# 1 Introduction

Throughout this book, we will work in $\Omega$ (except in sections 4.2 and 5.4), where

$$\Omega \subset \mathbf{R}^n \textbf{ is open and bounded.}$$

We denote by $\langle \cdot, \cdot \rangle$ the standard inner product in $\mathbf{R}^n$, and set $|x| = \sqrt{\langle x, x \rangle}$ for $x \in \mathbf{R}^n$. We use the standard notion of open balls: For $r > 0$ and $x \in \mathbf{R}^n$,

$$B_r(x) := \{y \in \mathbf{R}^n \mid |x - y| < r\}, \quad \text{and} \quad B_r := B_r(0).$$

For a function $u : \Omega \to \mathbf{R}$, we denote its gradient and Hessian matrix at $x \in \Omega$, respectively, by

$$Du(x) := \begin{pmatrix} \frac{\partial u(x)}{\partial x_1} \\ \vdots \\ \frac{\partial u(x)}{\partial x_n} \end{pmatrix},$$

$$D^2u(x) := \begin{pmatrix} \frac{\partial^2 u(x)}{\partial x_1^2} & \cdots & j\text{-th} & \cdots & \frac{\partial^2 u(x)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots & & \vdots \\ i\text{-th} & \cdots & \frac{\partial^2 u(x)}{\partial x_i \partial x_j} & \cdots & \vdots \\ \vdots & & \vdots & & \vdots \\ \frac{\partial^2 u(x)}{\partial x_n \partial x_1} & \cdots & \cdots & \cdots & \frac{\partial^2 u(x)}{\partial x_n^2} \end{pmatrix}.$$

Also, $S^n$ denotes the set of all real-valued $n \times n$ symmetric matrices. Note that if $u \in C^2(\Omega)$, then $D^2u(x) \in S^n$ for $x \in \Omega$.

We recall the standard ordering in $S^n$:

$$X \leq Y \quad \Longleftrightarrow \quad \langle X\xi, \xi \rangle \leq \langle Y\xi, \xi \rangle \qquad \text{for } \forall \xi \in \mathbf{R}^n.$$

We will also use the following notion in sections 6 and 7: For $\xi =^t (\xi_1, \ldots, \xi_n)$, $\eta =^t (\eta_1, \ldots, \eta_n) \in \mathbf{R}^n$, we denote by $\xi \otimes \eta$ the $n \times n$ matrix whose $(i, j)$-entry is $\xi_i \eta_j$ for $1 \leq i, j \leq n$;

$$\xi \otimes \eta = \begin{pmatrix} \xi_1\eta_1 & \cdots & j\text{-th} & \cdots & \xi_1\eta_n \\ \vdots & & \vdots & & \vdots \\ i\text{-th} & \cdots & \xi_i\eta_j & \cdots & \vdots \\ \vdots & & \vdots & & \vdots \\ \xi_n\eta_1 & \cdots & \cdots & \cdots & \xi_n\eta_n \end{pmatrix}.$$

We are concerned with general second-order partial differential equations (PDEs for short):

$$F(x, u(x), Du(x), D^2u(x)) = 0 \quad \text{in } \Omega. \tag{1.1}$$

We suppose (except in several sections) that

$$F : \Omega \times \mathbf{R} \times \mathbf{R}^n \times S^n \to \mathbf{R} \quad \text{is continuous}$$

with respect to all variables.

## 1.1 From classical solutions to weak solutions

As the first example of PDEs, we present the Laplace equation:

$$-\triangle u = 0 \quad \text{in } \Omega. \tag{1.2}$$

Here, we define $\triangle u := \text{trace}(D^2u)$. In the literature of the viscosity solution theory, we prefer to have the minus sign in front of $\triangle$.

Of course, since we do not require any boundary condition yet, all polynomials of degree one are solutions of (1.2). In many textbooks (particularly those for engineers), under certain boundary condition, we learn how to solve (1.2) when $\Omega$ has some special shapes such as cubes, balls, the half-space or the whole space $\mathbf{R}^n$. Here, "solve" means that we find an explicit formula of $u$ using elementary functions such as polynomials, trigonometric ones, etc.

However, the study on (1.2) in such special domains is not applicable because, for instance, solutions of equation (1.2) represent the density of a gas in a bottle, which is neither a ball nor a cube.

Unfortunately, in general domains, it seems impossible to find formulas for solutions $u$ with elementary functions. Moreover, in order to cover problems arising in physics, engineering and finance, we will have to study more general and complicated PDEs than (1.2). Thus, we have to deal with general PDEs (1.1) in general domains.

