J. Math. Soc. Japan
Vol. 70, No. 1 (2018) pp. 345–364
doi: 10.2969/jmsj/07017534

The selection problem for discounted Hamilton–Jacobi equations: some non-convex cases

By Diogo A. GOMES, Hiroyoshi MITAKE and Hung V. TRAN

(Received May 29, 2016)

Abstract. Here, we study the selection problem for the vanishing discount approximation of non-convex, first-order Hamilton–Jacobi equations. While the selection problem is well understood for convex Hamiltonians, the selection problem for non-convex Hamiltonians has thus far not been studied. We begin our study by examining a generalized discounted Hamilton–Jacobi equation. Next, using an exponential transformation, we apply our methods to strictly quasi-convex and to some non-convex Hamilton–Jacobi equations. Finally, we examine a non-convex Hamiltonian with flat parts to which our results do not directly apply. In this case, we establish the convergence by a direct approach.

1. Introduction.

Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ be the standard n-dimensional torus and fix a continuous Hamiltonian, $H : \mathbb{T}^n \times \mathbb{R}^d \to \mathbb{R}$. Here, we require H to be *coercive*; that is,

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

We do not, however, assume convexity. The *ergodic* Hamilton–Jacobi equation is the partial differential equation (PDE)

$$H(x, Du) = \overline{H} \qquad \text{in } \mathbb{T}^n, \tag{E}$$

and, for $\varepsilon > 0$, the corresponding *discounted problem* is

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

In (E), the unknown is a pair, $(u, \overline{H}) \in C(\mathbb{T}^n) \times \mathbb{R}$, whereas in (D$_\varepsilon$), the unknown is a function, $u^\varepsilon \in C(\mathbb{T}^n)$. In both (E) and (D$_\varepsilon$), we consider solutions in the viscosity sense. Here, we are interested in the *vanishing discount limit*, $\varepsilon \to 0$ in (D$_\varepsilon$), and in the characterization of the limit, u, of u^ε as a particular solution of (E).

The problem (D$_\varepsilon$) arises in optimal control theory and zero-sum differential game theory where ε is a discount factor. Moreover, (D$_\varepsilon$) plays an essential role in the homogenization of first-order Hamilton–Jacobi equations. For example, in the study of

---

2010 *Mathematics Subject Classification.* Primary 35B40; Secondary 37J50, 49L25.
Key Words and Phrases. nonconvex Hamilton–Jacobi equations, discounted approximation, ergodic problems, nonlinear adjoint methods.

homogenization in [**20**], the vanishing discount limit is used to construct solutions to the ergodic problem. The ergodic problem is sometimes called the *cell problem* or the *additive eigenvalue problem*. The PDE (D_ε) is also called the *discounted approximation* of the ergodic problem. Properties of the solutions of (E) are relevant in dynamical systems, namely in weak Kolmogorov–Arnold–Moser (KAM) theory (see [**11**]), and they have applications in the study of the long-time behavior of Hamilton–Jacobi equations [**12**], [**8**].

In recent years, there was significant progress in the analysis of non-convex Hamilton–Jacobi equations. Some remarkable results include the characterization of the shock structure of the gradient of solutions [**10**], construction of invariant measures in the spirit of weak KAM theory [**7**], and homogenization in random media [**2**], [**3**] (see also [**15**]). A better grasp of the vanishing discount problem for non-convex Hamiltonians is essential to improving our understanding of the nature of viscosity solutions of Hamilton–Jacobi equations.

Before we proceed, we recall some elementary properties of (E) and (D_ε). First, there exists a unique real constant, $\overline{H}$, such that (E) has viscosity solutions [**20**]. This constant is often called the *ergodic constant* or the *effective Hamiltonian*. However, in general, (E) does not have a unique solution, not even up to additive constants. The lack of uniqueness is a central issue in the study of the asymptotic behavior of u^ε as $\varepsilon \to 0$. As (D_ε) is strictly monotone with respect to u^ε for $\varepsilon > 0$, Perron's method gives the existence of a unique viscosity solution, u^ε. By the coercivity of the Hamiltonian, we have that

$$\|Du^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \le C \quad \text{for some } C > 0 \text{ independent of } \varepsilon. \tag{1.1}$$

