Chapter 1

Introduction to viscosity solutions for Hamilton–Jacobi equations

1.1. Introduction

Basic notation. Let $u : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. We have some basic notation as follows.

- $Du(x) = \nabla u(x) = \left(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_n}(x)\right)$.
- The Hessian of u at x is

$$D^2u(x) = \begin{pmatrix} \frac{\partial^2 u}{\partial x_1^2}(x) & \frac{\partial^2 u}{\partial x_1 \partial x_2}(x) & \ldots & \frac{\partial^2 u}{\partial x_1 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 u}{\partial x_n \partial x_1}(x) & \frac{\partial^2 u}{\partial x_n \partial x_2}(x) & \ldots & \frac{\partial^2 u}{\partial x_n^2}(x) \end{pmatrix}.$$

- The Laplacian of u at x is

$$\Delta u(x) = \mathrm{tr}(D^2u(x)) = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}(x).$$

For $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ smooth, we write

- $Du(x,t) = D_x u(x,t)$, and $u_t(x,t) = \frac{\partial u}{\partial t}(x,t)$;
- $D^2u(x,t) = D_x^2 u(x,t)$, and $\Delta u(x,t) = \Delta_x u(x,t)$.

The following equations are of interests.

Cauchy problem. We consider the initial value problem

$$\text{(C)}\qquad \begin{cases} u_t(x,t) + F(x, Du(x,t), D^2u(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n, \end{cases}$$

where $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is the unknown. Here, the initial data u_0 is given.

Static (stationary) problem. Given $\lambda \geq 0$, we consider the equation

$$(\mathrm{S}_\lambda)\qquad \lambda u + F(x, Du, D^2u) = 0 \qquad \text{in } \mathbb{R}^n.$$

Here $u : \mathbb{R}^n \to \mathbb{R}$ is the unknown. In both problems, $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}^n \to \mathbb{R}$ is a given function, where $\mathbb{S}^n$ is the set of all symmetric matrices of size n. These problems come from a lot of sources such as

- Hamilton–Jacobi equations (classical mechanics, n-body problems);
- optimal control theory;
- differential games (two-player zero-sum differential games);
- front propagation (the level set method).

Next, we present a few examples that lead to either Cauchy problems or static problems.

Example 1.1 (First-order front propagation)**.** Consider a surface $\Gamma_t \subset \mathbb{R}^n$ moving under a law of motion at time $t > 0$ with the initial profile Γ_0. The goal is to study how the family $\{\Gamma_t\}_{t\geq 0}$ evolves.

- The simplest example in Γ_0 is the unit sphere in $\mathbb{R}^n$, and every point is moving inward in the normal direction to the surface with constant (vector) speed 1. Then Γ_t remains a sphere for $t \in [0,1)$, and eventually shrinks into a point at $t = 1$, which is located at the center.
- In general, if each point on the surface Γ_t is moving with variable velocity, then the situation becomes more complicated. Osher, Sethian [**OS88**] introduced the level set method (numerically) to study this problem. The rigorous treatment was developed later by Evans, Spruck [**ES91**] and Chen, Giga, Goto [**CGG91**], independently.

Let us assume that Γ_t is the 0-level set of a function $u(x,t)$ for each $t \geq 0$; that is,

$$\Gamma_t = \{x \in \mathbb{R}^n \,:\, u(x,t) = 0\}.$$

Assume further that Γ_t is a closed hypersurface in $\mathbb{R}^n$. We set $u(x,t) > 0$ in the region enclosed by Γ_t and $u(x,t) < 0$ elsewhere. Suppose u and Γ_t are

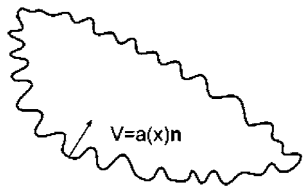

Figure 1.1. Front propagation of $\{\Gamma_t\}_{t\geq 0}$.

smooth and the given velocity at $x \in \Gamma_t$ is

$$V(x) = a(x)\mathbf{n}(x),$$

where $\mathbf{n}(x)$ is the inward normal vector to Γ_t at x. Here, $a : \mathbb{R}^n \to \mathbb{R}$ is a given function. See Figure 1.1. Let us then try to find a PDE for $u(x,t)$ based on this given law of motion.

For a point $x(0) \in \Gamma_0$, we keep track of its position $x(t) \in \Gamma_t$ for $t \geq 0$ under this front propagation problem. First of all, we have

$$x'(t) = a(x(t))\,\mathbf{n}(x(t)) = a(x(t))\,\frac{Du(x(t),t)}{|Du(x(t),t)|}.$$

Moreover, in light of the fact that $u(x(t),t) = 0$,

$$\frac{d}{dt}\left(u(x(t),t)\right) = u_t\,(x(t),t) + Du(x(t),t)\cdot x'(t) = 0,$$

which implies

$$u_t\,(x(t),t) + a(x(t))\,|Du(x(t),t)| = 0.$$

Thus, we obtain a PDE

$$u_t + a(x)|Du| = 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty),$$

which is a first-order Hamilton–Jacobi equation.

Example 1.2 (G-equation)**.** We assume the same settings as in Example 1.1. The law of motion is different here and is given as

$$V(x) = \mathbf{n}(x) + \mathbf{W}(x),$$

for each $x \in \Gamma_t$. Here, $\mathbf{n}(x)$ is the inward normal vector to Γ_t at x, and $\mathbf{W} : \mathbb{R}^n \to \mathbb{R}^n$ is a given vector field.

As above, for a point $x(0) \in \Gamma_0$, we keep of track of its position $x(t) \in \Gamma_t$ for $t \geq 0$ under this front propagation problem. Firstly,

$$x'(t) = \mathbf{n}(x(t)) + \mathbf{W}(x(t)) = \frac{Du(x(t),t)}{|Du(x(t),t)|} + \mathbf{W}(x(t)).$$

Besides, $u(x(t),t)=0$ gives

$$\frac{d}{dt}\left(u(x(t),t)\right) = u_t\,(x(t),t) + Du(x(t),t)\cdot x'(t) = 0,$$

which implies

$$u_t\,(x(t),t) + |Du(x(t),t)| + \mathbf{W}(x(t))\cdot Du(x(t),t) = 0.$$

Thus, we obtain a PDE

$$u_t + |Du| + \mathbf{W}(x)\cdot Du = 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty),$$

which is another first-order Hamilton–Jacobi equation. This equation is called a G-equation, which is very popular in the combustion science literature.

Example 1.3 (Level set mean curvature flow)**.** Let $\{\Gamma_t\}_{t\geq 0}$ be smooth surfaces in $\mathbb{R}^n$. Let $\kappa(x)$ be the summation of all principle curvatures at $x \in \Gamma_t$ of the surface Γ_t. By convention, we say that $\kappa(x)$ is the mean curvature at $x \in \Gamma_t$ of the surface Γ_t. For example, if Γ_t is a sphere of radius $R(t) > 0$, then for $x \in \Gamma_t$, $\kappa(x) = \frac{n-1}{R(t)}$.

Again, we assume that Γ_t is the 0-level set of some function $u(x,t)$; that is,

$$\Gamma_t = \{x \in \mathbb{R}^n \,:\, u(x,t) = 0\}\,.$$

Assume that Γ_t is a closed hypersurface in $\mathbb{R}^n$. Set $u(x,t) > 0$ in the region enclosed by Γ_t and $u(x,t) < 0$ elsewhere. Assume u and Γ_t are smooth and the given velocity at $x \in \Gamma_t$ is

$$V(x) = \kappa(x)\mathbf{n}(x),$$

where $\mathbf{n}(x)$ is the inward normal vector to Γ_t at x. As above, for a point $x(0) \in \Gamma_0$, we keep track of its position $x(t) \in \Gamma_t$ for $t \geq 0$ under this mean curvature flow motion. It is clear that

$$u_t(x(t),t) + Du(x(t),t)\cdot x'(t) = 0,$$

where

$$x'(t) = \kappa(x(t))\mathbf{n}(x(t)) = -\mathrm{div}\left(\frac{Du(x(t),t)}{|Du(x(t),t)|}\right)\frac{Du(x(t),t)}{|Du(x(t),t)|}.$$

Thus the level set mean curvature flow equation of interest is

$$u_t = |Du|\mathrm{div}\left(\frac{Du}{|Du|}\right) \qquad \text{in } \mathbb{R}^n \times (0,\infty).$$

Of course, the Cauchy problem (C) is a general form of all equations occurring in the above examples. From the PDE viewpoints, we focus on the following main issues:

(1) Well-posedness theory: Existence, uniqueness, and stability of solutions.
(2) The study of fine properties of solutions such as regularity, large time behavior, homogenization, dynamical properties of solutions.

Example 1.4 (One-dimensional eikonal equation)**.**

$$\begin{cases} |u'(x)| &= 1 \qquad \text{in } (-1,1), \\ u(-1) = u(1) &= 0. \end{cases}$$

It is not hard to see that there are infinitely many almost everywhere solutions to this equation. To design such a solution, one just needs to draw its graph which is zero at the two endpoints ± 1 and always has slope ± 1 in between. Here are some simple but important observations.

(1) This eikonal equation has no classical solution (C^1 solution).
(2) If u is an a.e. solution, then so is $-u$. In a sense, if we want to select only one solution (well-posedness goal), then we have to break down the symmetry. Besides, we might need to be careful with stability then.
(3) Clearly, we need to impose a bit more in order to get fewer solutions. This is typically the case in the theories of viscosity solutions, renormalized solutions, etc.

1.2. Vanishing viscosity method for first-order Hamilton–Jacobi equations

Let us look at the simple Cauchy problem for Hamilton–Jacobi equation

$$\begin{cases} u_t + H(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n, \end{cases} \tag{1.1}$$

where $H : \mathbb{R}^n \to \mathbb{R}$ is the given Hamiltonian and u_0 is the given initial data. Assume that H and u_0 are smooth enough. One way to study solutions of (1.1) is to use the idea of the vanishing viscosity procedure. For each $\varepsilon > 0$, we consider

$$\begin{cases} u^\varepsilon_t + H(Du^\varepsilon) &= \varepsilon \Delta u^\varepsilon \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{1.2}$$

Under some appropriate assumptions on H and u_0, (1.2) is a parabolic equation, which has a unique smooth solution u^ε. The question is what happens as $\varepsilon \to 0$? Do we have $u^\varepsilon \to u$ for some function u and in some sense? If it is the case, do we have that u solves (1.1) in some sense? This is the idea of a selection principle, which often appears when one introduces some approximation procedures to a nonlinear PDE.

Evans [**Eva80**] first showed that this process leads to $u^\varepsilon \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$ and u solves (1.1) in the viscosity sense, which will be defined later. Later on, Crandall and Lions [**CL83**] proved the uniqueness of viscosity solutions to (1.1); thus, they established the firm foundation for the theory of viscosity solutions to first-order equations. Roughly speaking, the procedure is carried out as follows:

- Equation (1.2) is a parabolic equation, and thus, it satisfies the maximum principle.
- Hamiltonian $H(p)$ is nonlinear in p in general (e.g., $H(p) = |p|^2$), so there is no way to use integration by parts technique to define weak solutions.
- There is an a priori estimate for $\{u^\varepsilon\}_{\varepsilon \in (0,1)}$: There exists a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that
$$\|u_t^\varepsilon\|_{L^\infty(\mathbb{R}^n \times [0,\infty))} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n \times [0,\infty))} \le C.$$
We will supply a proof of this later. Thus, $\{u^\varepsilon(x,t)\}_{\varepsilon \in (0,1)}$ is equicontinuous, and by the Arzelà–Ascoli theorem, there exists $\{\varepsilon_j\} \searrow 0$ such that $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$ as $j \to \infty$. We hence hope that u solves (1.1) naturally in some sense that fits well within the context of the maximum principle.

Let us now analyze further along this line for a possible definition of weak solutions to (1.1). Let $\varphi \in C^\infty(\mathbb{R}^n \times [0,\infty))$ be an arbitrary smooth test function. First, assume that $u^\varepsilon - \varphi$ has a strict maximum at $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$; then the maximum principle says that

$$\begin{cases} (u^\varepsilon - \varphi)_t(x_0,t_0) &= 0 \\ D(u^\varepsilon - \varphi)(x_0,t_0) &= 0 \\ \Delta(u^\varepsilon - \varphi)(x_0,t_0) &\le 0 \end{cases} \implies \varphi_t(x_0,t_0) + H(D\varphi(x_0,t_0)) \le \varepsilon \Delta\varphi(x_0,t_0).$$

In a sense, this is an L^∞-integration by parts trick, which kicks the derivatives of the solutions to our favorite (nice) test functions φ. Let us modify this argument a little bit to study u. Assume that $u - \varphi$ has a strict max at (x_0,t_0). Then, if $u^\varepsilon \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$, for ε small, $u^\varepsilon - \varphi$ has a max nearby at $(x_\varepsilon,t_\varepsilon)$, and $(x_\varepsilon,t_\varepsilon) \to (x_0,t_0)$ by passing to a subsequence if necessary as $\varepsilon \to 0+$. By the above analysis,

$$\varphi_t(x_\varepsilon,t_\varepsilon) + H(D\varphi(x_\varepsilon,t_\varepsilon)) \le \varepsilon \Delta\varphi(x_\varepsilon,t_\varepsilon).$$

Letting $\varepsilon \to 0+$, we arrive at

$$\varphi_t(x_0,t_0) + H(D\varphi(x_0,t_0)) \le 0.$$

Similarly, if $u-\psi$ has a strict min at $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$ for a given smooth test function ψ, then we get

$$\psi_t(x_0,t_0) + H(D\psi(x_0,t_0)) \ge 0.$$

The above two criteria seem natural from the viewpoint of the maximum principle, and indeed, they constitute the definition of viscosity solutions in the following.

1.2.1. Definition of viscosity solutions via touching functions. Let us denote

- $\mathrm{BUC}(\mathbb{R}^n)$ the space of bounded, uniformly continuous functions on $\mathbb{R}^n$;
- $\mathrm{Lip}(\mathbb{R}^n)$ the space of Lipschitz functions on $\mathbb{R}^n$.

For a given initial data $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$, we give the following definition, which was formulated by Crandall, Evans, Lions [**CEL84**].

Definition 1.5 (Viscosity solutions of (1.1))**.** For each time $T>0$, a function $u \in \mathrm{BUC}(\mathbb{R}^n \times [0,T))$ is called

(a) a viscosity subsolution of (1.1) if for any $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ such that $u(x_0,t_0) = \varphi(x_0,t_0)$ and $u-\varphi$ has a strict max at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$, then

$$\varphi_t(x_0,t_0) + H(D\varphi(x_0,t_0)) \le 0,$$

and $u(\cdot,0) \le u_0$;

(b) a viscosity supersolution of (1.1) if for any $\psi \in C^1(\mathbb{R}^n \times (0,T))$ such that $u(x_0,t_0) = \psi(x_0,t_0)$ and $u-\psi$ has a strict min at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$, then

$$\psi_t(x_0,t_0) + H(D\psi(x_0,t_0)) \ge 0,$$

and $u(\cdot,0) \ge u_0$;

(c) a viscosity solution of (1.1) if it is both a viscosity subsolution and a viscosity supersolution.

Remark 1.6. We actually do not need the condition $u(x_0,t_0) = \varphi(x_0,t_0)$ in the above definition since we can always add a constant to φ to adjust it appropriately. Requiring $u(x_0,t_0) = \varphi(x_0,t_0)$ means that φ touches u from above geometrically, which is quite helpful to think about the definition in geometric terms. See Figure 1.2.

Figure 1.2. An illustration of φ touches u from above at (x_0, t_0).

1.2.2. Problems.

Exercise 1. Consider the eikonal problem mentioned earlier:

$$\begin{cases} |u'(x)| &= 1 \qquad \text{in } (-1,1),\\ u(1) = u(-1) &= 0. \end{cases}$$

(a) Show that there is no C^1 solution.

(b) Show that all the continuous a.e. solutions with finitely many gradient jumps are mutually viscosity subsolutions.

Exercise 2. For each $\varepsilon > 0$, consider the equation

$$\begin{cases} |(u^\varepsilon)'| &= 1 + \varepsilon (u^\varepsilon)'' \qquad \text{in } (-1,1),\\ u^\varepsilon(1) = u^\varepsilon(-1) &= 0. \end{cases}$$

(a) Solve the equation to find u^ε for each $\varepsilon > 0$.

(b) Find the limit of u^ε as $\varepsilon \to 0$.

Exercise 3. Prove that in the above definition of viscosity solutions of (1.1), we can equivalently require the test functions $\varphi, \psi \in C^2(\mathbb{R}^n \times (0,\infty))$. The same holds when we require that $\varphi, \psi \in C^\infty(\mathbb{R}^n \times (0,\infty))$.

Exercise 4. Prove that in the above definition of viscosity subsolutions of (1.1), we can equivalently require that $u - \varphi$ has a local maximum at $(x_0, t_0) \in \mathbb{R}^n \times (0, T)$ (instead of strict maximum).

Exercises 3–4 show that the definition of viscosity solutions is rather flexible in terms of smoothness of test functions and requirements of local/strict maximum, minimum points. One can use any of these equivalent forms of definitions from now on.

