# Dynamical properties of Hamilton–Jacobi equations via the nonlinear adjoint method: Large time behavior and Discounted approximation

Hiroyoshi Mitake[1]

*Institute of Engineering, Division of Electrical, Systems and Mathematical Engineering, Hiroshima University*
*1-4-1 Kagamiyama, Higashi-Hiroshima-shi 739-8527, Japan*

Hung V. Tran[2]

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

December 4, 2016

[1]Email address: `hiroyoshi-mitake@hiroshima-u.ac.jp`
[2]Email address: `hung@math.wisc.edu`

# Contents

# Introduction

These notes are based on the two courses given by the authors at the summer school on "PDE and Applied Mathematics" at Vietnam Institute for Advanced Study in Mathematics (VIASM) from July 14 to July 25, 2014. The first course was about the basic theory of viscosity solutions, and and the second course was about asymptotic analysis of Hamilton–Jacobi equations. In particular, we focused on the large time asymptotics of solutions and the selection problem of the discounted approximation.

We study both the inviscid (or first-order) Hamilton–Jacobi equation

$$u_t(x,t) + H(x, Du(x,t)) = 0 \quad \text{for } x \in \mathbb{R}^n,\ t > 0, \tag{0.1}$$

and the viscous Hamilton–Jacobi equation

$$u_t(x,t) - \Delta u(x,t) + H(x, Du(x,t)) = 0 \quad \text{for } x \in \mathbb{R}^n,\ t > 0. \tag{0.2}$$

Here, $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is the unknown, and $u_t, Du, \Delta u$ denote the time derivative, the spatial gradient and the Laplacian of $u$, respectively. The Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a given continuous function. We will add suitable assumptions later when needed. At some points, we also consider the general (possibly degenerate) viscous Hamilton–Jacobi equation:

$$u_t(x,t) - \operatorname{tr}\big(A(x)D^2u(x,t)\big) + H(x, Du(x,t)) = 0 \quad \text{for } x \in \mathbb{R}^n,\ t > 0, \tag{0.3}$$

where $D^2u$ denotes the Hessian of $u$, and $A : \mathbb{R}^n \to \mathbb{M}^{n\times n}_{\mathrm{sym}}$ is a given continuous diffusion matrix, which is nonnegative definite and possibly degenerate. Here, $\mathbb{M}^{n\times n}_{\mathrm{sym}}$ is the set of $n \times n$ real symmetric matrices, and for $S \in \mathbb{M}^{n\times n}_{\mathrm{sym}}$, $\operatorname{tr}(S)$ denotes the trace of matrix $S$. The assumptions on $A$ will be specified later.

In the last decade, there has been much interest on dynamical properties of viscosity solutions of (0.1)–(0.3). Indeed, in view of the weak Kolmogorov–Arnold–Moser theory (weak KAM theory) established by Fathi (see [34]), the asymptotic analysis of solutions to Hamilton–Jacobi equation (0.1) with convex Hamiltonian $H$ has been dramatically developed. One of the features of this lecture note is to introduce a new way to investigate dynamical properties of solutions of (0.1)–(0.3) and related equations by using PDE methods. More precisely, we use the nonlinear adjoint method introduced by Evans [32] together with some new conserved quantities and estimates to study several type of asymptotic problems. The main point of this method is to look at the behavior of the solution of the regularized Hamilton–Jacobi equation combined with the adjoint equation of its linearized operator to derive new information about the solution, which could not be obtained by previous techniques. Evans introduced this method to study the gradient shock structures of the vanishing viscosity procedure of viscosity solutions. With Cagnetti, Gomes, the authors used this method to study the large-time behavior of solutions to (0.3). Another interesting topic is about the selection problem in the discounted approximation setting. This was studied by Davini, Fathi, Iturriaga, Zavidovique [26] by using a dynamical approach, and the authors [71] by using a nonlinear adjoint method.

The outline of the lecture notes is as follows. In Chapter 1, we investigate the ergodic problems associated with (0.1)–(0.3). In particular, we prove the existence of solutions to the ergodic problems. In Chapters 2 and 3, we study the large time behavior of solutions to (0.1)–(0.3), and the selection problem for the discounted approximation, respectively. To make the lecture notes self-contained, we prepare a brief introduction to the theory of viscosity solutions of first-order equations in Appendix. Appendix can be read independently from other chapters. Also, we note that Chapters 2 and 3 can be read independently.

It is worth pointing out that these lecture notes reflect the state of the art of the subject by the end of summer 2014. We will address some up-to-date developments at the end of Chapters 2 and 3.

**Acknowledgement.** The work of HM was partially supported by JSPS grants: KAKENHI #15K17574, #26287024, #16H03948, and the work of HT was partially supported by NSF grants DMS-1361236 and DMS-1615944. These lecture notes were started while the authors visited Vietnam Institute for Advanced Study in Mathematics (VIASM) in July 2014. The authors are extremely grateful for their hospitality. The authors thank the anonymous referee for carefully reading the previous version of the manuscript and valuable comments. The authors also thank Son Van for his thoughtful suggestions.

# Notations

- For $n \in \mathbb{N}$, $\mathbb{T}^n$ is the $n$-dimensional flat torus, that is, $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.
- For $x = (x_1, \dots, x_n)$, $y = (y_1, \dots y_n) \in \mathbb{R}^n$, $x \cdot y = x_1 y_1 + \cdots x_n y_n$ denotes the Euclidean inner product of $x$ and $y$.
- For $x \in \mathbb{R}^n$ and $r > 0$, $B(x, r)$ denotes the open ball with center $x$ and radius $r$, that is, $B(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$.
- $\mathbb{M}^{n\times n}_{\mathrm{sym}}$ is the set of $n \times n$ real symmetric matrices.
- $I_n$ denotes the identity matrix of size $n$.
- For $S \in \mathbb{M}^{n\times n}_{\mathrm{sym}}$, $\mathrm{tr}\,(S)$ denotes the trace of matrix $S$.
- For $A, B \in \mathbb{M}^{n\times n}_{\mathrm{sym}}$, $A \geq B$ (or $B \leq A$) means that $A - B$ is nonnegative definite.
- Given a metric space $X$, $C(X), \mathrm{USC}\,(X), \mathrm{LSC}\,(X)$ denote the space of all continuous, upper semicontinuous, lower semicontinuous functions in $X$, respectively. Let $C_c(X)$ denote the space of all continuous functions in $X$ with compact support.
- For any interval $J \subset \mathbb{R}$, $\mathrm{AC}\,(J, \mathbb{R}^m)$ is the set of all absolutely continuous functions in $J$ with value in $\mathbb{R}^m$.
- For $U \subset \mathbb{R}^n$ open, $k \in \mathbb{N}$ and $\alpha \in (0, 1]$, $C^k(U)$ and $C^{k,\alpha}(U)$ denote the space of all functions whose $k$-th order partial derivatives are continuous and Hölder continuous with exponent $\alpha$ in $U$, respectively. Also $C^\infty(U)$ is the set of all infinitely differentiable functions in $U$.
- For $U \subset \mathbb{R}^n$ open, $\mathrm{Lip}\,(U)$ is the set of all Lipschitz continuous function in $U$.
- $L^\infty$ norm of $u$ in $U$ is defined as

$$\|u\|_{L^\infty(U)} = \operatorname*{ess\,sup}_U |u|.$$

- For $u : \mathbb{R}^n \to \mathbb{R}$, we denote by $Du$ the gradient of $u$, that is,

$$Du = \nabla u = (u_{x_1}, \dots, u_{x_n}) = \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}\right).$$

- For $u : \mathbb{R}^n \to \mathbb{R}$, $D^2u$ denotes the Hessian matrix of $u$

$$D^2u = \left(u_{x_i x_j}\right)_{1 \le i,j \le n} = \left(\frac{\partial^2 u}{\partial x_i \partial x_j}\right)_{1 \le i,j \le n},$$

  and $\Delta u$ denotes the Laplacian of $u$

$$\Delta u = \operatorname{tr}\left(D^2 u\right) = \sum_{i=1}^{n} u_{x_i x_i}.$$

- We use the letter $C$ to denote any constant which can be explicitly computed in terms of known quantities. The exact value of $C$ could change from line to line in a given computation.

# Chapter 1

# Ergodic problems for Hamilton–Jacobi equations

## 1.1 Motivation

One of our main goals in the lecture note is to understand the large-time behavior of the solutions to various Hamilton–Jacobi type equations. We cover both the first-order and the second-order cases. The first-order equation is of the form

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

The viscous Hamilton–Jacobi equation is of the form

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

More generally, we consider the possibly degenerate viscous equation

$$\begin{cases} u_t - \operatorname{tr}\big(A(x)D^2u\big) + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ u(x, 0) = u_0(x) & \text{on } \mathbb{R}^n, \end{cases} \tag{1.3}$$

under rather general assumptions on the Hamiltonian $H$, and the diffusion $A$.

The problem of interest is the behavior of $u(x, t)$ as $t \to \infty$. In this section, we first give a heuristic (formal) argument to find out possible candidates for the limiting profiles. Let us work with (1.1) for now.

We *always* assume hereinafter the *coercivity* condition on $H$, that is,

$$H(x, p) \to \infty \quad \text{as } |p| \to \infty \ \text{ uniformly for } x \in \mathbb{R}^n. \tag{1.4}$$

It is often the case that we need to guess an expansion form of $u(x, t)$ when we do not know yet how it behaves as $t \to \infty$. Let us consider a *formal asymptotic expansion* of $u(x, t)$

$$u(x, t) = a_1(x)t + a_2(x) + a_3(x)t^{-1} + \dots,$$

where $a_i \in C^\infty(\mathbb{R}^n)$ for all $i \geq 1$. Plug this into equation (1.1) to yield

$$a_1(x) - a_3(x)t^{-2} + \ldots + H(x, Da_1(x)t + Da_2(x) + Da_3(x)t^{-1} + \ldots) = 0.$$

In view of (1.4), we should have $Da_1(x) \equiv 0$ as other terms are bounded with respect to $t$ as $t \to \infty$, which therefore implies that the function $a_1$ should be constant. Thus, there exists $c \in \mathbb{R}$ such that $a_1(x) \equiv -c$ for all $x \in \mathbb{R}^n$. Set $v(x) = a_2(x)$ for $x \in \mathbb{R}^n$. From this observation, we expect that the large-time behavior of the solution to (1.1) is

$$u(x,t) - (v(x) - ct) \to 0 \quad \text{locally uniformly for } x \in \mathbb{R}^n \text{ as } t \to \infty, \tag{1.5}$$

for some function $v$ and constant $c$. Moreover, if convergence (1.5) holds, then by the stability result of viscosity solutions (see Section 5.5), the pair $(v, c)$ satisfies

$$H(x, Dv) = c \quad \text{in } \mathbb{R}^n \text{ in the viscosity sense.}$$

Therefore, in order to investigate whether convergence (1.5) holds or not, we first need to study the well-posedness of the above problem. We call it an *ergodic problem* for Hamilton–Jacobi equations. This ergodic problem will be one of the main objects in the next section.

**Remark 1.1.** One may wonder why we do not consider terms like $b_i(x)t^i$ for $i \geq 2$ in the above formal asymptotic expansion of $u$. We will give a clear explanation at the end of this chapter.

## 1.2 Existence of solutions to ergodic problems

Henceforth, we consider the situation that everything is assumed to be $\mathbb{Z}^n$-periodic with respect to the spatial variable $x$. As it is equivalent to consider the equations in the $n$-dimensional torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, we always use this notation.

In this section, we consider *ergodic problems* for first-order and second-order Hamilton–Jacobi equations. The ergodic problem for the inviscid (first-order) case is the one addressed in the previous section

$$H(x, Dv) = c \quad \text{in } \mathbb{T}^n. \tag{1.6}$$

For second-order equations, we consider both the ergodic problem for the viscous case

$$-\Delta v + H(x, Dv) = c \quad \text{in } \mathbb{T}^n, \tag{1.7}$$

and, more generally, the ergodic problem for the possibly degenerate viscous case

$$-\mathrm{tr}\left(A(x)D^2 v(x)\right) + H(x, Dv) = c \quad \text{in } \mathbb{T}^n. \tag{1.8}$$

In all cases, we seek for a pair of unknowns $(v, c) \in C(\mathbb{T}^n) \times \mathbb{R}$ so that $v$ solves the corresponding equation in the viscosity sense.

We give three results on the existence of solutions $(v, c) \in C(\mathbb{T}^n) \times \mathbb{R}$ to (1.6)–(1.8). The last one includes the first two results, but we study all of them separately as each is important in its own right. Besides, the set of assumptions for each case is slightly different.

The first result concerns the inviscid case.

**Theorem 1.1.** *Assume that $H \in C(\mathbb{T}^n \times \mathbb{R}^n)$ and that $H$ satisfies* (1.4). *Then there exists a pair $(v, c) \in \mathrm{Lip}\,(\mathbb{T}^n) \times \mathbb{R}$ such that $v$ solves* (1.6) *in the viscosity sense.*

*Proof.* For $\delta > 0$, consider the following approximate problem

$$\delta v^\delta + H(x, Dv^\delta) = 0 \quad \text{ in } \mathbb{T}^n. \tag{1.9}$$

Setting $M := \max_{x\in\mathbb{T}^n} |H(x, 0)|$, we have $\pm M/\delta$ is a subsolution and supersolution of (1.9), respectively (see Section 4.2 for the definitions). By the Perron method in the theory of viscosity solutions (see Section 4.7), there exists a unique viscosity solution $v^\delta$ to (1.9) such that

$$|v^\delta| \le M/\delta,$$

which implies further that $H(x, Dv^\delta) \le M$. In view of coercivity assumption (1.4), we get

$$|Dv^\delta| \le C \quad \text{for some } C > 0 \text{ independent of } \delta. \tag{1.10}$$

Therefore, we obtain that $\left\{v^\delta(\cdot) - v^\delta(x_0)\right\}_{\delta>0}$ is equi-Lipschitz continuous for a fixed $x_0 \in \mathbb{T}^n$. Moreover, noting that

$$|v^\delta(x) - v^\delta(x_0)| \le \|Dv^\delta\|_{L^\infty(\mathbb{T}^n)}|x - x_0| \le C,$$

we see that $\left\{v^\delta(\cdot) - v^\delta(x_0)\right\}_{\delta>0}$ is uniformly bounded in $\mathbb{T}^n$. Thus, in light of the Arzelà–Ascoli theorem, there exists a subsequence $\{\delta_j\}_j$ converging to 0 so that $v^{\delta_j}(\cdot) - v^{\delta_j}(x_0) \to v$ uniformly on $\mathbb{T}^n$ as $j \to \infty$. Since $|\delta_j v^{\delta_j}(x_0)| \le M$, by passing to another subsequence if necessary, we obtain that

$$\delta_j v^{\delta_j}(x_0) \to -c \quad \text{ for some } \; c \in \mathbb{R}.$$

In view of the stability result of viscosity solutions, we get the conclusion. □

**Remark 1.2.** Let us notice that the approximation procedure above using (1.9) is called the *discounted approximation procedure*. It is a very natural procedure in many ways. Firstly, the approximation makes equation (1.9) strictly monotone in $v^\delta$, which fits perfectly in the well-posedness setting of viscosity solutions. See Section 4.1.2 for the formula of $v^\delta$ in terms of optimal control.

Secondly, for $w^\delta(x) = \delta v^\delta(x/\delta)$, $w^\delta$ solves

$$w^\delta + H\left(\frac{x}{\delta}, Dw^\delta\right) = 0 \quad \text{in } \mathbb{T}^n,$$

which is the setting to study an important phenomenon called *homogenization*.

The arguments in the proof of Theorem 1.1 are soft as we just use a priori estimate (1.10) on $|Dv^\delta|$ and the Arzelà–Ascoli theorem to get the result. In particular, from this argument, we only know convergence of $\left\{v^{\delta_j} - v^{\delta_j}(x_0)\right\}_j$ via the subsequence $\{\delta_j\}_j$. It is not clear at all at this moment whether $\left\{v^\delta - v^\delta(x_0)\right\}_{\delta>0}$ converges uniformly as $\delta \to 0$ or not. We will come back to this question and give a positive answer under some additional assumptions in Chapter 3.

Let us now provide the existence proof for the viscous case. To do this, we need a sort of superlinearity condition on $H$:

$$\lim_{|p|\to\infty} \left( \frac{1}{2n} H(x,p)^2 + D_x H(x,p)\cdot p \right) = +\infty \quad \text{uniformly for } x \in \mathbb{T}^n. \tag{1.11}$$

**Theorem 1.2.** *Assume that $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$ and that $H$ satisfies* (1.11)*. Then there exists a pair $(v,c) \in \mathrm{Lip}\,(\mathbb{T}^n) \times \mathbb{R}$ such that $v$ solves* (1.7) *in the viscosity sense.*

*Proof.* The proof is based on the standard Bernstein method. For $\delta > 0$, consider the approximate problem

$$\delta v^\delta - \Delta v^\delta + H(x, Dv^\delta) = 0 \quad \text{ in } \mathbb{T}^n. \tag{1.12}$$

By repeating the first step in the proof of Theorem 1.1, we obtain the existence of a solution $v^\delta$ to the above. Note that in this case, by the classical regularity theory for elliptic equations, $v^\delta$ is smooth due to the appearance of the diffusion $\Delta v^\delta$.

Differentiate (1.12) with respect to $x_i$ to get

$$\delta v^\delta_{x_i} - \Delta v^\delta_{x_i} + H_{x_i} + D_p H \cdot Dv^\delta_{x_i} = 0.$$

Multiplying this by $v^\delta_{x_i}$ and summing up with respect to $i$, we achieve that

$$\delta |Dv^\delta|^2 - \Delta v^\delta_{x_i} v^\delta_{x_i} + D_x H \cdot Dv^\delta + D_p H \cdot Dv^\delta_{x_i} v^\delta_{x_i} = 0.$$

Here we use Einstein's convention. Set $\varphi := |Dv^\delta|^2/2$. Noting that

$$\varphi_{x_j} = v^\delta_{x_i} v^\delta_{x_i x_j} \quad \text{and} \quad \varphi_{x_j x_j} = v^\delta_{x_i x_j} v^\delta_{x_i x_j} + v^\delta_{x_i} v^\delta_{x_i x_j x_j},$$

we obtain

$$\Delta \varphi = |D^2 v^\delta|^2 + \Delta v^\delta_{x_i} v^\delta_{x_i}.$$

Thus, $\varphi$ satisfies

$$2\delta\varphi - (\Delta\varphi - |D^2 v^\delta|^2) + D_x H \cdot Dv^\delta + D_p H \cdot D\varphi = 0.$$

Take a point $x_0 \in \mathbb{T}^n$ such that $\varphi(x_0) = \max_{\mathbb{T}^n} \varphi \geq 0$. As we have $D\varphi(x_0) = 0$, $D^2\varphi(x_0) \leq 0$, we obtain

$$|D^2 v^\delta(x_0)|^2 + D_x H \cdot Dv^\delta(x_0) \leq 0.$$

Noting furthermore that

$$|D^2 v^\delta(x_0)|^2 \geq \frac{1}{n}|\Delta v^\delta(x_0)|^2 \geq \frac{1}{2n} H(x_0, Dv^\delta(x_0))^2 - C$$

for some $C > 0$. Thus,

$$\frac{1}{2n} H(x_0, Dv^\delta(x_0))^2 + D_x H(x_0, Dv^\delta(x_0)) \cdot Dv^\delta(x_0) \leq C.$$

In light of (1.11), we get a priori estimate $\|Dv^\delta\|_{L^\infty(\mathbb{T}^n)} \leq C$. This is enough to get the conclusion as in the proof of Theorem 1.1. □

Here is a generalization of Theorems 1.1 and 1.2 to the degenerate viscous setting. We use the following assumptions:

(H1) $A(x) = (a^{ij}(x))_{1 \leq i,j \leq n} \in \mathbb{M}^{n\times n}_{\mathrm{sym}}$ with $A(x) \geq 0$, and $a^{ij} \in C^2(\mathbb{T}^n)$ for all $i, j \in \{1, \ldots, n\}$,

and there exists $\gamma > 1$ and $C > 0$ such that

$$\begin{cases} |D_x H(x,p)| \leq C(1 + |p|^\gamma) & \text{for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n, \\ \lim_{|p|\to\infty} \dfrac{H(x,p)^2}{|p|^{1+\gamma}} = +\infty & \text{uniformly for } x \in \mathbb{T}^n. \end{cases} \tag{1.13}$$

We remark that (1.13) is also a sort of superlinearity condition. It is clear that (1.13) is stronger than (1.11). We need to require a bit more than (1.11) as we deal with the general diffusion matrix $A$ here. We use the label (H1) for the assumptions on $A$ as it is one of the main assumptions in the lecture notes, and we will need to use it later.

**Theorem 1.3.** *Assume that $A$ satisfies* (H1). *Assume further that $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$ and $H$ satisfies* (1.13). *Then there exists a pair $(v, c) \in \mathrm{Lip}\,(\mathbb{T}^n) \times \mathbb{R}$ such that $v$ solves* (1.8) *in the viscosity sense.*

*Proof.* We first consider the discount approximation

$$\delta v^\delta - \mathrm{tr}\left(A(x) D^2 v^\delta\right) + H(x, Dv^\delta) = 0 \quad \text{in } \mathbb{T}^n \tag{1.14}$$

for $\delta > 0$.

Next, for $\alpha > 0$, consider the equation

$$\delta v^{\alpha,\delta} - \mathrm{tr}\left(A(x) D^2 v^{\alpha,\delta}\right) + H(x, Dv^{\alpha,\delta}) = \alpha \Delta v^{\alpha,\delta} \quad \text{in } \mathbb{T}^n. \tag{1.15}$$

Owing to the discount and viscosity terms, there exists a (unique) classical solution $v^{\alpha,\delta}$. By the comparison principle, it is again clear that $|\delta v^{\alpha,\delta}| \leq M$ for $M = \max_{x\in\mathbb{T}^n} |H(x, 0)|$.

We use the Bernstein method again. As in the proof of Theorem 1.2, differentiate (1.15) with respect to $x_i$, multiplying it by $v^{\alpha,\delta}_{x_i}$ and summing up with respect to $i$ to obtain

$$\begin{aligned} 2\delta\varphi - a^{ij}_{x_k} v^{\alpha,\delta}_{x_i x_j} v^{\alpha,\delta}_{x_k} - a^{ij}(\varphi_{x_i x_j} - v^{\alpha,\delta}_{x_i x_k} v^{\alpha,\delta}_{x_j x_k}) \\ + D_x H \cdot Dv^{\alpha,\delta} + D_p H \cdot D\varphi = \alpha(\Delta\varphi - |D^2 v^{\alpha,\delta}|^2), \end{aligned}$$

where $\varphi := |Dv^{\alpha,\delta}|^2/2$. Here we use Einstein's convention.

Take a point $x_0$ such that $\varphi(x_0) = \max_{\mathbb{T}^n} \varphi \geq 0$ and note that at that point

$$-a^{ij}_{x_k} v^{\alpha,\delta}_{x_i x_j} v^{\alpha,\delta}_{x_k} + D_x H \cdot Dv^{\alpha,\delta} + a^{ij} v^{\alpha,\delta}_{x_i x_k} v^{\alpha,\delta}_{x_j x_k} + \alpha |D^2 v^{\alpha,\delta}|^2 \leq 0. \tag{1.16}$$

The two terms $a^{ij} v^{\alpha,\delta}_{x_i x_k} v^{\alpha,\delta}_{x_j x_k}$ and $\alpha|D^2 v^{\alpha,\delta}|^2$ are the good terms, which will help us to control other terms and to deduce the result.

We first use the trace inequality (see [76, Lemma 3.2.3] for instance),

$$(\mathrm{tr}\,(A_{x_k} S))^2 \leq C \mathrm{tr}\,(SAS) \quad \text{for all } S \in \mathbb{M}^{n\times n}_{\text{sym}},\ 1 \leq k \leq n,$$

for some constant $C$ depending only on $n$ and $\|D^2 A\|_{L^\infty(\mathbb{T}^n)}$ to yield that, for some small constant $d > 0$,

$$\begin{aligned} a^{ij}_{x_k} v^{\alpha,\delta}_{x_i x_j} v^{\alpha,\delta}_{x_k} &= \mathrm{tr}\,(A_{x_k} D^2 v^{\alpha,\delta}) v^{\alpha,\delta}_{x_k} \leq d\big(\mathrm{tr}\,(A_{x_k} D^2 v^{\alpha,\delta})\big)^2 + \frac{C}{d}|Dv^{\alpha,\delta}|^2 \\ &\leq \frac{1}{2}\mathrm{tr}\,(D^2 v^{\alpha,\delta} A D^2 v^{\alpha,\delta}) + C|Dv^{\alpha,\delta}|^2 = \frac{1}{2} a^{ij} v^{\alpha,\delta}_{x_i x_k} v^{\alpha,\delta}_{x_j x_k} + C|Dv^{\alpha,\delta}|^2. \end{aligned} \tag{1.17}$$

Next, by using a modified Cauchy-Schwarz inequality for matrices (see Remark 1.3

$$(\mathrm{tr}\,AB)^2 \leq \mathrm{tr}\,(ABB)\mathrm{tr}\,A \quad \text{for all } A, B \in \mathbb{M}^{n\times n}_{\text{sym}},\ \text{with } A \geq 0 \tag{1.18}$$

we obtain

$$\begin{aligned} \left(a^{ij} v^{\alpha,\delta}_{x_i x_j}\right)^2 &= (\mathrm{tr}\,A(x) D^2 v^{\alpha,\delta})^2 \leq \mathrm{tr}\,(A(x) D^2 v^{\alpha,\delta} D^2 v^{\alpha,\delta}) \mathrm{tr}\,A(x) \\ &\leq C \mathrm{tr}\,\left(A(x) D^2 v^{\alpha,\delta} D^2 v^{\alpha,\delta}\right) = C a^{ik} v^{\alpha,\delta}_{x_i x_j} v^{\alpha,\delta}_{x_k x_j}. \end{aligned} \tag{1.19}$$

In light of (1.19), for some $c_0 > 0$ sufficiently small,

$$\begin{aligned} \frac{1}{2} a^{ij} v^{\alpha,\delta}_{x_i x_k} v^{\alpha,\delta}_{x_j x_k} + \alpha|D^2 v^{\alpha,\delta}|^2 &\geq 4c_0 \left( \left(a^{ij} v^{\alpha,\delta}_{x_i x_j}\right)^2 + (\alpha \Delta v^{\alpha,\delta})^2 \right) \\ &\geq 2c_0 \left(a^{ij} v^{\alpha,\delta}_{x_i x_j} + \alpha \Delta v^{\alpha,\delta}\right)^2 = 2c_0 \left(\delta v^{\alpha,\delta} + H(x, Dv^{\alpha,\delta})\right)^2 \\ &\geq c_0 H(x, Dv^{\alpha,\delta})^2 - C. \end{aligned} \tag{1.20}$$

Combining (1.16), (1.17) and (1.20) to achieve that

$$D_x H \cdot Dv^{\alpha,\delta} - C|Dv^{\alpha,\delta}|^2 + c_0 H(x, Dv^{\alpha,\delta})^2 \leq C.$$

We then use (1.13) in the above to get the existence of a constant $C > 0$ independent of $\alpha, \delta$ so that $|Dv^{\alpha,\delta}(x_0)| \leq C$. Therefore, as in the proof of Theorem 1.1, setting $w^{\alpha,\delta}(x) := v^{\alpha,\delta}(x) - v^{\alpha,\delta}(0)$, by passing some subsequences if necessary, we can send $\alpha \to 0$ and $\delta \to 0$ in this order to yield that $w^{\alpha,\delta}$ and $\delta v^{\alpha,\delta}$, respectively, converge uniformly to $v$ and $-c$ in $\mathbb{T}^n$. Clearly, $(v, c)$ satisfies (1.8) in the viscosity sense. $\square$

**Remark 1.3.** We give a simple proof of (1.18) here. By the Cauchy-Schwarz inequality, we always have

$$0 \le (\mathrm{tr}\,(ab))^2 \le \mathrm{tr}\,(a^2)\mathrm{tr}\,(b^2) \quad \text{for all } a, b \in \mathbb{M}^{n\times n}_{\mathrm{sym}}.$$

For $A, B \in \mathbb{M}^{n\times n}_{\mathrm{sym}}$ with $A \ge 0$, set $a := A^{1/2}$ and $b := A^{1/2}B$. Then,

$$(\mathrm{tr}\,(AB))^2 \le \mathrm{tr}\,(A)\mathrm{tr}\,(A^{1/2}BA^{1/2}B) = \mathrm{tr}\,(A)\mathrm{tr}\,(ABB).$$

**Definition 1.1.** *For a pair* $(v, c) \in C(\mathbb{T}^n) \times \mathbb{R}$ *solving one of the ergodic problems* (1.6)–(1.8)*, we call* $v$ *and* $c$ *an ergodic function and an ergodic constant, respectively.*

We now proceed to show that an ergodic constant $c$ is unique in all ergodic problems (1.6)–(1.8). It is enough to consider general case (1.8).

**Proposition 1.4.** *Assume that* $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$ *and that* (H1), (1.13) *hold. Then ergodic problem* (1.8) *admits the unique ergodic constant* $c \in \mathbb{R}$*, which is uniquely determined by* $A$ *and* $H$*.*

*Proof.* Suppose that there exist two pairs of solutions $(v_1, c_1)$, $(v_2, c_2) \in \mathrm{Lip}\,(\mathbb{T}^n) \times \mathbb{R}$ to (1.8) with $c_1 \ne c_2$. We may assume that $c_1 < c_2$ without loss of generality. Note that $v_1 - c_1 t - M$ and $v_2 - c_2 t + M$ are a subsolution and a supersolution to (1.3), respectively, for a suitably large $M > 0$. By the comparison principle for (1.3), we get

$$v_1 - c_1 t - M \le v_2 - c_2 t + M \quad \text{in } \mathbb{T}^n \times [0, \infty).$$

Thus, $(c_2 - c_1)t \le M'$ for some $M' > 0$ and all $t \in (0, \infty)$, which yields a contradiction. □

**Remark 1.4.** As shown in Proposition 1.4, an ergodic constant is unique but on the other hand, ergodic functions are not unique in general. It is clear that, if $v$ is an ergodic function, then $v + C$ is also an ergodic function for any $C \in \mathbb{R}$. But even up to additive constants, $v$ is not unique. See Section 3.1.1.

The ergodic constant $c$ is related to the *effective Hamiltonian* $\overline{H}$ in the homogenization theory (see Lions, Papanicolaou, Varadhan [61]). In fact, $c = \overline{H}(0)$. In general, for $p \in \mathbb{R}^n$, $\overline{H}(p)$ is the ergodic constant to

$$-\mathrm{tr}\,\left(A(x)D^2 v\right) + H(x, p + Dv) = \overline{H}(p) \quad \text{in } \mathbb{T}^n.$$

It is known that there are (abstract) formulas for the ergodic constant.

**Proposition 1.5.** *Assume that* $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$ *and that* (H1), (1.13) *hold. The ergodic constant of* (1.8) *can be represented by*

$$c = \inf\left\{a \in \mathbb{R} \,:\, \textit{there exists a solution to } -\mathrm{tr}\,\left(A(x)D^2 w\right) + H(x, Dw) \le a \textit{ in } \mathbb{T}^n\right\}.$$

*Moreover, if* $A \equiv 0$*, and* $p \mapsto H(x, p)$ *is convex for all* $x \in \mathbb{T}^n$*, then*

$$c = \inf_{\phi \in C^1(\mathbb{T}^n)} \sup_{x \in \mathbb{T}^n} H(x, D\phi(x)). \tag{1.21}$$

*Proof.* Let us denote by $d_1, d_2$ the first and the second formulas in statement of the proposition, respectively.

It is clear that $d_1 \le c$. Suppose that $c > d_1$. Then there exists a subsolution $(v_a, a)$ with $a < c$ to (1.8). By using the same argument as that of the proof of Proposition 1.4, we get $(c-a)t \le M'$ for some $M' > 0$ and all $t > 0$, which implies the contradiction. Therefore, $c = d_1$.

We proceed to prove the second part of the proposition. For any fixed $\phi \in C^1(\mathbb{T}^n)$,

$$H(x, D\phi) \le \sup_{x\in\mathbb{T}^n} H(x, D\phi(x)) \quad \text{in } \mathbb{T}^n \text{ in the classical sense,}$$

which implies that $d_1 \le \sup_{x\in\mathbb{T}^n} H(x, D\phi(x))$. Take infimum over $\phi \in C^1(\mathbb{T}^n)$ to yield that $d_1 \le d_2$.

Now let $v \in \mathrm{Lip}\,(\mathbb{T}^n)$ be a subsolution of (1.6). Take $\rho$ to be a standard mollifier and set $\rho_\varepsilon(\cdot) = \varepsilon^{-n}\rho(\cdot/\varepsilon)$ for any $\varepsilon > 0$. Denote by $v^\varepsilon = \rho_\varepsilon * v$. Then $v^\varepsilon \in C^\infty(\mathbb{T}^n)$. In light of the convexity of $H$ and Jensen's inequality,

$$\begin{aligned}
H(x, Dv^\varepsilon(x)) &= H\left(x, \int_{\mathbb{T}^n} \rho_\varepsilon(y) Dv(x-y)\,dy\right)\\
&\le \int_{\mathbb{T}^n} H(x, Dv(x-y))\rho_\varepsilon(y)\,dy\\
&\le \int_{\mathbb{T}^n} H(x-y, Dv(x-y))\rho_\varepsilon(y)\,dy + C\varepsilon \le c + C\varepsilon \quad \text{in } \mathbb{T}^n.
\end{aligned}$$

Thus, $d_2 \le \sup_{x\in\mathbb{T}^n} H(x, Dv^\varepsilon) \le c + C\varepsilon$. Letting $\varepsilon \to 0$ to get $d_2 \le c = d_1$. The proof is complete. □

**Remark 1.5.** Concerning formula (1.21), it is important pointing out that the approximation using mollification to a given subsolution $v$ plays an essential role. This only works for first-order convex Hamilton–Jacobi equations as seen in the proof of Proposition 1.5. If we consider first-order nonconvex Hamilton–Jacobi equations, then a smooth way to approximate a subsolution is not known. In light of the first formula of Proposition 1.5, we only have in case $A \equiv 0$ that

$$c = \inf_{\phi\in\mathrm{Lip}\,(\mathbb{T}^n)} \sup_{x\in\mathbb{T}^n} \sup_{p\in D^+\phi(x)} H(x,p),$$

where we denote by $D^+\phi(x)$ the *superdifferential* of $\phi$ at $x$ (see Section 4.2 for the definition). An analog to (1.21) in the general degenerate viscous case is not known yet even in the convex setting. See Section 3.4 for some further discussions.

In the end of this chapter, we give the results on the boundedness of solutions to (1.14), (1.3), and the asymptotic speed of solutions to (1.3). These are straightforward consequences of Theorem 1.3.

**Proposition 1.6.** *Assume that $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$ and that* (H1), (1.13) *hold. Let $v^\delta$ be the viscosity solution of* (1.14) *and $c$ be the associated ergodic constant. Then, there exists $C > 0$ independent of $\delta$ such that*

$$\left| v^\delta + \frac{c}{\delta} \right| \le C \quad \text{in } \mathbb{T}^n.$$

*Proof.* Let $(v,c)$ be a solution of (1.8). Take a suitably large constant $M>0$ so that $v-c/\delta\pm M$ are a subsolution and a supersolution of (1.14), respectively. In light of the comparison principle for (1.14), we get

$$v(x)-\frac{c}{\delta}-M\le v^\delta(x)\le v(x)-\frac{c}{\delta}+M\quad\text{for all } x\in\mathbb{T}^n,$$

which yields the conclusion. □

**Proposition 1.7.** *Assume that $H\in C^2(\mathbb{T}^n\times\mathbb{R}^n)$ and that* (H1), (1.13) *hold. Let $u$ be the viscosity solution of* (1.3) *with the given initial data $u_0\in\mathrm{Lip}\,(\mathbb{T}^n)$, and $c$ be the associated ergodic constant. Then,*

$$\begin{cases} u+ct \quad \textit{is bounded, and}\\ \dfrac{u(x,t)}{t}\to -c \quad \textit{uniformly on } \mathbb{T}^n \textit{ as } t\to\infty.\end{cases}$$

*Proof.* Let $(v,c)$ be a solution of (1.8). Take a suitably large constant $M>0$ so that $v-ct-M$, $v-ct+M$ are a subsolution and a supersolution of (1.3), respectively. In light of the comparison principle for (1.3), we get

$$v(x)-ct-M\le u(x,t)\le v-ct+M\quad\text{for all } (x,t)\in\mathbb{T}^n\times[0,\infty), \tag{1.22}$$

which implies the conclusion. □

**Remark 1.6.** A priori estimate (1.22) is the reason why we do not need to consider the terms like $b_i(x)t^i$ for $i\ge 2$ in the formal asymptotic expansion of $u$ in the introduction of this chapter.