We fix $x_0 \in \mathbb{T}^n$. The preceding estimate implies that

$$\{u^\varepsilon(\,\cdot\,) - u^\varepsilon(x_0)\}_{\varepsilon>0}$$

is uniformly bounded and equi-Lipschitz continuous in $\mathbb{T}^n$. Therefore, by the Arzelá-Ascoli theorem, there exists a subsequence, $\{\varepsilon_j\}_{j\in\mathbb{N}}$, with $\varepsilon_j \to 0$ as $j \to \infty$, a constant $\overline{H} \in \mathbb{R}$, and a function, $u \in C(\mathbb{T}^n)$, such that

$$\varepsilon_j u^{\varepsilon_j} \to -\overline{H}, \quad u^{\varepsilon_j} - u^{\varepsilon_j}(x_0) \to u \in C(\mathbb{T}^n), \tag{1.2}$$

uniformly in $\mathbb{T}^n$ as $j \to \infty$. By a standard viscosity solution argument, we see that $(u, \overline{H})$ solves (E). However, the convergence in (1.2) and the function u may depend on the choice of the subsequence $\{\varepsilon_j\}$ in this argument. Thus, the limit as $\varepsilon \to 0$ of u^ε may not exist.

Our primary goal is to study the *selection problem* for (D_ε); that is, we wish to understand whether or not the limit as $\varepsilon \to 0$ of u^ε exists and, if it does, what the characterization of this limit is. This problem was proposed in [**20**] (see also [**4**, Remark 1.2, page 400]). It remained unsolved for almost 30 years. Recently, there was substantial progress in the case of convex Hamiltonians. First, a partial characterization of the possible limits was given in [**16**] in terms of the Mather measures (see, for example, [**11**], [**21**], [**22**]). Then, the convergence to a unique limit and its characterization were established in [**9**] using weak KAM theory. Further selection problems including the case of degenerate viscous Hamilton–Jacobi equations were addressed using the nonlinear adjoint method in [**23**]. Finally, an analogous convergence result for the case of Neumann boundary conditions was examined in [**1**]. The selection problem for possibly degenerate, fully nonlinear Hamilton–Jacobi–Bellman equations was considered in [**17**], [**18**]. A related selection problem was addressed in [**5**], [**19**] and selection questions motivated by finite-difference schemes were examined in [**24**]. In all these papers, the convexity of the Hamiltonian was essential and no extensions to non-convex Hamiltonians were offered. Thus, the selection problem in the non-convex setting has yet to be studied.

Here, we develop methods to examine the selection problem for (D_ε) for non-convex Hamilton–Jacobi equations. Our main technical device is a selection theorem for a class of nonlinearly discounted Hamilton–Jacobi equations, Theorem 2.1. Although this theorem is of independent interest, we focus here on two main applications: the case of strictly quasiconvex Hamiltonians in Theorem 2.2 and the case of double-well problems in Theorem 2.3. These results and the main assumptions are stated in the next section. Next, in Section 3, we introduce a generalized discounted approximation, examine its convergence, and prove Theorem 2.1. Our proof is based on the method introduced in [**23**]. Then, in Section 4, we study strictly quasi-convex Hamiltonians and prove Theorem 2.2. Next, in Section 5, we consider the double-well Hamiltonian–Jacobi equation and prove Theorem 2.3. Finally, in Section 6, we examine the convergence for a quasi-convex Hamiltonian with flat parts. The results in this section do not follow from the general theory developed in Section 4 and they require a distinct approach. In this final section, we discuss maximal subsolutions and the Aubry set. In particular, we provide an answer to Question 12 in the list of open problems [**6**] from the conference "New connections between dynamical systems and PDEs" at the American Institute of Mathematics in 2003.

2. Assumptions and main results.

Here, we discuss the main assumptions used in the paper and present the main results.

Let $G \in C^1(\mathbb{T}^n \times \mathbb{R}^n)$ and $f \in C^2(\mathbb{T}^n \times \mathbb{R})$ satisfy

(A1) uniformly for $x \in \mathbb{T}^n$,

$$\lim_{|p|\to\infty}\left(\frac{1}{2|p|}G(x,p)^2 - |D_xG(x,p)|\right) = +\infty;$$

(A2) $p \mapsto G(x,p)$ is convex;

