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

Hiroyoshi Mitake

Institute for Sustainable Sciences and Development, Hiroshima University
1-4-1 Kagamiyama, Higashi-Hiroshima-shi 739-8527, Japan

Hung V. Tran

Department of Mathematics, The University of Chicago, 5734 S. University Avenue, Chicago, Illinois 60637, USA

July 21, 2015

1Email address: hiroyoshi-mitake@hiroshima-u.ac.jp
2Email address: hung@math.uchicago.edu
# Contents

1 Introduction .......................................................... 5

2 Ergodic problems ................................................... 7
   2.1 Motivation .................................................. 7
   2.2 Existence of ergodic problems ............................ 8

3 Large time asymptotics for HJ equations ....................... 15
   3.1 Separable Hamiltonians .................................. 16
   3.2 General Hamiltonians .................................... 18
      3.2.1 Formal calculation ................................. 19
      3.2.2 Regularizing Process ............................. 20
      3.2.3 Conservation of energy and a key observation .... 22
      3.2.4 Proof of key estimates ........................... 25
   3.3 Degenerate viscous Hamilton–Jacobi equations .......... 27
   3.4 Asymptotic profile ....................................... 31
   3.5 Viscous case ............................................. 35
   3.6 Some other directions and open questions ............... 36

4 Discounted approximation procedure ............................ 37
   4.1 Selection problems ....................................... 37
   4.2 Key ingredients and estimates ........................... 40
      4.2.1 Regularization process and construction of \( \mathcal{M} \) .... 40
      4.2.2 Key estimates ....................................... 43
   4.3 Proof of the commutation lemma .......................... 45
   4.4 Proof of Theorem 4.1 ..................................... 49
      4.4.1 Applications ....................................... 50
   4.5 Some other directions and open questions ............... 51

5 Appendix ............................................................. 53
   5.1 Motivation and Examples ................................ 54
      5.1.1 Front Propagation Problems ........................ 54
      5.1.2 Optimal Control Problems .......................... 55
   5.2 Definitions ................................................ 57
   5.3 Existence .................................................. 58
5.4 Consistency ................................................. 59
5.5 Stability ..................................................... 60
5.6 Uniqueness .................................................. 60
5.7 Lipschitz estimates ................................. 62
5.8 The Perron method ....................................... 63
Chapter 1

Introduction

This is a lecture note based on the two courses given by the two authors at the summer school on “PDE and Applied Mathematics” at Vietnam Institute for Advanced Study in Mathematics (VIASM) on 14–25/July/2014. In the first course, we studied a basic theory of viscosity solutions, and in the second course, we studied the asymptotic analysis of the Hamilton–Jacobi equation. Particularly, we focused on the large time asymptotics.

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

\[ u_t(x, t) + H(x, Du(x, t)) = 0 \quad \text{for } x \in \mathbb{R}^n, t > 0, \tag{1.1} \]

and the viscous Hamilton–Jacobi equation:

\[ u_t(x, t) - \Delta u(x, t) + H(x, Du(x, t)) = 0 \quad \text{for } x \in \mathbb{R}^n, t > 0, \tag{1.2} \]

where \( u_t, Du, D^2u \) denote the partial derivative with respect to \( t \), the spatial gradient and Hessian of the unknown \( u : T^n \times (0, \infty) \to \mathbb{R} \), respectively. The function \( H : T^n \times \mathbb{R}^n \to \mathbb{R} \) is a given continuous and coercive Hamiltonian. We will add suitable assumptions later when needed. At some points, we consider the general viscous Hamilton–Jacobi equation:

\[ u_t(x, t) - \text{tr} \left( A(x) D^2 u(x, t) \right) + H(x, Du(x, t)) = 0 \quad \text{for } x \in \mathbb{R}^n, t > 0, \tag{1.3} \]

where \( A : \mathbb{R}^n \to \mathcal{M}^{n \times n}_{\text{sym}} \) is a given continuous diffusive matrix, which is nonnegative and could be degenerate. Here \( \mathcal{M}^{n \times n}_{\text{sym}} \) is the set of \( n \times n \) real symmetric matrices.

In the last decade, there has been much interest on dynamical properties of viscosity solutions of (1.1)–(1.3). Indeed, in view of the weak Kolmogorov–Arnold–Moser (KAM) theory established by Fathi (see [26]), the asymptotic analysis on the Hamilton–Jacobi equation (1.1) with convex Hamiltonian has been dramatically developed. One of the feature of this lecture note is to introduce a new way to investigate dynamical properties of solutions of (1.1)–(1.3) and related equations by using PDE methods. More precisely, we use the nonlinear adjoint method introduced by Evans [25] together with some new conserved quantities and estimates. 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. The authors with Cagnetti, Gomes used this method to study the large-time behavior of solutions to (1.3). Another interesting topic is about selection problems in the discounted approximation setting. This was studied very recently by Davini, Fathi, Iturriaga, Zavidovique [22], and Mitake, Tran [56].

The outline of the lecture note is as follows: in Chapter 2 we investigate the ergodic problems associated with (1.1)–(1.3). In particular, we prove the existence of solutions. In Chapters 3 and 4, we study the large time behavior of solutions to (1.1)–(1.3), and the selection problem for the discounted approximation. To make the lecture note self-contained, we prepare a brief introduction of the theory of viscosity solutions in Appendix. Appendix can be read independently from other Chapters. Also, we note that Chapters 3 and 4 can be read independently.

Acknowledgement. The work of HM was partially supported by KAKENHI #15K17574, #26287024 and the work of HT was partially supported by NSF grant DMS-1361236. This lecture note was started while the authors visited Vietnam Institute for Advanced Study in Mathematics (VIASM) in July 2014. The authors are grateful for all of their hospitality.
Chapter 2

Ergodic problems for
Hamilton–Jacobi equations

2.1 Motivation

One of our main goals of this lecture note is to understand the large-time behavior of solution $u$ to

$$\begin{cases}
    u_t + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\
    u(x, 0) = u_0(x) & \text{on } \mathbb{R}^n,
\end{cases} \tag{2.1}$$

and

$$\begin{cases}
    u_t - \Delta u + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\
    u(x, 0) = u_0(x) & \text{on } \mathbb{R}^n,
\end{cases} \tag{2.2}$$

under rather general assumptions on $H$. More generally, we consider

$$\begin{cases}
    u_t - \text{tr} \left( A(x) D^2 u \right) + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\
    u(x, 0) = u_0(x) & \text{on } \mathbb{R}^n,
\end{cases} \tag{2.3}$$

We want to know what happens for $u(x, t)$ when we send $t \to \infty$. In this section we first give a heuristic formal argument to find what is a candidate of the large-time asymptotics of $u$. Let us work with (2.1) for now.

We always assume hereinafter the coercivity condition, i.e.,

$$H(x, p) \to \infty \quad \text{as } |p| \to \infty \text{ locally uniformly for } x \in \mathbb{R}^n. \tag{2.4}$$

Let us consider a formal asymptotic expansion:

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

with smooth functions $a_i$. Plug this into equation (2.1) 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 (2.4), we should have $Da_1 \equiv 0$, which therefore implies $a_1$ should be a constant. From this observation, we can expect that the large-time behavior of the solution to (2.1) is

$$u(x, t) - (v(x) - ct) \to 0 \quad \text{locally uniformly for } x \in \mathbb{R}^n \text{ as } t \to \infty \quad (2.5)$$

for some function $v$ and constant $c$. Moreover, if convergence (2.5) holds, then by the stability result of viscosity solutions, $(v, c)$ satisfies

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

Therefore, if we expect such a convergence (2.5), then we should first investigate the existence of $(v, c)$ satisfying the above problem, which will be obtained in the next section.

**Remark 2.1.** One may wonder why we do not consider terms like $a_i(x)t^i$ for $i \geq 2$ in the above asymptotic expansion. We will see the reason why in the end of this chapter.

### 2.2 Existence of ergodic problems

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

In this section, we consider the *ergodic problems* for inviscid (first order) Hamilton–Jacobi equation

$$H(x, Dv) = c \quad \text{in } \mathbb{T}^n, \quad (2.6)$$

and viscous Hamilton–Jacobi equation

$$-\Delta v + H(x, Dv) = c \quad \text{in } \mathbb{T}^n. \quad (2.7)$$

In both cases, we seek for a pair of unknowns $(v, c) \in C(\mathbb{T}^n) \times \mathbb{R}$. More generally, we consider

$$-\text{tr} \left( A(x) D^2 v(x) \right) + H(x, Dv) = c \quad \text{in } \mathbb{T}^n. \quad (2.8)$$

We give three results on the existence of (2.6)–(2.8). The last one includes the first and second results but we give all of them separately for readability.

**Theorem 2.1.** Assume that $H$ satisfies (2.4). There exists $(v, c) \in \text{Lip} (\mathbb{T}^n) \times \mathbb{R}$ which satisfies (2.6) in the viscosity solution sense.

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

$$\delta v^\delta + H(x, Dv^\delta) = 0 \quad \text{in } \mathbb{T}^n. \quad (2.9)$$

Setting $M := \max_{x \in \mathbb{T}^n} |H(x, 0)|$, we have $\pm M/\delta$ is a subsolution and supersolution of (2.9) respectively. By the Perron method in the theory of viscosity solutions, there exists a unique viscosity solution to (2.9) such that

$$|v^\delta| \leq M/\delta,$$
2.2. EXISTENCE OF ERGODIC PROBLEMS

which implies \( H(x, Dv^\delta) \leq M \). In view of the coercivity (2.4), we get \(|Dv^\delta| \leq C\) for some \( C > 0 \). Therefore, we obtain that \( \{v^\delta(\cdot) - v^\delta(x_0)\}_{\delta > 0} \) is uniformly bounded and equi-Lipschitz continuous for a fix \( x_0 \in \mathbb{T}^n \). Thus, in light of the Arzelà–Ascoli theorem, there exists a subsequence \( \{\delta_j\}_j \) so that \( v^\delta(\cdot) - v^\delta(x_0) \to v \) uniformly on \( \mathbb{T}^n \) as \( j \to \infty \). Also we have \( \delta_j v^\delta(x_0) \to -c \) for some \( c \in \mathbb{R} \). In view of the stability result of viscosity solutions, we get the conclusion.

\[ \square \]

**Remark 2.2.** Let us notice that the approximation procedure above using (2.9) is called the *discounted approximation procedure*. It is a very natural procedure in many ways. Firstly, the approximation makes (2.9) monotone in \( v^\delta \), which fits perfectly in the well-posedness setting of viscosity solutions. Secondly, we let \( w(x) = v^\delta(x/\delta) \) to yield that

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

which is precisely a homogenization problem.

The arguments in the proof of Theorem 2.1 are soft as we just use a priori estimates on \(|Dv^\delta|\) and the Arzelà–Ascoli theorem to get the result. In particular, from this argument, we only know the convergence of \( v^\delta_j - v^\delta(x_0) \) via the subsequence \( \{\delta_j\}_j \). It was not clear at all at this moment whether \( v^\delta - v^\delta(x_0) \) converges uniformly as \( \delta \to 0 \) or not. We will come back to this question and answer positively in Chapter 4.

Let us now provide the existence proof for the viscous case. We need a sort of superlinearity:

\[ \lim_{|p| \to \infty} \left( \frac{1}{2n} H(x, p)^2 + D_x H(x, p) \cdot p \right) = +\infty \quad \text{uniformly} \quad x \in \mathbb{T}^n, \quad (2.10) \]

**Theorem 2.2.** Assume that \( H \) satisfies (2.10). There exists \( (v, c) \in C^2(\mathbb{T}^n) \times \mathbb{R} \) which satisfies (2.7).

**Proof.** Consider the approximate problem

\[ \delta v^\delta - \Delta v^\delta + H(x, Dv^\delta) = 0 \quad \text{in} \quad \mathbb{T}^n \]

for \( \delta > 0 \). By the same argument in as in the proof of Theorem 2.1, there exists a solution \( v^\delta \) to the above. Note that in this case \( v^\delta \) is smooth due to the appearance of the diffusion \( \Delta v^\delta \).

Set \( \varphi := |Dv^\delta|^2/2 \). Then \( \varphi \) satisfies

\[ 2\delta \varphi - (\Delta \varphi - |D^2 v^\delta|^2) + D_x H \cdot Dv^{\alpha, \delta} + D_p H \cdot D \varphi = 0. \]

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

\[ |D^2 v^\delta|^2 + D_x H \cdot Dv^{\alpha, \delta} \leq 0. \]

Noting furthermore that

\[ |D^2 v^\delta|^2 \geq \frac{1}{n} |\Delta v^\delta|^2 \geq \frac{1}{2n} H(x, Dv^\delta)^2 - C \]
for some $C > 0$. Thus,
\[
\frac{1}{2n} H(x, Dv^\delta)^2 + D_x H(x, Dv^\delta) \cdot Dv^\delta \leq C.
\]
In light of (2.10), we get the conclusion. 

Here is a generalization of Theorems 2.1 and 2.2. We use the assumptions:

(H1) $A(x) = (a^{ij}(x)) \in M^{n \times n}_{\text{sym}}$ with $A(x) \geq 0$, and $A \in C^2(\mathbb{T}^n)$,

and there exists $\gamma > 1$ and $C > 0$ such that
\[
\begin{align*}
|D_x H(x, p)| &\leq C(1 + |p|^{\gamma}) \quad \text{for all } (x, p) \in \mathbb{R}^n \times \mathbb{R}^n, \\
\lim_{|p| \to \infty} \frac{|H(x, p)|^2}{|p|^{1+\gamma}} &= +\infty \quad \text{uniformly for } x \in \mathbb{T}^n.
\end{align*}
\]
(2.11)

We remark that (2.11) is also a sort of superlinearity condition. It is clear that (2.11) is stronger than (2.10). We need to require a bit more as we deal with the general diffusive matrix $A$ here.

**Theorem 2.3.** Assume that $A$ and $H$ satisfy (H1) and (2.11). There exists $(v, c) \in \text{Lip}(\mathbb{T}^n) \times \mathbb{R}$ which satisfies (2.8) in the viscosity solution sense.

**Proof.** The proof is based on the standard Bernstein method. For $\alpha > 0, \delta > 0$, consider the equation
\[
\alpha v^{\alpha, \delta} - \text{tr} \left( A(x) D^2 v^{\alpha, \delta} \right) + H(x, Dv^{\alpha, \delta}) = \delta \Delta v^{\alpha, \delta} \quad \text{in } \mathbb{T}^n.
\]
Owing to the discount and viscosity terms, there exists a (unique) classical solution $v^{\alpha, \delta}$. We easily see that $|\alpha v^{\alpha, \delta}| \leq C$ for $C > 0$ by the comparison principle. Set $\varphi := |Dv^{\alpha, \delta}|^2 / 2$ and then $\varphi$ satisfies
\[
2\alpha \varphi - a^{ij}_{x_k} v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_j} + \frac{1}{2} (\varphi_{x_i} - v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_i}) + D_x H \cdot Dv^{\alpha, \delta} + D_p H \cdot D\varphi = \delta (\Delta \varphi - |D^2 v^{\alpha, \delta}|^2).
\]

Take a point $x_0$ such that $\varphi(x_0) = \max_{\mathbb{T}^n} \varphi \geq 0$ and note that at that point
\[
- a^{ij}_{x_k} v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_j} + D_x H \cdot Dv^{\alpha, \delta} + a^{ij} v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_j} + \delta |D^2 v^{\alpha, \delta}|^2 \leq 0. \quad (2.12)
\]
The two terms $a^{ij} v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_j}$ and $\delta |D^2 v^{\alpha, \delta}|^2$ are the good terms, which will help us to control other terms and to deduce the result.

We first use the trace inequality (see [60, Lemma 3.2.3] for instance),
\[
(\text{tr} \left( A_{x_k} S \right))^2 \leq C \text{tr} \left( S^2 \right) \quad \text{for all } S \in M^{n \times n}_{\text{sym}}, \ 1 \leq k \leq n, \quad (2.13)
\]
for some constant $C$ depending only on $n$ and $\|D^2 A\|_{L^\infty(\mathbb{T}^n)}$ to yield that, for some small constant $c > 0$,
\[
a^{ij}_{x_k} v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_j} = \text{tr} \left( A_{x_k} D^2 v^{\alpha, \delta} \right) v^{\alpha, \delta}_{x_k} \leq c \left( \text{tr} \left( A_{x_k} D^2 v^{\alpha, \delta} \right) \right)^2 + \frac{C}{c} |Dv^{\alpha, \delta}|^2
\leq \frac{1}{2} \text{tr} \left( D^2 v^{\alpha, \delta} A D^2 v^{\alpha, \delta} \right) + |Dv^{\alpha, \delta}|^2 = \frac{1}{2} a^{ij}_{x_k} v^{\alpha, \delta}_{x_i} v^{\alpha, \delta}_{x_j} + C |Dv^{\alpha, \delta}|^2. \quad (2.14)
\]
2.2. EXISTENCE OF ERGODIC PROBLEMS

Next, since $A$ is a symmetric positive definite matrix, it can be diagonalized as $A = P^TDP$ where $D$ is the diagonal matrix, which could be written as $D = \text{diag}\{d^1, \ldots, d^n\}$ with $d^i \geq 0$, and $P^TP = I_n$. We have

$$\left( a_{ij} v_{x_ix_j}^{\alpha,\delta} \right)^2 = \left( p^{mi} p^{mj} d^{mk} v^{\alpha,\delta}_{x_ix_j} \right)^2 \leq \left( \sum_j C \left| p^{mi} d^{mk} v^{\alpha,\delta}_{x_ix_j} \right| \right)^2 \leq C \sum_{j,m} \left| \sum_i p^{mi} \sqrt{d^{mk}} v^{\alpha,\delta}_{x_ix_j} \right|^2 = C \sum_{j,m} p^{mi} \sqrt{d^{mk}} v^{\alpha,\delta}_{x_ix_j} p^{mk} \sqrt{d^{mk}} v^{\alpha,\delta}_{x_kx_j} = C p^{mi} p^{mk} d^{mk} v^{\alpha,\delta}_{x_ix_j} v^{\alpha,\delta}_{x_kx_j} = C a^{ik} v^{\alpha,\delta}_{x_ix_j} v^{\alpha,\delta}_{x_kx_j}. \quad (2.15)$$

In light of (2.15), for some $c_0 > 0$ sufficiently small,

$$\frac{1}{2} a_{ij} v^{\alpha,\delta}_{x_ix_k} v^{\alpha,\delta}_{x_jx_k} + \delta |D^2 v^{\alpha,\delta}|^2 \geq 4c_0 \left( \left( a_{ij} v^{\alpha,\delta}_{x_ix_j} \right)^2 + (\delta \Delta v^{\alpha,\delta})^2 \right) \geq 2c_0 \left( a_{ij} v^{\alpha,\delta}_{x_ix_j} + \delta \Delta v^{\alpha,\delta} \right)^2 = 2c_0(\alpha v^{\alpha,\delta} + H(x,Dv^{\alpha,\delta}))^2 \geq c_0 H(x,Dv^{\alpha,\delta})^2 - C. \quad (2.16)$$

Combining (2.12), (2.14), and (2.16) 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 (2.11) 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, setting $w^{\alpha,\delta}(x) := v^{\alpha,\delta}(x) - v^{\alpha,\delta}(0)$, by passing some subsequences if necessary, we can send $\delta$ and $\alpha$ to 0 in this order to yield that $w^{\alpha,\delta}$ and $\alpha v^{\alpha,\delta}$, respectively, uniformly converge $v$ and $-c$ which satisfies (2.8) in the viscosity sense.

