, uniqueness, and regularity of solutions for various initial-boundary value problems under several different hypotheses on the regularity of the data.

Chapter 6 contains a further investigation of weak solutions. The class of equations involved is much larger than that considered in Chapter 3 and the definition of weak solution must be appropriately expanded. The first part of the chapter is concerned with existence, uniqueness, and regularity questions for weak solutions in suitable spaces. The second part covers the pointwise properties of weak solutions, in particular with the (Hölder) continuity of these solutions. These properties are proved via estimates of the solutions in terms of weak information on the data of the problem; such estimates are an important part of our study of nonlinear equations.

Our study of linear equations is completed in Chapter 7, which is devoted to strong solutions of parabolic equations. These are functions having weak derivatives in L^p for some $p > 1$ and satisfying the equation only almost everywhere. We obtain Schauder-type estimates in L^p and pointwise estimates for strong solutions which are analogous to the estimates in Chapter 6 for weak solutions. We also prove a boundary Hölder estimate for the normal derivative of a solution of the Cauchy-Dirichlet problem in terms of pointwise estimates of the coefficients of the equation. This estimate, originally proved by Krylov [137], is an important element in the nonlinear theory.

We begin our study of nonlinear equations in Chapter 8, which introduces two fixed point theorems. The first one, due to Schauder [233], is an infinite dimensional version of the Brouwer fixed point theorem in $\mathbb{R}^n$ and is useful for studying the Cauchy-Dirichlet problem for quasilinear equations. The second method, due to Lieberman [169, 173], is a variant of a fixed point theorem of Kirk and Caristi [125] related to the contraction mapping principle and is useful for studying nonlinear oblique derivative problems with quasilinear equations and fully nonlinear equations (with either Cauchy-Dirichlet data or oblique derivative data). Our point of view on existence is generally to prove local existence first (that is, existence for a small time interval) and then global existence. Such an approach is especially important in blow-up theory. Local existence is often proved using relatively simple *a priori* estimates, but these estimates are the key to the global existence theory. The Cauchy-Dirichlet problem and the oblique derivative problem are considered separately.

The *a priori* estimates for the Cauchy-Dirichlet problem fall naturally into four basic types:

(1) the maximum of the absolute value of the solution,
(2) the maximum of the length of the gradient on the parabolic boundary,
(3) the maximum of the length of the gradient in the domain,
(4) a Hölder norm for the gradient.

These estimates are the topics of Chapters 9, 10, 11, and 12, respectively, and each presupposes the preceding ones.

Chapter 9 applies the maximum principle of Chapter 2 and the maximum estimates of Chapter 6 to get pointwise bounds on the solution. In this chapter, we also generate some comparison principles which are important for other, later estimates.

Chapter 10 is devoted to a boundary gradient bound, which turns out to be the key estimate in the existence theory. We develop parabolic analogs of the Serrin curvature conditions for elliptic equations [236], which relate the geometry of the boundary to the structure of the differential equation. As in [236], we also show that these conditions are necessary to solve the problem with arbitrary data. All results are based on the comparison principle.

The gradient estimate is the topic of Chapter 11. The basis for proving such an estimate goes back to Bernstein [16]: show that the square of the length of the gradient of the solution satisfies a suitable differential inequality. Unlike Chapters 9 and 10, this chapter uses quite detailed structure conditions, and often a change of dependent variable is useful.

Chapter 12 applies the Hölder estimates of Chapters 6 and 7 to the gradient of the solution. Assuming that a bound is known for the gradient of the solution, we prove the Hölder estimates under very weak assumptions on the equation in question. The chapter finishes with some examples which illustrate the various structure conditions.

The oblique derivative problem is examined in Chapter 13. Since the techniques are so similar to those for the Cauchy-Dirichlet problem, many proofs in this chapter are sketched or omitted. An interesting feature of our study of the oblique derivative problem is that we are not restricted to the conormal problem. For example, with the heat equation $-u_t + \Delta u = f(X)$, we may use the capillarity boundary condition $(1 + |Du|^2)^{1/2} Du \cdot \gamma + \psi(X) = 0$, where Du denotes the gradient of u and γ is the interior spatial normal. If f, ψ and the domain are smooth enough, we shall show that this problem has a solution for any smooth enough initial data (satisfying a certain compatibility condition) provided ψ satisfies the necessary condition for a smooth solution: $|\psi| < 1$.