1.2.3. Definition of viscosity solutions via generalized differentials.

Definition 1.7. Let u be a real-valued function defined on the open set $\Omega \subset \mathbb{R}^n$. For any $x \in \Omega$, the sets

$$D^- u(x) = \left\{ p \in \mathbb{R}^n \ : \ \liminf_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} \geq 0 \right\},$$
$$D^+ u(x) = \left\{ p \in \mathbb{R}^n \ : \ \limsup_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} \leq 0 \right\}$$

are called the (Frechét) subdifferential and superdifferential of u at x, respectively.

Theorem 1.8. *Let $\Omega \subset \mathbb{R}^n$ be an open set, and let $f : \Omega \to \mathbb{R}$ be a continuous function. Then, for $x \in \Omega$, $p \in D^+ f(x)$ if and only if there is a function $\varphi \in C^1(\Omega, \mathbb{R})$ such that $D\varphi(x) = p$ and $f - \varphi$ has a local max at x. The same claim holds if we replace superdifferential/max by subdifferential/min.*

Proof. We only need to prove "$\Longrightarrow$". Let $p \in D^+ f(x)$. If we have that

$$\limsup_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} < 0,$$

then we can find $r > 0$ such that $u(y) \leq u(x) + p \cdot (y - x)$ for all $y \in B_r(x)$. Simply set $\varphi(y) = u(x) + p \cdot (y - x) + C|y - x|^2$ for $C > 0$ sufficiently large to conclude.

We now consider the case that

$$\limsup_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} = 0.$$

There exists $\delta > 0$ such that $B_\delta(x) \subset \Omega$. Define $\sigma : (0, \delta] \to \mathbb{R}$ by

$$\sigma(r) = \sup_{y \in \overline{B}_r(x)} \frac{f(y) - f(x) - p \cdot (y - x)}{|y - x|} \qquad \Longrightarrow \qquad \lim_{r \to 0} \sigma(r) = \inf_{r > 0} \sigma(r) = 0.$$

Set $\sigma(0) = 0$. It is clear that σ is nondecreasing. It is not hard to check that σ is continuous as well. By the definition of σ,

$$f(y) \leq f(x) + p \cdot (y - x) + \sigma(|y - x|)|y - x| \qquad \text{for all} \qquad y \in \overline{B}_\delta(x).$$

Now define $\rho : \left[0, \frac{\delta}{2}\right] \to \mathbb{R}$ by

$$\rho(r) = \int_r^{2r} \sigma(s)\, ds.$$

It is clear that, for $r \in [0, \frac{\delta}{2}]$,

$$r\sigma(r) \leq \rho(r) \leq r\sigma(2r) \quad \Longrightarrow \quad \sigma(r) \leq \frac{\rho(r)}{r} \leq \sigma(2r). \tag{1.3}$$

Besides, ρ satisfies $\rho'(r) = 2\sigma(2r) - \sigma(r)$ for $r \in [0, \frac{\delta}{2}]$, and $\rho(0) = \rho'(0) = 0$. Now let us define for $y \in B_{\frac{\delta}{2}}(x)$

$$\varphi(y) = f(x) + p \cdot (y - x) + \rho(|y - x|).$$

We have $\varphi \in C^1\left(B_{\frac{\delta}{2}}(x)\right)$ and $\varphi(x) = f(x)$. Moreover, (1.3) yields that $D\varphi(x) = p$ since

$$\lim_{y \to x} \frac{\varphi(y) - \varphi(x) - p \cdot (y - x)}{|y - x|} = \lim_{y \to x} \frac{\rho(|y - x|)}{|y - x|} = 0.$$

Also, $u - \varphi$ has a local max at x since for $|y - x| < \frac{\delta}{2}$,

$$f(y) - f(x) \leq p \cdot (y - x) + \sigma(|y - x|)|y - x| \leq p \cdot (y - x) + \rho(|y - x|) = \varphi(y) - \varphi(x).$$

Finally, we can extend φ smoothly to Ω easily to complete the proof. □

Using the notation of subdifferentials and superdifferentials, one is able to give an equivalent definition of viscosity solutions, which is clear from the result of Theorem 1.8. Nevertheless, let us present this equivalent definition here for completeness and for later usage. In fact, it is important to keep in mind both of these definitions.

We consider the first-order static PDE

$$F\left(x, u(x), Du(x)\right) = 0 \qquad \text{in } \Omega. \tag{1.4}$$

Here, $\Omega \subset \mathbb{R}^n$ is a given open set, and $u : \Omega \to \mathbb{R}$ is an unknown. The function $F : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is a given continuous function.

Definition 1.9 (An equivalent definition of viscosity solutions to (1.4))**.** A function $u \in C(\Omega)$ is a viscosity subsolution of (1.4) if

$$F(x, u(x), p) \leq 0 \qquad \text{for every } x \in \Omega, p \in D^+u(x). \tag{1.5}$$

A function $u \in C(\Omega)$ is a viscosity supersolution of (1.4) if

$$F(x, u(x), p) \geq 0 \qquad \text{for every } x \in \Omega, p \in D^-u(x). \tag{1.6}$$

We say that u is a viscosity solution of (1.4) if it is both a viscosity subsolution and a viscosity supersolution of (1.4).

We have some basic properties of generalized differentials as follows.

Proposition 1.10. *Let $f : \Omega \to \mathbb{R}$ be a given function, and let $x \in \Omega$ be a given point. Then the following properties hold:*

(a) $D^+f(x) = -D^-(-f)(x)$.

(b) *$D^+f(x)$ and $D^-f(x)$ are convex (possibly empty).*

(c) *$D^+f(x)$ and $D^-f(x)$ are both nonempty if and only if f is differentiable at x. In this case, we have that $D^+f(x) = D^-f(x) = \{Df(x)\}$.*

(d) *If $f \in C(\Omega)$, the sets of points where a one-sided differential exists,*

$$\Omega^+ = \left\{x \in \Omega \,:\, D^+ f(x) \neq \emptyset\right\}, \qquad \Omega^- = \left\{x \in \Omega \,:\, D^- f(x) \neq \emptyset\right\},$$

are both nonempty. In fact, they are dense in Ω.

Proof. It is easy to see that (a) and (b) are obvious from the definitions. Let us proceed to prove the remaining two claims.

(c) If f is differentiable at x, then clearly $Df(x) \in D^+ f(x) \cap D^- f(x)$. Furthermore, if $p \in D^+ f(x)$, then there exists $\varphi \in C^1(\Omega)$ such that

$$\varphi(x) = f(x) \qquad \text{and} \qquad D\varphi(x) = p,$$

and $f - \varphi$ has a local maximum at x; hence $D(f - \varphi)(x) = 0$, and therefore $p = D\varphi(x) = Df(x)$. Doing similarly for $D^- f(x)$, we obtain $D^+ f(x) = D^- f(x) = \{Df(x)\}$.

For the converse, assume that $D^+ f(x)$ and $D^- f(x)$ are both nonempty. Pick any $p \in D^+ f(x)$ and $q \in D^- f(x)$; then there exist $\varphi, \psi \in C^1(\Omega)$ such that

$$\begin{cases} \varphi(x) = \psi(x) = f(x), \\ f - \varphi \text{ has local maximum at } x, \text{ and } D\varphi(x) = p, \\ f - \psi \text{ has local minimum at } x, \text{ and } D\psi(x) = q. \end{cases}$$

Therefore, in a neighborhood $B_\delta(x)$ for $\delta > 0$ sufficiently small, we have

$$\psi(y) \leq f(y) \leq \varphi(y) \qquad \text{for all } y \in B_\delta(x).$$

Since $\psi, \varphi \in C^1(\Omega)$, it is easy to see that f is also differentiable at x, and thus, $D^+ f(x) = D^- f(x) = \{Df(x)\}$.

(d) Let $x_0 \in \Omega$, and let $\varepsilon > 0$ be sufficiently small. We will show that there exists a function $\varphi \in C^1(\Omega)$ such that $f - \varphi$ has local maximum in $B(x_0, \varepsilon)$ at some point y in $B(x_0, \varepsilon)$. Consider a smooth function in $C^1(B(x_0, \varepsilon))$ given by

$$\varphi(x) = \frac{1}{\varepsilon^2 - |x - x_0|^2} \qquad \text{for all } x \in B(x_0, \varepsilon) \subset \Omega.$$

It is clear that

$$\varphi(x) \to +\infty \qquad \text{as } |x - x_0| \to \varepsilon - .$$

Since f is continuous, we have $f - \varphi$ has a local maximum in $B(x_0, \varepsilon)$, denoted by y. We conclude that $p = D\varphi(y) \in D^+ f(y)$, and therefore, Ω^+ is dense in Ω.

By a similar proof, Ω^- is also dense in Ω. □

Remark 1.11. It is worth noting that if $D^+u(x) = \emptyset$, then the viscosity subsolution test for u automatically holds there. Similarly, if $D^-u(x) = \emptyset$, then the viscosity supersolution test for u holds true at x.

Nevertheless, as $\Omega^\pm$ are dense in Ω, we surely need to check for the subsolution and supersolution tests for at least a.e. $x \in \Omega$. Later on, when we add more assumptions, we will typically have more regularity results on u (e.g., u is Lipschitz in Ω), and we will discuss this situation further later.

1.2.4. Problems.

Exercise 5. Let u be a viscosity solution of (1.4). Show the following:

(a) If u is differentiable at $y \in \Omega$, then $F(y, u(y), Du(y)) = 0$ in the classical sense.

(b) If $u \in C^1(\Omega)$, then u is a classical solution to (1.4).

Exercise 6. Let $u(x) = |x|$ for all $x \in B_1(0)$. Compute $D^\pm u(x)$ for all $x \in B_1(0)$. Then, show that u is not a viscosity solution to

$$|Du| = 1 \qquad \text{in } B_1(0).$$

Exercise 7. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, and let $F : \Omega \times \mathbb{R}^n \to \mathbb{R}$ be a continuous function. Assume that $u \in C(\Omega)$ is a viscosity solution to

$$F(y, Du(y)) = 0 \qquad \text{in } \Omega.$$

Show that $\tilde{u} = -u$ is a viscosity solution to

$$\tilde{F}(y, D\tilde{u}(y)) = 0 \qquad \text{in } \Omega,$$

where $\tilde{F}(y, p) = -F(y, -p)$ for $(y, p) \in \Omega \times \mathbb{R}^n$.

Exercise 8. Let $U \subset \mathbb{R}^n$ be a bounded domain with smooth boundary. Let $u(x) = \operatorname{dist}(x, \partial U)$ for $x \in \overline{U}$. Show that u is Lipschitz continuous and that u solves the following eikonal equation in the viscosity sense:

$$\begin{cases} |Du| = 1 & \text{in } U, \\ u = 0 & \text{on } \partial U. \end{cases}$$

1.3. Existence of viscosity solutions via the vanishing viscosity method

Let us look at the usual Cauchy problem that was discussed earlier:

$$\begin{cases} u_t + H(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0, \infty), \\ u(x, 0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{1.7}$$

Before going to the proof of the existence of viscosity solutions to (1.7), we need the following stability lemma.

Lemma 1.12 (Stability of maximum/minimum points)**.** *Let $u \in C(\mathbb{R}^n)$, and let $\varphi \in C^1(\mathbb{R}^n)$ be such that $u(x_0) = \varphi(x_0)$ for some $x_0 \in \mathbb{R}^n$, and assume $u - \varphi$ has a strict max (or strict min) at x_0. Assume $\{u^\varepsilon\}_{\varepsilon>0} \subset C(\mathbb{R}^n)$ converges to u locally uniformly on $\mathbb{R}^n$ as $\varepsilon \to 0+$. Then, for $\varepsilon > 0$ small enough, $u^\varepsilon - \varphi$ has a local max (or min) at x_ε nearby x_0, and there is a subsequence $\{\varepsilon_j\} \searrow 0$ such that $x_{\varepsilon_j} \to x_0$ as $j \to \infty$.*

Proof. Let $r > 0$ be sufficiently small such that $u(x) - \varphi(x) < 0$ for any $x \in B(x_0, 2r) \setminus \{x_0\}$. Since $\partial B(x_0, r)$ is compact, we note that

$$\alpha = \max\left\{u(x) - \varphi(x) \,:\, x \in \partial B(x_0, r)\right\} < 0.$$

Since $u^\varepsilon \to u$ uniformly on $\overline{B}(x_0, r)$, there exists $\lambda_r > 0$ such that, for any $\varepsilon < \lambda_r$,

$$\max_{\overline{B}(x_0,r)} |u^\varepsilon(x) - u(x)| < -\frac{\alpha}{2}.$$

From this fact, on $\partial B(x_0, r)$, we imply

$$\max_{\partial B(x_0,r)} \Big(u^\varepsilon(x) - \varphi(x)\Big) \le \max_{\overline{B}(x_0,r)} |u^\varepsilon(x) - u(x)| + \max_{\partial B(x_0,r)} \Big(u(x) - \varphi(x)\Big) < \frac{\alpha}{2}.$$

But $u^\varepsilon(x_0) - \varphi(x_0) = u^\varepsilon(x_0) - u(x_0) > \frac{\alpha}{2}$. Thus, $u^\varepsilon(x) - \varphi(x)$ must obtain its maximum over $\overline{B}(x_0, r)$ at some point $x_\varepsilon \in B(x_0, r)$. Finally, let $\varepsilon_1 < \lambda_1$, and construct by induction $\{\varepsilon_j\}$ as follows. Let $r = \frac{1}{j}$ for $j \ge 2$, and choose $\varepsilon_j < \min\left\{\lambda_{\frac{1}{j}}, \varepsilon_{j-1}\right\}$. By the above, we obtain $\{\varepsilon_j\} \searrow 0$ and $u^{\varepsilon_j} - \varphi$ achieves its local maximum over the closed ball $\overline{B}\left(x_0, \frac{1}{j}\right)$ at x_{ε_j} and $\left|x_{\varepsilon_j} - x_0\right| < \frac{1}{j}$. The proof is complete. □

Next is our existence result for viscosity solutions to (1.7). For now, we need to assume beforehand that (1.8) has a unique solution u^ε, and u^ε enjoys a priori estimates (1.9), which is independent of $\varepsilon \in (0, 1)$. These will be discussed and verified later.

Theorem 1.13 (Existence of viscosity solutions via the vanishing viscosity method)**.** *For each $\varepsilon > 0$, consider the equation*

$$\begin{cases} u^\varepsilon_t + H(Du^\varepsilon) &= \varepsilon\Delta u^\varepsilon \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \quad u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{1.8}$$

Here, the initial data $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ is given. Assume that (1.8) has a unique smooth solution u^ε for any $\varepsilon > 0$. Furthermore, we assume that there exists a constant $C > 0$ independent of $\varepsilon \in (0, 1)$ such that, for each $\varepsilon \in (0, 1)$,

$$|u^\varepsilon_t| + |Du^\varepsilon| \le C \qquad \text{on } \mathbb{R}^n \times [0,\infty). \tag{1.9}$$

Then, there exists a subsequence $\{\varepsilon_j\} \searrow 0$ *such that* $u^{\varepsilon_j} \to u$ *locally uniformly on* $\mathbb{R}^n \times [0,\infty)$ *for some function* $u \in C(\mathbb{R}^n \times [0,\infty))$. *Moreover,* u *is a viscosity solution of* (1.7).

Proof. Thanks to (1.9), by the Arzelà–Ascoli theorem, there exists a subsequence $\{\varepsilon_j\} \searrow 0$ such that $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$ for some function $u \in C(\mathbb{R}^n \times [0,\infty))$.

We show that u is a viscosity subsolution of (1.7). The proof that u is a viscosity supersolution of (1.7) is similar, hence omitted. Fix $T > 0$. By Exercise 3, we can choose test functions in $C^2(\mathbb{R}^n \times (0,T))$ or $C^\infty(\mathbb{R}^n \times (0,T))$. Pick $\varphi \in C^2(\mathbb{R}^n \times (0,T))$ such that $u - \varphi$ has a strict max at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. By Lemma 1.12, we may assume that $u^{\varepsilon_i} - \varphi$ has a local max at $(x_i,t_i) \in \mathbb{R}^n \times (0,T)$ for each $i \in \mathbb{N}$, and $(x_i,t_i) \to (x_0,t_0)$ as $i \to \infty$. Since $u^{\varepsilon_i} - \varphi$ has a local max at (x_i,t_i), we have

$$\begin{cases} D\big(u^{\varepsilon_i} - \varphi\big)(x_i,t_i) &= 0,\\ \big(u^{\varepsilon_i} - \varphi\big)_t(x_i,t_i) &= 0,\\ \Delta\big(u^{\varepsilon_i} - \varphi\big)(x_i,t_i) &\le 0. \end{cases}$$

Then, substituting these relations into (1.8), we obtain

$$\varphi_t(x_i,t_i) + H\big(D\varphi(x_i,t_i)\big) = \varepsilon_i \Delta u^{\varepsilon_i}(x_i,t_i) \le \varepsilon_i \Delta\varphi(x_i,t_i).$$

Let $i \to \infty$, to yield $\varphi_t(x_0,t_0) + H(D\varphi(x_0,t_0)) \le 0$, which concludes the proof. $\square$

Remark 1.14.