We call a function $v$ and a constant $c$ which we get in Theorems 2.1–2.3 an \textit{ergodic function} and an \textit{ergodic constant} for the inviscid (reps., viscous) Hamilton–Jacobi equation, respectively. We now proceed to show that ergodic constant $c$ is unique. It is enough to consider the most general case (2.8).

**Proposition 2.4.** Assume that (H1) and (2.11) hold. The ergodic problem (2.8) admits an unique ergodic constant $c$, which is uniquely determined by $A$ and $H$.

**Proof.** Suppose that there exist two solution $(v_1, c_1), (v_2, c_2)$ to (2.8) with $c_1 \neq c_2$. We may assume that $c_1 < c_2$. Note that $v_1 - c_1 t - M$ and $v_2 - c_2 t + M$ are a subsolution and a supersolution to (2.3) respectively for a suitable large $M > 0$. By the comparison principle for (2.3) we get

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

Thus, $(c_2 - c_1) t \leq M'$ for some $M' > 0$ and all $t \in (0, \infty)$, which yields a contradiction. \qed
Remark 2.3. As in Proposition 2.4, an ergodic constant is unique but on the other hand, an ergodic function is not unique in general. Look at Example 5.1 and Section 2.2.

The ergodic constant $c$ is very related to the effective Hamiltonian $H$ in the homogenization theory. In fact, $c = H(0)$.

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

Proposition 2.5. Assume that (H1) and (2.11) hold. The ergodic constant for (2.8) can be represented by

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

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

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

Proof. We only consider the inviscid case and assume $H$ is convex in $p$ here. We leave the proof of the general case for the readers. Let us denote by $d_1, d_2$ the two above formulas in statement of the proposition. For any fixed $\phi \in C^1(\mathbb{T}^n)$, $d_1 \leq \sup_x H(x, D\phi)$, which implies $d_1 \leq d_2$.

Now take $v$ to be a solution of (2.6). Take $\rho$ to be a standard mollifier and set $\rho_\varepsilon = \varepsilon^{-n} \rho(\cdot/\varepsilon)$ for any $\varepsilon > 0$. Denote by $v_\varepsilon = \rho_\varepsilon * v$. In light of the convexity of $H$,

$$H(x, Dv_\varepsilon) \leq c + C\varepsilon \text{ in } \mathbb{T}^n.$$ 

Letting $\varepsilon \to 0$ to get $d_2 \leq c$.

Finally suppose that $c > d_1$. Then there exists a subsolution $(v_\alpha, a)$ with $c > a$ to (2.6). By the same argument in the proof of Proposition 2.4, we get $-(a - c)t \leq M$ for some $M > 0$ and all $t$, which implies the contradiction. Therefore, $c = d_1 = d_2$. \hfill $\square$

Let us note that, to prove the first formula only, we do not need to use the regularization by a mollifier. Hence, the first formula is true in the general viscous case with non-convex Hamilton–Jacobi equations. On the other hand, the second formula is not trivial for such a setting. Indeed, it has not been known in a fully general setting yet.

In the end of this chapter, we give a result on the asymptotic speed of solutions to (2.3). This is a straightforward application of Theorem 2.3.

Proposition 2.6. Assume that (H1) and (2.11) hold. Let $u$ be the viscosity solution of (2.3) and $c$ be the associated ergodic constant. Then,

$$ \frac{u(x, t)}{t} \to -c \text{ uniformly on } \mathbb{T}^n \text{ as } t \to \infty.$$ 

Proof. Let $(v, c)$ be a solution of (2.8). Take a suitably large constant $M > 0$ so that $v - ct - M, v - ct + M$ are a subsolution and a supersolution of (2.3) respectively. In light of the comparison principle for (2.3), we get

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

which yields the conclusion. \hfill $\square$
This a priori estimate is the reason why we do not need to consider the terms like $a_i(x)t^i$ for $i \geq 2$ in the formal asymptotic expansion in the introduction of Chapter 2.

We also give a priori estimate of the Lipschitz bound of solutions to (2.3).

**Proposition 2.7.** Assume that (H1) and (2.11) hold. Assume further that $u_0 \in C^2(\mathbb{T}^n)$. Then the solution $u$ to (2.3) is globally Lipschitz continuous on $\mathbb{T}^n \times [0, \infty)$.

**Proof.** For a suitably large $M > 0$, $u_0 \pm Mt$ is a subsolution and a supersolution of (2.3) respectively. In light of the maximum principle we get $|u(x, t) - u(x, s)| \leq M|t - s|$ for any $x, t, s \in \mathbb{T}^n \times [0, \infty)$. Therefore, $|u_t| \leq M$. By using the same method as in the proof of Theorem 2.3, we get $|Du| \leq M'$ for some $M' > 0$.  

As a corollary of Propositions 2.6–2.7, 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

$$u(x, t_j) + ct_j \to v(x)$$

uniformly for $x \in \mathbb{T}^n$ as $j \to \infty$,

where $v$ is a solution of (2.8). However, we have to be careful about the fact that $v$ in the above depends on the choice of a subsequence. Whether this accumulation point is same for all of choices of sequences or not is a nontrivial fact and we will study this in the next chapter.
Chapter 3

Large time asymptotics of Hamilton–Jacobi equations

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 [57], under the assumptions: $p \mapsto H(x,p)$ is convex, $H(x, p) \geq H(x, 0)$ for all $(x, p) \in \mathbb{T}^n \times \mathbb{R}^n$, and $\max_{x \in \mathbb{T}^n} H(x, 0) = 0$. We will first discuss this setting in Section 3.1. In this setting, as the Hamiltonian has a simple structure, we can easily find a set 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 (2.5).

Fathi then gave a breakthrough in [26] by using a dynamical approach from the weak KAM theory. Contrary to [57], the results of [26] 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 [27] for a characterization of the Aubry set, which will be studied in Section 3.4. Afterwards, Davini and Siconolfi in [21] and Ishii in [38] 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 [10] obtained additional results, for possibly non-convex Hamiltonians, by using a PDE method in the context of viscosity solutions. Barles, Ishii and Mitake [7] simplified the ideas in [10] 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 time large in the context of the PDE approach.

In the uniformly parabolic setting, Barles and Souganidis [11] proved the long-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, the strong maximum principle holds, the ergodic problem has a
unique solution up to constants. The proof for the large-time convergence in [11] strongly depends on this fact. We will discuss this in Section 3.5.

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 3.3, 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 introduced a new method for the large-time behavior for Hamilton–Jacobi equation. In this method, the nonlinear adjoint method, which was introduced by Evans in [25], plays the essential role. In Section 3.2 we study this method.

### 3.1 Separable Hamiltonians

As mentioned in the end of Section 2.2, in general, (2.6) does not have unique solutions even up to additive constants. See Section 4.1 for the detail. This fact can be observed from Example 5.1 too. This requires a more delicate argument to prove the large-time convergence (2.5) for (2.1).

Before going to a general case, we first consider the case where the Hamiltonian is separable with respect to \(x, p\). Consider

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

where \(V \in C(\mathbb{T}^n)\) is a given function. See Example 5.2 in Appendix. In this case, since the structure of the Hamiltonian is very simple, we can easily prove (2.5). This was first done by Namah and Roquejoffre in [57]. First, let us find the ergodic constant in this case. As in Proposition 2.5, we have

\[
c = \inf_{\phi \in C^1(\mathbb{T}^n)} \sup_{x \in \mathbb{T}^n} (|D\phi(x)|^2 / 2 + V(x)).
\]

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

\[
c \geq \sup_{x \in \mathbb{T}^n} V(x) = \max_{x \in \mathbb{T}^n} V(x).
\]

Thus, we get

\[
c = \max_{x \in \mathbb{T}^n} V(x).
\]

Set \(u_\phi(x, t) := u(x, t) + ct\). Then

\[
(u_\phi)_t + \frac{1}{2} |Du_\phi|^2 = \max_{x \in \mathbb{T}^n} V(x) - V(x).
\]

Set

\[
\mathcal{A} := \{ x \in \mathbb{T}^n : V(x) = \max_{x \in \mathbb{T}^n} V(x) \}.
\]
Then we can easily see at least formally that for $x \in A$

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

which implies the monotonicity of $t \mapsto u_c(x, t)$. Thus, we get

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

where $\liminf_{t \to \infty}$ and $\limsup_{t \to \infty}$ are half-relaxed limits. In view of the stability result of viscosity solutions, $\limsup_{t \to \infty} u_c$ and $\liminf_{t \to \infty} u_c$ are a subsolution and a supersolution of (2.6) respectively. Also notice that the set $A$ is a uniqueness set (see Theorem 3.13 in Section 3.4). We will come back to this key point later. In light of this, we get

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

See [57] and also [53] for more details.

Next, let us consider

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

where $h \in C(\mathbb{T}^n)$ with $h(x) > 0$ for all $x \in \mathbb{T}^n$ is a given function. See Example 5.1 in Appendix. We obtain the ergodic constant first as follows:

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

Thus, we get

$$c = \sqrt{\max_{x \in \mathbb{T}^n} h(x)^2} = \max_{x \in \mathbb{T}^n} h(x). \quad \text{(3.1)}$$

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

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

In this case, we have

$$(u_c)_t = c - h(x)\sqrt{1 + |Du_c|^2} \leq 0 \quad \text{in } A_c := \{ x \in \mathbb{T}^n \mid h(x) = c \}.$$ 

Therefore, we get a similar type of monotonicity to the above example.

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

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

where $b : \mathbb{T}^n \to \mathbb{R}^n$ is a given vector valued function, then we cannot find such an easy structure of solutions. Therefore, we need more deep arguments to prove (2.5) in general case.
3.2 General Hamiltonians

In this section, we assume the followings:

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

(H3) \( D_{pp}^2 H(x, p) \geq 2\theta I_n \) 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) \( |D_x H(x, p)| \leq C(1 + |p|^2) \) for all \( x \in \mathbb{T}^n \) and \( p \in \mathbb{R}^n \).

We can easily check that if \( H \) satisfies (H3)–(H4), then it satisfies (2.11) hence also (2.10). 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 3.1.** Assume that (H2)–(H4) hold. Let \( u \) be the solution of (2.1) with given initial data \( u_0 \in W^{1,\infty}(\mathbb{T}^n) \). Then there exists \( (v, c) \in W^{1,\infty}(\mathbb{T}^n) \times \mathbb{R} \), a solution of ergodic problem (2.6), such that (2.5) holds.

We call \( v - ct \) obtained in Theorem 3.1 the asymptotic solution for (2.1).

**Remark 3.1.** It is worth pointing out delicate things on the convexity assumption here. Assumption (H3) is a uniformly convexity assumption. We can actually easily weaken this to a strictly convexity assumption, i.e., \( D_{pp}^2 H > 0 \), 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 strict convex one.

On the other hand, this “strictness” of convexity is very important to get convergence (2.5). 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 we can easily check 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 and convergence (2.5) does not hold. This was first pointed out by Barles and Souganidis in [10].

We also point out that the convexity is NOT a necessary condition either, since it is known that there is a result for a convergence for possibly non convex Hamilton–Jacobi equations in [10] (see also [7]). A typical example for non convex Hamiltonians is \( H(p) := (|p|^2 - 1)^2 - V(x) \).
3.2. GENERAL HAMILTONIANS

3.2.1 Formal calculation

In this subsection, we describe the idea in [14] in a heuristic way to get

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

(3.2)

where $c$ is the ergodic constant for (2.6). We call this an asymptotic monotone property for the solution (2.1). This is a much stronger result than that of Proposition 2.6. We “assume” that $u$ is smooth below in the derivation, even though we cannot expect a global smooth solution $u$ of Hamilton–Jacobi equations in general.

We consider the adjoint equation of the linearized equation of the Hamilton–Jacobi equation:

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

(3.3)

where $\delta_{x_0}$ is the Dirac delta measure at some point $x_0 \in \mathbb{T}^n$. Note that although (3.3) may have only a very singular solution, we do not mind as this is just a formal argument. It is clear that

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

(3.4)

Then we have the following energy conservation:

$$\frac{d}{dt} \int_{\mathbb{T}^n} H(x, Du(x,t))\sigma(x,t) \, dx$$

$$= \int_{\mathbb{T}^n} D_pH(x, Du) \cdot Du_t \sigma \, dx + \int_{\mathbb{T}^n} H(x, Du)\sigma_t \, dx$$

$$= -\int_{\mathbb{T}^n} \text{div} \left(D_pH(x, Du)\sigma\right) u_t \, dx - \int_{\mathbb{T}^n} u_t \sigma_t \, dx = 0,$$

which implies the new formula:

$$-u_t(x_0, T) = -\int_{\mathbb{T}^n} u_t \delta_{x_0} \, dx = \int_{\mathbb{T}^n} H(x, Du)\sigma \, dx \bigg|_{t=T} = \frac{1}{T} \int_0^T \int_{\mathbb{T}^n} H(x, Du)\sigma \, dx dt.$$

Noting (3.4) and $|Du(x,t)| \leq C$ by Proposition 2.7, in light of the Riesz theorem, there exists $\nu_T \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ such that

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

for all $\varphi \in C^1(\mathbb{T}^n \times \mathbb{R}^n)$.

Since

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} d\nu_T(x,p) = 1,$$
there exists a subsequence $T_j \to \infty$ as $j \to \infty$ so that

$$\nu_{T_j} \to \nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \quad \text{as} \quad j \to \infty \quad \text{(3.5)}$$

in terms of the measure. Then, we can expect the important facts

(i) $\nu$ is a Mather measure of (2.6),

(ii) $\text{supp} \nu \subset \{(x, p) \mid p = Dv(x)\}$, where $v$ is a viscosity solution to (2.6).

We will give the definition of the Mather measure in Chapter 4. Property (ii) in the above is called the graph theorem in the Hamiltonian dynamics, which is a very important theorem (see [50, 49] for the details). One way to look at (i) is the following: if we think $Du$ is a given function in (3.3), then (3.3) is a transport equation, and the characteristic ODE is given by

$$\begin{cases}
\dot{X}(t) = D_pH(X(t), Du(X(t), t)) & \text{for } t \in (0, T) \\
X(T) = x_0,
\end{cases} \quad \text{(3.6)}$$

which is formally equivalent to the Hamilton system.

If we admit these, then we obtain

$$-u_t(x_0, T_j) = \frac{1}{T_j} \int_0^{T_j} \int_{\mathbb{T}^n} H(x, Du) \sigma \, dx \, dt = \int \int_{\mathbb{T}^n \times \mathbb{R}^n} H(x, p) \, d\nu_{T_j}(x, p) \to c$$

as $j \to \infty$ for any subsequence $T_j$ satisfying (3.5).

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

(i) to introduce a regularizing process for (2.1),

(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 estimates,

which are discussed in details from next subsections.

### 3.2.2 Regularizing Process

In the following subsections, we make the formal argument in Section 3.2.1 rigorous by using a regularizing process and giving important estimates.

We only need to study the case where $c = 0$, where $c$ is the ergodic constant, thus 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 $c = 0$. 

3.2. GENERAL HAMILTONIANS

We first consider a rescaled problem. Setting \( u^\varepsilon(x, t) = u(x, t/\varepsilon) \) for \( \varepsilon > 0 \), where \( u \) is the solution of (2.1), one can easily check that \( u^\varepsilon \) satisfies

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

Notice however that in this way we do not have a priori uniform Lipschitz estimates on \( \varepsilon \), since the Lipschitz bounds on \( u \) give us that

\[
\|u^\varepsilon_t\|_{L^\infty(T^n \times [0,1])} \leq C/\varepsilon, \quad \|Du^\varepsilon\|_{L^\infty(T^n \times [0,1])} \leq C. \tag{3.7}
\]

In general, the function \( u^\varepsilon \) is only Lipschitz continuous. For this reason, we add a viscosity term to \( (C)_\varepsilon \), and we consider the regularized equation

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

We also consider the approximation for the ergodic problem (2.6):

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

By Theorem 2.2, the existence for \( (E)_\varepsilon \) holds. The advantage of considering \( (A)_\varepsilon \) and \( (E)_\varepsilon \) lies in the fact that the solutions of these equations are smooth, and this will allow us to use the nonlinear adjoint method in the next subsection.

**Proposition 3.2.** Assume that (2.10) and (H4) hold. There exists \( C > 0 \) independent of \( \varepsilon \) such that

\[
\|u^\varepsilon(\cdot, 1) - w^\varepsilon(\cdot, 1)\|_{L^\infty(T^n)} \leq C\varepsilon, \quad \|\overline{H}_\varepsilon\| \leq C\varepsilon^2.
\]

**Proof.** Let \( u^\varepsilon \) and \( w^\varepsilon \) be the solution of \( (C)_\varepsilon \) and \( (A)_\varepsilon \). We consider

\[
\max_{x,y \in T^n, t \in [0,1]} \{ u^\varepsilon(x, t) - w^\varepsilon(y, t) - \frac{|x-y|^2}{2\eta} - Kt \}
\]

for \( \eta > 0 \), which will be fixed later and let the maximum be achieved at \( (x_\eta, y_\eta, t_\eta) \in T^n \times T^n \times [0,1] \).

