Hamilton–Jacobi equations: viscosity solutions and applications

Hung Vinh Tran

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

First draft, April 2019

Contents

Preface

This is a first draft of my book of viscosity solutions and applications written in 2019. In this book, I intend to cover first the basic well-posedness theory of viscosity solutions. This is, by now, quite standard and there have been quite some great books on this matter since 1980s. 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 attention here to cover extensively about well-posedness of viscosity solutions for various different kinds of PDEs.

Then, I aim at discussing in deep the homogenization theory for Hamilton–Jacobi equations. Although this has always been a very active research topic since 1980s until this moment (2019), 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.

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.

I intend to keep the contents of various topics covered here as independent as possible so that interested readers are able to jump directly to a subject of interests in the book. Of course, the book can be used as a learning tool for new researchers in the field as well. In the latter case, the readers can follow the flow of the book from the beginning until the topics that the readers aim at.

My intention here 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.

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 first draft.

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.

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

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

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

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 (level set method).

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

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

- *The simplest example is Γ_0 is the unit sphere, and every point is moving inward with constant (vector) speed 1, then Γ_t is remain a sphere for $t \in [0, 1)$, and eventually shrinks into a point at $t = 1$, located at the center.*
- *If each point on the surface Γ_t is moving with variable velocity, then the situation becomes more complicated. Osher, Sethian [86] introduced the level set method (numerically) to study this problem. The rigorous treatment was developed later by Evans, Spruck [37] and Chen, Giga, Goto [20], independently.*

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

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

We 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 $V(x) = a(x)\mathbf{n}$, where $\mathbf{n}$ is the inward normal vector to Γ_t at x. Let us then try to find a PDE for $u(x, t)$ based on this given law of motions.

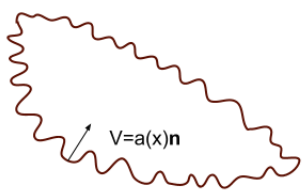


Figure 1.1: Front propagation of $\{\Gamma_t\}_{t\geq 0}$.

For a particle $x(0) \in \Gamma_0$, 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

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

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

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

which implies

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

Thus, we obtain a PDE $u_t+a(x)|Du|=0$*, which is a first-order Hamilton–Jacobi equation.*

Example 1.2 (Level set mean curvature flow)**.** *Let* $\kappa(x)$ *be the mean curvature at* $x\in\Gamma_t$ *of the surface* Γ_t*. For example, if* Γ_t *is a sphere of radius* $R(t)$*, 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,*

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

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

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

where

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

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 both above examples. From the PDE viewpoints, we focus on the following main issues

1. Well-posedness theory: Existence, uniqueness and stability of solutions;
2. Study fine properties of solutions such as large time behavior, homogenization, dynamical properties.

Example 1.3 (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 need 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

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

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

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

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 processes to a nonlinear PDE.

Evans [29] first showed that this procedure 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 [25] 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, thus it has 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 part technique to define weak solutions;

- There is a priori estimate for $\{u^\varepsilon\}_{\varepsilon>0}$: There exists a constant $C > 0$ independent of ε 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, $\{u^\varepsilon(x,t)\}_{\varepsilon>0}$ is equi-continuous and thus by the Arzelà-Ascoli theorem, there exists $\varepsilon_j \searrow 0$ such that $u^{\varepsilon_j} \to u$ locally uniformly on $\mathbb{R}^n \times [0,\infty)$. We hence hope that u solves (1.1) naturally in some sense that fits well with the context of maximum principle.

Let us now analyze further along this line for possible definition of weak solution 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 $u^\varepsilon \longrightarrow u$ 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 $(x_\varepsilon,t_\varepsilon) \longrightarrow (x_0,t_0)$ 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

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

For 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 [23].

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 definitions in geometric terms.

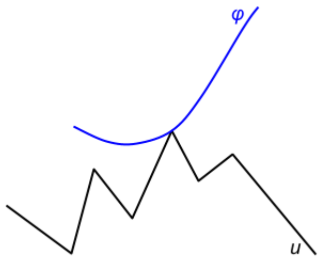


Figure 1.2: An illustration of φ touches u from above at (x_0,t_0).

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 a.e. solutions that we got are mutually viscosity subsolutions.

Exercise 2. *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 3. *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).*

The exercises 2–3 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.

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

Theorem 1.4. *Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \longrightarrow \mathbb{R}$ 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 to sup-differential/min.*

Proof. We only need to prove "⇒". 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 $u(y) \le u(x) + p\cdot(y-x)$ 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 $\sigma : (0,\delta] \longrightarrow \mathbb{R}$ 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,

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

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

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

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

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

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

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

Here, $\Omega \subset \mathbb{R}^n$ is a given open set, and $u : \Omega \to \mathbb{R}$ is a 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*

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

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

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

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*

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

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 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, $D^+f(x) = D^-f(x) = \{Df(x)\}$.

(e) 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. By (c), 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.
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 result on u (e.g., u is Lipschitz in Ω), and we will discuss this situation more later.

2.4 Problems

Exercise 4. *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 5. *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 6. *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 $(y,p) \in \Omega \times \mathbb{R}^n$.

3 Existence of viscosity solutions via the vanishing viscosity method

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

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

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 $\varepsilon \to 0+$. Prove that 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

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

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

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

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

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

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 independent of ε. 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*

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

Assume that (1.8) *has a unique solution u^ε for any $\varepsilon > 0$. Furthermore, we assume that there exists a constant $C > 0$ independent of ε such that*

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

Then, 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 $u \in C(\mathbb{R}^n \times [0,\infty))$. Then u is a viscosity solution of (1.7).

Proof. We show that u is a viscosity subsolution of (C). The viscosity supersolution test is similar, hence omitted. By Exercise 2, 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 obtain a subsequence $\{\varepsilon_i\} \searrow 0$ such that $u^{\varepsilon_i} - \varphi$ has a local max at (x_i,t_i) 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

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

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

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

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

Remark 1.10.

1. If $u - \varphi$ has a strict max at $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$, 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 priori estimate

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

   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

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

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 4).

Lemma 1.11. *Let Ω be an open subset of $\mathbb{R}^n$, and $u : \Omega \longrightarrow \mathbb{R}$ 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 $x \in \Omega$,

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

□

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*

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

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

$$\begin{cases} (u_k)_t + H_k(x, Du_k) &= 0 \quad \text{in } \mathbb{R}^n \times (0,\infty), \\ u_k(x,0) &= u_0(x) \quad \text{on } \mathbb{R}^n. \end{cases} \tag{1.11}$$

Then u is a viscosity solution to (1.10).

Proof. It is clear that u satisfies the initial condition in the classical sense. We show that u is a viscosity subsolution to (1.10). The supersolution 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 $(x_k,t_k)\longrightarrow(x_0,t_0)$ 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 $k\longrightarrow\infty$ 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 $x,y,p,q\in\mathbb{R}^n$ then:

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

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, $u(x)\le v(x)$ for any $x\in\mathbb{R}^n$.*

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

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

we consider the following auxiliary function

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

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

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

Recall that (1.14) gives

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

Plug it into (1.18) to yield

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

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

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

which is a contradiction. Thus, the proof is complete. □

Corollary 1.15 (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 1.16. 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.

6 The comparison principle and uniqueness result for Cauchy problem

We consider the following usual Cauchy problem

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

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

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

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

Lemma 1.17 (Extrema at terminal time $t = T$)**.** *Let u be a viscosity subsolution to* (1.22) *and $\varphi \in C^1(\mathbb{R}^n \times [0,\infty))$ such that $u - \varphi$ has a strict max at (x_0, t_0) over $(x,t) \in \mathbb{R}^n \times (0, T]$, then the subsolution test still holds, that is,*

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

Proof. It suffices to only consider the case $t_0 = T$. Define $\varphi_\varepsilon(x,t) = \varphi(x,t) + \frac{\varepsilon}{T-t}$ for any fixed $\varepsilon > 0$. Then for $\varepsilon > 0$ is small enough, $u - \varphi_\varepsilon$ has a local max at $(x_\varepsilon, t_\varepsilon)$, and $(x_\varepsilon, t_\varepsilon) \to (t_0, t_0)$ as $\varepsilon \to 0$ by passing to a subsequence if necessary (see Exercise 7 below for confirmation). As $u - \varphi_\varepsilon$ has a local max at $(x_\varepsilon, t_\varepsilon)$, by definition of viscosity subsolution, we have

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

which means

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

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

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

Theorem 1.18 (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, $u(x,t) \leq v(x,t)$ 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 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

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

STEP 3. Next, we claim that there exists a constant $\mu > 0$ independent to ε such that $t_\varepsilon, s_\varepsilon > \mu > 0$ for all $\varepsilon > 0$. It is important to have both $t_\varepsilon, s_\varepsilon$ bounded away from 0 in order to apply viscosity sub/supersolution tests.
To prove this claim, we use the uniform continuity of u, v and observe

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

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

STEP 4. The map $(x, t) \mapsto \Phi(x, y_\varepsilon, t, s_\varepsilon)$ has a max as $(x_\varepsilon, t_\varepsilon)$, 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

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

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

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

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

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

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

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

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

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

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

Corollary 1.19 (Uniqueness of viscosity solution of Cauchy problem (1.22))**.** *Assume* (1.13). *If u and v are viscosity solution 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 1.18. $\square$

6.1 Problems

Exercise 7. *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 $(x_\varepsilon, t_\varepsilon) \longrightarrow (x_0, T)$ up to a subsequence.*

Exercise 8. *Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ 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 (C) *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 the priori estimate 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 ε. 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

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

Theorem 1.20. *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_t^ε. Differentiate (1.25) in time,

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

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. Differentiate (1.25) in x_k, multiply the result by $u_{x_k}^\varepsilon$, then sum them up over $k=1,2,\ldots,n$ to obtain

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

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,

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

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

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

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

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

which implies that, at (x_0,t_0),

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

in light of assumption (1.27).

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

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

□

Remark 1.21. 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).

7.1 Problems

Exercise 9. *Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 Hamiltonian satisfying*

$$\begin{cases} \|H(\cdot,0)\|_{L^\infty(\mathbb{R}^n)} < +\infty, \\ \lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} \left(\frac{1}{2}H(x,p)^2 + D_xH(x,p)\cdot p\right) = +\infty. \end{cases} \quad (1.30)$$

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

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

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

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 1.22. *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 problem

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 1.23. *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 1.23 still holds true for u under the new definition, that is, u^*, its upper semicontinuous envelope, is a viscosity subsolution to (1.32).

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

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

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

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

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

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

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

The Perron method in the theory of viscosity solutions was first introduced by Ishii [58]. In the following, we give a variant of Ishii's argument in [58]. 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

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

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

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

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 1.23), 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

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

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

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

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

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

Now, set

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

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

□

As included in the proof, we obtain immediately the existence of a Lipschitz viscosity solution u to (1.32) under assumption (1.33). In fact, by Remark 1.16, 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 1.25. *Assume* (1.33). *Let u be defined as in Theorem 1.24. 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 1.26. *Assume* (1.33). *Let $u \in \mathrm{BUC}(\mathbb{R}^n)$ be a solution to* (1.32). *Then, u is Lipschitz in $\mathbb{R}^n$.*

Proof. As $u \in \mathrm{BUC}(\mathbb{R}^n)$, it is not hard to show that $-C_0 \le u \le C_0$ (this is being phrased as Exercise 10). By coercivity and the viscosity subsolution test, we get

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

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

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

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

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

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

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

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

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

Lemma 1.27. *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 10. *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*

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

Exercise 11. *Prove Lemma 1.27.*

8.3 Perron's method for Cauchy problem

Let us now focus on our usual Cauchy problem

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

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

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

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

- 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 1.28 (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 1.24. 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 1.29. Let us emphasize again that u defined in Theorem 1.28 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 $u^* \leq u_*$. Therefore, $u = u^* = u_*$, which means that u is continuous.

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

9 Lipschitz estimates for Cauchy problem 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

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

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

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

Theorem 1.30. *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*

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

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 1.28, 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 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 1.31. 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 $\|u_t\|_{L^\infty} \leq C_0$, which yields $H(x,Du) \leq C_0$. Thus, we are able to find $C > 0$ such that $\|u_t\|_{L^\infty} + \|Du\|_{L^\infty} \leq C$. In particular, information of $H(x,p)$ for $|p| \geq C$ does not matter. Define a new Hamiltonian $\tilde{H}$ such that

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

  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 $\|w_t\|_{L^\infty} + \|Dw\|_{L^\infty} \leq C$. It is clear then that w is a Lipschitz viscosity solution to (1.38). Then, Remark 1.16 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 1.32. *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*

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

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 12. *Give a detailed proof of Step 3 in the proof of Theorem 1.30.*

Exercise 13. *Write down a proof of Theorem 1.32.*

10 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

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

For each $\varepsilon > 0$, we consider

$$u^\varepsilon(x) + H(x, Du^\varepsilon(x)) = \varepsilon \Delta u^\varepsilon(x) \qquad \text{in } \mathbb{R}^n. \tag{1.42}$$

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 1.20) to obtain that (1.42) 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 (1.41). 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 (1.41). 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 1.33. *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* (1.42). *Let u be the unique Lipschitz viscosity solution of* (1.41). *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 [42] in the 1960s by using a differential game approach. Later on, within the framework of viscosity solutions, Crandall and Lions [26] proved Theorem 1.33 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 1.33 by using the nonlinear adjoint method was introduced by Evans [33] and Tran [91].

We give here in this section a proof based on the ideas of Crandall, Lions [26]. 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]

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

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

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

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

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

$$u^\varepsilon(x_\delta) + H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \leq \varepsilon\left(\frac{n}{\alpha} + \delta\Delta\mu(x_\delta)\right) \leq \varepsilon\left(\frac{n}{\alpha} + \delta\right). \tag{1.43}$$

STEP 2. As $y \mapsto \Phi^\delta(x_\delta, y)$ has a max at y_δ, $y \mapsto u(y) - \left[-\frac{|x_\delta - y|^2}{2\alpha} - \delta\mu(y)\right]$ has a min at y_δ. The supersolution test for (1.42) gives

$$u(y_\delta) + H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) \geq 0. \tag{1.44}$$

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

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

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

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

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

  Combine the above two inequalities to get

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

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

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

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

[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.

Thus,

$$H\left(y_\delta, \frac{x_\delta - y_\delta}{\alpha} - \delta D\mu(y_\delta)\right) - H\left(x_\delta, \frac{x_\delta - y_\delta}{\alpha} + \delta D\mu(x_\delta)\right) \le C\alpha + C\delta. \tag{1.45}$$

STEP 5. Combine the inequalities in (1.43), (1.44), and (1.45) to imply

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

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 (1.46). 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 1.34. 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

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

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

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

We consider the same situation like in the previous section. We are interested in the vanishing viscosity procedure for the usual static problem

$$u(x) + H(x, Du(x)) = 0 \qquad \text{in } \mathbb{R}^n. \tag{1.47}$$

For each $\varepsilon > 0$, we consider

$$u^\varepsilon(x) + H(x, Du^\varepsilon(x)) = \varepsilon \Delta u^\varepsilon(x) \qquad \text{in } \mathbb{R}^n. \tag{1.48}$$

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)

$$\begin{cases} H \in C^2(\mathbb{R}^n \times \mathbb{R}^n), H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) & \text{for each } R > 0, \\ |D_x H(x,p)| \le C(1 + |p|) & \text{for all } (x,p) \in \mathbb{R}^n \times \mathbb{R}^n, \\ \lim_{|p|\to\infty} \frac{H(x,p)}{|p|} = \infty & \text{uniformly for } x \in \mathbb{R}^n, \end{cases} \tag{1.49}$$

for some given $C > 0$.

Then, by Bernstein's method, (1.48) 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 (1.48), we have $F^\varepsilon[u^\varepsilon] = 0$. The linearized operator $\mathcal{L}^\varepsilon$ of F^ε about the solution u^ε is defined as, for $\varphi \in C_c^\infty(\mathbb{R}^n)$,

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

which gives

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

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

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

By integration by parts,

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

Thus,

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

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

$$\sigma^\varepsilon - \operatorname{div}(D_pH(x, Du^\varepsilon)\sigma^\varepsilon) - \varepsilon\Delta\sigma^\varepsilon = \delta_{x_0} \qquad \text{in } \mathbb{R}^n. \tag{1.50}$$

Here, δ_{x_0} is the Dirac delta at x_0. Let σ^ε be the unique solution to (1.50), 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 (1.50), introduced by Evans [33], Tran [91], 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 (1.48), and then integrating by parts with σ^ε.

Lemma 1.35. *Assume* (1.49). *For each $\varepsilon \in (0,1)$, let u^ε be the unique smooth solution to* (1.48)*, and let σ^ε be the unique solution to* (1.50) *for fixed $x_0 \in \mathbb{R}^n$. Then, there exists a constant C independent of ε such that*

$$\varepsilon \int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 \sigma^\varepsilon \, dx \le C. \tag{1.51}$$

Proof. Let $\varphi = \frac{1}{2}|Du^\varepsilon|^2$. By doing computations similar to these in the classical Bernstein method, we obtain from (1.48) that

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

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

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

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

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

Using the adjoint equation, we obtain

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

Thus,

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

Thus the proof is complete. ☐

Remark 1.36. It is important noting that (1.51) is one of the new key estimates in the development of the nonlinear adjoint method. Originally, if we look at (1.48), 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 (1.51) gives better control in the support of σ^ε, where we have, roughly speaking, $|D^2 u^\varepsilon| \le O(\frac{1}{\sqrt{\varepsilon}})$. This turns out to be quite useful in various situations.

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

Theorem 1.37. *Assume that H satisfies* (1.49). *For each $\varepsilon \in (0,1)$, let u^ε be the unique smooth solution to* (1.48). *Let u be the unique Lipschitz viscosity solution of* (1.47). *Then, there exists a constant $C > 0$ independent of ε such that*

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

Proof. We have that $\varepsilon \mapsto u^\varepsilon$ is smooth for $\varepsilon > 0$. Let us differentiate (1.48) 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 1.35 and Hölder's inequality, we obtain

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

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. □

11.1 Problems

Exercise 14. *In the general nonconvex setting, is the convergence rate $O(\sqrt{\varepsilon})$ of u^ε to u in Theorem 1.37 optimal?*

12 References

1. There have been many great textbooks in the study of viscosity solutions for Hamilton–Jacobi equations written by Bardi and Capuzzo-Dolcetta [9], Barles [10], Cannarsa, Sinestrari [17], Chapter 10 of Evans [32], Fabbri, Gozzi, Swiech [38], Fleming and Soner [43], Isaacs [57], Koike [69], Lions [73], Melikyan [80]. Besides, the user's guide written by Crandall, Ishii, and Lions [24] 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 [13], Calder [16], Le, Mitake, Tran [72].

3. The level set method was first introduced numerically by Osher, Sethian [86]. The rigorous treatment was developed later by Evans, Spruck [37] and Chen, Giga, Goto [20], independently. See the textbook of Giga [46] and the references therein for the developments of this direction.

4. Evans [29] first used the Minty trick to study the vanishing viscosity method and gave first definitions of possibly weak solutions. Crandall and Lions [25] 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 [71] in scalar conservation laws. Crandall and Lions chose the name "viscosity solutions" in honor of the vanishing viscosity technique.

5. Ishii [58] 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.

6. The nonlinear adjoint method was introduced first by Evans [33] to study the gradient shock structures of Cauchy problem for nonconvex Hamiltonians. The static cases were studied by Tran [91]. Recently, this method has been developed much further to study large time behaviors, selection problems, and dynamical properties of solutions to Hamilton–Jacobi equations. For this, see the lecture notes by Le, Mitake, Tran [72].

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 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. 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,

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

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. □

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

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

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

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

Thus,

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

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 to emphasize a difference between PDE and dynamical system viewpoints here.

- Dynamical system viewpoint: to 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

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

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

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

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

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

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

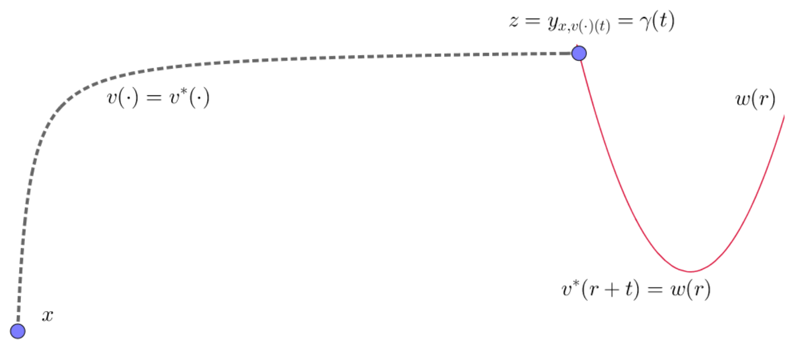


Figure 2.1: Connecting two controls $v(\cdot)$ and $w(\cdot)$ to form a new control $v^*(\cdot)$.

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

and

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

Thus, by combining these facts, we obtain

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

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. □

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, $\|u\|_{L^\infty(\mathbb{R}^n)} \le \frac{C}{\lambda}$. Furthermore, we have the following results.*

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

(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 15. *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 enable 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 $\psi \in C^2(\mathbb{R}^n)$ 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 $t > 0$,

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

Therefore,

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

where

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

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

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

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

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

and similarly,

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

as well. Hence,

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

Combine this with the above to deduce that

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

The proof is complete. □

4 Legendre's transform

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

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

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 $m\geq 1$.*

- *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.*

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*

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

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 $v \neq 0$, chose $p = s\frac{v}{|v|}$, 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 $s > 0$,

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

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

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

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 Alexandroff's theorem, for each $x \in \mathbb{R}^n$, $H(x,\cdot)$ is twice differentiable almost everywhere.

4.1 Problems

Exercise 16. *Compute the Legendre's 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 $V \in \mathrm{BUC}(\mathbb{R}^n)$.

Exercise 17. *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

$$\begin{cases} p \mapsto H(x,p) \text{ is convex,} \\ H(x,p) \text{ is superlinear in } p, \end{cases} \qquad \overset{\text{Legendre's transfrom}}{\longleftrightarrow} \qquad \begin{cases} v \mapsto L(x,v) \text{ is convex,} \\ L(x,v) \text{ is superlinear in } v. \end{cases}.$$

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

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

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 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 minimizer 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

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

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's transform of H, that is,

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

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

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

Besides, there exists $C > 0$ such that

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

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

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

□

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

Definition 2.18. *For each $(x,t) \in \mathbb{R}^n \times [0,\infty)$, we denote by*

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

Remark 2.19. 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 2.20 (Dynamic Programming Principle)**.** *The value function u defined above satisfies, for $(x,t) \in \mathbb{R}^n \times (0,\infty)$,*

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

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

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

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

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

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

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

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,

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

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 2.21. *Assume* (2.13). *Let u be defined as in* (2.16). *Using the Dynamic Programming Principle* (2.17) *to prove that 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. It is in fact an interesting exercise for interested readers.

5.3 Problems

Exercise 18. *Prove Theorem 2.21.*

5.4 The Hopf–Lax formula

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

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

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

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

Theorem 2.22 (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 $y\in\mathbb{R}^n$. □

Example 2.4. *We give here some famous 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*

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

Here, initial data v_0 is nice enough. Note that

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

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.

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

$$\lambda u + H(x, Du) = 0 \qquad \text{in } \mathbb{R}^n. \tag{2.20}$$

We assume throughout this section that

$$\begin{cases} H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R)) \text{ for all } R > 0, \\ p \mapsto H(x,p) \text{ is convex for all } x \in \mathbb{R}^n, \\ \lim\limits_{|p|\to\infty} \inf\limits_{x\in\mathbb{R}^n} H(x,p) = +\infty. \end{cases} \tag{2.21}$$

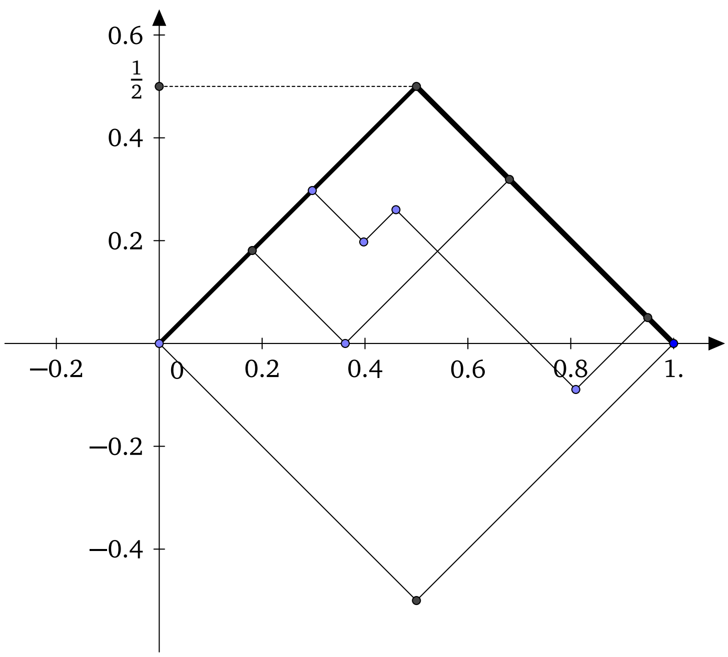
Remark 2.23. Under (2.21), for $\lambda > 0$, we apply the Perron method (Theorem 1.24) to imply that (2.20) has a unique Lipschitz viscosity solution $u \in \mathrm{Lip}(\mathbb{R}^n)$. Of course, this means that u is differentiable a.e. in $\mathbb{R}^n$.
If $\lambda = 0$, (2.20) is not monotone in u anymore, and anything can happen. For example, if $H(x,p) > 0$ for all $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$, then (2.20) does not have any solution. It could be also the case that (2.20) has infinitely many solutions, and we will discuss this later in this book.

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

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

$$\begin{cases} |u'(x)| &= 1 \qquad \textit{in } (0,1), \\ u(0) = u(1) &= 0. \end{cases} \tag{2.22}$$

Of course, here, $H(p) = |p|$, and $\lambda = 0$. The following graph describes various a.e. solutions to (2.22). *As discussed in Exercise 1, all 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 [11].

Theorem 2.24. *Assume $\lambda \geq 0$, and H satisfies* (2.21). *Then, the following claims are equivalent*

(i) $u \in \mathrm{Lip}(\mathbb{R}^n)$ *is viscosity subsolution of* (2.20).