1. If $u - \varphi$ has a strict max at $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$, then it does not mean that u touches φ from below at (x_0,t_0), but we can always add a constant to φ by

$$\overline{\varphi}(x,t) = \varphi(x,t) - \underbrace{\varphi(x_0,t_0) + u(x_0,t_0)}_{\text{a constant}}$$

so that u touches $\overline{\varphi}$ from below at (x_0,t_0). Geometrically, it is sometimes easier and more helpful to think about touching u by smooth test functions from above and below when performing sub/supersolution tests.

2. Note that by the vanishing viscosity method, we have the a priori estimate

$$|u_t(x,t)| + |Du(x,t)| \le C \qquad \text{in } \mathbb{R}^n \times [0,\infty),$$

which means that u is Lipschitz in space and time. Hence, by Rademacher's theorem, u is differentiable a.e. in $\mathbb{R}^n \times (0,\infty)$.

1.4. Consistency and stability of viscosity solutions

From the vanishing viscosity procedure, we obtain a viscosity solution $u \in \mathrm{Lip}(\mathbb{R}^n \times [0,\infty))$ to the Hamilton–Jacobi equation

$$\begin{cases} u_t + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{1.10}$$

It is worth noting that (1.10) is a bit more complicated than (1.7), but the procedure is the same. By Rademacher's theorem, u is differentiable a.e. in $\mathbb{R}^n \times (0,\infty)$. We show that indeed, if u is differentiable at (x_0, t_0), then u satisfies (1.10) in the usual sense at this point. Before showing that, we need the following lemma (compare this with Exercise 5).

Lemma 1.15. *Let Ω be an open subset of $\mathbb{R}^n$, and let $u : \Omega \to \mathbb{R}$ be a continuous function. If u is differentiable at $x_0 \in \Omega$, then there exist $\varphi, \psi \in C^1(\Omega)$ such that $\varphi(x_0) = u(x_0) = \psi(x_0)$ and $\varphi(x) < v(x) < \psi(x)$ for $x \in B_r(x_0) \setminus \{x_0\}$ for some $r > 0$ sufficiently small. As a consequence, $Du(x_0) = D\varphi(x_0) = D\psi(x_0)$.*

Proof. If u is differentiable at x_0, then $D^+u(x_0) = D^-u(x_0) = \{Du(x_0)\}$. There exist $\overline{\varphi}, \overline{\psi} \in C^1(\Omega)$ such that $\overline{\varphi}(x_0) = \overline{\psi}(x_0) = u(x_0)$, $D\overline{\varphi}(x_0) = D\overline{\psi}(x_0) = Du(x_0)$, and $u - \overline{\varphi}$ has a local minimum at x_0 and $u - \overline{\psi}$ has a local maximum at x_0. The proof is complete by setting, for $x \in \Omega$,

$$\varphi(x) = \overline{\varphi}(x) - |x - x_0|^2 \qquad \text{and} \qquad \psi(x) = \overline{\psi}(x) + |x - x_0|^2. \qquad \square$$

Theorem 1.16. *Let u be a viscosity solution of* (1.10) *constructed by the vanishing viscosity method. If u is differentiable at* $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$, *then*

$$u_t(x_0, t_0) + H(x, Du(x_0, t_0)) = 0.$$

Proof. Using the lemma above, there exist two test functions φ, $\psi \in C^1(\mathbb{R}^n \times (0,\infty))$ such that $u - \varphi$ has a strict minimum at (x_0, t_0), $u - \psi$ has a strict maximum at (x_0, t_0), and $u_t(x_0, t_0) = \varphi_t(x_0, t_0) = \psi_t(x_0, t_0)$, $Du(x_0, t_0) = D\varphi(x_0, t_0) = D\psi(x_0, t_0)$. Then, the viscosity subsolution and supersolution tests imply the result. $\square$

We now show that viscosity solutions are stable under locally uniform convergence.

Theorem 1.17 (Stability of viscosity solutions to (1.10))**.** *Assume that*

$$\begin{cases} H_k \to H & \text{locally uniformly in } \mathbb{R}^n \times \mathbb{R}^n, \\ u_{0,k} \to u_0 & \text{locally uniformly on } \mathbb{R}^n, \\ u_k \to u & \text{locally uniformly on } \mathbb{R}^n \times [0,\infty). \end{cases}$$

For each $k \in \mathbb{N}$, assume further that u_k is a viscosity solution to

$$\begin{cases} (u_k)_t + H_k(x, Du_k) &= 0 \qquad \text{in } \mathbb{R}^n \times (0, \infty), \\ u_k(x, 0) &= u_{0,k}(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{1.11}$$

Then u is a viscosity solution to (1.10).

Proof. It is clear that u satisfies the initial condition in the classical sense. We show that u is a viscosity subsolution to (1.10). The supersolution property follows in a similar way.

Take any C^1 test function φ such that $u - \varphi$ has strict max at $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$. Since $u_k \to u$ locally uniformly on $\mathbb{R}^n \times [0, \infty)$, for k large enough, $u_k - \varphi$ has a local max (x_k, t_k) near (x_0, t_0), and $(x_k, t_k) \to (x_0, t_0)$ up to passing to a subsequence if necessary. Since u_k is a viscosity solution of (1.11), we have

$$\varphi_t(x_k, t_k) + H_k\left(x_k, D\varphi(x_k, t_k)\right) \le 0.$$

Letting $k \to \infty$ and using the assumptions, we obtain

$$\varphi_t(x_0, t_0) + H\left(x_0, D\varphi(x_0, t_0)\right) \le 0.$$

The proof is complete. □

1.5. The comparison principle and uniqueness result for static problems

We consider the static problem

$$u(x) + H(x, Du(x)) = 0 \qquad \text{in } \mathbb{R}^n. \tag{1.12}$$

In this section we assume the following Lipschitz assumption on H. There exists a constant $C > 0$ such that, for all $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x, p) - H(y, p)| &\le C(1 + |p|)|x - y|, \\ |H(x, p) - H(x, q)| &\le C|p - q|. \end{cases} \tag{1.13}$$

The main result is the following comparison principle.

Theorem 1.18 (The comparison principle for static equation (1.12))**.** *Assume* (1.13)*. Assume that $u, v \in \mathrm{BUC}(\mathbb{R}^n)$ are a viscosity subsolution and a viscosity supersolution of* (1.12)*, respectively. Then, $u(x) \le v(x)$ for any $x \in \mathbb{R}^n$.*

Before writing down a proof, it is fair to say that condition (1.13) is a bit restrictive. It is fine to assume H is Lipschitz in x, but it is too strict to assume that H is global Lipschitz in p. For example, if one considers the classical mechanics Hamiltonian $H(x,p)=\frac{|p|^2}{2}+V(x)$, then (1.13) does not hold. This deserves some explanation after the proof of this comparison result.

Proof of Theorem 1.18. We give a proof by using the classical "doubling variables" method.

Since u, v are bounded in $\mathbb{R}^n$, assume by contradiction that

$$\sup_{x\in\mathbb{R}^n}\Big(u(x)-v(x)\Big)=\sigma>0.$$

Then, there exists $x_1\in\mathbb{R}^n$ such that $u(x_1)-v(x_1)>\frac{3\sigma}{4}$. For $\varepsilon>0$ such that

$$\varepsilon<\frac{\sigma}{8\left(1+|x_1|^2\right)}\quad\Longrightarrow\quad -2\varepsilon|x_1|^2>-\frac{\sigma}{4},$$

we consider the auxiliary function

$$\begin{aligned}\Phi^\varepsilon:\mathbb{R}^n\times\mathbb{R}^n&\longrightarrow\mathbb{R}\\ (x,y)&\longmapsto\Phi^\varepsilon(x,y)=u(x)-v(y)-\frac{|x-y|^2}{\varepsilon^2}-\varepsilon\big(|x|^2+|y|^2\big).\end{aligned}$$

Then Φ^ε is continuous, bounded above, and tends to $-\infty$ as either $|x|\to\infty$ or $|y|\to\infty$, and hence, it must achieve a global maximum at some point $(x_\varepsilon,y_\varepsilon)\in\mathbb{R}^{2n}$. Note first that

$$\Phi^\varepsilon\left(x_\varepsilon,y_\varepsilon\right)\geq\Phi^\varepsilon(x_1,x_1)=u(x_1)-v(x_1)-2\varepsilon|x_1|^2\geq\frac{3\sigma}{4}-\frac{\sigma}{4}=\frac{\sigma}{2}. \tag{1.14}$$

As this is the first time we present the doubling variables method, let us proceed gently by breaking the proofs into various simple steps as follows.

Step 1. We have $\Phi^\varepsilon\left(x_\varepsilon,y_\varepsilon\right)\geq\Phi^\varepsilon(0,0)$, which gives

$$u(x_\varepsilon)-v(y_\varepsilon)\geq u(0)-v(0)+\frac{|x_\varepsilon-y_\varepsilon|^2}{\varepsilon^2}+\varepsilon\big(|x_\varepsilon|^2+|y_\varepsilon|^2\big).$$

Letting $C=2(\|u\|_{L^\infty(\mathbb{R}^n)}+\|v\|_{L^\infty(\mathbb{R}^n)})$, we obtain

$$C\geq\frac{|x_\varepsilon-y_\varepsilon|^2}{\varepsilon^2}+\varepsilon\big(|x_\varepsilon|^2+|y_\varepsilon|^2\big).$$

This implies that $(x_\varepsilon-y_\varepsilon)\to 0$ as $\varepsilon\to 0$ and

$$|x_\varepsilon-y_\varepsilon|\leq C\varepsilon\qquad\text{and}\qquad|x_\varepsilon|+|y_\varepsilon|\leq\frac{C}{\sqrt{\varepsilon}}.$$

Step 2. We claim further that $|x_\varepsilon - y_\varepsilon| = o(\varepsilon)$; that is, $\frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \to 0$ as $\varepsilon \to 0$. Indeed, this follows by noting that

$$\Phi^\varepsilon(x_\varepsilon, y_\varepsilon) \geq \Phi^\varepsilon(x_\varepsilon, x_\varepsilon) \implies \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \leq v(x_\varepsilon) - v(y_\varepsilon) + \varepsilon\big(|x_\varepsilon|^2 - |y_\varepsilon|^2\big)$$
$$\implies \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \leq v(x_\varepsilon) - v(y_\varepsilon) + C\varepsilon^{3/2}$$

and that v is uniformly continuous in $\mathbb{R}^n$, which gives $\lim_{\varepsilon\to 0}(v(x_\varepsilon) - v(y_\varepsilon)) = 0$.

Step 3. Now $x \mapsto \Phi^\varepsilon(x, y_\varepsilon)$ has a max at x_ε, which means

$$x \longmapsto u(x) - \underbrace{\left(\frac{|x - y_\varepsilon|^2}{\varepsilon^2} + \varepsilon|x|^2\right)}_{\text{test function } \varphi(x)} \quad \text{has a max at } x_\varepsilon.$$

As u is a viscosity subsolution of (1.12), by the viscosity subsolution test, we have

$$u(x_\varepsilon) + H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) \leq 0. \tag{1.15}$$

Step 4. Next, as $y \mapsto \Phi^\varepsilon(x_\varepsilon, y)$ has a max at y_ε, we get

$$y \longmapsto v(y) - \underbrace{\left(-\frac{|x_\varepsilon - y|^2}{\varepsilon^2} - \varepsilon|y|^2\right)}_{\text{test function } \psi(y)} \quad \text{has a min at } y_\varepsilon.$$

Since v is a viscosity supersolution of (1.12), by the viscosity supersolution test, we obtain

$$v(y_\varepsilon) + H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) \geq 0. \tag{1.16}$$

Step 5. From (1.15) and (1.16), this implies

$$\begin{aligned} & u(x_\varepsilon) - v(y_\varepsilon) \\ & \quad \leq H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right). \end{aligned} \tag{1.17}$$

Now using the Lipschitz assumption (1.13) of H, we have

$$H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \leq 2C\varepsilon|y_\varepsilon|,$$

$$H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \leq C|x_\varepsilon - y_\varepsilon|\left(1 + \frac{2|x_\varepsilon - y_\varepsilon|}{\varepsilon^2}\right),$$

$$H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) \leq 2C\varepsilon|x_\varepsilon|.$$

Plugging all of these together, we obtain

$$H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right)$$
$$\leq 2C\left(\varepsilon\big(|x_\varepsilon| + |y_\varepsilon|\big) + \frac{|x_\varepsilon - y_\varepsilon|}{2} + \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2}\right).$$

Combine this with (1.17) to deduce that

$$u(x_\varepsilon) - v(y_\varepsilon) \leq 2C\left(\varepsilon\big(|x_\varepsilon| + |y_\varepsilon|\big) + \frac{|x_\varepsilon - y_\varepsilon|}{2} + \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2}\right). \tag{1.18}$$

Recall that (1.14) gives

$$u(x_\varepsilon) - v(y_\varepsilon) \geq \Phi^\varepsilon(x_\varepsilon, y_\varepsilon) \geq \frac{\sigma}{2}.$$

Plug it into (1.18) to yield

$$\frac{\sigma}{2} \leq 2C\left(\varepsilon\big(|x_\varepsilon| + |y_\varepsilon|\big) + \frac{|x_\varepsilon - y_\varepsilon|}{2} + \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2}\right).$$

Letting $\varepsilon \to 0$ and using results from Step 1 and Step 2, we get

$$0 < \frac{\sigma}{2} \leq 0,$$

which is a contradiction. The proof is complete. ☐

Remark 1.19. In the above proof via the doubling variable method, the following observation, which is elementary, plays a key role:

$$\frac{\partial}{\partial x}\left(\frac{|x - y_\varepsilon|^2}{\varepsilon^2}\right)\Big|_{x = x_\varepsilon} = \frac{\partial}{\partial y}\left(\frac{-|x_\varepsilon - y|^2}{\varepsilon^2}\right)\Big|_{y = y_\varepsilon} = \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}.$$

Corollary 1.20 (Uniqueness of viscosity solution of static equation (1.12))**.** *Assume* (1.13)*. If $u, v \in \mathrm{BUC}(\mathbb{R}^n)$ are the viscosity solution of* (1.12)*, then $u \equiv v$ in $\mathbb{R}^n$.*

Proof. Since u is a viscosity subsolution and v is a viscosity supersolution of (1.12), by the comparison principle above, we have $u \leq v$. Conversely, since v is a viscosity subsolution and u is a viscosity supersolution of (1.12), we deduce $v \leq u$. Thus, $u = v$. ☐

Remark 1.21. Let us discuss condition (1.13) further. In general, if we do not know anything further about the solutions, except that they are in $\mathrm{BUC}(\mathbb{R}^n)$, then it is hard to remove this condition. Still, from the proof, it is easy to see that (1.13) can be changed into the following weaker one: For all $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x,p) - H(y,p)| & \leq & \omega_H((1 + |p|)|x - y|), \\ |H(x,p) - H(x,q)| & \leq & \omega_H(|p - q|). \end{cases} \tag{1.19}$$

Here, $\omega_H : [0,\infty) \to [0,\infty)$ is a modulus of continuity corresponding to H; that is, $\lim_{r\to 0}\omega_H(r) = 0$. Still, a disadvantage of (1.19) is that these two inequalities have to hold for all $p, q \in \mathbb{R}^n$.

Nevertheless, we often have more information, such as the existence of a Lipschitz viscosity solution u to (1.12), and in such cases, (1.13) can be relaxed significantly. The following points are quite well-known to experts in the field, but sometimes, they are not written down and explained clearly.

1. It is typically the case that for a given nice H, we can obtain a Lipschitz viscosity solution u to (1.12) via some methods (e.g., the vanishing viscosity method, or the Perron method to be described later). It is then clear that information of H matters only for $(x,p) \in \mathbb{R}^n \times B(0,R)$ for $R = \|Dv\|_{L^\infty(\mathbb{R}^n)} + 1$. We then define a modification $\tilde{H}$ of H such that
$$\tilde{H}(x,p) = \begin{cases} H(x,p) & \text{for all } x \in \mathbb{R}^n,\ |p| \le R, \\ |p| & \text{for all } x \in \mathbb{R}^n,\ |p| \ge 2R, \end{cases}$$
and $\tilde{H}$ satisfies (1.13). Then, v is still a viscosity solution to (1.12) with $\tilde{H}$ in place of H. And, for this new equation with $\tilde{H}$ in place of H, we have the uniqueness of solutions. This technique of modifying H is used a lot in the theory of viscosity solutions whenever a priori estimates are available.
2. Again, under nice enough assumptions, let us assume that there is a Lipschitz viscosity solution u to (1.12). Here is a different way to look at the uniqueness proof by comparing every solution of (1.12) with u, which is already known to be Lipschitz. Let $v \in \mathrm{BUC}(\mathbb{R}^n)$ be another viscosity solution to (1.12). By looking back to Step 2 of the proof of Theorem 1.18, we have in addition that $|x_\varepsilon - y_\varepsilon| \le C\varepsilon^2$. Then, in order to have the uniqueness result, we are able to relax (1.13) a lot; for example, (1.13) can be replaced by the following:
$$\begin{cases} \text{For each } R > 0, \text{ there exists } C_R > 0 \text{ so that,} \\ \text{for } x, y \in \mathbb{R}^n,\ p, q \in B(0,R), \\ |H(x,p) - H(y,p)| \le C_R|x-y|, \\ |H(x,p) - H(x,q)| \le C_R|p-q|. \end{cases} \tag{1.20}$$
Actually, (1.13) can also be replaced by the following condition, which is much simpler and weaker than (1.20):
$$H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \qquad \text{for every } R > 0. \tag{1.21}$$
One can see that (1.13) and (1.20) have the same spirit. And, similarly, (1.19) and (1.21) are of the same type.

