1

Introduction

The Monge–Ampère equation is a fully nonlinear, degenerate elliptic equation that draws its name from its initial formulation in two dimensions, by the French mathematicians Monge [92] and Ampère [8], about two hundred years ago. The classical form of this equation is

$$\det D^2 u = f(x, u, \nabla u) \quad \text{in } \Omega, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is an open set, $u : \Omega \to \mathbb{R}$ is a convex function, and $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^+$ is given. As we shall explain below, the convexity of u is a necessary condition to make the equation elliptic and to hope for regularity results.

The prototypical place where the Monge–Ampère equation appears is the "prescribed Gaussian curvature equation", also known as the "Minkowski problem": if we take $f = K(x)(1 + |\nabla u|^2)^{(n+2)/2}$, then (1.1) corresponds to imposing that the Gaussian curvature of the graph of u at the point $(x, u(x))$ is equal to $K(x)$ (see Section 2.6).

Other classical appearances of the Monge–Ampère equation can be found in affine geometry (for instance, in the "affine sphere problem" and the "affine maximal surfaces" problem; see [25, 98, 27, 115, 116, 117] and the references in [118]) and in convex geometry (see, for instance, Section 4.4). More recently, the Monge–Ampère equation has found important applications in optimal transportation (see Section 4.6) and in meteorology (see Section 4.9). The goal of this book is to develop the existence, uniqueness, and regularity theory for (1.1), and to show how this equation appears in the above-mentioned problems.

1.1 On the degeneracy of the Monge–Ampère equation

Before entering into the theory of Monge–Ampère, we wish to discuss the terms "fully nonlinear" and "degenerate elliptic" that we have used above. Also, we want to explain why we are considering this equation only on convex functions.

1.1.1 A classification of second-order elliptic PDEs. The model second-order elliptic PDE is the Laplace equation

$$\Delta u = 0,$$

or, more generally, the Poisson equation

$$\Delta u = f,$$

where $f : \Omega \to \mathbb{R}$ is some given function. These equations are called *linear* as they depend linearly on the unknown function u.

Given a family of coefficients $a_{ij} : \Omega \to \mathbb{R}$, $b_i : \Omega \to \mathbb{R}$, $c : \Omega \to \mathbb{R}$, $i, j = 1, \dots, n$, the linear equation

$$\sum_{ij} a_{ij}\partial_{ij}u + \sum_i b_i\partial_i u + cu = f$$

is called *uniformly elliptic* provided the coefficients a_{ij} define positive-definite bounded matrices, that is,

$$\lambda|\xi|^2 \le \sum_{ij} a_{ij}(x)\xi^i\xi^j \le \Lambda|\xi|^2 \quad \forall\, \xi = (\xi^1, \dots, \xi^n) \in \mathbb{R}^n, \quad x \in \Omega, \tag{1.2}$$

for some constants $0 < \lambda \le \Lambda < \infty$, and *degenerate elliptic* if λ can be equal to zero or Λ can be equal to infinity. When (1.2) holds, we shall also write $\lambda\,\mathrm{Id} \le a_{ij} \le \Lambda\,\mathrm{Id}$.

When PDEs are nonlinear in the unknown u, one can classify them depending on the kind of nonlinear structure.

More precisely, if the leading-order term (the term involving the second derivatives of u) is linear in u, one says that the equation is *semilinear*; the model example is

$$\Delta u = f(x, u, \nabla u)$$

for some given function $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$.

Instead, if the leading-order term is linear in the Hessian of u but depends nonlinearly on u via lower-order terms, the equation is called *quasilinear*: elliptic equations of this form are

$$\sum_{ij} a_{ij}(x, u, \nabla u)\partial_{ij}u = f(x, u, \nabla u),$$

where the coefficients $a_{ij} = a_{ij}(x, z, p)$ satisfy (1.2) for all $(x, z, p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n$. A classical example is the p-Laplace equation.

Finally, a PDE is *fully nonlinear* if it is nonlinear in the Hessian of u. Model elliptic examples are the Bellman equation

$$0 = \sup_{\alpha \in A}\left\{\sum_{ij} a^\alpha_{ij}\partial_{ij}u + \sum_i b^\alpha_i\partial_i u + c^\alpha u - f^\alpha\right\}$$

and the Isaacs' equation

$$0 = \inf_{\beta \in B} \sup_{\alpha \in A} \left\{ \sum_{ij} a_{ij}^{\alpha,\beta} \partial_{ij} u + \sum_{i} b_i^{\alpha,\beta} \partial_i u + c^{\alpha,\beta} u - f^{\alpha,\beta} \right\},$$

where the coefficients $a_{ij}^{\alpha}, a_{ij}^{\alpha,\beta}$ satisfy (1.2) with constants independent of α and β.