(A3) $f_r(x,r) > 0$ for all $(x,r) \in \mathbb{T}^n \times \mathbb{R}$ and there exists $M > 0$ such that, for all $x \in \mathbb{T}^n$,

$$f(x,-M) \le -G(x,0) \le f(x,M).$$

We consider the following generalization of the discounted problem

$$f(x,\varepsilon v^\varepsilon) + G(x,Dv^\varepsilon) = 0 \quad \text{in } \mathbb{T}^n. \tag{GE_ε}$$

Because of (A3), for $\varepsilon > 0$, the left-hand side of (GE_ε) is strictly monotone in v^ε. Therefore, (GE_ε) has a comparison principle. Furthermore, the coercivity of G given by (A1) implies that $\|Dv^\varepsilon\|_{L^\infty} < C$ for some constant, C, independent of ε (see Lemma 3.1 below). Thus, arguing as before, we see that there exists a constant, $c \in \mathbb{R}$, such that $\varepsilon v^\varepsilon \to -c$ in $C(\mathbb{T}^n)$ as $\varepsilon \to 0$. Accordingly, we consider the following ergodic problem associated with (GE_ε):

$$f(x,-c) + G(x,Dv) = 0 \quad \text{in } \mathbb{T}^n. \tag{GE}$$

Without loss of generality, by replacing f by $f_c(x,r) = f(x,-c+r)$, if necessary, we can assume that $c = 0$. As in (1.2), for $x_0 \in \mathbb{T}^n$ fixed, we pick a subsequence $\varepsilon_j \to 0$ such that $\{v^{\varepsilon_j} - v^{\varepsilon_j}(x_0)\}$ converges to v uniformly in $\mathbb{T}^n$. Clearly, v is a solution to (GE) with $v(x_0) = 0$ and $\|Dv\|_{L^\infty} < C$. By (A3), $v - \|v\|_{L^\infty}$ and $v + \|v\|_{L^\infty}$ are a subsolution and a supersolution of (GE_ε), respectively. Hence,

$$v - \|v\|_{L^\infty} \le v^\varepsilon \le v + \|v\|_{L^\infty}.$$

Theorem 2.1. *Assume* (A1)–(A3) *hold and that* $c = 0$ *in* (GE)*. Let* v^ε *be the viscosity solution of* (GE_ε)*. Let* $\mathcal{M}$ *be the set of probability measures given by* (3.5)*. Let* $\mathcal{E}$ *be the family of subsolutions* w *of* (GE) *that satisfy*

$$\iint_{\mathbb{T}^n \times \mathbb{R}^n} f_r(x,0) w(x)\, d\mu \le 0 \qquad \text{for all } \mu \in \mathcal{M}. \tag{2.1}$$

Define

$$v^0(x) = \sup_{w \in \mathcal{E}} w(x).$$

Then, we have

$$v^\varepsilon(x) \to v^0(x), \quad \text{uniformly for } x \in \mathbb{T}^n \text{ as } \varepsilon \to 0. \tag{2.2}$$

A Hamiltonian, $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, is *strictly quasi-convex* if it satisfies the following assumption:

(A4) For any $a \in \mathbb{R}$ and $x \in \mathbb{T}^n$, the set $\{p \in \mathbb{R}^n : H(x,p) \le a\}$ is convex, and there exists a constant, $\lambda_0 > 0$, such that

$$\lambda_0^2 D_pH(x,p) \otimes D_pH(x,p) + \lambda_0 D^2_{pp}H(x,p) \ge 0 \quad \text{for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n.$$

If the preceding assumption holds, we have that $G(x,p) := e^{\lambda_0 H(x,p)}$ is a convex function of p. In addition to (A4), it is useful to introduce the following growth assumption on G.

(A5) $G(x,p) = e^{\lambda_0 H(x,p)}$ satisfies (A1).

Theorem 2.2. *Assume* (A4) *and* (A5) *hold. Let* u^ε *solve* (D_ε)*. Then,* u^ε *solves* (GE_ε) *for* $f(x,r) = -e^{-\lambda_0 r}$ *for* $(x,r) \in \mathbb{T}^n \times \mathbb{R}$*, and* G *as in Assumption* (A5)*. Moreover,* *as* $\varepsilon \to 0$, u^ε *converges uniformly to the function* u^0 *determined by the conditions in Theorem* 2.1.

Remark 1. While assumption (A4) is somewhat technical, it holds if, for each fixed $x \in \mathbb{T}^n$ and each $s > \min H(x,\,\cdot\,)$, the level set $\{p \in \mathbb{R}^n : H(x,p) = s\}$ is a closed $(n-1)$-dimensional manifold whose second fundamental form is strictly positive.

