HAMILTON–JACOBI EQUATIONS: THEORY AND APPLICATIONS

HUNG VINH TRAN

Department of Mathematics
University of Wisconsin Madison
Van Vleck hall, 480 Lincoln Drive, Madison, WI 53706, USA

Fall 2020

Contents

Preface

I introduce in this book an extensive survey of many important topics in the theory of Hamilton–Jacobi equations with particular emphasis on modern approaches and viewpoints.

Firstly, I cover the basic well-posedness theory of viscosity solutions for first-order Hamilton–Jacobi equations. This is, by now, quite standard and there have been some great books on this matter since 1980s in the literature. Nevertheless, it is important to have some key topics covered here in a self-contained way for the use throughout the book. It is not of our intention here to cover extensively about well-posedness of viscosity solutions for various different kinds of partial differential equations (PDEs).

Then, I aim at going beyond the well-posedness theory and studying further properties of viscosity solutions to Hamilton–Jacobi equations. Along this direction, I first discuss in deep the homogenization theory for Hamilton–Jacobi equations. Although this has always been a very active research topic since the late 1980s until this moment (2020), there has not been any standard textbook covering this. I am hopeful that this book will serve as a gentle introductory reference on this subject. Various connections between homogenization and other research subjects are discussed as well. I focus on the periodic and almost periodic settings in the book, and choose not to cover a more general and more complicated framework, which is the stationary ergodic setting.

Afterwards, dynamical properties, Aubry–Mather theory, and weak Kolmogorov–Arnold–Moser (KAM) theory are studied. These appear naturally in the study of first-order Hamilton–Jacobi equations when the Hamiltonian is convex in the momentum variable. I will introduce both dynamical and PDE approaches to study these theories. Then, I will discuss connections between homogenization and dynamical system, and optimal rate of convergence in homogenization theory as well.

Let me emphasize that this is a textbook, not a research monograph. My hope is that it can be used by advanced undergraduate students, first and second year graduate students, and new researchers entering the fields of Hamilton–Jacobi equations and viscosity solutions as a learning tool. In this case, the readers can follow the flow of the book from the beginning (Chapters 1 and 2), then jump to the topics that the readers aim at. Besides, I intend to keep the contents of various topics covered here as independent as possible so that other interested readers are able to jump directly to a subject of interests in the book.

My intention when writing this book is to present the essential ideas in the clearest possible ways, and thus, in various places, the assumptions/conditions imposed are not sharp. In many cases, the readers can improve the assumptions/conditions imposed right away. I will refer to a list of research articles and monographs at the end of each chapter that provide more general pictures of the situations.

Here is a quick outline of the content of the book. Chapter 1 contains the basic theory of viscosity solutions for Hamilton–Jacobi equations. This includes the well-posedness theory of viscosity solutions, the classical Bernstein method to obtain gradient bounds, Perron's method to prove existence of viscosity solutions, finite speed of propagation for Cauchy problems, and rate of convergence of the vanishing viscosity process via both the doubling variables method, and the nonlinear adjoint method. Chapter 2 is about Hamilton–Jacobi equations with convex Hamiltonians. We discuss the optimal control theory, Dynamical Programming Principle, Legendre's transform, the Lagrangian viewpoint, and the Hopf-Lax formula. We then study some further hidden convex structures, and also the maximal subsolutions with their representation formulas there.

Chapter 3 is concerned with Hamilton–Jacobi equations with possibly nonconvex Hamiltonians. We discuss two-player zero-sum differential games, the upper and lower values of the games, and the corresponding equations. We then give representation formulas including the Hopf formula of the solutions to these equations. Finally, we give a brief introduction to finite difference approximations to first-order Hamilton–Jacobi equations.

In Chapter 4, I cover the periodic homogenization theory for Hamilton–Jacobi equations. Homogenization results, cell problems, properties of the effective Hamiltonian in the convex and nonconvex settings, and some rates of convergence are studied. In a similar way, the almost periodic homogenization theory is discussed in Chapter 5 although much less is well understood here.

Chapter 6 is devoted to the analysis of convex Hamilton–Jacobi equations in a flat torus. We introduce new representation formulas for solutions to the discount problems and give some applications. The discount problems already appear in Chapter 2. Then, backward characteristics corresponding to the cell problems, and optimal rate of convergence in periodic homogenization are studied. This is related to the last part of Chapter 4. Besides, the backward characteristics provide a natural link between viscosity solutions and dynamical aspects of the corresponding Hamiltonian ODEs.

A gentle introduction to weak KAM theory is given in Chapter 7. Both Lagrangian methods and nonlinear PDE methods are presented. In particular, Mather measures, Mather set, and projected Aubry set are defined and analyzed. In Chapter 8, we study further properties of the effective Hamiltonians in the convex setting, which include strict convexity in certain directions, and the method of asymptotic expansion at infinity. Afterwards, the classical Hedlund example and its generalization are discussed.

The homework problems given in this book are of various level of difficulties. Most of the times, the exercises in corresponding sections are helpful for further understandings of relevant methods, ideas and techniques. Few of the problems are open ended and are related to some active research directions.

I would like to thank my Ph.D. student, Son Tu, who provided me the first draft of some of these notes based on a graduate topic course (Math 821) that I taught in Fall 2016 at UW Madison. Solutions to some problems were provided by him as well. I have been sitting on the notes for a long time before putting some real effort to have this book written.

Besides, I have also used some parts of my lecture notes taught at a topic course at University of Tokyo, Tokyo, Japan (September 2014), two topic courses at University of Science, Ho Chi Minh city, Vietnam (July 2015, July 2017) to form parts of this book. I would like to thank Professors Yoshikazu Giga, Hiroyoshi Mitake (University of Tokyo), Huynh Quang Vu (University of Science, Ho Chi Minh city) for their hospitalities.

I would like to thank my wife, Van Hai Van, and my daughter, An My Ngoc Tran, for their constant wonderful supports during the writing of this book. Besides, I am extremely grateful for the friendships and the supports from Wenjia Jing, Hiroyoshi Mitake, Yifeng Yu.

In the appendix, I include a characterization of the Legendre transform following Nam Le's very useful suggestion. I thank Nam much for this.

I appreciate Nattakorn Kittisut, Yeon-Eung Kim, Yuchen Mao, Loc Hoang Nguyen, Son Tu, Son Van, Lizhe Wan, Yifeng Yu for pointing out various typos and giving some great suggestions to improve the presentation of this book.

I am supported in part by NSF grant DMS-1664424 and NSF CAREER grant DMS-1843320 during the writing of this book.

Hung Vinh Tran
Madison, WI
Fall 2020.

Chapter 1

Introduction to viscosity solutions for Hamilton–Jacobi equations

1 Introduction

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

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

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

- The Laplacian $\Delta u(x) = \text{tr}(D^2u(x)) = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}(x)$ is the trace of $D^2u(x)$.

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

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

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:

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

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 players 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 is 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 is remain 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 [122] introduced the level set method (numerically) to study this problem. The rigorous treatment was developed later by Evans, Spruck [55] and Chen, Giga, Goto [32], 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 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 with 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 $u(x(t), t) = 0$,

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

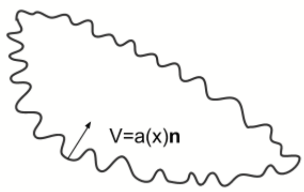


Figure 1.1: Front propagation of $\{\Gamma_t\}_{t\geq 0}$.

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 track with 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

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

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 with 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 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 breakdown 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 less solutions. This is typically the case in the theories of viscosity solution, renormalized solutions, etc.

2 Vanishing viscosity method for first-order Hamilton–Jacobi equations

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

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

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 the solution of (1.1) is using the idea of vanishing viscosity procedure. For each $\varepsilon > 0$, we consider

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

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 [46] 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 [39] proved the uniqueness of viscosity solutions to (1.1), thus, established the firm foundation for the theory of viscosity solutions to first-order equations. Roughly speaking, the procedure is carried out as following.

- 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 a priori estimate for $\{u^\varepsilon\}_{\varepsilon \in (0,1)}$: There exists a constant $C > 0$ independent of $\varepsilon \in (0,1)$ such that

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

  We will supply a proof of this later. Thus, $\{u^\varepsilon(x,t)\}_{\varepsilon \in (0,1)}$ is equi-continuous, 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

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

In a sense, this is a 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. By the above analysis,

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

Let $\varepsilon \to 0+$, 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.

2.1 Definition of viscosity solutions via touching functions

Let us denote

- BUC$(\mathbb{R}^n)$ the space of bounded, uniformly continuous functions on $\mathbb{R}^n$;
- 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 [37].

Definition 1.1 (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*

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

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*

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

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 1.2. 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).

2.2 Problems

Exercise 1. *Consider the eikonal problem mentioned earlier*

$$\begin{cases} |u'(x)| &= 1 \qquad \text{in } (-1,1), \\ u(1) = u(-1) &= 0. \end{cases} \tag{1.1}$$

(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*

$$\begin{cases} |(u^\varepsilon)'| &= 1 + \varepsilon (u^\varepsilon)'' \qquad \text{in } (-1,1), \\ u^\varepsilon(1) = u^\varepsilon(-1) &= 0. \end{cases} \tag{1.2}$$

(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 $\varphi, \psi \in C^2(\mathbb{R}^n \times (0,\infty))$. 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) (instead of strict maximum).*

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

2.3 Definition of viscosity solutions via generalized differentials

Definition 1.3. *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 1.4. *Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}$ 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 super-differential/max by sup-differential/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 non-decreasing. 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 $\varphi(x) = f(x)$, also from (1.3) we have $D\varphi(x) = p$ 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 $|y-x| < \frac{\delta}{2}$,

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

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

Using the notions of sub-differentials and super-differentials, one is able to give an equivalent definition of viscosity solution using somehow geometric interpretation of generalized differentials. This is clear from the result of Theorem 1.4. Nevertheless, let us present this equivalent definition here for completeness. In fact, it is important to keep in mind both of these definitions.

We consider the following first-order static PDE

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

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

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

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

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

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 following.

Proposition 1.6. *Let $f : \Omega \to \mathbb{R}$ and $x \in \Omega$, 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 an 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 non-empty. 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 $D(f - \varphi)(x) = 0$, 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

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

Since $\psi, \varphi \in C^1(\Omega)$, it's 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 $\varepsilon > 0$ 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 1.7. 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 put more assumptions, we will have typically more regularity results on u (e.g., u is Lipschitz in Ω), and we will discuss this situation more later.

2.4 Problems

Exercise 5. *Let u be a viscosity solution of* (1.4). *Show that*

(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 $|Du| = 1$ in $B_1(0)$.*

Exercise 7. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, and $F : \Omega \times \mathbb{R}^n \to \mathbb{R}$ 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 u solves the following eikonal equation in the viscosity sense*

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

3 Existence of viscosity solutions via the vanishing viscosity method

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

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

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

Lemma 1.8 (Stability of maximum/minimum points). *Let $u \in C(\mathbb{R}^n)$, and $\varphi \in C^1(\mathbb{R}^n)$ such that $u(x_0) = \varphi(x_0)$ for some $x_0 \in \mathbb{R}^n$, and $u - \varphi$ 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 $\varepsilon < \lambda_r$,

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

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 following. 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 before hand 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 1.9 (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 $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap Lip(\mathbb{R}^n)$ 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 viscosity supersolution test is similar, hence omitted. By Exercise 3, we can instead choose the test function $\varphi \in C^2(\mathbb{R}^n \times (0,T))$ (or $C^\infty(\mathbb{R}^n \times (0,T))$) such that $u - \varphi$ has a strict max at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. By Lemma 1.8, 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. □

Remark 1.10.

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
$$\overline{\varphi}(x,t)=\varphi(x,t)-\underbrace{\varphi(x_0,t_0)+u(x_0,t_0)}_{\text{a constant}}$$
to make 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 $\mathbb{R}^n\times(0,\infty)$.

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

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 a following lemma (compare this with Exercise 5).

Lemma 1.11. *Let Ω be an open subset of $\mathbb{R}^n$, and $u:\Omega\to\mathbb{R}$ be a continuous function. If u is differentiable $x_0\in\Omega$, 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, $u-\overline{\psi}$ has a local maximum at x_0. The proof is complete by setting, for $x\in\Omega$,
$$\varphi(x)=\overline{\varphi}(x)-|x-x_0|^2, \qquad \text{and} \qquad \psi(x)=\overline{\psi}(x)+|x-x_0|^2.$$
□

Theorem 1.12. *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 $\varphi, \psi \in C^1(\mathbb{R}^n \times (0,\infty))$ such that $u_t(x_0,t_0) = \varphi_t(x_0,t_0) = \psi_t(x_0,t_0)$, $Du(x_0,t_0) = D\varphi(x_0,t_0) = D\psi(x_0,t_0)$, and $u-\varphi$ has a strict minimum at (x_0,t_0), $u-\psi$ has a strict maximum at (x_0,t_0). Then, the viscosity subsolution and supersolution tests imply the result. □

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

Theorem 1.13 (Stability of viscosity solutions to (1.10))**.** *Assume that*

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

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

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

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 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

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

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

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

The proof is complete. □

5 The comparison principle and uniqueness result for static problem

We consider the following 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 1.14 (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,* $u(x) \leq v(x)$ *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 explanations after the proof of this comparison result.

Proof of Theorem 1.14. 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

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

we consider the following auxiliary function

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

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

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

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 following.

- **Step 1**. We have $\Phi^\varepsilon(x_\varepsilon, y_\varepsilon) \geq \Phi^\varepsilon(0,0)$, thus

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

  Let $C = 2(\|u\|_{L^\infty(\mathbb{R}^n)} + \|v\|_{L^\infty(\mathbb{R}^n)})$, 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, $\dfrac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \to 0$ as $\varepsilon \to 0$. Indeed, this follows by noting that

$$\begin{aligned}\Phi^\varepsilon(x_\varepsilon, y_\varepsilon) \geq \Phi^\varepsilon(x_\varepsilon, x_\varepsilon) \quad &\Longrightarrow \quad \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \leq v(x_\varepsilon) - v(y_\varepsilon) + \varepsilon\big(|x_\varepsilon|^2 - |y_\varepsilon|^2\big)\\ &\Longrightarrow \quad \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \leq v(x_\varepsilon) - v(y_\varepsilon) + C\varepsilon^{3/2},\end{aligned}$$

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, which yields

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

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), we imply

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

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

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

Plugging all of these together, we obtain

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

Combine this with (1.17) to deduce that

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

Recall that (1.14) gives

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

Plug it into (1.18) to yield

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

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

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

which is a contradiction. The proof is complete. □

Remark 1.15. In the above proof by 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 1.16 (Uniqueness of viscosity solution of static equation (1.12))**.** *Assume* (1.13). *If $u,v\in \mathrm{BUC}(\mathbb{R}^n)$ are 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 $u\le v$. Conversely, since v is a viscosity subsolution and u is a viscosity supersolution of (1.12), we deduce $v\le u$. Thus, $u=v$. □

Remark 1.17. Let us discuss further condition (1.13) here. 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 $x,y,p,q\in\mathbb{R}^n$,

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

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.

2. 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.

3. 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 into Step 2 of the proof of Theorem 1.14, we have in additional that $|x_\varepsilon - y_\varepsilon| \le C\varepsilon^2$. Then, in order to have the uniqueness result, we are able relax (1.13) a lot, for example, (1.13) can be replaced by the following

$$\begin{cases} \text{For each } R>0, \text{ there exists } C_R>0 \text{ so that, for } x,y \in \mathbb{R}^n,\ p,q \in B(0,R), \\ |H(x,p) - H(y,p)| \le C_R|x-y|, \\ |H(x,p) - H(x,q)| \le C_R|p-q|. \end{cases} \tag{1.20}$$

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.

5.1 Problems

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

6 The comparison principle and uniqueness result for Cauchy problem

We consider the following usual Cauchy problem

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

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 $x, y, p, q \in \mathbb{R}^n$,

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

The main result here is the comparison principle for (1.22), which is similar to Theorem 1.14. But before we proceed, we need the following simple lemma.

Lemma 1.18 (Extrema at terminal time $t = T$)**.** *Fix $T > 0$. Let u be a viscosity subsolution to* (1.22)*, and $\varphi \in C^1(\mathbb{R}^n \times [0,T])$ 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,*

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

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 $\varepsilon > 0$ is 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

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

which means

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

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

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

Theorem 1.19 (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 1.14, but it is worth presenting here since there is the time variable t that involves.

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 achieves 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 $\Phi(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon) \ge \Phi(0,0,0,0)$,

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

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

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

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

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

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

$$\lim_{\varepsilon\to0} \frac{|x_\varepsilon - y_\varepsilon|^2 + |s_\varepsilon - t_\varepsilon|^2}{\varepsilon^2} = 0.$$

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

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

where $\omega(\cdot)$ is a modulus of continuity, that is, $\lim_{r\to0} \omega(r) = 0$. 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,

$$(x,t) \longmapsto u(x,t) - \underbrace{\left[\frac{|x - y_\varepsilon|^2 + |t - s_\varepsilon|^2}{\varepsilon^2} + \varepsilon|x|^2 + \lambda t\right]}_{\varphi(x,t)} \qquad \text{has a max at } (x_\varepsilon, t_\varepsilon).$$

Since u is a viscosity subsolution to (1.22), the viscosity subsolution test gives

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

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

$$(y,s) \longmapsto v(y,s) - \underbrace{\left[-\frac{|x_\varepsilon - y|^2 + |t_\varepsilon - s|^2}{\varepsilon^2} - \varepsilon|y|^2 - \lambda s\right]}_{\psi(y,s)} \qquad \text{has a min at } (y_\varepsilon, s_\varepsilon).$$

The viscosity supersolution test yields

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

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

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

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. □

Corollary 1.20 (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 1.19. □

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 $x, y, p, q \in \mathbb{R}^n$,*

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

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 $t \geq 0$,

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

7 Introduction to the classical Bernstein method

For $\varepsilon > 0$, consider the following 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 $\|u^\varepsilon_t\|_{L^\infty} + \|Du^\varepsilon\|_{L^\infty} \leq C$ where $C > 0$ is independent of $\varepsilon \in (0,1)$. We put the following 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 [3, 65], [101, 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 1.21. *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 $(x,t) \in \mathbb{R}^n \times [0,\infty)$,*

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

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

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

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

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

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

Thus, we only need to bound $u_t^\varepsilon(\cdot,0)$. We build barriers to do this as following. 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 $x \in \mathbb{R}^n$.

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

$$\frac{d}{dt}\left(\frac{1}{2}\sum_{k=1}^n \left(u_{x_k}^\varepsilon\right)^2\right) + D_xH(x,Du^\varepsilon)\cdot Du^\varepsilon + D_pH(x,Du^\varepsilon)\cdot\left(\sum_{k=1}^n Du_{x_k}^\varepsilon u_{x_k}^\varepsilon\right) = \varepsilon\sum_{k=1}^n \Delta u_{x_k}^\varepsilon u_{x_k}^\varepsilon. \tag{1.28}$$

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

$$\sum_{k=1}^n Du_{x_k}^\varepsilon u_{x_k}^\varepsilon = D\left(\frac{1}{2}\sum_{k=1}^n \left(u_{x_k}^\varepsilon\right)^2\right) = D\psi \quad \text{and} \quad \sum_{k=1}^n \Delta u_{x_k}^\varepsilon u_{x_k}^\varepsilon = \Delta\psi - |D^2u^\varepsilon|^2.$$

Thus, (1.28) becomes

$$\psi_t + D_xH(x,Du^\varepsilon)\cdot Du^\varepsilon + D_pH(x,Du^\varepsilon)\cdot D\psi = \varepsilon\Delta\psi - \varepsilon|D^2u^\varepsilon|^2 \le \varepsilon\Delta\psi - \varepsilon\frac{(\Delta u^\varepsilon)^2}{n}.$$

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

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

Therefore,

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

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 $\|Du^\varepsilon\|_{L^\infty} \leq \|Du_0\|_{L^\infty} \leq C$, and the proof is complete. If $t_0 > 0$, then by the usual maximum principle, we have

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

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

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

which implies that, at (x_0, t_0),

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

in light of assumption (1.27).

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

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

□

Remark 1.22. 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 1.23. 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

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

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

7.1 Problems

Exercise 12. *Write down a detailed proof of the claim in Remark 1.23.*

Exercise 13. *Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 Hamiltonian satisfying*

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

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

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

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

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

Corollary 1.24. *Assume* (1.30). *Then, the static problem* (1.12) *has a Lipschitz viscosity solution u.*

8 Introduction to Perron's method

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 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 1.25. *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, definition for viscosity subsolutions to (1.32) can be given for upper semicontinuous functions in $\mathbb{R}^n$, $USC(\mathbb{R}^n)$, naturally. The result of Lemma 1.25 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 $0 \leq u(x_0) - u_n(x_0) \leq \frac{1}{n}$ for all $n \in \mathbb{N}$. For all $x \in \overline{B_r(x_0)}$, we have

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

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 viscosity subsolution of (1.32). □

The Perron method in the theory of viscosity solutions was first introduced by Ishii [82]. In the following, we give a variant of Ishii's argument in [82]. 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

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

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 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 1.26 (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 1.25), 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,

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

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

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

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 $\|D\overline{u}\|_{L^\infty(\mathbb{R}^n)} \leq C_1$. 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 1.17, we imply further that, under (1.33), u is actually is 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 Cauchy problem in the next section. Let us record what was discussed as a theorem here for later use.

Theorem 1.27. *Assume* (1.33). *Let u be defined as in Theorem 1.26. 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 1.28. *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 $-C_0 \leq u \leq C_0$ (this is being phrased as Exercise 14). By coercivity and the viscosity subsolution test, we get

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

We now show that u is Lipschitz with Lipschitz constant at most C_1. Given $\varepsilon > 0$ and $y \in \mathbb{R}^n$, consider $\varphi(x) = (C_1 + \varepsilon)|x - y|$, we have $\varphi \in C^\infty(\mathbb{R}^n \backslash \{y\})$. 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

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

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 1.29. *Let $u \in C(\mathbb{R}^n)$ 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.

8.2 Problems

Exercise 14. *Assume* (1.33). *Denote by $C_0 = \sup_{x \in \mathbb{R}^n} |H(x, 0)|$. 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 1.29.*

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,

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

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

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

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 1.26. 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

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

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 is 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 $\tilde{u}$ 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. □

Remark 1.31. Let us emphasize again that u defined in Theorem 1.30 might not be continuous. Besides (1.33) and (1.36), if we require in 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 $u^* \le u_*$. 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.

9 Lipschitz estimates for Cauchy problems using Perron's method

Let us continue focusing on the usual Cauchy problem

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

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

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

Theorem 1.32. *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 1.30, 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 1.33. We have two following comments.

- We use crucially the point that H is time independent in the above proof. In fact, if H is time dependent, then Step 2 above is completely broken. 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 of $H(x,p)$ 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

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

  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 1.17 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 get rid of (1.13) (or (1.19)) later. We record this important point in the following.

Theorem 1.34. *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.

9.1 Problems

Exercise 16. *Give a detailed proof of Step 3 in the proof of Theorem 1.32.*

Exercise 17. *Write down a proof of Theorem 1.34.*

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 of (1.41).

Theorem 1.35. *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 1.36. *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 1.19, 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 1.19), 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

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

Combining the two inequalities above to imply

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

We use (1.13) to deduce further that

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

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

To obtain Theorem 1.35, we now only need to show that $w(x,t) \leq 0$ for all $x \in B(0, R - Ct)$, and $t \leq \frac{R}{C}$.

Lemma 1.37. *Let w be a viscosity subsolution* (1.42). *Assume that $w(x,0) \leq 0$ for all $x \in B(0,R)$ for some given $R > 0$. Then,*

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

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 5.5 of Chapter 2.

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

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

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

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

Denote by

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

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 $x = 0$,

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

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

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

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 1.35. Let $w = u - v$. By using Lemmas 1.36 and 1.37, we immediately get the desired result. □

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,

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

Under this assumption, we use the classical Bernstein method (same arguments as in Theorem 1.21) 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

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

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

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

for some u satisfies $\|u\|_{L^\infty(\mathbb{R}^n)} + \|Du\|_{L^\infty(\mathbb{R}^n)} \le C$. 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 about 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 a 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 various interesting results in this direction, but still, the optimal rate for general case is not yet known. Up to now, for the general cases, the best known convergence rate is $O(\varepsilon^{1/2})$.

Theorem 1.38. *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*

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

This type of results with convergence rate $O(\varepsilon^{1/2})$ was first obtained by Fleming [62] in the 1960s by using a differential game approach. Later on, within the framework of viscosity solutions, Crandall and Lions [40] proved Theorem 1.38 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 1.38 by using the nonlinear adjoint method was introduced by Evans [50] and Tran [131].

We give here in this section a proof based on the ideas of Crandall, Lions [40]. The nonlinear adjoint method will be introduced in the next section.

Proof. By using the doubling variables method, consider the following 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]

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

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

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

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

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

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

[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.

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

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

- Similarly, $\Phi^\delta(y_\delta, y_\delta) \le \Phi^\delta(x_\delta, y_\delta)$ implies

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

  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

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

Thus,

$$H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \le C\alpha + C\delta. \tag{1.48}$$

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

$$\begin{aligned} u^\varepsilon(x_\delta) - u(y_\delta) &\le \varepsilon\left(\frac{n}{\alpha} + \delta\right) + H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \\ &\le \varepsilon\left(\frac{n}{\alpha} + \delta\right) + C\alpha + C\delta. \end{aligned} \tag{1.49}$$

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 $C = \max\{n, C\}$, we 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 1.39. 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.

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

12.1 General nonconvex Hamiltonians

We consider the same situation like 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)

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

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 following 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 $\varphi\in C_c^\infty(\mathbb{R}^n)$,

$$\mathcal{L}^\varepsilon[\varphi]=\lim_{t\to0}\frac{F^\varepsilon[u^\varepsilon+t\varphi]-F^\varepsilon[u^\varepsilon]}{t},$$

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 $x_0 \in \mathbb{R}^n$,

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

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 [50], Tran [131], 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 1.40. *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

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

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

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

Using the adjoint equation, we obtain

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

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 1.41. It is important noting 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 1.42. *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 1.40 and Hölder's inequality, we obtain

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

The above inequality yields

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

by the fundamental theorem of calculus. □

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

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

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

Theorem 1.43. *Assume that H satisfies* (1.52) *and* (1.56). *Let $r \in C_c^\infty(\mathbb{R}^n, [0,\infty))$ 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

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

Note that we abuse the notions 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 1.44. *Assume the settings in Theorem 1.43. Then, there exists a constant $C > 0$ independent of ε so that*

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

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

$$u^\varepsilon_{x_ix_i} + D_pH(x,Du^\varepsilon)\cdot Du^\varepsilon_{x_ix_i} + H_{x_ix_i} + 2H_{x_ip_k}u^\varepsilon_{x_ix_k} + H_{p_kp_l}u^\varepsilon_{x_ix_k}u^\varepsilon_{x_ix_l} = \varepsilon\Delta u^\varepsilon_{x_ix_i}.$$

Thanks to (1.56),

$$H_{p_kp_l}u^\varepsilon_{x_ix_k}u^\varepsilon_{x_ix_l} \geq \theta|Du^\varepsilon_{x_i}|^2 \qquad \text{and} \qquad 2\left|H_{x_ip_k}u^\varepsilon_{x_ix_k}\right| \leq \frac{\theta}{2}|Du^\varepsilon_{x_i}|^2 + C.$$

Thus,

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

Multiply the above by σ^ε and integrate to yield

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

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

We are now ready to prove Theorem 1.43.

Proof of Theorem 1.43. We proceed as in the first part of the proof of Theorem 1.42. We have that $\varepsilon \mapsto u^\varepsilon$ is smooth for $\varepsilon > 0$. Differentiate (1.51) with respect to ε to get

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

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

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

Multiply this by σ^ε and integrate by parts, one has

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

We then use the fundamental theorem in calculus to deduce that

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

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 one 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.

12.3 Problems

Exercise 18. *Give another proof of Theorem 1.42 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 1.42 optimal?*

Exercise 20. *Is the convergence rate $O(\varepsilon)$ in (1.57) of Theorem 1.43 optimal in the uniformly convex setting?*

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 [13], Barles [16], Cannarsa, Sinestrari [26], Chapter 10 of Evans [49], Fabbri, Gozzi, Swiech [56], Fleming and Soner [63], Isaacs [81], Koike [96], Lions [101], Melikyan [114]. Besides, the user's guide written by Crandall, Ishii, and Lions [38] is used extensively in the literature for second-order equations.

2. Besides these books, there are many interesting lecture notes available. Let me list few representative ones: Bressan [22], Calder [25], Crandall [36], Le, Mitake, Tran [100].

3. The level set method was first introduced numerically by Osher, Sethian [122]. The rigorous treatment was developed later by Evans, Spruck [55] and Chen, Giga, Goto [32], independently. See the textbook of Giga [68] and the references therein for the developments of this direction.

4. The G-equation is quite popular in the combustion science literature: see Markstein [111], Sivashinsky [126], Yakhot [136], and Denet [43]. We refer the readers to Cardaliaguet, Nolen, Souganidis [30], Xin, Yu [135], and Liu, Xin, Yu [106] for some recent important mathematical developments.

5. Evans [46] first used the Minty trick to study the vanishing viscosity method and gave first definitions of possibly weak solutions. Crandall and Lions [39] 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 "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 [98] in scalar conservation laws. Crandall and Lions chose the name "viscosity solutions" in honor of the vanishing viscosity technique. See also Friedman [65] for the vanishing viscosity process. We do not discuss the well-posedness of second-order equations here.

6. Ishii [82] 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 [50] to study the gradient shock structures of Cauchy problem for nonconvex Hamiltonians. The static cases were studied by Tran [131]. The result in Theorem 1.43 is new in the literature although the ideas in its proof are already in [50, 131]. 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, Tran [100]. We will employ this approach later on in the study of weak KAM theory.

8. For fundamental solutions of elliptic equations, see Littman, Stampacchia, Weinberger [105]. For fundamental solutions of parabolic equations, see Chapter 1 of Friedman [64].

9. We do not discuss viscosity solutions to boundary value problems here. See Appendix (Section 4) for some very brief discussions on this.

10. We do not discuss weak Bernstein's method, which is applicable directly to viscosity solutions here. See Barles [15], Armstrong, Tran [5], and the references therein.

Chapter 2

First-order Hamilton–Jacobi equations with convex Hamiltonians

Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a given Hamiltonian. Throughout this whole chapter, we always assume that $p \mapsto H(x,p)$ is convex for any given $x \in \mathbb{R}^n$. Usually, x represents the spatial variable (location), and p represents the momentum variable of a moving particle in $\mathbb{R}^n$. One important remark on the convexity assumption is that it is actually "one-sided" linearity. For each fixed $x \in \mathbb{R}^n$, we can always write

$$H(x,p) = \sup_{\alpha \in A_x} \{a_\alpha(x) \cdot p + b_\alpha(x)\},$$

where A_x is the collection of all planes $p \mapsto a_\alpha(x) \cdot p + b_\alpha(x)$ lie under the graph of $H(x,\cdot)$.

1 Introduction to the optimal control theory

Example 2.1 (Classical mechanics Hamiltonian)**.** *In this case, we assume that the mass of the particle is* 1 *(*$m = 1$*), and*

$$H(x,p) = \frac{1}{2}|p|^2 + V(x) \qquad \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Basically, $\frac{1}{2}|p|^2$ *is the kinetic energy, and* $V(x)$ *is the potential energy. It is not hard to check that*

$$H(x,p) = \sup_{q \in \mathbb{R}^n} \left\{ p \cdot q - \frac{1}{2}|q|^2 + V(x) \right\}.$$

The infinite horizon problem. Let us consider the following ODE, which represents the path of a moving person (or a particle)

$$\begin{cases} \gamma'(t) &= b(\gamma(t), v(t)) \qquad t > 0, \\ \gamma(0) &= x. \end{cases} \tag{2.1}$$

Here, we put the following assumptions.

- V is a given compact metric space, which is the control set.

- The vector field b is a map $b : \mathbb{R}^n \times V \to \mathbb{R}^n$ such that

$$\begin{cases} b \in C(\mathbb{R}^n \times V), \\ |b(x,v)| \le C & \text{for all } (x,v) \in \mathbb{R}^n \times V, \\ |b(x_1,v) - b(x_2,v)| \le C|x_1 - x_2| & \text{for all } x_1, x_2 \in \mathbb{R}^n,\ v \in V, \end{cases}$$

  for some $C > 0$.

- Every control $v(\cdot)$ is a measurable map $v : [0,\infty) \to V$. In principle, we are able to change this control as we wish.

Under the above assumptions, the ODE (2.1) has a unique solution, which is denoted by $y_{x,v(\cdot)}(\cdot)$. We write $y_{x,v(\cdot)}(\cdot)$ to emphasize that the path starts at x with the control $v(\cdot)$. For simplicity, we write $y_x(\cdot)$ instead of $y_{x,v(\cdot)}(\cdot)$ if there is no confusion. By being a solution to (2.1) here, we mean that

$$y_x(t) = x + \int_0^t b(y_x(s), v(s))\, ds \qquad \text{for all } t \ge 0.$$

We have the following lemma about the Lipschitz property of the trajectory.

Lemma 2.1. *The following claims hold.*

(a) For $t, s \ge 0$, $\left|y_{x,v(\cdot)}(t) - y_{x,v(\cdot)}(s)\right| \le C|t-s|$.

(b) Let $v(\cdot)$ be a control, and $y_x(\cdot)$, $y_z(\cdot)$ are corresponding trajectories starting from x, z, respectively. Then,

$$|y_x(t) - y_z(t)| \le e^{Ct}|x - z| \qquad \textit{for all } t > 0.$$

Proof. Claim (a) is obvious. To prove (b), we define $\varphi(s) = y_x(s) - y_z(s)$ for $s \ge 0$. Then, the Lipschitz continuity of b in the first variable gives $|\varphi'(s)| \le C|\varphi(s)|$ for $s \ge 0$. In particular, for any $t > 0$,

$$|\varphi(t)| = \left|\varphi(0) + \int_0^t \varphi'(s)\, ds\right| \le |\varphi(0)| + \int_0^t |\varphi'(s)|\, ds \le |x - z| + C\int_0^t |\varphi(s)|\, ds.$$

By Gronwall's inequality, we obtain

$$|\varphi(t)| = |y_x(t) - y_z(t)| \le e^{Ct}|x - z| \qquad \text{for all } t > 0,$$

and the proof is complete. $\square$

Cost functional. Fix $\lambda > 0$. For a given path $(y_x(\cdot), v(\cdot))$ of (2.1) we define the cost functional

$$J(x, v(\cdot)) = \int_0^\infty e^{-\lambda s} f(y_x(s), v(s))\, ds.$$

Here, $f : \mathbb{R}^n \times V \to \mathbb{R}$ is the running cost function, which satisfies

$$\begin{cases} f \in C(\mathbb{R}^n \times V), \\ |f(x,v)| \le C & \text{for all } (x,v) \in \mathbb{R}^n \times V, \\ |f(x_1,v) - f(x_2,v)| \le C|x_1 - x_2| & \text{for all } x_1, x_2 \in \mathbb{R}^n,\ v \in V, \end{cases}$$

for some $C > 0$.

The term $e^{-\lambda s}$ is called the discount factor. Technically, the discount factor helps to keep $\int_0^\infty e^{-\lambda s} f(y_x(s), v(s))\, ds$ finite as f is only bounded. More importantly, as we will see, this discount factor gives the appearance of the term λu in the static equation (2.3).

Main question. How to minimize the cost functional $J(x, v(\cdot))$ among all possible controls $v(\cdot)$? This type of questions appears a lot in Calculus of Variations. We define the cost value function as following. For $x \in \mathbb{R}^n$, set

$$u(x) = \inf_{v(\cdot)} J(x, v(\cdot)). \tag{2.2}$$

Basically, $u(x)$ is the minimum cost we must pay if we start at x. We now only study the cost function u, and ignore the underlying dynamics.

The following result is one of our main aims in this chapter.

Theorem 2.2. *Let u be defined as in* (2.2). *Then u is the unique viscosity solution to the following static equation*

$$\lambda u + H(x, Du) = 0 \qquad \text{in } \mathbb{R}^n. \tag{2.3}$$

Here, the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is determined by

$$H(x,p) = \sup_{v \in V} \Big(-b(x,v) \cdot p - f(x,v) \Big). \tag{2.4}$$

In order to prove the above theorem, we will obtain the following important identity for the value function u, whose proof is provided in the next section.

Dynamic Programming Principle (DPP). For any $x \in \mathbb{R}^n$ and $t > 0$, we have

$$u(x) = \inf_{v(\cdot)} \left(\int_0^t e^{-\lambda s} f(y_{x,v(\cdot)}(s), v(s))\, ds + e^{-\lambda t} u(y_{x,v(\cdot)}(t)) \right).$$

We summarize some useful properties of H defined in (2.4) in the following theorem.

Theorem 2.3. *Let H be defined as in* (2.4). *Then,*

(a) $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$, and $p \mapsto H(x,p)$ is convex for each $x \in \mathbb{R}^n$.

(b) *There exists* $C > 0$ *such that, for all* $x, y, p, q \in \mathbb{R}^n$,

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

Proof. For $v \in V$, let us denote $H_v(x,p) = -b(x,v) \cdot p - f(x,v)$ for all $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$. Then, $H(x,p) = \sup_{v \in V} H_v(x,p)$, and of course, H is convex in p.

Next, for $(x,p), (z,q) \in \mathbb{R}^n \times \mathbb{R}^n$, one has

$$\begin{aligned} |H_v(x,p) - H_v(z,q)| &= \left| \Big(b(x,v) - b(z,v) \Big) \cdot p + b(z,v) \cdot \Big(p - q \Big) + f(x,v) - f(z,v) \right| \\ &\leq C|p| \cdot |x-z| + C|p-q| + C|x-z|. \end{aligned}$$

Thus,

$$|H(x,p) - H(z,q)| \leq C\big(1 + |p|\big)|x-z| + C|p-q|.$$

The proof is complete. □

2 Dynamic Programming Principle

Let us recall quickly our setting. For each control $v(\cdot)$ and starting point $x \in \mathbb{R}^n$, the corresponding ODE is

$$\begin{cases} y_x'(t) & = b(y_x(t), v(t)) \qquad t > 0, \\ y_x(0) & = x. \end{cases} \tag{2.5}$$

Then, the value function u is defined as

$$u(x) = \inf_{v(\cdot)} J(x, v(\cdot)) = \inf_{v(\cdot)} \int_0^\infty e^{-\lambda s} f(y_x(s), v(s))\, ds.$$

Remark 2.4. It is worth emphasizing a difference between PDE and dynamical system viewpoints here.

- Dynamical system viewpoint: understand the behavior of minimizing paths.
- PDE viewpoint: forget about the underlying dynamics, only look at the value function u, and find out a PDE that u solves.

Before finding the PDE which u solves, we prove the Dynamic Programming Principle first.

Theorem 2.5 (Dynamic Programming Principle (DPP))**.** *Let u be defined as in* (2.2). *For any* $x \in \mathbb{R}^n$ *and* $t > 0$, *we have*

$$u(x) = \inf_{v(\cdot)} \left(\int_0^t e^{-\lambda s} f(y_x(s), v(s))\, ds + e^{-\lambda t} u(y_x(t)) \right). \tag{2.6}$$

Proof. Fix $x \in \mathbb{R}^n$ and $t > 0$. For each control $v(\cdot)$, let $\gamma(\cdot) = y_{x,v(\cdot)}$ be the solution to

$$\begin{cases} \gamma'(s) &= b(\gamma(s), v(s)) \qquad s > 0, \\ \gamma(0) &= x. \end{cases}$$

Denote by $\eta(s) = \gamma(s+t)$, $\tilde{v}(s) = v(s+t)$ for $s \geq 0$. Then $\tilde{v}$ is an admissible control, and η solves

$$\begin{cases} \eta'(s) &= b(\eta(s), \tilde{v}(s)) \qquad s > 0, \\ \eta(0) &= \gamma(t). \end{cases}$$

We easily deduce the following formula

$$\begin{aligned} J(x, v(\cdot)) &= \int_0^t e^{-\lambda s} f\,(\gamma(s), v(s))\; ds + e^{-\lambda t} J\big(\gamma(t), \tilde{v}(\cdot)\big) \\ &\geq \int_0^t e^{-\lambda s} f\,(\gamma(s), v(s))\; ds + e^{-\lambda t} u(\gamma(t)). \end{aligned}$$

Taking inf over all controls $v(\cdot)$, by definition of $u(x)$, we obtain LHS $\geq$ RHS in (2.6).

Conversely, with the previous control $v(\cdot)$ we have chosen at the beginning of the proof, given any $\varepsilon > 0$, let $w(\cdot)$ be a control such that

$$u(\gamma(t)) > J(\gamma(t), w(\cdot)) - \varepsilon = \int_0^\infty e^{-\lambda s} f\,\big(y_{\gamma(t), w(\cdot)}(s), w(s)\big)\; ds - \varepsilon.$$

Our goal is connect two controls $v(\cdot)$ from $[0, t]$ with $w(\cdot)$ on $[t, \infty)$ to form a new control. Let us define $z = y_{x,v(\cdot)}(t) = \gamma(t)$, and

$$\begin{cases} v^*(s) &= v(s) \qquad\qquad \text{if } s \in [0, t], \\ v^*(s) &= w(s-t) \qquad \text{if } s \in [t, \infty). \end{cases}$$

See Figure 2.1. Then, by the uniqueness of solution of (2.5), it is clear that

$$\begin{cases} y_{x,v^*(\cdot)}(s) \equiv y_{x,v(\cdot)}(s) & \text{for all } s \in [0, t], \\ y_{x,v^*(\cdot)}(s) \equiv y_{z,w(\cdot)}(s-t) & \text{for all } s \in [t, \infty). \end{cases}$$

Notice that

$$\int_0^t e^{-\lambda s} f\,(y_{x,v(\cdot)}(s), v(s))\, ds = \int_0^t e^{-\lambda s} f\,(y_{x,v^*(\cdot)}(s), v^*(s))\, ds,$$

and

$$\begin{aligned} e^{-\lambda t} u(y_{x,v(\cdot)}(t)) &\geq e^{-\lambda t} \int_0^\infty e^{-\lambda s} f\,\big(y_{\gamma(t), w(\cdot)}(s), w(s)\big)\; ds - e^{-\lambda t}\varepsilon \\ &= \int_t^\infty e^{-\lambda \zeta} f\Big(y_{x,v^*(\cdot)}(\zeta), v^*(\zeta)\Big)\, d\zeta - e^{-\lambda t}\varepsilon. \end{aligned}$$

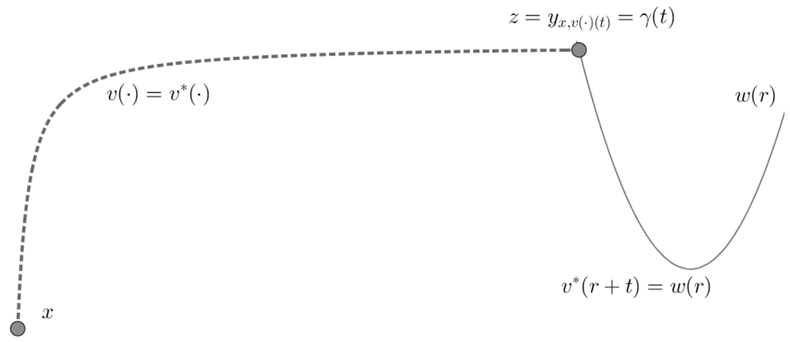


Figure 2.1: Connecting two controls $v(\cdot)$ and $w(\cdot)$ to form a new control $v^*(\cdot)$.

Thus, by combining these facts, we obtain

$$\begin{aligned}\int_0^t e^{-\lambda s} f(y_{x,v(\cdot)}(s), v(s))\, ds + e^{-\lambda t} u(y_{x,v(\cdot)}(t)) &\geq \int_0^\infty e^{-\lambda s} f\Big(y_{x,v^*(\cdot)}(s), v^*(s)\Big)\, ds - e^{-\lambda t}\varepsilon \\ &\geq u(x) - e^{-\lambda t}\varepsilon.\end{aligned}$$

Taking inf over all control $v(\cdot)$ we obtain RHS $\geq$ LHS $- e^{-\lambda t}\varepsilon$. Since this is true for all $\varepsilon > 0$, we deduce that RHS $\geq$ LHS, and the proof is complete. $\square$

Remark 2.6. It is worth noting that we require here that V is a compact metric space, and $b(x,v), f(x,v)$ are continuous, bounded, and Lipschitz in x. In particular, H is convex, and has linear growth in p.
For example, if $V = \overline{B(0,1)} \subset \mathbb{R}^n$, and $b(x,v) = v$, $f(x,v) = f(x)$ for all $(x,v) \in \mathbb{R}^n \times V$ with $f \in \mathrm{BUC}(\mathbb{R}^n)$, then

$$H(x,p) = \sup_{v \in \overline{B(0,1)}} [-v \cdot p - f(x)] = |p| - f(x).$$

We will come back to discuss this point, and relate the story between Lipschitz regularity of the viscosity solution and compactness of V.

Remark 2.7. Why DPP is good?

- Using DPP, we can find the corresponding PDE for $u(x)$.
- Using DPP, we are able to derive some first results on the regularity of $u(x)$.

Theorem 2.8 (Regularity of the value function based on DPP)**.** *Let u be defined as in* (2.2). *Set $\lambda_0 = \|D_x b(\cdot,\cdot)\|_{L^\infty(\mathbb{R}^n \times V)}$. Then, $\|u\|_{L^\infty(\mathbb{R}^n)} \leq \frac{C}{\lambda}$. Furthermore, we have the following results.*

(a) If $\lambda > \lambda_0$, then $u \in C^{0,1}(\mathbb{R}^n) = \mathrm{Lip}(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n)$.

(b) If $\lambda = \lambda_0$, then $u \in C^{0,\alpha}(\mathbb{R}^n)$ for any $0 < \alpha < 1$.

(c) If $0 < \lambda < \lambda_0$, then $u \in C^{0,\frac{\lambda}{\lambda_0}}(\mathbb{R}^n)$.

In particular, in all cases, $u \in \mathrm{BUC}(\mathbb{R}^n)$.

The proof of this theorem is rather clear and interesting, and we leave it as an exercise for the readers.

2.1 Problems

Exercise 21. *Prove Theorem 2.8 by using* (2.6).

3 Static Hamilton–Jacobi equation for the value function

Let us recall the definition of the value function u. For $x \in \mathbb{R}^n$,

$$u(x) = \inf_{v(\cdot)} \int_0^\infty e^{-\lambda s} f\Big(y_{x,v(\cdot)}(s), v(s)\Big)\, ds.$$

Besides, the Dynamic Programming Principle (DPP) reads

$$u(x) = \inf_{v(\cdot)} \left(\int_0^t e^{-\lambda s} f\Big(y_{x,v(\cdot)}(s), v(s)\Big)\, ds + e^{-\lambda t} u\Big(y_{x,v(\cdot)}(t)\Big) \right).$$

Remark 2.9. Recall that Theorem 2.8 gives us that $u \in \mathrm{BUC}(\mathbb{R}^n)$. This enables us to fit u well into the theory of continuous viscosity solutions.

Theorem 2.10. *The value function u is a viscosity solution of the following static Hamilton–Jacobi equation*

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

where, for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$H(x,p) = \sup_{v \in V} \Big(-b(x,v) \cdot p - f(x,v) \Big).$$

Proof. We divide the proof into two steps.

SUBSOLUTION TEST. Let $\varphi \in C^1(\mathbb{R}^n)$ such that $u - \varphi$ has a strict maximum at $x_0 \in \mathbb{R}^n$, and $u(x_0) = \varphi(x_0)$. Our goal is to show that

$$\lambda u(x_0) + H\Big(x_0, D\varphi(x_0)\Big) \le 0. \tag{2.7}$$

Pick a control $v(\cdot)$, and let $\gamma(\cdot) = y_{x_0,v(\cdot)}(\cdot)$ be the solution to $\gamma'(s) = b(\gamma(s), v(s))$ with $\gamma(0) = x_0$. For every $t > 0$ since $u(\gamma(t)) \le \varphi(\gamma(t))$, by DPP, we have

$$\begin{aligned} \varphi\big(\gamma(0)\big) = u(x_0) &\le \int_0^t e^{-\lambda s} f\big(\gamma(s), v(s)\big)\, ds + e^{-\lambda t} u\big(\gamma(t)\big) \\ &\le \int_0^t e^{-\lambda s} f\big(\gamma(s), v(s)\big)\, ds + e^{-\lambda t} \varphi\big(\gamma(t)\big). \end{aligned}$$

By the fundamental theorem of calculus for $s \mapsto e^{-\lambda s}\varphi(\gamma(s))$, the above can be written as

$$-\int_0^t \frac{d}{ds}\Big(e^{-\lambda s}\varphi\big(\gamma(s)\big)\Big)\, ds = \varphi(\gamma(0)) - e^{-\lambda t}\varphi\big(\gamma(t)\big) \le \int_0^t e^{-\lambda s} f\big(\gamma(s), v(s)\big)\, ds,$$

which is equivalent to

$$\int_0^t e^{-\lambda s}\Big(\lambda\varphi(\gamma(s)) + \Big[-b\big(\gamma(s), v(s)\big) \cdot D\varphi(\gamma(s)\big) - f\big(\gamma(s), v(s)\big) \Big] \Big)\, ds \le 0.$$

This holds for every control $v(\cdot)$ and every $t > 0$. Now pick the control $v(\cdot) \equiv v$ to be constant for all time for some $v \in V$, then the above formula gives

$$\frac{1}{t}\int_0^t e^{-\lambda s}\Big(\lambda\varphi(\gamma(s)) + \Big[-b\Big(\gamma(s), v\Big)\cdot D\varphi(\gamma(s)\Big) - f\Big(\gamma(s), v\Big)\Big]\Big)\, ds \le 0.$$

Let $t \to 0+$ to yield

$$\lambda\varphi(x_0) + \Big[-b(x_0, v)\cdot D\varphi(x_0) - f(x_0, v)\Big] \le 0.$$

Taking sup over all $v \in V$ in the above inequality to get (2.7).

SUPERSOLUTION TEST. Let $\psi \in C^1(\mathbb{R}^n)$ such that $u - \psi$ has a strict minimum at $x_0 \in \mathbb{R}^n$, and $u(x_0) = \psi(x_0)$. We aim at proving that

$$\lambda u(x_0) + H\Big(x_0, D\psi(x_0)\Big) \ge 0. \tag{2.8}$$

We note first that, for any $t > 0$,

$$\begin{aligned}\psi(x_0) = u(x_0) &= \inf_{v(\cdot)}\left\{\int_0^t e^{-\lambda s} f\Big(y_{x_0,v(\cdot)}(s), v(s)\Big)\, ds + e^{-\lambda t} u\Big(y_{x_0,v(\cdot)}(t)\Big)\right\}\\ &\ge \inf_{v(\cdot)}\left\{\int_0^t e^{-\lambda s} f\Big(y_{x_0,v(\cdot)}(s), v(s)\Big)\, ds + e^{-\lambda t} \psi\Big(y_{x_0,v(\cdot)}(t)\Big)\right\}.\end{aligned}$$

Therefore,

$$\begin{aligned}0 &\ge \inf_{v(\cdot)}\left\{\int_0^t e^{-\lambda s} f\Big(y_{x_0,v(\cdot)}(s), v(s)\Big)\, ds + e^{-\lambda t}\psi\Big(y_{x_0,v(\cdot)}(t)\Big) - \psi\Big(y_{x_0,v(\cdot)}(0)\Big)\right\}\\ &= -\sup_{v(\cdot)} \mathcal{K}_t[v(\cdot)],\end{aligned}$$

where

$$\begin{aligned}&\mathcal{K}_t[v(\cdot)]\\ &= \int_0^t e^{-\lambda s}\Big(\lambda\psi\Big(y_{x_0,v(\cdot)}(s)\Big) + \Big[-b\Big(y_{x_0,v(\cdot)}(s), v(s)\Big)\cdot D\psi\Big(y_{x_0,v(\cdot)}(s)\Big) - f\Big(y_{x_0,v(\cdot)}(s), v(s)\Big)\Big]\Big)\, ds\\ &\le \int_0^t e^{-\lambda s}\Big(\lambda\psi\Big(y_{x_0,v(\cdot)}(s)\Big) + H\Big(y_{x_0,v(\cdot)}(s), D\psi\Big(y_{x_0,v(\cdot)}(s)\Big)\Big)\Big)\, ds.\end{aligned}$$

By Lemma 2.1, for any control $v(\cdot)$, one has

$$\Big|y_{x_0,v(\cdot)}(t) - x_0\Big| \le Ct. \tag{2.9}$$

Thus, for $s \in [0, t]$,

$$\Big|\psi\Big(y_{x_0,v(\cdot)}(s)\Big) - \psi(x_0)\Big| \le C\,\Big|y_{x_0,v(\cdot)}(s) - x_0\Big| \le Cs \le Ct,$$

and similarly,

$$\Big|H\Big(y_{x_0,v(\cdot)}(s), D\psi(y_{x_0,v(\cdot)}(s))\Big) - H(x_0, D\psi(x_0))\Big| \le Cs \le Ct$$

as well. Hence,

$$\begin{aligned}\mathcal{K}_t[v(\cdot)] &\le \int_0^t e^{-\lambda s}\Big(\lambda\psi\big(y_{x_0,v(\cdot)}(s)\big)+H\Big(y_{x_0,v(\cdot)}(s),D\psi\big(y_{x_0,v(\cdot)}(s)\big)\Big)\Big)\,ds\\ &\le \int_0^t e^{-\lambda s}\Big(\lambda\psi(x_0)+H\big(x_0,D\psi(x_0)\big)\Big)\,ds+Ct\int_0^t e^{-\lambda s}\,ds.\end{aligned}$$

Combine this with the above to deduce that

$$\begin{aligned}0 &\le \lim_{t\to 0+}\left(\frac{1}{t}\sup_{v(\cdot)}\mathcal{K}_t[v(\cdot)]\right)\\ &\le \lim_{t\to 0+}\left(\frac{1}{t}\int_0^t e^{-\lambda s}\Big(\lambda\psi(x_0)+H\big(x_0,D\psi(x_0)\big)\Big)\,ds+C\int_0^t e^{-\lambda s}\,ds\right)\\ &= \lambda\psi(x_0)+H\big(x_0,D\psi(x_0)\big).\end{aligned}$$

The proof is complete. □

4 Legendre's transform

We consider the Hamiltonian $H(x,p):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ satisfying

$$\begin{cases} H\in C^1(\mathbb{R}^n\times\mathbb{R}^n),\ H\in \mathrm{BUC}(\mathbb{R}^n\times B(0,R)) \text{ for each } R>0,\\ p\mapsto H(x,p) \text{ is convex for all } x\in\mathbb{R}^n,\\ H \text{ is superlinear in } p, \text{ that is, } \lim\limits_{|p|\to\infty}\left(\inf\limits_{x\in\mathbb{R}^n}\dfrac{H(x,p)}{|p|}\right)=+\infty. \end{cases} \tag{2.10}$$

It is clear that superlinearity is stronger than coercivity.

Example 2.2. *Consider $H(x,p)=|p|^m+V(x)$ for all $(x,p)\in\mathbb{R}^n\times\mathbb{R}^n$ where $V\in\mathrm{BUC}(\mathbb{R}^n)$. Then H is convex in p if and only if $m\ge 1$.*

- *If $m>1$, then H is superlinear in p.*
- *If $m=1$, then H has linear growth. It is coercive, but not superlinear in p.*
- *If $m>2$, we say that H is superquadratic in p. And if $m<2$, we say that H is subquadratic in p. Of course, if $m=2$, then H is quadratic in p.*

Definition 2.11 (Legendre's transform)**.** *For the Hamiltonian $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$, we define its Legendre's transform $H^*=L:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ as*

$$L(x,v)=\sup_{p\in\mathbb{R}^n}\big(p\cdot v-H(x,p)\big).$$

Some further deep characterizations of the Legendre transform are given in Appendix.

Remark 2.12.

- In physics, we regard x as the position of a particle, and v as its corresponding velocity.

- We need to check the above definition is well-defined, that is, $L(x,v)$ is indeed finite.

Example 2.3. *For the classical mechanics Hamiltonian*

$$H(x,p) = \frac{1}{2}|p|^2 + V(x) \qquad \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n,$$

we have

$$\begin{aligned} L(x,v) &= \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - H(x,p)\Big) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - \frac{1}{2}|p|^2\Big) - V(x) \\ &= \sup_{p\in\mathbb{R}^n}\Big(\frac{1}{2}|v|^2 - \frac{1}{2}|p-v|^2\Big) - V(x) = \frac{1}{2}|v|^2 - V(x). \end{aligned}$$

Thus, $H^*(x,v) = L(x,v) = \frac{1}{2}|v|^2 - V(x)$. It is worth noting that in this case, H is the total energy, and L is the difference between kinetic energy and potential energy. We also observe that $H^{**} = L^* = H$.

We now have the following important result on convex duality via Legendre's transform.

Theorem 2.13. *Assume that H satisfies* (2.10)*. Then, the followings hold.*

(i) $L(x,v)$ is well-defined (finite), and $v \mapsto L(x,v)$ is convex and superlinear.

(ii) $L^* = H^{**} = H$.

In fact, the above theorem holds without the assumption that $H \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$. We just put it there to simplify our proof.

Proof. Let us proceed step by step.

(i) Fix $x, v \in \mathbb{R}^n$. Since H is superlinear in p, as $|p| \to \infty$, we have

$$p\cdot v - H(x,p) = |p|\left(\frac{p\cdot v}{|p|} - \frac{H(x,p)}{|p|}\right) \to -\infty,$$

which means that $\sup_{p\in\mathbb{R}^n}\Big(p\cdot q - H(x,p)\Big) = \max_{p\in\mathbb{R}^n}\Big(p\cdot q - H(x,p)\Big) < \infty$. It is clear that $v \mapsto L(x,v)$ is convex as it is a supremum of a family of affine functions in v.

Now, we prove that L is superlinear in v. For $v \neq 0$, choose $p = s\frac{v}{|v|}$, then for any $s > 0$, we have

$$L(x,v) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - H(x,p)\Big) \geq \left(s\frac{v}{|v|}\right)\cdot v - H\left(x, s\frac{v}{|v|}\right) \geq s|v| - \max_{|p|\leq s} H(x,p).$$

Thus, for any fixed $s > 0$,

$$\liminf_{|v|\to\infty}\frac{L(x,v)}{|v|} \geq s - \limsup_{|v|\to\infty}\left(\frac{1}{|v|}\max_{|p|\leq s}H(x,p)\right) = s \qquad \Longrightarrow \qquad \lim_{|v|\to\infty}\frac{L(x,v)}{|v|} = +\infty$$

uniformly for $x \in \mathbb{R}^n$.

(ii) We proceed to show that $L^* = H$. Note that

$$L(x,v) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - H(x,p)\Big) \geq p\cdot v - H(x,p) \qquad \text{for any } p\in\mathbb{R}^n.$$

This implies

$$H(x,p) + L(x,v) \geq p\cdot v \qquad \text{for all } x,p,q\in\mathbb{R}^n. \tag{2.11}$$

In particular,

$$H(x,p) \geq \sup_{v\in\mathbb{R}^n}(p\cdot v - L(x,v)) = L^*(x,p).$$

Thus $H \geq L^*$. Conversely, we have

$$\begin{aligned} L^*(x,p) &= \sup_{v\in\mathbb{R}^n}\Big(p\cdot v - L(x,v)\Big) = \sup_{v\in\mathbb{R}^n}\Big(p\cdot v - \sup_{r\in\mathbb{R}^n}\Big(r\cdot v - H(x,r)\Big)\Big) \\ &= \sup_{v\in\mathbb{R}^n}\inf_{r\in\mathbb{R}^n}\Big((p-r)\cdot v + H(x,r)\Big). \end{aligned}$$

Thus

$$L^*(x,p) \geq \inf_{r\in\mathbb{R}}\Big(H(x,r) - (r-p)\cdot v\Big) \qquad \text{for all } v\in\mathbb{R}^n.$$

Pick $v = D_pH(x,p)$. By the convexity of H in p,

$$H(x,r) - (r-p)\cdot v = H(x,r) - (r-p)\cdot D_pH(x,p) \geq H(x,p) \qquad \text{for all } r\in\mathbb{R}^n.$$

Therefore, $L^* \geq H$. We conclude that $L^* = H^{**} = H$.

□

Remark 2.14. We have some further comments about the convexity of H and L.

- (2.11) is an important inequality in the convex duality between H and L.
- In case that H is not C^1, we can always pick $v \in \mathbb{R}^n$ such that $v \in D_p^- H(x,p) = \partial_p H(x,p)$, which is the subgradient set of H in p at (x,p), in the last step of the above proof to finish.
- By Radamacher's theorem, as $p \mapsto H(x,p)$ is convex for each $x\in\mathbb{R}^n$, $H(x,\cdot)$ is also locally Lipschitz, hence is differentiable almost everywhere.
- Furthermore, by Alexandrov's theorem, for each $x\in\mathbb{R}^n$, $H(x,\cdot)$ is twice differentiable almost everywhere.

4.1 Problems

Exercise 22. *Compute the Legendre transform $L(x,v)$ of the Hamiltonian $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$, where*

$$H(x,p) = \frac{|p|^m}{m} + V(x) \qquad \textit{for all } (x,p)\in\mathbb{R}^n\times\mathbb{R}^n.$$

Here, $m \geq 1$ and $V \in \mathrm{BUC}(\mathbb{R}^n)$.

Exercise 23. *Find out when the equality in* (2.11) *holds.*

5 The optimal control formula from the Lagrangian viewpoint

5.1 New representation formula for the solution of the static equation based on the Lagrangian

We have the duality between H and L as following

$$\begin{cases} p \mapsto H(x,p) \text{ is convex,} \\ H(x,p) \text{ is superlinear in } p, \end{cases} \quad \xleftrightarrow{\text{Legendre's transform}} \quad \begin{cases} v \mapsto L(x,v) \text{ is convex,} \\ L(x,v) \text{ is superlinear in } v. \end{cases}.$$

Recall that

$$H(x,p) = \sup_{v\in\mathbb{R}^n}(p\cdot v - L(x,v)).$$

When $p \mapsto H(x,p)$ is convex, we are able to use Legendre's transform obtain the Lagrangian L, and get another representation formula (still optimal control formula) for the unique viscosity solution to the corresponding static equation. The new formula is defined in term of the Lagrangian, and not in term of the controls.

Theorem 2.15. *Fix $\lambda > 0$. Consider the following static Hamilton–Jacobi equation*

$$\lambda u + H(x,Du) = 0 \qquad \text{in } \mathbb{R}^n \tag{2.12}$$

Assume that the Hamiltonian H satisfies

$$\begin{cases} H \in C^2(\mathbb{R}^n\times\mathbb{R}^n),\ H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for each } R > 0, \\ p \mapsto H(x,p) \text{ is convex for all } x \in \mathbb{R}^n, \\ \displaystyle\lim_{|p|\to\infty}\left(\inf_{x\in\mathbb{R}^n}\frac{H(x,p)}{|p|}\right) = +\infty. \end{cases} \tag{2.13}$$

Then, the following function is a viscosity solution of (2.12)

$$u(x) = \inf\left\{\int_0^\infty e^{-\lambda s}L\Big(\gamma(s),-\gamma'(s)\Big)\,ds \,:\, \gamma(0)=x, \gamma'(\cdot)\in L^1([0,T]) \text{ for any } T>0\right\}. \tag{2.14}$$

We skip the proof of this theorem for now as it follows the same lines as that of Theorem 2.10. It is in fact interesting to go through its proof to compare the differences.

Remark 2.16. Some points are worth to be mentioned here.

- Where are the controls in (2.14)? In this representation formula, the controls are included in the Lagrangian $L(x,v)$, and they are basically the admissible velocities $\gamma'(\cdot)$ of admissible curves.

  In fact, this is the optimal control setting where $V = \mathbb{R}^n$, which is not compact, and

  $$b(x,v) = v, \qquad f(x,v) = L(x,-v) \qquad \text{for all } (x,v)\in\mathbb{R}^n\times\mathbb{R}^n.$$

- Under the dynamical system viewpoint, we are interested in finding the optimal paths γ so that

  $$u(x) = \int_0^\infty e^{-\lambda s}L\Big(\gamma(s),-\gamma'(s)\Big)\,ds.$$

  The existence of such minimizers comes from Calculus of Variations.

5.2 The representation formula for the solution of the Cauchy problem based on the Lagrangian

Consider the usual Cauchy problem

$$\begin{cases} u_t(x,t) + H(x,Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{2.15}$$

For the Hamiltonian H, we assume that it satisfies (2.13), that is,

$$\begin{cases} H \in C^2(\mathbb{R}^n \times \mathbb{R}^n),\ H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for each } R > 0, \\ p \mapsto H(x,p) \text{ is convex for all } x \in \mathbb{R}^n, \\ \lim_{|p|\to\infty} \left(\inf_{x\in\mathbb{R}^n} \dfrac{H(x,p)}{|p|} \right) = +\infty. \end{cases}$$

For the initial data u_0, we assume as usual that $u_0 \in \mathrm{BUC}(\mathbb{R}^n)$. As usual, let L be the Legendre transform of H, that is,

$$L(x,v) = \sup_{p\in\mathbb{R}^n} (p \cdot v - H(x,p)) \qquad \text{for all } (x,v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Lemma 2.17 (Properties of L)**.** *Assume* (2.13). *Then, L also satisfies*

$$\begin{cases} L \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \textit{ for each } R > 0, \\ v \mapsto L(x,v) \textit{ is convex for all } x \in \mathbb{R}^n, \\ \lim_{|v|\to\infty} \left(\inf_{x\in\mathbb{R}^n} \dfrac{L(x,v)}{|v|} \right) = +\infty. \end{cases}$$

Besides, there exists $C > 0$ such that

$$|\xi| \le C \qquad \textit{for all } \xi \in D_v^- L(x,0),\ x \in \mathbb{R}^n.$$

Proof. We only need to check the last claim. For each fix $x \in \mathbb{R}^n$, let $\xi \in D_v^- L(x,0)$. Then for all $v \in \mathbb{R}^n$,

$$L(x,v) \ge L(x,0) + \xi \cdot v.$$

Consider only $v \in \mathbb{R}^n$ with $|v| = 1$ to yield

$$|\xi| = \sup_{|v|=1} \xi \cdot v \le \sup_{|v|=1} L(x,v) - L(x,0) \le 2 \sup \left\{ |L(x,v)| \,:\, x \in \mathbb{R}^n, |v| \le 1 \right\} \le C.$$

□

Remark 2.18. Let us note that under assumption (2.13), we actually have furthermore that $L \in C^2(\mathbb{R}^n \times \mathbb{R}^n)$. In particular, $D_v^- L(x,v) = \{D_v L(x,v)\}$ for all $(x,v) \in \mathbb{R}^n \times \mathbb{R}^n$. As we do not need to use this fact here, we skip it.

We are now ready to define the following value function, which is of finite horizon type.

Definition 2.19. *For each $(x,t) \in \mathbb{R}^n \times [0,\infty)$, we denote by*

$$u(x,t) = \inf \left\{ \int_0^t L(\gamma(s), \gamma'(s))\, ds + u_0(\gamma(0)) \,:\, \gamma(t) = x, \gamma'(\cdot) \in L^1([0,t]) \right\}. \tag{2.16}$$

Remark 2.20. It is very important noticing that $\gamma'(\cdot)$ is integrable on $[0,t]$ is equivalent to the fact that $\gamma(\cdot)$ is absolutely continuous on $[0,t]$. Thus, the value function $u(x,t)$ is chosen as the infimum value of the above cost functional among all absolutely continuous paths $\gamma(\cdot)$ with endpoint $\gamma(t)=x$.

Similarly to the static case, we have the following Dynamic Programming Principle (DPP).

Theorem 2.21 (Dynamic Programming Principle)**.** *The value function u defined above satisfies, for* $(x,t)\in\mathbb{R}^n\times(0,\infty)$*,*

$$u(x,t)=\inf\left\{\int_s^t L(\gamma(r),\gamma'(r))\,dr+u(\gamma(s),s)\,:\,\gamma(t)=x,\gamma'(\cdot)\in L^1([0,t])\right\}, \tag{2.17}$$

for all $0\le s\le t$*.*

Proof. Fix $0\le s\le t$. Let $\xi(\cdot)$ be a path on $[s,t]$ with $\xi(t)=x$ and $\xi'(\cdot)\in L^1([s,t])$. Let $\gamma(\cdot)$ be an arbitrary path on $[0,s]$ with $\gamma(s)=\xi(s)$ and $\gamma'(\cdot)\in L^1([0,s])$. Then, define $\zeta:[0,t]\to\mathbb{R}^n$ as

$$\zeta(r)=\begin{cases}\gamma(r) & r\in[0,s],\\ \xi(r) & r\in[s,t].\end{cases}$$

It is clear that $\zeta(t)=x$ and $\zeta'(\cdot)\in L^1([0,t])$. By definition of u, we have

$$\begin{aligned}u(x,t)&\le\int_0^t L(\zeta(r),\zeta'(r))\,dr+u(\zeta(0),0)\\ &=\int_s^t L(\xi(r),\xi'(r))\,dr+\int_0^s L(\gamma(r),\gamma'(r))\,dr+u(\gamma(0),0)\end{aligned}$$

Taking the infimum in the above over all paths γ on $[0,s]$ with $\gamma(s)=\xi(s)$ and $\gamma'(\cdot)\in L^1([0,s])$ to imply

$$u(x,t)\le\int_s^t L(\xi(r),\xi'(r))\,dr+u(\xi(s),s).$$

Then, taking infimum over all path ξ on $[s,t]$ to obtain

$$u(x,t)\le\inf\left\{\int_s^t L(\gamma(r),\gamma'(r))\,dr+u(\gamma(s),s)\,:\,\gamma(t)=x,\gamma'(\cdot)\in L^1([0,t])\right\}.$$

Conversely, let γ be a path with $\gamma(t)=x$ and $\gamma(\cdot)\in L^1([0,t])$. We decompose γ into $\gamma_1(\cdot)=\gamma(\cdot)|_{[0,s]}$ and $\gamma_2(\cdot)=\gamma(\cdot)|_{[s,t]}$. Then,

$$\begin{aligned}&\int_0^t L(\gamma(r),\gamma'(r))\,dr+u(\gamma(0),0)\\ =&\int_s^t L(\gamma_2(r),\gamma_2'(r))\,dr+\int_0^s L(\gamma_1(r),\gamma_1'(r))\,r+u(\gamma_1(0),0)\\ \ge&\int_s^t L(\gamma_2(r),\gamma_2'(r))\,dr+u(\gamma_2(s),s)\\ \ge&\inf\left\{\int_s^t L(\gamma_2(r),\gamma_2'(r))\,dr+u(\gamma_2(s),s)\,:\,\gamma_2(t)=x,\gamma_2'(\cdot)\in L^1([s,t])\right\}.\end{aligned}$$

Taking infimum over all path γ in $[0,t]$ we obtain

$$u(x,t) \geq \inf\left\{ \int_s^t L(\gamma(r),\gamma'(r))\,dr + u(\gamma(s),s) \,:\, \gamma(t)=x, \gamma'(\cdot)\in L^1([0,t]) \right\}.$$

□

Theorem 2.22. *Assume* (2.13). *Let u be defined as in* (2.16). *Then, u is a viscosity solution to* (2.15).

The proof is omitted here as it follows the same lines as that of Theorem 2.10 by using the Dynamic Programming Principle (2.17). It is in fact an interesting exercise for interested readers.

5.3 Problems

Exercise 24. *Prove Theorem 2.22.*

Exercise 25. *Assume* (2.13). *Let L be the Legendre transform of H. Prove that*

$$L \in C^2(\mathbb{R}^n \times \mathbb{R}^n).$$

Exercise 26. *Assume* (2.13). *Let L be the Legendre transform of H. Fix $R > 0$. Show that there exists $C_R > 0$ such that*

$$L(x,v) = \max_{|p|\leq C_R} (p\cdot v - H(x,p)) \qquad \text{for all } (x,v)\in\mathbb{R}^n\times\overline{B}(0,R).$$

5.4 The Hopf–Lax formula

We now consider the spatially homogeneous Hamiltonian $H(x,p)=H(p)$ for $(x,p)\in\mathbb{R}^n\times\mathbb{R}^n$. Here, by spatially homogeneous H, we mean that it does not depend of the spatial variable (location) x.
We assume here that

$$\begin{cases} p\mapsto H(p) \text{ is convex,} \\ \lim\limits_{|p|\to\infty} \dfrac{H(p)}{|p|} = +\infty. \end{cases} \tag{2.18}$$

Let $L = L(v):\mathbb{R}^n\to\mathbb{R}$ be the corresponding Lagrangian, that is, $L=H^*$. Then clearly,

$$\begin{cases} v\mapsto L(v) \text{ is convex,} \\ \lim\limits_{|v|\to\infty} \dfrac{L(v)}{|v|} = +\infty. \end{cases}$$

Theorem 2.23 (The Hopf–Lax formula)**.** *Assume* (2.18). *Let u be the viscosity solution to*

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

Here, the initial data $u_0\in \mathrm{BUC}(\mathbb{R}^n)$. Then, u has the following representation formula. For $(x,t)\in\mathbb{R}^n\times(0,\infty)$,

$$u(x,t) = \inf_{y\in\mathbb{R}^n}\left\{ tL\left(\frac{x-y}{t}\right)+u_0(y)\right\} = \min_{y\in\mathbb{R}^n}\left\{ tL\left(\frac{x-y}{t}\right)+u_0(y)\right\}. \tag{2.19}$$

Formula (2.19) is known as the celebrated Hopf–Lax formula.

Proof. Fix $(x,t) \in \mathbb{R}^n \times (0,\infty)$. For each $y \in \mathbb{R}^n$, let us consider the path γ as the straight line segment connecting $(y,0)$ with (x,t), that is,

$$\gamma(s) = y + s\left(\frac{x-y}{t}\right) \qquad \text{for all } s \in [0,t].$$

The optimal control formula (2.16) gives

$$u(x,t) \le \int_0^t L(\gamma'(s))\,ds + u_0(\gamma(0)) = tL\left(\frac{x-y}{t}\right) + u_0(y),$$

and thus,

$$u(x,t) \le \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + u_0(y)\right\}.$$

On the other hand, if γ is any admissible path with $\gamma(t) = x$, then by Jensen's inequality, we get

$$L\left(\frac{1}{t}\int_0^t \gamma'(s)\,ds\right) \le \frac{1}{t}\int_0^t L(\gamma'(s))\,ds.$$

For $\gamma(0) = y$, notice that

$$\int_0^t \gamma'(s)\,ds = \gamma(t) - \gamma(0) = x - y,$$

and hence,

$$tL\left(\frac{x-y}{t}\right) + u_0(y) \le \int_0^t L(\gamma'(s))\,ds + u_0(\gamma(0)).$$

From this we get

$$\inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + u_0(y)\right\} \le u(x,t).$$

Therefore,

$$u(x,t) = \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + u_0(y)\right\}.$$

Finally, as $u_0 \in \mathrm{BUC}(\mathbb{R}^n)$, and L is superlinear, it is clear that inf on the right hand side above holds at a point $y \in \mathbb{R}^n$. $\square$

Example 2.4. *We give here some well-known examples in the literature.*

- *If $H(p) = \frac{|p|^2}{2}$ for $p \in \mathbb{R}^n$, then $L(v) = \frac{|v|^2}{2}$ for $v \in \mathbb{R}^n$. Then, the Hopf–Lax formula for solution u reads*

$$u(x,t) = \inf_{y\in\mathbb{R}^n}\left\{\frac{|x-y|^2}{2t} + u_0(y)\right\} = \min_{y\in\mathbb{R}^n}\left\{\frac{|x-y|^2}{2t} + u_0(y)\right\}.$$

- *In one dimension, let us consider the famous inviscid Burger equation*

$$\begin{cases} v_t(x,t) + v(x,t)v_x(x,t) &= 0 \qquad \textit{in } \mathbb{R} \times (0,\infty), \\ \qquad v(x,0) &= v_0(x) \qquad \textit{on } \mathbb{R}. \end{cases}$$

Here, initial data v_0 is nice enough. Note that

$$v_t + v v_x = 0 \qquad \Longleftrightarrow \qquad v_t + \left(\frac{v^2}{2}\right)_x = 0.$$

Take u so that $v = u_x$, then

$$u_{xt} + \left(\frac{(u_x)^2}{2}\right)_x = 0 \qquad \Longrightarrow \qquad u_t + \frac{(u_x)^2}{2} = C$$

for some constant C. Let $C = 0$. Then we are able to use the Hopf–Lax formula for u to obtain the formula for v as

$$v(x,t) = \frac{d}{dx}\left(\inf_{y\in\mathbb{R}}\left\{\frac{|x-y|^2}{2t} + u_0(y)\right\}\right).$$

Here, $u_0(y) = \int_0^y v_0(x)\,dx$ for all $y \in \mathbb{R}$. This formula for v turns out to be the Lax–Oleinik formula.

5.5 First-order front propagation problem

Let us now recall the first-order front propagation problem that was discussed in Example 1.1. The corresponding equation reads

$$\begin{cases} u_t(x,t) + a(x)|Du| &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{2.20}$$

Here, we assume that $a : \mathbb{R}^n \to \mathbb{R}$ is a given continuous function, and there exist $\alpha, \beta > 0$ such that

$$\alpha \le a(x) \le \beta \qquad \text{for all } x \in \mathbb{R}^n.$$

For the initial data u_0, we assume as usual that $u_0 \in \mathrm{BUC}(\mathbb{R}^n)$.

It is clear in this case that the Hamiltonian $H(x,p) = a(x)|p|$ for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$ is convex and uniformly coercive, but is not superlinear in p. Of course, (2.13) does not hold here. Nevertheless, we are still able to write down a representation formula for solution u of (2.20) based on the Lagrangian/optimal control formulation.

The Lagrangian $L(x,v)$ can be computed as following

$$\begin{aligned} L(x,v) &= \sup_{p\in\mathbb{R}^n}(p\cdot v - H(x,p)) = \sup_{p\in\mathbb{R}^n}(p\cdot v - a(x)|p|) \\ &= \begin{cases} 0 & \text{if } |v| \le a(x), \\ +\infty & \text{if } |v| > a(x). \end{cases} \end{aligned}$$

Although L is singular in the above, it is still convex in v. Then, in this setting, formula (2.16) is rewritten as

$$
\begin{aligned}
u(x,t) &= \inf\left\{\int_0^t L(\gamma(s),\gamma'(s))\,ds + u_0(\gamma(0))\,:\,\gamma(t)=x,\gamma'(\cdot)\in L^1([0,t])\right\}.\\
&= \inf\left\{u_0(\gamma(0))\,:\,\gamma(t)=x,|\gamma'(s)|\le a(\gamma(s)) \quad \text{for a.e. } s\in[0,t]\right\}.
\end{aligned}
$$

It turns out that this formula still gives the unique viscosity solution u to (2.20) as stated in the following theorem.

Theorem 2.24. *Let u be the unique viscosity solution to* (2.20) *with given conditions on a and u_0 as stated in this section. Then, u has the following representation formula, for $(x,t)\in \mathbb{R}^n\times(0,\infty)$,*

$$
u(x,t) = \min\left\{u_0(\gamma(0))\,:\,\gamma(t)=x,|\gamma'(s)|\le a(\gamma(s)) \quad \textit{for a.e. } s\in[0,t]\right\}.
$$

The proof of this theorem is left as an exercise. It is worth noting that, because of symmetry, by defining $\eta(s)=\gamma(t-s)$ for all $s\in[0,t]$, we can also write that

$$
\begin{aligned}
u(x,t) &= \min\left\{u_0(\gamma(0))\,:\,\gamma(t)=x,|\gamma'(s)|\le a(\gamma(s)) \quad \text{for a.e. } s\in[0,t]\right\}\\
&= \min\left\{u_0(\eta(t))\,:\,\eta(0)=x,|\eta'(s)|\le a(\eta(s)) \quad \text{for a.e. } s\in[0,t]\right\}.
\end{aligned}
$$

We now give an immediate consequence of the above theorem.

Corollary 2.25. *Assume $a(x)=\alpha$ for all $x\in\mathbb{R}^n$ for some given $\alpha>0$, and $u_0\in \mathrm{BUC}(\mathbb{R}^n)$. Then, the unique viscosity solution u to* (2.20) *has the following representation formula, for $(x,t)\in\mathbb{R}^n\times(0,\infty)$,*

$$
u(x,t) = \min\{u_0(y)\,:\,|y-x|\le \alpha t\} = \min\{u_0(y)\,:\,y\in \overline{B}(x,\alpha t)\}.
$$

Theorem 2.24 and Corollary 2.25 sometimes appear in the literature under the framework of reachable sets in front propagations.

5.6 Problems

Exercise 27. *Prove Theorem 2.24 by writing down its corresponding Dynamic Programming Principle.*

Exercise 28. *Let us consider the first-order front propagation problem that was discussed in Example 1.1. Assume $a(x)=\alpha$ for all $x\in\mathbb{R}^n$ for some given $\alpha>0$, and $\Gamma_0=\partial([-1,1]^n)$. Use Corollary 2.25 to describe the behavior of $\{\Gamma_t\}_{t\ge 0}$.*

6 A further hidden structure of convex first-order Hamilton–Jacobi equations

6.1 A characterization of subsolutions of convex first-order Hamilton–Jacobi equations

Fix $\lambda\ge 0$. We consider the following usual static problem

$$
\lambda u + H(x,Du) = 0 \qquad \text{in } \mathbb{R}^n. \tag{2.21}
$$

We assume throughout this section that

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for all } R > 0, \\ p \mapsto H(x,p) \text{ is convex for all } x \in \mathbb{R}^n, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} H(x,p) = +\infty. \end{cases} \tag{2.22}$$

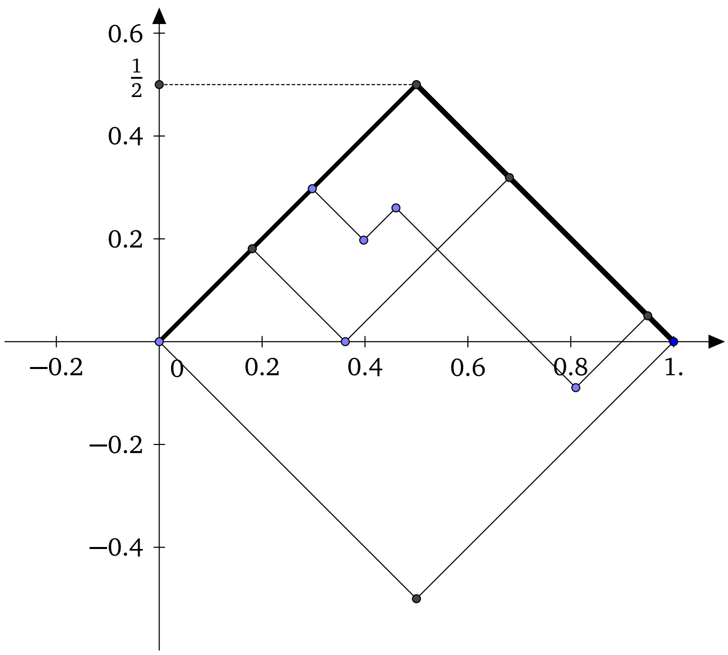
Remark 2.26. Under (2.22), for $\lambda > 0$, we apply the Perron method (Theorem 1.26) to imply that (2.21) has a unique Lipschitz viscosity solution $u \in \mathrm{Lip}(\mathbb{R}^n)$. Of course, this means that u is differentiable a.e. in $\mathbb{R}^n$.
If $\lambda = 0$, (2.21) is not monotone in u anymore, then anything can happen. For example, if $H(x,p) > 0$ for all $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$, then (2.21) does not have any solution. It could be also the case that (2.21) has infinitely many solutions, and we will discuss this point later in the book.

We focus here on viscosity subsolutions, and let us recall the following example.

Example 2.5. *Recall the eikonal equation in one dimension*

$$\begin{cases} |u'(x)| &= 1 \qquad \text{in } (0,1), \\ u(0) = u(1) &= 0. \end{cases} \tag{2.23}$$

Of course, here, $H(p) = |p|$*, and* $\lambda = 0$*. The following graph describes various a.e. solutions to* (2.23)*. As discussed in Exercise 1, such a.e. solutions are actually viscosity subsolutions. This fact can be checked quickly in a geometric way as following. Take one such a.e. solution* u*, whose graph consists of line segments of slopes* ± 1 *and corners from below and above. There is nothing to check at the corners from below as we cannot touch them from above by smooth functions. For the corners from above, every function that touches it from above there has slope between* -1 *and* 1*, and thus, the viscosity subsolution test is satisfied.*



Our goal now is to show that the above observation holds true for general convex cases. The following result is due to Barron and Jensen [18].

Theorem 2.27. *Assume $\lambda \geq 0$, and H satisfies* (2.22). *Then, the following claims are equivalent*

(i) $u \in \mathrm{Lip}(\mathbb{R}^n)$ *is viscosity subsolution of* (2.21).

(ii) $u \in \mathrm{Lip}(\mathbb{R}^n)$ *is an almost everywhere subsolution of* (2.21).

Proof. The implication (i) $\Longrightarrow$ (ii) was already done earlier.
For the converse, we need to smooth u up and use stability results of viscosity subsolutions. We use the convolution trick as following. Take η to be the standard mollifier, that is,

$$\eta \in C_c^\infty(\mathbb{R}^n, [0,\infty)), \qquad \mathrm{supp}(\eta) \subset B(0,1), \qquad \int_{\mathbb{R}^n} \eta(x)\, dx = 1.$$

For $\varepsilon > 0$, denote by $\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(\frac{x}{\varepsilon}\right)$ for all $x \in \mathbb{R}^n$. Set

$$u^\varepsilon(x) = (\eta_\varepsilon \star u)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)u(y)\, dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x-y)u(y)\, dy \quad \text{for } x \in \mathbb{R}^n.$$

Then $u^\varepsilon \in C^\infty(\mathbb{R}^n)$, and $u^\varepsilon \to u$ locally uniformly as $\varepsilon \to 0$. Since $u \in \mathrm{Lip}(\mathbb{R}^n)$ is an almost everywhere subsolution of (2.21), we multiply η_ε to both sides of (2.21) and integrate on $\mathbb{R}^n$ to yield

$$\lambda u^\varepsilon(x) + \int_{B(0,\varepsilon)} H(x-y, Du(x-y))\eta_\varepsilon(y)\, dy \leq 0$$

We need to fix x instead of $x-y$ in $H(x-y, Du(x-y))$. Denote ω_R to be the modulus of continuity of H on $\mathbb{R}^n \times B(0,R)$ where $R = \|Du\|_{L^\infty(\mathbb{R}^n)} + 1$. Then, a.e. in $B(0,\varepsilon)$, we have

$$\left|H(x-y, Du(x-y)) - H(x, Du(x-y))\right| \leq \omega_R(|y|) \leq \omega_R(\varepsilon).$$

This gives

$$\begin{aligned} 0 &\geq \lambda u^\varepsilon(x) + \int_{B(0,\varepsilon)} H(x-y, Du(x-y))\eta_\varepsilon(y)\, dy \\ &\geq \lambda u^\varepsilon(x) + \int_{B(0,\varepsilon)} \left(H(x, Du(x-y)) - \omega_R(\varepsilon)\right)\eta_\varepsilon(y)\, dy \\ &\geq \lambda u^\varepsilon(x) + H\left(x, \int_{B(0,\varepsilon)} Du(x-y)\eta_\varepsilon(y)\, dy\right) - \omega_R(\varepsilon) \\ &= \lambda u^\varepsilon(x) + H\left(x, Du^\varepsilon(x)\right) - \omega_R(\varepsilon). \end{aligned}$$

We used Jensen's inequality in the last inequality above. So, for each $\varepsilon > 0$, u^ε is a classical (smooth) subsolution to

$$\lambda u^\varepsilon(x) + H\left(x, Du^\varepsilon(x)\right) \leq \omega_R(\varepsilon) \qquad \text{in } \mathbb{R}^n.$$

We then let $\varepsilon \to 0$ and use stability results of viscosity subsolutions to conclude that u is a viscosity subsolution of (2.21). □

Remark 2.28. We have some further observations.

- Convolution with a standard mollifier is very important in the above proof, and to nonlinear PDEs in general. Whenever we need to find smooth approximations, this standard technique should be considered.
- We need some insights to deal with nonlinear terms, or terms with variable coefficients when doing convolutions. Many times, we need to handle the differences, and it is typically the case that certain commutator estimates appear naturally.

6.2 Characterization of viscosity solutions of convex first-order Hamilton–Jacobi equations

We now focus on viscosity subsolutions to (2.21).

Theorem 2.29. *Assume $\lambda \geq 0$, and H satisfies* (2.22)*. Then, the following claims are equivalent*

(i) $u \in \mathrm{Lip}(\mathbb{R}^n)$ *is viscosity solution of* (2.21).

(ii) $u \in \mathrm{Lip}(\mathbb{R}^n)$*, and for all* $x \in \mathbb{R}^n$*,* $p \in D^- u(x)$*,*

$$\lambda u(x) + H(x,p) = 0.$$

Proof. First of all, we have some elementary observations.

- If $p \mapsto H(x,p)$ is convex, then so is $p \mapsto H(x,-p)$.
- We have $q \in D^+ v(x)$ if and only if $-q = p \in D^- u(x)$ where $v = -u$.

Assume first that u is a Lipschitz viscosity solution of (2.21). For $x \in \mathbb{R}^n$ and $p \in D^- u(x)$, by supersolution test, $\lambda u(x) + H(x,p) \geq 0$. We need to show that $\lambda u(x) + H(x,p) = 0$.

As u is a Lipschitz a.e. solution of (2.21), by Rademacher's theorem, for $v = -u$, we have

$$-\lambda v(x) + H(x, -Dv(x)) = 0 \qquad \text{a.e. in } \mathbb{R}^n \qquad \Longleftrightarrow \qquad K(x, Dv(x)) = 0 \qquad \text{a.e. in } \mathbb{R}^n,$$

where $K(x,p) = -\lambda v(x) + H(x,-p)$ for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$. It is clear that K satisfies (2.22). Theorem 2.27 with $\lambda = 0$ concludes that v is a viscosity subsolution to $K(x, Dv(x)) = 0$. The viscosity subsolution test implies

$$\begin{aligned} -p \in D^+ v(x) \qquad &\Longrightarrow \qquad K(x,-p) \leq 0 \\ &\Longrightarrow \qquad -\lambda v(x) + H(x,p) \leq 0 \\ &\Longrightarrow \qquad \lambda u(x) + H(x,p) \leq 0 \qquad \Longrightarrow \qquad \lambda u(x) + H(x,p) = 0. \end{aligned}$$

Conversely, if $u \in \mathrm{Lip}(\mathbb{R}^n)$ such that for any $x \in \mathbb{R}^n$ and $p \in D^- u(x)$ then $\lambda u(x) + H(x,p) = 0$, then clearly by definition u is viscosity supersolution of (2.21). By Rademacher's theorem again, u is differentiable a.e. in $\mathbb{R}^n$, and thus

$$\lambda u(x) + H(x, Du(x)) = 0 \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

Theorem 2.27 implies automatically that u is a viscosity subsolution of (2.21), and the proof is complete. □

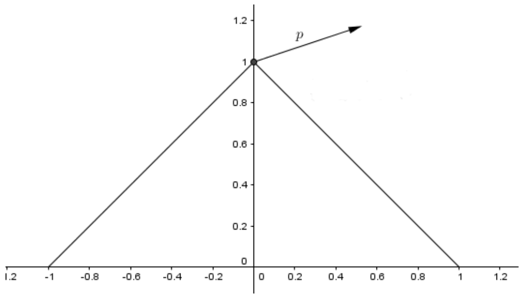


Figure 2.2: $p \in D^+u(0)$, no need to test.