In the case \( t_\eta > 0 \), in light of Ishii’s lemma, for any \( \rho > 0 \), there exist \( (p_\eta, X_\eta) \in T^{2,+}u^\varepsilon(x_\eta, t_\eta) \) and \( (p_\eta, Y_\eta) \in T^{2,-}w^\varepsilon(y_\eta, t_\eta) \) such that

\[
p_\eta = \frac{x_\eta - y_\eta}{\eta}, \quad \begin{pmatrix} x_\eta \\ Y_\eta \end{pmatrix} \leq A_\eta + \rho A^2_\eta, \tag{3.8}
\]

where

\[
A_\eta := \frac{1}{\eta} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}.
\]
The definition of viscosity solutions immediately implies the following inequality:

\[ \varepsilon K + H(x, p) - H(y, p) \leq \varepsilon^4 \text{tr} (X) \]

Note that

\[ \varepsilon^4 \text{tr} (X) = \sum_{i=1}^{n} \{ \langle X \varepsilon^2 e_i, \varepsilon^2 e_i \rangle - \langle Y 0 e_i, 0 e_i \rangle \} \]

\[ \leq \sum_{i=1}^{n} \left\{ A \left( \varepsilon^2 e_i \right) \cdot \left( \varepsilon^2 e_i \right) + \rho A^2 \left( \varepsilon^2 e_i \right) \cdot \left( \varepsilon^2 e_i \right) \right\} \]

\[ \leq \frac{C\varepsilon^4}{\eta} + O(\rho). \]

Note that by the definition of \( x, y, t, \eta \), we have

\[ u^\varepsilon(y, t) - u^\varepsilon(y, t) - K t \leq u^\varepsilon(y, t) - u^\varepsilon(y, t) - \frac{|x - y|^2}{2\eta} - K t, \]

which implies \( |p| \leq C \) for some \( C > 0 \) in view of the Lipschitz continuity of \( u^\varepsilon \). Thus, \( |x - y| \leq C \eta \). Therefore,

\[ |H(x, p) - H(y, p)| \leq C(1 + |p|^2)|x - y| \leq C \eta. \]

Combine the above to deduce \( \varepsilon K \leq C\varepsilon^4/\eta + C \eta + O(\rho) \). Sending \( \rho \to 0 \) and setting \( K := C\varepsilon^{-1}(\varepsilon^4/\eta + \eta) \), we necessarily have \( t = 0 \).

Thus, we get

\[ u^\varepsilon(x, 1) - u^\varepsilon(x, 1) \leq u_0(x) - u_0(y) + K \leq C \eta + C\varepsilon^{-1} \left( \frac{\varepsilon^4}{\eta} + \eta \right) = \frac{C\varepsilon^3}{\eta} + C \left( 1 + \frac{1}{\varepsilon} \right) \eta. \]

Setting

\[ \eta := \varepsilon^2, \]

we get the conclusion.

Noting that we are assuming the ergodic constant for (2.6) is \( c = 0 \) now, by a similar argument to the above, we get \( |\overline{H}_\varepsilon| \leq C\varepsilon^2 \).

### 3.2.3 Conservation of energy and a key observation

The adjoint equation of the linearized operator of \((A)_\varepsilon\) is

\[ (AJ)_\varepsilon \begin{cases} -\varepsilon \sigma_i^\varepsilon - \text{div}(D_p H(x, Dw^\varepsilon) \sigma^\varepsilon) = \varepsilon^4 \Delta \sigma^\varepsilon & \text{in } \mathbb{T}^n \times (0, 1), \\ \sigma^\varepsilon(x, 1) = \delta_{x_0} & \text{on } \mathbb{T}^n. \end{cases} \]

**Proposition 3.3** (Elementary Property of \( \sigma^\varepsilon \)). Assume that \((H2)\) holds. We have \( \sigma^\varepsilon \geq 0 \) and

\[ \int_{\mathbb{T}^n} \sigma^\varepsilon(x, t) \, dx = 1 \text{ for all } t \in [0, 1]. \]
3.2. GENERAL HAMILTONIANS

Proof. Since we have
\[
\frac{d}{dt} \int_{\mathbb{T}^n} \sigma^\varepsilon(\cdot, t) \, dx = \int_{\mathbb{T}^n} -\text{div}(D_p H(x, Dw^\varepsilon)) \sigma^\varepsilon - \eta \Delta \sigma^\varepsilon \, dx = 0,
\]
and \( \sigma^\varepsilon(\cdot, 1) = \delta_{x_0} \) for some \( x_0 \), we get the conclusion.

Lemma 3.4 (Conservation of Energy 1). Assume that (H2) hold. The followings hold:

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

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

We stress the fact that identity Lemma 3.4 (ii) is extremely important. If we scale back the time, the integral in 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.

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:

\[
\frac{d}{dt} \int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon) \sigma^\varepsilon \, dx = \int_{\mathbb{T}^n} (D_p H(x, Dw^\varepsilon) \cdot Dw^\varepsilon_t - \varepsilon^4 \Delta w^\varepsilon_t) \sigma^\varepsilon \, dx + \int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon) \sigma^\varepsilon_t \, dx = -\int_{\mathbb{T}^n} \left( \text{div} \left( D_p H(x, Dw^\varepsilon) \sigma^\varepsilon \right) + \varepsilon^4 \Delta \sigma^\varepsilon \right) w^\varepsilon_t \, dx - \int_{\mathbb{T}^n} \varepsilon w^\varepsilon_t \sigma^\varepsilon_t \, dx = 0. \]

Remark 3.2. We emphasize here that we do not use any specific structures of the equations up to now, and therefore this conservation law holds for much more general equations. To analyze more requires more specific analysis, but it is worth mentioning that in this reason this method for the large-time asymptotics for nonlinear equations is universal in principle.

Theorem 3.5. Assume that (H2)–(H4) hold. We have
\[
\lim_{\varepsilon \to 0} \varepsilon \| w^\varepsilon_t(\cdot, 1) \|_{L^\infty(\mathbb{T}^n)} = 0.
\]

More precisely, there exists a positive constant \( C \), independent of \( \varepsilon \), such that
\[
\varepsilon \| w^\varepsilon_t(\cdot, 1) \|_{L^\infty(\mathbb{T}^n)} = \| H(\cdot, Dw^\varepsilon(\cdot, 1)) - \varepsilon^4 \Delta w^\varepsilon(\cdot, 1) \|_{L^\infty(\mathbb{T}^n)} \leq C \varepsilon^{1/4}.
\]

To prove this, we use the following key estimates, which will be proved in the next subsection.
Lemma 3.6 (Key Estimates 1). Assume that (H2)–(H4) hold. There exists a positive constant \( C \), independent of \( \varepsilon \), such that the followings hold:

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

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

We now can give the proof of Theorem 3.5, which is the main principle to achieve large time asymptotics, by using the averaging action above and the key estimates in Lemma 3.6.

Proof of Theorem 3.5. Let us first choose \( x_0 \) such that

\[
|\varepsilon w^\varepsilon_1(x_0, 1)| = |H(x_0, Dw^\varepsilon(x_0, 1)) - \varepsilon^4 \Delta w^\varepsilon(x_0, 1)| = \|H(\cdot, Dw^\varepsilon(\cdot, 1)) - \varepsilon^4 \Delta w^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)}.
\]

Thanks to Lemma 3.4 and Proposition 3.2,

\[
\varepsilon \|w^\varepsilon_1(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = \|H(\cdot, Dw^\varepsilon(\cdot, 1)) - \varepsilon^4 \Delta w^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)}
\]

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

\[
\leq \int_0^1 \int_{\mathbb{T}^n} |(H(x, Dw^\varepsilon) - \varepsilon^4 \Delta w^\varepsilon) - (H(x, Dw^\varepsilon) - \varepsilon^4 \Delta v^\varepsilon)| \sigma^\varepsilon \, dx \, dt + |\mathcal{P}_\varepsilon|
\]

\[
\leq \int_0^1 \int_{\mathbb{T}^n} \left[ C |D(w^\varepsilon - v^\varepsilon)| + \varepsilon^4 |\Delta(w^\varepsilon - v^\varepsilon)| \right] \sigma^\varepsilon \, dx \, dt + C \varepsilon^2,
\]

where \( v^\varepsilon \) is a solution of \((E)_\varepsilon\). We finally use the Hölder inequality and Lemma 3.6 to get that

\[
\varepsilon \|w^\varepsilon_1(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} \leq C \left[ \left( \int_0^1 \int_{\mathbb{T}^n} |D(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} + \varepsilon^4 \left( \int_0^1 \int_{\mathbb{T}^n} |D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} \right] + \varepsilon^2
\]

\[
\leq C \varepsilon^{1/4}. \quad \square
\]

Proof of Theorem 3.1. Since we can prove the equi-Lipschitz continuity of \( \{w^\varepsilon(\cdot, 1)\} \) by a similar argument to that of Theorem 2.2, we can choose a sequence \( \{\varepsilon_m\} \to 0 \) such that \( \{w^{\varepsilon_m}(\cdot, 1)\} \) converges uniformly to a continuous function \( v \). In view of Theorem 3.5, \( v \) is a solution of \((E)\), and thus a (time independent) solution of the equation in \((C)_\varepsilon\). We let \( t_m = 1/\varepsilon_m \) and use Proposition 3.2 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 does not depend on the sequence \( \{t_m\}_{m \in \mathbb{N}} \). Now, for any \( x \in \mathbb{T}^n, \ t > 0 \) such that \( t_m \leq t < t_{m+1} \), we use the comparison principle to yield that

\[
|u(x,t) - v(x)| \leq \|u(\cdot, t_m + (t - t_m)) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)} \leq \|u(\cdot, t_m) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)}.
\]
3.2. GENERAL HAMILTONIANS

Thus,
\[ \lim_{t \to \infty} |u(x, t) - v(x)| \leq \lim_{m \to \infty} \|u(\cdot, t_m) - v(\cdot)\|_{L^\infty(T^n)} = 0, \]
which gives the conclusion. □

3.2.4 Proof of key estimates

Lemma 3.7. Assume that (H2) holds. There exists a constant \( C > 0 \) such that
\[ \int_0^1 \int_{T^n} \varepsilon^4 |D^2 w^\varepsilon|^2 \sigma^\varepsilon \, dx \, dt \leq C. \]

This is one of the key estimates which was first introduced by Evans [25] in the study of gradient shock structures of the vanishing viscosity procedure of first order Hamilton–Jacobi equations (see also [62]).

Proof. Let \( w^\varepsilon \) be the solution of \( (A)_\varepsilon \) and set \( \varphi(x, t) := |Dw^\varepsilon|^2 / 2. \) Then \( \varphi \) satisfies
\[ \varepsilon \varphi_t + D_p H \cdot D \varphi + D_x H \cdot Dw^\varepsilon = \varepsilon^4 (\Delta \varphi - |D^2 w^\varepsilon|^2) \]
Multiply the above by \( \sigma^\varepsilon \) and integrate over \([0, 1] \times T^n\) to yield the conclusion. □

Proof of Lemma 3.6 (i). Subtracting equation \( (A)_\varepsilon \) from \( (E)_\varepsilon, \) thanks to the uniform convexity of \( H, \) we get
\[ 0 = \varepsilon (v^\varepsilon - w^\varepsilon)_t + H(x, Dv^\varepsilon) - H(x, Dw^\varepsilon) - \varepsilon^4 \Delta (v^\varepsilon - w^\varepsilon) - \overline{H}_\varepsilon \]
\[ \geq \varepsilon (v^\varepsilon - w^\varepsilon)_t + D_p H(x, Dw^\varepsilon) \cdot D(v^\varepsilon - w^\varepsilon) + \theta |D(v^\varepsilon - w^\varepsilon)|^2 - \varepsilon^4 \Delta (v^\varepsilon - w^\varepsilon) - \overline{H}_\varepsilon. \]

Multiply the above inequality by \( \sigma^\varepsilon \) and integrate by parts on \([0, 1] \times T^n\) to deduce that
\[ \theta \int_0^1 \int_{T^n} |D(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \leq \overline{H}_\varepsilon - \int_0^1 \int_{T^n} \varepsilon ((v^\varepsilon - w^\varepsilon) \sigma^\varepsilon)_t \, dx \, dt \]
\[ + \int_0^1 \int_{T^n} \left[ \varepsilon \sigma^\varepsilon_t + \text{div} (D_p H(x, Dw^\varepsilon) \sigma^\varepsilon) + \varepsilon^4 \Delta \sigma^\varepsilon \right] (v - w) \, dx \, dt \]
\[ = \overline{H}_\varepsilon + \varepsilon \left[ \int_{T^n} (w^\varepsilon - v^\varepsilon) \sigma^\varepsilon \, dx \right]_{t=1}^{t=0} \]
\[ = \overline{H}_\varepsilon + \varepsilon (w^\varepsilon(x_0, 1) - v^\varepsilon(x_0)) - \varepsilon \int_{T^n} (u_0(x) - v^\varepsilon(x)) \sigma^\varepsilon(x, 0) \, dx \]
\[ = \overline{H}_\varepsilon + \varepsilon w^\varepsilon(x_0, 1) - \varepsilon \int_{T^n} (v^\varepsilon(x_0) - v^\varepsilon(x)) \sigma^\varepsilon(x, 0) \, dx - \varepsilon \int_{T^n} u_0(x) \sigma^\varepsilon(x, 0) \, dx \]
\[ \leq \overline{H}_\varepsilon + C \varepsilon + C \varepsilon \|Dv^\varepsilon\|_{L^\infty(T^n)} - \varepsilon \int_{T^n} u_0(x) \sigma^\varepsilon(x, 0) \, dx \leq C \varepsilon, \]
which implies the conclusion. □
Proof of Lemma 3.6 (ii). Subtract (A)\(\varepsilon\) from (E)\(\varepsilon\) and differentiate with respect to \(x_i\) to get
\[
\varepsilon(v^\varepsilon - w^\varepsilon)_{x_i,t} + D_p H(x, Dv^\varepsilon) \cdot Dv^\varepsilon_{x_i} - D_p H(x, Dw^\varepsilon) \cdot Dw^\varepsilon_{x_i} + H_{x_i}(x, Dv^\varepsilon) - H_{x_i}(x, Dw^\varepsilon) - \varepsilon^4 \Delta(v^\varepsilon - w^\varepsilon)_{x_i} = 0.
\]

Let \(\varphi(x, t) := |D(v^\varepsilon - w^\varepsilon)|^2/2\). Multiplying the last identity by \((v^\varepsilon - w^\varepsilon)\), and summing up with respect to \(i\), we achieve that
\[
\varepsilon \varphi_t + D_p H(x, Dv^\varepsilon) \cdot D\varphi + \left(\left(D_p H(x, Dv^\varepsilon) - D_p H(x, Dw^\varepsilon)\right) \cdot Dv^\varepsilon_{x_i}\right) (v^\varepsilon_{x_i} - w^\varepsilon_{x_i}) + \left(D_2 H(x, Dv^\varepsilon) - D_2 H(x, Dw^\varepsilon)\right) \cdot D(v^\varepsilon - w^\varepsilon) - \varepsilon^4 (\Delta \varphi - |D^2(v^\varepsilon - w^\varepsilon)|^2) = 0.
\]

By using the equi-Lipschitz continuity of \(v^\varepsilon, w^\varepsilon\) and (H4), we derive that
\[
\varepsilon \varphi_t + D_p H(x, Dv^\varepsilon) \cdot D\varphi - \varepsilon^4 \Delta \varphi + \varepsilon^4 |D^2(v^\varepsilon - w^\varepsilon)|^2 \leq C(|D^2v^\varepsilon| + 1)|D(v^\varepsilon - w^\varepsilon)|^2. \tag{3.9}
\]

The last term in the right hand side of (3.9) is a dangerous term. We now take advantage of Lemma 3.7 to handle it. Using the fact that \(|Dv^\varepsilon|_{L^\infty}\) and \(|Dw^\varepsilon|_{L^\infty}\) are bounded, we have
\[
C|D^2v^\varepsilon| \cdot |D(v^\varepsilon - w^\varepsilon)|^2 \leq C|D^2v^\varepsilon| \cdot |D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon| \cdot |D(v^\varepsilon - w^\varepsilon)|^2 \leq \frac{\varepsilon^4}{2}|D^2(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\varepsilon^4}|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|. \tag{3.10}
\]

Combine (3.9) and (3.10) to deduce
\[
\varepsilon \varphi_t + D_p H(x, Dv^\varepsilon) \cdot D\varphi - \varepsilon^4 \Delta \varphi + \frac{\varepsilon^4}{2}|D^2(v^\varepsilon - w^\varepsilon)|^2 \leq C|D(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\varepsilon^4}|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|. \tag{3.11}
\]

We multiply (3.11) by \(\sigma^\varepsilon\), integrate over \([0, 1] \times \mathbb{T}^n\), to yield that, in light of Lemma 3.7 and (i),
\[
\varepsilon \int_0^1 \int_{\mathbb{T}^n} |D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \leq C \varepsilon + \frac{C}{\varepsilon^3} \varepsilon + C \int_0^1 \int_{\mathbb{T}^n} |D^2w^\varepsilon| \sigma^\varepsilon \, dx \, dt \leq \frac{C}{\varepsilon^3} + C \leq \frac{C}{\varepsilon^2} \leq \frac{C}{\varepsilon^3}. \quad \Box
\]

Remark 3.3. Estimate Lemma 3.6 gives 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, a priori estimates 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 $\{\sigma^\varepsilon\}$ 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 worthwhile to mention that we eventually do not need to use the graph theorem which is a deep result in the dynamical system.

### 3.3 Degenerate viscous Hamilton–Jacobi equations

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

$$ u_t - \mathrm{tr} \left( A(x) D^2 u \right) + H(x, Du) = 0 \quad \text{in } \mathbb{T}^n \times (0, \infty). \tag{3.12} $$

Here is one of the main results of [14].