Since the Monge–Ampère equation depends nonlinearly on the Hessian of u, it falls into the last category.

1.1.2 Degeneracy and convexity. To understand the degenerate elliptic structure of Monge–Ampère, we consider $u : \Omega \to \mathbb{R}$ a smooth solution of (1.1) with $f = f(x) > 0$ smooth.

A standard technique to deal with nonlinear equations is to differentiate the equation solved by u to obtain a linear second-order equation for its derivatives. More precisely, fix a direction $e \in \mathbb{S}^{n-1}$ and differentiate (1.1) in the direction e. To simplify notation, we shall use subscripts to denote partial derivatives, that is, $u_e = \partial_e u$, $f_e = \partial_e f$, $u_{ij} = \partial_{ij} u$, etc.

Since $D^2u(x + \varepsilon e) = D^2u(x) + \varepsilon D^2u_e(x) + o(\varepsilon)$, we see that

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \det\big(D^2u(x + \varepsilon e)\big) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \det\big(D^2u(x) + \varepsilon D^2u_e(x)\big).$$

Then using (A.1), we get

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \det\big(D^2u(x + \varepsilon e)\big) = \det\big(D^2u(x)\big)\, \mathrm{tr}\big((D^2u(x))^{-1} D^2u_e(x)\big).$$

Hence, if we use u^{ij} to denote the inverse matrix of $u_{ij} = \partial_{ij}u$ and we use Einstein's convention of summing over repeated indices ($a_{ij}b_{ij} = \sum_{ij} a_{ij}b_{ij}$), we deduce that

$$(\det D^2u)u^{ij}\partial_{ij}u_e = f_e \quad \text{in } \Omega.$$

Recalling that $\det D^2u = f > 0$, we can rewrite the above equation as

$$u^{ij}\partial_{ij}u_e = \frac{f_e}{f} \quad \text{in } \Omega. \tag{1.3}$$

Thus, setting

$$a_{ij} := u^{ij}, \qquad v := u_e, \qquad g := \frac{f_e}{f},$$

we see that v solves the linear equation

$$a_{ij}\partial_{ij}v = g.$$

Now, if we want this equation to be uniformly elliptic, we need $a_{ij} = u^{ij}$ to be positive definite as in (1.2), which can be written in terms of its inverse $u_{ij} = \partial_{ij} u$ as follows:

$$\frac{1}{\Lambda}|\xi|^2 \le D^2 u(x)[\xi, \xi] \le \frac{1}{\lambda}|\xi|^2 \quad \forall\, \xi = (\xi^1, \dots, \xi^n) \in \mathbb{R}^n, \quad x \in \Omega.$$

So, u must be uniformly convex and $C^{1,1}$. In particular, in order for the coefficients a_{ij} to be at least nonnegative definite, we are forced to restrict our attention to convex functions. However, since a_{ij} may vanish or be unbounded at some points, the equation is *degenerate elliptic*.

Notice that if

$$\frac{1}{C}\,\mathrm{Id} \le D^2 u \le C\,\mathrm{Id} \quad \text{inside } \Omega \tag{1.4}$$

for some constant $C > 0$, then $\mathrm{Id}/C \le u^{ij} \le C\,\mathrm{Id}$ and the linearized equation (1.3) becomes uniformly elliptic. For this reason, proving the bound $\mathrm{Id}/C \le D^2 u \le C\,\mathrm{Id}$ is key for the regularity of solutions to (1.1).

