Preface

The main sources of nonlinear second-order elliptic and parabolic equations are the theory of differential equations and its numerous applications. At present, the class of quasilinear equations, which are linear with respect to the second-order derivatives of the unknown function, is studied almost as thoroughly as the class of linear equations. The results of these studies are summarized in many articles and books; see, for instance, Ladyzhenskaya and Ural'tseva [**86**], Ladyženskaja, Solonnikov, and Ural'ceva [**87**], Ivanov [**49**], Gilbarg and Trudinger [**45**], and Lieberman [**89**].

For the fully nonlinear equations, the main subject of this book, only sparse results were available for a long time, not allowing one to build any general theory. However, the equations on the plane were an exception. There, using some specific features available only for two-dimensional space, one was able to develop the solvability theory in $W^{1,2}_p$ and W^2_p for any p close to 2 (see, for example, [**53**]).

With no general theory at hand, some classes of fully nonlinear equations were investigated in the areas where they arose by using available methods in these areas. The first class of such equations is related to the famous Monge-Ampère equation which arose in geometry and was studied for quite some time by using the theory of convex surfaces. Aleksandrov [**1**] in 1958 introduced the notion of generalized solutions of Monge-Ampère equations. Its smoothness, before 1971, was proved only in dimension two (see Bakel'man [**7**]). Then, Pogorelov [**98, 99**] proved *interior* smoothness for any dimension (see also Cheng and Yau [**19**], who pointed out a vicious circle in some arguments of Pogorelov). In the multidimensional case, the smoothness up to the boundary was proved only after an approach to *general* nonlinear equations was developed in the early 1980s.

Another class of nonlinear elliptic and parabolic equations, so-called Bellman's equations, arose from the probabilistic theory of controlled stochastic diffusion processes. When the diffusion is not controlled, Bellman's equations become quasilinear (see Fleming and Rishel [**43**] and the reference therein). It turned out that by using probabilistic methods, it was possible to prove the solvability of general degenerate nonlinear elliptic Bellman's equations in the whole space in the class of functions with bounded second-order spacial derivatives. Some of these results are presented in [**57**]. Probably the latest results about general possibly degenerate Bellman's equations in smooth domains are presented in [**58**] and [**63**]. In [**63**], in particular, the solvability of possibly degenerate complex Monge-Ampère equations in strictly pseudo-convex domains is established (this result was earlier obtained by PDE methods in [**17**] for *nondegenerate* case). Part of [**63**] is translated into the pure PDE language in [**64**].

These probabilistic results enabled us to deal with a very large class of fully nonlinear equations. The technique was based to a large degree on the fact that the

Monge-Ampère equation is a particular case of Bellman's equations. However, this theory was very hard to comprehend for people from pure PDEs: it was all presented in probabilistic terms. Up until 1979 no other approaches were found. The probabilistic approach was developed by the author, Nisio, Pragarauskas, Safonov, Lions, Menaldi, and many others.

In 1979 Brézis and Evans [**11**] proved the existence of a $C^{2+\alpha}$-solution for Bellman's equation with two elliptic operators. The probabilistic approach only provided solutions with two *bounded* derivatives. The results of Evans and Friedman [**36**], Evans and Lions [**38**], and Lions [**91**] published in about 1979 were obtained by the PDE methods, but the solutions there had only bounded derivatives, and the only novelty in comparison with earlier results was that purely analytical methods were used.

A real breakthrough came in 1982 when Evans [**34**] and the author [**59**] independently proved the solvability in $C^{2+\alpha}$ of a wide class of fully nonlinear equations. In [**34**] elliptic equations are considered like

$$H\big(u(x), Du(x), D^2u(x), x\big) = 0, \tag{1}$$

and in [**59**] both elliptic equations like (1) and parabolic ones like

$$\partial_t u(t,x) + H\big(u(t,x), Du(t,x), D^2u(t,x), t, x\big) = 0 \tag{2}$$

are investigated. The proofs were presented in pure PDE terms. The starting point for these proofs was the fact that the solutions of *linear* equations with measurable coefficients are Hölder continuous. This result was obtained by Safonov and the author in [**83**] and [**84**]. We present the results of [**84**] in this book in almost the same way as it is done in [**61**]; however, we go slightly deeper, which enables us to obtain Fang-Hua Lin's type estimates for elliptic and parabolic equations. These estimates play a crucial role in our approach to the Sobolev space theory.