In Chapter 14, we consider a simple class of fully nonlinear equations, modeled on the parabolic Bellman equation:

$$-u_t + \inf_{\nu \in \mathcal{V}} \{a_\nu^{ij}(X) D_{ij} u\} = 0,$$

where (a_ν^{ij}) is a uniformly bounded, uniformly smooth family of uniformly positive definite matrices. Such equations were first studied systematically by Evans and Lenhart [57] and Krylov [136, 137]. The hypotheses of these authors are successfully weakened here by using simplifications due to Safonov [230, 231].

Finally, Chapter 15 looks at a class of nonuniformly parabolic fully nonlinear equations, modeled on the parabolic Monge-Ampère equation:

$$-u_t \det D^2 u = \psi(X, u, Du).$$

Elliptic versions of this equation and its generalizations were first studied by Caffarelli, Nirenberg, and Spruck [28, 29 ,30] and Ivochkina [97, 98, 99, 100, 101, 102]. Based on their work, we prove estimates and existence results for the parabolic equations. Our choice of parabolic Monge-Ampère equation was first identified by Krylov [134] as a suitable parabolc version of the equation $\det D^2u = \psi$, and the corresponding generalizations were recognized by Reye [229] as worthy of study. Other parabolic Monge-Ampère equations and their generalizations are discussed briefly. The most noteworthy is that of Ivochkina and Ladyzhenskaya: $-u_t + (\det D^2u)^{1/n} = \psi$. (See [103, 104].)

2. Further remarks

Although the notes in this book may seem quite extensive, we have really only scratched the surface of the subject of parabolic equations. An examination of any recent issue of Mathematical Reviews shows several dozen articles each month. Even restricting attention to single, nondegenerate equations (as we have here) leaves a staggering number of articles to read and assimilate. A complete bibliography would be much longer than this book and it would be obsolete before it appeared in print. Alternative sources of material which is directly relevant to the issues raised here are [147], [50], [139], [224].

There are also a number of subjects not covered in this book. Some of them are systems of parabolic equations and equations of higher order (see [147] and [4]), degenerate equations ([45]), and probabilistic aspects of parabolic equations ([139]).

In an effort to keep this book at a managable size, I have given a single point of view for most topics; for example, the Schauder estimates of Chapter 4 are proved using the Campanato method, which reappears in the nonlinear theory (although an alternative proof could be given using methods developed in Chapter 14). Many other proofs of these estimates are known; in fact, an entire book could be written just on the various means of proving them. Similarly, we have emphasized Moser iteration to prove most estimates for weak solutions; deGiorgi iteration is also an appropriate and frequently used technique, but it appears only briefly in Section 6.12. As a result of this single point of view, experts will certainly find a favorite technique missing. For example, I have avoided use of representation of solutions via Green's function and its cousin, potential representations. In addition, the method of viscosity solutions does not appear here.

3. Notation

In this book, we adopt the following conventions. First we use $X = (x,t)$ to denote a point in $\mathbb{R}^{n+1}$ with $n \geq 1$; x will always be a point in $\mathbb{R}^n$. We also write $Y = (y,s)$. Superscripts will be used to denote coordinates, so $x = (x^1, \ldots x^n)$.

Norms on $\mathbb{R}^n$ and $\mathbb{R}^{n+1}$ are given by

$$|x| = \left(\sum_{i=1}^{n}(x^i)^2\right)^{1/2} \quad \text{and} \quad |X| = \max\{|x|, |t|^{1/2}\},$$

respectively. A basis set of interest to us is the cylinder

$$Q(X_0, R) = Q(R) = \{X \in \mathbb{R}^{n+1} : |X - X_0| < R, t < t_0\}.$$

We also use the ball

$$B(x_0, R) = \{x \in \mathbb{R}^n : |x - x_0| < R\}.$$

We generally follow tensor notation. As previously indicated, superscripts denote coordinates of points in $\mathbb{R}^n$. Also, subscripts denote differentiation with respect to x. In particular,