1.5.1. Problems.

Exercise 9. Consider the setting in Exercise 8. Show that $u(x) = \operatorname{dist}(x, \partial U)$ for $x \in \overline{U}$ is the unique viscosity solution to the given eikonal equation.

1.6. The comparison principle and uniqueness result for Cauchy problems

We consider the usual Cauchy problem

$$\begin{cases} u_t + H(x, Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0, \infty), \\ u(x, 0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{1.22}$$

In this section, we still assume that H satisfies the Lipschitz assumption (1.13). For clarity, let us recall it here: There exists a constant $C > 0$ such that, for $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x,p) - H(y,p)| & \leq & C(1 + |p|)|x - y|, \\ |H(x,p) - H(x,q)| & \leq & C|p - q|. \end{cases}$$

The main result here is the comparison principle for (1.22), which is similar to Theorem 1.18. But before we proceed, we need the following simple lemma.

Lemma 1.22 (Extrema at terminal time $t = T$)**.** *Fix $T > 0$. Let u be a viscosity subsolution to* (1.22)*, and let $\varphi \in C^1(\mathbb{R}^n \times [0, T])$ be such that $u - \varphi$ has a strict max at (x_0, t_0) over $(x, t) \in \mathbb{R}^n \times (0, T]$. Then the subsolution test still holds; that is,*

$$\varphi_t(x_0, t_0) + H(x_0, D\varphi(x_0, t_0)) \leq 0.$$

Proof. It suffices to only consider the case $t_0 = T$. Define $\varphi_\varepsilon(x, t) = \varphi(x, t) + \frac{\varepsilon}{T - t}$ for any fixed $\varepsilon > 0$. Then for $\varepsilon > 0$ small enough, $u - \varphi_\varepsilon$ has a local max at $(x_\varepsilon, t_\varepsilon)$, and $(x_\varepsilon, t_\varepsilon) \to (x_0, t_0)$ as $\varepsilon \to 0$ by passing to a subsequence if necessary (see Exercise 10 below for confirmation). As $u - \varphi_\varepsilon$ has a local max at $(x_\varepsilon, t_\varepsilon)$, by the definition of viscosity subsolutions, we have

$$(\varphi_\varepsilon)_t (x_\varepsilon, t_\varepsilon) + H(x_\varepsilon, D\varphi_\varepsilon(x_\varepsilon, t_\varepsilon)) \leq 0,$$

which means

$$\varphi_t(x_\varepsilon, t_\varepsilon) + \frac{\varepsilon}{(T - t_\varepsilon)^2} + H(x_\varepsilon, D\varphi(x_\varepsilon, t_\varepsilon)) \leq 0.$$

Therefore,

$$\varphi_t(x_\varepsilon, t_\varepsilon) + H(x_\varepsilon, D\varphi(x_\varepsilon, t_\varepsilon)) \leq 0.$$

Let $\varepsilon \to 0$ to conclude. □

Here is our main result on the comparison principle for the Cauchy problem.

Theorem 1.23 (Comparison principle for Cauchy problem (1.22))**.** *Assume* (1.13)*. Fix* $T > 0$*. Assume* $u, v \in \mathrm{BUC}(\mathbb{R}^n \times [0,T])$ *are a viscosity subsolution and supersolution of* (1.22)*, respectively. Then,* $u(x,t) \le v(x,t)$ *on* $\mathbb{R}^n \times [0,T]$*.*

The proof is quite similar to that of Theorem 1.18, but it is worth presenting here since there is the time variable t involved.

Proof. We aim at proving that $u(x,t) \le v(x,t)$ for all $(x,t) \in \mathbb{R}^n \times (0,T]$. Since u, v are bounded, assume by contradiction that

$$\sup_{(x,t)\in\mathbb{R}^n\times[0,T]} \Big(u(x,t) - v(x,t)\Big) = \sigma > 0.$$

Then, there exists $(x_1, t_1) \in \mathbb{R}^n \times [0,T]$ so that $u(x_1,t_1) - v(x_1,t_1) > \frac{3\sigma}{4}$. It is clear that $t_1 > 0$. Let ε and λ be positive numbers such that

$$\varepsilon < \frac{\sigma}{16(|x_1|^2+1)} \quad \text{and} \quad \lambda < \frac{\sigma}{16(t_1+1)} \quad \Longrightarrow \quad 2\varepsilon|x_1|^2 + 2\lambda t_1 < \frac{\sigma}{4}.$$

For these ε, λ fixed, we consider the following auxiliary function $\Phi : \mathbb{R}^n \times \mathbb{R}^n \times [0,T] \times [0,T] \to \mathbb{R}$:

$$\Phi(x,y,t,s) = u(x,t) - v(y,s) - \frac{|x-y|^2 + |s-t|^2}{\varepsilon^2} - \varepsilon(|x|^2 + |y|^2) - \lambda(s+t).$$

Since Φ is continuous and bounded above, it must achieve its maximum at some point $(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon)$ on $\mathbb{R}^n \times [0,T]^2$. Note first that

$$\Phi(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon) \ge \Phi(x_1, x_1, t_1, t_1) > \frac{3\sigma}{4} - 2\varepsilon|x_1|^2 - 2\lambda t_1 > \frac{\sigma}{2}.$$

Again, we divide the proofs into various small steps.

Step 1. As $\Phi(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon) \ge \Phi(0,0,0,0)$,

$$\begin{aligned} &u(x_\varepsilon, t_\varepsilon) - v(y_\varepsilon, s_\varepsilon) \\ &\quad \ge u_0(0) - v_0(0) + \frac{|x_\varepsilon - y_\varepsilon|^2 + |s_\varepsilon - t_\varepsilon|^2}{\varepsilon^2} + \varepsilon(|x_\varepsilon|^2 + |y_\varepsilon|^2) + \lambda(s_\varepsilon + t_\varepsilon), \end{aligned}$$

which yields

$$C \ge \frac{|x_\varepsilon - y_\varepsilon|^2 + |s_\varepsilon - t_\varepsilon|^2}{\varepsilon^2} + \varepsilon(|x_\varepsilon|^2 + |y_\varepsilon|^2) + \lambda(s_\varepsilon + t_\varepsilon).$$

Thus, we obtain

$$(1.23) \qquad |x_\varepsilon - y_\varepsilon| + |t_\varepsilon - s_\varepsilon| \le C\varepsilon \qquad \text{and} \qquad |x_\varepsilon| + |y_\varepsilon| \le \frac{C}{\sqrt{\varepsilon}}.$$

Step 2. We use $\Phi(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon) \geq \Phi(x_\varepsilon, x_\varepsilon, t_\varepsilon, t_\varepsilon)$ to imply that

$$\begin{aligned}\frac{|x_\varepsilon - y_\varepsilon|^2 + |s_\varepsilon - t_\varepsilon|^2}{\varepsilon^2} &\leq v(x_\varepsilon, t_\varepsilon) - v(y_\varepsilon, s_\varepsilon) + \varepsilon\Big(|x_\varepsilon|^2 - |y_\varepsilon|^2\Big) + \lambda(t_\varepsilon - s_\varepsilon) \\ &\leq v(x_\varepsilon, t_\varepsilon) - v(y_\varepsilon, s_\varepsilon) + \varepsilon\frac{C}{\sqrt{\varepsilon}}C\varepsilon + C\varepsilon,\end{aligned}$$

which, together with the uniform continuity of v, yields further that

$$\lim_{\varepsilon\to 0} \frac{|x_\varepsilon - y_\varepsilon|^2 + |s_\varepsilon - t_\varepsilon|^2}{\varepsilon^2} = 0.$$

Step 3. Next, we claim that there exists a constant $\mu > 0$ independent of ε such that $t_\varepsilon, s_\varepsilon > \mu > 0$ for all $\varepsilon > 0$. It is important to have both $t_\varepsilon, s_\varepsilon$ bounded away from 0 in order to apply viscosity sub/supersolution tests.

To prove this claim, we use the uniform continuity of u, v and observe

$$\begin{aligned}\frac{\sigma}{2} &< u(x_\varepsilon, t_\varepsilon) - v(y_\varepsilon, s_\varepsilon) = \underbrace{u(x_\varepsilon, t_\varepsilon) - u(x_\varepsilon, 0)}_{\omega(t_\varepsilon)} + \underbrace{u(x_\varepsilon, 0) - v(x_\varepsilon, 0)}_{\leq 0 \text{ (by initial condition)}} \\ &\quad + \underbrace{v(x_\varepsilon, 0) - v(x_\varepsilon, t_\varepsilon)}_{\omega(t_\varepsilon)} + v(x_\varepsilon, t_\varepsilon) - v(y_\varepsilon, s_\varepsilon) \\ &\leq \omega(t_\varepsilon) + \omega(|x_\varepsilon - y_\varepsilon| + |t_\varepsilon - s_\varepsilon|),\end{aligned}$$

where $\omega(\cdot)$ is a modulus of continuity; that is, $\lim_{r\to 0}\omega(r) = 0$. Thus there exists $\mu > 0$ independent of ε such that $t_\varepsilon > \mu > 0$. By a similar argument, we also have $s_\varepsilon > \mu > 0$ for all $\varepsilon > 0$.

Step 4. The map $(x, t) \mapsto \Phi(x, y_\varepsilon, t, s_\varepsilon)$ has a max at $(x_\varepsilon, t_\varepsilon)$, and thus,

$$(x, t) \longmapsto u(x, t) - \underbrace{\left[\frac{|x - y_\varepsilon|^2 + |t - s_\varepsilon|^2}{\varepsilon^2} + \varepsilon|x|^2 + \lambda t\right]}_{\varphi(x,t)}$$

has a max at $(x_\varepsilon, t_\varepsilon)$. Since u is a viscosity subsolution to (1.22), the viscosity subsolution test gives

$$\frac{2(t_\varepsilon - s_\varepsilon)}{\varepsilon^2} + \lambda + H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) \leq 0.$$

Step 5. The map $(y, s) \mapsto \Phi(x_\varepsilon, y, t_\varepsilon, s)$ has a max at $(y_\varepsilon, s_\varepsilon)$; thus,

$$(y, s) \longmapsto v(y, s) - \underbrace{\left[-\frac{|x_\varepsilon - y|^2 + |t_\varepsilon - s|^2}{\varepsilon^2} - \varepsilon|y|^2 - \lambda s\right]}_{\psi(y,s)}$$

has a min at $(y_\varepsilon, s_\varepsilon)$. The viscosity supersolution test yields

$$\frac{2(t_\varepsilon - s_\varepsilon)}{\varepsilon^2} - \lambda + H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) \geq 0.$$

Step 6. We combine the inequalities in Step 4 and Step 5 to obtain

$$2\lambda \le H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right).$$

Using the Lipschitz assumption (1.13) on H, we have

$$H\left(y_\varepsilon, \frac{2\,(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) - H\left(y_\varepsilon, \frac{2\,(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \le 2C\varepsilon|y_\varepsilon|,$$

$$H\left(y_\varepsilon, \frac{2\,(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(x_\varepsilon, \frac{2\,(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \le C|x_\varepsilon - y_\varepsilon|\left(1 + \frac{2|x_\varepsilon - y_\varepsilon|}{\varepsilon^2}\right),$$

$$H\left(x_\varepsilon, \frac{2\,(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(x_\varepsilon, \frac{2\,(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) \le 2C\varepsilon|x_\varepsilon|.$$

Put all of the above inequalities in Step 6 together to imply

$$2\lambda \le 2C\varepsilon\Big(|x_\varepsilon| + |y_\varepsilon|\Big) + C|x_\varepsilon - y_\varepsilon| + \frac{2C|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2}.$$

Let $\varepsilon \to 0$ in the above to get a contradiction. The proof is complete. $\square$

Corollary 1.24 (Uniqueness of viscosity solution of Cauchy problem (1.22))**.** *Assume* (1.13). *If u and v are viscosity solutions of* (1.22), *then $u \equiv v$ in* $\mathbb{R}^n \times (0, \infty)$.

Proof. The proof follows immediately from the comparison principle in Theorem 1.23. $\square$

1.6.1. Problems.

Exercise 10. Let u, φ be two given continuous functions on $\mathbb{R}^n \times [0, T]$ for some $T > 0$ such that $u - \varphi$ has a strict max over $\mathbb{R}^n \times [0, T]$ at (x_0, T). For each $\varepsilon > 0$, let $\varphi_\varepsilon(x, t) = \varphi(x, t) + \frac{\varepsilon}{T-t}$ for all $(x, t) \in \mathbb{R}^n \times [0, T)$. Show that for $\varepsilon > 0$ small enough, $u - \varphi_\varepsilon$ has a local max at $(x_\varepsilon, t_\varepsilon) \in \mathbb{R}^n \times (0, T)$, and $(x_\varepsilon, t_\varepsilon) \to (x_0, T)$ up to passing to a subsequence.

Exercise 11. Let $H = H(x, p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a Hamiltonian satisfying that there exists $C > 0$ such that, for all $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x,p) - H(x,q)| & \le & C|p - q|, \\ |H(x,p) - H(y,q)| & \le & C(1 + |p|)|x - y|. \end{cases}$$

For $i = 1, 2$, let u^i be the viscosity solution to

$$\begin{cases} u^i_t + H(x, Du^i) & = & 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ u^i(x, 0) & = & g^i(x) & \text{on } \mathbb{R}^n, \end{cases} \tag{1.24}$$

where $g^i \in \mathrm{BUC}(\mathbb{R}^n)$ is given. Use the comparison principle for (1.24) to show the following L^∞ contraction property: For any $t \ge 0$,

$$\sup_{x \in \mathbb{R}^n} |u^1(x, t) - u^2(x, t)| \le \sup_{x \in \mathbb{R}^n} |g^1(x) - g^2(x)|.$$

1.7. Introduction to the classical Bernstein method

For $\varepsilon > 0$, consider the viscous Hamilton–Jacobi equation

$$\begin{cases} u^\varepsilon_t + H(x, Du^\varepsilon) = \varepsilon\Delta u^\varepsilon & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{1.25}$$

In this section, we introduce the classical Bernstein method to obtain a priori estimates for u^ε. Our aim is to get that $\|u^\varepsilon_t\|_{L^\infty} + \|Du^\varepsilon\|_{L^\infty} \le C$ where $C > 0$ is independent of $\varepsilon \in (0,1)$. We use the assumptions

$$u_0(x) \in C^2(\mathbb{R}^n) \quad \text{and} \quad \|u_0\|_{C^2(\mathbb{R}^n)} < \infty \tag{1.26}$$

and

$$\begin{cases} H \in C^2(\mathbb{R}^n \times \mathbb{R}^n), \\ H, D_pH \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for each } R > 0, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} \left(\frac{1}{2}H(x,p)^2 + D_xH(x,p)\cdot p\right) = +\infty. \end{cases} \tag{1.27}$$

By classical results (see [**AC78**, **Fri74**], [**Lio82**, Appendix] and the references therein), under (1.26)–(1.27), (1.25) has a unique solution u^ε which is smooth enough and its gradient is bounded, but of course, this bound might depend on ε. What is important in the following theorem is that we obtain a gradient bound for u^ε that is independent of $\varepsilon \in (0,1)$.

Theorem 1.25. *Assume* (1.26)–(1.27). *For each* $\varepsilon \in (0,1)$, *let* u^ε *be the unique solution to* (1.25). *Then, there exists a constant* $C > 0$ *independent of* $\varepsilon \in (0,1)$ *such that, for all* $(x,t) \in \mathbb{R}^n \times [0,\infty)$,

$$|u^\varepsilon_t(x,t)| + |Du^\varepsilon(x,t)| \le C.$$

Proof. We divide the proof into two steps as follows.