More precisely, the following assumption implies (A4):

(A4′) for each fixed $x \in \mathbb{T}^n$, and each $s > s_0 = \min H(x,\,\cdot\,)$, the level set $M_s = \{p \in \mathbb{R}^n : H(x,p) = s\}$ is a closed manifold of dimension $n-1$ and, for each $p \in M_s$, there exists $c = c(s) > 0$ such that

$$(B_p v) \cdot v \geq c|v|^2 \quad \text{for all } v \in T_p M_s,$$

where $T_p M_s$ is the tangent plane to M_s at p and $B_p : T_p M_s \times T_p M_s \to \mathbb{R}$ is the second fundamental form of M_s at p. Furthermore, there exists $\alpha > 0$ such that, for each $p \in \partial M_{s_0} = \partial\{p \in \mathbb{R}^n : H(x,p) = s_0\}$, we have

$$D^2_{pp} H(x,p) \geq \alpha I_n,$$

where I_n is the identity matrix of size n

(see [**7**, Section 9.7] for details). The preceding condition is satisfied by a broad class of quasi-convex Hamiltonians of which a typical example is

$$H(x,p) = K(|p|) + V(x),$$

where $K : [0,\infty) \to \mathbb{R}$ is of class C^2 and satisfies

$$K'(0) = 0, \; K''(0) > 0, \text{ and } K'(s) > 0 \text{ for } s > 0.$$

In Section 5, we consider an alternative approach to the non-convex, double-well Hamiltonian in one-dimensional space,

$$H(x,p) = (|p+P|^2 - 1)^2 - V(x), \tag{2.3}$$

where $P \in \mathbb{R}$ and $V : \mathbb{T} \to \mathbb{R}$ is a continuous function satisfying

$$\min_{\mathbb{T}} V = 0 \quad \text{and} \quad \max_{\mathbb{T}} V < 1. \tag{2.4}$$

Although this Hamiltonian does not satisfy (A4), we prove the following convergence result.

Theorem 2.3. *Let H be given by* (2.3). *Let u^ε be the corresponding solution of* (D_ε) *for a fixed $P \in \mathbb{R}$. Then, there exists a solution, $u^0 \in C(\mathbb{T})$, of*

$$(|P + Du^0|^2 - 1)^2 - V(x) = \overline{H}(P) \quad \text{in } \mathbb{T} \tag{2.5}$$

such that

$$\lim_{\varepsilon\to 0}\left(u^\varepsilon + \frac{\overline{H}(P)}{\varepsilon}\right) = u^0 \quad \textit{in } C(\mathbb{T}).$$

Remark 2. In the proof of the preceding theorem, we also obtain a characterization of limit u_0 that depends on the value of P (see Section 5).

3. A generalization of the discounted approximation.

Here, we use the nonlinear adjoint method [**10**] (see also [**25**]) and the strategy introduced in [**23**] for the study of (GE$_\varepsilon$) to investigate the limit $\varepsilon \to 0$.

3.1. A regularized problem and the construction of Mather measures.

To study (GE$_\varepsilon$), we introduce the following regularized problem. For each $\eta > 0$, we consider

$$f(x, \varepsilon v^{\varepsilon,\eta}) + G(x, Dv^{\varepsilon,\eta}) = \eta^2 \Delta v^{\varepsilon,\eta} \qquad \text{in } \mathbb{T}^n. \tag{A$^\eta_\varepsilon$}$$

Lemma 3.1. *Suppose that* (A1) *and* (A3) *hold. Then, there exists a constant, $C > 0$, independent of ε and η such that, for any solution $v^{\varepsilon,\eta}$ of* (A$^\eta_\varepsilon$)*, we have*

$$\|Dv^{\varepsilon,\eta}\|_{L^\infty(\mathbb{T}^n)} \le C. \tag{3.1}$$

Proof. Thanks to (A3), $-\varepsilon^{-1}M$ and $\varepsilon^{-1}M$ are a subsolution and a supersolution of (A$^\eta_\varepsilon$), respectively. We use the comparison principle to get

$$-\varepsilon^{-1}M \le v^{\varepsilon,\eta} \le \varepsilon^{-1}M \tag{3.2}$$

in $\mathbb{T}^n$. In particular, $|f(x, \varepsilon v^{\varepsilon,\eta})| \le C$ in $\mathbb{T}^n$ for $C = \max_{x\in\mathbb{T}^n,\ |r|\le M} |f(x,r)|$.