(ii) $u \in \mathrm{Lip}(\mathbb{R}^n)$ *is an almost everywhere subsolution of* (2.20).

Proof. The implication (i) ⇒ (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,

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

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 (2.20), we multiply η_ε to both sides of (2.20) 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 the classical (smooth) subsolution to

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

We then let $\varepsilon \to 0$ and use stability results of viscosity subsolutions to conclude that u is a subsolution of (2.20). □

Remark 2.25. We have some further observations.

- Normal convolution 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 (2.20).

Theorem 2.26. *Assume $\lambda \ge 0$, and H satisfies* (2.21). *Then, the following claims are equivalent*

(i) $u \in \mathrm{Lip}(\mathbb{R}^n)$ *is viscosity solution of* (2.20).

(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 (2.20). For $x \in \mathbb{R}^n$ and $p \in D^-u(x)$, by supersolution test, $\lambda u(x) + H(x, p) \ge 0$. We need to show that $\lambda u(x) + H(x, p) = 0$.

As u is a Lipschitz a.e. solution of (2.20), by Rademacher's theorem, for $v = -u$, we have

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

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 (2.21). Theorem 2.24 with $\lambda = 0$ concludes that v is a viscosity subsolution to $K(x, Dv(x)) = 0$. The viscosity subsolution test implies

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

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 (2.20). By Rademacher's theorem again, u is differentiable a.e. in $\mathbb{R}^n$, and thus

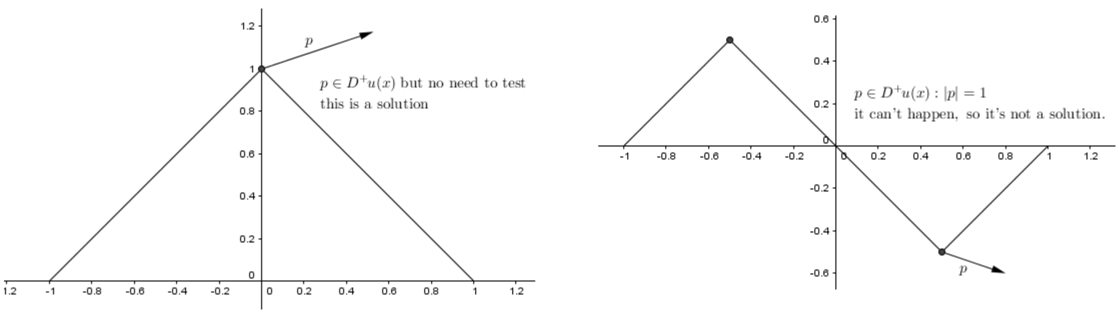
$$\lambda u(x) + H(x, Du(x)) = 0 \qquad \text{for a.e } x \in \mathbb{R}^n.$$

Theorem 2.24 implies automatically that u is a viscosity subsolution of (2.20), and the proof is complete. □

Remark 2.27. 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 $x \in \mathbb{R}^n$.
2. Criterion (ii) in Theorem 2.26 is quite important and useful. For example, we can use it to study the eikonal equation in one dimension again as following.

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



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

3. Theorems 2.24 and 2.26 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 use later.

Corollary 2.28. *Assume $\lambda \geq 0$, and H satisfies* (2.21). *Then, the followings hold.*

(i) If u_1, u_2 are Lipschitz solutions to (2.20)*, then* $\min\{u_1, u_2\}$ *is also a solution to* (2.20).

(ii) If $\{u_i\}_{i\in I}$ is a family of Lipschitz solutions to (2.20)*, then $u = \inf_{i\in I} u_i$ is also a solution to* (2.20) *provided u is finite.*

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 (2.20).

6.3 Problems

Exercise 19. *Prove Corollary 2.28.*

Exercise 20. *Show that the results of Theorems 2.24 and 2.26 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 21. *Consider the following viscous Hamilton-Jacobi equation in one dimensional space*

$$|u'|^3 - u'' - 1 = 0 \qquad \text{in } \mathbb{R}. \tag{2.23}$$

Clearly, $u_1(x) = x$ *and* $u_2(x) = -x$ *are two classical subsolutions of* (2.23). *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* (2.23). *Show however that* u_3 *is not a subsolution of* (2.23).

7 References

1. The optimal control theory part can be found in many references. For example, the readers can consult the book of Bardi, Capuzzo-Dolcetta [9], or Chapter 10 of Evans [32], or the book of Lions [73]. Lions [73] 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 also discussed in deep in Chapter 3 of Evans [32].

3. Theorems 2.24 and 2.26 are due to Barron, Jensen [11].

CHAPTER 3

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

$$\begin{cases} u^\varepsilon_t(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{3.1}$$

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 (3.1) 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. 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 (3.1), 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

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 so far since the equations are highly nonlinear.

1.2 Derivations

Our focus is on equation (3.1) 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.

$$y \mapsto H(y,p) \text{ is } \mathbb{Z}^n\text{-periodic, that is, } H(y,p) = H(y+k,p) \text{ for } k \in \mathbb{Z}^n, \tag{3.2}$$

and

$$\lim_{|p|\to\infty} H(y,p) = +\infty \text{ uniformly for } y \in \mathbb{R}^n. \tag{3.3}$$

We can think about our current problem (3.1) as a multi-scale problem

- x is the macroscopic scale or low scale.
- $y = \frac{x}{\varepsilon}$ is the microscopic scale or fast scale.

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 3.1. *Consider again the classical mechanics Hamiltonian*

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

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

$$\begin{cases} u_t^\varepsilon + \frac{1}{2}|Du^\varepsilon|^2 + V\left(\frac{x}{\varepsilon}\right) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty) \\ u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases}$$

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 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.*

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 (3.1) to get

$$u_t^0 + O(\varepsilon) + H\left(y, D_x u^0 + \frac{1}{\varepsilon} D_y u^0 + \varepsilon D_x u^1 + D_y u^1 + O(\varepsilon)\right) = 0. \tag{3.4}$$

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, (3.4) 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 (3.4) 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

$$u_t^0(x,t) + H\left(y, D_x u^0(x,t) + D_y u^1(x,y,t)\right) = 0.$$

Since $D_x u^1, u_t^1$ 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

$$H\left(y, D_x u^0(x,t) + D_y u^1(y)\right) = -u_t^0(x,t).$$

[1] *An ansatz* means *a formulation* or *an educated guess*

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

$$H\left(y, \underbrace{D_x u^0(x,t)}_{p\in\mathbb{R}^n} + D_y u^1(y)\right) = \underbrace{-u_t^0(x,t)}_{c\in\mathbb{R}} \qquad \text{in } \mathbb{R}^n.$$

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

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

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 3.1. The above derivation also works well for the general degenerate viscous Hamilton–Jacobi equation

$$w_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Dw^\varepsilon\right) = \varepsilon \operatorname{tr}\left(A\left(\frac{x}{\varepsilon}\right) D^2 w^\varepsilon\right) \qquad \text{in } \mathbb{R}^n \times (0,\infty).$$

Here, H satisfies (3.2) and (3.3). 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

$$H(y, p + Dw^1(y)) - \operatorname{tr}(A(y) D^2 w^1(y)) = c \qquad \text{in } \mathbb{R}^n.$$

We will revisit this second-order case later.

2 Cell problems and periodic homogenization of static Hamilton–Jacobi equations

Recall that, for (3.1), 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. Then,

$$u^0_t(x,t) + H\left(y, Du^0(x,t) + Du^1(y)\right) = 0 \qquad \text{in } \mathbb{R}^n \times [0,\infty).$$

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 $-c = u^0_t(x,t) \in \mathbb{R}$. 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 $p \in \mathbb{R}^n$.

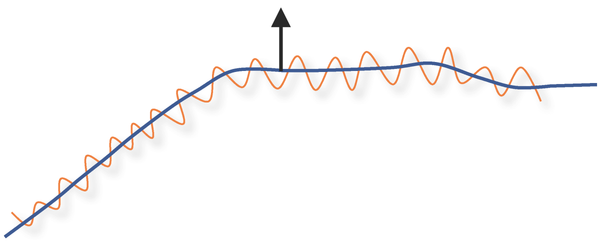


Figure 3.1: An example of graphs of u^ε and $u = u^0$ near (x,t)

2.1 Cell problems

In this section, we discuss the cell problems, which were studied first by Lions, Papanicolaou, and Varadhan [74].

Theorem 3.2. *Assume that H satisfies* (3.2) *and* (3.3). *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 3.3. *Assume that H satisfies* (3.2) *and* (3.3). *For each $p \in \mathbb{R}^n$, Theorem 3.2 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 3.2. For $\lambda > 0$, we consider the static equation

$$\lambda v^\lambda + H(y, p + Dv^\lambda) = 0 \qquad \text{in} \qquad \mathbb{R}^n. \tag{3.5}$$

By Theorem 1.25, we get that there exists a unique viscosity solution $v^\lambda \in \mathrm{Lip}(\mathbb{R}^n)$ of (3.5). 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 (3.5), hence $v^\lambda(y+k) = v^\lambda(y)$ for all $k \in \mathbb{Z}^n$ by uniqueness of (3.5). 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 (3.5), respectively, thus by the comparison principle, we have

$$\sup_{y\in\mathbb{T}^n} |\lambda v^\lambda(y)| \le C_0.$$

Plug it into (3.5) again, recall that $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ thus it is differentiable a.e., then in the a.e sense (3.5) becomes

$$|H(y, p + Dv^\lambda(y))| \le C_0 \qquad \text{for a.e. } y \in \mathbb{T}^n.$$

By coercivity of H we deduce that $\|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \le C_1$ independent of $\lambda > 0$. Note that the above estimates were already in Theorem 1.24. 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 $\sqrt{n}$,

$$\|w^\lambda\|_{L^\infty(\mathbb{T}^n)} \le \sqrt{n}\, \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \le C, \quad \text{and} \quad \|Dw^\lambda\|_{L^\infty(\mathbb{T}^n)} = \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \le C.$$

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

$$\begin{cases} w^{\lambda_j} = v^{\lambda_j}(\cdot) - v^{\lambda_j}(0) \to v(\cdot) & \text{uniformly on } \mathbb{T}^n, \\ \lambda v^{\lambda_j}(0) \to -c \in \mathbb{R} \end{cases}$$

for some $c \in \mathbb{R}$. It is clear that $\min_{\mathbb{T}^n} v = 0$ and $\|Dv\|_{L^\infty(\mathbb{T}^n)} \le C$. 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

$$H(y, p + Dv(y)) = c \qquad \text{in } \mathbb{T}^n. \tag{3.6}$$

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 1.26. Since v_1, v_2 are bounded in $\mathbb{T}^n$, we can find $\delta > 0$ sufficiently small such that[2]

$$\delta v_1(y) + H(y, p + Dv_1(y)) < \frac{c_1 + c_2}{2} < \delta v_1(y) + H(y, p + Dv_2(y)) \qquad \text{in } \mathbb{T}^n.$$

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 (3.6) 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 3.4. 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

$$\lambda v^\lambda(\cdot) \to -\overline{H}(p) \qquad \text{uniformly in } \mathbb{T}^n \text{ as } \lambda \longrightarrow 0.$$

   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 $v^\lambda(\cdot) - v^\lambda(x_0) \longrightarrow v(\cdot)$ 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 22. *Assume that H satisfies* (3.2) *and* (3.3). *Fix $p \in \mathbb{R}^n$, and we look at* (3.5). *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 23. *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* (3.2) *and* (3.3). *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 3.2. Recall the discounted problem (3.5), which can be viewed in terms of $y = \frac{x}{\lambda}$ as

$$\lambda v^\lambda\left(\frac{x}{\lambda}\right) + H\left(\frac{x}{\lambda}, p + Dv^\lambda\left(\frac{x}{\lambda}\right)\right) = 0 \quad \text{in } \mathbb{R}^n.$$

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

$$u^\lambda(x) + H\left(\frac{x}{\lambda}, p + Du^\lambda(x)\right) = 0 \qquad \text{in } \mathbb{R}^n. \tag{3.7}$$

Clearly, (3.7) 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 3.5. *Assume that H satisfies* (3.2) *and* (3.3). *Fix $p \in \mathbb{R}^n$, and we study the homogenization problem* (3.7). *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 (3.7) *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 [74], and Evans [31].

Theorem 3.6. *Assume that H satisfies* (3.2) *and* (3.3). *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*

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{3.8}$$

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

$$\begin{cases} u_t + \overline{H}(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{3.9}$$

We here introduce the perturbed test function method of Evans [31] 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

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n. \tag{3.10}$$

To make it clear the dependences, sometimes, we write $v = v(y, p)$, and clearly, 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 (3.10) are smooth. We will also see why the perturbed test function is needed.

A heuristic proof of Theorem 3.6. As usual, we break this heuristic proof into few steps.

1. We first obtain some a priori estimates for u^ε. By Theorem 1.32, 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 (3.9).

   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

   $$\varphi_t\left(x_j, t_j\right) + H\left(\frac{x_j}{\varepsilon_j}, D\varphi\left(x_j, t_j\right)\right) \le 0.$$

As $j \to \infty$ we have $\varphi_t\left(x_j, t_j\right) \to \varphi_t(x_0, t_0)$, but we do not have information about the second term $H\left(\frac{x_j}{\varepsilon_j}, D\varphi\left(x_j, t_j\right)\right)$ 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 (3.10) 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,

$$\psi_t\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) + H\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}, D\varphi\left(x_{\varepsilon_j}, t_{\varepsilon_j}\right) + Dv\left(\frac{x_{\varepsilon_j}}{\varepsilon_j}\right)\right) \leq 0. \tag{3.11}$$

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 (3.11) to conclude. The viscosity supersolution test follows in a similar way.

3. As $\overline{H}$ is continuous and coercive, (3.9) 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 3.7. 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 (3.9) 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 3.6. 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 (3.9).

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

$$\phi_t(x_0,t_0) + \overline{H}(D\phi(x_0,t_0)) \le 0.$$

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 (3.10) 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 $\Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, \frac{x_{\eta\varepsilon}}{\varepsilon}, t_{\eta\varepsilon}\right) \le \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$ for all $\eta > 0$, we obtain

$$\frac{1}{\eta}\left| y_{\eta\varepsilon} - \frac{x_{\eta\varepsilon}}{\varepsilon} \right|^2 \le 2\varepsilon \|v\|_{L^\infty(\mathbb{T}^n)} \qquad \Longrightarrow \qquad \lim_{\eta \to 0} y_{\eta\varepsilon} = \frac{x_\varepsilon}{\varepsilon}. \tag{3.12}$$

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 (3.8) gives

$$\phi_t\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + H\left(\frac{x_{\eta\varepsilon}}{\varepsilon}, D\phi\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon} \right) \right) \le 0. \tag{3.13}$$

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

$$-\overline{H}(p) + H\left(y_{\eta\varepsilon}, p + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon} \right) \right) \ge 0. \tag{3.14}$$

   Besides, as v is Lipschitz, we get

$$\left| \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon} \right) \right| \le C, \tag{3.15}$$

for some $C>0$ independent of η,ε. By compactness, we can assume (up to passing to a subsequence again) that

$$\lim_{\eta\to 0}\frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon}-y_{\eta\varepsilon}\right)=p_\varepsilon\in\mathbb{R}^n. \tag{3.16}$$

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 (3.16) 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 (3.13) and (3.14) 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 3.8. *Assume $H\in C(\mathbb{R}^n\times\mathbb{R}^n)$ satisfies* (3.2) *and* (3.3). *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 (3.10), 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

$$\min_{y\in\mathbb{T}^n} H(y,p) \le \overline{H}(p) \le \max_{y\in\mathbb{T}^n} H(y,p) \qquad \text{for all } p \in \mathbb{R}^n. \tag{3.17}$$

(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 (3.10) with $\min_{\mathbb{T}^n} v_k = 0$ and $p = p_k$ for all $k \in \mathbb{N}$. Note first that, in light of (3.17), 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

$$\lambda u^\lambda + H(y, q + Du^\lambda) = 0 \qquad \text{in } \mathbb{T}^n, \tag{3.18}$$

and

$$\lambda v^\lambda + H(y, p + Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n, \tag{3.19}$$

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 (3.19). 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 3.9. *Assume $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (3.2) *and* (3.3). *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 3.10. *Assume $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (3.2) *and* (3.3)*. 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 (3.10), 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 exists 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 3.8, 3.10, and Corollary 3.9 give us some good qualitative properties of the effective Hamiltonian $\overline{H}$. Most of these were already covered by Lions, Papanicolaou, Varadhan [74]. 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 (3.2) and (3.3). 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.

At this moment of 2019, computing $\overline{H}$ numerically is extremely challenging. The cell problem (3.10) 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 $p \neq q$ 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 use 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 3.11. *Assume that H satisfies* (3.2) *and* (3.3). *Fix $p \in \mathbb{R}^n$. Consider the following Cauchy problem*

$$\begin{cases} w_t + H(y, p + Dw) &= 0 \qquad \text{in } \mathbb{T}^n \times (0,\infty), \\ w(y,0) &= 0 \qquad \text{on } \mathbb{T}^n. \end{cases} \tag{3.20}$$

Let $w(y,t)$ be the unique viscosity solution to (4.21). *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 (3.10). We simply construct a separable subsolution and supersolution to (4.21), and use them to bound the actual solution $w(x,t)$.

Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to (3.10). Define:

$$\varphi(x,t) = v(x) - \overline{H}(p)t \qquad \text{for} \qquad (x,t) \in \mathbb{T}^n \times [0,\infty).$$

It is clear that φ is a separable solution to (4.21) 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 (4.21), respectively. By the comparison principle,

$$v(x) - \overline{H}(p)t - C \le w(x,t) \le v(x) - \overline{H}(p)t + C \qquad \text{for} \qquad (x,t) \in \mathbb{T}^n \times (0,\infty).$$

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. □

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 [48], 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 $C > 0$. 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 is subadditive, that is,

$$M(t) + M(s) \ge M(t+s) \qquad \text{for all } s,t \ge 0. \tag{3.21}$$

Indeed, fix $s \ge 0$. Set $\phi(y,t) = u(y,t+s) - M(s)$ for all $(x,t) \in \mathbb{T}^n \times [0,\infty)$. Then, ϕ solves (4.21) with initial data $\phi(y,0) = u(y,s) - M(s) \le 0$ for $x \in \mathbb{T}^n$. We use the comparison principle to get that $\phi \le u$. In particular,

$$M(t+s) - M(s) = \max_{y\in\mathbb{T}^n} \phi(y,t) \le \max_{y\in\mathbb{T}^n} u(y,t) = M(t).$$

Thus, (3.21) 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

4.3 Problems

Exercise 24. *Prove Corollary 3.9.*

Exercise 25. *Assume that H satisfies* (3.2) *and* (3.3). *Assume further that there exists $k > 0$ such that H is 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 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 3.12 (The inf-sup formula)**.** *Assume that H satisfies* (3.2) *and* (3.3). *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*

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \quad H\left(y, p + D\phi(y)\right). \tag{3.22}$$

Proof. Pick any $\varphi \in C^1(\mathbb{T}^n)$, by the representation formula in Theorem 3.10,

$$\overline{H}(p) \leq \max_{y \in \mathbb{T}^n} \quad H(y, p + D\varphi(y)),$$

and hence,

$$\overline{H}(p) \leq \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \quad H\left(y, p + D\phi(y)\right).$$

Conversely, given $\theta > 0$, we aim at proving that

$$\overline{H}(p) + \theta \geq \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \quad H\left(y, p + D\phi(y)\right).$$

Let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to (3.10), 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,

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

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 $\varepsilon \to 0$. 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 3.13. *Assume that H satisfies* (3.2) *and* (3.3). *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

$$\max_{x\in\mathbb{T}^n} H\left(x, \frac{p+q}{2} + D\left(\frac{\varphi+\psi}{2}\right)(x)\right) \leq \frac{1}{2}\left(\max_{x\in\mathbb{T}^n} H(x, p+D\varphi(x)) + \max_{x\in\mathbb{T}^n} H(x, q+D\psi(x))\right).$$

The inf-sup formula (3.12) 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 [11] in Theorem 2.24, we have another formula for $\overline{H}$ in the convex setting.

Corollary 3.14. *Assume that H satisfies* (3.2) *and* (3.3). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{T}^n$. Then, for each $p \in \mathbb{R}^n$,*

$$\overline{H}(p) = \inf\{c \in \mathbb{R} \,:\, \exists\, v \in \mathrm{Lip}(\mathbb{T}^n) : H(y, p + Dv(y)) \le c \text{ a.e. in } \mathbb{T}^n\}. \tag{3.23}$$

One can then use this Corollary to give another quick proof of Theorem 3.13. This proof is left as an exercise.

5.2 The large time average formula

We use Theorem 3.11 to give a large time average formula in the convex setting as following. This result was obtained first by Concordel [21].

Theorem 3.15. *Assume that H satisfies* (3.2)*, $p \mapsto H(y,p)$ 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 3.11 here. Let w be the solution to (4.21), 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 (4.21) 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 [21, 22] 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 [74]. According to the paper, Tartar was the one who provided this example.

Example 3.2. *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

$$\overline{H}(p) = \begin{cases} 0 & \text{for } |p| \le \langle\sqrt{V}\rangle, \\ \lambda & \text{for } |p| \ge \langle\sqrt{V}\rangle, \text{ where } \lambda \ge 0 \text{ is such that } |p| = \langle\sqrt{\lambda+V}\rangle. \end{cases} \tag{3.24}$$

Note that this formula only holds in one dimension. There is no such formula in multi dimensions.

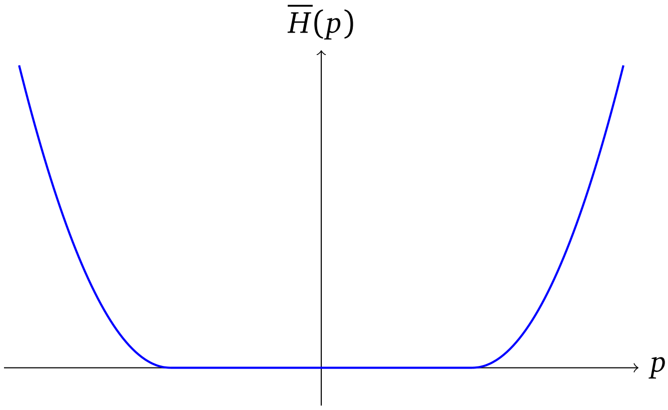


Figure 3.2: Graph of $\overline{H}$

Let us now prove the above formula.

Proof of formula (3.24). 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 $V \not\equiv 0$, 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 26. *Use Corollary 3.14 to give another quick proof of Theorem 3.13.*

Exercise 27. *Assume that H satisfies* (3.2) *and* (3.3). *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 $p \in \mathbb{R}^n$,*

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \quad H\left(y, p + D\phi(y)\right).$$

Exercise 28. *Assume that H satisfies* (3.2) *and* (3.3). *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 3.16. *Assume that H satisfies* (3.2) *and* (3.3). *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

$$\begin{aligned}\overline{H}(p) &= \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \quad H\left(y, p + D\phi(y)\right) \\ &= \inf_{\phi \in C^1(\mathbb{T}^n)} \max_{y \in \mathbb{T}^n} \quad H\left(y, -p + D(-\phi)(y)\right) = \overline{H}(-p).\end{aligned}$$

□

Since the inf-max formula still holds for the level-set quasiconvex case, we have the following corollary, which is quite useful.

Corollary 3.17. *Assume that H satisfies* (3.2) *and* (3.3). *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 3.18. 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 $\overline{H}$ is convex, but we want to know more about its behavior in this section.

Lemma 3.19. *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 (3.10), 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 3.20. *Assume that H satisfies* (3.2) *and* (3.3). *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 3.21. *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 [22]. 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 [81].

Proof. By Lemma 3.19, 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 $d = \min\{\text{dist}(U_0, \partial U_1), \text{dist}(U_1, \partial U_2)\} > 0$. 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 3.22. The proof of Concordel [22] 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 $k \ne j$,

$$\text{dist}(k + U_1, j + U_1) \ge d.$$

We show $\overline{H}(p) = 0$ for $|p| \le r = \frac{d\sqrt{\varepsilon_0}}{C}$. By Theorem 3.15, 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

$$\int_{t_1}^{t_2} \left(p \cdot \gamma'(s) - \frac{1}{2}|\gamma'(s)|^2 - V(\gamma(s)) \right) ds \le \int_{t_1}^{t_2} \left(p \cdot \gamma'(s) - \frac{1}{2}|\gamma'(s)|^2 - \varepsilon_0 \right) ds$$
$$\le \int_{t_1}^{t_2} \left(-\frac{1}{2}|\gamma'(s) - p|^2 + \frac{1}{2}|p|^2 - \varepsilon_0 \right) ds \le -\frac{1}{2} \int_{t_1}^{t_2} \left(|\gamma'(s) - p|^2 + \varepsilon_0 \right) ds,$$

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.

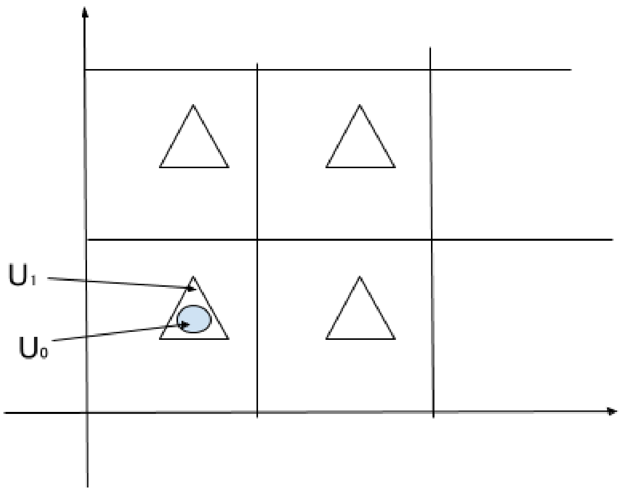


Figure 3.3: Periodic structures and $k + U_1$ for $k \in \mathbb{Z}^n$

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 3.3. *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 [22].