**Theorem 3.8.** Assume that (H1)–(H4) hold. Let $u$ be the solution of (3.12) with initial data $u(\cdot, 0) = u_0 \in W^{1,\infty}(\mathbb{T}^n)$. Then there exists $(v, c) \in W^{1,\infty}(\mathbb{T}^n) \times \mathbb{R}$ such that (2.5) holds, where the pair $(v, c)$ is a solution of the ergodic problem

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

For an easy explanation, we consider 1-dimensional case ($n = 1$) in this section. This makes the problem much easier but we do not lose a key point which comes from the degenerate viscous term. We repeat the same procedure in Sections 3.2.2, 3.2.3. Associated problems are now described below:

- **(C)** $u_t - a(x)u_{xx} + H(x, u_x) = 0$,
- **(A)$_\varepsilon$** $\varepsilon u_x^\varepsilon - a(x)w^\varepsilon_{xx} + H(x, w^\varepsilon_x) = \varepsilon^4 w^\varepsilon_{xx}$,
- **(A)$_J$** $-\sigma_{xx}^\varepsilon - (a(x)\sigma^\varepsilon)_{xx} - (H_p(x, w^\varepsilon_x)\sigma^\varepsilon)_x = \varepsilon^4 \sigma^\varepsilon_{xx},$
- **(E)$_\varepsilon$** $- a(x)v^\varepsilon_{xx} + H(x, v^\varepsilon_x) = \varepsilon^4 v^\varepsilon_{xx} + \overline{H}_\varepsilon.$

Here $a$ is a nonnegative $C^2(\mathbb{T}^n)$ function.

As pointed out in Remark 3.2 we have the same type conservation of energy.

**Lemma 3.9** (Conservation of Energy 2). Assume that (H1), (H2) hold. The following hold:

- (i) $\frac{d}{dt} \int_{\mathbb{T}} \left[ H(x, w^\varepsilon_x) - (a(x) + \varepsilon^4)w^\varepsilon_{xx} \right] \sigma^\varepsilon \, dx = 0$,
- (ii) $\varepsilon w^\varepsilon_t(x_0, 1) = \int_0^1 \int_{\mathbb{T}} \left[ H(x, w^\varepsilon_x) - (a(x) + \varepsilon^4)w^\varepsilon_{xx} \right] \sigma^\varepsilon \, dx \, dt.$
Now, we have
\[
\varepsilon \|w^\varepsilon (\cdot, 1)\|_{L^\infty(\mathbb{T})} = \|H(\cdot, w^\varepsilon (\cdot, 1)) - (\varepsilon^4 + a(x))w^\varepsilon_{xx}(\cdot, 1)\|_{L^\infty(\mathbb{T})}
\]
\[
= \left| \int_0^1 \int_{\mathbb{T}} \left[ H(x, w^\varepsilon_x) - (\varepsilon^4 + a(x))w^\varepsilon_{xx} \right] \sigma^\varepsilon \, dx \, dt \right|
\]
\[
\leq \int_0^1 \int_{\mathbb{T}} \left[ |H(x, w^\varepsilon_x) - (\varepsilon^4 + a(x))w^\varepsilon_{xx}| - |H(x, v^\varepsilon_x) - (\varepsilon^4 + a(x))v^\varepsilon_{xx}| \right] |\sigma^\varepsilon| \, dx \, dt + |H_c|
\]
\[
\leq C \left[ \left( \int_0^1 \int_{\mathbb{T}} |(w^\varepsilon - v^\varepsilon)_x|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} + \varepsilon^4 \left( \int_0^1 \int_{\mathbb{T}} |(w^\varepsilon - v^\varepsilon)_{xx}|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} \right]
\]
\[
+ \left( \int_0^1 \int_{\mathbb{T}} a(x)^2 |(w^\varepsilon - v^\varepsilon)_{xx}|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} + \varepsilon^2 \]
where \(v^\varepsilon\) is a solution of (E).  

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 3.10 (Key Estimates 2).** Assume that (H1)--(H4) hold. There exists a constant \(C > 0\), independent of \(\varepsilon\), such that

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

(ii) \(\int_0^1 \int_{\mathbb{T}} (a(x) + \varepsilon^4) |w^\varepsilon_{xx}|^2 \sigma^\varepsilon \, dx \, dt \leq C\),

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

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

**Proof.** The proof of (i) is the same as that of Lemma 3.6 (i) hence is omitted.

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

\[
\varepsilon w^\varepsilon_{xx} + H_p(x, w^\varepsilon_x) \cdot w^\varepsilon_x + H_x(x, w^\varepsilon_x) - (\varepsilon^4 + a)w^\varepsilon_{xxx} - a_x w^\varepsilon_{xx} = 0.
\]  
(3.13)

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

\[
\xi_t = w^\varepsilon_t w^\varepsilon_x, \quad \xi_x = w^\varepsilon_x w^\varepsilon_x, \quad \xi_{xx} = |w^\varepsilon|^2 + w^\varepsilon_x w^\varepsilon_{xx}.
\]

Multiply (3.13) by \(w^\varepsilon_x\) to arrive at

\[
\varepsilon \xi_t + H_p \cdot \xi_x + H_x \cdot w^\varepsilon_x = (\eta + a)(\xi_{xx} - |w^\varepsilon_{xx}|^2) + (a_x \cdot w^\varepsilon) w^\varepsilon_{xx}.
\]
3.3. DEGENERATE VISCOUS HAMILTON-JACOBI EQUATIONS

Notice first that we have \( a^2_x \leq Ca \) since \( a \in C^2(\mathbb{T}) \). Indeed, \( a \in C^2(\mathbb{T}) \) implies \( \sqrt{a} \in \text{Lip}(\mathbb{T}) \). Thus, \(|a_x| = 2(\sqrt{a})_x \sqrt{a} \leq C \sqrt{a} \). We next notice that for \( \delta > 0 \) small enough,

\[
a_x w^\varepsilon_x w^\varepsilon_{xx} \leq C |a_x||w^\varepsilon_x| \leq \frac{C}{\delta} + \frac{1}{2} a^2_x |w^\varepsilon_{xx}|^2 \leq C + \frac{1}{2} a |w^\varepsilon_x|^2.
\]  

(3.14)

Hence,

\[
\varepsilon \xi_t + H_p \cdot \xi_x + \frac{1}{2} (\varepsilon^4 + a)|w^\varepsilon_{xx}|^2 \leq (\varepsilon^4 + a)\xi_{xx} + C.
\]

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

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

\[
\varepsilon(v^\varepsilon - w^\varepsilon)_x t + H_p(x, v^\varepsilon_x) \cdot v^\varepsilon_x - H_p(x, w^\varepsilon_x) \cdot w^\varepsilon_x + H_x(x, v^\varepsilon_x) - H_x(x, w^\varepsilon_x) - (\varepsilon^4 + a)(v^\varepsilon - w^\varepsilon)_x x x - a_x(v^\varepsilon - w^\varepsilon)_x x = 0.
\]

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

\[
\varepsilon \varphi_t + H_p(x, w^\varepsilon_x) \cdot \varphi_x - (\varepsilon^4 + a(x)) \varphi_{xx} + \left( (\varepsilon^4 + a)\varphi_{xx}\right) (v^\varepsilon - w^\varepsilon)_x x
\]

\[
+ \left( H_x(x, v^\varepsilon_x) - H_x(x, w^\varepsilon_x) \right) \cdot (v^\varepsilon - w^\varepsilon)_x + (\varepsilon^4 + a(x))((v^\varepsilon - w^\varepsilon)_x x)^2 - \varphi_{xx}
\]

\[- a_x \cdot (v^\varepsilon - w^\varepsilon)_x x (v^\varepsilon - w^\varepsilon)_x = 0.
\]

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

\[
\left| a_x \cdot (v^\varepsilon - w^\varepsilon)_x x \right| (v^\varepsilon - w^\varepsilon)_x \leq \delta |a_x|^2 |(v^\varepsilon - w^\varepsilon)_x x| + \frac{1}{\delta} |(v^\varepsilon - w^\varepsilon)_x x|^2
\]

\[
\leq \frac{\alpha}{2} |(v^\varepsilon - w^\varepsilon)_x x|^2 + C |(v^\varepsilon - w^\varepsilon)_x x|^2.
\]

Thus,

\[
\varepsilon \varphi_t + H_p(x, w^\varepsilon_x) \cdot \varphi_x - (\varepsilon^4 + a(x)) \varphi_{xx} + \left( (\varepsilon^4 + a)\varphi_{xx}\right) |(v^\varepsilon - w^\varepsilon)_x x|^2
\]

\[
\leq C(1 + |v^\varepsilon_{xx}|) |(v^\varepsilon - w^\varepsilon)_x x|^2. \quad (3.15)
\]

By using the same trick as (3.10), we get

\[
\varepsilon \varphi_t + H_p(x, w^\varepsilon_x) \cdot \varphi_x - (\varepsilon^4 + a(x)) \varphi_{xx} + \left( \frac{\varepsilon^4}{2} + \frac{a}{2} \right) |(v^\varepsilon - w^\varepsilon)_x x|^2
\]

\[
\leq C |(v^\varepsilon - w^\varepsilon)_x x|^2 + C \varepsilon^4 |(v^\varepsilon - w^\varepsilon)_x x|^2 + C |w^\varepsilon_{xx}|
\]

We multiply the above by \( \sigma^\varepsilon \), integrate over \([0, 1] \times \mathbb{T}^n\), to yield that

\[
\int_0^1 \int_{\mathbb{T}^n} (\varepsilon^4 + a(x))\|(w^\varepsilon - v^\varepsilon)_x x\|^2 \sigma^\varepsilon dx dt \leq C \varepsilon + \frac{C}{\varepsilon^4} \varepsilon + C \int_0^1 \int_{\mathbb{T}^n} \|w^\varepsilon_{xx}\| \sigma^\varepsilon dx dt
\]

\[
\leq \frac{C}{\varepsilon^3} + C \left( \int_0^1 \int_{\mathbb{T}^n} |w^\varepsilon_{xx}|^2 \sigma^\varepsilon dx dt \right)^{1/2} \left( \int_0^1 \int_{\mathbb{T}^n} \sigma^\varepsilon dx dt \right)^{1/2} \leq \frac{C}{\varepsilon^3} + \frac{C}{\varepsilon^2} \leq \frac{C}{\varepsilon^3},
\]
which implies the conclusion of (iii).

Finally we prove (iv). Setting

\[ \psi(x, t) := a(x)\varphi(x, t) = \frac{a(x)(v^\varepsilon - w^\varepsilon)_{x}(x, t)}{2}, \]

and multiplying (3.15) by \( a(x) \), we get

\[ \varepsilon \psi_t + H_p(x, w^\varepsilon_x) \cdot (\psi_x - a_x \varphi) - (\varepsilon^4 + a(x))(\psi_{xx} - a_{xx} \varphi - 2a_x \cdot \varphi_x) \]

\[ + a(x)(\varepsilon^4 + \frac{a(x)}{2})|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \leq Ca(x)(|v^\varepsilon_{xx}| + 1)|(v^\varepsilon - w^\varepsilon)_x|^2. \]

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

\[ \varepsilon \psi_t + H_p(x, w^\varepsilon_x) \cdot \psi_x - (\varepsilon^4 + a(x))\psi_{xx} + a(x)(\varepsilon^4 + a(x)/2)|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \]

\[ \leq C \varphi(x) - 2(\varepsilon^4 + a(x))a_x \cdot \varphi_x + Ca(x)|v^\varepsilon_{xx}| |(v^\varepsilon - w^\varepsilon)_x|^2. \]

For \( \delta > 0 \) small enough

\[ 2(\varepsilon^4 + a(x))a_x \cdot \varphi_x \leq C(\varepsilon^4 + a(x))|a_x| |(v^\varepsilon - w^\varepsilon)_{xx}| |(v^\varepsilon - w^\varepsilon)_x| \]

\[ \leq \delta(\varepsilon^4 + a(x))|a_x|^2 |(v^\varepsilon - w^\varepsilon)_{xx}|^2 + \frac{C}{\delta} |(v^\varepsilon - w^\varepsilon)_x|^2 \]

\[ \leq \frac{1}{8}(\varepsilon^4 + a(x))a(x)|v^\varepsilon_{xx}|^2 + C|(v^\varepsilon - w^\varepsilon)_x|^2. \]

On the other hand,

\[ a(x)|v^\varepsilon_{xx}| |(v^\varepsilon - w^\varepsilon)_x|^2 \]

\[ \leq a(x)|w^\varepsilon_{xx}| |(v^\varepsilon - w^\varepsilon)_x|^2 + a(x)|v^\varepsilon_{xx}| |(v^\varepsilon - w^\varepsilon)_x| \]

\[ \leq \varepsilon^{1/2}a(x)|w^\varepsilon_{xx}|^2 + \frac{C}{\varepsilon^{1/2}} |(v^\varepsilon - w^\varepsilon)_x|^2 + \frac{a(x)^2}{8}|(v^\varepsilon - w^\varepsilon)_x|^2 + C|(v^\varepsilon - w^\varepsilon)_x|^2. \]

Therefore,

\[ \varepsilon \psi_t + H_p(x, w^\varepsilon_x) \cdot \psi_x - (\varepsilon^4 + a(x))\psi_{xx} + \frac{a(x)^2}{4}|(v^\varepsilon - w^\varepsilon)_{xx}|^2 \]

\[ \leq (C + C\varepsilon^{-1/2})|(v^\varepsilon - w^\varepsilon)_x|^2 + \varepsilon^{1/2}a(x)|w^\varepsilon_{xx}|^2. \]

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

Thus, we get

\[ \lim_{\varepsilon \to 0} \varepsilon \|w^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T})} = 0, \]

which directly implies the conclusion of Theorem 3.8 in \( n = 1 \). We refer to [14] and [53] for a general dimension.

Remark 3.4. If the equation in (C) is uniformly parabolic, i.e., \( a(x) > 0 \) for all \( x \in \mathbb{T} \), then estimate (iii) is not needed anymore as estimate (iv) is much stronger. On the other hand, if \( a \) is degenerate, then (iv) only provides estimation of \( |D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \) on the support of \( a \) and it is hence essential to use (iii) to control the part where \( a = 0 \).
3.4 Asymptotic profile

In this section, we investigate the first-order Hamilton–Jacobi equation (2.1) again, and specially consider the asymptotic profile. As we have already emphasized many times, because of the multiplicity of solutions to (2.6), the asymptotic profile \( v \) in Theorem 3.1 is completely decided through the initial data. 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 [21]. We use the following assumption:

\[(H3)’ \quad D_{pp}^2 H(x, p) \geq 0 \text{ for all } x \in \mathbb{T}^n, p \in \mathbb{R}^n.\]

We first introduce the notion of the Aubry set.

**Definition 3.1.** Let \( y \in \mathbb{T}^n \). We call \( y \) the element of the Aubry set \( A \) if the following

\[
\inf \left\{ \int_0^t L_c(\gamma(s), -\dot{\gamma}(s)) \, ds : t \geq \delta, \gamma \in AC(\mathbb{T}^n), \gamma(0) = \gamma(t) = y \right\} = 0
\tag{3.16}
\]

is satisfied for any \( \delta > 0 \), where \( L_c(x, v) := L(x, v) + c \) for any \( (x, v) \in \mathbb{T}^n \times \mathbb{R}^n \).

**Remark 3.5.** We can easily check that \( y \) is in \( A \) if and only if (3.16) holds only for some \( \delta_0 \). 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 \( \gamma_\varepsilon \in AC(\mathbb{T}^n) \) with \( \gamma(0) = \gamma(t_\varepsilon) = y \) such that

\[
\int_0^{t_\varepsilon} L_c(\gamma_\varepsilon(s), -\dot{\gamma}_\varepsilon(s)) \, ds < \varepsilon.
\]

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 \( \gamma_m(0) = \gamma_m(t_\varepsilon) = y \). We calculate that

\[
\int_0^{mt_\varepsilon} L_c(\gamma_m(s), -\dot{\gamma}_m(s)) \, ds = \sum_{j=0}^{m-1} \int_{(j-1)t_\varepsilon}^{jt_\varepsilon} L_c(\gamma(s - (j - 1)t_\varepsilon), -\dot{\gamma}(s - (j - 1)t_\varepsilon)) \, ds
\]

\[
= m \int_0^{t_\varepsilon} L_c(\gamma(s), -\dot{\gamma}(s)) \, ds < m\varepsilon,
\]

which implies

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

We define the function \( d : \mathbb{T}^n \times \mathbb{T}^n \to \mathbb{R} \) by

\[
d_c(x, y) := \inf \left\{ \int_0^t L_c(\gamma(s), -\dot{\gamma}(s)) \, ds : t > 0, \gamma \in AC(\mathbb{T}^n), \gamma(0) = x, \gamma(t) = y \right\}.
\tag{3.17}
\]

We gather some basic properties of the function \( d_c \).
Proposition 3.11. Assume that (H3)' holds. We have

(i) \( d_c(x, y) = \sup \{ v(x) - v(y) : v \) is a subsolution of (2.6) \} \).

(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) \( d_c(. , y) \) is a subsolution of (2.6) for all \( y \in \mathbb{T}^n \) and a solution of (2.6) 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 (2.6), and \( \gamma \in AC(\mathbb{T}^n) \) with \( \gamma(0) = x \) and \( \gamma(t) = y \). This is at least formally easy to prove. Indeed, let \( v \) be a smooth subsolution of (2.6). We have the following simple computations

\[
v(x) - v(y) = -\int_0^t \frac{d v(\gamma(s))}{ds} \, ds = \int_0^t D v(\gamma(s)) \cdot (-\dot{\gamma}(s)) \, ds
\]

\[
\leq \int_0^t L(\gamma(s), -\dot{\gamma}(s)) + H(\gamma(s), D v(\gamma(s))) \, ds \leq \int_0^t L_c(\gamma(s), -\dot{\gamma}(s)) \, ds.
\]

This immediately implies (i). We can also easily check that (i) implies (ii) and (iii) is a straightforward result of (i) and stability results of viscosity solutions. \( \square \)

Fathi and Siconolfi in [27] gave a beautiful characterization of the Aubry set as follows.

Theorem 3.12. Assume that (H3)' holds. The point \( y \in \mathbb{T}^n \) is in the Aubry set \( \mathcal{A} \) if and only if \( d_c( , y) \) is a solution of (2.6).

We omit the proof of the above. See [27, Proposition 5.8] and [38, Proposition A.3] for the proofs.

Theorem 3.13. Assume that (H3)' holds. The Aubry set \( \mathcal{A} \) is nonempty, compact, and a uniqueness set, i.e., if solutions \( v_1, v_2 \) to (2.6) coincide on \( \mathcal{A} \), then \( v_1 = v_2 \) on \( \mathbb{T}^n \).