The results in [**84**] opened the way to build a solvability theory in classes $C^{2+\alpha}$ of equations like (1) and (2) with H which is *convex or concave* with respect to D^2u and is in C^2 with respect to x or in $C^{1,2}$ with respect to (t,x). We refer the reader to [**45**] and [**61**] and the references therein for information on what was achieved in this area.

The next breakthrough came when Safonov ([**102**], [**103**], and [**104**]) came up with a technique allowing one to treat the H's that are only Hölder continuous with respect to the independent variables, which made the whole $C^{2+\alpha}$-theory of fully nonlinear equations as satisfactory as in the case of linear equations. Safonov's approach is also remarkable in the respect that it works equally well for linear and fully nonlinear equations. In our exposition here we also use these techniques.

Then, naturally arose the interest in equations when H is neither convex nor concave with respect to D^2u. These equations often appear as the Isaacs equations in the theory of stochastic differential games. By now we know that generally one can only hope that their solutions are in $C^{1+\alpha}$ for an $\alpha \in (0,1)$ (see [**96**]). In these circumstances the notion of the so-called viscosity solutions introduced by Lions [**92**] turned out to be quite effective. This notion does not require the solutions to have any derivatives. Viscosity solutions later turned out to exist and be unique in a wide variety of cases. At present the theory of viscosity solutions is quite a popular subject. The reader will find three chapters in this book devoted to this theory.

The next dramatic turn of the theory came from groundbreaking results by Caffarelli [**12**] in 1989 (see also [**14**]) about a priori estimates in Sobolev W^2_p classes. His discovery is comparable with the discoveries in [**34**] and [**59**], which were the starting points of the $C^{2+\alpha}$-theory. He works in the framework of the theory of viscosity solutions and proves that under some conditions any viscosity solution is in $W^2_{p,\text{loc}}$. The reader is referred to [**23**], [**51**], [**52**], and [**25**] for more detailed information about the results obtained by many authors in this theory. We also provide some additional information later in the text while comparing our results with the ones obtained by using the theory of viscosity solutions. The major difference is well seen when the equations are just linear. Then the main coefficients in the theory of viscosity solutions are assumed to be uniformly close to uniformly continuous functions (usual requirement in the classical L_p-theory for linear equations), and in our theory it quite suffices for them to only be in VMO or almost in VMO; that is, to just have sufficiently small integral oscillations over small balls (elliptic case) or cylinders (parabolic case).

The main object of studies in this book is the first boundary-value problem for general fully nonlinear second-order uniformly elliptic and parabolic equations like (1) or (2). Apart from the simplest degenerate Bellman's equations in the whole space with constant coefficients, like equations of Monge-Ampère type, which admit solutions with bounded second-order derivatives, we do not say much about any *specific* class of equation degenerate or not, although we provide quite a few examples of equations to illustrate our assumptions and results. These are given in 1.2.9, 4.1.20, 4.1.21, 6.6.11, 6.6.12, 10.1.1, 10.1.2, 10.1.5, 10.1.18 10.1.24, 10.1.25, 12.1.1, 12.1.14, 12.1.20, 16.2.10, 17.1.6, 18.1.4, C.1.3, C.1.4. Some examples are also given at the end of Section 1.2.

We are looking for solutions in Sobolev classes or for C-viscosity or L_p-viscosity solutions and present *basic solvability results* about such *general* equations. In particular, we do not discuss the remarkable results of Caffarelli [**13**] about Sobolev solutions of the Monge-Ampère equations continued by many researchers (see [**28**] and the references therein).

Most of the auxiliary results are taken from old sources ([**61**], [**103**], and [**104**]), and the core of the main new results was obtained in the last few years. Exposition of these results is based on a generalization of the Fefferman-Stein theorem, Fang-Hua Lin's type estimates, and the so-called "ersatz" existence theorems, saying that one can slightly modify "any" equation and get a "cut-off" equation that will have solutions with bounded derivatives. Here is an example of such "ersatz" existence theorem (cf. Theorem 6.6.3).

1. Assumption. We are given a domain $\Omega \subset \mathbb{R}^d$ which is bounded and which satisfies the exterior ball condition (see Definition 5.5.3). We are also given a function $g \in C^{1,1}(\mathbb{R}^d)$.