If we give up having formulas for solutions of (1.1), how do we investigate PDEs (1.1) ? In other words, what is the right question in the study of PDEs ?

In the literature of the PDE theory, the most basic questions are as follows:

| | | |
|---|---|---|
| **(1)** | **Existence:** | **Does there exist a solution ?** |
| **(2)** | **Uniqueness:** | **Is it the only solution ?** |
| **(3)** | **Stability:** | **If the PDE changes a little, does the solution change a little ?** |

The importance of the existence of solutions is trivial since, otherwise, the study on the PDE could be useless.

To explain the significance of the uniqueness of solutions, let us remember the reason why we study the PDE. Usually, we discuss PDEs or their solutions to understand some specific phenomena in nature, engineerings or economics etc. Particularly, people working in applications want to know how the solution looks like, moves, behaves etc. For this purpose, it might be powerful to use numerical computations. However, numerical analysis only shows us an "approximate" shapes, movements, etc. Thus, if there are more than one solution, we do not know which is approximated by the numerical solution.

Also, if the stability of solutions fails, we could not predict what will happen from the numerical experiments even though the uniqueness of solutions holds true.

Now, let us come back to the most essential question:

**What is the "solution" of a PDE ?**

For example, it is natural to call a function $u : \Omega \to \mathbf{R}$ a solution of (1.1) if there exist the first and second derivatives, $Du(x)$ and $D^2u(x)$, for all $x \in \Omega$, and (1.1) is satisfied at each $x \in \Omega$ when we plug them in the left hand side of (1.1). Such a function $u$ will be called a **classical solution** of (1.1).

However, unfortunately, it is difficult to seek for a classical solution because we have to verify that it is sufficiently differentiable and that it satisfies the equality (1.1) simultaneously.

Instead of finding a classical solution directly, we have decided to choose the following strategy:

**(A) Find a candidate of the classical solution,**
**(B) Check the differentiability of the candidate.**

In the standard books, the candidate of a classical solution is called a **weak solution**; if the weak solution has the first and second derivatives, then it becomes a classical solution. In the literature, showing the differentiability of solutions is called the study on the **regularity** of those.

Thus, with these terminologies, we may rewrite the above with mathematical terms:

**(A) Existence of weak solutions,**
**(B) Regularity of weak solutions.**

However, when we cannot expect classical solutions of a PDE to exist, what is the right candidate of solutions ?

We will call a function the candidate of solutions of a PDE if it is a "unique" and "stable" weak solution under a suitable setting. In section 2, we will define such a candidate named "viscosity solutions" for a large class of PDEs, and in the proceeding sections, we will extend the definition to more general (possibly discontinuous) functions and PDEs.

In the next subsection, we show a brief history on "weak solutions" to remind what was known before the birth of viscosity solutions.

## 1.2 Typical examples of weak solutions

In this subsection, we give two typical examples of PDEs to derive two kinds of weak solutions which are unique and stable.

### 1.2.1 Burgers' equation

We consider Burgers' equation, which is a model PDE in Fluid Mechanics:

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial (u^2)}{\partial x} = 0 \quad \text{in } \mathbf{R} \times (0, \infty) \tag{1.3}$$

under the initial condition:

$$u(x, 0) = g(x) \quad \text{for } x \in \mathbf{R}, \tag{1.4}$$

where $g$ is a given function.

In general, we cannot find classical solutions of (1.3)-(1.4) even if $g$ is smooth enough. See [8] for instance.

In order to look for the appropriate notion of weak solutions, we first introduce a function space $C_0^1(\mathbf{R} \times [0, \infty))$ as a "test function space":

$$C_0^1(\mathbf{R} \times [0, \infty)) := \left\{ \phi \in C^1(\mathbf{R} \times [0, \infty)) \;\middle|\; \begin{array}{c} \text{there is } K > 0 \text{ such that} \\ \mathit{supp}\ \phi \subset [-K, K] \times [0, K] \end{array} \right\}.$$

Here and later, we denote by *supp* $\phi$ the following set:

$$supp\ \phi := \overline{\{(x,t) \in \mathbf{R} \times [0,\infty) \mid \phi(x,t) \neq 0\}}.$$