Now, we prove the Lipschitz bound using Bernstein's method. First, we set $\phi := |Dv^{\varepsilon,\eta}|^2/2$. Differentiating (A$^\eta_\varepsilon$) in x and multiplying by $Dv^{\varepsilon,\eta}$, we get

$$\varepsilon f_r |Dv^{\varepsilon,\eta}|^2 + (D_x f + D_x G)\cdot Dv^{\varepsilon,\eta} + D_p G \cdot D\phi = \eta^2(\Delta\phi - |D^2 v^{\varepsilon,\eta}|^2).$$

Next, we choose $x_0 \in \mathbb{T}^n$ such that $\phi(x_0) = \max_{\mathbb{T}^n}\phi$. According to (A3), we obtain

$$(D_x f + D_x G)\cdot Dv^{\varepsilon,\eta} + \eta^2 |D^2 v^{\varepsilon,\eta}|^2 \le 0 \quad \text{at } x_0 \in \mathbb{T}^n.$$

For $\eta < n^{-1/2}$, we have

$$\eta^2 |D^2 v^{\varepsilon,\eta}|^2 \ge |\eta^2 \Delta v^{\varepsilon,\eta}|^2 = |f(x,\varepsilon v^{\varepsilon,\eta}) + G(x, Dv^{\varepsilon,\eta})|^2 \ge \frac{1}{2} G(x_0, Dv^{\varepsilon,\eta})^2 - C$$

for some $C > 0$. Therefore, we obtain

$$\frac{1}{2} G(x_0, Dv^{\varepsilon,\eta})^2 + (D_x f + D_x G)\cdot Dv^{\varepsilon,\eta} \le C,$$

which, together with (A1), yields (3.1). □

Due to (3.2), for every ε fixed, $\|v^{\varepsilon,\eta}\|_\infty$ is bounded. The Lipschitz bound (3.1) and the uniqueness of the solution of (GE_ε) give that $v^{\varepsilon,\eta} \to v^\varepsilon$ in $C(\mathbb{T}^n)$ as $\eta \to 0$.

Fix $x_0 \in \mathbb{T}^n$ and let δ_{x_0} denote the Dirac delta at x_0. Next, we consider the linearization of ($\mathrm{A}_\varepsilon^\eta$) and introduce the corresponding adjoint equation

$$\varepsilon f_r(x,0)\theta^{\varepsilon,\eta} - \operatorname{div}(D_pG(x,Dv^{\varepsilon,\eta})\theta^{\varepsilon,\eta}) = \eta^2\Delta\theta^{\varepsilon,\eta} + \varepsilon\delta_{x_0} \qquad \text{in } \mathbb{T}^n. \tag{$\mathrm{AJ}_\varepsilon^\eta$}$$

Integrating ($\mathrm{AJ}_\varepsilon^\eta$) in $\mathbb{T}^n$ and using the maximum principle, we get the following proposition.

Proposition 3.2. *Let $v^{\varepsilon,\eta}$ solve* ($\mathrm{A}_\varepsilon^\eta$) *and let $\theta^{\varepsilon,\eta}$ solve* ($\mathrm{AJ}_\varepsilon^\eta$). *Then, we have*

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

In light of Lemma 3.1 and of the Riesz representation theorem, there exists a nonnegative Radon measure, $\nu^{\varepsilon,\eta}$, on $\mathbb{T}^n \times \mathbb{R}^n$ such that

$$\int_{\mathbb{T}^n} \psi(x,Dv^{\varepsilon,\eta})\theta^{\varepsilon,\eta}(x)\,dx = \iint_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,p)\,d\nu^{\varepsilon,\eta}(x,p), \quad \forall\,\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n). \tag{3.3}$$

Because $f_r(x,0) > 0$ for all $x \in \mathbb{T}^n$ and because of Proposition 3.2, we have

$$\frac{1}{\max_{x\in\mathbb{T}^n} f_r(x,0)} \le \iint_{\mathbb{T}^n\times\mathbb{R}^n} d\nu^{\varepsilon,\eta} \le \frac{1}{\min_{x\in\mathbb{T}^n} f_r(x,0)}.$$