Theorem 3.23. *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 $\alpha=\lambda^2|p_0|^2>0$, 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 $\bar{t}_m=\frac{t_m}{\alpha}$ for $m\in\mathbb{N}$,

$$p\cdot\frac{\gamma(\bar{t}_m)-\gamma(0)}{\bar{t}_m}-\frac{1}{2}\alpha^2=p\cdot\frac{\xi(t_m)-\xi(0)}{\frac{t_m}{\alpha}}-\frac{1}{2}\alpha^2\longrightarrow\alpha\lambda^2|p_0|^2-\frac{1}{2}\alpha^2=\frac{1}{2}\lambda^4|p_0|^4,$$

as $m\to\infty$. Therefore, by Theorem 3.15,

$$\overline{H}(p)=\lim_{t\to\infty}\sup_{\gamma(\cdot)}\frac{1}{t}\int_0^t\left(p\cdot\gamma'(s)-\frac{1}{2}|\gamma'(s)|^2-V(\gamma(s))\right)ds\geq\frac{1}{2}\lambda^4|p_0|^4.$$

The proof is complete. □

Remark 3.24. 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

$$-\frac{d}{ds}(\gamma'(s)-p)+DV(\gamma(s))=0\quad\Rightarrow\quad\gamma''(s)=DV(\gamma(s)).$$

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 $\overline{H}$ on some cases. 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 [87].

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

$$\begin{cases} \min_{\mathbb{R}^n} H=0;\\ \text{there exists a bounded domain } U\subset\mathbb{R}^n \text{ such that } \{H=0\}=\partial U;\\ H \text{ is even, that is, } H(p)=H(-p) \text{ for all } p\in\mathbb{R}^n;\\ \text{there exist } H_1,H_2\in C(\mathbb{R}^n) \text{ such that } H=\max\{H_1,H_2\}. \end{cases} \tag{3.25}$$

Here, H_1,H_2 satisfy

$$\begin{cases} H_1 \text{ is coercive, level-set quasiconvex, even,}\\ \text{and } H_1=H \text{ in } \mathbb{R}^n\setminus U,\ H_1<0 \text{ in } U;\\ H_2 \text{ is level-set quasiconcave, even,}\\ \text{and } H_2=H \text{ in } U,\ H_2<0 \text{ in } \mathbb{R}^n\setminus U,\ \lim_{|p|\to\infty} H_2(p)=-\infty. \end{cases} \tag{3.26}$$

An example of H satisfying (3.25)–(3.26) is $H(p)=(|p|^2-2)^2$ as in the following graph. Below is the decomposition result for this simplest case.

Theorem 3.25. *Let $H\in C(\mathbb{R}^n)$ be a Hamiltonian satisfying* (3.25)–(3.26)*. 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. This is because of the nonconvex situation. See the discussion in Section 6.5 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 (3.10). Pick $y_0\in\mathbb{T}^n$ such that $\min_{\mathbb{T}^n} v=v(y_0)$. By the definition of viscosity supersolutions to (3.10), we get

$$H(p)\ge H(p)-V(y_0)\ge\overline{H}(p).$$

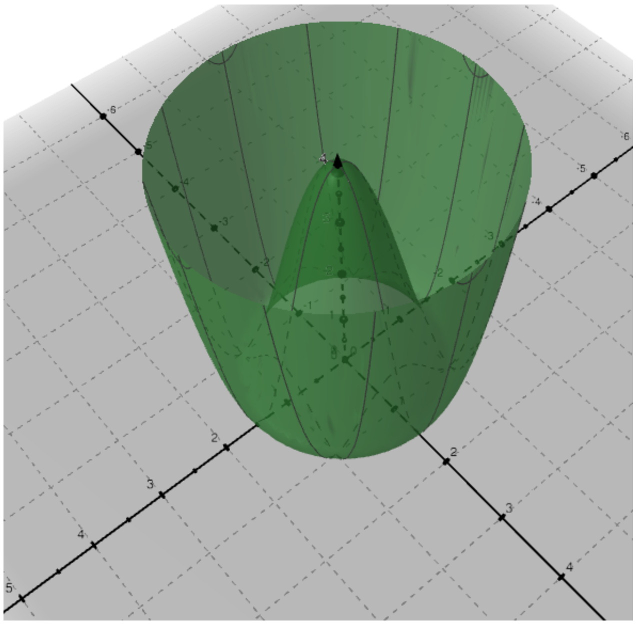


Figure 3.4: An example where $H(p) = (|p|^2 - 2)^2$

On the other hand, as (3.10) 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,

$$\overline{H}(p) = 0 \qquad \text{for all } p \in \partial U. \tag{3.27}$$

Besides, as $H_i \leq H$, we get $\overline{H}_i \leq \overline{H}$. Therefore,

$$\overline{H} \geq \max\left\{\overline{H}_1, \overline{H}_2, 0\right\}. \tag{3.28}$$

It remains to prove the reverse inequality of (3.28) 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 27) 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

$$H_1(-p + Dv(y, -p)) - V(y) = \overline{H}_1(-p) = \overline{H}_1(p) \qquad \text{in } \mathbb{T}^n. \tag{3.29}$$

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 (3.29) and the quasiconvexity of H_1 (Exercise 20),

$$\overline{H}_1(p) = H_1(-p-q) - V(y) = H_1(p+q) - V(y).$$

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 (3.26). 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 3.10 on a representation formula of $\overline{H}(p)$, $\overline{H}(p) \leq \overline{H}_1(p)$.

STEP 3. Assume now that $\overline{H}_2(p) \geq \max\left\{\overline{H}_1(p),\ 0\right\}$. By using 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 (3.27) 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

$$0 \leq \overline{H}^1 = \overline{H} \leq \overline{H}^{\sigma} \qquad \text{for all } \sigma \in [0,1]. \tag{3.30}$$

By repeating Steps 2 and 3 above, we get

$$\text{For } p \in \mathbb{R}^n \text{ and } \sigma \in [0,1], \text{ if } \max\left\{\overline{H}_1^{\sigma}(p), \overline{H}_2^{\sigma}(p)\right\} = 0, \text{ then } \overline{H}^{\sigma}(p) = 0. \tag{3.31}$$

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 (3.31), yield that $\overline{H}^s(p) = 0$. Combine this with (3.30) to finally get that $\overline{H}(p) = 0$.

□

Remark 3.26. 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 (3.25)–(3.26) are general and a bit complicated. A simple situation where (3.25)–(3.26) hold is a radially symmetric case where $H(p) = \psi(|p|)$, and $\psi \in C([0,\infty),\mathbb{R})$ satisfying

$$\begin{cases} \psi(0) > 0,\ \psi(1) = 0,\ \lim_{r\to\infty} \psi(r) = +\infty, \\ \psi \text{ is strictly decreasing in } (0,1), \text{ and is strictly increasing in } (1,\infty). \end{cases} \tag{3.32}$$

Let $\psi_1, \psi_2 \in C([0,\infty),\mathbb{R})$ be such that

$$\begin{cases} \psi_1 = \psi \text{ on } [1,\infty), \text{ and } \psi_1 \text{ is strictly increasing on } [0,1], \\ \psi_2 = \psi \text{ on } [0,1], \ \psi_2 \text{ is strictly decreasing on } [1,\infty), \text{ and } \lim_{r\to\infty} \psi_2(r) = -\infty. \end{cases} \tag{3.33}$$

See Figure 3.5. Set $H_i(p) = \psi_i(|p|)$ for $p \in \mathbb{R}^n$, and for $i = 1,2$. It is clear that (3.25)–(3.26) hold true provided that (3.32)–(3.33) hold.

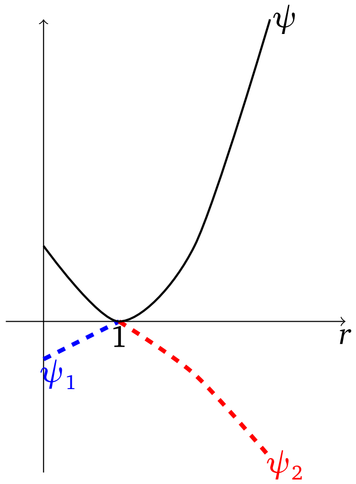


Figure 3.5: Graphs of ψ, ψ_1, ψ_2

An immediate consequence of Theorem 3.25 is the following result.

Corollary 3.27. *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* (3.32)–(3.33). *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 3.28. A special case of Corollary 3.27 is when

$$H(p) = \psi(|p|) = \left(|p|^2 - 1\right)^2 \qquad \text{for } p \in \mathbb{R}^n,$$

which was studied first by Armstrong, Tran and Yu [4]. Of course, Armstrong, Tran and Yu [4] 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 [4].

By using Corollary 3.27 and approximation, we get another representation formula for $\overline{H}$ which will be used later.

Corollary 3.29. *Assume that* (3.32)–(3.33) *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 3.6 for the graphs of $\psi, \tilde{\psi}_1, \psi_2$.

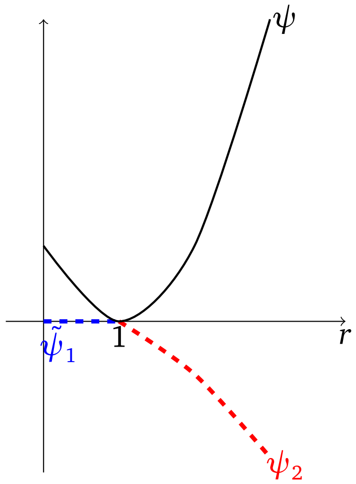


Figure 3.6: Graphs of $\psi, \tilde{\psi}_1, \psi_2$

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 3.30. *Let $H \in C(\mathbb{R}^n)$ be a coercive Hamiltonian satisfying* (3.25)–(3.26)*, 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 3.30 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,

$$\max\left\{\overline{H}_1,\ 0\right\} \le \overline{H} \le \overline{H}^+. \tag{3.34}$$

On the other hand, by Theorem 3.25, the representation formula for $\overline{H}^+$ is

$$\overline{H}^+ = \max\left\{\overline{H}_1,\ \overline{H}_2^+,\ 0\right\} = \max\{\overline{H}_1,\ 0\}, \tag{3.35}$$

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 (3.34) and (3.35) to get the conclusion. □

6.2 A more general case

We now proceed to give an extension of Theorem 3.25 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 (3.25)–(3.26).

Let $H : \mathbb{R}^n \to \mathbb{R}$ be such that

$$\begin{cases} H(p) = \varphi(|p|) \text{ for } p \in \mathbb{R}^n, \text{ where } \varphi \in C([0,\infty),\mathbb{R}) \text{ satisfies} \\ \varphi(0) > 0,\ \varphi(2) = 0,\ \lim_{r\to\infty} \varphi(r) = +\infty, \\ \varphi \text{ is strictly increasing on } [0,1] \text{ and } [2,\infty), \\ \text{and } \varphi \text{ is strictly decreasing on } [1,2]. \end{cases} \tag{3.36}$$

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

$$\begin{cases} \varphi_1 = \varphi \text{ on } [2,\infty),\ \varphi_1 \text{ is strictly increasing on } [0,2], \\ \varphi_2 = \varphi \text{ on } [1,2],\ \varphi_2 \text{ is strictly decreasing on } [0,1] \text{ and } [2,\infty),\ \lim_{r\to\infty} \varphi_2(r) = -\infty, \\ \varphi_3 = \varphi \text{ on } [0,1],\ \varphi_3 \text{ is strictly increasing on } [1,\infty),\ \text{and } \varphi_3 > \varphi \text{ in } (1,\infty). \end{cases} \tag{3.37}$$

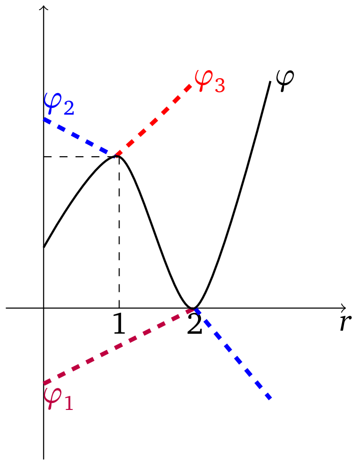


Figure 3.7: Graphs of $\varphi, \varphi_1, \varphi_2, \varphi_3$

Lemma 3.31. *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* (3.36)–(3.37). *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 3.32).

Proof. Considering $-K(-p)$, thanks to the representation formula and evenness from Theorem 3.25,

$$\overline{K} = \min\left\{\overline{H}_2, \overline{H}_3, \varphi(1) - \max_{\mathbb{T}^n} V\right\}.$$

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 3.29, we have another representation formula for $\overline{K}$ as following

$$\overline{K} = \min\left\{\overline{\tilde{H}}_2, \overline{H}_3\right\}. \tag{3.38}$$

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

$$\overline{H}(p) = 0 \qquad \text{for all } |p| = 2. \tag{3.39}$$

Besides, as $K, H_1 \le H$, we deduce furthermore that $\overline{K}, \overline{H}_1 \le \overline{H}$. Thus,

$$\overline{H} \ge \max\left\{0, \overline{H}_1, \overline{K}\right\} \tag{3.40}$$

We now show the reverse inequality of (3.40) 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 3.25 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 3.25 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

$$\tilde{H}_2(p + Dv(y, p)) - V(y) = \overline{\tilde{H}}_2(p) \ge 0 \qquad \text{in } \mathbb{T}^n. \tag{3.41}$$

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 3.10 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 (3.40), 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 (3.39) 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

$$0 \le \overline{H}^1 = \overline{H} \le \overline{H}^\sigma \qquad \text{for all } \sigma \in [0, 1]. \tag{3.42}$$

By repeating Steps 2 and 3 above, we get

$$\text{For } p \in \mathbb{R}^n \text{ and } \sigma \in [0, 1], \text{ if } \max\left\{\overline{H}_1^\sigma(p), \overline{K}^\sigma(p)\right\} = 0, \text{ then } \overline{H}^\sigma(p) = 0. \tag{3.43}$$

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 (3.42) and (3.43), 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 3.8 and 3.9. 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.

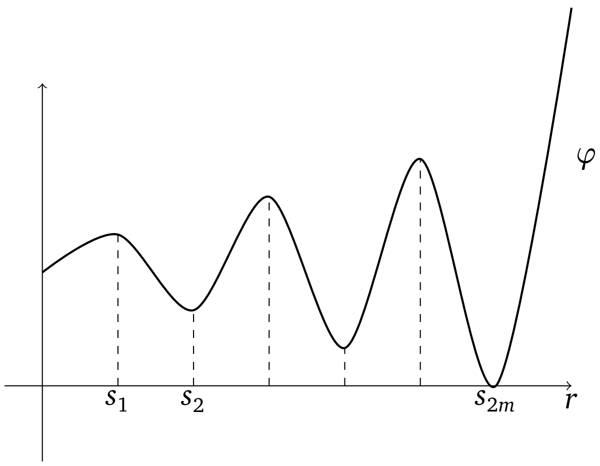


Figure 3.8: Graph of φ in first general case

In the first general case corresponding to Figure 3.8, we assume that

$$\begin{cases} \varphi \in C([0,\infty),\mathbb{R}) \text{ satisfies that} \\ \text{there exist } m \in \mathbb{N} \text{ and } 0 = s_0 < s_1 < \ldots s_{2m} < \infty = s_{2m+1} \text{ such that} \\ \varphi \text{ is strictly increasing in } (s_{2i}, s_{2i+1}), \text{ and is strictly decreasing in } (s_{2i+1}, s_{2i+2}), \\ \varphi(s_0) > \varphi(s_2) > \ldots > \varphi(s_{2m}), \text{ and } \varphi(s_1) < \varphi(s_3) < \ldots < \varphi(s_{2m+1}) = \infty. \end{cases} \tag{3.44}$$

Based on φ, we construct $\varphi_0, \ldots, \varphi_{2m}$ 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 3.32. *Assume that* (3.44) *holds for some* $m \in \mathbb{N}$. *Then,*

$$\overline{H}_m = \max\left\{\overline{K}_{m-1},\ \overline{\Phi}_{2m},\ \varphi(s_{2m}) - \min_{\mathbb{T}^n} V\right\}, \tag{3.45}$$

and

$$\overline{K}_{m-1} = \min\left\{\overline{H}_{m-1},\ \overline{\Phi}_{2m-1},\ \varphi(s_{2m-1}) - \max_{\mathbb{T}^n} V\right\}. \tag{3.46}$$

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 6.5 for this subtle issue.

Proof. We prove by induction.

The base case is when $m = 1$. The two formulas (3.45) and (3.46) follow immediately from Lemma 3.31 and Theorem 3.25.

Assume that (3.45) and (3.46) hold for $m \in \mathbb{N}$. We need to verify these equalities for $m+1$. Using similar arguments as those in the proof Lemma 3.31, 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 3.31, 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 3.33. Two comments are in order.

(i) By approximation, we see that representation formulas (3.45) and (3.46) still hold true if we relax (3.44) 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 3.30, 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

$$\overline{H}_m = \max\left\{\overline{\Phi}_{2m},\ \varphi(s_{2m}) - \min_{\mathbb{T}^n} V\right\}.$$

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 3.9 after normalization by a constant. By changing the notations appropriately, we obtain similar representation formulas as in Theorem 3.32. We omit the details here.

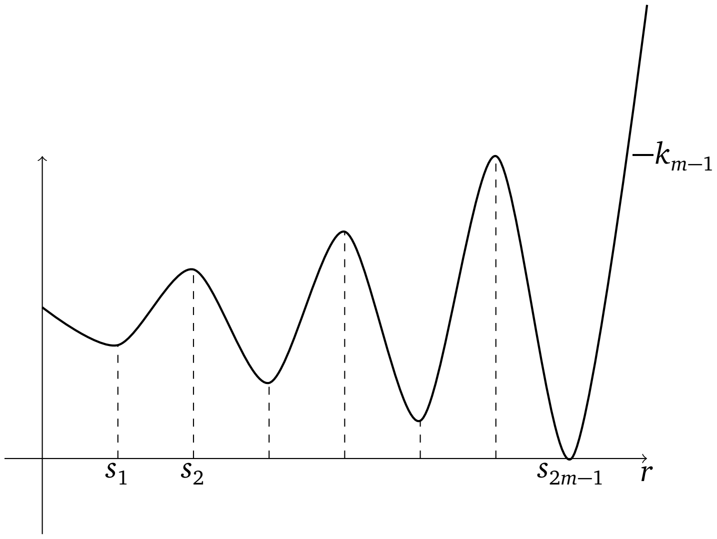


Figure 3.9: Graph of $-k_{m-1}$ in the second general case

6.4 Problems

Exercise 29. *Assume that H satisfies* (3.2) *and* (3.3)*. 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 30. *Assume $H(p) = -k_{m-1}(|p|)$ for all $p \in \mathbb{R}^n$ as described in Figure 3.9, and $V \in C(\mathbb{T}^n)$. Obtain the formula for $\overline{H}(p)$ of the Hamiltonian $H(p) - V(y)$.*

6.5 Loss of evenness and non-decomposable effective Hamiltonians

A natural question is whether we can extend Theorem 3.32 to other nonconvex H. 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 $\max V$:

$$\overline{H} = G(\overline{H}_1, \ldots, \overline{H}_m,\ \min V,\ \max V) \tag{3.47}$$

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 [74], 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 [87].

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 3.25.

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 3.27, Lemma 3.31, and Theorem 3.32).

3. However, in general, the evenness is lost as presented in [76, 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 3.10 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 (3.47) 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.

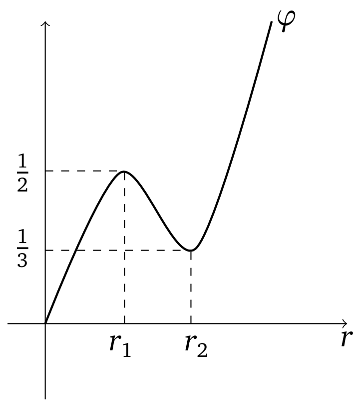


Figure 3.10: Graphs of φ

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.

7 Rates of convergence

7.1 The method of Capuzzo-Dolcetta and Ishii

We now address the results by Capuzzo-Dolcetta and Ishii [18]. Assume that H satisfies (3.2) and (3.3). 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 [18] 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 3.34. *Assume that $H \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (3.2) *and* (3.3)*. 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*

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{3.48}$$

And let u be the unique solution to the effective equation

$$\begin{cases} u_t + \overline{H}(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{3.49}$$

Then, for each $T > 0$, there exists a constant $C > 0$ dependent on H, u_0, and T such that

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C\varepsilon^{1/3}. \tag{3.50}$$

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

$$|H(y,p) - H(y,q)| \le C|p-q| \qquad \text{for all } y \in \mathbb{T}^n,\ p,q \in \mathbb{R}^n. \tag{3.51}$$

And of course, this additional condition does not change any generality of Theorem 3.34.

For each $p \in \mathbb{R}^n$, we first look back at the discount approximation of cell problem (3.10) as following. For each $\lambda > 0$, we consider the static equation

$$\lambda v^\lambda + H(y, p + Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n. \tag{3.52}$$

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 3.5 and part (c) of the proof of Theorem 3.8.

Lemma 3.35. *Assume that H satisfies* (3.2), (3.3), *and* (3.51)*. Then, the following claims hold.*

(i) There exists $C > 0$ independent of $\lambda > 0$ such that, for all $p, q \in \mathbb{R}^n$,

$$\lambda|v^\lambda(y,p) - v^\lambda(y,q)| \le C|p-q| \qquad \text{for all } y \in \mathbb{T}^n.$$

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 $p \in B(0,R)$,

$$|\lambda v^\lambda(y,p) + \overline{H}(p)| \le C\lambda \qquad \text{for all } y \in \mathbb{T}^n.$$

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 (3.52), respectively, thanks to (3.51). 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 (3.10) 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 (3.52), 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 3.34. Again, by the reduction step, we assume also (3.51).

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 [18, 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 $\Phi(\hat{x},\hat{x},\hat{t},\hat{t}) \leq \Phi(\hat{x},\hat{y},\hat{t},\hat{s})$, together with Lipschitz property of u and Lemma 3.35, implies

$$\begin{aligned}\frac{|\hat{x}-\hat{y}|^2+|\hat{t}-\hat{s}|^2}{2\varepsilon^\beta} &\leq u(\hat{y},\hat{s})-u(\hat{x},\hat{t})+\varepsilon\left(v^\lambda\left(\frac{\hat{x}}{\varepsilon},\frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)-v^\lambda\left(\frac{\hat{x}}{\varepsilon},0\right)\right)+K(\hat{s}-\hat{t})\\ &\leq C(|\hat{x}-\hat{y}|+|\hat{t}-\hat{s}|)+C\varepsilon\frac{1}{\lambda}\frac{|\hat{x}-\hat{y}|}{\varepsilon^\beta}\\ &\leq C(|\hat{x}-\hat{y}|+|\hat{t}-\hat{s}|)\end{aligned}$$

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

$$K+\frac{t_\alpha-\hat{s}}{\varepsilon^\beta}+H\left(\frac{x_\alpha}{\varepsilon},\frac{x_\alpha-\hat{y}}{\varepsilon^\beta}+\frac{(x_\alpha-\varepsilon\xi_\alpha)+(x_\alpha-z_\alpha)}{\alpha}\right)\leq 0,$$

and

$$\lambda v^\lambda\left(\xi_\alpha,\frac{z_\alpha-\hat{y}}{\varepsilon^\beta}\right)+H\left(\xi_\alpha,\frac{z_\alpha-\hat{y}}{\varepsilon^\beta}+\frac{x_\alpha-\varepsilon\xi_\alpha}{\alpha}\right)\geq 0.$$

Besides, since $\psi(x_\alpha,\xi_\alpha,z_\alpha,t_\alpha)\geq\psi(x_\alpha,\xi_\alpha,x_\alpha,t_\alpha)$,

$$\frac{|x_\alpha-z_\alpha|^2}{2\alpha}\leq\varepsilon\left(v^\lambda\left(\xi_\alpha,\frac{x_\alpha-\hat{y}}{\varepsilon^\beta}\right)-v^\lambda\left(\xi_\alpha,\frac{z_\alpha-\hat{y}}{\varepsilon^\beta}\right)\right)\leq\varepsilon^{1-\theta-\beta}|x_\alpha-z_\alpha|,$$

which yields $\frac{|x_\alpha-z_\alpha|}{\alpha}\leq C\varepsilon^{1-\theta-\beta}$. We now combine this with the two above inequalities on the sub/supersolution tests and let $\alpha\to 0+$ to deduce that

$$K+\frac{\hat{t}-\hat{s}}{\varepsilon^\beta}\leq\lambda v^\lambda\left(\frac{\hat{x}}{\varepsilon},\frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)+C\varepsilon^{1-\theta-\beta}\leq-\overline{H}\left(\frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)+C\varepsilon^\theta+C\varepsilon^{1-\theta-\beta},$$

and hence,

$$K+\frac{\hat{t}-\hat{s}}{\varepsilon^\beta}+\overline{H}\left(\frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)\leq C\varepsilon^\theta+C\varepsilon^{1-\theta-\beta}. \tag{3.53}$$

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

$$-K+\frac{\hat{t}-\hat{s}}{\varepsilon^\beta}+\overline{H}\left(\frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)+C\varepsilon^\theta+C\varepsilon^{1-\theta-\beta}\geq 0. \tag{3.54}$$

Combine (3.53) and (3.54) 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

$$\Phi(\hat{x}, \hat{y}, \hat{t}, \hat{s}) \le u^\varepsilon(\hat{x}, \hat{t}) - u(\hat{y}, \hat{s}) - \varepsilon v^\lambda\left(\frac{\hat{x}}{\varepsilon}, \frac{\hat{x} - \hat{y}}{\varepsilon^\beta}\right) \le C\varepsilon^{1/3}.$$

In particular, $\Phi(x, x, t, t) \le C\varepsilon^{1/3}$, which infers

$$u^\varepsilon(x, t) - u(x, t) \le C\varepsilon^{1/3} + \varepsilon v^\lambda\left(\frac{\hat{x}}{\varepsilon}, 0\right) + 2K_1\varepsilon^{1/3} t \le C(1 + T)\varepsilon^{1/3}.$$

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 3.36. 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 (3.10) 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 3.4). Instead, we work indirectly with v^λ for $\lambda = \varepsilon^\theta$, which has good regularity and stability estimates as stated in Lemma 3.35. 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.

$$\begin{cases} \text{For each } p \in \mathbb{R}^n, \text{ we are able to pick a solution } v(y, p) \text{ of (3.10) such that} \\ p \mapsto v(\cdot, p) \quad \text{is Lipschitz.} \end{cases} \tag{3.55}$$

Condition (3.55) is however a very strong and restrictive requirement. We will see that this does not hold in some examples later.