Step 1. We first obtain the boundedness of u^ε_t. Differentiate (1.25) in time,

$$(u^\varepsilon_t)_t + D_pH(x,Du^\varepsilon)\cdot Du^\varepsilon_t = \varepsilon\Delta u^\varepsilon_t \quad \Longrightarrow \quad \varphi_t + D_pH(x,Du^\varepsilon)\cdot D\varphi = \varepsilon\Delta\varphi,$$

where $\varphi = u^\varepsilon_t$. This is a linear parabolic equation for φ; thus, by the comparison principle for parabolic equations, we have for all $(x,t) \in \mathbb{R}^n \times [0,\infty)$,

$$\inf_{x\in\mathbb{R}^n} \varphi(x,0) \le \varphi(x,t) \le \sup_{x\in\mathbb{R}^n} \varphi(x,0).$$

Thus, we only need to bound $\varphi(\cdot,0) = u^\varepsilon_t(\cdot,0)$. We build barriers to do this as follows. For $C > 0$ large enough, set

$$\psi^\pm(x,t) = u_0(x) \pm Ct \qquad \text{for all } (x,t) \in \mathbb{R}^n \times [0,\infty).$$

Since $\|u_0\|_{C^2(\mathbb{R}^n)} < \infty$, we can find $C_0 > 0$ such that, for all $\varepsilon \in (0,1)$,

$$|H(x,Du_0) - \varepsilon\Delta u_0| \le |H(x,Du_0)| + |\Delta u_0| \le C_0.$$

Then, for $C > C_0$ and $\varepsilon \in (0,1)$,

$$\psi_t^\pm + H(x, D\psi^\pm) - \varepsilon\Delta\psi^\pm = \pm C + H(x, Du_0) - \varepsilon\Delta u_0 \gtrless 0 \quad \text{on } \mathbb{R}^n \times [0,\infty).$$

We conclude that $\psi^\pm$ are a supersolution and a subsolution to (1.25), respectively. Therefore, $\psi^- \le u^\varepsilon \le \psi^+$, which confirms that $|u_t^\varepsilon(x,0)| \le C$ for all $x \in \mathbb{R}^n$.

Step 2. Next, we show the boundedness of Du^ε, which is independent of ε. Differentiate (1.25) in x_k, multiply the result by $u^\varepsilon_{x_k}$, then sum them up over $k = 1, 2, \ldots, n$ to obtain

$$(1.28)\quad \frac{d}{dt}\left(\frac{1}{2}\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2\right) + D_x H(x, Du^\varepsilon)\cdot Du^\varepsilon + D_p H(x, Du^\varepsilon)\cdot\left(\sum_{k=1}^n Du^\varepsilon_{x_k} u^\varepsilon_{x_k}\right) = \varepsilon\sum_{k=1}^n \Delta u^\varepsilon_{x_k} u^\varepsilon_{x_k}.$$

Letting $\psi = \frac{1}{2}\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2 = \frac{1}{2}|Du^\varepsilon|^2$, we have

$$\sum_{k=1}^n Du^\varepsilon_{x_k} u^\varepsilon_{x_k} = D\left(\frac{1}{2}\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2\right) = D\psi, \qquad \sum_{k=1}^n \Delta u^\varepsilon_{x_k} u^\varepsilon_{x_k} = \Delta\psi - |D^2 u^\varepsilon|^2.$$

Thus, (1.28) becomes

$$\begin{aligned}\psi_t + D_x H(x, Du^\varepsilon)\cdot Du^\varepsilon + D_p H(x, Du^\varepsilon)\cdot D\psi &= \varepsilon\Delta\psi - \varepsilon|D^2u^\varepsilon|^2 \\ &\le \varepsilon\Delta\psi - \varepsilon\frac{(\Delta u^\varepsilon)^2}{n}.\end{aligned}$$

For each $\varepsilon < \frac{1}{n}$, we have $\frac{\varepsilon}{n} > \varepsilon^2$. Combine this with $|u_t^\varepsilon| \le C$ to get

$$\begin{aligned}\psi_t + D_x H(x, Du^\varepsilon)\cdot Du^\varepsilon + D_p H(x, Du^\varepsilon)\cdot D\psi &\le \varepsilon\Delta\psi - \Big(\varepsilon\Delta u^\varepsilon\Big)^2 \\ &= \varepsilon\Delta\psi - (u_t^\varepsilon + H(x, Du^\varepsilon))^2 \\ &\le \varepsilon\Delta\psi - \frac{1}{2}H(x, Du^\varepsilon)^2 + C.\end{aligned}$$

Therefore,

$$(1.29)\quad \Big(\psi_t + D_p H(x, Du^\varepsilon)\cdot D\psi - \varepsilon\Delta\psi\Big) + \left(\frac{1}{2}H(x, Du^\varepsilon)^2 + D_x H(x, Du^\varepsilon)\cdot Du^\varepsilon - C\right) \le 0.$$

Fix any $T > 0$. Assume that

$$\max_{\mathbb{R}^n\times[0,T]} \psi(x,t) = \psi(x_0, t_0)$$

for some $(x_0,t_0) \in \mathbb{R}^n \times [0,T]$. If $t_0 = 0$, then $\|Du^\varepsilon\|_{L^\infty} \le \|Du_0\|_{L^\infty} \le C$, and the proof is complete. If $t_0 > 0$, then by the usual maximum principle, we have

$$D\psi(x_0,t_0) = 0, \qquad \psi_t(x_0,t_0) \ge 0, \qquad \text{and} \qquad \Delta\psi(x_0,t_0) \le 0.$$

Using these facts in (1.29) evaluated at (x_0,t_0), we obtain

$$\underbrace{\Big(\psi_t + D_pH(x,Du^\varepsilon)\cdot D\psi - \varepsilon\Delta\psi\Big)}_{\ge 0} + \left(\frac{1}{2}H(x,Du^\varepsilon)^2 + D_xH(x,Du^\varepsilon)\cdot Du^\varepsilon - C\right) \le 0.$$

Therefore, at (x_0,t_0),

$$\frac{1}{2}H(x,Du^\varepsilon)^2 + D_xH(x,Du^\varepsilon)\cdot Du^\varepsilon \le C \quad \Longrightarrow \quad |Du^\varepsilon(x_0,t_0)| \le C,$$

in light of assumption (1.27).

Thus, we get the existence of a constant $C > 0$ independent of $\varepsilon \in (0,1)$ so that

$$\|u_t^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} \le C. \qquad \square$$

Remark 1.26. An important observation in the above proof is that as H is independent of time, $\varphi = u_t^\varepsilon$ solves the linearized equation, which is a nice linear parabolic equation. Therefore, boundedness of u_t^ε follows rather straightforwardly. If H is time dependent, then one needs to be careful in getting the bound for u_t^ε (for example, one has to have good control on H_t).

Remark 1.27. In the above proof, for each $\varepsilon \in (0,1)$ fixed, surely u_t^ε, Du^ε, and ψ are bounded, but in general, such a bound might depend on ε. The key point of the proof is that we obtain a bound on u_t^ε and Du^ε that is independent of $\varepsilon \in (0,1)$. In the last part of the proof, it might be the case that (x_0,t_0) does not exist. To overcome this difficulty, we consider, for each $\delta > 0$, the maximum point on $\mathbb{R}^n \times [0,T]$ of

$$(x,t) \mapsto \psi^\delta(x,t) = \Big(\psi(x,t) - \delta(1+|x|^2)^{1/2}\Big)$$

and use the maximum principle for ψ^δ at this point. Then, we let $\delta \to 0$ to obtain the desired result.

1.7.1. Problems.

Exercise 12. Write down a detailed proof of the claim in Remark 1.27.

Exercise 13. Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a Hamiltonian satisfying

$$\begin{cases} H \in C^2(\mathbb{R}^n \times \mathbb{R}^n), \\ H, D_pH \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for each } R > 0, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} \left(\frac{1}{2}H(x,p)^2 + D_xH(x,p)\cdot p\right) = +\infty. \end{cases} \tag{1.30}$$

For $\varepsilon \in (0,1)$, consider the static viscous Hamilton–Jacobi equation

$$u^\varepsilon + H\left(x, Du^\varepsilon\right) = \varepsilon \Delta u^\varepsilon \qquad \text{in } \mathbb{R}^n. \tag{1.31}$$

Let u^ε be the unique bounded, smooth solution to the above. Use the Bernstein method to show that there exists a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that $\|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n)} \le C$.

Letting $\varepsilon \to 0$ in the above and using the Arzelà–Ascoli theorem, we obtain the existence of a Lipschitz viscosity solution to the corresponding static problem.

Corollary 1.28. *Assume* (1.30). *Then, the static problem* (1.12) *has a Lipschitz viscosity solution* u.

1.8. Introduction to Perron's method

1.8.1. Perron's method for static problems. Recall the usual static problem

$$u + H(x, Du) = 0 \qquad \text{in } \mathbb{R}^n. \tag{1.32}$$

One simple observation we have is that if u_1, u_2 are subsolution of (1.32), then so is $\max\{u_1, u_2\}$. We generalized this into the following result.

Lemma 1.29. *Assume $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$. Let $\{u_i\}_{i\in I}$ be a family of (continuous) subsolutions to* (1.32). *Let*

$$u(x) = \sup_{i\in I} u_i(x) \qquad \text{for all } x \in \mathbb{R}^n.$$

Assume that u is finite and continuous. Then, u is also a viscosity subsolution to (1.32).

It is worth noting here that the assumption that u is finite is natural, but the assumption that u is continuous is not. We actually do not need it, but we put it here for simplicity. In general, we only expect that u is bounded, and in fact, a definition for viscosity subsolutions to (1.32) can be given for upper semicontinuous functions in $\mathbb{R}^n$, $\mathrm{USC}(\mathbb{R}^n)$, naturally. The result of Lemma 1.29 still holds true for u under the new definition; that is, u^*, its upper semicontinuous envelope, is a viscosity subsolution to (1.32).

Proof. Take $\varphi \in C^1(\mathbb{R}^n)$ such that $u - \varphi$ has a max at x_0 over $\overline{B}_r(x_0)$ and $u(x_0) = \varphi(x_0)$. Let $\psi(x) = \varphi(x) + |x - x_0|^2$. Then $u - \psi$ has a strict max over $\overline{B}_r(x_0)$. By definition, we can find a sequence (re-indexed) $\{u_n\}_{n\in\mathbb{N}} \subset \{u_i\}_{i\in I}$ such that $0 \le u(x_0) - u_n(x_0) \le \frac{1}{n}$ for all $n \in \mathbb{N}$. For all $x \in \overline{B}_r(x_0)$, we have

$$u_n(x) - \psi(x) \le u(x) - \varphi(x) - |x - x_0|^2 \le -|x - x_0|^2.$$

By compactness, we can assume $u_n - \psi$ has a max over $\overline{B}_r(x_0)$ at x_n, and thus,

$$\begin{aligned} u_n(x_0) - \varphi(x_0) &\le u_n(x_n) - \varphi(x_n) - |x_n - x_0|^2 \\ &\le \ u(x_n) \ - \varphi(x_n) - |x_n - x_0|^2 \le -|x_n - x_0|^2. \end{aligned}$$

From the above, we obtain $|x_n - x_0|^2 \le \frac{1}{n}$. Let $n \to \infty$ to yield that $x_n \to x_0$, and therefore, x_n is actually a local max of $u_n - \psi$ over $\mathbb{R}^n$ for n sufficiently large. As a consequence, $u_n(x_n) - \varphi(x_n) \to 0$ as $n \to \infty$. For n large, as u_n is a subsolution of (1.32), the subsolution test gives

$$u_n(x_n) + H(x_n, \varphi(x_n)) \le 0 \qquad \Longrightarrow \qquad \varphi(x_0) + H(x_0, D\varphi(x_0)) \le 0$$

by letting $n \to \infty$. Thus, u is a viscosity subsolution of (1.32). ☐

The Perron method in the theory of viscosity solutions was first introduced by Ishii [**Ish87**]. In the following, we give a variant of Ishii's argument in [**Ish87**]. Based on a coercivity assumption, we construct directly a Lipschitz viscosity solution, which was not written down explicitly by Ishii. Here is the assumption on H that we need:

$$\text{(1.33)} \qquad \begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) & \text{for all } R > 0, \\ \lim\limits_{|p|\to\infty} \inf\limits_{x\in\mathbb{R}^n} H(x,p) = +\infty. \end{cases}$$

Under this assumption, set $C_0 = \sup_{x\in\mathbb{R}^n} |H(x,0)|$. It is clear that C_0 and $-C_0$ are a viscosity supersolution and subsolution to (1.32), respectively. By coercivity of H, we are able to find $C_1 > 0$ such that

$$H(x,p) \le C_0 + 1 \quad \text{for some } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n \quad \Longrightarrow \quad |p| \le C_1.$$

Here is our main result in this section.

Theorem 1.30 (Perron's method for (1.32))**.** *Assume* (1.33)*. Define*

$$u(x) = \sup\Big\{v(x) \,:\, -C_0 \le v \le C_0,\ \|Dv\|_{L^\infty(\mathbb{R}^n)} \le C_1, \text{ and } v \text{ is a viscosity subsolution to } (1.32)\Big\}. \tag{1.34}$$

Then, u is a Lipschitz viscosity solution to (1.32)*.*

Proof. Of course, u is well-defined as $v \equiv -C_0$ itself is an admissible subsolution in the above formula. Furthermore, it is clear that u is Lipschitz in $\mathbb{R}^n$ and $\|Du\|_{L^\infty(\mathbb{R}^n)} \le C_1$. By the stability of viscosity subsolutions (Lemma 1.29), we imply first that u is a viscosity subsolution to (1.32).

Hence, we only need to show that u is a viscosity supersolution to (1.32). Assume by contradiction that this is not the case. Then, there exist a smooth test function $\phi \in C^1(\mathbb{R}^n)$ and a point $x_0 \in \mathbb{R}^n$ such that

$$\begin{cases} u(x_0) = \phi(x_0),\ u(x) > \phi(x) & \text{for all } x \in \mathbb{R}^n \setminus \{x_0\},\\ u(x_0) + H(x_0, D\phi(x_0)) = \phi(x_0) + H(x_0, D\phi(x_0)) < 0. \end{cases}$$

There are two cases to be considered here. The first case is when $u(x_0) = C_0$. This means that ϕ touches constant function C_0, a supersolution to (1.32), from below at x_0. By the definition of viscosity supersolutions,

$$\phi(x_0) + H(x_0, D\phi(x_0)) \ge 0,$$

which implies a contradiction immediately.

The second case is when $u(x_0) < C_0$. There exist $r, \varepsilon > 0$ sufficiently small such that

$$\begin{cases} u(x) < C_0 - \varepsilon & \text{for all } x \in B(x_0, r),\\ \phi(x) < u(x) - \varepsilon & \text{for all } x \in \partial B(x_0, r),\\ \phi(x) + H(x, D\phi(x)) < -2\varepsilon & \text{for all } x \in B(x_0, r),\\ |D\phi(x)| \le C_1 & \text{for all } x \in B(x_0, r). \end{cases}$$

Now, set

$$\overline{u}(x) = \begin{cases} \max\{u(x), \phi(x) + \varepsilon\} & \text{for all } x \in B(x_0, r),\\ u(x) & \text{for all } x \in \mathbb{R}^n \setminus B(x_0, r). \end{cases}$$

It is quite clear that $\overline{u}$ is a viscosity subsolution to (1.32) and $\|D\overline{u}\|_{L^\infty(\mathbb{R}^n)} \le C_1$. This again leads to a contradiction. The proof is complete. □

As included in the proof, we obtain immediately the existence of a Lipschitz viscosity solution u to (1.32) under assumption (1.33). In fact, by Remark 1.21, we imply further that, under (1.33), u is actually the unique viscosity solution to (1.32). This is quite interesting, and we completely bypass the need of the vanishing viscosity method to obtain a Lipschitz solution here. Of course, when we do not have coercivity, we would not be able to impose the Lipschitz constraint directly in the definition of u, and we will see that this is indeed the case for the Cauchy problem in the next section. Let us record what was discussed as a theorem here for later use.

Theorem 1.31. *Assume* (1.33)*. Let u be defined as in Theorem* 1.30*. Then, u is the unique Lipschitz viscosity solution to* (1.32)*.*

Let us now discuss further about solutions to (1.32) under (1.33). We show in the following that if we have a bounded uniformly continuous solution, then it is indeed Lipschitz.

Lemma 1.32. *Assume* (1.33)*. Let $u \in \mathrm{BUC}(\mathbb{R}^n)$ be a viscosity solution to* (1.32)*. Then, u is Lipschitz in $\mathbb{R}^n$.*

Proof. As $u \in \mathrm{BUC}(\mathbb{R}^n)$, it is not hard to show that $-C_0 \le u \le C_0$ (this is expressed in Exercise 14). By coercivity and the viscosity subsolution test, we get

$$|p| \le C_1 \quad \text{for all } x \in \mathbb{R}^n, p \in Du^+(x).$$

We now show that u is Lipschitz with Lipschitz constant at most C_1. Given $\varepsilon > 0$ and $y \in \mathbb{R}^n$, consider $\varphi(x) = (C_1 + \varepsilon)|x - y|$. Of course $\varphi \in C^\infty(\mathbb{R}^n \backslash \{y\})$. Since u is bounded, we have $u - \varphi$ has a max at some $x_\varepsilon \in \mathbb{R}^n$. If $x_\varepsilon \neq y$, then

$$D\varphi(x_\varepsilon) = (C_1 + \varepsilon)\left(\frac{x_\varepsilon - y}{|x_\varepsilon - y|}\right) \in D^+u(x_\varepsilon) \quad \Longrightarrow \quad |D\varphi(x_\varepsilon)| = C_1 + \varepsilon \le C_1,$$

which is a contradiction. Thus $x_\varepsilon = y$, which means that for all $x \in \mathbb{R}^n$,

$$u(x) - (C_1 + \varepsilon)|x - y| \le u(y) \qquad \Longrightarrow \qquad u(x) - u(y) \le (C_1 + \varepsilon)|x - y|.$$

By a symmetric argument, we obtain $|u(x) - u(y)| \le (C_1 + \varepsilon)|x - y|$ for all $x, y \in \mathbb{R}^n$. Finally, let $\varepsilon \to 0$ to imply our claim. ☐

In the above proof, there is one interesting point that if $u \in \mathrm{BUC}(\mathbb{R}^n)$ satisfies

$$|p| \le C_1 \quad \text{for all } x \in \mathbb{R}^n, p \in Du^+(x),$$

then u is Lipschitz with Lipschitz constant at most C_1. It is worth noting that we do not need boundedness of u to have this result.