Proof. The compactness of \( \mathcal{A} \) is a straightforward result of stability of viscosity solutions. To prove the property of a uniqueness set for \( \mathcal{A} \), we only need to prove that if a subsolution \( v \) and a supersolution \( w \) of (2.6) satisfy \( v \leq w \) 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 \( \mathcal{A} \subset U_\varepsilon \) and \( v \leq w + \varepsilon \). Set \( K_\varepsilon := \mathbb{T}^n \setminus U_\varepsilon \). Fix any \( z \in K_\varepsilon \). In light of Theorem 3.12, we may choose a constant \( r_z > 0 \) and a function \( \varphi_z \in C^1(\mathbb{T}^n) \) such that \( B(z, r_z) \subset \mathbb{T}^n \setminus \mathcal{A} \), \( H(x, D \varphi_z(x)) < 0 \) for all \( x \in B(z, r_z) \), \( \varphi_z(z) > 0 = d_c(z, z) \), and \( \varphi_z(x) < d_c(x, z) \) for all \( x \in \mathbb{T}^n \setminus B(z, r_z) \). We set \( \psi_z(x) = \max \{ d_H(x, z), \varphi_z(x) \} \) for \( x \in \mathbb{T}^n \) and observe that \( \psi_z \) is a subsolution of (2.6) 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 \( \{ z_j \}_{j=1}^J \) 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 (2.6) for some neighborhood \( V \).
3.4. ASYMPTOTIC PROFILE

of $K_{\varepsilon}$. Regularizing $\psi$ by mollification, if necessary, we may assume that $\psi \in C^1(V)$. Thus, we may apply the classical comparison result to conclude that $v \leq w + \varepsilon$ in $K_{\varepsilon}$. Sending $\varepsilon \to 0$ yields the conclusion.

**Theorem 3.14.** Assume that (H2)–(H4) holds. Let $u_\infty[u_0] - ct$ be the asymptotic solution for (2.1). Then we have for all $y \in \mathcal{A}$

$$u_\infty[u_0](y) = \min \{d_c(y, z) + u_0(z) : z \in \mathbb{T}^n \}$$

$$= \sup \{v(x) \mid v \text{ is a subsolution to (2.6) with } v \leq u_0 \text{ in } \mathbb{T}^n \}. \quad (3.18)$$

**Proof.** We write $v_{u_0}$ for the right hand side of (3.18). 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 $\xi_\varepsilon \in AC(\mathbb{T}^n)$ 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 $t_n \geq n$ and a curve $\delta_\varepsilon \in AC(\mathbb{T}^n)$ such that $\delta_\varepsilon(0) = \delta_\varepsilon(t_n) = y$, and

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

Define $\gamma_\varepsilon \in AC(\mathbb{T}^n)$ by

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

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

![Figure 3.1](image-url)
We observe that
\[ v_{u_0}(y) > \int_0^{t_n} L_c(\xi_s, -\dot{\xi}_s) \, ds + u_0(z_y) - \varepsilon \]
\[ > \int_0^{t_n} L_c(\delta_s(s), -\dot{\delta}_s(s)) \, ds + \int_0^{t_n} L_c(\xi_s, -\dot{\xi}_s) \, ds + u_0(z_y) - 2\varepsilon \]
\[ = \int_0^{t_n + t_0} L_c(\gamma_s, -\dot{\gamma}_s) \, ds + u_0(z_y) - 2\varepsilon \]
\[ \geq u_c(y, t_n + t_0) - 2\varepsilon. \]

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 light of the comparison principle for (2.1), we get \( v_{u_0}(x) - ct \leq u(x, t) \) for all \( x, t \). Thus, \( v_{u_0}(x) \leq \lim_{t \to -\infty} (u(x, t) + ct) = u_c^\infty(x) \).

In light of Proposition 3.11, Theorems 3.12, 3.13, 3.14, we get the asymptotic profile:

**Corollary 3.15.** Assume that (H2)–(H4) holds. Let \( u_c^\infty[u_0] - ct \) be the asymptotic solution for (2.1). Then we have the profile
\[ u_c^\infty[u_0](x) = \min \{ d_c(x, y) + d_c(y, z) + u_0(z) \mid y \in A, z \in \mathbb{T}^n \}. \]  

**Proof.** Note first that the right hand side of (3.19) coincides with \( \min \{ d_c(x, y) + v_{u_0}(y) \mid y \in A \} \), where \( v_{u_0} \) is the function which appears in the proof of Theorem 3.14. Also, this is a solution of (2.6) in view of Theorem 3.12, which immediately yields the conclusion.

**Example 3.1.** Now, let us consider the asymptotic profile for the Hamilton–Jacobi equation appearing in Example 5.1. As we observe in the beginning of Section 3.1, the associated ergodic problem is
\[ |Dv| = \sqrt{\frac{c^2 - h(x)^2}{h(x)^2}} \quad \text{in} \; \mathbb{T}^n, \]

where \( c \) is given by (3.1). We can easily check that we have the explicit formula for the Aubry set
\[ A := \{ x \in \mathbb{T}^n : h(x) = \max_{\mathbb{T}^n} h \} \]

from the definition of the Aubry set. Also, we have
\[ d(x, y) = \inf \left\{ \int_0^{t} \sqrt{\frac{c^2}{h(\gamma(s))^2} - 1} \, ds : t > 0, |\gamma| \leq 1, \gamma(0) = x, \gamma(t) = y \right\}. \]

From this, we can somehow understand how the asymptotic profile depends on the force term \( h \) and the initial data \( u_0 \) through Corollary 3.15.
3.5 Viscous case

We assume here for simplicity that \( u_0 \in C^2(\mathbb{T}^n) \).

**Theorem 3.16.** Assume that (2.10) holds and \( u_0 \in C^2(\mathbb{T}^n) \). Let \( u \) be the solution of (2.2). Then there exists a solution \((v, c) \in C^2(\mathbb{T}^n) \times \mathbb{R} \) of (2.7) such that, as \( t \to \infty \),

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

We call \( v - ct \) obtained in Theorem 3.16 the asymptotic solution for (2.2). The following proof is based on the argument in [11].

**Proof.** We normalize the ergodic constant \( c \) to be 0 by replacing \( H \) by \( H - c \). Let \( v \) be a solution of (2.7). 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 2.7, 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 \( u_\infty \in W^{1,\infty}(\mathbb{T}^n \times [0, \infty)) \) which may of course depend on the subsequence \( \{t_j\} \) up to now. By a standard stability result of viscosity solutions we see that \( u_\infty \) is a solution to the equation in (2.1). 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 the strong maximum principle, Proposition 3.17, 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 \( u_\infty(x, t) \equiv u_\infty(x) = v(x) + \overline{m} \). Noting that the right-hand side above does not depend on the choice of a subsequence now, we see that

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

**Proposition 3.17** (Strong Maximum Principle). Let \( U \) be a bounded domain in \( \mathbb{R}^n \) and set \( U_T := U \times (0, T] \) for some fixed times \( T > 0 \). Let \( u \) be a smooth subsolution of (2.2). If \( u \) attains its maximum over \( U_T \) at a point \((x_0, t_0) \in U_T\) then \( u \) is constant on \( U_{t_0} \).

See [24] for instance. If we do not have the regularity (smoothness) for solutions, then we have to be careful with the result in Proposition 3.17. A straightforward application of Proposition 3.17 is the uniqueness (up to additive constants) of solutions to (2.7). This is a crucial difference from that for the first-order Hamilton–Jacobi equation.
3.6 Some other directions and open questions

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

(i) Unbounded domain: If we consider the Cauchy problem in an unbounded domain (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 [38, 34] for the first order Hamilton–Jacobi equation and [28, 33, 35] for the viscous Hamilton–Jacobi equation.

(ii) Boundary value problems: If we consider several 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 types of difficulties. See [51] for the state constraint problem, [52, 61, 9] for the Dirichlet problem, [39, 8, 6] for the Neumann problem, and [55] for the obstacle problem.

(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 [53, 16, 54, 58] for the development on this direction. The profile of asymptotic limits is not solved yet.

(iv) Degenerate viscous Hamilton–Jacobi equations: In addition to the works [14, 55], we refer to [45] for this direction.

(v) Time-periodic Hamilton–Jacobi equations: There are only a few works in this direction. Only 1-dimensional case has been studied by [12], but for the multi-dimensional one, it is quite open. See also [44] for an interesting application in the modeling of traffic flows. Note here that we do not have conservation of energy anymore as $H$ depends on $t$.

(vi) Hamilton–Jacobi equations with mean curvature: This is also an interesting topic, since many important open questions still remain. See [17] for a result along this line.
Chapter 4

Selection problems in the discounted approximation procedure

4.1 Selection problems

Let us now revisit the ergodic problems and their solutions. In general, the solutions are not unique even up to additive constants. In the particular inviscid case (ergodic problem (2.6)), we can easily see the solutions is not unique even up to constants. Let us give first an explicit example along this line to see clearly the non-uniqueness issue.

Example 4.1. 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 \( x \in [0, 1] \). This is of the same setting as in Example 5.2 and the discussion at the beginning of Section 3.1.

\[
\begin{align*}
\text{Figure 4.1: Graph of } W \text{ on } [0, 1].
\end{align*}
\]

Then \( c = 0 \) and \( A = \{1/4, 3/4\} \). The ergodic problem becomes

\[
|u'|^2 = W(x)^2 \quad \text{in } \mathbb{R},
\]

(4.1)
where \( u \) is 1-periodic. For \( x \in [0, 1] \), set
\[
\begin{align*}
  u^b_1(x) := & \begin{cases} 
  (x - 1/4)^2 & \text{for } 0 \leq x \leq 1/2 \\
  \min \left\{ - (x - 3/4)^2 + 1/8, (x - 3/4)^2 + b \right\} & \text{for } 1/2 \leq x \leq 1 
  \end{cases} \\
  u^b_2(x) := & \begin{cases} 
  \min \left\{ (x - 1/4)^2 + b, -(x - 1/4)^2 + 1/8 \right\} & \text{for } 0 \leq x \leq 1/2 \\
  (x - 3/4)^2 & \text{for } 1/2 \leq x \leq 1 
  \end{cases}
\end{align*}
\]
and extend \( u^b_1, u^b_2 \) to \( \mathbb{R} \) periodically. We can easily check that \( u^b_1, u^b_2 \) are solutions to (4.1) for all \( b \in [0, 1/8] \).

![Figure 4.2: Graph of \( u^b_1 \) on \([0, 1]\) for \( b = 0, b = 1/16, b = 1/8 \).](image)

This of course shows that (4.1) has many solutions of different types.

Therefore, if we consider an approximation procedure for (2.6), then the question is highly nontrivial whether the approximation converges or not. Moreover, if it converges, then which solution is the limit (which solution is selected)? This type of question is called the selection problem.

Let us be more precise in addressing the above question by considering the discounted approximation procedure of the following possibly degenerate viscous Hamilton-Jacobi equation. For \( \varepsilon > 0 \), consider
\[
(E)_\varepsilon \quad \varepsilon u^\varepsilon + H(x, Du^\varepsilon) = a(x)\Delta u^\varepsilon \quad \text{in } \mathbb{T}^n,
\]
where we assume
\[
\text{(H5) } H \in C^2(\mathbb{T}^n \times \mathbb{R}^n), \quad p \mapsto H(x, p) \text{ is convex for each } x \in \mathbb{T}^n, \text{ and there exists } C > 0 \text{ so that }
\]
\[
|D_x H(x, p)| \leq C(1 + H(x, p)), \quad \text{for all } (x, p) \in \mathbb{T}^n \times \mathbb{R}^n,
\]
and
\[
\lim_{|p| \to +\infty} \frac{H(x, p)}{|p|} = +\infty, \quad \text{uniformly for } \ x \in \mathbb{T}^n,
\]
\[
\text{(H6) } a \geq 0 \text{ in } \mathbb{T}^n, \text{ and } a \in C^0(\mathbb{T}^n).
\]
4.1. SELECTION PROBLEMS

Note that the situation we consider is a special case of the general degenerate viscous case (2.8). For the general case, see Section 4.5.

Let us now repeat some of the arguments as in the proof of Theorem 2.3. Under assumptions (H5)–(H6) (or some other appropriate growth conditions), we can prove the following a priori estimate

\[ \| Du^\varepsilon \|_{L^\infty(\mathbb{T}^n)} \leq C \text{ for some } C > 0. \] (4.2)

Once (4.2) is achieved, we can easily see that

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

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

\[ \varepsilon_j u^{\varepsilon_j} \to -c \in \mathbb{R}, \quad u^{\varepsilon_j}(x_0) \to u \in C(\mathbb{T}^n) \text{ uniformly in } \mathbb{T}^n \text{ as } j \to \infty, \] (4.3)

where \( (u, c) \) is a solution of (2.8).

We assume without loss of generality that \( c = 0 \) henceforth. Let us notice that the procedure in (4.3) is a soft approach mainly using tools from functional analysis. In particular, the convergence (4.3) is just along subsequences. An important question to be studied is whether this convergence holds for the whole sequence \( \varepsilon \to 0 \) or not.

This question was first addressed by Gomes [31], Iturriaga and Sanchez-Morgado [42] under rather restricted assumptions few years ago. Recently, Davini, Fathi, Iturriaga and Zavidovique [22] gave a quite complete and positive answer for this question in case \( a \equiv 0 \) by using a dynamical system approach in light of weak KAM theory and characterizing the limit in terms of Mather measures. Also A.-Aidarous, Alzahrani, Ishii, and Younas [1] obtained a same type of convergence for the Neumann boundary problem by using an approach similar to that in [22]. It is important noticing that all the results mentioned here are for first order Hamilton–Jacobi equations \( (a \equiv 0) \) as the methods there use deep properties of extremal curves of optimal control theory formulae of solutions of \( (E)_\varepsilon \), \( (E) \), and minimizing properties of Mather measures.

We here prove the convergence result.

Theorem 4.1. Assume that (H5)–(H6). The following convergence holds

\[ u^\varepsilon(x) \to u^0(x) := \sup_{\phi \in \mathcal{E}} \phi(x) \text{ uniformly for } x \in \mathbb{T}^n \text{ as } \varepsilon \to 0, \] (4.4)

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

\[ \int \int_{\mathbb{T}^n \times \mathbb{R}^n} u \, d\mu \leq 0 \text{ for all } \mu \in \mathcal{M}. \] (4.5)

The set \( \mathcal{M} \) of probability measures on \( \mathbb{T}^n \times \mathbb{R}^n \), which are stochastic Mather measures, is defined in Section 4.2.1.
Remark 4.1. Note that 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 (not just subsolutions) in Theorem 4.1 which is different from that of [22].

Up to now, all of the obtained results require convexity of the Hamiltonians. We emphasize that the problem on convergence of solutions of the discounted Hamilton–Jacobi equation with non convex Hamiltonian remains rather open. In some nonconvex cases, an argument similar to the one which we will show here works and the convergence result holds (see the forthcoming work [32]).

4.2 Key ingredients and estimates

Recall that we assume that the ergodic constant is 0. The ergodic problem now becomes

$$\begin{align*}
\text{(E)} \quad H(x, Du) &= a(x)\Delta u \quad \text{in } \mathbb{T}^n.
\end{align*}$$

4.2.1 Regularization process and construction of $\mathcal{M}$

We denote by $\mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ the set of 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 $H$, i.e.,

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

By (H5), $L$ is finite on $\mathbb{T}^n \times \mathbb{R}^n$, of class $C^1$, and superlinear.

For each $\eta > 0$, we consider an approximation of $E_\varepsilon$ as following

$$\begin{align*}
\text{(A)}_\varepsilon^n \quad \varepsilon u^{\varepsilon, \eta} + H(x, Du^{\varepsilon, \eta}) &= (a(x) + \eta^2)\Delta u^{\varepsilon, \eta} \quad \text{in } \mathbb{T}^n.
\end{align*}$$

The following result is quite standard. It is of the same flavor as that of Proposition 3.2. See [25, 62, 15, 14] for instance.

Lemma 4.2. Assume (H5)–(H6). There exists a constant $C > 0$ independent of $\varepsilon$ and $\eta$ so that

$$\|u^{\varepsilon, \eta} - u^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \leq C\varepsilon^{-1}\eta.$$

We introduce the associated adjoint equation of the linearized operator of $A_\varepsilon^n$:

$$\begin{align*}
\text{(AJ)}_\varepsilon^n \quad \varepsilon \theta^{\varepsilon, \eta} - \text{div}(D_p H(x, Du^{\varepsilon, \eta})\theta^{\varepsilon, \eta}) &= \Delta(a(x)\theta^{\varepsilon, \eta}) + \eta^2 \Delta \theta^{\varepsilon, \eta} + \varepsilon \delta_{x_0} \quad \text{in } \mathbb{T}^n
\end{align*}$$

for some $x_0 \in \mathbb{T}^n$, where $\delta_{x_0}$ denotes the delta Dirac measure at $x_0$. Clearly, we have

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

For every $\varepsilon, \eta > 0$, let $\nu^{\varepsilon, \eta} \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be a probability measure satisfying

$$\int_{\mathbb{T}^n} \psi(x, Du^{\varepsilon, \eta})\theta^{\varepsilon, \eta}(x) \, dx = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, p) \, d\nu^{\varepsilon, \eta}(x, p) \quad (4.6)$$
4.2. KEY INGREDIENTS AND ESTIMATES

for all \( \psi \in C(\mathbb{T}^n \times \mathbb{R}^n) \). There exist two subsequences \( \varepsilon_j \to 0 \) and \( \eta_k \to 0 \) as \( j \to \infty, \) \( k \to \infty, \) respectively, and probability measures \( \nu^{\varepsilon_j}, \nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \) so that

\[
\begin{align*}
\nu^{\varepsilon_j, \eta_k} & \to \nu^{\varepsilon_j} \quad \text{as} \quad k \to \infty, \\
\nu^{\varepsilon_j} & \to \nu \quad \text{as} \quad j \to \infty,
\end{align*}
\]

(4.7)
in term of measures. Notice that the limit \( \nu \) might be different for different choices of subsequences \( \{\varepsilon_j\} \) and \( \{\eta_k\} \). In general, there could be many such limit \( \nu \). For each such \( \nu \), set \( \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \) to be a pushforward measure of \( \nu \) associated with \( \Phi(x, v) = (x, D_v L(x, v)), \) i.e., for all \( \psi \in C(\mathbb{T}^n \times \mathbb{R}^n) \),

\[
\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, p) \, d\nu(x, p) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, D_v L(x, v)) \, d\mu(x, v).
\]

(4.8)

We henceforth denote by \( \mathcal{M} \) the collection of all such measures \( \mu \). We give important properties of \( \nu \) and \( \mu \) in the following proposition.