$$D_i u = \frac{\partial u}{\partial x^i}, D_{ij} u = \frac{\partial^2 u}{\partial x^i \partial x^j}$$

for u a sufficiently smooth function and i and j integers between 1 and n, inclusive. We also write Du for the vector $(D_1 u, \ldots, D_n u)$ and $D^2 u$ for the matrix $(D_{ij} u)$. On the other hand, we write u_t (or occasionally $D_t u$) for $\partial u/\partial t$. In addition, we follow the summation convention that any term with a repeated index i is summed over $i = 1$ to n. For example,

$$b^i D_i u = \sum_{i=1}^{n} b^i D_i u.$$

To be formally correct, we should demand (as was the case in this example) that one occurrence of the index be a superscript and the other be a subscript, but we shall abuse this aspect of the summation convention when convenient.

A word of warning about superscripts and subscripts is also in order. Both will be used as generic indices and to indicate differentiation with respect to other variables. In addition, superscripts will also be used to indicate exponentiation. It should be clear from the context which meaning is intended.

We use Ω to denote a domain in $\mathbb{R}^{n+1}$, that is, Ω is an open connected subset of $\mathbb{R}^{n+1}$, so for any $X_0 \in \Omega$, there is a positive number R such that the space-time ball $\{X : |x - x_0|^2 + (t - t_0)^2 < R\}$ is a subset of Ω. For a fixed number t_0, we write $\omega(t_0)$ for the set of all points (x, t_0) in Ω, and we define $\gamma(X_0)$ to be the unit inner normal to $\omega(t_0)$ at X_0 provided $\omega(t_0)$ is not empty. We also write $I(\Omega)$ for the set of all t such that $\omega(t)$ is nonempty. Since Ω is connected, $I(\Omega)$ will be an open interval. As usual, $\partial\Omega$ denotes the topological boundary of Ω. The parabolic boundary $\mathcal{P}\Omega$ will be defined in Chapter 2.

We also use $|\Omega|$ to denote the Lebesgue measure of the set Ω, and we define $\operatorname{diam}\Omega = \sup\{|x - y| : X, Y \text{ in } \Omega\}$, so $\operatorname{diam}\Omega$ is not the usual diameter of a set. When Ω is a cylinder $\Omega = \omega \times (a, b)$, then $\operatorname{diam}\Omega$ is the usual diameter of the cross-section ω, but if Ω is noncylindrical, $\operatorname{diam}\Omega$ may be strictly larger than the diameter of any cross-section $\omega(t)$.

Points in $\mathbb{R}^n$ and $\mathbb{R}^{n+1}$ are given [illegible]

$$B(y,R) = \{x \in \mathbb{R}^n : |x-y| < R\} \quad \text{and} \quad Q(Y,R) = \{X \in \mathbb{R}^{n+1} : |X-Y| < R,\ t < s\}$$

respectively. A particular set of interest to us is the cylinder

$$Q(X_0,R) = [illegible]$$

We also use the ball

$$E(X_0,R) = \{X : |x - x_0| < R,\ [illegible]\}.$$

We generally follow tensor notation. As previously indicated, superscripts denote coordinates of points in $\mathbb{R}^n$. Also, subscripts denote differentiation with respect to coordinates. In particular,

$$D_i u = \frac{\partial u}{\partial x^i}, \quad D_{ij} u = \frac{\partial^2 u}{\partial x^i \partial x^j}$$

for u a sufficiently smooth function and i and j integers from 1 to n, inclusive. We also write Du for the vector $(D_1 u, \dots, D_n u)$ and $D^2 u$ for the matrix $(D_{ij} u)$. On the other hand, we write u_t (or, occasionally $D_t u$) for $\partial u/\partial t$. In addition, we follow the summation convention that any term with a repeated index is summed over the index. For example,

$$a^{ij} D_{ij} u = \sum_{i,j=1}^{n} a^{ij} D_{ij} u.$$

To be formally correct, we should distinguish [illegible] one occurrence of the index be a superscript and the other be a subscript, but we shall ignore this aspect of the summation convention when convenient.