Lemma 1.33. *Let $u \in C(\mathbb{R}^n)$ be such that, for all $p \in D^+u(x)$ for all $x \in \mathbb{R}^n$, we have $|p| \le C_1$. Then, u is Lipschitz with Lipschitz constant at most C_1.*

The proof of this is left as an exercise for the interested readers.

1.8.2. Problems.

Exercise 14. Assume (1.33). Denote $C_0 = \sup_{x\in\mathbb{R}^n} |H(x,0)|$. Let $u \in \mathrm{BUC}(\mathbb{R}^n)$ be a solution to (1.32). Show that

$$-C_0 \le u \le C_0.$$

Exercise 15. Prove Lemma 1.33.

1.8.3. Perron's method for Cauchy problems. Let us now focus on our usual Cauchy problem

$$\begin{cases} u_t + H(x,Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0 \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{1.35}$$

In order to apply the Perron method, we need the following assumptions:

- For H, we assume that it satisfies (1.33); that is,

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \qquad \text{for all } R>0, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} H(x,p) = +\infty. \end{cases}$$

- For initial data u_0, we assume

$$u_0 \in C^1(\mathbb{R}^n) \quad \text{and} \quad \|u_0\|_{C^1(\mathbb{R}^n)} < \infty. \tag{1.36}$$

By assumptions (1.33) and (1.36), we have $\|Du_0\|_{L^\infty(\mathbb{R}^n)} < \infty$ and $|H(x,Du_0(x))| \le C_0$ for all $x \in \mathbb{R}^n$ for some constant $C_0 > 0$. In particular:

- $\varphi_1(x,t) = u_0(x) - C_0 t$ is a classical subsolution to (1.35).
- $\varphi_2(x,t) = u_0(x) + C_0 t$ is a classical supersolution to (1.35).

Theorem 1.34 (Perron's method for (1.35))**.** *Assume* (1.33) *and* (1.36). *Denote, for* $(x,t) \in \mathbb{R}^n \times [0,\infty)$,

$$u(x,t) = \sup\left\{ \varphi(x,t) \in C(\mathbb{R}^n \times (0,\infty)) \ : \ \begin{cases} \varphi_1 \le \varphi \le \varphi_2, \\ \varphi \textit{ is a subsolution to } (1.35) \end{cases} \right\}.$$

Then, u is a viscosity solution of (1.35).

For the Cauchy problem, as there is the time variable t, we should think of the "overall Hamiltonian" as

$$\begin{aligned} F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} &\to \mathbb{R} \\ (x,p,p_{n+1}) &\mapsto F(x,p,p_{n+1}) = p_{n+1} + H(x,p). \end{aligned}$$

Here, p_{n+1} represents u_t. It is clear that F is not coercive in $p' = (p,p_{n+1})$, and hence, we cannot impose the a priori Lipschitz condition in the definition of u as in Theorem 1.30. In fact, in this case for (1.35), u defined above might be discontinuous. Further discussion on this and a priori estimates for u will be done in the next section.

Proof. For simplicity, we assume that u is continuous.

First of all, it is clear that u is a viscosity subsolution of (1.35). Now we prove that u is a viscosity supersolution of (1.35). Let $\psi \in C^1(\mathbb{R}^n \times (0,\infty))$ be a test function such that $u(x,t) - \psi(x,t)$ has a strict min at $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$ and $u(x_0,t_0) = \psi(x_0,t_0)$. We need to prove that

$$\psi_t(x_0,t_0) + H\big(x_0, D\psi(x_0,t_0)\big) \geq 0. \tag{1.37}$$

There are two cases to be considered here. The first case is when $\psi(x_0,t_0) = u(x_0,t_0) = \varphi_2(x_0,t_0)$. In this case, ψ touches φ_2 from below at (x_0,t_0). The viscosity supersolution test confirms that (1.37) is true.

The second case is when $\psi(x_0,t_0) = u(x_0,t_0) < \varphi_2(x_0,t_0)$. Assume by contradiction that (1.37) does not hold; that is,

$$\psi_t(x_0,t_0) + H\big(x_0, D\psi(x_0,t_0)\big) < 0.$$

We can find $\varepsilon, r > 0$ sufficiently small such that

$$\begin{cases} u(x,t) < \varphi_2(x,t) - \varepsilon & (x,t) \in \overline{B}\big(x_0,r\big) \times \big[t_0 - r, t_0 + r\big], \\ \psi(x,t) < u(x,t) - \varepsilon & (x,t) \in \partial\left(B\big(x_0,r\big) \times \big[t_0 - r, t_0 + r\big]\right), \\ \psi_t(x,t) + H(x, D\psi(x,t)) < -\varepsilon & (x,t) \in \overline{B}\big(x_0,r\big) \times \big[t_0 - r, t_0 + r\big]. \end{cases}$$

Now, we define

$$\tilde{u}(x,t) = \begin{cases} \max\{u(x,t), \psi(x,t) + \varepsilon\} & \text{if } (x,t) \in \overline{B}\big(x_0,r\big) \times \big[t_0 - r, t_0 + r\big], \\ u(x,t) & \text{if } (x,t) \notin \overline{B}\big(x_0,r\big) \times \big[t_0 - r, t_0 + r\big]. \end{cases}$$

It is not hard to check that $\tilde{u}$ is a viscosity subsolution to (1.35). This gives a contradiction as $\tilde{u}(x_0,t_0) > u(x_0,t_0)$. The proof is complete. $\square$

Remark 1.35. Let us emphasize again that u defined in Theorem 1.34 might not be continuous. Besides (1.33) and (1.36), if we require the additional condition (1.13), then we have the comparison principle to (1.35) and, hence, uniqueness of solutions to (1.35). Then, as u^* is a subsolution and u_* is a supersolution to (1.35), respectively, we get $u^* \leq u_*$. Therefore, $u = u^* = u_*$, which means that u is continuous.

In order to obtain Lipschitz bounds for u, we need a more complicated argument, since in this case we need to prove u_t is bounded first.

1.9. Lipschitz estimates for Cauchy problems using Perron's method

Let us continue focusing on the usual Cauchy problem

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{1.38}$$

We assume here (1.33), (1.36), and (1.13) to get Lipschitz estimates for the unique viscosity solution u to (1.38). Let us recall these assumptions here for clarity. Condition (1.13) is a structural one to get uniqueness of solutions

$$\begin{cases} |H(x,p) - H(y,p)| \le C(1+|p|)|x-y|, \\ |H(x,p) - H(x,q)| \le C|p-q|. \end{cases}$$

And conditions (1.33), (1.36) are for the use of Perron's method

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) & \text{for any } R > 0, \\ \lim_{|p|\to\infty} \left(\inf_{x\in\mathbb{R}^n} H(x,p) \right) = +\infty, \\ u_0 \in C^1(\mathbb{R}^n) \quad \text{and} \quad \|u_0\|_{C^1(\mathbb{R}^n)} < \infty. \end{cases}$$

Theorem 1.36. *Assume* (1.33), (1.36), *and* (1.13). *Then,* (1.38) *has a unique viscosity solution u, which is Lipschitz in both space and time. In particular, there exists a constant $C > 0$ such that*

$$|u_t(x,t)| + |Du(x,t)| \le C \qquad \textit{a.e. on } \mathbb{R}^n \times [0,\infty). \tag{1.39}$$

Proof. We show that u is Lipschitz in time; then coercivity of H implies that u is Lipschitz in space right away.

Step 1. We first show $t \mapsto u(x,t)$ is Lipschitz at $t = 0$. By Theorem 1.34, we have

$$u_0(x) - C_0 t \le u(x,t) \le u_0(x) - C_0 t \qquad \text{for all } (x,t) \in \mathbb{R}^n \times [0,\infty).$$

This implies that, for all $x \in \mathbb{R}^n$,

$$-C_0 \le \frac{u(x,t) - u(x,0)}{t} \le C_0 \qquad \Longrightarrow \qquad \sup_{t\ge 0} \left| \frac{u(x,t) - u(x,0)}{t} \right| \le C_0.$$

Step 2. We now show u is Lipschitz in time for all $t \ge 0$ with constant C_0. The key point here is that H is independent of t, which means that it is translation invariant in time. In particular, for fixed $s > 0$, $(x,t) \mapsto v(x,t) = u(x, s+t)$ is still a solution to (1.38) with different initial data $v_0(x) = v(x,0) = u(x,s)$ for $x \in \mathbb{R}^n$. As

$$v_0 - \|u_0 - v_0\|_{L^\infty(\mathbb{R}^n)} \le u_0 \le v_0 + \|u_0 - v_0\|_{L^\infty(\mathbb{R}^n)},$$

the usual comparison principle for (1.38) implies that

$$v(x,t) - \|u_0 - v_0\|_{L^\infty(\mathbb{R}^n)} \le u(x,t) \le v(x,t) + \|u_0 - v_0\|_{L^\infty(\mathbb{R}^n)}.$$

Thus, for all $(x,t) \in \mathbb{R}^n \times [0,\infty)$ and $s > 0$,

$$u(x,t+s) - \|u_0 - v_0\|_{L^\infty(\mathbb{R}^n)} \le u(x,t) \le u(x,t+s) + \|u_0 - v_0\|_{L^\infty(\mathbb{R}^n)},$$

which means

$$\left|\frac{u(x,t+s)-u(x,t)}{s}\right| \le \left\|\frac{u(\cdot,s)-u(\cdot,0)}{s}\right\|_{L^\infty(\mathbb{R}^n)} \le C_0,$$

thanks to Step 1. Thus, u is Lipschitz in time with constant C_0.

Step 3. Finally, we claim that u is Lipschitz in space. As its proof is rather standard, we omit it here and leave it as an exercise. ☐

Remark 1.37. We have the two following comments.

- It is crucial to use the point that H is time independent in the above proof. In fact, if H is time dependent, then Step 2 above completely breaks down. In such cases, in order to obtain Lipschitz estimates, one needs to do it in a very different way.
- Let us now assume only (1.33) and (1.36). We aim at finding a priori estimates to solution u of (1.38). By the above proof, we get first that $\|u_t\|_{L^\infty} \le C_0$, which yields $H(x,Du) \le C_0$. Thus, we are able to find $C > 0$ such that $\|u_t\|_{L^\infty} + \|Du\|_{L^\infty} \le C$. In particular, information about $H(x,p)$ for $|p| \ge C$ does not matter. Define a new Hamiltonian $\tilde{H}$ such that

$$\tilde{H}(x,p) = \begin{cases} H(x,p) & \text{for all } x \in \mathbb{R}^n,\ |p| \le C, \\ |p| & \text{for all } x \in \mathbb{R}^n,\ |p| \ge 2C, \end{cases}$$

and $\tilde{H}$ satisfies (1.33), (1.36), and (1.19). Recall that (1.19) is a replacement of (1.13). Then the Cauchy problem

$$\begin{cases} w_t(x,t) + \tilde{H}(x,Dw(x,t)) & = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty), \\ w(x,0) & = u_0(x) \quad \text{on } \mathbb{R}^n, \end{cases}$$

has a unique Lipschitz viscosity solution w, and $\|w_t\|_{L^\infty} + \|Dw\|_{L^\infty} \le C$. It is clear then that w is a Lipschitz viscosity solution to (1.38). Then, Remark 1.21 implies that $u = w$ is the unique Lipschitz viscosity solution to (1.38). This is an extremely important point as we are able to bypass the requirement of (1.13) (or (1.19)). Basically, we use a priori estimates to get gradient bounds on the solution first; then we eliminate (1.13) (or (1.19)) later. We record this important point in the following.

Theorem 1.38. *Assume* (1.33) *and* (1.36)*. Then,* (1.38) *has a unique viscosity solution u, which is Lipschitz in both space and time. In particular, there exists a constant $C > 0$ such that*

$$|u_t(x,t)| + |Du(x,t)| \le C \qquad \text{a.e. on } \mathbb{R}^n \times [0,\infty). \tag{1.40}$$

In fact, we only need to require that $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ in the above theorem.

1.9.1. Problems.

Exercise 16. Give a detailed proof of Step 3 in the proof of Theorem 1.36.

Exercise 17. Write down a proof of Theorem 1.38.

1.10. Finite speed of propagation for Cauchy problems

Our main focus in this section is still the usual Cauchy problem

$$u_t(x,t) + H(x, Du(x,t)) = 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty). \tag{1.41}$$

We do not impose yet the initial condition of (1.41). We assume (1.13), which is a structural condition to get uniqueness of solutions to (1.41). Let us recall it for clarity. There exists $C > 0$ such that, for all $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x,p) - H(y,p)| \le C(1+|p|)|x-y|, \\ |H(x,p) - H(x,q)| \le C|p-q|. \end{cases}$$

Here is the main result in this section on the finite speed of propagation for (1.41).

Theorem 1.39. *Assume* (1.13)*. Let u, v be a subsolution and a supersolution to* (1.41)*, respectively. Assume further that $u(x,0) \le v(x,0)$ for all $x \in B(0,R)$ for some given $R > 0$. Then,*

$$u(x,t) \le v(x,t) \qquad \text{for all } x \in B(0, R - Ct) \text{ and } t \le \frac{R}{C}.$$

To prove this theorem, we need the following preparation lemma.

Lemma 1.40. *Assume* (1.13)*. Let u, v be a subsolution and a supersolution to* (1.41)*, respectively. Let $w = u - v$. Then, w is a viscosity subsolution to*

$$w_t - C|Dw| = 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty). \tag{1.42}$$

Proof. Take a smooth test function φ such that $w - \varphi$ has a global strict maximum at $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$, $w(x_0,t_0) = \varphi(x_0,t_0)$, and $w - \varphi$ tends to $-\infty$ as $|x| \to \infty$ or $t \to \infty$. We consider the following auxiliary function $\Phi : \mathbb{R}^n \times \mathbb{R}^n \times [0,\infty) \times [0,\infty) \to \mathbb{R}$, where

$$\Phi(x,y,t,s) = u(x,t) - v(y,s) - \frac{|x-y|^2 + |t-s|^2}{\varepsilon^2} - \varphi(x,t).$$

It is clear that Φ achieves its maximum at some point $(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon)$ on $\mathbb{R}^n \times [0,\infty)^2$. By following the same arguments as in the proof of Theorem 1.23, we are able to obtain that $(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon) \to (x_0, x_0, t_0, t_0)$ as $\varepsilon \to 0$. Moreover,

$$\lim_{\varepsilon\to 0} \frac{|x_\varepsilon - y_\varepsilon|^2 + |t_\varepsilon - s_\varepsilon|^2}{\varepsilon^2} = 0. \tag{1.43}$$

By using the viscosity subsolution and supersolution tests as usual (same way as in the proof of Theorem 1.23), we get

$$\varphi_t(x_\varepsilon, t_\varepsilon) + \frac{2(t_\varepsilon - s_\varepsilon)}{\varepsilon^2} + H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + D\varphi(x_\varepsilon, t_\varepsilon)\right) \le 0$$

and

$$\frac{2(t_\varepsilon - s_\varepsilon)}{\varepsilon^2} + H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \ge 0.$$

Combining the two inequalities above implies

$$\varphi_t(x_\varepsilon, t_\varepsilon) + H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + D\varphi(x_\varepsilon, t_\varepsilon)\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \le 0.$$

We use (1.13) to deduce further that

$$\varphi_t(x_\varepsilon, t_\varepsilon) - C|D\varphi(x_\varepsilon, t_\varepsilon)| \le C|x_\varepsilon - y_\varepsilon| \left(1 + \frac{2|x_\varepsilon - y_\varepsilon|}{\varepsilon^2}\right).$$

Let $\varepsilon \to 0$ in the above and use (1.43) to conclude. □

To obtain Theorem 1.39, we now only need to show that $w(x,t) \le 0$ for all $x \in B(0, R - Ct)$ and $t \le \frac{R}{C}$.

Lemma 1.41. *Let w be a viscosity subsolution* (1.42). *Assume that $w(x,0) \le 0$ for all $x \in B(0,R)$ for some given $R > 0$. Then,*

$$w(x,t) \le 0 \qquad \text{for all } x \in B(0, R - Ct) \text{ and } t \le \frac{R}{C}.$$

Before giving a proof, let us mention here that (1.42) is in fact a first-order front propagation problem and is similar to what was discussed in Example 1.1. Another proof of this lemma can be found later in Section 2.5.5 of Chapter 2.