**Proposition 4.3.** Assume that (H5)–(H6). Let \( \nu \) and \( \mu \) be probability measures given by (4.7) and (4.8). Then we have

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

(ii) \( \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(x, p) \cdot D_\varphi - a(x)\Delta \varphi) \, d\nu(x, p) \)

\[= \int_{\mathbb{T}^n \times \mathbb{R}^n} (v \cdot D_\varphi - a(x)\Delta \varphi) \, d\mu(x, v) = 0 \quad \text{for any} \quad \varphi \in C^2(\mathbb{T}^n). \]

We notice that since we are assuming that the ergodic constant is 0, the right hand side of the identity of Proposition 4.3 (i) is 0. On the other hand, this normalization does not matter the right hand side of the identity of Proposition 4.3 (ii).

**Remark 4.2.** It is worthwhile to point out a delicate issue that we cannot replace \( C^2 \) test functions by \( C^{1,1} \) test functions in Proposition 4.3 (ii), since 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 have 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. We will clearly address this point in Section 4.3.

**Proof.** We rewrite (A) as

\[
\varepsilon u^{\varepsilon, \eta} + D_p H(x, Du^{\varepsilon, \eta}) \cdot Du^{\varepsilon, \eta} - (a(x) + \eta^2)\Delta u^{\varepsilon, \eta}
\]

\[= D_p H(x, Du^{\varepsilon, \eta}) \cdot Du^{\varepsilon, \eta} - H(x, Du^{\varepsilon, \eta}). \]

Multiply the above with \( \theta^{\varepsilon, \eta} \) and integrate over \( \mathbb{T}^n \) to yield

\[
\varepsilon u^{\varepsilon, \eta}(x_0) = \int_{\mathbb{T}^n} (D_p H(x, Du^{\varepsilon, \eta}) \cdot Du^{\varepsilon, \eta} - H(x, Du^{\varepsilon, \eta})) \theta^{\varepsilon, \eta} \, dx
\]

\[= \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(x, p) \cdot p - H(x, p)) \, d\nu^{\varepsilon, \eta}(x, p). \]
Choose $\varepsilon = \varepsilon_j$, $\eta = \eta_k$, and let $k \to \infty$, $j \to \infty$ in this order to derive that
\[
0 = \int \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(x, p) \cdot p - H(x, p)) \, dv(x, p)
\]
\[
= \int \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(x, D_v L(x, v)) \cdot D_v L(x, v) - H(x, D_v L(x, v))) \, d\mu(x, v)
\]
\[
= \int \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, v) \, d\mu(x, v)
\]
by the definition (4.8) of $\mu$, and the duality of convex functions.

Next, to prove (ii), we multiply $(AJ)_\varepsilon$ with any given $\varphi \in C^2(\mathbb{T}^n)$ and integrate over $\mathbb{T}^n$ to get
\[
\int_{\mathbb{T}^n} (D_p H(x, D_u^{\varepsilon, \eta}) \cdot D_\varphi - a(x) \Delta \varphi) \theta^{\varepsilon, \eta} = \eta^2 \int_{\mathbb{T}^n} \Delta \varphi \theta^{\varepsilon, \eta} \, dx + \varepsilon \varphi(x_0) - \varepsilon \int_{\mathbb{T}^n} \varphi^{\varepsilon, \eta} \, dx.
\]
By using (4.6) for $\varepsilon = \varepsilon_j$, $\eta = \eta_k$, and letting $k \to \infty$, we obtain
\[
\int \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(x, p) \cdot D_\varphi - a(x) \Delta \varphi) \, dv^{\varepsilon_j}(x, p) = \varepsilon_j \varphi(x_0) - \varepsilon_j \int \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x) \, dv^{\varepsilon_j}(x, p).
\]
Send $j \to \infty$ to arrive at the conclusion. \qed

Properties (i), (ii) in Proposition 4.3 of measure $\mu$ are essential ones to characterize a stochastic Mather measure, which will be defined in the following discussion. This idea was first discovered by Mañé [49], who relaxed the original idea of Mather [50]. See Fathi [26], Cagnetti, Gomes and Tran [15, Theorem 1.3] for some discussion about this. To be more precise, we can prove that each measure $\mu \in \mathcal{M}$ defined by (4.8) minimizes the action
\[
\min_{\mu \in \mathcal{F}} \int \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, v) \, d\mu(x, v) =: -c,
\]
where
\[
\mathcal{F} := \left\{ \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) : \int \int_{\mathbb{T}^n \times \mathbb{R}^n} (v \cdot D\phi - a(x) \Delta \phi) \, d\mu(x, v) = 0 \text{ for all } \phi \in C^2(\mathbb{T}^n) \right\}.
\]
Measures belonging to $\mathcal{F}$ are called holonomic measures. Note that $c$ coincides with the ergodic constant and we are assuming that the ergodic constant is 0 now.

**Definition 4.1.** We let $\widetilde{\mathcal{M}}$ to be the set of all minimizers of (4.9). Each measure in $\widetilde{\mathcal{M}}$ is called a stochastic Mather measure.

When $a \equiv 0$, this is precisely the definition of Mather measures for first order Hamilton–Jacobi equations by Mañé [49]. When $a \equiv 1$, this coincides with the definition of stochastic Mather measures for viscous Hamilton–Jacobi equations given by Gomes [29].
4.2. KEY INGREDIENTS AND ESTIMATES

Let us now give a proof of the assertion above. We use a commutation lemma, Lemma 4.4, below. For any $\eta > 0$, pick $w^n, S^n$ as defined in Lemma 4.4. For any $\mu \in \mathcal{F}$, one has

\[
\int \int_{T^n \times \mathbb{R}^n} S^n(x) d\mu(x, v) \geq \int \int_{T^n \times \mathbb{R}^n} (H(x, Dw^n) - a(x)\Delta w^n) d\mu(x, v)
\]

\[
\geq \int \int_{T^n \times \mathbb{R}^n} (-L(x, v) + (v \cdot Dw^n - a(x)\Delta w^n)) d\mu(x, v)
\]

\[
= - \int \int_{T^n \times \mathbb{R}^n} L(x, v) d\mu(x, v).
\]

Note that $|S^n| \leq C$ and $S^n \to 0$ pointwise in $T^n$ as $\eta \to 0$. Let $\eta \to 0$ and use the Lebesgue dominated convergence theorem to deduce that

\[
\int \int_{T^n \times \mathbb{R}^n} L(x, v) d\mu(x, v) \geq 0.
\]

Thus, in view of Proposition 4.3 (i), we can observe that any measure $\mu \in \mathcal{M}$ minimizes the action (4.9). Therefore, $\mathcal{M} \subset \mathcal{M}$.

4.2.2 Key estimates

Lemma 4.4 (A commutation lemma). Assume that (H5)--(H6). Assume that $w$ is a viscosity solution of (E). Let $\gamma \in C_c^\infty(\mathbb{R}^n)$ be a standard mollifier such that $\gamma \geq 0$, supp $\gamma \subset \overline{B}(0,1)$ and $\|\gamma\|_{L^1(\mathbb{R}^n)} = 1$. For each $\eta > 0$, set $\gamma^n(y) := \eta^{-n}\gamma(\eta^{-1}y)$ for $y \in \mathbb{R}^n$, and

\[
w^n(x) := \int_{\mathbb{R}^n} \gamma^n(y)w(x + y) dy.
\]

There exists a constant $C > 0$ and a continuous function $S^n : T^n \to \mathbb{R}$ such that

\[
|S^n(x)| \leq C \quad \text{and} \quad \lim_{\eta \to 0} S^n(x) = 0, \quad \text{for each } x \in T^n,
\]

and

\[
H(x, Dw^n) \leq a(x)\Delta w^n + S^n(x) \quad \text{in } T^n.
\]

Moreover, $|\eta^2\Delta w^n| \leq C\eta$.

We postpone the proof of the commutation lemma to the next section. Let us however mention here that this is a technical result but is very important in our analysis. Indeed, for each solution $w$ of (E) with some a priori bounds, we can construct a family of smooth approximated subsolutions $\{w^n\}$ of (E). In particular, for any $\eta > 0$, $w^n$ is $C^2$, which is good enough for us to use as test functions in Proposition 4.3 (ii). It is well-known that we can perform sup-convolutions of $w$, which was discovered by Jensen [43], to derive semi-convex approximated subsolutions of (E), but these are not smooth enough to use as test functions (see Remark 1). We also want to mention that a similar result was already discovered a long time ago by Lions [46]. However, Lions
only got convergence to 0 of $S^n$ in the *almost everywhere* sense, which is not enough for our purpose. The delicate point here is that, as each Mather measure $\mu$ can be very singular in $\mathbb{T}^n \times \mathbb{R}^n$, we need to have the convergence of $S^n$ everywhere. Moreover, we can actually show that $S^n$ converges to 0 uniformly on $\mathbb{T}^n$ with convergence rate $\eta^{1/2}$, which is necessary to prove Theorem 4.1.

**Lemma 4.5** *(Uniform convergence).* Assume that (H5)–(H6). There exists a universal constant $C > 0$ such that $\|S^n\|_{L^\infty(\mathbb{T}^n)} \leq C\eta^{1/2}$.

The proof of this Lemma is also postponed to the next section. The two following results provide the key estimates for our purpose, which are analogies of [22, Lemma 5.4, Proposition 5.2].

**Lemma 4.6.** Assume that (H5)–(H6). Let $w \in C(\mathbb{T}^n)$ be any solution of (E), and $w^n$ be the function given by (4.10) for $\eta > 0$. Then,

$$u^{\varepsilon,\eta}(x_0) \geq w^n(x_0) - \int_{\mathbb{T}^n} w^n(\theta^{\varepsilon,\eta}) dx - \frac{C\eta}{\varepsilon} - \frac{1}{\varepsilon} \int_{\mathbb{T}^n} S^n(\theta^{\varepsilon,\eta}) dx. \quad (4.11)$$

**Proof.** In view of Lemma 4.4, it is clear that $w^n$ satisfies

$$H(x, Dw^n) \leq (a(x) + \eta^2)\Delta w^n + C\eta + S^n(x) \quad \text{in } \mathbb{T}^n.$$ We subtract (A) from the above to get

$$\varepsilon w^{\eta} + C\eta + S^n(x)$$

$$\geq \varepsilon (w^n - u^{\varepsilon,\eta}) + H(x, Dw^n) - H(x, Du^{\varepsilon,\eta}) - (a(x) + \eta^2)\Delta (w^n - u^{\varepsilon,\eta})$$

$$\geq \varepsilon (w^n - u^{\varepsilon,\eta}) + D_p H(x, Du^{\varepsilon,\eta}) \cdot D(w^n - u^{\varepsilon,\eta}) - (a(x) + \eta^2)\Delta (w^n - u^{\varepsilon,\eta}),$$

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

Multiplying this with $\theta^{\varepsilon,\eta}$, integrating on $\mathbb{T}^n$, and using the integral by parts, we get

$$\int_{\mathbb{T}^n} \varepsilon w^n(\theta^{\varepsilon,\eta}) dx + C\eta + \int_{\mathbb{T}^n} S^n(x)\theta^{\varepsilon,\eta} dx \geq \varepsilon (w^n - u^{\varepsilon,\eta})(x_0).$$

Rearrange this to arrive at the conclusion. \qed

**Proposition 4.7.** Assume that (H5)–(H6). Let $u^{\varepsilon}$ be the solution of $(E)_{\varepsilon}$, and $\mu \in \mathcal{M}$. Then,

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} u^{\varepsilon}(x) d\mu(x, v) \leq 0 \quad \text{for any } \varepsilon > 0.$$ 

**Proof.** For each $\eta > 0$, define

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

By Lemma 4.4,

$$\varepsilon u^{\varepsilon} + H(x, D\psi^n) - a(x)\Delta \psi^n \leq S^n(x),$$
4.3. PROOF OF THE COMMUTATION LEMMA

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

By the convexity of \(H\), we have, for any \(v \in \mathbb{R}^n\),
\[
\varepsilon u^\varepsilon + v \cdot D\psi^n - L(x, v) - a(x)\Delta\psi^n \leq S^n(x).
\]
Thus, in light of properties (i), (ii) in Proposition 4.3 of \(\mu\), we yield that
\[
\int_{T^n \times \mathbb{R}^n} \varepsilon u^\varepsilon \, d\mu(x, v) \leq \int_{T^n \times \mathbb{R}^n} S^n(x) \, d\mu(x, v).
\]
Let \(\eta \to 0\) and use the Lebesgue dominated convergence theorem that to achieve the desired result. \(\square\)

4.3 Proof of the commutation lemma

We first observe an important a priori estimate of viscosity solutions to (E). In view of \([3, \text{Theorem 3.1}]\), we have a Lipschitz estimate for all viscosity solutions of (E). Therefore, we have
\[
-C \leq -a(x)\Delta w \leq C \quad \text{in the viscosity sense},
\]
for some \(C > 0\). Then by using the result of equivalence of viscosity solutions and solutions in the distribution sense by Ishii \([37]\), and also a simple structure of diffusion, we have
\[
\|Dw\|_{L^\infty(\mathbb{T}^n)} + \|a\Delta w\|_{L^\infty(\mathbb{T}^n)} \leq C \tag{4.12}
\]
for some constant \(C > 0\).

We now show that, \(w\) is actually a subsolution of (E) in the distributional sense based on the ideas in \([43]\). For each \(\delta > 0\), let \(w^\delta\) be the sup-convolution of \(w\), i.e.,
\[
\overline{w}^\delta(x) := \sup_{y \in \mathbb{R}^n} \left( w(y) - \frac{|x - y|^2}{2\delta} \right).
\]
It is clear from \([43, 19]\) that \(\overline{w}^\delta\) is semi-convex and a viscosity subsolution of
\[
H(x, D\overline{w}^\delta) \leq a(x)\Delta\overline{w}^\delta + \omega(\delta) \quad \text{in } \mathbb{T}^n, \tag{4.13}
\]
where \(\omega : (0, \infty) \to \mathbb{R}\) is a modulus of continuity, i.e., \(\lim_{\delta \to 0} \omega(\delta) = 0\). Since \(\overline{w}^\delta\) is a semi-convex function, it is twice differentiable almost everywhere and thus is also a solution of (4.13) in the almost everywhere sense. In view of (4.12), we deduce further that \(\overline{w}^\delta\) is a distributional solution of (4.13). By passing to a subsequence if necessary, we have
\[
\overline{w}^\delta \to w \quad \text{uniformly in } \mathbb{T}^n,
\]
\[
D\overline{w}^\delta \rightharpoonup Dw \quad \text{weakly in } L^\infty(\mathbb{T}^n).
\]
CHAPTER 4. DISCOUNTED APPROXIMATION PROCEDURE

For any test function $\phi \in C^2(\mathbb{T}^n)$ with $\phi \geq 0$, by convexity of $H$, one obtains that

$$
\int_{\mathbb{T}^n} (H(x, Dw)\phi - w\Delta(a(x)\phi)) \, dx
= \lim_{\delta \to 0} \int_{\mathbb{T}^n} (H(x, Dw)\phi + D_pH(x, Dw) \cdot D(\bar{w}^\delta - w)\phi - \bar{w}^\delta\Delta(a(x)\phi)) \, dx
\leq \lim_{\delta \to 0} \int_{\mathbb{T}^n} (H(x, D\bar{w}^\delta) - a\Delta\bar{w}^\delta)\phi \, dx
\leq \lim_{\delta \to 0} \int_{\mathbb{T}^n} \omega(\delta)\phi \, dx = 0.
$$

This confirms that $w$ is a subsolution of (E) in the distributional sense.

Set

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

In light of the above assertion that $w$ is a distributional subsolution of (E), it is clear that

$$
H(x, Dw^\eta) \leq a(x)\Delta w^\eta + R_1^\eta(x) + R_2^\eta(x) \quad \text{in } \mathbb{T}^n.
$$

We now need to estimate $R_1^\eta$ and $R_2^\eta$.

**Lemma 4.8.** Assume that (H5)–(H6). We have $R_1^\eta(x) \leq C\eta$ for all $x \in \mathbb{T}^n$ and $\eta > 0$, where $C > 0$ is some sufficiently large constant independent of $\eta$.

**Proof.** In view of (4.12) and (H1),

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

Thus, by convexity of $H$ and Jensen’s inequality, one can easily obtain

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

**Lemma 4.9.** Assume that (H5)–(H6). There exists a constant $C > 0$ independent of $\eta$ such that $|R_2^\eta(x)| \leq C$ for all $x \in \mathbb{T}^n$ and $\eta > 0$. Moreover, $\lim_{\eta \to 0} R_2^\eta(x) = 0$ for each $x \in \mathbb{T}^n$.

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

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

which immediately implies $\eta^2|\Delta w^\eta| \leq C\eta$. 

4.3. PROOF OF THE COMMUTATION LEMMA

We next show the boundedness of $R^3_2$ by the following simple computations:

$$|R^3_2(x)| = \left| \int_{\mathbb{R}^n} (a(x+y) - a(x)) \Delta w(x+y) \gamma^n(y) \, dy \right|$$

$$= \left| \int_{\mathbb{R}^n} \gamma^n(y) Da(x+y) \cdot Dw(x+y) \, dy + \int_{\mathbb{R}^n} (a(x+y) - a(x)) Dw(x+y) \cdot D\gamma^n(y) \, dy \right|$$

$$\leq C \int_{\mathbb{R}^n} (\gamma^n(y) + |y||D\gamma^n(y)|) \, dy \leq C.$$

We finally prove that $\lim_{\eta \to 0} R^3_2(x) = 0$ for each $x \in \mathbb{T}^n$. We consider two cases: (i) $a(x) = 0$, (ii) $a(x) > 0$.