A [illegible] superscript and subscript [illegible]. Both will be used as generic indices and to indicate differentiation with respect to other variables. In addition, [illegible] will also be used to indicate [illegible]. It should be clear from the context which meaning is intended.

We use Ω to denote a domain in $\mathbb{R}^{n+1}$, that is, Ω is an open connected subset of $\mathbb{R}^{n+1}$, so for any $X \in \Omega$, there is a positive number R such that [illegible] is a subset of Ω. For a fixed number t_0, we write $\Omega(t_0)$ for the set of all points (x, t_0) in Ω, and we define $I(\Omega)$ to be the set of all t such that $\Omega(t)$ is nonempty. Since Ω is connected, $I(\Omega)$ will be an open interval. As usual, $\partial\Omega$ denotes the topological boundary of Ω. The parabolic boundary $\mathcal{P}\Omega$ will be defined in Chapter 2.

We also use $|\Omega|$ to denote the Lebesgue measure of the set Ω, and we define $\operatorname{diam} \Omega = \sup\{|X - Y| : X, Y \in \Omega\}$, so $\operatorname{diam} \Omega$ is not the usual diameter of a set. When Ω is a cylinder [illegible], $\operatorname{diam} \Omega$ is the usual diameter of [illegible], but if Ω is not cylindrical, $\operatorname{diam} \Omega$ can be strictly larger than the diameter of any cross-section.

CHAPTER II

MAXIMUM PRINCIPLES

Introduction

An important tool in the theory of second order parabolic equations is the maximum principle, which asserts that the maximum of a solution to a homogeneous linear parabolic equation in a domain must occur on the boundary of that domain. In fact, this maximum must occur on a special subset of the boundary, called the parabolic boundary. The strong maximum principle asserts that the solution is constant (at least in a suitable subdomain) if the maximum occurs anywhere other than on the parabolic boundary. The maximum principle is used to prove uniqueness results for various boundary value problems, L^∞ bounds for solutions and their derivatives, and various continuity estimates as well.

1. The weak maximum principle

As noted in Chapter 1, we consider linear operators L defined by

$$Lu = a^{ij}(X)D_{ij}u + b^i(X)D_iu + c(X)u - u_t \tag{2.1}$$

in an $(n+1)$-dimensional domain Ω. We assume that L is weakly parabolic. In other words,

$$a^{ij}(X)\xi_i\xi_j \geq 0 \text{ for all } \xi \in \mathbb{R}^n \text{ and all } X \in \Omega. \tag{2.2}$$

We also write $\mathcal{T}$ for $\sum a^{ii}$, the trace of the matrix (a^{ij}).

For a domain $\Omega \subset \mathbb{R}^{n+1}$, we define the parabolic boundary $\mathcal{P}\Omega$ to be the set of all points $X_0 \in \partial\Omega$ such that for any $\varepsilon > 0$, the cylinder $Q(X_0, \varepsilon)$ contains points not in Ω. In the special case that $\Omega = D \times (0,T)$ for some $D \subset \mathbb{R}^n$ and $T > 0$, $\mathcal{P}\Omega$ is the union of the sets $B\Omega = D \times \{0\}$ (which is the *bottom* of Ω), $S\Omega = \partial D \times (0,T)$ (which is the *side* of Ω), and $C\Omega = \partial D \times \{0\}$ (which is the *corner* of Ω). These sets (and their analogs for more general domains) will play an important role in the theory of initial-boundary value problems.

The simplest version of the maximum principle is the following.

Lemma 2.1. *If $u \in C^{2,1}(\overline{\Omega} \setminus \mathcal{P}\Omega) \cap C^0(\overline{\Omega})$, if $Lu > 0$ in $\overline{\Omega} \setminus \mathcal{P}\Omega$ and if $u < 0$ on $\mathcal{P}\Omega$, then $u < 0$ in $\overline{\Omega}$.*

Proof. Suppose, to the contrary, that $u \geq 0$ somewhere in $\overline{\Omega}$. Let $t^* = \inf\{t : u(x,t) \geq 0 \text{ for some } X \in \overline{\Omega}\}$. By continuity and the assumption that $u < 0$ on