We also give here a result on the Lipschitz continuity of solutions to (1.3).

**Proposition 1.8.** *Assume that $H\in C^2(\mathbb{T}^n\times\mathbb{R}^n)$ and that* (H1), (1.13) *hold. Assume further that $u_0\in C^2(\mathbb{T}^n)$. Then the solution $u$ to* (1.3) *is globally Lipschitz continuous on $\mathbb{T}^n\times[0,\infty)$, i.e.,*

$$\|u_t\|_{L^\infty(\mathbb{T}^n\times(0,\infty))}+\|Du\|_{L^\infty(\mathbb{T}^n\times(0,\infty))}\le M\quad\textit{for some } M>0.$$

*Proof.* For a suitably large $M>0$, $u_0-Mt$ and $u_0+Mt$ are a subsolution and a supersolution of (1.3), respectively. By the comparison principle, we get $u_0(x)-Mt\le u(x,t)\le u_0(x)+Mt$ for any $(x,t)\in\mathbb{T}^n\times[0,\infty)$. We use the comparison principle again to get

$$|u(x,t+s)-u(x,t)|\le\max_{x\in\mathbb{T}^n}|u(x,s)-u_0(x)|\le Ms\quad\text{for all } x\in\mathbb{T}^n, t,s\ge 0.$$

Therefore, $|u_t|\le M$. By using the same method as that of the proof of Theorem 1.3, we get $|Du|\le M'$ for some $M'>0$. □

As a corollary of Propositions 1.7, 1.8, we can easily get that there exists a subsequence $\{t_j\}_{j\in\mathbb{N}}$ with $t_j \to \infty$ as $j \to \infty$ such that

$$u(x, t_j) + ct_j \to v(x) \quad \text{uniformly for } x \in \mathbb{T}^n \text{ as } j \to \infty,$$

where $v$ is a solution of (1.8). We call $v$ an *accumulation point*. However, we have to be careful about the fact that $v$ in the above may depend on the choice of a subsequence at this moment. The question whether this accumulation point is unique or not for all of choices of subsequences is nontrivial, and will be seriously studied in the next chapter.

# Chapter 2

# Large time asymptotics of Hamilton–Jacobi equations

## 2.1 A brief introduction

In the last decade, a number of authors have studied extensively the large time behavior of solutions of first-order Hamilton–Jacobi equations. Several convergence results have been established. The first general theorem in this direction was proven by Namah and Roquejoffre in [72], under the assumptions:

$$p \mapsto H(x,p) \text{ is convex}, H(x,p) \geq H(x,0) \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n, \ \max_{x \in \mathbb{T}^n} H(x,0) = 0.$$

We will first discuss this setting in Section 2.2. In this setting, as the Hamiltonian has a simple structure, we are able to find an explicit subset of $\mathbb{T}^n$ which has the monotonicity of solutions and the property of the uniqueness set. Therefore, we can relatively easily get a convergence result of the type (1.5), that is,

$$u(x,t) - (v(x) - ct) \to 0 \quad \text{uniformly for } x \in \mathbb{T}^n,$$

where $u$ is the solution of the initial value problem and $(v,c)$ is a solution to the associated ergodic problem.

Fathi then gave a breakthrough in [34] by using a dynamical approach from the weak KAM theory. Contrary to [72], the results of [34] use uniform convexity and smoothness assumptions on the Hamiltonian but do not require any structural conditions as above. These rely on a deep understanding of the dynamical structure of the solutions and of the corresponding ergodic problem. See also the paper of Fathi and Siconolfi [35] for a characterization of the Aubry set, which will be studied in Section 2.5. Afterwards, Davini and Siconolfi in [25] and Ishii in [49] refined and generalized the approach of Fathi, and studied the asymptotic problem for Hamilton–Jacobi equations on $\mathbb{T}^n$ and on the whole $n$-dimensional Euclidean space, respectively.

Besides, Barles and Souganidis [12] obtained additional results, for possibly non-convex Hamiltonians, by using a PDE method in the context of viscosity solutions.

Barles, Ishii and Mitake [8] simplified the ideas in [12] and presented the most general assumptions (up to now).

In general, these methods are based crucially on delicate stability results of extremal curves in the context of the dynamical approach in light of the finite speed of propagation, and of solutions for large time in the context of the PDE approach.

In the uniformly parabolic setting, Barles and Souganidis [13] proved the large-time convergence of solutions. Their proof relies on a completely distinct set of ideas from the ones used in the first-order case as the associated ergodic problem has a simpler structure. Indeed, since the strong maximum principle holds, the ergodic problem has a unique solution up to additive constants. The proof for the large-time convergence in [13] strongly depends on this fact. We will discuss this in Section 2.6.

It is clear that all the methods aforementioned are not applicable for the general degenerate viscous cases, which will be described in details in Section 2.4, because of the presence of the second order terms and the lack of both the finite speed of propagation as well as the strong comparison principle. Under these backgrounds, the authors with Cagnetti, Gomes [17] introduced a new method for the large-time behavior for general viscous Hamilton–Jacobi equations (1.3). In this method, the nonlinear adjoint method, which was introduced by Evans in [32], plays the essential role. In Section 2.3, we introduce this nonlinear adjoint method.

## 2.2 First-order case with separable Hamiltonians

As mentioned in the end of Section 1.2, in general, (1.6) does not have unique solutions even up to additive constants. See Section 3.1.1 for details. This fact can be observed from Example 4.1 too. This requires a more delicate and serious argument to prove the large-time convergence (1.5) for (1.1).

Before handling the general case, we first consider the case where the Hamiltonian is separable with respect to $x$ and $p$. We consider two representative examples here.

### 2.2.1 First example

Consider

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

where $V \in C(\mathbb{T}^n)$ is a given function. See Example 4.2 in Appendix. In this case, since the structure of the Hamiltonian is simple, we can easily prove (1.5), that is,

$$u(x,t) - (v(x) - ct) \to 0 \quad \text{uniformly for } x \in \mathbb{T}^n,$$

where $(v,c)$ satisfies

$$\frac{1}{2}|Dv|^2 + V = c \quad \text{in } \mathbb{T}^n. \tag{2.1}$$

This was first done by Namah and Roquejoffre in [72]. Firstly, let us find the ergodic constant $c$ in this case. By Proposition 1.5, we have

$$c = \inf_{\phi \in C^1(\mathbb{T}^n)} \sup_{x \in \mathbb{T}^n} \left( \frac{|D\phi(x)|^2}{2} + V(x) \right) \geq \sup_{x \in \mathbb{T}^n} V(x) = \max_{x \in \mathbb{T}^n} V(x).$$

On the other hand, if we take $\phi$ to be a constant function, then

$$c \leq \sup_{x \in \mathbb{T}^n} \left( \frac{|0|^2}{2} + V(x) \right) = \max_{x \in \mathbb{T}^n} V(x).$$

Thus, we get $c = \max_{x \in \mathbb{T}} V(x)$.

Set $u_c(x,t) := u(x,t) + ct$. Then

$$(u_c)_t + \frac{1}{2}|Du_c|^2 = \max_{x \in \mathbb{T}^n} V(x) - V(x).$$

Set

$$\mathcal{A} := \{x \in \mathbb{T}^n \,:\, V(x) = \max_{x \in \mathbb{T}^n} V(x)\}.$$

Then we observe, at least formally, that for $x \in \mathcal{A}$

$$(u_c)_t = -\frac{1}{2}|Du_c|^2 \leq 0,$$

which implies the monotonicity of $t \mapsto u_c(x,t)$ for $x \in \mathcal{A}$. Thus, we get

$$\liminf_{t \to \infty}{}_* u_c(x,t) = \limsup_{t \to \infty}{}^* u_c(x,t) \quad \text{for all } x \in \mathcal{A},$$

where, for $f \in C(\mathbb{T}^n \times [0,\infty))$, we set

$$\begin{aligned}
\limsup_{t \to \infty}{}^* f(x,t) &:= \lim_{t \to \infty} \sup \left\{ f(y,s) \,:\, |x-y| \leq 1/s,\ s \geq t \right\}, \\
\liminf_{t \to \infty}{}_* f(x,t) &:= \lim_{t \to \infty} \inf \left\{ f(y,s) \,:\, |x-y| \leq 1/s,\ s \geq t \right\}.
\end{aligned}$$

We call these limits $\liminf_{t \to \infty}{}_*$ and $\limsup_{t \to \infty}{}^*$ *the half-relaxed limits.* In view of the stability result of viscosity solutions (see Theorem 4.11 in Section 4.5), $\limsup_{t \to \infty}{}^* u_c$ and $\liminf_{t \to \infty}{}_* u_c$ are a subsolution and a supersolution of (2.1), respectively. Also notice that the set $\mathcal{A}$ is a uniqueness set of ergodic problem (2.1) (see Theorem 2.15 in Section 2.5), that is, for $v_1, v_2$ are a subsolution and a supersolution of (2.1), respectively,

$$\text{if } v_1 \leq v_2 \text{ on } \mathcal{A}, \text{ then } v_1 \leq v_2 \text{ on } \mathbb{T}^n.$$

This is an extremely important fact, and we will come back to it later. In light of this fact, we get $\limsup_{t \to \infty}{}^* u_c \leq \liminf_{t \to \infty}{}_* u_c$ on $\mathbb{T}^n$, and thus,

$$\liminf_{t \to \infty}{}_* u_c = \limsup_{t \to \infty}{}^* u_c \quad \text{on } \mathbb{T}^n,$$

which confirms (1.5) (see Proposition 4.13 in Section 4.5).

We encourage the interested readers to find a simple and direct PDE proof for the fact that $\mathcal{A}$ is the uniqueness set of (1.6) (see [72] and also [68] for more details).

### 2.2.2 Second example

Next, let us consider

$$\begin{cases} u_t + h(x)\sqrt{1+|Du|^2} = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ u(\cdot,0) = u_0 & \text{on } \mathbb{T}^n, \end{cases} \tag{2.2}$$

where $h \in C(\mathbb{T}^n)$ with $h(x) > 0$ for all $x \in \mathbb{T}^n$ is a given function. See Section 4.1.1 in Appendix. We obtain the ergodic constant $c$ first as follows:

$$h(x)\sqrt{1+|Dv|^2} = c \iff |Dv|^2 = \frac{c^2 - h(x)^2}{h(x)^2}. \tag{2.3}$$

Thus, we get

$$c = \sqrt{\max_{x\in\mathbb{T}^n} h(x)^2} = \max_{x\in\mathbb{T}^n} h(x). \tag{2.4}$$

Set $u_c := u + ct$ as above to get that

$$(u_c)_t + h(x)\sqrt{1+|Du_c|^2} = c.$$

In this case, we have

$$(u_c)_t = c - h(x)\sqrt{1+|Du_c|^2} \le 0 \quad \text{in } \mathcal{A} := \{x \in \mathbb{T}^n : h(x) = c\}. \tag{2.5}$$

Therefore, we get a similar type of monotonicity of $u_c$ in $\mathcal{A}$ as in the above example. Moreover, setting $V(x) := (c^2 - h(x)^2)/h(x)^2$, we see that

$$\mathcal{A} = \{x \in \mathbb{T}^n : V(x) = \min_{x\in\mathbb{T}^n} V(x) = 0\}.$$

Thus, we can see that $\mathcal{A}$ is a uniqueness set for (2.3) as in Section 2.2.1. The large time behavior result follows in a similar manner.

Moreover, we have the following proposition.

**Proposition 2.1.** *If the initial data $u_0$ is a subsolution of* (2.3)*, then $u(x,t)+ct = u_0(x)$ for all $(x,t) \in \mathcal{A} \times [0,\infty)$. In particular,*

$$\lim_{t\to\infty} (u(x,t) + ct) = u_0(x) \quad \textit{for all } x \in \mathcal{A}.$$

*Proof.* Since $u_0$ is a subsolution to (2.3), $u_0 - ct$ is also a subsolution to (2.2). Thus, by the comparison principle, we have $u_0(x) - ct \le u(x,t)$ for all $(x,t) \in \mathbb{T}^n \times [0,\infty)$. Combining this with (2.5), we obtain the conclusion. □

**Example 2.1.** Let us consider a more explicit example. Assume that $n = 1$, $h : \mathbb{R} \to \mathbb{R}$ is 1-periodic and

$$h(x) := \frac{2}{\sqrt{1+f(x)^2}}, \tag{2.6}$$

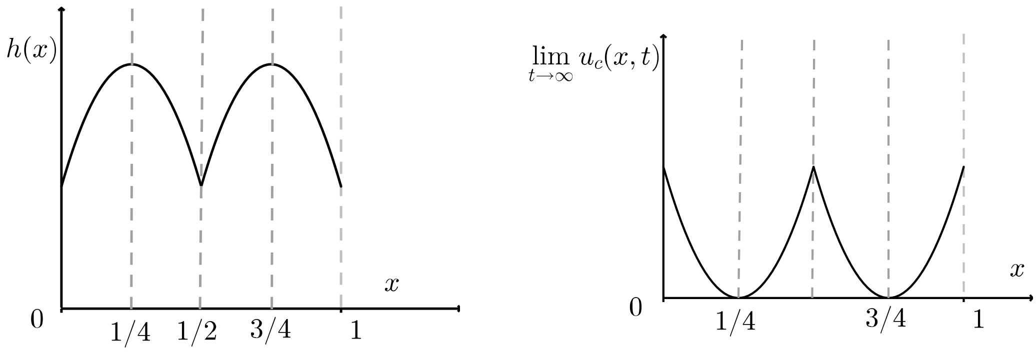


Figure 2.1: Graphs of $h$ and $\lim_{t\to\infty} u_c(x,t)$

where $f(x) := 2\min\{|x-1/4|, |x-3/4|\}$ for all $x \in [0,1]$. See Figure 2.1 for the graph of $h$. Consider $u_0 \equiv 0$ on $\mathbb{T}$. Since $u_0$ is a subsolution to (2.3), in light of Proposition 2.1, we obtain

$$\lim_{t\to\infty} (u(x,t) + ct) = u_0(x) = 0 \quad \text{for } x \in \mathcal{A} = \left\{\frac{1}{4}, \frac{3}{4}\right\},$$

which is enough to characterize the limit. See Figure 2.1. We will give further discussions on this example in Section 2.5.

Can we expect such a monotonicity in the general setting? The answer is NO. For instance, if we consider the Hamilton–Jacobi equation:

$$u_t + \frac{1}{2}|Du - \mathbf{b}(x)|^2 = |\mathbf{b}(x)|^2 \quad \text{in } \mathbb{T}^n \times (0,\infty),$$

where $\mathbf{b} : \mathbb{T}^n \to \mathbb{R}^n$ is a given smooth vector field, then we cannot find such an easy structure of solutions. Therefore, we need more profound arguments to prove (1.5) in the general case.

## 2.3 First-order case with general Hamiltonians

In this section, we assume the following conditions:

(H2) $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$,

(H3) $D^2_{pp}H(x,p) \geq 2\theta I_n$ for all $(x,p) \in \mathbb{T}^n \times \mathbb{R}^n$, and some $\theta > 0$, where $I_n$ is the identity matrix of size $n$,

(H4) $|D_xH(x,p)| \leq C(1+|p|^2)$ for all $x \in \mathbb{T}^n$ and $p \in \mathbb{R}^n$.

We see that if $H$ satisfies (H3), (H4), then it satisfies (1.13) hence also (1.11). Therefore, all of the results concerning ergodic problems in the previous chapter are valid here.

Our main goal in this section is to prove

**Theorem 2.2.** *Assume that* (H2)–(H4) *hold. Let $u$ be the solution of* (1.1) *with a given initial data $u_0 \in \mathrm{Lip}(\mathbb{T}^n)$. Then there exists $(v, c) \in \mathrm{Lip}(\mathbb{T}^n) \times \mathbb{R}$, a solution of ergodic problem* (1.6)*, such that* (1.5) *holds, that is,*

$$u(x,t) - (v(x) - ct) \to 0 \quad \textit{uniformly for } x \in \mathbb{T}^n.$$

We call $v - ct$ obtained in Theorem 2.2 the *asymptotic solution* for (1.1).

**Remark 2.1.** It is worth pointing out delicate things on the convexity assumption here. Assumption (H3) is a uniform convexity assumption. We can actually easily weaken this to a strictly convexity assumption, i.e., $D^2_{pp}H > 0$, since we do have an a priori estimate on the Lipschitz continuity of solutions. Therefore, we can construct a uniformly convex Hamilton–Jacobi equation which has the same solution as that of a strictly convex one.

On the other hand, this "strictness" of convexity is very important to get convergence (1.5). Consider the following explicit example:

$$u_t + |u_x - 1| = 1 \text{ in } \mathbb{R} \times (0, \infty), \quad u(\cdot, 0) = \sin(x).$$

Then, it is clear that

$$u(x,t) = \sin(x+t) \quad \text{for all } (x,t) \in \mathbb{R} \times [0, \infty)$$

is the solution of the above but convergence (1.5) does not hold. This was first pointed out by Barles and Souganidis in [12].

We also point out that the convexity is NOT a necessary condition either, since it is known that there are convergence results for possibly nonconvex Hamilton–Jacobi equations in [12, 8]. A typical example for nonconvex Hamiltonians is $H(x,p) := (|p|^2 - 1)^2 - V(x)$.

### 2.3.1 Formal calculation

In this subsection, we describe the idea in [17] in a heuristic way to get

$$u_t(\cdot, t) \to -c \quad \text{as } t \to \infty \text{ in the viscosity sense,} \tag{2.7}$$

where $c$ is the ergodic constant of (1.6). We call this an *asymptotic monotone property* of the solution to (1.1). This is a much stronger result than that of Proposition 1.7. We "assume" that $u$ is smooth below in the derivation. Notice that this is a completely formal assumption as we cannot expect a global smooth solution $u$ of Hamilton–Jacobi equations in general.

Let us first fix $T > 0$. We consider the adjoint equation of the linearized operator of the Hamilton–Jacobi equation:

$$\begin{cases} -\sigma_t - \operatorname{div}\big(D_pH(x, Du(x,t))\sigma\big) = 0 & \text{in } \mathbb{T}^n \times (0,T) \\ \sigma(x,T) = \delta_{x_0}(x) & \text{on } \mathbb{T}^n, \end{cases} \tag{2.8}$$

where $\delta_{x_0}$ is the Dirac delta measure at a fixed point $x_0 \in \mathbb{T}^n$. Note that although (2.8) may have only a very singular solution, we do not mind in this section as this is just a formal argument. It is clear that

$$\sigma(x,t) \geq 0 \quad \text{and} \quad \int_{\mathbb{T}^n} \sigma(x,t)\, dx = 1 \quad \text{for all } (x,t) \in \mathbb{T}^n \times [0,T]. \tag{2.9}$$

Then, we have the following conservation of energy:

$$\begin{aligned}
&\frac{d}{dt} \int_{\mathbb{T}^n} H(x, Du(x,t))\sigma(x,t)\, dx \\
&= \int_{\mathbb{T}^n} D_pH(x,Du) \cdot Du_t \sigma\, dx + \int_{\mathbb{T}^n} H(x,Du)\sigma_t\, dx \\
&= -\int_{\mathbb{T}^n} \operatorname{div}\big(D_pH(x,Du)\sigma\big) u_t\, dx - \int_{\mathbb{T}^n} u_t\sigma_t\, dx = 0,
\end{aligned}$$

which implies a new "formula":

$$\begin{aligned}
u_t(x_0,T) &= \int_{\mathbb{T}^n} u_t(\cdot,T)\delta_{x_0}\, dx = -\int_{\mathbb{T}^n} H(x,Du)\sigma\, dx\Big|_{t=T} \\
&= -\frac{1}{T}\int_0^T \int_{\mathbb{T}^n} H(x,Du)\sigma\, dxdt.
\end{aligned}$$

Noting (2.9) and $|Du(x,t)| \leq C$ by Proposition 1.8, in light of the Riesz theorem, there exists $\nu_T \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ such that

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x,p)\, d\nu_T(x,p) = \frac{1}{T}\int_0^T \int_{\mathbb{T}^n} \varphi(x,Du)\sigma\, dxdt$$

for all $\varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$. Here $\mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ is the set of all Radon probability measures on $\mathbb{T}^n \times \mathbb{R}^n$. Because of the gradient bound of $u$, we obtain that $\operatorname{supp}(\nu_T) \subset \mathbb{T}^n \times \overline{B}(0,C)$, where $\operatorname{supp}(\nu_T)$ denotes the support of $\nu_T$, that is,

$$\operatorname{supp}(\nu_T) = \big\{(x,p) \in \mathbb{T}^n \times \mathbb{R}^n \,:\, \nu_T\big(B((x,p),r)\big) > 0 \text{ for all } r > 0\big\}.$$

Since

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} d\nu_T(x,p) = 1,$$

there exists a subsequence $T_j \to \infty$ as $j \to \infty$ so that

$$\nu_{T_j} \rightharpoonup \nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \quad \text{as} \quad j \to \infty \tag{2.10}$$

in the sense of measures. Then, we can expect some important facts

(i) $\nu$ is a Mather measure associated with (1.6),

(ii) $\operatorname{supp}\nu \subset \{(x,p) \in \mathbb{T}^n \times \mathbb{R}^n \,:\, p = Dv(x)\}$, where $v$ is a viscosity solution to (1.6).

We do not give the proofs of these facts here. We will give the definition of Mather measures in Chapter 3. Property (ii) in the above is called the *graph theorem* in the Hamiltonian dynamics, which is an extremely important result (see [63, 62] for details). One way to look at (i) is the following: if we think of $Du$ as a given function in (2.8), then (2.8) is a transport equation, and the characteristic ODE is given by

$$\begin{cases} \dot{X}(t) = D_pH(X(t), Du(X(t), t)) & \text{for } t \in (0, T) \\ X(T) = x_0, \end{cases} \tag{2.11}$$

which is formally equivalent to the Hamiltonian system.

If we admit these, then we obtain

$$\begin{aligned} u_t(x_0, T_j) &= -\frac{1}{T_j}\int_0^{T_j}\int_{\mathbb{T}^n} H(x, Du)\sigma \, dxdt \\ &= -\iint_{\mathbb{T}^n\times\mathbb{R}^n} H(x,p)\, d\nu_{T_j}(x,p) \to -\iint_{\mathbb{T}^n\times\mathbb{R}^n} H(x,p)\, d\nu(x,p) = -c \end{aligned}$$

as $j \to \infty$ for any subsequence $T_j$ satisfying (2.10).

Now, we should ask ourselves how we can make this argument rigorous. Some important points are

(i) to introduce a regularizing process for (1.1),

(ii) to introduce a scaling process for $t$ as we need to look at both limits of a regularizing process and the large-time behavior, and

(iii) to give good estimates,

which are discussed in details in the next subsections.

### 2.3.2 Regularizing process

In the following subsections, we make the formal argument in Section 2.3.1 rigorous by using a regularizing process and giving important estimates.

We only need to study the case where the ergodic constant $c = 0$, and we always assume it henceforth. Indeed, by replacing, if necessary, $H$ and $u(x,t)$ by $H - c$ and $u(x,t) + ct$, respectively, we can always reduce the situation to the case that $c = 0$.

We first consider a rescaled problem. Let $u$ be the solution of (1.1). For $\varepsilon > 0$, set $u^\varepsilon(x,t) = u(x, t/\varepsilon)$ for $(x,t) \in \mathbb{T}^n \times [0,\infty)$. Then, $u^\varepsilon$ satisfies

$$(\mathrm{C})_\varepsilon \qquad \begin{cases} \varepsilon u^\varepsilon_t + H(x, Du^\varepsilon) = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{T}^n. \end{cases}$$

By repeating the proof of Proposition 1.8 with a small modification, we have the following a priori estimates

$$\|u^\varepsilon_t\|_{L^\infty(\mathbb{T}^n\times[0,1])} \le C/\varepsilon, \qquad \|Du^\varepsilon\|_{L^\infty(\mathbb{T}^n\times[0,1])} \le C. \tag{2.12}$$

for some constant $C > 0$ independent of $\varepsilon$. Notice that in general, the function $u^\varepsilon$ is only Lipschitz continuous.

For this reason, we add a viscosity term to $(\mathrm{C})_\varepsilon$, and consider the following regularized equation

$$(\mathrm{A})_\varepsilon \qquad \begin{cases} \varepsilon w^\varepsilon_t + H(x, Dw^\varepsilon) = \varepsilon^4 \Delta w^\varepsilon & \text{in } \mathbb{T}^n \times (0,\infty), \\ w^\varepsilon(x,0) = u_0(x) & \text{on } \mathbb{T}^n. \end{cases}$$

We also consider a corresponding approximation for the ergodic problem (1.6):

$$(\mathrm{E})_\varepsilon \qquad H(x, Dv^\varepsilon) = \varepsilon^4 \Delta v^\varepsilon + \overline{H}_\varepsilon \quad \text{in } \mathbb{T}^n.$$

By Theorem 1.2 and Proposition 1.4, the existence and uniqueness of the ergodic constant $\overline{H}_\varepsilon$ of $(\mathrm{E})_\varepsilon$ holds. Besides, there exists a smooth solution $v^\varepsilon$ to $(\mathrm{E})_\varepsilon$.

The advantage of considering $(\mathrm{A})_\varepsilon$ and $(\mathrm{E})_\varepsilon$ lies in the fact that the solutions of these equations are smooth, and this will allow us to use the nonlinear adjoint method to perform rigorous calculations in the next subsection.

**Proposition 2.3.** *Assume that* (1.11), (H2) *and* (H4) *hold, and the ergodic constant of* (1.6) *is* 0. *Let $u^\varepsilon$ and $w^\varepsilon$ be the solution of $(\mathrm{C})_\varepsilon$ and $(\mathrm{A})_\varepsilon$ with a given initial data $u_0 \in \mathrm{Lip}\,(\mathbb{T}^n)$, respectively. There exists $C > 0$ independent of $\varepsilon$ such that*

$$\begin{aligned} &\|u^\varepsilon(\cdot,1) - w^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} \le C\varepsilon, \\ &|\overline{H}_\varepsilon| \le C\varepsilon^2. \end{aligned}$$

*Proof.* We consider the function $\Phi : \mathbb{T}^n \times \mathbb{T}^n \times [0,1] \to \mathbb{R}$ defined by

$$\Phi(x,y,t) := u^\varepsilon(x,t) - w^\varepsilon(y,t) - \frac{|x-y|^2}{2\eta} - Kt$$

for $\eta > 0$ and $K > 0$ to be fixed later. Pick $(x_\eta, y_\eta, t_\eta) \in \mathbb{T}^n \times \mathbb{T}^n \times [0,1]$ such that

$$\Phi(x_\eta, y_\eta, t_\eta) = \max_{x,y\in\mathbb{T}^n,\, t\in[0,1]} \Phi.$$

In the case $t_\eta > 0$, in light of Ishii's lemma (see [23, Theorem 3.2.19]), for any $\rho \in (0,1)$, there exist $(a_\eta, p_\eta, X_\eta) \in \overline{J}^{2,+} u^\varepsilon(x_\eta, t_\eta)$ and $(b_\eta, p_\eta, Y_\eta) \in \overline{J}^{2,-} w^\varepsilon(y_\eta, t_\eta)$ such that

$$a_\eta - b_\eta = K, \; p_\eta = \frac{x_\eta - y_\eta}{\eta}, \; \begin{pmatrix} X_\eta & 0 \\ 0 & -Y_\eta \end{pmatrix} \le A_\eta + \rho A_\eta^2, \tag{2.13}$$

where

$$A_\eta := \frac{1}{\eta} \begin{pmatrix} I_n & -I_n \\ -I_n & I_n \end{pmatrix}.$$

Here, $\overline{J}^{2,\pm}$ denotes the super-semijet, and sub-semijet, respectively (see Section 4.2).

We need to be careful for the case $t_\eta = 1$, which is handled by Lemma 4.4. By the definition of viscosity solutions,

$$\varepsilon a_\eta + H(x_\eta, p_\eta) \le 0 \quad \text{and} \quad \varepsilon b_\eta + H(x_\eta, p_\eta) \ge \varepsilon^4 \mathrm{tr}\,(Y_\eta),$$

which implies

$$\varepsilon K + H(x_\eta, p_\eta) - H(y_\eta, p_\eta) \le -\varepsilon^4 \mathrm{tr}\,(Y_\eta).$$

Note that

$$\begin{aligned}
-\varepsilon^4 \mathrm{tr}\,(Y_\eta) &= \sum_{i=1}^n \big\{ \langle X_\eta 0 e_i, 0 e_i \rangle - \langle Y_\eta \varepsilon^2 e_i, \varepsilon^2 e_i \rangle \big\} \\
&\le \sum_{i=1}^n \left\{ A_\eta \begin{pmatrix} 0e_i \\ \varepsilon^2 e_i \end{pmatrix} \cdot \begin{pmatrix} 0e_i \\ \varepsilon^2 e_i \end{pmatrix} + \rho A_\eta^2 \begin{pmatrix} 0e_i \\ \varepsilon^2 e_i \end{pmatrix} \cdot \begin{pmatrix} 0e_i \\ \varepsilon^2 e_i \end{pmatrix} \right\} \\
&\le \frac{C\varepsilon^4}{\eta} + O(\rho).
\end{aligned}$$

Since $\Phi(y_\eta, y_\eta, t_\eta) \le \Phi(x_\eta, y_\eta, t_\eta)$, we have

$$u^\varepsilon(y_\eta, t_\eta) - w^\varepsilon(y_\eta, t_\eta) - Kt_\eta \le u^\varepsilon(x_\eta, t_\eta) - w^\varepsilon(y_\eta, t_\eta) - \frac{|x_\eta - y_\eta|^2}{2\eta} - Kt_\eta,$$

which implies $|p_\eta| \le C$ for some $C > 0$ in view of the Lipschitz continuity of $u^\varepsilon$. Thus, $|x_\eta - y_\eta| \le C\eta$. Therefore,

$$|H(x_\eta, p_\eta) - H(y_\eta, p_\eta)| \le C(1 + |p_\eta|^2)|x_\eta - y_\eta| \le C\eta.$$

Combine the above to deduce

$$\varepsilon K \le C\varepsilon^4/\eta + C\eta + O(\rho) \quad \text{as } \rho \to 0.$$

Sending $\rho \to 0$ and setting $K := C'\varepsilon^{-1}(\varepsilon^4/\eta + \eta)$ for $C' > C$, we necessarily have

$$t_\eta = 0.$$

Thus, we get, for all $x \in \mathbb{T}^n, t \in [0, 1]$,

$$\Phi(x, x, t) \le \Phi(x_\eta, y_\eta, t_\eta) = \Phi(x_\eta, y_\eta, 0),$$

which reads for $t = 1$,

$$\begin{aligned}
&u^\varepsilon(x, 1) - w^\varepsilon(x, 1) \le u_0(x_\eta) - u_0(y_\eta) + K \\
&\le \|Du_0\|_{L^\infty(\mathbb{T}^n)} \eta + C'\varepsilon^{-1}\Big(\frac{\varepsilon^4}{\eta} + \eta\Big) = C\left(\frac{\varepsilon^3}{\eta} + \Big(1 + \frac{1}{\varepsilon}\Big)\eta\right).
\end{aligned}$$

Setting

$$\eta := \varepsilon^2,$$

we get $u^\varepsilon(x,1) - w^\varepsilon(x,1) \le C\varepsilon^2$ for all $x \in \mathbb{T}^n$. By exchanging the role of $u^\varepsilon$ and $w^\varepsilon$ in $\Phi$ and repeating a similar argument, we obtain $\|u^\varepsilon(\cdot,1) - w^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} \le C\varepsilon^2$.

Let us now prove

$$\left|\overline{H}_\varepsilon\right| = \left|\overline{H}_\varepsilon - c\right| \le C\varepsilon^2$$

in a similar way. Set

$$\Psi(x,y) := v(x) - v^\varepsilon(y) - \frac{|x-y|^2}{2\eta},$$

where $v^\varepsilon$ and $v$ are solutions to $(\mathrm{E})_\varepsilon$ and (1.6), respectively. For a maximum point $(x_\eta, y_\eta)$ of $\Psi$ on $\mathbb{T}^n \times \mathbb{T}^n$, we have

$$H(x_\eta, p_\eta) \le 0 \quad \text{and} \quad H(y_\eta, p_\eta) \ge \varepsilon^4 \mathrm{tr}\,(Y_\eta) + \overline{H}_\varepsilon$$

for any $(p_\eta, X_\eta) \in \overline{J}^{2,+} v(x_\eta)$, $(p_\eta, Y_\eta) \in \overline{J}^{2,-} v^\varepsilon(y_\eta)$. Note here that we are assuming the ergodic constant of (1.6) is $c = 0$ now.

Therefore, similarly to the above,

$$\overline{H}_\varepsilon - 0 \le H(y_\eta, p_\eta) - H(x_\eta, p_\eta) - \varepsilon^4 \mathrm{tr}\,(Y_\eta) \le C\left(\eta + \frac{\varepsilon^4}{\eta}\right).$$

Setting $\eta = \varepsilon^2$, we obtain $\overline{H}_\varepsilon \le C\varepsilon^2$. Symmetrically, we can prove $\overline{H}_\varepsilon \ge -C\varepsilon^2$, which yields the conclusion. □

**Remark 2.2.** As seen in the proof, the vanishing viscosity method gives that the rate of convergence of $u^\varepsilon - w^\varepsilon$ is

$$\sqrt{\text{viscosity coefficient}}/(\text{the coefficient of } u_t^\varepsilon \text{ and } w_t^\varepsilon).$$

Because of this fact, we can choose $\varepsilon^\alpha$ for any $\alpha > 2$ as a coefficient of the viscosity terms in $(\mathrm{A})_\varepsilon$ and $(\mathrm{E})_\varepsilon$. We choose $\alpha = 4$ here just to make the computations nice and clear.