Theorem 3.37. *Assume that $H \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies* (3.2) *and* (3.3). *Let $\overline{H}$ be the corresponding effective Hamiltonian of H. Assume further that* (3.55) *holds. Assume $u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n)$. For $\varepsilon > 0$, let u^ε be the unique solution to* (3.48). *Also let u be the unique solution to* (3.49). *Then for each $T > 0$, there exists $C > 0$ dependent on H, u_0, and T such that*

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^n \times [0,T])} \le C\varepsilon^{1/2}. \tag{3.56}$$

Proof. Thanks to (3.55), 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 3.34, 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 [18, 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

$$\begin{aligned}\frac{|\hat{x}-\hat{y}|^2 + |\hat{t}-\hat{s}|^2}{2\varepsilon^\beta} &\le \left(u(\hat{x},\hat{t}) - u(\hat{y},\hat{s})\right) + \varepsilon\left(v\left(\frac{\hat{x}}{\varepsilon}, 0\right) - v\left(\frac{\hat{x}}{\varepsilon}, \frac{\hat{x}-\hat{y}}{\varepsilon^\beta}\right)\right) + K(\hat{t}-\hat{s}) \\ &\le C\left(|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}|\right) + C\varepsilon\frac{|\hat{x}-\hat{y}|}{\varepsilon^\beta} \le C\left(|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}|\right).\end{aligned}$$

Therefore,

$$|\hat{x}-\hat{y}| + |\hat{t}-\hat{s}| \le C\varepsilon^\beta. \tag{3.57}$$

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 (3.55) 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

$$|x_\alpha - z_\alpha| \le C\alpha\varepsilon^{1-\beta}. \tag{3.58}$$

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

$$|x_\alpha - \varepsilon\xi_\alpha| \le C\alpha. \tag{3.59}$$

By definition of viscosity solutions,

$$K + \frac{t_\alpha - \hat{s}}{\varepsilon^\beta} + H\left(\frac{x_\alpha}{\varepsilon}, \frac{x_\alpha-\hat{y}}{\varepsilon^\beta} + \frac{(x_\alpha - \varepsilon\xi_\alpha) + (x_\alpha - z_\alpha)}{\alpha}\right) \le 0, \tag{3.60}$$

and

$$H\left(\xi_\alpha, \frac{z_\alpha - \hat{y}}{\varepsilon^\beta} + \frac{x_\alpha - \varepsilon\xi_\alpha}{\alpha}\right) \geq \overline{H}\left(\frac{z_\alpha - \hat{y}}{\varepsilon^\beta}\right). \tag{3.61}$$

Combining (3.58)–(3.61) and letting $\alpha \to 0$ to yield that

$$K + \frac{\hat{t} - \hat{s}}{\varepsilon^\beta} + \overline{H}\left(\frac{\hat{x} - \hat{y}}{\varepsilon^\beta}\right) - C\varepsilon^{1-\beta} \leq 0. \tag{3.62}$$

By a similar procedure,

$$-K + \frac{\hat{t} - \hat{s}}{\varepsilon^\beta} + \overline{H}\left(\frac{\hat{x} - \hat{y}}{\varepsilon^\beta}\right) + C\varepsilon^{1-\beta} \geq 0. \tag{3.63}$$

Putting (3.62) and (3.63) together to get

$$K \leq C\varepsilon^{1-\beta}.$$

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 3.34. □

7.3 Problems

Exercise 31. *Let $n = 1$, and $H(y, p) = |p| - V(y)$ for some $V \in C(\mathbb{T})$. Show that* (3.55) *holds in this case.*

Exercise 32. *Let $n = 1$. Is it true that* (3.55) *always holds for H that satisfies* (3.2)*,* (3.3)*, and* (3.51)*?*

8 References

1. Periodic homogenization for Hamilton–Jacobi equations was first studied in the paper of Lions, Papanicolaou, Varadhan [74] circa 1987. The paper is still unpublished, but it has been extremely influential in this area.

2. Evans introduced the perturbed test function method [31] 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 [21, 22]. Some of the results are proven by using recent ideas of Mitake and Tran [81].

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 [87]. A special case was done earlier by Armstrong, Tran, Yu [4]. Then, Gao generalized [87] to general non separable Hamiltonians and obtained similar results in [45]. For one dimensional case, shape of $\overline{H}$ is well understood qualitatively by the results of Armstrong, Tran, Yu [5], and Gao [44]. Of course, quantitative and better understanding in one dimensional case are important to be studied in the near future.

5. The $O(\varepsilon^{1/3})$ rate of convergence was obtained by Capuzzo-Dolcetta and Ishii [18] about 15 years after the first qualitative homogenization result. The method introduced in [18] is also standard 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 4

Almost periodic homogenization theory for Hamilton–Jacobi equations

1 Introduction to almost periodic homogenization theory

1.1 Introduction

As in Chapter 3, our objects of interests are the same. The equations of interest are as following. For each $\varepsilon > 0$, we study

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad\qquad u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.1}$$

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. 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

$$\begin{cases} u_t + \overline{H}(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.2}$$

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. 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 [60], 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 4.1. *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 4.1. *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 4.2. *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.

$$H \text{ is almost periodic in } y \text{ in the sense of Definition 4.2,} \tag{4.3}$$

and

$$\lim_{|p| \to \infty} H(y,p) = +\infty \text{ uniformly for } y \in \mathbb{R}^n. \tag{4.4}$$

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

$$u^\varepsilon(x,t) = u(x,t) + \varepsilon v\left(\frac{x}{\varepsilon}\right) = u(x,t) + \varepsilon v(y).$$

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

$$\lim_{\varepsilon\to 0} \varepsilon v\left(\frac{x}{\varepsilon}\right) = 0. \tag{4.5}$$

Anyway, plug this expansion to (4.1) 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

$$H(y, p + Dv(y)) = c \qquad \text{in } \mathbb{R}^n. \tag{4.6}$$

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 (4.5) can be reformulated as

$$\lim_{|y|\to\infty} \frac{v(y)}{|y|} = 0, \tag{4.7}$$

which means that v is sublinear in $\mathbb{R}^n$. Hence, our task is to find $c \in \mathbb{R}$ so that (4.6) 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

$$\lambda v^\lambda(y) + H(y, p + Dv^\lambda(y)) = 0 \qquad \text{in } \mathbb{R}^n. \tag{4.8}$$

Our goal is to let $\lambda \to 0+$ to obtain $\overline{H}$. We have first the following proposition.

Proposition 4.3. *Assume* (4.3) *and* (4.4). *Fix $p \in \mathbb{R}^n$. For $\lambda > 0$, let v^λ be the viscosity solution to* (4.8). *Then,*

$$\lim_{\lambda\to 0+} \left(\lambda \sup_{y\in\mathbb{R}^n} |v^\lambda(y) - v^\lambda(0)| \right) = 0. \tag{4.9}$$

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 (4.3), 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 (4.8), respectively. Thus,

$$-\frac{C}{\lambda} \le v^\lambda \le \frac{C}{\lambda}.$$

Then, the coercivity of H gives us that $\|Dv^\lambda\|_{L^\infty(\mathbb{R}^n)} \le C$ for some $C > 0$ independent of $\lambda > 0$. Thus, for $R = C + |p| + 1$, one has $|p| + \|Dv^\lambda\|_{L^\infty(\mathbb{R}^n)} \le R$, and for $j, k \in \mathbb{N}$ large enough

$$|H(y + y_j, p) - H(y + y_k, p)| \le \frac{\delta}{4} \qquad \text{for all } y \in \mathbb{R}^n, p \in B(0,R). \tag{4.10}$$

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 (4.10), for $y \in \mathbb{R}^n$,

$$\lambda_j w_j(y) + H(y, p + Dw_j(y)) \le \lambda_j v^{\lambda_j}(y + y_j - y_1) + H(y + y_j - y_1, p + Dv^{\lambda_j}(y + y_j - y_1)) + \frac{\delta}{4} = \frac{\delta}{4},$$

and

$$\lambda_j w_j(y) + H(y, p + Dw_j(y)) \ge \lambda_j v^{\lambda_j}(y + y_j - y_1) + H(y + y_j - y_1, p + Dv^{\lambda_j}(y + y_j - y_1)) - \frac{\delta}{4} = -\frac{\delta}{4}.$$

Hence, by the usual comparison principle,

$$\lambda_j w_j(y) - \frac{\delta}{4} \le \lambda_j v^{\lambda_j}(y) \le \lambda_j w_j(y) + \frac{\delta}{4} \qquad \text{for } y \in \mathbb{R}^n.$$

Let $y = 0$ in the above to infer

$$\lambda_j |v^{\lambda_j}(y_j - y_1) - v^{\lambda_j}(0)| \le \frac{\delta}{4}.$$

We then use the Lipschitz bound on v^{λ_j} to imply further

$$\lambda_j |v^{\lambda_j}(y_j) - v^{\lambda_j}(0)| \le \frac{\delta}{4} + \lambda_j C|y_1| < \frac{\delta}{2},$$

for j sufficiently large. Thus, we get a contradiction. The proof is complete. □

Remark 4.4. It is extremely important for us to get (4.9) 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 (4.9).

Theorem 4.5. *Assume* (4.3) *and* (4.4). *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*

$$c - \delta \le H(y, p + Dw(y)) \le c + \delta \qquad \textit{in } \mathbb{R}^n. \tag{4.11}$$

Proof. We first prove the existence of c. For each $\lambda > 0$, let v^λ be the viscosity solution to (4.8). By the proof of Proposition 4.3, one has $|\lambda v^\lambda(0)| \leq C$ and (4.9). 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 $\|\lambda_j v^{\lambda_j} + c\|_{L^\infty(\mathbb{R}^n)} \leq \frac{\delta}{2}$. Let $w = v^{\lambda_j}$. It is clear that $w \in \mathrm{BUC}(\mathbb{R}^n)$, and w solves (4.11). 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

$$H(y, p + Dw_1(y)) \leq c_1 + \delta < c_2 - \delta \leq H(y, p + Dw_2(y)) \qquad \text{in } \mathbb{R}^n.$$

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, $w_1 \leq w_2$. By the same steps, $w_1 + C \leq w_2$ for any $C > 0$, which is absurd. Hence, the uniqueness of c is guaranteed. □

Definition 4.6. *Assume* (4.3) *and* (4.4)*. For each $p \in \mathbb{R}^n$, let c be the unique constant in Theorem 4.5. Denote by $\overline{H}(p) = c$. For each $\delta > 0$, let $w \in \mathrm{BUC}(\mathbb{R}^n)$ be a solution to* (4.11)*, that is, w solves*

$$\overline{H}(p) - \delta \leq H(y, p + Dw(y)) \leq \overline{H}(p) + \delta \qquad \text{in } \mathbb{R}^n. \tag{4.12}$$

We say that w is a δ-approximate corrector of the cell problem

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{R}^n. \tag{4.13}$$

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 (4.13). 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 (4.13), 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 (4.7).

3 Nonexistence of sublinear correctors

Let us now discuss about correctors, solutions to (4.13). In order for it to be useful, we need to require that correctors satisfy (4.7), 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 4.2. *Assume that $H(y, p) = H(p)$, where $H : \mathbb{R}^n \to \mathbb{R}$ is coercive. Let us study* (4.13) *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 4.7. *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 (4.13) *for $p = 0$ does not admit any sublinear solution.*

Proof. It is clear that H satisfies (4.3) and (4.4). 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,

$$y \mapsto v(y) + \eta(|y|^2 + 1)^{1/2}$$

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. 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 $y > \max\{y_0, 1\}$,

$$\frac{|v(y)|}{|y|} \ge \frac{1}{|y|}\left(\int_{y_0}^{y} V(s)\,ds - |v(y_0)|\right) \ge 2 - \frac{C}{|y|},$$

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 4.8. *Assume that H satisfies* (4.3) *and* (4.4). *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*

$$\begin{cases} u_t^\varepsilon(x,t) + H\left(\frac{x}{\varepsilon}, Du^\varepsilon(x,t)\right) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ \qquad\qquad u^\varepsilon(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.14}$$

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

$$\begin{cases} u_t + \overline{H}(Du) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) &= u_0(x) \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.15}$$

We present a proof of this theorem, which is basically a small modification to that of Theorem 3.6. 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, (4.15) has a unique Lipschitz solution u. As far as (4.14) is concerned, we have, as usual, the existence of a constant $C > 0$ independent of $\varepsilon > 0$ such that

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

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 (4.15).

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

$$\phi_t(x_0,t_0) + \overline{H}(D\phi(x_0,t_0)) \le 0.$$

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 (4.12) with this particular p where $\delta = \frac{\alpha}{2}$.

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 $\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right) \in B(x_0,r) \times \mathbb{R}^n \times (t_0 - r, t_0 + r)$ 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 $\Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, \frac{x_{\eta\varepsilon}}{\varepsilon}, t_{\eta\varepsilon}\right) \le \Phi^{\eta,\varepsilon}\left(x_{\eta\varepsilon}, y_{\eta\varepsilon}, t_{\eta\varepsilon}\right)$ for all $\eta > 0$, we obtain

$$\frac{1}{\eta}\left|y_{\eta\varepsilon} - \frac{x_{\eta\varepsilon}}{\varepsilon}\right|^2 \le 2\varepsilon \|v\|_{L^\infty(\mathbb{R}^n)} \qquad \Longrightarrow \qquad \lim_{\eta\to 0} y_{\eta\varepsilon} = \frac{x_\varepsilon}{\varepsilon}. \tag{4.16}$$

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 (4.14) gives

$$\phi_t\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + H\left(\frac{x_{\eta\varepsilon}}{\varepsilon}, D\phi\left(x_{\eta\varepsilon}, t_{\eta\varepsilon}\right) + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right) \le 0. \tag{4.17}$$

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

$$H\left(y_{\eta\varepsilon}, p + \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right) \ge \overline{H}(p) - \delta. \tag{4.18}$$

Besides, as v is Lipschitz, we infer

$$\left|\frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right)\right| \le C, \tag{4.19}$$

for some $C > 0$ independent of η, ε. By compactness, we can assume (up to passing to a subsequence again) that

$$\lim_{\eta\to 0} \frac{2}{\eta\varepsilon}\left(\frac{x_{\eta\varepsilon}}{\varepsilon} - y_{\eta\varepsilon}\right) = p_\varepsilon \in \mathbb{R}^n. \tag{4.20}$$

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 (4.20) to yield

$$u^\varepsilon(x,t) - \varepsilon v\left(\frac{x}{\varepsilon}\right) - \phi(x,t) \le u^\varepsilon\left(x_\varepsilon, t_\varepsilon\right) - \varepsilon v\left(\frac{x_\varepsilon}{\varepsilon}\right) - \phi\left(x_\varepsilon, t_\varepsilon\right)$$

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 (4.17) and (4.18) to get

$$\phi_t\left(x_\varepsilon, t_\varepsilon\right) + H\left(\frac{x_\varepsilon}{\varepsilon}, D\phi\left(x_\varepsilon, t_\varepsilon\right) + p_\varepsilon\right) \le 0,$$

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 4.9. 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 (4.16). The Lipschitz estimate of v was used to get (4.19) and (4.20).

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 3.10 in the periodic setting.

Theorem 4.10. *Assume that H satisfies* (4.3) *and* (4.4)*. Let $\overline{H}$ be its corresponding effective Hamiltonian. Then, for $p \in \mathbb{R}^n$,*

$$\begin{aligned}\overline{H}(p) &= \inf\{c \in \mathbb{R} \,:\, \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \le c \text{ in } \mathbb{R}^n \text{ in viscosity sense}\}\\ &= \sup\{c \in \mathbb{R} \,:\, \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \ge c \text{ in } \mathbb{R}^n \text{ in viscosity sense}\}.\end{aligned}$$

Since the proof of this is rather routine, we omit it. One can adapt the proof of Theorem 3.10 to this setting in a natural way. As H is coercive in p, one of the above formulas can also be written as

$$\overline{H}(p) = \inf\{c \in \mathbb{R} \,:\, \exists\, v \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n) : H(y, p + Dv(y)) \le c \text{ in } \mathbb{R}^n \text{ in viscosity sense}\}.$$

A consequence of this theorem is the following.

Corollary 4.11. *Assume that H satisfies* (4.3) *and* (4.4)*. 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

$$\inf_{y \in \mathbb{R}^n} H(y, p + D\phi(y)) \le \overline{H}(p) \le \sup_{y \in \mathbb{R}^n} H(y, p) \qquad \text{in } \mathbb{R}^n.$$

We apply Theorem 4.10 to conclude.

□

Theorem 4.12. *Assume that H satisfies* (4.3) *and* (4.4)*. 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

$$\left|\overline{H}(p)-\overline{H}(q)\right| \le \omega_R(|p-q|).$$

$\square$

It is clear from the above proof that the following corollary holds.

Corollary 4.13. *Assume that H satisfies* (4.3) *and* (4.4). *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 4.14. *Assume that H satisfies* (4.3) *and* (4.4). *Fix $p \in \mathbb{R}^n$. Consider the following Cauchy problem*

$$\begin{cases} w_t + H(y,p+Dw) &= 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty), \\ w(y,0) &= 0 \qquad \text{on } \mathbb{R}^n. \end{cases} \tag{4.21}$$

Let $w(y,t)$ be the unique viscosity solution to (4.21). *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 3.11 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 4.15 (The inf-sup formula)**.** *Assume that H satisfies* (4.3) *and* (4.4). *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*

$$\overline{H}(p) = \inf_{\phi \in C^1(\mathbb{R}^n)\cap \mathrm{BUC}(\mathbb{R}^n)} \sup_{y\in\mathbb{R}^n} H(y,p+D\phi(y)). \tag{4.22}$$

Proof. Pick any $\varphi \in C^1(\mathbb{R}^n)\cap \mathrm{BUC}(\mathbb{R}^n)$, by the representation formula in Theorem 4.10,

$$\overline{H}(p) \le \sup_{y\in\mathbb{R}^n} H(y,p+D\varphi(y)),$$

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 (4.12), 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 $\|Dv\|_{L^\infty(\mathbb{R}^n)} \leq C$, 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,

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

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 4.16. *Assume that H satisfies* (4.3) *and* (4.4). *Assume further that $p \mapsto H(y,p)$ is convex for every $y \in \mathbb{R}^n$. Then, for each $p \in \mathbb{R}^n$,*

$$\overline{H}(p) = \inf\{c \in \mathbb{R} \,:\, \exists\, v \in \mathrm{Lip}(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n) : H(y, p + Dv(y)) \leq c \text{ a.e. in } \mathbb{R}^n\}. \quad (4.23)$$

By using the above corollary, we deduce that $\overline{H}$ is also convex.

Theorem 4.17 (Convexity of $\overline{H}$). *Assume that H satisfies* (4.3) *and* (4.4). *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 4.18. *Assume that H satisfies* (4.3) *and* (4.4). *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 33. *Give a detailed proof of Theorem 4.10.*

Exercise 34. *Give a detailed proof of Theorem 4.14.*

Exercise 35. *Give a quick proof of Theorem 4.17.*

6 References

1. Almost periodic homogenization for Hamilton–Jacobi equations was studied first by Ishii [60].

2. The result on nonexistence of sublinear correctors was pointed out by Lions and Souganidis [75].

CHAPTER 5

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

$$\begin{cases} \lim_{|p|\to\infty} \left(\min_{y\in\mathbb{T}^n} H(y,p) \right) = +\infty, \\ p \mapsto H(y,p) \text{ is convex for all } y \in \mathbb{T}^n. \end{cases} \tag{5.1}$$

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

$$\lambda v^\lambda + H(y, Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n. \tag{5.2}$$

Of course, this equation has been one of the central objects of all previous chapters. In light of (5.1), (5.2) 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 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 (5.2) 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

$$\tilde{H}(y,p) = \max_{|v| \le h}\big(p \cdot v - \tilde{L}(y,v)\big) \qquad \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{5.3}$$

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 (5.3) 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,

$$H(y,p) = \max_{|v| \le h}(p \cdot v - L(y,v)) \qquad \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{5.4}$$

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 (5.3) 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 (5.1). 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 5.1. *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 5.2. *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 5.3. *Assume* (5.4) *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* (5.2)*. Then, for $z\in\mathbb{T}^n$,*

$$\lambda v^\lambda(z) = \min_{\mu\in\mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu(y,v). \tag{5.5}$$

Let us now proceed to prove the preparatory lemmas and this theorem. After our preparations in previous chapter, Lemmas 5.1 and 5.2 are not so hard to prove. Nevertheless, let us give complete proofs here.

Proof of Lemma 5.1. Pick $(\phi_1,u_1),(\phi_2,u_2)\in\mathcal{F}_\lambda$. For $i=1,2$, as H_{ϕ_i} satisfies (5.1), u_i is Lipschitz in $\mathbb{T}^n$. Moreover, in light of Theorem 2.24, $u_i\in\mathrm{Lip}(\mathbb{T}^n)$ 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. $y \in \mathbb{T}^n$.

$$
\begin{aligned}
&\lambda\frac{u_1(y)+u_2(y)}{2} + H_{\frac{\phi_1+\phi_2}{2}}\left(y, \frac{Du_1(y)+Du_2(y)}{2}\right)\\
&= \lambda\frac{u_1(y)+u_2(y)}{2} + \max_{|v|\le h}\left(\frac{Du_1(y)+Du_2(y)}{2}\cdot v - \frac{\phi_1(y,v)+\phi_2(y,v)}{2}\right)\\
&\le \frac{1}{2}\Big((u_1(y) + \max_{|v|\le h}(Du_1(y)\cdot v - \phi_1(y,v)) + (u_2(y) + \max_{|v|\le h}(Du_2(y)\cdot v - \phi_2(y,v)))\Big) \le 0.
\end{aligned}
$$

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 5.2. 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 5.1.

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 5.3. Firstly, as v^λ is the solution to (5.2), $(L, v^\lambda) \in \mathcal{F}_\lambda$. 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

$$
\lambda v^\lambda(z) + \varepsilon < \min_{\mu\in\mathcal{P}\cap\mathcal{G}'_{z,\lambda}} \int_{\mathbb{T}^n\times\overline{B}_h} L(y,v)\,d\mu(y,v). \tag{5.6}
$$

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 (5.6) 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 5.4. 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 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 (5.7) 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 (5.2). As usual, we assume (5.1). By the reduction step, we may assume that (5.4) holds true.

Let $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ be the unique solution to (5.2), that is,

$$\lambda v^\lambda + H(y, Dv^\lambda) = 0 \qquad \text{in } \mathbb{T}^n.$$

By Corollary 3.5 (or Lemma 3.52), 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 5.5. *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 5.6. *Assume* (5.4) *for some $h > 0$ and $L \in C(\mathbb{T}^n \times \overline{B}_h)$. Then,*

$$\min_{\mu \in \mathcal{P} \cap \mathcal{G}'_0} \int_{\mathbb{T}^n \times \overline{B}_h} L(y, v)\, d\mu(y, v) = -\overline{H}(0). \tag{5.7}$$

The proof of this is quite similar to that of Theorem 5.3. Let us sketch it here.

Proof. Firstly, let $w \in \mathrm{Lip}\,(\mathbb{T}^n)$ be a solution to the cell problem

$$H(y, Dw) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n. \tag{5.8}$$

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

$$-\overline{H}(0) + \varepsilon < \min_{\mu \in \mathcal{P} \cap \mathcal{G}_0'} \int_{\mathbb{T}^n \times \overline{B}_h} L(y, v)\, d\mu(y, v). \tag{5.9}$$

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 (5.9) 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 5.7. *Let $\mu \in \mathcal{G}_0'$. Then,*

$$\int_{\mathbb{T}^n \times \overline{B}_h} v \cdot D\psi(y)\, d\mu(y, v) = 0 \qquad \text{for all } \psi \in C^2(\mathbb{T}^n).$$

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. □

Next we show that we have stability of measures in the cones $\mathcal{G}_{z,\lambda}'$ as $\lambda \to 0$.

Lemma 5.8. *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 3.4), 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 (5.8). As (5.8) often has many solutions, 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 5.9. *Assume* (5.1). *For* $\lambda > 0$*, let* $v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ *be the unique solution to* (5.2). *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 (5.4) for some $h > 0$ and $L \in C(\mathbb{T}^n \times \overline{B}_h)$.

Since $\overline{H}(0) = 0$, we have that

$$\|v^\lambda\|_{L^\infty(\mathbb{T}^n)} + \|Dv^\lambda\|_{L^\infty(\mathbb{T}^n)} \leq C.$$

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 $u(z) \geq w(z)$ for each $z \in \mathbb{T}^n$. There exist $\{\lambda_j\} \to 0$ and $\{\delta_j\} \to 0$ such that $v^{\lambda_j} \to v$ and $v^{\delta_j} \to w$ in $C(\mathbb{T}^n)$ as $j \to \infty$. By Theorem 5.3, 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 5.8, $\mu \in \mathcal{G}'_0$. 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

$$0 \leq \langle \mu, L - \delta_j v^{\delta_j} \rangle = -\delta_j \langle \mu, v^{\delta_j} \rangle,$$

and

$$0 \leq \langle \mu_j, L + \lambda_j w - \lambda_j w(z) \rangle = \lambda_j (v^{\lambda_j}(z) - w(z)) + \lambda_j \langle \mu_j, w \rangle.$$

Therefore,

$$\langle \mu, v^{\delta_j} \rangle \leq 0 \quad \text{and} \quad v^{\lambda_j}(z) - w(z) + \langle \mu_j, w \rangle \geq 0.$$

Let $j \to \infty$ to deduce further that

$$\langle \mu, w \rangle \leq 0 \quad \text{and} \quad u(z) - w(z) + \langle \mu, w \rangle \geq 0,$$

which implies $u(z) \geq w(z)$. The proof is complete. □

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 (3.10), that is,

$$H(y, p + Dv(y)) = \overline{H}(p) \qquad \text{in } \mathbb{T}^n. \tag{5.10}$$

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 (5.10).

In this section, we assume a stronger condition that

$$\begin{cases} H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{pp}H(y,p) \le \theta^{-1} I_n \text{ for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{5.11}$$

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 5.10. *Assume* (5.11). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (5.10). *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*

$$p \cdot \xi(t_1) + v(\xi(t_1)) - p \cdot \xi(t_2) - v(\xi(t_2)) = \int_{t_2}^{t_1} \left(L(\xi(t), \xi'(t)) + \overline{H}(p) \right) dt \tag{5.12}$$

for all $t_2 < t_1 \le 0$.