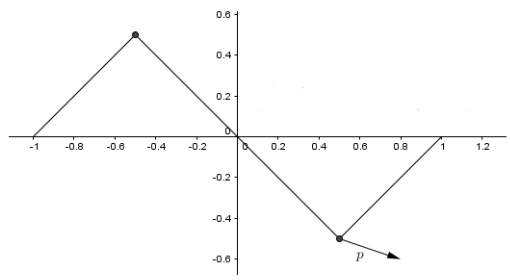


Figure 2.3: $p \in D^-u(1/2)$, and $|p| < 1$.

Remark 2.30. We have few further comments for first-order convex Hamilton–Jacobi equations.

1. There is no need to test for the supergradients $D^+u(x)$ for $x \in \mathbb{R}^n$. See Figure 2.2.

2. Criterion (ii) in Theorem 2.29 is quite important and useful. For example, we can use it to study the eikonal equation in one dimension again as following.

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

   It is clear that the function on the left $u(x) = 1-|x|$ (Figure 2.2) is the unique solution to the above. Besides, the function on the right (Figure 2.3) is not a solution as it fails (ii) at $x = 1/2$.

3. Theorems 2.27 and 2.29 only hold true for first-order equations in general. The similar results do not hold for second-order case. We will address this in an exercise later. For now, technically, we can see it as following. Let us consider

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

   This is an elliptic type problem with max principle. If we let $v = -u$, then v solves

$$H(x, -Dv(x)) + \Delta v(x) = 0 \qquad \text{in } \mathbb{R}^n,$$

   which is a wave type problem.

We have the following corollary, which is quite important for us to use later.

Corollary 2.31. *Assume $\lambda \geq 0$, and H satisfies* (2.22)*. Then, the followings hold.*

(i) If u_1, u_2 are Lipschitz solutions to (2.21)*, then* $\min\{u_1, u_2\}$ *is also a solution to* (2.21).

(ii) If $\{u_i\}_{i\in I}$ is a family of Lipschitz solutions to (2.21)*, then $u = \inf_{i\in I} u_i$ is also a solution to* (2.21) *provided u is finite and continuous.*

Note that normally (without convexity of H) we only have $\min\{u_1, u_2\}$ and $\inf_{i\in I} u_i$ are viscosity supersolutions to (2.21). It is important pointing out that the results of Theorems 2.27 and 2.29, and Corollary 2.31 hold naturally for Cauchy problems as well.

6.3 Problems

Exercise 29. *Prove Corollary 2.31.*

Exercise 30. *Show that the results of Theorems 2.27 and 2.29 still hold true if we replace the convexity of H by the level-set quasiconvexity of H. Here, by level-set quasiconvexity of H, we mean $\{p \in \mathbb{R}^n : H(x,p) \le s\}$ is convex in $\mathbb{R}^n$ for all $x \in \mathbb{R}^n$ and $s \in \mathbb{R}$.*

Exercise 31. *Consider the following viscous Hamilton-Jacobi equation in one dimensional space*

$$|u'|^3 - u'' - 1 = 0 \qquad \text{in } \mathbb{R}. \tag{2.24}$$

Clearly, $u_1(x) = x$ and $u_2(x) = -x$ are two classical subsolutions of (2.24). *They are actually two classical solutions. Set*

$$u_3(x) = \min\{u_1(x), u_2(x)\} = -|x| \qquad \text{for } x \in \mathbb{R}.$$

Of course u_3 is a supersolution of (2.24). *Show however that u_3 is not a subsolution of* (2.24).

Exercise 32. *Formulate corresponding versions of Theorems 2.27 and 2.29, and Corollary 2.31 for Cauchy problems and give the proofs.*

6.4 The Hopf–Lax formula revisited

We revisit and give another interpretation of the Hopf–Lax formula in light of the corresponding version of Corollary 2.31 for Cauchy problems. Assume the settings in Theorem 2.23, that is, the Hamiltonian $H = H(p) : \mathbb{R}^n \to \mathbb{R}$ is convex and superlinear. Let u be the viscosity solution to

$$\begin{cases} u_t(x,t) + H(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{2.25}$$

We assume the initial data $u_0 \in \mathrm{BUC}(\mathbb{R}^n)$. Then, for $(x,t) \in \mathbb{R}^n \times (0,\infty)$, the Hopf–Lax formula (Theorem 2.23) gives

$$\begin{aligned} u(x,t) &= \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + u_0(y) \right\} = \min_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + u_0(y) \right\} \\ &= \min_{y \in \mathbb{R}^n} \left\{ \max_{p \in \mathbb{R}^n} t\left(\frac{x-y}{t} \cdot p - H(p)\right) + u_0(y) \right\} \\ &= \min_{y \in \mathbb{R}^n} \max_{p \in \mathbb{R}^n} \{u_0(y) + (x-y)\cdot p - tH(p)\} = \min_{y \in \mathbb{R}^n} \max_{p \in \mathbb{R}^n} \phi^{y,p}(x,t). \end{aligned}$$

Here, for $(y,p) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\phi^{y,p}(x,t) = u_0(y) + (x-y)\cdot p - tH(p) \qquad \text{for } (x,t) \in \mathbb{R}^n \times [0,\infty).$$

Let us now give another way to look at the formula for $u(x,t)$. It is clear that $\phi^{y,p}$ is a separable (special) solution to (2.25) with initial data $\phi^{y,p}(x,0) = u_0(y) + (x-y)\cdot p$, which is affine, and $\phi^{y,p}(y,0) = u_0(y)$. For $y \in \mathbb{R}^n$, set

$$\phi^{y}(x,t) = \max_{p \in \mathbb{R}^n} \phi^{y,p}(x,t) \qquad \text{for } (x,t) \in \mathbb{R}^n \times [0,\infty).$$

Of course, $\phi^y(x,t)$ is finite for $t > 0$ because of the superlinearity of H. In fact, from the definition, for $(x,t) \in \mathbb{R}^n \times (0,\infty)$,

$$\phi^y(x,t) = tL\left(\frac{x-y}{t}\right) + u_0(y).$$

However, it is not hard to see that $\phi^y(x,0)$ is singular. More precisely,

$$\phi^y(x,0) = \begin{cases} u_0(y) & \text{for } x = y, \\ +\infty & \text{for } x \neq y. \end{cases}$$

We see that $\phi^y(x,0)$ is a convex, singular function, which is finite only at y with value $u_0(y)$. Let us assume that we still have the Hopf–Lax formula to (2.25) for this kind of singular initial data. This assumption can be verified rigorously by approximating $\phi^y(x,0)$ by nice and smooth functions. Then, ϕ^y is a solution to (2.25) with initial data $\phi^y(x,0)$. One can regard ϕ^y as a "fundamental solution" to (2.25).

Therefore, by using the corresponding version of Corollary 2.31 for Cauchy problems,

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \phi^y(x,t) = \min_{y\in\mathbb{R}^n} \phi^y(x,t), \qquad \text{for } (x,t) \in \mathbb{R}^n \times [0,\infty),$$

is automatically a solution to (2.25). Moreover, by using the formula of $\phi^y(x,0)$, one automatically get that $u(x,0) = u_0(x)$.

This interpretation does not give anything new, but indeed shows that the Hopf–Lax formula is rather natural. It is quite intuitive to see that once we have understandings on "fundamental solutions" to (2.25), we are able to understand all solutions thanks to the infimum stability result (the corresponding version of Corollary 2.31 for Cauchy problems). The following is an immediate corollary in this discussion.

Corollary 2.32. *Assume the settings in Theorem 2.23. For a given compact set $K \subset \mathbb{R}^n$, denote by*

$$u_0^K(x) = \begin{cases} u_0(x) & \text{for } x \in K, \\ +\infty & \text{for } x \notin K. \end{cases}$$

Then, the solution to (2.25) *with initial data u_0^K is*

$$u^K(x,t) = \min_{y\in K} \phi^y(x,t) \qquad \text{for } (x,t) \in \mathbb{R}^n \times [0,\infty).$$

Let us note that the above formula can also be seen directly from the Hopf–Lax formula as well.

7 Maximal subsolutions and their representation formulas

7.1 Maximal subsolutions and metric problems

We assume in this section that

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for all } R > 0, \\ p \mapsto H(x,p) \text{ is convex for all } x \in \mathbb{R}^n, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} H(x,p) = +\infty. \end{cases} \tag{2.26}$$

Our main object in this section is the maximal subsolution to the following equation

$$\begin{cases} H(x, Du) = \mu & \text{in } \mathbb{R}^n \setminus \{0\}, \\ u(0) = 0. \end{cases} \tag{2.27}$$

Here, $\mu \in \mathbb{R}$ is a given constant. To make sure that (2.27) admits some Lipschitz subsolutions, we define

$$\mu_* = \inf \{\mu \in \mathbb{R} \,:\, \text{there exists a viscosity subsolution } u \in \operatorname{Lip}(\mathbb{R}^n) \text{ to } (2.27)\}.$$

It is clear that (2.27) should only be considered for $\mu \geq \mu_*$.

Definition 2.33. *Fix $\mu \geq \mu_*$. The maximal subsolution is denoted by, for $x \in \mathbb{R}^n$,*

$$m_\mu(x) = \sup \{u(x) \,:\, u \in \operatorname{Lip}(\mathbb{R}^n) \text{ is a viscosity subsolution to } (2.27)\}.$$

Since H is convex in p, in light of Theorem 2.27, the above definition is equivalent to

$$m_\mu(x) = \sup \{u(x) \,:\, u \in \operatorname{Lip}(\mathbb{R}^n) \text{ is an a.e. subsolution to } (2.27)\}.$$

Here is the first result concerning m_μ.

Theorem 2.34. *Assume* (2.26)*. For $\mu \geq \mu_*$, $m_\mu \in \operatorname{Lip}(\mathbb{R}^n)$ and m_μ is a viscosity solution to* (2.27).

Proof. Since H is coercive in p uniformly in x, there exists $C_\mu > 0$ such that, for every $u \in \operatorname{Lip}(\mathbb{R}^n)$ being a subsolution to (2.27),

$$\|Du\|_{L^\infty(\mathbb{R}^n)} \leq C_\mu.$$

By changing C_μ to be a bigger constant if necessary, we assume further that

$$H\left(x, C_\mu e\right) \geq \mu + 1 \qquad \text{for all } e \in \mathbb{R}^n, |e| = 1. \tag{2.28}$$

By definition of m_μ, $\|Dm_\mu\|_{L^\infty(\mathbb{R}^n)} \leq C_\mu$. In particular, $m_\mu(0) = 0$, and

$$m_\mu(x) \leq C_\mu |x| \qquad \text{for all } x \in \mathbb{R}^n.$$

Note that $C_\mu|x|$ is a supersolution to (2.27) thanks to (2.28). By the Perron method (see Theorem 1.26), we conclude that m_μ is a viscosity solution to (2.27). It is also clear that m_μ is the maximal solution to (2.27). □

We have the following remark.

Remark 2.35. It is clear that if $u \in \operatorname{Lip}(\mathbb{R}^n)$ is a viscosity subsolution to (2.27), then it is also a global viscosity subsolution to

$$H(x, Dv) = \mu \qquad \text{in } \mathbb{R}^n. \tag{2.29}$$

On the other hand, for any global subsolution $v \in \operatorname{Lip}(\mathbb{R}^n)$ to (2.29), $v - v(0)$ is a viscosity subsolution to (2.27).

Here is an immediate corollary based on observations of the above remark.

Corollary 2.36. *Assume* (2.26). *For* $\mu \geq \mu_*$*,*

$$m_\mu(x) = \sup\{v(x) - v(0) \,:\, v \in \mathrm{Lip}(\mathbb{R}^n) \text{ is a subsolution to } (2.29)\}.$$

In particular, m_μ *is also a global subsolution to* (2.29).

Let us give one following explicit example to understand more about m_μ.

Example 2.6. *Assume that*

$$H(x,p) = |p| \qquad \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Of course, H satisfies (2.26). *Since* $H \geq 0$*,* $\mu_* \geq 0$*. We can then see that* $\mu_* = 0$ *as* $u \equiv 0$ *is a classical solution to* (2.27) *with* $\mu = 0$*.*

For $\mu \geq 0$*, it is not hard to see that*

$$m_\mu(x) = \mu|x| \qquad \text{for all } x \in \mathbb{R}^n.$$

In particular, $m_1(x) = |x|$*, which is precisely the Euclidean distance from* x *to* 0*. It is worth noting that* m_1 *is not a global viscosity solution to*

$$|Du| = 1 \qquad \text{in } \mathbb{R}^n$$

as it fails the supersolution test at $x = 0$*.*

As pointed out in Corollary 2.36 and Example 2.6, in general, for $\mu \geq \mu_*$, m_μ is a global viscosity subsolution but not a viscosity solution to (2.29). We discuss this point further in the following specific scenario.

Proposition 2.37. *Assume that* $H(x,p) = |p| - V(x)$ *for all* $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$*, and* $V \in \mathrm{BUC}(\mathbb{R}^n)$ *with* $V \geq 0$*. Then, for* $\mu = 0$*,* m_0 *is a viscosity solution to* (2.29) *if and only if* $V(0) = 0$*.*

Proof. We always fix $\mu = 0$ in this proof. It is clear that 0 is a global subsolution to (2.29), and hence, $m_0 \geq 0$ and $m_0(0) = 0$. This also gives us that $\mu_* \leq 0$.

We consider first the case that $V(0) > 0$. It is quite straightforward to show that m_0 is not a supersolution to (2.29) at $x = 0$ in this case. Indeed, $\phi \equiv 0$ touches m_0 from below at $x = 0$, but

$$|D\phi(0)| - V(0) = -V(0) < 0.$$

Thus, m_0 is not a solution to (2.29).

Next, we study the case that $V(0) = 0$ and show that m_0 is a solution to (2.29). Since $V \in \mathrm{BUC}(\mathbb{R}^n)$, there exists a modulus of continuity $\omega : [0,\infty) \to [0,\infty)$ with $\lim_{r\to 0}\omega(r) = 0$ such that, for each $r > 0$,

$$|V(x) - V(0)| = |V(x)| \leq \omega(r) \qquad \text{for all } x \in B(0,r).$$

As m_0 is a solution to (2.27), we use the above to yield that

$$|Dm_0(x)| \leq \omega(r) \qquad \text{for a.e. } x \in B(0,r).$$

In particular, for each $x \in B(0,r)$,

$$|m_0(x)| = |m_0(x) - m_0(0)| = \left| \int_0^1 Dm_0(sx) \cdot x \, ds \right| \le \omega(r)|x| \le \omega(r) r.$$

This means that m_0 is differentiable at 0 and $Dm_0(0) = 0$. Therefore, m_0 is a solution to (2.29) at $x = 0$. This completes our proof. □

In the general setting of H, it is much more complicated to determine whether m_μ satisfies the supersolution test at $x = 0$ or not.

Let us now discuss the metric problem for each fixed $\mu \ge \mu_*$.

Definition 2.38. *Fix $\mu \ge \mu_*$. For $x, y \in \mathbb{R}^n$, denote by*

$$m_\mu(x,y) = \sup \{ v(x) - v(y) \, : \, v \in \mathrm{Lip}(\mathbb{R}^n) \text{ is a subsolution to } (2.29) \}.$$

In particular, $m_\mu(x,0) = m_\mu(x)$ for $x \in \mathbb{R}^n$. When there is no confusion, we will use $m_\mu(x)$ (instead of $m_\mu(x,0)$) for short. In our notations here, we use the second slot in $m_\mu(\cdot,\cdot)$ as a fixed vertex, and geometrically, $x \mapsto m_\mu(x,y)$ looks like a bending upward cone with vertex y (see again Example 2.6 above). Sometimes, people would reverse the order of x and y in the literature.

We record important properties of $m_\mu(\cdot,\cdot)$ below.

Theorem 2.39. *Assume* (2.26). *For $\mu \ge \mu_*$, the following properties hold.*

(i) For each $y \in \mathbb{R}^n$, $x \mapsto m_\mu(x,y)$ is Lipschitz and is the maximal solution to

$$\begin{cases} H(x, Du(x)) = \mu & \text{in } \mathbb{R}^n \setminus \{y\}, \\ u(y) = 0. \end{cases} \tag{2.30}$$

In particular, $m_\mu(y,y) = 0$.

(ii) For $x, y, z \in \mathbb{R}^n$,

$$m_\mu(x,y) + m_\mu(y,z) \ge m_\mu(x,z). \tag{2.31}$$

Equation (7.24) is sometimes called a metric problem in the literature. Property (ii) in the above theorem means that m_μ is subadditive. We will see in the next sections that $m_\mu(x,y)$ represents a certain metric distance between y and x, and (7.25) is nothing but the usual triangle inequality for this metric.

Proof. First property (i) was already proved in Theorem 2.34.

Let us proceed to prove the second property. Fix $y, z \in \mathbb{R}^n$, and denote by

$$w(x) = m_\mu(x,z) - m_\mu(y,z) \qquad \text{for all } x \in \mathbb{R}^n.$$

It is clear that $w \in \mathrm{Lip}(\mathbb{R}^n)$ is a global subsolution to (2.29) since $m_\mu(y,z)$ is just a constant. At $x = y$,

$$w(y) = m_\mu(y,z) - m_\mu(y,z) = 0.$$

Hence, by the definition of $m_\mu(\cdot,y)$, for all $x \in \mathbb{R}^n$,

$$m_\mu(x,y) \ge w(x) - w(y) = w(x) = m_\mu(x,z) - m_\mu(y,z).$$

The proof is complete. □

Remark 2.40. We want to highlight here that all results in this section (in particular, Theorems 2.34, 2.39) still hold true if we replace the convexity of H in condition (2.26) by the level-set quasiconvexity of H, that is, (2.26) is replaced by

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for all } R > 0, \\ p \mapsto H(x,p) \text{ is level-set quasiconvex for all } x \in \mathbb{R}^n, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} H(x,p) = +\infty. \end{cases} \tag{2.32}$$

For Proposition 2.37, the conclusion is valid for the following general Hamiltonian

$$H(x,p) = K(p) - V(x) \qquad \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Here, $K : \mathbb{R}^n \to \mathbb{R}$ satisfies that $K(0) = 0$, K is coercive, level-set quasiconvex, and there exist $\alpha, R > 0$ such that

$$K(p) \geq \alpha|p| \qquad \text{for all } p \in B(0,R).$$

7.2 Representation formulas by using the Lagrangian

We intend to write down the optimal control formulation for m_μ defined in the previous section by using the Lagrangian. For this, we assume a bit more as following.

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for all } R > 0, \\ p \mapsto H(x,p) \text{ is convex for all } x \in \mathbb{R}^n, \\ \lim_{|p|\to\infty} \left(\inf_{x\in\mathbb{R}^n} \dfrac{H(x,p)}{|p|} \right) = +\infty. \end{cases} \tag{2.33}$$

Let L be the Lagrangian corresponding to this H. Here is the main result on the formula of m_μ in this section.

Theorem 2.41. *Assume* (2.33). *For* $\mu \geq \mu_*$ *and* $x \in \mathbb{R}^n$,

$$m_\mu(x) = \inf\left\{ \int_0^t \big(L(\gamma(s), \gamma'(s)) + \mu \big)\, ds \;:\; \gamma \in \mathrm{AC}([0,t], \mathbb{R}^n) \text{ for } t > 0, \gamma(0) = 0, \gamma(t) = x \right\}. \tag{2.34}$$

Proof. Denote by w the right hand side of (2.34). Our main goal is to show $m_\mu = w$.

Step 1. Let $u \in \mathrm{Lip}(\mathbb{R}^n)$ be a viscosity subsolution to (2.27). We first show that $u \leq w$. Since u is only Lipschitz, we need to smooth it up by using a standard mollifier η. For $\varepsilon > 0$, denote by $\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(\frac{x}{\varepsilon}\right)$ for all $x \in \mathbb{R}^n$. Set

$$u^\varepsilon(x) = (\eta_\varepsilon \star u)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)u(y)\,dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x-y)u(y)\,dy \quad \text{for } x \in \mathbb{R}^n.$$

Then $u^\varepsilon \in C^\infty(\mathbb{R}^n)$, and $u^\varepsilon \to u$ locally uniformly as $\varepsilon \to 0$. By repeating similar steps as in the proof of Theorem 2.27, we see that

$$H(x, Du^\varepsilon) \leq \mu + \omega(\varepsilon) \qquad \text{in } \mathbb{R}^n,$$

where $\omega(\varepsilon)$ is a modulus of continuity. By the Legendre transform and the above inequality,

$$\begin{aligned}\int_0^t \big(L(\gamma(s),\gamma'(s))+\mu\big)\,ds &\geq \int_0^t \big(L(\gamma(s),\gamma'(s))+H(\gamma(s),Du^\varepsilon(\gamma(s)))-\omega(\varepsilon)\big)\,ds\\ &\geq \int_0^t \big(Du^\varepsilon(\gamma(s))\cdot\gamma'(s)-\omega(\varepsilon)\big)\,ds\\ &= u^\varepsilon(\gamma(t))-u^\varepsilon(\gamma(0))-t\omega(\varepsilon) = u^\varepsilon(x)-u^\varepsilon(0)-t\omega(\varepsilon).\end{aligned}$$

Let $\varepsilon\to 0^+$ in the above to imply that

$$\int_0^t \big(L(\gamma(s),\gamma'(s))+\mu\big)\,ds \geq u(x)-u(0)\geq u(x),$$

which gives us further that $u\leq w$. Taking supremum over u to yield $m_\mu\leq w$.

Step 2. To finish the proof, we just need to show that w is a viscosity subsolution to (2.27) as this will give use directly that $w\leq m_\mu$. It is straightforward that $w(0)=0$. By the formula of w, for $x\neq 0$ and $r\in(0,|x|)$,

$$\begin{aligned}&w(x)=\\ &=\inf\left\{\int_0^t \big(L(\gamma(s),\gamma'(s))+\mu\big)\,ds+w(\gamma(0))\,:\,\gamma\in \mathrm{AC}([0,t],\mathbb{R}^n),\gamma(0)\in\partial B(x,r),\gamma(t)=x\right\}.\end{aligned}$$

This relation is precisely a Dynamic Programming Principle for w. We now use it to prove the subsolution test. Let $\phi\in C^\infty(\mathbb{R}^n)$ be a test function such that $w-\phi$ has a strict global maximum at x and $w(x)=\phi(x)$. For any $\gamma\in \mathrm{AC}([0,t],\mathbb{R}^n)$ such that $\gamma(0)\in\partial B(x,r)$, $\gamma(t)=x$, one has

$$\begin{aligned}\phi(x)=w(x)&\leq \int_0^t \big(L(\gamma(s),\gamma'(s))+\mu\big)\,ds+w(\gamma(0))\\ &\leq \int_0^t \big(L(\gamma(s),\gamma'(s))+\mu\big)\,ds+\phi(\gamma(0)).\end{aligned}$$

For each non-zero vector $e\in\mathbb{R}^n$, denote by

$$\gamma_e(s)=x-te+se \qquad \text{for } 0\leq s\leq t,$$

for $t>0$ sufficiently small. Then, for this path γ_e, we see that

$$\phi(x)\leq \int_0^t (L(x-te+se,e)+\mu)\,ds+\phi(x-te).$$

Hence,

$$\frac{\phi(x)-\phi(x-te)}{t}\leq \frac{1}{t}\int_0^t (L(x-te+se,e)+\mu)\,ds.$$

Let $t\to 0^+$ in the above to imply that

$$D\phi(x)\cdot e-L(x,e)\leq\mu.$$

Maximize this inequality over $e\in\mathbb{R}^n$ to get

$$H(x,D\phi(x))\leq\mu,$$

which confirms that w is a viscosity subsolution to (2.27). The proof is complete. □

Let us now give an application of Theorems 2.34 and 2.41.

Corollary 2.42. *Let $V \in \mathrm{BUC}(\mathbb{R}^n)$ be a given function such that $V \geq 0$. Then, the maximal viscosity solution v of the following equation*

$$\begin{cases} |Du| = V(x) & \text{in } \mathbb{R}^n \setminus \{0\}, \\ u(0) = 0, \end{cases} \tag{2.35}$$

has the formula

$$v(x) = \inf\left\{ \int_0^t \left(\frac{|\gamma'(s)|^2}{4} + V(\gamma(s)) \right) ds \,:\, \gamma \in \mathrm{AC}([0,t],\mathbb{R}^n) \text{ for } t > 0, \gamma(0) = 0, \gamma(t) = x \right\}.$$

Proof. Although the Hamiltonian in (2.35) does not have superlinear growth in p, we can rewrite (2.35) in an equivalent form to fix this issue. Indeed, (2.35) is equivalent to

$$\begin{cases} |Du|^2 = V(x)^2 & \text{in } \mathbb{R}^n \setminus \{0\}, \\ u(0) = 0. \end{cases}$$

The Hamiltonian of this PDE is $H(x,p) = |p|^2 - V(x)^2$, which satisfies (2.33), and its corresponding Lagrangian is $L(x,v) = \frac{|v|^2}{4} + V(x)^2$. We hence are able to apply Theorems 2.34 and 2.41 with $\mu = 0$ to conclude. □

Let us note that another representation formula of v will be given below in Remark 2.44.

7.3 Representation formulas for first-order front propagation problems

Our equation of interest here is

$$\begin{cases} a(x)|Du| = 1 & \text{in } \mathbb{R}^n \setminus \{0\}, \\ u(0) = 0. \end{cases} \tag{2.36}$$

Here, $a : \mathbb{R}^n \to \mathbb{R}$ satisfies

$$\begin{cases} a \in \mathrm{BUC}(\mathbb{R}^n), \\ \text{there exist } \alpha, \beta > 0 \text{ such that } \alpha \leq a(x) \leq \beta \text{ for all } x \in \mathbb{R}^n. \end{cases} \tag{2.37}$$

Note that (2.36) can be also written in the form of (2.35) with $V(x) = a(x)^{-1}$. Our goal is to write down an analog of Theorem 2.41 in the sense of front propagation problems.

It is clear in this case that the Hamiltonian $H(x,p) = a(x)|p|$ for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$ is convex and uniformly coercive, but is not superlinear in p. Of course, (2.33) does not hold here. Recall that the Lagrangian $L(x,v)$ is computed as following

$$\begin{aligned} L(x,v) &= \sup_{p\in\mathbb{R}^n} (p \cdot v - H(x,p)) = \sup_{p\in\mathbb{R}^n} (p \cdot v - a(x)|p|) \\ &= \begin{cases} 0 & \text{if } |v| \leq a(x), \\ +\infty & \text{if } |v| > a(x). \end{cases} \end{aligned}$$

Although L is singular in the above, it is still convex in v. Then, in this particular setting, formula (2.34) for $\mu = 1$ becomes

$$m_1(x) = \inf\left\{\int_0^t \big(L(\gamma(s),\gamma'(s))+1\big)\,ds\;:\;\gamma\in \mathrm{AC}([0,t],\mathbb{R}^n) \text{ for } t>0, \gamma(0)=0,\gamma(t)=x\right\}$$
$$= \inf\big\{t\;:\;\exists\gamma\in\mathrm{AC}([0,t],\mathbb{R}^n) \text{ s.t. } \gamma(0)=0,\gamma(t)=x,|\gamma'(s)|\le a(\gamma(s)) \text{ a.e.}\big\}.$$

Through the formula, we see that $m_1(x)$ is the minimal time to reach x from 0 under the velocity constraint of the paths, which cannot exceed a at any given point. In the literature, m_1 is sometimes called the minimal time function in this setting. Let us now state the precise result concerning maximal solution to (2.36).

Theorem 2.43. *Assume* (2.37). *Then the maximal solution* m_1 *to* (2.36) *has the following representation formula, for* $x \in \mathbb{R}^n$*,*

$$m_1(x) = \inf\big\{t\;:\;\exists\gamma\in\mathrm{AC}([0,t],\mathbb{R}^n) \textit{ s.t. } \gamma(0)=0,\gamma(t)=x,|\gamma'(s)|\le a(\gamma(s)) \textit{ a.e.}\big\}.$$

The proof of this theorem is similar to that of Theorem 2.41, so it is left as an exercise here.

Remark 2.44. Let us notice that, by a simple change of variable, we have another formula for m_1 in Theorem 2.43 as following

$$m_1(x) = \inf\left\{\int_0^T \frac{1}{a(\xi(s))}\,ds\;:\;\xi\in\mathrm{AC}([0,T],\mathbb{R}^n),\xi(0)=0,\xi(T)=x,|\xi'(s)|\le 1 \text{ a.e.}\right\}.$$

Basically, the only difference of this formula with the above in Theorem 2.43 is that we change the constraint on the admissible paths ξ such that they have at most unit velocity.

7.4 Generalizations

The maximal subsolution problem (2.27) can be generalized to various different situations. For example, it is reasonable sometimes to consider the problem in only a given bounded smooth domain $U \subset \mathbb{R}^n$. Another possibility is to consider the problem under a constraint on a given closed set $K \subset \mathbb{R}^n$ as following

$$\begin{cases} H(x,Du) = \mu & \text{in } \mathbb{R}^n\setminus K,\\ u\le 0 & \text{on } K. \end{cases} \tag{2.38}$$

Here, $\mu \in \mathbb{R}$ is a given constant. When $K = \{0\}$, (2.38) reduces to (2.27). We give some discussions on (2.38) here in this section. Let us assume (2.26). It turns out that all of the main results (Theorems 2.34, 2.41, and 2.43) still hold for this generalization when being adjusted appropriately.

Firstly, denote by

$$\mu_*(K) = \inf\{\mu\in\mathbb{R}\;:\;\text{there exists a viscosity subsolution } u\in\mathrm{Lip}(\mathbb{R}^n) \text{ to (2.38)}\}.$$

It is worth noting that $\mu_*(K)$ and μ_* defined earlier might not be the same. If one assumes further that K is compact, then $\mu_*(K) \le \mu_*$.

Definition 2.45. *Fix $\mu \geq \mu_*(K)$. The maximal subsolution of* (2.38) *is denoted by, for $x \in \mathbb{R}^n$,*

$$m_\mu(x,K) = \sup\left\{u(x) \,:\, u \in C(\mathbb{R}^n) \cap C^{0,1}(\mathbb{R}^n \setminus K) \text{ is a viscosity subsolution to } (2.38)\right\}.$$

Again, the above definition is equivalent to

$$m_\mu(x,K) = \sup\left\{u(x) \,:\, u \in C(\mathbb{R}^n) \cap C^{0,1}(\mathbb{R}^n \setminus K) \text{ is an a.e. subsolution to } (2.38)\right\}.$$

Here is the first main result concerning $m_\mu(\cdot,K)$.

Theorem 2.46. *Assume* (2.26). *For $\mu \geq \mu_*(K)$, $m_\mu(\cdot,K) \in C^{0,1}(\mathbb{R}^n \setminus K)$ and $m_\mu(\cdot,K)$ is a viscosity solution to* (2.38).

The proof of this theorem follows exactly that of Theorem 2.34, and hence, is omitted. We next give representation formulas to $m_\mu(\cdot,K)$ in two different situations. The first one is when H is superlinear in p.

Theorem 2.47. *Assume* (2.33). *For $\mu \geq \mu_*(K)$ and $x \in \mathbb{R}^n \setminus K$,*

$$\begin{aligned} &m_\mu(x,K) \\ &\quad = \inf\left\{\int_0^t \left(L(\gamma(s),\gamma'(s)) + \mu\right) ds \,:\, \gamma \in \mathrm{AC}([0,t],\mathbb{R}^n) \text{ for } t > 0, \gamma(0) \in K, \gamma(t) = x\right\}. \end{aligned}$$

The second one corresponds to the usual first-order front propagation problem.

Theorem 2.48. *Assume that $H(x,p) = a(x)|p|$ for all $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$, and a satisfies* (2.37). *Then the maximal solution $m_1(\cdot,K)$ to* (2.38) *with $\mu = 1$ has the following representation formula, for $x \in \mathbb{R}^n \setminus K$,*

$$m_1(x,K) = \inf\left\{t \,:\, \exists\, \gamma \in \mathrm{AC}([0,t],\mathbb{R}^n) \text{ s.t. } \gamma(0) \in K, \gamma(t) = x, |\gamma'(s)| \leq a(\gamma(s)) \text{ a.e.}\right\}.$$

Basically, $m_1(x,K)$ represents the minimal time to reach x from the given set K.

The proofs of Theorems 2.47 and 2.48 are also skipped here.

7.5 Problems

Exercise 33. *Give a detailed proof of Theorem 2.43.*

Exercise 34. *Prove Remark 2.44.*

8 References

1. The optimal control theory part can be found in many references. For example, the readers can consult the books of Bardi, Capuzzo-Dolcetta [13], Barles [16], or Chapter 10 of Evans [49], or the book of Lions [101]. Lions [101] observed the connection between the definition of viscosity solutions and the optimality conditions of optimal control theory.

2. The Hopf–Lax and Lax–Oleinik formulas are discussed in deep in Chapter 3 of Evans [49]. We do not give much discussion on the Lax–Oleinik formula here as it is not in the focus of the book.

3. Theorems 2.27 and 2.29 are due to Barron, Jensen [18].

4. Maximal subsolutions and their optimal control formulas were first discussed in the book of Lions mainly for bounded domains (see [101, Chapter 5]). We only cover the case of $\mathbb{R}^n$ here for simplicity.

CHAPTER 3

First-order Hamilton–Jacobi equations with possibly nonconvex Hamiltonians

Let $H = H(x, p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a given Hamiltonian. In this chapter, we consider a general setting where $p \mapsto H(x, p)$ might not be convex for given $x \in \mathbb{R}^n$, and hence, the theory developed in Chapter 2 is not applicable.

1 Introduction to two-player zero-sum differential games

1.1 Settings

We consider a zero-sum differential game played by two players I and II, who are both rational. In the game, player I aims at maximizing while player II aims at minimizing a certain payoff functional by controlling the dynamics of a particle in $\mathbb{R}^n$, which represents the location of the pair in the game.

Fix $T > 0$. Let A, B be two compact metric spaces. For $t \in [0, T)$, let

$$\mathcal{A}_t = \{a : [t, T] \to A \,:\, a \text{ is measurable}\},$$
$$\mathcal{B}_t = \{b : [t, T] \to B \,:\, b \text{ is measurable}\},$$

be the set of possible controls in time $[t, T]$ of players I and II, respectively. We henceforth identify any two controls which agree a.e.

Assume that the dynamics is given by an ordinary differential equation

$$\begin{cases} y_x'(s) = f(y_x(s), a(s), b(s)) & \text{for } s \in (t, T), \\ y_x(t) = x \in \mathbb{R}^n, \end{cases} \tag{3.1}$$

for given controls $a(\cdot) \in \mathcal{A}_t$ of player I, and $b(\cdot) \in \mathcal{B}_t$ of player II. Here, $f : \mathbb{R}^n \times A \times B \to \mathbb{R}^n$ is a given vector field satisfying: there exists $C > 0$ such that

$$\begin{cases} f \in C(\mathbb{R}^n \times A \times B), \\ |f(x, a, b)| \le C & \text{for all } x \in \mathbb{R}^n, a \in A, b \in B, \\ |f(x_1, a, b) - f(x_2, a, b)| \le C|x_1 - x_2| & \text{for all } x_1, x_2 \in \mathbb{R}^n, a \in A, b \in B. \end{cases}$$

Under the conditions on f, (3.1) has a unique solution. At any given time $s \in (t,T)$, $y_x(s)$ represents the location of the pair in the game. Associated with this (3.1) is the payoff functional

$$C_{x,t}(a(\cdot),b(\cdot)) = \int_t^T h(y_x(s),a(s),b(s))\,ds + g(y_x(T)),$$

where $h : \mathbb{R}^n \times A \times B \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are given functions satisfying: there exists $C > 0$ so that

$$\begin{cases} h \in C(\mathbb{R}^n \times A \times B), & \\ |h(x,a,b)| \leq C & \text{for all } x \in \mathbb{R}^n, a \in A, b \in B, \\ |h(x_1,a,b) - h(x_2,a,b)| \leq C|x_1 - x_2| & \text{for all } x_1, x_2 \in \mathbb{R}^n, a \in A, b \in B; \end{cases}$$

and

$$\begin{cases} |g(x)| \leq C & \text{for all } x \in \mathbb{R}^n, \\ |g(x_1) - g(x_2)| \leq C|x_1 - x_2| & \text{for all } x_1, x_2 \in \mathbb{R}^n. \end{cases}$$

The interpretation is that h is the running payoff and g is the terminal payoff. Of course, the goal of player I is to maximize the payoff functional $C_{x,t}(a(\cdot),b(\cdot))$. On the other hand, player II wants to minimize it (or to maximize $-C_{x,t}(a(\cdot),b(\cdot))$).
The set of strategies for player I beginning at time t is

$$\Sigma_t = \{\alpha : \mathcal{B}_t \to \mathcal{A}_t \text{ non-anticipating}\},$$

where non-anticipating means that, for all $b_1(\cdot), b_2(\cdot) \in \mathcal{B}_t$ and $s \in [t,T]$,

$$b_1(\cdot) = b_2(\cdot) \text{ on } [t,s) \qquad \Longrightarrow \qquad \alpha[b_1](\cdot) = \alpha[b_2](\cdot) \text{ on } [t,s).$$

Similarly, the set of strategies for player II beginning at time t is

$$\Gamma_t = \{\beta : \mathcal{A}_t \to \mathcal{B}_t \text{ non-anticipating}\}.$$

We call

$$\begin{aligned} V(x,t) &= \inf_{\beta \in \Gamma_t} \sup_{a(\cdot) \in \mathcal{A}_t} C_{x,t}(a(\cdot), \beta[a](\cdot)), \\ U(x,t) &= \sup_{\alpha \in \Sigma_t} \inf_{b(\cdot) \in \mathcal{B}_t} C_{x,t}(\alpha[b](\cdot), b(\cdot)), \end{aligned}$$

the lower value and the upper values of the game, respectively.

1.2 Viscosity solutions to terminal value problems

In the previous chapters, we have already defined and worked with viscosity solutions to initial value problems. In our current setting, it is more natural to work with the following terminal value problem

$$\begin{cases} u_t + H(x,Du) = 0 & \text{in } \mathbb{R}^n \times (0,T), \\ u(x,T) = g(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{3.2}$$

Here, $H \in C(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$, and $g \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ are given.

Definition 3.1 (viscosity solutions of (3.2)). *A function* $u \in \mathrm{BUC}(\mathbb{R}^n \times [0,T])$ *is called*

(a) *a viscosity subsolution of* (3.2) *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*

$$\varphi_t(x_0,t_0) + H(Du(x_0,t_0)) \geq 0,$$

and $u(\cdot,T) \leq g$*;*

(b) *a viscosity supersolution of* (3.2) *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*

$$\psi_t(x_0,t_0) + H(Du(x_0,t_0)) \leq 0,$$

and $u(\cdot,T) \geq g$*;*

(c) *a viscosity solution of* (3.2) *if it is both a viscosity subsolution and a viscosity supersolution.*

Remark 3.2. It is worth noting that for terminal value problems, the inequalities in (a) and (b) in the above definition for test functions are the reversed versions of those for initial value problems.

1.3 Upper and lower Hamiltonians of the game

For $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$, denote by

$$\begin{aligned} H^-(x,p) &= \max_{a\in A}\min_{b\in B}\{f(x,a,b)\cdot p + h(x,a,b)\}, \\ H^+(x,p) &= \min_{b\in B}\max_{a\in A}\{f(x,a,b)\cdot p + h(x,a,b)\}. \end{aligned}$$

We say that H^- and H^+ are the lower and upper Hamiltonians of the game, respectively.

Lemma 3.3. *Let* H^-, H^+ *be the functions defined above. Then,*

(i) *There exists* $C > 0$ *such that, for all* $x,y,p,q \in \mathbb{R}^n$*,*

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

(ii) *We always have that* $H^- \leq H^+$*.*

Proof. The proof of (i) is straightforward thanks to the properties of f,h, and hence is omitted. Let us give a proof of (ii). Since

$$f(x,a,b)\cdot p + h(x,a,b) \leq \max_{a\in A}\{f(x,a,b)\cdot p + h(x,a,b)\},$$

we can take minimum over $b \in B$ to yield that, for each $a \in A$,

$$\min_{b\in B}\{f(x,a,b)\cdot p + h(x,a,b)\} \leq \min_{b\in B}\max_{a\in A}\{f(x,a,b)\cdot p + h(x,a,b)\} = H^+(x,p).$$

Finally, take maximum over $a \in A$ in the above to conclude. □

We now state the main results for two-player zero-sum differential games.

Theorem 3.4. *Let V, U, H^-, H^+ be the functions defined above. Then,*

(i) V is the viscosity solution to the lower Isaacs equation

$$\begin{cases} V_t + H^-(x, DV) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,T), \\ V(x,T) &= g(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{3.3}$$

(ii) U is the viscosity solution to the upper Isaacs equation

$$\begin{cases} U_t + H^+(x, DU) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,T), \\ U(x,T) &= g(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{3.4}$$

The proof of this theorem will be given in the next sections. Since $H^- \leq H^+$, we get $V \leq U$ by using the comparison principle.

Definition 3.5. *We say that the two-player zero-sum differential game has a value if $H^- = H^+$, that is,*

$$\max_{a\in A}\min_{b\in B}\{f(x,a,b)\cdot p + h(x,a,b)\} = \min_{b\in B}\max_{a\in A}\{f(x,a,b)\cdot p + h(x,a,b)\}. \tag{3.5}$$

We say that (3.5) *is a minimax condition.*

We have immediately the following corollary.

Corollary 3.6. *Assume that* (3.5) *holds. Denote by*

$$H(x,p) = H^-(x,p) = H^+(x,p) \qquad \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Then, $U = V$ solves

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

Example 3.1. *Let $n = 2$, and for $p = (p_1, p_2) \in \mathbb{R}^2$, consider*

$$H(p) = H(p_1,p_2) = |p_1| - |p_2|.$$

Set $A = B = [-1,1] \subset \mathbb{R}$. It is clear that

$$\begin{aligned} H(p) = |p_1| - |p_2| &= \max_{|a|\leq 1}(ap_1) + \min_{|b|\leq 1}(bp_2) \\ &= \max_{a\in A}\min_{b\in B}(a,b)\cdot p = \min_{b\in B}\max_{a\in A}(a,b)\cdot p. \end{aligned}$$

Thus, in this specific situation, $f(x,a,b) = (a,b)$, and $h \equiv 0$, and the minimax condition (3.5) *holds true.*

1.4 Properties of the upper and lower values

We have the following Dynamic Programming Principles (DPP) for lower and upper values V, U.

Theorem 3.7 (DPP for V and U). *Let V, U be the functions defined above. Then, for each $0 \le t < t+\sigma \le T$, and $x \in \mathbb{R}^n$,*

$$V(x,t) = \inf_{\beta\in\Gamma_t} \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^{t+\sigma} h(y_x(s), a(s), \beta[a](s))\,ds + V(y_x(t+\sigma), t+\sigma) \right\},$$

$$U(x,t) = \sup_{\alpha\in\Sigma_t} \inf_{b(\cdot)\in\mathcal{B}_t} \left\{ \int_t^{t+\sigma} h(y_x(s), \alpha[b](s), b(s))\,ds + U(y_x(t+\sigma), t+\sigma) \right\}.$$

Proof. We only give the proof of the DPP for V, as the proof of the DPP for U is similar. Let

$$W(x,t) = \inf_{\beta\in\Gamma_t} \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^{t+\sigma} h(y_x(s), a(s), \beta[a](s))\,ds + V(y_x(t+\sigma), t+\sigma) \right\}.$$

Fix $\varepsilon > 0$. There exists $\delta \in \Gamma_t$ such that

$$W(x,t) > \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^{t+\sigma} h(y_x(s), a(s), \delta[a](s))\,ds + V(y_x(t+\sigma), t+\sigma) \right\} - \varepsilon.$$

Besides, for each $z \in \mathbb{R}^n$,

$$V(z, t+\sigma) = \inf_{\beta\in\Gamma_{t+\sigma}} \sup_{a(\cdot)\in\mathcal{A}_{t+\sigma}} C_{z,t+\sigma}(a(\cdot), \beta[a](\cdot)).$$

Thus, there exists $\delta_z \in \Gamma_{t+\sigma}$ so that

$$V(z, t+\sigma) \ge \sup_{a(\cdot)\in\mathcal{A}_{t+\sigma}} C_{z,t+\sigma}(a(\cdot), \delta_z[a](\cdot)) - \varepsilon.$$

We define $\beta \in \Gamma_t$ as following. For each $a(\cdot) \in \mathcal{A}_t$, set

$$\beta[a](s) = \begin{cases} \delta[a](s) & \text{for } t \le s < t+\sigma, \\ \delta_{y_x(t+\sigma)}[a](s) & \text{for } t+\sigma \le s \le T. \end{cases}$$

Then, for any $a(\cdot) \in \mathcal{A}_t$,

$$W(x,t) > \int_t^T h(y_x(s), a(s), \beta[a](s))\,ds + g(y_x(T)) - 2\varepsilon.$$

Hence,

$$W(x,t) \ge V(x,t) - 2\varepsilon. \tag{3.6}$$

Let us now proceed to obtain a reversed inequality. There exists $\beta \in \Gamma_t$ such that

$$V(x,t) > \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^T h(y_x(s), a(s), \beta[a](s))\,ds + g(y_x(T)) \right\} - \varepsilon.$$

Surely, for this fixed β,

$$W(x,t) \leq \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^{t+\sigma} h(y_x(s), a(s), \beta[a](s))\,ds + V(y_x(t+\sigma), t+\sigma) \right\}.$$

Consequently, there exists $a_1(\cdot) \in \mathcal{A}_t$ such that

$$W(x,t) \leq \int_t^{t+\sigma} h(y_x(s), a_1(s), \beta[a_1](s))\,ds + V(y_x(t+\sigma), t+\sigma) + \varepsilon.$$

Now, for each $a(\cdot) \in \mathcal{A}_{t+\sigma}$, denote $\tilde{a}(\cdot) \in \mathcal{A}_t$ by

$$\tilde{a}(s) = \begin{cases} a_1(s) & \text{for } t \leq s < t+\sigma, \\ a(s) & \text{for } t+\sigma \leq s \leq T. \end{cases}$$

Then, define $\tilde{\beta} \in \Gamma_{t+\sigma}$ as

$$\tilde{\beta}[a](s) = \beta[\tilde{a}](s) \qquad \text{for } t+\sigma \leq s \leq T.$$

By definition of V, for $z = y_x(t+\sigma)$,

$$V(y_x(t+\sigma), t+\sigma) = V(z, t+\sigma) \leq \sup_{a(\cdot)\in\mathcal{A}_{t+\sigma}} \left\{ \int_{t+\sigma}^{T} h(y_z(s), a(s), \tilde{\beta}[a](s))\,ds + g(y_z(T)) \right\},$$

and so there exists $a_2(\cdot) \in \mathcal{A}_{t+\sigma}$ for which

$$V(y_x(t+\sigma), t+\sigma) \leq \int_{t+\sigma}^{T} h(y_z(s), a_2(s), \tilde{\beta}[a_2](s))\,ds + g(y_z(T)) + \varepsilon.$$

Define $a(\cdot) \in \mathcal{A}_t$ as

$$a(s) = \begin{cases} a_1(s) & \text{for } t \leq s < t+\sigma, \\ a_2(s) & \text{for } t+\sigma \leq s \leq T. \end{cases}$$

Then, $y_x(s) = y_z(s)$ for $t+\sigma \leq s \leq T$, and

$$W(x,t) \leq \int_t^{T} h(y_x(s), a(s), \beta[a](s))\,ds + g(y_x(T)) + 2\varepsilon.$$

Therefore,

$$W(x,t) \leq V(x,t) + 3\varepsilon. \tag{3.7}$$

Let $\varepsilon \to 0$ and combine (3.6)–(3.7) to conclude. □

Next, we show that V, U are bounded and Lipschitz continuous on $\mathbb{R}^n \times [0,T]$.

Proposition 3.8. *Let V, U be the functions defined above. Then, there exists a constant $C = C(T) > 0$ such that*

$$\|V\|_{L^\infty(\mathbb{R}^n\times[0,T])} + \|V_t\|_{L^\infty(\mathbb{R}^n\times[0,T])} + \|DV\|_{L^\infty(\mathbb{R}^n\times[0,T])} \leq C,$$

and

$$\|U\|_{L^\infty(\mathbb{R}^n\times[0,T])} + \|U_t\|_{L^\infty(\mathbb{R}^n\times[0,T])} + \|DU\|_{L^\infty(\mathbb{R}^n\times[0,T])} \leq C.$$

Proof. We will only obtain the bounds for V. Since h and g are both bounded, it is clear that

$$|C_{x,t}(a(\cdot), b(\cdot))| \le C(T-t) + C \le C(T+1),$$

which implies right away that $\|V\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C(T+1)$. Let us now show that

$$\|DV\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C.$$

Fix $x_1, x_2 \in \mathbb{R}^n$ and $t \in [0,T)$. For each $\varepsilon > 0$, there exists $\delta \in \Gamma_t$ such that

$$V(x_1,t) > \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^T h(y_{x_1}(s), a(s), \delta[a](s))\, ds + g(y_{x_1}(T)) \right\} - \varepsilon.$$

For each $a(\cdot) \in \mathcal{A}_t$, recall that y_{x_1} solves

$$\begin{cases} y_{x_1}'(s) = f(y_{x_1}(s), a(s), \delta[a](s)) & \text{for } s \in (t,T), \\ y_{x_1}(t) = x_1. \end{cases}$$

Let y_{x_2} be the solution to

$$\begin{cases} y_{x_2}'(s) = f(y_{x_2}(s), a(s), \delta[a](s)) & \text{for } s \in (t,T), \\ y_{x_2}(t) = x_2. \end{cases}$$

By Gronwall's inequality, we have that

$$|y_{x_2}(s) - y_{x_1}(s)| \le C|x_2 - x_1| \qquad \text{for all } s \in [t,T].$$

Therefore,

$$\begin{aligned} V(x_1,t) &> \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^T h(y_{x_1}(s), a(s), \delta[a](s))\, ds + g(y_{x_1}(T)) \right\} - \varepsilon \\ &> \sup_{a(\cdot)\in\mathcal{A}_t} \left\{ \int_t^T h(y_{x_2}(s), a(s), \delta[a](s))\, ds + g(y_{x_2}(T)) - C|x_2 - x_1| \right\} - \varepsilon \\ &> V(x_2,t) - C|x_2 - x_1| - \varepsilon. \end{aligned}$$

Let $\varepsilon \to 0$ to yield

$$V(x_2,t) - V(x_1,t) \le C|x_2 - x_1|.$$

By a symmetric argument, we deduce

$$|V(x_2,t) - V(x_1,t)| \le C|x_2 - x_1|.$$

Thus, $\|DV\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C$. We skip the proof that $\|V_t\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C$, and leave it as an exercise. □

1.5 Proof of Theorem 3.4

Proof of Theorem 3.4. Let us provide the proof of the assertion for U only.

We first show that U is a viscosity subsolution to (3.4). Take $\phi \in C^1(\mathbb{R}^n \times (0,T))$ be a test function such that $U(x_0,t_0) = \phi(x_0,t_0)$, and $U-\phi$ has a strict maximum at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. We need to show that

$$\phi_t(x_0,t_0) + H^+(x_0, D\phi(x_0,t_0)) \geq 0.$$

Assume otherwise that there exists $\theta > 0$ such that

$$\phi_t(x_0,t_0) + H^+(x_0, D\phi(x_0,t_0)) < -\theta < 0. \tag{3.8}$$

By the definition of H^+ that

$$H^+(x,p) = \min_{b\in B} \max_{a\in A} \{f(x,a,b)\cdot p + h(x,a,b)\},$$

we can find $b(\cdot) \in \mathcal{B}_{t_0}$ and $\sigma > 0$ sufficiently small such that, for all $\alpha \in \Sigma_{t_0}$,

$$\int_{t_0}^{t_0+\sigma} \Big[h(y_{x_0}(s), \alpha[b](s), b(s)) + f(y_{x_0}(s), \alpha[b](s), b(s)) \cdot D\phi(y_{x_0}(s), s) + \phi_t(y_{x_0}(s), s)\Big]\, ds \leq -\frac{\theta\sigma}{2}.$$

Note that $y'_{x_0}(s) = f(y_{x_0}(s), \alpha[b](s), b(s))$, and so

$$\begin{aligned}
&\int_{t_0}^{t_0+\sigma} f(y_{x_0}(s), \alpha[b](s), b(s)) \cdot D\phi(y_{x_0}(s), s) + \phi_t(y_{x_0}(s), s)\, ds\\
&= \int_{t_0}^{t_0+\sigma} y'_{x_0}(s) \cdot D\phi(y_{x_0}(s), s) + \phi_t(y_{x_0}(s), s)\, ds\\
&= \phi(y_{x_0}(t_0+\sigma), t_0+\sigma) - \phi(x_0,t_0) \geq U(y_{x_0}(t_0+\sigma), t_0+\sigma) - U(x_0,t_0).
\end{aligned}$$

We combine this with the above inequality to yield

$$U(x_0,t_0) - \frac{\theta\sigma}{2} \geq \sup_{\alpha\in\Sigma_{t_0}} \left\{ \int_{t_0}^{t_0+\sigma} h(y_{x_0}(s), \alpha[b](s), b(s))\, ds + U(y_{x_0}(t_0+\sigma), t_0+\sigma) \right\},$$

which contradicts with the DPP for U that

$$U(x_0,t_0) = \sup_{\alpha\in\Sigma_{t_0}} \inf_{b(\cdot)\in\mathcal{B}_{t_0}} \left\{ \int_{t_0}^{t_0+\sigma} h(y_{x_0}(s), \alpha[b](s), b(s))\, ds + U(y_{x_0}(t_0+\sigma), t_0+\sigma) \right\}.$$

Hence, U is a viscosity subsolution to (3.4).

Next, we show that U is a viscosity supersolution to (3.4). Take $\phi \in C^1(\mathbb{R}^n \times (0,T))$ be a test function such that $U(x_0,t_0) = \phi(x_0,t_0)$, and $U-\phi$ has a strict minimum at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. We need to show that

$$\phi_t(x_0,t_0) + H^+(x_0, D\phi(x_0,t_0)) \leq 0.$$

Assume otherwise that there exists $\theta > 0$ such that

$$\phi_t(x_0, t_0) + H^+(x_0, D\phi(x_0, t_0)) > \theta > 0. \tag{3.9}$$

By the definition of H^+, we can find $\alpha \in \Sigma_{t_0}$ and $\sigma > 0$ sufficiently small such that, for all $b(\cdot) \in \mathcal{B}_{t_0}$,

$$\int_{t_0}^{t_0+\sigma} h(y_{x_0}(s), \alpha[b](s), b(s)) + f(y_{x_0}(s), \alpha[b](s), b(s)) \cdot D\phi(y_{x_0}(s), s) + \phi_t(y_{x_0}(s), s)\, ds \geq \frac{\theta\sigma}{2}.$$

Note again that $y'_{x_0}(s) = f(y_{x_0}(s), \alpha[b](s), b(s))$, and so

$$\begin{aligned}
&\int_{t_0}^{t_0+\sigma} f(y_{x_0}(s), \alpha[b](s), b(s)) \cdot D\phi(y_{x_0}(s), s) + \phi_t(y_{x_0}(s), s)\, ds \\
&= \int_{t_0}^{t_0+\sigma} y'_{x_0}(s) \cdot D\phi(y_{x_0}(s), s) + \phi_t(y_{x_0}(s), s)\, ds \\
&= \phi(y_{x_0}(t_0+\sigma), t_0+\sigma) - \phi(x_0, t_0) \leq U(y_{x_0}(t_0+\sigma), t_0+\sigma) - U(x_0, t_0).
\end{aligned}$$

We combine this with the above inequality to yield

$$U(x_0, t_0) + \frac{\theta\sigma}{2} \leq \inf_{b(\cdot)\in\mathcal{B}_{t_0}} \left\{ \int_{t_0}^{t_0+\sigma} h(y_{x_0}(s), \alpha[b](s), b(s))\, ds + U(y_{x_0}(t_0+\sigma), t_0+\sigma) \right\},$$

which contradicts with the DPP for U. The proof is complete. □

1.6 Problems

Exercise 35. *Give an example to show that the minimax condition* (3.5) *does not hold in general.*

Exercise 36. *Give the proof of the bound $\|V_t\|_{L^\infty(\mathbb{R}^n\times[0,T])} \leq C$ in Proposition 3.8.*

2 Representation formulas of solutions of Hamilton–Jacobi equations

We use the results for upper and lower values of two-player zero-sum differential games developed in the previous section to give representation formulas of solutions to Hamilton–Jacobi equations.

2.1 Terminal value problems

We focus on the following problem

$$\begin{cases} u_t + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0, T), \\ u(x, T) = g(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{3.10}$$

We put the following assumptions on H and g. Assume that there exists $C > 0$ such that, for all $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x,0)| \le C, \\ |H(x,p) - H(y,q)| \le C(|x-y| + |p-q|), \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} H(x,p) = +\infty; \end{cases} \tag{3.11}$$

and

$$\|g\|_{L^\infty(\mathbb{R}^n)} + \|Dg\|_{L^\infty(\mathbb{R}^n)} \le C. \tag{3.12}$$

By the results in Chapter 1 (e.g., Theorem 1.34), under assumptions (3.11)–(3.12), (3.10) has a unique viscosity solution $u \in \mathrm{Lip}(\mathbb{R}^n \times [0,T])$. Our goal is to find a representation formula for u. To do so, we aim at writing H as the max-min of appropriate affine functions first.

Lemma 3.9. *Assume* (3.11). *Fix $R > 0$. Let $A = \overline{B}(0,R)$, and $B = \overline{B}(0,1)$. Then, for $x \in \mathbb{R}^n$, and $p \in \overline{B}(0,R)$,*

$$H(x,p) = \max_{a\in A} \min_{b\in B} (H(x,a) + C(p-a)\cdot b).$$

Proof. Thanks to the Lipschitz assumption (3.11) on H, for $x \in \mathbb{R}^n$, and $p \in \overline{B}(0,R)$,

$$H(x,p) = \max_{a\in A} (H(x,a) - C|p-a|).$$

As $-|p-a| = \min_{b\in B}(p-a)\cdot b$, we conclude that

$$H(x,p) = \max_{a\in A} \min_{b\in B} (H(x,a) + C(p-a)\cdot b).$$

□

We are now ready to state our result on a representation formula for u, solution to (3.10).

Theorem 3.10. *Assume* (3.11)–(3.12). *Let u be the unique viscosity solution to* (3.10). *Pick $R > 0$ such that $\|Du\|_{L^\infty(\mathbb{R}^n\times[0,T])} \le R$. Let $A = \overline{B}(0,R)$, and $B = \overline{B}(0,1)$. Then, for $(x,t) \in \mathbb{R}^n \times (0,T)$,*

$$u(x,t) = \inf_{\beta\in\Gamma_t} \sup_{a(\cdot)\in\mathcal{A}_t} \left(\int_t^T (H(y_x(s),a(s)) - Ca(s)\cdot\beta[a](s))\,ds + g(y_x(T)) \right).$$

Here, y_x solves

$$\begin{cases} y_x'(s) = C\beta[a](s) & \text{for } t < s < T, \\ y_x(t) = x. \end{cases}$$

Proof. By Lemma 3.9, for $x \in \mathbb{R}^n$, and $p \in \overline{B}(0,R)$,

$$\begin{aligned} H(x,p) &= \max_{a\in A} \min_{b\in B} (H(x,a) + C(p-a)\cdot b) \\ &= \max_{a\in A} \min_{b\in B} (Cb\cdot p + H(x,a) - Ca\cdot b). \end{aligned}$$

Thus, in terms of two-player zero-sum games, we have $f(x,a,b) = Cb$, and $h(x,a,b) = H(x,a) - Ca\cdot b$. We then employ part (i) of Theorem 3.4 to conclude. □

Remark 3.11. It is worth noting that in the specific situation of the above proof, $a(\cdot) \in \mathcal{A}_t$ means exactly that $a(\cdot) \in L^\infty([t,T],\mathbb{R}^n)$ with $\|a\|_{L^\infty} \le R$.

2.2 Initial value problems

We focus on the following initial value problem

$$\begin{cases} u_t + H(x,Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,T), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{3.13}$$

Like the previous section, we put the following assumptions on H and u_0. Assume that there exists $C > 0$ such that, for all $x, y, p, q \in \mathbb{R}^n$,

$$\begin{cases} |H(x,0)| \leq C, \\ |H(x,p) - H(y,q)| \leq C(|x-y| + |p-q|), \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} H(x,p) = +\infty; \end{cases} \tag{3.14}$$

and

$$\|u_0\|_{L^\infty(\mathbb{R}^n)} + \|Du_0\|_{L^\infty(\mathbb{R}^n)} \leq C. \tag{3.15}$$

Under assumptions (3.14)–(3.15), (3.13) has a unique viscosity solution $u \in \mathrm{Lip}(\mathbb{R}^n \times [0,T])$. We now give a representation formula for u.

Theorem 3.12. *Assume* (3.14)–(3.15). *Let u be the unique viscosity solution to* (3.13). *Pick $R > 0$ such that $\|Du\|_{L^\infty(\mathbb{R}^n\times[0,T])} \leq R$. Let $A = \overline{B}(0,R)$, and $B = \overline{B}(0,1)$. Then, for $(x,t) \in \mathbb{R}^n \times (0,T)$,*

$$u(x,t) = \inf_{\beta\in\Gamma_{T-t}} \sup_{a(\cdot)\in\mathcal{A}_{T-t}} \left(\int_{T-t}^{T} (-H(y_x(s),a(s)) - Ca(s)\cdot\beta[a](s))\, ds + u_0(y_x(T)) \right).$$

Here, y_x solves

$$\begin{cases} y_x'(s) = C\beta[a](s) \quad \text{for } T-t < s < T, \\ y_x(T-t) = x. \end{cases}$$

Again, $a(\cdot) \in \mathcal{A}_{T-t}$ means exactly that $a(\cdot) \in L^\infty([T-t,T],\mathbb{R}^n)$ with $\|a\|_{L^\infty} \leq R$.

Proof. Denote by $v(x,t) = u(x,T-t)$ for all $(x,t) \in \mathbb{R}^n \times [0,T]$. Let $K(x,p) = -H(x,p)$ for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$. Then v solves

$$\begin{cases} v_t + K(x,Dv) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,T), \\ v(x,T) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases}$$

By Lemma 3.9, for $x \in \mathbb{R}^n$, and $p \in \overline{B}(0,R)$,

$$\begin{aligned} K(x,p) = -H(x,p) &= \max_{a\in A} \min_{b\in B} (-H(x,a) + C(p-a)\cdot b) \\ &= \max_{a\in A} \min_{b\in B} (Cb\cdot p - H(x,a) - Ca\cdot b). \end{aligned}$$

Thus, in terms of two-player zero-sum games, we have $f(x,a,b) = Cb$, and $h(x,a,b) = -H(x,a) - Ca\cdot b$. We then employ part (i) of Theorem 3.4 to conclude. ∎

3 The Hopf formula

We aim at deriving a formula of the solution to

$$\begin{cases} u_t + H(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,T), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{3.16}$$

Here, $H = H(p)$ depends only on p. Assume that there exists $C > 0$ such that, for all $x, y, p, q \in \mathbb{R}^n$,

$$|H(p) - H(q)| \le C|p - q|, \tag{3.17}$$

and

$$\|u_0\|_{L^\infty(\mathbb{R}^n)} + \|Du_0\|_{L^\infty(\mathbb{R}^n)} \le C. \tag{3.18}$$

Under assumptions (3.17)–(3.18), (3.16) has a unique viscosity solution $u \in \mathrm{Lip}(\mathbb{R}^n \times [0,T])$. Theorem 3.12 already gives a representation formula for u. It turns out that, if one has further that u_0 is convex, then u has a simpler representation formula, the Hopf formula.

Theorem 3.13 (The Hopf formula)**.** *Assume* (3.17)–(3.18)*. Assume further that u_0 is convex. Let u be the unique viscosity solution to* (3.16)*. Then, for $(x,T) \in \mathbb{R}^n \times (0,\infty)$,*

$$\begin{aligned} u(x,T) &= \sup_{z \in \mathbb{R}^n} \left(x \cdot z - u_0^*(z) - TH(z) \right) \\ &= \sup_{z \in \mathbb{R}^n} \inf_{y \in \mathbb{R}^n} \left(u_0(y) + (x - y) \cdot z - TH(z) \right). \end{aligned}$$

Here, u_0^* is the Legendre transform of u_0.

Proof. Fix $(x,T) \in \mathbb{R}^n \times (0,\infty)$. Let $Z = \{z \in \mathbb{R}^n : u_0^*(z) < +\infty\}$. Of course, $Z \neq \emptyset$ and

$$\sup_{z \in \mathbb{R}^n} \left(x \cdot z - u_0^*(z) - TH(z) \right) = \sup_{z \in Z} \left(x \cdot z - u_0^*(z) - TH(z) \right).$$

For $z \in Z$, denote by

$$\phi_z(x,t) = x \cdot z - u_0^*(z) - tH(z) \qquad \text{for all } (x,t) \in \mathbb{R}^n \times (0,\infty).$$

Clearly, ϕ_z is a classical solution to (3.16), and in light of the Legendre transform,

$$\phi_z(x,0) = x \cdot z - u_0^*(z) \le u_0(x).$$

Hence, $\phi_z \le u$, and in particular,

$$\sup_{z \in Z} \left(x \cdot z - u_0^*(z) - TH(z) \right) = \sup_{z \in Z} \phi_z(x,T) \le u(x,T).$$

We now prove the reverse inequality. Pick $R > 0$ such that $\|Du\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le R$. Let $A = \overline{B}(0,R)$, and $B = \overline{B}(0,1)$. By Theorem 3.12,

$$u(x,T) = \inf_{\beta \in \Gamma_0} \sup_{a(\cdot) \in \mathcal{A}_0} \left(\int_0^T \left(-H(a(s)) - Ca(s) \cdot \beta[a](s) \right) ds + u_0(y_x(T)) \right).$$

Here, y_x solves

$$\begin{cases} y_x'(s) = C\beta[a](s) & \text{for } 0 < s < T, \\ y_x(0) = x. \end{cases}$$

As u_0 is convex, we use Jensen's inequality to yield

$$u_0(y_x(T)) = u_0\left(x + \int_0^T y_x'(s)\,ds\right) = u_0\left(x + \frac{1}{T}\int_0^T Ty_x'(s)\,ds\right) \le \frac{1}{T}\int_0^T u_0(x + Ty_x'(s))\,ds.$$

Let $\beta^* \in \Gamma_0$ be such that $\beta^*[a](s) = \varphi(a(s))$ for a measurable function $\varphi : A \to B$ to be chosen. Then,

$$\begin{aligned} u(x,T) &\le \sup_{a(\cdot)\in\mathcal{A}_0}\left(\int_0^T [-H(a(s)) - a(s)\cdot C\varphi(a(s))]\,ds + u_0(y_x(T))\right) \\ &\le \sup_{a(\cdot)\in\mathcal{A}_0}\left(\frac{1}{T}\int_0^T [-TH(a(s)) - a(s)\cdot TC\varphi(a(s)) + u_0(x + TC\varphi(a(s)))]\,ds\right). \end{aligned}$$

Pick $\varphi : A \to B$ measurable such that, for $a \in A$,

$$x + TC\varphi(a) \in D^- u_0^*(a) = \partial u_0^*(a).$$

This function is well-defined by changing C to be a larger constant if needed. Then,

$$-a(s)\cdot TC\varphi(a(s)) + u_0(x + TC\varphi(a(s))) = a(s)\cdot x - u_0^*(a(s)).$$

We combine this with the above inequality to deduce that

$$\begin{aligned} u(x,T) &\le \sup_{a(\cdot)\in\mathcal{A}_0}\left(\frac{1}{T}\int_0^T [-TH(a(s)) - a(s)\cdot TC\varphi(a(s)) + u_0(x + TC\varphi(a(s)))]\,ds\right) \\ &= \sup_{a(\cdot)\in\mathcal{A}_0}\left(\frac{1}{T}\int_0^T \left[-TH(a(s)) + a(s)\cdot x - u_0^*(a(s))\right]ds\right) \\ &= \sup_{|z|\le C}\left(-TH(z) + z\cdot x - u_0^*(z)\right) \le \sup_{z\in\mathbb{R}^n}\left(-TH(z) + z\cdot x - u_0^*(z)\right). \end{aligned}$$

□

Remark 3.14. Thus far, we have obtained some representation formulas for the viscosity solution to (3.16) under assumptions (3.17)–(3.18). On the one hand, we always have a two-player zero-sum differential game representation formula as stated in Theorem 3.12, which is quite complicated and not easy to be analyzed. On the other hand, if we put an additional convexity assumption on either H or u_0, then we have furthermore the Hopf-Lax formula or the Hopf formula, respectively.

We recall here the Hopf-Lax formula and the Hopf formula for comparison. If H is convex, then the Hopf-Lax formula (Theorem 2.23) reads, for $(x,t) \in \mathbb{R}^n \times (0,\infty)$,

$$\begin{aligned} u(x,t) &= \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + u_0(y)\right\} = \inf_{y\in\mathbb{R}^n}\left\{\sup_{z\in\mathbb{R}^n} t\left(\frac{x-y}{t}\cdot z - H(z)\right) + u_0(y)\right\} \\ &= \inf_{y\in\mathbb{R}^n}\sup_{z\in\mathbb{R}^n}\{u_0(y) + (x-y)\cdot z - tH(z)\} = \inf_{y\in\mathbb{R}^n}\sup_{z\in\mathbb{R}^n}\phi^{y,z}(x,t). \end{aligned}$$

Here, for $(y,z) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\phi^{y,z}(x,t) = u_0(y) + (x-y)\cdot z - tH(z),$$

which is an affine solution to (3.16). If u_0 is convex, then the Lax formula (Theorem 3.13) gives

$$\begin{aligned} u(x,t) &= \sup_{z\in\mathbb{R}^n} \big(x\cdot z - u_0^*(z) - tH(z)\big) \\ &= \sup_{z\in\mathbb{R}^n} \inf_{y\in\mathbb{R}^n} (u_0(y) + (x-y)\cdot z - tH(z)) = \sup_{z\in\mathbb{R}^n} \inf_{y\in\mathbb{R}^n} \phi^{y,z}(x,t). \end{aligned}$$

Basically, we can see that the Hopf-Lax formula is only different from the Lax formula in the order of taking supremum and infimum. Both of the formulas are two-parameter envelopes constructed from $\{\phi^{y,z}\}_{y,z\in\mathbb{R}^n}$.

4 Finite difference approximations

We consider the following first-order equation

$$F(x,u(x),Du(x)) = 0 \qquad \text{in } U. \tag{3.19}$$

Here, $U \subset \mathbb{R}^n$ is a given open set, $F : U \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is a given continuous function, and $u : U \to \mathbb{R}$ is the unknown. Our aim is to provide a certain finite difference approximation to approximate u, a solution to (3.19), based on stability results of viscosity solutions.

4.1 Monotone and consistent schemes

Our generalized approximation schemes consist of a sequence of pairs $(X_k, M_k)_{k\in\mathbb{N}}$ such that $X_k \subset \mathbb{R}^n$ is locally finite, and $M_k : \mathbb{R}^{X_k} \to \mathbb{R}$. Here, $\mathbb{R}^{X_k}$ is the set of real-valued functions on X_k. In this general setting, we do not yet assume that X_k is of grid form in $\mathbb{R}^n$. We require the following assumptions.

- Density condition: There exists a sequence $\{\delta_k\} \to 0$ such that, for $k \in \mathbb{N}$,

$$\sup_{x\in U} \operatorname{dist}(x, X_k \cap U) < \delta_k. \tag{3.20}$$

- Locality condition: There exists a sequence $\{\varepsilon_k\} \to 0$ such that, for $k \in \mathbb{N}$,

$$\begin{cases} \text{for } u,v \in \mathbb{R}^{X_k}, \text{ and } x \in X_k \cap U, \\ \text{if } u = v \text{ on } X_k \cap B(x,\varepsilon_k), \text{ then } M_k u(x) = M_k v(x). \end{cases} \tag{3.21}$$

- Monotonicity condition:

$$\begin{cases} \text{for } u,v \in \mathbb{R}^{X_k}, \text{ and } x \in X_k \cap U, \\ \text{if } u \text{ touches } v \text{ from above at } x, \text{ then } M_k u(x) \le M_k v(x). \end{cases} \tag{3.22}$$

- Consistency condition: For any real sequence $\{s_k\} \to 0$,

$$\begin{cases} \text{for } \varphi \in C^1(U) \text{ and } x \in U, \text{ if } x_k \in X_k \text{ for } k \in \mathbb{N}, \text{ and } \lim_{k\to\infty} x_k = x, \\ \text{then } \lim_{k\to\infty} M_k(\varphi + s_k)(x_k) = F(x,\varphi(x),D\varphi(x)). \end{cases} \tag{3.23}$$

The density condition (3.20) can be stated in an equivalent way as following. There exist a sequence $\{\delta_k\} \to 0$ and a sequence of maps $\{\pi_k\}$ such that, for $k \in \mathbb{N}$, $\pi_k : U \to X_k \cap U$, and

$$\sup_{x \in U} |\pi_k(x) - x| < \delta_k. \tag{3.24}$$

Theorem 3.15. *Assume* (3.20)–(3.23). *Let $\{\pi_k\}$ be a sequence of maps as in* (3.24). *For each $k \in \mathbb{N}$, let $u_k \in \mathbb{R}^{X_k}$ be a solution to $M_k u_k = 0$ in $X_k \cap U$. Assume that $u_k \circ \pi_k \to u$ locally uniformly in U for some $u \in C(U)$. Then, u is a viscosity solution to* (3.19).

Proof. We only show that u is a viscosity subsolution to (3.19).
Pick $\phi \in C^1(U)$ such that $u - \phi$ has a strict maximum at $y \in U$ and $u(y) = \phi(y)$. As $u_k \circ \pi_k \to u$ locally uniformly in U, for $k \in \mathbb{N}$ sufficiently large, we imply that $u_k - \phi$ has a local maximum over $X_k \cap U$ at $x_k \in X_k \cap U$, and $\lim_{k\to\infty} x_k = y$. For $k \in \mathbb{N}$, let $s_k = u_k(x_k) - \phi(x_k)$. Clearly, $\lim_{k\to\infty} s_k = 0$. Then, $\phi + s_k$ touches u_k from above at $x_k \in X_k$, which yields

$$M_k(\phi + s_k)(x_k) \le M_k u_k(x_k) = 0.$$

Let $k \to \infty$ to get further that

$$F(y, u(y), D\phi(y)) = F(y, \phi(y), D\phi(y)) = \lim_{k\to\infty} M_k(\phi + s_k)(x_k) \le 0.$$

The proof is complete. □

4.2 Examples

In all the following examples, X_k are always chosen to be of grid form for all $k \in \mathbb{N}$. We first consider a simple transport equation.

Example 3.2 (Transport equation in one dimension)**.**

$$\begin{cases} u_t + c u_x = 0 & \text{in } \mathbb{R} \times (0, \infty), \\ u(x, 0) = u_0(x) & \text{on } \mathbb{R}. \end{cases} \tag{3.25}$$

Here, $c > 0$ is a positive constant, and $u_0 \in \mathrm{BUC}(\mathbb{R}) \cap C^1(\mathbb{R})$ is a given function. It is straightforward that the unique classical solution to (3.25) *is $u(x,t) = u_0(x - ct)$ for all $(x,t) \in \mathbb{R} \times [0, \infty)$. Let us consider a finite difference scheme for* (3.25) *despite this fact.*
For each $k \in \mathbb{N}$, fix mesh sizes $\Delta x, \Delta t > 0$. Denote by

$$X_k = \{(j\Delta x, n\Delta t) : j \in \mathbb{Z}, n \ge 0\}.$$

It is clear that (3.20) *holds provided that the mesh sizes tend to 0 as $k \to \infty$. For $U \in \mathbb{R}^{X_k}$, we write $U_j^n = U(j\Delta x, n\Delta t)$. To approximate u_t, we take*

$$u_t \approx \frac{U_j^{n+1} - U_j^n}{\Delta t}.$$

For u_x, we pick

$$u_x \approx \frac{U_j^n - U_{j-1}^n}{\Delta x}.$$

Our scheme is

$$\frac{U_j^{n+1}-U_j^n}{\Delta t}+c\frac{U_j^n-U_{j-1}^n}{\Delta x}=0.$$

By our choices, (3.21) *and* (3.23) *hold. The above can be rewritten as*

$$U_j^{n+1}=\left(1-c\frac{\Delta t}{\Delta x}\right)U_j^n+c\frac{\Delta t}{\Delta x}U_{j-1}^n. \tag{3.26}$$

In order to have (3.22)*, the monotonicity condition, we need to require that*

$$c\frac{\Delta t}{\Delta x}\leq 1 \qquad \Longleftrightarrow \qquad \frac{\Delta x}{\Delta t}\geq c, \tag{3.27}$$

which gives further that U_j^{n+1} *is a convex combination of* U_j^n *and* U_{j-1}^n*. The scheme* (3.26) *is called an upwind scheme in the literature. If* (3.27) *does not hold, then* $1-c\frac{\Delta t}{\Delta x}<0$ *in* (3.26)*, and the numerically computed solution might blow up.*
It is worth noting furthermore that our scheme can be rewritten in another way as

$$\frac{U_j^{n+1}-U_j^n}{\Delta t}+c\frac{U_{j+1}^n-U_{j-1}^n}{2\Delta x}=\frac{c\Delta x}{2}\frac{U_{j+1}^n-2U_j^n+U_{j-1}^n}{(\Delta x)^2},$$

which looks pretty much like a discretization of

$$u_t+cu_x=\frac{c\Delta x}{2}u_{xx}.$$

The term on the right hand side is called a numerical viscosity. It is clear that $c>0$ *is really needed in this scheme.*

Next, we study a first-order Hamilton–Jacobi equation in one dimension.

Example 3.3 (First-order Hamilton–Jacobi equation in one dimension)**.**

$$\begin{cases} u_t+H(u_x) &=0 \qquad \text{in } \mathbb{R}\times(0,\infty),\\ u(x,0) &=u_0(x) \qquad \text{on } \mathbb{R}. \end{cases} \tag{3.28}$$

Here, $H\in \mathrm{Lip}(\mathbb{R})$*, and* $u_0\in \mathrm{BUC}(\mathbb{R})\cap \mathrm{Lip}(\mathbb{R})$ *are given functions.*
For each $k\in\mathbb{N}$*, fix mesh sizes* $\Delta x,\Delta t>0$*. Set*

$$X_k=\{(j\Delta x,n\Delta t)\,:\,j\in\mathbb{Z},n\geq 0\}.$$

Assume that $\lambda_x=\frac{\Delta t}{\Delta x}$ *is fixed and is independent of* $k\in\mathbb{N}$*. It is clear that* (3.20) *holds provided that the mesh sizes tend to* 0 *as* $k\to\infty$*. For* $U\in\mathbb{R}^{X_k}$*, we write* $U_j^n=U(j\Delta x,n\Delta t)$*. Our scheme is*

$$\frac{U_j^{n+1}-U_j^n}{\Delta t}+H\left(\frac{U_{j+1}^n-U_{j-1}^n}{2\Delta x}\right)=\frac{\theta}{\lambda_x}\frac{U_{j+1}^n-2U_j^n+U_{j-1}^n}{\Delta x}.$$

The constant $\theta>0$ *will be chosen later. As above, based on the choice of our schemes,* (3.21) *and* (3.23) *hold naturally. We only need to check carefully the monotonicity condition* (3.22)*. The above relation can be rewritten as*

$$\begin{aligned} U_j^{n+1}&=U_j^n-\Delta t\,H\left(\frac{U_{j+1}^n-U_{j-1}^n}{2\Delta x}\right)+\theta(U_{j+1}^n-2U_j^n+U_{j-1}^n) \qquad (3.29)\\ &=(1-2\theta)U_j^n+\theta(U_{j+1}^n+U_{j-1}^n)-\Delta t\,H\left(\frac{U_{j+1}^n-U_{j-1}^n}{2\Delta x}\right). \end{aligned}$$

To have the monotonicity condition, we need to require that

$$0 < \theta < \frac{1}{2} \qquad \text{and} \qquad \theta > \frac{\lambda_x}{2} \|H'\|_{L^\infty(\mathbb{R})},$$

which means that

$$\frac{\lambda_x}{2} \|H'\|_{L^\infty(\mathbb{R})} < \theta < \frac{1}{2}. \tag{3.30}$$

Condition (3.30) *can be achieved by first choosing* $\theta \in (0, \frac{1}{2})$*, and then* $\lambda_x > 0$ *small enough. This scheme is analogous to the Lax-Friedrichs scheme for conservation laws.*

In fact, the ideas in Example 3.3 can be generalized to multi dimensional settings rather naturally. For simplicity of the presentation, we only consider a two dimensional example below.

Example 3.4 (First-order Hamilton–Jacobi equation in two dimensions)**.** *Here, we write* $(x, y) \in \mathbb{R}^2$ *as a variable. The equation of interests is*

$$\begin{cases} u_t + H(u_x, u_y) &= 0 \qquad \text{in } \mathbb{R}^2 \times (0, \infty), \\ u(x, y, 0) &= u_0(x, y) \qquad \text{on } \mathbb{R}^2. \end{cases} \tag{3.31}$$

The Hamiltonian $H \in \mathrm{Lip}(\mathbb{R}^2)$*, and the initial data* $u_0 \in \mathrm{BUC}(\mathbb{R}^2) \cap \mathrm{Lip}(\mathbb{R}^2)$ *are given functions.*

For each $k \in \mathbb{N}$*, fix mesh sizes* $\Delta x, \Delta y, \Delta t > 0$*. Set*

$$X_k = \{(j\Delta x, l\Delta y, n\Delta t) \, : \, j, l \in \mathbb{Z}, n \geq 0\}.$$

Assume that $\lambda_x = \frac{\Delta t}{\Delta x}$, $\lambda_y = \frac{\Delta t}{\Delta y}$ *are fixed and are independent of* $k \in \mathbb{N}$*. It is clear that* (3.20) *holds provided that the mesh sizes tend to* 0 *as* $k \to \infty$*. For* $U \in \mathbb{R}^{X_k}$*, we write* $U^n_{j,l} = U(j\Delta x, l\Delta y, n\Delta t)$*. Similar to the above, our scheme is*

$$\frac{U^{n+1}_{j,l} - U^n_{j,l}}{\Delta t} + H\left(\frac{U^n_{j+1,l} - U^n_{j-1,l}}{2\Delta x}, \frac{U^n_{j,l+1} - U^n_{j,l-1}}{2\Delta y} \right) = \frac{\theta}{\lambda_x} \frac{U^n_{j+1,l} - 2U^n_{j,l} + U^n_{j-1,l}}{\Delta x} + \frac{\theta}{\lambda_y} \frac{U^n_{j,l+1} - 2U^n_{j,l} + U^n_{j,l-1}}{\Delta y}.$$

In order to have a monotone and consistent scheme, we need to require that

$$\max\left\{ \frac{\lambda_x}{2} \|H_x\|_{L^\infty(\mathbb{R}^2)}, \frac{\lambda_y}{2} \|H_y\|_{L^\infty(\mathbb{R}^2)} \right\} < \theta < \frac{1}{4}. \tag{3.32}$$

Condition (3.32) *can be achieved by first choosing* $\theta \in (0, \frac{1}{4})$*, and then* $\lambda_x, \lambda_y > 0$ *small enough.*

5 References

1. Two-player zero-sum differential games in the context of viscosity solutions were first analyzed by Evans and Souganidis [54]. We refer the readers to the books of Bardi, Capuzzo-Dolcetta [13], Elliott [45], and the lecture notes by Cardaliaguet [29] for more details and references.

2. The Hopf formula was proposed by Hopf [80]. It was then rigorously verified in the context of viscosity solutions by Bardi and Evans [14]. For different proofs, see Lions, Rochet [103] and Cardaliaguet [29].

3. The framework of monotone and consistent schemes was proposed by Barles and Souganidis [17]. For finite difference approximations of first-order Hamilton–Jacobi equations, see Souganidis [130], Crandall, Lions [40] for further details. For the Lax-Friedrichs scheme for conservation laws, see Crandall and Majda [41]. We refer the readers to Chapter 5 of the book by Osher, Fedkiw [121] for further developments.

CHAPTER 4

Periodic homogenization theory for Hamilton–Jacobi equations

1 Introduction to periodic homogenization theory

1.1 Introduction

Homogenization theory has been blossoming in last couple of decades in various different directions for many kind of PDEs. In this chapter, we only focus on the periodic homogenization theory for Hamilton–Jacobi equations. The equations of interest are as following. For each $\varepsilon > 0$, we study

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{4.1}$$

Here, the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is continuous and satisfies some appropriate conditions to be addressed soon. We often assume that the initial data $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ unless otherwise specified.

In practice, $\varepsilon > 0$ is a fixed length scale, which is quite small. If we zoom in the system to the scale ε, we see the whole microstructure, and this is represented in (4.1) by the highly oscillatory variable $\frac{x}{\varepsilon}$. Of course, the Hamiltonian can be much more complex with various different scales such as $H = H(x, \frac{x}{\varepsilon}, \frac{x}{\varepsilon^{s_1}}, \ldots, \frac{x}{\varepsilon^{s_m}}, p)$ for given $s_1, \ldots, s_m > 0$, a typical multi-scale problem. We here focus on the simplest case $H = H(\frac{x}{\varepsilon}, p)$. Yet, dealing with this problem is already quite challenging, especially numerically as in order to be able to compute/approximate the solution accurately, one needs to have approximation schemes of sizes smaller than ε (or $O(\varepsilon)$). Otherwise, the microstructure will be missed.

Typically, the microstructure in the system is repeated somehow, and this gives hope for us to see (nonlinear) averaging effects. In this entire chapter, we assume that the microstructure is periodic, which is the most idealistic situation. Then, mathematically, we let $\varepsilon \to 0$ in (4.1), and we expect that u^ε converges to u as $\varepsilon \to 0$ in some sense, and u solves a certain averaging (effective) equation, which is simpler somewhat.

The above gives a minimalistic introduction to homogenization theory. Basic questions of interests are as follows.

1. Qualitative theory: Find out the effective equation, and show convergence of u^ε to u in some functional spaces.

2. Better understanding of the effective equation: Since the problem is nonlinear, it is extremely important to analyze the effective equation in various aspects.

3. Quantitative theory: Quantify the convergence of u^ε to u, and if possible, find optimal rate of convergence.

4. Numerics: Up to now, there have been very few results in this direction since the equations are highly nonlinear.

1.2 Derivations

Our focus is on equation (4.1) for each $\varepsilon > 0$. And our goal is to let $\varepsilon \to 0$ to observe a certain nonlinear averaging behavior.

Basic assumptions. Throughout this chapter, we assume the following two assumptions.

$$y \mapsto H(y,p) \text{ is } \mathbb{Z}^n\text{-periodic, that is, } H(y,p) = H(y+k,p) \text{ for } k \in \mathbb{Z}^n, \tag{4.2}$$

and

$$\lim_{|p|\to\infty} H(y,p) = +\infty \text{ uniformly for } y \in \mathbb{R}^n. \tag{4.3}$$

We can think about our current problem (4.1) as a multi-scale problem

- x is the macroscopic scale variable or low scale variable;
- $y = \frac{x}{\varepsilon}$ is the microscopic scale variable or fast scale variable.

The relation $x = \varepsilon y$ can be heuristically understood as when x changes a little bit of order $O(\varepsilon)$, we have y varies correspondingly a lot, that is, y sees the small changes in the environment. Conversely, when y changes a little (of order $O(1)$ or less), x does not see that essentially. Microscopically, the system is very complicated, even in the case we can use the optimal control representation formula as in the following example.

Example 4.1. *Consider again the classical mechanics Hamiltonian*

$$H(y,p) = \frac{1}{2}|p^2| + V(y) \qquad \textit{for all } (y,p) \in \mathbb{R}^n \times \mathbb{R}^n,$$

where $V \in C(\mathbb{R}^n)$ is $\mathbb{Z}^n$-periodic. Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ be the usual flat n-dimensional torus. We often write $V \in C(\mathbb{T}^n)$. Then the corresponding problem is

$$\begin{cases} u^\varepsilon_t + \frac{1}{2}|Du^\varepsilon|^2 + V\left(\frac{x}{\varepsilon}\right) & = 0 \qquad \textit{in } \mathbb{R}^n \times (0,\infty) \\ u^\varepsilon(x,0) & = u_0(x) \qquad \textit{on } \mathbb{R}^n. \end{cases}$$

Recall the Legendre transform $L(y,v) = \frac{1}{2}|v|^2 - V(y)$ for all $(y,v) \in \mathbb{R}^n \times \mathbb{R}^n$, we have the optimal control representation formula

$$u^\varepsilon(x,t) = \inf\left\{ \int_0^t \left[\frac{1}{2}|\gamma'(s)|^2 - V\left(\frac{\gamma(s)}{\varepsilon}\right)\right] ds \,:\, \gamma(t) = x, \gamma'(\cdot) \in L^1([0,t]) \right\}.$$

By change of variables,

$$u^\varepsilon(x,t) = \inf\left\{ \varepsilon \int_0^{t/\varepsilon} \left[\frac{1}{2}|\xi'(s)|^2 - V(\xi(s))\right] ds \,:\, \xi(t/\varepsilon) = x, \xi'(\cdot) \in L^1([0,t/\varepsilon]) \right\}.$$

This formula is extremely interesting and complicated at the same time. Basically, it is a large time average of some action functionals of the type $\frac{1}{T}\int_0^T L(\cdot)\,ds$ as $T = \varepsilon^{-1} \to +\infty$. And since we go for large time, the admissible paths ξ are able to explore all possible locations in the periodic environment, and thus, homogenization (large time average) should occur. It is however not clear at all what is the limit if there is any. The nonlinear dependence is quite twisted here in the formula between the two terms that makes it really hard to understand deeper. We will revisit this optimal control viewpoint for convex Hamiltonians later.

Heuristic arguments. We introduce the following ansatz[1] as an expansion of u^ε in ε

$$\begin{aligned} u^\varepsilon(x,t) &= u^0\left(x, \frac{x}{\varepsilon}, t\right) + \varepsilon u^1\left(x, \frac{x}{\varepsilon}, t\right) + \varepsilon^2 u^2\left(x, \frac{x}{\varepsilon}, t\right) + \dots \\ &= u^0(x,y,t) \ + \varepsilon u^1(x,y,t) \ + \varepsilon^2 u^2(x,y,t) \ + \dots \end{aligned}$$

Then,

$$\begin{aligned} u_t^\varepsilon(x,t) &= u_t^0(x,y,t) \ + \varepsilon u_t^1(x,y,t) \ + \varepsilon^2 u_t^2(x,y,t) + \dots \\ D_x u^\varepsilon(x,t) &= D_x u^0(x,y,t) + \frac{1}{\varepsilon} D_y u^0(x,y,t) + \varepsilon D_x u^1(x,y,t) + D_y u^1(x,y,t) + O(\varepsilon). \end{aligned}$$

Now, this is a crucial point. Think about x, y as independent variables, that is, x, y are unrelated. Although it is not true from the heuristic setting $x = \varepsilon y$, but from the explanation of separation of scales earlier (macroscopic variable x, and microscopic variable y), it sort of makes sense.