### 2.3.3 Conservation of energy and a key observation

The adjoint equation of the linearized operator of $(\mathrm{A})_\varepsilon$ is

$$(\mathrm{AJ})_\varepsilon \qquad \begin{cases} -\varepsilon\sigma_t^\varepsilon - \mathrm{div}(D_pH(x, Dw^\varepsilon)\sigma^\varepsilon) = \varepsilon^4 \Delta\sigma^\varepsilon & \text{in } \mathbb{T}^n \times (0,1), \\ \sigma^\varepsilon(x,1) = \delta_{x_0} & \text{on } \mathbb{T}^n. \end{cases}$$

**Proposition 2.4** (Elementary Property of $\sigma^\varepsilon$)**.** *Assume that* (H2)–(H4) *hold. We have* $\sigma^\varepsilon > 0$ *in* $\mathbb{T}^n \times (0,1)$, *and*

$$\int_{\mathbb{T}^n} \sigma^\varepsilon(x,t)\,dx = 1 \quad \textit{for all } t \in [0,1].$$

*Proof.* We have that $\sigma^\varepsilon > 0$ in $\mathbb{T}^n \times (0,1)$ by the strong maximum principle for $(\mathrm{AJ})_\varepsilon$. Since

$$\varepsilon \frac{d}{dt} \int_{\mathbb{T}^n} \sigma^\varepsilon(\cdot,t)\,dx = \int_{\mathbb{T}^n} \left(-\mathrm{div}(D_pH(x,Dw^\varepsilon)\sigma^\varepsilon) - \varepsilon^4 \Delta\sigma^\varepsilon\right)\,dx = 0,$$

we conclude

$$\int_{\mathbb{T}^n} \sigma^\varepsilon(x,t)\,dx = \int_{\mathbb{T}^n} \sigma^\varepsilon(x,1)\,dx = \int_{\mathbb{T}^n} \delta_{x_0}\,dx = 1$$

for all $t \in [0,1]$, which is the conclusion. □

**Lemma 2.5** (Conservation of Energy 1). *Assume that* (H2)–(H4) *hold. Then we have the following properties*

(i) $\displaystyle \frac{d}{dt}\int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon)\sigma^\varepsilon\,dx = 0,$

(ii) $\displaystyle \varepsilon w_t^\varepsilon(x_0,1) = -\int_0^1 \int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon)\sigma^\varepsilon\,dx\,dt.$

*Proof.* We only need to prove (i) as (ii) follows directly from (i). This is a straightforward result of adjoint operators and comes from a direct calculation:

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon)\sigma^\varepsilon\,dx \\
&= \int_{\mathbb{T}^n} (D_pH(x,Dw^\varepsilon)\cdot Dw_t^\varepsilon - \varepsilon^4 \Delta w_t^\varepsilon)\sigma^\varepsilon\,dx + \int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon)\sigma_t^\varepsilon\,dx \\
&= -\int_{\mathbb{T}^n} \Big(\mathrm{div}\left(D_pH(x,Dw^\varepsilon)\sigma^\varepsilon\right) + \varepsilon^4 \Delta\sigma^\varepsilon\Big) w_t^\varepsilon\,dx - \int_{\mathbb{T}^n} \varepsilon w_t^\varepsilon \sigma_t^\varepsilon\,dx = 0.
\end{aligned}$$

□

**Remark 2.3.**
(i) We stress the fact that identity (ii) in Lemma 2.5 is extremely important. If we scale back the time, the integral on the right hand side becomes

$$-\frac{1}{T}\int_0^T \int_{\mathbb{T}^n} \left[H(x,Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon\right] \sigma^\varepsilon(x,t)\,dx\,dt,$$

where $T = 1/\varepsilon \to \infty$. This is the averaging action as $t \to \infty$, which is a key observation. We observed this in a formal calculation in Section 2.3.1.
(ii) We emphasize here that we do not use any specific structure of the equations up to now, and therefore this conservation law holds in a much more general setting. To analyze further, we need to require more specific structures and perform some delicate analysis. But it is worth mentioning that, in this reason, this method for the large-time asymptotics for nonlinear equations is universal and robust in principle.

The following theorem is a rigorous interpretation of asymptotic monotone property (2.7) of the solution to (1.1), which is essential in the proof of Theorem 2.2.

**Theorem 2.6.** *Assume that* (H2)–(H4) *hold, and the ergodic constant of* (1.6) *is* 0. *We have*

$$\lim_{\varepsilon\to 0} \varepsilon \|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} = 0.$$

*More precisely, there exists a positive constant $C$, independent of $\varepsilon$, such that*

$$\varepsilon\|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} = \|H(\cdot, Dw^\varepsilon(\cdot,1)) - \varepsilon^4\Delta w^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} \le C\varepsilon^{1/2}.$$

To prove this, we use the following key estimates, which will be proved in the next subsection.

**Lemma 2.7** (Key Estimates 1)**.** *Assume that* (H2)–(H4) *hold, and the ergodic constant of* (1.6) *is* 0*. There exists a positive constant $C$, independent of $\varepsilon$, such that the followings hold*:

(i) $\displaystyle\int_0^1\int_{\mathbb{T}^n} |D(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt \le C\varepsilon,$

(ii) $\displaystyle\int_0^1\int_{\mathbb{T}^n} |D^2(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt \le C\varepsilon^{-7}.$

We now give the proof of Theorem 2.6 by using the averaging action above and the key estimates in Lemma 2.7.

*Proof of Theorem* 2.6. Let us first choose $x_0$, which may depend on $\varepsilon$, such that

$$|\varepsilon w_t^\varepsilon(x_0,1)| = \|\varepsilon w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} = \|H(\cdot, Dw^\varepsilon(\cdot,1)) - \varepsilon^4\Delta w^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)}.$$

Thanks to Lemma 2.5,

$$|\varepsilon w_t^\varepsilon(x_0,1)| = \left|\int_0^1\int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4\Delta w^\varepsilon)\sigma^\varepsilon\,dx\,dt\right|.$$

Let $v^\varepsilon$ be a solution of $(\mathrm{E})_\varepsilon$. By Proposition 2.3,

$$\begin{aligned}
&\varepsilon\|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)}\\
&= \left|\int_0^1\int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4\Delta w^\varepsilon)\sigma^\varepsilon\,dx\,dt\right|\\
&= \left|\int_0^1\int_{\mathbb{T}^n} (H(x,Dw^\varepsilon) - \varepsilon^4\Delta w^\varepsilon - (H(x,Dv^\varepsilon) - \varepsilon^4\Delta v^\varepsilon - \overline{H}_\varepsilon))\sigma^\varepsilon\,dx\,dt\right|\\
&\le \int_0^1\int_{\mathbb{T}^n} |H(x,Dw^\varepsilon) - H(x,Dv^\varepsilon)|\,\sigma^\varepsilon + \varepsilon^4|\Delta(w^\varepsilon - v^\varepsilon)|\sigma^\varepsilon\,dx\,dt + |\overline{H}_\varepsilon|\\
&\le C\int_0^1\int_{\mathbb{T}^n} \left[|D(w^\varepsilon - v^\varepsilon)| + \varepsilon^4|D^2(w^\varepsilon - v^\varepsilon)|\right]\sigma^\varepsilon\,dx\,dt + C\varepsilon^2.
\end{aligned}$$

We finally use the Hölder inequality and Lemma 2.7 to get that

$$\begin{aligned}
&\varepsilon\|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)}\\
&\le C\left[\left(\int_0^1\int_{\mathbb{T}^n} |D(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt\right)^{1/2} + \varepsilon^4\left(\int_0^1\int_{\mathbb{T}^n} |D^2(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt\right)^{1/2} + \varepsilon^2\right]\\
&\le C\varepsilon^{1/2}.
\end{aligned}$$

□

Let us now present the proof of the large time asymptotics of $u$, Theorem 2.2.

*Proof of Theorem* 2.2. Firstly, the equi-Lipschitz continuity of $\{w^\varepsilon(\cdot,1)\}_{\varepsilon>0}$ is obtained by an argument similar to that of the proof of Theorem 1.2. Therefore, we are able to choose a sequence $\varepsilon_m \to 0$ as $m \to \infty$ such that $\{w^{\varepsilon_m}(\cdot,1)\}_{m\in\mathbb{N}}$ converges uniformly to a continuous function $v$ in $\mathbb{T}^n$, which may depend on the choice of $\{\varepsilon_m\}_{m\in\mathbb{N}}$. We let $t_m := 1/\varepsilon_m$ for $m \in \mathbb{N}$, and use Proposition 2.3 to deduce that

$$\|u(\cdot,t_m) - v\|_{L^\infty(\mathbb{T}^n)} \to 0 \text{ as } m \to \infty.$$

Let us show that the limit of $u(\cdot,t)$ as $t \to \infty$ does not depend on the sequence $\{t_m\}_{m\in\mathbb{N}}$. In view of Theorem 2.6, which is one of our main results in this chapter, $v$ is a solution of (E), and thus a (time independent) solution of the equation in (C). Therefore, for any $x \in \mathbb{T}^n$, and $t > 0$ such that $t_m \le t < t_{m+1}$, we use the comparison principle for (C) to yield that

$$|u(x,t) - v(x)| \le \|u(\cdot,t_m + (t-t_m)) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)} \le \|u(\cdot,t_m) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)}.$$

Thus,

$$\lim_{t\to\infty} \|u(\cdot,t) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)} \le \lim_{m\to\infty} \|u(\cdot,t_m) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)} = 0,$$

which gives the conclusion. □

### 2.3.4 Proof of key estimates

A key idea to prove estimates in Lemma 2.7 is to use a combination of the Bernstein method and the adjoint technique.

**Lemma 2.8.** *Assume that* (H2), (H4) *hold. Let $w^\varepsilon$ be the solution of* $(\mathrm{A})_\varepsilon$. *There exists a constant $C > 0$ independent of $\varepsilon$ such that*

$$\int_0^1 \int_{\mathbb{T}^n} \varepsilon^4 |D^2 w^\varepsilon|^2 \sigma^\varepsilon \, dx \, dt \le C.$$

This is one of the key estimates which was first introduced by Evans [32] in the study of gradient shock structures of the vanishing viscosity procedure of nonconvex, first-order Hamilton–Jacobi equations. See also Tran [78]. The convexity of $H$ is not needed at all to get the conclusion of this lemma as can be seen in the proof.

*Proof.* By a computation similar to that in the proof of Theorem 1.2, for $\varphi(x,t) := |Dw^\varepsilon|^2/2$, we have

$$\varepsilon\varphi_t + D_pH \cdot D\varphi + D_xH \cdot Dw^\varepsilon = \varepsilon^4(\Delta\varphi - |D^2w^\varepsilon|^2).$$

Multiply the above by $\sigma^\varepsilon$ and integrate over $\mathbb{T}^n \times [0,1]$ to yield

$$\begin{aligned}\varepsilon^4 \int_0^1 \int_{\mathbb{T}^n} |D^2 w^\varepsilon|^2 \sigma^\varepsilon \, dxdt = &- \int_0^1 \int_{\mathbb{T}^n} \left(\varepsilon\varphi_t + D_pH \cdot D\varphi - \varepsilon^4 \Delta\varphi\right) \sigma^\varepsilon \, dxdt \\ &- \int_0^1 \int_{\mathbb{T}^n} D_xH \cdot Dw^\varepsilon \sigma^\varepsilon \, dxdt.\end{aligned}$$

Integrating by parts, we get

$$\begin{aligned}
&\int_0^1\int_{\mathbb{T}^n}\left(\varepsilon\varphi_t + D_pH\cdot D\varphi - \varepsilon^4\Delta\varphi\right)\sigma^\varepsilon\,dxdt\\
&=\int_{\mathbb{T}^n}[\varepsilon\varphi\sigma^\varepsilon]_{t=0}^{t=1}\,dx + \int_0^1\int_{\mathbb{T}^n}\left(-\varepsilon\sigma^\varepsilon_t - \operatorname{div}\,(D_pH\sigma^\varepsilon) - \varepsilon^4\Delta\sigma^\varepsilon\right)\varphi\,dxdt\\
&\le C\varepsilon.
\end{aligned}$$

Noting that $|D_xH\cdot Dw^\varepsilon|\le C$, we get the conclusion.

□

*Proof of Lemma* 2.7 (i). Subtracting equation $(\mathrm{A})_\varepsilon$ from $(\mathrm{E})_\varepsilon$, thanks to the uniform convexity of $H$, we get

$$\begin{aligned}
0 &= \varepsilon(v^\varepsilon - w^\varepsilon)_t + H(x,Dv^\varepsilon) - H(x,Dw^\varepsilon) - \varepsilon^4\Delta(v^\varepsilon - w^\varepsilon) - \overline{H}_\varepsilon\\
&\ge \varepsilon(v^\varepsilon - w^\varepsilon)_t + D_pH(x,Dw^\varepsilon)\cdot D(v^\varepsilon - w^\varepsilon) + \theta|D(v^\varepsilon - w^\varepsilon)|^2 - \varepsilon^4\Delta(v^\varepsilon - w^\varepsilon) - \overline{H}_\varepsilon.
\end{aligned}$$

Multiply the above inequality by $\sigma^\varepsilon$ and integrate by parts on $\mathbb{T}^n\times[0,1]$ to deduce that

$$\begin{aligned}
&\theta\int_0^1\int_{\mathbb{T}^n}|D(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt\\
&\le \overline{H}_\varepsilon - \int_0^1\int_{\mathbb{T}^n}\varepsilon((v^\varepsilon - w^\varepsilon)\sigma^\varepsilon)_t\,dxdt\\
&\quad + \int_0^1\int_{\mathbb{T}^n}\left[\varepsilon\sigma^\varepsilon_t + \operatorname{div}\,(D_pH(x,Dw^\varepsilon)\sigma^\varepsilon) + \varepsilon^4\Delta\sigma^\varepsilon\right](v^\varepsilon - w^\varepsilon)\,dxdt\\
&= \overline{H}_\varepsilon + \varepsilon\left[\int_{\mathbb{T}^n}(w^\varepsilon - v^\varepsilon)\sigma^\varepsilon\,dx\right]_{t=0}^{t=1}\\
&= \overline{H}_\varepsilon + \varepsilon(w^\varepsilon(x_0,1) - v^\varepsilon(x_0)) - \varepsilon\int_{\mathbb{T}^n}(u_0(x) - v^\varepsilon(x))\sigma^\varepsilon(x,0)\,dx\\
&= \overline{H}_\varepsilon + \varepsilon w^\varepsilon(x_0,1) - \varepsilon\int_{\mathbb{T}^n}(v^\varepsilon(x_0) - v^\varepsilon(x))\sigma^\varepsilon(x,0)\,dx - \varepsilon\int_{\mathbb{T}^n}u_0(x)\sigma^\varepsilon(x,0)\,dx.
\end{aligned}$$

Note here that $w^\varepsilon$ satisfies

$$\left\|w^\varepsilon + \frac{\overline{H}_\varepsilon t}{\varepsilon}\right\|_{L^\infty(\mathbb{T}^n\times(0,\infty))}\le C\quad\text{for some } C>0,$$

as we see in the proof of Proposition 1.6. Thus, $\overline{H}_\varepsilon + \varepsilon w^\varepsilon(x_0,1)\le C\varepsilon$ and

$$\theta\int_0^1\int_{\mathbb{T}^n}|D(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt\le\varepsilon(C + \|Dv^\varepsilon\|_{L^\infty(\mathbb{T}^n)} + \|u_0\|_{L^\infty(\mathbb{T}^n)})\le C\varepsilon,$$

which implies the conclusion.

□

*Proof of Lemma* 2.7 (ii). Subtract $(\mathrm{A})_\varepsilon$ from $(\mathrm{E})_\varepsilon$ and differentiate with respect to $x_i$ to get

$$\begin{aligned}\varepsilon(v^\varepsilon - w^\varepsilon)_{x_i t} + D_pH(x, Dv^\varepsilon)\cdot Dv^\varepsilon_{x_i} - D_pH(x, Dw^\varepsilon)\cdot Dw^\varepsilon_{x_i}\\ + H_{x_i}(x, Dv^\varepsilon) - H_{x_i}(x, Dw^\varepsilon) - \varepsilon^4\Delta(v^\varepsilon - w^\varepsilon)_{x_i} = 0.\end{aligned}$$

Let $\varphi(x,t) := |D(v^\varepsilon - w^\varepsilon)|^2/2$. Multiplying the last identity by $(v^\varepsilon - w^\varepsilon)_{x_i}$ and summing up with respect to $i$, we achieve that

$$\begin{aligned}&\varepsilon\varphi_t + D_pH(x, Dw^\varepsilon)\cdot D\varphi + \left[\Big(D_pH(x, Dv^\varepsilon) - D_pH(x, Dw^\varepsilon)\Big)\cdot Dv^\varepsilon_{x_i}\right](v^\varepsilon_{x_i} - w^\varepsilon_{x_i})\\ &+ \Big(D_xH(x, Dv^\varepsilon) - D_xH(x, Dw^\varepsilon)\Big)\cdot D(v^\varepsilon - w^\varepsilon) - \varepsilon^4\big(\Delta\varphi - |D^2(v^\varepsilon - w^\varepsilon)|^2\big) = 0.\end{aligned}$$

By using the equi-Lipschitz continuity of $v^\varepsilon, w^\varepsilon$ and (H4), we derive that

$$\begin{aligned}&\varepsilon\varphi_t + D_pH(x, Dw^\varepsilon)\cdot D\varphi - \varepsilon^4\Delta\varphi + \varepsilon^4|D^2(v^\varepsilon - w^\varepsilon)|^2\\ &\le C(|D^2v^\varepsilon| + 1)|D(v^\varepsilon - w^\varepsilon)|^2.\end{aligned} \tag{2.14}$$

The right hand side of (2.14) is a dangerous term because of the term $|D^2v^\varepsilon|$. We now take advantage of Lemma 2.8 to handle it. Using the fact that $\|Dv^\varepsilon\|_{L^\infty}$ and $\|Dw^\varepsilon\|_{L^\infty}$ are bounded, we have

$$\begin{aligned}C|D^2v^\varepsilon|\,|D(v^\varepsilon - w^\varepsilon)|^2 &\le C|D^2(v^\varepsilon - w^\varepsilon)|\,|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|\,|D(v^\varepsilon - w^\varepsilon)|^2\\ &\le \frac{\varepsilon^4}{2}|D^2(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\varepsilon^4}|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|.\end{aligned} \tag{2.15}$$

Combine (2.14) and (2.15) to deduce

$$\begin{aligned}&\varepsilon\varphi_t + D_pH(x, Dw^\varepsilon)\cdot D\varphi - \varepsilon^4\Delta\varphi + \frac{\varepsilon^4}{2}|D^2(v^\varepsilon - w^\varepsilon)|^2\\ &\qquad\le C|D(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\varepsilon^4}|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|.\end{aligned} \tag{2.16}$$

We multiply (2.16) by $\sigma^\varepsilon$, integrate over $\mathbb{T}^n \times [0,1]$, and use integration by parts to yield that, in light of Lemma 2.8 and (i),

$$\begin{aligned}&\varepsilon^4\int_0^1\int_{\mathbb{T}^n}|D^2(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon\,dx\,dt \le C\varepsilon + \frac{C}{\varepsilon^4}\varepsilon + C\int_0^1\int_{\mathbb{T}^n}|D^2w^\varepsilon|\sigma^\varepsilon\,dx\,dt\\ &\le \frac{C}{\varepsilon^3} + C\Big(\int_0^1\int_{\mathbb{T}^n}|D^2w^\varepsilon|^2\sigma^\varepsilon\,dx\,dt\Big)^{1/2}\Big(\int_0^1\int_{\mathbb{T}^n}\sigma^\varepsilon\,dx\,dt\Big)^{1/2} \le \frac{C}{\varepsilon^3} + \frac{C}{\varepsilon^2} \le \frac{C}{\varepsilon^3}. \qquad \square\end{aligned}$$

**Remark 2.4.** The estimates in Lemma 2.7 give us much better control of $D(w^\varepsilon - v^\varepsilon)$ and $D^2(w^\varepsilon - v^\varepsilon)$ on the support of $\sigma^\varepsilon$. More precisely, the classical a priori estimates by using the Bernstein method as in the proof of Theorem 1.2 only imply that $D(w^\varepsilon - v^\varepsilon)$ and $\varepsilon^4\Delta(w^\varepsilon - v^\varepsilon)$ are bounded.

By using the adjoint equation, we can get further formally that $\varepsilon^{-1/2}D(w^\varepsilon - v^\varepsilon)$ and $\varepsilon^{7/2}D^2(w^\varepsilon - v^\varepsilon)$ are bounded on the support of $\sigma^\varepsilon$. Clearly, these new estimates are much stronger than the known ones on the support of $\sigma^\varepsilon$. However, we must point out that, as $\varepsilon \to 0$, the supports of subsequential limits of $\{\sigma^\varepsilon\}_{\varepsilon>0}$ could be very singular. Understanding deeper about this point is essential in achieving further developments of this new approach in the near future.

It is also worth mentioning that we eventually do not need to use the graph theorem in the whole procedure above.

## 2.4 Degenerate viscous case

In this section, we consider a general possibly degenerate viscous Hamilton–Jacobi equation:

$$u_t - \operatorname{tr}\big(A(x)D^2u\big) + H(x, Du) = 0 \qquad \text{in } \mathbb{T}^n \times (0,\infty). \tag{2.17}$$

Here is one of the main results of [17].

**Theorem 2.9.** *Assume that* (H1)–(H4) *hold. Let $u$ be the solution of* (2.17) *with initial data $u(\cdot,0) = u_0 \in \operatorname{Lip}(\mathbb{T}^n)$. Then there exists $(v,c) \in \operatorname{Lip}(\mathbb{T}^n) \times \mathbb{R}$ such that* (1.5) *holds, that is,*

$$u(x,t) - (v(x) - ct) \to 0 \quad \textit{uniformly for } x \in \mathbb{T}^n \textit{ as } t \to \infty,$$

*where the pair $(v,c)$ is a solution of the ergodic problem*

$$-\operatorname{tr}\big(A(x)D^2v\big) + H(x, Dv) = c \quad \textit{in } \mathbb{T}^n.$$

For an easy explanation, we consider the 1-dimensional case (i.e., $n = 1$) in this section. This makes the problem much easier but we do not lose the key difficulties coming from the degenerate viscous term $-\operatorname{tr}\big(A(x)D^2u\big)$. We now assume that the ergodic constant $c$ for (1.8) is 0 as in Section 2.3.3.

We repeat the same procedures as those in Sections 2.3.2 and 2.3.3. Associated problems are now described below:

(C) $\quad u_t - a(x)u_{xx} + H(x,u_x) = 0$ in $\mathbb{T} \times (0,\infty)$, $\qquad u(x,0) = u_0(x)$ in $\mathbb{T}$,

$(\mathrm{A})_\varepsilon$ $\quad \varepsilon w^\varepsilon_t - a(x)w^\varepsilon_{xx} + H(x,w^\varepsilon_x) = \varepsilon^4 w^\varepsilon_{xx}$ in $\mathbb{T} \times (0,\infty)$, $\qquad w^\varepsilon(x,0) = u_0(x)$ in $\mathbb{T}$,

$(\mathrm{AJ})_\varepsilon$ $\quad -\varepsilon\sigma^\varepsilon_t - (a(x)\sigma^\varepsilon)_{xx} - (H_p(x,w^\varepsilon_x)\sigma^\varepsilon)_x = \varepsilon^4\sigma^\varepsilon_{xx}$ in $\mathbb{T} \times (0,1)$, $\quad \sigma^\varepsilon(x,1) = \delta_{x_0}$ in $\mathbb{T}$,

$(\mathrm{E})_\varepsilon$ $\quad -a(x)v^\varepsilon_{xx} + H(x,v^\varepsilon_x) = \varepsilon^4 v^\varepsilon_{xx} + \overline{H}_\varepsilon$ in $\mathbb{T}$.

Here, assumption (H1) means that $a \in C^2(\mathbb{T})$ is a nonnegative function.

As pointed out in Remark 2.3, we have the same type conservation of energy.

**Lemma 2.10** (Conservation of Energy 2)**.** *Assume that* (H1)–(H4) *hold, and the associated ergodic constant is zero. The following properties hold*:

(i) $\displaystyle \frac{d}{dt}\int_{\mathbb{T}} \left[H(x,w^\varepsilon_x) - (a(x)+\varepsilon^4)w^\varepsilon_{xx}\right]\sigma^\varepsilon\,dx = 0,$

(ii) $\displaystyle \varepsilon w^\varepsilon_t(x_0,1) = -\int_0^1\int_{\mathbb{T}} \left[H(x,w^\varepsilon_x) - (a(x)+\varepsilon^4)w^\varepsilon_{xx}\right]\sigma^\varepsilon\,dx\,dt.$

The proof of this lemma is similar to that of Lemma 2.5, hence is omitted.

Now, as in the proof of Theorem 2.6, we have

$$
\begin{aligned}
&\varepsilon\|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T})} = \|H(\cdot,w_x^\varepsilon(\cdot,1)) - (\varepsilon^4 + a(x))w_{xx}^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T})}\\
&= \left|\int_0^1\int_{\mathbb{T}} \left[H(x,w_x^\varepsilon) - (\varepsilon^4 + a(x))w_{xx}^\varepsilon\right]\sigma^\varepsilon\,dx\,dt\right|\\
&= \left|\int_0^1\int_{\mathbb{T}} \left[H(x,w_x^\varepsilon) - (\varepsilon^4 + a(x))w_{xx}^\varepsilon - (H(x,v_x^\varepsilon) - (\varepsilon^4 + a(x))v_{xx}^\varepsilon - \overline{H}_\varepsilon)\right]\sigma^\varepsilon\,dx\,dt\right|\\
&\le \int_0^1\int_{\mathbb{T}} \left(|H(x,w_x^\varepsilon) - H(x,v_x^\varepsilon)| + |(\varepsilon^4 + a(x))(w^\varepsilon - v^\varepsilon)_{xx}|\right)\sigma^\varepsilon\,dx\,dt + \left|\overline{H}_\varepsilon\right|\\
&\le C\Big[\left(\int_0^1\int_{\mathbb{T}} |(w^\varepsilon - v^\varepsilon)_x|^2\sigma^\varepsilon\,dx\,dt\right)^{1/2} + \varepsilon^4\left(\int_0^1\int_{\mathbb{T}} |(w^\varepsilon - v^\varepsilon)_{xx}|^2\sigma^\varepsilon\,dx\,dt\right)^{1/2}\\
&\quad + \left(\int_0^1\int_{\mathbb{T}} a(x)^2|(w^\varepsilon - v^\varepsilon)_{xx}|^2\sigma^\varepsilon\,dx\,dt\right)^{1/2} + \left|\overline{H}_\varepsilon\right|\Big],
\end{aligned}
$$

where $v^\varepsilon$ is a solution of $(\mathrm{E})_\varepsilon$ ($v^\varepsilon$ is unique up to an additive constant).

Since $c = 0$, we have $\left|\overline{H}_\varepsilon\right| \le C\varepsilon^2$. Therefore, in order to control $\varepsilon\|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T})}$, we basically need to bound three terms on the right hand side of the above. The first two already appear in the previous section, and the last term is a new term due to the appearance of the possibly degenerate diffusion $a(x)$. We now redo the same procedure to handle these three with great care as the possibly degenerate diffusion $a(x)$ is quite dangerous.

**Lemma 2.11** (Key Estimates 2)**.** *Assume that* (H1)–(H4) *hold, and the associated ergodic constant is zero. There exists a constant $C > 0$, independent of $\varepsilon$, such that*

(i) $\displaystyle\int_0^1\int_{\mathbb{T}} |(w^\varepsilon - v^\varepsilon)_x|^2\sigma^\varepsilon\,dx\,dt \le C\varepsilon,$

(ii) $\displaystyle\int_0^1\int_{\mathbb{T}} (a(x) + \varepsilon^4)|w_{xx}^\varepsilon|^2\sigma^\varepsilon\,dx\,dt \le C,$

(iii) $\displaystyle\int_0^1\int_{\mathbb{T}} |(w^\varepsilon - v^\varepsilon)_{xx}|^2\sigma^\varepsilon\,dx\,dt \le C\varepsilon^{-7},$

(iv) $\displaystyle\int_0^1\int_{\mathbb{T}} a^2(x)|(w^\varepsilon - v^\varepsilon)_{xx}|^2\sigma^\varepsilon\,dx\,dt \le C\sqrt{\varepsilon}.$

*Proof.* The proof of (i) is similar to that of Lemma 2.7 (i), hence is omitted. Notice that if we do not differentiate the equation, then we do not have any difficulty which comes from the diffusion term $a(x)$. On the other hand, once we differentiate the equation to obtain some estimates, then we face some of difficulties coming from the term $a$ as seen below.

We now prove (ii). Let $w^\varepsilon$ be the solution of $(\mathrm{A})_\varepsilon$. Differentiate $(\mathrm{A})_\varepsilon$ with respect to the $x$ variable to get

$$
\varepsilon w_{tx}^\varepsilon + H_p(x,w_x^\varepsilon)\cdot w_{xx}^\varepsilon + H_x(x,w_x^\varepsilon) - (\varepsilon^4 + a)w_{xxx}^\varepsilon - a_x w_{xx}^\varepsilon = 0. \tag{2.18}
$$

Here we write $H_p(x,p) = D_pH(x,p)$, $H_x(x,p) = D_xH(x,p)$ as we are in the 1-dimensional space.

Let $\xi(x,t) := |w_x^\varepsilon|^2/2$. Note that

$$\xi_t = w_x^\varepsilon w_{xt}^\varepsilon,\ \xi_x = w_x^\varepsilon w_{xx}^\varepsilon,\ \xi_{xx} = |w_{xx}^\varepsilon|^2 + w_x^\varepsilon w_{xxx}^\varepsilon.$$

Multiply (2.18) by $w_x^\varepsilon$ to arrive at

$$\varepsilon\xi_t + H_p \cdot \xi_x + H_x \cdot w_x^\varepsilon = (\varepsilon^4 + a)(\xi_{xx} - |w_{xx}^\varepsilon|^2) + (a_x \cdot w_x^\varepsilon)w_{xx}^\varepsilon.$$

We need to be careful for the last term which comes from the diffusion term. Notice first that we have

$$a_x^2(x) \le Ca(x) \quad \text{for all } x \in \mathbb{T} \tag{2.19}$$

since $a \in C^2(\mathbb{T})$. Indeed, $a \in C^2(\mathbb{T})$ implies $\sqrt{a} \in \mathrm{Lip}\,(\mathbb{T})$. Thus, $|a_x| = |2(\sqrt{a})_x\sqrt{a}| \le C\sqrt{a}$. We next notice that for $\delta > 0$ small enough,

$$a_x w_x^\varepsilon w_{xx}^\varepsilon \le C|a_x||w_{xx}^\varepsilon| \le \frac{C}{\delta} + \delta a_x^2|w_{xx}^\varepsilon|^2 \le C + \frac{1}{2}a|w_{xx}^\varepsilon|^2. \tag{2.20}$$

Hence,

$$\varepsilon\xi_t + H_p \cdot \xi_x - (\varepsilon^4 + a)\xi_{xx} + (\varepsilon^4 + \frac{a}{2})|w_{xx}^\varepsilon|^2 \le C.$$

Multiply the above by $\sigma^\varepsilon$, integrate over $\mathbb{T} \times [0,1]$, and use integration by parts to yield the conclusion of (ii).

Next, we prove (iii). Subtract $(\mathrm{A})_\varepsilon$ from $(\mathrm{E})_\varepsilon$ and differentiate with respect to the variable $x$ to get

$$\begin{aligned}\varepsilon(v^\varepsilon - w^\varepsilon)_{xt} &+ H_p(x,v_x^\varepsilon)\cdot v_{xx}^\varepsilon - H_p(x,w_x^\varepsilon)\cdot w_{xx}^\varepsilon \\ &+ H_x(x,v_x^\varepsilon) - H_x(x,w_x^\varepsilon) - (\varepsilon^4 + a)(v^\varepsilon - w^\varepsilon)_{xxx} - a_x(v^\varepsilon - w^\varepsilon)_{xx} = 0.\end{aligned}$$

Let $\varphi(x,t) := |(v^\varepsilon - w^\varepsilon)_x|^2/2$. Multiplying the last identity by $(v^\varepsilon - w^\varepsilon)_x$, we achieve that

$$\begin{aligned}\varepsilon\varphi_t + H_p(x,w_x^\varepsilon)\cdot\varphi_x &+ \left[\Big(H_p(x,v_x^\varepsilon) - H_p(x,w_x^\varepsilon)\Big)\cdot v_{xx}^\varepsilon\right](v^\varepsilon - w^\varepsilon)_x \\ + \Big(H_x(x,v_x^\varepsilon) - H_x(x,w_x^\varepsilon)\Big)&\cdot(v^\varepsilon - w^\varepsilon)_x + (\varepsilon^4 + a(x))(|(v^\varepsilon - w^\varepsilon)_{xx}|^2 - \varphi_{xx}) \\ &- [a_x\cdot(v^\varepsilon - w^\varepsilon)_{xx}]\,(v^\varepsilon - w^\varepsilon)_x = 0.\end{aligned}$$

We only need to be careful for the last term as in the above

$$\begin{aligned}\big|\big[a_x\cdot(v^\varepsilon - w^\varepsilon)_{xx}\big](v^\varepsilon - w^\varepsilon)_x\big| &\le \delta|a_x|^2|(v^\varepsilon - w^\varepsilon)_{xx}|^2 + \frac{1}{\delta}|(v^\varepsilon - w^\varepsilon)_x|^2 \\ \le \frac{a}{2}|(v^\varepsilon - w^\varepsilon)_{xx}|^2 + C|(v^\varepsilon - w^\varepsilon)_x|^2\end{aligned}$$

for $\delta > 0$ small enough. Thus,

$$\varepsilon\varphi_t + H_p(x, w^\varepsilon_x) \cdot \varphi_x - (\varepsilon^4 + a(x))\varphi_{xx} + \Big(\varepsilon^4 + \frac{a}{2}\Big)|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \\ \leq C(1 + |v^\varepsilon_{xx}|)|(v^\varepsilon - w^\varepsilon)_x|^2. \quad (2.21)$$

By using the same trick as (2.15), we get

$$\varepsilon\varphi_t + H_p(x, w^\varepsilon_x) \cdot \varphi_x - (\varepsilon^4 + a(x))\varphi_{xx} + \Big(\frac{\varepsilon^4}{2} + \frac{a}{2}\Big)|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \\ \leq C|(v^\varepsilon - w^\varepsilon)_x|^2 + \frac{C}{\varepsilon^4}|(v^\varepsilon - w^\varepsilon)_x|^2 + C|w^\varepsilon_{xx}|.$$

We multiply the above by $\sigma^\varepsilon$, integrate over $\mathbb{T} \times [0, 1]$, and use integration by parts to yield that

$$\int_0^1 \int_{\mathbb{T}} (\varepsilon^4 + a(x))|(w^\varepsilon - v^\varepsilon)_{xx}|^2 \sigma^\varepsilon \, dx \, dt \leq C\varepsilon + \frac{C}{\varepsilon^4}\varepsilon + C\int_0^1 \int_{\mathbb{T}} |w^\varepsilon_{xx}|\sigma^\varepsilon \, dx \, dt \\ \leq \frac{C}{\varepsilon^3} + C\Big(\int_0^1 \int_{\mathbb{T}} |w^\varepsilon_{xx}|^2 \sigma^\varepsilon \, dx \, dt\Big)^{1/2}\Big(\int_0^1 \int_{\mathbb{T}} \sigma^\varepsilon \, dx \, dt\Big)^{1/2} \leq \frac{C}{\varepsilon^3} + \frac{C}{\varepsilon^2} \leq \frac{C}{\varepsilon^3},$$

which implies the conclusion of (iii).