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 $\|\xi'\|_{L^\infty((-\infty,0])} \le C$. Furthermore, by the Dynamic Programming Principle,

$$v(\xi(-k+1)) - \overline{H}(0) = \int_s^1 L(\xi(r), \xi'(r))\, dr + v(\xi(-k+s)) - \overline{H}(0)s \quad \text{for all } k \in \mathbb{N},\, s \in [0,1].$$

Thus, for all $t_2 < t_1 \le 0$,

$$v(\xi(t_1)) - v(\xi(t_2)) = \int_{t_2}^{t_1} \left(L(\xi(t), \xi'(t)) + \overline{H}(0) \right) dt.$$

□

3.2 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 (5.10).

Theorem 5.11. *Assume* (5.11). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (5.10). *Fix $x \in \mathbb{T}^n$, and let ξ be a backward characteristic of v starting from x. Then, there exists 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 (5.11), 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 5.12. *Assume* (5.11). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (5.10). *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 5.11.

Lemma 5.13. *Assume* (5.11). *For a fixed $p \in \mathbb{R}^n$, let $v \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to* (5.10). *Let $\gamma : (-\infty, 0] \to \mathbb{R}^n$ be an arbitrary Lipschitz curve. Then, for every $T > 0$,*

$$\int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + \overline{H}(p) \right) dt \ge p \cdot (\gamma(0) - \gamma(-T)) + v(\gamma(0)) - v(\gamma(-T)).$$

Heuristically, if everything is smooth, then this result is not hard to prove. Indeed,

$$\int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + \overline{H}(p) \right) dt = \int_{-T}^{0} \left(L(\gamma(t), \gamma'(t)) + H(\gamma(t), p + Dv(\gamma(t))) \right) dt$$
$$\ge \int_{-T}^{0} \gamma'(t) \cdot (p + Dv(\gamma(t)))\, dt = p \cdot (\gamma(0) - \gamma(-T)) + v(\gamma(0)) - v(\gamma(-T)).$$

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,

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

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 2.24, we infer that v^ε satisfies

$$H(y, p + Dv^\varepsilon(y)) \le \overline{H}(p) + C\varepsilon \qquad \text{in } \mathbb{T}^n.$$

We can now perform a similar computation as the heuristic one above

$$\begin{aligned}
&\int_{-T}^0 \left(L(\gamma(t),\gamma'(t)) + \overline{H}(p)\right) dt \ge \int_{-T}^0 \left(L(\gamma(t),\gamma'(t)) + H(\gamma(t), p + Dv^\varepsilon(\gamma(t)) - C\varepsilon\right) dt \\
&\ge -CT\varepsilon + \int_{-T}^0 \gamma'(t)\cdot(p + Dv^\varepsilon(\gamma(t))\,dt \\
&= -CT\varepsilon + p\cdot(\gamma(0) - \gamma(-T)) + v^\varepsilon(\gamma(0)) - v^\varepsilon(\gamma(-T)).
\end{aligned}$$

Let $\varepsilon \to 0$ in the above to conclude. □

Remark 5.14. In fact, Lemma 5.13 holds if we only require that $v \in \operatorname{Lip}(\mathbb{T}^n)$ to be a subsolution to (5.10) 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 5.11.

Proof of Theorem 5.11. To make it clear, we write v_p to denote a solution to (5.10).

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 $v_{\tilde{p}} \in \operatorname{Lip}(\mathbb{T}^n)$ be a solution to the corresponding cell problem with $\min_{\mathbb{T}^n} v_{\tilde{p}} = 0$. Lemma 5.13 gives that

$$\tilde{p}\cdot(\xi(0) - \xi(t)) + v_{\tilde{p}}(\xi(0)) - v_{\tilde{p}}(\xi(t)) \le \int_t^0 \left(L(\xi(s),\xi'(s)) + \overline{H}(\tilde{p})\right) ds$$

Thus, for $\tilde{p} \in B(p,1)$,

$$\overline{H}(\tilde{p}) - \overline{H}(p) \ge (\tilde{p} - p)\cdot\frac{\xi(t) - \xi(0)}{t} - \frac{C}{|t|}. \tag{5.13}$$

Besides, the fact that $\|\xi'\|_{L^\infty((-\infty,0])} \le C$ implies

$$\left|\frac{\xi(t)-\xi(0)}{t}\right| \le C \qquad \text{for all } t<0.$$

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 $|q| \le C$. Plug this into (5.13) to yield

$$\overline{H}(\tilde{p}) - \overline{H}(p) \ge (\tilde{p}-p)\cdot q \qquad \text{for all } \tilde{p} \in B(p,1),$$

which means that $q \in D^-\overline{H}(p)$.

□

Remark 5.15. 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 5.16. *Assume* (5.11). *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* (5.10). *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*

$$\left|\frac{\xi(t)}{t} - D\overline{H}(p)\right| \le \frac{C}{|t|^{1/2}} \qquad \textit{for all } t<0.$$

Proof. This is essentially a quantitative version of Theorem 5.11. 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 (5.13), 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.

□

Here is an immediate corollary.

Corollary 5.17. *Assume* (5.11). *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* (5.10). *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*

$$\left|\frac{\xi(t)}{t}-D\overline{H}(p)\right|\le\frac{C}{|t|}\qquad \textit{for all } t<0.$$

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.

□

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 (5.1) here. Nevertheless, for simplicity, we assume that H satisfies (5.11) in this section. Let us recall quickly the homogenization problem.

For each $\varepsilon>0$, we study

$$\begin{cases} u^\varepsilon_t(x,t)+H\left(\frac{x}{\varepsilon},Du^\varepsilon(x,t)\right) &=0 \qquad \text{in } \mathbb{R}^n\times(0,\infty),\\ u^\varepsilon(x,0) &=u_0(x)\qquad \text{on } \mathbb{R}^n.\end{cases}\tag{5.14}$$

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 $\varepsilon\to0+$ and quantify the rate of convergence of u^ε to u, which solves a (simpler) effective equation

$$\begin{cases} u_t+\overline{H}(Du) &=0 \qquad \text{in } \mathbb{R}^n\times(0,\infty),\\ u(x,0) &=u_0(x)\qquad \text{on } \mathbb{R}^n.\end{cases}\tag{5.15}$$

Here is the main result of this section.

Theorem 5.18. *Assume* (5.11) *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* (5.14)*. Let* u *be the viscosity solution to* (5.15)*. 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,

$$u^\varepsilon(x,t) \geq u(x,t) - C\varepsilon \qquad \text{for all } (x,t) \in \mathbb{R}^n \times [0,\infty). \tag{5.16}$$

(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*

$$u^\varepsilon(x,t) \leq u(x,t) + C_p\sqrt{t\varepsilon} + C\varepsilon. \tag{5.17}$$

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*

$$u^\varepsilon(x,t) \leq u(x,t) + \widetilde{C}_p t\varepsilon + C\varepsilon. \tag{5.18}$$

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 Preparations

By the comparison principle, it is straightforward that

$$\|u_t^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} \leq C_0.$$

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 (5.11), we can make $\theta > 0$ smaller if needed to have

$$\frac{\theta}{2}|p|^2 - K_0 \leq H(y,p) \leq \frac{1}{2\theta}|p|^2 + K_0 \qquad \text{for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{5.19}$$

for some $K_0 > 1$. Then, we also have that

$$\frac{\theta}{2}|p|^2 - K_0 \leq \overline{H}(p) \leq \frac{1}{2\theta}|p|^2 + K_0 \qquad \text{for all } p \in \mathbb{R}^n. \tag{5.20}$$

We use (5.19) and (5.20) to get that, for each $v_p \in \mathrm{Lip}(\mathbb{T}^n)$ solving (5.10),

$$\|Dv_p\|_{L^\infty(\mathbb{T}^n)} \leq C(|p| + K_0).$$

In particular,

$$\max_{\mathbb{T}^n} v_p - \min_{\mathbb{T}^n} v_p \leq C\sqrt{n}(|p| + K_0) = C(|p| + K_0). \tag{5.21}$$

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

$$\frac{\theta}{2}|v|^2 - K_0 \leq L(y,v) \leq \frac{1}{2\theta}|v|^2 + K_0 \quad \text{for all } (y,v) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{5.22}$$

and

$$\frac{\theta}{2}|v|^2 - K_0 \leq \overline{L}(v) \leq \frac{1}{2\theta}|v|^2 + K_0 \quad \text{for all } q \in \mathbb{R}^n.$$

For $(x,t) \in \mathbb{R}^n \times (0,\infty)$, the optimal control formula for the solution to (5.14) implies

$$u^\varepsilon(x,t) = \inf_{\substack{\varepsilon\eta(0)=x \\ \eta\in \mathrm{AC}([-\varepsilon^{-1}t,0])}} \left\{ u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}t\right)\right) + \varepsilon \int_{-\varepsilon^{-1}t}^0 L(\eta(s),\eta'(s))\,ds \right\}. \tag{5.23}$$

4.2 Proof of Theorem 5.18

We divide the proof into two parts. We first derive the lower bound (5.16), which is of course optimal.

Proof of optimal lower bound (5.16). 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 (5.16) holds for $(x,t) = (0,1)$. In other words, we aim at showing

$$u^\varepsilon(0,1) - u(0,1) \geq -C\varepsilon. \tag{5.24}$$

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

$$|u_0(x)| \leq C|x| \qquad \text{for all } x \in \mathbb{R}^n. \tag{5.25}$$

The optimal control formula (5.23) 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 (5.22) 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 (5.25) 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,

$$u^\varepsilon(0,1) = \inf_{\substack{\eta(0)=0, \\ \varepsilon\left|\eta(-\varepsilon^{-1})\right|\leq C}} \left\{ u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) + \varepsilon \int_{-\varepsilon^{-1}}^0 L(\eta(t),\eta'(t))\,dt \right\}. \tag{5.26}$$

Clearly, there exists $C_1 > 0$ such that for any $|v| \leq C$,

$$\overline{L}(v) = \sup_{p\in\mathbb{R}^n} \left\{p \cdot v - \overline{H}(p)\right\} = \sup_{|p|\leq C_1} \{p \cdot v - \overline{H}(p)\}. \tag{5.27}$$

This is important as it means that we only need to deal with $|p| \le C_1$. For $p \in \mathbb{R}^n$, let $v_p \in \mathrm{Lip}(\mathbb{T}^n)$ be a viscosity solution to (5.10) such that $v_p(0) = 0$. Then for any Lipschitz continuous curve $\eta : [-\varepsilon^{-1}, 0] \to \mathbb{R}^n$, Lemma 5.13 gives

$$\int_{-\varepsilon^{-1}}^{0} \left(L(\eta(t), \eta'(t)) + \overline{H}(p)\right) dt \ge p \cdot \eta(0) - p \cdot \eta\left(-\varepsilon^{-1}\right) + v_p(\eta(0)) - v_p\left(\eta\left(-\varepsilon^{-1}\right)\right).$$

Therefore, if we assume further that $\eta(0) = 0$ and $\varepsilon \left|\eta(-\varepsilon^{-1})\right| \le C$, then we are able to combine the above with (5.21) and (5.27) to yield

$$\begin{aligned}
\varepsilon \int_{-\varepsilon^{-1}}^{0} (L(\eta(t), \eta'(t))\, dt &\ge \sup_{p \in \mathbb{R}^n} \left\{p \cdot \left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) - \overline{H}(p) + \varepsilon v_p(0) - \varepsilon v_p\left(\eta\left(-\varepsilon^{-1}\right)\right)\right\} \\
&\ge \sup_{|p| \le C_1} \left\{p \cdot \left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) - \overline{H}(p) + \varepsilon v_p(0) - \varepsilon v_p\left(\eta\left(-\varepsilon^{-1}\right)\right)\right\} \\
&\ge \overline{L}\left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) - C\varepsilon.
\end{aligned}$$

Plug this into (5.26) to imply

$$\begin{aligned}
u^\varepsilon(0,1) &\ge \inf_{\substack{\eta(0)=0, \\ \varepsilon\left|\eta(-\varepsilon^{-1})\right| \le C}} \left\{u_0\left(\varepsilon\eta\left(-\varepsilon^{-1}\right)\right) + \overline{L}\left(-\varepsilon\eta\left(-\varepsilon^{-1}\right)\right)\right\} - C\varepsilon \\
&\ge \inf_{y \in \mathbb{R}^n} \left\{u_0(y) + \overline{L}(-y)\right\} - C\varepsilon \\
&= u(0,1) - C\varepsilon.
\end{aligned}$$

The last equality in the above holds thanks to the Hopf–Lax formula for u. □

We now proceed to prove upper bounds (5.17) and (5.18). Again, this is just a conditionally optimal upper bound. The following lemma is a key step toward proving (5.17) and (5.18). Once it is proved, we can combine it with Theorem 5.16 to conclude right away.

Lemma 5.19. *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* (5.10) *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

$$u^\varepsilon(x,t) \le u(x,t) + CC_p t^{1-\alpha}\varepsilon^\alpha + C\varepsilon. \tag{5.28}$$

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

$$u^\varepsilon(x,t) \le u(x,t) + MC_p^2 t^{2(1-\alpha)}\varepsilon^{2\alpha} + C\varepsilon. \tag{5.29}$$

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 (5.29). 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 5.19 and approximations, we have that

Corollary 5.20. *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* (5.10) *and a backward characteristic* $\xi : (-\infty,0] \to \mathbb{R}^n$ *of* v_p *such that, for some* $C > 0$ *independent of* p,

$$\left|\frac{\xi(s)-\xi(0)}{s} - D\overline{H}(p)\right| \le \frac{C}{|s|} \qquad \textit{for all } s < 0.$$

Then

$$u^\varepsilon(x,t) \le u(x,t) + C\varepsilon \quad \textit{for all } (x,t) \in \mathbb{R}^n \times [0,\infty). \tag{5.30}$$

We are now ready to obtain (5.17) and (5.18).

Proof of upper bounds (5.17) *and* (5.18). Inequalities (5.17) and (5.18) follow immediately from Lemma 5.19 and Theorem 5.16. □

5 References

1. The new representation formula for solutions of the discount problems based on a duality method was derived by Ishii, Mitake, Tran [62, 63].

2. Sion's minimax theorem was derived by Sion [88]. We will give a detailed proof of this in Appendix.

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 [27] used weak KAM theory to obtain the result for first-order convex Hamilton–Jacobi equations. Mitake, Tran [82] used the nonlinear adjoint method to get the convergence for possibly degenerate viscous Hamilton–Jacobi equations. Al-Aidarous, Alzahrani, Ishii, Younas [1] studied the first-order problem with Neumann boundary condition. The proof presented here is completely different and due to Ishii, Mitake, Tran [62, 63]. This duality approach [62, 63] works for fully nonlinear second order equations as well. See the lecture notes of Le, Mitake, Tran [72] for an account of this via the nonlinear adjoint method. More references and related areas are provided there as well.

4. Backward characteristics of solutions to cell problems here is done based on the optimal control formula of a corresponding time-dependent problem. Fathi [40] gave a different approach based on the weak KAM theorem. See also the lecture notes of Ishii [61].

5. Rate of convergence of large time average of backward characteristics was taken from Gomes [49], Mitake, Tran, Yu [84].

6. The section on optimal rate of convergence in periodic homogenization theory was taken from Mitake, Tran, Yu [84]. See [84] for further results in one and two dimensional spaces that we do not cover here. For a more complicated situation in one dimension, see the work of Tu [93].

CHAPTER 6

Introduction to weak KAM theory

1 Introduction

In this chapter, we always assume that

$$\begin{cases} H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{pp} H(y,p) \le \theta^{-1} I_n \text{ for all } (y,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{6.1}$$

Let $L = L(y,v)$ be the corresponding Lagrangian. By changing $\theta > 0$ to be smaller if needed, we may also assume that

$$\begin{cases} L \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{vv} L(y,v) \le \theta^{-1} I_n \text{ for all } (y,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{6.2}$$

Let us give a minimalistic type introduction to this subject. We are concerned with the following Hamiltonian system

$$\begin{cases} x'(t) = D_p H(x(t), p(t)), \\ p'(t) = -D_x H(x(t), p(t)). \end{cases} \tag{6.3}$$

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

$$H(x, P + Dv(x,P)) = \overline{H}(P) \qquad \text{in } \mathbb{T}^n. \tag{6.4}$$

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 (6.3) into the following integrable system

$$\begin{cases} X'(t) = D\overline{H}(P(t)), \\ P'(t) = 0. \end{cases} \tag{6.5}$$

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, (6.4) 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 (6.1), 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 (6.4), and it is extremely important to come up with weak interpretations of the classical program briefly mentioned above. Various great works of Aubry [6], Mather [78, 79], Mañé [77], Fathi [39, 40], E [28], Evans, Gomes [36] show that some solutions of (6.3), 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 (6.4) 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 (6.4) 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 [40]. 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 6.1. *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 6.2. *Assume* (6.1). *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

$$\frac{d}{dt}\left(D_v L(\xi(t),\xi'(t))\right) = D_x L(\xi(t),\xi'(t)) \qquad \text{for all } t \in [0,T].$$

Denote by $x(t) = \xi(t)$, and $p(t) = D_v L(\xi(t),\xi'(t))$ for $t \in [0,T]$. Then (x,p) solves the following Hamiltonian system

$$\begin{cases} x'(t) = D_p H(x(t),p(t)), \\ p'(t) = -D_x H(x(t),p(t)), \end{cases} \qquad \text{for } t \in [0,T].$$

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

$$\frac{d}{dt}H(x(t),p(t)) = D_x H(x(t),p(t))\cdot x'(t) + D_p H(x(t),p(t))\cdot p'(t) = 0.$$

In particular, this allows us to get that $H(x(t),p(t)) \le C_T$, which implies $|p(t)| \le C_T$, and also $|\xi'(t)| \le C_T$ for all $t \in [0,T]$. $\square$

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 6.3. *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 6.4 (Properties of the Lax–Oleinik semigroup). *Assume* (6.1). *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 [40] via the method of finding a fixed point for the Lax–Oleinik semigroup.

Theorem 6.5 (Weak KAM theorem). *Assume* (6.1). *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 5.10 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 (6.4), 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 5.10, 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

$$v(\xi(t_1)) - v(\xi(t_2)) = \int_{t_2}^{t_1} \left(L(\xi(t), \xi'(t)) + \overline{H}(0) \right) dt \tag{6.6}$$

for all $t_2 < t_1 \leq 0$. We show that v is differentiable at $\xi(t)$ for $t < 0$.

Theorem 6.6. *Assume* (6.1). *Let* $P = 0$, *and* $v = v(x, 0) \in \mathrm{Lip}(\mathbb{T}^n)$ *be a solution of the corresponding cell problem* (6.4). *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 5.13, $\phi \geq v$, 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 5.13, $\psi \leq v$, 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 $Dv(y) = D_v L(\xi(t), \xi'(t))$. □

Remark 6.7. 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 5.11 and Corollary 5.12, we know that if $\overline{H}$ is differentiable at P, then for a backward characteristic v of (6.4),

$$\lim_{t \to -\infty} \frac{\xi(t)}{t} = D\overline{H}(P).$$

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 [56]. See also Bangert [8], E [28], Mitake, Tran, Yu [84]. We will discuss about 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

$$\begin{cases} \frac{d}{dt}\left(D_v L(x(t), x'(t))\right) = D_x L(x(t), x'(t)), \\ x(0) = x, x'(0) = v. \end{cases} \tag{6.7}$$

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 6.8. *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 6.9. *Assume* (6.1). *Then,*

$$\overline{H}(0) = -\inf\left\{\int_{\mathbb{T}^n\times\mathbb{R}^n} L(x, v)\, d\mu(x, v) \,:\, \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \text{ is flow invariant}\right\}. \tag{6.8}$$

This result is of course quite similar to Theorem 5.6 in the previous chapter. We will go back to this point later.

Proof. Take v_- to be a solution to (6.4) with $P = 0$. Or in other words, v_- is taken from Theorem 6.5. By Lemma 5.13, for $x(\cdot)$ solves (6.7),

$$v_-(x(1)) - v_-(x(0)) \le \int_0^1 \left(L(x(s), x'(s)) + \overline{H}(0)\right) ds.$$

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 $\text{spt}(\mu_t) \subset \mathbb{T}^n \times \overline{B}(0,C)$ 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 $\text{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. 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 6.10. 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.

2.3 Mather measures and Mather set

We are now ready to define Mather measures and Mather set.

Definition 6.11. *Each measure μ that minimizes* (6.14) *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 6.12. *Assume* (6.1). *Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a subsolution to* (6.4) *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 5.14,

$$u(\pi(\Phi_{t'}(x, v))) - u(\pi(\Phi_t(x, v))) \le \int_t^{t'} \left(L(\Phi_s(x, v)) + \overline{H}(0)\right) ds. \tag{6.9}$$

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 (6.9) 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 6.13. *Assume* (6.1). *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 (6.4) with $P = 0$. We use Lemma 6.12 and repeat Theorem 6.6 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 (6.4) with $P = 0$.

Theorem 6.14 (Uniqueness set for (6.4) with $P = 0$)**.** *Assume* (6.1)*. Let $u_1, u_2 \in \mathrm{Lip}(\mathbb{T}^n)$ be two solutions to* (6.4) *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 $t < 0$,

$$u_1(x) - u_1(\xi(t)) = \int_t^0 \big(L(\xi(s), \xi'(s)) + \overline{H}(0)\big)\, ds,$$

and

$$u_2(x) - u_2(\xi(t)) \le \int_t^0 \big(L(\xi(s), \xi'(s)) + \overline{H}(0)\big)\, ds.$$

Combine these two to infer that

$$u_2(x) - u_1(x) \le u_2(\xi(t)) - u_1(\xi(t)) \qquad \text{for all } t \le 0.$$

Let us now use the construction in the later part of the proof of Theorem 6.9 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

$$u_2(x) - u_1(x) \le \langle \mu_t, (u_2 - u_1) \circ \pi \rangle = \frac{1}{|t|} \int_t^0 (u_2 - u_1)(\pi \circ (\xi(s), \xi'(s)))\, ds.$$

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

$$u_2(x) - u_1(x) \le \langle \mu, (u_2 - u_1) \circ \pi \rangle = \int_{\mathbb{T}^n \times \mathbb{R}^n} (u_2 - u_1)(x)\, d\mu(x, v) = 0,$$

by our hypothesis. Thus, $u_2(x) \le u_1(x)$. By a symmetric argument, $u_1(x) \le u_2(x)$, and hence, $u_1(x) = u_2(x)$. □

2.4 Lipschitz graph theorem

Theorem 6.15. *Assume* (6.1)*. Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a subsolution to* (6.4) *for $P = 0$. There exists $C > 0$ depending only on H such that, for all $x \in \mathcal{M}_0$ and $h \in \mathbb{R}^n$,*

$$|u(x+h) + u(x-h) - 2u(x)| \le C|h|^2.$$

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 6.12,

$$u(x(1)) - u(x(0)) = \int_0^1 \big(L(x(s), x'(s)) + \overline{H}(0)\big)\, ds, \tag{6.10}$$

and

$$u(x(0))-u(x(-1))=\int_{-1}^{0}\left(L(x(s),x'(s))+\overline{H}(0)\right)ds. \tag{6.11}$$

Let us obtain first the lower bound. By Lemma 5.13,

$$u(x(1))-u(x(0)+h)\le\int_{0}^{1}\left(L(x(s)+(1-s)h,x'(s)-h)+\overline{H}(0)\right)ds,$$

and

$$u(x(1))-u(x(0)-h)\le\int_{0}^{1}\left(L(x(s)-(1-s)h,x'(s)+h)+\overline{H}(0)\right)ds.$$

Combine these two inequalities with (6.10) to get

$$\begin{aligned}
&u(x+h)+u(x-h)-2u(x)\\
&\ge\int_{0}^{1}\left(2L(x(s),x'(s))-L(x(s)+(1-s)h,x'(s)-h)-L(x(s)-(1-s)h,x'(s)+h)\right)ds\\
&\ge -C|h|^2.
\end{aligned}\tag{6.12}$$

On the other hand, use Lemma 5.13 again to yield

$$u(x(0)+h)-u(x(-1))\le\int_{-1}^{0}\left(L(x(s)+(1+s)h,x'(s)+h)+\overline{H}(0)\right)ds,$$

and

$$u(x(0)-h)-u(x(-1))\le\int_{-1}^{0}\left(L(x(s)-(1+s)h,x'(s)-h)+\overline{H}(0)\right)ds.$$

The above two inequalities, together with (6.11), imply

$$\begin{aligned}
&u(x+h)+u(x-h)-2u(x)\\
&\le\int_{0}^{1}\left(L(x(s)+(1+s)h,x'(s)+h)-L(x(s)-(1+s)h,x'(s)-h)-2L(x(s),x'(s))\right)ds\\
&\le C|h|^2.
\end{aligned}\tag{6.13}$$

The lower bound (6.12) and the upper bound (6.13) give us the desired result. □

The following Lipschitz graph theorem is due to Mather.

Theorem 6.16. *Assume* (6.1). *Let* $u\in\mathrm{Lip}(\mathbb{T}^n)$ *be a subsolution to* (6.4) *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,

$$\begin{aligned}
u(x+h)-u(x) &\geq \int_0^1 \big(L(x(s),x'(s))-L(x(s)+(1-s)h,x'(s)-h)\big)\,ds\\
&\geq \int_0^1 \big(D_xL(x(s),x'(s))\cdot(s-1)h+D_vL(x(s),x'(s))\cdot h\big)\,ds - C|h|^2\\
&= \int_0^1 \left(\frac{d}{ds}\big(D_vL(x(s),x'(s))\big)\cdot(s-1)h+D_vL(x(s),x'(s))\cdot h\right)ds - C|h|^2\\
&= \int_0^1 \frac{d}{ds}\big(D_vL(x(s),x'(s))\cdot(s-1)h\big)\,ds - C|h|^2\\
&= D_vL(x(0),x'(0))\cdot h - C|h|^2 = Du(x)\cdot h - C|h|^2.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
u(x+h)-u(x) &\leq \int_{-1}^0 \big(L(x(s)+(1+s)h,x'(s)+h)-L(x(s),x'(s))\big)\,ds\\
&\leq \int_{-1}^0 \big(D_xL(x(s),x'(s))\cdot(s+1)h+D_vL(x(s),x'(s))\cdot h\big)\,ds + C|h|^2\\
&= \int_{-1}^0 \left(\frac{d}{ds}\big(D_vL(x(s),x'(s))\big)\cdot(s+1)h+D_vL(x(s),x'(s))\cdot h\right)ds + C|h|^2\\
&= \int_{-1}^0 \frac{d}{ds}\big(D_vL(x(s),x'(s))\cdot(s+1)h\big)\,ds + C|h|^2\\
&= D_vL(x(0),x'(0))\cdot h + C|h|^2 = Du(x)\cdot h + C|h|^2.
\end{aligned}$$

Thus,

$$|u(x+h)-u(x)-Du(x)\cdot h|\leq C|h|^2,$$

which completes part (i). For part (ii), note that, for $x,y\in\mathcal{M}_0$,

$$|u(y)-u(x)-Du(x)\cdot(y-x)|\leq C|y-x|^2,$$

and

$$|u(x)-u(y)-Du(y)\cdot(x-y)|\leq C|y-x|^2.$$

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.