Therefore, there are two subsequences, ε_j and η_k, with $\varepsilon_j \to 0$ and $\eta_k \to 0$ as $j,k \to \infty$ and corresponding probability measures, $\nu^{\varepsilon_j}, \nu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$, also exist such that

$$\begin{aligned} &\nu^{\varepsilon_j,\eta_k} \rightharpoonup \nu^{\varepsilon_j} && \text{as } k \to \infty, \\ &\nu^{\varepsilon_j} \rightharpoonup \nu && \text{as } j \to \infty, \end{aligned} \tag{3.4}$$

weakly in the sense of measures. The limit ν depends on x_0 and on the subsequences $\{\varepsilon_j\}$ and $\{\eta_k\}$. Thus, when we need to highlight this explicit dependence, we write it as $\nu = \nu(x_0,\{\varepsilon_j\},\{\eta_k\})$. Next, we define the family of measures, $\mathcal{M} \subset \mathcal{P}$, as

$$\mathcal{M} = \bigcup_{x_0\in\mathbb{T}^n,\{\varepsilon_j\},\{\eta_k\}} \nu(x_0,\{\varepsilon_j\},\{\eta_k\}). \tag{3.5}$$

Proposition 3.3. *Suppose that* (A1) and (A3) *hold. Then, for any $\nu \in \mathcal{M}$, we have*

(i) $$\iint_{\mathbb{T}^n\times\mathbb{R}^n} \big(D_pG(x,p)\cdot p - G(x,p)\big)\,d\nu(x,p) = \iint_{\mathbb{T}^n\times\mathbb{R}^n} f(x,0)\,d\nu(x,p),$$

(ii) $$\iint_{\mathbb{T}^n\times\mathbb{R}^n} D_pG(x,p)\cdot D\varphi\,d\nu(x,p) = 0 \quad \text{for any } \varphi \in C^1(\mathbb{T}^n).$$

Proof. We first prove (i). Note that ($\mathrm{AJ}_\varepsilon^\eta$) can be rewritten as

$$f(x,\varepsilon v^{\varepsilon,\eta}) + D_pG(x,Dv^{\varepsilon,\eta})\cdot Dv^{\varepsilon,\eta} - \eta^2\Delta v^{\varepsilon,\eta} = D_pG(x,Dv^{\varepsilon,\eta})\cdot Dv^{\varepsilon,\eta} - G(x,Dv^{\varepsilon,\eta}).$$

Let $\theta^{\varepsilon,\eta}$ solve (AJ$^\eta_\varepsilon$). Multiplying the previous equation by $\theta^{\varepsilon,\eta}$, integrating on $\mathbb{T}^n$, and using integration by parts, we get

$$\begin{aligned}&\int_{\mathbb{T}^n} (D_pG(x, Dv^{\varepsilon,\eta}) \cdot Dv^{\varepsilon,\eta} - G(x, Dv^{\varepsilon,\eta}))\theta^{\varepsilon,\eta}\, dx \\ &\quad= \int_{\mathbb{T}^n} f(x, \varepsilon v^{\varepsilon,\eta})\theta^{\varepsilon,\eta}\, dx - \int_{\mathbb{T}^n} \left(\operatorname{div}(D_pG\theta^{\varepsilon,\eta}) + \eta^2\Delta\theta^{\varepsilon,\eta}\right) v^{\varepsilon,\eta}\, dx \\ &\quad= \int_{\mathbb{T}^n} (f(x, \varepsilon v^{\varepsilon,\eta}) - \varepsilon f_r(x, 0)v^{\varepsilon,\eta})\theta^{\varepsilon,\eta}\, dx + \varepsilon v^{\varepsilon,\eta}(x_0).\end{aligned}$$

We use (3.3), set $\eta = \eta_k$, and let $k \to \infty$. Finally, we set $\varepsilon = \varepsilon_j$ and let $j \to \infty$ to get (i).

Next, we multiply (AJ$^\eta_\varepsilon$) by $\varphi \in C^1(\mathbb{T}^n)$. Then, we integrate on $\mathbb{T}^n$, use integration by parts, and set $\eta = \eta_k$ and $\varepsilon = \varepsilon_j$. Finally, we take the limit $k \to \infty$ and then $j \to \infty$ to obtain (ii). $\square$

3.2. Key estimates.

Next, we use the nonlinear adjoint method to establish estimates for the solutions of (GE$_\varepsilon$). These estimates are essential ingredients of our convergence result for (GE$_\varepsilon$).