Finally we prove (iv). Setting

$$\psi(x, t) := a(x)\varphi(x, t) = \frac{a(x)|(v^\varepsilon - w^\varepsilon)_x(x, t)|^2}{2},$$

and multiplying (2.21) by $a(x)$, we get

$$\varepsilon\psi_t + H_p(x, w^\varepsilon_x) \cdot (\psi_x - a_x\varphi) - (\varepsilon^4 + a(x))(\psi_{xx} - a_{xx}\varphi - 2a_x \cdot \varphi_x) \\ + a(x)\left(\varepsilon^4 + \frac{a(x)}{2}\right)|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \leq Ca(x)(|v^\varepsilon_{xx}| + 1)|(v^\varepsilon - w^\varepsilon)_x|^2.$$

Note that $a_x,\ a_{xx}$ are bounded. Then,

$$\varepsilon\psi_t + H_p(x, w^\varepsilon_x) \cdot \psi_x - (\varepsilon^4 + a(x))\psi_{xx} + a(x)\left(\varepsilon^4 + \frac{a(x)}{2}\right)|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \\ \leq C\varphi(x) - 2(\varepsilon^4 + a(x))a_x \cdot \varphi_x + Ca(x)|v^\varepsilon_{xx}|\ |(v^\varepsilon - w^\varepsilon)_x|^2.$$

For $\delta > 0$ small enough

$$2|(\varepsilon^4 + a(x))a_x \cdot \varphi_x| \leq C(\varepsilon^4 + a(x))|a_x|\ |(v^\varepsilon - w^\varepsilon)_{xx}|\ |(v^\varepsilon - w^\varepsilon)_x| \\ \leq \delta(\varepsilon^4 + a(x))|a_x|^2|(v^\varepsilon - w^\varepsilon)_{xx}|^2 + \frac{C}{\delta}|(v^\varepsilon - w^\varepsilon)_x|^2 \\ \leq \frac{1}{8}(\varepsilon^4 + a(x))a(x)|(v^\varepsilon - w^\varepsilon)_{xx}|^2 + C|(v^\varepsilon - w^\varepsilon)_x|^2$$

by using (2.19) again. Moreover,

$$
\begin{aligned}
&a(x)|v^\varepsilon_{xx}|\,|(v^\varepsilon - w^\varepsilon)_x|^2\\
\le\,& a(x)|w^\varepsilon_{xx}|\,|(v^\varepsilon - w^\varepsilon)_x|^2 + a(x)|(v^\varepsilon - w^\varepsilon)_{xx}|\,|(v^\varepsilon - w^\varepsilon)_x|^2\\
\le\,& \varepsilon^{1/2}a(x)|w^\varepsilon_{xx}|^2 + \frac{C}{\varepsilon^{1/2}}|(v^\varepsilon - w^\varepsilon)_x|^2 + \frac{a(x)^2}{8}|(v^\varepsilon - w^\varepsilon)_{xx}|^2 + C|(v^\varepsilon - w^\varepsilon)_x|^2.
\end{aligned}
$$

Combining everything, we obtain

$$
\begin{aligned}
\varepsilon\psi_t + H_p(x, w^\varepsilon_x)\cdot\psi_x - (\varepsilon^4 + a(x))\psi_{xx} + \frac{a(x)^2}{4}|(v^\varepsilon - w^\varepsilon)_{xx}|^2&\\
\le (C + C\varepsilon^{-1/2})|(v^\varepsilon - w^\varepsilon)_x|^2 + \varepsilon^{1/2}a(x)|w^\varepsilon_{xx}|^2&.
\end{aligned}
$$

We multiply the above inequality by $\sigma^\varepsilon$, integrate over $\mathbb{T}\times[0,1]$ and use (i), (ii) to yield (iv). □

Thanks to Lemmas 2.10, 2.11, we obtain

**Theorem 2.12.** *Assume that* (H1)–(H4) *hold, and the associated ergodic constant is zero. Let* $w^\varepsilon$ *be the solution of* $(\mathrm{A})_\varepsilon$ *with initial data* $u(\cdot,0) = u_0 \in \mathrm{Lip}\,(\mathbb{T})$*. Then,*

$$
\varepsilon\|w^\varepsilon_t(\cdot,1)\|_{L^\infty(\mathbb{T})} \le C\varepsilon^{1/4} \quad \textit{for some } C > 0,
$$

Theorem 2.9 in the case $n = 1$ is a straightforward result of Theorem 2.12 as seen in the proof of Theorem 2.2. We refer to [17] and [70] for the multi-dimensional setting.

**Remark 2.5.** If the equation in (C) is uniformly parabolic, that is, $a(x) > 0$ for all $x \in \mathbb{T}$, then estimate (iii) in Lemma 2.11 is not needed anymore as estimate (iv) in Lemma 2.11 is much stronger.

On the other hand, if $a$ is degenerate, then (iv) in Lemma 2.11 only provides the estimate of $|D^2(w^\varepsilon - v^\varepsilon)|^2\sigma^\varepsilon$ on the support of $a$, and it is hence essential to use (iii) in Lemma 2.11 to control the part where $a = 0$.

## 2.5 Asymptotic profile of the first-order case

In this section, we investigate the first-order Hamilton–Jacobi equation (1.1) again, and specifically focus on the *asymptotic profile*, which is

$$
u^\infty_c[u_0](x) := \lim_{t\to\infty}\left(u(x,t) + ct\right),
$$

where $u$ is the solution to (1.1) and $c$ is the ergodic constant of (1.6). Due to Theorem 2.2, this limit exists. As we have already emphasized many times, because of the multiplicity of solutions to (1.6), the asymptotic profile $v$ in Theorem 2.2 is completely decided through the initial data for $H$ fixed. In this section, we try to make clear how the asymptotic profile depends on the initial data, which is based on the argument by Davini, Siconolfi [25]. We use the following assumption

(H3)' $H : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$\begin{cases} \lim_{|p|\to\infty} \dfrac{H(x,p)}{|p|} = +\infty & \text{uniformly for } x \in \mathbb{T}^n, \\ D^2_{pp}H(x,p) \geq 0 & \text{for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n, \end{cases}$$

which is weaker than (H3).

Let $L : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ be the Legendre transform of $H$, that is,

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

The function $L$ is called the *Lagrangian* in the literature.

We first introduce the notion of the Aubry set.

**Definition 2.1.** *Let $c$ be the ergodic constant of* (1.6) *and set $L_c(x,v) := L(x,v)+c$ for any $(x,v) \in \mathbb{T}^n \times \mathbb{R}^n$. We call $y \in \mathbb{T}^n$ the element of the Aubry set $\mathcal{A}$ if the following*

$$\inf\left\{ \int_0^t L_c(\gamma(s), -\dot\gamma(s))\, ds \;:\; t \geq \delta, \gamma \in \mathrm{AC}\,([0,t], \mathbb{T}^n), \gamma(0) = \gamma(t) = y \right\} = 0 \quad (2.22)$$

*is satisfied for any fixed $\delta > 0$.*

Let us define the function $d_c : \mathbb{T}^n \times \mathbb{T}^n \to \mathbb{R}$ by

$$\begin{aligned} &d_c(x,y) \\ &\quad := \inf\left\{ \int_0^t L_c(\gamma(s), -\dot\gamma(s))\, ds \;:\; t > 0,\ \gamma \in \mathrm{AC}\,([0,t], \mathbb{T}^n), \gamma(0) = x, \gamma(t) = y \right\}. \end{aligned} \quad (2.23)$$

The function $d_c$ plays a role of a fundamental solution for Hamilton–Jacobi equations. We gather some basic properties of the function $d_c$.

**Proposition 2.13.** *Assume that* (H3)' *holds. We have*
(i) $d_c(x,y) = \sup\{v(x) - v(y) : v$ *is a subsolution of* (1.6)$\}$,
(ii) $d_c(x,x) = 0$ *and* $d_c(x,y) \leq d_c(x,z) + d_c(z,y)$ *for any* $x,y,z \in \mathbb{T}^n$,
(iii) $d_c(\cdot,y)$ *is a subsolution of* (1.6) *for all* $y \in \mathbb{T}^n$ *and a solution of* (1.6) *in* $\mathbb{T}^n \setminus \{y\}$ *for all* $y \in \mathbb{T}^n$.

*Proof.* We first prove

$$v(x) - v(y) \leq \int_0^t L_c(\gamma, -\dot\gamma)\, ds$$

for all $x,y \in \mathbb{T}^n$, any subsolution $v$ of (1.6), and $\gamma \in \mathrm{AC}\,([0,t], \mathbb{T}^n)$ with $\gamma(0) = x$ and $\gamma(t) = y$. This is at least formally easy to prove. Indeed, let $v$ be a smooth subsolution of (1.6). We have the following simple computations

$$\begin{aligned} v(x) - v(y) &= -\int_0^t \frac{dv(\gamma(s))}{ds}\, ds = \int_0^t Dv(\gamma(s)) \cdot (-\dot\gamma(s))\, ds \\ &\leq \int_0^t (L(\gamma(s), -\dot\gamma(s)) + c) + (H(\gamma(s), Dv(\gamma(s))) - c)\ ds \leq \int_0^t L_c(\gamma(s), -\dot\gamma(s))\, ds. \end{aligned}$$

This immediately implies (i). Due to the convexity assumption on $H$, we can obtain an approximated smooth subsolution by using mollification as in the proof of Proposition 1.5. By using this approximation, we can make this argument rigorous. We ask the interested readers to fulfill the details here.

It is straightforward to check that (i) implies (ii), and (iii) is a consequence of (i) and stability results of viscosity solutions (see Proposition 4.10 in Section 4.5) . $\square$

**Remark 2.6.** We can easily check that $y$ is in $\mathcal{A}$ if and only if (2.22) holds only for some $\delta_0 > 0$. Indeed, for any $\delta > 0$, we only need to consider the case where $\delta > \delta_0$. Fix $\varepsilon > 0$ and then there exist $t_\varepsilon \geq \delta_0$ and $\gamma_\varepsilon \in \mathrm{AC}\,([0,t_\varepsilon],\mathbb{T}^n)$ with $\gamma(0) = \gamma(t_\varepsilon) = y$ such that

$$0 = v(\gamma_\varepsilon(0)) - v(\gamma_\varepsilon(t_\varepsilon)) \leq \int_0^{t_\varepsilon} L_c(\gamma_\varepsilon(s), -\dot{\gamma}_\varepsilon(s))\, ds < \varepsilon.$$

We choose $m \in \mathbb{N}$ such that $mt_\varepsilon \geq \delta$ and set

$$\gamma_m(s) := \gamma(s - (j-1)t_\varepsilon) \quad \text{for } s \in [(j-1)t_\varepsilon, jt_\varepsilon],\ j = 1, \ldots, m.$$

Then $\gamma_m(0) = \gamma_m(mt_\varepsilon) = y$. We calculate that

$$\begin{aligned} 0 \leq & \int_0^{mt_\varepsilon} L_c(\gamma_m(s), -\dot{\gamma}_m(s))\, ds = \sum_{j=1}^m \int_{(j-1)t_\varepsilon}^{jt_\varepsilon} L_c(\gamma(s-(j-1)t_\varepsilon), -\dot{\gamma}(s-(j-1)t_\varepsilon))\, ds \\ = & \, m \int_0^{t_\varepsilon} L_c(\gamma(s), -\dot{\gamma}(s))\, ds < m\varepsilon. \end{aligned}$$

Sending $\varepsilon \to 0$ yields

$$\inf\left\{ \int_0^t L_c(\gamma(s), -\dot{\gamma}(s))\, ds \,:\, t \geq \delta, \gamma \in \mathrm{AC}\,([0,t],\mathbb{T}^n), \gamma(0) = \gamma(t) = y \right\} = 0$$

for any $\delta > 0$.

Fathi and Siconolfi in [35] gave a beautiful characterization of the Aubry set as follows.

**Theorem 2.14.** *Assume that* (H3)' *holds. A point $y \in \mathbb{T}^n$ is in the Aubry set $\mathcal{A}$ if and only if $d_c(\cdot, y)$ is a solution of* (1.6).

We refer to [35, Proposition 5.8] and [49, Proposition A.3] for the proofs.

**Theorem 2.15.** *Assume that* (H3)' *holds. Then the Aubry set $\mathcal{A}$ is nonempty, compact, and a uniqueness set of* (1.6)*, that is, if $v$ and $w$ are solutions of* (1.6)*, and $v = w$ on $\mathcal{A}$, then $v = w$ on $\mathbb{T}^n$.*

*Proof.* Let us first proceed to prove that $\mathcal{A}$ is a uniqueness set of (1.6). It is enough to show that if $v \leq w$ on $\mathcal{A}$, then $v \leq w$ on $\mathbb{T}^n$. For any small $\varepsilon > 0$, there exists an open set $U_\varepsilon$ such that $\mathcal{A} \subset U_\varepsilon$ with $\overline{\cap_{\varepsilon>0} U_\varepsilon} = \mathcal{A}$, and $v \leq w + \varepsilon$ in $U_\varepsilon$. Set $K_\varepsilon := \mathbb{T}^n \setminus U_\varepsilon$. Fix any $z \in K_\varepsilon$. Since $z \notin \mathcal{A}$, $d_c(\cdot, z)$ is not a supersolution at $x = z$ in light of

Proposition 2.13 (iii) and Theorem 2.14. Then, there exist a constant $r_z > 0$ and a function $\varphi_z \in C^1(\mathbb{T}^n)$ such that $B(z, r_z) \subset \mathbb{T}^n \setminus \mathcal{A}$,

$$\begin{aligned} &H(x, D\varphi_z(x)) < 0 \text{ for all } x \in B(z, r_z), \\ &\varphi_z(z) > 0 = d_c(z, z), \text{ and } \varphi_z(x) < d_c(x, z) \text{ for all } x \in \mathbb{T}^n \setminus B(z, r_z). \end{aligned}$$

See the proof of Theorem 4.19 for details. We set $\psi_z(x) = \max\{d_c(x, z), \varphi_z(x)\}$ for $x \in \mathbb{T}^n$ and observe that $\psi_z$ is a subsolution of (1.6) in light of Proposition 4.10, and that $H(x, D\psi_z(x)) < 0$ in a neighborhood $V_z$ of $z$ in the classical sense.

By the compactness of $K_\varepsilon$, there is a finite sequence of points $\{z_j\}_{j=1}^J$ such that $K_\varepsilon \subset \bigcup_{j=1}^J V_{z_j}$. We define the function $\psi \in C(\mathbb{T}^n)$ by $\psi(x) = (1/J) \sum_{j=1}^J \psi_{z_j}(x)$ and observe by convexity (H3)' that $\psi$ is a strict subsolution to (1.6) for some neighborhood $V$ of $K_\varepsilon$. Regularizing $\psi$ by mollification, if necessary, we may assume that $\psi \in C^1(V)$. Thus, we may apply the comparison result (see Theorem 4.8 in Section 4.4) to conclude that $v \le w + \varepsilon$ in $K_\varepsilon$. Sending $\varepsilon \to 0$ yields $v \le w$ in $\mathbb{T}^n \setminus \mathcal{A}$, which implies the conclusion.

To prove that $\mathcal{A} \neq \emptyset$, suppose that for all $y \in \mathbb{T}^n$, $d_c(\cdot, y)$ is not a solution to (1.6). By the above argument, for each $z \in \mathbb{T}^n$, $\psi_z$ is a subsolution of (1.6), and $H(x, D\psi_z(x)) < 0$ in a neighborhood $V_z$ of $z$ in the classical sense. By the compactness of $\mathbb{T}^n$, there is a finite sequence $\{y_i\}_{i=1}^N \subset \mathbb{T}^n$ such that $\mathbb{T}^n = \bigcup_{i=1}^N V_{y_i}$. We set $w(x) := (1/N) \sum_{i=1}^N \psi_{y_i}(x)$ for all $x \in \mathbb{T}^n$ and $\delta := (1/N) \min_{i=1,\ldots,N} \delta_i$. By the convexity of $H(x, \cdot)$, we have $H(x, Dw(x)) \le c - \delta$ in $\mathbb{T}^n$ in the viscosity sense, which contradicts the first formula of $c$ in Proposition 1.5.

The compactness of $\mathcal{A}$ is a straightforward result of stability of viscosity solutions (see Proposition 4.9 in Section 4.5). □

**Theorem 2.16.** *Assume that* (H2)–(H4) *hold. Let $u_c^\infty[u_0] - ct$ be the asymptotic solution for* (1.1)*, that is, $u_c^\infty[u_0](x) := \lim_{t\to\infty}(u(x, t) + ct)$, where $u$ is the solution to* (1.1)*. Then we have, for all $y \in \mathcal{A}$,*

$$\begin{aligned} u_c^\infty[u_0](y) &= \min\{d_c(y, z) + u_0(z) \,:\, z \in \mathbb{T}^n\} \\ &= \sup\{v(y) \,:\, v \textit{ is a subsolution to } (1.6) \textit{ with } v \le u_0 \textit{ in } \mathbb{T}^n\}. \end{aligned} \tag{2.24}$$

*Proof.* We write $v_{u_0}$ for the right hand side of (2.24). Let $y \in \mathcal{A}$ and choose $z_y \in \mathbb{T}^n$ so that

$$v_{u_0}(y) = d_c(y, z_y) + u_0(z_y).$$

By the definition of the function $d_c$, for any $\varepsilon > 0$, there exists $t_\varepsilon > 0$ and a curve $\xi_\varepsilon \in \mathrm{AC}([0, t_\varepsilon], \mathbb{T}^n)$ with $\xi_\varepsilon(0) = y, \xi_\varepsilon(t_\varepsilon) = z_y$ such that

$$d_c(y, z_y) > \int_0^{t_\varepsilon} L_c(\xi_\varepsilon, -\dot{\xi}_\varepsilon)\, ds - \varepsilon.$$

By the definition of the Aubry set, for any $n \in \mathbb{N}$, there exists a sequence $t_n \ge n$ and a curve $\delta_\varepsilon \in \mathrm{AC}([0, t_n], \mathbb{T}^n)$ such that $\delta_\varepsilon(0) = \delta_\varepsilon(t_n) = y$, and

$$\int_0^{t_n} L_c(\delta_\varepsilon(s), -\dot{\delta}_\varepsilon(s))\, ds < \epsilon.$$

Define $\gamma_\varepsilon \in \mathrm{AC}\,([0, t_n + t_\varepsilon], \mathbb{T}^n)$ by

$$\gamma_\varepsilon(s) = \begin{cases} \delta_\varepsilon(s) & \text{for} \quad s \in [0, t_n], \\ \xi_\varepsilon(s - t_n) & \text{for} \quad s \in [t_n, t_n + t_\varepsilon]. \end{cases}$$

Note that $\gamma_\varepsilon(0) = y$ and $\gamma_\varepsilon(t_n + t_\varepsilon) = z_y$.

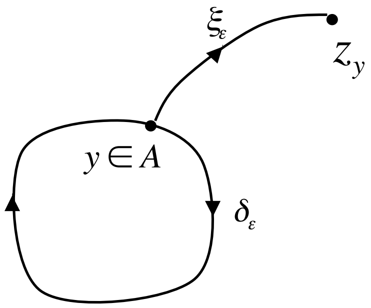


Figure 2.2

We observe that

$$\begin{aligned} v_{u_0}(y) &> \int_0^{t_\varepsilon} L_c(\xi_\varepsilon, -\dot{\xi}_\varepsilon)\, ds + u_0(z_y) - \varepsilon \\ &> \int_0^{t_n} L_c(\delta_\varepsilon(s), -\dot{\delta}_\varepsilon(s))\, ds + \int_0^{t_\varepsilon} L_c(\xi_\varepsilon, -\dot{\xi}_\varepsilon)\, ds + u_0(z_y) - 2\varepsilon \\ &= \int_0^{t_n + t_\varepsilon} L_c(\gamma_\varepsilon, -\dot{\gamma}_\varepsilon)\, ds + u_0(z_y) - 2\varepsilon \\ &\geq u_c(y, t_n + t_\varepsilon) - 2\varepsilon, \end{aligned}$$

where $u_c(x,t) := u(x,t) + ct$ for $(x,t) \in \mathbb{T}^n \times [0,\infty)$. Thus, sending $n \to \infty$ and $\varepsilon \to 0$ in this order yields $v_{u_0}(y) \geq u_c^\infty(y)$.

By the definition of $v_{u_0}$, we can easily check $v_{u_0} \leq u_0$ on $\mathbb{T}^n$ in view of Proposition 2.13 (ii). Note that $v_{u_0}$ is a subsolution to (1.6) in view of Proposition 2.13 (iii) and Corollary 4.16 (i). Thus, in light of the comparison principle for (1.1), we get $v_{u_0}(x) - ct \leq u(x,t)$ for all $(x,t) \in \mathbb{T}^n \times [0,\infty)$. Thus, $v_{u_0}(x) \leq \lim_{t\to\infty}(u(x,t) + ct) = u_c^\infty(x)$.

The second equality is a straightforward result of Proposition (2.13) with the observation $v_{u_0} \leq u_0$ on $\mathbb{T}^n$. □

In light of Proposition 2.13, Theorems 2.14, 2.15 and 2.16, we get the following representation formula for the asymptotic profile:

**Corollary 2.17.** *Assume that* (H2)–(H4) *holds. Let* $u_c^\infty[u_0] - ct$ *be the asymptotic solution for* (1.1). *Then we have the representation formula for the asymptotic profile*

$u_c^\infty[u_0]$ *as*

$$\begin{aligned} u_c^\infty[u_0](x) &= \min\{d_c(x,y) + v_{u_0}(y) \,:\, y \in \mathcal{A}\} \\ &= \inf\{v(x) \,:\, v \text{ is a solution to } (1.6) \text{ with } v \geq v_{u_0} \text{ in } \mathbb{T}^n\}, \end{aligned} \tag{2.25}$$

*where*

$$v_{u_0}(x) = \min\{d_c(x,z) + u_0(z) \,:\, z \in \mathbb{T}^n\} \quad \text{for all } x \in \mathbb{T}^n.$$

*Proof.* We denote by $w_{u_0}$ the right hand side in (2.25). Note first that this is a solution of (1.6) in view of Theorem 2.14 and Corollary 4.16. Moreover, we can check that

$$w_{u_0}(x) = \min\{d_c(x,y) + v_{u_0}(y) \,:\, y \in \mathcal{A}\} = v_{u_0}(x) \quad \text{for all } x \in \mathcal{A}$$

by Proposition 2.13 (i), (ii). Thus, $u_c^\infty[u_0] \equiv v_{u_0}$ on $\mathcal{A}$.

In light of a property of a uniqueness set of $\mathcal{A}$, Theorem 2.15, we obtain $u_c^\infty[u_0] \equiv v_{u_0}$ on $\mathbb{T}^n$, which is the conclusion. $\square$

**Example 2.2.** Now, let us consider the asymptotic profile for the Hamilton–Jacobi equation appearing in Example 4.1. As we observe in the beginning of Section 2.2, the associated ergodic problem is

$$|Dv| = \sqrt{\frac{c^2 - h(x)^2}{h(x)^2}} \quad \text{in } \mathbb{T}^n,$$

where

$$c := \max_{x \in \mathbb{T}^n} h(x).$$

We can easily check that we have the explicit formula for the Aubry set

$$\mathcal{A} := \{x \in \mathbb{T}^n \,:\, h(x) = \max_{\mathbb{T}^n} h\}$$

from the definition of the Aubry set. Also, we have

$$d_c(x,y) = \inf\left\{\int_0^t \sqrt{\frac{c^2}{h(\gamma(s))^2} - 1}\, ds \,:\, t > 0, |\dot\gamma| \leq 1, \gamma(0) = x, \gamma(t) = y\right\}.$$

From this, we somehow have a better understanding on how the asymptotic profile depends on the force term $h$ and the initial data $u_0$ through Corollary 2.17.

**Example 2.3.** We consider Example 2.2 in a more explicit setting which we discussed in Example 2.1. Let $n = 1$ and $h$ be the function given by (2.6). Our goal is to derive the asymptotic profiles by using the formula given in Corollary 2.17 for some given initial data $u_0$.

In this setting, we have $\mathcal{A} = \{1/4, 3/4\}$. Thus, letting $u_c^\infty[u_0] := \lim_{t\to\infty}(u(x,t) + ct)$, we obtain by Corollary 2.17,

$$u_c^\infty[u_0](x) = \min\left\{d_c\left(x, \frac{1}{4}\right) + v_{u_0}\left(\frac{1}{4}\right), d_c\left(x, \frac{3}{4}\right) + v_{u_0}\left(\frac{3}{4}\right)\right\}.$$

We are able to compute $d_c(\cdot, 1/4), d_c(\cdot, 3/4)$ explicitly as

$$d_c\left(x, \frac{1}{4}\right) = \begin{cases} \left(x - \frac{1}{4}\right)^2 & \text{for } 0 \le x \le \frac{1}{2}, \\ -\left(x - \frac{3}{4}\right)^2 + \frac{1}{8} & \text{for } \frac{1}{2} \le x \le 1, \end{cases}$$

$$d_c\left(x, \frac{3}{4}\right) = \begin{cases} -\left(x - \frac{1}{4}\right)^2 + \frac{1}{8} & \text{for } 0 \le x \le \frac{1}{2}, \\ \left(x - \frac{3}{4}\right)^2 & \text{for } \frac{1}{2} \le x \le 1. \end{cases}$$

Firstly, let $u_0 = u_0^1 \equiv 0$ as in Example 2.1. Then, we can check that $v_{u_0^1}(x) \equiv 0$. Thus, we conclude that $u_c^\infty[u_0^1]$ coincides with the one which we got in Example 2.1.

Next, let us consider another case where $u_0 = u_0^2$ and $u_0^2 : \mathbb{R} \to \mathbb{R}$ is 1-periodic,

$$u_0^2(x) = \begin{cases} -2\left(x - \frac{1}{4}\right)^2 + \frac{1}{8} & \text{for } 0 \le x \le \frac{1}{2}, \\ 0 & \text{for } \frac{1}{2} \le x \le 1. \end{cases}$$

Notice that $u_0^2$ is not a subsolution to (2.3). Thus, Proposition 2.1 does not hold.

In this case, we need to find $v_{u_0^2}$. In this setting, it is not hard to see that

$$v_{u_0^2}(x) = \begin{cases} -\left(x - \frac{1}{4}\right)^2 + \frac{1}{16} & \text{for } 0 \le x \le \frac{1}{2}, \\ 0 & \text{for } \frac{1}{2} \le x \le 1. \end{cases}$$

Thus, we derive that

$$u_c^\infty[u_0^2](x) = \min\left\{ d_c\left(x, \frac{1}{4}\right) + \frac{1}{16}, d_c\left(x, \frac{3}{4}\right) \right\}.$$

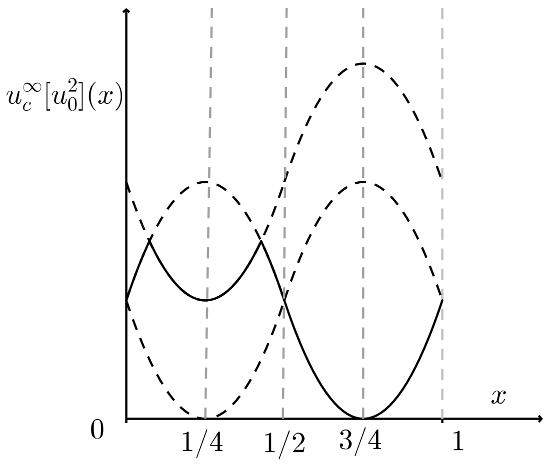


Figure 2.3: Graph of $u_c^\infty[u_0^2]$

.

## 2.6 Viscous case

In this subsection, we give a proof of the convergence result (1.5) for (1.2) based on the strong maximum principle, which is rather simple. It is worth emphasizing that this argument works only for uniformly parabolic equations including (1.2).

**Theorem 2.18.** *Assume that $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, (1.11) holds and $u_0 \in C^2(\mathbb{T}^n)$. Let $u$ be the solution of (1.2). Then there exists a solution $(v, c) \in C^2(\mathbb{T}^n) \times \mathbb{R}$ of (1.7) such that, as $t \to \infty$,*

$$u(x,t) - (v(x) - ct) \to 0 \quad \textit{uniformly on } \mathbb{T}^n.$$

The following proof is based on the argument in [13].

*Proof.* We normalize the ergodic constant $c$ to be 0 as usual.

Let $v$ be a solution of (1.7). By the maximum principle we see that

$$m(t) := \max_{x \in \mathbb{T}^n} (u(x,t) - v(x))$$

is nonincreasing. Therefore we have $m(t) \to \overline{m} \in \mathbb{R}$ as $t \to \infty$. By the global Lipschitz regularity result, Proposition 1.8, there exists a sequence $\{t_j\}_{j\in\mathbb{N}}$ with $t_j \to \infty$ such that

$$u(x, t + t_j) \to u_\infty(x,t) \quad \text{locally uniformly on } \mathbb{T}^n \times [0, \infty)$$

as $j \to \infty$ for some $u_\infty \in \mathrm{Lip}\,(\mathbb{T}^n \times [0,\infty))$ which may depend on the subsequence $\{t_j\}_{j\in\mathbb{N}}$ at this moment. By a standard stability result of viscosity solutions we see that $u_\infty$ is a solution to the equation in (1.1). Noting that $m(t+t_j) = \max_{x\in\mathbb{T}^n} (u(x, t+t_j) - v(x))$, we obtain

$$\overline{m} = \max_{x\in\mathbb{T}^n} (u_\infty(x,t) - v(x)) \quad \text{for all } t \geq 0.$$

By Proposition 2.19 (the strong maximum principle) below, we obtain

$$\overline{m} = u_\infty(x,t) - v(x) \quad \text{for all } (x,t) \in \mathbb{T}^n \times [0,\infty),$$

which implies that $u_\infty(x,t) \equiv u_\infty(x) = v(x) + \overline{m}$. As the right hand side above does not depend on the choice of subsequences anymore, we see that

$$u(x,t) \to v(x) + \overline{m} \quad \text{uniformly on } \mathbb{T}^n \text{ as } t \to \infty. \qquad \square$$

**Proposition 2.19** (Strong maximum principle)**.** *Let $U$ be a bounded domain in $\mathbb{T}^n$ and set $U_T := U \times (0, T]$ for each time $T > 0$. Let $u, v$ be a smooth subsolution and a smooth supersolution to (1.2), respectively. If, for some given $T > 0$, $u - v$ attains its maximum over $\overline{U}_T$ at a point $(x_0, t_0) \in U_T$, then*

$$u - v \textit{ is constant on } U_{t_0}.$$

See [31] for instance. If we do not have the regularity (smoothness) for solutions, then we need to be careful with the result in Proposition 2.19. A straightforward application of Proposition 2.19 is the uniqueness (up to additive constants) of solutions to (1.7). This uniqueness result is a crucial difference from that of the first-order Hamilton–Jacobi equation.

## 2.7 Some other directions and open questions

In this section, we present other developments in the study of large time behaviors of solutions to Hamilton–Jacobi equations or related ones very briefly.

(i) Unbounded domains: If we consider the Cauchy problem in unbounded domains (for instance, the whole space $\mathbb{R}^n$), then the behavior of the solution at infinity in $x$ may be quite complicated as it involves some compactness issues. Therefore, some compactness conditions are often required and the analysis along this direction is much more complicated. For this, there are several results: see Barles, Roquejoffre [11], Ishii [49], Ichihara, Ishii [45] for first-order Hamilton–Jacobi equations and Fujita, Ishii, Loreti [37], Ichihara [44], Ichihara, Sheu [46] for viscous Hamilton–Jacobi equations.

(ii) Boundary value problems: If we consider different types of optimal control problems (e.g., state constraint, exit-time problem, reflection problem, stopping time problem), then we need to consider several types of boundary value problems for Hamilton–Jacobi equations, which cause various kinds of difficulties. See Mitake [64], Barles, Ishii, Mitake [7] for state constraint problems, Mitake [65], Tchamba [77], Barles, Porretta, Tchamba [10] , Barles, Ishii, Mitake [7] for Dirichlet problems, Ishii [50], Barles, Mitake [9], Barles, Ishii, Mitake [7] for Neumann problems, and Mitake, Tran [70] for obstacle problems.

(iii) Weakly coupled systems: If we consider an optimal control problem which appears in the dynamic programming for the system whose states are governed by random changes (jumps), then we can naturally derive the weakly coupled system of Hamilton–Jacobi equations. See Cagnetti, Gomes, Mitake Tran [17], Mitake, Tran [68, 69], Camilli, Ley, Loreti, Nguyen [19], Nguyen [73], Davini, Zavidovique [27], Mitake, Siconolfi, Tran, Yamada [66] for developments on this direction. The profile of asymptotic limits is not solved yet.

(iv) Degenerate viscous Hamilton–Jacobi equations: In addition to the works [17, 70], we refer to Ley, Nguyen [58] for this direction. Also, not much is known about the limiting profiles.

(v) Time-periodic Hamilton–Jacobi equations: This is the case when $H = H(x, t, p)$ : $\mathbb{T}^n \times \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}$. There are only a few works in this direction. Only 1-dimensional case has been studied by Bernard, Roquejoffre [15]. See Jin, Yu [56] for an interesting application in the modeling of traffic flows. The multi-dimensional case ($n \geq 2$) is completely open. Note here that we do not have conservation of energy anymore as $H$ depends on $t$. It is also known that there is a time-periodic solution of the associated ergodic problem, whose minimum time period is 2. This fact makes the analysis of the large time behavior complicated.

(vi) Hamilton–Jacobi equations with mean curvature terms: This is an interesting topic, and many questions still remain open. See Cesaroni, Novaga [21] for a result along this line.

# Chapter 3

# Selection problems in the discounted approximation procedure

## 3.1 Selection problems

In this chapter, we consider the following ergodic problem

$$-a(x)\Delta u(x) + H(x, Du) = c \quad \text{in } \mathbb{T}^n, \tag{3.1}$$

where $(u, c) \in C(\mathbb{T}^n) \times \mathbb{R}$ so that $u$ solves the corresponding equation in the viscosity sense. This is a special case of (1.8). We will give precise assumptions on the Hamiltonian $H : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ and the nonnegative diffusion coefficient $a : \mathbb{T}^n \to [0, \infty)$ in Section 3.1.2.

We emphasize first that in general, solutions to ergodic problem (3.1) are not unique even up to additive constants. This can be seen via several examples below. Therefore, if we consider an approximation procedure for ergodic problem (3.1), then two natural questions appear:

(i) Does the whole family of approximate solutions converges?

(ii) If it converges, then which solution of the corresponding ergodic problem is the limit (which solution is selected)?

This type of questions is called a *selection problem* for ergodic problem (3.1).

### 3.1.1 Examples on nonuniqueness of ergodic problems

Let us give first two explicit examples for the inviscid case (ergodic problem (1.6)) to show the nonuniqueness issue.

**Example 3.1.** Let $n = 1$, $H(x, p) = |p|^2 - W(x)^2$, where $W : \mathbb{R} \to \mathbb{R}$ is 1-periodic, and $W(x) = 2\min\{|x - 1/4|, |x - 3/4|\}$ for all $x \in [0, 1]$. We identify the torus $\mathbb{T}$ as the interval $[0, 1]$ here.

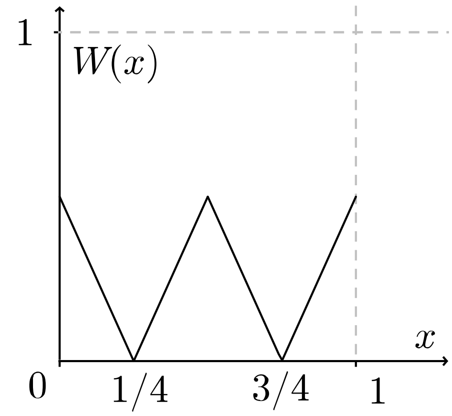


Figure 3.1: Graph of $W$ on $[0,1]$.

Then the ergodic constant $c = \max_{x\in\mathbb{T}}(-W(x)^2) = 0$, and the Aubry set

$$\mathcal{A} = \left\{ x \in [0,1] \,:\, -W(x)^2 = \max_{x\in\mathbb{T}}(-W(x)^2) \right\} = \left\{ \frac{1}{4}, \frac{3}{4} \right\}.$$

The ergodic problem becomes

$$|u'|^2 = W(x)^2 \quad \text{in } \mathbb{R}, \tag{3.2}$$

where $u$ is 1-periodic. For $x \in [0,1]$, set

$$u_1^b(x) := \begin{cases} \left(x - \frac{1}{4}\right)^2 & \text{for } 0 \le x \le \frac{1}{2}, \\ \min\left\{ -\left(x - \frac{3}{4}\right)^2 + \frac{1}{8}, \left(x - \frac{3}{4}\right)^2 + b \right\} & \text{for } \frac{1}{2} \le x \le 1, \end{cases}$$

$$u_2^b(x) := \begin{cases} \min\left\{ \left(x - \frac{1}{4}\right)^2 + b, -\left(x - \frac{1}{4}\right)^2 + \frac{1}{8} \right\} & \text{for } 0 \le x \le \frac{1}{2}, \\ \left(x - \frac{3}{4}\right)^2 & \text{for } \frac{1}{2} \le x \le 1, \end{cases}$$

and extend $u_1^b, u_2^b$ to $\mathbb{R}$ periodically. It is straightforward to check that all functions $u_1^b, u_2^b$ for any $b \in [0, 1/8]$ are solutions to (3.2).

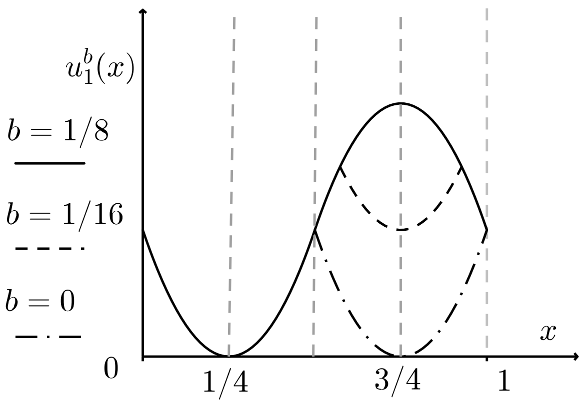

Figure 3.2: Graph of $u_1^b$ on $[0,1]$ for $b=0$, $b=1/16$, $b=1/8$

.