2.5 A relaxed problem

A disadvantage of the flow invariant property (Definition 6.8) is that it is a nonlinear constraint that depends on L (and hence H). For this reason, Mañé [77] 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 6.17. *Assume* (6.1). *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 6.9

Theorem 6.18. *Assume* (6.1). *Then,*

$$\overline{H}(0) = -\min_{\nu \in \mathcal{F}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) \, d\nu(x,v). \tag{6.14}$$

Roughly speaking, this is very close to Theorem 5.6 in the previous chapter.

Proof. Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (6.4) 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's 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

$$\min_{\nu \in \mathcal{F}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\nu(x,v) \geq -\overline{H}(0).$$

The reverse inequality follows immediately from the later part of the proof of Theorem 6.9 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 $t < 0$,

$$u(\xi(0)) - u(\xi(t)) = \int_t^0 \big(L(\xi(s), \xi'(s)) + \overline{H}(0)\big)\, ds.$$

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 $\|\xi'\|_{L^\infty((-\infty,0])} \leq C$. 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

$$\begin{aligned}
&\int_{\mathbb{T}^n \times \mathbb{R}^n} v \cdot D\varphi(x)\, d\mu(x,v) = \lim_{k\to\infty} \frac{1}{|t_k|} \int_{t_k}^0 \xi'(s) \cdot D\varphi(\xi(s))\, ds \\
= &\lim_{k\to\infty} \frac{1}{|t_k|} \left(\varphi(\xi(0)) - \varphi(\xi(t_k))\right) = 0.
\end{aligned}$$

□

Theorem 6.19. *Assume* (6.1). *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 6.20. *Assume* (6.1). *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 6.21. *Assume the settings in Theorem 6.19. Let* $u \in \mathrm{Lip}(\mathbb{T}^n)$ *be a solution to* (6.15)*. Then, for* $(x,v) \in \mathrm{spt}(\nu)$*, we have* u *is differentiable at* x *and* $Du(x) = D_vL(x,v)$*. Moreover,* ν *is supported on a Lipschitz graph in* $\mathbb{T}^n \times \mathbb{R}^n$*.*

Proof. By using Lemma 6.20 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 6.12 holds for (x,v). Then, repeat Theorem 6.6 to deduce further that u is differentiable at x, and $Du(x) = D_vL(x,v)$, which means $v = D_pH(x,Du(x))$. By abuse of notions, we write $v(x) = D_pH(x,Du(x))$ for $(x,v) \in \mathrm{spt}(\nu)$.

Next, repeating the results in Section 2.4, we obtain that ν is also supported on a Lipschitz graph. Indeed, for $(x,v(x)),(y,v(y)) \in \mathrm{spt}(\nu)$,

$$|Du(x) - Du(y)| \le C|x-y|,$$

which also means that

$$|Dv(x) - Dv(y)| \le C|x-y|.$$

□

We are now ready to prove that ν is a Mather measure thanks to the graph theorem above. This proof is taken from Evans [35].

Sketch of proof of Theorem 6.19. Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ be a solution to (6.15).

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.

□

3 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 (6.4) at $P = 0$

$$H(x, Du(x)) = \overline{H}(0) \qquad \text{in } \mathbb{T}^n. \tag{6.15}$$

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

$$-\mathrm{div}(D_pH(x,Du(x))\sigma) = 0 \qquad \text{in } \mathbb{T}^n. \tag{6.16}$$

Therefore, to think about weak KAM theory, a correct way is to think of a system of two equations (6.15) and (6.16). 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 (6.15) 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, (6.16) 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 [36], allows us to introduce nonlinear PDE methods to weak KAM theory to read off more information.

There have been many different ways to approximate (6.15) and (6.16) 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.

3.1 Vanishing viscosity approximations

Here, we aim at approximating (6.15) by adding a small viscosity term. For each $\varepsilon > 0$, we consider

$$H(x, Du^\varepsilon(x)) = \varepsilon \Delta u^\varepsilon + \overline{H}^\varepsilon(0) \qquad \text{in } \mathbb{T}^n. \tag{6.17}$$

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 6.22. *Assume* (6.1). *For every $\varepsilon > 0$, there exists a unique constant $\overline{H}^\varepsilon(0) \in \mathbb{R}$ such that* (6.17) *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 (6.15).

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

$$-\|H(\cdot, 0)\|_{L^\infty(\mathbb{T}^n)} \leq \lambda v^\lambda \leq \|H(\cdot, 0)\|_{L^\infty(\mathbb{T}^n)}.$$

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

$$2\lambda w^\lambda + D_p H(x, Dv^\lambda) \cdot Dw^\lambda + D_x H(x, Dv^\lambda) \cdot Dv^\lambda = \varepsilon \Delta w^\lambda - \varepsilon |D^2 v^\lambda|^2.$$

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 (6.1) 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. □

The linearized operator of (6.17) around u^ε is

$$\mathcal{L}^\varepsilon[\phi] = D_p H(x, Du^\varepsilon) \cdot D\phi - \varepsilon \Delta \phi \qquad \text{for all } \phi \in C^2(\mathbb{T}^n).$$

This allows us to consider the adjoint equation to this linearized operator as

$$(\mathcal{L}^\varepsilon)^*[\sigma^\varepsilon] = -\mathrm{div}(D_p H(x, Du^\varepsilon)\sigma^\varepsilon) - \varepsilon \Delta \sigma^\varepsilon = 0 \qquad \text{in } \mathbb{T}^n. \tag{6.18}$$

It is quite clear that 0 is the principal eigenvalue of $(\mathcal{L}^\varepsilon)^*$, and so, (6.18) 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 $\operatorname{spt}(\mu^\varepsilon) \subset \mathbb{T}^n \times \overline{B}(0,C)$. 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, $\operatorname{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. 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

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\, d\mu(x,p) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x, D_v L(x,v))\, d\nu(x,v)$$

for all bounded continuous functions ψ.

Theorem 6.23. *Assume* (6.1). *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 (6.18) 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 (6.18) 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 (6.17) 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 6.24. *Assume* (6.1). *Then, there exists* $C > 0$ *independent of* ε *such that*

$$\int_{\mathbb{T}^n} |D^2 u^\varepsilon|^2 \sigma^\varepsilon \, dx \leq C.$$

Proof. For each $1 \leq i \leq n$, differentiate (6.17) with respect to x_i twice to get

$$D_p H \cdot D u^\varepsilon_{x_i x_i} + H_{p_k p_l} u^\varepsilon_{x_k x_i} u^\varepsilon_{x_l x_i} + H_{x_i x_i} + 2 H_{x_i p_k} u^\varepsilon_{x_k x_i} = \varepsilon \Delta u^\varepsilon_{x_i x_i}.$$

By the uniform convexity of H in p (assumption (6.1)), we simplify the above as

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

Multiply this inequality with σ^ε and integrate over $\mathbb{T}^n$ to deduce

$$\frac{\theta}{2} \int_{\mathbb{T}^n} |D u^\varepsilon_{x_i}|^2 \sigma^\varepsilon \, dx \leq C.$$

Sum this over $i = 1, 2, \ldots, n$ to conclude. □

3.2 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 (6.15). Our aim is to use large time average of solutions to derive Mather measures. Consider

$$\begin{cases} \varphi_t + H(x, D\varphi) = 0 & \text{in } \mathbb{T}^n \times (0, \infty), \\ \varphi(x, 0) = u(x) & \text{on } \mathbb{T}^n. \end{cases}$$

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

$$\begin{cases} \varepsilon w_t + H(x, Dw) = 0 & \text{in } \mathbb{T}^n \times (0, \infty), \\ w(x, 0) = u(x) & \text{on } \mathbb{T}^n. \end{cases} \tag{6.19}$$

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,

$$\|u^\delta - u\|_{L^\infty(\mathbb{T}^n)} \leq C\delta,$$

and

$$\|Du^\delta\|_{L^\infty(\mathbb{T}^n)} + \delta \|D^2 u^\delta\|_{L^\infty(\mathbb{T}^n)} \leq C.$$

Let us consider the following Cauchy problems

$$\begin{cases} \varepsilon w^\varepsilon_t + H(x, Dw^\varepsilon) = \varepsilon^4 \Delta w^\varepsilon & \text{in } \mathbb{T}^n \times (0,1), \\ w^\varepsilon(x,0) = u^{\varepsilon^4}(x) & \text{on } \mathbb{T}^n, \end{cases} \tag{6.20}$$

and

$$\begin{cases} \varepsilon \phi^\varepsilon_t + H(x, D\phi^\varepsilon) = \varepsilon^4 \Delta \phi^\varepsilon & \text{in } \mathbb{T}^n \times (0,1), \\ \phi^\varepsilon(x,0) = u(x) & \text{on } \mathbb{T}^n. \end{cases} \tag{6.21}$$

Here, u^{ε^4} is u^δ with $\delta = \varepsilon^4$. As $\|u^{\varepsilon^4} - u\|_{L^\infty(\mathbb{T}^n)} \leq C\varepsilon^4$, it is straightforward that

$$\|w^\varepsilon - \phi^\varepsilon\|_{L^\infty(\mathbb{T}^n \times [0,1])} \leq C\varepsilon^4.$$

The next result concerns gradient bound of w^ε.

Lemma 6.25. *Assume* (6.1). *There is a constant $C > 0$ independent of $\varepsilon > 0$ such that*

$$\varepsilon \|w^\varepsilon_t\|_{L^\infty(\mathbb{T}^n \times [0,1])} + \|Dw^\varepsilon\|_{L^\infty(\mathbb{T}^n \times [0,1])} \leq C.$$

Proof. Denote by

$$\varphi^\pm(x,t) = w^\varepsilon(x,0) \pm \frac{C}{\varepsilon} t \quad \text{for all } (x,t) \in \mathbb{T}^n \times [0,1].$$

Then, φ^-, φ^+ are, respectively, a subsolution, and a supersolution to (6.20). Hence, by the comparison principle,

$$\varphi^- \leq w^\varepsilon \leq \varphi^+ \quad \Rightarrow \quad \|w^\varepsilon(\cdot,s) - w^\varepsilon(\cdot,0)\|_{L^\infty} \leq \frac{Cs}{\varepsilon}.$$

Note next that both w^ε and $w^\varepsilon(\cdot,\cdot+s)$ solve (6.20) with initial data $w^\varepsilon(\cdot,0)$ and $w^\varepsilon(\cdot,s)$, respectively. By the comparison principle,

$$\|w^\varepsilon(\cdot,\cdot+s) - w^\varepsilon\|_{L^\infty} \leq \|w^\varepsilon(\cdot,s) - w^\varepsilon(\cdot,0)\|_{L^\infty} \leq \frac{Cs}{\varepsilon} \quad \Rightarrow \quad \varepsilon\|w^\varepsilon_t\|_{L^\infty(\mathbb{T}^n)} \leq C.$$

To prove the spatial gradient bound, we use the usual Bernstein method. Let $\psi(x,t) = \frac{|Dw^\varepsilon|^2}{2}$. Then ψ satisfies

$$\varepsilon\psi_t + D_pH \cdot D\psi + D_xH \cdot Dw^\varepsilon = \varepsilon^4 \Delta\psi - \varepsilon^4 |D^2 w^\varepsilon|^2.$$

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,

$$D_xH \cdot Dw^\varepsilon + \varepsilon^4 |D^2 w^\varepsilon|^2 \leq 0 \quad \text{at } (x_0,t_0).$$

For $\varepsilon < n^{-1}$, we have

$$\varepsilon^4 |D^2 w^\varepsilon|^2 \geq (\varepsilon^4 \Delta w^\varepsilon)^2 = (\varepsilon w^\varepsilon_t + H(x,Dw^\varepsilon))^2 \geq \frac{1}{2} H(x,Dw^\varepsilon)^2 - C.$$

Therefore,

$$\frac{1}{2} H(x,Dw^\varepsilon)^2 + D_xH \cdot Dw^\varepsilon \leq C \quad \text{at } (x_0,t_0),$$

which, together with (6.1), yields the desired result. ∎

Lemma 6.26. *Assume* (6.1). *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 1.33. 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 (6.1) 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 6.27. All the above steps are mainly to show that instead of working with (6.19) directly, we can work with (6.20), 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 (6.15) via the nonlinear adjoint method described below.

The linearized operator of (6.20) about the solution w^ε is

$$\mathcal{L}^\varepsilon[\phi] = \varepsilon \phi_t + D_p H(x, Dw^\varepsilon) \cdot D\phi - \varepsilon^4 \Delta \phi.$$

The corresponding adjoint equation is

$$\begin{cases} -\varepsilon \sigma^\varepsilon_t - \operatorname{div}(D_p H(x, Dw^\varepsilon)\sigma^\varepsilon) = \varepsilon^4 \Delta \sigma^\varepsilon & \text{in } \mathbb{T}^n \times (0,1), \\ \sigma^\varepsilon(x,1) = \delta_{x_0}. \end{cases} \tag{6.22}$$

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 6.28. *The following holds*

$$\int_{\mathbb{T}^n} \sigma^\varepsilon(x,t)\, dx = 1 \qquad \textit{for all } t \in (0,1).$$

Proof. For $t \in (0,1)$, integrate (6.22) on $\mathbb{T}^n$ to yield

$$\varepsilon \frac{d}{dt} \int_{\mathbb{T}^n} \sigma^\varepsilon \, dx = \int_{\mathbb{T}^n} -\operatorname{div}(D_p H(x, w^\varepsilon, Dw^\varepsilon)\sigma^\varepsilon) - \varepsilon^4 \Delta \sigma^\varepsilon \, dx = 0,$$

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, $\operatorname{spt}(\mu^\varepsilon) \subset \mathbb{T}^n \times \overline{B}(0,C)$. 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 $\operatorname{spt}(\mu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. 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 6.29. *Assume* (6.1). *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 6.23. First, we prove $\nu \in \mathcal{F}$. Multiply (6.22) with $\phi \in C^2(\mathbb{T}^n)$ and integrate to imply

$$\varepsilon \int_{\mathbb{T}^n} \phi(x)\sigma^\varepsilon(x,0)\,dx - \varepsilon\phi(x_0) + \int_0^1 \int_{\mathbb{T}^n} D_pH(x,Dw^\varepsilon)\cdot D\phi(x)\sigma^\varepsilon(x,t)\,dxdt$$
$$= \varepsilon^4 \int_0^1 \int_{\mathbb{T}^n} \Delta\phi(x)\sigma^\varepsilon(x,t)\,dxdt.$$

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 (6.20) by σ^ε, multiply (6.22) by w^ε, combine them and integrate to infer

$$\varepsilon w^\varepsilon(x_0,1) - \varepsilon \int_{\mathbb{T}^n} w^\varepsilon(x,0)\sigma^\varepsilon(x,0)\,dx$$
$$= \int_0^1 \int_{\mathbb{T}^n} (D_pH(x,Dw^\varepsilon)\cdot Dw^\varepsilon - H(x,Dw^\varepsilon))\sigma^\varepsilon(x,t)\,dxdt.$$

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 present here an application of this PDE approach.

Theorem 6.30. *Assume* (6.1). *Normalize so that* $\overline{H}(0) = 0$. *Let* $u, \bar{u} \in \mathrm{Lip}(\mathbb{T}^n)$ *be two solutions to* (6.15). *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 6.14. As we see right away in the proof below, the approach here is quite different. This uniqueness result is taken from Mitake, Tran [83].

Proof. Consider (6.20) and (6.22) as above with solutions w^ε and σ^ε, respectively. Let $\bar{w}^\varepsilon$ be the solution to (6.20) 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, $\bar{u}\le u$. □

4 References

1. For further developments in weak KAM theory, see Fathi's book [40]. We do not cover many topics here such as the Aubry set, regularity of subsolutions on the Aubry set, large time behavior. See also the books of Gomes [53], Sorrentino [89].

2. There are many excellent survey papers and lecture notes in this area: see Evans [34, 35], Ishii [61], Kaloshin [68], and the references therein.

3. We do not cover the two dimensional Aubry–Mather theory here.

4. 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 [15].

5. 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 [49], Cagnetti, Mitake, Gomes, Tran [14], Mitake, Tran [82, 83], Ishii, Mitake, Tran [62, 63].

CHAPTER 7

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

$$\begin{cases} H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{pp} H(x,p) \le \theta^{-1} I_n \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{7.1}$$

Let $L = L(y,v)$ be the corresponding Lagrangian. By changing $\theta > 0$ to be smaller if needed, we may also assume that

$$\begin{cases} L \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \\ \text{there exists } \theta > 0 \text{ such that } \theta I_n \le D^2_{vv} L(x,v) \le \theta^{-1} I_n \text{ for all } (x,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{7.2}$$

Here is the main result in this section.

Theorem 7.1. *Assume* (7.1). *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 [36]. 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 6.16), μ is supported on a Lipschitz graph, and for $x \in \operatorname{spt}(\sigma)$, u is differentiable at x, and

$$\begin{aligned} &H(x, P + tR + D\tilde{u}^\delta(x)) + H(x, P - tR + D\hat{u}^\delta(x)) - 2H(x, P + Du(x)) \\ &\qquad\qquad\qquad \le \overline{H}(P + tR) + \overline{H}(P - tR) - 2\overline{H}(P) + C\delta. \end{aligned}$$

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,

$$\int_{\mathbb{T}^n} \left(|tR + D(\tilde{u}^\delta - u)|^2 + |-tR + D(\hat{u}^\delta - u)|^2\right) d\sigma \le C\left(\overline{H}(P + tR) + \overline{H}(P - tR) - 2\overline{H}(P) + C\delta\right). \tag{7.3}$$

On the other hand,

$$\overline{H}(P)-\overline{H}(P+tR)\le\int_{\mathbb{T}^n}(H(x,P+Du)-H(x,P+tR+D\tilde{u}^{\delta}))\,d\sigma+C\delta$$
$$\le C\left(\int_{\mathbb{T}^n}|-tR+D(u-\tilde{u}^{\delta})|^2\,d\sigma\right)^{1/2}+C\delta, \tag{7.4}$$

and

$$\overline{H}(P)-\overline{H}(P-tR)\le\int_{\mathbb{T}^n}(H(x,P+Du)-H(x,P-tR+D\hat{u}^{\delta}))\,d\sigma+C\delta$$
$$\le C\left(\int_{\mathbb{T}^n}|tR+D(u-\hat{u}^{\delta})|^2\,d\sigma\right)^{1/2}+C\delta. \tag{7.5}$$

Combine (7.3)–(7.5) 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 7.2. 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

$$\begin{cases} V(x) = \lambda_0 + \sum_{j=1}^m (\lambda_j e^{i2\pi k_j \cdot x} + \overline{\lambda_j} e^{-i2\pi k_j \cdot x}), \\ \text{where } \lambda_0 \in \mathbb{R},\ \{\lambda_j\}_{j=1}^m \subset \mathbb{C} \text{ and } \{k_j\}_{j=1}^m \subset \mathbb{Z}^n \setminus \{0\} \text{ are given.} \end{cases} \tag{7.6}$$

Recall that $\overline{\lambda_j}$ is the complex conjugate of λ_j for $1 \leq j \leq m$. 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

$$\frac{1}{2}|Q + \sqrt{\varepsilon} Dv^\varepsilon|^2 + \varepsilon V(x) = \varepsilon \overline{H}\left(\frac{Q}{\sqrt{\varepsilon}}\right) =: \overline{H}^\varepsilon(Q) \qquad \text{in } \mathbb{T}^n. \tag{7.7}$$

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 (7.7) 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

$$Dv_1 = -\sum_{j=1}^m (\lambda_j e^{i2\pi k_j \cdot x} + \overline{\lambda_j} e^{-i2\pi k_j \cdot x}) \frac{k_j}{k_j \cdot Q}, \tag{7.8}$$

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

$$a_2 = \sum_{j=1}^{m} \frac{|\lambda_j|^2 |k_j|^2}{|k_j \cdot Q|^2}. \tag{7.9}$$

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

$$Dv_2 = -\frac{1}{2} \sum_G \frac{\lambda_{j_1}^{\pm} \lambda_{j_2}^{\pm} k_{j_1} \cdot k_{j_2}}{(k_{j_1} \cdot Q)(k_{j_2} \cdot Q)} e^{i2\pi(\pm k_{j_1} \pm k_{j_2})\cdot x} \frac{\pm k_{j_1} \pm k_{j_2}}{(\pm k_{j_1} \pm k_{j_2}) \cdot Q}. \tag{7.10}$$

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

$$\begin{aligned} Dv_3 = -\frac{1}{2} \sum_G & \frac{\lambda_{j_1}^{\pm} \lambda_{j_2}^{\pm} \lambda_{j_3}^{\pm} (k_{j_1} \cdot k_{j_2})(\pm k_{j_1} \pm k_{j_2}) \cdot k_{j_3}}{(k_{j_1} \cdot Q)(k_{j_2} \cdot Q)(k_{j_3} \cdot Q)(\pm k_{j_1} \pm k_{j_2}) \cdot Q} \times \\ & \times e^{i2\pi(\pm k_{j_1} \pm k_{j_2} \pm k_{j_3})\cdot x} \frac{\pm k_{j_1} \pm k_{j_2} \pm k_{j_3}}{(\pm k_{j_1} \pm k_{j_2} \pm k_{j_3}) \cdot Q}. \end{aligned} \tag{7.11}$$

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

$$Q \cdot Dv_4 = a_4 - Dv_1 \cdot Dv_3 - \frac{1}{2}|Dv_2|^2. \tag{7.12}$$

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 7.3. *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 (7.6). 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 (7.8), (7.10), (7.11), (7.12), 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 (7.7). We now use ϕ, which is smooth, as a test function for (7.7). 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 7.4. 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 [56]. See Bangert [8, Section 5] and E [28] for more modern accounts of this example.

Let us consider the simplest case in three dimensions ($n = 3$) with Hamiltonian

$$H(y,p) = \frac{1}{a(y)}|p| \qquad \text{for all } (y,p) \in \mathbb{R}^3 \times \mathbb{R}^3, \tag{7.13}$$

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 7.5. *Assume that H is of the form* (7.13). *Let $\overline{H}$ be its corresponding effective Hamiltonian. Then,*

$$\overline{H}(p) = \frac{1}{\delta}\max\{|p_1|, |p_2|, |p_3|\} \qquad \textit{for } p = (p_1, p_2, p_3) \in \mathbb{R}^3.$$

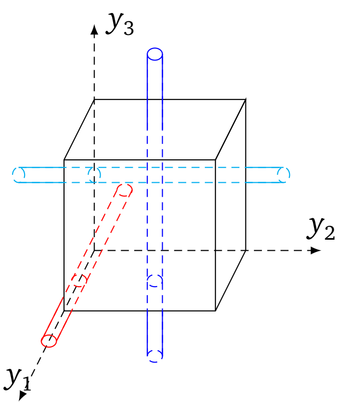


Figure 7.1: Shape of $U_\delta(\mathcal{L})$

Theorem 7.5 was proved by Bangert [8] 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 7.6. 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 $1 \leq i \leq 3$, 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 about the Aubry set here, the above proof also gives that the 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 6.7.

This Hedlund example also explains a weakness of weak KAM theory in dimensions three and higher, where the 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 References

1. Strict convexity of the effective Hamiltonian in certain directions is taken from Evans, Gomes [36].

2. Asymptotic expansion at infinity is taken from Luo, Tran, Yu [76]. See also Jing, Tran, Yu [67], and Tran, Yu [92]. The method of asymptotic expansion at infinity was used in [76, 67, 92] 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 [56]. Then, Bangert [8] and E [28] 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 [84]).

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 [70], 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

$$u(x_1, T) = \inf_{x\in\mathbb{T}^n} (V(x) + g(x)).$$

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 $k \in \mathbb{N}$,

$$\int_0^T \frac{\theta}{2}|\gamma_k'(s)|^2\,ds \le \int_0^T \big(L(\gamma_k(s),\gamma_k'(s)) + K_0\big)\,ds \le C + K_0 T \le C.$$

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

$$\int_0^T L(\gamma_{k_j}(s),\gamma_{k_j}'(s))\,ds \ge \int_0^T \Big(L(\gamma_{k_j}(s),\gamma'(s)) + D_vL(\gamma_{k_j}(s),\gamma'(s))\cdot(\gamma_{k_j}'(s)-\gamma'(s))\Big)\,ds.$$

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 [17], Evans [32], Ishii [61].

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

$$\frac{d}{dt}\big(D_vL(\gamma(t),\gamma'(t))\big) = D_xL(\gamma(t),\gamma'(t)) \qquad \text{on } [0,T].$$

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 $t \in [0,T]$. $\square$

Solutions to some exercises

Solutions to some exercises here were provided to me by Son Tu. I intend to keep them mostly as they are here without much editing for this first draft. More exercises and perhaps their solutions will be added later.

Exercise 1. *Consider the eikonal problem in one dimension*

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

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

(b) Show that all the a.e. solutions that we got in class are mutually viscosity subsolutions.