Let $\mathbb{S}$ be the set of symmetric $d \times d$ matrices. Suppose that we are given a Borel measurable function $H(\mathsf{u}, x)$,

$$\mathsf{u} = (\mathsf{u}', \mathsf{u}''), \quad \mathsf{u}' = (\mathsf{u}'_0, \mathsf{u}'_1, ..., \mathsf{u}'_d) \in \mathbb{R}^{d+1}, \quad \mathsf{u}'' \in \mathbb{S}, \quad x \in \mathbb{R}^d.$$

2. Assumption. (i) The function $H(\mathsf{u}, x)$ is Lipschitz continuous with respect to u'', and at all points of differentiability of H with respect to u'' we have $D_{\mathsf{u}''}H \in \mathbb{S}_\delta$ (a bounded set of strictly positive symmetric $d \times d$ matrices),

(ii) For a constant $K_0 \in [0,\infty)$, the number

$$\bar{H} := \sup_{\mathsf{u}',x} \big(\big|H(\mathsf{u}',0,x)\big| - K_0|\mathsf{u}'|\big)$$

is finite,

(iii) The function $H(\mathsf{u},x)$ is continuous with respect to u for any x,

(iv) The function $H(\mathsf{u},x)$ is nonincreasing with respect to u_0'.

3. THEOREM. *Under the above assumptions, there exists a function $P(\mathsf{u}'')$ which is positive homogeneous of order one and convex with respect to u'' such that for any $K>0$ the equation*

$$\max\Big[H\big(u(x),Du(x),D^2u(x),x\big),P\big(D^2u(x)\big)-K\Big]=0$$

(a.e.) in Ω with boundary condition $u=g$ on $\partial\Omega$ has a solution which is continuous in $\bar{\Omega}$, has bounded gradient, and whose second-order derivatives are locally bounded in Ω.

Theorems of that kind allow us to prove the solvability in Sobolev classes for equations that are quite far from the ones that are convex or concave with respect to the Hessian of the unknown functions. They play the role of starting points, which in the $C^{2+\alpha}$-theory play linear equations allowing one by using the method of continuity to go from linear to fully nonlinear equations. We will see how this works in Chapter 13. The method of continuity consists of moving step-by-step along the parameter connecting a linear equation to the one of interest. It does not work in Sobolev space setting, and the results like Theorem 3 allow one not to use any steps but just send K to infinity and having a priori estimates arrive at solutions of (1).

In the framework of viscosity solutions these approximation theorems also allow us to deal with classical approximating solutions, thus avoiding sometimes very involved constructions from the usual theory of viscosity solutions.

Comments on the book's structure

One of fundamental results of the book is the above Theorem 3 and its parabolic counterpart, Theorem 8.3.3. Their particular cases, when H depends not on the full Hessian D^2u but on finitely many pure second-order derivatives in fixed directions, are given as Theorems 6.1.5 and 8.1.4, respectively. For understanding the proofs of these particular cases the reader need not know *anything* from the theory of linear or nonlinear partial differential equations. Everything is based on rather simple arithmetical manipulations with finite-differences presented in Chapters 5 and 7. Therefore, there was a strong temptation to start the book from scratch with Chapter 5 and in some 35 pages to arrive at a solvability result like Theorem 3 for particular operators H.

Actually, it could even be a good idea for the reader to start with Chapter 5 in order to see why the previous material is needed if one wants to go to more general H.

After we prove the solvability of "cut-off" equations, we send K to infinity, but first we need to obtain some estimates of solutions uniform with respect to K. First, in Chapter 9, come C^α estimates of Krylov-Safonov and Fang-Hua Lin's type estimates of the integrals of small powers of the magnitude of the second-order derivatives of solutions. The Krylov-Safonov estimates and the fact that we know the solvability of "cut-off" equations allow us in Chapter 10 to prove Evans-Krylov $C^{2+\alpha}$-estimates for model equations like $F(D^2u) = 0$ and further, by combining them with Fang-Hua Lin's type estimates from Section 9.4 and a generalization of the Fefferman-Stein theorem from Appendix C for not necessarily even locally summable functions, to establish in Theorem 10.1.14 the solvability in $W^2_{p,\text{loc}}$ for elliptic equations under rather weak conditions without assuming that H is convex or concave with respect to D^2u. This line of arguments:

Evans-Krylov $C^{2+\alpha}$-estimates, Fang-Hua Lin's
type estimates, Fefferman-Stein theorem

is repeated several times like in rondo in different circumstances, each time requiring some new initial information.