Put the above expansions into (4.1) to get

$$u_t^0 + O(\varepsilon) + H\left(y, D_x u^0 + \frac{1}{\varepsilon} D_y u^0 + \varepsilon D_x u^1 + D_y u^1 + O(\varepsilon)\right) = 0. \tag{4.4}$$

Heuristically, if $|D_y u^0| \neq 0$ then $\frac{1}{\varepsilon}|D_y u^0| \to \infty$ as $\varepsilon \to 0$, thus it forces

$$H\left(y, D_x u^0 + \frac{1}{\varepsilon} D_y u^0 + \varepsilon D_x u^1 + D_y u^1 + O(\varepsilon)\right) \to \infty \qquad \text{as } \varepsilon \to 0,$$

by the coercivity of H, and hence, (4.4) does not hold. Thus, we must have $D_y u^0 \equiv 0$, that is, $u^0(x,y,t) \equiv u^0(x,t)$, and (4.4) becomes

$$u_t^0 + O(\varepsilon) + H\left(y, D_x u^0 + D_y u^1 + O(\varepsilon)\right) = 0.$$

Let $\varepsilon \to 0$ to yield further that

$$u_t^0(x,t) + H\left(y, D_x u^0(x,t) + D_y u^1(x,y,t)\right) = 0.$$

[1]*An ansatz* means *a formulation* or *an educated guess*

Since $D_x u^1, u^1_t$ only play a role at $O(\varepsilon)$ level of expansions, let us take $u^1(x,y,t) \equiv u^1(y)$, then we get

$$H\left(y, D_x u^0(x,t) + D_y u^1(y)\right) = -u^0_t(x,t).$$

Recall that we have assumed that x and y are unrelated. Fix $(x,t) \in \mathbb{R}^n \times [0,\infty)$, and think of y as the only running variable, then we arrive at an equation for $y \mapsto u^1(y)$ as

$$H\left(y, \underbrace{D_x u^0(x,t)}_{p\in\mathbb{R}^n} + D_y u^1(y)\right) = \underbrace{-u^0_t(x,t)}_{c\in\mathbb{R}} \qquad \text{in } \mathbb{R}^n.$$

Let us recast it as following. Fix $p \in \mathbb{R}^n$, we would like to solve

$$H(y, p + Du^1(y)) = c \qquad \text{in } \mathbb{R}^n.$$

As H is periodic in y, we can think of the above problem in $\mathbb{T}^n$ as well. If it is solvable, and if we are able to find a unique constant $c \in \mathbb{R}$ so that it has a solution u^1, then denote by $\overline{H}(p) = c$. It is not trivial and clear at all if we are able to show this, but let us take it for granted for now.

It is then clear from the ansatz that $u^\varepsilon(x,t) \approx u^0(x,t) + \varepsilon u^1(y) \to u^0(x,t)$ as $\varepsilon \to 0$, and u^0 solves

$$\begin{cases} u^0_t(x,t) + \overline{H}\left(Du^0(x,t)\right) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad\qquad u^0(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases}$$

This is an effective equation, and clearly, homogenization was achieved at the heuristic level. Of course, there were many heuristic ideas in the above derivation (including the facts that we have asymptotic expansions, we have x and y are unrelated, and we have the existence and uniqueness of constant c above). We need somehow to verify these at the rigorous level, and we will see that not all are that clear.

Remark 4.1. The above derivation also works well for the following general degenerate viscous Hamilton–Jacobi equation

$$w^\varepsilon_t(x,t) + H\left(\frac{x}{\varepsilon}, Dw^\varepsilon\right) = \varepsilon \operatorname{tr}\left(A\left(\frac{x}{\varepsilon}\right) D^2 w^\varepsilon\right) \qquad \text{in } \mathbb{R}^n \times (0,\infty).$$

Here, H satisfies (4.2) and (4.3). The diffusion matrix $A(y)$ is a symmetric, nonnegative definite matrix of size n for all $y \in \mathbb{R}^n$. Besides, the map $y \mapsto A(y)$ is $\mathbb{Z}^n$-periodic, Lipschitz. There are two points to note here. First, $A(\cdot)$ might be degenerate in some directions or all directions at various locations, so the diffusion is not helpful in general. Second, as we put the factor ε in front of the diffusion, its effect, if there is any, vanishes anyhow as $\varepsilon \to 0$. In other words, in the limit, we should only see the effective equation of first-order type.

Following the above derivation, we think of

$$w^\varepsilon(x,t) = w^0(x,t) + \varepsilon w^1(y) + \dots.$$

Then, for fixed $p \in \mathbb{R}^n$, we solve

$$H(y, p + Dw^1(y)) - \operatorname{tr}(A(y) D^2 w^1(y)) = c \qquad \text{in } \mathbb{R}^n.$$

Here, $c \in \mathbb{R}$ is an unknown constant.

2 Cell problems and periodic homogenization of static Hamilton–Jacobi equations

Recall that, for (4.1), we introduced the ansatz $u^\varepsilon(x,t) \approx u^0(x,t) + \varepsilon u^1(y)$ where $y = \frac{x}{\varepsilon}$. Here, x is the macroscopic variable, and y is the microscopic variable. See Figure 4.1. Then,

$$u_t^0(x,t) + H\left(y, Du^0(x,t) + Du^1(y)\right) = 0 \qquad \text{in } \mathbb{R}^n \times [0,\infty).$$

Fix $(x,t) \in \mathbb{R}^n \times [0,\infty)$, and think of y as a variable. Let $p = Du^0(x,t) \in \mathbb{R}^n$, and $-c = u_t^0(x,t) \in \mathbb{R}$. Then, we have the following PDE for u^1

$$H(y, p + Du^1(y)) = c \qquad \text{in } \mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n. \tag{E_p}$$

We call (E_p) the cell problem corresponding to $p \in \mathbb{R}^n$. In the literature, it is also called the ergodic problem or the corrector problem corresponding to $p \in \mathbb{R}^n$.

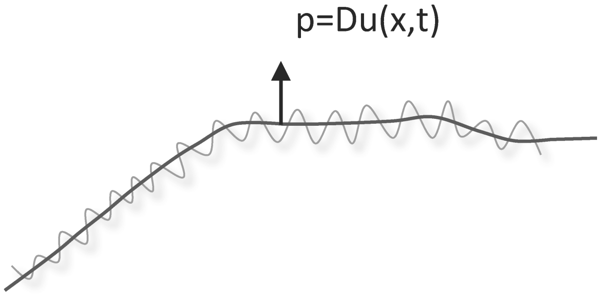


Figure 4.1: An example of graphs of u^ε and $u = u^0$ near (x,t).

2.1 Cell problems

In this section, we discuss the cell problems, which were studied first by Lions, Papanicolaou, and Varadhan [102].

Theorem 4.2. *Assume that H satisfies* (4.2) *and* (4.3). *Fix $p \in \mathbb{R}^n$. There exists a unique constant $c \in \mathbb{R}$ such that the cell problem (E_p) has a viscosity solution $v \in \mathrm{Lip}(\mathbb{T}^n)$.*

Definition 4.3. *Assume that H satisfies* (4.2) *and* (4.3). *For each $p \in \mathbb{R}^n$, Theorem 4.2 gives us the existence and uniqueness of a constant $c \in \mathbb{R}$ such that the cell problem (E_p) has a viscosity solution $v \in \mathrm{Lip}(\mathbb{T}^n)$. We denote by $\overline{H}(p) = c$. We call $\overline{H} : \mathbb{R}^n \to \mathbb{R}$ the effective Hamiltonian.*

It is worth noting right away that as (E_p) is nonlinear, behavior $\overline{H}$ is very complicated and does not depend on H in a linear way. In particular, there is no explicit formula for $\overline{H}$. We will study properties of $\overline{H}$ soon. There are, however, many open questions along this direction.

Proof of theorem 4.2. For $\lambda > 0$, we consider the static equation

$$\lambda v^\lambda + H(y, p + Dv^\lambda) = 0 \qquad \text{in } \mathbb{R}^n. \tag{4.5}$$

By Theorem 1.27, we get that there exists a unique viscosity solution $v^\lambda \in \mathrm{Lip}(\mathbb{R}^n)$ of (4.5). We prove that indeed v^λ is $\mathbb{Z}^n$-periodic. For each $k \in \mathbb{Z}^n$,

$$\lambda v^\lambda(y+k) + H(y+k, p+Dv^\lambda(y+k)) = 0 \qquad \Longrightarrow \qquad \lambda v^\lambda(y+k) + H\big(y, p+Dv^\lambda(y+k)\big) = 0$$

since $y \mapsto H(y,p)$ is $\mathbb{Z}^n$-periodic. Thus, $y \mapsto v(y+k)$ is also a (viscosity) solution to (4.5), and hence, $v^\lambda(y+k) = v^\lambda(y)$ for all $k \in \mathbb{Z}^n$ by the uniqueness of solutions to (4.5). In particular, we can think of $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ now.

Next, take $C_0 = \max_{y \in \mathbb{T}^n} |H(y,p)|$. It is clear that $\frac{C_0}{\lambda}$ and $-\frac{C_0}{\lambda}$ are a viscosity supersolution and subsolution of (4.5), respectively, thus by the comparison principle, we have

$$\sup_{y \in \mathbb{T}^n} |\lambda v^\lambda(y)| \leq C_0.$$

Plug it into (4.5) again, recall that $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ thus it is differentiable a.e., then in the a.e. sense (4.5) becomes

$$|H(y, p + Dv^\lambda(y))| \leq C_0 \qquad \text{for a.e. } y \in \mathbb{T}^n.$$

By coercivity of H we deduce that $\|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \leq C_1$ independent of $\lambda > 0$. Note that the above estimates were already in Theorem 1.26. We redo them here for clarity.

For each $\lambda > 0$, denote by

$$w^\lambda(y) = v^\lambda(y) - v^\lambda(0) \qquad \text{for all } y \in \mathbb{T}^n.$$

Then, as the diameter of $[0,1]^n$ is $\sqrt{n}$,

$$\|w^\lambda\|_{L^\infty(\mathbb{T}^n)} \leq \sqrt{n}\,\|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \leq C, \qquad \text{and} \qquad \|Dw^\lambda\|_{L^\infty(\mathbb{T}^n)} = \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \leq C.$$

In particular, $\{w^\lambda\}_{\lambda>0}$ is equi-continuous on $\mathbb{T}^n$. By the Arzelà–Ascoli theorem, there exists a subsequence $\{\lambda_j\} \to 0$ such that

$$\begin{cases} w^{\lambda_j} = v^{\lambda_j}(\cdot) - v^{\lambda_j}(0) \to v(\cdot) & \text{uniformly on } \mathbb{T}^n, \\ \lambda_j v^{\lambda_j}(0) \to -c \in \mathbb{R} \end{cases}$$

for some $c \in \mathbb{R}$. It is clear that $\min_{\mathbb{T}^n} v = 0$ and $\|Dv\|_{L^\infty(\mathbb{T}^n)} \leq C$. Note that w^λ solves the following equation in the viscosity sense

$$\lambda w^\lambda(y) + H\left(y, p + Dw^\lambda(y)\right) = -\lambda v^\lambda(0) \qquad \text{in } \mathbb{T}^n.$$

By stability results for viscosity solutions, one has that v solves

$$H(y, p + Dv(y)) = c \qquad \text{in } \mathbb{T}^n. \tag{4.6}$$

Thus we obtain a pair $(v, c) \in \mathrm{Lip}(\mathbb{T}^n) \times \mathbb{R}$, which solves the cell problem.

What is left is to prove that c is unique. Indeed, assume that $(v_1,c_1),(v_2,c_2)\in C(\mathbb{T}^n)\times\mathbb{R}$ with $c_1<c_2$ are both solutions to the cell problem. Then,

$$H(y,p+Dv_1(y))=c_1<c_2=H(y,p+Dv_2(y)) \qquad \text{in } \mathbb{T}^n.$$

Note that we have right away that $v_1,v_2\in \mathrm{Lip}(\mathbb{T}^n)$ by Lemma 1.28. Since v_1,v_2 are bounded in $\mathbb{T}^n$, we can find $\delta>0$ sufficiently small such that[2]

$$\delta v_1(y)+H(y,p+Dv_1(y))<\frac{c_1+c_2}{2}<\delta v_2(y)+H(y,p+Dv_2(y)) \qquad \text{in } \mathbb{T}^n.$$

Thus v_1 and v_2 are a subsolution and a supersolution to $\delta w+H(y,p+Dw)=\frac{1}{2}(c_1+c_2)$ in $\mathbb{T}^n$, respectively. By the usual comparison principle for this static problem we obtain $v_1\le v_2$. As (v_1+C,c_1) is also a pair solution to the cell problem (4.6) for any $C>0$, by repeating the above steps, we also get $v_1+C\le v_2$, which is a contradiction. Thus, we must have $c_1=c_2$ and hence the constant $c=\overline{H}(p)$ is unique. □

Remark 4.4. Some comments are in order.

1. It is worth noting first that (E_p) is not monotone in v, and solutions $v\in \mathrm{Lip}(\mathbb{T}^n)$ to (E_p) are not unique. In fact, if $v\in \mathrm{Lip}(\mathbb{T}^n)$ is a solution, then so is $v+C$ for any constant $C\in\mathbb{R}$. In many cases, there are other family of nontrivial solutions to (E_p). This is a very important phenomenon, which deserves further and deeper analysis. For now, the convex case is handled, but not so much is known for nonconvex cases.

2. As $\|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)}\le C$ independent of λ, and $\lim_{\lambda\to 0}\lambda v^\lambda(0)=-\overline{H}(p)$, we get

$$\lambda v^\lambda(\cdot)\to -\overline{H}(p) \qquad \text{uniformly in } \mathbb{T}^n \text{ as } \lambda\to 0.$$

   In the following exercise, we can see that this convergence has rate $O(\lambda)$. But it is important pointing out that it does not give any detailed information about $\overline{H}$.

3. In the above proof, we only achieve the convergence of $v^\lambda(\cdot)-v^\lambda(x_0)\to v(\cdot)$ along a subsequence $\{\lambda_j\}\to 0$. The question on whether or not one has this convergence for the whole sequence $\lambda\to 0$ is extremely interesting, and it is basically a selection problem on vanishing discount.

2.2 Problems

Exercise 37. *Assume that H satisfies* (4.2) *and* (4.3)*. Fix $p\in\mathbb{R}^n$, and we look at* (4.5)*. Show that there exists a constant $C>0$ independent of $\lambda>0$ such that, for any $\lambda>0$, we have*

$$\left\|\lambda v^\lambda(\cdot)+\overline{H}(p)\right\|_{L^\infty(\mathbb{T}^n)}\le C\lambda.$$

Exercise 38. *Let $\psi\in C^1(\mathbb{T}^n)$ be given, and $H(y,p)=p\cdot(p-D\psi(y))$ for $(y,p)\in\mathbb{T}^n\times\mathbb{R}^n$. It is clear that H satisfies* (4.2) *and* (4.3)*. Find $\overline{H}(0)$ and various solutions to (E_p) with $p=0$.*

[2]Indeed, δ can be chosen such that $\delta\left\{\max_{y\in\mathbb{T}^n}|v_1(y)|,\max_{y\in\mathbb{T}^n}|v_2(y)|\right\}<\frac{c_2-c_1}{2}$.

2.3 Periodic homogenization of static Hamilton–Jacobi equations

Let us now prove the periodic homogenization of static Hamilton–Jacobi equations. This is just a simple consequence of Theorem 4.2. Recall the discounted problem (4.5), which can be viewed in terms of $y = \frac{x}{\lambda}$ as

$$\lambda v^\lambda\left(\frac{x}{\lambda}\right) + H\left(\frac{x}{\lambda}, p + Dv^\lambda\left(\frac{x}{\lambda}\right)\right) = 0 \qquad \text{in } \mathbb{R}^n.$$

Let $u^\lambda(x) = \lambda v^\lambda\left(\frac{x}{\lambda}\right)$, then $Du^\lambda(x) = Dv^\lambda\left(\frac{x}{\lambda}\right)$. The above equation becomes

$$u^\lambda(x) + H\left(\frac{x}{\lambda}, p + Du^\lambda(x)\right) = 0 \qquad \text{in } \mathbb{R}^n. \tag{4.7}$$

Clearly, (4.7) is a homogenization problem for static Hamilton–Jacobi equations. We already knew that $u^\lambda \to -\overline{H}(p)$ uniformly in $\mathbb{R}^n$. But let us pretend that we do not have this, and only expect that $u^\lambda \to u$ locally uniformly in $\mathbb{R}^n$, and if homogenization holds, we have that u solves

$$u + \overline{H}(p + Du) = 0 \qquad \text{in } \mathbb{R}^n.$$

A bit of analysis shows that the unique solution to the above is $u \equiv -\overline{H}(p)$, and therefore, everything is consistent. Let us record this here as a corollary.

Corollary 4.5. *Assume that H satisfies* (4.2) *and* (4.3). *Fix $p \in \mathbb{R}^n$, and we study the homogenization problem* (4.7). *As $\lambda \to 0$, $u^\lambda \to u \equiv -\overline{H}(p)$ uniformly in $\mathbb{R}^n$. In fact, there is a constant $C > 0$ independent of λ such that*

$$\|u^\lambda + \overline{H}(p)\|_{L^\infty(\mathbb{R}^n)} \leq C\lambda.$$

Thus, homogenization for (4.7) *holds.*

3 Periodic homogenization for Cauchy problems

Let us state right away the main result in this section, which was proved by Lions, Papanicolaou, Varadhan [102], and Evans [48].

Theorem 4.6. *Assume that H satisfies* (4.2) *and* (4.3). *Assume $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$. For each $\varepsilon > 0$, let u^ε be the unique viscosity solution of*

$$\begin{cases} u^\varepsilon_t(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad\qquad\qquad u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.8}$$

Then, as $\varepsilon \to 0$, u^ε converges to u locally uniformly on $\mathbb{R}^n \times [0,\infty)$, and u solves the effective equation

$$\begin{cases} u_t + \overline{H}(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.9}$$

We here introduce the perturbed test function method of Evans [48] to prove the above theorem. Roughly speaking, the perturbed test function method is a way to make the formal ansatz rigorous. One needs to be extremely careful here as if we recall, for $p \in \mathbb{R}^n$, the corresponding cell problem is

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n. \tag{4.10}$$

To make it clear the dependences, sometimes, we write $v = v(y, p)$, and in fact, v depends on p in a very nonlinear way. It is worth mentioning here that $\overline{H} : \mathbb{R}^n \to \mathbb{R}$ continuous and coercive. To focus on the homogenization results, we postpone the proof of this fact until the next section.

The ansatz we found was that for each $p = Du(x,t)$, $v(y,p) = v(y, Du(x,t))$ is a corresponding corrector, and our asymptotic expansion around $(x,t) \in \mathbb{R}^n \times (0,\infty)$ looks like

$$u^\varepsilon(x,t) \approx u(x,t) + \varepsilon v(y,p) = u(x,t) + \varepsilon v\left(\frac{x}{\varepsilon}, Du(x,t)\right).$$

The last term in the above is quite problematic because of two issues. First, u is often only Lipschitz, and not C^1, which means that $Du(x,t)$ is only defined a.e., and there is no continuity property with respect to (x,t). Second, we do not know well the dependence $p \mapsto v(y,p)$. Of course, these two issues come from the highly nonlinear feature of our PDE, and they need to be handled appropriately.

3.1 A heuristic proof

We first give a heuristic proof of the homogenization result by the perturbed test function method of Evans. As one will see, the first difficulty is handled by kicking the gradient Du to the test functions as often seen in the theory of viscosity solutions. The proof is not yet rigorous as we assume that solutions to (4.10) are smooth. We will also see why the perturbed test function is needed.

A heuristic proof of Theorem 4.6. As usual, we break this heuristic proof into few steps.

1. We first obtain some a priori estimates for u^ε. By Theorem 1.34, we have the existence of $C > 0$ independent of $\varepsilon > 0$ such that

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

   By the Arzelà–Ascoli theorem, there exists a subsequence $\{\varepsilon_j\} \to 0$ such that $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$.

2. We now prove that u solves the effective equation (4.9).

   First, we perform the subsolution test. If $\varphi \in C^1\big(\mathbb{R}^n \times (0,\infty)\big)$ is such that $u - \varphi$ has strict max at (x_0, t_0), then we plan to show that $\varphi_t(x_0,t_0) + \overline{H}(D\varphi(x_0,t_0)) \le 0$.

   It is natural to try first the usual approach. As $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$, we may assume that $u^{\varepsilon_j} - \varphi$ has max at (x_j, t_j) and $(x_j,t_j) \to (x_0,t_0)$ as $j \to \infty$. The viscosity subsolution test gives

$$\varphi_t(x_j,t_j) + H\left(\frac{x_j}{\varepsilon_j}, D\varphi(x_j,t_j)\right) \le 0.$$

As $j \to \infty$ we have $\varphi_t(x_j, t_j) \to \varphi_t(x_0, t_0)$, but we do not have information about the second term $H\left(\frac{x_j}{\varepsilon_j}, D\varphi(x_j, t_j)\right)$ since φ does not oscillate around (x_0, t_0).

In order to capture the oscillating behavior, we use Evans's perturbed test function method. Let us denote $p = D\varphi(x_0, t_0)$, and consider

$$\psi^{\varepsilon}(x,t) = \varphi(x,t) + \varepsilon v\left(\frac{x}{\varepsilon}, p\right)$$

where $v \in \mathrm{Lip}(\mathbb{T}^n)$ is the viscosity solution of the cell problem (4.10) with this particular p. We assume here that v is smooth enough so that $\psi \in C^1$. Note that ψ^{ε} is just a perturbation of φ, hence the name "perturbed test function method". We may assume that $u^{\varepsilon_j} - \psi^{\varepsilon_j}$ has a local max at $(x_{\varepsilon_j}, t_{\varepsilon_j})$, and $(x_{\varepsilon_j}, t_{\varepsilon_j}) \to (x_0, t_0)$ as $j \to \infty$. By the viscosity subsolution test,

$$\psi_t^{\varepsilon_j}\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) + H\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}, D\varphi\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) + Dv\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}\right)\right) \le 0. \tag{4.11}$$

As $D\varphi\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) \to p$ as $j \to \infty$,

$$\lim_{j\to\infty}\left(H\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}, D\varphi\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) + Dv\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}\right)\right) - H\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}, p + Dv\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}\right)\right)\right) = 0,$$

which means

$$\lim_{j\to\infty}\left(H\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}, D\varphi\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) + Dv\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}\right)\right) - \overline{H}(p)\right) = 0.$$

Combine this with (4.11) to conclude. The viscosity supersolution test follows in a similar way.

3. As $\overline{H}$ is continuous and coercive, (4.9) has a unique Lipschitz solution u. Therefore, we conclude that $u^{\varepsilon} \to u$ locally uniformly in $\mathbb{R}^n \times [0, \infty)$ as $\varepsilon \to 0$.

□

Remark 4.7. In the above heuristic proof, Steps 1 and 3 are actually rigorous. The only heuristic part is Step 2, in which we assume that $y \mapsto v(y,p)$ for $p = D\varphi(x_0, t_0)$ is C^1. This is of course not realistic, and we need to fix it in our rigorous proof. Our goal of giving this heuristic proof is to show clearly the key point of the perturbed test function method without clouded technicalities.

The convergence of $u^{\varepsilon} \to u$ as $\varepsilon \to 0$ for full sequence is based on the fact that the limiting equation (4.9) has a unique Lipschitz solution u. This is essentially a compactness step, and it does not give a quantitative result on how fast u^{ε} converges to u. We will revisit this point later.

3.2 A rigorous proof by using Evans's perturbed test function method

Let us now give a rigorous proof of the homogenization for the Cauchy problem.

Proof of Theorem 4.6. We reuse Steps 1 and 3 in the heuristic proof above. There exists a subsequence $\{\varepsilon_j\} \to 0$ such that $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0, \infty)$. In fact, by abuse of notions, we assume $u^\varepsilon \to u$ locally uniformly on $\mathbb{R}^n \times [0, \infty)$ as $\varepsilon \to 0$. All we need to do is to prove that u solves the effective equation (4.9).

We will perform only the subsolution test since the argument for supersolution test is similar. For $\phi \in C^1(\mathbb{R}^n \times (0, \infty))$ such that $u - \phi$ has a global strict max at $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$, we aim at proving

$$\phi_t(x_0, t_0) + \overline{H}(D\phi(x_0, t_0)) \leq 0.$$

Let $p = D\phi(x_0, t_0) \in \mathbb{R}^n$, and let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be the viscosity solution of (4.10) with this particular p. Let us assume further that $u(x_0, t_0) = \phi(x_0, t_0)$, and for some $r \in (0, t_0/2)$,

$$u(x,t) - \phi(x,t) < -(\|v\|_{L^\infty(\mathbb{T}^n)} + 1) \qquad \text{for all } (x,t) \notin B(x_0, r) \times [t_0 - r, t_0 + r].$$

In order to overcome the lack of smoothness of v, we use the doubling variables method. We divide the proof into several steps.

1. Fix $T > 2t_0$. For each $\varepsilon, \eta > 0$ we consider the auxiliary function

$$\Phi^{\eta,\varepsilon}(x,y,t) : \mathbb{R}^n \times \mathbb{R}^n \times [0,T] \to \mathbb{R}$$
$$(x,y,t) \mapsto u^\varepsilon(x,t) - \left(\phi(x,t) + \varepsilon v(y) + \frac{\left|y - \frac{x}{\varepsilon}\right|^2}{\eta} \right).$$

   For $\varepsilon > 0$ sufficiently small, it is clear that $\Phi^{\eta,\varepsilon}$ has a max at $\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right) \in B(x_0, r) \times \mathbb{R}^n \times [t_0 - r, t_0 + r]$. As $\eta \to 0$, by compactness $\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) \to (x_\varepsilon, t_\varepsilon)$ up to a subsequence. We claim that $y_{\eta\varepsilon} \to \frac{x_\varepsilon}{\varepsilon}$ as $\eta \to 0$. Since $\Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, \frac{x_{\eta\varepsilon}}{\varepsilon}, t_{\eta\varepsilon}\right) \leq \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$ for all $\eta > 0$, we obtain

$$\frac{1}{\eta}\left|y_{\eta\varepsilon} - \frac{x_{\eta\varepsilon}}{\varepsilon}\right|^2 \leq 2\varepsilon \|v\|_{L^\infty(\mathbb{T}^n)} \qquad \Longrightarrow \qquad \lim_{\eta \to 0} y_{\eta\varepsilon} = \frac{x_\varepsilon}{\varepsilon}. \tag{4.12}$$

2. As $(x,t) \mapsto \Phi^{\eta,\varepsilon}\left(x, y_{\eta\varepsilon}, t\right)$ has max at $\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$, we imply that $u^\varepsilon - \phi - \frac{1}{\eta}\left|y_{\eta\varepsilon} - \frac{x}{\varepsilon}\right|^2$ has max at $\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$. The subsolution test of (4.8) gives

$$\phi_t\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + H\left(\frac{x_{\eta\varepsilon}}{\varepsilon}, D\phi\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right) \leq 0. \tag{4.13}$$

3. Next, $y \mapsto \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y, t_{\eta\varepsilon}\right)$ has max at $y_{\eta\varepsilon}$, thus $v(y) - \frac{-1}{\eta\varepsilon}\left|y - \frac{x_{\eta\varepsilon}}{\varepsilon}\right|^2$ has min at $y_{\eta\varepsilon}$, and hence, the supersolution test of the cell problem gives us

$$-\overline{H}(p) + H\left(y_{\eta\varepsilon}, p + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right) \geq 0. \tag{4.14}$$

   Besides, as v is Lipschitz, we get

$$\left|\frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right| \leq C, \tag{4.15}$$

for some $C > 0$ independent of η, ε. By compactness, we can assume (up to passing to a subsequence again) that

$$\lim_{\eta\to 0} \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right) = p_\varepsilon \in \mathbb{R}^n. \tag{4.16}$$

4. Note that $\Phi^{\eta,\varepsilon}\left(x, \frac{x}{\varepsilon}, t\right) \le \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$. Let $\eta \to 0$ in this relation and use (4.16) to yield

$$u^\varepsilon(x,t) - \varepsilon v\left(\frac{x}{\varepsilon}\right) - \phi(x,t) \le u^\varepsilon(x_\varepsilon, t_\varepsilon) - \varepsilon v\left(\frac{x_\varepsilon}{\varepsilon}\right) - \phi(x_\varepsilon, t_\varepsilon)$$

   for all $(x,t) \in \mathbb{R}^n \times [0,T]$. That means $(x,t) \mapsto u^\varepsilon(x,t) - \varepsilon v\left(\frac{x}{\varepsilon}\right) - \phi(x,t)$ has max at $(x_\varepsilon, t_\varepsilon)$. Again, by passing to a subsequence if needed, $(x_\varepsilon, t_\varepsilon) \to (x_0, t_0)$ as $\varepsilon \to 0$.

5. Let $\eta \to 0$ in (4.13) and (4.14) to get

$$\phi_t(x_\varepsilon, t_\varepsilon) + H\left(\frac{x_\varepsilon}{\varepsilon}, D\phi(x_\varepsilon, t_\varepsilon) + p_\varepsilon\right) \le 0,$$

   and

$$-\overline{H}(p) + H\left(\frac{x_\varepsilon}{\varepsilon}, p + p_\varepsilon\right) \ge 0.$$

   Combine the above two and let $\varepsilon \to 0$ to conclude that

$$\phi_t(x_0, t_0) + \overline{H}(p) \le 0.$$

□

4 Some first properties of the effective Hamiltonian

4.1 Simple qualitative properties of $\overline{H}$

We start with some preliminary properties of $\overline{H}$.

Theorem 4.8. *Assume $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (4.2) *and* (4.3). *Then $\overline{H} : \mathbb{R}^n \to \mathbb{R}$ is also continuous and coercive.*
Furthermore, if $p \mapsto H(y,p)$ is Lipschitz for all $y \in \mathbb{T}^n$ with Lipschitz constant at most $C > 0$, then $p \mapsto \overline{H}(p)$ is also Lipschitz.

Proof. We present here the proof using the discounted approximation of the cell problem, and the cell problem.

(a) We first show that $\overline{H}$ is coercive, which is rather simple. Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (4.10), that is,

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n.$$

   Observe that since $v \in C(\mathbb{T}^n)$, it has maximum at some point $x_0 \in \mathbb{T}^n$ and that this point, we must have $0 \in D^+v(x_0)$, thus the subsolution test at x_0 shows

$$\min_{\mathbb{T}^n} H(y,p) \le H(x_0, p) \le \overline{H}(p),$$

which implies

$$\lim_{|p|\to\infty} \overline{H}(p) = \lim_{|p|\to\infty} \left(\min_{y\in\mathbb{T}^n} H(y,p) \right) = +\infty.$$

Actually, it is useful to know that

$$\min_{y\in\mathbb{T}^n} H(y,p) \le \overline{H}(p) \le \max_{y\in\mathbb{T}^n} H(y,p) \qquad \text{for all } p \in \mathbb{R}^n. \tag{4.17}$$

(b) We now show that $\overline{H}$ is continuous. Pick an arbitrary sequence $\{p_k\} \subset \mathbb{R}^n$ such that $\{p_k\} \to p$ and $\{\overline{H}(p_k)\} \to c \in \mathbb{R}$. We just need to show that $\overline{H}(p) = c$. Let $v_k \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (4.10) with $\min_{\mathbb{T}^n} v_k = 0$ and $p = p_k$ for all $k \in \mathbb{N}$. Note first that, in light of (4.17), we are able to find $C > 0$ such that, for all $k \in \mathbb{N}$,

$$H(y, p_k + Dv_k(y)) = \overline{H}(p_k) \le \max_{y\in\mathbb{T}^n} H(y,p_k) \le C \qquad \text{in } \mathbb{T}^n.$$

Hence, coercivity of H yields the existence of $C_1 > 0$ such that

$$\|Dv_k\|_{L^\infty(\mathbb{T}^n)} \le C_1.$$

By the Arzelà–Ascoli theorem, by passing to a subsequence if necessary, we get that $v_k \to v$ uniformly in $\mathbb{T}^n$ for some $v \in \mathrm{Lip}(\mathbb{T}^n)$. The usual stability results imply that v is a solution to

$$H(y, p + Dv(y)) = c \qquad \text{in } \mathbb{T}^n,$$

which means that $\overline{H}(p) = c$.

(c) We now assume $p \mapsto H(y,p)$ is Lipschitz for all $y \in \mathbb{T}^n$ with Lipschitz constant at most $C > 0$. Fix $p, q \in \mathbb{R}^n$. For each $\lambda > 0$, let $u^\lambda, v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ be the solutions to

$$\lambda u^\lambda + H(y, q + Du^\lambda) = 0 \qquad \text{in } \mathbb{T}^n, \tag{4.18}$$

and

$$\lambda v^\lambda + H(y, p + Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n, \tag{4.19}$$

respectively. We now use the comparison principle to obtain needed estimates. It is not hard to see that $u^\lambda + \frac{C|p-q|}{\lambda}$ is a supersolution, and $u^\lambda - \frac{C|p-q|}{\lambda}$ is a subsolution to (4.19). Therefore,

$$u^\lambda - \frac{C|p-q|}{\lambda} \le v^\lambda \le u^\lambda + \frac{C|p-q|}{\lambda}.$$

Multiply the above by λ and let $\lambda \to 0$ to deduce

$$\overline{H}(q) - C|p-q| \le \overline{H}(p) \le \overline{H}(q) + C|p-q|.$$

□

In fact, from part (c) in the above proof, we have the following immediate corollary.

Corollary 4.9. *Assume $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is locally Lipschitz uniformly in $y \in \mathbb{T}^n$. Then $p \mapsto \overline{H}(p)$ is also locally Lipschitz.*

We now introduce some elementary representation formulas for $\overline{H}$.

Theorem 4.10. *Assume $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (4.2) *and* (4.3)*. Then, for $p \in \mathbb{R}^n$,*

$$\begin{aligned}\overline{H}(p) = &\ \inf\{c \in \mathbb{R} : \exists\, v \in C(\mathbb{T}^n) : H(y, p + Dv(y)) \le c \text{ in } \mathbb{T}^n \text{ in viscosity sense}\} \\ &= \sup\{c \in \mathbb{R} : \exists\, v \in C(\mathbb{T}^n) : H(y, p + Dv(y)) \ge c \text{ in } \mathbb{T}^n \text{ in viscosity sense}\}.\end{aligned}$$

Proof. Let us define

$$\begin{aligned}\mathcal{A} &= \{c \in \mathbb{R} : \exists\, v \in C(\mathbb{T}^n) : H(y, p + Dv(y)) \le c \text{ in } \mathbb{T}^n \text{ in viscosity sense}\} \\ \mathcal{B} &= \{c \in \mathbb{R} : \exists\, v \in C(\mathbb{T}^n) : H(y, p + Dv(y)) \ge c \text{ in } \mathbb{T}^n \text{ in viscosity sense}\}.\end{aligned}$$

Recall that from the cell problem there exists $v \in \mathrm{Lip}(\mathbb{T}^n)$ solves (4.10), thus,

$$\inf \mathcal{A} \le \overline{H}(p) \le \sup \mathcal{B}.$$

Next, we show that $\inf \mathcal{A} = \overline{H}(p)$. The other part follows in a similar way. Assume by contradiction that $\inf \mathcal{A} < \overline{H}(p)$. Then, there exist some $c_1 \in \mathcal{A}$ and $v_1 \in C(\mathbb{T}^n)$ such that $\inf \mathcal{A} < c_1 < \overline{H}(p)$, while $H(y, p + Dv_1(y)) \le c_1$ in $\mathbb{T}^n$ in the viscosity sense. Since v, v_1 are bounded, there exists $\delta > 0$ so that

$$\delta v_1 + H(y, p + Dv_1(y)) < \frac{c_1 + \overline{H}(p)}{2} < \delta v + H(y, p + Dv(y)) \qquad \text{in } \mathbb{T}^n.$$

The usual comparison principle implies $v_1 \le v$. By same steps, we obtain that $v_1 \le v - C$ for any constant $C > 0$, which is absurd. Therefore, $\inf \mathcal{A} = \overline{H}(p)$. □

We can see that Theorems 4.8, 4.10, and Corollary 4.9 give us some good qualitative properties of the effective Hamiltonian $\overline{H}$. Most of these were already covered by Lions, Papanicolaou, Varadhan [102]. Thus, theoretically, we can claim that homogenization holds, and we have certain understandings about $\overline{H}$. In other words, well-posedness of periodic homogenization of Hamilton–Jacobi equations is done.

Yet, for further understandings in both theoretical and numerical viewpoints, if we would like to know more about $\overline{H}$ such as its shape, its formula, its differentiability, the above results do not give us any hint. In fact, not so much is known about $\overline{H}$ if we are given a general H which satisfies (4.2) and (4.3). It is therefore extremely important to go beyond the well-posedness theory to understand better about $\overline{H}$, about the limiting solution u, and about the rate of convergence of u^ε to u.

So far, computing $\overline{H}$ numerically is extremely challenging. The cell problem (4.10) for each $p \in \mathbb{R}^n$ is already highly nonlinear, and it takes much time to compute a single $\overline{H}(p)$. It seems that there is not yet a way to relate $\overline{H}(p)$ with $\overline{H}(q)$ for $p \ne q$ through the cell problems. And hence, to get a good approximation of $\overline{H}$, one needs to compute $\overline{H}(p)$ at many different values of p, each of which is already costly, and uses interpolation to get such approximation.

4.2 Large time average and $\overline{H}$

We give in the following a large time average result, which is often used to compute $\overline{H}(p)$ for each fixed $p \in \mathbb{R}^n$. Although it is very simple, up to now, it seems to be the most effective one to compute $\overline{H}$ in the general (possibly nonconvex) setting.

Theorem 4.11. *Assume that H satisfies* (4.2) *and* (4.3). *Fix $p \in \mathbb{R}^n$. Consider the following Cauchy problem*

$$\begin{cases} w_t + H(y, p + Dw) &= 0 \qquad \text{in } \mathbb{T}^n \times (0, \infty), \\ w(y, 0) &= 0 \qquad \text{on } \mathbb{T}^n. \end{cases} \tag{4.20}$$

Let $w(y,t)$ be the unique viscosity solution to (4.20). *Then,*

$$\lim_{t\to\infty} \frac{w(y,t)}{t} = -\overline{H}(p) \qquad \text{uniformly for } y \in \mathbb{T}^n.$$

First proof. We give the first proof by using the cell problem (4.10). We simply construct a separable subsolution and supersolution to (4.20), respectively, and use them to bound the actual solution $w(x,t)$.

Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to (4.10). Define:

$$\varphi(x,t) = v(x) - \overline{H}(p)t \qquad \text{for } (x,t) \in \mathbb{T}^n \times [0, \infty).$$

It is clear that φ is a separable solution to (4.20) with initial data $\varphi(\cdot, 0) = v$. Let $C = \|v\|_{L^\infty(\mathbb{T}^n)}$. Then $\varphi(x,t) - C$ and $\varphi(x,t) + C$ is a viscosity subsolution and supersolution to (4.20), respectively. By the comparison principle,

$$v(x) - \overline{H}(p)t - C \le w(x,t) \le v(x) - \overline{H}(p)t + C \qquad \text{for } (x,t) \in \mathbb{T}^n \times (0, \infty).$$

Therefore,

$$\frac{v(x) - C}{t} - \overline{H}(p) \le \frac{w(x,t)}{t} \le \frac{v(x) + C}{t} - \overline{H}(p),$$

which gives us the desired result. Moreover, the rate of convergence is $O(\frac{1}{t})$, which is quite good. $\square$

As seen many times throughout this chapter, one key point to grasp is that homogenization is equivalent to large time average. In the proof above, we utilize strongly the cell problem. A natural question to ask is what happens in case one does not have such cell problems. We present next a second proof, which does not need to use the cell problems. This is based on the ideas in Giga, Mitake, Ohtsuka, and Tran [70], which utilize subadditivity instead.

Second proof. In this second proof, we will show that there exists $c \in \mathbb{R}$ such that

$$\lim_{t\to\infty} \frac{w(y,t)}{t} = c \qquad \text{uniformly for } y \in \mathbb{T}^n.$$

It is clear that w is Lipschitz on $\mathbb{T}^n \times [0, \infty)$ with a Lipschitz constant at most $C > 0$. Denote by $M(t) = \max_{y \in \mathbb{T}^n} w(y,t)$ for each $t \ge 0$. Then, $|M(t)| \le Ct$. We claim that $M(\cdot)$ is subadditive, that is,

$$M(t) + M(s) \ge M(t+s) \qquad \text{for all } s, t \ge 0. \tag{4.21}$$

Indeed, fix $s \ge 0$. Set $\phi(y,t) = w(y, t+s) - M(s)$ for all $(x,t) \in \mathbb{T}^n \times [0, \infty)$. Then, ϕ solves (4.20) with initial data $\phi(y, 0) = w(y,s) - M(s) \le 0$ for $x \in \mathbb{T}^n$. We use the comparison principle to get that $\phi \le w$. In particular,

$$M(t+s) - M(s) = \max_{y \in \mathbb{T}^n} \phi(y,t) \le \max_{y \in \mathbb{T}^n} w(y,t) = M(t).$$

Thus, (4.21) holds. By Fekete's lemma, there exists $c \in \mathbb{R}$ such that

$$\lim_{t\to\infty} \frac{M(t)}{t} = \inf_{t>0} \frac{M(t)}{t} = c.$$

Finally, we use the Lipschitz regularity of w and the above to conclude. □

This second proof to get large time average result is quite general, and is applicable to a lot of different settings.

4.3 Problems

Exercise 39. *Prove Corollary 4.9.*

Exercise 40. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that there exists $k > 0$ such that H is positively k-homogeneous in p, that is, $H(y, sp) = s^k H(y, p)$ for all $(y, p) \in \mathbb{T}^n \times \mathbb{R}^n$, and $s \geq 0$. Show that $\overline{H}$ is positively k-homogeneous as well.*

5 Further properties of the effective Hamiltonian in the convex setting

In this section, we always assume that $p \mapsto H(y, p)$ is convex for every $y \in \mathbb{T}^n$.

5.1 The inf-sup formula

Theorem 4.12 (The inf-sup formula)**.** *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y, p)$ is convex for every $y \in \mathbb{T}^n$. Then, for fixed $p \in \mathbb{R}^n$, we have*

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} H\left(y, p + D\phi(y)\right). \tag{4.22}$$

Proof. Pick any $\varphi \in C^1(\mathbb{T}^n)$, by the representation formula in Theorem 4.10,

$$\overline{H}(p) \leq \max_{y \in \mathbb{T}^n} H(y, p + D\varphi(y)),$$

and hence,

$$\overline{H}(p) \leq \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} H\left(y, p + D\phi(y)\right).$$

Conversely, given $\theta > 0$, we aim at proving that

$$\overline{H}(p) + \theta \geq \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} H\left(y, p + D\phi(y)\right).$$

Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to (4.10), that is,

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n.$$

It is clear that v is differentiable and solves the above a.e. in $\mathbb{T}^n$. We need to smooth v up, and we use the convolution trick as earlier. Take η to be the standard mollifier, that is,

$$\eta \in C_c^\infty(\mathbb{R}^n, [0, \infty)), \qquad \mathrm{supp}(\eta) \subset B(0, 1), \qquad \int_{\mathbb{R}^n} \eta(x)\, dx = 1.$$

For $\varepsilon > 0$, denote by $\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(\frac{x}{\varepsilon}\right)$ for all $x \in \mathbb{R}^n$. Set

$$v^\varepsilon(x) = (\eta_\varepsilon \star v)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)v(y)\,dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x-y)v(y)\,dy \qquad \text{for } x \in \mathbb{R}^n.$$

Then $v^\varepsilon \in C^\infty(\mathbb{T}^n)$, and $v^\varepsilon \to v$ uniformly in $\mathbb{T}^n$ as $\varepsilon \to 0$. We compute, for every fixed $x \in \mathbb{T}^n$,

$$\begin{aligned}
\overline{H}(p) &= \int_{\mathbb{R}^n} H\Big(x-y, p+Dv(x-y)\Big)\eta_\varepsilon(y)\,dy = \int_{B(0,\varepsilon)} H\Big(x-y, p+Dv(x-y)\Big)\eta_\varepsilon(y)\,dy \\
&\geq \int_{B(0,\varepsilon)} \Big(H\Big(x, p+Dv(x-y)\Big) - \omega(\varepsilon)\Big)\eta_\varepsilon(y)\,dy \\
&= \int_{B(0,\varepsilon)} H\Big(x, p+Dv(x-y)\Big)\eta_\varepsilon(y)\,dy - \omega(\varepsilon) \\
&\geq H\left(x, \int_{B(0,\varepsilon)} (p+Dv(x-y))\eta_\varepsilon(y)\,dy\right) - \omega(\varepsilon) = H(x, p+Dv^\varepsilon(x)) - \omega(\varepsilon).
\end{aligned}$$

Thus, v^ε satisfies

$$\max_{x\in\mathbb{T}^n} H(x, p+Dv^\varepsilon(x)) \leq \overline{H}(p) + \omega(\varepsilon).$$

Pick $\varepsilon > 0$ sufficiently small so that $\omega(\varepsilon) < \theta$ to conclude.

□

The following theorem is an immediate consequence of the inf-sup (or inf-max) formula.

Theorem 4.13. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{T}^n$. Then, $\overline{H}$ is convex.*

Proof. Fix $p, q \in \mathbb{R}^n$. We need to show

$$\overline{H}\left(\frac{p+q}{2}\right) \leq \frac{1}{2}\Big(\overline{H}(p) + \overline{H}(q)\Big).$$

For $\varphi, \psi \in C^1(\mathbb{T}^n)$, the convexity of $p \mapsto H(x,p)$ implies that, for $x \in \mathbb{T}^n$,

$$H\left(x, \frac{p+q}{2} + D\left(\frac{\varphi+\psi}{2}\right)(x)\right) \leq \frac{1}{2}\left(H\left(x, p+D\varphi(x)\right) + H\left(x, q+D\psi(x)\right)\right),$$

and so

$$\max_{x\in\mathbb{T}^n} H\left(x, \frac{p+q}{2} + D\left(\frac{\varphi+\psi}{2}\right)(x)\right) \leq \frac{1}{2}\Big(\max_{x\in\mathbb{T}^n} H(x, p+D\varphi(x)) + \max_{x\in\mathbb{T}^n} H(x, q+D\psi(x))\Big).$$

The inf-sup formula (4.12) implies that $\overline{H}\left(\frac{p+q}{2}\right) \leq \frac{1}{2}\Big(\overline{H}(p) + \overline{H}(q)\Big)$, and the proof is complete. □

It is worth pointing out that by using the idea of Barron, Jensen [18] in Theorem 2.27, we have another formula for $\overline{H}$ in the convex setting.

Corollary 4.14. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{T}^n$. Then, for each $p \in \mathbb{R}^n$,*

$$\overline{H}(p) = \inf\{c \in \mathbb{R} \,:\, \exists\, v \in \mathrm{Lip}(\mathbb{T}^n) : H(y, p + Dv(y)) \le c \text{ a.e. in } \mathbb{T}^n\}. \tag{4.23}$$

One can then use this Corollary to give another quick proof of Theorem 4.13. This proof is left as an exercise.

5.2 The large time average formula

We use Theorem 4.11 to give a large time average formula in the convex setting as following. This result was obtained first by Concordel [33].

Theorem 4.15. *Assume that H satisfies* (4.2)*, $p \mapsto H(y,p)$ is convex and superlinear for each $y \in \mathbb{T}^n$. Fix $p \in \mathbb{R}^n$. Then,*

$$\overline{H}(p) = \lim_{t\to\infty} \sup_{\gamma(\cdot)} \frac{1}{t} \int_0^t \big(p \cdot \gamma'(s) - L(\gamma(s), \gamma'(s))\big)\, ds.$$

Proof. We just need to apply the result of Theorem 4.11 here. Let w be the solution to (4.20), then we have that

$$\lim_{t\to\infty} \frac{w(y,t)}{t} = -\overline{H}(p) \qquad \text{uniformly for } y \in \mathbb{T}^n.$$

The Lagrangian corresponding to $H(\cdot, p + \cdot)$ is $(x,v) \mapsto L(x,v) - p \cdot v$. We apply the optimal control formula for Cauchy problem to (4.20) to get that, for $(y,t) \in \mathbb{T}^n \times (0,\infty)$,

$$w(y,t) = \inf_{\gamma(t)=y} \int_0^t \big(L(\gamma(s), \gamma'(s)) - p \cdot \gamma'(s)\big)\, ds$$

Combine the two identities above to complete the proof. □

Concordel [33, 34] used this formula to study properties of $\overline{H}$, especially whether $\overline{H}$ has a flat part or not. We will address this in the next section.

5.3 An one dimensional example

We give in the following an one dimensional example that was introduced by Lions, Papanicolaou, Varadhan [102]. According to the paper, Tartar was the one who provided this example.

Example 4.2. *Assume that $n = 1$, and $H(y,p) = |p|^2 - V(y)$, where $V \in C(\mathbb{T})$ with $\min_{\mathbb{T}} V = 0$. We intend to give a formula for $\overline{H}$ here.*

For any 1-periodic integrable function ϕ, denote by $\langle \phi \rangle$ its average, that is, $\langle \phi \rangle = \int_0^1 \phi(y)\, dy$. We claim that

$$\overline{H}(p) = \begin{cases} 0 & \text{for } |p| \le \langle \sqrt{V} \rangle, \\ \lambda & \text{for } |p| \ge \langle \sqrt{V} \rangle, \text{ where } \lambda \ge 0 \text{ is such that } |p| = \langle \sqrt{\lambda + V} \rangle. \end{cases} \tag{4.24}$$

Note that this formula only holds in one dimension. There is no such formula in multi dimensions.

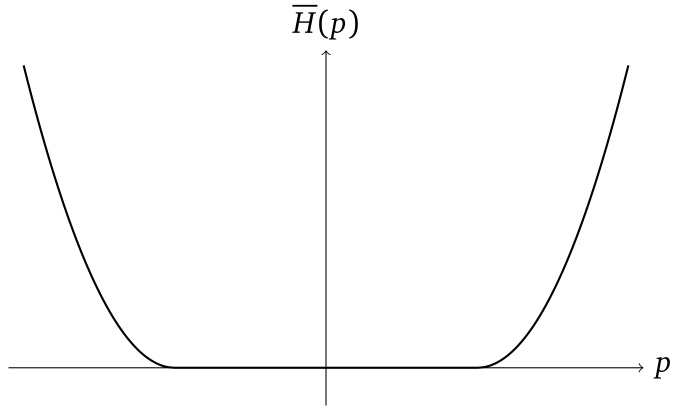


Figure 4.2: Graph of $\overline{H}$

Let us now prove the above formula.

Proof of formula (4.24). Pick $y_0 \in [0,1]$ such that $V(y_0)=0$. For $|p| \le \langle \sqrt{V} \rangle$, we can find $y_1 \in [y_0, y_0+1]$ such that

$$\int_{y_0}^{y_1} \left(-p+\sqrt{V(s)}\right) ds = \int_{y_1}^{y_0+1} \left(p+\sqrt{V(s)}\right) ds,$$

which means that

$$p = \int_{y_0}^{y_1} \sqrt{V(s)}\, ds - \int_{y_1}^{y_0+1} \sqrt{V(s)}\, ds.$$

Let $v : [y_0, y_0+1] \to \mathbb{R}$ be such that

$$v'(y) = \begin{cases} -p+\sqrt{V(y)} & \text{for } y_0 \le y < y_1, \\ -p-\sqrt{V(y)} & \text{for } y_1 < y \le y_0+1. \end{cases}$$

By the choice of y_1, $v(y_0) = v(y_0+1)$. Extend v to $\mathbb{R}$ in a periodic way. It is clear then that v is a viscosity solution to

$$|p+v'|^2 - V(y) = 0 \qquad \text{in } \mathbb{T}.$$

Indeed, $v \in C^1(\mathbb{T} \setminus \{y_1\})$ and solves the equation in the classical sense in $\mathbb{T} \setminus \{y_1\}$. At y_1, v has a corner from above, so there is nothing to check. Thus, $\overline{H}(p) = 0$ for $|p| \le \langle \sqrt{V} \rangle$.

Now, for $p > \langle \sqrt{V} \rangle$, we are able to find $\lambda > 0$ such that $p = \langle \sqrt{\lambda + V} \rangle$. Let $v : [y_0, y_0+1] \to \mathbb{R}$ be such that

$$v'(y) = -p + \sqrt{\lambda + V(y)} \qquad \text{for } y_0 \le y \le y_0+1.$$

By the choice of λ, $v(y_0) = v(y_0+1)$. Extend v to $\mathbb{R}$ in a periodic way. One can see that v is a classical solution to

$$|p+v'|^2 - V(y) = \lambda \qquad \text{in } \mathbb{T},$$

which yields that $\overline{H}(p) = \lambda$. □

It is interesting to see that if $V \neq 0$, then $\overline{H}$ is not uniformly convex, and $\{\overline{H} = 0\}$ is a symmetric line segment around 0. We will address this point more systematically in the section about flat parts of $\overline{H}$.

5.4 Problems

Exercise 41. *Use Corollary 4.14 to give another quick proof of Theorem 4.13.*

Exercise 42. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is level-set quasiconvex for every $y \in \mathbb{T}^n$. Show that the inf-sup formula still holds, that is, for $p \in \mathbb{R}^n$,*

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} H(y, p + D\phi(y)).$$

Exercise 43. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is level-set quasiconvex for every $y \in \mathbb{T}^n$. Show that $\overline{H}$ is level-set quasiconvex.*

5.5 Qualitative properties of $\overline{H}$ in the convex setting

We first show that evenness is preserved.

Theorem 4.16. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is convex and even for every $y \in \mathbb{T}^n$. Then, $p \mapsto \overline{H}(p)$ is also convex and even.*

Proof. Of course, we only need to show that $\overline{H}$ is even. Using the inf-sup formula, we have

$$\begin{aligned}\overline{H}(p) &= \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} H(y, p + D\phi(y)) \\ &= \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} H(y, -p + D(-\phi)(y)) = \overline{H}(-p).\end{aligned}$$

□

Since the inf-max formula still holds for the level-set quasiconvex case, we have the following corollary, which is quite useful.

Corollary 4.17. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is level-set quasiconvex and even for every $y \in \mathbb{T}^n$. Then, $p \mapsto \overline{H}(p)$ is also level-set quasiconvex and even.*

Remark 4.18. It is important noting that evenness is not preserved in the nonconvex setting. We will address this point later.

5.6 Flat parts of $\overline{H}$

We come back to the classical mechanics Hamiltonian

$$H(y,p) = \frac{1}{2}|p|^2 - V(y) \qquad \text{for } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n.$$

Here $V \in C(\mathbb{T}^n)$ is a given potential energy. Of course, the corresponding effective Hamiltonian $\overline{H}$ is convex, but we want to know more about its behavior in this section.

Lemma 4.19. *Assume that $H(y,p) = \frac{1}{2}|p|^2 - V(y)$ for $(y,p) \in \mathbb{T}^n \times \mathbb{R}^n$, where $V \in C(\mathbb{T}^n)$ with* $\min_{\mathbb{T}^n} V = 0$. *Then* $\min_{\mathbb{R}^n} \overline{H} = 0$.

Proof. Note first that, for $p = 0$ and $\phi \equiv 0$, the inf-sup formula gives

$$\overline{H}(0) = \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \left(\frac{1}{2}|D\phi(y)|^2 - V(y) \right) \le \max_{y \in \mathbb{T}^n}(-V(y)) \le 0.$$

On the other hand, for each $p \in \mathbb{R}^n$, let v be a Lipschitz solution to (4.10), that is,

$$\frac{1}{2}|p + Dv|^2 - V = \overline{H}(p) \qquad \text{in } \mathbb{T}^n.$$

Surely, v solves the above a.e. in $\mathbb{T}^n$. Pick y_0 such that $V(y_0) = 0$. Then, we are able to find a sequence $\{y_k\} \to y_0$ such that v is differentiable at y_k for $k \in \mathbb{N}$, and classically,

$$\frac{1}{2}|p + Dv(y_k)|^2 - V(y_k) = \overline{H}(p).$$

Therefore,

$$\overline{H}(p) \ge \lim_{k \to \infty} (-V(y_k)) = 0.$$

We obtain that $\overline{H}(0) = \min_{\mathbb{R}^n} \overline{H} = 0$. □

Let us give a clear definition for flat parts of $\overline{H}$ before we move on.

Definition 4.20. *Assume that H satisfies* (4.2) *and* (4.3). *Assume further that $p \mapsto H(y,p)$ is convex. If the set $\left\{p \in \mathbb{R}^n : \overline{H}(p) = \min_{\mathbb{R}^n} \overline{H}\right\}$ has nonempty interior, we say that $\overline{H}$ has a flat part at its minimum value.*

We now show that, in many situations, $\overline{H}$ corresponding to the classical mechanics Hamiltonian has a flat part at its minimum value. This is quite surprising as although we start with a nice, uniformly convex Hamiltonian, the homogenization process gives back the effective Hamiltonian with a flat part at its minimum value, and of course, is not uniformly convex anymore. This tells us that there is a strong interplay between the kinetic and potential energies, and the potential energy V plays a crucial role in forming the shape of $\overline{H}$.

Let us state the first result along this line. By abuse of notions, we often identify $\mathbb{T}^n$ with the unit cell $Y = [0,1]^n$.

Theorem 4.21. *Assume that $H(y,p) = \frac{1}{2}|p|^2 - V(y)$ for $(y,p) \in \mathbb{T}^n \times \mathbb{R}^n$, where $V \in C(\mathbb{T}^n)$ with $\min_{\mathbb{T}^n} V = 0$. Assume further that $\{V = 0\} \subset\subset (0,1)^n$. Then, $\overline{H}$ has a flat part at its minimum value* 0*.*

This result was first proved by Concordel [34]. Of course, one can state it in a bit more general setting, but we choose to make it simple this way with the requirement that $\{V = 0\} \subset\subset (0,1)^n$. Geometrically, this means that $\{V = 0\}$ is isolated in each cell of unit size $k + [0,1]^n$ for $k \in \mathbb{Z}^n$, and this isolation is sort of a trapping effect. Here, we follow a different approach by using the inf-sup formula (or equivalently, constructions of smooth subsolutions). This was done by Mitake and Tran [115].

Proof. By Lemma 4.19, we already have

$$\overline{H}(0) = \min_{\mathbb{R}^n} \overline{H} = 0.$$

We identify $\mathbb{T}^n$ with the unit cell $Y = [0,1]^n$. Denote by $U_0 = \{V = 0\} \subset\subset (0,1)^n$. We are able to find two open sets U_1, U_2 such that

$$U_0 \subset\subset U_1 \subset\subset U_2 \subset\subset (0,1)^n.$$

Let $d = \min\{\operatorname{dist}(U_0, \partial U_1), \operatorname{dist}(U_1, \partial U_2)\} > 0$. By definition, we can find $\varepsilon_0 > 0$ such that

$$V(y) > \varepsilon_0 > 0 \qquad \text{for all } y \in Y \setminus U_1.$$

For $p \in \mathbb{R}^n$ to be chosen, we define a smooth function $\phi : Y \to \mathbb{R}$ such that

$$\begin{cases} \phi(y) = -p \cdot y & \text{for } y \in U_1, \\ \phi(y) = 0 & \text{for } y \in Y \setminus U_2, \\ |D\phi(y)| \le \dfrac{C|p|}{d} & \text{for } y \in Y. \end{cases}$$

We compute that

$$\frac{1}{2}|p + D\phi(y)|^2 - V(y) = \begin{cases} -V(y) \le 0 & \text{for } y \in U_1, \\ \le \dfrac{C|p|^2}{d^2} - \varepsilon_0 & \text{for } y \in Y \setminus U_1. \end{cases}$$

Hence, for $|p| \le r = \frac{d\sqrt{\varepsilon_0}}{C}$,

$$\frac{1}{2}|p + D\phi(y)|^2 - V(y) \le 0 \qquad \text{in } \mathbb{T}^n,$$

which means that $\overline{H}(p) \le 0$ correspondingly. We thus derive that $B(0,r) \subset \{\overline{H} = 0\}$. □

Remark 4.22. The proof of Concordel [34] is quite complicated, but geometrically intuitive. Let us describe the key points of her proof here. We use the same setting as in the above proof, and we assume further that, for any $k, j \in \mathbb{Z}^n$ with $k \ne j$,

$$\operatorname{dist}(k + U_1, j + U_1) \ge d.$$

See Figure 4.3. We show $\overline{H}(p) = 0$ for $|p| \le r = \frac{d\sqrt{\varepsilon_0}}{C}$. By Theorem 4.15, we have the formula

$$\overline{H}(p) = \lim_{t\to\infty} \sup_{\gamma(\cdot)} \frac{1}{t} \int_0^t \left(p \cdot \gamma'(s) - \frac{1}{2}|\gamma'(s)|^2 - V(\gamma(s)) \right) ds.$$

On one hand, we can pick $\gamma_1(s) = \gamma_1(0) \in U_0$ for all $s \ge 0$ to get that $\overline{H}(p) \ge 0$ always. On the other hand, we need to show that $\overline{H}(p) \le 0$ for $|p| \le r$ as well. The idea is to show that an optimal path γ to the above formula is trapped in one of the copies of $k + U_1$ for $k \in \mathbb{Z}^n$. Indeed, if γ travels outside of $k + U_1$ for $k \in \mathbb{Z}^n$, the action functional is quite negative there. More precisely, assume $\gamma([t_1, t_2]) \subset \mathbb{R}^n \setminus \bigcup_{k\in\mathbb{Z}^n}(k + U_1)$ for some $t_1 < t_2$, then

$$\begin{aligned} &\int_{t_1}^{t_2} \left(p \cdot \gamma'(s) - \frac{1}{2}|\gamma'(s)|^2 - V(\gamma(s)) \right) ds \le \int_{t_1}^{t_2} \left(p \cdot \gamma'(s) - \frac{1}{2}|\gamma'(s)|^2 - \varepsilon_0 \right) ds \\ \le &\int_{t_1}^{t_2} \left(-\frac{1}{2}|\gamma'(s) - p|^2 + \frac{1}{2}|p|^2 - \varepsilon_0 \right) ds \le -\frac{1}{2} \int_{t_1}^{t_2} \left(|\gamma'(s) - p|^2 + \varepsilon_0 \right) ds, \end{aligned}$$

which gives us the intuition why γ should not travel outside of $k + U_1$ for $k \in \mathbb{Z}^n$. Of course, one needs to be careful in the analysis here, but this is basically the heart of Concordel's arguments.

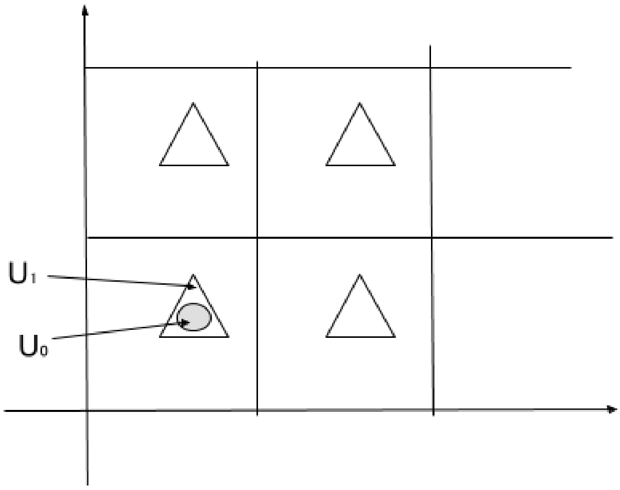


Figure 4.3: Periodic structures and $k+U_1$ for $k\in\mathbb{Z}^n$

The condition that we put in the above theorem is in fact optimal. If it does not hold, that is, $\{V=0\}$ is not trapped, then $\overline{H}$ might not have a flat part at its minimum value. Let us give now a simple example to demonstrate this.

Example 4.3. *Assume that $n=2$, $H(y,p)=\frac{1}{2}|p|^2-V(y)$ for $(y,p)\in\mathbb{T}^2\times\mathbb{R}^2$. Again, we identify $\mathbb{T}^2$ with $[0,1]^2$, and $\mathbb{T}$ with $[0,1]$. For $y=(y_1,y_2)\in\mathbb{T}^2$, the potential energy V satisfies that $V(y_1,y_2)=\tilde{V}(y_1)$, where $\min_{\mathbb{T}}\tilde{V}=0$ and $\{\tilde{V}=0\}=\left\{\frac{1}{2}\right\}$. Then,*

$$\{V=0\}=\left\{\frac{1}{2}\right\}\times[0,1],$$

which is not compactly supported in $(0,1)^2$. Let us now find the formula for $\overline{H}$. Let $\overline{K}$ be the effective Hamiltonian corresponding to $K(y_1,p_1)=\frac{1}{2}|p_1|^2-\tilde{V}(y_1)$ for all $(y_1,p_1)\in\mathbb{T}\times\mathbb{R}$. We know that $\min_{\mathbb{R}}\overline{K}=0=\overline{K}(0)$. Moreover, it is clear that

$$\overline{H}(p_1,p_2)=\overline{K}(p_1)+\frac{1}{2}|p_2|^2 \qquad \text{for all } (p_1,p_2)\in\mathbb{R}^2.$$

In this case, $\overline{H}$ does not have a flat part at its 0 level-set.

We now give a more general result, in which case $\overline{H}$ does not have a flat part at its 0 level-set. This is a result taken from Concordel [34].

Theorem 4.23. *Assume that $H(y,p)=\frac{1}{2}|p|^2-V(y)$ for $(y,p)\in\mathbb{T}^n\times\mathbb{R}^n$, where $V\in C(\mathbb{T}^n)$ with $\min_{\mathbb{T}^n}V=0$. Assume that there exist a C^1 curve $\xi:[0,\infty)\to\mathbb{R}^n$, a sequence $\{t_m\}\to\infty$, and a vector $p_0\neq 0$ such that*

$$\begin{cases} |\xi'(s)|=1 & \text{for all } s\geq 0,\\ V(\xi(s))=0 & \text{for all } s\geq 0,\\ \lim_{m\to\infty}\dfrac{\xi(t_m)}{t_m}=p_0\neq 0. \end{cases}$$

Then, $\overline{H}$ does not have a flat part at its 0 level-set.

As it is clear in the statement of this theorem, the curve ξ makes the set $\{V=0\}$ not being trapped in the unit cell, and one can use ξ to form the needed paths in the formula of $\overline{H}$.

Proof. Fix $\lambda>0$ and let $p=\lambda p_0$. We will show that $\overline{H}(p)>0$.

Let $\alpha=\lambda|p_0|^2>0$, and denote by $\gamma(s)=\xi(\alpha s)$ for all $s\geq 0$. Then, $|\gamma'(s)|=\alpha$, and

$$\frac{1}{t}\int_0^t\left(p\cdot\gamma'(s)-\frac{1}{2}|\gamma'(s)|^2-V(\gamma(s))\right)ds=p\cdot\frac{\gamma(t)-\gamma(0)}{t}-\frac{1}{2}\alpha^2.$$

At $\overline{t}_m=\frac{t_m}{\alpha}$ for $m\in\mathbb{N}$,

$$p\cdot\frac{\gamma(\overline{t}_m)-\gamma(0)}{\overline{t}_m}-\frac{1}{2}\alpha^2=p\cdot\frac{\xi(t_m)-\xi(0)}{\frac{t_m}{\alpha}}-\frac{1}{2}\alpha^2\longrightarrow\alpha\lambda|p_0|^2-\frac{1}{2}\alpha^2=\frac{1}{2}\lambda^2|p_0|^4,$$

as $m\to\infty$. Therefore, by Theorem 4.15,

$$\overline{H}(p)=\lim_{t\to\infty}\sup_{\gamma(\cdot)}\frac{1}{t}\int_0^t\left(p\cdot\gamma'(s)-\frac{1}{2}|\gamma'(s)|^2-V(\gamma(s))\right)ds\geq\frac{1}{2}\lambda^2|p_0|^4.$$

The proof is complete. □

Remark 4.24. Assume that $H(y,p)=\frac{1}{2}|p|^2-V(y)$ for $(y,p)\in\mathbb{T}^n\times\mathbb{R}^n$, where $V\in C(\mathbb{T}^n)$ with $\min_{\mathbb{T}^n}V=0$. We first note that the formula of $\overline{H}(p)$ can be rewritten as

$$\begin{aligned}\overline{H}(p)&=\lim_{t\to\infty}\sup_{\gamma(\cdot)}\frac{1}{t}\int_0^t\left(p\cdot\gamma'(s)-\frac{1}{2}|\gamma'(s)|^2-V(\gamma(s))\right)ds\\&=\frac{1}{2}|p|^2-\lim_{t\to\infty}\inf_{\gamma(\cdot)}\frac{1}{t}\int_0^t\left(\frac{1}{2}|\gamma'(s)-p|^2+V(\gamma(s))\right)ds.\end{aligned}$$

If we assume further that $V\in C^{1,1}(\mathbb{T}^n)$, then for each finite time $t>0$, an optimal path to the minimizing problem

$$\inf_{\gamma(\cdot)}\frac{1}{t}\int_0^t\left(\frac{1}{2}|\gamma'(s)-p|^2+V(\gamma(s))\right)ds$$

satisfies the Euler–Lagrange equation

$$-\frac{d}{ds}(\gamma'(s)-p)+DV(\gamma(s))=0\quad\Longrightarrow\quad\gamma''(s)=DV(\gamma(s)).$$

In particular, $s\mapsto\frac{1}{2}|\gamma'(s)|^2-V(\gamma(s))$ is constant, which gives the boundedness of the traveling speed $|\gamma'(s)|$ for $s\geq 0$. Then, we have the following refined formula for $\overline{H}(p)$

$$\overline{H}(p)=\frac{1}{2}|p|^2-\lim_{t\to\infty}\inf_{v\in\mathbb{R}^n}\frac{1}{t}\int_0^t\left(\frac{1}{2}|\gamma'(s)-p|^2+V(\gamma(s))\right)ds,$$

where for each $v\in\mathbb{R}^n$, $\gamma(\cdot)$ is the solution to

$$\begin{cases}\gamma''(s)=DV(\gamma(s)) & \text{for } s>0,\\ \gamma(0)=0,\ \gamma'(0)=v.\end{cases}$$

6 Some representation formulas of the effective Hamiltonian in nonconvex settings

As we have seen above, even for the convex setting, we do not yet have much deep knowledge about the shape of $\overline{H}$. In this section, we present some new results on formulas of $\overline{H}$. The Hamiltonians considered in this section are always of separable forms of $H(p)-V(y)$. By abuse of notions, sometimes, we still write $H(y,p)=H(p)-V(y)$, where $H:\mathbb{R}^n\to\mathbb{R}$ is continuous and coercive, and $V\in C(\mathbb{T}^n)$. This is simply to avoid using too many notions. The results here are taken from Qian, Tran, Yu [124].

6.1 The simplest case

The setting is this. Let $H=H(p):\mathbb{R}^n\to\mathbb{R}$ be a continuous, coercive Hamiltonian such that

$$\begin{cases} \min_{\mathbb{R}^n} H = 0; \\ \text{there exists a bounded domain } U\subset\mathbb{R}^n \text{ such that } \{H=0\}=\partial U; \\ H \text{ is even, that is, } H(p)=H(-p) \text{ for all } p\in\mathbb{R}^n; \\ \text{there exist } H_1,H_2\in C(\mathbb{R}^n) \text{ such that } H=\max\{H_1,H_2\}. \end{cases} \tag{4.25}$$

Here, H_1,H_2 satisfy

$$\begin{cases} H_1 \text{ is coercive, level-set quasiconvex, even,} \\ \text{and } H_1=H \text{ in } \mathbb{R}^n\setminus U,\ H_1<0 \text{ in } U; \\ H_2 \text{ is level-set quasiconcave, even,} \\ \text{and } H_2=H \text{ in } U,\ H_2<0 \text{ in } \mathbb{R}^n\setminus U,\ \lim_{|p|\to\infty} H_2(p)=-\infty. \end{cases} \tag{4.26}$$

An example of H satisfying (4.25)–(4.26) is $H(p)=(|p|^2-2)^2$ as in Figure 4.4.
Below is the decomposition result for this simplest case.

Theorem 4.25. *Let $H\in C(\mathbb{R}^n)$ be a Hamiltonian satisfying (4.25)–(4.26). Let $V\in C(\mathbb{T}^n)$ be given such that $\min_{\mathbb{T}^n} V=0$.*

Assume that $\overline{H}$ is the effective Hamiltonian corresponding to $H(p)-V(y)$. Assume also that $\overline{H}_i$ is the effective Hamiltonian corresponding to $H_i(p)-V(y)$ for $i=1,2$. Then,

$$\overline{H}=\max\left\{\overline{H}_1,\overline{H}_2,0\right\}.$$

In particular, $\overline{H}$ is even.

We would like to point out that the evenness of $\overline{H}$ will be used later and is not obvious at all although H is even. The nonconvex situation makes things much more complicated. See the discussion in Section 6.6 for this subtle issue.

Proof. We proceed in few steps.

Step 1. It is straightforward that $0\le\overline{H}(p)\le H(p)$ for all $p\in\mathbb{R}^n$. Indeed, for each fixed $p\in\mathbb{R}^n$, the corresponding cell problem is (4.10). Pick $y_0\in\mathbb{T}^n$ such that $\min_{\mathbb{T}^n} v=v(y_0)$. By the definition of viscosity supersolutions to (4.10), we get

$$H(p)\ge H(p)-V(y_0)\ge\overline{H}(p).$$

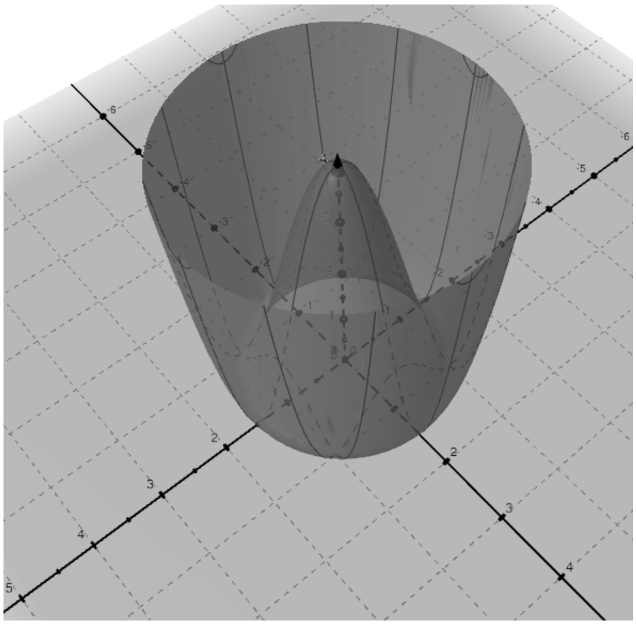


Figure 4.4: An example where $H(p) = (|p|^2 - 2)^2$

On the other hand, as (4.10) holds in the almost everywhere sense, we take essential supremum of its sides to imply

$$\overline{H}(p) = \operatorname*{ess\,sup}_{y\in\mathbb{T}^n} (H(p + Dv(y)) - V(y)) \geq \operatorname*{ess\,sup}_{y\in\mathbb{T}^n}(-V(y)) = 0.$$

In particular,

$$\overline{H}(p) = 0 \qquad \text{for all } p \in \partial U. \tag{4.27}$$

Besides, as $H_i \leq H$, we get $\overline{H}_i \leq \overline{H}$. Therefore,

$$\overline{H} \geq \max\left\{\overline{H}_1, \overline{H}_2, 0\right\}. \tag{4.28}$$

It remains to prove the reverse inequality of (4.28) in order to get the conclusion.

Step 2. Fix $p \in \mathbb{R}^n$. Assume now that $\overline{H}_1(p) \geq \max\left\{\overline{H}_2(p), 0\right\}$. In particular, $\overline{H}_1(p) \geq 0$. We will show that $\overline{H}_1(p) \geq \overline{H}(p)$.

Since H_1 is quasiconvex and even, we use the inf-sup (or inf-max) representation formula for $\overline{H}_1$ (Exercise 42) to get that

$$\begin{aligned}
\overline{H}_1(p) &= \inf_{\phi\in C^1(\mathbb{T}^n)} \max_{y\in\mathbb{T}^n} (H_1(p + D\phi(y)) - V(y)) \\
&= \inf_{\phi\in C^1(\mathbb{T}^n)} \max_{y\in\mathbb{T}^n} (H_1(-p - D\phi(y)) - V(y)) \\
&= \inf_{\psi\in C^1(\mathbb{T}^n)} \max_{y\in\mathbb{T}^n} (H_1(-p + D\psi(y)) - V(y)) = \overline{H}_1(-p).
\end{aligned}$$

Thus, $\overline{H}_1$ is even. Let $v(y, -p)$ be a solution to the cell problem

$$H_1(-p + Dv(y, -p)) - V(y) = \overline{H}_1(-p) = \overline{H}_1(p) \qquad \text{in } \mathbb{T}^n. \tag{4.29}$$

Let $w(y) = -v(y, -p)$. For any $y \in \mathbb{T}^n$ and $q \in D^+w(y)$, we have $-q \in D^-v(y, -p)$ and hence, in light of (4.29) and the quasiconvexity of H_1 (Exercise 30),

$$\overline{H}_1(p) = H_1(-p-q) - V(y) = H_1(p+q) - V(y).$$

We thus get $H_1(p+q) = \overline{H}_1(p) + V(y) \geq 0$ as $\overline{H}_1(p) \geq 0$, and therefore, $H(p+q) = H_1(p+q) \geq 0$ in light of (4.26). This yields that w is a viscosity subsolution to

$$H(p + Dw) - V(y) = \overline{H}_1(p) \qquad \text{in } \mathbb{T}^n.$$

Hence, by Theorem 4.10 on a representation formula of $\overline{H}(p)$, we imply $\overline{H}(p) \leq \overline{H}_1(p)$.

STEP 3. Assume now that $\overline{H}_2(p) \geq \max\left\{\overline{H}_1(p),\ 0\right\}$. By employing similar arguments as those in the previous step (except that we use $v(y, p)$ directly here instead of $v(y, -p)$ due to the quasiconcavity of H_2), we deduce that $\overline{H}_2(p) \geq \overline{H}(p)$.

STEP 4. What is left is the case that $\max\left\{\overline{H}_1(p), \overline{H}_2(p)\right\} < 0$. We now show that $\overline{H}(p) = 0$ in this case. Thanks to (4.27) in Step 1, we may assume that $p \notin \partial U$.

We now introduce an idea that is quite close to the continuation method. For $\sigma \in [0,1]$ and $i = 1, 2$, let $\overline{H}^{\sigma}$, $\overline{H}_i^{\sigma}$ be the effective Hamiltonians corresponding to $H(p) - \sigma V(y)$, $H_i(p) - \sigma V(y)$, respectively. It is clear that

$$0 \leq \overline{H}^1 = \overline{H} \leq \overline{H}^{\sigma} \qquad \text{for all } \sigma \in [0,1]. \tag{4.30}$$

By repeating Steps 2 and 3 above, we get

$$\text{For } p \in \mathbb{R}^n \text{ and } \sigma \in [0,1], \text{ if } \max\left\{\overline{H}_1^{\sigma}(p), \overline{H}_2^{\sigma}(p)\right\} = 0, \text{ then } \overline{H}^{\sigma}(p) = 0. \tag{4.31}$$

We only need consider the case $p \notin \overline{U}$ here as the case $p \in U$ is analogous. Let us notice that

$$H(p) = H_1(p) = \overline{H}_1^0(p) > 0 \quad \text{and} \quad \overline{H}_1(p) = \overline{H}_1^1(p) < 0.$$

By the continuity of $\sigma \mapsto \overline{H}_1^{\sigma}(p)$, there exists $s \in (0,1)$ such that $\overline{H}_1^s(p) = 0$. Note furthermore that, as $p \notin \overline{U}$, $\overline{H}_2^s(p) \leq H_2(p) < 0$. These, together with (4.31), yield that $\overline{H}^s(p) = 0$. Combine this with (4.30) to finally get that $\overline{H}(p) = 0$.

□

Remark 4.26. We emphasize that Step 4 in the above proof is extremely important. It plays the role of a "patching" step, which helps glue $\overline{H}_1$ and $\overline{H}_2$ together. So far, this kind of ideas has not been used so much in the theory of viscosity solutions, and probably it is not needed in the well-posedness theory. Nevertheless, to go beyond the well-posedness theory to understand more about $\overline{H}$ and properties of solutions, it is important to develop this systematically.

Assumptions (4.25)–(4.26) are general and a bit complicated. A simple situation where (4.25)–(4.26) hold is a radially symmetric case where $H(p) = \psi(|p|)$, and $\psi \in C([0,\infty), \mathbb{R})$ satisfying

$$\begin{cases} \psi(0) > 0,\ \psi(1) = 0,\ \lim_{r\to\infty} \psi(r) = +\infty, \\ \psi \text{ is strictly decreasing in } (0,1), \text{ and is strictly increasing in } (1,\infty). \end{cases} \tag{4.32}$$

Let $\psi_1, \psi_2 \in C([0,\infty), \mathbb{R})$ be such that

$$\begin{cases} \psi_1 = \psi \text{ on } [1,\infty), \text{ and } \psi_1 \text{ is strictly increasing on } [0,1], \\ \psi_2 = \psi \text{ on } [0,1], \ \psi_2 \text{ is strictly decreasing on } [1,\infty), \text{ and } \lim_{r\to\infty} \psi_2(r) = -\infty. \end{cases} \tag{4.33}$$

See Figure 4.5. Set $H_i(p) = \psi_i(|p|)$ for $p \in \mathbb{R}^n$, and for $i = 1,2$. It is clear that (4.25)–(4.26) hold true provided that (4.32)–(4.33) hold.

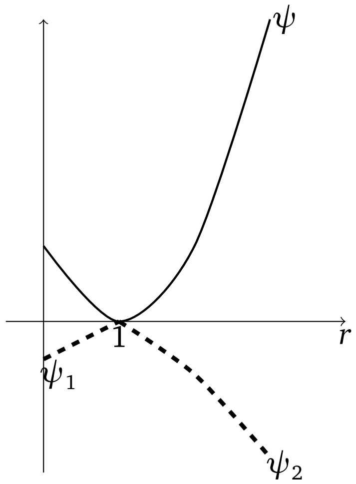


Figure 4.5: Graphs of ψ, ψ_1, ψ_2

An immediate consequence of Theorem 4.25 is the following result.

Corollary 4.27. *Let $H(p) = \psi(|p|)$, $H_i(p) = \psi_i(|p|)$ for $i = 1,2$ and $p \in \mathbb{R}^n$, where ψ, ψ_1, ψ_2 satisfy* (4.32)–(4.33). *Let $V \in C(\mathbb{T}^n)$ be a potential energy with $\min_{\mathbb{T}^n} V = 0$.*

Assume that $\overline{H}$ is the effective Hamiltonian corresponding to $H(p) - V(y)$. Assume also that $\overline{H}_i$ is the effective Hamiltonian corresponding to $H_i(p) - V(y)$ for $i = 1,2$. Then

$$\overline{H} = \max\left\{\overline{H}_1, \overline{H}_2, 0\right\}.$$

Remark 4.28. A special case of Corollary 4.27 is when

$$H(p) = \psi(|p|) = \left(|p|^2 - 1\right)^2 \quad \text{for } p \in \mathbb{R}^n,$$

which was studied first by Armstrong, Tran and Yu [6]. Of course, Armstrong, Tran and Yu [6] dealt with stochastic (random) homogenization, but their results can be casted in term of periodic homogenization as well. The method here is much simpler and more robust than that in [6].

By using Corollary 4.27 and approximation, we get another representation formula for $\overline{H}$ which will be used later.

Corollary 4.29. *Assume that* (4.32)–(4.33) *hold. Set*

$$\tilde{\psi}_1(r) = \max\{\psi_1, 0\} = \begin{cases} 0 & \text{for } 0 \le r \le 1, \\ \psi(r) & \text{for } r > 1. \end{cases}$$

Let $H(p) = \psi(|p|)$, $\tilde{H}_1(p) = \tilde{\psi}_1(p|)$ and $H_2(p) = \psi_2(|p|)$ for $p \in \mathbb{R}^n$. Let $V \in C(\mathbb{T}^n)$ be a potential energy with $\min_{\mathbb{T}^n} V = 0$.
Assume that $\overline{H}, \overline{\tilde{H}}_1, \overline{H}_2$ are the effective Hamiltonian corresponding to $H(p) - V(y), \tilde{H}_1(p) - V(y), H_2(p) - V(y)$, respectively. Then

$$\overline{H} = \max\left\{\overline{\tilde{H}}_1, \overline{H}_2\right\}.$$

See Figure 4.6 for the graphs of $\psi, \tilde{\psi}_1, \psi_2$.

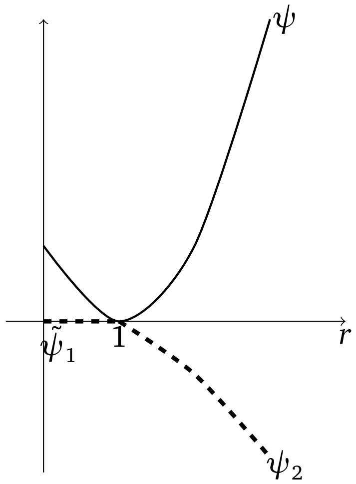


Figure 4.6: Graphs of $\psi, \tilde{\psi}_1, \psi_2$

When the oscillation of V is large enough, it turns out that $\overline{H}$ is level-set quasiconvex. This is the content of the next result.

Corollary 4.30. *Let $H \in C(\mathbb{R}^n)$ be a coercive Hamiltonian satisfying* (4.25)–(4.26)*, except that we do not require H_2 to be quasiconcave. Assume that*

$$\operatorname{osc}_{\mathbb{T}^n} V = \max_{\mathbb{T}^n} V - \min_{\mathbb{T}^n} V \geq \max_{\overline{U}} H = \max_{\mathbb{R}^n} H_2.$$

Then

$$\overline{H} = \max\left\{\overline{H}_1, \ -\min_{\mathbb{T}^n} V\right\}.$$

In particular, $\overline{H}$ is quasiconvex in this situation.

It is worth noting that the result of Corollary 4.30 is interesting in the sense that we do not require any structure of H in U except that $H > 0$ there. In earlier results in this section, we needed to assume that H is quasiconcave in U, but when $\operatorname{osc}_{\mathbb{T}^n} V$ is large enough, we do not need it. Roughly speaking, when $\operatorname{osc}_{\mathbb{T}^n} V$ is large, V has enough power to iron out all the ripples in the graph of H in U to get a nice $\overline{H}$. It is, in fact, quite unexpected that $\overline{H}$ behaves better than H. It is often known in the literature earlier that $\overline{H}$ always behaves worse than H (see discussions in Section 5.6). This is one of the first instance showing that it is otherwise provided that $\operatorname{osc}_{\mathbb{T}^n} V$ is large.

Proof. Without loss of generality, we assume that $\min_{\mathbb{T}^n} V = 0$. Choose an even, quasiconcave function $H_2^+ \in C(\mathbb{R}^n)$ such that

$$\begin{cases} \{H = 0\} = \{H_2^+ = 0\} = \partial U, \\ H \le H_2^+ \text{ in } U, \text{ and } \max_{\overline{U}} H = \max_{\mathbb{R}^n} H_2^+, \\ \lim_{|p|\to\infty} H_2^+(p) = -\infty. \end{cases}$$

Denote $H^+ \in C(\mathbb{R}^n)$ as

$$H^+(p) = \max\{H, H_2^+\} = \begin{cases} H_1(p) & \text{for } p \in \mathbb{R}^n \backslash U, \\ H_2^+(p) & \text{for } p \in \overline{U}. \end{cases}$$

Let $\overline{H}^+$ and $\overline{H}_2^+$ be the effective Hamiltonians associated with $H^+(p) - V(y)$ and $H_2^+(p) - V(y)$, respectively. Apparently,

$$\max\left\{\overline{H}_1,\ 0\right\} \le \overline{H} \le \overline{H}^+. \tag{4.34}$$

On the other hand, by Theorem 4.25, the representation formula for $\overline{H}^+$ is

$$\overline{H}^+ = \max\left\{\overline{H}_1,\ \overline{H}_2^+,\ 0\right\} = \max\{\overline{H}_1,\ 0\}, \tag{4.35}$$

where the second equality is due to the fact that

$$\overline{H}_2^+ \le \max_{\mathbb{R}^n} H_2^+ - \max_{\mathbb{T}^n} V = \max_{\overline{U}} H - \max_{\mathbb{R}^n} V \le 0.$$

We combine (4.34) and (4.35) to get the conclusion. □

6.2 A more general case

We now proceed to give an extension of Theorem 4.25 to a case which is a bit more general. To avoid unnecessary technicalities, we only consider radially symmetric cases from now on in this section. The results still hold true for general Hamiltonians (without the radially symmetric assumption) under corresponding appropriate conditions, which are similar to (4.25)–(4.26).

Let $H : \mathbb{R}^n \to \mathbb{R}$ be such that

$$\begin{cases} H(p) = \varphi(|p|) \text{ for } p \in \mathbb{R}^n, \text{ where } \varphi \in C([0,\infty),\mathbb{R}) \text{ satisfies} \\ \varphi(0) > 0,\ \varphi(2) = 0,\ \lim_{r\to\infty} \varphi(r) = +\infty, \\ \varphi \text{ is strictly increasing on } [0,1] \text{ and } [2,\infty), \\ \text{and } \varphi \text{ is strictly decreasing on } [1,2]. \end{cases} \tag{4.36}$$

Now, we denote by $H_i(p) = \varphi_i(|p|)$ for $p \in \mathbb{R}^n$ and $1 \le i \le 3$, where $\varphi_1, \varphi_2, \varphi_3 \in C([0,\infty),\mathbb{R})$ are such that

$$\begin{cases} \varphi_1 = \varphi \text{ on } [2,\infty),\ \varphi_1 \text{ is strictly increasing on } [0,2], \\ \varphi_2 = \varphi \text{ on } [1,2],\ \varphi_2 \text{ is strictly decreasing on } [0,1] \text{ and } [2,\infty),\ \lim_{r\to\infty}\varphi_2(r) = -\infty, \\ \varphi_3 = \varphi \text{ on } [0,1],\ \varphi_3 \text{ is strictly increasing on } [1,\infty), \text{ and } \varphi_3 > \varphi \text{ in } (1,\infty). \end{cases} \tag{4.37}$$

See Figure 4.7.

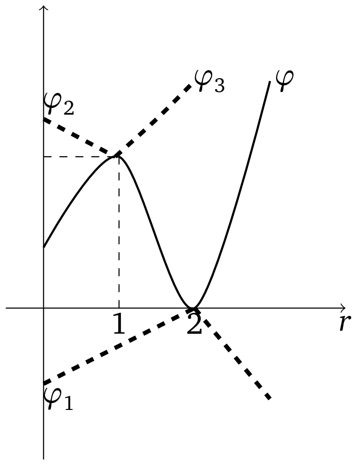


Figure 4.7: Graphs of $\varphi, \varphi_1, \varphi_2, \varphi_3$

Lemma 4.31. *Let $H(p) = \varphi(|p|)$, $H_i(p) = \varphi_i(|p|)$ for $1 \le i \le 3$ and $p \in \mathbb{R}^n$, where $\varphi, \varphi_1, \varphi_2, \varphi_3$ satisfy* (4.36)–(4.37). *Let $V \in C(\mathbb{T}^n)$ be a potential energy with $\min_{\mathbb{T}^n} V = 0$. Assume that $\overline{H}$ is the effective Hamiltonian corresponding to $H(p) - V(y)$. Assume also that $\overline{H}_i$ is the effective Hamiltonian corresponding to $H_i(p) - V(y)$ for $1 \le i \le 3$. Then*

$$\begin{aligned}\overline{H} &= \max\left\{0, \overline{H}_1, \overline{K}\right\} \\ &= \max\left\{0, \overline{H}_1, \min\left\{\overline{H}_2, \overline{H}_3, \varphi(1) - \max_{\mathbb{T}^n} V\right\}\right\}.\end{aligned}$$

Here $\overline{K}$ is the effective Hamiltonian corresponding to $K(p) - V(y)$, where $K : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$K(p) = \min\{\varphi_2(|p|), \varphi_3(|p|)\} = \begin{cases} \varphi(|p|) & \text{if } |p| \le 2, \\ \varphi_2(|p|) & \text{if } |p| \ge 2. \end{cases}$$

In particular, both $\overline{H}$ and $\overline{K}$ are even.