This shows that inviscid ergodic problem (3.2) has many solutions of different types, which confirms that solutions of (1.6) are not unique even up to additive constants in general. This example was known from the beginning of the theory of viscosity solutions (see [60, Proposition 5.4]).

**Example 3.2.** Let $H(x,p)=p\cdot(p-Df)$, where $f\in C^1(\mathbb{T}^n)$ is given. Then, we can easily see that

$$Du\cdot(Du-Df)=0 \quad \text{in } \mathbb{T}^n \tag{3.3}$$

has two solutions $u(x)\equiv C_1$ and $u(x)=f(x)+C_2$ for any $C_1,C_2\in\mathbb{R}$ fixed. Thus, $c=0$. Moreover, in view of the stability result for convex Hamilton–Jacobi equations (see Corollary 4.16 in Section 4.5), we see that

$$\tilde{u}(x)=\min\{C_1,f(x)+C_2\}, \qquad \text{for all } x\in\mathbb{T}^n,$$

are also solutions to ergodic problem (3.3) for all $C_1,C_2\in\mathbb{R}$.

We now give two examples on the nonuniqueness issue for the degenerate viscous case (ergodic problem (3.1)).

**Example 3.3.** Let $n=1$, $H(x,p)=|p|^2-V(x)$, where $V:\mathbb{R}\to\mathbb{R}$ is 1-periodic and

$$V(x)=\begin{cases} x^4\left(x-\dfrac{1}{4}\right)^4 & \text{for } 0\le x\le \dfrac{1}{4},\\ 0 & \text{for } \dfrac{1}{4}\le x\le\dfrac{3}{4},\\ \left(x-\dfrac{3}{4}\right)^4(x-1)^4 & \text{for } \dfrac{3}{4}\le x\le 1. \end{cases}$$

The diffusion $a:\mathbb{R}\to\mathbb{R}$ is an 1-periodic, $C^2$ function such that

$$\begin{cases} a(x)=0 & \text{for } 0\le x\le\dfrac{1}{4} \text{ and } \dfrac{3}{4}\le x\le 1,\\ a(x)\ge 0 & \text{for } \dfrac{1}{4}\le x\le\dfrac{3}{4}. \end{cases}$$

See Figure 3.3. Here, notice that we do not require much on the behavior of $a$ in $(1/4,3/4)$.

We identify the torus $\mathbb{T}$ as the interval $[0,1]$ here. The ergodic problem in this setting is

$$|u'|^2=V(x)+a(x)u''+c \qquad \text{in } \mathbb{R}, \tag{3.4}$$

where $u$ is 1-periodic.

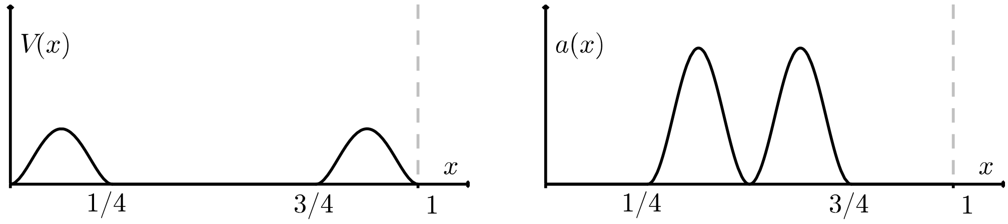


Figure 3.3: Graphs of $V$ and $a$

Set

$$\overline{u}(x) := \begin{cases} \int_0^x y^2 \left(y - \frac{1}{4}\right)^2 dy & \text{for } 0 \leq x \leq \frac{1}{4}, \\ \alpha & \text{for } \frac{1}{4} \leq x \leq \frac{3}{4}, \\ \alpha - \int_{3/4}^x \left(y - \frac{3}{4}\right)^2 (y-1)^2 \, dy & \text{for } \frac{3}{4} \leq x \leq 1, \end{cases}$$

where

$$\alpha := \int_0^{1/4} y^2 \left(y - \frac{1}{4}\right)^2 dy = \frac{1}{30720}. \tag{3.5}$$

Extend $\overline{u}$ in a periodic way to $\mathbb{R}$. It is not hard to see that $\pm\overline{u}$ are solutions to (3.4) with $c = 0$. Moreover, for $\beta \geq 0$, define

$$u_\beta(x) := \min\{\overline{u}(x), -\overline{u}(x) + \beta\} \quad \text{for } x \in \mathbb{R}. \tag{3.6}$$

Note that $u_\beta(x) = \min\{\alpha, -\alpha + \beta\}$ for $x \in [1/4, 3/4]$. By checking carefully, we see that $u_\beta$ is also a solution of (3.4) with $c = 0$.

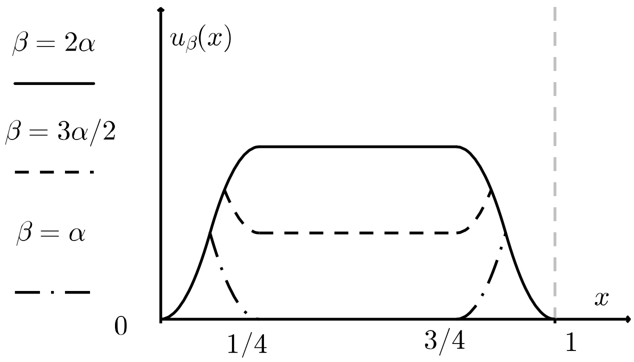


Figure 3.4: Graph of $u_\beta$ on $[0, 1]$ for $\beta = 2\alpha$, $\beta = 3\alpha/2$, $\beta = \alpha$

.

This again demonstrates that ergodic problem for degenerate viscous Hamilton–Jacobi equations (3.1) has many solutions of different types in general.

**Example 3.4.** Let $n = 1$, $H(x,p) = |p|^2 - V(x)$, where $V : \mathbb{R} \to \mathbb{R}$ is 1-periodic and

$$V(x) = \begin{cases} x^4 \left(x - \frac{1}{4}\right)^4 & \text{for } 0 \le x \le \frac{1}{4}, \\ \left(x - \frac{1}{4}\right)^4 \left(x - \frac{1}{2}\right)^4 & \text{for } \frac{1}{4} \le x \le \frac{1}{2}, \\ -\frac{32}{3}\left(x - \frac{1}{2}\right)^7 (x-1)^7 & \text{for } \frac{1}{2} \le x \le 1. \end{cases}$$

Here $a : \mathbb{R} \to \mathbb{R}$ is an 1-periodic, $C^2$ function such that

$$a(x) = \begin{cases} 0 & \text{for for } 0 \le x \le \frac{1}{2}, \\ \frac{4}{3}\left(x - \frac{1}{2}\right)^4 (x-1)^4 & \text{for } \frac{1}{2} \le x \le 1. \end{cases}$$

See Figure 3.5.

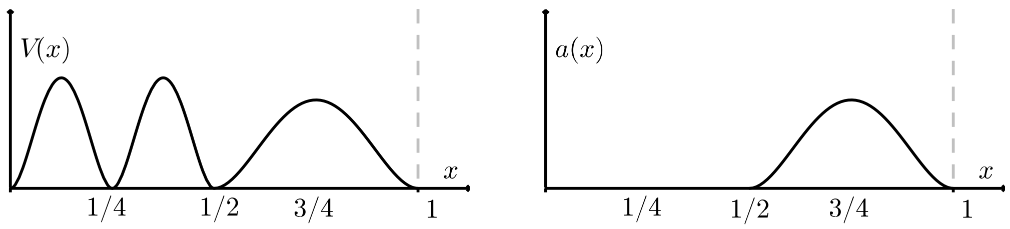


Figure 3.5: Graphs of $V$ and $a$

For $x \in [0, 1/2]$, set

$$\overline{u}(x) := \begin{cases} \int_0^x y^2 \left(y - \frac{1}{4}\right)^2 dy & \text{for } 0 \le x \le \frac{1}{4}, \\ \alpha - \int_{1/4}^x \left(y - \frac{1}{4}\right)^2 \left(y - \frac{1}{2}\right)^2 dy & \text{for } \frac{1}{4} \le x \le \frac{1}{2}, \end{cases}$$

where $\alpha$ is given by (3.5). For $\beta \ge 0$, define the functions $u_\beta$ by

$$u_\beta(x) := \begin{cases} \min\{\overline{u}(x), -\overline{u}(x) + \beta\} & \text{for } 0 \le x \le \frac{1}{2}, \\ \left(x - \frac{1}{2}\right)^4 (x-1)^4 & \text{for } \frac{1}{2} \le x \le 1. \end{cases} \tag{3.7}$$

Extend $u_\beta$ in a periodic way to $\mathbb{R}$. Then we can check that $u_\beta$ is a solution to (3.4) with $c = 0$ for each $\beta \ge 0$.

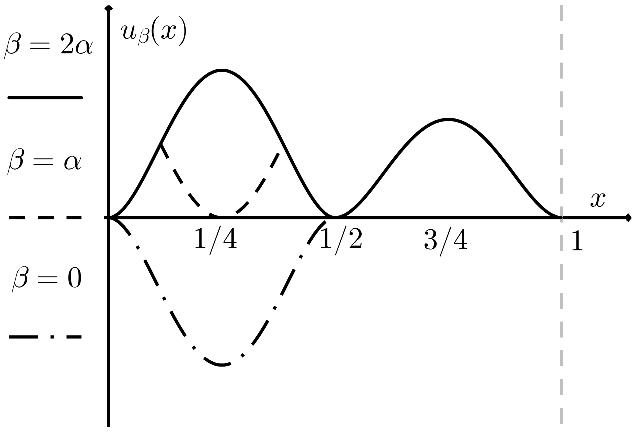


Figure 3.6: Graph of $u_\beta$ on $[0,1]$ for $\beta = 2\alpha$, $\beta = \alpha$, $\beta = 0$

.

### 3.1.2 Discounted approximation

There are several methods to construct or approximate viscosity solutions of ergodic problem (3.1). One natural way in terms of the well-posedness theory and a priori estimates is the following discounted approximation procedure. For $\varepsilon > 0$, consider

$$(\mathrm{D})_\varepsilon \qquad \varepsilon u^\varepsilon - a(x)\Delta u^\varepsilon + H(x, Du^\varepsilon) = 0 \qquad \text{in } \mathbb{T}^n.$$

Problem $(\mathrm{D})_\varepsilon$ is uniquely solvable as we see in Section 1.2 because of the term "$\varepsilon u^\varepsilon$", which is sometimes called a *discount factor* in optimal control theory.

The main assumptions in this chapter are:

(H5) $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, $p \mapsto H(x,p)$ is convex for each $x \in \mathbb{T}^n$, and there exists $C > 0$ so that

$$|D_x H(x,p)| \le C(1 + H(x,p)), \qquad \text{for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n,$$

$$\lim_{|p|\to+\infty} \frac{H(x,p)}{|p|} = +\infty, \qquad \text{uniformly for } x \in \mathbb{T}^n,$$

(H6) $a \ge 0$ in $\mathbb{T}^n$, and $a \in C^2(\mathbb{T}^n)$.

Note that the situation we consider is a special case of general degenerate viscous case (1.14). For the general case, see the discussion in Section 3.6.

Let us now repeat some of arguments in the proof of Theorem 1.3. Under assumptions (H5), (H6) (or some other appropriate growth conditions), we derive the following a priori estimate

$$\|\varepsilon u^\varepsilon\|_{L^\infty(\mathbb{T}^n)} + \|Du^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \le C \quad \text{for some } C > 0. \tag{3.8}$$

See the proof of Theorem 1.3. Once (3.8) is achieved, we obtain that

$$\{u^\varepsilon(\cdot) - u^\varepsilon(x_0)\}_{\varepsilon>0} \quad \text{is uniformly bounded and equi-Lipschitz continuous in } \mathbb{T}^n,$$

for some fixed $x_0 \in \mathbb{T}^n$. Therefore, in view of the Arzelà-Ascoli theorem, there exists a subsequence $\{\varepsilon_j\}_{j\in\mathbb{N}}$ with $\varepsilon_j \to 0$ as $j \to \infty$ such that

$$\varepsilon_j u^{\varepsilon_j} \to -c \in \mathbb{R}, \quad u^{\varepsilon_j} - u^{\varepsilon_j}(x_0) \to u \in C(\mathbb{T}^n) \quad \text{uniformly in } \mathbb{T}^n \text{ as } j \to \infty, \tag{3.9}$$

where $(u, c)$ is a solution of ergodic problem (3.1).

Moreover, Proposition 1.6 gives

$$\left\{u^\varepsilon + \frac{c}{\varepsilon}\right\}_{\varepsilon>0} \quad \text{is uniformly bounded and equi-Lipschitz continuous in } \mathbb{T}^n.$$

Thus, by using the Arzelà-Ascoli theorem again, there exists a subsequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ with $\varepsilon_k \to 0$ as $k \to \infty$ such that

$$u^{\varepsilon_k} + \frac{c}{\varepsilon_k} \to \tilde{u} \in C(\mathbb{T}^n) \quad \text{uniformly in } \mathbb{T}^n \text{ as } k \to \infty, \tag{3.10}$$

where $(\tilde{u}, c)$ is a solution of ergodic problem (3.1).

Up to now, in convergence (3.10), we only use a soft approach mainly based on tools from functional analysis. As explained in Introduction of this chapter, our main question in this chapter is the selection problem concerning $(\mathrm{D})_\varepsilon$, that is, whether convergence (3.10) holds for the whole sequence $\varepsilon \to 0$ or not.

This problem was proposed by Lions, Papanicolaou, Varadhan [61] (see also Bardi, Capuzzo-Dolcetta [5, Remark 1.2, page 400]). It remained unsolved for almost 30 years. Recently, there was substantial progress in the case of convex Hamiltonians. First, a partial characterization of the possible limits was given by Gomes [41] in terms of the Mather measures. Iturriaga and Sanchez-Morgado [53] then studied this under rather restricted assumptions. Davini, Fathi, Iturriaga, Zavidovique [26] and Mitake, Tran [71] gave a positive answer for this question in case $a \equiv 0$ and $a \geq 0$, respectively, by using a dynamical approach and the nonlinear adjoint method. These approaches are based on the weak KAM theory. By characterizing the limit in terms of Mather measures, the convergence for the whole sequence is proven. A selection problem for Neumann boundary problems conditions was examined by Al-Aidarous, Alzahrani, Ishii, Younas [1].

We state here the main result in this chapter.

**Theorem 3.1.** *Assume* (H5), (H6) *hold. For each $\varepsilon > 0$, let $u^\varepsilon$ be the solution to* $(\mathrm{D})_\varepsilon$. *Then, we have that, as $\varepsilon \to 0$,*

$$u^\varepsilon(x) + \frac{c}{\varepsilon} \to u^0(x) := \sup_{\phi\in\mathcal{E}} \phi(x) \quad \textit{uniformly for } x \in \mathbb{T}^n, \tag{3.11}$$

*where we denote by $\mathcal{E}$ the family of solutions $u$ of* (3.1) *satisfying*

$$\iint_{\mathbb{T}^n\times\mathbb{R}^n} u\, d\mu \leq 0 \qquad \textit{for all } \mu \in \mathcal{M}. \tag{3.12}$$

*The set $\mathcal{M}$, which is a family of probability measures on $\mathbb{T}^n \times \mathbb{R}^n$, is defined in Section* 3.2.1.

**Remark 3.1.** (i) It is worth emphasizing that all of the above results strongly require the convexity of the Hamiltonians. On the other hand, to obtain a priori estimate (3.8), we only need the superlinearity of $H$, and in particular, we do NOT need the convexity assumption. Thus, the question whether convergence of $u^\varepsilon + c/\varepsilon$ as $\varepsilon \to 0$ without the convexity assumption holds or not remains. Indeed, selection problems for Hamilton–Jacobi equations with nonconvex Hamiltonians remain rather open. See Section 3.6 for some further discussions on more recent developments.
(ii) Note also that, in the above theorem, the first-order case and the second-order case are quite different because of the appearance of the diffusion term, which is delicate to be handled. In particular, $\mathcal{E}$ is a family of solutions of (3.1) (not just subsolutions), which is different from that of [26]. We will address this matter clearly later.

## 3.2 Regularizing process, stochastic Mather measures and key estimates

Hereinafter, we assume that the ergodic constant $c$ of (3.1) is zero. Indeed, by replacing, if necessary, $H$ and $u^\varepsilon$ by $H - c$ and $u^\varepsilon + c/\varepsilon$, respectively, we can always reduce the situation to the case that $c = 0$.

Thus, the ergodic problem is

$$\text{(E)} \qquad H(x, Du) = a(x)\Delta u \qquad \text{in } \mathbb{T}^n.$$

Since $u^\varepsilon, u$ are not smooth in general, in order to perform our analysis, we need a regularizing process as in the previous chapter.

### 3.2.1 Regularizing process and construction of $\mathcal{M}$

We denote by $\mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ the set of Radon probability measures on $\mathbb{T}^n \times \mathbb{R}^n$. Let the function $L : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ be the Legendre transform of $H$. It is worth recalling the formula of $L$

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

By (H5), $L$ is well-defined, that is, $L(x, v)$ is finite for each $(x, v) \in \mathbb{T}^n \times \mathbb{R}^n$. Furthermore, $L$ is of class $C^1$, convex with respect to $v$, and superlinear.

For each $\varepsilon, \eta > 0$, we study

$$\text{(A)}_\varepsilon^\eta \qquad \varepsilon u^{\varepsilon,\eta} + H(x, Du^{\varepsilon,\eta}) = (a(x) + \eta^2)\Delta u^{\varepsilon,\eta} \qquad \text{in } \mathbb{T}^n,$$

which is an approximation of $\text{(D)}_\varepsilon$. Due to the appearance of viscosity term $\eta^2 \Delta u^{\varepsilon,\eta}$, $\text{(A)}_\varepsilon^\eta$ has a (unique) smooth solution $u^{\varepsilon,\eta}$. The following result on the rate of convergence of $u^{\varepsilon,\eta}$ to $u^\varepsilon$ as $\eta \to 0$ is standard. It is of the same flavor as that of Proposition 2.3, and we omit its proof.

**Lemma 3.2.** *Assume* (H5), (H6). *Then there exists a constant $C > 0$ independent of $\varepsilon$ and $\eta$ so that*

$$\|Du^{\varepsilon,\eta}\|_{L^\infty(\mathbb{T}^n)} \le C,$$
$$\|u^{\varepsilon,\eta} - u^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \le \frac{C\eta}{\varepsilon}.$$

It is time to use the nonlinear adjoint method to construct the set $\mathcal{M} \subset \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ in Theorem 3.1. For $x_0 \in \mathbb{T}^n$ fixed, we consider an adjoint equation of the linearized operator of $(\mathrm{A})_\varepsilon^\eta$:

$$(\mathrm{AJ})_\varepsilon^\eta \qquad \varepsilon\theta^{\varepsilon,\eta} - \mathrm{div}(D_pH(x, Du^{\varepsilon,\eta})\theta^{\varepsilon,\eta}) = \Delta(a(x)\theta^{\varepsilon,\eta}) + \eta^2\Delta\theta^{\varepsilon,\eta} + \varepsilon\delta_{x_0} \qquad \text{in } \mathbb{T}^n,$$

where $\delta_{x_0}$ denotes the Dirac delta measure at $x_0$. By the maximum principle and integrating $(\mathrm{AJ})_\varepsilon^\eta$ on $\mathbb{T}^n$, we obtain

$$\theta^{\varepsilon,\eta} > 0 \text{ in } \mathbb{T}^n \setminus \{x_0\}, \quad \text{and} \quad \int_{\mathbb{T}^n} \theta^{\varepsilon,\eta}(x)\,dx = 1.$$

In light of the Riesz theorem, for every $\varepsilon, \eta > 0$, there exists a probability measure $\nu^{\varepsilon,\eta} \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ satisfying

$$\int_{\mathbb{T}^n} \psi(x, Du^{\varepsilon,\eta})\theta^{\varepsilon,\eta}(x)\,dx = \iint_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\nu^{\varepsilon,\eta}(x,p) \tag{3.13}$$

for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$. It is clear that $\mathrm{supp}\,(\nu^{\varepsilon,\eta}) \subset \mathbb{T}^n \times \overline{B}(0,C)$ for some $C > 0$ due to Lemma 3.2. Since

$$\iint_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\nu^{\varepsilon,\eta}(x,p) = 1 \quad \text{for all } \varepsilon > 0, \eta > 0,$$

due to the compactness of weak convergence of measures, there exist two subsequences $\eta_k \to 0$ and $\varepsilon_j \to 0$ as $k \to \infty$, $j \to \infty$, respectively, and probability measures $\nu^{\varepsilon_j}, \nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ (see [30, Theorem 4] for instance) so that

$$\begin{aligned} &\nu^{\varepsilon_j,\eta_k} \rightharpoonup \nu^{\varepsilon_j} && \text{as } k \to \infty, \\ &\nu^{\varepsilon_j} \rightharpoonup \nu && \text{as } j \to \infty, \end{aligned} \tag{3.14}$$

in the sense of measures. We also have that $\mathrm{supp}\,(\nu^{\varepsilon_j})$, $\mathrm{supp}\,(\nu) \subset \mathbb{T}^n \times \overline{B}(0,C)$. For each such $\nu$, set $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ so that the pushforward measure of $\mu$ associated with $\Phi(x,v) = (x, D_vL(x,v))$ is $\nu$, that is, for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$,

$$\iint_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\nu(x,p) = \iint_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x, D_vL(x,v))\,d\mu(x,v). \tag{3.15}$$

We denote the pushforward measure of $\mu$ by $\Phi_\#\mu$.

Notice that the measure $\mu$ constructed by the above process depends on the choice of $x_0, \{\eta_k\}_k, \{\varepsilon_j\}_j$, and when needed, we write $\mu = \mu(x_0, \{\eta_k\}_k, \{\varepsilon_j\}_j)$ to demonstrate the clear dependence. In general, there could be many such limit $\mu$ for different choices of $x_0$, $\{\eta_k\}_k$ or $\{\varepsilon_j\}_j$. We define the set $\mathcal{M} \subset \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ by

$$\mathcal{M} := \bigcup_{x_0 \in \mathbb{T}^n, \{\eta_k\}_k, \{\varepsilon_j\}_j} \mu(x_0, \{\eta_k\}_k, \{\varepsilon_j\}_j).$$

The following simple proposition records important properties of $\nu$ and $\mu$.

**Proposition 3.3.** *Assume that* (H5), (H6) *hold and the ergodic constant of* (3.1) *is* 0. *Let $\nu$ and $\mu$ be probability measures given by* (3.14) *and* (3.15). *Then,*

(i) $$\iint_{\mathbb{T}^n \times \mathbb{R}^n} (D_pH(x,p) \cdot p - H(x,p))\, d\nu(x,p) = \iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\mu(x,v) = 0,$$

(ii) $$\begin{aligned} &\iint_{\mathbb{T}^n \times \mathbb{R}^n} (D_pH(x,p) \cdot D\varphi - a(x)\Delta\varphi)\, d\nu(x,p) \\ &= \iint_{\mathbb{T}^n \times \mathbb{R}^n} (v \cdot D\varphi - a(x)\Delta\varphi)\, d\mu(x,v) = 0 \quad \textit{for any } \varphi \in C^2(\mathbb{T}^n). \end{aligned}$$

*Proof.* Equation $(\mathrm{A})_\varepsilon^\eta$ can be rewritten as

$$\begin{aligned} &\varepsilon u^{\varepsilon,\eta} + D_pH(x, Du^{\varepsilon,\eta}) \cdot Du^{\varepsilon,\eta} - (a(x) + \eta^2)\Delta u^{\varepsilon,\eta} \\ &= D_pH(x, Du^{\varepsilon,\eta}) \cdot Du^{\varepsilon,\eta} - H(x, Du^{\varepsilon,\eta}). \end{aligned}$$

Multiply this by $\theta^{\varepsilon,\eta}$ and integrate on $\mathbb{T}^n$ to get

$$\begin{aligned} &\int_{\mathbb{T}^n} \left(\varepsilon u^{\varepsilon,\eta} + D_pH(x, Du^{\varepsilon,\eta}) \cdot Du^{\varepsilon,\eta} - (a(x) + \eta^2)\Delta u^{\varepsilon,\eta}\right) \theta^{\varepsilon,\eta}\, dx \\ &= \int_{\mathbb{T}^n} \left(\varepsilon\theta^{\varepsilon,\eta} - \mathrm{div}\,(D_pH(x, Du^{\varepsilon,\eta})\theta^{\varepsilon,\eta}) - \Delta\left((a(x) + \eta^2)\theta^{\varepsilon,\eta}\right)\right) u^{\varepsilon,\eta}\, dx \\ &= \int_{\mathbb{T}^n} \varepsilon\delta_{x_0} u^{\varepsilon,\eta}\, dx = \varepsilon u^{\varepsilon,\eta}(x_0). \end{aligned}$$

Moreover,

$$\begin{aligned} &\int_{\mathbb{T}^n} (D_pH(x, Du^{\varepsilon,\eta}) \cdot Du^{\varepsilon,\eta} - H(x, Du^{\varepsilon,\eta}))\, \theta^{\varepsilon,\eta}\, dx \\ &= \iint_{\mathbb{T}^n \times \mathbb{R}^n} (D_pH(x,p) \cdot p - H(x,p))\, d\nu^{\varepsilon,\eta}(x,p). \end{aligned}$$

Set $\eta = \eta_k$, $\varepsilon = \varepsilon_j$, and let $k \to \infty$, $j \to \infty$ in this order to yield

$$\begin{aligned} 0 &= \iint_{\mathbb{T}^n \times \mathbb{R}^n} (D_pH(x,p) \cdot p - H(x,p))\, d\nu(x,p) \\ &= \iint_{\mathbb{T}^n \times \mathbb{R}^n} (D_pH(x, D_vL(x,v)) \cdot D_vL(x,v) - H(x, D_vL(x,v)))\, d\mu(x,v) \\ &= \iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\mu(x,v), \end{aligned}$$

by (3.15) and the duality of convex functions. Note in the above computation that we have $\lim_{j\to\infty} \varepsilon_j u^{\varepsilon_j}(x_0) = 0$ because of the assumption that $c = 0$.

We now proceed to prove the second part. Fix $\varphi \in C^2(\mathbb{T}^n)$. Multiply $(\mathrm{AJ})^\eta_\varepsilon$ by $\varphi$ and integrate on $\mathbb{T}^n$ to get

$$\begin{aligned}&\int_{\mathbb{T}^n} (D_pH(x, Du^{\varepsilon,\eta}) \cdot D\varphi - a(x)\Delta\varphi)\, \theta^{\varepsilon,\eta}\, dx \\ &= \eta^2 \int_{\mathbb{T}^n} \Delta\varphi\theta^{\varepsilon,\eta}\, dx + \varepsilon\varphi(x_0) - \varepsilon \int_{\mathbb{T}^n} \varphi\theta^{\varepsilon,\eta}\, dx.\end{aligned}$$

We use (3.13) for $\varepsilon = \varepsilon_j$, $\eta = \eta_k$, and let $k \to \infty$ to obtain

$$\begin{aligned}&\iint_{\mathbb{T}^n\times\mathbb{R}^n} (D_pH(x, p) \cdot D\varphi - a(x)\Delta\varphi)\, d\nu^{\varepsilon_j}(x, p) \\ &= \varepsilon_j\varphi(x_0) - \varepsilon_j \iint_{\mathbb{T}^n\times\mathbb{R}^n} \varphi(x)\, d\nu^{\varepsilon_j}(x, p).\end{aligned}$$

Finally, let $j \to \infty$ to complete the proof. □

**Remark 3.2.** It is worth emphasizing a delicate issue that we cannot replace $C^2$ test functions by $C^{1,1}$ test functions in item (ii) of Proposition 3.3. This is because each measure $\mu \in \mathcal{M}$ can be quite singular and it can see the jumps of $\Delta\varphi$ in case $\varphi$ is $C^{1,1}$ but not $C^2$. This issue actually complicates our analysis later on as we need to build $C^2$-approximated subsolutions of (E), which is not quite standard in the theory of viscosity solutions to second-order degenerate elliptic or parabolic equations. This point will be addressed in Section 3.4.

Properties (i), (ii) in Proposition 3.3 of measure $\mu$ are essential ones to characterize a stochastic Mather measure, which will be defined in the following section. This idea was first discovered by Mañé [62], who relaxed the original idea of Mather [63]. See Fathi [34], Cagnetti, Gomes and Tran [18, Theorem 1.3] for some discussions on this.

### 3.2.2 Stochastic Mather measures

We are concerned with the following minimization problem

$$\min_{\mu\in\mathcal{F}} \iint_{\mathbb{T}^n\times\mathbb{R}^n} L(x, v)\, d\mu(x, v), \tag{3.16}$$

where

$$\mathcal{F} := \left\{ \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \,:\, \iint_{\mathbb{T}^n\times\mathbb{R}^n} (v \cdot D\phi - a(x)\Delta\phi)\, d\mu(x, v) = 0 \quad \text{for all } \phi \in C^2(\mathbb{T}^n) \right\}.$$

Measures belonging to $\mathcal{F}$ are called *holonomic measures* or *closing measures* associated with (3.1). By (ii) of Proposition 3.3, $\mathcal{M} \subset \mathcal{F}$.

**Definition 3.1.** *We let* $\widetilde{\mathcal{M}}$ *to be the set of all minimizers of* (3.16). *Each measure in* $\widetilde{\mathcal{M}}$ *is called a stochastic Mather measure.*

When $a \equiv 0$, holonomic condition is equivalent to the invariance condition under the *Euler–Lagrange flow*

$$\frac{d}{ds} D_v L(\gamma(s), \dot{\gamma}(s)) = D_x L(\gamma(s), \dot{\gamma}(s)).$$

This idea was first discovered by Mañé [62], who relaxed the original idea of Mather [63]. Minimizers of the minimizing problem (3.16) are precisely Mather measures for first-order Hamilton–Jacobi equations.

When $a \equiv 1$, this coincides with the definition of stochastic Mather measures for viscous Hamilton–Jacobi equations given by Gomes [39]. This means that this definition is quite natural for the current degenerate viscous case, and it covers both the first-order and the viscous case. Gomes [40, 41] also introduced the notion of generalized Mather measures by using the duality principle.

**Proposition 3.4.** *Fix $\mu \in \mathcal{M}$. Then $\mu$ is a minimizer of* (3.16)*.*

This proposition clearly implies that $\mathcal{M} \subset \widetilde{\mathcal{M}}$.

**Lemma 3.5.** *Assume that* (H5), (H6) *hold and the ergodic constant of* (3.1) *is* 0*. We have*

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x, v)\, d\mu(x, v) \geq 0 \quad \textit{for all } \mu \in \mathcal{F}. \tag{3.17}$$

*Furthermore,*

$$\min_{\mu \in \mathcal{F}} \iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x, v)\, d\mu(x, v) = 0.$$

Since a solution $w$ of ergodic problem (E) is not smooth in general, in order to use the admissible condition in $\mathcal{F}$, we need to find a family of smooth approximations of $w$, which are approximate subsolutions to (E). A natural way to perform this task is to use the usual convolution technique. More precisely, for each $\eta > 0$, let

$$w^\eta(x) := \gamma^\eta * w(x) = \int_{\mathbb{R}^n} \gamma^\eta(y) w(x + y)\, dy, \tag{3.18}$$

where $\gamma^\eta(y) = \eta^{-n}\gamma(\eta^{-1}y)$ (here $\gamma \in C_c^\infty(\mathbb{R}^n)$ is a standard symmetric mollifier such that $\gamma \geq 0$, $\operatorname{supp} \gamma \subset \overline{B}(0, 1)$ and $\|\gamma\|_{L^1(\mathbb{R}^n)} = 1$).

In the first-order case, it is quite simple to show that $\{w^\eta\}_{\eta>0}$ indeed are approximate subsolutions to (E) (see the second part of Proposition 1.5). In the current degenerate viscous setting, it is much more complicated because of the appearance of the possibly degenerate viscous term $a(x)\Delta w$. To prove that $\{w^\eta\}_{\eta>0}$ are approximate subsolutions to (E), we need to be able to control the commutation term

$$\gamma^\eta * (a\Delta w) - a(\gamma^\eta * \Delta w).$$

We give a commutation lemma, which itself is interesting and important:

**Lemma 3.6** (A commutation lemma)**.** *Assume* (H5), (H6) *hold. Assume that $w$ is a viscosity solution of* (E) *and $w^\eta$ be the function defined by* (3.18) *for $\eta > 0$. There exists a constant $C > 0$ and a continuous function $S^\eta : \mathbb{T}^n \to \mathbb{R}$ such that*

$$|S^\eta(x)| \le C \quad \textit{and} \quad \lim_{\eta \to 0} S^\eta(x) = 0, \quad \textit{for each } x \in \mathbb{T}^n, \tag{3.19}$$

*and*

$$H(x, Dw^\eta) \le a(x)\Delta w^\eta + S^\eta(x) \qquad \textit{in } \mathbb{T}^n.$$

Moreover, we can actually show that $S^\eta$ converges to 0 uniformly on $\mathbb{T}^n$ with convergence rate $\eta^{1/2}$, which is necessary in the proof of Theorem 3.1.

**Lemma 3.7** (Uniform convergence)**.** *Assume* (H5), (H6) *hold. Then there exists a universal constant $C > 0$ such that $\|S^\eta\|_{L^\infty(\mathbb{T}^n)} \le C\eta^{1/2}$.*

The proofs of Lemmas 3.6 and 3.7 are postponed to Section 3.4. By using the commutation lemma, Lemma 3.6, we give a proof of Lemma 3.5.

*Proof of Lemma* 3.5 *and Proposition* 3.4. Let $w$ be a solution of ergodic problem (E). By Lemma 3.6, we have that

$$H(x, Dw^\eta) \le a(x)\Delta w^\eta + S^\eta(x) \quad \text{in } \mathbb{T}^n,$$

where $S^\eta$ is an error term and we have a good control (3.19).

For any $\mu \in \mathcal{F}$, one has

$$\begin{aligned}
\iint_{\mathbb{T}^n \times \mathbb{R}^n} S^\eta(x)\, d\mu(x,v) &\ge \iint_{\mathbb{T}^n \times \mathbb{R}^n} (H(x, Dw^\eta) - a(x)\Delta w^\eta)\, d\mu(x,v) \\
&\ge \iint_{\mathbb{T}^n \times \mathbb{R}^n} (-L(x,v) + (v \cdot Dw^\eta - a(x)\Delta w^\eta))\ d\mu(x,v) \\
&= -\iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\mu(x,v),
\end{aligned}$$

where we use the admissible condition of $\mu \in \mathcal{F}$ to go from the second line to the last line. Thanks to (3.19), we let $\eta \to 0$ and use the Lebesgue dominated convergence theorem to deduce that

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\mu(x,v) \ge 0.$$

Thus, item (i) in Proposition 3.3 confirms that any measure $\mu \in \mathcal{M}$ minimizes the action (3.16). This is equivalent to the fact that $\mathcal{M} \subset \widetilde{\mathcal{M}}$. □

**Remark 3.3.** In general, if we do not assume that $c = 0$, then

$$\min_{\mu \in \mathcal{F}} \iint_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v)\, d\mu(x,v) = -c.$$

### 3.2.3 Key estimates

In this section, we give two important estimates.