In case (i), noting that $a(x) = 0 = \min_{\mathbb{T}^n} a$, we also have $Da(x) = 0$. Therefore,

$$|R^3_2(x)| = \int_{\mathbb{R}^n} a(x+y) \Delta w(x+y) \gamma^n(y) \, dy$$

$$= \left| \int_{\mathbb{R}^n} Dw(x+y) \cdot Da(x+y) \gamma^n(y) \, dy + \int_{\mathbb{R}^n} Dw(x+y) \cdot D\gamma^n(y)a(x+y) \, dy \right|$$

$$\leq C \int_{\mathbb{R}^n} (|Da(x+y)|\gamma^n(y) + a(x+y)|D\gamma^n(y)|) \, dy$$

$$= C \int_{\mathbb{R}^n} (|Da(x+y) - Da(x)|\gamma^n(y) + (a(x+y) - a(x) - Da(x) \cdot y)|D\gamma^n(y)|) \, dy$$

$$\leq C \int_{\mathbb{R}^n} (|y|\gamma^n(y) + |y|^2|D\gamma^n(y)|) \, dy \leq C \eta.$$

In the case that $a(x) > 0$, then we can 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 (4.12), we deduce further that

$$|\Delta w(z)| \leq \frac{C}{c_x} : C_x \quad \text{for a.e. } z \in B(x, \eta_0). \quad (4.14)$$

Thus, for $\eta < \eta_0$, we have

$$|R^3_2(x)| = \left| \int_{\mathbb{R}^n} (a(x+y) - a(x)) \Delta w(x+y) \gamma^n(y) \, dy \right|$$

$$\leq C_x \int_{\mathbb{R}^n} |a(x+y) - a(x)| \gamma^n(y) \, dy \leq C_x \int_{\mathbb{R}^n} |y|\gamma^n(y) \, dy \leq C_x \eta.$$

In both cases, we can conclude that $\lim_{\eta \to 0} |R^3_2(x)| = 0$. Note however that the bound for $|R^3_2(x)|$ is dependent on $x$. \qed

Letting $S^n(x) := C \eta + R^3_2(x)$, we achieve the result of Lemma 4.4.

**Remark 4.3.** We want to emphasize that we need (4.12) for the establishment of Lemma 4.9. That is the main reason why we require $w$ to be a solution instead of just a subsolution of (E) so that (4.12) holds automatically. In fact, (4.12) does not hold for subsolutions of (E) in general. This point is one of the main differences between first and second order Hamilton–Jacobi equations, as we do have the estimate (4.12) even
just for subsolutions in case \( a \equiv 0 \), which is the case of first order Hamilton–Jacobi equations.

We also want to comment a bit more on the rate of convergence of \( R_2^\delta \) 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

\[
|R_2^\delta(x)| \leq C(\delta)\eta \quad \text{for all } x \in U^\delta.
\]

We however do not know the rate of convergence of \( R_2^\delta \) in \( \mathbb{T}^n \setminus U^\delta \) through the above proof yet.

With a more careful computation, we can get a uniform convergence of \( R_2^\delta \) with a rate \( \eta^{1/2} \), which is the same as the convergence rate of \( S^n \).

**Proof of Lemma 4.5.** Fix \( x \in \mathbb{T}^n \). We consider two cases: (i) \( \min_{B(x, \eta)} a \leq \eta \), (ii) \( \min_{B(x, \eta)} a > \eta \).

In case (i), there exists \( \bar{x} \in B(x, \eta) \) such that \( a(\bar{x}) \leq \eta \). Then, in light of [14, Lemma 2.6], there exists a constant \( C > 0 \) such that,

\[
|Da(\bar{x})| \leq Ca(\bar{x})^{1/2} \leq C\eta^{1/2}.
\]

For any \( z \in B(x, \eta) \) we have the following estimates

\[
|Da(z)| \leq |Da(z) - Da(\bar{x})| + |Da(\bar{x})| \leq C\eta + C\eta^{1/2} \leq C\eta^{1/2},
\]

and

\[
|a(z) - a(x)| \leq |a(z) - a(\bar{x})| + |a(x) - a(\bar{x})|
\]

\[
\leq |Da(\bar{x})|(|z - \bar{x}| + |x - \bar{x}|) + C(|z - \bar{x}|^2 + |x - \bar{x}|^2) \leq C\eta^{3/2} + C\eta^2 \leq C\eta^{3/2}.
\]

In light of the two estimates above, we can bound \( R_2^\delta \) as

\[
|R_2^\delta(x)| = \left| \int_{\mathbb{R}^n} (a(x + y) - a(x)) \Delta w(x + y) \gamma^n(y) \, dy \right|
\]

\[
= \left| \int_{\mathbb{R}^n} Dw(x + y) \cdot Da(x + y) \gamma^n(y) \, dy + \int_{\mathbb{R}^n} Dw(x + y) \cdot D\gamma^n(y)(a(x + y) - a(x)) \, dy \right|
\]

\[
\leq C \int_{\mathbb{R}^n} \left( \eta^{1/2} \gamma^n(y) + \eta^{3/2} |D\gamma^n(y)| \right) \, dy \leq C\eta^{1/2}.
\]

In case (ii), we can directly estimate as

\[
|R_2^\delta(x)| = \left| \int_{\mathbb{R}^n} (a(x + y) - a(x)) \Delta w(x + y) \gamma^n(y) \, dy \right|
\]

\[
\leq C \int_{\mathbb{R}^n} \frac{|a(x + y) - a(x)|}{a(x + y)} \gamma^n(y) \, dy \leq C \int_{\mathbb{R}^n} \frac{|Da(x + y)|}{a(x + y)} \gamma^n(y) \, dy
\]

\[
\leq C \int_{\mathbb{R}^n} \frac{|y|}{a(x + y)^{1/2}} \gamma^n(y) \, dy \leq C \int_{\mathbb{R}^n} \frac{|y|}{\eta^{1/2}} \gamma^n(y) \, dy \leq C\eta^{1/2}.
\]
The commutation lemma, Lemma 4.4, is independently an interesting result. For instance, we can immediately get an equivalence of viscosity subsolutions of (E) and subsolutions of (E) in the almost everywhere sense as a result of Lemmas 4.4 and 4.5.

**Proposition 4.10.** Let \( w \in C(\mathbb{T}^n) \) satisfy (4.12). 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 4.4, \( w \) is a subsolution of (E) in the distribution sense. In light of (4.12), \( w \) 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 (4.10). In view of Lemmas 4.5, and the stability result of viscosity solutions, we obtain that \( w \) is a viscosity subsolution of (E).

\[ \]

### 4.4 Proof of Theorem 4.1

**Proposition 4.11.** Assume that (H5)--(H6). We have \( \liminf_{\varepsilon \to 0} u^\varepsilon(x) \geq u^0(x) \).

**Proof.** Let \( \phi \in \mathcal{E} \), i.e., a solution of (E) satisfying (4.4), and \( \phi^\eta \) be the function defined by (4.10). Fix \( x_0 \in \mathbb{T}^n \).

Take two subsequences \( \varepsilon_j \to 0 \) and \( \eta_k \to 0 \) so that (4.7) 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 to satisfy \( \nu = \Phi_\#\mu \). Sending \( k \to \infty \) in (4.11), we get

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

in view of Lemma 4.5. Let \( j \to \infty \) in the above inequality to deduce further that

\[ \lim_{j \to \infty} u^{\varepsilon_j}(x_0) \geq \phi(x_0) - \iint_{\mathbb{T}^n \times \mathbb{R}^n} \phi(x) d\nu(x, p) = \phi(x_0) - \iint_{\mathbb{T}^n \times \mathbb{R}^n} \phi(x) d\mu(x, v) \geq \phi(x_0), \]

which implies the conclusion.

**Proposition 4.12.** Assume (H5)--(H6). 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, \( \limsup_{\varepsilon \to 0} u^\varepsilon(x) \leq u^0(x) \).

**Proof.** In view of Proposition 4.7, 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 \( \lim_{j \to \infty} u^{\varepsilon_j}(x) \leq u^0(x) \).

It is clear now that Theorem 4.1 is a straightforward consequence of Propositions 4.11, 4.12.
Remark 4.4. We finally want to address three important points. Firstly, \( M \) is the collection of stochastic Mather measures that can be derived from the solutions of the adjoint equations \( \{ \theta^{\epsilon,\eta} \} \). It should be made clear that we do not collect all minimizing measures of (4.9) in \( M \). Also we do not need to use the minimizing properties of stochastic Mather measures (4.9) in our analysis. Of course we still derived it for the sake of completeness.

Secondly, by repeating the whole proof, we obtain that

\[
\phi^\epsilon(x) \rightarrow \tilde{u}^0(x) := \sup_{\phi \in \tilde{E}} \phi(x) \quad \text{uniformly for } \ x \in \mathbb{T}^n \text{ as } \epsilon \rightarrow 0,
\]

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

\[
\iint_{\mathbb{T}^n \times \mathbb{R}^n} u \, d\mu \leq 0 \quad \text{for all } \mu \in \tilde{M}.
\]

Thus, \( u^0 = \tilde{u}^0 \).

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

\[
\iint_{\mathbb{T}^n \times \mathbb{R}^n} H(x,p) \, d\nu(x,p) = 0 \quad \text{for all } \nu \text{ given by (4.7)}.
\]

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

4.4.1 Applications

Let us now apply Theorem 4.1 to find the limit of \( \phi^\epsilon \) for Example 4.1 in Section 4.1. By Theorem 4.1 and the above remark, we have

\[
\phi^\epsilon(x) \rightarrow \tilde{u}^0(x)
\]

\[
= \sup \left\{ w(x) : w \text{ is a solution to (4.1) s.t. } \iint_{\mathbb{T}^n \times \mathbb{R}^n} w \, d\mu(x,v) \leq 0, \ \forall \mu \in \tilde{M} \right\}
\]

uniformly for \( x \in \mathbb{T}^n \) as \( \epsilon \rightarrow 0 \). In this specific case, we can easily see

\[
\{ \delta_1/4 \} \cup \{ \delta_3/4 \} \subset \tilde{M}.
\]

Thus,

\[
\tilde{u}^0(x) \leq \sup \left\{ w(x) : w \text{ is a solution to (4.1) s.t. } \iint_{\mathbb{T}^n \times \mathbb{R}^n} w \, d\mu(x,v) \leq 0, \ \forall \mu \in \{ \delta_1/4 \} \cup \{ \delta_3/4 \} \right\}
\]

\[
= \sup \left\{ w(x) : w \text{ is a solution to (4.1) s.t. } w(1/4) \leq 0, w(3/4) \leq 0 \right\},
\]
which implies $\bar{u}^0(1/4) \leq 0$ and $\bar{u}^0(3/4) \leq 0$. On the other hand, noting that 0 is a subsolution of $(E)_\varepsilon$, by the comparison principle, we have $u^\varepsilon \geq 0$ in $\mathbb{R}$, which implies $\bar{u}^0 \geq 0$ in $\mathbb{R}$. Thus, we obtain $\bar{u}^0(1/4) = 0$ and $\bar{u}^0(3/4) = 0$, and therefore $u^0 = u_1^0 = u_2^0$, where $u_1^0, u_2^0$ are the functions defined in the example in Introduction.

![Graph of $u^0$ on [0, 1]](image)

## 4.5 Some other directions and open questions

In this section, we present other developments in the study of selection problems for Hamilton–Jacobi equations. If we consider a different type of approximation for (2.6), then the selection procedure could be quite different. Therefore, a different type of difficulties may appear in general. Let us describe briefly these directions as well as questions here in the followings.

### Discounted approximation procedure

(i) A general setting (e.g. general possibly degenerate viscous diffusions, Neumann boundary conditions): Ishii, Mitake, Tran [40] works in progress.

(ii) Aubry (uniqueness) set: The structure of solutions of (2.8) 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. In a general viscous case where the diffusion could be degenerate, there has not been any similar notions/results on the uniqueness (Aubry) set up to now.

(iii) Selection problem for nonconvex Hamilton–Jacobi equations: Most problems are open. In some examples, invariant measures and invariant sets do not exist. It is therefore extremely challenging to establish general convergence result and to describe the limit if it exists. In some specific inviscid cases, convergence result is proved (see [32]).

(iv) Applications: Theorem 4.1 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

$$H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon + \overline{H}^\varepsilon \text{ in } \mathbb{T}^n$$

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 proved. See [13, 2]. In the general setting, there are still many questions which are not solved yet. See also [63].

(ii) Finite difference approximation: In [59] the selection problem which appears in the finite difference procedure was first formulated, and the convergence was also proved there.
Chapter 5

Appendix

The reader 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 [20] (see also Crandall, Evans, and Lions [18]). 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 [24]. We refer the readers to the user guide of Crandall, Ishii, and Lions [19] for the theory of viscosity solutions of second order Hamilton–Jacobi equations. Let us for simplicity only consider the initial-value problem of Hamilton–Jacobi equation:

\[
\begin{aligned}
\text{(C)} & \quad \left\{ 
    \begin{array}{ll}
    u_t + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\
    u(x, 0) = g(x) & \text{on } \mathbb{R}^n,
    \end{array}
\right.
\end{aligned}
\]

where the Hamiltonians \( H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) is given, as is the initial function \( g : \mathbb{R}^n \to \mathbb{R} \). We will need to assume some conditions on \( H \) and \( g \) to have some a priori estimates, which will be addressed precisely later.

The original approach [20, 18] is to consider the following approximated equation

\[
\begin{aligned}
\text{(C)}_\varepsilon & \quad \left\{ 
    \begin{array}{ll}
    u_\varepsilon + H(x, Du_\varepsilon) = \varepsilon \Delta u_\varepsilon & \text{in } \mathbb{R}^n \times (0, \infty), \\
    u_\varepsilon(x, 0) = g(x) & \text{on } \mathbb{R}^n,
    \end{array}
\right.
\end{aligned}
\]

for \( \varepsilon > 0 \). The term \( \varepsilon \Delta \) in \( \text{(C)}_\varepsilon \) regularizes the Hamilton–Jacobi equations, and this is 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 locally equicontinuous on \( \mathbb{R}^n \times [0, \infty) \). We hence can use the Arzela-Ascoli theorem to deduce that

\[
u_\varepsilon \to u, \quad \text{locally uniformly in } \mathbb{R}^n \times [0, \infty),
\]

for some subsequence \( \{u_\varepsilon\} \) and some limit function \( u \in C(\mathbb{R}^n \times [0, \infty)) \). We expect that \( u \) is some kind of solution of \( \text{(C)} \) but we only have that \( u \) is continuous and absolutely no information about \( Du \) and \( u_t \). Also as \( \text{(C)} \) is fully nonlinear and not of the divergence structure, we cannot use integration by parts and weak convergence techniques to justify that \( u \) is the weak solution in such sense. We instead use the maximum principle to obtain the notion of weak solution, which is viscosity solution.
CHAPTER 5. APPENDIX

The term *viscosity solutions* is used in honor of the vanishing viscosity technique. In the modern approach, the existence of viscosity solutions can be obtained by using Perron’s method (for both first order and second order equations.) 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.

5.1 Motivation and Examples

5.1.1 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:

\[
V(y, t) = -h(y) \quad \text{on } \Gamma(t),
\]

where \( V \) is the normal velocity at each point on \( \Gamma(t) \), and \( h \in C(\mathbb{R}^{n+1}) \) 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 given real-valued auxiliary function \( u : \mathbb{R}^n \to \mathbb{R} \).

![Figure 5.1](image)

Figure 5.1 shows a picture of the situation. We note that the direction of \( x_{n+1} \) 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.

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 to the direction $x_{n+1}$. Plug this into (5.1) and we get the Hamilton–Jacobi equation

$$u_t + h(x)\sqrt{1 + |Du|^2} = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty).$$

**Example 5.1.** We consider the simplest case where $n = 1$, $h(x) \equiv 0$ and two initial data: (i) a line in Figure 5.2, (ii) a curve in Figure 5.3.

![Figure 5.2](image1)

![Figure 5.3](image2)

### 5.1.2 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 $\lambda \in [0, 1]$,

$$L(x, v) \geq \frac{L(x, v)}{|v|} \to \infty \quad \text{locally uniformly for } x \in \mathbb{R}^n \text{ as } |v| \to \infty,$$

$$L(x, \lambda v_1 + (1 - \lambda)v_2) \leq \lambda L(x, v_1) + (1 - \lambda)L(x, v_2).$$

**Inviscid cases.**

We consider the optimal control problem, for fixed $(x, t) \in \mathbb{R}^n \times [0, \infty)$,

$$\text{Minimize } \int_0^t L(\gamma(s), -\dot{\gamma}(s)) \, ds + u_0(\gamma(t))$$
over all controls $\gamma \in AC([0, t])$ with $\gamma(0) = x$. Here $u_0 \in \text{BUC}(\mathbb{R}^n)$ is given. We denote by $u(x, t)$ the minimum cost. It can be proved that $u$ solves the Cauchy problem for the Hamilton–Jacobi equation

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

where the Hamiltonian $H$ is the Fenchel-Legendre transforms of the Lagrangian:

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

Let us show a quick formal proof of this. Note first that $u$ satisfies the so-called dynamic programming principle, for $h > 0$,

$$
u(x, t + h) = \inf \left\{ \int_0^h L(\gamma(s), -\dot{\gamma}(s)) \, ds + u(\gamma(h), t) : \gamma(0) = x \right\}. \tag{5.2}
$$

The dynamic programming principle can be checked in a rough way as following

$$
u(x, t + h) = \inf \left\{ \int_0^h L(\gamma(s), -\dot{\gamma}(s)) \, ds + u(\gamma(t + h)) : \gamma(0) = x \right\} \\ \approx \int_0^h L(\gamma(s), -\dot{\gamma}(s)) \, ds + \int_0^t L(\delta(s), -\dot{\delta}(s)) \, ds + u_0(\delta(t)) \\ \approx \int_0^h L(\gamma(s), -\dot{\gamma}(s)) \, ds + u(\gamma(h), t),
$$

where we denote by $\delta(s) := \gamma(s + h)$ for $s \in [-h, t]$. Therefore, (5.2) holds. 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

$$u_t + Du \cdot \left( -\dot{\delta}(0) \right) - L \left( x, -\dot{\delta}(0) \right) \approx 0,
$$

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.