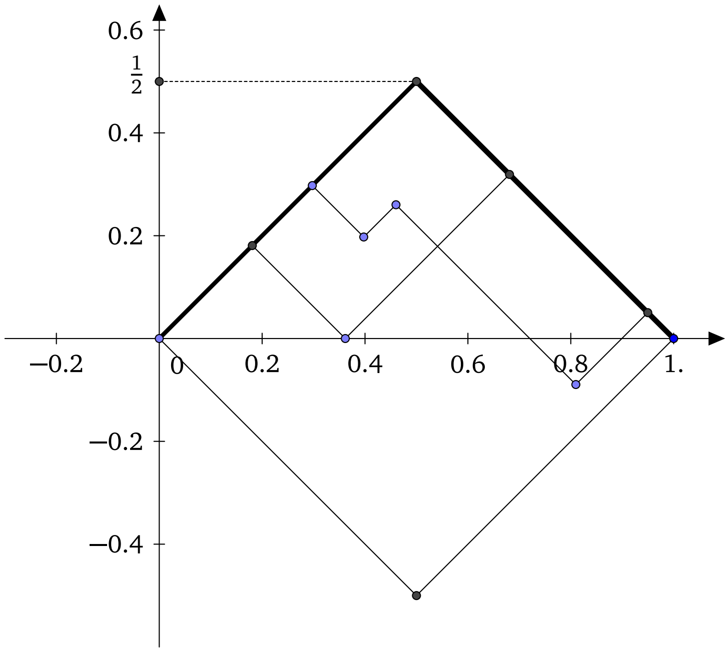
Proof.

(a) Assume that there exists a C^1 solution $u : [-1,1] \longrightarrow \mathbb{R}$ satisfies (1.1), then $x \longmapsto u'(x)$ must be continuous, so by mean value theorem there exists some $c \in (-1,1)$ such that:

$$0 = u(1) - u(-1) = u'(c)\Big(1-(-1)\Big) = 2u'(c) \qquad \Longrightarrow \qquad u'(c) = 0.$$

That is a contradiction since $|u'(c)| = 1$, thus (1.1) has no C^1 solution.

(b) Generally all the solutions we got in class has the following form:



They are combining of segments with slope 1 or -1. As we can see, $u'(x)$ is exist a.e, so we only need to check if they are subsolution at points where $u'(x)$ is not well defined, i.e at the vertex.

- 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 $|u'(x)| \leq 1$ in the viscosity sense holds true.
- 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) can varies with $\varphi'(x) \in [-1,1]$, since $u(z)$ with $z < x$ is a segment with slope 1, and $u(z)$ with $z > x$ is a segment with slope -1 . That means the condition $|u'(x)| \leq 1$ in the viscosity sense holds true.

Thus we have any solution we have in the above form is a viscosity subsolution of (1.1).

□

Exercise 2. *Show that in the definition of viscosity solutions (for first-order equations), it is equivalent to assume that the test functions $\varphi, \psi \in C^1$ or the test functions $\varphi, \psi \in C^2$.*

Proof. Recall the definition of viscosity solution for first order equations:

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

Let's define definition 1 to be the definition with C^1-test functions, while definition 2 will be the definition with C^1-test functions. It is easy to see that definition 1 implies definition 2, since any C^2 function if-self is a C^1 function. Now we will prove the opposite, assume u is subsolution of (C) by definition 2, then $u(x,0) \leq u_0(x)$. 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:

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

Let extend $\varphi(x,t)=0$ for any $t\notin(0,T)$. Let $\{\eta_\varepsilon\}_\varepsilon\subset C_c^\infty(\mathbb{R}^{n+1})$ be the standard mollifiers, that is $\eta_\varepsilon(x)=\frac{1}{\varepsilon^{n+1}}\eta\left(\frac{x}{\varepsilon}\right)$ where $\eta(x)=Ce^{\frac{1}{|x|^2-1}}\chi_{B(0,1)}(x)$ where C is a constant be chosen such that $\int_{\mathbb{R}^{n+1}}\eta\,d\mu=1$. Then

$$\eta_\varepsilon\in C_c^\infty\left(\mathbb{R}^{n+1}\right),\qquad \operatorname{supp}\eta_\varepsilon\subset B(0,\varepsilon),\qquad\text{and}\qquad\int_{\mathbb{R}^{n+1}}\eta_\varepsilon\,d\mu=1.$$

Now for every $\varepsilon>0$ we consider the convolution:

$$\begin{aligned}\varphi^\varepsilon(x,t)=\left(\eta_\varepsilon\star\varphi\right)(x,t)&=\int_{\mathbb{R}^{n+1}}\eta_\varepsilon\Big((x,t)-(y,s)\Big)\varphi(y,s)\,dyds\\&=\int_{B(0,\varepsilon)}\eta_\varepsilon(y,s)\varphi\Big((x,t)-(y,s)\Big)\,dyds=\left(\varphi\star\eta_\varepsilon\right)(x,t).\end{aligned}$$

- **Claim 1.** $\varphi^\varepsilon\longrightarrow\varphi$ locally uniformly on $\mathbb{R}^{n+1}$ and $\varphi_t^\varepsilon\longrightarrow\varphi_t$ locally uniformly on $\mathbb{R}^n\times(0,T)$.

  We have $\varphi^\varepsilon\in C^\infty(\mathbb{R}^{n+1})$ and $\varphi^\varepsilon\longrightarrow\varphi$ locally uniformly on $\mathbb{R}^{n+1}$. Since $\varphi\in C^1\left(\mathbb{R}^n\times(0,T)\right)$, if we denote $\frac{\partial}{\partial x_{n+1}}=\frac{\partial}{\partial t}$ then for any $(x,t)\in\mathbb{R}^n\times(0,T)$ we have:

  $$\varphi_t^\varepsilon(x,t)=\frac{\partial}{\partial t}\varphi^\varepsilon(x,t)=\frac{\partial}{\partial t}\Big(\eta_\varepsilon\star\varphi\Big)(x,t)=\left(\eta_\varepsilon\star\frac{\partial}{\partial t}\varphi\right)(x,t)=(\eta_\varepsilon\star\varphi_t)(x,t)$$

  Since φ_t is continuous, we have $(\eta_\varepsilon\star\varphi_t)\longrightarrow\varphi_t$ locally uniformly on $\mathbb{R}^n\times(0,T)$.

- **Claim 2.** For any ε small enough, then $u-\varphi^\varepsilon$ has a max at $(x_\varepsilon,t_\varepsilon)$ near (x_0,t_0) and we can choose a decreasing subsequence $\varepsilon_i\searrow 0$ such that $\left(x_{\varepsilon_i},t_{\varepsilon_i}\right)\longrightarrow(x_0,t_0)$ as $i\longrightarrow\infty$.

Now using these result, we have $u-\varphi^\varepsilon$ has max at $\left(x_{\varepsilon_i},t_{\varepsilon_i}\right)$, and φ^ε is smooth in $\mathbb{R}^n\times(0,T)$, so:

$$\varphi^\varepsilon\left(x_{\varepsilon_i},t_{\varepsilon_i}\right)+H\left(D\varphi^\varepsilon\left(x_{\varepsilon_i},t_{\varepsilon_i}\right)\right)\leq 0.$$

Let $\varepsilon_i\longrightarrow 0$ and using the fact that $\left(x_{\varepsilon_i},t_{\varepsilon_i}\right)\longrightarrow(x_0,t_0)$, H is continuous and $\varphi_t^\varepsilon\longrightarrow\varphi_t$ uniformly we obtain the result:

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

Thus u satisfies the definition 1. The argument for case supersolution is similar. □

Exercise 3. *Let u,φ be continuous functions on $\mathbb{R}^n\times[0,\infty)$ such that $u-\varphi$ has a strict max over $\mathbb{R}^n\times[0,T]$ at (x_0,T). For each $\varepsilon>0$ let $\varphi_\varepsilon(x,t)=\varphi(x,t)+\frac{\varepsilon}{T-t}$, show that for $\varepsilon>0$ small enough, $u-\varphi_\varepsilon$ has a local max at $(x_\varepsilon,t_\varepsilon)$ and $(x_\varepsilon,t_\varepsilon)\longrightarrow(x_0,T)$ up to subsequence.*

Proof. Fix $0<r<T$ and let $\Omega_r=\overline{B_r(x_0)}\times[T-r,T)$, we have $u-\varphi<0$ for all $(x,t)\in\Omega_r$. For $\varepsilon>0$ let's define $\varphi_\varepsilon:\mathbb{R}^n\times[0,T)\longrightarrow\mathbb{R}$ by

$$\varphi_\varepsilon(x,t)=\varphi(x,t)+\frac{\varepsilon}{T-t}\quad\Longrightarrow\quad u(x,t)-\varphi_\varepsilon(x,t)\leq-\frac{\varepsilon}{T-t}<0\tag{A.4}$$

Thus $u-\varphi_\varepsilon$ has a max over $\overline{B_r(x_0)}\times[T-r,T)$ at $(x_\varepsilon,t_\varepsilon)$ [1]. We show that for $r>0$, there $\varepsilon=\varepsilon(r)$ 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[2]

$$\partial\Omega_r=\underbrace{\left(\overline{B(x_0,r)}\times\{T-r\}\right)}_{\text{bottome of cylinder}}\cup\underbrace{\left(\partial B(x_0,r)\times(T-r,T)\right)}_{\text{surface between the bottom and the top}}.$$

Let

$$\alpha=\sup_{\partial\Omega_r}(u-\varphi)(x,t)<0.$$

By continuity, there exists $0<\delta<r$ such that $|(u-\varphi)(x_0,s)|<-\frac{\alpha}{2}$ for all $s\in[T-\delta,T]$, thus

$$(u-\varphi)(x,t)<\frac{\alpha}{2}+(u-\varphi)(x_0,T-\delta)\qquad\text{for all}\quad (x,t)\in\partial\Omega_r.$$

Thus for all $(x,t)\in\partial\Omega_r$ then

$$(u-\varphi_\varepsilon)<(u-\varphi)(x_0,T-\delta)+\frac{\alpha}{2}-\frac{\varepsilon}{T-t}<(u-\varphi)(x_0,T-\delta)+\frac{\alpha}{2}-\frac{\varepsilon}{r}.$$

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 max $(x_\varepsilon,t_\varepsilon)$ of $u-\varphi_\varepsilon$ cannot achieves on $\partial\Omega_r$. Our claim is proven with $\varepsilon(r)=-\frac{\alpha_r}{2}\left(\frac{1}{\delta_r}-\frac{1}{r}\right)^{-1}$. Now let $r=\frac{1}{n}$ and construct ε_n decreasing by induction we obtain $(x_\varepsilon,t_\varepsilon)\longrightarrow(x_0,T)$ since Ω_r shrinks to (x_0,T) as $r\longrightarrow 0$. □

Exercise 4. *Let $H=H(x,p):\mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R}$ be a Hamiltonian satisfying:*

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

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

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

where $g^i\in \mathrm{BUC}(\mathbb{R}^n)$. Use the comparison principle for (C) *to show the following contraction property: For any $t\ge 0$ then:*

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

Proof. Denote $C=|g^1-g^2|_\infty=\sup_{x\in\mathbb{R}^n}|g^1(x)-g^2(x)|$.

[1] 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 $u(x_j,t_j)-\varphi_\varepsilon(x_j,t_j)\longrightarrow\alpha$. By compactness we have $(x_j,t_j)\longrightarrow(\overline{x},\overline{t})\in\Omega_r$ up to subsequence. If $\overline{t}=T$ then from (A.4) we have $\alpha=-\infty$, which is a contradiction.

[2] It is not actually the boundary of Ω_r under the Euclid metric, but we don't have to worry about the top of the cylinder.

- Let $\zeta(x,t) = u^2(x,t) + C$ then it is a viscosity supersolution of (C) with the initial data $g^1(x)$, since:
  - For $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ such that $\zeta - \varphi$ has a min at (x_0,t_0) then $u^2 - \varphi$ also has a min at (x_0,t_0), so:
  $$\varphi_t(x,t) + H(x, D\varphi(x,t)) \geq 0.$$
  - $\zeta(x,0) = u^2(x,0) + C = g^2(x) + C \geq g^1(x)$.

  By comparison principle for (C), $\zeta(x,t) \geq u^1(x,t)$, i.e $u^2(x,t) + C \geq u^1(x,t)$ 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 (C) with the initial data $g^1(x)$, since:
  - For $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ such that $\delta - \varphi$ has a max at (x_0,t_0) then $u^2 - \varphi$ also has a max at (x_0,t_0), so:
  $$\varphi_t(x,t) + H(x, D\varphi(x,t)) \leq 0.$$
  - $\delta(x,0) = u^2(x,0) - C = g^2(x) - C \leq g^1(x)$.

  By comparison principle for (C), $\delta(x,t) \leq u^1(x,t)$, i.e $u^2(x,t) - C \leq u^1(x,t)$ for $(x,t) \in \mathbb{R}^n \times (0,\infty)$.

Thus $u^2(x,t) - C \leq u^1(x,t) \leq u^2(x,t) + C$, i.e $|u^1(x,t) - u^2(x,t)| \leq C$ for all $(x,t) \in \mathbb{R}^n \times (0,\infty)$. □

Exercise 5. *Let $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ be a C^1 Hamiltonian satisfying $|H(x,0)| \leq C$ and*
$$\lim_{|p|\longrightarrow\infty} \inf_{x\in\mathbb{R}^n} \left(\frac{1}{2} H(x,p)^2 + D_x H(x,p) \cdot p \right) = +\infty.$$
Consider the following vanishing viscosity problem:
$$u^\varepsilon + H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon \qquad in\ \mathbb{R}^n. \tag{S_ε}$$
Use Bernstein method to show that there exists a constant $C > 0$ independent of ε such that $|Du^\varepsilon| \leq C$.

Proof. For $k = 1,2,\dots,n$, differentiate (S_ε) with respect to x_k we have:
$$u^\varepsilon_{x_k} + H_{x_k}(x, Du^\varepsilon) + D_p(x, Du^\varepsilon) \cdot Du^\varepsilon_{x_k} = \varepsilon \Delta u^\varepsilon_{x_k}.$$
Multiplying two side with $u^\varepsilon_{x_k}$ and taking the sum over $i = 1,2,\dots,n$ we have:
$$\sum_{k=1}^n \left(u^\varepsilon_{x_k}\right)^2 + H_x(x, Du^\varepsilon) \cdot Du^\varepsilon + D_p(x, Du^\varepsilon) \cdot \sum_{k=1}^n Du^\varepsilon_{x_k} u^\varepsilon_{x_k} = \varepsilon \sum_{k=1}^n \Delta u^\varepsilon_{x_k} u^\varepsilon_{x_k}. \tag{1.5.1}$$

- $Du^\varepsilon_{x_k} u^\varepsilon_{x_k} = \left(u^\varepsilon_{x_1 x_k} u^\varepsilon_{x_k}, \dots, u^\varepsilon_{x_n x_k} u^\varepsilon_{x_k}\right) = \frac{1}{2}\left(\frac{\partial}{\partial x_1}\left(u^\varepsilon_{x_k}\right)^2, \dots, \frac{\partial}{\partial x_n}\left(u^\varepsilon_{x_k}\right)^2\right) = \frac{1}{2} D\left(u^\varepsilon_{x_k}\right)^2$.
- $\Delta\left(u^\varepsilon_{x_k}\right)^2 = 2\sum_{i=1}^n \left(\frac{\partial^2}{\partial x_i^2} u^\varepsilon_{x_k}\right) u^\varepsilon_{x_k} + \frac{1}{2}\sum_{i=1}^n \left(\frac{\partial u^\varepsilon_{x_k}}{\partial x_i}\right)^2$, hence $\sum_{k=1}^n \Delta u^\varepsilon_{x_k} u^\varepsilon_{x_k} = \frac{1}{2}\Delta \sum_{k=1}^n \left(u_{x_k}\right)^2 - \left|D^2 u^\varepsilon\right|^2$.

From this (1.5.1) becomes:

$$\sum_{k=1}^{n}\left(u^{\varepsilon}_{x_k}\right)^2+H_x(x,Du^{\varepsilon})\cdot Du^{\varepsilon}+D_p(x,Du^{\varepsilon})\cdot D\left(\frac{1}{2}\sum_{k=1}^{n}\left(u^{\varepsilon}_{x_k}\right)^2\right)=\varepsilon\Delta\left(\frac{1}{2}\sum_{k=1}^{n}\left(u^{\varepsilon}_{x_k}\right)^2\right)-\varepsilon\left|D^2u^{\varepsilon}\right|^2.$$

Set $\psi^{\varepsilon}(x,t)=\frac{1}{2}\sum_{k=1}^{n}\left(u^{\varepsilon}_{x_k}\right)^2=\frac{1}{2}|Du^{\varepsilon}|^2\geq 0$ then:

$$\left(2\psi^{\varepsilon}+D_p(x,Du^{\varepsilon})\cdot D\psi^{\varepsilon}-\varepsilon\Delta\psi^{\varepsilon}\right)+H_x(x,Du^{\varepsilon})\cdot Du^{\varepsilon}+\varepsilon\left|D^2u^{\varepsilon}\right|^2=0. \tag{1.5.2}$$

If $\varepsilon<\frac{1}{n}$ then

$$\varepsilon\left|D^2u^{\varepsilon}\right|^2\geq\varepsilon\sum_{i=1}^{n}\left(u^{\varepsilon}_{x_ix_i}\right)^2\geq\frac{\varepsilon}{n}\left(\sum_{i=1}^{n}u^{\varepsilon}_{x_ix_i}\right)=\frac{\varepsilon}{n}(\Delta u^{\varepsilon})^2\geq\left(\varepsilon\Delta u^{\varepsilon}\right)^2=\left(u^{\varepsilon}+H(x,Du^{\varepsilon})\right)^2.$$

Assume u^{ε} achieves its maximum and minimum at some points x_1,x_2 respectively, then $Du^{\varepsilon}(x_1)=Du^{\varepsilon}(x_2)=0$ and $\Delta u(x_1)\leq 0\leq\Delta u(x_2)$, thus: [3]

$$-C\leq -H(x_2,0)\leq u^{\varepsilon}(x_2)\leq u^{\varepsilon}(x)\leq u^{\varepsilon}(x_1)\leq -H(x_1,0)\leq C\Longrightarrow |u^{\varepsilon}(x)|\leq C$$

for all $x\in\mathbb{R}^n$ and $\varepsilon>0$. Thus $\left(u^{\varepsilon}+H(x,Du^{\varepsilon})\right)^2\geq\frac{1}{2}H(x,Du^{\varepsilon})^2-C$ for some constant C independent to ε. Now using this fact in (1.5.2) we have:

$$\left(2\psi^{\varepsilon}+D_p(x,Du^{\varepsilon})\cdot D\psi^{\varepsilon}-\varepsilon\Delta\psi^{\varepsilon}\right)+\frac{1}{2}H(x,Du^{\varepsilon})^2+H_x(x,Du^{\varepsilon})\cdot Du^{\varepsilon}\leq C. \tag{1.5.3}$$

Now let assume that ψ^{ε} achieves its max on $\mathbb{R}^n$ at x_{ε}, then $D\psi^{\varepsilon}(x_{\varepsilon})=0$ and $\Delta\psi^{\varepsilon}(x_{\varepsilon})\leq 0$, at x_{ε} from (1.5.3) we have:

$$\frac{1}{2}H\left(x_{\varepsilon},Du^{\varepsilon}(x_{\varepsilon})\right)^2+H_x\left(x_{\varepsilon},Du^{\varepsilon}(x_{\varepsilon})\right)\cdot Du^{\varepsilon}(x_{\varepsilon})\leq C.$$

This is true for all $\varepsilon>0$, by coercivity assumption, we must have $|Du^{\varepsilon}(x_{\varepsilon})|\leq C$ for all $\varepsilon>0$. It follows that:

$$|Du^{\varepsilon}(x)|\leq|Du^{\varepsilon}(x_{\varepsilon})|\leq C$$

for all $x\in\mathbb{R}^n$ since $\psi^{\varepsilon}(x)=\frac{1}{2}|Du^{\varepsilon}(x)|^2$. Thus $|Du^{\varepsilon}|\leq C$ for all $\varepsilon>0$ small enough. □

Exercise 6 (Regularity of the value function based on DPP). *Assume that the cost function satisfies*

$$\begin{cases} f\in C(\mathbb{R}^n\times V), & |f(x,v)|\leq C \quad \text{for all } (x,v)\in\mathbb{R}^n\times V.\\ |f(y_1,v)-f(y_2,v)|\leq \mathrm{Lip}(b)|y_1-y_2|. \end{cases}$$

Assume that $b(\cdot,v)$ is Lipschitz in the first variable for all v, i.e

$$|b(y_1,v)-b(y_2,v)|\leq C|y_1-y_2|$$

for all $y_1,y_2\in\mathbb{R}^n$ and $v\in V$. Set $\lambda_0=|D_yb(\cdot,\cdot)|_{L^{\infty}(\mathbb{R}^n\times V)}$, i.e the best constant Lip(b) *in the above inequality is λ_0. Prove that:*

[3] If we can use the maximum principle here, then it will be more easy. Since $|H(x,0)|\leq C$, we obtain $\varphi=C$ and $\psi=-C$ are viscosity supersolution and viscosity subsolution of (S_{ε}) respectively, so by comparison principle we have $|u^{\varepsilon}(x)|\leq C$ for all x.

(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. Let $v(\cdot)$ be a control and $y_x(\cdot)$ and $y_z(\cdot)$ are trajectories solution to

$$\begin{cases} y_x'(s) & = b(y_x(s), v(s)) \\ y_x(0) & = x \end{cases} \qquad \text{and} \qquad \begin{cases} y_z'(s) & = b(y_z(s), v(s)) \\ y_x(0) & = z. \end{cases}$$

respectively. Then for all $s > 0$ we have

$$|y_x'(s) - y_z'(s)| \leq \lambda_0 |y_x(s) - y_z(s)| \qquad \Longrightarrow \qquad |\varphi'(s)| \leq \lambda_0 |\varphi(s)|$$

where $\varphi(s) = y_x(s) - y_z(s)$. Since b is bounded, we have

$$|\varphi(t)| = \left| \varphi(0) + \int_0^t \varphi'(s)\, ds \right| \leq |\varphi(0)| + \int_0^t |\varphi'(s)|\, ds \leq |x - z| + \lambda_0 \int_0^t |\varphi(s)|\, ds.$$

By Gronwall's inequality[4] we obtain

$$|\varphi(t)| = |y_x(t) - y_z(t)| \leq e^{\lambda_0 t} |x - z| \qquad \text{for all } t > 0. \tag{A.5}$$

(a) For (a) we don't need to use DPP, indeed for any control $v(\cdot)$ we have:

$$\begin{aligned} |J(x, v(\cdot)) - J(z, v(\cdot))| &= \left| \int_0^\infty e^{-\lambda s} f\left(y_x(s), v(s)\right)\, ds - \int_0^\infty e^{-\lambda s} f\left(y_z(s), v(s)\right)\, ds \right| \\ &= \int_0^\infty e^{-\lambda s} \left| f\left(y_x(s), v(s)\right) - f\left(y_z(s), v(s)\right) \right|\, ds \\ &\leq \int_0^\infty C e^{(-\lambda + \lambda_0) s} |x - z|\, ds = \frac{C}{\lambda - \lambda_0} |x - z| = C_0 |x - z|. \end{aligned}$$

From that we have

$$\begin{array}{rcl} J(x, v(\cdot)) \leq C_0 |x - z| + J(z, v(\cdot)) & \Longrightarrow & u(x) \leq C_0 |x - z|^\alpha + J(z, v(\cdot)) \\ \text{(take inf over } v(\cdot)) & \Longrightarrow & u(x) \leq C_0 |x - z|^\alpha + u(z) \\ J(z, v(\cdot)) \leq C_0 |x - z|^\alpha + J(x, v(\cdot)) & \Longrightarrow & u(z) \leq C_0 |x - z| + J(x, v(\cdot)) \\ \text{(take inf over } v(\cdot)) & \Longrightarrow & u(z) \leq C_0 |x - z| + u(x). \end{array}$$

---

[4]**Gronwall's inequality:** Let φ continuous $[a, b]$, $\varphi \geq 0$, assume there exists constants K, C s.t

$$0 \leq \varphi(t) \leq K + C \int_a^t \varphi(s)\, ds \qquad \forall\ t \in [a, b]$$

then

$$0 \leq \varphi(t) \leq K e^{c(t-a)} \qquad \forall\ t \in [a, b]$$

(b) We define

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

for all $t>0$ by dynamic programming principle. Also, since we assume $|f(x,v)| \le C$ for all $(x,v) \in \mathbb{R}^n \times V$, clearly for an arbitrary control $v(\cdot)$ then by definition:

$$|u(x)| \le \int_0^\infty C e^{-\lambda s}\, ds = \frac{C}{\lambda}. \tag{A.6}$$

From (A.6) and (A.5) 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 in t, then the minimum of the right hand side will be obtained at t such that $F'(t)=0$ where

$$\begin{aligned} F(t) = |x-z|t + \frac{e^{-\lambda t}}{\lambda} \quad &\Longrightarrow \quad F'(t) = |x-z| - e^{-\lambda t} \\ &\Longrightarrow \quad F'(t) = 0 \quad \text{iff} \quad t = \frac{1}{\lambda} \log\left(\frac{1}{|x-z|}\right). \end{aligned}$$

We consider the case $0 < |x-z| < 1$ first so that this 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 $\lim_{s\longrightarrow 0} G_\beta(s) = 0$, $\lim_{s \longrightarrow 1} G_\beta(s) = 1$ thus G_β is bounded by some constant C_α and hence we are done. Finally for any $\alpha \in (0,1)$ and $|x-z|<1$ then

$$\begin{aligned} K(t,x,v(\cdot)) \le C_\alpha |x-z|^\alpha + K(t,z,v(\cdot)) \quad &\Longrightarrow \quad u(x) \le C_\alpha |x-z|^\alpha + K(t,z,v(\cdot)) \\ \text{(take inf over } v(\cdot)\text{)} \quad &\Longrightarrow \quad u(x) \le C_\alpha |x-z|^\alpha + u(z) \\ K(t,z,v(\cdot)) \le C_\alpha |x-z|^\alpha + K(t,x,v(\cdot)) \quad &\Longrightarrow \quad u(z) \le C_\alpha |x-z|^\alpha + K(t,x,v(\cdot)) \\ \text{(take inf over } v(\cdot)\text{)} \quad &\Longrightarrow \quad u(z) \le C_\alpha |x-z|^\alpha + u(x). \end{aligned}$$

thus $|u(x) - u(z)| \le C_\alpha |x-z|^\alpha$ 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.6) we have

$$\frac{|u(x)-u(z)|}{|x-z|^\alpha} \le \frac{2C}{\lambda} \qquad \Longrightarrow \qquad \frac{|u(x)-u(z)|}{|x-z|^\alpha} \le \max\left\{C_\alpha, \frac{2C}{\lambda}\right\}$$

for any $x \ne z$, and for any $\alpha \in (0,1)$.