We want to note that the proof below does not depend on the quasiconvexity of $\overline{H}_3$. As $H_3 \ge H$, we only use the simple fact that $\overline{H}_3 \ge \overline{H}$. This point is essential for us to prove the most general result later (see Theorem 4.32).

Proof. Considering $-K(-p)$, thanks to the representation formula and evenness from Theorem 4.25,

$$\overline{K} = \min\left\{\overline{H}_2, \overline{H}_3, \varphi(1) - \max_{\mathbb{T}^n} V\right\}.$$

Define $\tilde{\varphi}_2 = \min\{\varphi_2, \varphi(1)\}$. Let $\tilde{H}_2(p) = \tilde{\varphi}_2(|p|)$ and $\overline{\tilde{H}}_2$ be the effective Hamiltonian corresponding to $\tilde{H}_2(p) - V(y)$. Then, by Corollary 4.29, we have another representation formula for $\overline{K}$ as following

$$\overline{K} = \min\left\{\overline{\tilde{H}}_2, \overline{H}_3\right\}. \tag{4.38}$$

Our goal is then to show that $\overline{H} = \max\left\{0, \overline{H}_1, \overline{K}\right\}$. To do this, we again divide the proof into few steps for clarity.

STEP 1. First of all, it is clear that $0 \le \overline{H} \le H$. This implies further that

$$\overline{H}(p) = 0 \qquad \text{for all } |p| = 2. \tag{4.39}$$

Besides, as $K, H_1 \le H$, we deduce furthermore that $\overline{K}, \overline{H}_1 \le \overline{H}$. Thus,

$$\overline{H} \ge \max\left\{0, \overline{H}_1, \overline{K}\right\} \tag{4.40}$$

We now show the reverse inequality of (4.40) to finish the proof.

STEP 2. Fix $p \in \mathbb{R}^n$. Assume that $\overline{H}_1(p) \ge \max\left\{0, \overline{K}(p)\right\}$. Since H_1 is quasiconvex, we follow exactly the same lines of Step 2 in the proof of Theorem 4.25 to deduce that $\overline{H}_1(p) \ge \overline{H}(p)$.

STEP 3. Assume that $\overline{K}(p) \ge \max\left\{0, \overline{H}_1(p)\right\}$. Since K is not quasiconvex or quasiconcave, we cannot directly use Step 2 or Step 3 in the proof of Theorem 4.25 to conclude. Instead, there are two cases that need to be considered.

Firstly, we consider the case that $\overline{K}(p) = \overline{\tilde{H}}_2(p) \le \overline{H}_3(p)$. Let $v(y,p)$ be a solution to the cell problem

$$\tilde{H}_2(p + Dv(y,p)) - V(y) = \overline{\tilde{H}}_2(p) \ge 0 \qquad \text{in } \mathbb{T}^n. \tag{4.41}$$

Since $\tilde{H}_2$ is quasiconcave, for any $y \in \mathbb{T}^n$ and $q \in D^+v(y,p)$, we have

$$\tilde{H}_2(p+q) - V(y) = \overline{\tilde{H}}_2(p) \ge 0,$$

which gives that $\tilde{H}_2(p+q) \ge 0$, and hence, $\tilde{H}_2(p+q) \ge H(p+q)$. Therefore, $v(y,p)$ is a viscosity subsolution to

$$H(p + Dv(y,p)) - V(y) = \overline{\tilde{H}}_2(p) \qquad \text{in } \mathbb{T}^n.$$

This, together with Theorem 4.10 on a representation formula of $\overline{H}(p)$, implies that $\overline{K}(p) = \overline{\tilde{H}}_2(p) \ge \overline{H}(p)$.

Secondly, assume that $\overline{K}(p) = \overline{H}_3(p) \le \overline{\tilde{H}}_2(p)$. Since $\varphi_3 \ge \varphi$, $\overline{H}_3(p) \ge \overline{H}(p)$. Combining with $\overline{H}(p) \ge \overline{K}(p)$ in (4.40), we obtain $\overline{K}(p) = \overline{H}(p)$ in this step.

STEP 4. Assume that $0 > \max\left\{\overline{H}_1(p), \overline{K}(p)\right\}$. Our goal now is to show $\overline{H}(p) = 0$. Thanks to (4.39) in Step 1, we may assume that $|p| \ne 2$.

For $\sigma \in [0,1]$, let $\overline{H}^\sigma, \overline{H}_1^\sigma$, $\overline{K}^\sigma$ be the effective Hamiltonians corresponding to $H(p) - \sigma V(y), H_1(p) - \sigma V(y)$, $K(p) - \sigma V(y)$, respectively. It is clear that

$$0 \le \overline{H}^1 = \overline{H} \le \overline{H}^\sigma \quad \text{ for all } \sigma \in [0,1]. \tag{4.42}$$

By repeating Steps 2 and 3 above, we get

$$\text{For } p \in \mathbb{R}^n \text{ and } \sigma \in [0,1], \text{ if } \max\left\{\overline{H}_1^\sigma(p), \overline{K}^\sigma(p)\right\} = 0, \text{ then } \overline{H}^\sigma(p) = 0. \tag{4.43}$$

It is enough to consider the case $|p| < 2$ here as the case $|p| > 2$ is analogous. Notice that

$$H(p) = K(p) = \overline{K}^0(p) > 0 \quad \text{and} \quad \overline{K}(p) = \overline{K}^1(p) < 0.$$

By the continuity of $\sigma \mapsto \overline{K}^\sigma(p)$, there exists $s \in (0,1)$ such that $\overline{K}^s(p) = 0$. Note furthermore that, as $|p| < 2$, $\overline{H}_1^s(p) \le H_1(p) < 0$. These, together with (4.42) and (4.43), yield the desired result. □

6.3 General cases

By using induction, we are able to obtain min-max (max-min) formulas for $\overline{H}$ in case $H(p) = \varphi(|p|)$ where φ satisfies some certain conditions described below. The approach is essentially the same as in the above two sections provided that we are careful enough with the iterations.

We consider two such cases corresponding to Figures 4.8 and 4.9. Roughly speaking, in both cases, the graph of φ has a finite number of oscillations starting from 0, and geometrically, the magnitudes of oscillations of $\varphi(s)$ increase as s increases.

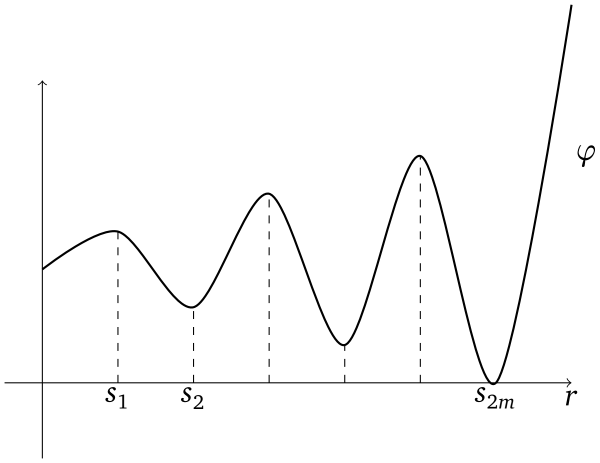


Figure 4.8: Graph of φ in first general case

In the first general case corresponding to Figure 4.8, we assume that

$$\begin{cases} \varphi \in C([0,\infty),\mathbb{R}) \text{ satisfies that} \\ \text{there exist } m \in \mathbb{N} \text{ and } 0 = s_0 < s_1 < \dots s_{2m} < \infty = s_{2m+1} \text{ such that} \\ \varphi \text{ is strictly increasing in } (s_{2i}, s_{2i+1}), \text{ and is strictly decreasing in } (s_{2i+1}, s_{2i+2}), \\ \varphi(s_0) > \varphi(s_2) > \dots > \varphi(s_{2m}), \text{ and } \varphi(s_1) < \varphi(s_3) < \dots < \varphi(s_{2m+1}) = \infty. \end{cases} \tag{4.44}$$

Based on φ, we construct $\varphi_0, \dots, \varphi_{2m}$ as following.

- For $0 \le i \le m$, let $\varphi_{2i} : [0,\infty) \to \mathbb{R}$ be a continuous, strictly increasing function such that $\varphi_{2i} = \varphi$ on $[s_{2i}, s_{2i+1}]$ and $\lim_{s\to\infty} \varphi_{2i}(s) = \infty$. Besides, we construct so that $\varphi_{2i} \ge \varphi_{2i+2}$ for $0 \le i \le m-1$.
- For $0 \le i \le m-1$, let $\varphi_{2i+1} : [0,\infty) \to \mathbb{R}$ be a continuous, strictly decreasing function such that $\varphi_{2i+1} = \varphi$ on $[s_{2i+1}, s_{2i+2}]$ and $\lim_{s\to\infty} \varphi_{2i+1}(s) = -\infty$. Besides, we construct so that $\varphi_{2i+1} \le \varphi_{2i+3}$ for $0 \le i \le m-2$.

Define

$$H_{m-1}(p) = \max\{\varphi(|p|),\ \varphi_{2m-2}(|p|)\} = \begin{cases} \varphi(|p|) & \text{for } |p| \le s_{2m-1}, \\ \varphi_{2m-2}(|p|) & \text{for } |p| > s_{2m-1} \end{cases}$$

and

$$k_{m-1}(s)=\min\{\varphi(s),\ \varphi_{2m-1}(s)\}=\begin{cases}\varphi(s) & \text{for } s\le s_{2m},\\ \varphi_{2m-1}(s) & \text{for } s>s_{2m}.\end{cases}$$

Denote $\overline{H}_{m-1}$, $\overline{H}_m$, $\overline{K}_{m-1}$, $\overline{\Phi}_j$ as the effective Hamiltonians associated with the Hamiltonians $H_{m-1}(p)-V(y)$, $\varphi(|p|)-V(y)$, $k_{m-1}(|p|)-V(y)$ and $\varphi_j(|p|)-V(y)$ for $0\le j\le 2m$, respectively.

This is the main decomposition result of $\overline{H}$ in this section.

Theorem 4.32. *Assume that* (4.44) *holds for some* $m\in\mathbb{N}$. *Then,*

$$\overline{H}_m=\max\left\{\overline{K}_{m-1},\ \overline{\Phi}_{2m},\ \varphi(s_{2m})-\min_{\mathbb{T}^n}V\right\},\tag{4.45}$$

and

$$\overline{K}_{m-1}=\min\left\{\overline{H}_{m-1},\ \overline{\Phi}_{2m-1},\ \varphi(s_{2m-1})-\max_{\mathbb{T}^n}V\right\}.\tag{4.46}$$

In particular, $\overline{H}_m$ *and* $\overline{K}_{m-1}$ *are both even.*

We stress again that the evenness of $\overline{H}_m$ and $\overline{K}_{m-1}$ is far from being obvious although H_m and K_m are both even. See the discussion in Section 6.6 for this subtle issue.

Proof. We prove by induction.

The base case is when $m=1$. The two formulas (4.45) and (4.46) follow immediately from Lemma 4.31 and Theorem 4.25.

Assume that (4.45) and (4.46) hold for $m\in\mathbb{N}$. We need to verify these equalities for $m+1$. Using similar arguments as those in the proof Lemma 4.31, and noting the statement right before its proof, we first get that

$$\overline{K}_m=\min\left\{\overline{H}_m,\ \overline{\Phi}_{2m+1},\ \varphi(s_{2m+1})-\max_{\mathbb{T}^n}V\right\}.$$

Then again, by basically repeating the proof of Lemma 4.31, we obtain

$$\overline{H}_{m+1}=\max\left\{\overline{K}_m,\ \overline{\Phi}_{2m+2},\ \varphi(s_{2m+2})-\min_{\mathbb{T}^n}V\right\}.$$

□

Remark 4.33. Two comments are in order.

(i) By approximation, we see that representation formulas (4.45) and (4.46) still hold true if we relax (4.44) a bit, that is, we only require that φ satisfies

$$\begin{cases}\varphi \text{ is increasing in } (s_{2i},s_{2i+1}), \text{ and is decreasing in } (s_{2i+1},s_{2i+2}),\\ \varphi(s_0)\ge\varphi(s_2)\ge\ldots\ge\varphi(s_{2m}), \text{ and } \varphi(s_1)\le\varphi(s_3)\le\ldots<\varphi(s_{2m+1})=\infty.\end{cases}$$

(ii) According to Corollary 4.30, if $\operatorname{osc}_{\mathbb{T}^n}V=\max_{\mathbb{T}^n}V-\min_{\mathbb{T}^n}V\ge\varphi(s_{2m-1})-\varphi(s_{2m})$, then $\overline{H}$ is quasiconvex and

$$\overline{H}=\max\left\{\overline{\Phi}_{2m},\ \varphi(s_{2m})-\min_{\mathbb{T}^n}V\right\}.$$

The second general case corresponds to the case where $H(p)=-k_{m-1}(|p|)$ for all $p\in\mathbb{R}^n$ as described in Figure 4.9 after normalization by a constant. By changing the notations appropriately, we obtain similar representation formulas as in Theorem 4.32. We omit the details here.

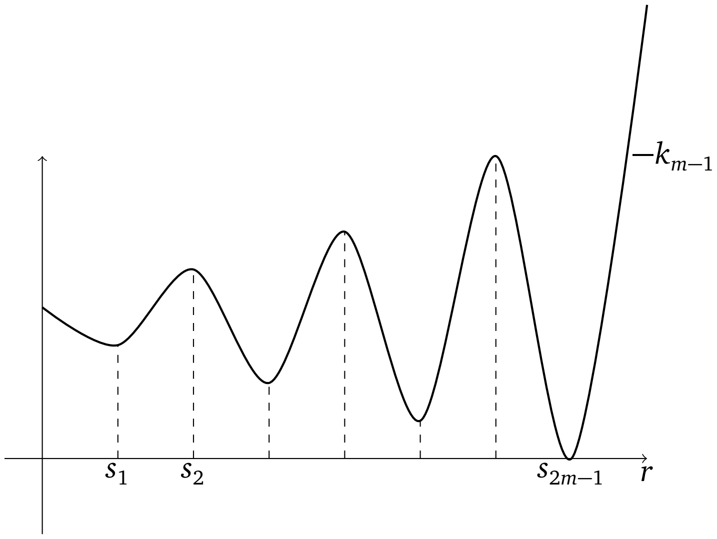


Figure 4.9: Graph of $-k_{m-1}$ in the second general case

6.4 Quasiconvexification effect

This quasiconvexification effect was discussed in the previous section already. We just want to emphasize again clearly this very interesting phenomenon here.

By Corollary 4.30, if $\operatorname{osc}_{\mathbb{T}^n} V = \max_{\mathbb{T}^n} V - \min_{\mathbb{T}^n} V \geq \max_{\overline{U}} H = \max_{\mathbb{R}^n} H_2$, then $\overline{H}$ is quasiconvex, which means that $\overline{H}$ behaves better than the original Hamiltonian H before the homogenization process. This goes against the earlier belief in the literature of homogenization theory that $\overline{H}$ always behaves worse than H (see discussions in Section 5.6). In fact, we have an explicit representation formula for $\overline{H}$ as

$$\overline{H} = \max\left\{\overline{H}_1,\ -\min_{\mathbb{T}^n} V\right\}.$$

In case $\min_{\mathbb{T}^n} V = 0$, then we require that $\operatorname{osc}_{\mathbb{T}^n} V = \max_{\mathbb{T}^n} V \geq \max_{\overline{U}} H = \max_{\mathbb{R}^n} H_2$ to have

$$\overline{H} = \max\left\{\overline{H}_1,\ 0\right\}.$$

Roughly speaking, when $\operatorname{osc}_{\mathbb{T}^n} V$ is large, potential energy V has enough power to iron out all the ripples in the graph of H in U to get a nice $\overline{H}$. See Figures 4.10–4.11 for two one dimensional examples of H and $\overline{H}$. Note that $\overline{H}$ is not even in Figure 4.11.

This quasiconvexification phenomenon also holds for a more general setting in Theorem 4.32. Here, if $H(p) = \varphi(|p|)$, and φ satisfies (4.44) (that is, φ has the graph as in Figure 4.8), and $\operatorname{osc}_{\mathbb{T}^n} V = \max_{\mathbb{T}^n} V - \min_{\mathbb{T}^n} V \geq \varphi(s_{2m-1}) - \varphi(s_{2m})$, then $\overline{H}$ is quasiconvex and

$$\overline{H} = \max\left\{\overline{\Phi}_{2m},\ \varphi(s_{2m}) - \min_{\mathbb{T}^n} V\right\}.$$

It is clear from Theorem 4.32 that $\varphi(s_{2m-1}) - \varphi(s_{2m})$ is the optimal lower bound for $\operatorname{osc}_{\mathbb{T}^n} V$ to see the quasiconvexification effect.

In general, if $H(p) = \varphi(|p|)$ for some $\varphi : [0,\infty) \to \mathbb{R}$, which is coercive but does not necessarily satisfy (4.44), then it is not clear yet whether this quasiconvexification phenomenon happens or not. Some further analysis and discussions on this can be found in [124].

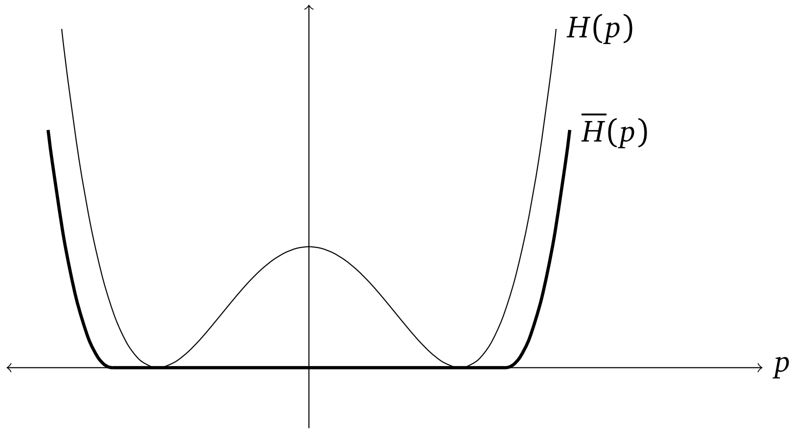


Figure 4.10: Quasiconvexification of $\overline{H}$ in one dimension - first example.

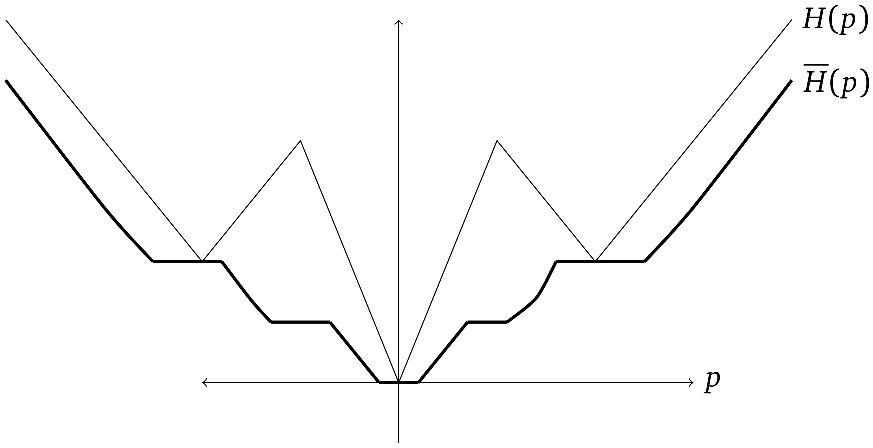


Figure 4.11: Quasiconvexification of $\overline{H}$ in one dimension - second example.

6.5 Problems

Exercise 44. *Assume that H satisfies* (4.2) *and* (4.3). *Let $G(y,p) = -H(y,-p)$ for $(y,p) \in \mathbb{T}^n \times \mathbb{R}^n$. Show that $\overline{G}(p) = -\overline{H}(-p)$ for all $p \in \mathbb{R}^n$.*

Exercise 45. *Assume $H(p) = -k_{m-1}(|p|)$ for all $p \in \mathbb{R}^n$ as described in Figure 4.9, and $V \in C(\mathbb{T}^n)$. Obtain the formula for $\overline{H}(p)$ of the Hamiltonian $H(p) - V(y)$.*

6.6 Loss of evenness and non-decomposable effective Hamiltonians

A natural question is whether we can extend Theorem 4.32 to other nonconvex H or not. That is, if H can be decomposed into m nice quasiconvex/concave Hamiltonians H_i $(1 \le i \le m)$, then can we have that $\overline{H}$ is given by a decomposition formula (e.g., min-max type) involving $\overline{H}_i$, $\min V$ and $\max V$:

$$\overline{H} = G(\overline{H}_1, ..., \overline{H}_m, \ \min V, \ \max V) \tag{4.47}$$

for any $V \in C(\mathbb{T}^n)$? Here $\overline{H}$ and $\overline{H}_i$ are effective Hamiltonians associated with $H - V$ and $H_i - V$, respectively.

Note that for quasiconvex/concave function $F : \mathbb{R}^n \to \mathbb{R}$, using the inf-sup formula, it is easy to see that the effective Hamiltonians associated with $F(p)-V(y)$ and $F(p)-V(-y)$ are the same. Hence if such a decomposition formula indeed exists for a specific nonconvex H, effective Hamiltonians associated with $H(p)-V(y)$ and $H(p)-V(-y)$ have to be identical as well. In particular, if H is even in p, then we may assume that H_i $(1 \le i \le m)$ are even in p as well. The question of interest then is whether $\overline{H}$ is even too?

Although this is a simple and natural question, it has not been studied much in the literature. In [102], it was briefly discussed that if H is even in p, then so is $\overline{H}$. However, this turns out to be false in some cases. We give below some answers and discussions to this simple point following the results in [124].

1. If H is quasiconvex, the answer is of course affirmative due to the inf-sup formula

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{T}^n)} \sup_{y \in \mathbb{T}^n} (H(p + D\phi(y)) - V(y))$$

   as shown in the proof of Theorem 4.25.

2. For genuinely nonconvex H, if $\overline{H}$ can be written as a min-max formula involving effective Hamiltonians of even quasiconvex (or quasiconcave) Hamiltonians, then $\overline{H}$ is still even (e.g., see Corollary 4.27, Lemma 4.31, and Theorem 4.32).

3. However, in general, the evenness is lost as presented in [107, Remark 1.2]. Let us quickly recall the setting there.

   We consider the one dimensional case $(n = 1)$, and choose $H(p) = \varphi(|p|)$ for $p \in \mathbb{R}$, where φ satisfies

$$\begin{cases} \varphi \in C([0,\infty),[0,\infty)), \text{ and there exist } 0 < r_1 < r_2 \text{ so that} \\ \varphi(0) = 0, \varphi(r_1) = \frac{1}{2}, \varphi(r_2) = \frac{1}{3}, \lim_{r\to\infty} \varphi(r) = +\infty, \\ \varphi \text{ is strictly increasing on } [0,r_1] \text{ and } [r_2,\infty), \\ \varphi \text{ is strictly decreasing on } [r_1,r_2]. \end{cases}$$

   See Figure 4.12 below. Fix $s \in (0,1)$, and set $V_s(y) = \min\left\{\frac{y}{s}, \frac{1-y}{1-s}\right\}$ for $y \in [0,1]$. Extend V to $\mathbb{R}$ in a periodic way. Then $\overline{H}$ is not even unless $s = \frac{1}{2}$. In particular, this implies that a decomposition formula for $\overline{H}$ of the form (4.47) does not exist. This lack of evenness is natural if we think of the fact that viscosity solutions select gradient jumps in a non-symmetric way. Nevertheless, this also means that much needs to be studied in order to have more systematic understandings of this kind of Hamiltonians.

4. It is extremely interesting if we can point out some further general requirements on H and V in the genuinely nonconvex setting, under which $\overline{H}$ is even. The interplay between H and V plays a crucial role here as we have seen many times in this section and the earlier ones.

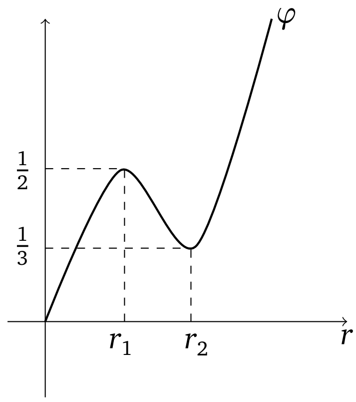


Figure 4.12: Graphs of φ

7 Rates of convergence

7.1 The method of Capuzzo-Dolcetta and Ishii

We now address the results by Capuzzo-Dolcetta and Ishii [27]. Assume that H satisfies (4.2) and (4.3). Our goal here is to show that the rate of convergence of u^ε to u is $O(\varepsilon^{1/3})$. Capuzzo-Dolcetta and Ishii [27] studied homogenizations for static Hamilton–Jacobi equations, but their approach can be easily adjusted to handle the Cauchy problem as well. Here is the main result.

Theorem 4.34. *Assume that $H \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (4.2) *and* (4.3)*. Let $\overline{H}$ be the corresponding effective Hamiltonian of H. Assume $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$. For each $\varepsilon > 0$, let u^ε be the unique viscosity solution of*

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{4.48}$$

And let u be the unique solution to the effective equation

$$\begin{cases} u_t + \overline{H}(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{4.49}$$

Then, for each $T > 0$, there exists a constant $C > 0$ dependent on H, u_0, and T such that

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C\varepsilon^{1/3}. \tag{4.50}$$

We first make some observations and reductions. Under our assumptions, we can find $C > 0$, which depends only on H and u_0, such that

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

Therefore, behavior of $H(y,p)$ for $|p| > C+1$ does not matter. We thus can modify $H(y,p)$ for $|p| > C+1$ so that H is always Lipschitz in p. In other words, we impose the following additional assumption in this section from now on: There exists $C > 0$ such that

$$|H(y,p) - H(y,q)| \le C|p-q| \qquad \text{for all } y \in \mathbb{T}^n,\ p,q \in \mathbb{R}^n. \tag{4.51}$$

And of course, this additional condition does not change any generality of Theorem 4.34.

For each $p \in \mathbb{R}^n$, we first look back at the discount approximation of cell problem (4.10) as following. For each $\lambda > 0$, we consider the static equation

$$\lambda v^\lambda + H(y, p + Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n. \tag{4.52}$$

To make it clear, we write the unique solution to the above as $v^\lambda = v^\lambda(y,p)$. Let us summarize some needed results here, which were covered already in Corollary 4.5 and part (c) of the proof of Theorem 4.8.

Lemma 4.35. *Assume that H satisfies* (4.2)*,* (4.3)*, and* (4.51). *Then, the following claims hold.*

(i) There exists $C > 0$ independent of $\lambda > 0$ such that, for all $p, q \in \mathbb{R}^n$,

$$\lambda |v^\lambda(y,p) - v^\lambda(y,q)| \le C|p-q| \qquad \textit{for all } y \in \mathbb{T}^n.$$

In particular, $|\overline{H}(p) - \overline{H}(q)| \le C|p-q|$.

(ii) For each $R > 0$, there exists a constant $C = C(R) > 0$ independent of $\lambda > 0$ such that, for all $p \in B(0,R)$,

$$|\lambda v^\lambda(y,p) + \overline{H}(p)| \le C\lambda \qquad \textit{for all } y \in \mathbb{T}^n.$$

Proof. Part (i) is quite straightforward as we see that $v^\lambda(y,q) \pm \frac{C}{\lambda}|p-q|$ are a supersolution and a subsolution to (4.52), respectively, thanks to (4.51). Therefore,

$$v^\lambda(\cdot,q) - \frac{C}{\lambda}|p-q| \le v^\lambda(\cdot,p) \le v^\lambda(\cdot,q) + \frac{C}{\lambda}|p-q|.$$

Then, let $\lambda \to 0$ to get $|\overline{H}(p) - \overline{H}(q)| \le C|p-q|$.

To prove (ii), let v be a solution to (4.10) with $\min_{\mathbb{T}^n} v = 0$, that is, v solves

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n.$$

Fix $R > 0$. For $|p| < R$, $\overline{H}(p) \le \overline{H}(0) + CR$. This, together with the coercivity of H, implies that there exists $C = C(R)$ such that $\|Dv\|_{L^\infty(\mathbb{T}^n)} \le C(R)$. Hence,

$$\|v\|_{L^\infty(\mathbb{T}^n)} = \max_{\mathbb{T}^n} v \le \min_{\mathbb{T}^n} v + \sqrt{n}\,\|Dv\|_{L^\infty(\mathbb{T}^n)} \le C(R).$$

We now note that $-\frac{\overline{H}(p)}{\lambda} + v \pm \|v\|_{L^\infty(\mathbb{T}^n)}$ are a supersolution and a subsolution to (4.52), respectively. The usual comparison principle gives

$$-\frac{\overline{H}(p)}{\lambda} + v - \|v\|_{L^\infty(\mathbb{T}^n)} \le v^\lambda \le -\frac{\overline{H}(p)}{\lambda} + v + \|v\|_{L^\infty(\mathbb{T}^n)},$$

which means

$$\|\lambda v^\lambda + \overline{H}(p)\|_{L^\infty(\mathbb{T}^n)} \le 2\lambda \|v\|_{L^\infty(\mathbb{T}^n)} \le C(R)\lambda.$$

□

We are now ready to prove the $O(\varepsilon^{1/3})$ rate of convergence.

Proof of Theorem 4.34. Again, by the reduction step, we assume also (4.51).

We consider the following auxiliary function

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

where $\lambda = \varepsilon^\theta$, and $\beta, \theta \in (0,1)$ and $K > 0$ are to be chosen later. Assume that Φ admits a strict global maximum at $(\hat{x}, \hat{y}, \hat{t}, \hat{s})$ on $\mathbb{R}^{2n} \times [0,T]^2$ for simplicity (for rigorous proof, we need to add the term $-\gamma|x|^2$ to Φ for $\gamma > 0$ (see [27, Theorem 1.1]).

Let us consider first the case that $\hat{t}, \hat{s} > 0$. We claim that if $0 < \theta < 1-\beta$, then there exists $C > 0$ such that

$$|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}| \le C\varepsilon^\beta.$$

Indeed, the fact that $\Phi(\hat{x},\hat{x},\hat{t},\hat{t}) \le \Phi(\hat{x},\hat{y},\hat{t},\hat{s})$, together with Lipschitz property of u and Lemma 4.35, implies

$$\begin{aligned}
\frac{|\hat{x}-\hat{y}|^2+|\hat{t}-\hat{s}|^2}{2\varepsilon^\beta} &\le u(\hat{y},\hat{s}) - u(\hat{x},\hat{t}) + \varepsilon\left(v^\lambda\left(\frac{\hat{x}}{\varepsilon}, \frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right) - v^\lambda\left(\frac{\hat{x}}{\varepsilon}, 0\right)\right) + K(\hat{s}-\hat{t}) \\
&\le C(|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}|) + C\varepsilon\frac{1}{\lambda}\frac{|\hat{x}-\hat{y}|}{\varepsilon^\beta} \\
&\le C(|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}|)
\end{aligned}$$

as $\lambda = \varepsilon^\theta$ with $0 < \theta < 1-\beta$. Thus, our claim holds true.

Notice that $(x,t) \mapsto \Phi(x,t,\hat{y},\hat{s})$ has a maximum at $(\hat{x},\hat{t})$. For $\alpha > 0$, set

$$\psi(x,\xi,z,t) = u^\varepsilon(x,t) - \varepsilon v^\lambda\left(\xi, \frac{z-\hat{y}}{\varepsilon^\beta}\right) - \frac{|x-\hat{y}|^2+|t-\hat{s}|^2}{2\varepsilon^\beta} - \frac{|x-\varepsilon\xi|^2+|x-z|^2}{2\alpha} - Kt.$$

Assume ψ has a maximum at $(x_\alpha, \xi_\alpha, z_\alpha, t_\alpha)$ and we can assume by passing to a subsequence if necessary that $(x_\alpha, \xi_\alpha, z_\alpha, t_\alpha) \to (\hat{x}, \hat{x}/\varepsilon, \hat{x}, \hat{t})$ as $\alpha \to 0$. By the definition of viscosity solutions, we have

$$K + \frac{t_\alpha - \hat{s}}{\varepsilon^\beta} + H\left(\frac{x_\alpha}{\varepsilon}, \frac{x_\alpha-\hat{y}}{\varepsilon^\beta} + \frac{(x_\alpha-\varepsilon\xi_\alpha)+(x_\alpha-z_\alpha)}{\alpha}\right) \le 0,$$

and

$$\lambda v^\lambda\left(\xi_\alpha, \frac{z_\alpha-\hat{y}}{\varepsilon^\beta}\right) + H\left(\xi_\alpha, \frac{z_\alpha-\hat{y}}{\varepsilon^\beta} + \frac{x_\alpha-\varepsilon\xi_\alpha}{\alpha}\right) \ge 0.$$

Besides, since $\psi(x_\alpha,\xi_\alpha,z_\alpha,t_\alpha) \ge \psi(x_\alpha,\xi_\alpha,x_\alpha,t_\alpha)$,

$$\frac{|x_\alpha-z_\alpha|^2}{2\alpha} \le \varepsilon\left(v^\lambda\left(\xi_\alpha, \frac{x_\alpha-\hat{y}}{\varepsilon^\beta}\right) - v^\lambda\left(\xi_\alpha, \frac{z_\alpha-\hat{y}}{\varepsilon^\beta}\right)\right) \le \varepsilon^{1-\theta-\beta}|x_\alpha-z_\alpha|,$$

which yields $\frac{|x_\alpha-z_\alpha|}{\alpha} \le C\varepsilon^{1-\theta-\beta}$. We now combine this with the two above inequalities on the sub/supersolution tests and let $\alpha \to 0+$ to deduce that

$$K + \frac{\hat{t}-\hat{s}}{\varepsilon^\beta} \le \lambda v^\lambda\left(\frac{\hat{x}}{\varepsilon}, \frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right) + C\varepsilon^{1-\theta-\beta} \le -\overline{H}\left(\frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right) + C\varepsilon^\theta + C\varepsilon^{1-\theta-\beta},$$

and hence,

$$K+\frac{\hat{t}-\hat{s}}{\varepsilon^{\beta}}+\overline{H}\left(\frac{\hat{x}-\hat{y}}{\varepsilon^{\beta}}\right)\leq C\varepsilon^{\theta}+C\varepsilon^{1-\theta-\beta}. \tag{4.53}$$

Next, we use the fact that $(y,s)\mapsto\Phi(\hat{x},\hat{t},y,s)$ has a maximum at $(\hat{y},\hat{s})$, and perform a similar procedure to the above to obtain

$$-K+\frac{\hat{t}-\hat{s}}{\varepsilon^{\beta}}+\overline{H}\left(\frac{\hat{x}-\hat{y}}{\varepsilon^{\beta}}\right)+C\varepsilon^{\theta}+C\varepsilon^{1-\theta-\beta}\geq 0. \tag{4.54}$$

Combine (4.53) and (4.54) to imply

$$2K\leq C(\varepsilon^{\theta}+\varepsilon^{1-\theta-\beta}).$$

Choose $\theta=\beta=\frac{1}{3}$ and $K=K_1\varepsilon^{1/3}$ for K_1 sufficiently large to get a contradiction. Therefore, either $\hat{t}=0$ or $\hat{s}=0$. Then, either $u^{\varepsilon}(\hat{x},\hat{t})=u_0(\hat{x})$ or $u(\hat{y},\hat{s})=u_0(\hat{y})$, and

$$\Phi(\hat{x},\hat{y},\hat{t},\hat{s})\leq u^{\varepsilon}(\hat{x},\hat{t})-u(\hat{y},\hat{s})-\varepsilon v^{\lambda}\left(\frac{\hat{x}}{\varepsilon},\frac{\hat{x}-\hat{y}}{\varepsilon^{\beta}}\right)\leq C\varepsilon^{1/3}.$$

In particular, $\Phi(x,x,t,t)\leq C\varepsilon^{1/3}$, which infers

$$u^{\varepsilon}(x,t)-u(x,t)\leq C\varepsilon^{1/3}+\varepsilon v^{\lambda}\left(\frac{\hat{x}}{\varepsilon},0\right)+2K_1\varepsilon^{1/3}t\leq C(1+T)\varepsilon^{1/3}.$$

By a symmetric argument, we get the desired result. It is worth noting here that the constant C depends on T in a linear way. □

Remark 4.36. Few comments are in order.

1. Firstly, as u^{ε} and u are not smooth enough, it is natural to use the doubling variables method. However, as this is a homogenization problem, one needs to take a corresponding corrector into account and also use the perturbed test function method together with the doubling variables method. Here, we use $\varepsilon v^{\lambda}\left(\frac{x}{\varepsilon},p\right)$ with $\lambda=\varepsilon^{\theta}$ and $p=\frac{x-y}{\varepsilon^{\beta}}$. The choice of this p is suitable with the doubling variables as intuitively speaking
$$p=\frac{x-y}{\varepsilon^{\beta}}=Du^{\varepsilon}(x,t).$$

2. We do not deal directly with the cell problems and their solutions in the proof. The reason is that (4.10) has many solutions in general, and we do not know if we can have a good selection of solution $v(y,p)$ for $y\in\mathbb{T}^n$ and $p\in\mathbb{R}^n$ so that $v(y,p)$ depends on p in a nice way (see also Remark 4.4). We will discuss the nonuniqueness phenomenon in the following section. Instead, we work indirectly with v^{λ} for $\lambda=\varepsilon^{\theta}$, which has good regularity and stability estimates as stated in Lemma 4.35. Of course, as we introduce two new parameters $\theta,\beta\in(0,1)$ in the proof, we need to optimize them, and as the result, we only get rate of convergence $O(\varepsilon^{1/3})$. It seems that this rate $O(\varepsilon^{1/3})$ is not optimal. Nevertheless, this method of Capuzzo-Dolcetta and Ishii is quite general, and it works for various different situations.

3. Based on the formal asymptotic expansion, the optimal rate of convergence should be $O(\varepsilon)$. This is, however, extremely challenging to be obtained. We will discuss this point later.

7.2 An improvement

Next, we show that if we have a bit better understanding of solutions to cell problems, then we have better rate of convergence of our homogenization problem.

$$\begin{cases} \text{For each } p \in \mathbb{R}^n, \text{ we are able to pick a solution } v(y,p) \text{ of (4.10) such that} \\ p \mapsto v(\cdot,p) \quad \text{is Lipschitz.} \end{cases} \tag{4.55}$$

Condition (4.55) is however a very strong and restrictive requirement. We will see that this does not hold in some examples later.

Theorem 4.37. *Assume that $H \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies (4.2) and (4.3). Let $\overline{H}$ be the corresponding effective Hamiltonian of H. Assume further that (4.55) holds. Assume $u_0 \in$ BUC$(\mathbb{R}^n) \cap$ Lip$(\mathbb{R}^n)$. For $\varepsilon > 0$, let u^ε be the unique solution to (4.48). Also let u be the unique solution to (4.49). Then for each $T > 0$, there exists $C > 0$ dependent on H, u_0, and T such that*

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C\varepsilon^{1/2}. \tag{4.56}$$

Proof. Thanks to (4.55), we use directly the correctors in our test function. We consider the auxiliary function

$$\Phi(x,y,t,s) = u^\varepsilon(x,t) - u(y,s) - \varepsilon v\left(\frac{x}{\varepsilon}, \frac{x-y}{\varepsilon^\beta}\right) - \frac{|x-y|^2 + |t-s|^2}{2\varepsilon^\beta} - K(t+s)$$

where $\beta \in (0,1)$ and $K > 0$ to be chosen later. Note that this auxiliary function looks pretty much like that in the proof of Theorem 4.34, but we use v instead of v^λ for $\lambda = \varepsilon^\theta$. This way, we introduce only one parameter $\beta \in (0,1)$ in our auxiliary function instead of two.

Assume that Φ admits a global maximum at $(\hat{x}, \hat{y}, \hat{t}, \hat{s})$ on $\mathbb{R}^{2n} \times [0,T]^2$ for simplicity (for rigorous proof, we need to add the term $-\gamma|x|^2$ to Φ for $\gamma > 0$ (see [27, Theorem 1.1]).

Consider first the case that $\hat{t}, \hat{s} > 0$. By using the fact that $\Phi(\hat{x}, \hat{y}, \hat{t}, \hat{s}) \ge \Phi(\hat{x}, \hat{x}, \hat{t}, \hat{t})$, we deduce that

$$\begin{aligned} \frac{|\hat{x}-\hat{y}|^2 + |\hat{t}-\hat{s}|^2}{2\varepsilon^\beta} &\le \big(u(\hat{x},\hat{t}) - u(\hat{y},\hat{s})\big) + \varepsilon\left(v\left(\frac{\hat{x}}{\varepsilon}, 0\right) - v\left(\frac{\hat{x}}{\varepsilon}, \frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)\right) + K(\hat{t}-\hat{s}) \\ &\le C\big(|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}|\big) + C\varepsilon\frac{|\hat{x}-\hat{y}|}{\varepsilon^\beta} \le C\big(|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}|\big). \end{aligned}$$

Therefore,

$$|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}| \le C\varepsilon^\beta. \tag{4.57}$$

Notice that $(x,t) \mapsto \Phi(x,t,\hat{y},\hat{s})$ has a maximum at $(\hat{x},\hat{t})$. For $\alpha > 0$, set

$$\psi(x,\xi,z,t) = u^\varepsilon(x,t) - \varepsilon v\left(\xi, \frac{z-\hat{y}}{\varepsilon^\beta}\right) - \frac{|x-\hat{y}|^2 + |t-\hat{s}|^2}{2\varepsilon^\beta} - \frac{|x-\varepsilon\xi|^2 + |x-z|^2}{2\alpha} - Kt.$$

Assume ψ has a maximum at $(x_\alpha, \xi_\alpha, z_\alpha, t_\alpha)$ and we can assume by passing to a subsequence if necessary that $(x_\alpha, \xi_\alpha, z_\alpha, t_\alpha) \to (\hat{x}, \hat{x}/\varepsilon, \hat{x}, \hat{t})$ as $\alpha \to 0$.
By using (4.55) and the fact that $\psi(x_\alpha, \xi_\alpha, z_\alpha, t_\alpha) \ge \psi(x_\alpha, \xi_\alpha, x_\alpha, t_\alpha)$,

$$\frac{|x_\alpha - z_\alpha|^2}{2\alpha} \le \varepsilon\left(v\left(\xi_\alpha, \frac{x_\alpha - \hat{y}}{\varepsilon^\beta}\right) - v\left(\xi_\alpha, \frac{z_\alpha - \hat{y}}{\varepsilon^\beta}\right)\right) \le C\varepsilon^{1-\beta}|x_\alpha - z_\alpha|,$$

and hence

$$|x_\alpha - z_\alpha| \le C\alpha\varepsilon^{1-\beta}. \tag{4.58}$$

The same argument for $\psi(x_\alpha, \xi_\alpha, z_\alpha, t_\alpha) \ge \psi(x_\alpha, x_\alpha/\varepsilon, x_\alpha, t_\alpha)$ gives further

$$|x_\alpha - \varepsilon\xi_\alpha| \le C\alpha. \tag{4.59}$$

By definition of viscosity solutions,

$$K + \frac{t_\alpha - \hat{s}}{\varepsilon^\beta} + H\left(\frac{x_\alpha}{\varepsilon}, \frac{x_\alpha - \hat{y}}{\varepsilon^\beta} + \frac{(x_\alpha - \varepsilon\xi_\alpha) + (x_\alpha - z_\alpha)}{\alpha}\right) \le 0, \tag{4.60}$$

and

$$H\left(\xi_\alpha, \frac{z_\alpha - \hat{y}}{\varepsilon^\beta} + \frac{x_\alpha - \varepsilon\xi_\alpha}{\alpha}\right) \ge \overline{H}\left(\frac{z_\alpha - \hat{y}}{\varepsilon^\beta}\right). \tag{4.61}$$

Combining (4.58)–(4.61) and letting $\alpha \to 0$ to yield that

$$K + \frac{\hat{t} - \hat{s}}{\varepsilon^\beta} + \overline{H}\left(\frac{\hat{x} - \hat{y}}{\varepsilon^\beta}\right) - C\varepsilon^{1-\beta} \le 0. \tag{4.62}$$

By a similar procedure,

$$-K + \frac{\hat{t} - \hat{s}}{\varepsilon^\beta} + \overline{H}\left(\frac{\hat{x} - \hat{y}}{\varepsilon^\beta}\right) + C\varepsilon^{1-\beta} \ge 0. \tag{4.63}$$

Putting (4.62) and (4.63) together to get

$$K \le C\varepsilon^{1-\beta}.$$

Choose $\beta = 1/2$ and $K = K_1\varepsilon^{1/2}$ for $K_1 \gg 1$ to get a contradiction.

Thus, either $\hat{t} = 0$ or $\hat{s} = 0$. The proof is hence completed by following the last step in the proof of Theorem 4.34. □

7.3 Problems

Exercise 46. *Let $n = 1$, and $H(y,p) = |p| - V(y)$ for some $V \in C(\mathbb{T})$. Show that* (4.55) *holds in this case.*

Exercise 47. *Let $n = 1$. Is it true that* (4.55) *always holds for H that satisfies* (4.2), (4.3), *and* (4.51)*?*

8 Nonuniqueness of solutions to the cell problems

Let us recall the cell problem (4.10) at a given $p \in \mathbb{R}^n$

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n.$$

We have already shown that $\overline{H}(p)$ is unique, and there exists a solution $v \in \mathrm{Lip}(\mathbb{T}^n)$ to the above. In this section, we discuss the nonuniqueness of v in various situations.

First of all, as already pointed out in Remark 4.4, if v is a solution to the above, then $v + C$ is also a solution for any given $C \in \mathbb{R}$. Thus, (4.10) always has infinitely many viscosity solutions. A natural question then is whether we have uniqueness for solutions to (4.10) up to additive constants or not. In the following assorted collection of examples, we show that nonuniqueness (even up to additive constants) still appears.

Example 4.4. *Assume that*

$$H(y,p) = p\cdot(p - D\varphi(y)) \qquad \text{for all } (y,p)\in\mathbb{T}^n\times\mathbb{R}^n,$$

where $\varphi \in C^1(\mathbb{T}^n)$ is a given function and $\varphi \neq 0$. The cell problem at 0 reads

$$Dv\cdot(Dv - D\varphi) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n. \tag{4.64}$$

It is not hard to see that $\overline{H}(0) = 0$ as $v = C_1$ for a given constant $C_1 \in \mathbb{R}$ is a corresponding solution to the above. Besides, $v = \varphi + C_2$ for each $C_2 \in \mathbb{R}$ is also a classical solution. These are two different families of solutions to (4.64).

Moreover, as H is convex in p, Corollary 2.31 yields further that $v = \min\{C_1, \varphi + C_2\}$ is another solution to (4.64) for each fixed $C_1, C_2 \in \mathbb{R}$. Thus, (4.64) has infinity many solutions of different types. It is worth noting that (4.64) might have other solutions that are not listed here as well.

If H is convex in p, the minimum stability result in Corollary 2.31 allows us to create new solutions out of given solutions as seen above. This means that in general, the structures of solutions to cell problems are very complicated, and it is not easy to characterize all possible solutions even in the convex setting. The problem of characterization of all solutions is of course much harder in the nonconvex settings. Here is another example where we have different families of solutions.

Example 4.5. *Assume that $n = 1$, and*

$$H(y,p) = |p| - V(y) \qquad \text{for all } (y,p)\in\mathbb{T}\times\mathbb{R},$$

where $V \in C(\mathbb{T})$ such that

$$V(y) = 1 - \cos(4\pi y) \qquad \text{for } y \in [0,1].$$

As usual, we identify $\mathbb{T}$ as $[0,1]$. It is clear that $V(y) = 0$ for $y = 0, \frac{1}{2}$.

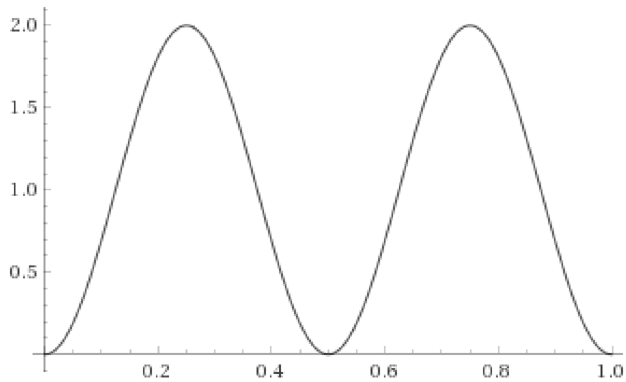


Figure 4.13: Graph of V.

The cell problem at 0 reads

$$|v'| - V(y) = \overline{H}(0) \qquad \text{in } \mathbb{T}. \tag{4.65}$$

We claim that $\overline{H}(0)=0$ *by constructing solutions to* (4.65). *Denote by*

$$v_1(y)=\begin{cases}\int_0^y V(x)\,dx = y-\frac{\sin(4\pi y)}{4\pi} & \text{for } y\in[0,\frac{1}{2}],\\ -\int_0^y V(x)\,dx=-y+\frac{\sin(4\pi y)}{4\pi} & \text{for } y\in[-\frac{1}{2},0].\end{cases}$$

Extend v_1 *to* $\mathbb{R}$ *in the periodic way. It is not hard to check that* v_1 *is* C^1. *It is hence straightforward that* v_1 *is a solution to* (4.65) *with* $\overline{H}(0)=0$.

Besides, set

$$v_2(y)=v_1\left(y+\frac{1}{2}\right)\qquad \text{for all } y\in\mathbb{R}.$$

Since $V(y)=V\left(y+\frac{1}{2}\right)$ *for* $y\in\mathbb{R}$, *we deduce that* v_2 *is also a solution to* (4.65) *with* $\overline{H}(0)=0$. *Finally, because of the convexity of* H *in* p,

$$v_3=\min\{v_1+C_1, v_2+C_2\}$$

solves (4.65) *with* $\overline{H}(0)=0$ *as well for any given* $C_1,C_2\in\mathbb{R}$.

Next is an example of nonuniqueness of the cell problem at $p\neq 0$.

Example 4.6. *We consider the same settings of Example 4.65. Then,* $\langle V\rangle=\int_0^1 V(y)\,dy=1$. *Let us fix* $p\in(0,1)$. *The corresponding cell problem is*

$$|p+v'|-V=\overline{H}(p)\qquad \text{in } \mathbb{T}. \tag{4.66}$$

We claim that $\overline{H}(p)=0$ *by constructing solutions to the above. Pick* $\bar{y}\in(\frac{1}{2},1)$ *such that*

$$p=\int_0^{\bar{y}} V(y)\,dy-\int_{\bar{y}}^1 V(y)\,dy.$$

Denote by

$$v_1(y)=\begin{cases}\int_0^y V(x)\,dx-py & \text{for } y\in[0,\bar{y}],\\ -\int_0^y V(x)\,dx-py & \text{for } y\in[\bar{y}-1,0].\end{cases}$$

Extend v_1 *to* $\mathbb{R}$ *in the periodic way. It is clear that* $v_1\in C^1(\mathbb{T}\setminus\{\bar{y}\})$ *and* v_1 *has a corner from above at* $\bar{y}$. *Therefore,* v_1 *is a viscosity solution to* (4.66) *with* $\overline{H}(p)=0$.

By the same logic as in the previous example,

$$v_2(y)=v_1\left(y+\frac{1}{2}\right)\qquad \text{for all } y\in\mathbb{R}$$

is also a solution to (4.66) *with* $\overline{H}(p)=0$. *Lastly,*

$$v_3=\min\{v_1+C_1, v_2+C_2\}$$

solves (4.66) *as well for any given* $C_1,C_2\in\mathbb{R}$.

We now show that, in the situation in Example 4.65, if $\overline{H}(p)>0=\min_{\mathbb{R}^n}\overline{H}$, then the corresponding cell problem has a unique solution (up to additive constants). This is quite easy to show, and we include it here to have a clear picture of this specific one dimensional convex case.

Proposition 4.38. *Assume that* $n = 1$*, and*

$$H(y,p) = |p| - V(y) \qquad \textit{for all } (y,p) \in \mathbb{T} \times \mathbb{R},$$

where $V \in C(\mathbb{T})$ *such that* $\min_{\mathbb{T}} V = 0$*. Let* $\langle V \rangle = \int_0^1 V(y)\,dy$*. Then,* $\overline{H}$ *has the following formula*

$$\overline{H}(p) = \begin{cases} 0 & \textit{for } |p| \le \langle V \rangle, \\ |p| - \langle V \rangle & \textit{for } |p| \ge \langle V \rangle. \end{cases} \tag{4.67}$$

Moreover, for $|p| \ge \langle V \rangle$*, the corresponding cell problem* (4.10) *has a unique solution (up to additive constants).*

Proof. We skip the proof of the representation formula of $\overline{H}$ in (4.67) as it is quite similar to that of formula (4.24) earlier. We leave it as an exercise.

Fix $p \in \mathbb{R}$ such that $|p| \ge \langle V \rangle$. Without loss of generality, assume $p \ge \langle V \rangle$. The corresponding cell problem is

$$|p + v'(y)| - V(y) = p - \langle V \rangle \qquad \text{in } \mathbb{T}.$$

Of course v is differentiable a.e. and this equation holds also in the a.e. sense. Integrate it over $\mathbb{T}$ and use the usual triangle inequality to yield

$$p = \int_{\mathbb{T}} |p + v'(y)|\,dy \ge \left| \int_{\mathbb{T}} (p + v'(y))\,dy \right| = p.$$

Thus, equality in the above must appear, and therefore, $p + v'(y) \ge 0$ for a.e. $y \in \mathbb{T}$. This allows us to conclude that in fact $p + v'(y) \ge 0$ for all $y \in \mathbb{T}$, and

$$v'(y) = V(y) - \langle V \rangle \qquad \text{in } \mathbb{T}.$$

In particular, v is unique up to additive constants. □

8.1 Problems

Exercise 48. *Give a proof of the representation formula of* $\overline{H}$ *in* (4.67).

Exercise 49. *Assume that* $n = 1$*, and*

$$H(y,p) = |p| - V(y) \qquad \textit{for all } (y,p) \in \mathbb{T} \times \mathbb{R},$$

where $V \in C(\mathbb{T})$ *such that* $\min_{\mathbb{T}} V = 0$*, and* $\{V = 0\} = \{z\}$ *for a given point* $z \in \mathbb{T}$*. Check to see whether the cell problem at each* $p \in \mathbb{R}$ *has a unique solution (up to additive constants) or not.*

9 References

1. Periodic homogenization for Hamilton–Jacobi equations was first studied in the paper of Lions, Papanicolaou, Varadhan [102] circa 1987. The paper is still unpublished, but it has been extremely influential in this area.

2. Evans introduced the perturbed test function method [48] few years later. This method becomes a standard tool for people to use. Basically, the method gives a robust way to use solutions of cell problems to prove qualitative/quantitative homogenization results.

3. Properties of effective Hamiltonians are still not being investigated much. First works along this direction for convex Hamiltonians were done by Concordel [33, 34]. Some of the results are proven by using recent ideas of Mitake and Tran [115]. The inf-sup (inf-max) formula in the convex setting was derived by Contreras, Iturriaga, Paternain, and Paternain [35] and Gomes [72]. The proof we give here is a simple and direct PDE proof.

4. For nonconvex Hamiltonians, there have been some recent interesting developments. In multi dimensions, the decomposition formulas and some further properties of $\overline{H}$, such as evenness, in case $H(y,p) = H(p) - V(y)$ were taken from Qian, Tran, Yu [124]. A special case was done earlier by Armstrong, Tran, Yu [6]. Then, Gao generalized [124] to general non separable Hamiltonians and obtained similar results in [67]. For one dimensional case, shape of $\overline{H}$ is well understood qualitatively by the results of Armstrong, Tran, Yu [7], and Gao [66]. Of course, quantitative and better understandings in one dimensional case are important to be studied in the near future.

5. Further analysis and discussions on the quasiconvexification effect can be found in Qian, Tran, Yu [124]. Much needs to be studied along this direction.

6. The first numerical computation of effective Hamiltonians is due to Qian [123] based on the so called big-T method (see Theorem 4.11). See also Qian, Tran, Yu [124] for some recent examples of $\overline{H}$. For other numerical schemes, we refer to Gomes, Oberman [78], Falcone, Rorro [57], Achdou, Camilli, Capuzzo-Dolcetta [1], Oberman, Takei, Vladimirsky [120], Luo, Yu, Zhao [108] and the references therein.

7. The $O(\varepsilon^{1/3})$ rate of convergence was obtained by Capuzzo-Dolcetta and Ishii [27] about 15 years after the first qualitative homogenization result. The method introduced in [27] is quite standard and robust and is being used a lot for other related problems. Later on, we will give new results on optimal rate of convergence in the convex setting.

CHAPTER 5

Almost periodic homogenization theory for Hamilton–Jacobi equations

1 Introduction to almost periodic homogenization theory

1.1 Introduction

As in Chapter 4, our objects of interests are the same. The equations of interest are as following. For each $\varepsilon > 0$, we study

$$\begin{cases} u^\varepsilon_t(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{5.1}$$

Here, the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is continuous and satisfies some appropriate conditions to be addressed soon. We often assume that the initial data $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ unless otherwise specified. Our goal is to let $\varepsilon \to 0+$ and we hope to see that the homogenization effect happens, that is, u^ε converges to u locally uniformly on $\mathbb{R}^n \times [0,\infty)$, and u solves a (simpler) effective equation

$$\begin{cases} u_t + \overline{H}(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{5.2}$$

To have this in the previous chapter, we assume that $H(y,p)$ is $\mathbb{Z}^n$-periodic in y, and uniformly coercive in p. As we have seen, coercivity of H gives us good uniform Lipschitz estimates on u^ε for all $\varepsilon > 0$, and we will keep this assumption in this chapter. The periodicity of H might be viewed as a bit too restrictive. One might argue that we do see repeated structures in practice, but it is often the case that these repeated structures are not as perfect as the periodic structure. This is often the case in composite materials. For example, we may have that $H(y,p) = |p|^2 + V(y)$, where V is the sum of many functions which are periodic of different periods, that is,

$$V(y) = V_1(y) + V_2(y) + \cdots + V_k(y) \qquad \text{for } y \in \mathbb{R}^n.$$

Here, for $1 \le i \le k$, V_i is $(s_i\mathbb{Z})^n$-periodic where $s_i > 0$ is a given number. In this case, we say that V is quasi periodic.

As such, our goal in this chapter is to study homogenization under a slightly more general assumption that $y \mapsto H(y,p)$ is almost periodic. This was first studied by Ishii [84], and we will follow his approach here to obtain homogenization results. Of course, Ishii's result was for the static case, and we adapt it to the Cauchy problem.

1.2 Derivations

Let us first give a definition of almost periodic function.

Definition 5.1. *Let $f \in \mathrm{BUC}(\mathbb{R}^n)$. We say that f is almost periodic if the family of functions*

$$\{f(\cdot + z) \, : \, z \in \mathbb{R}^n\}$$

is relatively compact in $\mathrm{BUC}(\mathbb{R}^n)$.

Example 5.1. *Let us give few elementary examples of almost periodic functions below.*

1. *If $V \in \mathrm{BUC}(\mathbb{R}^n)$ is $\mathbb{Z}^n$-periodic, then V is also almost periodic. Indeed, for any sequence $\{z_k\} \subset \mathbb{R}^n$, we write $z_k = r_k + s_k$ where $r_k \in \mathbb{Z}^n$ and $s_k \in [0,1)^n$. Then,*

$$V(\cdot + z_k) = V(\cdot + s_k) \qquad \text{for all } k \in \mathbb{N}.$$

Moreover, there exists a subsequence $\{s_{k_j}\}$ of $\{s_k\}$ that converges to $s \in [0,1]^n$ as $j \to \infty$. Thus, as $j \to \infty$,

$$V(\cdot + z_{k_j}) = V(\cdot + s_{k_j}) \to V(\cdot + s) \qquad \text{in } \mathrm{BUC}(\mathbb{R}^n).$$

2. *Assume that V is the sum of finitely many functions which are periodic of different periods, that is,*

$$V(y) = V_1(y) + V_2(y) + \cdots + V_k(y) \qquad \text{for } y \in \mathbb{R}^n.$$

Here, for $1 \le i \le k$, V_i is $(s_i\mathbb{Z})^n$-periodic where $s_i > 0$ is a given number. Then, by using a similar argument as the above one, we also get that V is almost periodic.

Next, to make things precise, we give a definition for almost periodic Hamiltonians.

Definition 5.2. *Let $H = H(y,p) \in C(\mathbb{R}^n \times \mathbb{R}^n)$. We say that H is almost periodic in y if for each $R > 0$, the family of functions*

$$\{H(\cdot + z, \cdot) \, : \, z \in \mathbb{R}^n\}$$

is relatively compact in $\mathrm{BUC}(\mathbb{R}^n \times B(0,R))$.

Basic assumptions. Throughout this chapter, we assume the following two assumptions.

$$H \text{ is almost periodic in } y \text{ in the sense of Definition 5.2,} \tag{5.3}$$

and

$$\lim_{|p| \to \infty} H(y,p) = +\infty \text{ uniformly for } y \in \mathbb{R}^n. \tag{5.4}$$

Example 5.2. *Let $n = 1$, and*

$$H(y,p) = |p| - (2 - \cos y - \cos(\sqrt{2}y)) \qquad \text{for all } (y,p) \in \mathbb{R} \times \mathbb{R}.$$

Then H is coercive in p, and quasi periodic hence almost periodic in y. Surely, H satisifies (5.3)–(5.4). *We will investigate this example further in Theorem 5.7.*

Formally, one can repeat the whole derivations as done in the previous chapter to obtain homogenization results. Let us give a minimalistic recap here. Recall that x is the macroscopic variable, and $y = \frac{x}{\varepsilon}$ is the microscopic variable. A correct ansatz for asymptotic expansion of u^ε around (x,t) is

$$u^\varepsilon(x,t) \approx u(x,t) + \varepsilon v\left(\frac{x}{\varepsilon}\right) = u(x,t) + \varepsilon v(y).$$

It is important noting that $\varepsilon v\left(\frac{x}{\varepsilon}\right)$ is a small perturbation term, and we will need to pay attention to this point later. Let us remark it here that we need

$$\lim_{\varepsilon \to 0} \varepsilon v\left(\frac{x}{\varepsilon}\right) = 0. \tag{5.5}$$

Anyway, plug this expansion to (5.1) to get

$$u_t(x,t) + H(y, Du(x,t) + Dv(y)) = 0.$$

As usual, we assume that x and y are unrelated. Fix $(x,t) \in \mathbb{R}^n \times (0,\infty)$, denote by $p = Du(x,t) \in \mathbb{R}^n$, and $c = -u_t(x,t) \in \mathbb{R}$, we arrive at the usual cell problem

$$H(y, p + Dv(y)) = c \qquad \text{in } \mathbb{R}^n. \tag{5.6}$$

Of course, a key different between this cell problem and the earlier one in the periodic setting is that it is defined in the whole $\mathbb{R}^n$, and in general, it cannot be reduced to the n-dimensional torus. It is not hard to see that (5.5) can be reformulated as

$$\lim_{|y| \to \infty} \frac{v(y)}{|y|} = 0, \tag{5.7}$$

which means that v is sublinear in $\mathbb{R}^n$. Hence, our task is to find $c \in \mathbb{R}$ so that (5.6) has a sublinear viscosity solution v. Formally, if there exists such a unique constant $c \in \mathbb{R}$, we denote by $\overline{H}(p) = c$, and thus, $\overline{H}$ is well-defined. Let us now proceed to identify $\overline{H}$ in a rigorous way.

2 Vanishing discount problems and identification of the effective Hamiltonian

As in the previous chapter, we use the vanishing discount problems to identify $\overline{H}$. Fix $p \in \mathbb{R}^n$. For $\lambda > 0$, consider the following static equation

$$\lambda v^\lambda(y) + H(y, p + Dv^\lambda(y)) = 0 \qquad \text{in } \mathbb{R}^n. \tag{5.8}$$

Our goal is to let $\lambda \to 0+$ to obtain $\overline{H}$. We have first the following proposition.

Proposition 5.3. *Assume* (5.3) *and* (5.4). *Fix* $p \in \mathbb{R}^n$. *For* $\lambda > 0$, *let* v^λ *be the viscosity solution to* (5.8). *Then,*

$$\lim_{\lambda \to 0+} \left(\lambda \sup_{y \in \mathbb{R}^n} |v^\lambda(y) - v^\lambda(0)| \right) = 0. \tag{5.9}$$

Proof. We argue by contradiction. Suppose that there are $\delta > 0$, $\{\lambda_j\} \to 0$, and $\{y_j\} \subset \mathbb{R}^n$ such that

$$\lambda_j |v^{\lambda_j}(y_j) - v^{\lambda_j}(0)| \geq \delta \qquad \text{for all } j \in \mathbb{N}.$$

In light of (5.3), we may assume that there exists a function $G \in C(\mathbb{R}^n \times \mathbb{R}^n)$ such that $H(\cdot + y_j, \cdot) \to G$ uniformly on $\mathbb{R}^n \times \overline{B(0,R)}$ for all $R > 0$.

Besides, set $C = \|H(\cdot, p)\|_{L^\infty(\mathbb{R}^n)}$. Then, $\pm\frac{C}{\lambda}$ are a viscosity supersolution and subsolution to (5.8), respectively. Thus,

$$-\frac{C}{\lambda} \leq v^\lambda \leq \frac{C}{\lambda}.$$

Then, the coercivity of H gives us that $\|Dv^\lambda\|_{L^\infty(\mathbb{R}^n)} \leq C$ for some $C > 0$ independent of $\lambda > 0$. Thus, for $R = C + |p| + 1$, one has $|p| + \|Dv^\lambda\|_{L^\infty(\mathbb{R}^n)} \leq R$, and for $j, k \in \mathbb{N}$ large enough

$$|H(y + y_j, p) - H(y + y_k, p)| \leq \frac{\delta}{4} \qquad \text{for all } y \in \mathbb{R}^n, p \in B(0,R). \tag{5.10}$$

By relabeling $\{y_j\}$ if needed, assume that the above holds for all $j, k \in \mathbb{N}$. For $j \in \mathbb{N}$, denote by

$$w_j(y) = v^{\lambda_j}(y + y_j - y_1) \qquad \text{for all } y \in \mathbb{R}^n.$$

In light of (5.10), for $y \in \mathbb{R}^n$,

$$\lambda_j w_j(y) + H(y, p + Dw_j(y)) \leq \lambda_j v^{\lambda_j}(y + y_j - y_1) + H(y + y_j - y_1, p + Dv^{\lambda_j}(y + y_j - y_1)) + \frac{\delta}{4} = \frac{\delta}{4},$$

and

$$\lambda_j w_j(y) + H(y, p + Dw_j(y)) \geq \lambda_j v^{\lambda_j}(y + y_j - y_1) + H(y + y_j - y_1, p + Dv^{\lambda_j}(y + y_j - y_1)) - \frac{\delta}{4} = -\frac{\delta}{4}.$$

Hence, by the usual comparison principle,

$$\lambda_j w_j(y) - \frac{\delta}{4} \leq \lambda_j v^{\lambda_j}(y) \leq \lambda_j w_j(y) + \frac{\delta}{4} \qquad \text{for } y \in \mathbb{R}^n.$$

Let $y = 0$ in the above to infer

$$\lambda_j |v^{\lambda_j}(y_j - y_1) - v^{\lambda_j}(0)| \leq \frac{\delta}{4}.$$

We then use the Lipschitz bound on v^{λ_j} and the above inequality to imply further

$$\lambda_j |v^{\lambda_j}(y_j) - v^{\lambda_j}(0)| \leq \frac{\delta}{4} + \lambda_j C |y_1| < \frac{\delta}{2},$$

for j sufficiently large. Thus, we get a contradiction. The proof is complete. □

Remark 5.4. It is extremely important for us to get (5.9) in the above proof. One can see clearly that the almost periodic assumption is essentially a compactness assumption that allows us to control nicely the oscillation of λv^λ as $\lambda \to 0$. The proof is of course a proof by contradiction proof, and we have no control on $\{y_j\} \subset \mathbb{R}^n$. In particular, it is unclear if there is any quantitative version of (5.9).

Theorem 5.5. *Assume* (5.3) *and* (5.4). *Fix $p \in \mathbb{R}^n$. There is a unique constant $c \in \mathbb{R}$ such that for each $\delta > 0$, we are able to find a solution $w \in \mathrm{BUC}(\mathbb{R}^n)$ such that w solves*

$$c - \delta \le H(y, p + Dw(y)) \le c + \delta \qquad \text{in } \mathbb{R}^n. \tag{5.11}$$

Proof. We first prove the existence of c. For each $\lambda > 0$, let v^λ be the viscosity solution to (5.8). By the proof of Proposition 5.3, one has $|\lambda v^\lambda(0)| \le C$ and (5.9). Thus, there exist a sequence $\{\lambda_j\} \to 0$ and $c \in \mathbb{R}$ such that

$$\lim_{j\to\infty} \lambda_j v^{\lambda_j}(y) = -c \qquad \text{uniformly for } y \in \mathbb{R}^n.$$

Now, for each $\delta > 0$, pick $j \in \mathbb{N}$ sufficiently large so that $\|\lambda_j v^{\lambda_j} + c\|_{L^\infty(\mathbb{R}^n)} \le \frac{\delta}{2}$. Let $w = v^{\lambda_j}$. It is clear that $w \in \mathrm{BUC}(\mathbb{R}^n)$, and w solves (5.11). The existence of $c \in \mathbb{R}$ is confirmed.

Next, we show the uniqueness of c, which is quite a standard step. Assume otherwise that there exist two such constants $c_1, c_2 \in \mathbb{R}$ with $c_1 < c_2$. Fix $\delta \in (0, \frac{1}{4}(c_2 - c_1))$. There exist $w_1, w_2 \in \mathrm{BUC}(\mathbb{R}^n)$ such that

$$H(y, p + Dw_1(y)) \le c_1 + \delta < c_2 - \delta \le H(y, p + Dw_2(y)) \qquad \text{in } \mathbb{R}^n.$$

As w_1 and w_2 are both bounded, there exists $\lambda > 0$ sufficiently small such that

$$\lambda w_1 + H(y, p + Dw_1) < \frac{c_1 + c_2}{2} < \lambda w_2 + H(y, p + Dw_2) \qquad \text{in } \mathbb{R}^n.$$

By the usual comparison principle, $w_1 \le w_2$. By the same steps, $w_1 + C \le w_2$ for any $C > 0$, which is absurd. Hence, the uniqueness of c is guaranteed. □

Definition 5.6. *Assume* (5.3) *and* (5.4). *For each $p \in \mathbb{R}^n$, let c be the unique constant in Theorem 5.5. Denote by $\overline{H}(p) = c$. For each $\delta > 0$, let $w \in \mathrm{BUC}(\mathbb{R}^n)$ be a solution to* (5.11), *that is, w solves*

$$\overline{H}(p) - \delta \le H(y, p + Dw(y)) \le \overline{H}(p) + \delta \qquad \text{in } \mathbb{R}^n. \tag{5.12}$$

We say that w is a δ-approximate corrector of the cell problem

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{R}^n. \tag{5.13}$$

The definition of $\overline{H}$ is essentially the same as that in the periodic case. However, it is very important noting here that we have not discussed about the correctors, solutions to (5.13). In the above definition, we introduce a new object, δ-approximate correctors, for $\delta > 0$. Although a δ-approximate corrector w does not solve precisely (5.13), it is enough to be employed for arguments with certain room to play with by choosing $\delta > 0$ sufficiently small. Furthermore, $w \in \mathrm{BUC}(\mathbb{R}^n)$, hence is obvious sublinear, that is, w satisfies (5.7).

3 Nonexistence of sublinear correctors

Let us now discuss the correctors, solutions to (5.13). In order for it to be useful, we need to require that correctors satisfy (5.7), that is, they are sublinear. This requirement is clearly needed for us to obtain homogenization result as discussed earlier in the derivations. Furthermore, without sublinearity requirement, the problem might be strange as in the following example.

Example 5.3. *Assume that $H(y,p)=H(p)$, where $H:\mathbb{R}^n\to\mathbb{R}$ is coercive. Let us study* (5.13) *for $p=0$, which is*

$$H(Dv(y))=c \qquad \text{in } \mathbb{R}^n.$$

Then, for any $q\in\mathbb{R}^n$, $v^q(y)=q\cdot y$ for $y\in\mathbb{R}^n$ is a solution to the above with $c=H(q)$. Thus, if we do not require sublinearity of v, then c is not unique.

Of course, among all those v^q, only v^0 is sublinear, and therefore, it is natural to see that the if we put forth the sublinearity assumption, $c=H(0)$ should be the unique constant.

Let us now discuss a simple situation where we cannot expect to have sublinear correctors.

Theorem 5.7. *Assume that $n=1$, and*

$$H(y,p)=|p|-(2-\cos y-\cos(\sqrt{2}y)) \qquad \text{for all } (y,p)\in\mathbb{R}\times\mathbb{R}.$$

Then, $\overline{H}(0)=0$, and (5.13) *for $p=0$ does not admit any sublinear solution.*

Proof. It is clear that H satisfies (5.3) and (5.4). Let us first compute $\overline{H}(0)$. For each $\lambda>0$, we consider

$$\lambda v^\lambda+|Dv^\lambda|-(2-\cos y-\cos(\sqrt{2}y))=0 \qquad \text{in } \mathbb{R}.$$

As the above also holds in the a.e. sense, we imply

$$\lambda v^\lambda(y)\le 2-\cos y-\cos(\sqrt{2}y) \qquad \text{for all } y\in\mathbb{R},$$

and in particular, $\lambda v^\lambda(0)\le 0$. Let $\lambda\to 0+$ to yield that $\overline{H}(0)\ge 0$.

On the other hand, for $\eta>0$, as v^λ is bounded,

$$y\mapsto v^\eta(y)+\eta(|y|^2+1)^{1/2}$$

has a minimum at $y_\eta\in\mathbb{R}$. By the supersolution test,

$$\lambda v^\lambda(y_\eta)\ge -\eta\frac{|y_\eta|}{(|y_\eta|^2+1)^{1/2}}+(2-\cos y_\eta-\cos(\sqrt{2}y_\eta))\ge -\eta.$$

Let $\eta\to 0$, and $\lambda\to 0$ in this order to obtain that $\overline{H}(0)\le 0$. Combine the two inequalities to get $\overline{H}(0)=0$.

Now, let us look at the cell problem at $p=0$

$$|v'(y)|=2-\cos y-\cos(\sqrt{2}y)=:V(y) \qquad \text{for all } y\in\mathbb{R}.$$

This is a convex Hamilton–Jacobi equation. Let v be a viscosity solution to the above. Here, $V\ge 0$ always, and $V(y)=0$ if and only if $y=0$. Therefore, geometrically, the graph of v

cannot have corners from below at points $y \neq 0$. This implies further that the graph of v cannot have more than two corners from above for $y > 0$. In particular, there exists $y_0 \in \mathbb{R}$ such that $v'(y)$ does not change sign for $y > y_0$. That is, either $v'(y) = V(y)$ for all $y > y_0$ or $v'(y) = -V(y)$ for all $y > y_0$. Hence, for $y > \max\{y_0, 1\}$,

$$\frac{|v(y)|}{|y|} \geq \frac{1}{|y|}\left(\int_{y_0}^{y} V(s)\,ds - |v(y_0)|\right) \geq 2 - \frac{C}{|y|},$$

which means that v is not sublinear. □

This result demonstrates that in general, we cannot hope for existence of sublinear correctors, and thus, cannot use them to prove homogenization results. As it turns out, to obtain homogenization, it is enough for us to use approximate correctors.

4 Homogenization for Cauchy problems

Here is our main result.

Theorem 5.8. *Assume that H satisfies* (5.3) *and* (5.4). *Assume $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$. For each $\varepsilon > 0$, let u^ε be the unique viscosity solution of*

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{5.14}$$

Then, as $\varepsilon \to 0$, u^ε converges to u locally uniformly on $\mathbb{R}^n \times [0,\infty)$, and u solves the effective equation

$$\begin{cases} u_t + \overline{H}(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{5.15}$$

We present a proof of this theorem, which is basically a small modification to that of Theorem 4.6. Nevertheless, it is important to present it here for the sake of clarity and completeness.

Proof. We will show later in the next section that $\overline{H}$ is continuous and coercive. Hence, (5.15) has a unique Lipschitz solution u. As far as (5.14) is concerned, we have, as usual, the existence of a constant $C > 0$ independent of $\varepsilon > 0$ such that

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

There exists a subsequence $\{\varepsilon_j\} \to 0$ such that $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$ thanks to the Arzelà–Ascoli theorem. In fact, by abuse of notions, we assume $u^\varepsilon \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$ as $\varepsilon \to 0$. All we need to do to finish the proof is to prove that u solves the effective equation (5.15).

We perform only the subsolution test since the argument for supersolution test is similar. For $\phi \in C^1(\mathbb{R}^n \times (0,\infty))$ such that $u - \phi$ has a global strict max at $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$ with $u(x_0,t_0) = \phi(x_0,t_0)$, we aim at proving

$$\phi_t(x_0,t_0) + \overline{H}(D\phi(x_0,t_0)) \leq 0.$$

Let $p = D\phi(x_0,t_0) \in \mathbb{R}^n$. We prove the above by contradiction. Assume that there exists $\alpha > 0$ such that

$$\phi_t(x_0,t_0) + \overline{H}(p) > \alpha.$$

Let $v \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ be a δ-approximate corrector of (5.12) with this particular p where $\delta = \frac{\alpha}{2}$.

For each $\varepsilon, \eta > 0$ we consider the auxiliary function

$$\begin{aligned}\Phi^{\eta,\varepsilon}(x,y,t) : \mathbb{R}^n \times \mathbb{R}^n \times [0,T] &\to \mathbb{R}\\ (x,y,t) &\mapsto u^\varepsilon(x,t) - \left(\phi(x,t) + \varepsilon v(y) + \frac{\left|y - \frac{x}{\varepsilon}\right|^2}{\eta}\right).\end{aligned}$$

For $\varepsilon > 0$ sufficiently small, it is clear that $\Phi^{\eta,\varepsilon}$ has a max at $\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right) \in B(x_0,r) \times \mathbb{R}^n \times (t_0 - r, t_0 + r)$ for some fixed $r > 0$. As $\eta \to 0$, by compactness $\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) \to (x_\varepsilon, t_\varepsilon)$ up to a subsequence. We claim that $y_{\eta\varepsilon} \to \frac{x_\varepsilon}{\varepsilon}$ as $\eta \to 0$. Since $\Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, \frac{x_{\eta\varepsilon}}{\varepsilon}, t_{\eta\varepsilon}\right) \leq \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$ for all $\eta > 0$, we obtain

$$\frac{1}{\eta}\left|y_{\eta\varepsilon} - \frac{x_{\eta\varepsilon}}{\varepsilon}\right|^2 \leq 2\varepsilon \|v\|_{L^\infty(\mathbb{R}^n)} \qquad \Longrightarrow \qquad \lim_{\eta\to 0} y_{\eta\varepsilon} = \frac{x_\varepsilon}{\varepsilon}. \tag{5.16}$$

As $(x,t) \mapsto \Phi^{\eta,\varepsilon}\left(x, y_{\eta\varepsilon}, t\right)$ has max at $\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$, we imply that $u^\varepsilon - \phi - \frac{1}{\eta}\left|y_{\eta\varepsilon} - \frac{x}{\varepsilon}\right|^2$ has max at $\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$. The subsolution test of (5.14) gives

$$\phi_t\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + H\left(\frac{x_{\eta\varepsilon}}{\varepsilon}, D\phi\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right) \leq 0. \tag{5.17}$$

Next, $y \mapsto \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y, t_{\eta\varepsilon}\right)$ has max at $y_{\eta\varepsilon}$, thus $v(y) - \frac{-1}{\eta\varepsilon}\left|y - \frac{x_{\eta\varepsilon}}{\varepsilon}\right|^2$ has min at $y_{\eta\varepsilon}$, and hence, the supersolution test gives us

$$H\left(y_{\eta\varepsilon}, p + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right) \geq \overline{H}(p) - \delta. \tag{5.18}$$

Besides, as v is Lipschitz, we infer

$$\left|\frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right| \leq C, \tag{5.19}$$

for some $C > 0$ independent of η, ε. By compactness, we can assume (up to passing to a subsequence again) that

$$\lim_{\eta\to 0} \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right) = p_\varepsilon \in \mathbb{R}^n. \tag{5.20}$$

Note that $\Phi^{\eta,\varepsilon}\left(x, \frac{x}{\varepsilon}, t\right) \leq \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$. Let $\eta \to 0$ in this relation and use (5.20) to yield

$$u^\varepsilon(x,t) - \varepsilon v\left(\frac{x}{\varepsilon}\right) - \phi(x,t) \leq u^\varepsilon(x_\varepsilon, t_\varepsilon) - \varepsilon v\left(\frac{x_\varepsilon}{\varepsilon}\right) - \phi(x_\varepsilon, t_\varepsilon)$$

for all $(x,t) \in \mathbb{R}^n \times [0,\infty)$. That means $(x,t) \mapsto u^\varepsilon(x,t) - \varepsilon v\left(\frac{x}{\varepsilon}\right) - \phi(x,t)$ has max at $(x_\varepsilon, t_\varepsilon)$. Again, by passing to a subsequence if needed, $(x_\varepsilon, t_\varepsilon) \to (x_0, t_0)$ as $\varepsilon \to 0$.

Let $\eta \to 0$ in (5.17) and (5.18) to get

$$\phi_t(x_\varepsilon, t_\varepsilon) + H\left(\frac{x_\varepsilon}{\varepsilon}, D\phi(x_\varepsilon, t_\varepsilon) + p_\varepsilon\right) \le 0,$$

and

$$H\left(\frac{x_\varepsilon}{\varepsilon}, p + p_\varepsilon\right) \ge \overline{H}(p) - \delta.$$

Combine the above two and let $\varepsilon \to 0$ to conclude that

$$\phi_t(x_0, t_0) + \overline{H}(p) \le \delta = \frac{\alpha}{2},$$

which is absurd. The proof is complete.

□

Remark 5.9. In the above proof, we use strongly the fact that δ-approximate corrector v is bounded and Lipschitz. Without the boundedness of v, we need to be extremely careful with handling the auxiliary function $\Phi^{\eta,\varepsilon}$ and obtaining (5.16). The Lipschitz estimate of v was used to get (5.19) and (5.20).

5 Properties of the effective Hamiltonians

5.1 Basic properties of $\overline{H}$

We first present the following representation formulas of $\overline{H}$, which is an analog of Theorem 4.10 in the periodic setting.

Theorem 5.10. *Assume that H satisfies* (5.3) *and* (5.4). *Let $\overline{H}$ be its corresponding effective Hamiltonian. Then, for $p \in \mathbb{R}^n$,*

$$\begin{aligned}\overline{H}(p) = &\inf\{c \in \mathbb{R} : \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \le c \textit{ in } \mathbb{R}^n \textit{ in viscosity sense}\} \\ = &\sup\{c \in \mathbb{R} : \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \ge c \textit{ in } \mathbb{R}^n \textit{ in viscosity sense}\}.\end{aligned}$$

One can adapt the proof of Theorem 4.10 to this setting in a natural way. As H is coercive in p, one of the above formulas can also be written as

$$\overline{H}(p) = \inf\{c \in \mathbb{R} : \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n) : H(y, p + Dv(y)) \le c \text{ in } \mathbb{R}^n \text{ in viscosity sense}\}.$$

Proof. Let us define

$$\begin{aligned}\mathcal{A} &= \{c \in \mathbb{R} : \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \le c \text{ in } \mathbb{R}^n \text{ in viscosity sense}\} \\ \mathcal{B} &= \{c \in \mathbb{R} : \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \ge c \text{ in } \mathbb{R}^n \text{ in viscosity sense}\}.\end{aligned}$$

Thanks to Theorem 5.5, we have the existence of δ-approximate correctors for all $\delta > 0$, and hence,

$$\inf \mathcal{A} \le \overline{H}(p) \le \sup \mathcal{B}.$$

Next, we show that $\inf \mathcal{A} = \overline{H}(p)$. The other part follows in an analogous way. Assume by contradiction that $\inf \mathcal{A} < \overline{H}(p)$. Then, there exist some $c_1 \in \mathcal{A}$ and $v_1 \in \mathrm{BUC}(\mathbb{R}^n)$ such that $\inf \mathcal{A} < c_1 < \overline{H}(p)$, while $H(y, p + Dv_1(y)) \le c_1$ in $\mathbb{R}^n$ in the viscosity sense. Let

$\delta = \frac{\overline{H}(p)-c_1}{4} > 0$, and $v \in \mathrm{BUC}(\mathbb{R}^n)$ be a δ-approximate corrector. Since v, v_1 are bounded on $\mathbb{R}^n$, there exists $\lambda > 0$ small enough so that

$$\lambda v_1 + H(y, p + Dv_1(y)) < \frac{c_1 + \overline{H}(p)}{2} < \lambda v + H(y, p + Dv(y)) \qquad \text{in } \mathbb{R}^n.$$

The usual comparison principle implies $v_1 \le v$. By same steps, we obtain that $v_1 \le v - C$ for any constant $C > 0$, which is absurd. Therefore, $\inf \mathcal{A} = \overline{H}(p)$. □

A consequence of Theorem 5.10 is the following.

Corollary 5.11. *Assume that H satisfies* (5.3) *and* (5.4). *Let $\overline{H}$ be its corresponding effective Hamiltonian. Then, for each $p \in \mathbb{R}^n$,*

$$\inf_{y \in \mathbb{R}^n} H(y, p) \le \overline{H}(p) \le \sup_{y \in \mathbb{R}^n} H(y, p).$$

In particular, $\overline{H}$ is coercive.

Proof. Take $\phi \equiv 0$, then ϕ is a classical solution to

$$\inf_{y \in \mathbb{R}^n} H(y, p) \le H(y, p + D\phi) \le \sup_{y \in \mathbb{R}^n} H(y, p) \qquad \text{in } \mathbb{R}^n.$$

We apply Theorem 5.10 to conclude.

□

Theorem 5.12. *Assume that H satisfies* (5.3) *and* (5.4). *Let $\overline{H}$ be its corresponding effective Hamiltonian. Then, $\overline{H}$ is continuous.*

Proof. Fix $R > 0$, and $p, q \in B(0, R)$. For each $\delta \in (0, 1)$, let $w \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ be a δ-approximate corrector of

$$\overline{H}(p) - \delta \le H(y, p + Dw(y)) \le \overline{H}(p) + \delta \qquad \text{in } \mathbb{R}^n.$$

The coercivity of H implies that there exists $C = C(R) > 0$ such that $\|Dw\|_{L^\infty(\mathbb{R}^n)} \le C(R)$. Therefore, by the fact that $H \in \mathrm{BUC}(\mathbb{R}^n \times B(0, R + C(R) + 1))$, there is a modulus of continuity ω_R such that w is also a subsolution to

$$H(y, q + Dw(y)) \le \overline{H}(p) + \delta + \omega_R(|p - q|) \qquad \text{in } \mathbb{R}^n.$$

This implies

$$\overline{H}(q) \le \overline{H}(p) + \delta + \omega_R(|p - q|).$$

Let $\delta \to 0$ and use a symmetric argument to deduce that

$$\left|\overline{H}(p) - \overline{H}(q)\right| \le \omega_R(|p - q|).$$

□

It is clear from the above proof that the following corollary holds.

Corollary 5.13. *Assume that H satisfies* (5.3) *and* (5.4). *Assume further that for each $R > 0$, there exists $C_R > 0$ such that*

$$|H(y,p) - H(y,q)| \le C_R|p-q| \qquad \text{for all } y \in \mathbb{R}^n, p,q \in B(0,R).$$

Let $\overline{H}$ be its corresponding effective Hamiltonian. Then, $\overline{H}$ is locally Lipschitz.

Next is the usual large time average result to compute $\overline{H}(p)$.

Theorem 5.14. *Assume that H satisfies* (5.3) *and* (5.4). *Fix $p \in \mathbb{R}^n$. Consider the following Cauchy problem*

$$\begin{cases} w_t + H(y, p + Dw) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ w(y,0) &= 0 \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{5.21}$$

Let $w(y,t)$ be the unique viscosity solution to (5.21). *Then,*

$$\lim_{t\to\infty} \frac{w(y,t)}{t} = -\overline{H}(p) \qquad \text{uniformly for } y \in \mathbb{R}^n.$$

The proof of this theorem is similar to that of Theorem 4.11 by using δ-approximate correctors (instead of actual correctors). We therefore leave it as an exercise.

5.2 Representation formula of $\overline{H}$ in the convex setting

In this section, we always assume that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{R}^n$.

Theorem 5.15 (The inf-sup formula). *Assume that H satisfies* (5.3) *and* (5.4). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{R}^n$. Then, for fixed $p \in \mathbb{R}^n$, we have*

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{R}^n)\cap \mathrm{BUC}(\mathbb{R}^n)} \sup_{y\in\mathbb{R}^n} H(y, p + D\phi(y)). \tag{5.22}$$

Proof. Pick any $\varphi \in C^1(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n)$, by the representation formula in Theorem 5.10,

$$\overline{H}(p) \le \sup_{y\in\mathbb{R}^n} H(y, p + D\varphi(y)),$$

and hence,

$$\overline{H}(p) \le \inf_{\phi \in C^1(\mathbb{R}^n)\cap \mathrm{BUC}(\mathbb{R}^n)} \sup_{y\in\mathbb{R}^n} H(y, p + D\phi(y)).$$

Conversely, given $\theta > 0$, we aim at proving that

$$\overline{H}(p) + \theta \ge \inf_{\phi \in C^1(\mathbb{R}^n)\cap \mathrm{BUC}(\mathbb{R}^n)} \sup_{y\in\mathbb{R}^n} H(y, p + D\phi(y)).$$

Let $v \in \mathrm{Lip}(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n)$ be a $(\theta/2)$-approximate corrector to (5.12), that is,

$$\overline{H}(p) - \frac{\theta}{2} \le H(y, p + Dv(y)) \le \overline{H}(p) + \frac{\theta}{2} \qquad \text{in } \mathbb{R}^n.$$

It is clear that $\|Dv\|_{L^\infty(\mathbb{R}^n)} \le C$, v is differentiable and solves the above a.e. in $\mathbb{R}^n$. As usual, we smooth v up by using the convolution trick. Take η to be the standard mollifier, that is,

$$\eta \in C_c^\infty(\mathbb{R}^n, [0,\infty)), \qquad \mathrm{supp}(\eta) \subset B(0,1), \qquad \int_{\mathbb{R}^n} \eta(x)\,dx = 1.$$

For $\varepsilon > 0$, denote by $\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(\frac{x}{\varepsilon}\right)$ for all $x \in \mathbb{R}^n$. Set

$$v^\varepsilon(x) = (\eta_\varepsilon \star v)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)v(y)\,dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x-y)v(y)\,dy \quad \text{for } x \in \mathbb{R}^n.$$

Then $v^\varepsilon \in C^\infty(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n)$, and $v^\varepsilon \to v$ uniformly in $\mathbb{R}^n$ as $\varepsilon \to 0$. For every fixed $x \in \mathbb{R}^n$, we compute that

$$\begin{aligned}
\overline{H}(p) + \frac{\theta}{2} &\geq \int_{\mathbb{R}^n} H\Big(x-y, p + Dv(x-y)\Big)\eta_\varepsilon(y)\,dy \\
&\geq \int_{B(0,\varepsilon)} \Big(H\Big(x, p + Dv(x-y)\Big) - \omega(\varepsilon)\Big)\,\eta_\varepsilon(y)\,dy \\
&= \int_{B(0,\varepsilon)} H\Big(x, p + Dv(x-y)\Big)\eta_\varepsilon(y)\,dy - \omega(\varepsilon) \\
&\geq H\left(x, \int_{B(0,\varepsilon)} (p + Dv(x-y))\eta_\varepsilon(y)\,dy\right) - \omega(\varepsilon) = H(x, p + Dv^\varepsilon(x)) - \omega(\varepsilon).
\end{aligned}$$

Thus, v^ε satisfies

$$\sup_{x\in\mathbb{R}^n} H(x, p + Dv^\varepsilon(x)) \leq \overline{H}(p) + \frac{\theta}{2} + \omega(\varepsilon).$$

Pick $\varepsilon > 0$ sufficiently small so that $\omega(\varepsilon) < \frac{\theta}{2}$ to conclude.