For instance, Chapter 11, "Nonlinear elliptic equations in $C^{2+\alpha}_{\text{loc}}(\Omega) \cap C(\bar{\Omega})$", appeared, in particular, because of the need to obtain Evans-Krylov $C^{2+\alpha}$-estimates for model parabolic equations like $\partial_t u + F(D^2u) = 0$. There were two possibilities: either to reprove Evans-Krylov $C^{2+\alpha}$-estimates for parabolic equations (following what was done in [**61**]), or to rewrite the equation as $F(D^2u) = -\partial_t u$, first estimate the C^α-norm of $\partial_t u$, and then consider the equation for any fixed t as an elliptic one. We preferred the second choice because it is more elegant and richer in ideas, but then we needed a theory of solvability of equations $F(D^2u) = f$ in $C^{2+\alpha}$-spaces, which is insightful in its own rights.

After that in Chapter 12 we are able to basically repeat the reasoning in Chapter 10 and develop the solvability theory for *parabolic* equations in $W^{1,2}_{p,\text{loc}}$.

Then comes the need for solvability theory in global Sobolev spaces W^2_p for elliptic and $W^{1,2}_p$ for parabolic equations. Everything here depends on a priori estimates near the boundary, so model equations should be considered in Hölder classes of functions in half balls or half cylinders, and this gave rise to a comparatively long Chapter 13, "Elements of the $C^{2+\alpha}$-theory of fully nonlinear elliptic and parabolic equations".

Once the boundary a priori estimates are obtained for model equations, again mimicking the line of arguments in Chapter 10, we can then prove the solvability results in W^2_p spaces in Chapter 14 and in $W^{1,2}_p$ spaces in Chapter 15.

The remaining chapters contain applications of the parabolic counterpart of Theorem 3 to various issues in the theory of viscosity solutions for parabolic equations. It would be interesting to know if all of them, especially Theorem 17.1.5 on the $C^{1+\alpha}$-regularity of the L_p-viscosity solutions for Isaacs equations with almost VMO leading coefficients, can be proved by means of the theory of viscosity solution.

There are no historical notes or comments after each chapter. These are spread throughout the text and only a few times presented in a more condensed form in Remarks 3.0.3, 3.1.7, 3.3.8, 4.2.10, 6.6.22, 9.1.3, 9.1.7, 9.4.8, 9.6.9, 10.1.23, 10.1.26, 10.2.5, 11.2.5, 12.1.12, 12.1.21, 14.1.9, 15.1.6, 16.2.9, 17.1.7, 18.1.3, 18.3.6, and 18.4.3; and Examples 10.1.24, 16.2.10, and 18.1.4.

Comments on the exposition

The theorems, lemmas, remarks, and such which are part of the main body of the text are numbered serially in a single system that proceeds by section. Theorem 1.7.6 is the sixth numbered unit in the seventh section in the first chapter. In the course of Chapter 1, this theorem is referred to as Theorem 7.6, and in the course of the seventh section of Chapter 1, it is referred to as Theorem 6. Similarly and independently of these units formulas are numbered and cross-referenced.

In the proofs of various results in this book we use the symbol N sometimes with indices to denote finite nonnegative constants which may change from one occurrence to another, and we do not always specify on which data these constants depend. In these cases the reader should remember that, if in the statement of a result there are constants called N which are claimed to depend only on certain parameters, then in the proof of the result the constants N also depend only on the same parameters unless specifically stated otherwise. Of course, if we write

$$N = N(...),$$

this means that N depends only on what is inside the parentheses.

A complete reference list of notation can be found in the index at the end of the book.

In the case of elliptic equations we concentrate only on equations in the Euclidean space

$$\mathbb{R}^d = \{x = (x^1, ..., x^d) : x^i \in (-\infty, \infty)\}$$

or in domains $\Omega \subset \mathbb{R}^d$. By domains we mean general open sets. If Γ is a measurable subset of $\mathbb{R}^d$, by $|\Gamma|$ we denote its Lebesgue measure.

The scalar product in $\mathbb{R}^d$ is denoted by

$$\langle \cdot, \cdot \rangle.$$

If a is a matrix of any dimension (for instance, column-vector), we set

$$|a| = [\operatorname{tr} aa^*]^{1/2},$$

where a^* is the transpose of a.