Proof. Let $T = \frac{R}{C}$, and let

$$M = \max_{\overline{B}(0,R)\times[0,T]} w.$$

We construct supersolutions to (1.42) and use the comparison principle to get the desired conclusion. For each $\varepsilon > 0$ sufficiently small, we design a smooth cut-off function $\xi_\varepsilon : \mathbb{R} \to \mathbb{R}$ such that ξ_ε is nondecreasing and

$$\begin{cases} \xi_\varepsilon(s) = 0 & \text{for } s \le R - \varepsilon, \\ \xi_\varepsilon(s) = M & \text{for } s \ge R. \end{cases}$$

Denote

$$\psi^\varepsilon(x,t) = \xi_\varepsilon(|x| + Ct) \qquad \text{for } x \in \mathbb{R}^n,\ 0 \le t < T_\varepsilon = \frac{R-\varepsilon}{C}.$$

We claim that ψ^ε is a classical solution to (1.42) in $\mathbb{R}^n \times (0, T_\varepsilon)$. Indeed, ψ^ε is smooth, and at $x = 0$,

$$\psi^\varepsilon_t(0,t) = 0,\ D\psi^\varepsilon(0,t) = 0 \qquad \text{for all } 0 \le t < T_\varepsilon = \frac{R-\varepsilon}{C},$$

so there is nothing to check here. For $x \neq 0$ and $t \in (0, T_\varepsilon)$, we compute

$$\psi^\varepsilon_t(x,t) = C\xi'_\varepsilon(|x| + Ct), \quad D\psi^\varepsilon(x,t) = \xi'_\varepsilon(|x| + Ct)\frac{x}{|x|},$$

which immediately gives that

$$\psi^\varepsilon_t(x,t) - C|D\psi^\varepsilon(x,t)| = C\xi'_\varepsilon(|x| + Ct) - C\xi'_\varepsilon(|x| + Ct) = 0.$$

Besides, from the definition of ψ^ε and ξ_ε,

$$w(x,t) \le M = \psi^\varepsilon(x,t) \qquad \text{for all } (x,t) \in \partial B(0,R) \times [0, T_\varepsilon].$$

By the comparison principle for (1.42), we get that $w \le \psi^\varepsilon$ on $\overline{B}(0,R) \times [0, T_\varepsilon]$. Let $\varepsilon \to 0$ to get the conclusion. □

We are now ready to prove the main theorem in this section.

Proof of Theorem 1.39. Let $w = u - v$. By using Lemmas 1.40 and 1.41, we immediately get the desired result. □

1.11. Rate of convergence of the vanishing viscosity process for static problems via the doubling variables method

Let us recall the vanishing viscosity procedure for the usual static problem

$$u(x) + H(x, Du(x)) = 0 \qquad \text{in } \mathbb{R}^n. \tag{1.44}$$

For each $\varepsilon > 0$, we consider

$$u^\varepsilon(x) + H(x, Du^\varepsilon(x)) = \varepsilon \Delta u^\varepsilon(x) \qquad \text{in } \mathbb{R}^n. \tag{1.45}$$

We assume that H satisfies (1.27); that is,

$$\begin{cases} H \in C^2(\mathbb{R}^n \times \mathbb{R}^n), \\ H, D_pH \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for each } R > 0, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} \left(\frac{1}{2}H(x,p)^2 + D_xH(x,p)\cdot p\right) = +\infty. \end{cases}$$

Under this assumption, we use the classical Bernstein method (same arguments as in Theorem 1.25) to obtain that (1.45) has a unique smooth solution u^ε. Moreover, there exists a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that

$$\|u^\varepsilon\|_{L^\infty(\mathbb{R}^n)} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n)} \leq C \qquad \text{for all } \varepsilon \in (0,1).$$

In light of this estimate, $\{u^\varepsilon\}_{\varepsilon\in(0,1)}$ is locally equicontinuous in $\mathbb{R}^n$. By the Arzelà–Ascoli theorem, for each sequence $\{\varepsilon_k\} \searrow 0$, there exists a subsequence $\{\varepsilon_{k_i}\} \searrow 0$ such that

$$u^{\varepsilon_{k_j}} \to u \qquad \text{locally uniformly in } \mathbb{R}^n \text{ as } j \to \infty,$$

for some u that satisfies $\|u\|_{L^\infty(\mathbb{R}^n)} + \|Du\|_{L^\infty(\mathbb{R}^n)} \leq C$. Thus, we deduce that u is the unique Lipschitz viscosity solution of (1.44). Because of the uniqueness of the limiting function u, we imply that $u^\varepsilon \to u$ locally uniformly as $\varepsilon \searrow 0$.

It is actually very important to understand more about this vanishing viscosity process. A pretty much open problem is to understand the gradient shock structures of u, the unique Lipschitz viscosity solution of (1.44). It is typically the case that u is Lipschitz, but not C^1, and the behaviors of the singularities of Du (e.g., the corners of the graph of u) are determined by the viscosity sub/supersolution tests. However, we do not have clear knowledge about these singularities in general, especially when H is not convex in p, at this moment. This topic should be one of the most important directions to study in the field in the future.

Another point, which is simpler, is to study the rate of convergence of $\{u^\varepsilon\}_{\varepsilon>0}$ to u as $\varepsilon \to 0$. There have been many interesting results in this direction, but still, the optimal rate for various cases is not yet known. Up to now, for the general cases, the best known convergence rate is $O(\varepsilon^{1/2})$.

Theorem 1.42. *Assume that H satisfies* (1.27). *Assume further that* $H \in \mathrm{Lip}(\mathbb{R}^n \times B(0,R))$ *for each $R > 0$. For each $\varepsilon \in (0,1)$, let u^ε be the unique smooth solution to* (1.45). *Let u be the unique Lipschitz viscosity solution of* (1.44). *Then, there exists a constant $C > 0$ independent of ε such that*

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n)} \leq C\sqrt{\varepsilon}.$$

This type of results with convergence rate $O(\varepsilon^{1/2})$ was first obtained by Fleming [**Fle64**] in the 1960s by using a differential game approach. Later on, within the framework of viscosity solutions, Crandall and Lions [**CL84**] proved Theorem 1.42 by using the doubling variables method. Of course, the approach of Crandall and Lions is quite general and can be adapted to many other situations. Another proof of Theorem 1.42 by using the nonlinear adjoint method was introduced by Evans [**Eva10a**] and Tran [**Tra11**].

We give here in this section a proof based on the ideas of Crandall and Lions [**CL84**]. The nonlinear adjoint method will be introduced in the next section.

Proof. By using the doubling variables method, consider the auxiliary function

$$\Phi^{\delta}(x,y) = u^{\varepsilon}(x) - u(y) - \frac{|x-y|^2}{2\alpha} - \delta\Big(\mu(x) + \mu(y)\Big),$$

where $\delta, \alpha > 0$ are to be chosen and $\mu \in C^2(\mathbb{R}^n, \mathbb{R})$ satisfies[1]

$$\begin{cases} \mu(0) = 0, \ \ \mu(x) \geq 0 & \text{for all } x \in \mathbb{R}^n, \\ \lim_{|x|\to\infty} \mu(x) = +\infty, \\ |D\mu(x)| + |D^2\mu(x)| \leq 1 & \text{in } \mathbb{R}^n. \end{cases}$$

Since u^{ε} and u are continuous and bounded, we can assume that

$$\max_{\mathbb{R}^n \times \mathbb{R}^n} \Phi^{\delta}(x,y) = \Phi^{\delta}(x_\delta, y_\delta),$$

for some $(x_\delta, y_\delta) \in \mathbb{R}^n \times \mathbb{R}^n$.

Step 1. Since $x \mapsto \Phi^{\delta}(x, y_\delta)$ has a max at x_δ, $x \mapsto u^{\varepsilon}(x) - \left[\frac{|x-y_\delta|^2}{2\alpha} + \delta\mu(x)\right]$ has a max at x_δ. Therefore,

(1.46)
$$u^{\varepsilon}(x_\delta) + H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \leq \varepsilon\left(\frac{n}{\alpha} + \delta\Delta\mu(x_\delta)\right) \leq \varepsilon\left(\frac{n}{\alpha} + \delta\right).$$

Step 2. As $y \mapsto \Phi^{\delta}(x_\delta, y)$ has a max at y_δ, $y \mapsto u(y) - \left[-\frac{|x_\delta - y|^2}{2\alpha} - \delta\mu(y)\right]$ has a min at y_δ. The supersolution test for (1.45) gives

$$u(y_\delta) + H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) \geq 0. \tag{1.47}$$

Step 3. We have in the following some simple observations.

- We use the fact that $\Phi^{\delta}(x_\delta, x_\delta) \leq \Phi^{\delta}(x_\delta, y_\delta)$ to yield

$$\frac{|x_\delta - y_\delta|^2}{2\alpha} \leq u(x_\delta) - u(y_\delta) + \delta\Big(\mu(x_\delta) - \mu(y_\delta)\Big).$$

- Similarly, $\Phi^{\delta}(y_\delta, y_\delta) \leq \Phi^{\delta}(x_\delta, y_\delta)$ implies

$$\frac{|x_\delta - y_\delta|^2}{2\alpha} \leq u^{\varepsilon}(x_\delta) - u^{\varepsilon}(y_\delta) + \delta\Big(\mu(y_\delta) - \mu(x_\delta)\Big).$$

[1] An example for such a function like this is $\mu(x) = c\left(\sqrt{1+|x|^2} - 1\right)$ for $c > 0$ small enough.

Combine the above two inequalities to get

$$\frac{|x_\delta - y_\delta|^2}{\alpha} \le u(x_\delta) - u(y_\delta) + u^\varepsilon(x_\delta) - u^\varepsilon(y_\delta) \le 2C|x_\delta - y_\delta|,$$

and therefore, $|x_\delta - y_\delta| \le C\alpha$.

Step 4. By the assumption that $H \in \mathrm{Lip}(\mathbb{R}^n \times B(0,R))$ for each $R > 0$, if we pick $\delta \in (0,1)$, then we have

$$H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) \le C\alpha,$$

$$H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \le C\delta.$$

Thus,

(1.48)

$$H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \le C(\alpha + \delta).$$

Step 5. Combine the inequalities in (1.46), (1.47), and (1.48) to imply

(1.49)

$$\begin{aligned} &u^\varepsilon(x_\delta) - u(y_\delta) \\ &\le \varepsilon\left(\frac{n}{\alpha} + \delta\right) + H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \\ &\le \varepsilon\left(\frac{n}{\alpha} + \delta\right) + C\alpha + C\delta. \end{aligned}$$

Now, for any $x \in \mathbb{R}^n$, we have $\Phi^\delta(x,x) \le \Phi^\delta(x_\delta, y_\delta) \le u^\varepsilon(x_\delta) - u(y_\delta)$, and hence,

$$u^\varepsilon(x) - u(x) - 2\delta\mu(x) \le u^\varepsilon(x_\delta) - u(y_\delta) \le \varepsilon\left(\frac{n}{\alpha} + \delta\right) + C\alpha + C\delta$$

by (1.49). Let $\delta \to 0$ and $C = \max\{n, C\}$ to obtain

$$u^\varepsilon(x) - u(x) \le C\left(\frac{\varepsilon}{\alpha} + \alpha\right).$$

Choose $\alpha = \sqrt{\varepsilon}$. We then get $u^\varepsilon(x) - u(x) \le C\sqrt{\varepsilon}$ for all $x \in \mathbb{R}^n$. By repeating the above, we obtain the other inequality in a similar way. The proof is complete. □

Remark 1.43. In fact, Step 3 in the above proof is often used in the viscosity solution theory to get a bound of $|x_\delta - y_\delta|$. Another way, which is quicker in this situation, to bound $|x_\delta - y_\delta|$ is already hidden in Step 1. Indeed, we note that

$$Du^\varepsilon(x_\delta) = \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta) \quad \Longrightarrow \quad \frac{|x_\delta - y_\delta|}{\alpha} \le |Du^\varepsilon(x_\delta)| + \delta \le C,$$

for $\delta \in (0,1)$. Thus, Step 3 is obtained.

1.12. Rate of convergence of the vanishing viscosity process for static problems via the nonlinear adjoint method

1.12.1. General nonconvex Hamiltonians. We consider the same situation as in the previous section. We are interested in the vanishing viscosity procedure for the usual static problem

$$u(x) + H(x, Du(x)) = 0 \qquad \text{in } \mathbb{R}^n. \tag{1.50}$$

For each $\varepsilon > 0$, we consider

$$u^\varepsilon(x) + H(x, Du^\varepsilon(x)) = \varepsilon \Delta u^\varepsilon(x) \qquad \text{in } \mathbb{R}^n. \tag{1.51}$$

We aim at proving $\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n)} \le C\sqrt{\varepsilon}$ by a different method via the nonlinear adjoint method to be described soon. Here is the assumption that we require, which is quite similar to (1.27):

(1.52)
$$\begin{cases} H \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n), H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) & \text{for each } R > 0, \\ |D_x H(x,p)| \le C(1+|p|) & \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n, \\ \lim_{|p|\to\infty} \frac{H(x,p)}{|p|} = \infty & \text{uniformly for } x \in \mathbb{R}^n, \end{cases}$$

for some given $C > 0$.

Then, by Bernstein's method, (1.51) has a unique smooth solution u^ε, and there is a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that

$$\|u^\varepsilon\|_{L^\infty(\mathbb{R}^n)} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n)} \le C.$$

Everything is set for us to study the convergence rate of u^ε to u.

Let us now give a gentle introduction to the nonlinear adjoint method. For $\varepsilon > 0$, consider the operator

$$\begin{aligned} F^\varepsilon : C^2(\mathbb{R}^n) &\longrightarrow C(\mathbb{R}^n) \\ \varphi(x) &\longmapsto F^\varepsilon[\varphi](x) = \varphi(x) + H(x, D\varphi(x)) - \varepsilon\Delta\varphi(x). \end{aligned}$$

Then from (1.51), we have $F^\varepsilon[u^\varepsilon] = 0$. The linearized operator $\mathcal{L}^\varepsilon$ of F^ε about the solution u^ε is defined as, for $\varphi \in C_c^\infty(\mathbb{R}^n)$,

$$\mathcal{L}^\varepsilon[\varphi] = \lim_{t\to 0} \frac{F^\varepsilon[u^\varepsilon + t\varphi] - F^\varepsilon[u^\varepsilon]}{t},$$

which gives

$$\mathcal{L}^\varepsilon[\varphi](x) = \varphi(x) + D_pH(x, Du^\varepsilon(x)) \cdot D\varphi(x) - \varepsilon\Delta\varphi(x).$$

We denote by $(\mathcal{L}^\varepsilon)^*$ the adjoint operator of $\mathcal{L}^\varepsilon$, which means

$$\int_{\mathbb{R}^n} \mathcal{L}^\varepsilon[\varphi]\sigma \, dx = \int_{\mathbb{R}^n} (\mathcal{L}^\varepsilon)^*[\sigma]\varphi \, dx \qquad \text{for all } \sigma \in C_c^\infty(\mathbb{R}^n).$$

By integration by parts,

$$\begin{aligned}\int_{\mathbb{R}^n} \mathcal{L}^\varepsilon[\varphi]\sigma\, dx &= \int_{\mathbb{R}^n} \Big(\varphi + D_pH(x, Du^\varepsilon)\cdot D\varphi - \varepsilon\Delta\varphi\Big)\sigma\, dx \\ &= \int_{\mathbb{R}^n} \Big(\sigma - \operatorname{div}\Big(D_pH(x, Du^\varepsilon)\sigma\Big) - \varepsilon\Delta\sigma\Big)\varphi\, dx \\ &= \int_{\mathbb{R}^n} (\mathcal{L}^\varepsilon)^*[\sigma]\varphi\, dx.\end{aligned}$$

Thus,

$$(\mathcal{L}^\varepsilon)^*[\sigma] = \sigma - \operatorname{div}\Big(D_pH(x, Du^\varepsilon)\sigma\Big) - \varepsilon\Delta\sigma.$$

Based on the adjoint operator $(\mathcal{L}^\varepsilon)^*$, we consider the following adjoint equation: For each $x_0 \in \mathbb{R}^n$,

$$\sigma^\varepsilon - \operatorname{div}\left(D_pH(x, Du^\varepsilon)\sigma^\varepsilon\right) - \varepsilon\Delta\sigma^\varepsilon = \delta_{x_0} \qquad \text{in } \mathbb{R}^n. \tag{1.53}$$

Here, δ_{x_0} is the Dirac delta at x_0. Let σ^ε be the unique solution to (1.53), which is basically its fundamental solution. Then, we have the following properties:

(1) $\sigma^\varepsilon \in C^\infty(\mathbb{R}^n\backslash\{x_0\})$,
(2) $\sigma^\varepsilon > 0$ in $\mathbb{R}^n \setminus \{x_0\}$,
(3) $\int_{\mathbb{R}^n} \sigma^\varepsilon\, dx = 1$.

Equation (1.53), introduced by Evans [**Eva10a**] and Tran [**Tra11**], is a new object in the study of viscosity solutions. The goal now is to find new estimates by doing various kinds of linearizations to the PDE (1.51) and then integrating by parts with σ^ε.