(c) From (A.6) and (A.5) we have

$$|K(t,x,v(\cdot))-K(t,z,v(\cdot))| \le \int_0^t e^{-\lambda s}|f(y_x(s),v(s))-f(y_z(s),v(s))|\, ds + e^{-\lambda t}|u(y_x(t))-u(y_z(t))|$$
$$\le C|x-z|\int_0^t e^{(\lambda_0-\lambda)s}\, ds + \frac{2C}{\lambda}e^{-\lambda t}$$
$$= C|x-z|\frac{e^{(\lambda_0-\lambda)t}-1}{\lambda_0-\lambda} + \frac{2C}{\lambda}e^{-\lambda t} \le \frac{C}{\lambda_0-\lambda}|x-z|e^{(\lambda_0-\lambda)t} + \frac{2C}{\lambda}e^{-\lambda t}.$$

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

$$F(t) = \frac{C}{\lambda_0-\lambda}|x-z|e^{(\lambda_0-\lambda)t} + \frac{2C}{\lambda}e^{-\lambda t} \quad \Longrightarrow \quad F'(t) = C|x-z|e^{(\lambda_0-\lambda)t} - 2Ce^{-\lambda t}$$
$$\Longrightarrow \quad F'(t) = 0 \quad \text{iff} \quad t = \frac{1}{\lambda_0}\log\left(\frac{2}{|x-z|}\right).$$

We consider the case $0 < |x-z| < 2$ first so that this value t above is indeed positive, then

$$|K(t,x,v(\cdot))-K(t,z,v(\cdot))| \le \frac{C}{\lambda_0-\lambda}|x-z|2^{\frac{\lambda_0-\lambda}{\lambda_0}}|x-z|^{\frac{\lambda-\lambda_0}{\lambda_0}} + \frac{2C}{\lambda}2^{\frac{-\lambda}{\lambda_0}}|x-z|^{\frac{\lambda}{\lambda_0}}.$$
$$= \left(2^{\frac{\lambda_0-\lambda}{\lambda_0}}\frac{C}{\lambda_0-\lambda} + 2^{\frac{\lambda_0-\lambda}{\lambda_0}}\frac{C}{\lambda}\right)|x-z|^{\frac{\lambda}{\lambda_0}} \le C_1|x-z|^{\frac{\lambda}{\lambda_0}}.$$

Doing similarly to (b), by DPP we have $|u(x)-u(z)| \le C_1|x-z|^{\frac{\lambda}{\lambda_0}}$ whenever $|x-z| < 2$, i.e $\frac{|u(x)-u(z)|}{|x-z|^{\frac{\lambda}{\lambda_0}}} \le C_1$. If $|x-z| \ge 2$, then from (A.6) we have

$$\frac{|u(x)-u(z)|}{|x-z|^{\frac{\lambda}{\lambda_0}}} \le \frac{2C}{\lambda 2^{\frac{\lambda}{\lambda_0}}} = 2^{1-\frac{\lambda}{\lambda_0}}\frac{C}{\lambda} = C_2 \qquad \Longrightarrow \qquad \frac{|u(x)-u(z)|}{|x-z|^{\frac{\lambda}{\lambda_0}}} \le \max\{C_1, C_2\} = C_3$$

for any $x \ne z$. Thus $|u(x)-u(z)| \le C_3|x-z|^{\frac{\lambda}{\lambda_0}}$.

□

Exercise 7. *Compute the Legendre's transform $L(x,q)$ for $H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ maps $(x,p) \longmapsto \frac{|p|^m}{m} + V(x)$, where $m > 1$ and $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous.*

Proof. We have

$$L(x,q) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot q - H(x,p)\Big) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot q - \frac{|p|^m}{m}\Big) - V(x).$$

The map $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ maps $p \longmapsto p\cdot q - \frac{|p|^m}{m}$ is continuous and

$$\lim_{|p|\longrightarrow\infty} f(p) = \lim_{|p|\longrightarrow\infty}|p|\left(\frac{p\cdot q}{|p|} - \frac{|p|^{m-1}}{m}\right) = -\infty$$

thus f achieves maximum on $\mathbb{R}^n$ at p^* such that $\nabla f(p^*) = 0$. We have

$$\nabla f(p) = q - p|p|^{m-2} = 0 \quad \Longleftrightarrow \quad p^*|p^*|^{m-2} = q \quad \Longrightarrow \quad f(p^*) = |q|^{\frac{m}{m-1}} - \frac{1}{m}|q|^{\frac{m}{m-1}} = \frac{m-1}{m}|q|^{\frac{m}{m-1}}.$$

Thus we have the Legendre's transform if $L(x,q) = \frac{m-1}{m}|q|^{\frac{m}{m-1}} - V(x)$. □

Exercise 8. *Consider the Cauchy problem:*

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

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 (C).

Proof. The initial condition is obviously true.
<u>SUBSOLUTION.</u> Take $\varphi \in C^1(\mathbb{R}^n \times [0,\infty))$ such that $u - \varphi$ has a strict max at (x_0,t_0), and $u(x_0,t_0) = \varphi(x_0,t_0)$, we need to prove

$$\varphi_t(x_0,t_0) + \sup_{q\in\mathbb{R}^n}\Big(q \cdot D\varphi(x_0,t_0) - L(x_0,q)\Big) \le 0.$$

Pick a path $\gamma(\cdot)$ with $\gamma(t_0) = x_0$, for $s < t_0$ we have

$$u(\gamma(t_0),t_0) - u(\gamma(s),s) \ge \varphi(\gamma(t_0),t_0) - \varphi(\gamma(s),s) = \int_s^{t_0}\Big(\varphi_t(\gamma(r),r) + \gamma'(r)\cdot D\varphi(\gamma(r),r)\Big)\,dr. \tag{A.7}$$

By dynamic programming principle (DPP) we have

$$\int_s^{t_0}\Big(L(\gamma(r),\gamma'(r))\Big)\,dr \ge u(\gamma(t_0),t_0) - u(\gamma(s),s). \tag{A.8}$$

From (A.7) and (A.8) 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 $s \longrightarrow t_0$ 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

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

since we can design the path $\gamma(\cdot)$ such that $\gamma'(t_0) = q$ for any $q \in \mathbb{R}^n$.
<u>SUPERSOLUTION.</u> Take $\varphi \in C^1(\mathbb{R}^n \times [0,\infty))$ such that $u - \varphi$ has a strict min at (x_0,t_0), and $u(x_0,t_0) = \varphi(x_0,t_0)$, we need to prove

$$\varphi_t(x_0,t_0) + \sup_{q\in\mathbb{R}^n}\Big(q \cdot D\varphi(x_0,t_0) - L(x_0,q)\Big) \ge 0.$$

For any $s \in (0, t_0)$ we have

$$u(\gamma(t_0), t_0) - u(\gamma(s), s) \le \varphi(\gamma(t_0), t_0) - \varphi(\gamma(s), s) = \int_s^{t_0} \Big(\varphi_t(\gamma(r), r) + \gamma'(r) \cdot D\varphi(\gamma(r), r) \Big)\, dr.$$

Subtracts two sides for $L(\gamma(r), \gamma'(r))$ we obtain

$$\begin{aligned} u(x_0, t_0) - \left(\int_s^{t_0} L(\gamma(r), \gamma'(r))\, dr + u(\gamma(s), s) \right) &\le \int_s^{t_0} \Big(\varphi_t(\gamma(r), r) + \gamma'(r) \cdot D\varphi(\gamma(r), r) - L(\gamma(r), \gamma'(r)) \Big)\, dr \\ &\le \int_s^{t_0} \Big(\varphi_t(\gamma(r), r) + H(\gamma(r), D\varphi(\gamma(r), r)) \Big)\, dr. \end{aligned}$$

Take the inf over all paths $\gamma(\cdot) \in \mathcal{A}$, thanks to lemma **??** we obtain

$$0 \le \sup_{\substack{\gamma(t_0) = x_0 \\ \gamma \in \mathcal{A}}} \underbrace{\int_s^{t_0} \Big(\varphi_t(\gamma(r), r) + H(\gamma(r), D\varphi(\gamma(r), r)) \Big)\, dr}_{\mathcal{K}[\gamma(\cdot)]} \tag{A.9}$$

where $\mathcal{A}$ to be the set of all "almost-admissible" paths with $\gamma(t_0) = x_0$, i.e $\gamma(\cdot)$ such that

$$\int_0^{t_0} L(\gamma(s), \gamma'(s))\, ds + u_0(\gamma(0)) < u(x_0, t_0) + 1.$$

Now for $\gamma(\cdot) \in \mathcal{A}$ we have

$$\begin{aligned} \mathcal{K}[\gamma(\cdot)] &= (t_0 - s)\Big(\varphi_t(x_0, t_0) + H(x_0, D\varphi(x_0, t_0)) \Big) \\ &\quad + \int_s^{t_0} \Big[\Big(\varphi_t(\gamma(r), r) - \varphi_t(x_0, t_0) \Big) + \Big(H(\gamma(r), D\varphi(\gamma(r), r)) - H(x_0, D\varphi(x_0, t_0)) \Big) \Big]\, dr. \end{aligned} \tag{A.10}$$

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 **??** we know that $|\gamma(r)|$ is bounded independent to $\gamma \in \mathcal{A}$ and $r < t_0$, thus since u is locally bounded, we can get $|u(\gamma(r), r)| \le C = C(x_0, t_0)$ for all $r \in [s, t_0]$. Thus given $\delta > 0$, by super-linearity we can choose M large so that

$$\inf_{x \in \mathbb{R}^n} \left(\frac{L(x, q)}{|q|} \right) > \frac{2(2C + 1)}{\delta} \qquad \text{for all} \qquad |q| \ge M. \tag{A.11}$$

3. Let $\varepsilon > 0$, by (DPP) with the version in lemma **??** we can find $\gamma \in \mathcal{A}$ such that ($\varepsilon \ll 1$)

$$\int_s^{t_0} L(\gamma(r), \gamma'(r))\, dr \le u(x_0, t_0) - u(\gamma(s), s) + \varepsilon \le 2C + 1. \tag{A.12}$$

- Estimate for the first term is easy:

$$\int_{\{r\in[s,t_0]:|\gamma'(r)|\le M\}} |\gamma'(r)|\,dr \le M(t_0-s). \tag{A.13}$$

- Estimate for the second term:

$$\int_{\{r\in[s,t_0]:|\gamma'(r)|\le M\}} L(\gamma(r),\gamma'(r))\,dr \ge -\left(\sup_{\substack{x\in\mathbb{R}^n\\|q|\le M}} L(x,q)\right) =: -C_M.$$

and

$$\begin{aligned}\int_{\{r\in[s,t_0]:|\gamma'(r)|\ge M\}} L(\gamma(r),\gamma'(r))\,dr &= \int_{\{r\in[s,t_0]:|\gamma'(r)|\ge M\}} \left(\frac{L(\gamma(r),\gamma'(r))}{|\gamma'(r)|}\right)|\gamma'(r)|\,dr\\ &\ge \left(\inf_{\substack{x\in\mathbb{R}^n\\|q|\ge M}} \frac{L(x,q)}{|q|}\right)\int_{\{r\in[s,t_0]:|\gamma'(r)|\ge M\}} |\gamma'(r)|\,dr.\end{aligned}$$

From this we obtain

$$\int_s^{t_0} L(\gamma(r),\gamma'(r))\,dr + C_M \ge \left(\inf_{\substack{x\in\mathbb{R}^n\\|q|\ge M}} \frac{L(x,q)}{|q|}\right)\int_{\{r\in[s,t_0]:|\gamma'(r)|\ge M\}} |\gamma'(r)|\,dr.$$

From (A.12) and (A.11) we obtain

$$\int_{\{r\in[s,t_0]:|\gamma'(r)|\ge M\}} |\gamma'(r)|\,dr \le \left(\inf_{\substack{x\in\mathbb{R}^n\\|q|\ge M}} \frac{L(x,q)}{|q|}\right)^{-1}(2C+1) \le \frac{\delta}{2}. \tag{A.14}$$

4. With M in step 2, (A.13) and (A.14) yields

$$\begin{aligned}\sup_{r\in[s,t_0]} |\gamma(r)-x_0| \le \int_s^t |\gamma'(r)|\,dr &\le \int_{\{r\in[s,t]:|\gamma'(r)|\le M\}} |\gamma'(r)|\,dr + \int_{\{r\in[s,t]:|\gamma'(r)|\ge M\}} |\gamma'(r)|\,dr\\ &\le M(t_0-s)+\frac{\delta}{2}.\end{aligned}$$

All of these steps are true for every $s\in[0,t]$, thus choosing s closed to t_0 such that $M(t_0-s)<\frac{\delta}{2}$, then

$$\sup_{r\in[s,t_0]} |\gamma(r)-x_0| < \delta$$

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.10) we obtain for any given η then there exists $s\in[0,t_0]$ such that

$$\mathcal{K}[\gamma(\cdot) \le (t_0-s)\Big(\varphi_t(x_0,t_0)+H(x_0,D\varphi(x_0,t_0))\Big)+2\eta(t_0-s).$$

Taking sup over all path $\gamma(\cdot)\in\mathcal{A}$ and divide both sides by $t_0-s>0$ we obtain

$$\varphi_t(x_0,t_0)+H(x_0,D\varphi(x_0,t_0))+2\eta \ge 0.$$

Finally since η is arbitrary, $\varphi_t(x_0,t_0)+H(x_0,D\varphi(x_0,t_0))\ge 0$ and the proof is complete. □

Exercise 9. *There exists a constant $C > 0$ such that for any $\lambda > 0$ we have*

$$\left|\lambda v^\lambda(\cdot) + \overline{H}(p)\right|_{C(\mathbb{T}^n)} \le C\lambda.$$

To be precise, if $|Dv^\lambda(\cdot)| \le C$ in the viscosity sense, then

$$\left|\lambda v^\lambda(\cdot) + \overline{H}(p)\right|_{C(\mathbb{T}^n)} \le \left(C\sqrt{n}\right)\lambda.$$

Proof. Recall that if we let $C = \max_{y\in\mathbb{T}^n} H(y,p)$ then by comparison principle $\sup_{y\in\mathbb{T}^n} |\lambda v^\lambda(\lambda)| \le C$. Coercivity implies that $\sup_{y\in\mathbb{T}^n} |Dv^\lambda(y)| \le C_1$, and for all $y, x_0 \in \mathbb{T}^n$ we have

$$\begin{aligned}|v^\lambda(y) - v^\lambda(x_0)| \le C_1\sqrt{n} \quad &\Longrightarrow \quad \lambda v^\lambda(x_0) - \lambda C_1\sqrt{n} \le \lambda v^\lambda(y) \le \lambda v^\lambda(x_0) + \lambda C_1\sqrt{n}\\ &\Longrightarrow \quad \lambda \sup_{\mathbb{T}^n} v^\lambda(\cdot) - \lambda C_1\sqrt{n} \le \lambda v^\lambda(y) \le \lambda \inf_{\mathbb{T}^n} v^\lambda(\cdot) + \lambda C_1\sqrt{n}\end{aligned}$$

since it is true for all $x_0 \in \mathbb{T}^n$. From that it is suffices to prove that:

$$\lambda \inf_{\mathbb{T}^n} v^\lambda(\cdot) \le -\overline{H}(p) \le \lambda \sup_{\mathbb{T}^n} v^\lambda(\cdot) \quad \Longleftrightarrow \quad \underbrace{-\lambda \sup_{\mathbb{T}^n} v^\lambda(\cdot)}_{\alpha} \le \overline{H}(p) \le \underbrace{-\lambda \inf_{\mathbb{T}^n} v^\lambda(\cdot)}_{\beta}. \tag{A.15}$$

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 for all $y \in \mathbb{T}^n$ we have

$$H(y, p + Dv(y)) = \overline{H}(p) > \alpha \ge -\lambda v^\lambda(y) = H\left(y, p + Dv^\lambda(y)\right).$$

Since $v, v^\lambda \in \mathrm{Lip}(\mathbb{T}^n)$ are bounded, we can choose $\delta > 0$ such that[5]

$$\delta v(y) + H(y, p + Dv(y)) > \frac{\overline{H}(p) + \alpha}{2} > \delta v^\lambda(y) + H\left(y, p + Dv^\lambda(y)\right)$$

in the viscosity sense. Then $v(\cdot)$ and $v^\lambda(\cdot)$ are viscosity supersolution and subsolution to the problem $\delta w + H(y, p + Dw) = \frac{1}{2}(\overline{H}(p) + \alpha)$ respectively, thus by comparison principle $v \ge v^\lambda$. This is a contradiction since $v - C$ is also a viscosity solution to the cell problem for any constant C. Doing similarly for β we have (A.15) is true, and thus we obtain the rate of convergence is $O(\lambda)$. □

[5]Precisely, we can choose δ to be $\delta \max\left\{\max_{\mathbb{T}^n} |v(\cdot)|, \max_{\mathbb{T}^n} |v^\lambda(\cdot)|\right\} \le \frac{\overline{H}(p) - \delta}{2}$.

Bibliography

[1] 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.

[2] 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.

[3] S. N. Armstrong, H. V. Tran, *Viscosity solutions of general viscous Hamilton–Jacobi equations*, Mathematische Annalen, 361 (2015), no. 3, 647–687.

[4] 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.

[5] 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.

[6] S. Aubry, *The twist map, the extended Frenkel–Kantorova model and the devil's staircase*, Physica D 7 (1983), 240–258.

[7] V. Bangert, *Mather Sets for Twist Maps and Geodesics on Tori*, Dynamics Reported, Volume 1, 1988.

[8] V. Bangert, *Minimal geodesics*, Ergod. Th. & Dynam. Sys. (1989), 10, 263–286.

[9] 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.

[10] G. Barles, Solutions de viscosité des équations de Hamilton-Jacobi, Mathématiques & Applications (Berlin), 17, Springer-Verlag, Paris, 1994.

[11] 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.

[12] U. Bessi, *Aubry-Mather theory and Hamilton-Jacobi equations*, Comm. Math. Phys. 235 (2003), no. 3, 495–511.

[13] A. Bressan, Viscosity Solutions of Hamilton-Jacobi Equations and Optimal Control Problems, lecture notes on Bressan's webpage.

[14] 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.

[15] 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.

[16] J. Calder, Lecture notes on viscosity solutions, lecture notes on Calder's website.

[17] 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.

[18] 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.

[19] M. J. Carneiro, *On minimizing measures of the action of autonomous Lagrangians*, Nonlinearity 8 (1995) 1077–1085.

[20] 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.

[21] M. C. Concordel, *Periodic homogenization of Hamilton–Jacobi equations: additive eigenvalues and variational formula*, Indiana Univ. Math. J. 45 (1996), no. 4, 1095–1117.

[22] M. C. Concordel, *Periodic homogenisation of Hamilton–Jacobi equations. II. Eikonal equations*, Proc. Roy. Soc. Edinburgh Sect. A 127 (1997), no. 4, 665–689.

[23] 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.

[24] 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.

[25] M. G. Crandall, P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), 1–42.

[26] M. G. Crandall and P-L. Lions, *Two approximations of solutions of Hamilton-Jacobi equations*, Math. Comp. 43 (1984), 1–19.

[27] A. Davini, A. Fathi, R. Iturriaga, M. Zavidovique, *Convergence of the solutions of the discounted equation*, Invent. Math. 206 (1) (2016) 29–55.

[28] W. E, *Aubry-Mather theory and periodic solutions of the forced Burgers equation*, Comm. Pure Appl. Math., 52(7): 811–828, 1999.

[29] L. C. Evans, *On solving certain nonlinear partial differential equations by accretive operator methods*, Israel J. Math. 86 (1980), 225–247.

[30] L. C. Evans, Weak convergence methods for nonlinear partial differential equations. CBMS 74, American Mathematical Society, 1990.

[31] L. C. Evans, *Periodic homogenisation of certain fully nonlinear partial differential equations*, Proceedings of the Royal Society of Edinburgh, 120A, 245–265, 1992.

[32] L. C. Evans, Partial differential equations. Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010.

[33] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton–Jacobi PDE*, Archive for Rational Mechanics and Analysis 197 (2010), 1053–1088.

[34] 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.

[35] L. C. Evans, *A survey of partial differential equations methods in weak KAM theory*, Comm. Pure Appl. Math. 57 (2004), no. 4, 445–480.

[36] L. C. Evans, D. Gomes, *Effective Hamiltonians and Averaging for Hamiltonian Dynamics. I*, Arch. Ration. Mech. Anal. 157 (2001), no. 1, 1–33.

[37] L. C. Evans, and J. Spruck, *Motion of level sets by mean curvature. I.* J. Diff. Geom. **33**, 1991, No. 3, 635–681.

[38] 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.

[39] 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.

[40] A. Fathi, Weak KAM Theorem in Lagrangian Dynamics, to appear in Cambridge Studies in Advanced Mathematics.

[41] A. Fathi, A. Siconolfi, *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, Invent. Math. 155 (2004), no. 2, 363–388.

[42] W. Fleming, *The convergence problem for differential games II*, Advances in Game Theory, 1964, Princeton Univ. Press, 195–210.

[43] 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.

[44] H. Gao, *Random homogenization of coercive Hamilton-Jacobi equations in 1d*, Calc. Var. Partial Differential Equations, 55 (2016), no. 2, 1–39.

[45] H. Gao, *Stochastic homogenization of certain nonconvex Hamilton-Jacobi equations*, arXiv:1803.08633 [math.AP].

[46] Y. Giga, Surface Evolution Equations. A Level Set Approach, Monographs in Mathematics, 99. Birkhäuser, Basel-Boston-Berlin (2006), xii+264pp.

[47] 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.

[48] 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*, accepted.

[49] D. A. Gomes, *Viscosity solutions of Hamilton-Jacobi equations, and asymptotics for Hamiltonian systems*, Calc. Var. 14, 345–357 (2002).

[50] D. A. Gomes, *A stochastic analogue of Aubry-Mather theory*, Nonlinearity 15 (2002), no. 3, 581–603.

[51] 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.

[52] D. A. Gomes, *Generalized Mather problem and selection principles for viscosity solutions and Mather measures*, Adv. Calc. Var., 1 (2008), 291–307.

[53] 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.

[54] 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.

[55] 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.

[56] G. A. Hedlund, *Geodesics on a two-dimensional Riemannian manifold with periodic coefficients*, Ann. of Math. 33 (1932), 719–739.

[57] 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.

[58] H. Ishii, *Perron's method for Hamilton–Jacobi equations*, Duke Math Journal, 55 (1987), no. 2, 369–384.

[59] H. Ishii, *On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions*, Funkcial. Ekvac. 38 (1995), no. 1, 101–120.

[60] 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.

[61] H. Ishii, Lecture notes on the weak KAM theorem, lecture notes available on Ishii's website.

[62] 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.

[63] 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.

[64] R. Iturriaga, H. Sanchez-Morgado, *On the stochastic Aubry-Mather theory*, Bol. Soc. Mat. Mexicana (3) 11 (2005), no. 1, 91–99.

[65] R. Iturriaga, H. Sanchez-Morgado, *Limit of the infinite horizon discounted Hamilton–Jacobi equation*, Discrete Contin. Dyn. Syst. Ser. B, 15 (2011), 623–635.

[66] R. Jensen, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rat. Mech. Anal. 101 (1988), 1–27.

[67] 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.

[68] V. Kaloshin, *Mather theory, weak KAM, and viscosity solutions of Hamilton–Jacobi PDE's*, EQUADIFF 2003, 39–48, World Sci. Publ., Hackensack, NJ, 2005.

[69] S. Koike, A Beginner's Guide to the Theory of Viscosity Solutions, MSJ Memoir, 13, Math. Soc. Japan, Tokyo, 2004.

[70] H. Komiya, *Elementary proof for Sion's minimax theorem*, Kodai Mathematical Journal. 11 (1988) (1): 5–7.

[71] 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.

[72] 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.

[73] P.-L. Lions, Generalized solutions of Hamilton-Jacobi equations, Research Notes in Mathematics, Vol. 69, Pitman, Boston, Masso. London, 1982.

[74] P.-L. Lions, G. Papanicolaou, S. R. S. Varadhan, *Homogenization of Hamilton–Jacobi equations*, unpublished work (1987).

[75] 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.

[76] 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.

[77] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity 9 (1996), no. 2, 273–310.

[78] J. N. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems,* Math. Z. 207 (1991), no. 2, 169–207.

[79] J. N. Mather, *Differentiability of the minimal average action as a function of the rotation number,* Bol. Soc. Bras. Math 21 (1990), 59–70.

[80] A. Melikyan, Generalized characteristics of first order PDEs. Applications in optimal control and differential games. Birkhäuser Boston, Inc., Boston, MA, 1998. xiv+310 pp.

[81] 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.

[82] H. Mitake, H. V. Tran, *Selection problems for a discount degenerate viscous Hamilton–Jacobi equation* , Adv. Math., 306 (2017), 684–703.

[83] H. Mitake, H. V. Tran, *On uniqueness sets of additive eigenvalue problems and applications*, Proc. Amer. Math. Soc., 146, no 11, 4813–4822.

[84] H. Mitake, H. V. Tran, Y. Yu, *Rate of convergence in periodic homogenization of Hamilton-Jacobi equations: the convex setting*, arXiv:1801.00391 [math.AP], Arch. Ration. Mech. Anal., accepted.

[85] 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.

[86] 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.

[87] 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.

[88] M. Sion, *On general minimax theorems,* Pacific Journal of Mathematics. 8 (1958) (1): 171–176.

[89] A. Sorrentino, Action-Minimizing Methods in Hamiltonian Dynamics. An Introduction to Aubry-Mather Theory, Mathematical Notes Series Vol. 50 (Princeton University Press), 2015.

[90] P. E. Souganidis, *Approximation schemes for viscosity solutions of Hamilton-Jacobi equations*, J. Differential Equations 59 (1985), no. 1, 1–43.

[91] H. V. Tran, *Adjoint methods for static Hamilton-Jacobi equations*, Calc. Var. Partial Differential Equations 41 (2011), 301–319.

[92] H. V. Tran, Y. Yu, A rigidity result for effective Hamiltonians with 3-mode periodic potentials, *Advances in Math.*, 334, 300–321.

[93] S. N.T. Tu, *Rate of convergence for periodic homogenization of convex Hamilton-Jacobi equations in one dimension*, arXiv:1808.06129 [math.AP]

[94] 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.