Suppose that $u$ satisfies (1.3). Multiplying (1.3) by $\phi \in C_0^1(\mathbf{R} \times [0,\infty))$ and then, using integration by parts, we have

$$\int_{\mathbf{R}} \int_0^\infty \left( u\frac{\partial \phi}{\partial t} + \frac{u^2}{2}\frac{\partial \phi}{\partial x} \right)(x,t)dtdx + \int_{\mathbf{R}} u(x,0)\phi(x,0)dx = 0.$$

Since there are no derivatives of $u$ in the above, this equality makes sense if $u \in \cup_{K>0} L^1((-K,K)\times(0,K))$. Hence, we may adapt the following property as the definition of weak solutions of (1.3)-(1.4).

$$\begin{cases} \displaystyle\int_{\mathbf{R}} \int_0^\infty \left( u\frac{\partial \phi}{\partial t} + \frac{u^2}{2}\frac{\partial \phi}{\partial x} \right)(x,t)dtdx + \int_{\mathbf{R}} g(x)\phi(x,0)dx = 0 \\ \qquad\qquad\qquad \text{for all } \phi \in C_0^1(\mathbf{R} \times [0,\infty)). \end{cases}$$

We often call this a <u>weak solution in the distribution sense</u>. As you noticed, we derive this notion by an essential use of integration by parts. We say that a PDE is in **divergence form** when we can adapt the notion of weak solutions in the distribution sense. When the PDE is not in divergence form, we say that it is in **nondivergence form**.

We note that the solution of (1.3) may have singularities even though the initial value $g$ belongs to $C^\infty$ by an observation via "characteristic method". From the definition of weak solutions, we can derive the so-called Rankine-Hugoniot condition on the set of singularities.

On the other hand, unfortunately, we cannot show the uniqueness of weak solutions of (1.3)-(1.4) in general while we know the famous Lax-Oleinik formula (see [8] for instance), which is the "expected" solution.

In order to obtain the uniqueness of weak solutions, for the definition, we add the following property (called "entropy condition") which holds for the expected solution given by the Lax-Oleinik formula: There is $C > 0$ such that

$$u(x+z,t) - u(x,t) \leq \frac{Cz}{t}$$

for all $(x,t,z) \in \mathbf{R} \times (0,\infty) \times (0,\infty)$. We call $u$ an entropy solution of (1.3) if it is a weak solution satisfying this inequality. It is also known that such a weak solution has a certain stability property.

We note that this entropy solution satisfies the above mentioned important properties; "existence, uniqueness and stability". Thus, it must be a right definition for weak solutions of (1.3)-(1.4).

### 1.2.2 Hamilton-Jacobi equations

Next, we shall consider general Hamilton-Jacobi equations, which arise in Optimal Control and Classical Mechanics:

$$\frac{\partial u}{\partial t} + H(Du) = 0 \quad \text{in } (x,t) \in \mathbf{R}^n \times (0,\infty) \tag{1.5}$$

under the same initial condition (1.4).

In this example, we suppose that $H : \mathbf{R}^n \to \mathbf{R}$ is convex, *i.e.*

$$H(\theta p + (1-\theta)q) \le \theta H(p) + (1-\theta)H(q) \tag{1.6}$$

for all $p, q \in \mathbf{R}^n, \theta \in [0,1]$.

*Remark.* Since a convex function is locally Lipschitz continuous in general, we do not need to assume the continuity of $H$.

*Example.* In Classical Mechanics, we often call this $H$ a "Hamiltonian".

As a simple example of $H$, we have $H(p) = |p|^2$.

Notice that we **cannot** adapt the weak solution in the distribution sense for (1.5) since we **cannot** use the integration by parts.

We next introduce the Lagrangian $L : \mathbf{R}^n \to \mathbf{R}$ defined by

$$L(q) = \sup_{p \in \mathbf{R}^n} \{\langle p, q\rangle - H(p)\}.$$

When $H(p) = |p|^2$, it is easy to verify that the maximum is attained in the right hand side of the above.

It is surprising that we have a neat formula for the expected solution (called Hopf-Lax formula) presented by

$$u(x,t) = \min_{y \in \mathbf{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\}. \tag{1.7}$$

More precisely, it is shown that the right hand side of (1.7) is differentiable and satisfies (1.5) almost everywhere.

Thus, we could call $u$ a weak solution of (1.5)-(1.4) when $u$ satisfies (1.5) almost everywhere. However, if we decide to use this notion as a weak solution, the uniqueness of those fails in general. We will see an example in the next section.