**Lemma 3.8.** *Assume that* (H5), (H6) *hold and the ergodic constant of* (3.1) *is* 0*. Let* $w \in C(\mathbb{T}^n)$ *be any solution of* (E)*, and, for* $\varepsilon, \eta > 0$*,* $w^\eta$ *and* $\theta^{\varepsilon,\eta}$ *be, respectively, the function given by* (3.18) *and the solution to* $(\mathrm{AJ})^\eta_\varepsilon$ *for some* $x_0 \in \mathbb{T}^n$*.*

*Then,*

$$u^{\varepsilon,\eta}(x_0) \geq w^\eta(x_0) - \int_{\mathbb{T}^n} w^\eta \theta^{\varepsilon,\eta}\, dx - \frac{C\eta}{\varepsilon} - \frac{1}{\varepsilon}\int_{\mathbb{T}^n} S^\eta \theta^{\varepsilon,\eta}\, dx, \tag{3.20}$$

*where* $S^\eta(x)$ *is the function given in Lemma* 3.6*.*

*Proof.* We first calculate, for every $x \in \mathbb{T}^n$,

$$\begin{aligned}
|\Delta w^\eta(x)| &\leq \int_{\mathbb{R}^n} |D\gamma^\eta(y) \cdot Dw(x+y)|\, dy \\
&\leq \frac{C}{\eta^{n+1}} \int_{\mathbb{R}^n} |D\gamma(\eta^{-1}y)|\, dy = \frac{C}{\eta}\int_{\mathbb{R}^n} |D\gamma(z)|\, dz \leq \frac{C}{\eta},
\end{aligned}$$

which immediately implies $\eta^2|\Delta w^\eta| \leq C\eta$. Combing this with Lemma 3.6, we see that $w^\eta$ satisfies

$$H(x, Dw^\eta) \leq (a(x) + \eta^2)\Delta w^\eta + C\eta + S^\eta(x) \qquad \text{in } \mathbb{T}^n.$$

Subtract $(\mathrm{A})^\eta_\varepsilon$ from the above inequality to yield

$$\begin{aligned}
&\varepsilon w^\eta + C\eta + S^\eta(x) \\
\geq\, &\varepsilon(w^\eta - u^{\varepsilon,\eta}) + H(x, Dw^\eta) - H(x, Du^{\varepsilon,\eta}) - (a(x)+\eta^2)\Delta(w^\eta - u^{\varepsilon,\eta}) \\
\geq\, &\varepsilon(w^\eta - u^{\varepsilon,\eta}) + D_pH(x, Du^{\varepsilon,\eta}) \cdot D(w^\eta - u^{\varepsilon,\eta}) - (a(x)+\eta^2)\Delta(w^\eta - u^{\varepsilon,\eta}),
\end{aligned}$$

where we use the convexity of $H$ in the last inequality.

Then, multiplying this by $\theta^{\varepsilon,\eta}$, integrating on $\mathbb{T}^n$, and using the integration by parts, we get

$$\begin{aligned}
&\int_{\mathbb{T}^n} (\varepsilon w^{\varepsilon,\eta} + C\eta + S^\eta(x))\theta^{\varepsilon,\eta}\, dx \\
\geq\, &\varepsilon \int_{\mathbb{T}^n} (w^\eta - u^{\varepsilon,\eta})\theta^\varepsilon\, dx + \int_{\mathbb{T}^n} \left(D_pH(x, Du^{\varepsilon,\eta}) \cdot D(w^\eta - u^{\varepsilon,\eta}) - a(x)\Delta(w^\eta - u^{\varepsilon,\eta})\right)\theta^{\varepsilon,\eta}\, dx \\
=\, &\varepsilon \int_{\mathbb{T}^n} (w^\eta - u^{\varepsilon,\eta})\theta^{\varepsilon,\eta}\, dx - \int_{\mathbb{T}^n} \left(\operatorname{div}(D_pH(x, Du^{\varepsilon,\eta})\theta^{\varepsilon,\eta}) + \Delta(a(x)\theta^{\varepsilon,\eta})\right)(w^\eta - u^{\varepsilon,\eta})\, dx \\
=\, &\varepsilon \int_{\mathbb{T}^n} (w^\eta - u^{\varepsilon,\eta})\theta^{\varepsilon,\eta}\, dx - \int_{\mathbb{T}^n} (\varepsilon\theta^{\varepsilon,\eta} - \varepsilon\delta_{x_0})(w^\eta - u^{\varepsilon,\eta})\, dx \\
=\, &\varepsilon(w^\eta - u^{\varepsilon,\eta})(x_0)
\end{aligned}$$

which, after a rearrangement, implies (3.20). □

**Proposition 3.9.** *Assume that* (H5), (H6) *hold and the ergodic constant of* (3.1) *is* 0. *Let* $u^\varepsilon$ *be the solution of* $(\mathrm{E})_\varepsilon$*, and* $\mu \in \mathcal{M}$*. Then, for any* $\varepsilon > 0$*,*

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} u^\varepsilon(x)\, d\mu(x, v) \le 0.$$

*Proof.* For each $\eta > 0$, we set

$$\psi^\eta(x) := \int_{\mathbb{R}^n} \gamma^\eta(y) u^\varepsilon(x + y)\, dy.$$

Thanks to Lemma 3.6, $\psi^\eta$ satisfies

$$\varepsilon u^\varepsilon + H(x, D\psi^\eta) - a(x)\Delta\psi^\eta \le S^\eta(x),$$

where $|S^\eta(x)| \le C$ in $\mathbb{T}^n$ for some $C > 0$ independent of $\eta$, and $S^\eta \to 0$ pointwise in $\mathbb{T}^n$ as $\eta \to 0$.

For any vector $v \in \mathbb{R}^n$, we use the convexity of $H$ that $H(x, D\psi^\eta(x)) + L(x, v) \ge v \cdot D\psi^\eta(x)$ to obtain

$$\varepsilon u^\varepsilon + v \cdot D\psi^\eta - L(x, v) - a(x)\Delta\psi^\eta \le S^\eta(x).$$

Thus, in light of properties (i), (ii) in Proposition 3.3 of $\mu$, we integrate the above inequality with respect to $d\mu(x, v)$ on $\mathbb{T}^n \times \mathbb{R}^n$ to imply

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} \varepsilon u^\varepsilon\, d\mu(x, v) \le \iint_{\mathbb{T}^n \times \mathbb{R}^n} S^\eta(x)\, d\mu(x, v).$$

Let $\eta \to 0$ and use the Lebesgue dominated convergence theorem for the integral on the right hand side of the above to complete the proof. $\square$

We remark that the key idea of Proposition 3.9 was first observed in [41, Corollary 4].

We suggest readers give the statements and the proofs of heuristic versions of Lemma 3.8 and Proposition 3.9, in which we "assume" $w, u^\varepsilon \in C^\infty(\mathbb{T}^n)$, where $w$ and $u^\varepsilon$ are solutions of (E), $(\mathrm{D})_\varepsilon$, respectively. By doing so, one will be able to see the clear intuitions behind the complicated technicalities. To make it rigorous, as we see in the proofs of Lemma 3.8 and Proposition 3.9, the regularizing process and the commutation lemma in Section 3.4 play essential roles.

## 3.3 Proof of Theorem 3.1

Theorem 3.1 is a straightforward consequence of the following two propositions.

**Proposition 3.10.** *Assume that* (H5), (H6) *hold and the ergodic constant of* (3.1) *is* 0*. Then*

$$\liminf_{\varepsilon \to 0} u^\varepsilon(x) \ge u^0(x).$$

*Proof.* Let $\phi \in \mathcal{E}$, that is, $\phi$ is a solution of (E) satisfying (3.12). Let $\phi^\eta = \gamma^\eta * \phi$ for $\eta > 0$.

Fix $x_0 \in \mathbb{T}^n$. Take two subsequences $\eta_k \to 0$ and $\varepsilon_j \to 0$ so that (3.14) holds, and $\lim_{j\to\infty} u^{\varepsilon_j}(x_0) = \liminf_{\varepsilon\to 0} u^\varepsilon(x_0)$. Let $\mu$ be the corresponding measure satisfying $\nu = \Phi_{\#}\mu$. In view of Lemmas 3.8, 3.7,

$$\begin{aligned} u^{\varepsilon_j,\eta_k}(x_0) &\geq \phi^{\eta_k}(x_0) - \int_{\mathbb{T}^n} \phi^{\eta_k}\theta^{\varepsilon_j,\eta_k}\,dx - \frac{C\eta_k}{\varepsilon_j} - \frac{1}{\varepsilon_j}\int_{\mathbb{T}^n} S^{\eta_k}\theta^{\varepsilon_j,\eta_k}\,dx \\ &\geq \phi^{\eta_k}(x_0) - \int_{\mathbb{T}^n} \phi^{\eta_k}\theta^{\varepsilon_j,\eta_k}\,dx - \frac{C\eta_k}{\varepsilon_j} - \frac{C\eta_k^{1/2}}{\varepsilon_j}. \end{aligned}$$

Let $k \to \infty$ to imply

$$u^{\varepsilon_j}(x_0) \geq \phi(x_0) - \iint_{\mathbb{T}^n\times\mathbb{R}^n} \phi(x)\,d\nu^{\varepsilon_j}(x,p).$$

Let $j \to \infty$ in the above inequality to deduce further that

$$\begin{aligned} \liminf_{\varepsilon\to 0} u^\varepsilon(x_0) = \lim_{j\to\infty} u^{\varepsilon_j}(x_0) &\geq \phi(x_0) - \iint_{\mathbb{T}^n\times\mathbb{R}^n} \phi(x)\,d\nu(x,p) \\ &= \phi(x_0) - \iint_{\mathbb{T}^n\times\mathbb{R}^n} \phi(x)\,d\mu(x,v) \geq \phi(x_0), \end{aligned}$$

which implies the conclusion. □

**Proposition 3.11.** *Assume that* (H5), (H6) *hold and the ergodic constant of* (3.1) *is* 0. *Let* $\{\varepsilon_j\}_{j\in\mathbb{N}}$ *be any subsequence converging to* 0 *such that* $u^{\varepsilon_j}$ *uniformly converges to a solution* $u$ *of* (E) *as* $j \to \infty$. *Then the limit* $u$ *belongs to* $\mathcal{E}$. *In particular,*

$$\limsup_{\varepsilon\to 0} u^\varepsilon(x) \leq u^0(x),$$

*where* $u^0$ *is the function defined in Theorem* 3.1.

*Proof.* In view of Proposition 3.9, it is clear that any uniform limit along subsequences belongs to $\mathcal{E}$. By the definition of the function $u^0$, it is also obvious that $\lim_{j\to\infty} u^{\varepsilon_j}(x) \leq u^0(x)$. □

**Remark 3.4.** We discuss here four important points.

The first point is a technical one appearing in the proof of Proposition 3.10. In order to show that

$$\lim_{k\to\infty} \frac{1}{\varepsilon_j}\int_{\mathbb{T}^n} S^{\eta_k}\theta^{\varepsilon_j,\eta_k}\,dx = 0,$$

we needed to use the estimate $\|S^{\eta_k}\|_{L^\infty} \leq C\eta_k^{1/2}$ in Lemma 3.7. The pointwise convergence of $S^{\eta_k}$ to 0 in Lemma 3.6 is not enough.

Secondly, $\mathcal{M}$ is the collection of stochastic Mather measures that can be derived from the solutions $\theta^{\varepsilon,\eta}$ of the adjoint equations. It should be made clear that we do not collect all minimizing measures of (3.16) in $\mathcal{M}$ in general. However, we do not know whether $\mathcal{M} \subsetneq \widetilde{\mathcal{M}}$ is true or not, where $\widetilde{\mathcal{M}}$ is the set of all stochastic Mather measures defined in Definition 3.1. This is an interesting question (though technical) worth to be studied.

Thirdly, by repeating the whole proof, we obtain that

$$u^\varepsilon(x) \to \widetilde{u}^0(x) := \sup_{\phi \in \widetilde{\mathcal{E}}} \phi(x) \quad \text{uniformly for } x \in \mathbb{T}^n \text{ as } \varepsilon \to 0, \tag{3.21}$$

where we denote by $\widetilde{\mathcal{E}}$ the family of solutions $u$ of (E) satisfying

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} u \, d\mu \le 0 \qquad \text{for all } \mu \in \widetilde{\mathcal{M}}.$$

Thus, $u^0 = \widetilde{u}^0$. We will use this point later in Section 3.5.

Finally, as we only assume here that $H$ is convex, and not uniformly convex in general, we cannot expect to get deeper properties of Mather measures like Lipschitz graph property and such. For instance, we cannot expect in our setting

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} H(x,p) \, d\nu(x,p) = 0 \quad \text{for all } \nu \text{ given by (3.14)}.$$

It would be extremely interesting to investigate this kind of property for a degenerate viscous Hamilton–Jacobi equation in case $H$ is uniformly convex.

## 3.4 Proof of the commutation lemma

We will give a proof of the commutation lemma, Lemma 3.6, which is a technical result, but plays a very important role in our analysis. Indeed, for each solution $w$ of (E) with some a priori bounds, we can construct a family of smooth approximated subsolutions $\{w^\eta\}_{\eta>0}$ of (E). In particular, for any $\eta > 0$, $w^\eta$ is in $C^2(\mathbb{T}^n)$, which is good enough for us to use as test functions in Proposition 3.3 (ii). We have already seen this idea in the proof of Propositions 3.4, 3.9.

It is well-known that we can perform sup-convolutions of $w$, which was discovered by Jensen [55], to derive semi-convex approximated subsolutions of (E), but these are not smooth enough to use as test functions (see Remark 3.2). It is worth pointing out that a similar result was already discovered a long time ago by Lions [59]. However, Lions only got that $S^\eta$ converges to 0 in the *almost everywhere* sense, which is not enough for our purpose. This is because each Mather measure $\mu$ can be very singular in $\mathbb{T}^n \times \mathbb{R}^n$, and the almost everywhere sense may miss some points on the support of $\mu$. We need to have the convergence of $S^\eta$ everywhere to perform our analysis (e.g., the last step in the proof of Proposition 3.4).

The results related to the commutation lemma 3.6 may be of independent interests elsewhere, and that is the reason why we present it separately here in this section. See Section 3.6 for some further comments.

*Proof of Lemma* 3.6. It is important noting that, in view of Theorem 1.3 (see also [4, Theorem 3.1]), all viscosity solutions of (E) are Lipschitz continuous with a universal Lipschitz constant $C$. Therefore, we have

$$-C \le -a(x)\Delta w \le C \quad \text{in } \mathbb{T}^n$$

in the viscosity sense. The result of Ishii [48] on the equivalence of viscosity solutions and solutions in the distribution sense for linear elliptic equations, and the simple structure of $a(x)$ allow us to conclude further that

$$\|Dw\|_{L^\infty(\mathbb{T}^n)} + \|a\Delta w\|_{L^\infty(\mathbb{T}^n)} \le C \tag{3.22}$$

for some constant $C > 0$.

Let us next show that $w$ is actually a subsolution of (E) in the distributional sense based on the ideas in [55]. For each $\delta > 0$, let $\overline{w}^\delta$ be the sup-convolution of $w$, that is,

$$\overline{w}^\delta(x) := \sup_{y\in\mathbb{R}^n}\left(w(y) - \frac{|x-y|^2}{2\delta}\right).$$

Thanks to [55, 23], $\overline{w}^\delta$ is semi-convex and is a viscosity subsolution of

$$-a(x)\Delta\overline{w}^\delta + H(x, D\overline{w}^\delta) \le \omega(\delta) \qquad \text{in } \mathbb{T}^n. \tag{3.23}$$

Here, $\omega : (0,\infty) \to \mathbb{R}$ is a modulus of continuity, that is, $\lim_{\delta\to 0}\omega(\delta) = 0$. Since $\overline{w}^\delta$ is a semi-convex function, it is twice differentiable almost everywhere and thus is also a solution of (3.23) in the almost everywhere sense. We use (3.22) to deduce further that $\overline{w}^\delta$ is a distributional subsolution of (3.23). By passing to a subsequence if necessary, we obtain the following convergence

$$\begin{aligned}
&\overline{w}^\delta \to w && \text{uniformly in } \mathbb{T}^n,\\
&D\overline{w}^\delta \overset{*}{\rightharpoonup} Dw && \text{weakly in } L^\infty(\mathbb{T}^n),
\end{aligned}$$

as $\delta \to 0$. Take an arbitrary test function $\phi \in C^2(\mathbb{T}^n)$ with $\phi \ge 0$. We use the convexity of $H$ to yield that

$$\begin{aligned}
&\int_{\mathbb{T}^n} \left(H(x, Dw)\phi - w\Delta(a(x)\phi)\right) dx\\
&= \lim_{\delta\to 0}\int_{\mathbb{T}^n} \big(H(x, Dw)\phi + D_pH(x, Dw)\cdot D(\overline{w}^\delta - w)\phi - \overline{w}^\delta\Delta(a(x)\phi)\big)\, dx\\
&\le \lim_{\delta\to 0}\int_{\mathbb{T}^n} \big(H(x, D\overline{w}^\delta) - a\Delta\overline{w}^\delta\big)\phi\, dx \le \lim_{\delta\to 0}\int_{\mathbb{T}^n}\omega(\delta)\phi\, dx = 0.
\end{aligned}$$

Therefore, $w$ is a subsolution of (E) in the distributional sense. For each $\eta > 0$, we multiply (E) by $\gamma^\eta$ and integrate on $\mathbb{T}^n$ to get

$$-a(x)\Delta w^\eta + H(x, Dw^\eta) \le R_1^\eta(x) + R_2^\eta(x) \qquad \text{in } \mathbb{T}^n,$$

where

$$\begin{aligned} R_1^\eta(x) &:= H(x, Dw^\eta(x)) - \int_{\mathbb{R}^n} H(x+y, Dw(x+y))\gamma^\eta(y)\,dy, \\ R_2^\eta(x) &:= \int_{\mathbb{R}^n} a(x+y)\Delta w(x+y)\gamma^\eta(y)\,dy - a(x)\Delta w^\eta(x). \end{aligned}$$

We will provide treatments for $R_1^\eta$ and $R_2^\eta$ separately in Lemmas 3.12 and 3.13 below. Note that $R_2^\eta$ is exactly the commutation term mentioned in Section 3.2.2.

Basically, Lemma 3.12 gives that $R_1^\eta(x) \le C\eta$ for all $x \in \mathbb{T}^n$ and $\eta > 0$. Lemma 3.13 confirms that $|R_2^\eta(x)| \le C$ for all $x \in \mathbb{T}^n$ and $\eta > 0$, and $\lim_{\eta\to 0} R_2^\eta(x) = 0$ for each $x \in \mathbb{T}^n$.

We thus set $S^\eta(x) := C\eta + R_2^\eta(x)$ to finish the proof. □

**Lemma 3.12.** *Assume that* (H5), (H6) *hold. Then there exists* $C > 0$ *independent of* $\eta$ *such that*

$$R_1^\eta(x) \le C\eta \quad \textit{for all } x \in \mathbb{T}^n \textit{ and } \eta > 0.$$

The proof goes essentially in the same way as that of the second part of Proposition 1.5. Nevertheless, we repeat it here to remind the readers of this simple but important technique.

*Proof.* In view of (3.22) and (H5) that $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, we have

$$|H(x+y, Dw(x+y)) - H(x, Dw(x+y))| \le C\eta \quad \text{for a.e. } y \in B(x,\eta).$$

We then use the convexity of $H$ and Jensen's inequality to yield

$$\begin{aligned} R_1^\eta(x) &\le H\left(x, \int_{\mathbb{R}^n} \gamma^\eta(y) Dw(x+y)\,dy\right) - \int_{\mathbb{R}^n} H(x, Dw(x+y))\gamma^\eta(y)\,dy + C\eta \\ &\le C\eta. \end{aligned}$$

□

**Lemma 3.13.** *Assume that* (H5), (H6) *hold. Then there exists a constant* $C > 0$ *independent of* $\eta$ *such that* $|R_2^\eta(x)| \le C$ *for all* $x \in \mathbb{T}^n$ *and* $\eta > 0$. *Moreover,*

$$\lim_{\eta\to 0} R_2^\eta(x) = 0 \qquad \textit{for each } x \in \mathbb{T}^n.$$

*Proof.* We first show the boundedness of $R_2^\eta$. By using the integration by parts,

$$\begin{aligned} |R_2^\eta(x)| &= \left| \int_{\mathbb{R}^n} (a(x+y) - a(x))\Delta w(x+y)\gamma^\eta(y)\,dy \right| \\ &= \left| \int_{\mathbb{R}^n} \gamma^\eta(y) Da(x+y)\cdot Dw(x+y)\,dy + \int_{\mathbb{R}^n} (a(x+y)-a(x)) Dw(x+y)\cdot D\gamma^\eta(y)\,dy \right| \\ &\le C\int_{\mathbb{R}^n} \left(\gamma^\eta(y) + |y|\cdot|D\gamma^\eta(y)|\right)\,dy \le C. \end{aligned}$$

Next, we prove the last claim that, for each $x \in \mathbb{T}^n$, $\lim_{\eta\to 0} R_2^\eta(x) = 0$. There are two cases to be considered

$$\text{(i) } a(x) = 0, \quad \text{and} \quad \text{(ii) } a(x) > 0.$$

We handle case (i) first. Since $a(x) = 0 = \min_{\mathbb{T}^n} a$, we also have $Da(x) = 0$. Therefore,

$$\begin{aligned}
|R_2^\eta(x)| &= \left| \int_{\mathbb{R}^n} a(x+y)\Delta w(x+y)\gamma^\eta(y)\,dy \right| \\
&= \left| \int_{\mathbb{R}^n} Dw(x+y)\cdot Da(x+y)\gamma^\eta(y)\,dy + \int_{\mathbb{R}^n} Dw(x+y)\cdot D\gamma^\eta(y)a(x+y)\,dy \right| \\
&\leq C\int_{\mathbb{R}^n} \left(|Da(x+y)|\gamma^\eta(y) + a(x+y)|D\gamma^\eta(y)|\right)\,dy \\
&= C\int_{\mathbb{R}^n} \left(|Da(x+y)-Da(x)|\gamma^\eta(y) + (a(x+y)-a(x)-Da(x)\cdot y)|D\gamma^\eta(y)|\right)\,dy \\
&\leq C\int_{\mathbb{R}^n} \left(|y|\gamma^\eta(y) + |y|^2|D\gamma^\eta(y)|\right)\,dy \leq C\eta.
\end{aligned}$$

The use of Taylor's expansion of $a(\cdot) \in C^2(\mathbb{T}^n)$ around $x$ is important in the above computation.

Let us now study case (ii), in which $a(x) > 0$. We choose $\eta_0 > 0$ sufficiently small such that $a(z) \geq c_x > 0$ for $|z - x| \leq \eta_0$ for some $c_x > 0$. In view of (3.22), we deduce further that

$$|\Delta w(z)| \leq \frac{C}{c_x} =: C_x \qquad \text{for a.e. } z \in B(x,\eta_0). \tag{3.24}$$

Note that $\eta_0$ depends on $x$. For $\eta < \eta_0$, we have

$$\begin{aligned}
|R_2^\eta(x)| &= \left| \int_{\mathbb{R}^n} (a(x+y)-a(x))\Delta w(x+y)\gamma^\eta(y)\,dy \right| \\
&\leq C_x \int_{\mathbb{R}^n} |a(x+y)-a(x)|\gamma^\eta(y)\,dy \leq C_x\int_{\mathbb{R}^n} |y|\gamma^\eta(y)\,dy \leq C_x\eta.
\end{aligned}$$

In both cases, we can conclude that $\lim_{\eta\to 0}|R_2^\eta(x)| = 0$. Note however that the bound for $|R_2^\eta(x)|$ is dependent on $x$. □

**Remark 3.5.** In the proof of Lemma 3.13, estimate (3.22) plays an extremely crucial role. That is the main reason why we require $w$ to be a solution instead of just a subsolution of (E) so that (3.22) holds automatically. In fact, (3.22) does not hold for subsolutions of (E) in general. This point is one of the main difference between first-order and second-order Hamilton–Jacobi equations. For first-order Hamilton–Jacobi equations, that is, the case $a \equiv 0$, estimate (3.22) holds automatically even just for subsolutions thanks to the coercivity of $H$.

We also want to comment a bit more on the rate of convergence of $R_2^\eta$ in the above proof. For each $\delta > 0$, set $U^\delta := \{x \in \mathbb{T}^n \,:\, a(x) = 0 \text{ or } a(x) > \delta\}$. Then there exists a constant $C = C(\delta) > 0$ such that

$$|R_2^\eta(x)| \leq C(\delta)\eta \quad \text{for all } x \in U^\delta.$$

We however do not know the rate of convergence of $R_2^\eta$ in $\mathbb{T}^n \setminus U^\delta$ through the above proof yet.

With a more careful analysis, we are indeed able to improve the convergence rate of $R_2^\eta$ to $\eta^{1/2}$ in Lemma 3.7 by a more careful analysis. We do not know whether this rate is optimal or not, but for our purpose, it is good enough. See the proof of Proposition 3.10 and the first point in Remark 3.4.

*Proof of Lemma* 3.7. Fix $x \in \mathbb{T}^n$ and $\eta > 0$. We consider two cases

$$\text{(i)} \min_{y \in \overline{B(x,\eta)}} a(y) \le \eta, \quad \text{and} \quad \text{(ii)} \min_{y \in \overline{B(x,\eta)}} a > \eta.$$

In case (i), there exists $\bar{x} \in \overline{B(x,\eta)}$ such that $a(\bar{x}) \le \eta$. Then, in light of (2.19) (see also [17, Lemma 2.6]), there exists a constant $C > 0$ such that,

$$|Da(\bar{x})| \le Ca(\bar{x})^{1/2} \le C\eta^{1/2}.$$

For any $z \in \overline{B(x,\eta)}$ we have the following estimates

$$|Da(z)| \le |Da(z) - Da(\bar{x})| + |Da(\bar{x})| \le C\eta + C\eta^{1/2} \le C\eta^{1/2}.$$

Moreover, by using Taylor's expansion,

$$\begin{aligned}
&|a(z) - a(x)| \le |a(z) - a(\bar{x})| + |a(x) - a(\bar{x})| \\
&\le |Da(\bar{x})|(|z - \bar{x}| + |x - \bar{x}|) + C(|z - \bar{x}|^2 + |x - \bar{x}|^2) \le C\eta^{3/2} + C\eta^2 \le C\eta^{3/2}.
\end{aligned}$$

We use the two above inequalities to control $R_2^\eta$ as

$$\begin{aligned}
&|R_2^\eta(x)| = \left| \int_{\mathbb{R}^n} (a(x+y) - a(x))\Delta w(x+y)\gamma^\eta(y)\, dy \right| \\
&= \left| \int_{\mathbb{R}^n} Dw(x+y) \cdot Da(x+y)\gamma^\eta(y)\, dy + \int_{\mathbb{R}^n} Dw(x+y) \cdot D\gamma^\eta(y)(a(x+y) - a(x))\, dy \right| \\
&\le C \int_{\mathbb{R}^n} \left(\eta^{1/2}\gamma^\eta(y) + \eta^{3/2}|D\gamma^\eta(y)|\right)\, dy \le C\eta^{1/2}.
\end{aligned}$$

Let us now consider case (ii), in which $\min_{B(x,\eta)} a > \eta$. A direct computation shows

$$\begin{aligned}
&|R_2^\eta(x)| \le \int_{\mathbb{R}^n} |(a(x+y) - a(x))|\, |\Delta w(x+y)|\, \gamma^\eta(y)\, dy \\
&\le C \int_{\mathbb{R}^n} \frac{|a(x+y) - a(x)|}{a(x+y)} \gamma^\eta(y)\, dy \le C \int_{\mathbb{R}^n} \frac{|Da(x+y)| \cdot |y|}{a(x+y)} \gamma^\eta(y)\, dy + C\eta \\
&\le C \int_{\mathbb{R}^n} \frac{|y|}{a(x+y)^{1/2}} \gamma^\eta(y)\, dy + C\eta \le C \int_{\mathbb{R}^n} \frac{|y|}{\eta^{1/2}} \gamma^\eta(y)\, dy + C\eta \le C\eta^{1/2}.
\end{aligned}$$

Combining these estimates we get the conclusion. □

An immediate consequence of the above lemmas is the following.

**Lemma 3.14.** *Let $w \in C(\mathbb{T}^n)$ satisfy* (3.22). *Then, $w$ is a viscosity subsolution of* (E) *if and only if $w$ is a subsolution of* (E) *in the almost everywhere sense.*

*Proof.* Assume first that $w$ be a viscosity subsolution of (E). Then by the first part of the proof of Lemma 3.6, $w$ is a subsolution of (E) in the distribution sense. In light of (3.22), $w$ is furthermore a subsolution of (E) in the almost everywhere sense.

On the other hand, assume that $w$ is a subsolution of (E) in the almost everywhere sense. For each $\eta > 0$, let $w^\eta$ be the function defined by (3.18). In view of Lemmas 3.7, and the stability result of viscosity solutions, we obtain that $w$ is a viscosity subsolution of (E). □

Another consequence of Lemmas 3.6 and 3.7 is a representation formula for ergodic constant $c$ in this setting. If we repeat the argument in the proof of Proposition 1.5 by using Lemmas 3.6 and 3.7, we obtain

**Proposition 3.15.** *Assume* (H5), (H6) *hold. Let $c$ be the ergodic constant of* (3.1). *Then,*

$$c = \inf_{\phi \in C^2(\mathbb{T}^n)} \max_{x \in \mathbb{T}^n} \left(-a(x)\Delta\phi(x) + H(x, D\phi(x))\right).$$

See Section 3.6 for some further discussions.

## 3.5 Applications

Let us now discuss the limit of $u^\varepsilon$ in Examples 3.1, 3.3 and 3.4 in Section 3.1.

### Limit of $u^\varepsilon$ in Example 3.1

In this example, the equation for $u^\varepsilon$ is

$$\varepsilon u^\varepsilon + |(u^\varepsilon)'|^2 - W(x)^2 = 0 \quad \text{in } \mathbb{T}.$$

By Theorem 3.1 and Remark 3.4, we have

$$\begin{aligned} &u^\varepsilon(x) \to \widetilde{u}^0(x) \\ &\qquad = \sup\left\{ w(x) \,:\, w \text{ is a solution to (3.2) s.t. } \iint_{\mathbb{T}^n\times\mathbb{R}^n} w\, d\mu(x,v) \le 0,\ \forall\, \mu \in \widetilde{\mathcal{M}} \right\} \end{aligned}$$

uniformly for $x \in \mathbb{T}$ as $\varepsilon \to 0$. In this specific case, we have that

$$\left\{\delta_{\{1/4\}\times\{0\}}\right\} \cup \left\{\delta_{\{3/4\}\times\{0\}}\right\} \subset \widetilde{\mathcal{M}}.$$

Thus,

$$\begin{aligned} \widetilde{u}^0(x) &\le \sup\Big\{ w(x) \,:\, w \text{ is a solution to (3.2) s.t.} \\ &\qquad\qquad \iint_{\mathbb{T}^n\times\mathbb{R}^n} w\, d\mu(x,v) \le 0,\ \forall\, \mu \in \left\{\delta_{\{1/4\}\times\{0\}}\right\} \cup \left\{\delta_{\{3/4\}\times\{0\}}\right\} \Big\} \\ &= \sup\left\{ w(x) \,:\, w \text{ is a solution to (3.2) s.t. } w\big(1/4\big) \le 0, w\big(3/4\big) \le 0 \right\}, \end{aligned}$$

which implies $\widetilde{u}^0(1/4) \leq 0$ and $\widetilde{u}^0(3/4) \leq 0$. On the other hand, noting that 0 is a subsolution of $(\mathrm{D})_\varepsilon$, by the comparison principle, we have $u^\varepsilon \geq 0$ in $\mathbb{T}$, which implies $\widetilde{u}^0 \geq 0$ in $\mathbb{R}$. Thus, we obtain $\widetilde{u}^0(1/4) = 0$ and $\widetilde{u}^0(3/4) = 0$, and therefore

$$\widetilde{u}^0 = u^0 = u_1^0 = u_2^0,$$

where $u_1^0, u_2^0$ are the functions defined in Example 3.1.

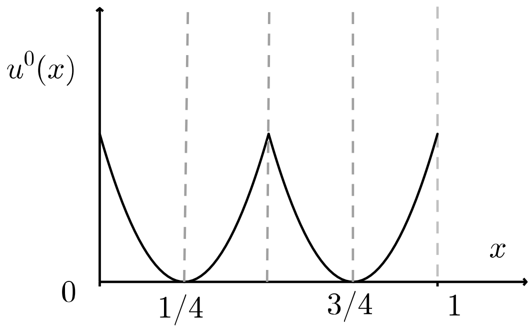


Figure 3.7: Graph of $u^0$ on $[0,1]$

.

Let us consider a slightly more general case:

$$\varepsilon u^\varepsilon + |(u^\varepsilon)' + P|^2 - V(x) = 0 \quad \text{in } \mathbb{T},$$

for $P \in \mathbb{R}$ fixed, where $V \in C^2(\mathbb{T}^n)$ with $V \geq 0$. Associated ergodic problem is

$$|u' + P|^2 - V(x) = \overline{H}(P) \quad \text{in } \mathbb{T}, \tag{3.25}$$

where we denote the ergodic constant by $\overline{H}(P)$ instead of $c$. This ergodic problem and $\overline{H}(P)$ are called, respectively, the *cell problem* and the *effective Hamiltonian* in the context of periodic homogenization of Hamilton–Jacobi equations. It is well-known that the structure of solutions to (3.25) strongly depends on $P$. Indeed, if

$$|P| \geq P_0 := \int_{\mathbb{T}} \sqrt{V(y)}\, dy,$$

then solutions of (3.25) are unique (up to additive constants). On the other hand, if $|P| < P_0$, then solutions of (3.25) are not unique in general. See [61] for more details.

Let us therefore only consider the case where $|P| < P_0$ here. In this case, it is known that

$$\delta_{\{x_0\}\times\{0\}} \in \mathcal{M} \quad \text{if and only if} \quad x_0 \in \{V = 0\} := \{x \in \mathbb{T} \,:\, V(x) = 0\}. \tag{3.26}$$

See [2, 42] for instance. Therefore, by Theorem 3.1 and Remark 3.4, we obtain

$$u^\varepsilon(x) \to u^0(x) = \sup\Big\{ w(x) \,:\, w \text{ is a solution to (3.25) s.t. } w \leq 0 \text{ on } \{V = 0\} \Big\}$$

uniformly for $x \in \mathbb{T}$ as $\varepsilon \to 0$.

We emphasize here that the characterization of Mather measures is very hard in general. Indeed, it is still not known yet whether characterization (3.26) holds or not in the multi-dimensional cases even in this specific form. See Section 3.6 for some further discussions.

### Limit of $u^\varepsilon$ in Examples 3.3, 3.4

In these examples, the equation for $u^\varepsilon$ is

$$\varepsilon u^\varepsilon - a(x)(u^\varepsilon)'' + |(u^\varepsilon)'|^2 - V(x) = 0 \quad \text{in } \mathbb{T},$$

where $V$ and $a$ are given functions in Examples 3.3, 3.4.

Firstly, let us consider Example 3.3. By Theorem 3.1 and Remark 3.4, we know that $u^\varepsilon$ uniformly converges to $u^0$ given in (3.21). Moreover, noting that

$$\delta_{\{0\}\times\{0\}}, \delta_{\{1/4\}\times\{0\}}, \delta_{\{3/4\}\times\{0\}} \in \widetilde{\mathcal{M}},$$

we obtain

$$u^\varepsilon \to u_\alpha \quad \text{uniformly for } x \in \mathbb{T} \text{ as } \varepsilon \to 0,$$

where $u_\alpha$ and $\alpha$ are the function and the constant given by (3.6) and (3.5), respectively. See Figure 3.4.

Similarly, we can characterize the limit of the discount approximation for Example 3.4. Noting that

$$\delta_{\{0\}\times\{0\}}, \delta_{\{1/4\}\times\{0\}}, \delta_{\{1/2\}\times\{0\}} \in \widetilde{\mathcal{M}},$$

we obtain

$$u^\varepsilon \to u_\alpha \quad \text{uniformly for } x \in \mathbb{T} \text{ as } \varepsilon \to 0,$$

where $u_\alpha$ and $\alpha$ are the function and the constant given by (3.7) and (3.5), respectively. See Figure 3.6.

## 3.6 Some other directions and open questions

In this section, we present recent developments in the study of selection problems for Hamilton–Jacobi equations. There are other methods to construct or approximate viscosity solutions of the ergodic problem for Hamilton–Jacobi equations such as the vanishing viscosity method, a finite difference approximation. If we consider a different type of approximation for (1.6), then the selection procedure could be rather different. Therefore, different types of difficulties may appear in general. Let us describe briefly these directions as well as some open questions.