□

Here is an immediate consequence of the inf-sup formula above.

Corollary 5.16. *Assume that H satisfies* (5.3) *and* (5.4). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{R}^n$. Then, for each $p \in \mathbb{R}^n$,*

$$\overline{H}(p) = \inf\{c \in \mathbb{R} : \exists\, v \in \mathrm{Lip}(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \leq c \text{ a.e. in } \mathbb{R}^n\}. \quad (5.23)$$

By using the above corollary, we deduce that $\overline{H}$ is also convex.

Theorem 5.17 (Convexity of $\overline{H}$)**.** *Assume that H satisfies* (5.3) *and* (5.4). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{R}^n$. Then, $\overline{H}$ is convex.*

Another immediate consequence of the inf-sup formula is as following.

Corollary 5.18. *Assume that H satisfies* (5.3) *and* (5.4). *Assume further that $p \mapsto H(y,p)$ is convex and even for every $y \in \mathbb{R}^n$. Then, $\overline{H}$ is also even.*

5.3 Problems

Exercise 50. *Give another example of a Hamiltonian H satisfying* (5.3) *and* (5.4) *so that* (5.13) *does not admit a sublinear solution for some $p \in \mathbb{R}^n$.*

Exercise 51. *Give a detailed proof of Theorem 5.14.*

Exercise 52. *Give a quick proof of Theorem 5.17.*

6 References

1. Almost periodic homogenization for Hamilton–Jacobi equations was studied first by Ishii [84].

2. The result on nonexistence of sublinear correctors was pointed out by Lions and Souganidis [104].

3. So far, there has not been any quantitative result on the rate of convergence of u^ε to u in this almost periodic setting. Besides, deeper properties of $\overline{H}$ are not yet explored.

CHAPTER 6

First-order convex Hamilton–Jacobi equations in a torus

In this chapter, we revisit first-order convex Hamilton–Jacobi equations in the flat n-dimensional torus $\mathbb{T}^n$. We always assume that the Hamiltonian $H = H(y,p) \in C(\mathbb{T}^n \times \mathbb{R}^n)$, and

$$\begin{cases} \lim_{|p|\to\infty} \left(\min_{y\in\mathbb{T}^n} H(y,p) \right) = +\infty, \\ p \mapsto H(y,p) \text{ is convex for all } y \in \mathbb{T}^n. \end{cases} \tag{6.1}$$

Later on, further assumptions on the smoothness of H and uniform convexity of H will be put based on topics that we deal with. Our aim here is to study further properties of solutions to the discount problems and the cell problems.

1 New representation formulas for solutions of the discount problems

Fix $\lambda > 0$. The focus of this section is the following discount problem

$$\lambda v^\lambda + H(y, Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n. \tag{6.2}$$

Of course, this equation has been one of the central objects of all previous chapters. In light of (6.1), (6.2) has a unique Lipschitz solution $v^\lambda \in \mathrm{Lip}\,(\mathbb{T}^n)$. Let us recall some estimates on v^λ. First, the comparison principle gives

$$-\max_{y\in\mathbb{T}^n} |H(y,0)| \le \lambda v^\lambda \le \max_{y\in\mathbb{T}^n} |H(y,0)|.$$

Then, the coercivity of H infers the existence of $C > 0$ independent of $\lambda > 0$ such that

$$\|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \le C.$$

Besides, if H is superlinear in p, that is,

$$\lim_{|p|\to\infty} \left(\min_{y\in\mathbb{T}^n} \frac{H(y,p)}{|p|} \right) = +\infty,$$

then v^λ has an optimal control formula based on the Lagrangian $L = L(y, v)$, the Legendre transform of H. For $y \in \mathbb{T}^n$,

$$v^\lambda(y) = \inf\left\{ \int_0^\infty e^{-\lambda s} L(\gamma(s), -\gamma'(s))\, ds \,:\, \gamma \in \mathrm{AC}([0,\infty), \mathbb{T}^n),\, \gamma(0) = y \right\}.$$

We here aim at getting another representation formula for v^λ based on a duality method. We will compare the two formulas later.

1.1 Reduction to optimal control with a compact control set

Before stating the formula for v^λ, let us do some reductions/simplifications first. From the a priori estimates on v^λ, information of $H(y,p)$ for $|p| > C$ does not matter. Let us now provide a modification of H as following.

Pick two constants $h_0, h_1 \in \mathbb{R}$ such that $h_0 < h_1$ and

$$\begin{cases} H(y,p) > h_0 & \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n, \\ H(y,p) < h_1 & \text{for all } (y,p) \in \mathbb{T}^n \times B(0, C+1). \end{cases}$$

Denote by $H_0 : \mathbb{R}^n \to \mathbb{R}$ such that

$$H_0(p) = h_0 + (h_1 - h_0)(|p| - C) \qquad \text{for } p \in \mathbb{R}^n.$$

It is clear that $H_0(p) \le h_0$ for $|p| \le C$, and $H_0(p) \ge h_1$ for $|p| \ge C+1$. Set $\tilde{H} : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ as

$$\tilde{H}(y,p) = \begin{cases} \max\{H(y,p), H_0(p)\} & \text{for } y \in \mathbb{T}^n, |p| \le C+1, \\ H_0(p) & \text{for } y \in \mathbb{T}^n, |p| \ge C+1. \end{cases}$$

Then, $\tilde{H}$ is continuous, convex in p, and $\tilde{H}(y,p) = H(y,p)$ for $|p| \le C$. This means that we can replace H by $\tilde{H}$ in the study of (6.2) without changing anything. The key point of using $\tilde{H}$ is that it has a linear growth rate in p as $|p| \to \infty$. More precisely, for $h = h_1 - h_0 > 0$, we are able to write

$$\tilde{H}(y,p) = \max_{|v| \le h} \big(p \cdot v - \tilde{L}(y,v)\big) \qquad \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{6.3}$$

where $\tilde{L}$ is continuous on $\mathbb{T}^n \times \overline{B}_h$ and is given by

$$\tilde{L}(y,v) = \sup_{p \in \mathbb{R}^n} \big(p \cdot v - \tilde{H}(y,p)\big) \qquad \text{for all } (y,v) \in \mathbb{T}^n \times \overline{B}_h.$$

The point of (6.3) is that we are now in the situation of optimal control with compact control set $\overline{B}_h$, which is convenient to use. Without loss of generality, we now assume H also has this form, that is,

$$H(y,p) = \max_{|v| \le h} (p \cdot v - L(y,v)) \qquad \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{6.4}$$

where $L \in C(\mathbb{T}^n \times \overline{B}_h)$.

1.2 New representation formula

By the reduction step, we may assume H satisfies (6.3) for $L \in C(\mathbb{T}^n \times \overline{B}_h)$ for some fixed $h > 0$ as discussed above. For any $\phi \in C(\mathbb{T}^n \times \overline{B}_h)$, we also denote by

$$H_\phi(y,p) = \max_{|v|\le h}(p\cdot v - \phi(y,v)) \qquad \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n.$$

Of course, H_ϕ satisfies (6.1). Define $\mathcal{F}_\lambda \subset C(\mathbb{T}^n \times \overline{B}_h) \times C(\mathbb{T}^n)$ as

$$\mathcal{F}_\lambda = \left\{(\phi,u) \in C(\mathbb{T}^n \times \overline{B}_h) \times C(\mathbb{T}^n) : u \text{ solves } \lambda u + H_\phi(y,Du) \le 0 \text{ in } \mathbb{T}^n\right\}.$$

Lemma 6.1. *For $\lambda > 0$, the set $\mathcal{F}_\lambda$ is convex.*

For $(z,\lambda) \in \mathbb{T}^n \times (0,\infty)$, we define the evaluation cone $\mathcal{G}_{z,\lambda} \subset C(\mathbb{T}^n \times \overline{B}_h)$ by

$$\mathcal{G}_{z,\lambda} = \{\phi - \lambda u(z) : (\phi,u) \in \mathcal{F}_\lambda\}.$$

Lemma 6.2. *For $(z,\lambda) \in \mathbb{T}^n \times (0,\infty)$, $\mathcal{G}_{z,\lambda}$ is a convex cone in $C(\mathbb{T}^n \times \overline{B}_h)$ with vertex at the origin.*

Denote by $\mathcal{R}$ the space of Radon measures on $\mathbb{T}^n \times \overline{B}_h$, and $\mathcal{P}$ the space of Radon probability measures on $\mathbb{T}^n \times \overline{B}_h$. The Riesz representation theorem ensures us that the dual space of $C(\mathbb{T}^n \times \overline{B}_h)$ identified with $\mathcal{R}$. In this aspect, we write

$$\langle \mu, f\rangle = \int_{\mathbb{T}^n\times\overline{B}_h} f(y,v)\,d\mu(y,v) \qquad \text{for } f \in C(\mathbb{T}^n \times \overline{B}_h), \mu \in \mathcal{R}.$$

Let $\mathcal{G}'_{z,\lambda}$ denote the dual cone of $\mathcal{G}_{z,\lambda}$, that is,

$$\mathcal{G}'_{z,\lambda} = \left\{\mu \in \mathcal{R} : \langle \mu, f\rangle \ge 0 \quad \text{for all } f \in \mathcal{G}_{z,\lambda}\right\}.$$

Let us remark that measures in $\mathcal{G}'_{z,\lambda}$ are nonnegative measures. Indeed, pick any $\mu \in \mathcal{G}'_{z,\lambda}$. For every $\phi \in C(\mathbb{T}^n \times \overline{B}_h)$ such that $\phi \ge 0$, we have $(\phi,0) \in \mathcal{F}_\lambda$, and so, $\langle \mu, \phi\rangle \ge 0$, which gives us that μ is a nonnegative measure.

Here is the new representation formula for v^λ.

Theorem 6.3. *Assume* (6.4) *for some $h > 0$ and $L \in C(\mathbb{T}^n \times \overline{B}_h)$. For $\lambda > 0$, let v^λ be the unique solution to* (6.2). *Then, for $z \in \mathbb{T}^n$,*

$$\lambda v^\lambda(z) = \min_{\mu\in\mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu(y,v). \tag{6.5}$$

Let us now proceed to prove the preparatory lemmas and this theorem. After our preparations in previous chapter, Lemmas 6.1 and 6.2 are not so hard to prove. Nevertheless, let us give complete proofs here.

Proof of Lemma 6.1. Pick $(\phi_1,u_1),(\phi_2,u_2) \in \mathcal{F}_\lambda$. For $i = 1,2$, as H_{ϕ_i} satisfies (6.1), u_i is Lipschitz in $\mathbb{T}^n$. Moreover, in light of Theorem 2.27, $u_i \in \mathrm{Lip}(\mathbb{T}^n)$ is a viscosity solution to

$$\lambda u_i + H_{\phi_i}(y,Du_i) \le 0 \qquad \text{in } \mathbb{T}^n$$

if and only if $u_i \in \mathrm{Lip}(\mathbb{T}^n)$ is an a.e. solution to the above. Thus, for a.e. $y \in \mathbb{T}^n$.

$$\begin{aligned}
&\lambda \frac{u_1(y)+u_2(y)}{2} + H_{\frac{\phi_1+\phi_2}{2}}\left(y, \frac{Du_1(y)+Du_2(y)}{2}\right) \\
&= \lambda \frac{u_1(y)+u_2(y)}{2} + \max_{|v|\le h}\left(\frac{Du_1(y)+Du_2(y)}{2}\cdot v - \frac{\phi_1(y,v)+\phi_2(y,v)}{2}\right) \\
&\le \frac{1}{2}\left((u_1(y) + \max_{|v|\le h}(Du_1(y)\cdot v - \phi_1(y,v)) + (u_2(y) + \max_{|v|\le h}(Du_2(y)\cdot v - \phi_2(y,v)))\right) \le 0.
\end{aligned}$$

Hence, $\left(\frac{\phi_1+\phi_2}{2}, \frac{u_1+u_2}{2}\right) \in \mathcal{F}_\lambda$, which means that $\mathcal{F}_\lambda$ is convex. The proof is complete. □

Next, we show that $\mathcal{G}_{z,\lambda}$ is a convex cone with vertex at the origin.

Proof of Lemma 6.2. First of all, it is clear that $\mathcal{G}_{z,\lambda}$ is a convex set in $C(\mathbb{T}^n \times \overline{B}_h)$ as $\mathcal{F}_\lambda$ is convex by Lemma 6.1.

Next, as $(0,0) \in \mathcal{F}_\lambda$, we infer that $0 \in \mathcal{G}_{z,\lambda}$. Finally, we need to show that $\mathcal{G}_{z,\lambda}$ is a cone. Pick any $(\phi, u) \in \mathcal{F}_\lambda$. It is not hard to see that $s(\phi, u) \in \mathcal{F}_\lambda$ as well for any $s \ge 0$. Thus, if $\phi - \lambda u(z) \in \mathcal{G}_{z,\lambda}$, then $s(\phi - \lambda u(z)) \in \mathcal{G}_{z,\lambda}$ for all $s \ge 0$. The proof is done. □

The convex cone structure of $\mathcal{G}_{z,\lambda}$ is extremely important for us to use later on. We are now ready to prove our main result in this section.

Proof of Theorem 6.3. Firstly, as v^λ is the solution to (6.2), $(L, v^\lambda) \in \mathcal{F}_\lambda$. In particular, $L - \lambda v^\lambda(z) \in \mathcal{G}_{z,\lambda}$. By the definition of the dual cone $\mathcal{G}'_{z,\lambda}$,

$$\langle \mu, L - \lambda v^\lambda(z)\rangle \ge 0 \qquad \text{for all } \mu \in \mathcal{G}'_{z,\lambda},$$

which gives

$$\lambda v^\lambda(z) \le \min_{\mu \in \mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v).$$

To conclude, we need to obtain the converse inequality. We prove this by contradiction. Assume otherwise that there exists $\varepsilon > 0$ such that

$$\lambda v^\lambda(z) + \varepsilon < \min_{\mu \in \mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v). \tag{6.6}$$

Since $\mathcal{G}_{z,\lambda}$ is a convex cone with vertex at the origin, we deduce that

$$\inf_{f \in \mathcal{G}_{z,\lambda}} \langle \mu, f\rangle = \begin{cases} 0 & \text{if } \mu \in \mathcal{P}\cap\mathcal{G}'_{z,\lambda}, \\ -\infty & \text{if } \mu \in \mathcal{P}\setminus\mathcal{G}'_{z,\lambda}. \end{cases}$$

Accordingly,

$$\begin{aligned}
\inf_{\mu \in \mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \langle \mu, L\rangle &= \inf_{\mu\in\mathcal{P}}\left(\langle \mu, L\rangle - \inf_{f\in\mathcal{G}_{z,\lambda}} \langle \mu, f\rangle\right) \\
&= \inf_{\mu\in\mathcal{P}} \sup_{f\in\mathcal{G}_{z,\lambda}} \langle \mu, L - f\rangle.
\end{aligned}$$

Observe that $\mathcal{P}$ is a compact convex subset of $\mathcal{R}$ with topology of weak convergence of measures, and $\mathcal{G}_{z,\lambda}$ is a convex subset of $C(\mathbb{T}^n \times \overline{B}_h)$. Our functional $\mu \mapsto \langle \mu, L - f\rangle$ is continuous and linear on $\mathcal{R}$ with topology of weak convergence of measures for any fixed $f \in C(\mathbb{T}^n \times \overline{B}_h)$, and $f \mapsto \langle \mu, L - f \rangle$ is continuous and affine on $C(\mathbb{T}^n \times \overline{B}_h)$ for any $\mu \in \mathcal{R}$. By Sion's minimax theorem, we are able to interchange the order of infimum and supremum in the above, that is,

$$\inf_{\mu \in \mathcal{P}} \sup_{f \in \mathcal{G}_{z,\lambda}} \langle \mu, L - f \rangle = \sup_{f \in \mathcal{G}_{z,\lambda}} \inf_{\mu \in \mathcal{P}} \langle \mu, L - f \rangle.$$

See Appendix for a proof of Sion's minimax theorem. Combine this with (6.6) to imply that

$$\lambda v^\lambda(z) + \varepsilon < \inf_{\mu \in \mathcal{P}} \langle \mu, L - \phi + \lambda u(z) \rangle$$

for some $(\phi, u) \in \mathcal{F}_\lambda$. Since the Dirac delta measure $\delta_{(y,v)} \in \mathcal{P}$ for each $(y, v) \in \mathbb{T}^n \times \overline{B}_h$, we deduce further that

$$\lambda v^\lambda(z) + \varepsilon < L(y, v) - \phi(y, v) + \lambda u(z) \qquad \text{for all } (y, v) \in \mathbb{T}^n \times \overline{B}_h.$$

Thus, for all $(y, p) \in \mathbb{T}^n \times \mathbb{R}^n$,

$$\begin{aligned} H(y, p) = \sup_{|v| \le h} (p \cdot v - L(y, v)) &\le \sup_{|v| \le h} (p \cdot v - \phi(y, v)) + \lambda(u - v^\lambda)(z) - \varepsilon \\ &= H_\phi(y, p) + \lambda(u - v^\lambda)(z) - \varepsilon. \end{aligned}$$

In particular, we infer that v^λ solves

$$\lambda v^\lambda + H_\phi(y, Dv^\lambda) + \lambda(u - v^\lambda)(z) - \varepsilon \ge 0 \qquad \text{in } \mathbb{T}^n.$$

In other words, $w = v^\lambda + (u - v^\lambda)(z) - \varepsilon/\lambda$ is a supersolution to

$$\lambda w + H_\phi(y, Dw) = 0 \qquad \text{in } \mathbb{T}^n.$$

As u is a subsolution to the above, the comparison principle gives that $w \ge u$. At z, $w(z) \ge u(z)$ implies $-\varepsilon/\lambda > 0$, which is absurd. Therefore,

$$\lambda v^\lambda(z) = \min_{\mu \in \mathcal{P} \cap \mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n \times \overline{B}_h} L(y, v)\, d\mu(y, v).$$

□

Remark 6.4. It is now time to compare this newly obtained formula with the classical optimal control formula. Each one has its own advantages.

On the one hand, the optimal control formula allows us to go further to investigate the optimal paths, which minimize the action functional. But as we deal with paths in $\mathrm{AC}([0, \infty), \mathbb{T}^n)$, we need to be careful with issues related to compactness and stability of these curves. Note further that as

$$\int_0^\infty \lambda e^{-\lambda s}\, ds = 1,$$

we are able to write

$$\int_0^\infty \lambda e^{-\lambda s} L(\gamma(s), -\gamma'(s))\, ds = \int_{\mathbb{T}^n \times \mathbb{R}^n} L(y, v)\, d\mu_\gamma(y, v)$$

for a corresponding probability measure $\mu_\gamma \in \mathcal{P}$.

On the other hand, the new formula (6.7) deals with minimizing the action functional against probability measures in the convex cone $\mathcal{G}'_{z,\lambda}$, which does not give any understanding of the optimal paths. But as $\mathcal{P}$ is a compact convex subset of $\mathcal{R}$ with topology of weak convergence of measures, it is quite convenient to be used when studying compactness and stability problems. We will see this aspect in the next section.

2 New representation formula for the effective Hamiltonian and applications

2.1 New representation formula for $\overline{H}(0)$

We are still interested in studying (6.2). As usual, we assume (6.1). By the reduction step, we may assume that (6.4) holds true.

Let $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ be the unique solution to (6.2), that is,

$$\lambda v^\lambda + H(y, Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n.$$

By Corollary 4.5 (or Lemma 4.52), we know that $\lambda v^\lambda \to -\overline{H}(0)$, and furthermore,

$$\|\lambda v^\lambda + \overline{H}(0)\|_{L^\infty(\mathbb{T}^n)} \le C\lambda,$$

for some constant $C > 0$ independent of $\lambda > 0$. Let us now give a new representation formula for $\overline{H}(0)$ based on the duality method in the previous section.

As it turns out, most of the frameworks in the previous section can be repeated for $\lambda = 0$. Define $\mathcal{F}_0 \subset C(\mathbb{T}^n \times \overline{B}_h) \times C(\mathbb{T}^n)$ as

$$\mathcal{F}_0 = \left\{(\phi, u) \in C(\mathbb{T}^n \times \overline{B}_h) \times C(\mathbb{T}^n) : u \text{ solves } H_\phi(y, Du) \le 0 \text{ in } \mathbb{T}^n\right\}.$$

Then, define the cone $\mathcal{G}_0 \subset C(\mathbb{T}^n \times \overline{B}_h)$ by

$$\mathcal{G}_0 = \{\phi : (\phi, u) \in \mathcal{F}_0\}.$$

The following result is quite straightforward, and we omit its proof.

Lemma 6.5. *The set $\mathcal{F}_0$ is convex. Besides, $\mathcal{G}_0$ is a convex cone in $C(\mathbb{T}^n \times \overline{B}_h)$ with vertex at the origin.*

Let $\mathcal{G}'_0$ denote the dual cone of $\mathcal{G}_0$, that is,

$$\mathcal{G}'_0 = \{\mu \in \mathcal{R} : \langle \mu, f\rangle \ge 0 \quad \text{for all } f \in \mathcal{G}_0\}.$$

By using a same argument as in the previous section, we get that $\mathcal{G}'_0$ contains only nonnegative measures. Here is the new representation formula for $\overline{H}(0)$.

Theorem 6.6. *Assume* (6.4) *for some $h > 0$ and $L \in C(\mathbb{T}^n \times \overline{B}_h)$. Then,*

$$\min_{\mu \in \mathcal{P} \cap \mathcal{G}'_0} \int_{\mathbb{T}^n \times \overline{B}_h} L(y, v)\, d\mu(y, v) = -\overline{H}(0). \tag{6.7}$$

The proof of this is quite similar to that of Theorem 6.3. Let us sketch it here.

Proof. Firstly, let $w \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to the cell problem

$$H(y, Dw) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n. \tag{6.8}$$

Then, $(L + \overline{H}(0), w) \in \mathcal{F}_0$, and $L + \overline{H}(0) \in \mathcal{G}_0$. By the definition of the dual cone $\mathcal{G}_0'$,

$$-\overline{H}(0) \le \min_{\mu \in \mathcal{P} \cap \mathcal{G}_0'} \int_{\mathbb{T}^n \times \overline{B}_h} L(y, v)\, d\mu(y, v).$$

We now prove the converse inequality to conclude by contradiction. Assume otherwise that there exists $\varepsilon > 0$ such that

$$-\overline{H}(0) + \varepsilon < \min_{\mu \in \mathcal{P} \cap \mathcal{G}_0'} \int_{\mathbb{T}^n \times \overline{B}_h} L(y, v)\, d\mu(y, v). \tag{6.9}$$

Since $\mathcal{G}_0$ is a convex cone with vertex at the origin, we deduce that

$$\inf_{f \in \mathcal{G}_0} \langle \mu, f \rangle = \begin{cases} 0 & \text{if } \mu \in \mathcal{P} \cap \mathcal{G}_0', \\ -\infty & \text{if } \mu \in \mathcal{P} \setminus \mathcal{G}_0'. \end{cases}$$

Accordingly,

$$\begin{aligned} \inf_{\mu \in \mathcal{P} \cap \mathcal{G}_0'} \langle \mu, L \rangle &= \inf_{\mu \in \mathcal{P}} \left(\langle \mu, L \rangle - \inf_{f \in \mathcal{G}_0} \langle \mu, f \rangle \right) \\ &= \inf_{\mu \in \mathcal{P}} \sup_{f \in \mathcal{G}_0} \langle \mu, L - f \rangle. \end{aligned}$$

We again apply Sion's minimax theorem to interchange the order of infimum and supremum in the above

$$\inf_{\mu \in \mathcal{P}} \sup_{f \in \mathcal{G}_0} \langle \mu, L - f \rangle = \sup_{f \in \mathcal{G}_0} \inf_{\mu \in \mathcal{P}} \langle \mu, L - f \rangle.$$

Combine this with (6.9) to imply that

$$-\overline{H}(0) + \varepsilon < \inf_{\mu \in \mathcal{P}} \langle \mu, L - \phi \rangle$$

for some $(\phi, u) \in \mathcal{F}_0$. Since the Dirac delta measure $\delta_{(y,v)} \in \mathcal{P}$ for each $(y, v) \in \mathbb{T}^n \times \overline{B}_h$, we deduce further that

$$-\overline{H}(0) + \varepsilon < L(y, v) - \phi(y, v) \qquad \text{for all } (y, v) \in \mathbb{T}^n \times \overline{B}_h.$$

Thus, for all $(y, p) \in \mathbb{T}^n \times \mathbb{R}^n$,

$$H(y, p) = \sup_{|v| \le h} (p \cdot v - L(y, v)) \le \sup_{|v| \le h} (p \cdot v - \phi(y, v)) - \varepsilon = H_\phi(y, p) - \varepsilon.$$

In particular, we infer that

$$H_\phi(y, Dw) \ge \varepsilon > 0 \ge H_\phi(y, Du) \qquad \text{in } \mathbb{T}^n.$$

By the usual trick of adding a small monotone term, we use the comparison principle to imply that $w \geq u$. By the same steps, we obtain as well that $w - C \geq u$ for any $C > 0$, which gives a contradiction. Hence,

$$\min_{\mu \in \mathcal{P} \cap \mathcal{G}_0'} \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v) = -\overline{H}(0).$$

□

We show that measures in $\mathcal{G}_0'$ has a further nice property.

Proposition 6.7. *Let $\mu \in \mathcal{G}_0'$. Then,*

$$\int_{\mathbb{T}^n \times \overline{B}_h} v \cdot D\psi(y)\, d\mu(y,v) = 0 \qquad \text{for all } \psi \in C^2(\mathbb{T}^n). \tag{6.10}$$

Proof. Fix $\psi \in C^2(\mathbb{T}^n)$. Let $\phi(y,v) = v \cdot D\psi(y)$ for $(y,v) \in \mathbb{T}^n \times \overline{B}_h$, then it is clear that $(\phi, \psi) \in \mathcal{F}_0$. It is also clear that $(-\phi, -\psi) \in \mathcal{F}_0$ as well. Therefore, $\pm\phi \in \mathcal{G}_0$, and

$$\langle \mu, \pm\phi \rangle \geq 0,$$

which gives us the conclusion. □

We will see later on that (6.10) essentially says that μ is a holonomic measure. Next we show that we have stability of measures in the cones $\mathcal{G}_{z,\lambda}'$ as $\lambda \to 0$.

Lemma 6.8. *Fix $z \in \mathbb{T}^n$. Let $\{\lambda_j\} \subset (0,\infty)$ be a sequence convergent to 0. For each $j \in \mathbb{N}$, pick $\mu_j \in \mathcal{G}_{z,\lambda_j}'$. Assume that $\mu_j \to \mu$ weakly in the sense of measures for some $\mu \in \mathcal{R}$. Then, $\mu \in \mathcal{G}_0'$.*

Proof. Pick any $(\phi, u) \in \mathcal{F}_0$. Then, $(\phi + \lambda_j u, u) \in \mathcal{F}_{\lambda_j}$, which means that

$$\langle \mu_j, \phi + \lambda_j (u - u(z)) \rangle \geq 0.$$

Thus,

$$\langle \mu, \phi \rangle = \lim_{j\to\infty} \langle \mu_j, \phi \rangle \geq \lim_{j\to\infty} \lambda_j \langle \mu_j, u(z) - u \rangle = 0.$$

Hence, $\mu \in \mathcal{G}_0'$. □

2.2 Applications

We now use the new representation formulas obtained above to study the vanishing discount problem, that is, the asymptotic behavior of v^λ as $\lambda \to 0$. As noted much earlier (see for example Remark 4.4), in general, for fixed $x_0 \in \mathbb{T}^n$, we only have that there exists a subsequence $\{\lambda_j\} \to 0$ such that

$$v^{\lambda_j} - v^{\lambda_j}(x_0) \to w \qquad \text{uniformly in } \mathbb{T}^n,$$

and w is a solution to the cell problem (6.8). As (6.8) often has many solutions as discussed in Chapter 4, it is not clear whether we have the convergence of the whole family $v^\lambda - v^\lambda(x_0)$ as $\lambda \to 0$ or not. This is called a selection problem.

We show that we do have convergence of the whole family of v^λ (after appropriate normalizations) in the convex setting.

Theorem 6.9. *Assume* (6.1). *For $\lambda > 0$, let $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ be the unique solution to* (6.2). *Then, the family $\left\{v^\lambda + \lambda^{-1}\overline{H}(0)\right\}_{\lambda>0}$ is convergent in $C(\mathbb{T}^n)$ as $\lambda \to 0$.*

Proof. By subtracting to a constant from H, we assume first without loss of generality that $\overline{H}(0) = 0$. Again, by the reduction step earlier, we may assume further that H satisfies (6.4) for some $h > 0$ and $L \in C(\mathbb{T}^n \times \overline{B}_h)$.

Since $\overline{H}(0) = 0$, we have that

$$\|v^\lambda\|_{L^\infty(\mathbb{T}^n)} + \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \le C.$$

Let $\mathcal{U}$ be the set of accumulation points in $C(\mathbb{T}^n)$, as $\lambda \to 0$, of $\{v^\lambda\}_{\lambda>0}$. Obviously, $\mathcal{U} \neq \emptyset$. To complete our theorem, we need to show that $\mathcal{U}$ is a singleton. Pick any $u, w \in \mathcal{U}$. We aim at showing that $u(z) \ge w(z)$ for each $z \in \mathbb{T}^n$. There exist $\{\lambda_j\} \to 0$ and $\{\delta_j\} \to 0$ such that $v^{\lambda_j} \to u$ and $v^{\delta_j} \to w$ in $C(\mathbb{T}^n)$ as $j \to \infty$. By Theorem 6.3, we are able to find a sequence of measures $\{\mu_j\} \subset \mathcal{P}$ such that, for $j \in \mathbb{N}$, $\mu_j \in \mathcal{P} \cap \mathcal{G}'_{z,\lambda_j}$, and

$$\lambda_j v^{\lambda_j}(z) = \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu_j(y,v) = \min_{\mu \in \mathcal{P} \cap \mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v).$$

We may assume by passing to a subsequence of $\{\mu_j\}$ that $\mu_j \to \mu$ weakly in the sense of measures for some $\mu \in \mathcal{P}$. By Lemma 6.8, $\mu \in \mathcal{G}'_0$. Let $j \to \infty$ in the above to obtain

$$0 = \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v) = \min_{\mu \in \mathcal{P} \cap \mathcal{G}'_0} \int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v).$$

Next, we combine $(L - \delta_j v^{\delta_j}, v^{\delta_j}) \in \mathcal{F}_0$ and $(L + \lambda_j w, w) \in \mathcal{F}_{\lambda_j}$ with the above identities to yield

$$0 \le \langle \mu, L - \delta_j v^{\delta_j} \rangle = -\delta_j \langle \mu, v^{\delta_j} \rangle,$$

and

$$0 \le \langle \mu_j, L + \lambda_j w - \lambda_j w(z) \rangle = \lambda_j (v^{\lambda_j}(z) - w(z)) + \lambda_j \langle \mu_j, w \rangle.$$

Therefore,

$$\langle \mu, v^{\delta_j} \rangle \le 0 \quad \text{and} \quad v^{\lambda_j}(z) - w(z) + \langle \mu_j, w \rangle \ge 0.$$

Let $j \to \infty$ to deduce further that

$$\langle \mu, w \rangle \le 0 \quad \text{and} \quad u(z) - w(z) + \langle \mu, w \rangle \ge 0,$$

which implies $u(z) \ge w(z)$. The proof is complete. □

In fact, we are able to characterize the limit of $v^\lambda + \lambda^{-1}\overline{H}(0)$ as $\lambda \to 0$ as well. We provide here another version of Theorem 6.9 with this characterization, which might be helpful for further analysis later. Denote by $\mathcal{M}_0$ the set of all measures $\mu \in \mathcal{P} \cap \mathcal{G}'_0$ such that

$$\int_{\mathbb{T}^n \times \overline{B}_h} L(y,v)\, d\mu(y,v) = -\overline{H}(0).$$

We say that $\mathcal{M}_0$ is the set of minimizing measures corresponding to the cell problem (6.8).

Theorem 6.10. *Assume* (6.1). *For $\lambda > 0$, let $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ be the unique solution to* (6.2). *Then, the family $\left\{v^\lambda + \lambda^{-1}\overline{H}(0)\right\}_{\lambda>0}$ is convergent in $C(\mathbb{T}^n)$ as $\lambda \to 0$ to v^0, where*

$$v^0 = \sup_{v\in\mathcal{E}} v.$$

Here, $\mathcal{E}$ denotes the family of subsolutions v to the cell problem (6.8) *such that*

$$\langle \mu, v\rangle \le 0 \qquad \text{for all } \mu \in \mathcal{M}_0.$$

It is quite interesting to notice that although v^0 is a subsolution to (6.8) by stability of viscosity solutions, it is not clear at all from the definition whether v^0 is a solution to (6.8) or not. This nice and subtle point is included the the proof of this theorem, which shares a same philosophy as that of Theorem 6.9. We give a complete proof of Theorem 6.10 here as we believe that it gives another viewpoint of this vanishing discount problem.

Proof. By subtracting to a constant from H, we assume first without loss of generality that $\overline{H}(0) = 0$. Again, by the reduction step earlier, we may assume further that H satisfies (6.4) for some $h > 0$ and $L \in C(\mathbb{T}^n \times \overline{B}_h)$.

Since $\overline{H}(0) = 0$, we have that

$$\|v^\lambda\|_{L^\infty(\mathbb{T}^n)} + \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \le C.$$

Let $\mathcal{U}$ be the set of accumulation points in $C(\mathbb{T}^n)$, as $\lambda \to 0$, of $\{v^\lambda\}_{\lambda>0}$. Obviously, $\mathcal{U} \neq \emptyset$. We aim at showing that $\mathcal{U} = \{v^0\}$ to conclude.

Pick any $u \in \mathcal{U}$. There exist $\{\lambda_j\} \to 0$ such that $v^{\lambda_j} \to u$. We first show that $u \le v^0$. Indeed, $(L - \lambda_j v^{\lambda_j}, v^{\lambda_j}) \in \mathcal{F}_0$ for all $j \in \mathbb{N}$. For every $\mu \in \mathcal{M}_0$, we use the definition of $\mathcal{M}_0$ to imply

$$0 \le \langle \mu, L - \lambda_j v^{\lambda_j}\rangle = -\lambda_j \langle \mu, v^{\lambda_j}\rangle.$$

Thus, $\langle \mu, v^{\lambda_j}\rangle \le 0$. Let $j \to \infty$ to get that $\langle \mu, u\rangle \le 0$ for all $\mu \in \mathcal{M}_0$. Therefore, $u \le v^0$.

Next, we show that $u \ge v^0$ by showing that $u \ge v$ for any $v \in \mathcal{E}$. Fix $z \in \mathbb{T}^n$. By Theorem 6.3, we are able to find a sequence of measures $\{\mu_j\} \subset \mathcal{P}$ such that, for $j \in \mathbb{N}$, $\mu_j \in \mathcal{P} \cap \mathcal{G}'_{z,\lambda_j}$, and

$$\lambda_j v^{\lambda_j}(z) = \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu_j(y,v) = \min_{\mu\in\mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu(y,v).$$

We may assume by passing to a subsequence of $\{\mu_j\}$ that $\mu_j \to \mu_0$ weakly in the sense of measures for some $\mu_0 \in \mathcal{P}$. By Lemma 6.8, $\mu_0 \in \mathcal{G}'_0$. Let $j \to \infty$ in the above to obtain

$$0 = \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu_0(y,v) = \min_{\mu\in\mathcal{P}\cap\mathcal{G}'_0} \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu(y,v).$$

Thus, $\mu_0 \in \mathcal{M}_0$. Next, we combine $(L + \lambda_j v, v) \in \mathcal{F}_{\lambda_j}$ with the above identities to yield

$$0 \le \langle \mu_j, L + \lambda_j v - \lambda_j v(z)\rangle = \lambda_j(v^{\lambda_j}(z) - v(z)) + \lambda_j\langle \mu_j, v\rangle.$$

Therefore,

$$v^{\lambda_j}(z) - v(z) + \langle \mu_j, v\rangle \ge 0.$$

Let $j \to \infty$ to deduce further that

$$u(z) \geq v(z) - \langle \mu_0, v \rangle \geq v(z).$$

We use the fact that $\langle \mu_0, v \rangle \leq 0$ in the last inequality above as $v \in \mathcal{E}$ and $\mu_0 \in \mathcal{M}_0$. Thus, $u \geq v$ for any $v \in \mathcal{E}$, which gives further that $u \geq v^0$. The proof is complete. □

Although Theorem 6.9 and Theorem 6.10 give the same convergence result, they are quite different their approaches and each has its own advantages. In particular, the proof of Theorem 6.9 is simpler in a way, and there is no need of using the minimizing measures $\mathcal{M}_0$. The proof of Theorem 6.10 is a bit more complicated (and seemingly ad hoc), but it gives a nice characterization of the limit v^0. In practice, depending on the situations, one can be flexible in using either one of these two theorems.

2.3 Problems

Exercise 53. *Formulate and give a proof for an analogous result to Theorem 6.9 (or Theorem 6.10) for the family* $\left\{v^\lambda - v^\lambda(0)\right\}_{\lambda>0}$ *in place of* $\left\{v^\lambda + \lambda^{-1}\overline{H}(0)\right\}_{\lambda>0}$.

3 Cell problems, backward characteristics, and applications

We recall the cell problems of interests here. For each $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to the cell problem (4.10), that is,

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n. \tag{6.11}$$

Whenever needed, we write $v = v_p$ or $v = v(\cdot, p)$ to demonstrate clear dependence on p. We aim at studying backward characteristics of solutions to (6.11).

In this section, we assume a stronger condition that

$$\begin{cases} H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \leq D^2_{pp}H(y,p) \leq \theta^{-1} I_n \text{ for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{6.12}$$

Here, I_n is the identity matrix of size n. We say that H is C^2 and is uniformly convex in p. Let $L = L(y, v)$ be the usual Lagrangian. Then, $L \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$ and L is also uniformly convex in v.

3.1 Backward characteristics

Here is our result on backward characteristics.

Theorem 6.11. *Assume* (6.12)*. For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (6.11)*. Then, for every $x \in \mathbb{T}^n$, there exists a C^1 curve $\xi : (-\infty, 0] \to \mathbb{R}^n$ such that $\xi(0) = x$, and*

$$p \cdot \xi(t_1) + v(\xi(t_1)) - p \cdot \xi(t_2) - v(\xi(t_2)) = \int_{t_2}^{t_1} \left(L(\xi(t), \xi'(t)) + \overline{H}(p)\right) dt \tag{6.13}$$

for all $t_2 < t_1 \leq 0$.

We say that ξ is a backward characteristic of v starting from x.

Proof. For simplicity of notions, let us assume $p = 0$.

We consider the following Cauchy problem

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

The unique solution to the above is $u(y, t) = v(y) - \overline{H}(0)t$ for $(y, t) \in \mathbb{R}^n \times [0, \infty)$.

We construct ξ by on $[-k, -k+1]$ iteratively for $k \in \mathbb{N}$ as following. Of course, we are given that $\xi(0) = x$. For $k \in \mathbb{N}$, by the optimal control formula,

$$u(\xi(-k+1), 1) = \inf\left\{ \int_0^1 L(\gamma(s), \gamma'(s))\, ds + v(\gamma(0)) \,:\, \gamma \in \mathrm{AC}([0,1], \mathbb{R}^n), \gamma(1) = \xi(-k+1) \right\}.$$

Since L is C^2 and is uniformly convex in v, there exists a C^1 minimizer $\eta \in C^1([0,1], \mathbb{R}^n)$ with $\eta(1) = \xi(-k+1)$ to the above. See Appendix for a detailed proof of this point. Denote by

$$\xi(-k+s) = \eta(s) \qquad \text{for } s \in [0,1].$$

By this iteration, we get that ξ is defined on $(-\infty, 0]$, $\xi(0) = x$. It is clear that ξ is C^1, and $\|\xi'\|_{L^\infty((-\infty,0])} \leq C$. Furthermore, by the Dynamic Programming Principle,

$$v(\xi(-k+1)) - \overline{H}(0) = \int_s^1 L(\xi(r), \xi'(r))\, dr + v(\xi(-k+s)) - \overline{H}(0)s \qquad \text{for all } k \in \mathbb{N},\, s \in [0,1].$$

Thus, for all $t_2 < t_1 \leq 0$,

$$v(\xi(t_1)) - v(\xi(t_2)) = \int_{t_2}^{t_1} \left(L(\xi(t), \xi'(t)) + \overline{H}(0) \right) dt.$$

□

3.2 Problems

Exercise 54. *Give another proof of Theorem 6.11 by constructing optimal paths $\xi_k : [-k, 0] \to \mathbb{R}^n$ with $\xi_k(0) = x$ to the Cauchy problem for $k \in \mathbb{N}$. Then, use compactness of $\{\xi_k\}$ and a diagonal argument to pass to the limit to get a backward characteristic.*

3.3 Large time average of backward characteristics

We are now concerned with the behavior of $\frac{\xi(t)}{t}$ as $t \to -\infty$, where ξ is a backward characteristic of v, solution to (6.11).

Theorem 6.12. *Assume* (6.12). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (6.11). *Fix $x \in \mathbb{T}^n$, and let ξ be a backward characteristic of v starting from x. Then, there exist a subsequence $\{t_k\} \to -\infty$ and a vector $q \in D^-\overline{H}(p)$ such that*

$$\lim_{k\to\infty} \frac{\xi(t_k)}{t_k} = q \in D^-\overline{H}(p).$$

We need to do some preparations before proving this theorem. But let us give a quick comment first. As H satisfies (6.12), we have that $\overline{H}$ is convex and coercive. Therefore, for each $p \in \mathbb{R}^n$, $D^-\overline{H}(p) \neq \emptyset$. Of course, if $\overline{H}$ is differentiable at p, then $D^-\overline{H}(p) = \{D\overline{H}(p)\}$, and we have the following direct consequence of the above theorem.

Corollary 6.13. *Assume* (6.12). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (6.11). *Assume further that $\overline{H}$ is differentiable at p. Fix $x \in \mathbb{T}^n$, and let ξ be a backward characteristic of v starting from x. Then,*

$$\lim_{t \to -\infty} \frac{\xi(t)}{t} = D\overline{H}(p).$$

The following is an important lemma toward proving Theorem 6.12.

Lemma 6.14. *Assume* (6.12). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (6.11). *Let $\gamma : (-\infty, 0] \to \mathbb{R}^n$ be an arbitrary Lipschitz curve. Then, for every $T > 0$,*

$$\int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + \overline{H}(p) \right) dt \geq p \cdot (\gamma(0) - \gamma(-T)) + v(\gamma(0)) - v(\gamma(-T)).$$

Heuristically, if everything is smooth, then this result is not hard to prove. Indeed,

$$\begin{aligned}
&\int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + \overline{H}(p) \right) dt = \int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + H(\gamma(t), p + Dv(\gamma(t))) \right) dt \\
\geq &\int_{-T}^{0} \gamma'(t) \cdot (p + Dv(\gamma(t)))\, dt = p \cdot (\gamma(0) - \gamma(-T)) + v(\gamma(0)) - v(\gamma(-T)).
\end{aligned}$$

Of course, as v is only Lipschitz, we need to be careful. As usual, to overcome this difficulty, we perform a convolution trick to smooth v up.

Proof. Take η to be the standard mollifier, that is,

$$\eta \in C_c^\infty(\mathbb{R}^n, [0, \infty)), \qquad \mathrm{supp}(\eta) \subset B(0,1), \qquad \int_{\mathbb{R}^n} \eta(x)\, dx = 1.$$

For $\varepsilon > 0$, denote by $\eta_\varepsilon(x) = \varepsilon^{-n} \eta\left(\frac{x}{\varepsilon}\right)$ for all $x \in \mathbb{R}^n$. Set

$$v^\varepsilon(x) = (\eta_\varepsilon \star v)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x - y) v(y)\, dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x - y) v(y)\, dy \quad \text{for } x \in \mathbb{R}^n.$$

Then $v^\varepsilon \in C^\infty(\mathbb{T}^n)$, and $v^\varepsilon \to v$ uniformly in $\mathbb{T}^n$ as $\varepsilon \to 0$. As $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, by repeating the proof of Theorem 2.27, we infer that v^ε satisfies

$$H(y, p + Dv^\varepsilon(y)) \leq \overline{H}(p) + C\varepsilon \qquad \text{in } \mathbb{T}^n.$$

We can now perform a similar computation as the heuristic one above

$$\begin{aligned}
&\int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + \overline{H}(p) \right) dt \geq \int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + H(\gamma(t), p + Dv^\varepsilon(\gamma(t)) - C\varepsilon \right) dt \\
\geq & -CT\varepsilon + \int_{-T}^{0} \gamma'(t) \cdot (p + Dv^\varepsilon(\gamma(t))\, dt \\
= & -CT\varepsilon + p \cdot (\gamma(0) - \gamma(-T)) + v^\varepsilon(\gamma(0)) - v^\varepsilon(\gamma(-T)).
\end{aligned}$$

Let $\varepsilon \to 0$ in the above to conclude. $\square$

Remark 6.15. In fact, Lemma 6.14 holds if we only require that $v \in \mathrm{Lip}(\mathbb{T}^n)$ to be a subsolution to (6.11) instead of a solution. This can be seen directly from the proof above as we only use the subsolution property.

We utilize the above lemma to prove Theorem 6.12.

Proof of Theorem 6.12. To make it clear, we write v_p to denote a solution to (6.11).

As ξ is a backward characteristic of $v = v_p$ starting from x, for every $t < 0$,

$$p \cdot (\xi(0) - \xi(t)) + v_p(\xi(0)) - v_p(\xi(t)) = \int_t^0 \left(L(\xi(s), \xi'(s)) + \overline{H}(p) \right) ds.$$

On the other hand, for any $\tilde{p} \in \mathbb{R}^n$, let $v_{\tilde{p}} \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to the corresponding cell problem with $\min_{\mathbb{T}^n} v_{\tilde{p}} = 0$. Lemma 6.14 gives that

$$\tilde{p} \cdot (\xi(0) - \xi(t)) + v_{\tilde{p}}(\xi(0)) - v_{\tilde{p}}(\xi(t)) \leq \int_t^0 \left(L(\xi(s), \xi'(s)) + \overline{H}(\tilde{p}) \right) ds$$

Thus, for $\tilde{p} \in B(p, 1)$,

$$\overline{H}(\tilde{p}) - \overline{H}(p) \geq (\tilde{p} - p) \cdot \frac{\xi(t) - \xi(0)}{t} - \frac{C}{|t|}. \tag{6.14}$$

Besides, the fact that $\|\xi'\|_{L^\infty((-\infty,0])} \leq C$ implies

$$\left| \frac{\xi(t) - \xi(0)}{t} \right| \leq C \qquad \text{for all } t < 0.$$

Therefore, there exists a sequence $\{t_k\} \to -\infty$ such that $\frac{\xi(t_k)}{t_k} \to q \in \mathbb{R}^n$ as $k \to \infty$ with $|q| \leq C$. Plug this into (6.14) to yield

$$\overline{H}(\tilde{p}) - \overline{H}(p) \geq (\tilde{p} - p) \cdot q \qquad \text{for all } \tilde{p} \in B(p, 1),$$

which means that $q \in D^- \overline{H}(p)$.

□

Remark 6.16. Of course, the above proof is a qualitative proof based on a compactness argument. It is not clear at this moment if $\overline{H}$ is not differentiable at p, that is, $D^- \overline{H}(p)$ is not a singleton, then whether one can find two different sequences $\{t_k\} \to -\infty$ and $\{s_k\} \to -\infty$ such that

$$\lim_{k \to \infty} \frac{\xi(t_k)}{t_k} = q_1 \neq q_2 = \lim_{k \to \infty} \frac{\xi(s_k)}{s_k}$$

or not.

It is surely important to quantify, if possible, the rate of convergence of $\frac{\xi(t)}{t}$ to $D\overline{H}(p)$ as $t \to -\infty$ in case that $\overline{H}$ is differentiable at p. In general, this is not a simple question as we do not have much information about $\overline{H}$ as discussed earlier in previous chapters.

As $\overline{H}$ is convex, it is twice differentiable almost everywhere, thanks to Alexandrov's theorem. It turns out that if $\overline{H}$ is twice differentiable at p, then we are able to obtain a rate of convergence $O(|t|^{-1/2})$ of $\frac{\xi(t)}{t}$ to $D\overline{H}(p)$ as $t \to -\infty$. Here is a precise statement.

Theorem 6.17. *Assume* (6.12)*. Fix $p \in \mathbb{R}^n$, and assume $\overline{H}$ is twice differentiable at this p. Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (6.11)*. Fix $x \in \mathbb{T}^n$, and let ξ be a backward characteristic of v starting from x. Then, there exists a constant $C = C(p) > 0$ depending on $H, \overline{H}, p$ such that*

$$\left| \frac{\xi(t)}{t} - D\overline{H}(p) \right| \le \frac{C}{|t|^{1/2}} \qquad \text{for all } t < 0.$$

Proof. This is essentially a quantitative version of Theorem 6.12. It is enough to prove the result for $t < -1$. Let

$$w = \frac{\xi(t) - \xi(0)}{t} - D\overline{H}(p).$$

Recall that we have (6.14), that is, for $\tilde{p} \in B(p,1)$,

$$\overline{H}(\tilde{p}) - \overline{H}(p) \ge (\tilde{p} - p) \cdot \frac{\xi(t) - \xi(0)}{t} - \frac{C}{|t|}.$$

Since $\overline{H}$ is twice differentiable at p, there is a constant $C = C(p) > 0$ such that, for $\tilde{p} \in B(p,1)$,

$$\overline{H}(\tilde{p}) \le \overline{H}(p) + D\overline{H}(p) \cdot (\tilde{p} - p) + C|\tilde{p} - p|^2.$$

Combine the two inequalities to deduce that, for $\tilde{p} \in B(p,1)$,

$$C|\tilde{p} - p|^2 \ge (\tilde{p} - p) \cdot \left(\frac{\xi(t) - \xi(0)}{t} - D\overline{H}(p) \right) - \frac{C}{|t|}.$$

If $w = 0$, then there is nothing to prove. Else, choose $\tilde{p} = p + \frac{1}{|t|^{1/2}} \frac{w}{|w|}$ to conclude. $\square$

Here is an immediate corollary.

Corollary 6.18. *Assume* (6.12)*. Fix $p \in \mathbb{R}^n$, and assume $\overline{H}$ is linear in a neighborhood of p. Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (6.11)*. Fix $x \in \mathbb{T}^n$, and let ξ be a backward characteristic of v starting from x. Then, there exists a constant $C = C(p) > 0$ depending on $H, \overline{H}, p$ such that*

$$\left| \frac{\xi(t)}{t} - D\overline{H}(p) \right| \le \frac{C}{|t|} \qquad \text{for all } t < 0.$$

Proof. It is enough to prove the result for $t < -1$. Let

$$w = \frac{\xi(t) - \xi(0)}{t} - D\overline{H}(p).$$

Again, for $\tilde{p} \in B(p,1)$,

$$\overline{H}(\tilde{p}) - \overline{H}(p) \ge (\tilde{p} - p) \cdot \frac{\xi(t) - \xi(0)}{t} - \frac{C}{|t|}.$$

Since $\overline{H}$ is linear in a neighborhood of p, we can find $r \in (0,1)$ so that, for $\tilde{p} \in B(p,r)$,

$$\overline{H}(\tilde{p}) - \overline{H}(p) = D\overline{H}(p) \cdot (\tilde{p} - p).$$

Combine the two above to infer that, for $\tilde{p} \in B(p,r)$,

$$\frac{C}{|t|} \ge (\tilde{p} - p) \cdot \left(\frac{\xi(t) - \xi(0)}{t} - D\overline{H}(p) \right).$$

If $w = 0$, then there is nothing to prove. Otherwise, pick $\tilde{p} = p + r\frac{w}{|w|}$ to finish the proof. $\square$

4 Optimal rate of convergence in periodic homogenization theory

We now apply what we just developed to study the rate of convergence problem in periodic homogenization theory under an additional assumption that H is convex in p. It is enough to assume (6.1) here. Nevertheless, for simplicity, we assume that H satisfies (6.12) in this section. Let us recall quickly the homogenization problem.

For each $\varepsilon > 0$, we study

$$\begin{cases} u^\varepsilon_t(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{6.15}$$

We often assume that the initial data $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ unless otherwise specified. Our goal is to let $\varepsilon \to 0+$ and quantify the rate of convergence of u^ε to u, which solves a (simpler) effective equation

$$\begin{cases} u_t + \overline{H}(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{6.16}$$

4.1 The general case

Here is the main result of this section.

Theorem 6.19. *Assume* (6.12) *and* $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$. *For* $\varepsilon > 0$, *let* u^ε *be the viscosity solution to* (6.15). *Let* u *be the viscosity solution to* (6.16). *Then, there exists a constant* $C > 0$ *dependent only on* H *and* $\|Du_0\|_{L^\infty(\mathbb{R}^n)}$ *such that the following claims hold.*

(i) The lower bound is always optimal, that is,

$$u^\varepsilon(x,t) \ge u(x,t) - C\varepsilon \qquad \text{for all } (x,t) \in \mathbb{R}^n \times [0,\infty). \tag{6.17}$$

(ii) For fixed $(x,t) \in \mathbb{R}^n \times (0,\infty)$, *if* u *is differentiable at* (x,t) *and* $\overline{H}$ *is twice differentiable at* $p = Du(x,t)$, *then*

$$u^\varepsilon(x,t) \le u(x,t) + C_p\sqrt{t\varepsilon} + C\varepsilon. \tag{6.18}$$

Here $C_p > 0$ *is a constant depending on* $H, \overline{H}, p$ *and* $\|Du_0\|_{L^\infty(\mathbb{R}^n)}$.

If we further assume that the initial data $u_0 \in C^2(\mathbb{R}^n)$ *with* $\|u_0\|_{C^2(\mathbb{R}^n)} < \infty$, *then*

$$u^\varepsilon(x,t) \le u(x,t) + \widetilde{C}_p t\varepsilon + C\varepsilon. \tag{6.19}$$

Here $\widetilde{C}_p$ *is a constant depending on* $H, \overline{H}, p$ *and* $\|u_0\|_{C^2(\mathbb{R}^n)}$.

It is worth noting that if $u_0 \in C^2(\mathbb{R}^n)$ with $\|u_0\|_{C^2(\mathbb{R}^n)} < \infty$, then the upper bound in the theorem is only conditionally optimal. As u is Lipschitz in (x,t), it is differentiable almost everywhere. Also $\overline{H}$ is twice differentiable almost everywhere because of the convexity of $\overline{H}$. It is therefore natural to require that u is differentiable or $\overline{H}$ is twice differentiable at a particular point. However, it is quite restrictive if we require that u is differentiable at (x,t), and $\overline{H}$ is twice differentiable at exactly $p = Du(x,t)$.

Before presenting a proof of the above theorem, let us recall various important facts that we need in the following.

4.1.1 Preparations

By the comparison principle, it is straightforward that

$$\|u_t^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))}+\|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))}\le C_0.$$

Here $C_0>0$ is a constant depending only on H and $\|Du_0\|_{L^\infty(\mathbb{R}^n)}$. Same bound holds for u. By (6.12), we can make $\theta>0$ smaller if needed to have

$$\frac{\theta}{2}|p|^2-K_0\le H(y,p)\le\frac{1}{2\theta}|p|^2+K_0\qquad\text{for all }(y,p)\in\mathbb{T}^n\times\mathbb{R}^n,\tag{6.20}$$

for some $K_0>1$. Then, we also have that

$$\frac{\theta}{2}|p|^2-K_0\le\overline{H}(p)\le\frac{1}{2\theta}|p|^2+K_0\qquad\text{for all }p\in\mathbb{R}^n.\tag{6.21}$$

We use (6.20) and (6.21) to get that, for each $v_p\in\mathrm{Lip}(\mathbb{T}^n)$ solving (6.11),

$$\|Dv_p\|_{L^\infty(\mathbb{T}^n)}\le C(|p|+K_0).$$

In particular,

$$\max_{\mathbb{T}^n}v_p-\min_{\mathbb{T}^n}v_p\le C\sqrt{n}(|p|+K_0)=C(|p|+K_0).\tag{6.22}$$

Let $L(y,v)$ and $\overline{L}(v)$ be the Lagrangians (Legendre transforms) of the Hamiltonians $H(y,p)$ and $\overline{H}(p)$, respectively. It is clear that

$$\frac{\theta}{2}|v|^2-K_0\le L(y,v)\le\frac{1}{2\theta}|v|^2+K_0\quad\text{for all }(y,v)\in\mathbb{T}^n\times\mathbb{R}^n,\tag{6.23}$$

and

$$\frac{\theta}{2}|v|^2-K_0\le\overline{L}(v)\le\frac{1}{2\theta}|v|^2+K_0\quad\text{for all }q\in\mathbb{R}^n.$$

For $(x,t)\in\mathbb{R}^n\times(0,\infty)$, the optimal control formula for the solution to (6.15) implies

$$u^\varepsilon(x,t)=\inf_{\substack{\varepsilon\eta(0)=x\\ \eta\in\mathrm{AC}([-\varepsilon^{-1}t,0])}}\left\{u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}t\right)\right)+\varepsilon\int_{-\varepsilon^{-1}t}^0 L(\eta(s),\eta'(s))\,ds\right\}.\tag{6.24}$$

4.1.2 Proof of Theorem 6.19

We divide the proof into two parts. We first derive the lower bound (6.17), which is of course optimal.

Proof of optimal lower bound (6.17). To get this, we only need $u_0\in\mathrm{BUC}(\mathbb{R}^n)\cap\mathrm{Lip}(\mathbb{R}^n)$.

By scaling and translation, it suffices to prove that (6.17) holds for $(x,t)=(0,1)$. In other words, we aim at showing

$$u^\varepsilon(0,1)-u(0,1)\ge -C\varepsilon.\tag{6.25}$$

Without loss of generality, we may assume that $u_0(0)=0$ by considering $\tilde{u}_0=u_0-u_0(0)$. Hence, the Lipschitz of u_0 gives

$$|u_0(x)|\le C|x|\qquad\text{for all }x\in\mathbb{R}^n.\tag{6.26}$$

The optimal control formula (6.24) gives us that

$$u^\varepsilon(0,1) = \inf_{\substack{\eta(0)=0\\ \eta\in \mathrm{AC}([-\varepsilon^{-1},0])}} \left\{ u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) + \varepsilon \int_{-\varepsilon^{-1}}^{0} L(\eta(t),\eta'(t))\,dt \right\}.$$

Due to (6.23) and Jensen's inequality,

$$\varepsilon \int_{-\varepsilon^{-1}}^{0} L(\eta(t),\eta'(t))\,dt \geq \varepsilon \int_{-\varepsilon^{-1}}^{0} \left(\theta \frac{|\eta'(t)|^2}{2} - K_0\right) dt \geq \frac{\theta}{2}\varepsilon^2 \left|\eta(-\varepsilon^{-1})\right|^2 - K_0.$$

Combine this with (6.26) to imply that there exists $C > 0$ such that minimization in the formula of $u^\varepsilon(0,1)$ happens when $\varepsilon\left|\eta(-\varepsilon^{-1})\right| \leq C$, that is,

$$u^\varepsilon(0,1) = \inf_{\substack{\eta(0)=0,\\ \varepsilon\left|\eta(-\varepsilon^{-1})\right|\leq C}} \left\{ u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) + \varepsilon \int_{-\varepsilon^{-1}}^{0} L(\eta(t),\eta'(t))\,dt \right\}. \tag{6.27}$$

Clearly, there exists $C_1 > 0$ such that for any $|v| \leq C$,

$$\overline{L}(v) = \sup_{p\in\mathbb{R}^n} \left\{p\cdot v - \overline{H}(p)\right\} = \sup_{|p|\leq C_1} \{p\cdot v - \overline{H}(p)\}. \tag{6.28}$$

This is important as it means that we only need to deal with $|p| \leq C_1$. For $p \in \mathbb{R}^n$, let $v_p \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to (6.11) such that $v_p(0) = 0$. Then for any Lipschitz continuous curve $\eta : [-\varepsilon^{-1}, 0] \to \mathbb{R}^n$, Lemma 6.14 gives

$$\int_{-\varepsilon^{-1}}^{0} \left(L(\eta(t),\eta'(t)) + \overline{H}(p)\right) dt \geq p\cdot\eta(0) - p\cdot\eta\left(-\varepsilon^{-1}\right) + v_p(\eta(0)) - v_p\left(\eta\left(-\varepsilon^{-1}\right)\right).$$

Therefore, if we assume further that $\eta(0) = 0$ and $\varepsilon\left|\eta(-\varepsilon^{-1})\right| \leq C$, then we are able to combine the above with (6.22) and (6.28) to yield

$$\begin{aligned}
\varepsilon \int_{-\varepsilon^{-1}}^{0} (L(\eta(t),\eta'(t))\,dt &\geq \sup_{p\in\mathbb{R}^n} \left\{p\cdot\left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) - \overline{H}(p) + \varepsilon v_p(0) - \varepsilon v_p\left(\eta\left(-\varepsilon^{-1}\right)\right)\right\}\\
&\geq \sup_{|p|\leq C_1} \left\{p\cdot\left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) - \overline{H}(p) + \varepsilon v_p(0) - \varepsilon v_p\left(\eta\left(-\varepsilon^{-1}\right)\right)\right\}\\
&\geq \overline{L}\left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) - C\varepsilon.
\end{aligned}$$

Plug this into (6.27) to imply

$$\begin{aligned}
u^\varepsilon(0,1) &\geq \inf_{\substack{\eta(0)=0,\\ \varepsilon\left|\eta(-\varepsilon^{-1})\right|\leq C}} \left\{u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) + \overline{L}\left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right)\right\} - C\varepsilon\\
&\geq \inf_{y\in\mathbb{R}^n} \left\{u_0(y) + \overline{L}(-y)\right\} - C\varepsilon\\
&= u(0,1) - C\varepsilon.
\end{aligned}$$

The last equality in the above holds thanks to the Hopf–Lax formula for u. □

We now proceed to prove upper bounds (6.18) and (6.19). Again, this is just a conditionally optimal upper bound. The following lemma is a key step toward proving (6.18) and (6.19). Once it is proved, we can combine it with Theorem 6.17 to conclude right away.

Lemma 6.20. *Fix $(x,t) \in \mathbb{R}^n \times (0,\infty)$. Assume that u is differentiable at (x,t) and $\overline{H}$ is differentiable at p for $p = Du(x,t)$. Suppose that there exist a viscosity solution $v_p \in \mathrm{Lip}(\mathbb{T}^n)$ of (6.11) and a backward characteristic $\xi : (-\infty, 0] \to \mathbb{R}^n$ of v_p such that, for some given $C_p > 0$ and $\alpha \in (0,1]$,*

$$\left| \frac{\xi(s) - \xi(0)}{s} - D\overline{H}(p) \right| \le \frac{C_p}{|s|^\alpha} \qquad \text{for all } s < 0.$$

Then

$$u^\varepsilon(x,t) \le u(x,t) + CC_p t^{1-\alpha} \varepsilon^\alpha + C\varepsilon. \tag{6.29}$$

If we further assume that the initial data $u_0 \in C^2(\mathbb{R}^n)$ with $M = \|D^2 u_0\|_{C(\mathbb{R}^n)} < \infty$, then the above bound can be improved to

$$u^\varepsilon(x,t) \le u(x,t) + MC_p^2 t^{2(1-\alpha)} \varepsilon^{2\alpha} + C\varepsilon. \tag{6.30}$$

Proof. Note that $\|Du\|_{L^\infty(\mathbb{R}^n \times [0,\infty))} \le \|Du_0\|_{L^\infty(\mathbb{R}^n)}$. It suffices to prove the above for $(x,t) = (0,t)$. By the Hopf–Lax formula,

$$\begin{aligned} u(0,t) &= \min_{y \in \mathbb{R}^n} \left\{ u_0(y) + t\overline{L}(-t^{-1}y) \right\} \\ &= u_0(y_0) + t\overline{L}(-t^{-1}y_0) \end{aligned}$$

for some $y_0 \in \mathbb{R}^n$. Then $p = Du(0,t) \in \partial \overline{L}(-t^{-1}y_0)$. The Legendre transform also tells us that $-t^{-1}y_0 = D\overline{H}(p)$, and

$$t\overline{L}(-t^{-1}y_0) = -y_0 \cdot p - t\overline{H}(p).$$

Let v_p and ξ be the viscosity solution and its backward characteristic from the assumption. By periodicity, we may assume that $\xi(0) \in Y = [0,1]^n$. By our assumption,

$$|y_0 - \varepsilon\xi(-\varepsilon^{-1}t) + \varepsilon\xi(0)| \le C_p t^{1-\alpha} \varepsilon^\alpha,$$

and hence

$$|y_0 - \varepsilon\xi(-\varepsilon^{-1}t)| \le C_p t^{1-\alpha}\varepsilon^\alpha + C\varepsilon.$$

We use the above and optimal control formula of $u^\varepsilon(0,t)$ to compute that

$$\begin{aligned} u^\varepsilon(0,t) &\le u^\varepsilon(\varepsilon\xi(0),t) + C\varepsilon \le u_0\left(\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) + \varepsilon \int_{-\varepsilon^{-1}t}^{0} L(\xi(s),\xi'(s))\,ds + C\varepsilon \\ &= u_0\left(\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) - t\overline{H}(p) + p \cdot \left(-\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) + p \cdot (\varepsilon\xi(0)) \\ &\qquad - \varepsilon v_p\left(\xi\left(-\varepsilon^{-1}t\right)\right) + \varepsilon v_p(\xi(0)) + C\varepsilon \\ &\le u_0(y_0) + (-y_0) \cdot p - t\overline{H}(p) + CC_p t^{1-\alpha}\varepsilon^\alpha + C\varepsilon \\ &= u(0,t) + CC_p t^{1-\alpha}\varepsilon^\alpha + C\varepsilon. \end{aligned}$$

Next we prove (6.30). If $u_0 \in C^2(\mathbb{R}^n)$, then $p = Du_0(y_0)$. Accordingly, we are able to refine the above calculation as following

$$
\begin{aligned}
u^\varepsilon(0,t) &\le u_0\left(\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) - t\overline{H}(p) + p\cdot\left(-\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) + p\cdot(\varepsilon\xi(0)) \\
&\qquad - \varepsilon v_p\left(\xi\left(-\varepsilon^{-1}t\right)\right) + \varepsilon v_p(\xi(0)) + C\varepsilon \\
&\le u_0\left(\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) + Du_0(y_0)\cdot\left(-\varepsilon\xi\left(-\varepsilon^{-1}t\right)\right) - t\overline{H}(p) + C\varepsilon \\
&\le u_0(y_0) + Du_0(y_0)\cdot(-y_0) + \frac{M}{2}|y_0 - \varepsilon\xi(-\varepsilon^{-1}t)|^2 - t\overline{H}(p) + C\varepsilon \\
&\le u_0(y_0) + p\cdot(-y_0) - t\overline{H}(p) + MC_p^2 t^{2(1-\alpha)}\varepsilon^{2\alpha} + C\varepsilon \\
&= u(0,t) + MC_p^2 t^{2(1-\alpha)}\varepsilon^{2\alpha} + C\varepsilon.
\end{aligned}
$$

□

Since $\|Du\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} = \|Du_0\|_{L^\infty(\mathbb{R}^n)}$ and u is differentiable a.e. in $\mathbb{R}^n\times(0,\infty)$, by Lemma 6.20 and approximations, we have the following corollary.

Corollary 6.21. *Assume that $\overline{H}\in C^1(\mathbb{R}^n)$. Assume further that for every $|p|\le\|Du_0\|_{L^\infty(\mathbb{R}^n)}$, there exist a viscosity solution $v_p\in\mathrm{Lip}(\mathbb{T}^n)$ of* (6.11) *and a backward characteristic $\xi:(-\infty,0]\to\mathbb{R}^n$ of v_p such that, for some $C>0$ independent of p,*

$$
\left|\frac{\xi(s)-\xi(0)}{s} - D\overline{H}(p)\right| \le \frac{C}{|s|} \qquad \text{for all } s<0.
$$

Then

$$
u^\varepsilon(x,t) \le u(x,t) + C\varepsilon \quad \text{for all } (x,t)\in\mathbb{R}^n\times[0,\infty). \tag{6.31}
$$

We are now ready to obtain (6.18) and (6.19).

Proof of upper bounds (6.18) *and* (6.19). Inequalities (6.18) and (6.19) follow immediately from Lemma 6.20 and Theorem 6.17. □

4.2 The one dimensional setting

In one dimension, we have unconditional optimal convergence rate $O(\varepsilon)$ as in the following theorem.

Theorem 6.22. *Let $n=1$. Assume* (6.12) *and $u_0\in\mathrm{BUC}(\mathbb{R})\cap\mathrm{Lip}(\mathbb{R})$. For $\varepsilon>0$, let u^ε be the viscosity solution to* (6.15). *Let u be the viscosity solution to* (6.16). *Then, there exists a constant $C>0$ dependent only on H and $\|Du_0\|_{L^\infty(\mathbb{R})}$ such that*

$$
\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}\times[0,\infty))} \le C\varepsilon. \tag{6.32}
$$

Proof. Thanks to (6.17), the lower bound is always optimal, that is,

$$
u^\varepsilon(x,t) \ge u(x,t) - C\varepsilon \qquad \text{for all } (x,t)\in\mathbb{R}\times[0,\infty).
$$

Here, $C>0$ dependent only on H and $\|Du_0\|_{L^\infty(\mathbb{R})}$.

We now prove the optimal upper bound. In one dimension, $\overline{H}$ has an explicit formula. In particular, we have that $\overline{H}\in C^1(\mathbb{R})$ (see Exercise 55 or [20]). Thanks to Corollary 6.21 and Lemma 6.23 right below, we get the desired conclusion.

□

As discussed above, $\overline{H} \in C^1(\mathbb{R})$. We now show that the assumption in Corollary 6.21 holds.

Lemma 6.23. *For $p \in \mathbb{R}$, let v be a viscosity solution to*

$$H(y, p + v') = \overline{H}(p) \qquad \text{in } \mathbb{T}.$$

Then, for every backward characteristic $\xi : (-\infty, 0] \to \mathbb{R}$ of v, we have

$$\left| \frac{\xi(t) - \xi(0)}{t} - \overline{H}'(p) \right| \leq \frac{1}{|t|} \qquad \text{for all } t < 0. \tag{6.33}$$

Proof. Fix $p \in \mathbb{R}$. There are two cases to be considered.

Case 1. $\overline{H}(p) = \min \overline{H}$. Then $\overline{H}'(p) = 0$.

Let ξ be a backward characteristic of v with $\xi(0) = 0$. Since ξ cannot intersect itself, we have either $\xi((-\infty, 0]) \subset [0, \infty)$ or $\xi((-\infty, 0]) \subset (-\infty, 0]$. Without loss of generality, we assume that $\xi((-\infty, 0]) \subset [0, \infty)$, that is, ξ is nonincreasing on $(-\infty, 0]$. It is clear that we utilize much the one dimensional structure here. Note that, ξ satisfies

$$\xi'(t) = D_p H(\xi(t), p + v'(\xi(t))) \qquad \text{for all } t \leq 0.$$

We claim that

$$\xi((-\infty, 0]) \subset [0, 1). \tag{6.34}$$

Assume otherwise that (6.34) does not hold. Then $\xi(T) = 1$ for some $T < 0$, and we deduce also that $v \in C^1(\mathbb{T})$. By periodicity, $\xi(mT) = m$ for all $m \in \mathbb{N}$. Therefore,

$$\lim_{t \to -\infty} \frac{\xi(t)}{t} = \frac{1}{T} \neq 0 = \overline{H}'(p),$$

which is a contradiction with our assumption. Thus, (6.34) holds, which means that ξ is a bounded orbit. Surely, (6.33) holds true.

Case 2. $\overline{H}(p) > \min \overline{H}$.

Without loss of generality, we assume $\overline{H}'(p) > 0$. Let ξ be a backward characteristic of v with $\xi(0) = 0$. Then $\xi((-\infty, 0]) \subset (-\infty, 0]$, $v \in C^{1,1}(\mathbb{T})$ and

$$\xi'(t) = D_p H(\xi(t), p + v'(\xi(t))) > 0 \qquad \text{for all } t \leq 0.$$

Then, by changing of variables $x = \xi(s)$, we imply

$$|t| = \int_t^0 ds = \int_t^0 \frac{\xi'(s)}{\xi'(s)}\, ds = \int_{\xi(t)}^0 \frac{1}{F_1(x)}\, dx,$$

where $F_1(x) = D_p H(x, p + v'(x))$ for $x \in \mathbb{R}$. Of course, F_1 is 1-periodic (or we write $F_1 \in C(\mathbb{T})$). Accordingly, for $t < 0$,

$$\frac{t}{\xi(t)} = \frac{1}{|\xi(t)|} \int_{\xi(t)}^0 \frac{1}{F_1(x)}\, dx = \left(1 + \frac{E_t}{\xi(t)}\right) \int_0^1 \frac{1}{F_1(x)}\, dx,$$

where E_t is an error term satisfying $|E_t| \le 1$ thanks to Lemma 6.24 below. Then

$$\left| \frac{\xi(t)}{t} - \left(\int_0^1 \frac{1}{F_1(x)}\, dx \right)^{-1} \right| = \frac{|E_t|}{|t|} \le \frac{1}{|t|}.$$

The proof is complete and we get in addition that

$$\overline{H}'(p) = \left(\int_0^1 \frac{1}{F_1(x)}\, dx \right)^{-1}.$$

□

Lemma 6.24. *Assume that $f \in C(\mathbb{T}, [0,\infty))$ and $L > 0$ are given. Then*

$$\left| \int_0^L f\, dy - L \int_0^1 f\, dy \right| \le \int_0^1 f\, dy. \tag{6.35}$$

We can view this lemma as a quantitative version of the ergodic theorem for periodic functions in one dimension. It is also not so hard to see that inequality (6.35) is sharp.

Proof. For a given real number $s \in \mathbb{R}$, denote by $[s]$ its integer part. We have

$$\begin{aligned}
\left| \int_0^L f\, dy - L \int_0^1 f\, dy \right| &= \left| \int_0^{[L]} f\, dy + \int_{[L]}^L f\, dy - L \int_0^1 f\, dy \right| \\
&= \left| ([L] - L) \int_0^1 f\, dy + \int_{[L]}^L f\, dy \right| \\
&\le \max\left\{ (L - [L]) \int_0^1 f\, dy, \int_{[L]}^L f\, dy \right\} \le \int_0^1 f\, dy.
\end{aligned}$$

We use the fact that $f \ge 0$ in the last line above. □

Remark 6.25. It is worth noting that (6.33) is sharp, and $\frac{1}{|t|}$ is the best possible bound that we can obtain. This can been seen rather clearly from the proof of Lemma 6.23.

Besides, the proof of Lemma 6.23 gives us another way to obtain some properties of $\overline{H}$ in one dimension.

4.3 The two dimensional setting

Here is our main result in two dimensions.

Theorem 6.26. *Let $n = 2$. Assume (6.12) and $g \in \mathrm{BUC}(\mathbb{R}^2) \cap \mathrm{Lip}(\mathbb{R}^2)$. Assume further that H is positively homogeneous of degree k in p for some $k \ge 1$, that is, $H(y, \lambda p) = \lambda^k H(y,p)$ for all $(\lambda, y, p) \in [0,\infty) \times \mathbb{T}^2 \times \mathbb{R}^2$. Then,*

$$|u^\varepsilon(x,t) - u(x,t)| \le C\varepsilon \qquad \text{for all } (x,t) \in \mathbb{R}^2 \times [0,\infty). \tag{6.36}$$

Here $C > 0$ is a constant depending only on H and $\|Dg\|_{L^\infty(\mathbb{R}^2)}$.

Of course, when $k = 1$, H is positively homogeneous of degree 1, which corresponds to a front propagation problem that has been discussed few times in the book. This is probably one of the most physically relevant situations in the homogenization theory.

The proof of Theorem 6.26 is rather involved, and is outside of the scope of this book. As a matter of fact, one needs to use two dimensional Aubry–Mather theory here. We therefore skip its proof, and refer the readers to Mitake, Tran, Yu [118] for details.

4.4 Problems

Exercise 55. *Let* $n = 1$. *Assume* (6.12). *Show that* $\overline{H} \in C^1(\mathbb{R})$.

5 Equivalent characterizations of Lipschitz viscosity subsolutions

5.1 Characterizations of Lipschitz subsolutions

Let us now give characterizations of Lipschitz subsolutions to the cell problems. This is an upgraded version of Theorem 2.27. We note that the problem can be phrased in a more general domain ($\mathbb{R}^n$ or bounded domain U) as well. Similar characterizations hold for static problems and Cauchy problems (see the exercises below). The problem of interest is (6.11), that is,

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n.$$

Theorem 6.27. *Assume* (6.12). *Fix* $p \in \mathbb{R}^n$. *Let* $v \in \mathrm{Lip}(\mathbb{T}^n)$. *Then, the following claims are equivalent*

(i) v *is a viscosity subsolution to* (6.11)*;*

(ii) v *is an a.e. subsolution to* (6.11)*;*

(iii) for any arbitrary Lipschitz curve $\gamma : [-T, 0] \to \mathbb{T}^n$ *for some* $T > 0$*,*

$$\int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + \overline{H}(p)\right) dt \geq p \cdot (\gamma(0) - \gamma(-T)) + v(\gamma(0)) - v(\gamma(-T)).$$

Proof. We note first that (i) and (ii) are equivalent thanks to Theorem 2.27. The new point here is characterization (iii).

It is clear that Lemma 6.14 and Remark 6.15 give us that "(i) $\Longrightarrow$ (iii)". To finish off the proof, we need to show that "(iii) $\Longrightarrow$ (ii)". Indeed, fix a point $x \in \mathbb{T}^n$ which is a differentiable point of v. Fix a direction $e \in \mathbb{R}^n$, and denote by

$$\gamma(s) = x + se \qquad \text{for all } s \leq 0.$$

By the hypothesis in (iii), for each $T > 0$,

$$\int_{-T}^{0} \left(L(x + te, e) + \overline{H}(p)\right) dt \geq Tp \cdot e + v(x) - v(x - Te).$$

Divide both sides of the above by T and let $T \to 0+$ to infer

$$L(x,e) + \overline{H}(p) \geq p \cdot e + Dv(x) \cdot e,$$

which means that

$$e \cdot (p + Dv(x)) - L(x,e) \leq \overline{H}(p).$$

Take the supremum of the above over $e \in \mathbb{R}^n$ to conclude that

$$H(x, p + Dv(x)) \leq \overline{H}(p).$$

The proof is complete. □

Remark 6.28. Let us recall that if $v \in C(\mathbb{T}^n)$ is a subsolution to (6.11), then Lemma 1.28 implies immediately that $v \in \mathrm{Lip}(\mathbb{T}^n)$. This is just to show that the assumption that $v \in \mathrm{Lip}(\mathbb{T}^n)$ in the above theorem is not quite needed for viscosity subsolutions. Nevertheless, it is needed for (ii) and (iii) in the theorem.

5.2 Problems

Exercise 56. *Formulate and give a proof for an analogous result to Theorem 6.27 for the discount problem*

$$\lambda v^\lambda + H(y, Dv^\lambda) = 0 \qquad \text{in } \mathbb{R}^n.$$

Here, $\lambda > 0$ is given.

Exercise 57. *Formulate and give a proof for an analogous result to Theorem 6.27 for the Cauchy problem*

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

Here, $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$ is given.

6 References

1. The new representation formula for solutions of the discount problems based on a duality method was derived by Ishii, Mitake, Tran [87, 88]. See [87, 88] for representation formulas of solutions to other boundary problems (state-constraint, Dirichlet, and Neumann problems). The methods in [87, 88] are quite robust and are applicable to second-order equations as well.

2. Sion's minimax theorem was derived by Sion [125]. We give a detailed proof of this minimax theorem (Theorem A.1) in Appendix. There, we follow a proof by Komiya [97], which is quite elementary.

3. The selection problem for the vanishing discount problem has been studied extensively recently. There are various different approaches to prove convergence for the convex setting. Davini, Fathi, Iturriaga, Zavidovique [42] used weak KAM theory to obtain the result for first-order convex Hamilton–Jacobi equations. Mitake, Tran [116] used

the nonlinear adjoint method to get the convergence for possibly degenerate viscous Hamilton–Jacobi equations. Al-Aidarous, Alzahrani, Ishii, Younas [2] studied the first-order problem with Neumann boundary condition. The proofs presented here are completely different and due to Ishii, Mitake, Tran [87, 88]. This duality approach [87, 88] works for fully nonlinear second order equations as well. We give two proofs of convergence for the vanishing discount problem. The first proof, Theorem 6.9, is taken from [87]. The second proof, Theorem 6.10, follows the philosophy in [42, 116]. All the results in [42, 116, 2, 87, 88] require convexity of the Hamiltonians. Some nonconvex cases were studied by Gomes, Mitake, Tran [77]. See the lecture notes of Le, Mitake, Tran [100] for an account of this vanishing discount problem via the nonlinear adjoint method, which is quite different from what is provided here. More references and related areas are provided there as well.

4. Backward characteristics of solutions to cell problems here are done based on the optimal control formula of a corresponding time-dependent problem. This formulation is quite natural from the viewpoint of the optimal control formula for Cauchy problems. Fathi [59] gave a different approach based on the weak KAM theorem. See also the lecture notes of Ishii [85].

5. Rate of convergence of large time average of backward characteristics was taken from Gomes [71], Mitake, Tran, Yu [118]. This is still a largely unexplored topic, and not so much is known.

6. The section on optimal rate of convergence in periodic homogenization theory was taken from Mitake, Tran, Yu [118]. See [118] for the proof of Theorem 6.26 that we do not cover here. Other aspects and open problems are also discussed in [118]. For a more complicated situation with multi-scales in one dimension, see the work of Tu [133].

7. Although quite simple, the equivalent characterizations of Lipschitz viscosity subsolutions in Theorem 6.27 are quite useful in various situations. Of course, similar characterizations hold for static problems and Cauchy problems. It should be noted that the results are still valid for a general domain $U \subset \mathbb{R}^n$.

CHAPTER 7

Introduction to weak KAM theory

1 Introduction

In this chapter, we always assume that

$$\begin{cases} H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{pp}H(y,p) \le \theta^{-1} I_n \text{ for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{7.1}$$

Let $L = L(y,v)$ be the corresponding Lagrangian. By changing $\theta > 0$ to be smaller if needed, we may also assume that

$$\begin{cases} L \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{vv}L(y,v) \le \theta^{-1} I_n \text{ for all } (y,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{7.2}$$

Let us give a minimalistic type introduction to this subject. We are concerned with the following Hamiltonian system

$$\begin{cases} x'(t) = D_pH(x(t),p(t)), \\ p'(t) = -D_xH(x(t),p(t)). \end{cases} \tag{7.3}$$

In general, this Hamiltonian system is complicated to be studied deeply, and a natural idea is to find generating functions and do canonical changes of variables to arrive at an integrable system, which is solvable. Heuristically, the generating functions and canonical changes of variables are strongly tied to the cell problems that we discussed in previous chapters. Recall that, for $P \in \mathbb{R}^n$, our cell problem is

$$H(x, P + Dv(x,P)) = \overline{H}(P) \qquad \text{in } \mathbb{T}^n. \tag{7.4}$$

Assume for now that both $v(x,P)$ and $\overline{H}(P)$ are smooth functions. Then, if the relation

$$\begin{cases} X = x + D_P v(x,P), \\ p = P + D_x v(x,P), \end{cases}$$

defines a smooth and invertible change of variables, then we can transform (7.3) into the following integrable system

$$\begin{cases} X'(t) = D\overline{H}(P(t)), \\ P'(t) = 0. \end{cases} \tag{7.5}$$

In terms of mechanics, P is called an action, and X is called an angle or rotation variable.

However, in general, this classical procedure cannot be carried out because of various reasons. First of all, (7.4) does not have smooth solutions $v(x,P)$ in general. In fact, $v(\cdot,P)$ is often only Lipschitz in x. The dependence of v on P is even worse, and we will see later that there are cases that this dependence is even discontinuous. Second of all, $\overline{H}$ is convex because of the convexity of H in assumption (7.1), but it is not known to be smooth. Of course, there are examples that $\overline{H}$ is not C^1. To date, very little is known about deep properties of $\overline{H}$ as explained in previous chapters. Finally, the canonical transformation $(x,p)\mapsto(X,P)$, even if can be defined locally, is not usually globally defined.

Nevertheless, there is a rich underlying structure in (7.4), and it is extremely important to come up with weak interpretations of the classical program briefly mentioned above. Various great works of Aubry [9], Mather [112, 113], Mañé [109], Fathi [58, 59], E [44], Evans, Gomes [53] show that some solutions of (7.3), which correspond to appropriate minimizers of the action functionals, see some kind of "integrable structures" within the full dynamics. Weak KAM theory, which was named by Fathi, is an attempt to bring PDE techniques to analyze more (7.4) and their underlying dynamics in multi dimensions.

It is important emphasizing that weak KAM is different from conventional KAM theory as it is not a perturbative theory. Here, our Hamiltonian H is not a perturbation of an integrable Hamiltonian. As already explained, we see that solutions $v(x,P)$ of (7.4) are only Lipschitz in x, and are not dependent in P in a nice way, and so, we need to be careful with interpretations and usages of these viscosity (generalized) solutions.

One final point is that in dimension three or higher, the minimizing trajectories might occupy just a small part of the torus, and hence, might not give us much information.

There are often two kinds of approaches to study weak KAM: the Lagrangian (dynamical system) methods, and the nonlinear PDE methods. Let us go first into the Lagrangian method.

2 Lagrangian methods in weak KAM theory

This section is inspired by the book of Fathi [59]. Many of the results are taken from there. Some are presented in a different way that are more of my personal taste.

2.1 The weak KAM theorem

Given $\gamma\in \mathrm{AC}([0,T],\mathbb{T}^n)$ for some $T>0$, we define the action functional corresponding to γ to be

$$A_T[\gamma]=\int_0^T L(\gamma(s),\gamma'(s))\,ds.$$

Definition 7.1. *Let $\xi\in \mathrm{AC}([0,T],\mathbb{T}^n)$ for some given $T>0$. We say that ξ is a minimizer of $A_T[\cdot]$ if*

$$A_T[\xi]\le A_T[\gamma]$$

for all $\gamma\in \mathrm{AC}([0,T],\mathbb{T}^n)$ with $\gamma(0)=\xi(0)$, $\gamma(T)=\xi(T)$.

Lemma 7.2. *Assume* (7.1). *Let $\xi \in \mathrm{AC}([0,T],\mathbb{T}^n)$ be a minimizer of $A_T[\cdot]$. Then, there exists $C_T > 0$ such that*

$$\max_{t\in[0,T]} |\xi'(t)| \le C_T.$$

Proof. It is clear that ξ satisfies an Euler–Lagrange equation

$$\frac{d}{dt}\big(D_vL(\xi(t),\xi'(t))\big) = D_xL(\xi(t),\xi'(t)) \qquad \text{for all } t\in[0,T].$$

Denote by $x(t)=\xi(t)$, and $p(t)=D_vL(\xi(t),\xi'(t))$ for $t\in[0,T]$. Then (x,p) solves the following Hamiltonian system

$$\begin{cases} x'(t) = D_pH(x(t),p(t)), \\ p'(t) = -D_xH(x(t),p(t)), \end{cases} \qquad \text{for } t\in[0,T].$$

As $H\in C^2(\mathbb{T}^n\times\mathbb{R}^n)$, we get that $x\in C^2([0,T])$, which means $\xi\in C^2([0,T])$.

Furthermore, it is worth noting here that we have conservation of energy, that is, $t\mapsto H(x(t),p(t))$ is constant on $[0,T]$. This can be easily checked as

$$\frac{d}{dt}H(x(t),p(t)) = D_xH(x(t),p(t))\cdot x'(t) + D_pH(x(t),p(t))\cdot p'(t) = 0.$$

In particular, this allows us to get that $H(x(t),p(t))\le C_T$, which implies $|p(t)|\le C_T$, and also $|\xi'(t)|\le C_T$ for all $t\in[0,T]$. ☐

For given $u_0\in C(\mathbb{T}^n)$, we consider the usual Cauchy problem

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

The optimal control formula for u gives, for $(x,t)\in\mathbb{T}^n\times[0,\infty)$,

$$\begin{aligned} u(x,t) &= \inf\left\{\int_0^t L(\gamma(s),\gamma'(s))\,ds + u_0(\gamma(0)) \,:\, \gamma\in\mathrm{AC}([0,t],\mathbb{T}^n), \gamma(t)=x\right\} \\ &= \inf\left\{A_t[\gamma] + u_0(\gamma(0)) \,:\, \gamma\in\mathrm{AC}([0,t],\mathbb{T}^n), \gamma(t)=x\right\}. \end{aligned}$$

Definition 7.3. *We define*

$$T_t^- u_0(x) = u(x,t) = \inf\left\{\int_0^t L(\gamma(s),\gamma'(s))\,ds + u_0(\gamma(0)) \,:\, \gamma\in\mathrm{AC}([0,t],\mathbb{T}^n), \gamma(t)=x\right\}.$$

We call $\{T_t^-\}_{t\ge 0}$ the Lax–Oleinik semigroup.

As shown in Section 2 in Appendix, $u(x,t)=T_t^-u_0(x)$ admits a minimizer in the formula, that is, there exists $\xi\in\mathrm{AC}([0,t],\mathbb{T}^n)$ such that $\xi(t)=x$, and

$$u(x,t) = T_t^-u_0(x) = \int_0^t L(\xi(s),\xi'(s))\,ds + u_0(\xi(0)).$$

As we have developed the theory for viscosity solutions of Cauchy problem, various properties of the Lax–Oleinik semigroup $\{T_t^-\}_{t\ge 0}$ hold accordingly. Let us record them here.