**Example 5.2** (Classical mechanics). We consider the case that $L$ is the difference between kinetic energy and a potential energy, i.e., $L(x, v) := |v|^2/2 - V(x)$ for a given function $V \in C(\mathbb{R}^n)$. Then,

$$
u(x, t) = \inf \left\{ \int_0^t \frac{1}{2} |\dot{\gamma}(s)|^2 - V(\gamma(s)) \, ds + u_0(\gamma(t)) : \gamma(0) = x \right\}
$$

solves the following Cauchy problem

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

Example 5.3 (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 problem. In short, we have

\[
u(x, t) = \inf \left\{ \int_0^t L(-\dot{\gamma}(s)) \, ds + u_0(\gamma(t)) : \gamma(0) = x \right\}
\]

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


Viscous cases.

We consider the stochastic optimal control problem

\[
\begin{align*}
\text{Minimize} & \quad E_{x,t} \left[ \int_0^t L(v(s), X(s)) \, ds + u_0(X(t)) \right] \\
\text{subject to} & \quad X^v = -\int_0^t v(s) \, ds + \sqrt{2} \int_0^t \sigma(X^v(s)) \, dW_s,
\end{align*}
\]

over all controls \( v \) from some admissible class, where \( \sigma(\mathbb{R}^n, \mathbb{M}^{m \times n}) \) for some \( m \in \mathbb{N} \) is a given matrix-valued function, and \( W_s \) denotes a standard \( d \)-dimensional Brownian motion with value in \( \mathbb{M}^{n \times m} \). Let \( u(x, t) \) be the corresponding minimum cost.

We can prove that the function \( u \) solves the Cauchy problem for the following general viscous Hamilton–Jacobi equation by using the dynamic programming principle and the Ito formula

\[
\begin{align*}
v_t - \text{tr} (A(x)D^2 u) + H(x, Du) &= 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \\
u(x, 0) &= u_0(x) \quad \text{on } \mathbb{R}^n,
\end{align*}
\]

where \( A(x) := \sigma(x)\sigma^T(x) \).

5.2 Definitions

Definition 5.1 (Viscosity subsolutions, supersolutions, solutions). A bounded, uniformly continuous function \( u \) is called a viscosity subsolution of the initial-value problem (C) provided that

- \( u(\cdot, 0) = g \) on \( \mathbb{R}^n \).
- For each \( v \in C^1(\mathbb{R}^n \times (0, \infty)) \), if \( u - v \) has a local maximum at \( (x_0, t_0) \in \mathbb{R}^n \times (0, \infty) \) then
  \[
v_t(x_0, t_0) + H(x_0, Dv(x_0, t_0)) \leq 0.
  \]

A bounded, uniformly continuous function \( u \) is called a viscosity supersolution of the initial-value problem (C) provided that
• \(u(\cdot, 0) = g\) on \(\mathbb{R}^n\).

• For each \(v \in C^1(\mathbb{R}^n \times (0, \infty))\), if \(u - v\) has a local minimum at \((x_0, t_0) \in \mathbb{R}^n \times (0, \infty)\) then
  \[v_t(x_0, t_0) + H(x_0, Dv(x_0, t_0)) \geq 0.\]

A bounded, uniformly continuous function \(u\) is called a viscosity solution of the initial-value problem (C) if \(u\) is both a subsolution, and a supersolution of (C).

**Remark 5.1.** In Definition 5.1, 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.

### 5.3 Existence

**Theorem 5.1.** Let \(u^\varepsilon\) be the solution of \((C)_{\varepsilon}\) for \(\varepsilon > 0\). Assume that there exists a subsequence \(\{u^{\varepsilon_j}\}\) such that
  \[u^{\varepsilon_j} \to u, \quad \text{locally uniformly in } \mathbb{R}^n \times [0, \infty)\]
for some \(u \in C(\mathbb{R}^n \times [0, \infty))\) bounded and uniformly continuous. Then \(u\) is a viscosity solution of (C).

**Proof.** It is enough to prove that \(u\) is a viscosity subsolution of (C). Take any \(v \in C^\infty(\mathbb{R}^n \times (0, \infty))\) and assume that \(u - v\) 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, \(u^{\varepsilon_j} - v\) has a local maximum at \((x_j, t_j)\) and
  \[(x_j, t_j) \to (x_0, t_0), \quad \text{as } j \to \infty.\]

We have \(Du^{\varepsilon_j}(x_j, t_j) = Dv(x_j, t_j),\) \(u^{\varepsilon_j}_t(x_j, t_j) = v_t(x_j, t_j),\) and \(-\Delta u^{\varepsilon_j}(x_j, t_j) \geq -\Delta v(x_j, t_j).\) Hence,
  \[v_t(x_j, t_j) + H(x_j, Dv(x_j, t_j)) = u^{\varepsilon_j}_t(x_j, t_j) + H(x_j, Du^{\varepsilon_j}(x_j, t_j))\]
  \[= \varepsilon \Delta u^{\varepsilon_j}(x_j, t_j) \leq \varepsilon_j \Delta v(x_j, t_j).\]

Let \(j \to \infty\) to imply that
  \[v_t(x_0, t_0) + H(x_0, Dv(x_0, t_0)) \leq 0.\]

**Remark 5.2.** Let us emphasize that obtaining viscosity solutions through the vanishing viscosity approach is the classical approach. This method does not work for second order equations. In general, we can use Perron’s method to prove the existence of viscosity solutions. We will present some key ideas of both arguments later here.
5.4 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 \((C)\) and \( u \) is also bounded and continuous, 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 5.2** (Touching by a \( C^1 \) function). Assume \( u : \mathbb{R}^n \to \mathbb{R} \) is continuous and differentiable at some point \( x_0 \). There exists \( v \in C^1(\mathbb{R}^n) \) such that \( u(x_0) = v(x_0) \) and \( u - v \) has a strict local maximum at \( x_0 \).

*Proof.* Without loss of generality, we assume first that \( x_0 = 0, \ u(0) = 0, \) and \( Du(0) = 0 \) (5.3).

We use (5.3) and the differentiability of \( u \) at 0 to deduce that

\[
u(x) = |x|\omega(x)
\]

where \( \omega : \mathbb{R}^n \to \mathbb{R} \) is continuous with \( \omega(0) = 0 \). For each \( r > 0 \), we define

\[
\rho(r) = \max_{x \in B_r(0)} |\omega(x)|.
\]

We see that \( \rho : [0, \infty) \to [0, \infty) \) is continuous, increasing, and \( \rho(0) = 0 \).

We define

\[
v(x) = \int_{|x|}^{2|x|} \rho(r)dr + |x|^2, \quad \text{for} \ x \in \mathbb{R}^n.
\]

It is clear that \( |v(x)| \leq |x|\rho(2|x|) + |x|^2 \), which implies

\[
v(0) = 0, \ Dv(0) = 0.
\]

Besides, for \( x \neq 0 \), explicit computations give us that

\[
Dv(x) = \frac{2x}{|x|}\rho(|x|) - \frac{x}{|x|}\rho(|x|) + 2x,
\]

and hence \( v \in C^1(\mathbb{R}^n) \).

Finally for every \( x \neq 0 \),

\[
u(x) - v(x) = |x|\omega(x) - \int_{|x|}^{2|x|} \rho(r)dr - |x|^2
\]

\[
\leq |x|\rho(|x|) - |x|\rho(|x|) - |x|^2 < 0 = u(0) - v(0).
\]

The proof is complete. \qed

Lemma 5.2 immediately implies the following.

**Theorem 5.3** (Consistency of viscosity solutions). Let \( u \) be a viscosity solution of \((C)\) and suppose that \( u \) is differentiable at \( (x_0, t_0) \in \mathbb{T}^n \times (0, \infty) \), then

\[
u_t(x_0, t_0) + H(x_0, Du(x_0, t_0)) = 0.
\]
5.5 Stability

It is really important to mention that viscosity solutions remain stable under the $L^\infty$-norm. The following proposition shows this basic fact.

**Proposition 5.4.** 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 $\mathbb{R}^n \times \mathbb{R}^n$ 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 initial-value Hamilton–Jacobi equations corresponding to $H_k$ 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 $u$ bounded and uniformly continuous. 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

$$v_t(x_0,t_0) + H(x_0,D\phi(x_0,t_0)) \leq 0.$$

We let $k \to \infty$ to obtain the result. \hfill \qed

5.6 Uniqueness

We now establish the uniqueness of a viscosity solution of (C).

**Lemma 5.5 (Extrema at a terminal time).** Assume that $u$ is a viscosity subsolution (resp., supersolution) of (C) and $u - v$ has a local maximum (resp., minimum) at a point $(x_0,t_0) \in \mathbb{R}^n \times (0,T]$ for some $v \in C^1(\mathbb{T}^n \times [0,\infty))$. Then

$$v_t(x_0,t_0) + H(x_0,Dv(x_0,t_0)) \leq 0(\geq 0).$$

The point here is that terminal time $t_0 = T$ is allowed.

**Proof.** We just need to verify the case of subsolution. Assume $u - v$ has a strict maximum at $(x_0,T)$. We define

$$\bar{v}(x,t) = v(x,t) + \frac{\varepsilon}{T-t}, \quad \text{for } (x,t) \in \mathbb{T}^n \times (0,\infty).$$

For $\varepsilon > 0$ small enough, $u - \bar{v}$ has a local maximum at $(x_\varepsilon,t_\varepsilon) \in \mathbb{T}^n \times (0,\infty)$ and $(x_\varepsilon,t_\varepsilon) \to (x_0,T)$ as $\varepsilon \to 0$. By definition of viscosity subsolutions, we have

$$v_t(x_\varepsilon,t_\varepsilon) + H(x_\varepsilon,D\bar{v}(x_\varepsilon,t_\varepsilon)) \leq 0$$

which is equivalent to

$$v_t(x_\varepsilon,t_\varepsilon) + \frac{\varepsilon}{(T-t_\varepsilon)^2} + H(x_\varepsilon,Dv(x_\varepsilon,t_\varepsilon)) \leq 0.$$
5.6. UNIQUENESS

Hence

$$v_t(x, t) + H(x, Dv(x, t)) \leq 0.$$  

We let $\varepsilon \to 0$ to achieve the result. \hfill $\Box$

We now assume further that the Hamiltonian $H$ satisfies

(A1) There exist a positive constant $C$ such that

$$|H(x, p) - H(x, q)| \leq C|p - q|,$$  

$$|H(x, p) - H(y, p)| \leq C|x - y|(1 + |p|), \quad \text{for } (x, y, p, q) \in (\mathbb{R}^n)^4.$$  

**Theorem 5.6** (Comparison Principle for $(C)$). Assume that (A1) holds. If $u$, $\tilde{u}$ are a viscosity subsolution, and supersolution of $(C)$ respectively, then $u \leq \tilde{u}$.

**Proof.** We assume by contradiction that

$$\sup_{\mathbb{R}^n \times [0, \infty)} (u - \tilde{u}) = \sigma > 0.$$  

For $\varepsilon, \lambda \in (0, 1)$, we define

$$\Phi(x, y, t, s) = u(x, t) - \tilde{u}(y, s) - \lambda(t+s) - \frac{1}{\varepsilon^2}(|x-y|^2 + (t-s)^2) - \varepsilon(|x|^2 + |y|^2), \quad \text{for } x, y \in \mathbb{R}^n, \ t, s \geq 0.$$  

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 $\Phi(x_0, y_0, t_0, s_0) \geq \sigma/2$.

We use $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(0, 0, 0, 0)$ to get

$$\lambda(t_0+s_0) + \frac{1}{\varepsilon^2}(|x_0-y_0|^2 + (t_0-s_0)^2) + \varepsilon(|x_0|^2 + |y_0|^2) \leq u(x_0, t_0) - \tilde{u}(y_0, s_0) - u(0, 0) + \tilde{u}(0, 0) \leq C.$$  

(5.6)

Hence

$$|x_0 - y_0| + |t_0 - s_0| \leq C\varepsilon, \quad |x_0| + |y_0| \leq \frac{C}{\varepsilon^{1/2}}.$$  

(5.7)

We next use $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(x_0, x_0, t_0, t_0)$ to deduce that

$$\frac{1}{\varepsilon^2}(|x_0 - y_0|^2 + (t_0 - s_0)^2) \leq \tilde{u}(x_0, t_0) - \tilde{u}(y_0, s_0) + \lambda(t_0 - s_0) + \varepsilon(x_0 - y_0) \cdot (x_0 + y_0).$$

In view of (5.7) and the uniformly continuity of $\tilde{u}$, we get

$$|x_0 - y_0| + |t_0 - s_0| = o(\varepsilon).$$  

(5.8)

By (5.7) and (5.8), we can take $\varepsilon > 0$ small enough so that $s_0, t_0 \geq \mu > 0$ 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 \(\Phi\), \(u - v\) has a maximum at \((x_0, t_0)\) for
\[
v(x, t) = \tilde{u}(y_0, s_0) + \lambda(t + s_0) + \frac{1}{\varepsilon^2}(|x - y_0|^2 + (t - s_0)^2) + \varepsilon(|x|^2 + |y_0|^2).
\]
By definition of viscosity subsolutions,
\[
\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} + H(x_0, \frac{2(x_0 - y_0)}{\varepsilon^2} + 2\varepsilon x_0) \leq 0. \tag{5.9}
\]
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
\[
-\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} + H(y_0, \frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0) \geq 0. \tag{5.10}
\]
Subtract (5.10) from (5.9)
\[
2\lambda \leq H(y_0, \frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0) - H(x_0, \frac{2(x_0 - y_0)}{\varepsilon^2} + 2\varepsilon x_0) \leq C\varepsilon(|x_0| + |y_0|) + C|x_0 - y_0|. \tag{5.11}
\]
We let \(\varepsilon \to 0\) to discover that \(\lambda \leq 0\), which is the contradiction.

By using the comparison principle above, we obtain the following uniqueness result immediately.

**Theorem 5.7 (Uniqueness of viscosity solution).** Under assumption (A1) there exists at most one viscosity solution of (C).

### 5.7 Lipschitz estimates

We provide here a way to obtain Lipschitz estimates (a priori estimates) for \(u^\varepsilon\), which is the solution of \((C)_\varepsilon\). Assume for simplicity the followings

(A2) \[
\lim_{|p| \to \infty} \inf_{x \in \mathbb{R}^n} \left( H(x, p)^2 - D_x H(x, p) \cdot p \right) = +\infty.
\]

(A3) \(u_0 \in C^2(\mathbb{R}^n)\) and \(\|u_0\|_{C^2(\mathbb{R}^n)} \leq C < +\infty\).

**Theorem 5.8.** Assume that (A2)–(A3) hold. There exists a constant \(C > 0\) independent of \(\varepsilon\) such that
\[
\|u^\varepsilon_t\|_{L^\infty} + \|Du^\varepsilon\|_{L^\infty} \leq C.
\]

**Sketch of proof.** We first note that for \(C > 0\) sufficiently large, \(u_0 \pm Ct\) are respectively a supersolution and a subsolution of \((C)_\varepsilon\). By the comparison principle, we get
\[
u_0 - Ct \leq u^\varepsilon \leq u_0 + Ct.
\]
This, together with the comparison principle once more, yields that \(\|u_t\|_{L^\infty} \leq C\).
Next, set $\phi = |D u^\varepsilon|^2/2$. We have
\[ \phi_t + D_p H \cdot D \phi + D_2 H \cdot Du^\varepsilon = \varepsilon \Delta \phi - \varepsilon |D^2 u^\varepsilon|^2. \]
For $\varepsilon > 0$ sufficiently small, one has
\[ \varepsilon |D^2 u^\varepsilon|^2 \geq 2 (\varepsilon \Delta u^\varepsilon)^2 = 2(u_t^\varepsilon + H(x, Du^\varepsilon))^2 \geq H(x, Du^\varepsilon)^2 - C. \]
Thus,
\[ \phi_t + D_p H \cdot D \phi + (H(x, Du^\varepsilon)^2 + D_2 H \cdot Du^\varepsilon - C) \leq \varepsilon \Delta \phi. \]
By the maximum principle and (A2), we get the desired result. \hfill \Box

5.8 The Perron method

Theorem 5.9. Let $f$ and $g$ be a subsolution and a supersolution of $(C)$, respectively. Assume that $f \leq g$ in $\mathbb{R}^n \times (0, \infty)$. Then, the function $u : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ defined by
\[ u(x, t) = \sup \{ v(x, t) \mid f \leq v \leq g \text{ in } \mathbb{R}^n \times [0, \infty), v \text{ is a subsolution of } (C) \} \]
is a solution of $(C)$. Moreover, $f \leq u \leq g$ in $\mathbb{R}^n \times [0, \infty)$.

The above construction of solutions is called Perron’s method. The use of this method in the area of viscosity solutions was introduced by H. Ishii [36]. For simplicity in this proof, we will assume $u$ is continuous.

Sketch of proof. Set
\[ S^- := \{ v : \mathbb{R}^n \times [0, \infty) \to \mathbb{R} \mid f \leq v \leq g \text{ in } \mathbb{R}^n \times [0, \infty), v \text{ is a subsolution of } (C) \}. \]
Since $f \in S^-$, $S^- \neq \emptyset$. It is clear that $f \leq u \leq g$ in $\mathbb{R}^n \times [0, \infty)$. Thus, $u$ is locally bounded in $\mathbb{R}^n \times [0, \infty)$ and a subsolution of $(C)$.

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 $u \leq g$, we deduce $\phi \leq g$ in $\mathbb{R}^n \times (0, \infty)$. Assume that $u(y, s) = g(y, s)$. Then $g - \phi$ attains a minimum at $(y, s)$. Therefore, we have $0 \leq \phi_t(y, s) + H(y, D\phi(y, s))$, 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
\[ g(x, t) \geq \phi(x, t) + \frac{\varepsilon_0}{2} \quad \text{for all } (x, t) \in B(y, r) \times (s - r, s + r), \]
\[ \phi_t + H(x, D\phi) \leq 0 \quad \text{for all } (x, t) \in B(y, r) \times (s - r, s + r). \]
Set $U = B(y, r) \times (s - r, s + r)$ and $\varepsilon = \frac{1}{2} \min\{\varepsilon_0, \min_{\partial U} (u - \phi)(x)\} > 0$. 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 $u \leq \tilde{u}$ and $f \leq \tilde{u} \leq g$ 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 $u$. \qed
Bibliography


[40] H. Ishii, H. Mitake, H. V. Tran, in preparation.