Lemma 1.44. *Assume* (1.52). *For each $\varepsilon \in (0,1)$, let u^ε be the unique smooth solution to* (1.51), *and let σ^ε be the unique solution to* (1.53) *for fixed $x_0 \in \mathbb{R}^n$. Then, there exists a constant C independent of ε such that*

$$\varepsilon\int_{\mathbb{R}^n} |D^2u^\varepsilon|^2\sigma^\varepsilon\, dx \le C. \tag{1.54}$$

Proof. Let $\varphi = \frac{1}{2}|Du^\varepsilon|^2$. By doing computations similar to these in the classical Bernstein method, we obtain from (1.51) that

$$2\varphi + D_xH(x, Du^\varepsilon)\cdot Du^\varepsilon + D_pH(x, Du^\varepsilon)\cdot D\varphi = \varepsilon\Delta\varphi - \varepsilon|D^2u^\varepsilon|^2.$$

By the Bernstein method, $2\varphi = |Du^\varepsilon|^2 \le C$; thus from the assumption that $|D_xH(x,p)| \le C(1+|p|)$, we get

$$\begin{aligned}&\Big(\varphi + D_pH(x,Du^\varepsilon)\cdot D\varphi - \varepsilon\Delta\varphi\Big) + \varepsilon|D^2u^\varepsilon|^2\\ &= -\underbrace{\Big(\varphi + D_xH(x,Du^\varepsilon)\cdot Du^\varepsilon\Big)}_{\text{bounded}}.\end{aligned}$$

Multiplying both sides by σ^ε and taking integration over $\mathbb{R}^n$, we obtain

$$\begin{aligned}&\int_{\mathbb{R}^n}\Big(\varphi + D_pH(x,Du^\varepsilon)\cdot D\varphi - \varepsilon\Delta\varphi\Big)\sigma^\varepsilon\,dx + \varepsilon\int_{\mathbb{R}^n}|D^2u^\varepsilon|^2\sigma^\varepsilon\,dx\\ = &-\int_{\mathbb{R}^n}\Big(\varphi + D_xH(x,Du^\varepsilon)\cdot Du^\varepsilon\Big)\sigma^\varepsilon\,dx \le C.\end{aligned}$$

Using the adjoint equation, we obtain

$$\int_{\mathbb{R}^n}\underbrace{\Big(\sigma^\varepsilon - \operatorname{div}\,(D_pH(x,Du^\varepsilon)\sigma^\varepsilon) - \varepsilon\Delta\sigma^\varepsilon\Big)}_{\delta_{x_0}}\varphi\,dx + \varepsilon\int_{\mathbb{R}^n}|D^2u^\varepsilon|^2\sigma^\varepsilon\,dx \le C.$$

Thus,

$$\varepsilon\int_{\mathbb{R}^n}|D^2u^\varepsilon|^2\sigma^\varepsilon\,dx \le C - \varphi(x_0) \le C.$$

The proof is complete. □

Remark 1.45. It is important to note that (1.54) is one of the new key estimates in the development of the nonlinear adjoint method. Originally, if we look at (1.51), we are only able to get that $\varepsilon|\Delta u^\varepsilon| \le C$, which means that $|\Delta u^\varepsilon| \le O(\frac{1}{\varepsilon})$ in $\mathbb{R}^n$. The new estimate (1.54) gives better control of D^2u^ε on the support of σ^ε, where we have, roughly speaking, $|D^2u^\varepsilon| \le O(\frac{1}{\sqrt{\varepsilon}})$. This turns out to be quite useful in various situations.

We are ready to state and prove our rate of convergence result.

Theorem 1.46. *Assume that H satisfies* (1.52)*. For each $\varepsilon \in (0,1)$, let u^ε be the unique smooth solution to* (1.51)*. Let u be the unique Lipschitz viscosity solution of* (1.50)*. Then, there exists a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that*

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n)} \le C\sqrt{\varepsilon}. \tag{1.55}$$

Proof. We have that $\varepsilon \mapsto u^\varepsilon$ is smooth for $\varepsilon > 0$. Let us differentiate (1.51) with respect to ε to get

$$u^\varepsilon_\varepsilon + D_pH(x,Du^\varepsilon)\cdot Du^\varepsilon_\varepsilon = \Delta u^\varepsilon + \varepsilon\Delta u^\varepsilon_\varepsilon.$$

Here, we write $u^\varepsilon_\varepsilon = \frac{\partial u^\varepsilon}{\partial \varepsilon}$. In terms of the linearized operator $\mathcal{L}^\varepsilon$, we can rewrite the above equation as

$$\begin{aligned}\mathcal{L}^\varepsilon[u^\varepsilon_\varepsilon] = \Delta u^\varepsilon \quad &\Longrightarrow \quad \int_{\mathbb{R}^n} \mathcal{L}^\varepsilon[u^\varepsilon_\varepsilon]\sigma^\varepsilon \, dx = \int_{\mathbb{R}^n} \Delta u^\varepsilon \sigma^\varepsilon \, dx \\ &\Longrightarrow \quad u^\varepsilon_\varepsilon(x_0) = \int_{\mathbb{R}^n} (\mathcal{L}^\varepsilon)^* [\sigma^\varepsilon] u^\varepsilon_\varepsilon \, dx = \int_{\mathbb{R}^n} \Delta u^\varepsilon \sigma^\varepsilon \, dx.\end{aligned}$$

Now using Lemma 1.44 and Hölder's inequality, we obtain

$$\begin{aligned}|u^\varepsilon_\varepsilon(x_0)| = \left| \int_{\mathbb{R}^n} \Delta u^\varepsilon \sigma^\varepsilon \, dx \right| &\leq \left(\int_{\mathbb{R}^n} |\Delta u^\varepsilon|^2 \sigma^\varepsilon \, dx \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^n} \sigma^\varepsilon \, dx \right)^{\frac{1}{2}} \\ &\leq C \left(\int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 \sigma^\varepsilon \, dx \right)^{\frac{1}{2}} \leq \frac{C}{\sqrt{\varepsilon}}.\end{aligned}$$

The above inequality yields

$$|u^\varepsilon(x_0) - u(x_0)| = \left| \int_0^\varepsilon \frac{\partial u^\delta(x_0)}{\partial \delta} \, d\delta \right| \leq C \int_0^\varepsilon \frac{C}{\sqrt{\delta}} \, d\delta = C\sqrt{\varepsilon}$$

by the fundamental theorem of calculus. □

1.12.2. Uniformly convex Hamiltonians. Next, we show that in the case where H is uniformly convex in p, then we have some further estimates. In addition to (1.52), we assume that

(1.56)
$$\begin{cases} D^2_{pp} H(x,p) \geq \theta I_n & \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n, \\ D^2_{xx} H, D^2_{xp} H, D^2_{pp} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) & \text{for each } R > 0, \end{cases}$$

for some given $\theta > 0$. Here, I_n is the identity matrix of size n.

Theorem 1.47. *Assume that H satisfies* (1.52) *and* (1.56). *We fix $r \in C^\infty_c(\mathbb{R}^n, [0,\infty))$ such that $\int_{\mathbb{R}^n} r(x)\, dx = 1$. For each $\varepsilon \in (0,1)$, let u^ε be the unique smooth solution to* (1.51). *Let u be the unique Lipschitz viscosity solution of* (1.50). *Then, there exists a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that, for every $y \in \mathbb{R}^n$,*

$$\left| \int_{\mathbb{R}^n} (u^\varepsilon(x) - u(x)) r(x+y)\, dx \right| \leq C \left(1 + \|Dr\|_{L^1(\mathbb{R}^n)}\right)^{\frac{1}{2}} \varepsilon. \tag{1.57}$$

Before proving this theorem, let us give a new estimate in this uniformly convex setting. For this case, we consider the following adjoint equation: For each $y \in \mathbb{R}^n$, let σ^ε be the solution to

$$\sigma^\varepsilon - \mathrm{div}\,(D_p H(x, Du^\varepsilon)\sigma^\varepsilon) - \varepsilon \Delta \sigma^\varepsilon = r(\cdot + y) \qquad \text{in } \mathbb{R}^n. \tag{1.58}$$

Note that we abuse the notation here as we use the same σ^ε in (1.53) and (1.58). It is clear that σ^ε satisfies

(1) $\sigma^\varepsilon \in C^\infty(\mathbb{R}^n, (0,\infty))$,

(2) $\int_{\mathbb{R}^n} \sigma^\varepsilon \, dx = 1$.

Lemma 1.48. *Assume the settings in Theorem* 1.47*. Then, there exists a constant $C > 0$ independent of ε so that*

$$\int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 \sigma^\varepsilon \, dx \le C \left(1 + \|Dr\|_{L^1(\mathbb{R}^n)}\right). \tag{1.59}$$

Proof. Differentiate (1.51) twice with respect to x_i for $1 \le i \le n$ to obtain

$$u^\varepsilon_{x_i x_i} + D_p H(x, Du^\varepsilon) \cdot Du^\varepsilon_{x_i x_i} + H_{x_i x_i} + 2H_{x_i p_k} u^\varepsilon_{x_i x_k} + H_{p_k p_l} u^\varepsilon_{x_i x_k} u^\varepsilon_{x_i x_l} = \varepsilon \Delta u^\varepsilon_{x_i x_i}.$$

Thanks to (1.56),

$$H_{p_k p_l} u^\varepsilon_{x_i x_k} u^\varepsilon_{x_i x_l} \ge \theta |Du^\varepsilon_{x_i}|^2 \qquad \text{and} \qquad 2\left|H_{x_i p_k} u^\varepsilon_{x_i x_k}\right| \le \frac{\theta}{2} |Du^\varepsilon_{x_i}|^2 + C.$$

Thus,

$$\mathcal{L}^\varepsilon[u^\varepsilon_{x_i x_i}] + \frac{\theta}{2} |Du^\varepsilon_{x_i}|^2 \le C.$$

Multiply the above by σ^ε and integrate to yield

$$\begin{aligned} \frac{\theta}{2} \int_{\mathbb{R}^n} |Du^\varepsilon_{x_i}|^2 \sigma^\varepsilon \, dx &\le C - \int_{\mathbb{R}^n} u^\varepsilon_{x_i x_i}(x) r(x+y) \, dx \\ &= C + \int_{\mathbb{R}^n} u^\varepsilon_{x_i}(x) r_{x_i}(x+y) \, dx \le C \left(1 + \|Dr\|_{L^1(\mathbb{R}^n)}\right). \end{aligned}$$

Sum the above inequality over i to complete the proof. □

We are now ready to prove Theorem 1.47.

Proof of Theorem 1.47. We proceed as in the first part of the proof of Theorem 1.46. We have that $\varepsilon \mapsto u^\varepsilon$ is smooth for $\varepsilon > 0$. Differentiate (1.51) with respect to ε to get

$$u^\varepsilon_\varepsilon + D_p H(x, Du^\varepsilon) \cdot Du^\varepsilon_\varepsilon = \Delta u^\varepsilon + \varepsilon \Delta u^\varepsilon_\varepsilon.$$

Recall that $u^\varepsilon_\varepsilon = \frac{\partial u^\varepsilon}{\partial \varepsilon}$. In terms of the linearized operator $\mathcal{L}^\varepsilon$, we can rewrite the above equation as

$$\mathcal{L}^\varepsilon[u^\varepsilon_\varepsilon] = \Delta u^\varepsilon.$$

Multiplying this by σ^ε and integrating by parts, one has

$$\begin{aligned}\left|\int_{\mathbb{R}^n} u^\varepsilon_\varepsilon(x) r(x+y)\,dx\right| &= \left|\int_{\mathbb{R}^n} \Delta u^\varepsilon \sigma^\varepsilon\,dx\right| \\ &\leq \left(\int_{\mathbb{R}^n} |\Delta u^\varepsilon|^2 \sigma^\varepsilon\,dx\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^n} \sigma^\varepsilon\,dx\right)^{\frac{1}{2}} \\ &\leq C\left(\int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 \sigma^\varepsilon\,dx\right)^{\frac{1}{2}} \leq C\left(1+\|Dr\|_{L^1(\mathbb{R}^n)}\right)^{\frac{1}{2}}.\end{aligned}$$

We then use the fundamental theorem in calculus to deduce that

$$\begin{aligned}\left|\int_{\mathbb{R}^n} (u^\varepsilon(x)-u(x)) r(x+y)\,dx\right| &= \left|\int_0^\varepsilon \int_{\mathbb{R}^n} \frac{\partial u^\delta(x)}{\partial \delta} r(x+y)\,dx d\delta\right| \\ &\leq C\left(1+\|Dr\|_{L^1(\mathbb{R}^n)}\right)^{\frac{1}{2}} \varepsilon.\end{aligned}$$

The proof is complete. □

It is clear that (1.57) gives a better rate of convergence $O(\varepsilon)$ compared to the rate $O(\sqrt{\varepsilon})$ in (1.55). One technical point here that we would like to address is that (1.57) is an average estimate, not a pointwise estimate like (1.55). This comes from the fact that in order to control $\int_{\mathbb{R}^n} u^\varepsilon_{x_i x_i}(x) r(x+y)\,dx$, we need to use integration by parts and $\|Dr\|_{L^1(\mathbb{R}^n)}$. Nevertheless, (1.57) is a natural estimate that one would expect in the uniformly convex setting.

1.12.3. Problems.

Exercise 18. Give another proof of Theorem 1.46 by using directly the usual maximum principle (without using the nonlinear adjoint method) for

$$\psi = \sqrt{\varepsilon} u^\varepsilon_\varepsilon + |Du^\varepsilon|^2.$$

Exercise 19. In the general nonconvex setting, is the convergence rate $O(\sqrt{\varepsilon})$ of u^ε to u in Theorem 1.46 optimal?

Exercise 20. Is the convergence rate $O(\varepsilon)$ in (1.57) of Theorem 1.47 optimal in the uniformly convex setting?

1.13. References

(1) There have been many great textbooks in the study of viscosity solutions for Hamilton–Jacobi equations written by Bardi and Capuzzo-Dolcetta [**BCD97**], Barles [**Bar94**], Cannarsa and Sinestrari [**CS04**], Evans [**Eva10b**, Chapter 10], Fabbri, Gozzi, and Święch [**FGS17**], Fleming and Soner [**FS06**], Isaacs [**Isa65**], Koike

[**Koi04**], Lions [**Lio82**], and Melikyan [**Mel98**]. Besides, the user's guide written by Crandall, Ishii, and Lions [**CIL92**] is used extensively in the literature for second-order equations.

(2) Besides these books, there are many interesting lecture notes available. Let me list a few representative ones: Bressan [**Bre**], Calder [**Cal**], Crandall [**Cra97**], and Le, Mitake, and Tran [**LMT17**].

(3) The level set method was first introduced numerically by Osher and Sethian [**OS88**]. The rigorous treatment was developed later by Evans and Spruck [**ES91**] and Chen, Giga, and Goto [**CGG91**], independently. See the textbook of Giga [**Gig06**] and the references therein for the developments of this direction.

(4) The G-equation is quite popular in the combustion science literature: See Markstein [**Mar**], Sivashinsky [**Siv88**], Yakhot [**Yak88**], and Denet [**Den99**]. We refer the readers to Cardaliaguet, Nolen, and Souganidis [**CNS11**], Xin and Yu [**XY10**], and Liu, Xin, and Yu [**LXY11**] for some recent important mathematical developments.

(5) Evans [**Eva80**] first used the Minty trick to study the vanishing viscosity method and gave definitions of possibly weak solutions. Crandall and Lions [**CL83**] proved the uniqueness of viscosity solutions to (1.1), thus established the firm foundation for the theory of viscosity solutions to first-order equations. In the literature, people often call this "the Crandall–Lions theory of viscosity solutions". The key new idea introduced by Crandall and Lions is the doubling variables method, which was inspired by an idea of Kružkov [**Kru67**] in scalar conservation laws. Crandall and Lions chose the name "viscosity solutions" in honor of the vanishing viscosity technique. See also Friedman [**Fri74**] for the vanishing viscosity process. We do not discuss the well-posedness of second-order equations here.

(6) Ishii [**Ish87**] introduced the Perron method to the theory of viscosity solutions, and since then, it has been used extensively in the literature to establish existence of viscosity solutions. The advantage of this approach is that one does not need to go through the vanishing viscosity method to get existence of solutions.

(7) The nonlinear adjoint method was introduced first by Evans [**Eva10a**] to study the gradient shock structures of the Cauchy problem for nonconvex Hamiltonians. The static cases were studied by Tran [**Tra11**]. Much needs to be studied in the direction of the gradient shock structures. The result in Theorem 1.47 is new in the literature although the ideas in its proof are already in

[**Eva10a**, **Tra11**]. Recently, this method has been developed much further to study large time behaviors, selection problems, and dynamical properties of solutions to Hamilton–Jacobi equations in the convex setting. For this, see the lecture notes by Le, Mitake, and Tran [**LMT17**]. We will employ this approach later on in the study of weak KAM theory.

(8) For fundamental solutions of elliptic equations, see Littman, Stampacchia, and Weinberger [**LSW63**]. For fundamental solutions of parabolic equations, see Chapter 1 of Friedman [**Fri64**].

(9) We do not discuss viscosity solutions to boundary value problems here. See Appendix E for some very brief discussions on this.

(10) We do not discuss the weak Bernstein method, which is applicable directly to viscosity solutions here. See Barles [**Bar91**], Armstrong and Tran [**AT15**], and the references therein.