Lemma 7.4 (Properties of the Lax–Oleinik semigroup)**.** *Assume* (7.1). *Then, the following properties hold.*

- $\{T_t^-\}_{t\geq 0}$ *is a semigroup, that is,* $T_{t+s}^- = T_t^- \circ T_s^-$ *for all* $t,s \geq 0$.
- *For* $v,w \in C(\mathbb{T}^n)$ *with* $v \leq w$, $T_t^- v \leq T_t^- w$ *for all* $t \geq 0$.
- *For* $v \in C(\mathbb{T}^n)$ *and* $c \in \mathbb{R}$, $T_t^-(v+c) = T_t^- v + c$ *for all* $t \geq 0$.
- *For* $v \in C(\mathbb{T}^n)$, $\lim_{t\to 0+} T_t^- v = v$ *in* $C(\mathbb{T}^n)$.
- *For* $v \in C(\mathbb{T}^n)$, $t \mapsto T_t^- v$ *is uniformly continuous.*

Here is the weak KAM theorem that was done by Fathi [59] via the method of finding a fixed point for the Lax–Oleinik semigroup.

Theorem 7.5 (Weak KAM theorem)**.** *Assume* (7.1). *There exists a function* $v_- \in C(\mathbb{T}^n)$ *and a constant* $c \in \mathbb{R}$ *such that*

$$T_t^- v_- + ct = v_- \qquad \text{for all } t \geq 0.$$

In fact, this theorem can be derived quickly from the cell problems, and it already appears in previous chapters (in the proof of Theorem 6.11 for example). Let us recall it here for clarity.

Proof. Let $P = 0$, and $v = v(x,0) \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution of the corresponding cell problem (7.4), that is,

$$H(x, Dv(x)) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n.$$

Then, $u(x,t) = T_t^- v(x) = v(x) - \overline{H}(0)t$ for all $(x,t) \in \mathbb{T}^n \times [0,\infty)$. The proof is complete with $v_- = v$ and $c = \overline{H}(0)$. □

Let us now proceed to understand further about properties of v. Recall the backward characteristics of v that we develop in the previous chapter. By Theorem 6.11, for every $x \in \mathbb{T}^n$, there exists a C^1 backward characteristic $\xi : (-\infty, 0] \to \mathbb{T}^n$ such that $\xi(0) = x$, and

$$v(\xi(t_1)) - v(\xi(t_2)) = \int_{t_2}^{t_1} \left(L(\xi(t), \xi'(t)) + \overline{H}(0) \right) dt \tag{7.6}$$

for all $t_2 < t_1 \leq 0$. We show that v is differentiable at $\xi(t)$ for $t < 0$.

Theorem 7.6. *Assume* (7.1). *Let* $P = 0$, *and* $v = v(x,0) \in \mathrm{Lip}(\mathbb{T}^n)$ *be a solution of the corresponding cell problem* (7.4). *For* $x \in \mathbb{T}^n$, *let* ξ *be a backward characteristic of* v *starting from* x. *Then,* v *is differentiable at* $\xi(t)$ *for all* $t < 0$, *and*

$$Dv(\xi(t)) = D_v L(\xi(t), \xi'(t)).$$

Proof. Fix $z \in \mathbb{T}^n$, and let ξ be a backward characteristic of v starting from z. Fix $t < 0$, and denote by $y = \xi(t)$. We aim at showing that v is differentiable at y, and $Dv(y) = D_v L(\xi(t), \xi'(t))$.

For every $x \in \mathbb{T}^n$, define $\xi_x : [2t, t] \to \mathbb{T}^n$ as

$$\xi_x(s) = \xi(s) + \frac{2t-s}{t}(x-y) \qquad \text{for } s \in [2t, t].$$

Then we have that $\xi_x(2t) = \xi(2t)$, and $\xi_x(t) = \xi(t) + (x-y) = x$. Set

$$\begin{aligned}\phi(x) &= v(\xi(2t)) + \int_{2t}^{t} L(\xi_x(s), \xi_x'(s))\, ds \\ &= v(\xi(2t)) + \int_{2t}^{t} L\left(\xi(s) + \frac{2t-s}{t}(x-y), \xi'(s) - \frac{x-y}{t}\right) ds\end{aligned}$$

It is clear that ϕ is smooth, and by Lemma 6.14, $\phi \ge v$, and $\phi(y) = v(y)$. In other words, ϕ touches v from above at y. By computations and the Euler–Lagrange equations, we see that

$$\begin{aligned}D\phi(y) &= \int_{2t}^{t} \left(\frac{2t-s}{t} D_x L(\xi(s), \xi'(s)) - \frac{1}{t} D_v L(\xi(s), \xi'(s))\right) ds \\ &= \int_{2t}^{t} \left(\frac{2t-s}{t} \frac{d}{ds}\left(D_v L(\xi(s), \xi'(s))\right) - \frac{1}{t} D_v L(\xi(s), \xi'(s))\right) ds \\ &= \int_{2t}^{t} \frac{d}{ds}\left(\left(2 - \frac{s}{t}\right) D_v L(\xi(s), \xi'(s))\right) ds = D_v L(\xi(t), \xi'(t)).\end{aligned}$$

Next, for $x \in \mathbb{T}^n$, define $\xi_x : [t, 0] \to \mathbb{T}^n$ as

$$\xi_x(s) = \xi(s) + \frac{s}{t}(x-y) \qquad \text{for } s \in [t, 0].$$

By abuse of notions, we still use ξ_x here. Note that $\xi_x(t) = x$, and $\xi_x(0) = \xi(0)$. Set

$$\begin{aligned}\psi(x) &= v(\xi(0)) - \int_{t}^{0} L(\xi_x(s), \xi_x'(s))\, ds \\ &= v(\xi(0)) - \int_{t}^{0} L\left(\xi(s) + \frac{s}{t}(x-y), \xi'(s) + \frac{x-y}{t}\right) ds\end{aligned}$$

Again, we see that ψ is smooth, and by Lemma 6.14, $\psi \le v$, and $\psi(y) = v(y)$. In other words, ψ touches v from below at y. A similar computation to the above gives

$$D\psi(y) = D_v L(\xi(t), \xi'(t)).$$

Thus, v is differentiable at y and $Dv(y) = D_v L(\xi(t), \xi'(t))$. $\square$

Remark 7.7. It is important to see that v is differentiable $\xi(t)$ for $t < 0$. Of course, we want to study further the properties of these backward characteristics $\xi(t)$ as $t \to -\infty$. By Theorem 6.12 and Corollary 6.13, we know that if $\overline{H}$ is differentiable at P, then for a backward characteristic v of (7.4),

$$\lim_{t \to -\infty} \frac{\xi(t)}{t} = D\overline{H}(P).$$

If $\overline{H}$ is not differentiable at P, then we only have that there exists a sequence $\{t_k\} \to -\infty$ so that

$$\lim_{k\to\infty} \frac{\xi(t_k)}{t_k} = q \in D^{-}\overline{H}(P).$$

There are several weaknesses here. First, we do not know precisely what is q in general. Second, we do not know if different subsequences of $\frac{\xi(t)}{t}$ converge to different limits yet. Finally, a natural question to ask is that if we are given a vector $V \in \mathbb{R}^n$, then is there any ξ such that

$$\lim_{t\to-\infty} \frac{\xi(t)}{t} = V?$$

However, in general, the answer to this question is negative. This is shown by a famous example of Hedlund [79]. See also Bangert [12], E [44], Mitake, Tran, Yu [118]. We will discuss this matter later.

Therefore, this is a strong need to relax this question a bit to study further. In the following, we introduce one such relaxation.

2.2 Flow invariance and another characterization of $\overline{H}(0)$

Let us now consider the initial-value problem for the Euler–Lagrange equation

$$\begin{cases} \frac{d}{dt}\left(D_v L(x(t), x'(t))\right) = D_x L(x(t), x'(t)), \\ x(0) = x, x'(0) = v. \end{cases} \tag{7.7}$$

Let $v(t) = x'(t)$ for $t \in \mathbb{R}$. Define the flow map $\{\Phi_t\}_{t\in\mathbb{R}}$ as

$$\Phi_t(x,v) = (x(t), v(t)) \qquad \text{for all } t \in \mathbb{R}.$$

Definition 7.8. *A Radon probability measure $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ is said to be flow invariant if*

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(\Phi_t(x,v))\, d\mu(x,v) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,v)\, d\mu(x,v)$$

for every bounded continuous function ψ.

Here is another characterization of $\overline{H}(0)$.

Theorem 7.9. *Assume* (7.1). *Then,*

$$\overline{H}(0) = -\inf\left\{ \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\, d\mu(x,v) \,:\, \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \text{ is flow invariant} \right\}. \tag{7.8}$$

This result is of course quite similar to Theorem 6.6 in the previous chapter. We will go back to this point later.

Proof. Take v_- to be a solution to (7.4) with $P = 0$. Or in other words, v_- is taken from Theorem 7.5. By Lemma 6.14, for $x(\cdot)$ solves (7.7),

$$v_-(x(1)) - v_-(x(0)) \le \int_0^1 \left(L(x(s), x'(s)) + \overline{H}(0) \right) ds.$$

Integrate this inequality with respect to $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ which is flow invariant to imply

$$0 = \int_{\mathbb{T}^n\times\mathbb{R}^n} (v_-(x(1)) - v_-(x)))\, d\mu(x,v) \le \int_0^1 \int_{\mathbb{T}^n\times\mathbb{R}^n} \left(L(x(s),x'(s)) + \overline{H}(0)\right) d\mu(x,v) ds,$$

which yields further that

$$-\overline{H}(0) \le \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\, d\mu(x,v).$$

Take infimum over all such μ to get

$$-\overline{H}(0) \le \inf\left\{ \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\, d\mu(x,v) \,:\, \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \text{ is flow invariant} \right\}.$$

We now prove the converse. Fix $x \in \mathbb{T}^n$, and take ξ to be a backward characteristic of v_- starting from x. We have that, for $t < 0$,

$$v_-(\xi(0)) - v_-(\xi(t)) = \int_t^0 \left(L(\xi(s),\xi'(s)) + \overline{H}(0)\right) ds.$$

Define $\mu_t \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ as

$$\langle \mu_t, \psi \rangle = \frac{1}{|t|} \int_t^0 \psi(\xi(s),\xi'(s))\, ds$$

for every bounded continuous function ψ. It is very important noting that $\mathrm{spt}(\mu_t) \subset \mathbb{T}^n \times \overline{B}(0,C)$ for $C > 0$ sufficiently large because of the fact that $\|\xi'\|_{L^\infty((-\infty,0])} \le C$. Then,

$$\frac{v_-(x) - v_-(\xi(t))}{|t|} = \langle \mu_t, L \rangle + \overline{H}(0).$$

By compactness, we are able to find a sequence $\{t_k\} \to \infty$ such that $\mu_{t_k} \to \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ weakly in the sense of measures, and $\mathrm{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. The above equality infers that

$$-\overline{H}(0) = \langle \mu, L \rangle = \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\, d\mu(x,v).$$

We only need to verify that μ is flow invariant to complete the proof. Indeed, for each bounded continuous function ψ and each $t > 0$,

$$\begin{aligned}
&\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(\Phi_t(x,v))\, d\mu(x,v) = \lim_{k\to\infty} \frac{1}{|t_k|} \int_{t_k}^0 \psi \circ \Phi_t(\xi(s),\xi'(s))\, ds \\
&= \lim_{k\to\infty} \frac{1}{|t_k|} \int_{t_k}^0 \psi(\xi(s+t),\xi'(s+t))\, ds \\
&= \lim_{k\to\infty} \frac{1}{|t_k|} \left(\int_{t_k}^0 \psi(\xi(s),\xi'(s))\, ds + \int_0^t \psi(\xi(s),\xi'(s))\, ds - \int_{t_k}^{t_k+t} \psi(\xi(s),\xi'(s))\, ds \right) \\
&= \lim_{k\to\infty} \frac{1}{|t_k|} \int_{t_k}^0 \psi(\xi(s),\xi'(s))\, ds = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,v)\, d\mu(x,v).
\end{aligned}$$

□

Remark 7.10. In the later part of the above proof, we construct minimizing measure μ as a large time average (via a subsequence) of the uniform distribution on the trajectory $\{(\xi(s), \xi'(s)) : s \in (-\infty, 0]\}$. Automatically, $\mathrm{spt}(\mu)$ is a subset of the α-limit set of this trajectory.

3 Mather measures and Mather set

We are now ready to define Mather measures and Mather set based on the minimization problem (7.8).

Definition 7.11. *Each measure μ that minimizes* (7.8) *is called a Mather measure. Denote the Mather set by*

$$\widetilde{\mathcal{M}}_0 = \overline{\bigcup_{\mu} \mathrm{spt}(\mu)},$$

where the union above is over all minimizing measures. Let π be the natural projection from $\mathbb{T}^n \times \mathbb{R}^n$ to $\mathbb{T}^n$, that is, $\pi(x,v) = x$ for all $(x,v) \in \mathbb{T}^n \times \mathbb{R}^n$. Then, the projected Mather set is defined as

$$\mathcal{M}_0 = \pi(\widetilde{\mathcal{M}}_0).$$

We have the following property of $\widetilde{\mathcal{M}}_0$.

Lemma 7.12. *Assume* (7.1). *Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a subsolution to* (7.4) *for $P = 0$. Pick $(x,v) \in \widetilde{\mathcal{M}}_0$. Then, for each $t \le t'$,*

$$u(\pi(\Phi_{t'}(x,v))) - u(\pi(\Phi_t(x,v))) = \int_t^{t'} \left(L(\Phi_s(x,v)) + \overline{H}(0) \right) ds.$$

Proof. Let $(x,v) \in \mathrm{spt}(\mu)$ for a minimizing measure μ. Firstly, by Remark 6.15,

$$u(\pi(\Phi_{t'}(x,v))) - u(\pi(\Phi_t(x,v))) \le \int_t^{t'} \left(L(\Phi_s(x,v)) + \overline{H}(0) \right) ds. \tag{7.9}$$

Integrate the above over $d\mu(x,v)$, use the invariant property and the minimizing measure property to infer

$$\begin{aligned} 0 = \int_{\mathbb{T}^n \times \mathbb{R}^n} u \circ \pi \, d\mu - \int_{\mathbb{T}^n \times \mathbb{R}^n} u \circ \pi \, d\mu &= \int_{\mathbb{T}^n \times \mathbb{R}^n} (u(\pi(\Phi_{t'}(x,v))) - u(\pi(\Phi_t(x,v)))) \, d\mu(x,v) \\ &\le \int_t^{t'} \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(L(\Phi_s(x,v)) + \overline{H}(0) \right) d\mu(x,v) ds = 0. \end{aligned}$$

Thus, the above inequality (7.9) must be an equality, which concludes our proof. □

It is not hard to see that in fact $\widetilde{\mathcal{M}}_0$ lies in the energy level $\overline{H}(0)$ of the Hamiltonian.

Lemma 7.13. *Assume* (7.1). *Then,*

$$\widetilde{\mathcal{M}}_0 \subset \left\{ (x,v) \in \mathbb{T}^n \times \mathbb{R}^n : H(x, D_v L(x,v)) = \overline{H}(0) \right\}.$$

Proof. Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (7.4) with $P = 0$. We use Lemma 7.12 and repeat Theorem 7.6 to see that, for $(x, v) \in \widetilde{\mathcal{M}}_0$, u is differentiable at x, and $Du(x) = D_v L(x, v)$. Therefore,

$$H(x, Du(x)) = H(x, D_v L(x, v)) = \overline{H}(0).$$

□

Let us now show that $\mathcal{M}_0$ serves as a uniqueness set for the cell problem (7.4) with $P = 0$. Note again that (7.4) may have infinitely many solutions (see Chapter 4, and Le, Mitake, Tran [100, Chapter 6] for such examples), and it is therefore important to obtain that $\mathcal{M}_0$ is a uniqueness set for (7.4) with $P = 0$.

Theorem 7.14 (Uniqueness set for (7.4) with $P = 0$)**.** *Assume* (7.1). *Let* $u_1, u_2 \in \mathrm{Lip}(\mathbb{T}^n)$ *be two solutions to* (7.4) *with* $P = 0$. *Assume that* $u_1 = u_2$ *on* $\mathcal{M}_0$. *Then* $u_1 = u_2$.

Proof. Fix $x \in \mathbb{T}^n$. Let ξ be a backward characteristic of u_1 starting from x. Then, for any $t < 0$,

$$u_1(x) - u_1(\xi(t)) = \int_t^0 \left(L(\xi(s), \xi'(s)) + \overline{H}(0) \right) ds,$$

and

$$u_2(x) - u_2(\xi(t)) \leq \int_t^0 \left(L(\xi(s), \xi'(s)) + \overline{H}(0) \right) ds.$$

Combine these two to infer that

$$u_2(x) - u_1(x) \leq u_2(\xi(t)) - u_1(\xi(t)) \qquad \text{for all } t \leq 0.$$

Let us now use the construction in the later part of the proof of Theorem 7.9 to construct a Mather measure μ to conclude. By the construction, for each $\mu_t \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ for $t < 0$, it is clear that

$$u_2(x) - u_1(x) \leq \langle \mu_t, (u_2 - u_1) \circ \pi \rangle = \frac{1}{|t|} \int_t^0 (u_2 - u_1)(\pi \circ (\xi(s), \xi'(s)))\, ds.$$

As $\mu_{t_k} \to \mu$ weakly in the sense of measures as $k \to \infty$, and μ is a Mather measure, we deduce that

$$u_2(x) - u_1(x) \leq \langle \mu, (u_2 - u_1) \circ \pi \rangle = \int_{\mathbb{T}^n \times \mathbb{R}^n} (u_2 - u_1)(x)\, d\mu(x, v) = 0,$$

by our hypothesis. Thus, $u_2(x) \leq u_1(x)$. By a symmetric argument, $u_1(x) \leq u_2(x)$, and hence, $u_1(x) = u_2(x)$. □

3.1 Lipschitz graph theorem

Theorem 7.15. *Assume* (7.1). *Let* $u \in \mathrm{Lip}(\mathbb{T}^n)$ *be a subsolution to* (7.4) *for* $P = 0$. *There exists* $C > 0$ *depending only on* H *such that, for all* $x \in \mathcal{M}_0$ *and* $h \in \mathbb{R}^n$,

$$|u(x + h) + u(x - h) - 2u(x)| \leq C|h|^2.$$

Proof. Let $(x,v) \in \widetilde{\mathcal{M}}_0$. For $t \in \mathbb{R}$, write $\Phi_t(x,v) = (x(t), x'(t))$ for clarity. Of course, $x(0) = x$. By Lemma 7.12,

$$u(x(1)) - u(x(0)) = \int_0^1 \big(L(x(s), x'(s)) + \overline{H}(0)\big)\, ds, \tag{7.10}$$

and

$$u(x(0)) - u(x(-1)) = \int_{-1}^0 \big(L(x(s), x'(s)) + \overline{H}(0)\big)\, ds. \tag{7.11}$$

Let us obtain first the lower bound. By Lemma 6.14,

$$u(x(1)) - u(x(0)+h) \le \int_0^1 \big(L(x(s)+(1-s)h, x'(s)-h) + \overline{H}(0)\big)\, ds,$$

and

$$u(x(1)) - u(x(0)-h) \le \int_0^1 \big(L(x(s)-(1-s)h, x'(s)+h) + \overline{H}(0)\big)\, ds.$$

Combine these two inequalities with (7.10) to get

$$\begin{aligned} &u(x+h) + u(x-h) - 2u(x) \\ \ge &\int_0^1 \big(2L(x(s), x'(s)) - L(x(s)+(1-s)h, x'(s)-h) - L(x(s)-(1-s)h, x'(s)+h)\big)\, ds \\ \ge &-C|h|^2. \end{aligned} \tag{7.12}$$

On the other hand, use Lemma 6.14 again to yield

$$u(x(0)+h) - u(x(-1)) \le \int_{-1}^0 \big(L(x(s)+(1+s)h, x'(s)+h) + \overline{H}(0)\big)\, ds,$$

and

$$u(x(0)-h) - u(x(-1)) \le \int_{-1}^0 \big(L(x(s)-(1+s)h, x'(s)-h) + \overline{H}(0)\big)\, ds.$$

The above two inequalities, together with (7.11), imply

$$\begin{aligned} &u(x+h) + u(x-h) - 2u(x) \\ \le &\int_0^1 \big(L(x(s)+(1+s)h, x'(s)+h) - L(x(s)-(1+s)h, x'(s)-h) - 2L(x(s), x'(s))\big)\, ds \\ \le &\, C|h|^2. \end{aligned} \tag{7.13}$$

The lower bound (7.12) and the upper bound (7.13) give us the desired result. □

The following Lipschitz graph theorem is due to Mather.

Theorem 7.16. *Assume* (7.1). *Let* $u \in \mathrm{Lip}(\mathbb{T}^n)$ *be a subsolution to* (7.4) *for* $P = 0$. *Then, there exists* $C > 0$ *depending only on* H *such that* *(i) for all $x \in \mathcal{M}_0$ and $y \in \mathbb{T}^n$,*

$$|u(y) - u(x) - Du(x)\cdot(y-x)| \le C|y-x|^2;$$

(ii) For all $x, y \in \mathcal{M}_0$,

$$|Du(x) - Du(y)| \le C|x-y|.$$

Proof. Let $(x,v) \in \widetilde{\mathcal{M}}_0$. Note that u is differentiable at x and $Du(x) = D_vL(x,v)$. We utilize various inequalities and identities in the above proof to prove (i) first. Fix $h \in \mathbb{T}^n$. On one hand,

$$\begin{aligned}
u(x+h) - u(x) &\ge \int_0^1 \big(L(x(s),x'(s)) - L(x(s)+(1-s)h, x'(s)-h)\big)\, ds \\
&\ge \int_0^1 \big(D_xL(x(s),x'(s))\cdot(s-1)h + D_vL(x(s),x'(s))\cdot h\big)\, ds - C|h|^2 \\
&= \int_0^1 \left(\frac{d}{ds}\big(D_vL(x(s),x'(s))\big)\cdot(s-1)h + D_vL(x(s),x'(s))\cdot h\right) ds - C|h|^2 \\
&= \int_0^1 \frac{d}{ds}\big(D_vL(x(s),x'(s))\cdot(s-1)h\big)\, ds - C|h|^2 \\
&= D_vL(x(0),x'(0))\cdot h - C|h|^2 = Du(x)\cdot h - C|h|^2.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
u(x+h) - u(x) &\le \int_{-1}^0 \big(L(x(s)+(1+s)h, x'(s)+h) - L(x(s),x'(s))\big)\, ds \\
&\le \int_{-1}^0 \big(D_xL(x(s),x'(s))\cdot(s+1)h + D_vL(x(s),x'(s))\cdot h\big)\, ds + C|h|^2 \\
&= \int_{-1}^0 \left(\frac{d}{ds}\big(D_vL(x(s),x'(s))\big)\cdot(s+1)h + D_vL(x(s),x'(s))\cdot h\right) ds + C|h|^2 \\
&= \int_{-1}^0 \frac{d}{ds}\big(D_vL(x(s),x'(s))\cdot(s+1)h\big)\, ds + C|h|^2 \\
&= D_vL(x(0),x'(0))\cdot h + C|h|^2 = Du(x)\cdot h + C|h|^2.
\end{aligned}$$

Thus,

$$|u(x+h) - u(x) - Du(x)\cdot h| \le C|h|^2,$$

which completes part (i). For part (ii), note that, for $x, y \in \mathcal{M}_0$,

$$|u(y) - u(x) - Du(x)\cdot(y-x)| \le C|y-x|^2,$$

and

$$|u(x) - u(y) - Du(y)\cdot(x-y)| \le C|y-x|^2.$$

Combine these two and use triangle inequality to conclude. □

From the above theorem, we see that the map $\pi|_{\widetilde{\mathcal{M}}_0} : \widetilde{\mathcal{M}}_0 \to \mathcal{M}_0$ is injective, and its inverse is Lipschitz.

3.2 A relaxed problem

A disadvantage of the flow invariant property (Definition 7.8) is that it is a nonlinear constraint that depends on L (and hence H). For this reason, Mañé [109] proposed a relaxed problem as following

$$\min_{\nu\in\mathcal{F}} \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\,d\nu(x,v),$$

where

$$\mathcal{F} = \left\{ \nu \in \mathcal{P}(\mathbb{T}^n\times\mathbb{R}^n) : \int_{\mathbb{T}^n\times\mathbb{R}^n} v\cdot D\varphi(x)\,d\nu(x,v) = 0 \text{ for every } \varphi\in C^1(\mathbb{T}^n) \right\}.$$

Measures belonging to $\mathcal{F}$ are called *holonomic* measures. Of course, the constraint in $\mathcal{F}$ is a linear constraint, and it is independent of L and H. We first show that $\mathcal{F}$ is a bigger class than flow invariant probability measures.

Lemma 7.17. *Assume* (7.1). *Then, if* $\mu\in\mathcal{P}(\mathbb{T}^n\times\mathbb{R}^n)$ *is a flow invariant measure,* $\mu\in\mathcal{F}$.

Proof. Let $\mu\in\mathcal{P}(\mathbb{T}^n\times\mathbb{R}^n)$ be a flow invariant measure. Fix $\varphi\in C^1(\mathbb{T}^n)$. By the flow invariant property,

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \varphi(\pi\circ\Phi_t(x,v))\,d\mu(x,v) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \varphi(x)\,d\mu(x,v).$$

Thus,

$$\frac{d}{dt}\int_{\mathbb{T}^n\times\mathbb{R}^n} \varphi(\pi\circ\Phi_t(x,v))\,d\mu(x,v) = 0.$$

Note that $\frac{d}{dt}\varphi(x(t)) = D\varphi(x(t))\cdot x'(t)$. Let $t=0$ in the above relation to deduce

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} v\cdot D\varphi(x)\,d\mu(x,v) = 0,$$

which implies that $\mu\in\mathcal{F}$. □

We now show that although $\mathcal{F}$ is bigger than the class of flow invariant probability measures, we still have the same result in the minimization problem as in Theorem 7.9

Theorem 7.18. *Assume* (7.1). *Then,*

$$\overline{H}(0) = -\min_{\nu\in\mathcal{F}} \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\,d\nu(x,v). \tag{7.14}$$

Roughly speaking, this is very close to Theorem 6.6 in the previous chapter.

Proof. Let $u\in \mathrm{Lip}\,(\mathbb{T}^n)$ be a solution to (7.4) with $P=0$. Let η be a standard mollifier, and for $\varepsilon>0$, let $\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(\frac{x}{\varepsilon}\right)$ for $x\in\mathbb{R}^n$. Denote by

$$u^\varepsilon(x) = (\eta_\varepsilon * u)(x) \qquad \text{for } x\in\mathbb{T}^n.$$

Then, $u^\varepsilon\in C^\infty(\mathbb{T}^n)$, $u^\varepsilon\to u$ uniformly in $\mathbb{T}^n$ as $\varepsilon\to 0$, and u^ε satisfies

$$H(x,Du^\varepsilon(x)) \le \overline{H}(0) + C\varepsilon \qquad \text{in } \mathbb{T}^n.$$

By the Legendre transform,

$$v \cdot Du^\varepsilon(x) - L(x,v) \le H(x, Du^\varepsilon(x)) \le \overline{H}(0) + C\varepsilon \qquad \text{for all } x \in \mathbb{T}^n, v \in \mathbb{R}^n.$$

Integrate this with respect to $d\nu$ for any $\nu \in \mathcal{F}$ to get

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\nu(x,v) \ge -\overline{H}(0) - C\varepsilon.$$

Let $\varepsilon \to 0$ to imply first that

$$\min_{\nu \in \mathcal{F}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\nu(x,v) \ge -\overline{H}(0).$$

The reverse inequality follows immediately from the later part of the proof of Theorem 7.9 as $\mathcal{F}$ is bigger than the class of flow invariant probability measures. Nevertheless, let us still repeat the construction here as it is quite important and natural. Fix $x \in \mathbb{T}^n$, and take ξ to be a backward characteristic of u starting from x. We have that, for $t < 0$,

$$u(\xi(0)) - u(\xi(t)) = \int_t^0 \left(L(\xi(s), \xi'(s)) + \overline{H}(0) \right) ds.$$

Define $\mu_t \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ as

$$\langle \mu_t, \psi \rangle = \frac{1}{|t|} \int_t^0 \psi(\xi(s), \xi'(s))\, ds$$

for every bounded continuous function ψ. It is very important noting that $\operatorname{spt}(\mu_t) \subset \mathbb{T}^n \times \overline{B}(0,C)$ for $C > 0$ sufficiently large because of the fact that $\|\xi'\|_{L^\infty((-\infty,0])} \le C$. Then,

$$\frac{u(x) - u(\xi(t))}{|t|} = \langle \mu_t, L \rangle + \overline{H}(0).$$

By compactness, we are able to find a sequence $\{t_k\} \to \infty$ such that $\mu_{t_k} \to \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ weakly in the sense of measures, and $\operatorname{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. The above equality infers that

$$-\overline{H}(0) = \langle \mu, L \rangle = \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\mu(x,v).$$

Let us verify quickly that $\mu \in \mathcal{F}$. For $\varphi \in C^1(\mathbb{T}^n)$, let $\psi(x,v) = v \cdot D\varphi(x)$, and note that

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} v \cdot D\varphi(x)\, d\mu(x,v) = \lim_{k \to \infty} \frac{1}{|t_k|} \int_{t_k}^0 \xi'(s) \cdot D\varphi(\xi(s))\, ds$$
$$= \lim_{k \to \infty} \frac{1}{|t_k|} (\varphi(\xi(0)) - \varphi(\xi(t_k))) = 0.$$

□

Theorem 7.19. *Assume* (7.1). *Let* $\nu \in \mathcal{F}$ *be such that*

$$\overline{H}(0) = -\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, v)\, d\nu(x, v).$$

Then, ν is a Mather measure.

The proof of this is actually quite complicated. Let us give here an outline of the proof. We need the following results.

Lemma 7.20. *Assume* (7.1). *Let* $\nu \in \mathcal{F}$*. Then, for each* $f \in C^1(\mathbb{T}^n)$*,*

$$t \mapsto \int_{\mathbb{T}^n \times \mathbb{R}^n} f(\pi(\Phi_t(x, v))\, d\nu(x, v) \qquad \textit{is constant}.$$

Proof. We note that

$$\frac{d}{dt}(f(\pi(\Phi_t(x, v))))|_{t=0} = \frac{d}{dt}(f(x(t)))|_{t=0} = Df(x(0)) \cdot x'(0) = Df(x) \cdot v.$$

As $\nu \in \mathcal{F}$, we get the desired conclusion. □

Theorem 7.21. *Assume the settings in Theorem 7.19. Let* $u \in \mathrm{Lip}(\mathbb{T}^n)$ *be a solution to* (7.15)*. Then, for* $(x, v) \in \mathrm{spt}(\nu)$*, we have u is differentiable at x and* $Du(x) = D_v L(x, v)$*. Moreover, ν is supported on a Lipschitz graph in* $\mathbb{T}^n \times \mathbb{R}^n$*.*

Proof. By using Lemma 7.20 and approximations, we see that it stills hold for $f \in C(\mathbb{T}^n)$, and in particular,

$$t \mapsto \int_{\mathbb{T}^n \times \mathbb{R}^n} u(\pi(\Phi_t(x, v))\, d\nu(x, v) \qquad \text{is constant.}$$

Let $(x, v) \in \mathrm{spt}(\nu)$, then we use the above to imply that Lemma 7.12 holds for (x, v). Then, repeat Theorem 7.6 to deduce further that u is differentiable at x, and $Du(x) = D_v L(x, v)$, which means $v = D_p H(x, Du(x))$. By abuse of notions, we write $v(x) = D_p H(x, Du(x))$ for $(x, v) \in \mathrm{spt}(\nu)$.

Next, repeating the results in Section 3.1, we obtain that ν is also supported on a Lipschitz graph. Indeed, for $(x, v(x)), (y, v(y)) \in \mathrm{spt}(\nu)$,

$$|Du(x) - Du(y)| \leq C|x - y|,$$

which also means that

$$|Dv(x) - Dv(y)| \leq C|x - y|.$$

□

We are now ready to prove that ν is a Mather measure thanks to the graph theorem above. This proof is taken from Evans [52].

Sketch of proof of Theorem 7.19. Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (7.15).

So far, we have been working with configuration space of (x,v)-variables. For this proof, it is simpler to work with state space of (x,p)-variables. Let $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be such that

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,v)\, d\nu(x,v) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x, D_pH(x,p))\, d\mu(x,p)$$

for all bounded continuous functions ψ. We need to show that μ is flow invariant, that is,

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \{\psi, H\}\, d\mu(x,p) = 0$$

for all smooth bounded functions ψ. Here, $\{\psi, H\}$ denotes the Poisson bracket between ψ and H, that is,

$$\{\psi, H\} = D_p\psi(x,p)\cdot D_xH(x,p) - D_x\psi(x,p)\cdot D_pH(x,p).$$

Let $\phi(x) = \psi(x, Du(x))$. Then, ϕ is Lipschitz on the support of μ. Let us assume ϕ is C^1 for simplicity (else, do the usual convolution trick). As we are only concerned with ϕ and its first-order derivative on spt(μ), everything is fine.

We have $D\phi(x) = D_x\psi + D_p\psi D^2u$. Besides, as $H(x,Du(x)) = \overline{H}(0)$, one gets further that $D_xH + D_pHD^2u = 0$. Thus,

$$\begin{aligned} 0 &= \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH(x,p)\cdot D\phi\, d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH\cdot(D_x\psi + D_p\psi D^2u)\, d\mu(x,p) \\ &= \int_{\mathbb{T}^n\times\mathbb{R}^n} (D_pH\cdot D_x\psi - D_p\psi\cdot D_xH)\, d\mu(x,p). \end{aligned}$$

The proof is complete.

□

4 Nonlinear PDE methods in weak KAM theory

One key point that we see from weak KAM theory is the appearance of Mather measures. We show now that, at least heuristically, Mather measures give rise to a new PDE, which is coupled with our usual cell problem. Recall the cell problem (7.4) at $P = 0$

$$H(x, Du(x)) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n. \tag{7.15}$$

Let $u \in \mathrm{Lip}\,(\mathbb{T}^n)$ be a solution to the above. Let $\mu \in \mathcal{P}(\mathbb{T}^n\times\mathbb{R}^n)$ be a Mather measure, and $\sigma = \pi\circ\mu$, its projection to $\mathbb{T}^n$. Of course, $\mu \in \mathcal{F}$. For $(x,v) \in \mathrm{spt}(\mu)$, we know from the previous section that u is differentiable at x, and $Du(x) = D_vL(x,v)$. Thus, $v = D_pH(x, Du(x))$, and for any test function $\varphi \in C^1(\mathbb{T}^n)$,

$$\begin{aligned} 0 &= \int_{\mathbb{T}^n\times\mathbb{R}^n} v\cdot D\varphi(x)\, d\mu(x,v) \\ &= \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH(x,Du(x))\cdot D\varphi(x)\, d\mu(x,v) = \int_{\mathbb{T}^n} D_pH(x,Du(x))\cdot D\varphi(x) d\sigma(x). \end{aligned}$$

This means that the measure σ is a weak solution of the following transport type equation

$$-\operatorname{div}(D_pH(x,Du(x))\sigma) = 0 \qquad \text{in } \mathbb{T}^n. \tag{7.16}$$

Therefore, to think about weak KAM theory, a correct way is to think of a system of two equations (7.15) and (7.16). Moreover, let us point out here that this is closely related to the nonlinear adjoint method. Indeed, assuming that u is smooth, then the linearized operator of (7.15) around u is

$$\mathcal{L}[\phi](x) = D_pH(x,Du(x)) \cdot D\phi(x) \qquad \text{for all } \phi \in C^1(\mathbb{T}^n).$$

Then, (7.16) is nothing but the adjoint equation to this linearized operator $\mathcal{L}$. Surely, we need to be extremely careful with smoothness issues when handling and interpreting this system, but this important viewpoint, observed by Evans, Gomes [53], allows us to introduce nonlinear PDE methods to weak KAM theory to read off more information.

There have been many different ways to approximate (7.15) and (7.16) and pass to the limits to obtain Mather measures rigorously. We will employ the nonlinear adjoint method here to introduce few such approximations. As this is an introductory chapter, we only introduce some approaches here and do not go too deeply into further aspects of weak KAM theory.

4.1 Vanishing viscosity approximations

Here, we aim at approximating (7.15) by adding a small viscosity term. For each $\varepsilon > 0$, we consider

$$H(x,Du^\varepsilon(x)) = \varepsilon\Delta u^\varepsilon + \overline{H}^\varepsilon(0) \qquad \text{in } \mathbb{T}^n. \tag{7.17}$$

In the equation above, the pair of unknown is $\left(u^\varepsilon, \overline{H}^\varepsilon(0)\right) \in C(\mathbb{T}^n) \times \mathbb{R}$.

Theorem 7.22. *Assume* (7.1). *For every $\varepsilon > 0$, there exists a unique constant $\overline{H}^\varepsilon(0) \in \mathbb{R}$ such that* (7.17) *has a solution $u^\varepsilon \in C(\mathbb{T}^n)$. In fact, u^ε is smooth, and is unique up to additive constants. Furthermore, as $\varepsilon \to 0$,*

$$\lim_{\varepsilon\to 0} \overline{H}^\varepsilon(0) = \overline{H}(0),$$

and there exists a subsequence $\{\varepsilon_k\} \to 0$ such that

$$u^{\varepsilon_k} - \min_{\mathbb{T}^n} u^{\varepsilon_k} \to u \qquad \text{in } C(\mathbb{T}^n),$$

for some $u \in C(\mathbb{T}^n)$, which solves (7.15).

Proof. The existence and uniqueness of $\overline{H}^\varepsilon(0)$ are similar to those of $\overline{H}(0)$. Let us present only the existence of $\overline{H}^\varepsilon(0)$ here as its uniqueness proof follows exactly the same lines of that for $\overline{H}(0)$.

Fix $\varepsilon > 0$. For $\lambda > 0$, we consider

$$\lambda v^\lambda + H(x,Dv^\lambda) = \varepsilon\Delta v^\lambda \qquad \text{in } \mathbb{T}^n.$$

It is clear that the above has a unique smooth solution v^λ, and the comparison principle gives

$$-\|H(\cdot,0)\|_{L^\infty(\mathbb{T}^n)} \le \lambda v^\lambda \le \|H(\cdot,0)\|_{L^\infty(\mathbb{T}^n)}.$$

Let us now obtain bound for Dv^λ via the classical Bernstein method. Let $w^\lambda = \frac{|Dv^\lambda|^2}{2}$, then w^λ satisfies

$$2\lambda w^\lambda + D_pH(x, Dv^\lambda) \cdot Dw^\lambda + D_xH(x, Dv^\lambda) \cdot Dv^\lambda = \varepsilon \Delta w^\lambda - \varepsilon |D^2 v^\lambda|^2.$$

Pick $x_0 \in \mathbb{T}^n$ such that $w^\lambda(x_0) = \max_{\mathbb{T}^n} w^\lambda \geq 0$. Then, by the maximum principle, at x_0,

$$2\lambda w^\lambda + \varepsilon |D^2 v^\lambda|^2 + D_x H \cdot Dv^\lambda \leq 0.$$

For $\varepsilon < n^{-1}$, note that

$$\varepsilon |D^2 v^\lambda|^2 \geq (\varepsilon \Delta v^\lambda)^2 = (\lambda v^\lambda + H(x, Dv^\lambda))^2 \geq \frac{1}{2} H(x, Dv^\lambda)^2 - C.$$

Combine the above two inequalities to yield, at x_0,

$$\frac{1}{2} H(x_0, Dv^\lambda)^2 + D_x H \cdot Dv^\lambda \leq C.$$

Employ (7.1) to imply that $|Dv^\lambda(x_0)| \leq C$. Thus,

$$\|\lambda v^\lambda\|_{L^\infty(\mathbb{T}^n)} + \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \leq C.$$

By the Arzelà–Ascoli theorem, we obtain a sequence $\{\lambda_k\} \to 0$ and $u^\varepsilon \in \mathrm{Lip}\,(\mathbb{T}^n)$ such that, as $k \to \infty$,

$$\begin{cases} v^{\lambda_k} - v^{\lambda_k}(0) \to u^\varepsilon \text{ in } C(\mathbb{T}^n), \\ \lambda_k v^{\lambda_k}(0) \to -c \in \mathbb{R}. \end{cases}$$

By stability of viscosity solutions, u^ε solves

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

As explained, we get further that c is unique, and we denote by $\overline{H}^\varepsilon(0) = c$. Of course, u^ε is smooth, unique up to additive constants, and moreover, $\|Du^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \leq C$. For $\overline{H}^\varepsilon(0)$, we have a clear bound

$$-\|H(\cdot, 0)\|_{L^\infty(\mathbb{T}^n)} \leq \overline{H}^\varepsilon(0) \leq \|H(\cdot, 0)\|_{L^\infty(\mathbb{T}^n)}.$$

Let us now let $\varepsilon \to 0$ to get the second part of the theorem. By the Arzelà–Ascoli theorem again, we obtain a sequence $\{\varepsilon_k\} \to 0$ and $u \in \mathrm{Lip}\,(\mathbb{T}^n)$ such that, as $k \to \infty$,

$$\begin{cases} u^{\varepsilon_k} - \min_{\mathbb{T}^n} u^{\varepsilon_k} \to u \text{ in } C(\mathbb{T}^n), \\ \overline{H}^\varepsilon(0) \to -c \in \mathbb{R}. \end{cases}$$

Use stability of viscosity solutions again to yield that u solves

$$H(x, Du) = c \qquad \text{in } \mathbb{T}^n.$$

Thus, $c = \overline{H}(0)$, which is unique. This means that $\overline{H}^\varepsilon(0) \to \overline{H}(0)$ as $\varepsilon \to 0$ for a full sequence. $\square$

The linearized operator of (7.17) around u^ε is

$$\mathcal{L}^\varepsilon[\phi] = D_pH(x, Du^\varepsilon) \cdot D\phi - \varepsilon\Delta\phi \qquad \text{for all } \phi \in C^2(\mathbb{T}^n).$$

This allows us to consider the adjoint equation to this linearized operator as

$$(\mathcal{L}^\varepsilon)^*[\sigma^\varepsilon] = -\mathrm{div}(D_pH(x, Du^\varepsilon)\sigma^\varepsilon) - \varepsilon\Delta\sigma^\varepsilon = 0 \qquad \text{in } \mathbb{T}^n. \tag{7.18}$$

It is quite clear that 0 is the principal eigenvalue of $(\mathcal{L}^\varepsilon)^*$, and so, (7.18) admits a unique nonnegative solution σ^ε with

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

Denote by $\mu^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ the unique measure such that

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\mu^\varepsilon(x,p) = \int_{\mathbb{T}^n} \psi(x, Du^\varepsilon)\sigma^\varepsilon\,dx$$

for all bounded continuous functions ψ. Note that it is more convenient here for us to work with measures on phase space of (x,p)-variables. Our goal is to let $\varepsilon \to 0$ to obtain Mather measures. Since $\|Du^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \le C$, we get $\mathrm{spt}(\mu^\varepsilon) \subset \mathbb{T}^n \times \overline{B}(0,C)$. So, by compactness, there exists a sequence $\{\varepsilon_k\} \to 0$ such that $\mu^{\varepsilon_k} \to \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ weakly in the sense of measures. Of course, $\mathrm{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. To switch from state space of (x,p)-variables to configuration space of (x,v)-variables, we let $\nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be such that

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x, D_vL(x,v))\,d\nu(x,v)$$

for all bounded continuous functions ψ.

Theorem 7.23. *Assume* (7.1). *Let* $\nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ *be defined as in the procedure above. Then,* ν *is a Mather measure.*

Proof. We first show that $\nu \in \mathcal{F}$. Multiply (7.18) by a test function $\phi \in C^2(\mathbb{T}^n)$ and integrate to have

$$\varepsilon\int_{\mathbb{T}^n} \Delta\phi\,\sigma^\varepsilon\,dx = \int_{\mathbb{T}^n} D_pH(x, Du^\varepsilon)\cdot D\phi(x)\sigma^\varepsilon(x)\,dx = \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH(x,p)\cdot D\phi(x)\,d\mu^\varepsilon(x,p).$$

Let $\varepsilon = \varepsilon_k$ and $k \to \infty$ to yield further that

$$0 = \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH(x,p)\cdot D\phi(x)\,d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} v\cdot D\phi(x)\,d\nu(x,v).$$

By approximations, we get that the above holds for all $\phi \in C^1(\mathbb{T}^n)$. Thus, $\nu \in \mathcal{F}$. We show next that

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\,d\nu(x,v) = -\overline{H}(0).$$

Multiply (7.18) by u^ε and integrate to have

$$\varepsilon\int_{\mathbb{T}^n} \Delta u^\varepsilon\sigma^\varepsilon\,dx = \int_{\mathbb{T}^n} D_pH(x, Du^\varepsilon)\cdot Du^\varepsilon\sigma^\varepsilon(x)\,dx = \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH(x,p)\cdot p\,d\mu^\varepsilon(x,p).$$

Next, multiply (7.17) by σ^ε and integrate

$$\overline{H}^\varepsilon(0) = \int_{\mathbb{T}^n} (H(x, Du^\varepsilon) - \varepsilon\Delta u^\varepsilon)\sigma^\varepsilon(x)\,dx = \int_{\mathbb{T}^n\times\mathbb{R}^n} H(x,p)\,d\mu^\varepsilon(x,p) - \int_{\mathbb{T}^n} \varepsilon\Delta u^\varepsilon\sigma^\varepsilon(x)\,dx.$$

Combine the two above to imply

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} (D_pH(x,p)\cdot p - H(x,p))\,d\mu^\varepsilon(x,p) = -\overline{H}^\varepsilon(0).$$

Note that $\overline{H}^\varepsilon(0) \to \overline{H}(0)$ as $\varepsilon \to 0$. By letting $\varepsilon = \varepsilon_k$ and $k \to \infty$, we deduce that

$$-\overline{H}(0) = \int_{\mathbb{T}^n\times\mathbb{R}^n} (D_pH(x,p)\cdot p - H(x,p))\,d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\,d\nu(x,v).$$

□

We have furthermore the following estimate. This is a L^2 version of the Lipschitz graph theorem.

Lemma 7.24. *Assume* (7.1). *Then, there exists* $C > 0$ *independent of* ε *such that*

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

Proof. For each $1 \le i \le n$, differentiate (7.17) with respect to x_i twice to get

$$D_pH\cdot Du^\varepsilon_{x_ix_i} + H_{p_kp_l}u^\varepsilon_{x_kx_i}u^\varepsilon_{x_lx_i} + H_{x_ix_i} + 2H_{x_ip_k}u^\varepsilon_{x_kx_i} = \varepsilon\Delta u^\varepsilon_{x_ix_i}.$$

By the uniform convexity of H in p (assumption (7.1)), we simplify the above as

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

Multiply this inequality with σ^ε and integrate over $\mathbb{T}^n$ to deduce

$$\frac{\theta}{2}\int_{\mathbb{T}^n} |Du^\varepsilon_{x_i}|^2\sigma^\varepsilon\,dx \le C.$$

Sum this over $i = 1, 2, \ldots, n$ to conclude. □

4.2 Large time average approximations and applications

We present here another way to obtain Mather measures and give an application.

Let $u \in \mathrm{Lip}\,(\mathbb{T}^n)$ be a solution to (7.15). Our aim is to use large time average of solutions to derive Mather measures. Consider

$$\begin{cases} \varphi_t + H(x, D\varphi) &= 0 \qquad \text{in } \mathbb{T}^n\times(0,\infty),\\ \varphi(x,0) &= u(x) \qquad \text{on } \mathbb{T}^n. \end{cases}$$

Then, $\varphi(x,t) = u(x) - \overline{H}(0)t$ is the unique solution to the above. Instead of letting $t \to \infty$ directly in the above, we rescale the problem as $w(x,t) = \varphi(x, \frac{t}{\varepsilon})$ for $(x,t) \in \mathbb{T}^n \times [0,\infty)$ and $\varepsilon > 0$. Then, w solves

$$\begin{cases} \varepsilon w_t + H(x, Dw) &= 0 \qquad \text{in } \mathbb{T}^n \times (0,\infty), \\ w(x,0) &= u(x) \qquad \text{on } \mathbb{T}^n. \end{cases} \tag{7.19}$$

It is clear that $w(x,t) = u(x) - \frac{\overline{H}(0)t}{\varepsilon}$ for $(x,t) \in \mathbb{T}^n \times [0,\infty)$. Our goal is to let $\varepsilon \to 0$ to see the large time average of φ to get Mather measures. For simplicity, let us normalize to have $\overline{H}(0) = 0$ always in this section. Then, $w(x,t) = u(x)$ for $(x,t) \in \mathbb{T}^n \times [0,\infty)$.
As u is not smooth, we first smooth it up as usual. Let $\rho \in C_c^\infty(\mathbb{R}^n, [0,\infty))$ be a standard mollifier. For $\delta > 0$, let $\rho^\delta(x) = \delta^{-n}\rho(\delta^{-1}x)$ for all $x \in \mathbb{R}^n$. Denote by $u^\delta = \rho^\delta * u$. Then,

$$\|u^\delta - u\|_{L^\infty(\mathbb{T}^n)} \le C\delta,$$

and

$$\|Du^\delta\|_{L^\infty(\mathbb{T}^n)} + \delta \|D^2 u^\delta\|_{L^\infty(\mathbb{T}^n)} \le C.$$

Let us consider the following Cauchy problems

$$\begin{cases} \varepsilon w^\varepsilon_t + H(x, Dw^\varepsilon) = \varepsilon^4 \Delta w^\varepsilon \qquad \text{in } \mathbb{T}^n \times (0,1), \\ w^\varepsilon(x,0) = u^{\varepsilon^4}(x) \qquad \text{on } \mathbb{T}^n, \end{cases} \tag{7.20}$$

and

$$\begin{cases} \varepsilon \phi^\varepsilon_t + H(x, D\phi^\varepsilon) = \varepsilon^4 \Delta \phi^\varepsilon \qquad \text{in } \mathbb{T}^n \times (0,1), \\ \phi^\varepsilon(x,0) = u(x) \qquad \text{on } \mathbb{T}^n. \end{cases} \tag{7.21}$$

Here, u^{ε^4} is u^δ with $\delta = \varepsilon^4$. As $\|u^{\varepsilon^4} - u\|_{L^\infty(\mathbb{T}^n)} \le C\varepsilon^4$, it is straightforward that

$$\|w^\varepsilon - \phi^\varepsilon\|_{L^\infty(\mathbb{T}^n \times [0,1])} \le C\varepsilon^4.$$

The next result concerns gradient bound of w^ε.

Lemma 7.25. *Assume* (7.1). *There is a constant $C > 0$ independent of $\varepsilon > 0$ such that*

$$\varepsilon \|w^\varepsilon_t\|_{L^\infty(\mathbb{T}^n \times [0,1])} + \|Dw^\varepsilon\|_{L^\infty(\mathbb{T}^n \times [0,1])} \le C.$$

Proof. Denote by

$$\varphi^\pm(x,t) = w^\varepsilon(x,0) \pm \frac{C}{\varepsilon} t \qquad \text{for all } (x,t) \in \mathbb{T}^n \times [0,1].$$

Then, φ^-, φ^+ are, respectively, a subsolution, and a supersolution to (7.20). Hence, by the comparison principle,

$$\varphi^- \le w^\varepsilon \le \varphi^+ \quad \Longrightarrow \quad \|w^\varepsilon(\cdot,s) - w^\varepsilon(\cdot,0)\|_{L^\infty} \le \frac{Cs}{\varepsilon}.$$

Note next that both w^ε and $w^\varepsilon(\cdot, \cdot + s)$ solve (7.20) with initial data $w^\varepsilon(\cdot,0)$ and $w^\varepsilon(\cdot,s)$, respectively. By the comparison principle,

$$\|w^\varepsilon(\cdot, \cdot + s) - w^\varepsilon\|_{L^\infty} \le \|w^\varepsilon(\cdot,s) - w^\varepsilon(\cdot,0)\|_{L^\infty} \le \frac{Cs}{\varepsilon} \quad \Longrightarrow \quad \varepsilon \|w^\varepsilon_t\|_{L^\infty(\mathbb{T}^n)} \le C.$$

To prove the spatial gradient bound, we use the usual Bernstein method. Let $\psi(x,t) = \frac{|Dw^\varepsilon|^2}{2}$. Then ψ satisfies

$$\varepsilon\psi_t + D_pH \cdot D\psi + D_xH \cdot Dw^\varepsilon = \varepsilon^4\Delta\psi - \varepsilon^4|D^2w^\varepsilon|^2.$$

Assume that $\max_{\mathbb{T}^n\times[0,1]}\psi = \psi(x_0,t_0)$. If $t_0 = 0$, then we are done. If $t_0 > 0$, then by the maximum principle,

$$D_xH \cdot Dw^\varepsilon + \varepsilon^4|D^2w^\varepsilon|^2 \le 0 \quad \text{at } (x_0,t_0).$$

For $\varepsilon < n^{-1}$, we have

$$\varepsilon^4|D^2w^\varepsilon|^2 \ge (\varepsilon^4\Delta w^\varepsilon)^2 = (\varepsilon w^\varepsilon_t + H(x,Dw^\varepsilon))^2 \ge \frac{1}{2}H(x,Dw^\varepsilon)^2 - C.$$

Therefore,

$$\frac{1}{2}H(x,Dw^\varepsilon)^2 + D_xH \cdot Dw^\varepsilon \le C \quad \text{at } (x_0,t_0),$$

which, together with (7.1), yields the desired result. □

Lemma 7.26. *Assume* (7.1). *Normalize so that* $\overline{H}(0) = 0$. *We have*

$$\|w^\varepsilon - u\|_{L^\infty(\mathbb{T}^n\times[0,1])} + \|\phi^\varepsilon - u\|_{L^\infty(\mathbb{T}^n\times[0,1])} \le C\varepsilon.$$

The proof of this is similar to that of Theorem 1.38. As we have not presented such proofs for Cauchy problems, let us give it here.

Proof. We only need to show that $\|w^\varepsilon - u\|_{L^\infty(\mathbb{T}^n\times[0,1])} \le C\varepsilon$. Let us first get an upper bound for $w^\varepsilon - u$. Define an auxiliary function

$$\Phi(x,y,t) = w^\varepsilon(x,t) - u(y) - \frac{|x-y|^2}{2\varepsilon^2} - K\varepsilon t \quad \text{for } (x,y,t) \in \mathbb{T}^n \times \mathbb{T}^n \times [0,1],$$

where $K > 0$ is to be chosen. Pick $(x_\varepsilon, y_\varepsilon, t_\varepsilon) \in \mathbb{T}^n \times \mathbb{T}^n \times [0,1]$ so that

$$\Phi(x_\varepsilon,y_\varepsilon,t_\varepsilon) = \max_{\mathbb{T}^n\times\mathbb{T}^n\times[0,1]} \Phi.$$

If $\Phi(x_\varepsilon,y_\varepsilon,t_\varepsilon) \le 0$, then we are done as

$$w^\varepsilon(x,t) - u(x) = \Phi(x,x,t) + K\varepsilon t \le K\varepsilon.$$

Therefore, we can assume $\Phi(x_\varepsilon,y_\varepsilon,t_\varepsilon) > 0$. This gives that $w^\varepsilon(x_\varepsilon,t_\varepsilon) > u(y_\varepsilon)$.
Let us consider first the case that $t_\varepsilon > 0$. Since w^ε and u are Lipschitz in space with constant C, by comparing $\Phi(x_\varepsilon,y_\varepsilon,t_\varepsilon)$ with $\Phi(y_\varepsilon,y_\varepsilon,t_\varepsilon)$, we deduce first that

$$|x_\varepsilon - y_\varepsilon| \le C\varepsilon^2.$$

By the viscosity subsolution and supersolution tests, we have

$$K\varepsilon^2 + H\left(x_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon^2}\right) \le \varepsilon^4\frac{n}{\varepsilon^2} = n\varepsilon^2,$$

and

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

Combine these two inequalities, and use (7.1) to imply

$$\begin{aligned} K\varepsilon^2 &\le n\varepsilon^2 + H\left(y_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon^2}\right) - H\left(x_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon^2}\right) \\ &\le n\varepsilon^2 + C|y_\varepsilon - x_\varepsilon| \le (C+n)\varepsilon^2. \end{aligned}$$

By picking $K = C + n + 1$, we conclude that t_ε cannot be positive. Thus, $t_\varepsilon = 0$, and

$$\Phi(x_\varepsilon, y_\varepsilon, t_\varepsilon) \le u^{\varepsilon^4}(x_\varepsilon) - u(y_\varepsilon) \le C\varepsilon^4 + C|x_\varepsilon - y_\varepsilon| \le C\varepsilon^2.$$

Then, for $(x,t) \in \mathbb{T}^n \times [0,1]$,

$$w^\varepsilon(x,t) - u(x) = \Phi(x,x,t) + K\varepsilon t \le C\varepsilon^2 + K\varepsilon \le C\varepsilon.$$

To get the other bound, we need to get an upper bound of $u - w^\varepsilon$. This can be done by repeating the above steps carefully for another auxiliary function

$$\Psi(x,y,t) = u(x) - w^\varepsilon(y,t) - \frac{|x-y|^2}{2\varepsilon^2} - K\varepsilon t \quad \text{for } (x,y,t) \in \mathbb{T}^n \times \mathbb{T}^n \times [0,1],$$

where $K > 0$ is to be chosen. We omit the proof of this part here. □

Remark 7.27. All the above steps are mainly to show that instead of working with (7.19) directly, we can work with (7.20), which has the unique smooth solution w^ε for each $\varepsilon > 0$. The fact that w^ε stays close to u means that there is no complication here, and as we let $\varepsilon \to 0$, we are able to obtain Mather measures for (7.15) via the nonlinear adjoint method described below.

The linearized operator of (7.20) about the solution w^ε is

$$\mathcal{L}^\varepsilon[\phi] = \varepsilon\phi_t + D_pH(x, Dw^\varepsilon) \cdot D\phi - \varepsilon^4 \Delta\phi.$$

The corresponding adjoint equation is

$$\begin{cases} -\varepsilon\sigma^\varepsilon_t - \operatorname{div}(D_pH(x,Dw^\varepsilon)\sigma^\varepsilon) = \varepsilon^4\Delta\sigma^\varepsilon & \text{in } \mathbb{T}^n \times (0,1), \\ \sigma^\varepsilon(x,1) = \delta_{x_0}. \end{cases} \tag{7.22}$$

Here, δ_{x_0} is the Dirac delta measure at $x_0 \in \mathbb{T}^n$. It is clear that $\sigma^\varepsilon > 0$ in $\mathbb{T}^n \times (0,1)$. Basically, σ^ε is the fundamental solution to the above backward parabolic equation in $\mathbb{T}^n \times (0,1)$.

Lemma 7.28. *The following holds*

$$\int_{\mathbb{T}^n} \sigma^\varepsilon(x,t)\,dx = 1 \qquad \text{for all } t \in (0,1).$$

Proof. For $t \in (0,1)$, integrate (7.22) on $\mathbb{T}^n$ to yield

$$\varepsilon \frac{d}{dt} \int_{\mathbb{T}^n} \sigma^\varepsilon\,dx = \int_{\mathbb{T}^n} -\operatorname{div}(D_pH(x, w^\varepsilon, Dw^\varepsilon)\sigma^\varepsilon) - \varepsilon^4 \Delta\sigma^\varepsilon\,dx = 0,$$

which gives the result. □

For each σ^ε, there exists a unique measure $\mu^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ satisfying

$$\int_0^1 \int_{\mathbb{T}^n} \psi(x, Du^\varepsilon)\sigma^\varepsilon(x,t)\,dx\,dt = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\mu^\varepsilon(x,p)$$

for all bounded continuous functions ψ. By a priori estimates, $\mathrm{spt}(\mu^\varepsilon) \subset \mathbb{T}^n \times \overline{B}(0,C)$. We are able to pick a subsequence $\{\varepsilon_j\} \to 0$ such that $\mu^{\varepsilon_j} \to \mu$ as $j \to \infty$ weakly in the sense of measures. Surely, $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ and $\mathrm{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. Then, as above, let $\nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be such that

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x, D_v L(x,v))\,d\nu(x,v)$$

for all bounded continuous functions ψ.

Theorem 7.29. *Assume* (7.1). *Normalize so that* $\overline{H}(0) = 0$. *Let* ν *be constructed as above. Then,* ν *is a Mather measure.*

Proof. The proof is quite similar to that of Theorem 7.23. First, we prove $\nu \in \mathcal{F}$. Multiply (7.22) with $\phi \in C^2(\mathbb{T}^n)$ and integrate to imply

$$\varepsilon \int_{\mathbb{T}^n} \phi(x)\sigma^\varepsilon(x,0)\,dx - \varepsilon\phi(x_0) + \int_0^1\int_{\mathbb{T}^n} D_pH(x,Dw^\varepsilon)\cdot D\phi(x)\sigma^\varepsilon(x,t)\,dx\,dt$$
$$= \varepsilon^4 \int_0^1 \int_{\mathbb{T}^n} \Delta\phi(x)\sigma^\varepsilon(x,t)\,dx\,dt.$$

Let $\varepsilon = \varepsilon_j$ and $j \to \infty$, then

$$0 = \int_{\mathbb{T}^n\times\mathbb{R}^n} D_pH(x,p)\cdot D\phi(x)\,d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} v\cdot D\phi(x)\,d\nu(x,v).$$

We then use usual approximations to get that the above holds for all $\phi \in C^1(\mathbb{T}^n)$, and so, $\nu \in \mathcal{F}$.

Next, multiply (7.20) by σ^ε, multiply (7.22) by w^ε, combine them and integrate to infer

$$\varepsilon w^\varepsilon(x_0,1) - \varepsilon \int_{\mathbb{T}^n} w^\varepsilon(x,0)\sigma^\varepsilon(x,0)\,dx$$
$$= \int_0^1\int_{\mathbb{T}^n} (D_pH(x,Dw^\varepsilon)\cdot Dw^\varepsilon - H(x,Dw^\varepsilon))\sigma^\varepsilon(x,t)\,dx\,dt.$$

Again, let $\varepsilon = \varepsilon_j$ and $j \to \infty$, then

$$0 = \int_{\mathbb{T}^n\times\mathbb{R}^n} (D_pH(x,p)\cdot p - H(x,p))\,d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)\,d\nu(x,v).$$

□

We see that Mather measures were constructed quite naturally through the above two different viewpoints. It is surely the case that one needs to handle approximations carefully and rigorously, but other than that, the nonlinear adjoint method gives Mather measures quite straightforwardly in the limits. Since this chapter is of introductory type, we will not go deeper to study further properties of approximated solutions and corresponding measures. Instead, we present here a quick application of this PDE approach.

Theorem 7.30. *Assume* (7.1). *Normalize so that* $\overline{H}(0) = 0$. *Let* $u, \bar{u} \in \mathrm{Lip}(\mathbb{T}^n)$ *be two solutions to* (7.15). *Assume further that*

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \bar{u}\, d\nu(x,v) \le \int_{\mathbb{T}^n \times \mathbb{R}^n} u\, d\nu(x,v)$$

for all Mather measures ν. *Then,* $\bar{u} \le u$.

This theorem is a variant of Theorem 7.14. As we see right away in the proof below, the approach here is quite different. This uniqueness result is taken from Mitake, Tran [117].

Proof. Consider (7.20) and (7.22) as above with solutions w^ε and σ^ε, respectively. Let $\bar{w}^\varepsilon$ be the solution to (7.20) with initial data $\bar{w}^\varepsilon(x,0) = \bar{u}^{\varepsilon^4}$. Compare $\bar{w}^\varepsilon$ with w^ε and use convexity of H to get that

$$\mathcal{L}^\varepsilon[\bar{w}^\varepsilon - w^\varepsilon] \le 0.$$

Multiply this by σ^ε and integrate to yield

$$\frac{d}{dt}\int_{\mathbb{T}^n} (\bar{w}^\varepsilon - w^\varepsilon)\sigma^\varepsilon\, dx \le 0.$$

Thus,

$$(\bar{w}^\varepsilon - w^\varepsilon)(x_0, 1) \le \int_0^1 \int_{\mathbb{T}^n} (\bar{w}^\varepsilon - w^\varepsilon)\sigma^\varepsilon\, dx dt.$$

Let $\varepsilon = \varepsilon_j$ and $j \to \infty$, we obtain

$$\bar{u}(x_0) - u(x_0) \le \int_{\mathbb{T}^n \times \mathbb{R}^n} (\bar{u} - u)\, d\nu(x,v) \le 0.$$

Hence, $\bar{u}(x_0) \le u(x_0)$. As x_0 is arbitrary, $\bar{u} \le u$. □

5 The projected Aubry set

5.1 The PDE viewpoint

We now use the maximal subsolutions to define the projected Aubry set and study its properties. Maximal subsolutions were already studied in Section 7 of Chapter 2 in the general convex setting in $\mathbb{R}^n$. We here focus on the periodic setting, that is, our equations and maximal subsolutions are considered in $\mathbb{T}^n$.

Let us recall the cell problem (7.4) at $P = 0$

$$H(x, Du(x)) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n. \tag{7.23}$$

For $x, y \in \mathbb{T}^n$, denote by

$$S(x,y) = \sup\{v(x) - v(y) : v \in \mathrm{Lip}(\mathbb{T}^n) \text{ is a subsolution to (7.23)}\}.$$

We use S here instead of m_μ earlier. In our notations, we use the second slot in $S(\cdot,\cdot)$ as a fixed vertex, and geometrically, $x \mapsto S(x,y)$ looks like a bending upward cone with vertex y for x close to y. Of course, $x \mapsto S(x,y)$ does not look like a global cone as it is periodic in x. Sometimes, people would reverse the order of x and y in the literature.

Let us recall the results in Theorem 2.39.

Theorem 7.31. *Assume* (7.1). *The following properties hold.*

(i) For each $y \in \mathbb{T}^n$, $x \mapsto S(x,y)$ is Lipschitz and is the maximal solution to

$$\begin{cases} H(x, Du(x)) = \overline{H}(0) & \text{in } \mathbb{T}^n \setminus \{y\}, \\ u(y) = 0. \end{cases} \tag{7.24}$$

In particular, $S(y,y) = 0$.

(ii) For $x, y, z \in \mathbb{T}^n$,

$$S(x,y) + S(y,z) \geq S(x,z). \tag{7.25}$$

Of course, we have also discussed that $x \mapsto S(x,y)$ needs not be a solution to (7.23) and it might fail the viscosity supersolution test at the vertex y. This leads us to the following definition.

Definition 7.32 (Projected Aubry set)**.** *Denote by*

$$\mathcal{A} = \{y \in \mathbb{T}^n : x \mapsto S(x,y) \text{ is a solution to (7.23)}\}.$$

We say that $\mathcal{A}$ is the projected Aubry set corresponding to $P = 0$.

Roughly speaking, $\mathcal{A}$ contains all the good vertices y at which the viscosity supersolution test for $S(x,y)$ holds. We first need to show that $\mathcal{A}$ is not empty.

Proposition 7.33. *Assume* (7.1). *Then, $\mathcal{A} \neq \emptyset$.*

We use some ideas of the Perron method in the proof.

Proof. Assume by contradiction that $\mathcal{A} = \emptyset$. Then, for each $y \in \mathbb{T}^n$, the viscosity supersolution test for $S(x,y)$ fails at $x = y$. This means that we can find a smooth test function ϕ with $\phi(y) = 0$, $x \mapsto S(x,y) - \phi(x)$ has a strict minimum at y, and

$$H(y, D\phi(y)) < \overline{H}(0).$$

There exist $r, \varepsilon_y > 0$ sufficiently small such that

$$\begin{cases} \phi(x) < S(x,y) - \varepsilon_y & \text{for all } x \in \partial B(y,r), \\ H(x, D\phi(x)) < \overline{H}(0) - \varepsilon_y & \text{for all } x \in B(y,r). \end{cases}$$

Denote by

$$\psi_y(x) = \begin{cases} \max\left\{S(x,y), \phi(x) + \varepsilon_y\right\} & \text{for all } x \in B(y,r), \\ S(x,y) & \text{for all } x \in \mathbb{T}^n \setminus B(y,r) \end{cases}$$

Clearly, $\psi_y \in \mathrm{Lip}(\mathbb{T}^n)$, ψ_y is a subsolution to (7.23), and there exists $r_y \in (0, r)$ such that

$$H(x, D\psi_y(x)) < \overline{H}(0) - \varepsilon_y \qquad \text{for all } x \in B(y, r_y). \tag{7.26}$$

Of course,

$$\mathbb{T}^n \subset \bigcup_{y \in \mathbb{T}^n} B(y, r_y).$$

By the compactness of $\mathbb{T}^n$, we are able to find $y_1, \ldots, y_k \in \mathbb{T}^n$ such that

$$\mathbb{T}^n \subset \bigcup_{i=1}^{k} B(y_i, r_{y_i}).$$

We then set

$$\psi = \frac{1}{k}\sum_{i=1}^{k} \psi_{y_i}, \qquad \varepsilon = \frac{1}{k}\min_{1\le i\le k} \varepsilon_{y_i}.$$

In light of (7.26) and the convexity of H in p, we have that

$$H(x, D\psi(x)) \le \overline{H}(0) - \varepsilon \qquad \text{in } \mathbb{T}^n,$$

which gives a contradiction to the representation formula of $\overline{H}(0)$ in Theorem 4.10. We therefore conclude that $\mathcal{A} \neq \emptyset$. □

Remark 7.34. We give some comments about the proof of Proposition 7.33. One can see the ideas of the Perron method used in the construction of ψ_y for $y \in \mathbb{T}^n$ above quite clearly. We then use the compactness of $\mathbb{T}^n$ crucially in the next step.

In term of assumptions on H, we actually do not need to assume (7.1) fully. We only need that $H \in C(\mathbb{T}^n \times \mathbb{R}^n)$, and $p \mapsto H(x, p)$ is convex, and coercive uniformly for $x \in \mathbb{T}^n$.

Theorem 7.35. *Assume* (7.1). *Then, $\mathcal{A}$ is a nonempty, compact subset of $\mathbb{T}^n$.*

Proof. By Proposition 7.33, we already have that $\mathcal{A}$ is not empty.

To finish off, we only need to show that $\mathcal{A}$ is closed, which is a rather straightforward from the stability of viscosity solutions. Indeed, pick a sequence $\{y_k\} \subset \mathcal{A} \subset \mathbb{T}^n$. There exists a subsequence $\{y_{k_j}\}$ of $\{y_k\}$ that converges to some $y \in \mathbb{T}^n$. We have that $S(\cdot, y_{k_j})$ is a viscosity solution of (7.23) and $S(y_{k_j}, y_{k_j}) = 0$. By the coercivity of H, we can find a constant $C > 0$ independent of $j \in \mathbb{N}$ such that

$$\|S(\cdot, y_{k_j})\|_{L^\infty(\mathbb{T}^n)} + \|DS(\cdot, y_{k_j})\|_{L^\infty(\mathbb{T}^n)} \le C.$$

By the usual Arzelà-Ascoli theorem, by passing to another subsequence if necessary, we might assume that $S(\cdot, y_{k_j})$ converges to some $w \in \mathrm{Lip}(\mathbb{T}^n)$ uniformly as $j \to \infty$. It is clear that $w(y) = 0$ and w is a viscosity solution of (7.23). Moreover, as

$$S(x, y_{k_j}) \ge S(x, y) - S(y_{k_j}, y),$$

we let $j \to \infty$ to deduce that $w(x) \ge S(x, y)$ for all $x \in \mathbb{T}^n$. By the definitions of $S(\cdot, y)$ and the projected Aubry set, we see that $w = S(\cdot, y)$ and also that $y \in \mathcal{A}$. The proof is complete. □

Let us now give one example in which we know precisely what is the projected Aubry set.

Example 7.1. *Assume that*

$$H(x,p) = |p| - V(x) \qquad \text{for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n,$$

where $V \in C(\mathbb{T}^n)$ is given such that $\min_{\mathbb{T}^n} V = 0$. This Hamiltonian does not satisfy (7.1)*, but it is enough here as $p \mapsto H(x,p)$ is convex, and coercive uniformly for $x \in \mathbb{T}^n$.*

We have shown that $\overline{H}(0) = 0$. By Proposition 2.37, $S(\cdot, y)$ is a viscosity solution of (7.23) *if and only if $V(y) = 0$. It is therefore clear that*

$$\mathcal{A} = \left\{ y \in \mathbb{T}^n \, : \, V(y) = \min_{\mathbb{T}^n} V = 0 \right\}.$$

We have another characterization of $\mathcal{A}$ as following.

Proposition 7.36. *Assume* (7.1). *Then, for $y \in \mathbb{T}^n$, $y \notin \mathcal{A}$ if and only if there exists a subsolution $w \in \mathrm{Lip}(\mathbb{T}^n)$ to* (7.23) *which is strict at y, that is, there exists $q \in D^- w(y)$ such that $H(y,q) < \overline{H}(0)$.*

Proof. Firstly, if $y \notin \mathcal{A}$, then by the first part of the proof of Proposition 7.33, we let $w = \psi_y$ to conclude.

Let us now assume that there is a subsolution $w \in \mathrm{Lip}(\mathbb{T}^n)$ to (7.23) which is strict at y for some given $y \in \mathbb{T}^n$. Then, there exists $q \in D^- w(y)$ such that $H(y,q) < \overline{H}(0)$. By the definition of $S(\cdot, y)$, we see that

$$S(x,y) \geq w(x) - w(y) \qquad \text{for all } y \in \mathbb{T}^n,$$

which yields that $q \in D^- S(y,y)$. This means that $S(\cdot, y)$ is not a solution to (7.23), and therefore, $y \notin \mathcal{A}$. □

Next, we show that we are able to construct a subsolution to (7.23) which is strict outside of the projected Aubry set.

Proposition 7.37. *Assume* (7.1). *Then, there exists a subsolution $w \in \mathrm{Lip}(\mathbb{T}^n)$ to* (7.23) *which is strict in $\mathbb{T}^n \setminus \mathcal{A}$. More precisely, for each open set U such that $U \subset\subset \mathbb{T}^n \setminus \mathcal{A}$, there exists $\varepsilon_U > 0$ such that*

$$H(x, Dw(x)) \leq \overline{H}(0) - \varepsilon_U \qquad \text{in } U.$$

Proof. Part of the proof here was already done in the proof of Proposition 7.33. For each $y \in \mathbb{T}^n \setminus \mathcal{A}$, there are $\psi_y \in \mathrm{Lip}(\mathbb{T}^n)$ and $r_y, \varepsilon_y > 0$ such that ψ_y is a subsolution to (7.23), and $B(y, r_y) \subset \mathbb{T}^n \setminus \mathcal{A}$, and

$$H(x, D\psi_y(x)) < \overline{H}(0) - \varepsilon_y \qquad \text{for all } x \in B(y, r_y).$$

Since

$$\mathbb{T}^n \setminus \mathcal{A} \subset \bigcup_{y \in \mathbb{T}^n \setminus \mathcal{A}} B(y, r_y),$$

we are able to find a sequence $\{y_k\}_{k \in \mathbb{N}} \subset \mathbb{T}^n \setminus \mathcal{A}$ such that

$$\mathbb{T}^n \setminus \mathcal{A} \subset \bigcup_{i=1}^{\infty} B(y_i, r_{y_i}).$$

Set
$$w(x) = \sum_{i=1}^{\infty} 2^{-i} \psi_{y_i}(x) \qquad \text{for } x \in \mathbb{T}^n.$$
It is straightforward that $w \in \mathrm{Lip}(\mathbb{T}^n)$ is a viscosity subsolution to (7.23) which is strict in $\mathbb{T}^n \setminus \mathcal{A}$. □

We are now ready to present another comparison result for solutions to (7.23). In a way, this is quite similar to Theorems 7.14 and 7.30.

Theorem 7.38. *Assume* (7.1). *Let $u_1, u_2 \in \mathrm{Lip}(\mathbb{T}^n)$ be a subsolution and a supersolution to* (7.23)*, respectively. Assume further that $u_1 \le u_2$ on $\mathcal{A}$. Then, $u_1 \le u_2$.*

Proof. To obtain the result, we show that $u_1 \le u_2 + \delta$ for each given $\delta > 0$.

Let $\bar{u}_2 = u_2 + \delta$, and w be as in Proposition 7.37. Since $u_1 \le u_2$ on $\mathcal{A}$, there exists an open set U such that $U \subset\subset \mathbb{T}^n \setminus \mathcal{A}$ such that
$$u_1 \le u_2 + \frac{\delta}{4} = \bar{u}_2 - \frac{3\delta}{4} \qquad \text{on } \mathbb{T}^n \setminus U.$$
For $s \in (0,1)$, denote by $\bar{u}_1 = s u_1 + (1-s) w$. For s quite close to 1, we have that
$$\bar{u}_1 \le \bar{u}_2 - \frac{\delta}{2} \qquad \text{on } \mathbb{T}^n \setminus U. \tag{7.27}$$
Besides,
$$H(x, D\bar{u}_1(x)) \le sH(x, Du_1(x)) + (1-s)H(x, Dw(x)) \le \overline{H}(0) - (1-s)\varepsilon_U \qquad \text{in } U. \tag{7.28}$$
Thanks to (7.27) and (7.28), we can find $\lambda > 0$ sufficiently small such that
$$\lambda \bar{u}_1 + H(x, D\bar{u}_1) \le \lambda \bar{u}_2 + H(x, D\bar{u}_2) \qquad \text{in } \mathbb{T}^n.$$
Thus, $\bar{u}_1 \le \bar{u}_2$ in light of the usual comparison principle for this static Hamilton–Jacobi equation. Let $s \to 1$ and $\delta \to 0$ in this order to conclude. □

5.2 A representation formula for solutions to (7.23)

We have shown in Theorem 7.38 that $\mathcal{A}$ is a uniqueness set for (7.23). Let us now proceed further to give a new representation formula to solutions to (7.23) based on data on $\mathcal{A}$.

Theorem 7.39. *Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (7.23). *Then, for every $x \in \mathbb{T}^n$,*
$$u(x) = \min_{y \in \mathcal{A}} (u(y) + S(x,y)).$$

Proof. Let $v(x) = \min_{y \in \mathcal{A}} (u(y) + S(x,y))$ for $x \in \mathbb{T}^n$. As $x \mapsto u(y) + S(x,y)$ is a solution to (7.23) for each $y \in \mathcal{A}$, we imply that v is also a solution to (7.23) thanks to Corollary 2.31.

Moreover, by the definition of $S(x,y)$, we see that
$$u(x) - u(y) \le S(x,y),$$
which means
$$u(x) \le \min_{y \in \mathcal{A}} (u(y) + S(x,y)) = v(x).$$
In particular, for $x \in \mathcal{A}$, as $u(x) + S(x,x) = u(x)$, we deduce that $u(x) = v(x)$. Thus,
$$u = v \qquad \text{on } \mathcal{A}.$$
By Theorem 7.38, we conclude that $u = v$. □

5.3 The Lagrangian viewpoint

We use the representation formula for $S(x,y)$ studied in Section 7 of Chapter 2 to discuss the Lagrangian viewpoint of the projected Aubry set $\mathcal{A}$. Let L be the Lagrangian corresponding to this H. Here is the formula of $S(x,y)$ thanks to Theorem 2.41.

Theorem 7.40. *Assume* (7.1). *For* $x,y \in \mathbb{T}^n$,

$$\begin{aligned} &S(x,y) \\ &= \inf\left\{\int_0^t \left(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\right) ds \,:\, \gamma \in \mathrm{AC}([0,t],\mathbb{T}^n) \text{ for } t>0, \gamma(0)=y, \gamma(t)=x\right\}. \end{aligned} \tag{7.29}$$

In the weak KAM theory literature, S is also called the critical Mañé potential. We give the following equivalent characterization of points in the Aubry set. This characterization is rather important geometrically.

Theorem 7.41. *Assume* (7.1). *Then,* $y \in \mathcal{A}$ *if and only if*

$$\inf\left\{\int_0^t \left(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\right) ds \,:\, \gamma \in \mathrm{AC}([0,t],\mathbb{T}^n) \text{ for } t>\delta, \gamma(0)=\gamma(t)=y\right\} = 0 \tag{7.30}$$

for any fixed $\delta > 0$.

Proof. This proof is rather long and we divide it into two steps.

Step 1. We first assume that (7.30) holds for each $\delta > 0$ and show that $y \in \mathcal{A}$. Assume by contradiction that $y \notin \mathcal{A}$. Then by the first part of the proof of Proposition 7.33, there are $\psi_y \in \mathrm{Lip}(\mathbb{T}^n)$ and $r_y, \varepsilon_y > 0$ such that ψ_y is a subsolution to (7.23), and $B(y,r_y) \subset \mathbb{T}^n \setminus \mathcal{A}$, and

$$H(x, D\psi_y(x)) < \overline{H}(0) - \varepsilon_y \qquad \text{for all } x \in B(y,r_y).$$

Fix $\delta = 1$. For each $\varepsilon > 0$, there exist $T > 1$ and $\xi \in \mathrm{AC}([0,T],\mathbb{T}^n)$ such that

$$\begin{aligned} &\inf\left\{\int_0^T \left(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\right) ds \,:\, \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n), \gamma(0)=\gamma(T)=y\right\} \\ &\qquad\qquad = \int_0^T \left(L(\xi(s),\xi'(s)) + \overline{H}(0)\right) ds \le \varepsilon. \end{aligned}$$

By Theorem A.9 in Appendix, we see that $\xi \in C^2([0,T],\mathbb{T}^n)$ and there exists $C > 0$ independent of $T > 1$ such that $\|\xi'\|_{L^\infty} \le C$. In particular,

$$\xi(s) \in B(y,r_y) \qquad \text{for all } s \in [0, r_y/C]. \tag{7.31}$$

In the following computations, we assume ψ_y is smooth enough for simplicity. To make things rigorous, we just need to do the usual trick of convolution with the standard mollifier.

We use (7.31) to compute that

$$
\begin{aligned}
\varepsilon &\geq \int_0^T \left(L(\xi(s),\xi'(s))+\overline{H}(0)\right)ds \\
&= \int_0^{\frac{r_y}{C}} \left(L(\xi(s),\xi'(s))+\overline{H}(0)\right)ds + \int_{\frac{r_y}{C}}^T \left(L(\xi(s),\xi'(s))+\overline{H}(0)\right)ds \\
&\geq \int_0^{\frac{r_y}{C}} \left(L(\xi(s),\xi'(s))+H(\xi(s),D\psi_y(\xi(s)))+\varepsilon_y\right)ds \\
&\qquad\qquad + \int_{\frac{r_y}{C}}^T \left(L(\xi(s),\xi'(s))+H(\xi(s),D\psi_y(\xi(s)))\right)ds \\
&= \int_0^T \left(L(\xi(s),\xi'(s))+H(\xi(s),D\psi_y(\xi(s)))\right)ds + \frac{\varepsilon_y r_y}{C} \\
&\geq \int_0^T \xi'(s)\cdot D\psi_y(\xi(s)))\,ds + \frac{\varepsilon_y r_y}{C} = \psi_y(\xi(T))-\psi_y(\xi(0)) + \frac{\varepsilon_y r_y}{C} = \frac{\varepsilon_y r_y}{C}.
\end{aligned}
$$

We then get a contradiction by simply choosing $\varepsilon = \frac{\varepsilon_y r_y}{2C}$. Thus, $y \in \mathcal{A}$, and the proof of the first claim is complete.

Step 2. Next, we assume that $y \in \mathcal{A}$. We need to show that (7.30) holds for each $\delta > 0$. Assume by contradiction that (7.30) fails for some $\delta > 0$. Take $\gamma(s) = y$ for all $s \in [0,\delta+1]$, we see that $L(y,0)+\overline{H}(0) > 0$. By the Legendre transform, this gives

$$\min_{p\in\mathbb{R}^n} H(y,p) < \overline{H}(0).$$

Thus, there is $q \in \mathbb{R}^n$ such that $H(y,q) < \overline{H}(0)$. Then, there exist two constants $C, r > 0$ such that

$$
\begin{cases}
S(x,y) < 1 & \text{for all } x \in B(y,r), \\
H(x,q) < \overline{H}(0) & \text{for all } x \in B(y,r), \\
L(x,v) \leq C & \text{for all } (x,v) \in B(y,r)\times B(0,r).
\end{cases}
\tag{7.32}
$$

Let $r_1 \in (0,r)$ be a radius to be fixed later. Pick $x \in B(y,r_1)\setminus\{y\}$. Pick a path $\gamma \in \mathrm{AC}([0,t],\mathbb{T}^n)$ such that $\gamma(0) = y, \gamma(t) = x$, and

$$\int_0^t \left(L(\gamma(s),\gamma'(s))+\overline{H}(0)\right)ds \leq 1.$$

By using the fact that $L(x,v) \geq \frac{\theta}{2}|v|^2 - C$, we deduce

$$\int_0^t |\gamma'(s)|^2\,ds \leq C(1+t).$$

Then, by the usual Cauchy-Schwarz inequality,

$$\int_0^t |\gamma'(s)|\,ds \leq \left(\int_0^t |\gamma'(s)|^2\,ds\right)^{1/2}\left(\int_0^t 1\,ds\right)^{1/2} \leq Ct^{1/2}(1+t)^{1/2}. \tag{7.33}$$

Let $\sigma = r^2/C$. Thanks to (7.33), if $t \le \sigma$, then $\gamma(s) \in B(y,r)$ for all $s \in [0,t]$. By adjusting σ, we might assume that $\delta = k\sigma$ for some $k \in \mathbb{N}$. Note then that

$$\begin{aligned} &k \inf\left\{\int_0^T \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds \,:\, \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n), T \ge \sigma, \gamma(0)=\gamma(T)=y\right\} \\ &\ge \inf\left\{\int_0^T \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds \,:\, \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n), T \ge \delta, \gamma(0)=\gamma(T)=y\right\} > 0, \end{aligned}$$

we see that there exists $a > 0$ such that

$$\inf\left\{\int_0^T \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds \,:\, \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n), T \ge \sigma, \gamma(0)=\gamma(T)=y\right\} > a.$$

We now consider two cases. The first case is when $t \le \sigma$. Then, as noted above, $\gamma(s) \in B(y,r)$ for all $s \in [0,t]$, and

$$\begin{aligned} \int_0^t \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds &\ge \int_0^t \big(L(\gamma(s),\gamma'(s)) + H(\gamma(s),q)\big)\,ds \\ &\ge \int_0^t \gamma'(s)\cdot q\,ds = q\cdot(x-y). \end{aligned}$$

The second case is when $t > \sigma$. We use γ to create a loop starting from y as following. Let $\eta : [0, t + |x-y|/r] \to \mathbb{T}^n$ be such that

$$\eta(s) = \begin{cases} \gamma(s) & \text{for } s \in [0,t], \\ x + (s-t)r\frac{y-x}{|y-x|} & \text{for } s \in [t, t+|x-y|/r]. \end{cases}$$

Then,

$$\begin{aligned} a &< \int_0^{t+|x-y|/r} \big(L(\eta(s),\eta'(s)) + \overline{H}(0)\big)\,ds \\ &= \int_0^t \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds + \int_t^{t+|x-y|/r} \left(L\left(\eta(s), r\frac{y-x}{|y-x|}\right) + \overline{H}(0)\right)ds \\ &\le \int_0^t \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds + \frac{C|x-y|}{r} \\ &\le \int_0^t \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds + \frac{Cr_1}{r} \end{aligned}$$

We now choose r_1 such that $\frac{Cr_1}{r} \le \frac{a}{2}$ to deduce that

$$\int_0^t \big(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\big)\,ds \ge \frac{a}{2}.$$

Combining the two cases, we yield that, for $x \in B(y,r_1)\setminus\{y\}$,

$$S(x,y) \ge \min\left\{q\cdot(x-y), \frac{a}{2}\right\}. \tag{7.34}$$

Thanks to (7.34), we have that $q \in D^-S(y,y)$. However, as $H(y,q) < \overline{H}(0)$, the supersolution test for $S(x,y)$ fails at the vertex $x = y$. This gives that $y \notin \mathcal{A}$, which is absurd. $\square$

Remark 7.42. One important point in the second step of the proof above is if $\delta = k\sigma$ for some $k \in \mathbb{N}$, then

$$k \inf\left\{ \int_0^T \left(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\right) ds \,:\, \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n), T \ge \sigma, \gamma(0)=\gamma(T)=y \right\}$$
$$\ge \inf\left\{ \int_0^T \left(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\right) ds \,:\, \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n), T \ge \delta, \gamma(0)=\gamma(T)=y \right\}.$$

This is quite clear to see as for any given path γ in the admissible class of the left hand side above, we let $\bar{\gamma}$ be the curve γ with multiplicity k, then $\bar{\gamma}$ is in the admissible class of the right hand side above.

Thus, we see that (7.30) holds for all $\delta > 0$ if and only if (7.30) holds for some $\delta > 0$.

We are now ready to give an equivalent definition of the projected Aubry set.

Definition 7.43 (An equivalent definition of the projected Aubry set). *For $y \in \mathbb{T}^n$, we say that $y \in \mathcal{A}$ if*

$$\inf\left\{ \int_0^t \left(L(\gamma(s),\gamma'(s)) + \overline{H}(0)\right) ds \,:\, \gamma \in \mathrm{AC}([0,t],\mathbb{T}^n) \text{ for } t > \delta, \gamma(0)=\gamma(t)=y \right\} = 0 \tag{7.35}$$

for some fixed $\delta > 0$.

Roughly speaking, this definition gives a nice geometric interpretation of points in the projected Aubry set as following. Fix $\delta > 0$. Then, a point $y \in \mathbb{T}^n$ is in the projected Aubry set if one is able to find loops containing y with length at least δ such that it costs almost nothing (with the precise cost functional in (7.35)) to travel on these loops.

We next show that the projected Mather set is a subset of the projected Aubry set. Again, we always fix $P = 0$ here.

Theorem 7.44. *Assume* (7.1). *Then*

$$\mathcal{M}_0 \subset \mathcal{A}.$$

Proof. We use some ideas in the proof of Theorem 7.9 and Remark 7.10.

Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (7.23) and $y \in \mathcal{M}_0$. Let ξ be a backward characteristic of u starting from y. Then, as $y \in \mathcal{M}_0$, y is in the α-limit set of the trajectory $\{\xi(s) : s \in (-\infty, 0]\}$. In other words, there exists a sequence $\{s_k\} \to -\infty$ such that $\xi(s_k) \to y$, and

$$u(y) - u(\xi(s_k)) = \int_{s_k}^0 \left(L(\xi(s),\xi'(s)) + \overline{H}(0)\right) ds$$

We use ξ to create a loop starting from y as following. For each $k \in \mathbb{N}$, denote by $\eta_k : [s_k - |\xi(s_k) - y|, 0] \to \mathbb{T}^n$ the following curve

$$\eta_k(s) = \begin{cases} \xi(s) & \text{for } s \in [s_k, 0], \\ y + (s - s_k + |\xi(s_k) - y|)\frac{\xi(s_k)-y}{|\xi(s_k)-y|} & \text{for } s \in [s_k - |\xi(s_k)-y|, s_k]. \end{cases}$$

It is then clear that

$$\lim_{k\to\infty}\int_{s_k-|\xi(s_k)-y|}^{0}\left(L(\eta_k(s),\eta_k'(s))+\overline{H}(0)\right)ds=0,$$

which gives us that $y\in\mathcal{A}$. □

In the following, we show that if $y\in\mathcal{A}$, then $S(x,y)$ is differentiable at y.

Proposition 7.45. *Assume* (7.1) *and* $y\in\mathcal{A}$. *Then,* $x\mapsto S(x,y)$ *is differentiable at* y.