As was shown for Burgers' equation, in order to say that the "unique weak" solution is given by (1.7), we have to add one more property for the definition of weak solutions: There is $C > 0$ such that

$$u(x+z,t) - 2u(x,t) + u(x-z,t) \leq C|z|^2 \tag{1.8}$$

for all $x, z \in \mathbf{R}, t > 0$. This is called the "semi-concavity" of $u$.

We note that (1.8) is a hypothesis on the one-sided bound of second derivatives of functions $u$.

In 60s, Kruzkov showed that the limit function of approximate solutions by the vanishing viscosity method (see the next section) has this property (1.8) when $H$ is convex. He named $u$ a "generalized" solution of (1.5) when it satisfies (1.5) almost everywhere and (1.8).

To my knowledge, between Kruzkov's works and the birth of viscosity solutions, there had been no big progress in the study of first-order PDEs in nondivergence form.

*Remark.* The convexity (1.6) is a natural hypothesis when we consider only optimal control problems where one person intends to minimize some "costs" ("energy" in terms of Physics). However, when we treat game problems (one person wants to minimize costs while the other tries to maximize them), we meet non-convex and non-concave (*i.e.* "fully nonlinear")

Hamiltonians. See section 4.2.

In this book, since we are concerned with viscosity solutions of PDEs in nondivergence form, for which the integration by parts argument cannot be used to define the notion of weak solutions in the distribution sense, we shall give typical examples of such PDEs.

*Example.* (Bellman and Isaacs equations)

We first give Bellman equations and Isaacs equations, which arise in (stochastic) optimal control problems and differential games, respectively. As will be seen, those are extensions of linear PDEs.

Let $A$ and $B$ be sets of parameters. For instance, we suppose $A$ and $B$ are (compact) subsets in $\mathbf{R}^m$ (for some $m \geq 1$). For $a \in A$, $b \in B$, $x \in \Omega$,

$r \in \mathbf{R}$, $p = {}^t(p_1, \ldots, p_n) \in \mathbf{R}^n$, and $X = (X_{ij}) \in S^n$, we set

$$\begin{aligned} L^a(x,r,p,X) &:= -\mathrm{trace}(A(x,a)X) + \langle g(x,a), p\rangle + c(x,a)r, \\ L^{a,b}(x,r,p,X) &:= -\mathrm{trace}(A(x,a,b)X) + \langle g(x,a,b), p\rangle + c(x,a,b)r. \end{aligned}$$

Here $A(\cdot,a), A(\cdot,a,b), g(\cdot,a), g(\cdot,a,b), c(\cdot,a)$ and $c(\cdot,a,b)$ are given functions for $(a,b) \in A \times B$.

For inhomogeneous terms, we consider functions $f(\cdot,a)$ and $f(\cdot,a,b)$ in $\Omega$ for $a \in A$ and $b \in B$.

We call the following PDEs Bellman equations:

$$\sup_{a\in A}\{L^a(x,u(x),Du(x),D^2u(x)) - f(x,a)\} = 0 \quad \text{for } x \in \Omega. \tag{1.9}$$

Notice that the supremum over $A$ is taken at each point $x \in \Omega$.

Taking account of one more parameter set $B$, we call the following PDEs Isaacs equations:

$$\sup_{a\in A}\inf_{b\in B}\{L^{a,b}(x,u(x),Du(x),D^2u(x)) - f(x,a,b)\} = 0 \quad \text{for } x \in \Omega \tag{1.10}$$

and

$$\inf_{b\in B}\sup_{a\in A}\{L^{a,b}(x,u(x),Du(x),D^2u(x)) - f(x,a,b)\} = 0 \quad \text{for } x \in \Omega. \tag{1.10$'$}$$

*Example.* ("Quasi-linear" equations)

We say that a PDE is quasi-linear if the coefficients of $D^2u$ contains $u$ or $Du$. Although we will not study quasilinear PDEs in this book, we give some of those which are in nondivergence form.

We first give the PDE of mean curvature type:

$$F(x,p,X) := -\left(|p|^2\mathrm{trace}(X) - \langle Xp, p\rangle\right).$$

Notice that this $F$ is independent of $x$-variables. We refer to [12] for applications where this kind of operators appears.

Next, we show a relatively "new" one called $L^\infty$-Laplacian:

$$F(x,p,X) := -\langle Xp, p\rangle.$$

Again, this $F$ does not contain $x$-variables. We refer to Jensen's work [16], where he first studied the PDE "$-\langle D^2uDu, Du\rangle = 0$ in $\Omega$" via the viscosity solution approach.