### Discounted approximation procedure

(i) General convex settings (e.g., fully nonlinear, degenerate elliptic PDEs under various type of boundary conditions such as periodic condition, state constraint condition, Dirichlet condition, Neumann condition): Ishii, Mitake, Tran [51, 52] obtained convergence results in 2016. The proofs in [51, 52] are based on a variational approach and a duality principle, which are completely different from the ones presented here. They nevertheless share the same philosophy.

(ii) Selection problems for nonconvex Hamilton–Jacobi equations: Most problems are open. In some examples, invariant measures and invariant sets do not exist (see Cagnetti, Gomes, Tran [18] for the discussion on Mather measures, and Gomes, Mitake, Tran [43] for the discussion on Aubry set). It is therefore extremely challenging to establish general convergence results and to describe the limits if they exist. Gomes, Mitake, Tran [43] proved convergence results for some special nonconvex first-order cases in 2016.

(iii) Rate of convergence: It is quite challenging to obtain some rates of the convergence (quantitative results) of Theorem 3.1. Mitake, Soga [67] studied this for some special first-order situations in 2016. It is demonstrated there that error estimates would depend highly on dynamics of the corresponding dynamical systems in general.

(iv) Aubry (uniqueness) set: The structure of solutions of (1.8) is poorly understood. For instance, in the case of the inviscid (first-order) equation, the Aubry set plays a key role as a uniqueness set for the ergodic problem. In a general viscous case where the diffusion could be degenerate, there has not been any similar notions/results on the uniqueness (Aubry) set for (1.8) up to now.

(v) Commutation lemma 3.6: Another way to perform this task is to do sup-inf-convolution first, and usual convolution later. Ishii, Mitake, Tran did this in a unpublished note first before finding the new variational approach in [51]. Are these useful in other contexts?

(vi) Applications: Theorem 3.1 is very natural in its own right. It is therefore extremely interesting to use it to get some further PDE results and to find connections to dynamical systems.

## Vanishing viscosity procedure

(i) Vanishing viscosity procedure: For $\varepsilon > 0$, consider the following problem

$$H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon + \overline{H}^\varepsilon \quad \text{in } \mathbb{T}^n$$

where $\overline{H}^\varepsilon$ is the corresponding ergodic constant. The question of interest is to investigate the limit of $u^\varepsilon$ as $\varepsilon \to 0$. Under relatively restrictive assumptions on the Aubry set, the convergence is proven. See Bessi [16], Anantharaman, Iturriaga, Padilla, Sanchez-Morgado [3]. In the general setting, there are still many questions which are not solved yet. See also E [29] and Jauslin-Kreiss-Moser [54] for related works on entropy solutions.

(ii) Finite difference approximation: In [75], the selection problem which appears in the finite difference procedure was first formulated by Soga, and the convergence was also proven there in a similar setting to that of the vanishing viscosity procedure.

## Selection of Mather measures

The fact that ergodic problems (1.6) and (1.8) have many solutions of different types is strongly related to the multiplicity of Mather measures. Each approximation of the ergodic problem has associated generalized Mather measures. Thus, the selection problem for Mather measures appears. Many questions still remain open. See Anantharaman [2], Evans [33], Gomes [39], Gomes, Iturriaga, Sanchez-Morgado, Yu [42], Yu [79], Mitake, Soga [67] for this direction.

# Chapter 4

# Appendix

The readers can read Appendix independently from other chapters. In Appendix, we give a short introduction to the theory of viscosity solutions of first-order Hamilton–Jacobi equations, which was introduced by Crandall and Lions [24] (see also Crandall, Evans, and Lions [22]). The readers can use this as a starting point to learn the theory of viscosity solutions. Some of this short introduction is taken from the book of Evans [31]. Let us for simplicity focus on the initial-value problem of first-order (inviscid) Hamilton–Jacobi equations

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

where the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and the initial function $u_0 : \mathbb{R}^n \to \mathbb{R}$ are given. We will give precise assumptions on $H$ and $u_0$ when necessary.

The original approach [24, 22, 60] is to consider the following approximated equation

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

for $\varepsilon > 0$. The term $\varepsilon \Delta u^\varepsilon$ in $\text{(C)}_\varepsilon$ regularizes the Hamilton–Jacobi equation, and this is called the method of *vanishing viscosity*. We then let $\varepsilon \to 0$ and study the limit of the family $\{u^\varepsilon\}_{\varepsilon>0}$. It is often the case that, in light of a priori estimates, $\{u^\varepsilon\}_{\varepsilon>0}$ is bounded and equicontinuous on compact subsets of $\mathbb{R}^n \times [0, \infty)$. We hence can use the Arzelà-Ascoli theorem to deduce that, there exists a subsequence $\{\varepsilon_j\}_j$ converging to $0$ as $j \to \infty$ such that,

$$u^{\varepsilon_j} \to u, \quad \text{locally uniformly in } \mathbb{R}^n \times [0, \infty),$$

for some limit function $u \in C(\mathbb{R}^n \times [0, \infty))$. We expect that $u$ is some kind of solution of (C), but we only have that $u$ is continuous and absolutely no information about $Du$ and $u_t$. Also as (C) is fully nonlinear in $Du$ and not of the divergence structure, we cannot use integration by parts and weak convergence techniques to justify that $u$ is a weak solution in such sense. We instead use the maximum principle to obtain the notion of weak solution, which is viscosity solution.

The terminology *viscosity solutions* is used in honor of the vanishing viscosity technique (see the proof of Theorem 4.1 in Section 4.2). We can see later that the definition of viscosity solutions does not involve viscosity of any kind but the name remains because of the history of the subject. We refer to [6, 23, 31] for general theory of viscosity solutions.

# 4.1 Motivation and Examples

In this section we give some examples to explain motivations to study (C).

## 4.1.1 Front propagation problems

We consider a *surface evolution equation* as follows. Let $n \in \mathbb{N}$ and $\{\Gamma(t)\}_{t\geq 0}$ be a given family of hypersurfaces embedded in $\mathbb{R}^n$ parametrized by time $t$. Assume that the surface evolves in time according to the law:

$$V(x, x_{n+1}, t) = -h(x) \quad \text{on } \Gamma(t), \tag{4.1}$$

where $V$ is the normal velocity at each point on $\Gamma(t)$, and $h \in C(\mathbb{R}^n)$ is a given positive function. In this section, we consider the case where $\Gamma(t)$ is described by the following graph

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

for a real-valued auxiliary function $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$.

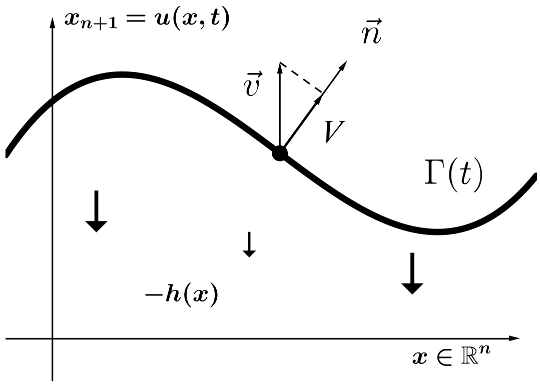


Figure 4.1

Figure 4.1 shows an example of $\Gamma(t)$ and how it evolves. We note that the direction $x_{n+1}$ in the picture shows the positive direction of $V$. The function $h$ is decided by the phenomenon which we want to consider and it sometimes depends on the curvatures, the time, etc. We simply consider the situation that $h$ depends only on the $x$ variable here. We refer to [38] for many interesting applications appearing in front propagation problems.

Suppose that everything is smooth, and then by elementary calculations, we get

$$V = \vec{v} \cdot \vec{n} = \begin{pmatrix} 0 \\ u_t \end{pmatrix} \cdot \frac{1}{\sqrt{1+|Du|^2}} \begin{pmatrix} -Du \\ 1 \end{pmatrix} = \frac{u_t}{\sqrt{1+|Du|^2}},$$

where $\vec{v}$ denotes the velocity in the direction $x_{n+1}$. Plug this into (4.1), we get that $u$ is a solution to the Hamilton–Jacobi equation

$$u_t + h(x)\sqrt{1+|Du|^2} = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty).$$

**Example 4.1.** We consider the simplest case where $n = 1$, $h(x) \equiv 1$ and two initial data: (i) a line in Figure 4.2, (ii) a curve in Figure 4.3.

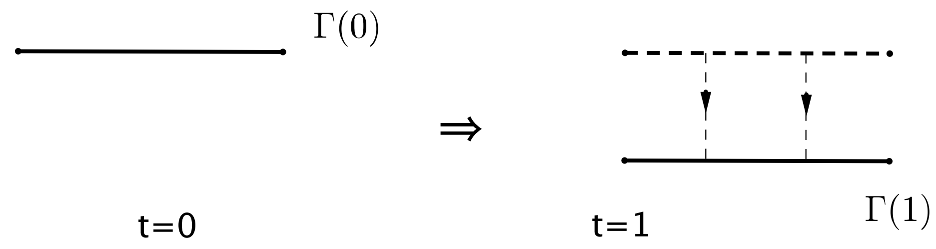


Figure 4.2

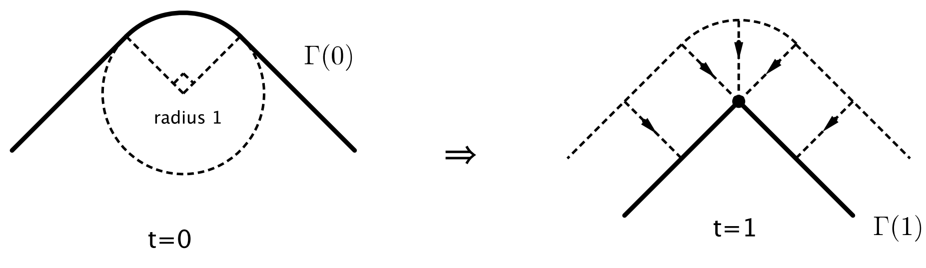


Figure 4.3

In the context of large time behavior (Chapter 2), the large time limit (asymptotic profile), if exists, is a solution to the associated ergodic problem. We also observe that it depends on the initial data as demonstrated in Figures 4.2, 4.3. In general, it is highly nontrivial to characterize this dependence as we deal with nonlinear equations. Section 2.5 somehow gives an answer to this question (see Examples 2.2, 2.3 in Section 2.5).

### 4.1.2 Optimal control problems

Let $L(x,v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a superlinear and convex Lagrangian with respect to the variable $v$, that is, for all $x, v_1, v_2 \in \mathbb{R}^n$, and $\lambda \in [0,1]$,

$$\begin{cases} \dfrac{L(x,v)}{|v|} \to \infty \quad \text{locally uniformly for } x \in \mathbb{R}^n \text{ as } |v| \to \infty, \\ L(x, \lambda v_1 + (1-\lambda)v_2) \le \lambda L(x,v_1) + (1-\lambda)L(x,v_2). \end{cases}$$

**Inviscid cases.**

We consider the optimal control problem, for fixed $(x,t) \in \mathbb{R}^n \times [0,\infty)$,

$$\text{Minimize} \quad \int_0^t L(\gamma(s), -\dot{\gamma}(s))\, ds + u_0(\gamma(t))$$

over all controls $\gamma \in \mathrm{AC}\,([0,t],\mathbb{R}^n)$ with $\gamma(0) = x$. Here $u_0$ is a given bounded uniformly continuous function on $\mathbb{R}^n$. We denote by $u(x,t)$ the minimum cost. It can be proven that $u$ solves the following Cauchy problem:

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

where the Hamiltonian $H$ is the Legendre transform of the Lagrangian $L$, that is,

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

Let us show a quick formal proof of this. Note first that $u$ satisfies the so-called *dynamic programming principle*, that is, for any $h > 0$,

$$u(x,t+h) = \inf \left\{ \int_0^h L(\gamma(s), -\dot{\gamma}(s))\, ds + u(\gamma(h), t) \,:\, \gamma(0) = x \right\}. \tag{4.2}$$

The dynamic programming principle can be checked in a rough way as following:

$$\begin{aligned} u(x,t+h) &= \int_0^{t+h} L(\gamma^*(s), -\dot{\gamma}^*(s))\, ds + u_0(\gamma^*(t+h)) \\ &= \int_0^h L(\gamma^*(s), -\dot{\gamma}^*(s))\, ds + \int_0^t L(\delta(s), -\dot{\delta}(s))\, ds + u_0(\delta(t)) \\ &= \int_0^h L(\gamma^*(s), -\dot{\gamma}^*(s))\, ds + u(\gamma^*(h), t), \end{aligned}$$

where we denote a minimizer of the minimizing problem for $u(x,t+h)$ by $\gamma^*$, and set $\delta(s) := \gamma^*(s+h)$ for $s \in [-h,t]$, and we used the Bellman principle.

We rewrite it as

$$\frac{u(\delta(-h), t+h) - u(\delta(0), t)}{h} = \frac{1}{h} \int_0^h L(\gamma(s), -\dot{\gamma}(s))\, ds.$$

Sending $h \to 0$ yields

$$u_t + Du \cdot \left(-\dot{\delta}(0)\right) - L\left(x, -\dot{\delta}(0)\right) = 0,$$

which more or less implies the conclusion. We can use this formal idea to give a rigorous proof by performing careful computations and using the notion of viscosity solutions. We refer to [31, 5] for details for instance.

**Example 4.2** (Classical mechanics)**.** We consider the case that $L$ is the difference between a kinetic energy and a potential energy, i.e., $L(x,v) := |v|^2/2 - V(x)$ for a given function $V$ which is uniformly bounded continuous on $\mathbb{R}^n$. Then,

$$u(x,t) = \inf\left\{\int_0^t \left[\frac{1}{2}|\dot\gamma(s)|^2 - V(\gamma(s))\right] ds + u_0(\gamma(t)) \,:\, \gamma \in \mathrm{AC}\left([0,t],\mathbb{R}^n\right), \gamma(0) = x\right\}$$

solves the following Cauchy problem

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

In this case, $H(x,p) = |p|^2/2 + V(x)$.

**Example 4.3** (Hopf–Lax formula)**.** If we consider a Lagrangian which is independent of the $x$ variable, then we can get the *Hopf–Lax formula* from the optimal control formula. In short, we have

$$\begin{aligned} u(x,t) &= \inf\left\{\int_0^t L(-\dot\gamma(s))\, ds + u_0(\gamma(t)) \,:\, \gamma \in \mathrm{AC}\left([0,t],\mathbb{R}^n\right), \gamma(0) = x\right\} \\ &= \inf_{y \in \mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + u_0(y)\right\}. \end{aligned}$$

It is quite straightforward to prove that the first line implies the second line in the above by using Jensen's inequality. We leave it to the interested readers. See [31] for instance. We also refer to [5, 20] for the connections between the theory of viscosity solutions and optimal control theory.

**Example 4.4** (Discounted approximation)**.** Fix $\delta > 0$. For $x \in \mathbb{R}^n$, define

$$v^\delta(x) = \inf\left\{\int_0^\infty e^{-\delta s} L(\gamma(s), -\dot\gamma(s))\, ds \,:\, \gamma \in \mathrm{AC}\left([0,\infty),\mathbb{R}^n\right), \gamma(0) = x\right\}.$$

This is an *infinite horizon problem* in optimal control theory. The function $v^\delta$ satisfies the dynamic programming principle

$$v^\delta(x) = \inf\left\{\int_0^h e^{-\delta s} L(\gamma(s), -\dot\gamma(s))\, ds + e^{-\delta h} v^\delta(\gamma(h)) \,:\, \gamma \in \mathrm{AC}\left([0,h],\mathbb{R}^n\right), \gamma(0) = x\right\}$$

for any $h > 0$. We can use this to check that $v^\delta$ solves the following discount Hamilton–Jacobi equation

$$\delta v^\delta + H(x, Dv^\delta) = 0 \quad \text{in } \mathbb{R}^n.$$

In the formula of $v^\delta$, the function $e^{-\delta s}$ plays a role of discount, and therefore, the constant $\delta$ in the above formula is called the *discount factor* in optimal control theory.

**Viscous cases.**

We consider the stochastic optimal control problem

$$\begin{aligned} &\text{Minimize} \quad \mathbb{E}\left[\int_0^t L(X^v(s), v(s))\,ds + u_0(X^v(t))\right] \\ &\text{subject to} \quad X^v = x - \int_0^t v(s)\,ds + \sqrt{2}\int_0^t \sigma(X^v(s))\,dW_s \end{aligned}$$

for $x \in \mathbb{R}^n$, over all controls $v$ in some admissible class, where $m \in \mathbb{N}$ and $\sigma : \mathbb{R}^n \to \mathbb{M}^{n\times m}$ is a given matrix-valued function which is Lipschitz, and $W_s$ denotes a standard $m$-dimensional Brownian motion. Here, $\mathbb{M}^{n\times m}$ denotes the set of $n$-by-$m$ matrices. Let $u(x,t)$ be the corresponding minimum cost.

We can prove in appropriate settings that the function $u$ solves the Cauchy problem for the general viscous Hamilton–Jacobi equation

$$\begin{cases} u_t - \operatorname{tr}\left(A(x)D^2u\right) + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n, \end{cases}$$

where $A(x) := \sigma(x)\sigma^T(x)$, by using the dynamic programming principle, the Itô formula and the notion of viscosity solutions.

We refer to [36, 74] for the connections between the theory of viscosity solutions and stochastic optimal control theory.

## 4.2 Definitions

Let us now introduce the definitions of viscosity subsolutions, supersolutions, and solutions. These definitions are encoded naturally in the vanishing viscosity method (see the proof of Theorem 4.1 below).

**Definition 4.1** (Viscosity subsolutions, supersolutions, solutions). *An upper semicontinuous function $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is called a viscosity subsolution of the intial-value problem* (C) *provided that*

- $u(\cdot, 0) \le u_0$ *on* $\mathbb{R}^n$,
- *for each* $\varphi \in C^1(\mathbb{R}^n \times (0,\infty))$, *if* $u - \varphi$ *has a local maximum at* $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$ *then*
$$\varphi_t(x_0,t_0) + H(x_0, D\varphi(x_0,t_0)) \le 0.$$

*A lower semicontinuous function $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is called a viscosity supersolution of the intial-value problem* (C) *provided that*

- $u(\cdot, 0) \ge u_0$ *on* $\mathbb{R}^n$,

- *for each* $\varphi \in C^1(\mathbb{R}^n \times (0,\infty))$*, if* $u - \varphi$ *has a local minimum at* $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$ *then*
$$\varphi_t(x_0,t_0) + H(x_0, D\varphi(x_0,t_0)) \geq 0.$$

*A function* $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ *is called a viscosity solution of the intial-value problem* (C) *if* $u$ *is both a viscosity subsolution, and a viscosity supersolution* (*hence continuous*) *of* (C).

**Remark 4.1.** (i) In Definition 4.1, a local maximum (resp., minimum) can be replaced by a maximum (resp., minimum) or even by a strict maximum (resp., minimum). Besides, a $C^1$ test function $v$ can be replaced by a $C^\infty$ test function $v$ as well.

(ii) For $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$, we set

$$\begin{aligned} D^+u(x_0,t_0) := \{(\varphi_t(x_0,t_0), D\varphi(x_0,t_0)) \ :& \\ &\varphi \in C^1 \text{ and } u - \varphi \text{ has a local maximum at } (x_0,t_0)\}, \\ D^-u(x_0,t_0) := \{(\varphi_t(x_0,t_0), D\varphi(x_0,t_0)) \ :& \\ &\varphi \in C^1 \text{ and } u - \varphi \text{ has a local minimum at } (x_0,t_0)\}. \end{aligned}$$

The sets $D^+u(x_0,t_0)$, $D^-u(x_0,t_0)$ are called the *superdifferential* and *subdifferential* of $u$ at $(x_0,t_0)$, respectively. We can rewrite the definitions of viscosity subsolutions and supersolutions by using the superdifferential and subdifferential, respectively (see [22]).

We also give the definitions of viscosity subsolutions, supersolutions, and solutions to the following second order equation

$$\begin{cases} u_t + F(x, Du, D^2u) = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{4.3}$$

**Definition 4.2.** *An upper semicontinuous function* $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ *is called a viscosity subsolution of the intial-value problem* (4.3) *provided that*

- $u(\cdot,0) \leq u_0$ *on* $\mathbb{R}^n$,
- *for any* $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$ *and* $(a,p,X) \in J^{2,+}u(x_0,t_0)$,
$$a + F(x_0,p,X) \leq 0,$$
*where*
$$\begin{aligned} J^{2,+}u(x_0,t_0) := \{(\varphi_t(x_0,t_0), D\varphi(x_0,t_0), D^2\varphi(x_0,t_0)) \ :& \\ &\varphi \in C^2 \textit{ and } u - \varphi \textit{ has a local maximum at } (x_0,t_0)\}. \end{aligned}$$

*A lower semicontinuous function* $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ *is called a viscosity supersolution of the intial-value problem* (4.3) *provided that*

- $u(\cdot,0) \geq u_0$ *on* $\mathbb{R}^n$,

- *for any* $(x_0,t_0)\in\mathbb{R}^n\times(0,\infty)$ *and* $(a,p,X)\in J^{2,-}u(x_0,t_0)$,

$$a+F(x_0,p,X)\geq 0,$$

*where*

$$\begin{aligned}J^{2,-}u(x_0,t_0):=\{(\varphi_t(x_0,t_0),D\varphi(x_0,t_0),D^2\varphi(x_0,t_0)):\\ \varphi\in C^2 \text{ and } u-\varphi \text{ has a local minimum at } (x_0,t_0)\}.\end{aligned}$$

*A function* $u:\mathbb{R}^n\times[0,\infty)\to\mathbb{R}$ *is called a viscosity solution of the intial-value problem* (4.3) *if* $u$ *is both a viscosity subsolution, and a viscosity supersolution* (*hence continuous*) *of* (4.3).

We call $J^{2,+}u(x_0,t_0)$ and $J^{2,-}u(x_0,t_0)$ the *super-semijet* and *sub-semijet* of $u$ at $(x_0,t_0)$, respectively.

**Remark 4.2.** In Definition 4.2, $J^{2,\pm}u(x_0,t_0)$ can be replaced by the closure of these sets, which are defined as

$$\begin{aligned}\overline{J}^{2,\pm}u(x_0,t_0):=\{(a,p,X)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{M}^{n\times n}_{\mathrm{sym}}:\exists\,(x_k,t_k,a_k,p_k,X_k)\text{ s.t.}\\ (a_k,p_k,X_k)\in J^{2,\pm}u(x_k,t_k)\text{ and }(x_k,t_k,a_k,p_k,X_k)\to(x_0,t_0,a,p,X)\text{ as }k\to\infty\}.\end{aligned}$$

We here give a precise result concerning the vanishing viscosity method explained in the introduction of this appendix. It shows that Definition 4.1 arises naturally in light of this procedure and the maximum principle. We will verify the assumption in this theorem in Section 4.6 below.

**Theorem 4.1** (Vanishing viscosity method). *Let* $u^\varepsilon$ *be the smooth solution of* $(\mathrm{C})_\varepsilon$ *for* $\varepsilon>0$. *Assume that there exists a subsequence* $\{u^{\varepsilon_j}\}_j$ *such that*

$$u^{\varepsilon_j}\to u,\quad \text{locally uniformly in } \mathbb{R}^n\times[0,\infty),$$

*for some* $u\in C(\mathbb{R}^n\times[0,\infty))$ *as* $j\to\infty$. *Then* $u$ *is a viscosity solution of* (C).

*Proof.* We only prove that $u$ is a viscosity subsolution of (C) as similarly we can prove that it is a viscosity supersolution of (C). Take any $\varphi\in C^\infty(\mathbb{R}^n\times(0,\infty))$ and assume that $u-\varphi$ has a strict maximum at $(x_0,t_0)\in\mathbb{R}^n\times(0,\infty)$.

Recall that $u^{\varepsilon_j}\to u$ locally uniformly as $j\to\infty$. For $j$ large enough, $u^{\varepsilon_j}-\varphi$ has a local maximum at some point $(x_j,t_j)$ and

$$(x_j,t_j)\to(x_0,t_0),\quad \text{as } j\to\infty.$$

We have

$$u^{\varepsilon_j}_t(x_j,t_j)=\varphi_t(x_j,t_j),\ Du^{\varepsilon_j}(x_j,t_j)=D\varphi(x_j,t_j),\ -\Delta u^{\varepsilon_j}(x_j,t_j)\geq-\Delta\varphi(x_j,t_j).$$

Hence,

$$\begin{aligned}\varphi_t(x_j,t_j)+H(x_j,D\varphi(x_j,t_j)) &= u_t^{\varepsilon_j}(x_j,t_j)+H(x_j,Du^{\varepsilon_j}(x_j,t_j))\\ &= \varepsilon_j\Delta u^{\varepsilon_j}(x_j,t_j)\le \varepsilon_j\Delta\varphi(x_j,t_j).\end{aligned}$$

Let $j\to\infty$ to imply that

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

**Remark 4.3.**
(i) Let us emphasize that obtaining viscosity solutions through the vanishing viscosity approach is the classical approach. This method does not work for general second-order equations. In general, we can use Perron's method to prove the existence of viscosity solutions.
(ii) As seen in the proof of Theorem 4.1, we lose the information of $Du^\varepsilon$ and $\Delta u^\varepsilon$ as $\varepsilon\to 0$ in this argument. Evans [32] introduced the *nonlinear adjoint method* to understand these in the vanishing viscosity procedure. In particular, his aim is to understand gradient shock structures in the nonconvex setting.

## 4.3 Consistency

We here prove that the notion of viscosity solutions is consistent with that of classical solutions.

Firstly, it is quite straightforward to see that if $u\in C^1(\mathbb{R}^n\times[0,\infty))$ solves (C) in the classical sense, then $u$ is a viscosity solution of (C). Next, we show that if a viscosity solution is differentiable at some point, then it solves (C) there. We need the following lemma.

**Lemma 4.2** (Touching by a $C^1$ function)**.** *Let $m\in\mathbb{N}$. Assume $u:\mathbb{R}^m\to\mathbb{R}$ is continuous in $\mathbb{R}^m$, and is differentiable at some point $x_0$. There exists $\varphi\in C^1(\mathbb{R}^m)$ such that $u(x_0)=\varphi(x_0)$ and $u-\varphi$ has a strict local maximum at $x_0$.*

*Proof.* Without loss of generality, we may assume that

$$x_0=0,\ u(0)=0,\ \text{and } Du(0)=0. \tag{4.4}$$

We use (4.4) and the differentiability of $u$ at 0 to deduce that

$$u(x)=|x|\omega(x), \tag{4.5}$$

where $\omega:\mathbb{R}^m\to\mathbb{R}$ is continuous with $\omega(0)=0$. For each $r>0$, we define

$$\rho(r)=\max_{x\in B(0,r)}|\omega(x)|.$$

We see that $\rho:[0,\infty)\to[0,\infty)$ is continuous, increasing, and $\rho(0)=0$.

We define

$$\varphi(x) = \int_{|x|}^{2|x|} \rho(r)dr + |x|^2, \quad \text{for } x \in \mathbb{R}^m. \tag{4.6}$$

It is clear that $|\varphi(x)| \le |x|\rho(2|x|) + |x|^2$, which implies

$$\varphi(0) = 0, \; D\varphi(0) = 0.$$

Besides, for $x \neq 0$, explicit computations give us that

$$D\varphi(x) = \frac{2x}{|x|}\rho(2|x|) - \frac{x}{|x|}\rho(|x|) + 2x,$$

and hence $\varphi \in C^1(\mathbb{R}^m)$.

Finally, for every $x \neq 0$,

$$\begin{aligned} u(x) - \varphi(x) &= |x|\omega(x) - \int_{|x|}^{2|x|} \rho(r)dr - |x|^2 \\ &\le |x|\rho(|x|) - |x|\rho(|x|) - |x|^2 < 0 = u(0) - \varphi(0). \end{aligned}$$

The proof is complete. □

Lemma 4.2 immediately implies the following.

**Theorem 4.3** (Consistency of viscosity solutions)**.** *Let $u$ be a viscosity solution of* (C)*, and suppose that $u$ is differentiable at $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$. Then*

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

## 4.4 Comparison principle and Uniqueness

In this section, we establish the comparison principle for (C). Let us first prepare a useful lemma.

**Lemma 4.4** (Extrema at a terminal time)**.** *Fix $T > 0$. Assume that $u$ is a viscosity subsolution* (*resp., supersolution*) *of* (C)*. Assume further that, on $\mathbb{R}^n \times (0, T]$, $u - \varphi$ has a local maximum* (*resp., minimum*) *at a point $(x_0, t_0) \in \mathbb{R}^n \times (0, T]$, for some $\varphi \in C^1(\mathbb{R}^n \times [0, \infty))$. Then*

$$\varphi_t(x_0, t_0) + H(x_0, D\varphi(x_0, t_0)) \le 0 \quad (\textit{resp.}, \; \ge 0).$$

The point here is that terminal time $t_0 = T$ is allowed.

*Proof.* We just need to verify the case of subsolution. Assume $u - \varphi$ has a strict maximum at $(x_0, T)$. We define

$$\overline{\varphi}(x, t) = \varphi(x, t) + \frac{\varepsilon}{T - t} \quad \text{for } (x, t) \in \mathbb{R}^n \times (0, \infty)$$

for $\varepsilon > 0$. If $\varepsilon > 0$ is small enough, then $u - \overline{\varphi}$ has a local maximum at $(x_\varepsilon, t_\varepsilon) \in \mathbb{R}^n \times (0, T)$ and $(x_\varepsilon, t_\varepsilon) \to (x_0, T)$ as $\varepsilon \to 0$. By definition of viscosity subsolutions, we have

$$\overline{\varphi}_t(x_\varepsilon, t_\varepsilon) + H(x_\varepsilon, D\overline{\varphi}(x_\varepsilon, t_\varepsilon)) \leq 0$$

which is equivalent to

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

Hence,

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

We let $\varepsilon \to 0$ to achieve the result. □

We fix $T > 0$ now and consider (C) in $\mathbb{R}^n \times [0, T]$ only, i.e.,

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

Let us first give a formal argument to see how the comparison principle works. Let $u, v$ be a smooth subsolution and a smooth supersolution to (C), respectively, with the same initial data. Our goal is to prove that $u \leq v$ on $\mathbb{R}^n \times [0, T]$. We argue by contradiction, and therefore we suppose that

$$\max_{\mathbb{R}^n \times [0,T]} \{(u - v)(x, t) - \lambda t\} > 0$$

for a small $\lambda > 0$. Suppose formally that the maximum is attained at $(x_0, t_0) \in \mathbb{R}^n \times [0, T)$. Because of the initial data, we have $t_0 > 0$. Then,

$$u_t(x_0, t_0) = v_t(x_0, t_0) + \lambda, \quad Du(x_0, t_0) = Dv(x_0, t_0).$$

Thus,

$$0 \geq u_t(x_0, t_0) + H(x_0, Du(x_0, t_0)) = v_t(x_0, t_0) + \lambda + H(x_0, Dv(x_0, t_0)) \geq \lambda > 0,$$

which is a contradiction.

We now establish the comparison principle (hence uniqueness) for (C) rigorously by using the so-called *doubling variable* argument, which was originally introduced by Kružkov [57].

We assume further that the Hamiltonian $H$ satisfies

(A1) There exist a positive constant $C$ such that

$$\begin{aligned} &|H(x, p) - H(x, q)| \leq C|p - q|, \\ &|H(x, p) - H(y, p)| \leq C|x - y|(1 + |p|), \qquad \text{for } (x, y, p, q) \in (\mathbb{R}^n)^4. \end{aligned}$$

**Theorem 4.5** (Comparison Principle for (C))**.** *Assume that* (A1) *holds. If $u$, $\tilde{u}$ are a bounded uniformly continuous viscosity subsolution, and supersolution of* (C) *on* $\mathbb{R}^n \times [0,T]$*, respectively, then* $u \le \tilde{u}$ *on* $\mathbb{R}^n \times [0,T]$.

*Proof.* We assume by contradiction that

$$\sup_{\mathbb{R}^n\times[0,T]} (u-\tilde{u}) = \sigma > 0.$$

For $\varepsilon, \lambda \in (0,1)$, we define

$$\Phi(x,y,t,s) = u(x,t) - \tilde{u}(y,s) - \lambda(t+s) - \frac{1}{\varepsilon^2}(|x-y|^2 + (t-s)^2) - \varepsilon(|x|^2+|y|^2)$$

for $x, y \in \mathbb{R}^n$, $t, s \ge 0$. Since $u, \tilde{u}$ are bounded, there exists a point $(x_0, y_0, t_0, s_0) \in \mathbb{R}^{2n} \times [0,T]^2$ such that

$$\Phi(x_0,y_0,t_0,s_0) = \max_{\mathbb{R}^{2n}\times[0,T]^2} \Phi(x,y,t,s).$$

For $\varepsilon, \lambda$ small enough, we have $\Phi(x_0,y_0,t_0,s_0) \ge \sigma/2$.

We use $\Phi(x_0,y_0,t_0,s_0) \ge \Phi(0,0,0,0)$ to get

$$\begin{aligned}&\lambda(t_0+s_0) + \frac{1}{\varepsilon^2}(|x_0-y_0|^2 + (t_0-s_0)^2) + \varepsilon(|x_0|^2+|y_0|^2)\\ \le\, &u(x_0,t_0) - \tilde{u}(y_0,s_0) - u(0,0) + \tilde{u}(0,0) \le C.\end{aligned} \tag{4.7}$$

Hence,

$$|x_0-y_0| + |t_0-s_0| \le C\varepsilon, \quad |x_0| + |y_0| \le \frac{C}{\varepsilon^{1/2}}. \tag{4.8}$$

We next use $\Phi(x_0,y_0,t_0,s_0) \ge \Phi(x_0,x_0,t_0,t_0)$ to deduce that

$$\frac{1}{\varepsilon^2}(|x_0-y_0|^2 + (t_0-s_0)^2) \le \tilde{u}(x_0,t_0) - \tilde{u}(y_0,s_0) + \lambda(t_0-s_0) + \varepsilon(x_0-y_0)\cdot(x_0+y_0).$$

In view of (4.8) and the uniformly continuity of $\tilde{u}$, we get

$$\frac{|x_0-y_0|^2 + (t_0-s_0)^2}{\varepsilon^2} \to 0, \quad |x_0-y_0| + |t_0-s_0| = o(\varepsilon) \text{ as } \varepsilon \to 0. \tag{4.9}$$

By (4.8) and (4.9), we can take $\varepsilon > 0$ small enough so that $s_0, t_0 \ge \mu > 0$ for some $\mu > 0$.

Notice that $(x,t) \mapsto \Phi(x,y_0,t,s_0)$ has a maximum at $(x_0,t_0)$. In view of the definition of $\Phi$, $u - \varphi$ has a maximum at $(x_0,t_0)$ for

$$\varphi(x,t) := \tilde{u}(y_0,s_0) + \lambda(t+s_0) + \frac{1}{\varepsilon^2}(|x-y_0|^2 + (t-s_0)^2) + \varepsilon(|x|^2+|y_0|^2).$$

By definition of viscosity subsolutions,

$$\lambda + \frac{2(t_0-s_0)}{\varepsilon^2} + H\left(x_0, \frac{2(x_0-y_0)}{\varepsilon^2} + 2\varepsilon x_0\right) \le 0. \tag{4.10}$$

Similarly, by using the fact that $(y,s) \mapsto \Phi(x_0, y, t_0, s)$ has a maximum at $(y_0, s_0)$, we obtain that

$$-\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} + H\left(y_0, \frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0\right) \geq 0. \tag{4.11}$$

Subtracting (4.11) from (4.10), and using (4.8) and (A1) to get

$$\begin{aligned} 2\lambda &\leq H\left(y_0, \frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0\right) - H\left(x_0, \frac{2(x_0 - y_0)}{\varepsilon^2} + 2\varepsilon x_0\right) \\ &\leq C|x_0 - y_0| \left(1 + \left|\frac{2(x_0 - y_0)}{\varepsilon^2}\right| + 2\varepsilon |y_0|\right) + C\varepsilon |x_0 - y_0|. \end{aligned}$$

In view of (4.9), we let $\varepsilon \to 0$ to discover that $\lambda \leq 0$, which is a contradiction. □

**Remark 4.4.** In Theorem 4.5, we assume that $u, \tilde{u}$ are uniformly continuous just to make the proof simple and clean. In fact, the comparison principle holds for the general case that $u, \tilde{u}$ are a bounded viscosity subsolution in $\mathrm{USC}(\mathbb{R}^n \times [0,T])$, and supersolution in $\mathrm{LSC}(\mathbb{R}^n \times [0,T])$ of (C) on $\mathbb{R}^n \times [0,T]$, respectively. The proof for the general case follows the same philosophy as the above one. We leave this to the interested readers to complete.