Proof. As $y\in\mathcal{A}$, we are able to find a sequence of C^2 curves $\gamma_n:[0,t_n]\to\mathbb{T}^n$ such that

$$\begin{cases}\gamma_n(0)=\gamma_n(t_n)=y,\ t_n\to\infty,\ \|\gamma_n'\|_{L^\infty([0,t_n])}\le C,\\ \int_0^{t_n}\left(L(\gamma_n(s),\gamma_n'(s))+\overline{H}(0)\right)ds\to 0.\end{cases}\tag{7.36}$$

Note that γ_n is an orbit of the Euler-Lagrange flow for each $n\in\mathbb{N}$. By extracting a subsequence if necessary, we may assume that $\{\gamma_n\}$ converges to $\gamma:[0,\infty)\to\mathbb{T}^n$ locally uniformly in the C^1 topology.

Fix $t\in(0,\infty)$. For $n\in\mathbb{N}$ sufficiently large such that $t_n>t$, let $d_n=|\gamma_n(t)-\gamma(t)|$. Similar to the construction of η_k in the proof of Theorem 7.44 above, we construct a curve $\tilde\gamma_n:[t-d_n,t_n]\to\mathbb{T}^n$ from $\gamma(t)$ to y by joining together a unit speed line segment from $\gamma(t)$ to $\gamma_n(t)$ with γ_n on $[t,t_n]$. It is clear that

$$S(y,\gamma(t))\le\int_{t-d_n}^{t_n}L(\tilde\gamma_n(s),\tilde\gamma_n'(s))\,ds\le Cd_n+\int_t^{t_n}L(\gamma_n(s),\gamma_n'(s))\,ds.$$

Hence,

$$S(y,\gamma(t))+\int_0^t\left(L(\gamma_n(s),\gamma_n'(s))+\overline{H}(0)\right)ds\le Cd_n+\int_0^{t_n}L(\gamma_n(s),\gamma_n'(s))\,ds.$$

Let $n\to\infty$ and use (7.36) to yield that

$$S(y,\gamma(t))+\int_0^t\left(L(\gamma(s),\gamma'(s))+\overline{H}(0)\right)ds\le 0.\tag{7.37}$$

On the other hand,

$$S(\gamma(t),y)\le\int_0^t\left(L(\gamma(s),\gamma'(s))+\overline{H}(0)\right)ds.\tag{7.38}$$

Combine the two inequalities above to imply

$$0=S(y,y)\le S(y,\gamma(t))+S(\gamma(t),y)\le 0$$

Thus, equalities in (7.37) and (7.38) must happen. In a similar way, we can construct $\gamma:(-\infty,0]\to\mathbb{T}^n$.

We therefore have $\gamma:(-\infty,\infty)\to\mathbb{T}^n$ with $\gamma(0)=y$, and, for $t\in\mathbb{R}$,

$$S(\gamma(t),y)=-S(y,\gamma(t))=\int_0^t\left(L(\gamma(s),\gamma'(s))+\overline{H}(0)\right)ds.\tag{7.39}$$

By Theorem 7.6, $S(\cdot, y)$ is differentiable at $\gamma(t)$ for all $t \in \mathbb{R}$, and

$$DS(\gamma(t), y) = D_v L(\gamma(t), \gamma'(t)).$$

In particular, $S(\cdot, y)$ is differentiable at $x = y$. □

Remark 7.46. In the above proof, we have actually shown that for each $y \in \mathcal{A}$, there is a two-sided minimizer $\gamma : (-\infty, \infty) \to \mathbb{T}^n$ with $\gamma(0) = y$ satisfying (7.39). In the literature, γ is also called a calibrated curve on $\mathbb{R}$.

In fact, for any subsolution $u \in \mathrm{Lip}(\mathbb{T}^n)$ of (7.23), we have

$$u(y) - u(\gamma(t)) \leq S(y, \gamma(t)) \qquad \text{and} \qquad u(\gamma(t)) - u(y) \leq S(\gamma(t), y).$$

Thus, (7.37) and (7.38) can be changed to

$$u(y) - u(\gamma(t)) + \int_0^t \big(L(\gamma(s), \gamma'(s)) + \overline{H}(0)\big)\, ds \leq 0.$$

On the other hand,

$$u(\gamma(t)) - u(y) \leq \int_0^t \big(L(\gamma(s), \gamma'(s)) + \overline{H}(0)\big)\, ds.$$

This means that we also must have equalities happen in the two above. In particular, for each $t \in \mathbb{R}$,

$$u(\gamma(t)) - u(y) = \int_0^t \big(L(\gamma(s), \gamma'(s)) + \overline{H}(0)\big)\, ds.$$

By Theorem 7.6, u is differentiable at y, and

$$Du(y) = DS(y, y) = D_v L(\gamma(0), \gamma'(0)).$$

6 References

1. For further developments in weak KAM theory, see Fathi's book [59]. We do not cover many topics here such as the Peierls barrier, large time behaviors of solutions to Cauchy problems. See also the books of Gomes [75], Sorrentino [129].

2. There are many excellent survey papers and lecture notes in this area: see Evans [51, 52], Ishii [85], Kaloshin [95], and the references therein.

3. We only cover basic aspects of the projected Aubry set here. For further aspects and developments (e.g., existence of C^1 subsolutions), see Fathi's book [59], Fathi, Siconolfi [60, 61], Bernard [21], and Ishii [86]. As mentioned, the maximal subsolution S is also called the critical Mañé potential in the weak KAM theory literature (see Mañé [110]).

4. We do not cover the two dimensional Aubry–Mather theory here.

5. We only deal with the convex case here. Nonconvex Aubry–Mather theory was studied by using the vanishing viscosity approximations addressed above by Cagnetti, Gomes, Tran [24].

6. The PDE approach via nonlinear adjoint method here has an advantage that it works well for general viscous Hamilton–Jacobi equations. We do not cover the viscous cases here. See Gomes [71], Cagnetti, Mitake, Gomes, Tran [23], Mitake, Tran [116, 117], Ishii, Mitake, Tran [87, 88].

7. For various examples on non-uniqueness of solutions to the cell problems, see Chapter 4, and Le, Mitake, Tran [100, Chapter 6].

CHAPTER 8

Further properties of the effective Hamiltonians in the convex setting

In this chapter, we aim at studying further properties of $\overline{H}$ in case that $H = H(x, p)$ is convex in p. As mentioned repeatedly in previous chapters, not so much of deep properties of H is known at the moment. Nevertheless, with the developments of weak KAM theory in the previous chapter, we are able to understand a bit more about $\overline{H}$. We will address appropriate assumptions that we need in each section below. The results in the sections in this chapter are rather disjoint.

1 Strict convexity of the effective Hamiltonian in certain directions

In this section, we always assume that

$$\begin{cases} H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{pp}H(x,p) \le \theta^{-1} I_n \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{8.1}$$

Let $L = L(y, v)$ be the corresponding Lagrangian. By changing $\theta > 0$ to be smaller if needed, we may also assume that

$$\begin{cases} L \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{vv}L(x,v) \le \theta^{-1} I_n \text{ for all } (x,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{8.2}$$

Here is the main result in this section.

Theorem 8.1. *Assume* (8.1). *Then, there exists a positive constant C such that for each $R \in \mathbb{R}^n$, we have*

$$-R \cdot \tilde{Q},\, R \cdot \hat{Q} \le C \left(\liminf_{t \to 0+} \frac{\overline{H}(P+tR) + \overline{H}(P-tR) - 2\overline{H}(P)}{t^2} \right)^{1/2},$$

for some $\tilde{Q}, \hat{Q} \in D^-\overline{H}(P)$. In particular, if $\overline{H}$ is twice differentiable at P, then

$$|D\overline{H}(P) \cdot R| \le C (R \cdot D^2\overline{H}(P) R)^{1/2}$$

for each $R \in \mathbb{R}^n$.

This result is taken from Evans, Gomes [53]. We also follow their approach to give a proof here.

Proof. Fix $R \in \mathbb{R}^n$. Denote by

$$\tilde{u} = u(\cdot, P + tR) \quad \text{and} \quad \hat{u} = u(\cdot, P - tR)$$

solutions to the cell problems at $P + tR$ and $P - tR$, respectively. As $\tilde{u}, \hat{u}$ are not smooth, we smooth them up as usual. Let $\rho \in C_c^\infty(\mathbb{R}^n, [0, \infty))$ be a standard mollifier. For $\delta > 0$, let $\rho^\delta(x) = \delta^{-n}\rho(\delta^{-1}x)$ for all $x \in \mathbb{R}^n$. Denote by

$$\tilde{u}^\delta = \rho^\delta * \tilde{u} \quad \text{and} \quad \hat{u}^\delta = \rho^\delta * \hat{u}.$$

Then, of course, $\tilde{u}^\delta, \hat{u}^\delta \in C^\infty(\mathbb{T}^n)$, $\tilde{u}^\delta \to \tilde{u}$, $\hat{u}^\delta \to \hat{u}$ in $C(\mathbb{T}^n)$ as $\delta \to 0$. Moreover,

$$\begin{cases} H(x, P + tR + D\tilde{u}^\delta) \le \overline{H}(P + tR) + C\delta, \\ H(x, P - tR + D\hat{u}^\delta) \le \overline{H}(P - tR) + C\delta, \end{cases} \qquad \text{in } \mathbb{T}^n.$$

We simply write $u = u(\cdot, P)$ as a solution to the cell problem at P. Let μ be a Mather measure corresponding to u, and $\sigma = \pi \circ \mu$, its projection to $\mathbb{T}^n$. By the Lipschitz graph theorem (Theorem 7.16), μ is supported on a Lipschitz graph, and for $x \in \operatorname{spt}(\sigma)$, u is differentiable at x, and

$$\begin{aligned} &H(x, P + tR + D\tilde{u}^\delta(x)) + H(x, P - tR + D\hat{u}^\delta(x)) - 2H(x, P + Du(x)) \\ &\qquad\qquad \le \overline{H}(P + tR) + \overline{H}(P - tR) - 2\overline{H}(P) + C\delta. \end{aligned}$$

By the uniform convexity of H,

$$\begin{aligned} &H(x, P + tR + D\tilde{u}^\delta(x)) - H(x, P + Du(x)) \\ &\qquad \ge D_pH(x, P + Du(x)) \cdot (tR + (D\tilde{u}^\delta(x) - Du(x))) + \frac{\theta}{2}|tR + (D\tilde{u}^\delta(x) - Du(x))|^2, \end{aligned}$$

and

$$\begin{aligned} &H(x, P - tR + D\hat{u}^\delta(x)) - H(x, P + Du(x)) \\ &\qquad \ge D_pH(x, P + Du(x)) \cdot (-tR + (D\hat{u}^\delta(x) - Du(x))) + \frac{\theta}{2}|-tR + (D\hat{u}^\delta(x) - Du(x))|^2. \end{aligned}$$

Combine the two inequalities above to imply

$$\begin{aligned} &H(x, P + tR + D\tilde{u}^\delta(x)) + H(x, P - tR + D\hat{u}^\delta(x)) - 2H(x, P + Du(x)) \\ &\ge D_pH(x, P + Du(x)) \cdot D(\tilde{u}^\delta + \hat{u}^\delta - 2u) + \frac{\theta}{2}|tR + D(\tilde{u}^\delta - u)(x)|^2 + \frac{\theta}{2}|-tR + D(\hat{u}^\delta - u)(x)|^2. \end{aligned}$$

Note that

$$\int_{\mathbb{T}^n} D_pH(x, P + Du(x)) \cdot D(\tilde{u}^\delta + \hat{u}^\delta - 2u)\, d\sigma(x) = 0.$$

Therefore,

$$\int_{\mathbb{T}^n} \left(|tR + D(\tilde{u}^\delta - u)|^2 + |-tR + D(\hat{u}^\delta - u)|^2\right) d\sigma \le C\left(\overline{H}(P + tR) + \overline{H}(P - tR) - 2\overline{H}(P) + C\delta\right). \tag{8.3}$$

On the other hand,

$$
\begin{aligned}
\overline{H}(P)-\overline{H}(P+tR) &\le \int_{\mathbb{T}^n} (H(x,P+Du)-H(x,P+tR+D\tilde{u}^{\delta}))\,d\sigma + C\delta \\
&\le C\left(\int_{\mathbb{T}^n} |-tR+D(u-\tilde{u}^{\delta})|^2\,d\sigma\right)^{1/2} + C\delta, \qquad (8.4)
\end{aligned}
$$

and

$$
\begin{aligned}
\overline{H}(P)-\overline{H}(P-tR) &\le \int_{\mathbb{T}^n} (H(x,P+Du)-H(x,P-tR+D\hat{u}^{\delta}))\,d\sigma + C\delta \\
&\le C\left(\int_{\mathbb{T}^n} |tR+D(u-\hat{u}^{\delta})|^2\,d\sigma\right)^{1/2} + C\delta. \qquad (8.5)
\end{aligned}
$$

Combine (8.3)–(8.5) and let $\delta \to 0$ to observe that

$$
\begin{cases}
\overline{H}(P)-\overline{H}(P+tR) \le C\left(\overline{H}(P+tR)+\overline{H}(P-tR)-2\overline{H}(P)\right)^{1/2}, \\
\overline{H}(P)-\overline{H}(P-tR) \le C\left(\overline{H}(P+tR)+\overline{H}(P-tR)-2\overline{H}(P)\right)^{1/2}.
\end{cases}
$$

Thus, for any $\tilde{Q}(t) \in D^-\overline{H}(P+tR)$, and $\hat{Q}(t) \in D^-\overline{H}(P-tR)$,

$$
-t\tilde{Q}(t)\cdot R,\ t\hat{Q}(t)\cdot R \le C\left(\overline{H}(P+tR)+\overline{H}(P-tR)-2\overline{H}(P)\right)^{1/2}.
$$

Hence, we can find a sequence $\{t_k\} \to 0+$ so that $\tilde{Q}(t_k) \to \tilde{Q}$, $\hat{Q}(t_k) \to \hat{Q}$ with $\tilde{Q}, \hat{Q} \in D^-\overline{H}(P)$ such that

$$
-R\cdot\tilde{Q},\ R\cdot\hat{Q} \le C\left(\liminf_{t\to 0+} \frac{\overline{H}(P+tR)+\overline{H}(P-tR)-2\overline{H}(P)}{t^2}\right)^{1/2},
$$

Of course, if $\overline{H}$ is twice differentiable at P, we have last claim in the theorem automatically. □

Remark 8.2. By the above theorem, we see that if $\overline{H}$ is differentiable at P, then it is strictly convex in each direction R which is not tangent to the level set $\left\{\overline{H} = \overline{H}(P)\right\}$.

For $\overline{H}(P) > \min_{\mathbb{R}^n} \overline{H}$, then $0 \notin D^-\overline{H}(P)$. This implies that there is an open convex cone of directions R in which $\overline{H}$ is strictly convex at P. In particular, we conclude that $\overline{H}$ can only have flat parts at its minimum value. Of course, we have seen earlier in some examples that $\overline{H}$ indeed has a flat part there at $\min_{\mathbb{R}^n} H$, and this theorem confirms that this is the only possible flat part.

2 Asymptotic expansion at infinity

We assume here that

$$
H(x,p) = \frac{1}{2}|p|^2 + V(x) \qquad \text{for } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n.
$$

Here, we consider a very simple setting where the potential energy $V \in C(\mathbb{T}^n)$ is a trigonometric polynomial, that is, V satisfies that

$$\begin{cases} V(x) = \lambda_0 + \sum_{j=1}^m (\lambda_j e^{i2\pi k_j \cdot x} + \overline{\lambda_j} e^{-i2\pi k_j \cdot x}), \\ \text{where } \lambda_0 \in \mathbb{R},\ \{\lambda_j\}_{j=1}^m \subset \mathbb{C} \text{ and } \{k_j\}_{j=1}^m \subset \mathbb{Z}^n \setminus \{0\} \text{ are given.} \end{cases} \tag{8.6}$$

Recall that $\overline{\lambda_j}$ is the complex conjugate of λ_j for $1 \le j \le m$. Our aim, of course, is to understand $\overline{H}$ better. It turns out that we are able to read off certain information of $\overline{H}(p)$ as $|p| \to \infty$.

2.1 The method of asymptotic expansion at infinity

Let us explain first what is this method heuristically. For a given vector $Q \neq 0$ and $\varepsilon > 0$, set $p = \frac{Q}{\sqrt{\varepsilon}}$. The cell problem for this vector p is

$$\frac{1}{2}\left|\frac{Q}{\sqrt{\varepsilon}} + Dv^\varepsilon\right|^2 + V(x) = \overline{H}\left(\frac{Q}{\sqrt{\varepsilon}}\right) \qquad \text{in } \mathbb{T}^n.$$

Here, $v^\varepsilon \in C(\mathbb{T}^n)$ is a solution to the above. Multiply both sides by ε to yield

$$\frac{1}{2}|Q + \sqrt{\varepsilon} Dv^\varepsilon|^2 + \varepsilon V(x) = \varepsilon \overline{H}\left(\frac{Q}{\sqrt{\varepsilon}}\right) =: \overline{H}^\varepsilon(Q) \qquad \text{in } \mathbb{T}^n. \tag{8.7}$$

To understand the asymptotic behavior of $\overline{H}$ in the direction Q at infinity (more precisely, at $\frac{Q}{\sqrt{\varepsilon}}$ as $\varepsilon \to 0$), we aim at finding asymptotic expansion of $\overline{H}^\varepsilon(Q)$. Let us first use a formal asymptotic expansion to do computations. We use an ansatz as following

$$\begin{cases} \sqrt{\varepsilon} v^\varepsilon(x) = \varepsilon v_1(x) + \varepsilon^2 v_2(x) + \varepsilon^3 v_3(x) + \cdots, \\ \overline{H}^\varepsilon(Q) = a_0 + \varepsilon a_1 + \varepsilon^2 a_2 + \varepsilon^3 a_3 + \cdots. \end{cases}$$

Plug these into (8.7) to imply

$$\frac{1}{2}|Q + \varepsilon Dv_1 + \varepsilon^2 Dv_2 + \cdots|^2 + \varepsilon V = \overline{H}^\varepsilon(Q) = a_0 + \varepsilon a_1 + \varepsilon^2 a_2 + \cdots \qquad \text{in } \mathbb{T}^n.$$

We first compare the $O(1)$ terms in both sides of the above equality to get

$$a_0 = \frac{1}{2}|Q|^2.$$

By using $O(\varepsilon)$, we get

$$Q \cdot Dv_1 + V = a_1 \qquad \text{in } \mathbb{T}^n.$$

Hence, $a_1 = \int_{\mathbb{T}^n} V\,dx = \lambda_0$ and

$$Dv_1 = -\sum_{j=1}^m (\lambda_j e^{i2\pi k_j \cdot x} + \overline{\lambda_j} e^{-i2\pi k_j \cdot x}) \frac{k_j}{k_j \cdot Q}, \tag{8.8}$$

provided that we do not divide by zero. Next, using $O(\varepsilon^2)$,

$$\frac{1}{2}|Dv_1|^2 + Q \cdot Dv_2 = a_2 \qquad \text{in } \mathbb{T}^n,$$

we achieve that

$$a_2 = \sum_{j=1}^{m} \frac{|\lambda_j|^2 |k_j|^2}{|k_j \cdot Q|^2}. \tag{8.9}$$

Plug this back to get an equation for v_2 as

$$\begin{aligned} Q \cdot Dv_2 &= a_2 - \frac{1}{2}|Dv_1|^2 \\ &= -\frac{1}{2} \sum_{\pm k_j \pm k_l \neq 0} \frac{\lambda_j^{\pm} \lambda_l^{\pm} k_j \cdot k_l}{(k_j \cdot Q)(k_l \cdot Q)} e^{i2\pi(\pm k_j \pm k_l)\cdot x}. \end{aligned}$$

Here for convenience, for $1 \le j \le m$, we denote by

$$\lambda_j^+ = \lambda_j \quad \text{and} \quad \lambda_j^- = \overline{\lambda_j}.$$

Thus,

$$Dv_2 = -\frac{1}{2} \sum_{\pm k_j \pm k_l \neq 0} \frac{\lambda_j^{\pm} \lambda_l^{\pm} k_j \cdot k_l}{(k_j \cdot Q)(k_l \cdot Q)} e^{i2\pi(\pm k_j \pm k_l)\cdot x} \frac{\pm k_j \pm k_l}{(\pm k_j \pm k_l) \cdot Q}.$$

Let us now switch to a symbolic way of writing to keep track with all terms. We write $\sum_G$ to mean that it is a good sum where all terms are well-defined, that is, all denominators of the fractions in the sum are not zero. We have

$$Dv_2 = -\frac{1}{2} \sum_{G} \frac{\lambda_{j_1}^{\pm} \lambda_{j_2}^{\pm} k_{j_1} \cdot k_{j_2}}{(k_{j_1} \cdot Q)(k_{j_2} \cdot Q)} e^{i2\pi(\pm k_{j_1} \pm k_{j_2})\cdot x} \frac{\pm k_{j_1} \pm k_{j_2}}{(\pm k_{j_1} \pm k_{j_2}) \cdot Q}. \tag{8.10}$$

Next, $O(\varepsilon^3)$ gives us the following relation

$$Q \cdot Dv_3 = a_3 - Dv_1 \cdot Dv_2.$$

Hence,

$$a_3 = \int_{\mathbb{T}^n} Dv_1 \cdot Dv_2 \, dx,$$

and

$$\begin{aligned} Dv_3 = -\frac{1}{2} \sum_{G} & \frac{\lambda_{j_1}^{\pm} \lambda_{j_2}^{\pm} \lambda_{j_3}^{\pm} (k_{j_1} \cdot k_{j_2})(\pm k_{j_1} \pm k_{j_2}) \cdot k_{j_3}}{(k_{j_1} \cdot Q)(k_{j_2} \cdot Q)(k_{j_3} \cdot Q)(\pm k_{j_1} \pm k_{j_2}) \cdot Q} \times \\ & \times e^{i2\pi(\pm k_{j_1} \pm k_{j_2} \pm k_{j_3})\cdot x} \frac{\pm k_{j_1} \pm k_{j_2} \pm k_{j_3}}{(\pm k_{j_1} \pm k_{j_2} \pm k_{j_3}) \cdot Q}. \end{aligned} \tag{8.11}$$

The $O(\varepsilon^4)$ term yields

$$Dv_1 \cdot Dv_3 + \frac{1}{2}|Dv_2|^2 + Q \cdot Dv_4 = a_4.$$

Integrate to get

$$a_4 = \frac{1}{2} \int_{\mathbb{T}^2} |Dv_2|^2 \, dx + \int_{\mathbb{T}^2} Dv_1 \cdot Dv_3 \, dx.$$

Of course, v_4 satisfies

$$Q \cdot Dv_4 = a_4 - Dv_1 \cdot Dv_3 - \frac{1}{2}|Dv_2|^2. \tag{8.12}$$

It can be seen that although we have formulas for a_3 and a_4, they are already quite complicated to be written down explicitly in general. By computing in an iterative way, we can get formulas of a_l and v_l for all $l \in \mathbb{N}$. Clearly, these formulas are extremely involved and hard to be use. Nevertheless, they do contain important information about how V influences $\overline{H}^{\varepsilon}(Q)$. It is necessary to come up with correct ways to read off the information.

2.2 Rigorous expansion

It turns out that the above formal asymptotic expansion of $\overline{H}^{\varepsilon}(Q)$ holds true rigorously. As we stop at a_4, let us verify the result up to five terms in the asymptotic expansion.

Theorem 8.3. *Assume that $H(x,p) = \frac{1}{2}|p|^2 + V(x)$ for all $(x,p) \in \mathbb{T}^n \times \mathbb{R}^n$, where V satisfies* (8.6). *Let $\overline{H}$ be the effective Hamiltonian corresponding to H. Let $Q \neq 0$ be a vector in $\mathbb{R}^n$ such that Q is not perpendicular to each nonzero vector of k_{j_1}, $\pm k_{j_1} \pm k_{j_2}$, $\pm k_{j_1} \pm k_{j_2} \pm k_{j_3}$, and $\pm k_{j_1} \pm k_{j_2} \pm k_{j_3} \pm k_{j_4}$ for $1 \le j_1, j_2, j_3, j_4 \le m$.*

For $\varepsilon > 0$, set $\overline{H}^{\varepsilon}(Q) = \varepsilon \overline{H}_1\left(\frac{Q}{\sqrt{\varepsilon}}\right)$. Then we have that, as $\varepsilon \to 0$,

$$\overline{H}^{\varepsilon}(Q) = \frac{1}{2}|Q|^2 + \varepsilon a_1 + \varepsilon^2 a_2 + \varepsilon^3 a_3 + \varepsilon^4 a_4 + O(\varepsilon^5),$$

where a_1, a_2, a_3, a_4 are the constants derived in the previous section. Here, the error term satisfies $|O(\varepsilon^5)| \le K\varepsilon^5$ for some K depending only on Q, $\{\lambda_j\}_{j=1}^m$ and $\{k_j\}_{j=1}^m$.

The proof of this turns out to be quite simple. We just need to use the viscosity solution techniques to show that our expansion, which is smooth, approximates pretty well $\overline{H}^{\varepsilon}(Q)$. It is worth mentioning that the error term $O(\varepsilon^5)$ does depend on the position of Q.

Proof. Let v_1, v_2, v_3, v_4 be solutions to (8.8), (8.10), (8.11), (8.12), respectively. Let $\phi = \varepsilon v_1 + \varepsilon^2 v_2 + \varepsilon^3 v_3 + \varepsilon^4 v_4$. Then ϕ is of course smooth, and ϕ satisfies

$$\frac{1}{2}|Q + D\phi|^2 + \varepsilon V = \frac{1}{2}|Q|^2 + \varepsilon a_1 + \varepsilon^2 a_2 + \varepsilon^3 a_3 + \varepsilon^4 a_4 + O(\varepsilon^5) \qquad \text{in } \mathbb{T}^n.$$

Here, the error term $O(\varepsilon^5)$ can be seen explicitly in the computations as

$$O(\varepsilon^5) = \varepsilon^5\left(Dv_1 \cdot Dv_4 + Dv_2 \cdot Dv_3\right) + \varepsilon^6\left(Dv_2 \cdot Dv_4 + \frac{1}{2}|Dv_3|^2\right) + \varepsilon^7\left(Dv_3 \cdot Dv_4\right) + \varepsilon^8 \frac{|Dv_4|^2}{2}.$$

It is clear that $|O(\varepsilon^5)| \le K\varepsilon^5$ for some K depending only on Q, $\{\lambda_j\}_{j=1}^m$ and $\{k_j\}_{j=1}^m$.

Recall that $w = \sqrt{\varepsilon} v_1^{\varepsilon}$ is a solution to (8.7). We now use ϕ, which is smooth, as a test function for (8.7). By looking at the places where $w - \phi$ attains its maximum and minimum in $\mathbb{T}^n$ and using the definition of viscosity subsolutions and supersolutions, respectively, we arrive at the conclusion that

$$\overline{H}^{\varepsilon}(Q) = \frac{1}{2}|Q|^2 + \varepsilon a_1 + \varepsilon^2 a_2 + \varepsilon^3 a_3 + \varepsilon^4 a_4 + O(\varepsilon^5).$$

□

Remark 8.4. Let $t = \varepsilon^{-1/2}$. Then, from the above theorem, we get that

$$\frac{\overline{H}(tQ)}{t^2} = \frac{1}{2}|Q|^2 + \frac{1}{t^2}a_1 + O\left(\frac{1}{t^4}\right) = \frac{1}{2}|Q|^2 + \frac{1}{t^2}\int_{\mathbb{T}^n} V\,dx + O\left(\frac{1}{t^4}\right),$$

which tells us that at infinity, we see the average of V as the next order term after $\frac{1}{2}|Q|^2$. This is quite interesting as this term is independent of Q. The next term in the expansion

$$\frac{1}{t^4}a_2 = \frac{1}{t^4}\sum_{j=1}^{m}\frac{|\lambda_j|^2|k_j|^2}{|k_j\cdot Q|^2}$$

is clearly dependent on Q.

3 The classical Hedlund example

In dimensions three or higher ($n \geq 3$), it is quite hard to understand deeply about $\overline{H}$. We discuss now a classical and famous example pointed out by Hedlund [79]. See Bangert [12, Section 5] and E [44] for more modern accounts of this example.

Let us consider the simplest case in three dimensions ($n = 3$) with Hamiltonian

$$H(y,p) = \frac{1}{a(y)}|p| \qquad \text{for all } (y,p) \in \mathbb{R}^3 \times \mathbb{R}^3, \tag{8.13}$$

where $a : \mathbb{R}^3 \to [\delta, 1+\delta]$ is a smooth $\mathbb{Z}^3$-periodic function satisfying

(i) $a \geq 1$ outside $U_\delta(\mathcal{L})$ and $\min_{\mathbb{R}^3} a = \delta$;

(ii) $a(y) = \delta$ if and only if $y \in \mathcal{L}$.

Here,

$$\mathcal{L} = \bigcup_{i=1}^{3}\left(l_i + \mathbb{Z}^3\right)$$

where $l_1 = \mathbb{R}\times\{0\}\times\{0\}$, $l_2 = \{0\}\times\mathbb{R}\times\{\frac{1}{2}\}$ and $l_3 = \{\frac{1}{2}\}\times\{\frac{1}{2}\}\times\mathbb{R}$ are straight lines in $\mathbb{R}^3$. The constant $\delta \in (0, 10^{-2})$ is fixed, and U_δ is the Euclidean δ-neighborhood of $\mathcal{L}$, which is basically the union of tubes. For $1 \leq i \leq 3$, each l_i is a minimizing geodesic for the Riemannian metric $ds^2 = a(y)^2\sum_{i=1}^{3} dy_i^2$. It is important noting that the tubes around l_i for $1 \leq i \leq 3$ stay away from each other.

Of course, we can think of H as $H \in C(\mathbb{T}^3 \times \mathbb{R}^3)$. It is clear here that H is convex, but not uniformly convex in p, and it corresponds to a front propagation problem, which is extremely important in practice. This Hamiltonian is often called a metric Hamiltonian in the literature. It turns out that in this case, we have an explicit formula for $\overline{H}$ as following.

Theorem 8.5. *Assume that H is of the form* (8.13). *Let $\overline{H}$ be its corresponding effective Hamiltonian. Then,*

$$\overline{H}(p) = \frac{1}{\delta}\max\{|p_1|,|p_2|,|p_3|\} \qquad \textit{for } p = (p_1,p_2,p_3) \in \mathbb{R}^3.$$

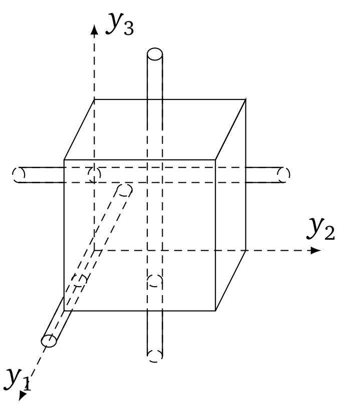


Figure 8.1: Shape of $U_\delta(\mathcal{L})$

Theorem 8.5 was proved by Bangert [12] in the dual form of the stable norm. We give here a purely PDE proof.

Proof. By the inf-sup (or inf-max) formula, we have, for $p \in \mathbb{R}^3$,

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{T}^3)} \max_{y \in \mathbb{T}^3} \frac{1}{a(y)} |p + D\phi(y)|.$$

It is clear that $\overline{H}$ is positively 1-homogeneous. Fix $p \in \mathbb{R}^3$ with $|p| \geq 1$. Without loss of generality, let us assume $|p_1| \geq |p_2|, |p_3|$. For each $\phi \in C^1(\mathbb{T}^3)$, on l_1, $y_1 \mapsto \phi(y_1, 0, 0)$ has a minimum at some $\bar{y} = (\bar{y}_1, 0, 0) \in \mathbb{T}^3$. This implies

$$\max_{y \in \mathbb{T}^3} \frac{1}{a(y)} |p + D\phi(y)| \geq \frac{1}{a(\bar{y})} |p + D\phi(\bar{y})| \geq \frac{1}{\delta} |p_1|.$$

Thus,

$$\overline{H}(p) \geq \frac{1}{\delta} \max\{|p_1|, |p_2|, |p_3|\}.$$

To prove the converse, we construct a corresponding subsolution $\varphi \in C^1(\mathbb{T}^3)$ so that

$$\varphi(y) = \begin{cases} -(p_2 y_2 + p_3 y_3) & \text{for } y \in U_\delta(l_1) \cap \mathbb{T}^3, \\ -(p_1 y_1 + p_3(y_3 - \frac{1}{2})) & \text{for } y \in U_\delta(l_2) \cap \mathbb{T}^3, \\ -(p_1(y_1 - \frac{1}{2}) + p_2(y_2 - \frac{1}{2})) & \text{for } y \in U_\delta(l_3) \cap \mathbb{T}^3, \\ 0 & \text{for } y \in \mathbb{T}^3 \setminus U_{2\delta}(l_1 \cup l_2 \cup l_3), \end{cases}$$

and $|D\varphi| \leq C$ in $\mathbb{T}^3$. This is possible as $|\varphi(y)| \leq C\delta$ for $y \in U_\delta(l_1 \cup l_2 \cup l_3) \cap \mathbb{T}^3$. Then, it is quite straightforward to check that

$$\frac{1}{a(y)} |p + D\varphi(y)| \leq \frac{1}{\delta} \max\{|p_1|, |p_2|, |p_3|\} \qquad \text{in } \mathbb{T}^3.$$

In fact, the above inequality is strict for all $y \in \mathbb{T}^3 \setminus (l_1 \cup l_2 \cup l_3)$. The proof is complete. □

Remark 8.6. Let us discuss more about the Hedlund example here. As

$$\overline{H}(p) = \frac{1}{\delta}\max\{|p_1|,|p_2|,|p_3|\} \qquad \text{for } p=(p_1,p_2,p_3)\in\mathbb{R}^3,$$

it is clear that $\overline{H}$ is only Lipschitz, not differentiable in $\mathbb{R}^3$, and its level sets are concentric cubes in $\mathbb{R}^3$. Moreover, if $\overline{H}$ is differentiable at p, then $D\overline{H}(p)\in\left\{\frac{e_1}{\delta},\frac{e_2}{\delta},\frac{e_3}{\delta}\right\}$. If $D\overline{H}(p)=\frac{e_i}{\delta}$ for some $1\le i\le 3$, then a corresponding backward characteristic is l_i. It is not hard to show that l_i is the unique trajectory of the projected Mather set at p.

Although we do not discuss in deep the projected Aubry set here, the above proof also gives that the projected Aubry set at each $p\in\mathbb{R}^3$ can contain at most $l_1\cup l_2\cup l_3$. And as the Aubry set is bigger than the projected Mather set, this also means that the projected Mather set is always a subset of $l_1\cup l_2\cup l_3$. Thus, classically, to obtain rotation vectors from backward characteristics, we are only able to get three rotation vectors $\left\{\frac{e_1}{\delta},\frac{e_2}{\delta},\frac{e_3}{\delta}\right\}$. This gives a detailed explanation for Remark 7.7.

This Hedlund example also explains a weakness of weak KAM theory in dimensions three and higher, where the projected Aubry and projected Mather sets might only occupy a tiny part of $\mathbb{T}^n$, and do not give us much information. Notice that a solution $u\in \mathrm{Lip}(\mathbb{T}^n)$ to our cell problem is differentiable almost everywhere, and thus, the set of differentiable points of u is much richer than Aubry and projected Mather sets in this situation.

4 A generalization of the classical Hedlund example

In this section, we assume that $n\ge 3$, and we provide a generalization of the classical Hedlund example as following.

Theorem 8.7. *Assume that $n\ge 3$. Let $P\subset\mathbb{R}^n$ be a centrally symmetric polytope with rational vertices and nonempty interior. Then, we are able to construct explicitly a function $a\in C^\infty(\mathbb{T}^n,(0,\infty))$ such that, for $H(y,p)=\frac{1}{a(y)}|p|$ for $(y,p)\in\mathbb{T}^n\times\mathbb{R}^n$, then the corresponding $\overline{H}$ is*

$$\overline{H}(p)=\max_{q\in P} p\cdot q \qquad \text{for all } p\in\mathbb{R}^n.$$

We present here a proof following Jing, Tran, Yu [93], which is quite simple and has similar flavors as that of Theorem 8.5. This result was also presented in Babenko, Balacheff [10], Jotz [94] in the equivalent form of stable norms.

Proof. Assume that the vertices of P are $\pm q_1,\pm q_2,\dots,\pm q_m$, which are rational vectors in $\mathbb{R}^n$. Denote by $L_i=\{tq_i\,:\,t\in\mathbb{R}\}$ for $1\le i\le m$. Since P is convex and has nonempty interior, $q_1,q_2,\dots,q_m$ are mutually non-parallel and

$$\mathrm{span}\{q_1,q_2,\dots,q_m\}=\mathbb{R}^n.$$

As a result,

$$\theta=\min_{|p|=1}\max_{1\le i\le m}|p\cdot q_i|>0.$$

As usual, we divide the proof into several steps for clarity.

Step 1. Construction of a_A. Let $y_1 = 0$. For $k \le m-1$, choose inductively that

$$y_{k+1} \in (0,1)^n \setminus \bigcup_{i=1}^{k} \{y_i + sq_{k+1} + tq_i + \mathbb{Z}^n : s,t \in \mathbb{R}\}.$$

Then for such selected points $y_1, y_2, \ldots, y_m \in (0,1)^n$, we have that for $i \neq j$,

$$(y_i + L_i + \mathbb{Z}^n) \cap (y_j + L_j + \mathbb{Z}^n) = \emptyset. \tag{8.14}$$

Due to the fact that $\{q_i\}$ are rational vectors, the projection of $\{y_i + L_i + \mathbb{Z}^n\}$ to the flat torus $\mathbb{T}^n$ is a closed orbit for each $1 \le i \le m$. As a result, we can choose a sufficiently small number $\delta \in (0, 1/3)$ so that, for $i \neq j$,

$$U_{\delta,i} \cap U_{\delta,j} = \emptyset,$$

where

$$U_{\delta,i} = \{y \in \mathbb{R}^n : \operatorname{dist}(y, y_i + L_i + \mathbb{Z}^n) \le \delta\}.$$

Choose a smooth $\mathbb{Z}^n$-periodic function $a_A : \mathbb{R}^n \to (0, \infty)$ such that

$$\begin{cases} a_A(y) = \frac{1}{A|q_i|} & \text{on } y_i + L_i + \mathbb{Z}^n \text{ for } 1 \le i \le m, \\ \frac{1}{A|q_i|} \le a_A(y) \le 1 & \text{on } U_{\delta,i} \text{ for } 1 \le i \le m, \\ a_A(y) = 1 & \text{on } \mathbb{R}^n \setminus \bigcup_{i=1}^m U_{\delta,i}. \end{cases}$$

Here, $A > 0$ is a large positive constant to be determined later.

Next, for every unit vector $|p| = 1$ and $1 \le i \le m$, write

$$p_i^\perp = p - \frac{(p \cdot q_i)}{|q_i|^2} q_i,$$

which is the projection of p onto the $(n-1)$-dimensional Euclidean subspace of $\mathbb{R}^n$ that is perpendicular to q_i. Apparently, we can construct a smooth $\mathbb{Z}^n$-periodic function ϕ satisfying that

$$D\phi(y) = -p_i^\perp \qquad \text{in } U_{\delta,i} \text{ for all } 1 \le i \le m,$$

and

$$\|D\phi\|_{L^\infty} \le C_\delta,$$

for a constant $C_\delta > 0$ depending only on δ, and $q_1, q_2, \ldots, q_m$. We now pick A such that

$$A \ge \max\left\{ \frac{1 + C_\delta}{\theta}, \frac{1}{\min_{1 \le i \le m} |q_i|} \right\}.$$

Step 2. Characterization of the effective Hamiltonian. Let $\overline{H}_A$ be the effective Hamiltonian corresponding to the Hamiltonian $a_A(y)|p|$. We claim that

$$\overline{H}_A(p) = A \max_{1 \le i \le m} |p \cdot q_i| \qquad \text{for all } p \in \mathbb{R}^n. \tag{8.15}$$

We only need to prove this claim for unit vectors $|p| = 1$. Let us fix such a p. Firstly, by using ϕ and the choice of A above, it is clear that

$$\begin{aligned}\overline{H}_A(p) &\le \max_{y\in\mathbb{R}^n} \frac{1}{a_A(y)}|p + D\phi(y)| \\ &\le \max\left\{A \max_{1\le i\le m} |p\cdot q_i|, 1 + C_\delta\right\} = A \max_{1\le i\le m} |p\cdot q_i|.\end{aligned}$$

Secondly, let $v = v_p$ be a solution of the corresponding ergodic problem

$$\frac{1}{a_A(y)}|p + Dv(y)| = \overline{H}_A(p) \qquad \text{in } \mathbb{T}^n.$$

For simplicity, we assume $v \in C^1(\mathbb{T}^n)$ (to make this rigorous, one needs to do convolution with a standard mollifier, but we omit it here as this was done in various earlier proofs already). Then, for each $1 \le i \le m$,

$$A|q_i| \cdot |p + Dv(y)| = \overline{H}_A(p) \qquad \text{for } y \in y_i + L_i + \mathbb{Z}^n.$$

Denote by $u(y) = p\cdot y + v(y)$ for $y \in \mathbb{R}^n$. Choose $m \in \mathbb{Z}$ such that $mq_i \in \mathbb{Z}^n$. Thanks to the periodicity of v,

$$u(y_i + mq_i) - u(y_i) = mp \cdot q_i.$$

Since $|u(y_i + mq_i) - u(y_i)| \le m|q_i| \max_{y\in x_i+L_i} |Du(y)|$, we deduce that

$$\overline{H}_A(p) \ge A|p\cdot q_i|.$$

Therefore, (8.15) holds true.

Step 3. Construction of a. Let $a(y) = Aa_A(y)$ for $y \in \mathbb{R}^n$. Then by scaling the result of Step 2, the effective Hamiltonian $\overline{H}(p)$ is

$$\overline{H}(p) = \frac{\overline{H}_A(p)}{A} = \max_{1\le i\le m} |p\cdot q_i| = \max_{q\in\{\pm q_1,\cdots,\pm q_m\}} p\cdot q = \max_{q\in P} p\cdot q, \qquad \text{for all } p \in \mathbb{R}^n.$$

This completes the proof. Basically, it means that $\overline{H}$ is a convex, positively 1-homogeneous function with the support set P. □

Remark 8.8. From the constructions in the above proof, we observe two following simple but important points.

- This kind of construction does not work in two dimensions. Indeed, in two dimensions, there is no room for us in order to have that (8.14) holds.
- By properly choosing the rational vectors $\{q_i\}_{1\le i\le m}$ and δ, it is not hard to construct a sequence $\{a_m(\cdot)\} \subset C^\infty(\mathbb{T}^n)$ such that

  $$0 < a_m \le 1, \qquad \lim_{m\to\infty} a_m(y) = 0 \qquad \text{for a.e. } y \in \mathbb{T}^n,$$

  and

  $$\lim_{m\to\infty} \overline{H}_m(p) = |p| \qquad \text{locally uniformly in } \mathbb{R}^n.$$

  Here, $\overline{H}_m$ is the effective Hamiltonian corresponding to the Hamiltonian $a_m(y)|p|$.

5 References

1. Strict convexity of the effective Hamiltonian in certain directions is taken from Evans, Gomes [53].

2. Asymptotic expansion at infinity is taken from Luo, Tran, Yu [107]. See also Jing, Tran, Yu [92], and Tran, Yu [132]. The method of asymptotic expansion at infinity was used in [107, 92, 132] to study inverse problems on how V affects $\overline{H}$. This can be seen also from the above expansion of $\overline{H}$.

3. The classical Hedlund example was studied by Hedlund [79]. Then, Bangert [12] and E [44] give connections of this example to Aubry–Mather theory and weak KAM theory. Still, optimal rate of convergence of homogenization holds for this Hamiltonian (see Mitake, Tran, Yu [118]).

4. The proof of Theorem 8.7 is taken from Jing, Tran, Yu [93]. This result was also presented in Babenko, Balacheff [10], Jotz [94] in the equivalent form of stable norms with more complicated proofs.

Appendix

In Appendix, we cover some important results that we need in the book.

1 Sion's minimax theorem

Here is the statement of the theorem.

Theorem A.1 (Sion's minimax theorem)**.** *Let X be a compact convex subset of a linear topological space, and Y be a convex subset of a linear topological space. Let $f : X \times Y \to \mathbb{R}$ be a function such that*

(i) $f(x, \cdot)$ is upper semicontinuous and quasiconcave on Y for each $x \in X$;

(ii) $f(\cdot, y)$ is lower semicontinuous and quasiconvex on X for each $y \in Y$.

Then,

$$\min_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \min_{x \in X} f(x, y).$$

We always assume the settings of Theorem A.1 in this section. We follow here a proof by Komiya [97], which is quite elementary. Here are two preparatory lemmas.

Lemma A.2. *Assume that there are $y_1, y_2 \in Y$ and $a \in \mathbb{R}$ such that*

$$a < \min_{x \in X} \max\{f(x, y_1), f(x, y_2)\}.$$

Then, there exists $y_0 \in Y$ such that

$$a < \min_{x \in X} f(x, y_0).$$

Proof. Assume by contradiction that $a \geq \min_{x \in X} f(x, y)$ for all $y \in Y$. Pick $b \in \mathbb{R}$ such that

$$a < b < \min_{x \in X} \max\{f(x, y_1), f(x, y_2)\}.$$

Denote by $[y_1, y_2]$ the line segment connecting y_1 and y_2. For each $z \in [y_1, y_2]$, set

$$C_z = \{x \in X : f(x, z) \leq a\} \quad \text{and} \quad D_z = \{x \in X : f(x, z) \leq b\}.$$

Let $A = D_{y_1}$ and $B = D_{y_2}$. It is clear that C_z, D_z, A, B are all nonempty, convex closed sets in X because of the lower semicontinuity and quasiconvexity of $f(\cdot, z)$. In particular, C_z, D_z, A, B are all connected sets. Moreover, by our hypothesis, $A \cap B = \emptyset$.

On the other hand, the quasiconcavity of $f(x, \cdot)$ gives, for $z \in [y_1, y_2]$,

$$f(x,z) \geq \min\{f(x, y_1), f(x, y_2)\},$$

which yields $D_z \subset A \cup B$. The connectedness of D_z yields that

$$C_z \subset D_z \subset A \quad \text{or} \quad C_z \subset D_z \subset B.$$

Denote by

$$I = \{z \in [y_1, y_2] : C_z \subset A\} \quad \text{and} \quad J = \{z \in [y_1, y_2] : C_z \subset B\}.$$

Then, of course, $I, J \neq \emptyset$, $I \cap J = \emptyset$, and $I \cup J = [y_1, y_2]$. As $[y_1, y_2]$ is connected, we will show that I, J are both closed to arrive at a contradiction. It is enough to show that I is closed. Take a sequence $\{z_k\} \subset I$ such that $z_k \to z \in [y_1, y_2]$ as $k \to \infty$. Pick $x \in C_z$, then $f(x,z) \leq a$. By the upper semicontinuity of $f(x, \cdot)$,

$$\limsup_{k \to \infty} f(x, z_k) \leq f(x,z) \leq a.$$

Hence, we can find $k \in \mathbb{N}$ sufficiently large such that $f(x, z_k) < b$, which means that $x \in D_{z_k} \subset A$ by the fact that $\{z_k\} \subset I$. Thus, $C_z \subset A$, and $z \in I$. The proof is complete. □

We apply induction to the above lemma to have the following.

Lemma A.3. *Assume that there are $y_1, y_2, \ldots, y_n \in Y$ and $a \in \mathbb{R}$ such that*

$$a < \min_{x \in X} \max\{f(x, y_i) : 1 \leq i \leq n\}.$$

Then, there exists $y_0 \in Y$ such that

$$a < \min_{x \in X} f(x, y_0).$$

Proof. We prove by induction. There is nothing to prove if $n = 1$. Assume that the lemma holds for $n = m - 1$ for $m \geq 2$. We show that it holds for $n = m$. Let

$$X' = \{x \in X : f(x, y_m) \leq a\}.$$

If $X' = \emptyset$, then choose $y_0 = y_m$ to conclude. Otherwise, X' is a nonempty, convex, compact set. Of course, we have

$$a < \min_{x \in X'} \max\{f(x, y_i) : 1 \leq i \leq m - 1\}.$$

By the induction hypothesis, there exists $y_0' \in Y$ such that $\min_{x \in X'} f(x, y_0') > a$, which implies

$$a < \min_{x \in X} \max\{f(x, y_0'), f(x, y_m)\}.$$

Apply Lemma A.2 to conclude. □

We are now ready to prove Sion's minimax theorem.

Proof of Theorem A.1. It is always the case that

$$\sup_{y\in Y}\min_{x\in X} f(x,y) \le \min_{x\in X}\sup_{y\in Y} f(x,y).$$

To complete, we need to prove the converse. Pick an arbitrary $a\in\mathbb{R}$ such that

$$a < \min_{x\in X}\sup_{y\in Y} f(x,y).$$

For $y\in Y$, let $X_y=\{x\in X\,:\, f(x,y)\le a\}$. Then $\bigcap_{y\in Y} X_y=\emptyset$, and the compactness of X infers that there are $y_1,\ldots,y_n\in Y$ such that $\bigcap_{i=1}^n X_{y_n}=\emptyset$. Therefore,

$$a < \min_{x\in X}\max\{f(x,y_i)\,:\,1\le i\le n\}.$$

By Lemma A.3, we find that there is $y_0\in Y$ so that $a<\min_{x\in X} f(x,y_0)$, which yields

$$\sup_{y\in Y}\min_{x\in X} f(x,y)\ge a.$$

Hence,

$$\sup_{y\in Y}\min_{x\in X} f(x,y) \ge \min_{x\in X}\sup_{y\in Y} f(x,y).$$

□

2 Existence and regularity of minimizers for action functionals

In this section, we study the existence and regularity of minimizers for action functionals. Let $L=L(y,v):\mathbb{T}^n\times\mathbb{R}^n\to\mathbb{R}$ be the usual Lagrangian. For our purpose, we only consider the spatial variable y in the flat n-dimensional torus instead of $\mathbb{R}^n$. We always assume in this section the following

$$\begin{cases} L\in C^2(\mathbb{T}^n\times\mathbb{R}^n),\\ \text{there exists } \theta>0 \text{ such that } \theta I_n\le D^2_{vv}L(y,v)\le\theta^{-1}I_n \text{ for all } (y,v)\in\mathbb{T}^n\times\mathbb{R}^n.\end{cases} \tag{A.1}$$

It is straightforward to see that (A.1) gives us nice bounds of L and D_vL as following. Firstly, it is clear that there exists $C>0$ such that

$$|D_vL(y,v)|\le C(1+|v|) \qquad \text{for all } (y,v)\in\mathbb{T}^n\times\mathbb{R}^n.$$

Secondly, by making $\theta>0$ smaller if needed, we have

$$\frac{\theta}{2}|v|^2-K_0\le L(y,v)\le\frac{1}{2\theta}|v|^2+K_0 \qquad \text{for all } (y,v)\in\mathbb{T}^n\times\mathbb{R}^n,$$

for some $K_0>0$.

Let $g\in \mathrm{Lip}(\mathbb{T}^n)$ be a given function. Fix $T>0$ and $x_1\in\mathbb{T}^n$. Consider the following variational problem

$$u(x_1,T)=\inf\left\{\int_0^T L(\gamma(s),\gamma'(s))\,ds+g(\gamma(0))\,:\,\gamma\in \mathrm{AC}([0,T],\mathbb{T}^n),\ \gamma(T)=x_1\right\}. \tag{A.2}$$

2.1 Existence of minimizers

Here is our theorem on existence of minimizers for the above action functional.

Theorem A.4. *Assume* (A.1). *Then* (A.2) *admits a minimizer* $\gamma \in \mathrm{AC}([0,T],\mathbb{T}^n)$.

We need various preparations before proving this result. Firstly, we need the following result, which is a classical result in Calculus of Variations on the existence of a minimizer with fixed endpoints.

Lemma A.5. *Fix* $x_0 \in \mathbb{T}^n$. *Define*

$$V(x_0) = \inf\left\{\int_0^T L(\gamma(s),\gamma'(s))\,ds \;:\; \gamma \in \mathrm{AC}([0,T],\mathbb{T}^n),\, \gamma(0)=x_0, \gamma(T)=x_1\right\}.$$

Then there is a minimizer for $V(x_0)$.

We note first that V is surely always bounded. Fix $x_0 \in \mathbb{T}^n$. On one hand, as $L(y,v) \geq -K_0$ for all $(y,v) \in \mathbb{T}^n \times \mathbb{R}^n$, $V(x_0) \geq -K_0 T$. On the other hand, for $\gamma_0 : [0,T] \to \mathbb{T}^n$ such that $\gamma_0(s) = x_0 + \frac{s}{T}(x_1 - x_0)$ for $0 \leq s \leq T$, we have

$$V(x_0) \leq \int_0^T L(\gamma_0(s),\gamma_0'(s))\,ds \leq \left(\frac{|x_1-x_0|^2}{2\theta T^2} + K_0\right)T \leq \left(\frac{n}{2\theta T^2} + K_0\right)T \leq C.$$

Next is a key point to prove Lemma A.5 and Theorem A.4.

Lemma A.6. *Let* $\{\gamma_k\} \subset \mathrm{AC}([0,T],\mathbb{T}^n)$ *with* $\gamma_k(T) = x_1$ *for all* $k \in \mathbb{N}$. *Assume that there is a constant* $C > 0$ *such that*

$$\int_0^T L(\gamma_k(s),\gamma_k'(s))\,ds \leq C \qquad \text{for all } k \in \mathbb{N}.$$

Then, there exist a subsequence $\{\gamma_{k_j}\}$ *of* $\{\gamma_k\}$ *and* $\gamma \in \mathrm{AC}([0,T],\mathbb{T}^n)$ *such that*

$$\gamma_{k_j} \to \gamma \qquad \text{uniformly on } [0,T],$$

as $j \to \infty$, *and*

$$\int_0^T L(\gamma(s),\gamma'(s))\,ds \leq \liminf_{k\to\infty} \int_0^T L(\gamma_k(s),\gamma_k'(s))\,ds.$$

Basically, this is a result on compactness and lower semicontinuity of the action functional. We postpone the proof of Lemma A.6 for later. Let us now use it to prove Lemma A.5 and Theorem A.4 first.

Proof of Lemma A.5. Fix $x_0 \in \mathbb{T}^n$. As explained earlier, $V(x_0)$ is bounded. Pick a minimizing sequence $\{\gamma_k\} \subset \mathrm{AC}([0,T],\mathbb{T}^n)$ for $V(x_0)$ with $\gamma_k(0) = x_0$, $\gamma_k(T) = x_1$ such that

$$\int_0^T L(\gamma_k(s),\gamma_k'(s))\,ds \leq V(x_0) + \frac{1}{k} \leq C + 1 \qquad \text{for all } k \in \mathbb{N}.$$

Thanks to Lemma A.6, we find a subsequence $\{\gamma_{k_j}\} \subset \mathrm{AC}([0,T],\mathbb{T}^n)$ and $\gamma \in \mathrm{AC}([0,T],\mathbb{T}^n)$ such that

$$\gamma_{k_j} \to \gamma \qquad \text{uniformly on } [0,T],$$

as $j \to \infty$, and

$$\int_0^T L(\gamma(s), \gamma'(s))\, ds \le \liminf_{k\to\infty} \int_0^T L(\gamma_k(s), \gamma_k'(s))\, ds \le V(x_0).$$

The uniform convergence of $\{\gamma_{k_j}\}$ to γ on $[0, T]$ also gives that $\gamma(0) = x_0$ and $\gamma(T) = x_1$. Thus, γ is a minimizer for $V(x_0)$. □

We have in addition that V is lower semicontinuous in $\mathbb{T}^n$.

Lemma A.7. *The function V is lower semicontinuous in $\mathbb{T}^n$.*

Proof. Pick a sequence $\{y_k\} \subset \mathbb{T}^n$ that converges to $x_0 \in \mathbb{T}^n$. We aim at showing

$$V(x_0) \le \liminf_{k\to\infty} V(y_k).$$

For each $k \in \mathbb{N}$, we can find $\gamma_k \in \mathrm{AC}([0, T], \mathbb{T}^n)$ such that $\gamma_k(0) = y_k$, $\gamma_k(T) = x_1$, and

$$\int_0^T L(\gamma_k(s), \gamma_k'(s))\, ds = V(y_k) \le C.$$

We use Lemma A.6 again to find a subsequence $\{\gamma_{k_j}\} \subset \mathrm{AC}([0, T], \mathbb{T}^n)$ and $\gamma \in \mathrm{AC}([0, T], \mathbb{T}^n)$ such that

$$\gamma_{k_j} \to \gamma \qquad \text{uniformly on } [0, T],$$

as $j \to \infty$, and

$$\int_0^T L(\gamma(s), \gamma'(s))\, ds \le \liminf_{k\to\infty} \int_0^T L(\gamma_k(s), \gamma_k'(s))\, ds = \liminf_{k\to\infty} V(y_k).$$

As $\gamma(0) = x_0$ and $\gamma(T) = x_1$, we conclude that

$$V(x_0) \le \int_0^T L(\gamma(s), \gamma'(s))\, ds \le \liminf_{k\to\infty} \int_0^T L(\gamma_k(s), \gamma_k'(s))\, ds = \liminf_{k\to\infty} V(y_k).$$

□

Proof of Theorem A.4. Recall that, by definition of $u(x_1, T)$ in (A.2), we have

$$u(x_1, T) = \inf_{x\in\mathbb{T}^n} \left(V(x) + g(x)\right).$$

As $V + g$ is lower semicontinuous in $\mathbb{T}^n$, it attains its minimum at a point $x_0 \in \mathbb{T}^n$. By Lemma A.5, there is a minimizer γ for $V(x_0)$, and therefore, γ is also a minimizer for $u(x_1, T)$. □

Let us now proceed to prove Lemma A.6.

Lemma A.8. *Assume the settings in Lemma A.6. Then, the sequence $\{\gamma_k\}$ is equi-absolutely continuous on $[0, T]$.*

Proof. By our assumption (A.1) on L, for all $k \in \mathbb{N}$,

$$\int_0^T \frac{\theta}{2}|\gamma_k'(s)|^2\,ds \le \int_0^T \left(L(\gamma_k(s),\gamma_k'(s)) + K_0\right)ds \le C + K_0 T \le C.$$

Thus, for any Borel set $B \subset [0,T]$ and any $k \in \mathbb{N}$,

$$\int_B |\gamma_k'(s)|\,ds \le \left(\int_B |\gamma_k'(s)|^2\,ds\right)^{1/2}\left(\int_B 1\,ds\right)^{1/2} \le C|B|^{1/2}.$$

Here, $|B|$ denotes the Lebesgue measure of B. The above implies the conclusion. □

Proof of Lemma A.6. By Lemma A.8, we are able to find a subsequence $\{\gamma_{k_j}\}$ of $\{\gamma_k\}$ and $\gamma \in \mathrm{AC}([0,T],\mathbb{T}^n)$ such that

$$\begin{cases} \gamma_{k_j} \to \gamma \text{ uniformly on } [0,T], \\ \gamma_{k_j}' \rightharpoonup \gamma' \text{ weakly in } L^2([0,T]). \end{cases} \tag{A.3}$$

Note that, the convexity of L yields

$$\int_0^T L(\gamma_{k_j}(s),\gamma_{k_j}'(s))\,ds \ge \int_0^T \left(L(\gamma_{k_j}(s),\gamma'(s)) + D_vL(\gamma_{k_j}(s),\gamma'(s))\cdot(\gamma_{k_j}'(s)-\gamma'(s))\right)ds.$$

By using the bounds on L, D_vL and (A.3), we obtain

$$\lim_{j\to\infty}\int_0^T L(\gamma_{k_j}(s),\gamma'(s))\,ds = \int_0^T L(\gamma(s),\gamma'(s))\,ds,$$

and

$$\lim_{j\to\infty}\int_0^T D_vL(\gamma_{k_j}(s),\gamma'(s))\cdot(\gamma_{k_j}'(s)-\gamma'(s))\,ds = 0.$$

The proof is complete. □

For more complicated situations about existence of minimizers, see Cannarsa, Sinestrari [26], Evans [49], Ishii [85].

2.2 Regularity of minimizers

Theorem A.9. *Assume* (A.1). *Let* $\gamma \in \mathrm{AC}([0,T],\mathbb{T}^n)$ *be a minimizer in* (A.2). *Then* $\gamma \in C^2([0,T])$.

Sketch of proof. By the calculus of variation theory, γ solves the following Euler–Lagrange equation

$$\frac{d}{dt}\left(D_vL(\gamma(t),\gamma'(t))\right) = D_xL(\gamma(t),\gamma'(t)) \qquad \text{on } [0,T].$$

Denote by $X(t) = \gamma(t)$, and $P(t) = D_vL(\gamma(t),\gamma'(t))$ for $t \in [0,T]$. Then (X,P) solves the following Hamiltonian system

$$\begin{cases} X'(t) = D_pH(X(t),P(t)), \\ P'(t) = -D_xH(X(t),P(t)), \end{cases} \qquad \text{for } t \in [0,T].$$

As $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, we get that $X \in C^2([0,T])$, which means $\gamma \in C^2([0,T])$.

Furthermore, it is worth noting here that we have conservation of energy, that is, $t \mapsto H(X(t),P(t))$ is constant. This can be easily checked as

$$\frac{d}{dt}H(X(t),P(t)) = D_x H(X(t),P(t)) \cdot X'(t) + D_p H(X(t),P(t)) \cdot P'(t) = 0.$$

In particular, this allows us to get that $|P(t)| \leq C$, and also $|\gamma'(t)| \leq C$ for all $t \in [0,T]$. $\square$

3 Characterization of the Legendre transform

This is taken from the paper of Artstein-Avidan, Milman [8]. Let us first provide the setting.

Denote the class of lower semi-continuous convex functions $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ by $\mathrm{Cvx}(\mathbb{R}^n)$. Clearly, the only function in $\mathrm{Cvx}(\mathbb{R}^n)$ that attains the value $-\infty$ is the constant $-\infty$ function. Recall that, for $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$, its Legendre transform ϕ^* is defined as

$$\phi^*(x) = \sup_{y \in \mathbb{R}^n} (y \cdot x - \phi(y)).$$

And moreover, $(\phi^*)^* = \phi$. It is straightforward from these that the Legendre transform has two basic properties

$$\begin{cases} \text{For } \phi \in \mathrm{Cvx}(\mathbb{R}^n),\ (\phi^*)^* = \phi, \\ \text{For } \phi, \psi \in \mathrm{Cvx}(\mathbb{R}^n) \text{ so that } \phi \geq \psi, \text{ then } \phi^* \leq \psi^*. \end{cases}$$

In the following, we show that if a transformation from $\mathrm{Cvx}(\mathbb{R}^n)$ to $\mathrm{Cvx}(\mathbb{R}^n)$ that respects the above two properties is essentially the Legendre transform.

Theorem A.10. *Assume that $T : \mathrm{Cvx}(\mathbb{R}^n) \to \mathrm{Cvx}(\mathbb{R}^n)$ is a transformation satisfying*

- $T(T\phi) = \phi$,
- $\phi \geq \psi$ *implies* $T\phi \leq T\psi$.

Then, T is essentially the Legendre transformation, that is, there exist $c_0 \in \mathbb{R}$, $v_0 \in \mathbb{R}^n$, and an invertible symmetric matrix B of size n such that

$$(T\phi)(x) = \phi^*(Bx + v_0) + v_0 \cdot x + c_0.$$

Let us now proceed to prove this main theorem, which was obtained by Artstein-Avidan, Milman [8]. We always assume the settings of Theorem A.10 in this section. For a family $\{f_\alpha\}_{\alpha \in A} \subset \mathrm{Cvx}(\mathbb{R}^n)$, we have that $\sup_\alpha f_\alpha \in \mathrm{Cvx}(\mathbb{R}^n)$. It is not always the case that $\inf_\alpha f_\alpha$ is convex. We denote by $\check{\inf}_\alpha f_\alpha$ the convex envelope of $\inf_\alpha f_\alpha$.

3.1 Preliminaries

We have the following preparatory results.

Lemma A.11. *Fix a family* $\{f_\alpha\}_{\alpha\in\mathcal{A}} \subset \mathrm{Cvx}(\mathbb{R}^n)$ *The following identities hold.*

$$T(\check{\inf}_\alpha f_\alpha) = \sup_\alpha T(f_\alpha) \quad \textit{and} \quad \check{\inf}_\alpha T(f_\alpha) = T(\sup_\alpha f_\alpha).$$

Proof. We only prove the first identity as the second one follows in an analogous way. First of all, it is clear that $T(\check{\inf}_\alpha f_\alpha) \geq T(f_\alpha)$ for each $\alpha \in \mathcal{A}$. Therefore,

$$T(\check{\inf}_\alpha f_\alpha) \geq \sup_\alpha T(f_\alpha).$$

On the other hand, as T is surjective, there is $g \in \mathrm{Cvx}(\mathbb{R}^n)$ such that $\sup_\alpha T(f_\alpha) = Tg$. Then, $g \leq f_\alpha$ for all $\alpha \in \mathcal{A}$, and as a result, $g \leq \inf_\alpha f_\alpha$. By the definition of convex envelopes, $g \leq \check{\inf}_\alpha f_\alpha$. Hence,

$$\sup_\alpha T(f_\alpha) = Tg \geq T(\check{\inf}_\alpha f_\alpha).$$

□

Next, we see that it is enough to understand the analysis for affine and delta type functions in order to get the conclusion. For $z \in \mathbb{R}^n$, denote by

$$D_z(x) = \begin{cases} 0 & \text{for } x = z, \\ +\infty & \text{for } x \neq z. \end{cases}$$

Of course $D_z + c \in \mathrm{Cvx}(\mathbb{R}^n)$ for all $z \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Besides, for any function f, we can always express that

$$f(x) = \inf_{y\in\mathbb{R}^n} \left(D_y(x) + f(y)\right) \qquad \text{for } x \in \mathbb{R}^n.$$

Lemma A.12. *Assume that there exist* $c_0 \in \mathbb{R}, c_1 > 0$*, an invertible matrix* B *of size* n*, and* $v_0, v_1 \in \mathbb{R}^n$ *such that*

$$T(D_z + c)(x) = (Bz + v_1)\cdot x + v_0 \cdot z - c_1 c + c_0.$$

Then, for all $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$*,*

$$(T\phi)(x) = c_1\phi^*(\overline{B}x + \overline{v}_0) + v_1 \cdot x + c_0.$$

Here, $\overline{B} = B^T/c_1$*, and* $\overline{v}_0 = v_0/c_1$*.*

Proof. Fix $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$. Recall that

$$\phi(x) = \inf_{y\in\mathbb{R}^n}\left(D_y(x) + \phi(y)\right) = \check{\inf}_{y\in\mathbb{R}^n}\left(D_y(x) + \phi(y)\right).$$

Therefore,

$$\begin{aligned}
(T\phi)(x) &= \sup_{y\in\mathbb{R}^n}\left((By + v_1)\cdot x + v_0\cdot y - c_1\phi(y) + c_0\right) \\
&= \sup_{y\in\mathbb{R}^n}\left((By + v_1)\cdot x + v_0\cdot y - c_1\phi(y) + c_0\right) \\
&= \sup_{y\in\mathbb{R}^n}\left(y\cdot(B^T x + v_0) - c_1\phi(y)\right) + v_1\cdot x + c_0 \\
&= c_1\phi^*(\overline{B}x + \overline{v}_0) + v_1\cdot x + c_0.
\end{aligned}$$

□

Next, we conclude that we must have B is symmetric, $c_1 = 1$, and $v_0 = v_1$.

Lemma A.13. *Assume that there exist $c_0 \in \mathbb{R}, c_1 > 0$, an invertible matrix B of size n, and $v_0, v_1 \in \mathbb{R}^n$ such that*

$$T(D_z + c)(x) = (Bz + v_1) \cdot x + v_0 \cdot z - c_1 c + c_0.$$

Then, B is symmetric, $c_1 = 1$, and $v_0 = v_1$. Moreover, for all $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$,

$$(T\phi)(x) = \phi^*(Bx + v_0) + v_0 \cdot x + c_0.$$

Proof. By the previous lemma, we already have, for all $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$,

$$(T\phi)(x) = c_1 \phi^*(\overline{B}x + \overline{v}_0) + v_1 \cdot x + c_0, \tag{A.4}$$

where $\overline{B} = B^T/c_1$, and $\overline{v}_0 = v_0/c_1$.

We note that for $\phi \equiv C$, we have $\phi^* = D_0 - C$. Plug this into (A.4) carefully to derive that

$$T\phi = D_{-(B^T)^{-1}v_0} + c_0 - Cc_1 - v_1 \cdot ((B^T)^{-1}v_0).$$

Then, by the fact that $T(T\phi) = \phi$, we deduce

$$\begin{aligned} C &= (-B(B^T)^{-1}v_0 + v_1) \cdot x + v_0 \cdot (-(B^T)^{-1}v_0) - c_1\left(c_0 - Cc_1 - v_1 \cdot ((B^T)^{-1}v_0)\right) + c_0 \\ &= (-B(B^T)^{-1}v_0 + v_1) \cdot x + (c_1 v_1 - v_0) \cdot (-(B^T)^{-1}v_0) + c_0(1 - c_1) + Cc_1^2. \end{aligned}$$

Since the above holds true for all $C \in \mathbb{R}$ and $x \in \mathbb{R}^n$, we yield that $c_1 = 1$, and $v_1 = B(B^T)^{-1}v_0$. Thus, (A.4) is simplified as, for $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$,

$$(T\phi)(x) = \phi^*(Bx + v_0) + (B(B^T)^{-1}v_0) \cdot x + c_0.$$

Use the identity $T(T\phi) = \phi$ once more to deduce that $(B^T)^{-1}B = I_n$, which gives us that $B = B^T$. ☐

3.2 Affine functions and delta type functions

By the preliminaries, in order to prove Theorem A.10, we just need to verify that there exist $c_0 \in \mathbb{R}, c_1 > 0$, an invertible matrix B of size n, and $v_0, v_1 \in \mathbb{R}^n$ such that

$$T(D_z + c)(x) = (Bz + v_1) \cdot x + v_0 \cdot z - c_1 c + c_0 \qquad \text{for all } z \in \mathbb{R}^n. \tag{A.5}$$

This is a much simpler task since we only need to interact with affine and delta type functions.

Proposition A.14. *Assume the settings in Theorem A.10. Then, there exist $c_0 \in \mathbb{R}, c_1 > 0$, an invertible matrix B of size n, and $v_0, v_1 \in \mathbb{R}^n$ such that* (A.5) *holds.*

To make the proof clear, we break it into various parts.

Lemma A.15. *The map T maps delta type functions to affine functions, and affine functions to delta type functions.*

Proof. First, fix $z \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Let $\phi \in \mathrm{Cvx}(\mathbb{R}^n)$ so that $T\phi = D_z + c$. We need to show that ϕ is affine, that is,

$$\phi(x) = a \cdot x + b$$

for some $a \in \mathbb{R}^n$, $b \in \mathbb{R}$. Assume by contradiction that ϕ is not affine, then we are able to find two affine functions ϕ_i for $i = 1, 2$ such that

$$\phi_i(x) = a_i \cdot x + b_i,$$

$\phi_i \leq \phi$, and furthermore $a_1 \neq a_2$. Then, $T\phi_i \geq T\phi = D_z + c$. This means that $T\phi_i = D_z + c_i$ for some constants $c_i \in \mathbb{R}$ for $i = 1, 2$, and $c_1 \neq c_2$. Without loss of generality, assume $c_1 > c_2$. Then,

$$T(\max(\phi_1, \phi_2)) = T(\phi_2) = D_z + c_2,$$

which is absurd as $\max(\phi_1, \phi_2) \neq \phi_2$.

Next, let ϕ be an affine function. We need to show that $T\phi$ is a delta type function. Assume again by contradiction that $T\phi$ is not a delta type function. Then, there exist $y, z \in \mathbb{R}^n$ with $y \neq z$ such that $T(\phi)(y), T(\phi)(z) \leq c < +\infty$ for some $c \in \mathbb{R}$. Let $\psi_1, \psi_2 \in \mathrm{Cvx}(\mathbb{R}^n)$ be such that $T\psi_1 = D_y + c$, $T\psi_2 = D_z + c$. Then, as $T\phi \leq T\psi_1, T\psi_2$, we infer that $\psi_1, \psi_2 \leq \phi$. This means that both ψ_1, ψ_2 are affine functions, and their graphs are parallel to that of ϕ. Without loss of generality, assume then that $\psi_1 \leq \psi_2$. This yields that

$$T\psi_1 = D_y + c \geq T\psi_2 = D_z + c,$$

which is a contradiction.

As the conditions on T and T^{-1} are the same, we get right away the desired result. Note moreover that the correspondence between delta type functions and affine functions of T is one-to-one and onto. □

Remark A.16. From the above proof, we get furthermore that, for $z \in \mathbb{R}^n$,

$$T\{D_z + c \,:\, c \in \mathbb{R}\} = \{TD_z + c \,:\, c \in \mathbb{R}\}.$$

Besides, for $a \in \mathbb{R}^n$, denote by $l_a(x) = a \cdot x$ for $x \in \mathbb{R}^n$, then

$$T\{l_a + s \,:\, c \in \mathbb{R}\} = \{Tl_a + s \,:\, c \in \mathbb{R}\}.$$

Next, we define $G_1, G_2 : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ as following. For G_1, denote by

$$G_1(z, c) = (a, s) \qquad \text{provided that } T(D_z + c) = l_a + s.$$

For G_2, set

$$G_2(a, s) = (z, c) \qquad \text{provided that } T(l_a + s) = D_z + c.$$

Lemma A.17. *For $i = 1, 2$, G_i maps an interval in $\mathbb{R}^{n+1}$ to an interval in $\mathbb{R}^{n+1}$.*

Proof. It is enough to show the proof for G_1. Fix $z_1, z_2 \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$. Assume that $G_1(z_1, c_1) = (a_1, s_1)$ and $G_1(z_2, c_2) = (a_2, s_2)$. Let L be the interval joining (z_1, c_1) and (z_2, c_2), that is,

$$L = \{\lambda(z_1, c_1) + (1 - \lambda)(z_2, c_2) \,:\, \lambda \in [0, 1]\}.$$

We aim at showing that

$$T(L) = \left\{ \mu(l_{a_1} + s_1) + (1-\mu)(l_{a_2} + s_2) \,:\, \mu \in [0,1] \right\}.$$

Indeed, we have that

$$T(\check{\min}(D_{z_1} + c_1, D_{z_2} + c_2)) = \max\{l_{a_1} + s_1, l_{a_2} + s_2\}.$$

Here, $\check{\min}(D_{z_1} + c_1, D_{z_2} + c_2)$ is convex, which is linear on L, and $+\infty$ elsewhere. Fix $\lambda \in [0,1]$, and let $(z,c) = \lambda(z_1, c_1) + (1-\lambda)(z_2, c_2)$. It is clear that $D_z + c \geq \check{\min}(D_{z_1} + c_1, D_{z_2} + c_2)$, and so,

$$T(D_z + c) = l_a + s \leq \max\{l_{a_1} + s_1, l_{a_2} + s_2\}.$$

For the affine function $l_a + s$ to lie below $l_{a_1} + s_1$ and $l_{a_2} + s_2$, we need to have that $a \in [a_1, a_2]$, that is,

$$a = \mu a_1 + (1-\mu) a_2$$

for some $\mu \in [0,1]$. If the graph of $l_a + s$ touches the graph of $\max\{l_{a_1} + s_1, l_{a_2} + s_2\}$, then we are done. Otherwise, there is $\delta_1 > 0$ such that

$$T(D_z + c) + \delta_1 = l_a + s + \delta_1 \leq \max\{l_{a_1} + s_1, l_{a_2} + s_2\}.$$

Then, one is able to find $\delta_2 > 0$ so that

$$D_z + c - \delta_2 \geq \check{\min}(D_{z_1} + c_1, D_{z_2} + c_2),$$

which is absurd. The proof is complete. □

Clearly, G_1, G_2 map straight lines to straight lines in $\mathbb{R}^{n+1}$ from this result. The following result is the fundamental fact of affine geometry.

Lemma A.18. *Let $m \geq 2$, and $G : \mathbb{R}^m \to \mathbb{R}^m$ be an injective map which maps all straight lines to straight lines. Then, G is affine, that is,*

$$G(x) = G(0) + Bx \qquad \text{for all } x \in \mathbb{R}^m,$$

for some invertible matrix B of size m.

We will not give a proof of this result. See [8] for some discussions and references there. We are ready to prove Proposition A.14, which in turns gives the conclusion of Theorem A.10 right away.

Proof of Proposition A.14. By the above, G_1, G_2 are affine. We can write

$$G_1(z,c) = B_1(z,c) + V_1 \quad \text{and} \quad G_2(a,s) = B_2(a,s) + V_2,$$

where B_1, B_2 are invertible matrices of size $n+1$ and $V_1, V_2 \in \mathbb{R}^{n+1}$.

By Remark A.16, B_1, B_2 have zeros in all the entries of their last columns except for the $(n+1)$-entry. Let B be the first $n \times n$ block of B_1, and $(v_0, -c_1) \in \mathbb{R}^n \times \mathbb{R}$ be its $(n+1)$-th row. Of course B is invertible itself. Write $V_1 = (v_1, c_0) \in \mathbb{R}^n \times \mathbb{R}$. Then,

$$G_1(z,c) = (Bz, v_0 \cdot z - c_1 c) + (v_1, c_0) = (Bz + v_1, v_0 \cdot z - c_1 c + c_0).$$

This implies that

$$T(D_z + c)(x) = (Bz + v_1) \cdot x + v_0 \cdot z - c_1 c + c_0.$$

Since $c \mapsto T(D_z + c)$ is strictly decreasing, we deduce that $c_1 > 0$. Therefore, (A.5) holds, and our proof is complete.

□

4 Boundary value problems

Let us only focus on static (time-independent) problems here. Throughout this book, we only deal with equations in the whole $\mathbb{R}^n$, or equations in the periodic setting, which can be formulated as equations on $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. We here give some basic and brief discussions on boundary value problems for first-order equations and present some examples.

Let $U \subset \mathbb{R}^n$ be an open, bounded domain with smooth boundary. In a general form, the boundary value problem reads

$$\begin{cases} F(x,u,Du) = 0 & \text{in } U, \\ B(x,u,Du) = 0 & \text{on } \partial U. \end{cases} \tag{A.6}$$

Here, $F : U \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ and $B : \partial U \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ are given continuous functions. The unknown in (A.6) is u. Of course, the second equality in (A.6) represents a general boundary condition. We give first a general definition of viscosity solutions to (A.6).

Definition A.19 (viscosity solutions of (A.6))**.** *Let $u \in C(\overline{U})$.*

(a) We say that u is a viscosity subsolution to (A.6) *if for any test function $\varphi \in C^1(\overline{U})$ such that $u - \varphi$ has a strict maximum $x_0 \in \overline{U}$, then*

$$F(x_0, u(x_0), D\varphi(x_0)) \le 0 \qquad \textit{if } x_0 \in U,$$

or

$$\min\{F(x_0,u(x_0),D\varphi(x_0)), B(x_0,u(x_0),D\varphi(x_0))\} \le 0 \qquad \textit{if } x_0 \in \partial U.$$

(b) We say that u is a viscosity supersolution to (A.6) *if for any test function $\varphi \in C^1(\overline{U})$ such that $u - \varphi$ has a strict minimum $x_0 \in \overline{U}$, then*

$$F(x_0, u(x_0), D\varphi(x_0)) \ge 0 \qquad \textit{if } x_0 \in U,$$

or

$$\max\{F(x_0,u(x_0),D\varphi(x_0)), B(x_0,u(x_0),D\varphi(x_0))\} \ge 0 \qquad \textit{if } x_0 \in \partial U.$$

(c) We say that u is a viscosity solution to (A.6) *if it is both a viscosity subsolution and a viscosity supersolution to* (A.6).

It is clear from the definition above that boundary conditions in the viscosity sense do not hold in the classical way. This definition arises naturally from the usual vanishing viscosity process, but we omit this discussion here. Based on the definition, u is a viscosity solution to (A.6) if it satisfies in the viscosity sense

$$\begin{cases} F(x,u,Du) = 0 & \text{in } U, \\ \min\{F(x,u,Du), B(x,u,Du)\} \le 0 & \text{on } \partial U, \\ \max\{F(x,u,Du), B(x,u,Du)\} \ge 0 & \text{on } \partial U. \end{cases} \tag{A.7}$$

This equation makes clearer the meaning of being a viscosity solution in the boundary value problem. We now discuss further some specific situations.

4.1 State-constraint problems

A state-constraint boundary problem has the following form

$$\begin{cases} F(x,u,Du) \le 0 & \text{in } U, \\ F(x,u,Du) \ge 0 & \text{on } \overline{U}. \end{cases} \tag{A.8}$$

This equation can be written in an equivalent way as

$$\begin{cases} F(x,u,Du) = 0 & \text{in } U, \\ F(x,u,Du) \ge 0 & \text{on } \partial U. \end{cases}$$

One can see that this equation is of a special form of (A.7) but it is clearly simpler. As a matter of fact, we only require the supersolution property $F(x,u,Du) \ge 0$ on the boundary of U. We give here an example for problems of this type.

Example A.1. *Assume that* $n = 1$, $U = (-1,1)$*, and*

$$F(x,z,p) = z + |p-1| - 1 \qquad \text{for all } (x,z,p) \in [-1,1] \times \mathbb{R} \times \mathbb{R}.$$

The corresponding state-constraint problem becomes

$$\begin{cases} u(x) + |u'(x)-1| - 1 \le 0 & \text{in } (-1,1), \\ u(x) + |u'(x)-1| - 1 \ge 0 & \text{on } [-1,1]. \end{cases} \tag{A.9}$$

Firstly, it is clear that $v \equiv 0$ *on* $[-1,1]$ *solves the equation*

$$v(x) + |v'(x)-1| - 1 = 0 \qquad \text{in } (-1,1).$$

We now argue that v however is not a solution to (A.9) *as it does not satisfy the state-constraint boundary condition. Indeed, let* $\varphi(x) = x - 1$ *for all* $x \in [-1,1]$*. Then* $v - \varphi$ *has a strict minimum at* 1 *on* $[-1,1]$*, but*

$$v(1) + |\varphi'(1) - 1| - 1 = -1 < 0.$$

This shows that the state-constraint boundary condition plays an essential role in the problem. Secondly, denote by

$$u(x) = e^{x-1} \qquad \text{for } x \in [-1,1].$$

We claim that u is a solution to (A.9)*. It is clear that u satisfies the equation for* $x \in (-1,1)$ *in the classical sense, and we only need to verify the boundary condition. Let us only check the supersolution condition at* $x_0 = 1$ *as the supersolution condition at* $x_0 = -1$ *can be checked in a similar manner. Let* $\varphi \in C^1([-1,1])$ *be a test function such that* $u - \varphi$ *has a strict minimum at* $x_0 = 1$*, and* $u(1) = \varphi(1)$*. Then,* $\varphi'(1) \ge u'(1) = 1$*, which means that*

$$u(1) + |\varphi'(1) - 1| - 1 = 1 + (\varphi'(1) - 1) - 1 = \varphi'(1) - 1 \ge 0.$$

It turns out that this u is the unique viscosity solution to (A.9).

We do not discuss further about well-posedness of solutions to (A.8) here. Let us give a representation formula of the solution in the convex setting.

Theorem A.20. *Assume that*

$$F(x,z,p) = z + H(x,p) \qquad \textit{for all } (x,z,p) \in \overline{U} \times \mathbb{R} \times \mathbb{R}^n,$$

where $H \in C^1(\overline{U} \times \mathbb{R}^n)$ *satisfies that* $p \mapsto H(x,p)$ *is convex for* $x \in \overline{U}$*, and*

$$\lim_{|p|\to\infty} \left(\min_{x\in\overline{U}} \frac{H(x,p)}{|p|} \right) = +\infty.$$

Let L be the corresponding Legendre transform of H. Then, the unique viscosity solution u to (A.8) *has the following representation formula*

$$u(x) = \inf \left\{ \int_0^\infty e^{-s} L\big(\gamma(s), -\gamma'(s)\big)\, ds \,:\, \gamma(0) = x, \gamma([0,\infty)) \subset \overline{U}, \gamma \in \mathrm{AC}([0,\infty), \mathbb{R}^n) \right\}.$$

A key feature of the above representation formula is that all admissible paths are running on $\overline{U}$, which explains intuitively the keyword "state-constraint". For further discussions on state-constraint problems, see Soner [127, 128], Capuzzo-Dolcetta, Lions [28].

4.2 Dirichlet problems

A Dirichlet boundary problem has the following form

$$\begin{cases} F(x,u,Du) = 0 & \text{in } U, \\ u = g & \text{on } \partial U. \end{cases} \tag{A.10}$$

Here, $g \in C(\partial U)$ is given. This is of course a special case of (A.6) where $B(x,z,p) = z - g(x)$. As discussed earlier, u is a viscosity solution to (A.10) if it satisfies in the viscosity sense

$$\begin{cases} F(x,u,Du) = 0 & \text{in } U, \\ \min\{F(x,u,Du), u - g(x)\} \le 0 & \text{on } \partial U, \\ \max\{F(x,u,Du), u - g(x)\} \ge 0 & \text{on } \partial U. \end{cases}$$

Let us give a simple example to show that this Dirichlet boundary condition does not hold in the classical sense.

Example A.2. *Assume that* $n = 1$, $U = (0,1)$*, and*

$$F(x,z,p) = p - 1 \qquad \textit{for all } (x,z,p) \in [0,1] \times \mathbb{R} \times \mathbb{R}.$$

Assume further that $g \equiv 0$ *on* $[0,1]$*. The corresponding Dirichlet problem is*

$$\begin{cases} u'(x) - 1 = 0 & \textit{in } (0,1), \\ u(0) = u(1) = 0. \end{cases} \tag{A.11}$$

On the first hand, it is quite straightforward to see that (A.11) *does not admit any classical solution with* $u(0) = u(1) = 0$*. On the other hand, we claim that*

$$u(x) = x \qquad \textit{for } x \in [0,1]$$

is a viscosity solution to (A.11). *Of course, u satisfies the equation for $x \in (0,1)$ in the classical sense, and we only need to verify the Dirichlet boundary condition. The Dirichlet boundary condition at $x_0 = 0$ satisfies classically, so there is nothing to check. At $x_0 = 1$, the supersolution test holds automatically as $u(1) = 1 > 0$. To check the subsolution property, take $\varphi \in C^1([0,1])$ such that $u - \varphi$ has a strict maximum at $x_0 = 1$, and $u(1) = \varphi(1)$. Then, $\varphi'(1) \le u'(1) = 1$, and therefore,*

$$\varphi'(1) - 1 \le 0.$$

We do not discuss further about well-posedness of solutions to (A.10) here. It is important pointing out that, in general, it is still an open problem to determine in which parts of the boundary of U that one has $u = g$ in the classical sense. This of course has a strong relation to the method of characteristics. Let us give a representation formula of the solution in the convex setting.

Theorem A.21. *Assume that*

$$F(x,z,p) = z + H(x,p) \qquad \text{for all } (x,z,p) \in \overline{U} \times \mathbb{R} \times \mathbb{R}^n,$$

where $H \in C^1(\overline{U} \times \mathbb{R}^n)$ satisfies that $p \mapsto H(x,p)$ is convex for $x \in \overline{U}$, and

$$\lim_{|p| \to \infty} \left(\min_{x \in \overline{U}} \frac{H(x,p)}{|p|} \right) = +\infty.$$

Let L be the corresponding Legendre transform of H. Then, the unique viscosity solution u to (A.10) *has the following representation formula*

$$u(x) = \inf \left\{ \int_0^{\tau_x} e^{-s} L\big(\gamma(s), -\gamma'(s)\big)\, ds + e^{-\tau_x} g(\gamma(\tau_x)) \,:\, \gamma(0) = x, \gamma \in \mathrm{AC}([0,\infty), \overline{U}) \right\}.$$

Here,

$$\tau_x = \tau_x(\gamma) = \min\{t > 0 \,:\, \gamma(t) \in \partial U\},$$

which is the first exit time from U of the path γ. Of course, in case $\{t > 0 : \gamma(t) \in \partial U\}$ is empty, $\tau_x = \infty$ and $e^{-\tau_x} g(\gamma(\tau_x)) = 0$.

4.3 Neumann problems

A Neumann boundary problem has the following form

$$\begin{cases} F(x,u,Du) = 0 & \text{in } U, \\ Du(x) \cdot \gamma(x) = g(x) & \text{on } \partial U. \end{cases} \tag{A.12}$$

Here, $g \in C(\partial U)$ is given, and $\gamma \in C(\partial U, \mathbb{R}^n)$ is a given vector field such that

$$n(x) \cdot \gamma(x) > 0 \qquad \text{for } x \in \partial U,$$

where $n(x)$ denotes the outer unit normal vector to U at x. This is again a special case of (A.6) where $B(x,z,p) = p \cdot \gamma(x) - g(x)$. The boundary condition in (A.12) is called the inhomogeneous linear Neumann boundary condition. We note that for $\phi \in C^1(\overline{U})$, we can write

$$D\phi(x) \cdot \gamma(x) = \frac{\partial \phi}{\partial \gamma}(x) = \lim_{s \to 0} \frac{\phi(x + s\gamma(x)) - \phi(x)}{s} \qquad \text{for } x \in \partial U.$$

Let us also give a simple example demonstrating that the Neumann boundary condition does not satisfy in the classical sense.

Example A.3. *Assume that $n = 1$, $U = (0,1)$, and*

$$F(x,z,p) = z + p - x - 1 \qquad \text{for all } (x,z,p) \in [0,1] \times \mathbb{R} \times \mathbb{R}.$$

Assume further that $g \equiv 0$ on $[0,1]$. The Neumann problem of interests is

$$\begin{cases} u(x) + u'(x) - x - 1 = 0 & \text{in } (0,1), \\ u'(0) = u'(1) = 0. \end{cases} \tag{A.13}$$

We claim that

$$u(x) = x + e^{-x} \qquad \text{for all } x \in [0,1]$$

is a viscosity solution to (A.13). *By direct computations, we see that $u \in C^1([0,1])$ satisfies the equation in $(0,1)$ classically. Besides, $u'(0) = 0$, which means that the Neumann boundary condition holds classically at* 0.
However, $u'(1) = 1 - e^{-1} > 0$, which means that the subsolution property at $x_0 = 1$ does not hold in the strong sense. Take $\varphi \in C^1([0,1])$ such that $u - \varphi$ has a strict maximum at $x_0 = 1$, and $u(1) = \varphi(1)$. Then, $\varphi'(1) \le u'(1) = 1 - e^{-1}$, which means that

$$u(1) + \varphi'(1) - 2 \le u(1) + u'(1) - 2 = 0,$$

and hence, the subsolution test holds true at $x_0 = 1$.

There is also an analog of Theorems A.20 and A.21 for (A.12), but it is slightly more complicated, and we omit it here.

5 Sup-convolutions

Sup-convolutions and inf-convolutions are basic and very important tools to regularize viscosity solutions. These approximations were first realized by Jensen [91]. We give here some properties of sup-convolutions.