We shall use a crucial observation in the sequel:

Remark 1.1. Let u solve (1.1) with $f \ge a_0 > 0$, and assume that

$$\|D^2 u(x)\| := \sup_{e \in \mathbb{S}^{n-1}} \partial_{ee} u(x) \le A \quad \forall\, x \in \Omega.$$

Then (1.4) holds. Indeed, given $x \in \Omega$, we can choose a system of coordinates so that $D^2 u(x)$ is a diagonal matrix with eigenvalues $(\lambda_1, \dots, \lambda_n)$. Since $\det D^2 u(x) = \prod_{i=1}^n \lambda_i$, it follows that

$$\prod_{i=1}^{n} \lambda_i = f(x) \ge a_0 \qquad \text{and} \qquad \max_{1 \le k \le n} \lambda_k \le A,$$

and we get that

$$\lambda_i = \frac{\prod_j \lambda_j}{\prod_{k \ne i} \lambda_k} \ge \frac{a_0}{A^{n-1}} \quad \forall\, i = 1, \dots, n,$$

which proves (1.4) with $C := \max\{A, A^{n-1}/a_0\}$.

1.2 Some history

The first notable results on the Monge–Ampère equation are due to Minkowski [90, 91]. At the end of the 19th and the beginning of the 20th century, he proved the existence of weak solutions to the "prescribed Gaussian curvature problem" (this is now called the "Minkowski problem"): Given a function K on the sphere, find a convex surface whose Gaussian curvature in polar coordinates is equal to K (see Section 2.6).

Using convex polyhedra with given generalized curvatures at the vertices, forty years later, Alexandrov proved the existence of a weak solution in all dimensions, as well as the C^1 smoothness of solutions in two dimensions [1, 2, 3, 4]. Then, based on these results, Alexandrov [5] (and also Bakelman [9] in two dimensions) introduced the notion of a generalized solution to the Monge–Ampère equation, and proved existence and uniqueness of solutions to the Dirichlet problem (see Chapter 2). Their treatment also led to the Alexandrov–Bakelman maximum principle which plays a fundamental role in the study of non-divergence elliptic equations (see, for instance, [58, Section 9.8]).

The notion of weak solutions introduced by Alexandrov (now called "Alexandrov solutions") has often been used in recent years, and a lot of attention has been given to proving smoothness of Alexandrov solutions under suitable assumptions on the right-hand side and the boundary data.

The regularity of Alexandrov solutions in higher dimensions is a very delicate problem. In the 1960s, Pogorelov found a convex function which is not of class C^2 but satisfies the Monge–Ampère equation (1.1) inside $B_{1/2}$ with positive analytic right-hand side (see Section 3.2). As we shall explain in detail in Chapter 4, the main obstacle to regularity is the presence of a line segment in the graph of u (in other words, u is not strictly convex). Indeed, Calabi [24] and Pogorelov [97] were able to prove a priori interior second- and third-derivative estimates for strictly convex solutions, in turn proving the smoothness of strictly convex Alexandrov solutions (see Section 3.3, where instead of Calabi's estimates we use the interior regularity theory for fully nonlinear uniformly elliptic equations established by Evans [41] and Krylov [75] in the 1980s).

Later on, using the continuity method, Ivochkina [65], Krylov [76], and Caffarelli–Nirenberg–Spruck [23] were able to show the existence of globally smooth solutions to the Dirichlet problem. In particular, Alexandrov solutions are smooth up to the boundary provided all given data are smooth (cf. Section 3.1).

In all the situations mentioned above one assumes that the right-hand side f is positive and sufficiently smooth, but the many applications of Monge–Ampère motivated the development of a regularity theory under weaker assumptions on f.

In the 1990s, under only the hypothesis that f is bounded away from zero and infinity, Caffarelli proved the $C^{1,\alpha}$ regularity of strictly convex solutions [15]. Then, if in addition f is continuous (resp. $C^{0,\alpha}$), using a perturbation argument, Caffarelli proved an interior $W^{2,p}$ estimate for any $p > 1$ (resp. $C^{2,\alpha}$ interior estimates) [14]. More recently, the author and De Philippis proved interior $L \log L$ estimates on D^2u

when f is merely bounded away from zero and infinity [30], and together with Savin they improved this result showing that $u \in W^{2,1+\varepsilon}_{\mathrm{loc}}$ [36] (see also [106]). All these results, as well as some selected applications, are described in Chapter 4.

We also mention that these interior regularity results have a natural counterpart at the boundary, which is briefly described in Section 5.1.2.

In recent years, applications to optimal transportation and antenna design problems have motivated the study of a much more general class of Monge–Ampère-type equations, as well as their boundary regularity and their linearization (cf. Chapter 5).

It is important to remark that, despite all these recent developments, several important questions on the regularity of solutions to Monge–Ampère are still open, and the Monge–Ampère equation and its applications remain very active areas of research.