Lemma 3.4. *Suppose that* (A1)–(A3) *hold. Let* v^ε *solve* (GE$_\varepsilon$). *Then, as* $\varepsilon \to 0$,

$$\iint_{\mathbb{T}^n\times\mathbb{R}^n} f_r(x, 0)v^\varepsilon(x)\, d\nu(x, p) \le o(1) \quad \textit{for all } \nu \in \mathcal{M}.$$

Proof. Let $\gamma \in C_c^\infty(\mathbb{R}^n)$ be a standard mollifier; that is, $\gamma \ge 0$, $\operatorname{supp}\gamma \subset \overline{B}(0, 1)$ and $\|\gamma\|_{L^1(\mathbb{R}^n)} = 1$. For each $\eta > 0$, set $\gamma^\eta(y) := \eta^{-n}\gamma(\eta^{-1}y)$ for $y \in \mathbb{R}^n$ and define

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

Then

$$\begin{aligned}G(x, D\psi^\eta) &= G(x, Dv^\varepsilon * \gamma^\eta) = G\left(x, \int_{\mathbb{R}^n} Dv^\varepsilon(x - y)\gamma^\eta(y)\, dy\right) \\ &\le \int_{\mathbb{R}^n} G(x, Dv^\varepsilon(x - y))\gamma^\eta(y)\, dy \le \int_{\mathbb{R}^n} G(x - y, Dv^\varepsilon(x - y))\gamma^\eta(y)\, dy + C\eta \\ &= -\int_{\mathbb{R}^n} f(x - y, \varepsilon v^\varepsilon(x - y))\gamma^\eta(y)\, dy + C\eta \\ &\le -\int_{\mathbb{R}^n} f(x, \varepsilon v^\varepsilon(x))\gamma^\eta(y)\, dy + C(\eta + \varepsilon\eta) \le G(x, Dv^\varepsilon) + C\eta \quad \text{for } \varepsilon < 1.\end{aligned} \tag{3.6}$$

We have used the fact that $\|Dv^\varepsilon\|_{L^\infty}$ is bounded uniformly in ε, and the Jensen inequality in the above computation.

Besides, we use Taylor's expansion to get

$$f(x, \varepsilon v^\varepsilon) = f(x, 0) + \varepsilon f_r(x, 0)v^\varepsilon + o(\varepsilon) \quad \text{as } \varepsilon \to 0. \tag{3.7}$$

Using (3.6), (3.7), and the convexity of G, we obtain

$$\begin{aligned}0 &= f(x,\varepsilon v^\varepsilon) + G(x,Dv^\varepsilon) \geq f(x,0) + \varepsilon f_r(x,0)v^\varepsilon + G(x,D\psi^\eta) - C\eta - o(\varepsilon)\\ &\geq f(x,0) + \varepsilon f_r(x,0)v^\varepsilon + G(x,p) + D_pG(x,p)\cdot(D\psi^\eta - p) - C\eta - o(\varepsilon)\end{aligned}$$

for $p \in \mathbb{R}^n$. Next, we integrate the preceding inequality with respect to $d\nu(x,p)$ for $\nu \in \mathcal{M}$ and use properties (i) and (ii) of Proposition 3.3 to conclude that

$$\iint_{\mathbb{T}^n\times\mathbb{R}^n} f_r(x,0)v^\varepsilon(x)\,d\nu(x,p) \leq \frac{C\eta}{\varepsilon} + o(1).$$

Finally, we let $\eta \to 0$ to achieve the desired result. $\square$

Lemma 3.5. *Suppose that* (A1)–(A3) *hold. Let $w \in C(\mathbb{T}^n)$ be a subsolution of* (GE). *Then, we have*

$$v^\varepsilon(x_0) \geq w(x_0) - \iint_{\mathbb{T}^n\times\mathbb{R}^n} f_r(x,0)w(x)\,d\nu^\varepsilon(x,p) + o(1) \quad \textit{as } \varepsilon \to 0$$

for all $x_0 \in \mathbb{T}^n$, where ν^ε is a weak limit in the sense of measures of a subsequence of $\nu^{\varepsilon,\eta}$ as $\eta \to 0$.