Definition A.22. *Let $u : \mathbb{R}^n \to \mathbb{R}$ be a given bounded function, and $\varepsilon > 0$ be a parameter. The sup-convolution $u^\varepsilon : \mathbb{R}^n \to \mathbb{R}$ and inf-convolution $u_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ are defined as*

$$u^\varepsilon(x) = \sup_{y \in \mathbb{R}^n} \left(u(y) - \frac{|y-x|^2}{2\varepsilon} \right) \qquad \text{for } x \in \mathbb{R}^n,$$

and

$$u_\varepsilon(x) = \inf_{y \in \mathbb{R}^n} \left(u(y) + \frac{|y-x|^2}{2\varepsilon} \right) \qquad \text{for } x \in \mathbb{R}^n.$$

It is worth noting that, for $x \in \mathbb{R}^n$,

$$u_\varepsilon(x) = -\sup_{y \in \mathbb{R}^n} \left(-u(y) - \frac{|y-x|^2}{2\varepsilon} \right) = -(-u)^\varepsilon(x).$$

This relation allows us to interpret properties of sup-convolutions into the corresponding ones of inf-convolutions automatically and vice versa. Here is a main result on the properties of sup-convolution u^ε.

Proposition A.23. *Assume that u is upper semicontinuous in $\mathbb{R}^n$. Assume further that there exists $M > 0$ such that $\|u\|_{L^\infty(\mathbb{R}^n)} \le M$. Then, the following properties hold.*

(i) We have

$$-M \le u(x) \le u^\varepsilon(x) \le M \qquad \text{for all } x \in \mathbb{R}^n.$$

(ii) The function $x \mapsto u^\varepsilon(x) + \frac{|x|^2}{2\varepsilon}$ is convex in $\mathbb{R}^n$.

(iii) If $p \in D^+u^\varepsilon(x)$ for some $x \in \mathbb{R}^n$, then

$$|p| \le 2\sqrt{\frac{M}{\varepsilon}} \qquad \text{and} \qquad p \in D^+u(x + \varepsilon p).$$

Proof. Claim (i) is quite clear as

$$u^\varepsilon(x) = \sup_{y\in\mathbb{R}^n}\left(u(y) - \frac{|y-x|^2}{2\varepsilon}\right) \le \sup_{y\in\mathbb{R}^n} u(y) \le M,$$

and if we choose $y = x$ in the above formula,

$$u^\varepsilon(x) \ge u(x) \ge -M.$$

To prove claim (ii), we rewrite the formula of $u^\varepsilon(x)$ as

$$u^\varepsilon(x) + \frac{|x|^2}{2\varepsilon} = \sup_{y\in\mathbb{R}^n}\left(u(y) - \frac{|y|^2}{2\varepsilon} + \frac{y\cdot x}{\varepsilon}\right).$$

The right hand side above is the supremum of a family of affine functions in x, which is surely a convex function in x.

Let us now prove assertion (iii). Assume that $p \in D^+u^\varepsilon(\bar{x})$ for some $\bar{x} \in \mathbb{R}^n$. By Theorem 1.4, there exists a function $\phi \in C^1(\mathbb{R}^n)$ such that $D\phi(\bar{x}) = p$, and $u^\varepsilon - \phi$ has a global strict maximum at $\bar{x}$. Besides, as u is upper semicontinuous, we can find $\bar{y} \in \mathbb{R}^n$ such that

$$u^\varepsilon(\bar{x}) = \sup_{y\in\mathbb{R}^n}\left(u(y) - \frac{|y-\bar{x}|^2}{2\varepsilon}\right) = u(\bar{y}) - \frac{|\bar{y}-\bar{x}|^2}{2\varepsilon}.$$

In particular,

$$\frac{|\bar{y}-\bar{x}|^2}{2\varepsilon} = u(\bar{y}) - u^\varepsilon(\bar{x}) \le 2M. \tag{A.14}$$

Consider an auxiliary function $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as

$$(x, y) \mapsto \Phi(x, y) = u(y) - \frac{|y-x|^2}{2\varepsilon} - \phi(x).$$

It is clear that Φ has a global maximum at $(\bar{x}, \bar{y})$. In particular, $x \mapsto \Phi(x, \bar{y})$ has a global maximum at $\bar{x}$, and $y \mapsto \Phi(\bar{x}, y)$ has a global maximum at $\bar{y}$. These allow us to imply that

$$p = D\phi(\bar{x}) = \frac{\bar{y}-\bar{x}}{\varepsilon} \qquad \text{and} \qquad \frac{\bar{y}-\bar{x}}{\varepsilon} \in D^+u(\bar{y}).$$

Thus, $\bar{y} = \bar{x} + \varepsilon p$, and in light of (A.14),

$$|p| \le 2\sqrt{\frac{M}{\varepsilon}} \qquad \text{and} \qquad p \in D^+u(\bar{x} + \varepsilon p).$$

□

In this book, we have not used sup-convolutions and inf-convolutions to regularize and analyze viscosity solutions (more generally, subsolutions and supersolutions) as we typically deal with nice enough solutions already. Let us give here a prototypical example of their usage.

Example A.4. *Let $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ be a given Hamiltonian. Let $u : \mathbb{R}^n \to \mathbb{R}$ be a bounded and upper semicontinuous function. Pick $M > 0$ such that $\|u\|_{L^\infty(\mathbb{R}^n)} \le M$. Assume that u is a viscosity subsolution to*

$$H(x, Du(x)) \le 0 \qquad \text{in } \mathbb{R}^n. \tag{A.15}$$

For each $\varepsilon > 0$, let u^ε be the sup-convolution of u. Let $\delta = 2\sqrt{M\varepsilon}$. By Proposition A.23, we see that u^ε is a viscosity subsolution to both

$$H(x + \varepsilon Du^\varepsilon(x), Du^\varepsilon(x)) \le 0 \qquad \text{in } \mathbb{R}^n, \tag{A.16}$$

and

$$|Du^\varepsilon(x)| \le \frac{\delta}{\varepsilon} \qquad \text{in } \mathbb{R}^n.$$

Thus, u^ε is Lipschitz in $\mathbb{R}^n$ with Lipschitz constant δ/ε. Define $\widetilde{H} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as

$$\widetilde{H}(x, p) = \min_{|z| \le \delta} H(x + z, p) \qquad \text{for all } (x, p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Then, thanks to (A.16)*, we see that u^ε is a viscosity subsolution to*

$$\widetilde{H}(x, Du^\varepsilon(x)) \le 0 \qquad \text{in } \mathbb{R}^n.$$

We now argue that the inf-convolution u_ε is actually quite a familiar object. Indeed, consider the following Cauchy problem

$$\begin{cases} v_t + \frac{|Dv|^2}{2} = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ v(x, 0) = u(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{A.17}$$

Let us not worry much about the regularity of initial data u here. By the Hopf-Lax formula, we have, for $(x, t) \in \mathbb{R}^n \times (0, \infty)$,

$$v(x, t) = \inf_{y \in \mathbb{R}^n} \left(u(y) + \frac{|y - x|^2}{2t} \right).$$

Consequently, for $t = \varepsilon$, we see that

$$u_\varepsilon(x) = v(x, \varepsilon) \qquad \text{for all } x \in \mathbb{R}^n.$$

This shows that Cauchy problem (A.17) has a natural regularizing effect. In fact, this regularizing effect holds for solutions of Cauchy problems with general convex, superlinear Hamiltonians.

Remark A.24. It is important pointing out that we can do inf-sup convolutions to regularize more a given function. Indeed, for $\varepsilon, \alpha > 0$, it is quite clear that $w = (u^{\varepsilon+\alpha})_\varepsilon$ is both semiconvex and semiconcave, and hence, $w \in C^{1,1}(\mathbb{R}^n)$. We refer to Lasry, Lions [99] for some applications on this.

6 Notations

We list here notations that are used in the book.

6.1 Notation for sets and spaces

- $n \in \mathbb{N}$ is often used to denote the dimensions.
- $\mathbb{R}^n = n$-dimensional real Euclidean space; $\mathbb{R} = \mathbb{R}^1$.
- e_i is the i-th vector in the canonical basis of $\mathbb{R}^n$ for $1 \le i \le n$, that is,

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0),$$

  where 1 occurs in the i-th position.
- A typical point in $\mathbb{R}^n$ is often denoted by $x = (x_1, \ldots, x_n)$. Depending on different situations, we might regard x as a row vector or a column vector.
- For $x, y \in \mathbb{R}^n$ with $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$, write

$$x \cdot y = \sum_{i=1}^{n} x_i y_i \quad \text{and} \quad |x| = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

- A typical point in $\mathbb{R}^n \times [0, \infty)$ is often denoted by $(x, t) = (x_1, \ldots, x_n, t)$, where t often stands for the time variable.
- For a given real number $s \in \mathbb{R}$, denote by $[s]$ its integer part.
- $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ is the usual n-dimensional flat torus. When there is no confusion, we identify $\mathbb{T}^n$ with the unit cell $Y = [0, 1]^n$ with periodic boundary condition on Y.
- For an open set $U \subset \mathbb{R}^n$, we write ∂U to denote its boundary, and $\overline{U} = U \cup \partial U$ to denote its closure.
- For U, V open sets in $\mathbb{R}^n$, we write

$$U \subset\subset V$$

  if $U \subset \overline{U} \subset V$, and $\overline{U}$ is compact, and say that U is compactly supported in V.
- For $x \in \mathbb{R}^n$ and $r > 0$, we denote by $B(x, r)$ the open ball in $\mathbb{R}^n$ with center x, radius r, that is,

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

  Denote by $\overline{B(x, r)}$ or $\overline{B}(x, r)$ the closed ball with center x, radius r, that is,

$$\overline{B(x, r)} = \overline{B}(x, r) = \{y \in \mathbb{R}^n : |y - x| \le r\}.$$

  We also write $B(x, r)$, $\overline{B}(x, r)$ as $B_r(x)$, $\overline{B}_r(x)$, respectively. When $x = 0$, we simply write $B_r = B_r(0)$, $\overline{B}_r = \overline{B}_r(0)$.

6.2 Notation for functions

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

- $Du(x) = \nabla u(x) = \left(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_n}(x)\right)$.
- $D^2u(x) = \text{Hessian of } u \text{ at } x = \begin{pmatrix} \frac{\partial^2 u}{\partial x_1^2}(x) & \frac{\partial^2 u}{\partial x_1 \partial x_2}(x) & \ldots & \frac{\partial^2 u}{\partial x_1 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 u}{\partial x_n \partial x_1}(x) & \frac{\partial^2 u}{\partial x_n \partial x_2}(x) & \ldots & \frac{\partial^2 u}{\partial x_n^2}(x) \end{pmatrix}$.
- The Laplacian $\Delta u(x) = \operatorname{tr}(D^2u(x)) = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}(x)$ is the trace of $D^2u(x)$.

In this book, we use the notion $Du(x)$ instead of $\nabla u(x)$. We usually write u_{x_i} for $\frac{\partial u}{\partial x_i}$.

When u is not smooth, we have the following definition for subdifferential and superdifferential of u at x.

- The subdifferential of u at x is denoted by $D^-u(x)$, where

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

- The superdifferential of u at x is denoted by $D^+u(x)$, where

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

If u is differentiable at x then

$$D^-u(x) = D^+u(x) = \{Du(x)\}.$$

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 $\Delta u(x,t) = \Delta_x u(x,t)$.

Besides, we use the following for a given function $u : \mathbb{R}^n \to \mathbb{R}$.

- Set $u^+ = \max\{u, 0\}$, and $u^- = -\min\{u, 0\}$. Surely, $u = u^+ - u^-$, and $|u| = u^+ + u^-$.
- If u is compactly supported, then the support of u is denoted by $\operatorname{spt}(u)$.
- If u is $\mathbb{Z}^n$-periodic, then we can think of u as a function from $\mathbb{T}^n$ to $\mathbb{R}$ as well, and vice versa. In the book, we switch freely between the two interpretations.

For a smooth path $\gamma : \mathbb{R} \to \mathbb{R}^n$ and $t \in \mathbb{R}$, we write

$$\gamma'(t) = \frac{d}{dt}\gamma(t)$$

In many occasions, we use a modulus of continuity ω. By this, we mean $\omega : [0,\infty) \to [0,\infty)$ is a continuous function such that $\omega(0) = 0 = \lim_{r\to 0} w(r)$.

The following convolution trick is used quite often throughout the book. Take η to be the standard mollifier, that is,

$$\eta \in C_c^\infty(\mathbb{R}^n, [0,\infty)), \qquad \mathrm{supp}(\eta) \subset B(0,1), \qquad \int_{\mathbb{R}^n} \eta(x)\,dx = 1.$$

For $\varepsilon > 0$, denote by $\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(\frac{x}{\varepsilon}\right)$ for all $x \in \mathbb{R}^n$. Let $u : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. Set

$$u^\varepsilon(x) = (\eta_\varepsilon \star u)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)u(y)\,dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x-y)u(y)\,dy \quad \text{ for } x \in \mathbb{R}^n.$$

Then $u^\varepsilon \in C^\infty(\mathbb{R}^n)$, and $u^\varepsilon \to u$ locally uniformly as $\varepsilon \to 0$. If needed, one can assume further that η is symmetric or radially symmetric.

6.3 Notation for function spaces

- $C(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is continuous}\}$.
- $B(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is bounded}\}$.
- $\mathrm{BC}(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is bounded, and continuous}\}$.
- $\mathrm{BUC}(\mathbb{R}^n) = \{u \in C(\mathbb{R}^n) : u \text{ is bounded, and uniformly continuous}\}$.
- For $k \in \mathbb{N}$, $C^k(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is } k\text{-times continuously differentiable}\}$.
- $C^\infty(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is infinitely differentiable}\}$. For $u \in C^\infty(\mathbb{R}^n)$, we say that u is smooth.
- $C_c^k(\mathbb{R}^n)$, $C_c^\infty(\mathbb{R}^n)$ denote the space of functions in $C^k(\mathbb{R}^n)$, $C^\infty(\mathbb{R}^n)$ that have compact supports, respectively.
- $\mathrm{Lip}(\mathbb{R}^n) = \{u \in C(\mathbb{R}^n) : \exists\, C > 0 \text{ so that } |u(x)-u(y)| \le C|x-y| \text{ for all } x, y \in \mathbb{R}^n\}$.

  We write
  $$\mathrm{Lip}[u] = \sup_{\substack{x,y\in\mathbb{R}^n\\ x\neq y}} \frac{|u(x)-u(y)|}{|x-y|},$$
  and say that $\mathrm{Lip}[u]$ is the Lipschitz constant of u.
- For $\alpha \in (0,1]$, we say that $u \in C(\mathbb{R}^n)$ is Hölder continuous with exponent α if there exists $C > 0$ such that
  $$|u(x)-u(y)| \le C|x-y|^\alpha \qquad \text{for all } x, y \in \mathbb{R}^n.$$
  In this case, the α-th Hölder seminorm of u is
  $$[u]_{C^{0,\alpha}(\mathbb{R}^n)} = \sup_{\substack{x,y\in\mathbb{R}^n\\ x\neq y}} \frac{|u(x)-u(y)|}{|x-y|^\alpha}.$$

If we have in addition that u is bounded, then we define the α-th Hölder norm of u to be

$$\|u\|_{C^{0,\alpha}(\mathbb{R}^n)} = \|u\|_{C(\mathbb{R}^n)} + [u]_{C^{0,\alpha}(\mathbb{R}^n)}.$$

Then, the Hölder space $C^{0,\alpha}(\mathbb{R}^n)$ is defined as

$$C^{0,\alpha}(\mathbb{R}^n) = \left\{u \in C(\mathbb{R}^n) : \|u\|_{C^{0,\alpha}(\mathbb{R}^n)} < +\infty\right\}.$$

- $L^\infty(\mathbb{R}^n) = \left\{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is Lebesgue measurable and } \|u\|_{L^\infty(\mathbb{R}^n)} < +\infty\right\}$, where

$$\|u\|_{L^\infty(\mathbb{R}^n)} = \operatorname*{ess\,sup}_{\mathbb{R}^n} |u|.$$

- It is clear that $C^{0,1}(\mathbb{R}^n) = L^\infty(\mathbb{R}^n) \cap \mathrm{Lip}\,(\mathbb{R}^n)$, and $\mathrm{Lip}\,[u] = [u]_{C^{0,1}(\mathbb{R}^n)}$.
- In a same way, one can define $C^{k,\alpha}(\mathbb{R}^n)$ for $k \in \mathbb{N}$ and $\alpha \in (0,1]$.
- $\mathrm{USC}\,(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is upper semicontinuous}\}$.
- $\mathrm{LSC}\,(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} : u \text{ is lower semicontinuous}\}$.
- For a function $u : \mathbb{R}^n \to \mathbb{R}$ that is bounded, we denote by

$$u^*(x) = \limsup_{y \to x} u(y) \qquad \text{for all } x \in \mathbb{R}^n,$$

  and

$$u_*(x) = \liminf_{y \to x} u(y) \qquad \text{for all } x \in \mathbb{R}^n.$$

  It is clear that $u^* \in \mathrm{USC}\,(\mathbb{R}^n)$, $u_* \in \mathrm{LSC}\,(\mathbb{R}^n)$. We say that u^*, u_* are the upper semicontinuous envelope, and the lower semicontinuous envelope of u, respectively. One has that u is continuous in $\mathbb{R}^n$ if and only if $u^* = u_*$.
- Let $U \subset \mathbb{R}^n$ be a given open set. All above function spaces can be defined in U and $\overline{U}$ in place of $\mathbb{R}^n$ in a similar way.
- For given $T > 0$, $\mathrm{AC}\,([0,T],\mathbb{R}^n)$ denotes the space of all absolutely continuous curves from $[0,T]$ to $\mathbb{R}^n$.
- $\mathrm{Cvx}\,(\mathbb{R}^n)$ denotes the class of lower semi-continuous convex functions $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$.

6.4 Notation for estimates

- The constants in the estimates are often denoted by C (and C_1, C_2, etc.), which might change from line to line in a given computation. This makes our presentation clearer without keeping track with various factors in each step. Of course, we specify clearly the dependence of these constants on specific parameters.
- (Big-oh notation) For two given functions f, h, we write $f = O(h)$ as $x \to y$ if there exists $C > 0$ such that

$$|f(x)| \le C|h(x)| \qquad \text{for all } x \text{ sufficiently close to } y.$$

- (Little-oh notation) For two given functions f, h, we write $f = o(h)$ as $x \to y$ if

$$\lim_{x \to y} \frac{|f(x)|}{|h(x)|} = 0.$$

  In particular, when $h \equiv 1$, we have the notions of $O(1)$ and $o(1)$, respectively.

Solutions to some exercises

Solutions to some exercises in the book were provided to me by Son Tu.

Exercise 1. *Consider the eikonal problem in one dimension*

$$\begin{cases} |u'(x)| &= 1 \qquad in\ (-1,1), \\ u(1)=u(-1) &= 0. \end{cases} \tag{A.18}$$

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

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

Proof of Exercise 1.

(a) Assume that there exists a C^1 solution $u:[-1,1]\to\mathbb{R}$ satisfies (A.18), then $x\mapsto u'(x)$ must be continuous. By the mean value theorem, there exists some $c\in(-1,1)$ such that $0=u(1)-u(-1)=2u'(c)$, and thus, $u'(c)=0$. This is a contradiction since one should have $|u'(c)|=1$. Therefore, (A.18) has no C^1 solution.

(b) Generally, a continuous a.e. solution with finitely many gradient jumps must have the form as in figure A.2.

It is clear that a graph of such a solution is a combination of line segments with slope 1 or -1. As we can see, $u'(x)$ exists a.e., so we only need to check if they are subsolution at points where $u'(x)$ is not well defined (at the vertices).

- If x is the vertex of the shape $\bigvee$, then there is no C^1 function φ that can touch u from above at x (in the sense that $u-\varphi$ has a strict max at x). That means the condition $|u'(x)|\le 1$ in the viscosity sense holds true automatically.
- If x is the vertex of the shape $\bigwedge$, then any C^1 function φ that can touch u from above at x (in the sense that $u-\varphi$ has a strict max at x) must have $\varphi'(x)\in[-1,1]$. That means the condition $|u'(x)|\le 1$ in the viscosity sense holds true.

□

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

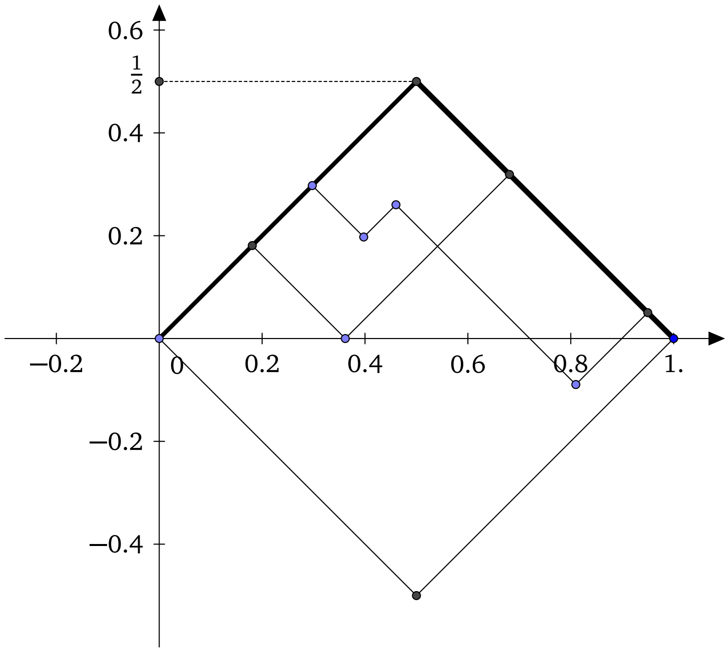


Figure A.2: Typical continuous a.e. solutions with finitely many gradient jumps.

Proof of Exercise 3 . Recall the definition of viscosity solution for first-order equations:

$$\begin{cases} u_t(x,t) + H(Du(x,t)) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{A.19}$$

In Definition 1.1, let us modify a little bit. A function is a

- viscosity solution with C^1 test functions if it satisfies Definition 1.1 with C^1 test functions,
- viscosity solution with C^2 test functions if it satisfies Definition 1.1 with C^2 test functions.

It is clear that a viscosity solution with C^1 test functions is also a viscosity solution with C^2 test functions. For the converse, assume that u is a viscosity subsolution of (A.19) with C^2 test functions, then $u(x,0) \le u_0(x)$. Let $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ such that $u(x_0,t_0) = \varphi(x_0,t_0)$ and $u - \varphi$ has a strict max at $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$, we need to prove that

$$\varphi_t(x_0,t_0) + H(D\varphi(x_0,t_0)) \le 0. \tag{A.20}$$

For simplicity, let us extend φ to $\mathbb{R}^n \times \mathbb{R}$ so that it has compact support. Let $\{\eta_\varepsilon\}_{\varepsilon>0} \subset C_c^\infty(\mathbb{R}^{n+1})$ be the standard mollifiers, that is, $\eta_\varepsilon(x) = \varepsilon^{-(n+1)}\eta\left(\varepsilon^{-1}x\right)$ where $\eta \in C_c^\infty(B(0,1))$ with

$$0 \le \eta \le 1, \qquad \operatorname{supp} \eta \subset B(0,1), \qquad \text{and} \qquad \int_{\mathbb{R}^{n+1}} \eta(x)\,dx = 1.$$

For $\varepsilon > 0$ we let $\varphi^\varepsilon(x,t) = \big(\eta_\varepsilon \star \varphi\big)(x,t)$, then it is clear that $\varphi^\varepsilon \to \varphi$, $\varphi^\varepsilon_t \to \varphi_t$, and $D\varphi^\varepsilon \to D\varphi$ locally uniformly on $\mathbb{R}^n \times (0,T)$. Also by stability of viscosity solutions (see

Lemma 1.8), we can choose a decreasing subsequence $\{\varepsilon_i\} \searrow 0$ such that $\left(x_{\varepsilon_i}, t_{\varepsilon_i}\right) \to (x_0, t_0)$ as $i \to \infty$ and $u - \varphi^{\varepsilon_i}$ has a local max at $\left(x_{\varepsilon_i}, t_{\varepsilon_i}\right)$, thus

$$\varphi_t^{\varepsilon_i}\left(x_{\varepsilon_i}, t_{\varepsilon_i}\right) + H\left(D\varphi^{\varepsilon_i}\left(x_{\varepsilon_i}, t_{\varepsilon_i}\right)\right) \leq 0.$$

Let $\varepsilon_i \to 0$ and using the facts that $\left(x_{\varepsilon_i}, t_{\varepsilon_i}\right) \to (x_0, t_0)$, H is continuous, and $\varphi_t^\varepsilon \to \varphi_t$, $D\varphi^\varepsilon \to D\varphi$ locally uniformly, we obtain (A.20). Thus u is a viscosity subsolution with C^1 test functions. The argument for supersolution test is similar. □

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 a subsequence.*

Proof of Exercise 10. Without loss of generality, we assume that $u(x_0, T) = \varphi(x_0, T)$.

Fix $0 < r < T$ and let $\Omega_r = \overline{B_r(x_0)} \times [T - r, T)$, we have $u - \varphi < 0$ for all $(x, t) \in \Omega_r$. It is clear that

$$u(x, t) - \varphi_\varepsilon(x, t) \leq -\frac{\varepsilon}{T - t} < 0. \tag{A.21}$$

We claim that $u - \varphi_\varepsilon$ has a local max over $\overline{B_r(x_0)} \times [T - r, T)$ at $(x_\varepsilon, t_\varepsilon)$. Indeed, let $\zeta = \sup_{\overline{B_r(x_0)} \times [T-r,T)}(u - \varphi_\varepsilon)$ and $(x_j, t_j) \in \Omega_r$ such that $u(x_j, t_j) - \varphi_\varepsilon(x_j, t_j) \to \alpha$. By compactness we have $(x_j, t_j) \to (\overline{x}, \overline{t}) \in \overline{\Omega}_r$ up to subsequence. If $\overline{t} = T$ then from (A.21) we have $\alpha = -\infty$, which is a contradiction.

We show that for $r > 0$, there exists $\varepsilon = \varepsilon(r) > 0$ small enough so that $(x_\varepsilon, t_\varepsilon) \in \operatorname{int}(\Omega_r)$, which implies that $u - \varphi_\varepsilon$ has a local max at $(x_\varepsilon, t_\varepsilon)$. Since $t_\varepsilon < T$ for all $\varepsilon > 0$, it is suffices to consider (we do not have to worry about the top of the cylinder)

$$\partial \Omega_r = \underbrace{\left(\overline{B(x_0, r)} \times \{T - r\}\right)}_{\text{bottom of the cylinder}} \cup \underbrace{\left(\partial B(x_0, r) \times (T - r, T)\right)}_{\text{the surface between the bottom and the top}} .$$

Let $\alpha = \sup_{\partial\Omega_r}(u - \varphi)(x, t) < 0$. There exists $0 < \delta < r$ such that $|(u - \varphi)(x_0, s)| < -\frac{\alpha}{2}$ for all $s \in [T - \delta, T]$, thus

$$(u - \varphi)(x, t) < \frac{\alpha}{2} + (u - \varphi)(x_0, T - \delta)$$

for all $(x, t) \in \partial\Omega_r$. Therefore

$$(u - \varphi_\varepsilon) < (u - \varphi)(x_0, T - \delta) + \frac{\alpha}{2} - \frac{\varepsilon}{T - t} < (u - \varphi)(x_0, T - \delta) + \frac{\alpha}{2} - \frac{\varepsilon}{r}$$

for all $(x, t) \in \partial\Omega_r$. Choose ε such that $\varepsilon\left(\frac{1}{\delta} - \frac{1}{r}\right) < -\frac{\alpha}{2}$, we obtain

$$(u - \varphi_\varepsilon)(x, t) < (u - \varphi_\varepsilon)(x_0, T - \delta)$$

for all $(x, t) \in \partial\Omega_r$. Thus the maximum point $(x_\varepsilon, t_\varepsilon)$ of $u - \varphi_\varepsilon$ cannot belong to $\partial\Omega_r$. Our claim is proven with $\varepsilon(r) = -\frac{\alpha_r}{2}\left(\frac{1}{\delta_r} - \frac{1}{r}\right)^{-1}$. Now let $r = \frac{1}{n}$ we obtain a sequence $\varepsilon_n \to 0$ and thus $\left(x_{\varepsilon_n}, t_{\varepsilon_n}\right) \to (x_0, T)$ since Ω_r shrinks to (x_0, T) as $r \to 0$. □

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 $x, y, p, q \in \mathbb{R}^n$,*

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

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

$$\begin{cases} u^i_t + H(x, Du^i) & = & 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^i(x,0) & = & g^i(x) & \text{on } \mathbb{R}^n, \end{cases} \tag{A.22}$$

where $g^i \in \mathrm{BUC}(\mathbb{R}^n)$ is given. Use the comparison principle for (A.22) *to show the following L^∞ contraction property: For any $t \geq 0$,*

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

Proof of Exercise 11. Denote $C = \|g^1 - g^2\|_{L^\infty(\mathbb{R}^n)} = \sup_{x\in\mathbb{R}^n} |g^1(x) - g^2(x)|$.

- Let $\zeta(x,t) = u^2(x,t) + C$, then it is a viscosity supersolution of (A.22) with the initial data $g^1(x)$, since:
  - If $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ such that $\zeta - \varphi$ has a local min at (x_0, t_0) then $u^2 - \varphi$ also has a local min at (x_0, t_0), hence $\varphi_t(x,t) + H(x, D\varphi(x,t)) \geq 0$.
  - $\zeta(x,0) = u^2(x,0) + C = g^2(x) + C \geq g^1(x)$.

  By comparison principle for (A.22), $\zeta(x,t) \geq u^1(x,t)$, that is, $u^2(x,t) + C \geq u^1(x,t)$ for $(x,t) \in \mathbb{R}^n \times (0,\infty)$.

- Let $\delta(x,t) = u^2(x,t) - C$, then it is a viscosity subsolution of (A.22) with the initial data $g^1(x)$, since:
  - If $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ such that $\delta - \varphi$ has a local max at (x_0, t_0) then $u^2 - \varphi$ also has a local max at (x_0, t_0), hence $\varphi_t(x,t) + H(x, D\varphi(x,t)) \leq 0$.
  - $\delta(x,0) = u^2(x,0) - C = g^2(x) - C \leq g^1(x)$.

  By comparison principle for (A.22), $\delta(x,t) \leq u^1(x,t)$, that is, $u^2(x,t) - C \leq u^1(x,t)$ for $(x,t) \in \mathbb{R}^n \times (0,\infty)$.

Therefore $u^2(x,t) - C \leq u^1(x,t) \leq u^2(x,t) + C$, which yields that

$$|u^1(x,t) - u^2(x,t)| \leq C \qquad \text{for all } (x,t) \in \mathbb{R}^n \times (0,\infty).$$

□

Exercise 12. *Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 Hamiltonian satisfying*

$$\begin{cases} H, D_pH \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for each } R > 0, \\ \displaystyle\lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} \left(\frac{1}{2} H(x,p)^2 + D_xH(x,p)\cdot p \right) = +\infty. \end{cases} \tag{A.23}$$

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

$$u^\varepsilon + H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon \qquad \text{in } \mathbb{R}^n. \tag{A.24}$$

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

Proof of Exercise 13. For $k = 1, 2, \ldots, n$, differentiate (A.24) with respect to x_k, we have

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

Multiplying two sides by $u^\varepsilon_{x_k}$ and taking the sum over $k = 1, 2, \ldots, n$, we imply

$$\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2 + D_xH(x, Du^\varepsilon) \cdot Du^\varepsilon + D_pH(x, Du^\varepsilon) \cdot \sum_{k=1}^n Du^\varepsilon_{x_k} u^\varepsilon_{x_k} = \varepsilon \sum_{k=1}^n \Delta u^\varepsilon_{x_k} u^\varepsilon_{x_k}. \tag{A.25}$$

- $Du^\varepsilon_{x_k} u^\varepsilon_{x_k} = \left(u^\varepsilon_{x_1x_k} u^\varepsilon_{x_k}, \ldots, u^\varepsilon_{x_nx_k} u^\varepsilon_{x_k}\right) = \frac{1}{2}\left(\frac{\partial}{\partial x_1}\left(u^\varepsilon_{x_k}\right)^2, \ldots, \frac{\partial}{\partial x_n}\left(u^\varepsilon_{x_k}\right)^2\right) = \frac{1}{2} D\left(u^\varepsilon_{x_k}\right)^2.$
- $\Delta\left(u^\varepsilon_{x_k}\right)^2 = 2\sum_{i=1}^n \left(\frac{\partial^2}{\partial x_i^2} u^\varepsilon_{x_k}\right) u^\varepsilon_{x_k} + \frac{1}{2}\sum_{i=1}^n \left(\frac{\partial u^\varepsilon_{x_k}}{\partial x_i}\right)^2$, hence

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

Using these equations, (A.25) becomes

$$\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2 + D_xH(x, Du^\varepsilon) \cdot Du^\varepsilon + D_pH(x, Du^\varepsilon) \cdot D\left(\frac{1}{2}\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2\right) = \varepsilon\Delta\left(\frac{1}{2}\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2\right) - \varepsilon\left|D^2u^\varepsilon\right|^2.$$

Set $\psi^\varepsilon(x, t) = \frac{1}{2}\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2 = \frac{1}{2}|Du^\varepsilon|^2 \ge 0$. Then,

$$\left(2\psi^\varepsilon + D_pH(x, Du^\varepsilon) \cdot D\psi^\varepsilon - \varepsilon\Delta\psi^\varepsilon\right) + D_xH(x, Du^\varepsilon) \cdot Du^\varepsilon + \varepsilon\left|D^2u^\varepsilon\right|^2 = 0. \tag{A.26}$$

If $\varepsilon < \frac{1}{n}$, then

$$\varepsilon\left|D^2u^\varepsilon\right|^2 \ge \varepsilon \sum_{i=1}^n \left(u^\varepsilon_{x_ix_i}\right)^2 \ge \frac{\varepsilon}{n}\left(\sum_{i=1}^n u^\varepsilon_{x_ix_i}\right)^2 = \frac{\varepsilon}{n}(\Delta u^\varepsilon)^2 \ge \left(\varepsilon\Delta u^\varepsilon\right)^2 = \left(u^\varepsilon + H(x, Du^\varepsilon)\right)^2.$$

Assume u^ε achieves its maximum and minimum at x_1 and x_2, respectively, then $Du^\varepsilon(x_1) = Du^\varepsilon(x_2) = 0$ and $\Delta u^\varepsilon(x_1) \le 0 \le \Delta u^\varepsilon(x_2)$. Thus,

$$-C \le -H(x_2, 0) \le u^\varepsilon(x_2) \le u^\varepsilon(x) \le u^\varepsilon(x_1) \le -H(x_1, 0) \le C \implies |u^\varepsilon(x)| \le C$$

for all $x \in \mathbb{R}^n$ and $\varepsilon > 0$. In particular, $\left(u^\varepsilon + H(x, Du^\varepsilon)\right)^2 \ge \frac{1}{2}H(x, Du^\varepsilon)^2 - C$ for some constant C independent of ε. Now, using this fact in (A.26), we have:

$$\left(2\psi^\varepsilon + D_p(x, Du^\varepsilon) \cdot D\psi^\varepsilon - \varepsilon\Delta\psi^\varepsilon\right) + \frac{1}{2}H(x, Du^\varepsilon)^2 + D_xH(x, Du^\varepsilon) \cdot Du^\varepsilon \le C. \tag{A.27}$$

Now, let us assume that ψ^ε achieves its max on $\mathbb{R}^n$ at x_ε, then $D\psi^\varepsilon(x_\varepsilon) = 0$ and $\Delta\psi^\varepsilon(x_\varepsilon) \le 0$, at x_ε. Plug these into (A.27) to yield

$$\frac{1}{2}H\Big(x_\varepsilon, Du^\varepsilon(x_\varepsilon)\Big)^2 + D_x H\Big(x_\varepsilon, Du^\varepsilon(x_\varepsilon)\Big) \cdot Du^\varepsilon(x_\varepsilon) \le C.$$

This is true for all $\varepsilon > 0$, by coercivity assumption we must have $|Du^\varepsilon(x_\varepsilon)| \le C$ for all $\varepsilon > 0$. It follows that

$$|Du^\varepsilon(x)| \le |Du^\varepsilon(x_\varepsilon)| \le C$$

for all $x \in \mathbb{R}^n$ since $\psi^\varepsilon(x) = \frac{1}{2}|Du^\varepsilon(x)|^2$. Thus $|Du^\varepsilon| \le C$ for all $\varepsilon > 0$ small enough. □

Exercise 19. *Assume that the cost function satisfies*

$$\begin{cases} f \in C(\mathbb{R}^n \times V), \qquad |f(x,v)| \le C \qquad \text{for all } (x,v) \in \mathbb{R}^n \times V. \\ |f(y_1,v) - f(y_2,v)| \le \mathrm{Lip}(b)|y_1 - y_2|. \end{cases}$$

Assume that $b(\cdot, v)$ is Lipschitz in the first variable for all v, i.e.,

$$|b(y_1,v) - b(y_2,v)| \le C|y_1 - y_2|$$

for all $y_1, y_2 \in \mathbb{R}^n$ and $v \in V$. Set $\lambda_0 = \|D_y b(\cdot,\cdot)\|_{L^\infty(\mathbb{R}^n \times V)}$, i.e., the best constant Lip(b) *in the above inequality is λ_0. Prove that*

(a) If $\lambda > \lambda_0$, then $u \in C^{0,1}(\mathbb{R}^n) = \mathrm{Lip}(\mathbb{R}^n) = W^{1,\infty}(\mathbb{R}^n)$.

(b) If $\lambda = \lambda_0$, then $u \in C^{0,\alpha}(\mathbb{R}^n)$ for any $0 < \alpha < 1$.

(c) If $0 < \lambda < \lambda_0$, then $u \in C^{0,\frac{\lambda}{\lambda_0}}(\mathbb{R}^n)$.

Proof of Exercise 21. Let $v(\cdot)$ be a control, and $y_x(\cdot)$ and $y_z(\cdot)$ be solutions to

$$\begin{cases} y_x'(s) &= b(y_x(s), v(s)), \\ y_x(0) &= x, \end{cases} \qquad \text{and} \qquad \begin{cases} y_z'(s) &= b(y_z(s), v(s)), \\ y_z(0) &= z, \end{cases}$$

respectively. Then, for all $s > 0$, we have

$$|y_x'(s) - y_z'(s)| \le \lambda_0 |y_x(s) - y_z(s)| \qquad \Longrightarrow \qquad |\varphi'(s)| \le \lambda_0 |\varphi(s)|$$

where $\varphi(s) = y_x(s) - y_z(s)$. By the fundamental theorem of calculus,

$$|\varphi(t)| = \left|\varphi(0) + \int_0^t \varphi'(s)\, ds\right| \le |\varphi(0)| + \int_0^t |\varphi'(s)|\, ds \le |x - z| + \lambda_0 \int_0^t |\varphi(s)|\, ds.$$

By Gronwall's inequality, we obtain

$$|\varphi(t)| = |y_x(t) - y_z(t)| \le e^{\lambda_0 t}|x - z| \qquad \text{for all } t > 0. \tag{A.28}$$

(a) For (a) there is no need to use DPP, indeed, for any control $v(\cdot)$ we have:

$$\begin{aligned}
|J(x,v(\cdot))-J(z,v(\cdot))| &= \left|\int_0^\infty e^{-\lambda s} f\left(y_x(s),v(s)\right)\,ds - \int_0^\infty e^{-\lambda s} f\left(y_z(s),v(s)\right)\,ds\right| \\
&\le \int_0^\infty e^{-\lambda s}\left|f\left(y_x(s),v(s)\right)-f\left(y_z(s),v(s)\right)\right|\,ds \\
&\le \int_0^\infty Ce^{(-\lambda+\lambda_0)s}|x-z|\,ds = \frac{C}{\lambda-\lambda_0}|x-z| = C_0|x-z|.
\end{aligned}$$

From that we have

$$\begin{aligned}
J(x,v(\cdot)) \le C_0|x-z|+J(z,v(\cdot)) &\implies u(x)\le C_0|x-z|+J(z,v(\cdot)) \\
\text{(take inf over } v(\cdot)) &\implies u(x)\le C_0|x-z|+u(z) \\
J(z,v(\cdot)) \le C_0|x-z|+J(x,v(\cdot)) &\implies u(z)\le C_0|x-z|+J(x,v(\cdot)) \\
\text{(take inf over } v(\cdot)) &\implies u(z)\le C_0|x-z|+u(x).
\end{aligned}$$

(b) We define

$$K(t,x,v(\cdot)) = \int_0^t e^{-\lambda s} f\left(y_{x,v(\cdot)}(s),v(s)\right)\,ds + e^{-\lambda t}u\left(y_{x,v(\cdot)}(t)\right).$$

Then,

$$u(x) = \inf_{v(\cdot)} K(t,x,v(\cdot))$$

for all $t>0$ by dynamic programming principle. Besides, as $|f(x,v)|\le C$ for all $(x,v)\in\mathbb{R}^n\times V$, it is clear that

$$|u(x)| \le \int_0^\infty Ce^{-\lambda s}\,ds = \frac{C}{\lambda}. \tag{A.29}$$

From (A.29) and (A.28), we have

$$\begin{aligned}
&|K(t,x,v(\cdot))-K(t,z,v(\cdot))| \\
&\le \int_0^t e^{-\lambda s}\left|f\left(y_x(s),v(s)\right)-f\left(y_z(s),v(s)\right)\right|\,ds + e^{-\lambda t}\left|u\left(y_x(t)\right)-u\left(y_z(t)\right)\right| \\
&\le C|x-z|\int_0^t e^{(\lambda_0-\lambda)s}\,ds + \frac{2C}{\lambda}e^{-\lambda t} \\
&= C|x-z|t + \frac{2C}{\lambda}e^{-\lambda t} \le 2C\left(|x-z|t+\frac{e^{-\lambda t}}{\lambda}\right).
\end{aligned}$$

This is true for all $t>0$, thus we can see it as a function of t, then the minimum of the right hand side will be obtained at t such that $F'(t)=0$, where

$$\begin{aligned}
F(t)=|x-z|t+\frac{e^{-\lambda t}}{\lambda} &\implies F'(t)=|x-z|-e^{-\lambda t} \\
&\implies F'(t)=0 \quad \text{iff} \quad t=\frac{1}{\lambda}\log\left(\frac{1}{|x-z|}\right).
\end{aligned}$$

We consider the case $0 < |x-z| < 1$ first so that the value t above is indeed positive. Then,

$$|K(t,x,v(\cdot)) - K(t,z,v(\cdot))| \le \frac{2C}{\lambda}\left(|x-z|\log\left(\frac{1}{|x-z|}\right) + |x-z|\right).$$

Setting $G(s) = s\left(\log\left(\frac{1}{s}\right)+1\right) = s(1-\log(s))$, we prove that there exists $C_\alpha > 0$ such that $G(s) \le C_\alpha s^\alpha$ on $(0,1)$ for any $0<\alpha<1$. Indeed, for $\beta = 1-\alpha \in (0,1)$, we have

$$G_\beta(s) = \frac{s(1-\log(s))}{s^\alpha} = s^\beta(1-\log(s))$$

is continuous on $(0,1)$ and $\lim_{s\to 0} G_\beta(s) = 0$, $\lim_{s\to 1} G_\beta(s) = 1$, thus, G_β is bounded by some constant C_α. If $\alpha \in (0,1)$ and $|x-z|<1$, then

$$\begin{aligned}
K(t,x,v(\cdot)) \le C_\alpha|x-z|^\alpha + K(t,z,v(\cdot)) &\implies u(x) \le C_\alpha|x-z|^\alpha + K(t,z,v(\cdot))\\
\text{(take inf over } v(\cdot)) &\implies u(x) \le C_\alpha|x-z|^\alpha + u(z)\\
K(t,z,v(\cdot)) \le C_\alpha|x-z|^\alpha + K(t,x,v(\cdot)) &\implies u(z) \le C_\alpha|x-z|^\alpha + K(t,x,v(\cdot))\\
\text{(take inf over } v(\cdot)) &\implies u(z) \le C_\alpha|x-z|^\alpha + u(x).
\end{aligned}$$

Therefore $|u(x)-u(z)| \le C_\alpha|x-z|^\alpha$ whenever $|x-z|<1$, i.e., $\frac{|u(x)-u(z)|}{|x-z|^\alpha} \le C_\alpha$ if $0<|x-z|<1$. If $|x-z|\ge 1$, then from (A.29) we have

$$\frac{|u(x)-u(z)|}{|x-z|^\alpha} \le \frac{2C}{\lambda} \qquad \implies \qquad \frac{|u(x)-u(z)|}{|x-z|^\alpha} \le \max\left\{C_\alpha, \frac{2C}{\lambda}\right\}$$

for any $x \ne z$ and for any $\alpha \in (0,1)$.

(c) From (A.29) and (A.28), we have

$$\begin{aligned}
&|K(t,x,v(\cdot)) - K(t,z,v(\cdot))|\\
&\le \int_0^t e^{-\lambda s}\,|f(y_x(s),v(s)) - f(y_z(s),v(s))|\,ds + e^{-\lambda t}\,|u(y_x(t)) - u(y_z(t))|\\
&\le C|x-z|\int_0^t e^{(\lambda_0-\lambda)s}\,ds + \frac{2C}{\lambda}e^{-\lambda t}\\
&= C|x-z|\frac{e^{(\lambda_0-\lambda)t}-1}{\lambda_0-\lambda} + \frac{2C}{\lambda}e^{-\lambda t} \le \frac{C}{\lambda_0-\lambda}|x-z|e^{(\lambda_0-\lambda)t} + \frac{2C}{\lambda}e^{-\lambda t}.
\end{aligned}$$

This is true for all $t>0$, thus we can see it as a function in t, then the minimum of the right hand side will be obtained at t such that $F'(t)=0$, where

$$\begin{aligned}
F(t) = \frac{C}{\lambda_0-\lambda}|x-z|e^{(\lambda_0-\lambda)t} + \frac{2C}{\lambda}e^{-\lambda t} \quad &\implies \quad F'(t) = C|x-z|e^{(\lambda_0-\lambda)t} - 2Ce^{-\lambda t}\\
&\implies \quad F'(t) = 0 \quad \text{iff} \quad t = \frac{1}{\lambda_0}\log\left(\frac{2}{|x-z|}\right).
\end{aligned}$$

We consider the case $0<|x-z|<2$ first so that the value t above is indeed positive. Then,

$$\begin{aligned}
|K(t,x,v(\cdot)) - K(t,z,v(\cdot))| &\le \frac{C}{\lambda_0-\lambda}|x-z|2^{\frac{\lambda_0-\lambda}{\lambda_0}}|x-z|^{\frac{\lambda-\lambda_0}{\lambda_0}} + \frac{2C}{\lambda}2^{\frac{-\lambda}{\lambda_0}}|x-z|^{\frac{\lambda}{\lambda_0}}.\\
&= \left(2^{\frac{\lambda_0-\lambda}{\lambda_0}}\frac{C}{\lambda_0-\lambda} + 2^{\frac{\lambda_0-\lambda}{\lambda_0}}\frac{C}{\lambda}\right)|x-z|^{\frac{\lambda}{\lambda_0}} \le C_1|x-z|^{\frac{\lambda}{\lambda_0}}.
\end{aligned}$$

By using a similar argument to the latter part of (b) and DPP, we have $|u(x)-u(z)| \leq C_1|x-z|^{\frac{\lambda}{\lambda_0}}$ whenever $|x-z|<2$, i.e., $\frac{|u(x)-u(z)|}{|x-z|^{\frac{\lambda}{\lambda_0}}} \leq C_1$. If $|x-z| \geq 2$, then from (A.29), we have

$$\frac{|u(x)-u(z)|}{|x-z|^{\frac{\lambda}{\lambda_0}}} \leq \frac{2C}{\lambda 2^{\frac{\lambda}{\lambda_0}}} = 2^{1-\frac{\lambda}{\lambda_0}}\frac{C}{\lambda} = C_2 \qquad \Longrightarrow \qquad \frac{|u(x)-u(z)|}{|x-z|^{\frac{\lambda}{\lambda_0}}} \leq \max\{C_1, C_2\} = C_3$$

for any $x \neq z$. Thus, $|u(x)-u(z)| \leq C_3|x-z|^{\frac{\lambda}{\lambda_0}}$.

□

Exercise 20. *Compute the Legendre transform $L(x,v)$ of the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, where*

$$H(x,p) = \frac{|p|^m}{m} + V(x) \qquad \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Here, $m \geq 1$ and $V \in \mathrm{BUC}(\mathbb{R}^n)$.

Proof of Exercise 22. We have

$$L(x,v) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - H(x,p)\Big) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - \frac{|p|^m}{m}\Big) - V(x).$$

The map $f : \mathbb{R}^n \to \mathbb{R}$ maps $p \mapsto p\cdot v - \frac{|p|^m}{m}$ is continuous, and

$$\lim_{|p|\to\infty} f(p) = \lim_{|p|\to\infty} |p|\left(\frac{p\cdot v}{|p|} - \frac{|p|^{m-1}}{m}\right) = -\infty.$$

Therefore, f achieves maximum on $\mathbb{R}^n$ at p^* such that $Df(p^*) = 0$. We have

$$Df(p) = v - p|p|^{m-2} = 0 \quad \Longleftrightarrow \quad p^*|p^*|^{m-2} = v,$$

which gives

$$f(p^*) = |v|^{\frac{m}{m-1}} - \frac{1}{m}|v|^{\frac{m}{m-1}} = \frac{m-1}{m}|v|^{\frac{m}{m-1}}.$$

Thus, for $(x,v) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$L(x,v) = \frac{m-1}{m}|v|^{\frac{m}{m-1}} - V(x).$$

□

Exercise 22. *Consider the Cauchy problem*

$$\begin{cases} u_t(x,t) + H(x,Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad (x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{A.30}$$

For $(x,t) \in \mathbb{R}^n \times (0,\infty)$, let

$$u(x,t) = \inf\left\{\int_0^t L(\gamma(s),\gamma'(s))\,ds + u_0(\gamma(0)) \,:\, \gamma(t) = x, \gamma(0) \in \mathbb{R}^n, \gamma' \in L^1([0,t])\right\}.$$

Using the Dynamic programming principle (DPP)

$$u(x,t) = \inf\left\{\int_s^t L(\gamma(r),\gamma'(r))\,dr + u(\gamma(s),s) \,:\, \gamma(t) = x, \gamma' \in L^1([s,t])\right\} \tag{DPP}$$

to prove that u is a viscosity solution to (A.30).

Proof of Exercise 24. The initial condition is obviously true. The subsolution test is pretty simple. Take $\varphi \in C^1(\mathbb{R}^n \times [0,\infty))$ such that $u-\varphi$ has a strict local maximum at $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$, and $u(x_0,t_0) = \varphi(x_0,t_0)$, we need to prove

$$\varphi_t(x_0,t_0) + H(x_0, D\varphi(x_0,t_0)) \le 0. \tag{A.31}$$

Pick a path $\gamma(\cdot)$ with $\gamma(t_0) = x_0$, then for $s < t_0$, we have

$$\begin{aligned} u(\gamma(t_0),t_0) - u(\gamma(s),s) &\ge \varphi(\gamma(t_0),t_0) - \varphi(\gamma(s),s) \\ &= \int_s^{t_0} \Big(\varphi_t(\gamma(r),r) + \gamma'(r)\cdot D\varphi(\gamma(r),r)\Big)\,dr. \end{aligned} \tag{A.32}$$

By dynamic programming principle (DPP), we have

$$\int_s^{t_0} \Big(L(\gamma(r),\gamma'(r))\Big)\,dr \ge u(\gamma(t_0),t_0) - u(\gamma(s),s). \tag{A.33}$$

From (A.32) and (A.33) we have

$$0 \ge \frac{1}{t_0 - s}\int_s^{t_0} \Big(\varphi_t(\gamma(r),r) + \gamma'(r)\cdot D\varphi(\gamma(r),r) - L(\gamma(r),\gamma'(r))\Big)\,dr.$$

Since the function inside the integral sign is continuous, taking $s \to t_0$, we obtain

$$\varphi_t(\gamma(t_0),t_0) + \gamma'(t_0)\cdot D\varphi(\gamma(t_0),t_0) - L(\gamma(t_0),\gamma'(t_0)) \le 0$$

and thus (A.31) is true since we can design the path $\gamma(\cdot)$ such that $\gamma'(t_0) = v$ for any $v \in \mathbb{R}^n$.

Now we perform the supersolution test. Take $\varphi \in C^1(\mathbb{R}^n \times [0,\infty))$ such that $u-\varphi$ has a strict local minimum at (x_0,t_0), and $u(x_0,t_0) = \varphi(x_0,t_0)$, we need to prove

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

For any $s \in (0,t_0)$ we have

$$\begin{aligned} u(\gamma(t_0),t_0) - u(\gamma(s),s) &\le \varphi(\gamma(t_0),t_0) - \varphi(\gamma(s),s) \\ &= \int_s^{t_0} \Big(\varphi_t(\gamma(r),r) + \gamma'(r)\cdot D\varphi(\gamma(r),r)\Big)\,dr. \end{aligned}$$

Let us subtract from two sides by $L(\gamma(r),\gamma'(r))$ we obtain

$$\begin{aligned} u(x_0,t_0) - &\left(\int_s^{t_0} L(\gamma(r),\gamma'(r))\,dr + u(\gamma(s),s)\right) \\ &\le \int_s^{t_0} \Big(\varphi_t(\gamma(r),r) + H(\gamma(r),D\varphi(\gamma(r),r))\Big)\,dr. \end{aligned} \tag{A.34}$$

Define $\mathcal{A}$ to be the set of all “almost-admissible" paths with $\gamma(t_0) = x_0$, i.e., $\gamma(\cdot)$ such that

$$\int_0^{t_0} L(\gamma(s),\gamma'(s))\,ds + u_0(\gamma(0)) < u(x_0,t_0) + 1.$$

It is easy to see that the Dynamic Programming Principle remains true with the new admissible set $\mathcal{A}$. Take the infimum over all paths $\gamma(\cdot)\in\mathcal{A}$ in (A.34), we obtain

$$0 \leq \sup_{\substack{\gamma(t_0)=x_0\\ \gamma\in\mathcal{A}}} \underbrace{\int_s^{t_0} \Big(\varphi_t(\gamma(r),r) + H(\gamma(r), D\varphi(\gamma(r),r))\Big)\,dr}_{\mathcal{K}[\gamma(\cdot)]}. \tag{A.35}$$

Now for $\gamma(\cdot)\in\mathcal{A}$ we have

$$\begin{aligned}\mathcal{K}[\gamma(\cdot)] &= (t_0-s)\Big(\varphi_t(x_0,t_0) + H(x_0, D\varphi(x_0,t_0))\Big)\\ &\quad + \int_s^{t_0}\Big[\Big(\varphi_t(\gamma(r),r)-\varphi_t(x_0,t_0)\Big) + \Big(H(\gamma(r),D\varphi(\gamma(r),r)) - H(x_0,D\varphi(x_0,t_0))\Big)\Big]\,dr.\end{aligned} \tag{A.36}$$

1. Now given $\eta>0$, since φ is smooth and H is continuous at (x_0,t_0), there exists $\delta>0$ such that

$$|(y,s)-(x_0,t_0)|<\delta \qquad \Longrightarrow \qquad \begin{cases}|\varphi_t(y,s)-\varphi_t(x_0,t_0)|<\eta\\ |H(y,D\varphi(y,s)) - H(x_0,D\varphi(x_0,t_0))|<\eta.\end{cases}$$

2. By Lemma A.25 we know that $|\gamma(r)|$ is bounded independent of $\gamma\in\mathcal{A}$ and $r<t_0$, thus since u is locally bounded, we can get $|u(\gamma(r),r)|\leq C = C(x_0,t_0)$ for all $r\in[s,t_0]$. Thus given $\delta>0$, by super-linearity we can choose M large so that

$$\inf_{x\in\mathbb{R}^n}\left(\frac{L(x,v)}{|v|}\right) > \frac{2(2C+1)}{\delta} \qquad \text{for all} \qquad |v|\geq M. \tag{A.37}$$

3. Let $\varepsilon>0$, by (DPP) we can find $\gamma\in\mathcal{A}$ such that ($\varepsilon\ll 1$)

$$\int_s^{t_0} L(\gamma(r),\gamma'(r))\,dr \leq u(x_0,t_0) - u(\gamma(s),s) + \varepsilon \leq 2C+1. \tag{A.38}$$

   - Estimate for the first term is easy:

$$\int_{\{r\in[s,t_0]:|\gamma'(r)|\leq M\}} |\gamma'(r)|\,dr \leq M(t_0-s). \tag{A.39}$$

   - Estimate for the second term:

$$\int_{\{r\in[s,t_0]:|\gamma'(r)|\leq M\}} L(\gamma(r),\gamma'(r))\,dr \geq -\left(\sup_{\substack{x\in\mathbb{R}^n\\ |v|\leq M}} L(x,v)\right) =: -C_M.$$

   and

$$\begin{aligned}\int_{\{r\in[s,t_0]:|\gamma'(r)|\geq M\}} L(\gamma(r),\gamma'(r))\,dr &= \int_{\{r\in[s,t_0]:|\gamma'(r)|\geq M\}} \left(\frac{L(\gamma(r),\gamma'(r))}{|\gamma'(r)|}\right)|\gamma'(r)|\,dr\\ &\geq \left(\inf_{\substack{x\in\mathbb{R}^n\\ |v|\geq M}} \frac{L(x,v)}{|v|}\right)\int_{\{r\in[s,t_0]:|\gamma'(r)|\geq M\}} |\gamma'(r)|\,dr.\end{aligned}$$

From this we obtain

$$\int_s^{t_0} L(\gamma(r),\gamma'(r))\,dr + C_M \geq \left(\inf_{\substack{x\in\mathbb{R}^n\\|v|\geq M}} \frac{L(x,v)}{|v|}\right)\int_{\{r\in[s,t_0]:|\gamma'(r)|\geq M\}} |\gamma'(r)|\,dr.$$

From (A.38) and (A.37) we obtain

$$\int_{\{r\in[s,t_0]:|\gamma'(r)|\geq M\}} |\gamma'(r)|\,dr \leq \left(\inf_{\substack{x\in\mathbb{R}^n\\|v|\geq M}} \frac{L(x,v)}{|v|}\right)^{-1}(2C+1) \leq \frac{\delta}{2}. \tag{A.40}$$

4. With M in step 2, (A.39) and (A.40) yield

$$\begin{aligned}\sup_{r\in[s,t_0]} |\gamma(r)-x_0| &\leq \int_s^t |\gamma'(r)|\,dr\\ &\leq \int_{\{r\in[s,t]:|\gamma'(r)|\leq M\}} |\gamma'(r)|\,dr + \int_{\{r\in[s,t]:|\gamma'(r)|\geq M\}} |\gamma'(r)|\,dr \leq M(t_0-s)+\frac{\delta}{2}.\end{aligned}$$

Choose s closed to t_0 such that $M(t_0-s)<\frac{\delta}{2}$, we obtain

$$\sup_{r\in[s,t_0]} |\gamma(r)-x_0| < \delta$$

which implies that

$$\begin{cases} |\varphi_t(\gamma(r),r)-\varphi_t(x_0,t_0)| < \eta\\ |H(\gamma(r),D\varphi(\gamma(r),r))-H(x_0,D\varphi(x_0,t_0))| < \eta.\end{cases}$$

Using these facts in (A.36), we obtain that for any given η, there exists $s\in[0,t_0]$ such that

$$\mathcal{K}[\gamma(\cdot) \leq (t_0-s)\Big(\varphi_t(x_0,t_0)+H(x_0,D\varphi(x_0,t_0))\Big)+2\eta(t_0-s).$$

Taking sup over all path $\gamma(\cdot)\in\mathcal{A}$ and divide both sides by $t_0-s>0$, we obtain

$$\varphi_t(x_0,t_0)+H(x_0,D\varphi(x_0,t_0))+2\eta \geq 0.$$

Finally since η is arbitrary, $\varphi_t(x_0,t_0)+H(x_0,D\varphi(x_0,t_0))\geq 0$, and the proof is complete. □

Lemma A.25. *u is locally bounded on $\mathbb{R}^n\times[0,\infty)$, i.e., if $(x,t)\in B_R(0)\times[0,T]$ then there exists a positive constant $C_{R,T}$ such that $|u(x,t)|\leq C_{R,T}$.*

Proof. It is easy to see that $u(x,t)$ locally bounded from above on $\mathbb{R}^n\times[0,\infty)$. To prove $u(x,t)$ is locally bounded from below, let us fix $(x,t)\in B_R(0)\times[0,T]$ and $\gamma(\cdot)\in\mathcal{A}$, then for $(x,t)\in\mathbb{R}^n\times[0,T]$ we have:

1. For every path $\gamma(\cdot)\in\mathcal{A}$ then

$$\int_0^t L(\gamma(s),\gamma'(s))\,ds \leq T\left(\sup_{x\in\mathbb{R}^n}|L(x,0)|\right)+2\|u_0\|_{L^\infty}+1 =: C_T^1.$$

2. By the superlinearity of L, there exists $M > 0$ to such that

$$\inf_{x\in\mathbb{R}^n} \frac{L(x,v)}{|v|} \geq 1 \qquad \text{for all} \qquad |v| \geq M.$$

3. The $L^1([0,t])$ norm of $\gamma'(\cdot)$ is uniformly bounded. Indeed,

$$\begin{aligned} C_T^1 \geq \int_{\{s\in[0,t]:|\gamma'(s)|\geq M\}} L(\gamma(s),\gamma'(s))\,ds &= \int_{\{s\in[0,t]:|\gamma'(s)|\geq M\}} \left(\frac{L(\gamma(s),\gamma'(s))}{|\gamma'(s)|}\right)|\gamma'(s)|\,ds \\ &\geq \int_{\{s\in[0,t]:|\gamma'(s)|\geq M\}} \left(\inf_{x\in\mathbb{R}^n} \frac{L(x,\gamma'(s))}{|\gamma'(s)|}\right)|\gamma'(s)|\,ds \\ &\geq \int_{\{s\in[0,t]:|\gamma'(s)|\geq M\}} |\gamma'(s)|\,ds. \end{aligned} \tag{A.41}$$

And

$$\int_{\{s\in[0,t]:|\gamma'(s)|\leq M\}} |\gamma'(s)|\,ds \leq M(s-t) \leq MT. \tag{A.42}$$

which implies that for all $\gamma \in \mathcal{A}$ then

$$\int_0^t |\gamma'(s)|\,ds \leq C_T^1 + MT.$$

4. From the above result we have the bound for $|\gamma(\cdot)| \in \mathcal{A}$ as

$$|x-\gamma(s)| \leq \int_0^t |\gamma'(r)|\,dr \leq T(C_T + MT) \qquad \Longrightarrow \qquad |\gamma(s)| \leq R + T(C_T^1 + MT) = C_T^2.$$

5. From (A.41) we obtain for all $\gamma \in \mathcal{A}$ then

$$\int_{\{s\in[s,t]:|\gamma'(s)|\geq M\}} L(\gamma(s),\gamma'(s))\,ds \geq 0. \tag{A.43}$$

While on $\{s \in [0,t] : |\gamma'(s)| \leq M\}$ we can use the continuity of L to estimate

$$\begin{aligned} |L(\gamma(s),\gamma'(s)) - L(x,\gamma'(s))| &\leq \omega_{\max\{M,C_T^2\}}\big(|x-\gamma(s)|\big) \\ &\leq \omega_{\max\{M,C_T^2\}}\Big(T(C_T^1 + MT)\Big) = C_T^3 \end{aligned}$$

where $\omega(\cdot)$ is a modulus of continuity. Thus

$$L(\gamma(s),\gamma'(s)) \geq L(x,\gamma'(s)) - C_T^3 \qquad \text{where} \qquad |\gamma'(s)| \leq M.$$

6. Using the convexity of $v \mapsto L(x,v)$ at $v = 0$, there exists some $\xi \in D_v^- L(x,0)$, then

$$L(x,\gamma'(s)) \geq L(x,0) + \gamma'(s)\cdot\xi \geq L(x,0) - |\xi|\cdot|\gamma'(s)|$$

and thus from (A.42) and Lemma 2.17 we obtain

$$
\begin{aligned}
&\int_{\{s\in[s,t]:|\gamma'(s)|\leq M\}} L(\gamma(s),\gamma'(s))\,ds \geq \int_{s\in[s,t]:|\gamma'(s)|\leq M} L(x,\gamma'(s))\,ds - C_T^3 t\\
&\geq \int_{\{s\in[s,t]:|\gamma'(s)|\leq M\}} \Big(L(x,0)-|\xi|\cdot|\gamma'(s)|\Big)\,ds - C_T^3 T\\
&\geq -T\left(\sup_{x\in\mathbb{R}^n}|L(x,0)|\right) - |\xi|\int_{\{s\in[s,t]:|\gamma'(s)|\leq M\}} |\gamma'(s)|\,ds - C_T^3 T\\
&\geq -T\left(\sup_{x\in\mathbb{R}^n}|L(x,0)|\right) - |\xi| MT - C_T^3 T\\
&\geq -T\left(\sup_{x\in\mathbb{R}^n}|L(x,0)|\right) - \left(\sup_{|x|\leq R}|D_v^- L(x,0)|\right) MT - C_T^3 T = C_T^4.
\end{aligned}
$$

Finally from (A.43) and the previous step we obtain for $\gamma(\cdot)\in\mathcal{A}$ then

$$u(x,t)+\|u_0\|_{L^\infty}+1 \geq \int_0^t L(\gamma(s),\gamma'(s))\,ds \geq C_T^4$$

which implies

$$u(x,t) \geq C_T^4 - \|u\|_{L^\infty} - 1.$$

Thus u is locally bounded. □

Exercise 28. *Assume that H satisfies* (4.2) *and* (4.3). *Fix $p\in\mathbb{R}^n$, and we look at* (4.5). *Show that there exists a constant $C>0$ independent of $\lambda>0$ such that, for any $\lambda>0$, we have*

$$\left\|\lambda v^\lambda(\cdot)+\overline{H}(p)\right\|_{L^\infty(\mathbb{T}^n)} \leq C\lambda.$$

Proof of Exercise 37. Let $C=\max_{y\in\mathbb{T}^n} H(y,p)$. Then, by the comparison principle, we have

$$\sup_{y\in\mathbb{T}^n}|\lambda v^\lambda(y)| \leq C.$$

The coercivity of H implies that $\sup_{y\in\mathbb{T}^n}|Dv^\lambda(y)|\leq C_1$, and for all $y,x_0\in\mathbb{T}^n$, we have

$$
\begin{aligned}
|v^\lambda(y)-v^\lambda(x_0)|\leq C_1\sqrt{n} \quad &\Longrightarrow \quad \lambda v^\lambda(x_0)-\lambda C_1\sqrt{n} \leq \lambda v^\lambda(y) \leq \lambda v^\lambda(x_0)+\lambda C_1\sqrt{n}\\
&\Longrightarrow \quad \lambda\sup_{\mathbb{T}^n} v^\lambda(\cdot)-\lambda C_1\sqrt{n} \leq \lambda v^\lambda(y) \leq \lambda\inf_{\mathbb{T}^n} v^\lambda(\cdot)+\lambda.C_1\sqrt{n}
\end{aligned}
$$

From the above, it is suffices to prove that

$$\lambda\inf_{\mathbb{T}^n} v^\lambda(\cdot) \leq -\overline{H}(p) \leq \lambda\sup_{\mathbb{T}^n} v^\lambda(\cdot) \quad\Longleftrightarrow\quad \underbrace{-\lambda\sup_{\mathbb{T}^n} v^\lambda(\cdot)}_{\beta} \leq \overline{H}(p) \leq \underbrace{-\lambda\inf_{\mathbb{T}^n} v^\lambda(\cdot)}_{\alpha}. \tag{A.44}$$

Let $v\in \mathrm{Lip}(\mathbb{T}^n)$ be any viscosity solution to the cell problem $H(y,p+Dv)=\overline{H}(p)$. If $\overline{H}(p)>\alpha$, then in the viscosity sense, we have

$$H(y,p+Dv(y))=\overline{H}(p)>\alpha\geq -\lambda v^\lambda(y) = H\left(y,p+Dv^\lambda(y)\right) \qquad \text{in } \mathbb{T}^n.$$

Since $v, v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ are bounded, we can choose $\delta > 0$ such that

$$\delta v(y) + H(y, p + Dv(y)) > \frac{\overline{H}(p) + \alpha}{2} > \delta v^\lambda(y) + H\left(y, p + Dv^\lambda(y)\right)$$

in the viscosity sense. Then, $v(\cdot)$ and $v^\lambda(\cdot)$ are a viscosity supersolution and subsolution to the problem $\delta w + H(y, p + Dw) = \frac{1}{2}(\overline{H}(p) + \alpha)$, respectively. Thus, by the comparison principle, $v \geq v^\lambda$. This is a contradiction since $v - C$ is also a viscosity solution to the cell problem for any constant $C \in \mathbb{R}$. Performing a similar procedure for β, we deduce that (A.44) is true, and thus, we obtain the rate of convergence is $O(\lambda)$. □

Bibliography

[1] Y. Achdou, F. Camilli and I. Capuzzo-Dolcetta, *Homogenization of Hamilton-Jacobi equations: numerical methods*, Math. Models Methods Appl. Sci. 18 (2008), no. 7, 1115–1143.

[2] E. S. Al-Aidarous, E. O. Alzahrani, H. Ishii, A. M. M. Younas, *A convergence result for the ergodic problem for Hamilton–Jacobi equations with Neumann type boundary conditions*, Proc. Roy. Soc. Edinburgh Sect. A 146 (2016) 225–242.

[3] H. Amann, M. G. Crandall, *On some existence theorems for semi-linear elliptic equations* Indiana University Mathematics Journal, Vol. 27, No. 5 (1978), 779–790.

[4] N. Anantharaman, R. Iturriaga, P. Padilla, H. Sanchez-Morgado, *Physical solutions of the Hamilton-Jacobi equation*, Discrete Contin. Dyn. Syst. Ser. B 5 (2005), no. 3, 513–528.

[5] S. N. Armstrong, H. V. Tran, *Viscosity solutions of general viscous Hamilton–Jacobi equations*, Mathematische Annalen, 361 (2015), no. 3, 647–687.

[6] S. N. Armstrong, H. V. Tran, Y. Yu, *Stochastic homogenization of a nonconvex Hamilton-Jacobi equation*, Calculus of Variations and PDE (2015), no. 2, 1507–1524.

[7] S. N. Armstrong, H. V. Tran, Y. Yu, *Stochastic homogenization of nonconvex Hamilton-Jacobi equations in one space dimension*, J. Differential Equations 261 (2016), 2702–2737.

[8] S. Artstein-Avidan, V. Milman, *The concept of duality in convex analysis, and the characterization of the Legendre transform*, Annals of Mathematics, 169(2009), 661–674.

[9] S. Aubry, *The twist map, the extended Frenkel–Kantorova model and the devil's staircase*, Physica D 7 (1983), 240–258.

[10] I. Babenko, F. Balacheff, *Sur la forme de la boule unité de la norme stable unidimensionnelle*, Manuscripta Math., 119(3):347–358, 2006. ISSN 0025–2611.

[11] V. Bangert, *Mather Sets for Twist Maps and Geodesics on Tori*, Dynamics Reported, Volume 1, 1988.

[12] V. Bangert, *Minimal geodesics*, Ergod. Th. & Dynam. Sys. (1989), 10, 263–286.

[13] M. Bardi, I. Capuzzo-Dolcetta, Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations, Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1997.

[14] M. Bardi and L. C. Evans, *On Hopf's formulas for solutions of Hamilton-Jacobi equations*, Nonlinear Analysis 8 (1984), 1373–1381.

[15] G. Barles, *A weak Bernstein method for fully nonlinear elliptic equations*, Differential Integral Equations 4 (1991), no. 2, 241–262.

[16] G. Barles, Solutions de viscosité des équations de Hamilton-Jacobi, Mathématiques & Applications (Berlin), 17, Springer-Verlag, Paris, 1994.

[17] G. Barles and P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic analysis, 4(3): 271–283, 1991.

[18] E. N. Barron, R. Jensen, *Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex Hamiltonians*, Comm. Partial Differential Equations 15 (1990), no. 12, 1713–1742.

[19] U. Bessi, *Aubry-Mather theory and Hamilton-Jacobi equations*, Comm. Math. Phys. 235 (2003), no. 3, 495–511.

[20] P. Bernard, *The asymptotic behaviour of solutions of the forced Burgers equation on the circle*, Nonlinearity 18 (2005) 101–124.

[21] P. Bernard, *Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds*, Ann. Sci. École Norm. Sup. (4) 40(3), 445–452 (2007).

[22] A. Bressan, Viscosity Solutions of Hamilton-Jacobi Equations and Optimal Control Problems, lecture notes on Bressan's webpage.

[23] F. Cagnetti, D. Gomes, H. Mitake, H. V. Tran, *A new method for large time behavior of convex Hamilton–Jacobi equations: degenerate equations and weakly coupled systems*, Annales de l'Institut Henri Poincaré - Analyse non linéaire 32 (2015), 183–200.

[24] F. Cagnetti, D. Gomes, H. V. Tran, *Aubry-Mather measures in the non convex setting*, SIAM J. Math. Anal. 43 (2011), no. 6, 2601–2629.

[25] J. Calder, Lecture notes on viscosity solutions, lecture notes on Calder's website.

[26] P. Cannarsa and C. Sinestrari, Semiconcave functions, Hamilton-Jacobi equations, and optimal control, Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston, Inc., Boston, 2002.

[27] I. Capuzzo-Dolcetta, H. Ishii, *On the rate of convergence in homogenization of Hamilton–Jacobi equations*, Indiana Univ. Math. J. 50 (2001), no. 3, 1113–1129.

[28] I. Capuzzo-Dolcetta, P.-L. Lions, *Hamilton-Jacobi equations with state constraints*, Trans. Am. Math. Soc. 318(2), 643–683 (1990).

[29] P. Cardaliaguet, Introduction to differential games, Lecture notes on Cardaliaguet's website.

[30] P. Cardaliaguet, J. Nolen, P. E. Souganidis, *Homogenization and Enhancement for the G–Equation*, Arch. Ration. Mech. Anal. 199, 527–561 (2011).

[31] M. J. Carneiro, *On minimizing measures of the action of autonomous Lagrangians*, Nonlinearity 8 (1995) 1077–1085.

[32] Y. G. Chen, Y. Giga, and S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, J. Diff. Geom., **33**, 1991, 749–786.

[33] M. C. Concordel, *Periodic homogenization of Hamilton–Jacobi equations: additive eigenvalues and variational formula*, Indiana Univ. Math. J. 45 (1996), no. 4, 1095–1117.

[34] M. C. Concordel, *Periodic homogenisation of Hamilton–Jacobi equations. II. Eikonal equations*, Proc. Roy. Soc. Edinburgh Sect. A 127 (1997), no. 4, 665–689.

[35] G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain, *Lagrangian graphs, minimizing measures and Mãné's critical values*, Geom. Funct. Anal., 8(5):788–809, 1998.

[36] M. G. Crandall, Viscosity solutions: a primer. Viscosity solutions and applications (Montecatini Terme, 1995), 1–43, Lecture Notes in Math., 1660, Fond. CIME/CIME Found. Subser., Springer, Berlin, 1997.

[37] M. G. Crandall, L. C. Evans, P.-L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 282 (1984), 487–502.

[38] M. G. Crandall, H. Ishii, and P.-L. Lions. *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27(1):1–67, 1992.

[39] M. G. Crandall, P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), 1–42.

[40] M. G. Crandall and P-L. Lions, *Two approximations of solutions of Hamilton-Jacobi equations*, Math. Comp. 43 (1984), 1–19.

[41] M. G. Crandall, A. Majda, *Monotone difference approximations for scalar conservation laws*, Math. Comp., vol. 34, 1980, pp. 1–21.

[42] A. Davini, A. Fathi, R. Iturriaga, M. Zavidovique, *Convergence of the solutions of the discounted equation*, Invent. Math. 206 (1) (2016) 29–55.

[43] B. Denet, *Possible role of temporal correlations in the bending of turbulent flame velocity*, Combust Theory Modelling, 3, 585–589, 1999.

[44] W. E, *Aubry-Mather theory and periodic solutions of the forced Burgers equation*, Comm. Pure Appl. Math., 52(7): 811–828, 1999.

[45] R. J. Elliott, Viscosity solutions and optimal control. Pitman Research Notes in Mathematics Series, 165. Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York (1987).

[46] L. C. Evans, *On solving certain nonlinear partial differential equations by accretive operator methods*, Israel J. Math. 86 (1980), 225–247.

[47] L. C. Evans, Weak convergence methods for nonlinear partial differential equations. CBMS 74, American Mathematical Society, 1990.

[48] L. C. Evans, *Periodic homogenisation of certain fully nonlinear partial differential equations*, Proceedings of the Royal Society of Edinburgh, 120A, 245–265, 1992.

[49] L. C. Evans, Partial differential equations. Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010.

[50] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton–Jacobi PDE*, Archive for Rational Mechanics and Analysis 197 (2010), 1053–1088.

[51] L. C. Evans, *Weak KAM theory and partial differential equations*, Calculus of variations and nonlinear partial differential equations, 123–154, Lecture Notes in Math., 1927, Springer, Berlin, 2008.

[52] L. C. Evans, *A survey of partial differential equations methods in weak KAM theory*, Comm. Pure Appl. Math. 57 (2004), no. 4, 445–480.

[53] L. C. Evans, D. Gomes, *Effective Hamiltonians and Averaging for Hamiltonian Dynamics. I*, Arch. Ration. Mech. Anal. 157 (2001), no. 1, 1–33.

[54] L. C. Evans, P. E. Souganidis, *Differential games and representation formulas for solutions of Hamilton-Jacobi Equations*, Indiana Univ. Math. J., 282, pp. 487–502.

[55] L. C. Evans, and J. Spruck, *Motion of level sets by mean curvature. I.* J. Diff. Geom. **33**, 1991, No. 3, 635–681.

[56] G. Fabbri, F. Gozzi and A. Swiech, Stochastic Optimal Control in Infinite Dimension: Dynamic Programming and HJB Equations, with a contribution by M. Fuhrman and G. Tessitore, Probability Theory and Stochastic Modeling, vol. 82, Springer, 2017.

[57] M. Falcone, M. Rorro, *On a variational approximation of the effective Hamiltonian*, Numerical Mathematics and Advanced Applications, Springer-Verlag, Berlin, 2008, 719–726.

[58] A. Fathi, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sér. I Math. 327 (1998), no. 3, 267–270.

[59] A. Fathi, Weak KAM Theorem in Lagrangian Dynamics, to appear in Cambridge Studies in Advanced Mathematics.

[60] A. Fathi, A. Siconolfi, *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, Invent. Math. 155 (2004), no. 2, 363–388.

[61] A. Fathi and A. Siconolfi, *PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians*, Calc. Var. Partial Differential Equations 22 (2005), no. 2, 185–228.

[62] W. Fleming, *The convergence problem for differential games II*, Advances in Game Theory, 1964, Princeton Univ. Press, 195–210.

[63] W. H. Fleming, H. M. Soner, Controlled Markov processes and viscosity solutions. Second edition. Stochastic Modelling and Applied Probability, 25. Springer, New York, 2006. xviii+429 pp.

[64] A. Friedman, Partial differential equations of parabolic type. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964 xiv+347 pp.

[65] A. Friedman, *The Cauchy Problem for First Order Partial Differential Equations,* Indiana University Mathematics Journal, Vol. 23, No. 1 (1973), pp. 27–40.

[66] H. Gao, *Random homogenization of coercive Hamilton-Jacobi equations in 1d*, Calc. Var. Partial Differential Equations, 55 (2016), no. 2, 1–39.

[67] H. Gao, *Stochastic homogenization of certain nonconvex Hamilton-Jacobi equations*, Journal of Differential Equations, Vol. 267, no. 5 (2019), 2918–2949.

[68] Y. Giga, Surface Evolution Equations. A Level Set Approach, Monographs in Mathematics, 99. Birkhäuser, Basel-Boston-Berlin (2006), xii+264pp.

[69] Y. Giga, H. Mitake, H. V. Tran, *On asymptotic speed of solutions to level-set mean curvature flow equations with driving and source terms*, SIAM J. Math. Anal. 48 (5), 3515–3546.

[70] Y. Giga, H. Mitake, T. Ohtsuka, H. V. Tran, Existence of asymptotic speed of solutions to birth and spread type nonlinear partial differential equations, *Indiana University Math Journal*, to appear.

[71] D. A. Gomes, *Viscosity solutions of Hamilton-Jacobi equations, and asymptotics for Hamiltonian systems*, Calc. Var. 14, 345–357 (2002).

[72] D. A. Gomes, *A stochastic analogue of Aubry-Mather theory*, Nonlinearity 15 (2002), no. 3, 581–603.

[73] D. A. Gomes, *Duality principles for fully nonlinear elliptic equations*, Trends in partial differential equations of mathematical physics, 125–136, Progr. Nonlinear Differential Equations Appl., 61, Birkhauser, Basel, 2005.

[74] D. A. Gomes, *Generalized Mather problem and selection principles for viscosity solutions and Mather measures*, Adv. Calc. Var., 1 (2008), 291–307.

[75] D. A. Gomes, Viscosity solutions of Hamilton-Jacobi equations, IMPA Mathematical Publications, 27th Brazilian Mathematics Colloquium, Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2009. ii+210 pp.

[76] D. Gomes, R. Iturriaga, H. Sanchez-Morgado, Y. Yu, *Mather measures selected by an approximation scheme*, Proc. Amer. Math. Soc. 138 (2010), no. 10, 3591–3601.

[77] D. Gomes, H. Mitake, H.V. Tran, *The Selection problem for discounted Hamilton-Jacobi equations: some non-convex cases*, Journal of the Mathematical Society of Japan, 70, no 1 (2018), 345–364.

[78] D. A. Gomes, A. M. Oberman, *Computing the effective Hamiltonian using a variational formula*, SIAM J. Control Optim. 43 (2004), pp. 792–812.

[79] G. A. Hedlund, *Geodesics on a two-dimensional Riemannian manifold with periodic coefficients*, Ann. of Math. 33 (1932), 719–739.

[80] E. Hopf, *Generalized solutions of nonlinear equations of first order*, J. Math. Mech. 14 (1965), 951–973.

[81] R. Isaacs, Differential games. A mathematical theory with applications to warfare and pursuit, control and optimization. John Wiley & Sons, Inc., New York-London-Sydney 1965 xvii+384 pp.

[82] H. Ishii, *Perron's method for Hamilton–Jacobi equations*, Duke Math Journal, 55 (1987), no. 2, 369–384.

[83] H. Ishii, *On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions*, Funkcial. Ekvac. 38 (1995), no. 1, 101–120.

[84] H. Ishii, *Almost periodic homogenization of Hamilton-Jacobi equations*, International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999), 600–605, World Sci. Publ., River Edge, NJ, 2000.

[85] H. Ishii, Lecture notes on the weak KAM theorem, lecture notes available on Ishii's website.

[86] H. Ishii, *Asymptotic solutions for large-time of Hamilton–Jacobi equations in Euclidean n space*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 25 (2008), no 2, 231–266.

[87] H. Ishii, H. Mitake, H. V. Tran, *The vanishing discount problem and viscosity Mather measures. Part 1: the problem on a torus*, Journal Math. Pures Appl., 108 (2017), no. 2, 125–149.

[88] H. Ishii, H. Mitake, H. V. Tran, *The vanishing discount problem and viscosity Mather measures. Part 2: boundary value problems*, Journal Math. Pures Appl., 108 (2017), no. 3, 261–305.

[89] R. Iturriaga, H. Sanchez-Morgado, *On the stochastic Aubry-Mather theory*, Bol. Soc. Mat. Mexicana (3) 11 (2005), no. 1, 91–99.

[90] R. Iturriaga, H. Sanchez-Morgado, *Limit of the infinite horizon discounted Hamilton–Jacobi equation*, Discrete Contin. Dyn. Syst. Ser. B, 15 (2011), 623–635.

[91] R. Jensen, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rat. Mech. Anal. 101 (1988), 1–27.

[92] W. Jing, H. V. Tran, Y. Yu, *Inverse problems, non-roundedness and flat pieces of the effective burning velocity from an inviscid quadratic Hamilton-Jacobi model*, Nonlinearity, 30 (2017) 1853–1875.

[93] W. Jing, H. V. Tran, Y. Yu, *Effective fronts of polytope shapes*, Minimax Theory and its Applications, 05 (2020), No. 2, 347–360.

[94] M. Jotz, *Hedlund metrics and the stable norm*, Differential Geometry and its Applications, Volume 27, Issue 4, August 2009, Pages 543–550.

[95] V. Kaloshin, *Mather theory, weak KAM, and viscosity solutions of Hamilton–Jacobi PDE's*, EQUADIFF 2003, 39–48, World Sci. Publ., Hackensack, NJ, 2005.

[96] S. Koike, A Beginner's Guide to the Theory of Viscosity Solutions, MSJ Memoir, 13, Math. Soc. Japan, Tokyo, 2004, viii+123 pp.

[97] H. Komiya, *Elementary proof for Sion's minimax theorem*, Kodai Mathematical Journal. 11 (1988) (1): 5–7.

[98] S. N. Kružkov, *Generalized solutions of nonlinear equations of the first order with several independent variables. II,* (Russian) Mat. Sb. (N.S.) 72 (114) 1967 108–134.

[99] J. M. Lasry and P. L. Lions, *A remark on regularization in Hilbert spaces*, Israel J. Math 55 (1986), 257–266.

[100] N. Q. Le, H. Mitake, H.V. Tran, Dynamical and Geometric Aspects of Hamilton-Jacobi and Linearized Monge-Ampère Equations, Lecture Notes in Mathematics 2183, Springer.

[101] P.-L. Lions, Generalized solutions of Hamilton-Jacobi equations, Research Notes in Mathematics, Vol. 69, Pitman, Boston, Masso. London, 1982.

[102] P.-L. Lions, G. Papanicolaou, S. R. S. Varadhan, *Homogenization of Hamilton–Jacobi equations*, unpublished work (1987).

[103] P.-L. Lions and J.-C. Rochet, *Hopf's formula and multitime Hamilton-Jacobi equations*, Proc. Amer. Math. Soc. 96 (1986), 79–84.

[104] P.-L. Lions, P. E. Souganidis, *Correctors for the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting*, Comm. Pure Appl. Math. 56 (2003), no. 10, 1501–1524.

[105] W. Littman, G. Stampacchia, H. F. Weinberger, *Regular points for elliptic equations with discontinuous coefficients*, Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 3e série, tome 17, no 1-2 (1963), p. 43–77.

[106] Y.-Y. Liu, J. Xin, Y. Yu, *Asymptotics for turbulent flame speeds of the viscous G-equation enhanced by cellular and shear flows*, Arch. Ration. Mech. Anal. 202 (2011), no. 2, 461–492.

[107] S. Luo, H. V. Tran, Y. Yu, *Some inverse problems in periodic homogenization of Hamilton-Jacobi equations*, Arch. Ration. Mech. Anal. 221 (2016), no. 3, 1585–1617.

[108] S. Luo, Y. Yu, H. Zhao, *A new approximation for effective Hamiltonians for homogenization of a class of Hamilton-Jacobi equations*, Multiscale Model. Simul. 9 (2011), no. 2, 711–734.

[109] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity 9 (1996), no. 2, 273–310.

[110] R. Mañé, *Lagrangian flows: the dynamics of globally minimizing orbits*, Bol. Soc. Brasil. Mat. (N.S.), 28(2):141–153, 1997.

[111] G. Markstein, Nonsteady Flame Propagation, Pergamon Press, Oxford, 1964.

[112] J. N. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. 207 (1991), no. 2, 169–207.

[113] J. N. Mather, *Differentiability of the minimal average action as a function of the rotation number*, Bol. Soc. Bras. Math 21 (1990), 59–70.

[114] A. Melikyan, Generalized characteristics of first order PDEs. Applications in optimal control and differential games. Birkhäuser Boston, Inc., Boston, MA, 1998. xiv+310 pp.

[115] H. Mitake, H. V. Tran, *Homogenization of weakly coupled systems of Hamilton–Jacobi equations with fast switching rates*, Arch. Ration. Mech. Anal. 211 (2014), no. 3, 733–769.

[116] H. Mitake, H. V. Tran, *Selection problems for a discount degenerate viscous Hamilton–Jacobi equation*, Adv. Math., 306 (2017), 684–703.

[117] H. Mitake, H. V. Tran, *On uniqueness sets of additive eigenvalue problems and applications*, Proc. Amer. Math. Soc., 146, no 11, 4813–4822.

[118] H. Mitake, H. V. Tran, Y. Yu, *Rate of convergence in periodic homogenization of Hamilton-Jacobi equations: the convex setting*, Arch. Ration. Mech. Anal., 233, 901–934 (2019).

[119] G. Namah, J.-M. Roquejoffre, *Remarks on the long time behaviour of the solutions of Hamilton–Jacobi equations*, Comm. Partial Differential Equations 24 (1999), no. 5-6, 883–893.

[120] A. M. Oberman, R. Takei and A. Vladimirsky, *Homogenization of metric Hamilton-Jacobi equations*, Multiscale Model. Simul. 8 (2009), pp. 269–295.

[121] S. Osher, R. Fedkiw, Level Set Methods and Dynamic Implicit Surfaces, Applied Mathematical Sciences 153, Springer-Verlag New York, 2003.

[122] S. Osher, J. Sethian, *Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations*. J. Comput. Phys. 79 (1988), no. 1, 12–49.

[123] J.-L. Qian, *Two Approximations for Effective Hamiltonians Arising from Homogenization of Hamilton-Jacobi Equations*, UCLA CAM report 03-39, University of California, Los Angeles, CA, 2003.

[124] J. Qian, H. V. Tran, Y. Yu, *Min-max formulas and other properties of certain classes of nonconvex effective Hamiltonians*, Math. Ann. (2018) 372: 91.

[125] M. Sion, *On general minimax theorems*, Pacific Journal of Mathematics. 8 (1958) (1): 171–176.

[126] G. Sivashinsky, *Cascade-renormalization theory of turbulent flame speed*, Combust. Sci. Tech., 62, 77–96, 1988.

[127] H. Soner, *Optimal control with state-space constraint I*, SIAM Journal on Control and Optimization 24, 3 (1986), 552–561.

[128] H. Soner, *Optimal control with state-space constraint. II*, SIAM Journal on Control and Optimization 24, 6 (1986), 1110–1122.

[129] A. Sorrentino, Action-Minimizing Methods in Hamiltonian Dynamics. An Introduction to Aubry-Mather Theory, Mathematical Notes Series Vol. 50 (Princeton University Press), 2015.

[130] P. E. Souganidis, *Approximation schemes for viscosity solutions of Hamilton-Jacobi equations*, J. Differential Equations 59 (1985), no. 1, 1–43.

[131] H. V. Tran, *Adjoint methods for static Hamilton-Jacobi equations*, Calc. Var. Partial Differential Equations 41 (2011), 301–319.

[132] H. V. Tran, Y. Yu, A rigidity result for effective Hamiltonians with 3-mode periodic potentials, *Advances in Math.*, 334, 300–321.

[133] S. N. T. Tu, *Rate of convergence for periodic homogenization of convex Hamilton-Jacobi equations in one dimension*, Asymptotic Analysis, vol. Pre-press, no. Pre-press, pp. 1–24, 2020.

[134] S. N. T. Tu, S. P. T. Van, *Homogenization of Hamilton-Jacobi equations: Notes and remarks*, available on the authors' websites.

[135] J. Xin, Y. Yu, *Periodic homogenization of the inviscid G-equation for incompressible flows*, Commun. Math. Sci., Vol. 8, No. 4 (2010), 1067–1078.

[136] V. Yakhot, *Propagation velocity of premixed turbulent flames*, Combust. Sci. Tech., 60, 191–241, 1988.

[137] Y. Yu, *A remark on the semi-classical measure from* $-\frac{h^2}{2}\Delta+V$ *with a degenerate potential* V, Proc. Amer. Math. Soc. 135 (2007), no. 5, 1449–1454.