On some occasions, we allow ourselves to use different symbols for the same objects, for example,

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

$$u_x = Du = (D_1 u, ..., D_d u), \quad u_{xx} = D^2 u = (D_{ij} u)_{i,j=1}^d.$$

Somewhat ambiguously we set

$$D_i uv := (D_i u) v.$$

In the case of parabolic equations we work with

$$\mathbb{R}^{d+1} = \{(t, x) : t \in \mathbb{R}, x \in \mathbb{R}^d\}.$$

If Γ is a measurable subset of $\mathbb{R}^{d+1}$, by $|\Gamma|$ we denote its Lebesgue measure. If $z_1 = (t_1, x_1), z_2 = (t_2, x_2) \in \mathbb{R}^{d+1}$, we set

$$\rho(z_1, z_2) = |x_1 - x_2| + |t_1 - t_2|^{1/2}$$

and call $\rho(z_1, z_2)$ the parabolic distance between z_1 and z_2.

For functions $u(t, x)$ given on subdomains of $\mathbb{R}^{d+1}$ we use the above notation only for the derivatives in x; sometimes we write $D_x u$, $D_x^2 u$ to emphasize this, and denote

$$\partial_t u = \frac{\partial u}{\partial t} = u_t, \quad \partial_t D_k u = u_{tx^k} = u_{x^k t},$$

and so on.

In various parts of the book we speak about measurable functions and sets: Borel measurable and Lebesgue measurable, and if we say that a particular function or a set is *measurable*, we always mean that it is Lebesgue measurable. On many occasions we consider functions like $F(\mathsf{u}, x)$, where u belongs to a Euclidean space, say $\mathfrak{U}$, and $x \in \mathbb{R}^d$, and require them to be Borel measurable in (u, x). This is done for simplicity of presentation. Actually, what is needed is for them to be measurable with respect to the product of the Borel σ-field in $\mathfrak{U}$ and the Lebesgue σ-field in $\mathbb{R}^d$. Generally, it is useful to keep in mind that for any Lebesgue measurable function there exists a Borel measurable one coinciding with it almost everywhere.

If Ω is a domain in $\mathbb{R}^d$ and $p \in [1, \infty]$, by $L_p(\Omega)$ we mean the set of all (Lebesgue) measurable real-valued functions f for which

$$\|f\|_{L_p(\Omega)} := \Big(\int_\Omega \big|f(x)\big|^p \, dx \Big)^{1/p} < \infty$$

with the standard extension of this formula if $p = \infty$. One knows that $L_p(\Omega)$ is a Banach space.

By $C_0^\infty(\Omega)$ we mean the set of all infinitely differentiable functions on Ω with compact support (contained) in Ω. By the *support* of a function we mean the closure of the set where the function is different from zero. We call a subset of $\mathbb{R}^d$ *compact* if it is closed and bounded. Of course, saying "compact support" is the same as saying "bounded support"; we keep "compact" just to remind us that we are talking about *closed* sets. We set

$$C_0^\infty = C_0^\infty(\mathbb{R}^d).$$

Similar notation is introduced for domains in $\mathbb{R}^{d+1}$ and for functions of (t, x) rather than x alone.

We always use the summation convention with respect to repeated indices no matter whether they are at the same level or at different ones, and if there are signs of inf or sup followed by the sums of products of terms with repeated indices, then the summation is performed first, inside the signs of inf or sup, for instance,

$$\inf_\alpha \big[a_i^\alpha u_i + b_j^\alpha v_j\big] := \inf_\alpha \Big[\sum_i a_i^\alpha u_i + \sum_j b_j^\alpha v_j\Big].$$

By $\mathbb{S}$ we mean the set of $d \times d$ symmetric matrices. If $a \in \mathbb{S}$ we usually represent its entries by a^{ij}. However, on many occasions a itself has superscripts, and then we use a_{ij} for the entries of a. In the same vein for numbers u we use the notation

$$u^\pm = u_\pm = (1/2)\big(|u| \pm u\big).$$

We remind the reader that

$$a \wedge b = \min(a, b), \quad a \vee b = \max(a, b).$$

One more general stipulation is that

positive=nonnegative, negative=nonpositive,

decreasing=nonincreasing, increasing=nondecreasing.

Acknowledgments

My deepest gratitude is due to Sungbin Lim and Hyunwoo Kwon, who on their own initiative typeset in LaTeX my book [**61**] and sent me the TeX file, allowing me to use it if I ever go back to the subject of [**61**]. I used their file in a few places. I am also grateful to the NSF for supporting me in the early stages of writing the book through grant DMS-1160569. Finally, the referees made some comments, which certainly helped improve the presentation.

Nicolai Krylov
Minneapolis, MN
August 2017