By using the comparison principle above, we obtain the following uniqueness result immediately.

**Theorem 4.6** (Uniqueness of viscosity solution)**.** *Under assumption* (A1) *there exists at most one bounded uniformly continuous viscosity solution of* (C) *on* $\mathbb{R}^n \times [0,T]$.

We state here the comparison principles for stationary problems.

**Theorem 4.7.** *Assume that* (A1) *holds. If* $v$, $\tilde{v}$ *are a bounded uniformly continuous viscosity subsolution, and supersolution of*

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

*respectively, then* $v \leq \tilde{v}$ *on* $\mathbb{R}^n$.

**Theorem 4.8.** *Assume that* (A1) *holds. If* $v$, $\tilde{v}$ *are, respectively, a bounded uniformly continuous viscosity subsolution, and supersolution of*

$$H(x, Dv) \leq -\delta, \qquad H(x, D\tilde{v}) \geq 0 \quad \text{in } \mathbb{R}^n$$

*for* $\delta > 0$ *given, then* $v \leq \tilde{v}$ *on* $\mathbb{R}^n$.

Since the proofs of Theorems 4.7, 4.8 are similar to that of Theorem 4.5, we omit them.

## 4.5 Stability

It is really important mentioning that viscosity solutions remain stable under the $L^\infty$-norm. The following proposition shows this basic fact.

**Proposition 4.9.** *Let $\{H_k\}_{k\in\mathbb{N}} \subset C(\mathbb{R}^n \times \mathbb{R}^n)$ and $\{g_k\}_{k\in\mathbb{N}} \subset C(\mathbb{R}^n)$. Assume that $H_k \to H$, $g_k \to g$ locally uniformly in $\mathbb{R}^n \times \mathbb{R}^n$ and in $\mathbb{R}^n$, respectively, as $k \to \infty$ for some $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$ and $g \in C(\mathbb{R}^n)$. Let $\{u_k\}_{k\in\mathbb{N}}$ be viscosity solutions of the Hamilton–Jacobi equations corresponding to $\{H_k\}_{k\in\mathbb{N}}$ with $u_k(\cdot, 0) = g_k$. Assume furthermore that $u_k \to u$ locally uniformly in $\mathbb{R}^n \times [0,\infty)$ as $k \to \infty$ for some $u \in C(\mathbb{R}^n \times [0,\infty))$. Then $u$ is a viscosity solution of* (C).

*Proof.* It is enough to prove that $u$ is a viscosity subsolution of (C). Take $\phi \in C^1(\mathbb{R}^n \times [0,\infty))$ and assume that $u - \phi$ has a strict maximum at $(x_0, t_0) \in \mathbb{R}^n \times (0,\infty)$. By the hypothesis, for $k$ large enough, $u_k - \phi$ has a maximum at some point $(x_k, t_k) \in \mathbb{R}^n \times (0,\infty)$ and $(x_k, t_k) \to (x_0, t_0)$ as $k \to \infty$. By definition of viscosity subsolutions, we have

$$\phi_t(x_k, t_k) + H_k(x_k, D\phi(x_k, t_k)) \le 0.$$

We let $k \to \infty$ to obtain the result. $\square$

We also give useful stability results on supremum of subsolutions and infimum of supersolutions. Let us consider

$$H(x, Dv) = 0 \quad \text{in } \mathbb{R}^n \tag{4.12}$$

for simplicity.

**Proposition 4.10.**
(i) *Let $\mathcal{S}^-$ be a collection of subsolutions of* (4.12). *Define the function $u$ on $\mathbb{R}^n$ by*

$$u(x) := \sup\{v(x) \,:\, v \in \mathcal{S}^-\}.$$

*Assume that $u$ is upper semicontinuous on $\mathbb{R}^n$. Then $u$ is a subsolution of* (4.12).
(ii) *Let $\mathcal{S}^+$ be a collection of supersolutions of* (4.12). *Define the function $u$ on $\mathbb{R}^n$ by*

$$u(x) := \inf\{v(x) \,:\, v \in \mathcal{S}^+\}.$$

*Assume that $u$ is lower semicontinuous on $\mathbb{R}^n$. Then $u$ is a supersolution of* (4.12).

*Proof.* We only prove (ii) since we can prove (i) similarly. Let $\phi \in C^1(\mathbb{R}^n)$, $x_0 \in \mathbb{R}^n$ and $k \in \mathbb{N}$. Assume that $u - \phi$ has a strict minimum at $x_0$. From the definition of $u$, there exists $v_k \in \mathcal{S}^+$ such that $v_k(x_0) - 1/k < u(x_0)$. Due to the lower semicontinuity of $u$, there exists $\delta_k > 0$ such that $u(x_0) < u(x) + 1/k$ for any $x \in B(x_0, \delta_k)$.

Choose a sequence $\{x_k\}_{k\in\mathbb{N}} \subset B(x_0, \delta_k)$ so that $x_k \to x_0$ and $u(x_k) \to u(x_0)$. We have $v_k(x_0) < u(x_k) + 2/k$ for any $k \in \mathbb{N}$. Fix $r > 0$. Let $y_k \in \overline{B}(x_0, r)$ be a minimum point of $v_k - \phi$ over $\overline{B}(x_0, r)$ for each $k \in \mathbb{N}$. Then, we have

$$y_k \to x_0 \quad \text{and} \quad v_k(y_k) \to u(x_0) \quad \text{as } k \to \infty.$$

Indeed, we observe

$$\begin{aligned} u(x_k) + \frac{2}{k} - \phi(x_0) &> (v_k - \phi)(x_0) \ge (v_k - \phi)(y_k) \\ &\ge (u - \phi)(y_k) \ge (u - \phi)(x_0). \end{aligned}$$

From the above, we have $(v_k - \phi)(y_k) \to (u - \phi)(x_0)$ and $(u - \phi)(y_k) \to (u - \phi)(x_0)$ as $k \to \infty$. We consider any convergent subsequence $\{y_{k_j}\}_{j\in\mathbb{N}}$ and $y_0$ denotes its limit point. Noting that $u$ is lower semicontinuous, $(u - \phi)(x_0) = \liminf_{k\to\infty}(u - \phi)(y_k) \ge (u - \phi)(y_0)$, which guarantees $y_k \to x_0$ as $k \to \infty$. Moreover, we get $v_k(y_k) \to u(x_0)$ as $k \to \infty$.

Now, by definition of viscosity supersolutions, we have

$$H(y_k, D\phi(y_k)) \ge 0 \quad \text{for any } k \in \mathbb{N} \text{ large enough.}$$

Sending $k \to \infty$ yields the conclusion. □

Let $\{u_\alpha\}_{\alpha\in\mathbb{R}}$ be a family of locally bounded functions on $\mathbb{R}^n$, and define functions $\overline{u}, \underline{u}$ on $\mathbb{R}^n$ by

$$\overline{u}(x) = \limsup_{\alpha\to\infty}{}^{*} u_\alpha(x) := \lim_{\alpha\to\infty} \sup\{u_\beta(y) \,:\, |x - y| \le 1/\beta,\ \beta \ge \alpha\}, \tag{4.13}$$

$$\underline{u}(x) = \liminf_{\alpha\to\infty}{}_{*} u_\alpha(x) := \lim_{\alpha\to\infty} \inf\{u_\beta(y) \,:\, |x - y| \le 1/\beta,\ \beta \ge \alpha\}. \tag{4.14}$$

We call $\overline{u}$ and $\underline{u}$ the *upper half-relaxed limit* and the *lower half-relaxed limit* of $u_\alpha$ as $\alpha \to \infty$, respectively. Note that $\overline{u}$ and $\underline{u}$ are upper and lower semicontinuous, respectively. We show some stability properties of $\overline{u}$ and $\underline{u}$.

**Theorem 4.11.** *Let $\{H_\alpha\}_{\alpha\in\mathbb{R}} \subset C(\mathbb{R}^n \times \mathbb{R}^n)$. Assume that $H_\alpha \to H$ locally uniformly in $\mathbb{R}^n \times \mathbb{R}^n$ as $\alpha \to \infty$ for some $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$. Let $\{u_\alpha\}_{\alpha\in\mathbb{R}} \subset C(\mathbb{R}^n)$ be a family of locally uniformly bounded functions, which are solutions of* (4.12). *Then the half-relaxed limits $\overline{u}$ and $\underline{u}$ are a subsolution and a supersolution of* (4.12)*, respectively.*

**Lemma 4.12.** *Let $\{u_\alpha\}_{\alpha\in\mathbb{R}} \subset C(\mathbb{R}^n)$ be a family of locally uniformly bounded functions, $\overline{u}$, $\underline{u}$ be the functions defined by* (4.13) *and* (4.14)*, respectively. Assume that $\overline{u} - \varphi$ takes a strict maximum (resp., $\underline{u} - \varphi$ takes a strict minimum) at some $x_0 \in \mathbb{R}^n$ and $\varphi \in C^1(\mathbb{R}^n)$. Then there exist $\{x_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ converging to $x_0$ and $\{\alpha_k\}_{k\in\mathbb{N}}$ converging to infinity such that $u_{\alpha_k} - \varphi$ attains a local maximum (resp., minimum) at $x_k \in \mathbb{R}^n$, and $u_{\alpha_k}(x_k) \to \overline{u}(x_0)$ (resp., $u_{\alpha_k}(x_k) \to \underline{u}(x_0)$) as $k \to \infty$.*

*Proof.* We only deal with the case of $\overline{u}$. Choose $\{y_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ so that $y_k \to x_0$ and $u_{\alpha_k}(y_k) \to \overline{u}(x_0)$. Let $x_k \in \overline{B}(x_0, r)$ be a maximum point of $u_{\alpha_k} - \varphi$ on $\overline{B}(x_0, r)$ for $r > 0$. By replacing the sequence by its subsequence if necessary, we may assume that $x_{k_j} \to \overline{x} \in \overline{B}(x_0, r)$ and $u_{\alpha_{k_j}}(x_{k_j}) \to a \in \mathbb{R}$ as $j \to \infty$. Noting that $u_{\alpha_{k_j}}(y_{k_j}) \le u_{\alpha_{k_j}}(x_{k_j})$, sending $j \to \infty$ yields $(\overline{u} - \varphi)(x_0) \le a - \varphi(\overline{x}) \le (\overline{u} - \varphi)(\overline{x})$. Since $x_0$ is a strict maximum point of $\overline{u} - \varphi$ on $\mathbb{R}^n$, we see that $\overline{x} = x_0$. Moreover, we see that $a = \overline{u}(x_0)$. Therefore, we obtain $u_{\alpha_{k_j}} - \varphi$ attains a local maximum at $x_{k_j}$, $x_{k_j} \to x_0$ and $u_{\alpha_{k_j}}(x_{k_j}) \to \overline{u}(x_0)$ as $j \to \infty$. □

*Proof of Theorem* 4.11. We only prove that $\overline{u}$ is a subsolution of (4.12). Let $\varphi \in C^1(\mathbb{R}^n)$ and $x_0 \in \mathbb{R}^n$. We assume that $\overline{u} - \varphi$ attains a strict maximum at $x_0$ and let $\{x_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, $\{\alpha_k\}_{k\in\mathbb{N}}$ be the sequences obtained in Lemma 4.12. Then by definition of viscosity solutions, we have

$$H(x_k, D\varphi(x_k)) \le 0 \quad \text{for all} \ \ k \in \mathbb{N}.$$

Sending $k \to \infty$, we get $H(x_0, D\varphi(x_0)) \le 0$. □

**Proposition 4.13.** *Let $\{u_\alpha\}_{\alpha\in\mathbb{R}} \subset C(\mathbb{R}^n)$ be uniformly bounded in $\mathbb{R}^n$. Assume that $\overline{u} = \underline{u} =: u$ on $K$ for a compact set $K \subset \mathbb{R}^n$. Then, $u \in C(K)$ and $u_\alpha \to u$ uniformly on $K$ as $\alpha \to \infty$.*

*Proof.* It is clear that $u \in C(K)$. Suppose that $u_\alpha$ does not uniformly converges to $u$ on $K$. Then, there exist $\varepsilon_0 > 0$, a sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ converging to infinity, and a sequence $\{x_k\}_{k\in\mathbb{N}} \subset K$ such that $|u_{\alpha_k}(x_k) - u(x_k)| > \varepsilon_0$ for any $k \in \mathbb{N}$. Since $K$ is a compact set, we can extract a subsequence $\{x_{k_j}\}_{j\in\mathbb{N}}$ such that $x_{k_j} \to x_0 \in K$ as $j \to \infty$. Sending $j \to \infty$ to get

$$\begin{aligned} \varepsilon_0 &\le \limsup_{j\to\infty} (u_{\alpha_{k_j}} - u)(x_{k_j})) \le (\overline{u} - u)(x_0) = 0 \quad \text{or} \\ -\varepsilon_0 &\ge \liminf_{j\to\infty} (u_{\alpha_{k_j}} - u)(x_{k_j})) \ge (\underline{u} - u)(x_0) = 0, \end{aligned}$$

which is a contradiction. □

**Remark 4.5.** The idea of using half-relaxed limits arises naturally when attempting to pass to the limits with maxima and minima. This result is in particular powerful when it can be used with a comparison principle. If the comparison principle holds for the limit equation, then a straightforward consequence of Theorem 4.11, Proposition 4.13 is that $u_\alpha \to u$ locally uniformly on $\mathbb{R}^n$ as $\alpha \to \infty$ for some $u \in C(\mathbb{R}^n)$.

If we have further that $H$ is convex in $p$, then we are able to obtain more stability results. Assume that

(A2) $p \mapsto H(x,p)$ is convex for all $x \in \mathbb{R}^n$.

**Proposition 4.14.** *Assume that* (A2) *holds. Let $u$ be a Lipschitz continuous function on $\mathbb{R}^n$. Then, $u$ is a viscosity subsolution of* (4.12) *if and only if $u$ satisfies $H(x, Du(x)) \le 0$ for almost every $x \in \mathbb{R}^n$.*

*Proof.* In light of Theorem 4.3 and Rademacher's theorem, we can easily see that if $u$ is a viscosity subsolution of (4.12), then $u$ satisfies $H(x, Du(x)) \le 0$ for a.e. $x \in \mathbb{R}^n$. Conversely, if $u$ satisfies $H(x, Du(x)) \le 0$ for a.e. $x \in \mathbb{R}^n$, then by mollification, for each $\varepsilon > 0$, we can construct a smooth function $u^\varepsilon$ satisfying $H(x, Du^\varepsilon) \le C\varepsilon$ in $\mathbb{R}^n$ as in the proof of Proposition 1.5. Furthermore, $u^\varepsilon \to u$ locally uniformly in $\mathbb{R}^n$ as $\varepsilon \to 0$. Thus, in light of the stability result, Proposition 4.9, we obtain the conclusion. □

**Corollary 4.15.** *Assume that* (A2) *holds. Let $u$ be a Lipschitz continuous function on $\mathbb{R}^n$. Then, $u$ is a viscosity solution of* (4.12) *if and only if $H(x,p)=0$ for any $x\in\mathbb{R}^n$, $p\in D^-u(x)$.*

*Proof.* We only need to prove that $u$ is a viscosity subsolution of (4.12) if and only if $H(x,p)\le 0$ for any $x\in\mathbb{R}^n$, $p\in D^-u(x)$. By Proposition 4.14, we have

$$\begin{aligned}
& H(x,Du(x))\le 0 \quad \text{in } \mathbb{R}^n \text{ in the viscosity sense.}\\
\Longleftrightarrow\quad & H(x,Du(x))\le 0 \quad \text{for almost every } x\in\mathbb{R}^n.\\
\Longleftrightarrow\quad & H(x,-Dv(x))\le 0 \quad \text{for almost every } x\in\mathbb{R}^n,\\
& \text{where } v(x)=-u(x) \text{ for all } x\in\mathbb{R}^n.\\
\Longleftrightarrow\quad & H(x,-Dv(x))\le 0 \quad \text{in } \mathbb{R}^n \text{ in the viscosity sense.}\\
\Longleftrightarrow\quad & H(x,-p)\le 0 \quad \text{for any } x\in\mathbb{R}^n, p\in D^+v(x).\\
\Longleftrightarrow\quad & H(x,q)\le 0 \quad \text{for any } x\in\mathbb{R}^n, q\in D^-u(x).
\end{aligned}$$

□

**Corollary 4.16.** *Assume that* (A2) *holds.*
(i) *Let $\mathcal{S}^-$ be a collection of subsolutions of* (4.12)*. Define the function $u$ on $\mathbb{R}^n$ by*

$$u(x):=\inf\{v(x)\,:\,v\in\mathcal{S}^-\}.$$

*Assume that $u$ is Lipschitz continuous on $\mathbb{R}^n$. Then $u$ is a subsolution of* (4.12)*.*
(ii) *Let $\mathcal{S}$ be a collection of solutions of* (4.12)*. Define the function $u$ on $\mathbb{R}^n$ by*

$$u(x):=\inf\{v(x)\,:\,v\in\mathcal{S}\}.$$

*Assume that $u$ is Lipschitz continuous on $\mathbb{R}^n$. Then $u$ is a solution of* (4.12)*.*

Taking Corollary 4.15 into account in the proof of Proposition 4.10, we are able to prove Corollary 4.16 in a similar way. Thus, we omit the proof. Corollaries 4.15, 4.16 were first observed by Barron, Jensen [14].

## 4.6 Lipschitz estimates

We provide here a way to obtain Lipschitz estimates (a priori estimates) for $u^\varepsilon$, which is the solution of $(\mathrm{C})_\varepsilon$ in the introduction of Appendix. Assume for simplicity the followings

(A3) $\displaystyle\lim_{|p|\to\infty}\inf_{x\in\mathbb{R}^n}\big(H(x,p)^2+D_xH(x,p)\cdot p\big)=+\infty.$

(A4) $u_0\in C^2(\mathbb{R}^n)$ and $\|u_0\|_{C^2(\mathbb{R}^n)}\le C<+\infty$.

**Lemma 4.17.** *Assume that* (A3), (A4) *hold. There exists a constant $C>0$ independent of $\varepsilon$ such that*

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

*Sketch of proof.* We first note that for $C>0$ sufficiently large, $u_0 \pm Ct$ are, respectively, a supersolution and a subsolution of $(\mathrm{C})_\varepsilon$. By the comparison principle, we get

$$u_0 - Ct \le u^\varepsilon \le u_0 + Ct.$$

This, together with the comparison principle once more, yields that $\|u_t\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} \le C$.

Next, set $\phi := |Du^\varepsilon|^2/2$. We have

$$\phi_t + D_pH \cdot D\phi + D_xH \cdot Du^\varepsilon = \varepsilon\Delta\phi - \varepsilon|D^2u^\varepsilon|^2.$$

For $\varepsilon>0$ sufficiently small, one has

$$\varepsilon|D^2u^\varepsilon|^2 \ge 2\left(\varepsilon\Delta u^\varepsilon\right)^2 = 2(u_t^\varepsilon + H(x,Du^\varepsilon))^2 \ge H(x,Du^\varepsilon)^2 - C.$$

Thus,

$$\phi_t + D_pH \cdot D\phi + (H(x,Du^\varepsilon)^2 + D_xH\cdot Du^\varepsilon - C) \le \varepsilon\Delta\phi.$$

By the maximum principle and (A3), we get the desired result. □

**Theorem 4.18.** *Assume that* (A1), (A3), (A4) *hold. Then, we obtain*

$$u^\varepsilon \to u \quad \textit{locally uniformly on } \mathbb{R}^n\times[0,\infty)$$

*for* $u \in C(\mathbb{R}^n\times[0,\infty))$*, where* $u$ *is the unique viscosity solution of* (C).

*Proof.* In view of Lemma 4.17, there exists a subsequence $\{u^{\varepsilon_j}\}_{j\in\mathbb{N}}$ such that

$$u^{\varepsilon_j} \to u \quad \text{locally uniformly on } \mathbb{R}^n\times[0,\infty)$$

for some $u \in C(\mathbb{R}^n\times[0,\infty))$, which is Lipschitz. In particular, $u$ is bounded uniformly continuous on $\mathbb{R}^n\times[0,T]$ for each $T>0$.

By Theorems 4.1 and 4.6, we see that $u$ is the unique viscosity solution of (C). This implies further that $u^\varepsilon \to u$ locally uniformly on $\mathbb{R}^n\times[0,\infty)$ as $\varepsilon\to 0$. □

## 4.7 The Perron method

**Theorem 4.19.** *Let* $f$ *and* $g$ *be a subsolution and a supersolution of* (C)*, respectively. Assume that* $f \le g$ *on* $\mathbb{R}^n\times[0,\infty)$*. Then, the function* $u:\mathbb{R}^n\times[0,\infty)\to\mathbb{R}$ *defined by*

$$u(x,t) = \sup\{v(x,t) \,:\, f\le v\le g \text{ on } \mathbb{R}^n\times[0,\infty), v \text{ is a subsolution of (C)}\}$$

*is a solution of* (C)*. Moreover,* $f\le u\le g$ *on* $\mathbb{R}^n\times[0,\infty)$.

The above construction of solutions is called Perron's method. The use of this method in the area of viscosity solutions was introduced by Ishii [47]. For simplicity in this proof, we will assume $u$ is continuous.

*Sketch of proof.* Set

$$\mathcal{S}^- := \{v : \mathbb{R}^n \times [0,\infty) \to \mathbb{R} \,:\, f \le v \le g \text{ in } \mathbb{R}^n \times [0,\infty), v \text{ is a subsolution of (C)}\}.$$

Since $f \in \mathcal{S}^-$, $\mathcal{S}^- \neq \emptyset$. It is clear that $f \le u \le g$ in $\mathbb{R}^n \times [0,\infty)$. Thus, $u$ is locally bounded in $\mathbb{R}^n \times [0,\infty)$ and a subsolution of (C) by Proposition 4.10.

The proof is completed by showing that $u$ is a supersolution of (C). To do this, we argue by contradiction, and therefore we suppose that $u$ is not a supersolution of (C). Then we may choose a function $\phi \in C^1(\mathbb{R}^n \times (0,\infty))$ such that $u - \phi$ attains a strict minimum at some $(y,s) \in \mathbb{R}^n \times (0,\infty)$ and $\phi_t(y,s) + H(y, D\phi(y,s)) < 0$. We may assume that $(u-\phi)(y,s) = 0$ by adding some constant to $\phi$.

We now prove that $u(y,s) = \phi(y,s) < g(y,s)$. Noting that $u \le g$, we deduce $\phi \le g$ in $\mathbb{R}^n \times (0,\infty)$. Assume that $u(y,s) = g(y,s)$. Then $g - \phi$ attains a minimum at $(y,s)$. Noting that $g$ is a supersolution of (C), we obtain $0 \le \phi_t(y,s) + H(y, D\phi(y,s))$, which contradicts the above.

Set $\varepsilon_0 = (g-\phi)(y,s) > 0$. By the continuity of $g - \phi$ and $H$, there exists $r > 0$ such that

$$\begin{aligned}
&g(x,t) \ge \phi(x,t) + \frac{\varepsilon_0}{2} &&\text{for all } (x,t) \in B(y,r) \times (s-r, s+r),\\
&\phi_t(x,t) + H(x, D\phi(x,t)) \le 0 &&\text{for all } (x,t) \in B(y,r) \times (s-r, s+r).
\end{aligned}$$

Set $U = B(y,r) \times (s-r,s+r)$ and $\varepsilon = \frac{1}{2}\min\{\varepsilon_0, \min_{\partial U}(u-\phi)(x,t)\} > 0$. We define the function $\tilde{u} : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ by

$$\tilde{u}(x,t) := \begin{cases} \max\{u(x,t), \phi(x,t) + \varepsilon\} & \text{for } (x,t) \in B(y,r) \times (s-r,s+r), \\ u(x,t) & \text{otherwise.} \end{cases}$$

It is clear that $\tilde{u} \in C(\mathbb{R}^n \times [0,\infty))$, $u \le \tilde{u}$ and $f \le \tilde{u} \le g$ in $\mathbb{R}^n \times [0,\infty)$. Besides, $\tilde{u}(y,s) > u(y,s)$ and $\tilde{u}$ is a subsolution of (C), which contradicts the definition of $u$. $\square$

# 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 eqautions with Neumann type boundary conditions*, Proc. Roy. Soc. Edinburgh Sect. A **146** (2016) 225–242.

[2] N. Anantharaman, *On the zero-temperature or vanishing viscosity limit for certain Markov processes arising from Lagrangian dynamics*, J. Eur. Math. Soc. **6** (2004), no. 2, 207–276.

[3] 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.

[4] S. N. Armstrong, H. V. Tran, *Viscosity solutions of general viscous Hamilton–Jacobi equations*, Mathematische Annalen, **361** (2015), no. 3, 647–687.

[5] 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.

[6] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Mathématiques & Applications (Berlin), **17**, Springer-Verlag, Paris, 1994.

[7] G. Barles, H. Ishii, H. Mitake, *On the large time behavior of solutions of Hamilton–Jacobi equations associated with nonlinear boundary conditions*, Arch. Ration. Mech. Anal. **204** (2012), no. 2, 515–558.

[8] G. Barles, H. Ishii, H. Mitake, *A new PDE approach to the large time asymptotics of solutions of Hamilton-Jacobi equation*, Bull. Math. Sci. **3** (2013), 363–388.

[9] G. Barles, H. Mitake, *A PDE approach to large-time asymptotics for boundary-value problems for nonconvex Hamilton-Jacobi equations*, Comm. Partial Differential Equations **37** (2012), no. 1, 136–168.

[10] G. Barles, A. Porretta, T. T. Tchamba, *On the large time behavior of solutions of the Dirichlet problem for subquadratic viscous Hamilton-Jacobi equations*, J. Math. Pures Appl. (9) **94** (2010), no. 5, 497–519.

[11] G. Barles, J.-M. Roquejoffre, *Ergodic type problems and large time behaviour of unbounded solutions of Hamilton-Jacobi equations*, Comm. Partial Differential Equations **31** (2006), no. 7-9, 1209–1225.

[12] G. Barles, P. E. Souganidis, *On the large time behavior of solutions of Hamilton–Jacobi equations*, SIAM J. Math. Anal. **31** (2000), no. 4, 925–939.

[13] G. Barles, P. E. Souganidis, *Space-time periodic solutions and long-time behavior of solutions to quasi-linear parabolic equations*, SIAM J. Math. Anal. **32** (2001), no. 6, 1311–1323.

[14] 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.

[15] P. Bernard, J.-M. Roquejoffre, *Convergence to time-periodic solutions in time-periodic Hamilton-Jacobi equations on the circle*, Comm. Partial Differential Equations **29** (2004), no. 3-4, 457–469.

[16] U. Bessi, *Aubry-Mather theory and Hamilton-Jacobi equations*, Comm. Math. Phys. **235** (2003), no. 3, 495–511.

[17] 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.

[18] 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.

[19] F. Camilli, O. Ley, P. Loreti, and V. D. Nguyen, *Large time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations*, NoDEA Nonlinear Differential Equations Appl., **19** (6), (2012), 719–749.

[20] 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.

[21] A. Cesaroni, M. Novaga, *Long-time behavior of the mean curvature flow with periodic forcing*, Comm. Partial Differential Equations **38** (2013), no. 5, 780–801.

[22] 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.

[23] 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.

[24] M. G. Crandall, P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **277** (1983), 1–42.

[25] A. Davini, A. Siconolfi, *A generalized dynamical approach to the large-time behavior of solutions of Hamilton–Jacobi equations*, SIAM J. Math. Anal. **38** (2006), no. 2, 478–502.

[26] A. Davini, A. Fathi, R. Iturriaga, M. Zavidovique, *Convergence of the solutions of the discounted equation*, Invent. Math. **206** (1) (2016) 29–55.

[27] A. Davini, M. Zavidovique, *Aubry sets for weakly coupled systems of Hamilton-Jacobi equations*, SIAM J. Math. Anal. 46 (2014), no. 5, 3361–3389.

[28] R. J. Di Perna, P.-L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math. **98**, 511–547 (1989).

[29] W. E, *Aubry-Mather theory and periodic solutions of the forced Burgers equation*, Comm. Pure Appl. Math. **52** (1999), no. 7, 811–828.

[30] L. C. Evans, *Weak convergence methods for nonlinear partial differential equations.* CBMS 74, American Mathematical Society, 1990.

[31] L. C. Evans, *Partial differential equations. Second edition.* Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010.

[32] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton–Jacobi PDE*, Archive for Rational Mechanics and Analysis **197** (2010), 1053–1088.

[33] L. C. Evans, *Towards a quantum analog of weak KAM theory*, Comm. Math. Phys. **244** (2004), no. 2, 311–334.

[34] A. Fathi, *Weak KAM Theorem in Lagrangian Dynamics*, to appear in Cambridge Studies in Advanced Mathematics.

[35] A. Fathi, A. Siconolfi, *Existence of $C^1$ critical subsolutions of the Hamilton-Jacobi equation*, Invent. Math. **155** (2004), no. 2, 363–388.

[36] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Stochastic Modelling and Applied Probability, **25**. Springer, New York, 2006.

[37] Y. Fujita, H. Ishii, P. Loreti, *Asymptotic solutions of viscous Hamilton-Jacobi equations with Ornstein-Uhlenbeck operator*, Comm. Partial Differential Equations **31** (2006), no. 4-6, 827–848.

[38] Y. Giga, *Surface Evolution Equations. A Level Set Approach*, Monographs in Mathematics, **99**. Birkhäuser, Basel-Boston-Berlin (2006), xii+264pp.

[39] D. A. Gomes, *A stochastic analogue of Aubry-Mather theory*, Nonlinearity **15** (2002), no. 3, 581–603.

[40] 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.

[41] D. A. Gomes, *Generalized Mather problem and selection principles for viscosity solutions and Mather measures*, Adv. Calc. Var., 1 (2008), 291–307.

[42] 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.

[43] D. Gomes, H. Mitake, H. V. Tran, *The Selection problem for discounted Hamilton-Jacobi equations: some non-convex cases*, arXiv:1605.07532 [math.AP], submitted.

[44] N. Ichihara *Large time asymptotic problems for optimal stochastic control with superlinear cost*, Stochastic Process. Appl. **122** (4) (2012), 1248–1275.

[45] N. Ichihara, H. Ishii, *Long-time behavior of solutions of Hamilton-Jacobi equations with convex and coercive Hamiltonians*, Arch. Ration. Mech. Anal. **194** (2009), no. 2, 383–419.

[46] N. Ichihara, S.-J. Sheu, *Large time behavior of solutions of Hamilton-Jacobi-Bellman equations with quadratic nonlinearity in gradients*, SIAM J. Math. Anal. **45** (2013), no. 1, 279–306.

[47] H. Ishii, *Perron's method for Hamilton–Jacobi equations*, Duke Math Journal, **55** (1987), no. 2, 369–384.

[48] H. Ishii, *On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions*, Funkcial. Ekvac. **38** (1995), no. 1, 101–120.

[49] H. Ishii, *Asymptotic solutions for large-time of Hamilton–Jacobi equations in Euclidean n space*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **25** (2008), no 2, 231–266.

[50] H. Ishii, *Long-time asymptotic solutions of convex Hamilton-Jacobi equations with Neumann type boundary conditions*, Calc. Var. Partial Differential Equations **42** (2011), no. 1-2, 189–209.

[51] 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., to appear.

[52] H. Ishii, H. Mitake, H. V. Tran, *The vanishing discount problem and viscosity Mather measures. Part 2: boundary value problems*, Journal Math. Pures Appl., to appear.

[53] R. Iturriaga, H. Sanchez-Morgado, *Limit of the infinite horizon discounted Hamilton–Jacobi equation*, Discrete Contin. Dyn. Syst. Ser. B, **15** (2011), 623–635.

[54] H. R. Jauslin, H. O. Kreiss and J. Moser, *On the forced Burgers equation with periodic boundary conditions*, Proc. Symp. Pure Math. **65** (1999), 133–153.

[55] R. Jensen, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rat. Mech. Anal. **101** (1988), 1–27.

[56] W.-L. Jin, Y. Yu, *Asymptotic solution and effective Hamiltonian of a Hamilton–Jacobi equation in the modeling of traffic flow on a homogeneous signalized road*, J. Math. Pures Appl. (9) **104** (2015), no. 5, 982–1004.

[57] 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.

[58] O. Ley, V. D. Nguyen, *Large time behavior for some nonlinear degenerate parabolic equations*, J. Math. Pures Appl. **102** (2014), 293–314.

[59] P.-L. Lions, *Control of Diffusion Processes in* $\mathbb{R}^N$, Comm. Pure Appl. Math. **34** (1981), 121–147.

[60] P.-L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics, Vol. **69**, Pitman, Boston, Masso. London, 1982.

[61] P.-L., Lions, G. Papanicolaou, S. R. S. Varadhan, *Homogenization of Hamilton–Jacobi equations*, unpublished work (1987).

[62] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996), no. 2, 273–310.

[63] J. N. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. **207** (1991), no. 2, 169–207.

[64] H. Mitake, *Asymptotic solutions of Hamilton-Jacobi equations with state constraints*, Appl. Math. Optim. **58** (2008), no. 3, 393–410.

[65] H. Mitake, *The large-time behavior of solutions of the Cauchy-Dirichlet problem for Hamilton-Jacobi equations*, NoDEA Nonlinear Differential Equations App. **15** (2008), no. 3, 347–362.

[66] H. Mitake, A. Siconolfi, H. V. Tran, N. Yamada, *A Lagrangian approach to weakly coupled Hamilton-Jacobi systems*, SIAM J. Math. Anal. 48 (2016), no. 2, 821–846.

[67] H. Mitake, K. Soga, *Weak KAM theory for discount Hamilton-Jacobi equations and its application*, preprint.

[68] H. Mitake, H. V. Tran, *Remarks on the large time behavior of viscosity solutions of quasi-monotone weakly coupled systems of Hamilton–Jacobi equations*, Asymptot. Anal. **77** (2012), 43–70.

[69] H. Mitake, H. V. Tran, *A dynamical approach to the large-time behavior of solutions to weakly coupled systems of Hamilton–Jacobi equations*, J. Math. Pures Appl. **101** (2014) 76–93.

[70] H. Mitake, H. V. Tran, *Large-time behavior for obstacle problems for degenerate viscous Hamilton-Jacobi equations*, Calc. Var. Partial Differential Equations **54** (2015) 2039–2058.

[71] H. Mitake, H. V. Tran, *Selection problems for a discount degenerate viscous Hamilton–Jacobi equation* , Adv. Math., **306** (2017), 684–703.

[72] 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.

[73] V. D. Nguyen, *Some results on the large time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations*, J. Evol. Equ. **14** (2014), 299–331.

[74] M. Nisio, *Stochastic Control Theory: Dynamic Programming Principle*, Probability Theory and Stochastic Modelling, vol. 72. Springer, Tokyo, 2015. xvi+250 pp.

[75] K. Soga, *Selection problems of $\mathbb{Z}^2$-periodic entropy solutions and viscosity solutions*, preprint (arXiv:1501.03594).

[76] D. W. Stroock and S. R. S. Varadhan, Multidimensional diffusion processes. Reprint of the 1997 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2006. xii+338 pp.

[77] T. T. Tchamba, *Large time behavior of solutions of viscous Hamilton-Jacobi equations with superquadratic Hamiltonian*, Asymptot. Anal. **66** (2010), no. 3-4, 161–186.

[78] H. V. Tran, *Adjoint methods for static Hamilton-Jacobi equations*, Calc. Var. Partial Differential Equations **41** (2011), 301–319.

[79] 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.