Proof. For $\eta > 0$, let $w^\eta := w * \gamma^\eta$, where γ^η is a mollifier as in the proof of Lemma 3.4. Because $\eta^2|\Delta w^\eta| \leq C\eta$, we use an argument similar to the one in (3.6) to obtain

$$f(x,0) + G(x,Dw^\eta) \leq \eta^2\Delta w^\eta + C\eta \quad \text{in } \mathbb{T}^n. \tag{3.8}$$

Now, using (3.7), we rewrite $(\mathrm{A}_\varepsilon^\eta)$ as

$$f(x,0) + \varepsilon f_r(x,0)v^{\varepsilon,\eta} + o(\varepsilon) + G(x,Dv^{\varepsilon,\eta}) = \eta^2\Delta v^{\varepsilon,\eta}. \tag{3.9}$$

Next, we subtract (3.8) from (3.9) to get

$$\begin{aligned}C\eta + \varepsilon f_r(x,0)v^{\varepsilon,\eta} + o(\varepsilon) &\geq G(x,Dw^\eta) - G(x,Dv^{\varepsilon,\eta}) - \eta^2\Delta(w^\eta - v^{\varepsilon,\eta})\\ &\geq D_pG(x,Dv^{\varepsilon,\eta})\cdot D(w^\eta - v^{\varepsilon,\eta}) - \eta^2\Delta(w^\eta - v^{\varepsilon,\eta}).\end{aligned}$$

Multiplying the preceding inequality by a solution, $\theta^{\varepsilon,\eta}$, of $(\mathrm{AJ})_\varepsilon^\eta$, integrating on $\mathbb{T}^n$, and using integration by parts, we get

$$\begin{aligned}&\int_{\mathbb{T}^n} (C\eta + \varepsilon f_r(x,0)v^{\varepsilon,\eta} + o(\varepsilon))\theta^{\varepsilon,\eta}\,dx\\ &\qquad\geq -\int_{\mathbb{T}^n}\big(\operatorname{div}(D_pG\theta^{\varepsilon,\eta}) + \eta^2\Delta\theta^{\varepsilon,\eta}\big)(w^\eta - v^{\varepsilon,\eta})\,dx\\ &\qquad= \varepsilon\int_{\mathbb{T}^n}(\delta_{x_0} - f_r(x,0)\theta^{\varepsilon,\eta})(w^\eta - v^{\varepsilon,\eta})\,dx.\end{aligned}$$

Next, we rearrange the previous estimate and get

$$v^{\varepsilon,\eta}(x_0) \geq w^\eta(x_0) - \iint_{\mathbb{T}^n\times\mathbb{R}^n} f_r(x,0)w^\eta\, d\nu^{\varepsilon,\eta}(x,p) - \frac{C\eta+o(\varepsilon)}{\varepsilon}\iint_{\mathbb{T}^n\times\mathbb{R}^n} d\nu^{\varepsilon,\eta}(x,p).$$

Finally, we set $\eta=\eta_k\to 0$. Thus, $\nu^{\varepsilon,\eta_k} \rightharpoonup \nu^\varepsilon$ as measures. Taking the limit in the preceding inequality ends the proof. □

3.3. Convergence.

Here, we prove the selection theorem for (GE$_\varepsilon$), Theorem 2.1. This theorem substantially extends the existing results for convex Hamiltonians and is the key technical device in the study of quasi-convex and double-well Hamiltonians.

Proof of Theorem 2.1. Let $\{\varepsilon_j\}_{j\in\mathbb{N}}$ be any subsequence converging to 0 such that v^{ε_j} converges uniformly to a solution of (GE) as $j\to\infty$. In view of Lemma 3.4 and the definition of v^0, we have that

$$v^0 \geq \lim_{j\to\infty} v^{\varepsilon_j}. \tag{3.10}$$

Moreover, by Lemma 3.5, we get

$$\lim_{j\to\infty} v^{\varepsilon_j}(x) \geq w(x) - \iint_{\mathbb{T}^n\times\mathbb{R}^n} f_r(x,0)w(x)\, d\nu(x,p)$$

for any subsolution, w, of (GE). In particular, we take the supremum of all $w\in\mathcal{E}$ in the above inequality to get that

$$\lim_{j\to\infty} v^{\varepsilon_j}(x) \geq \sup_{w\in\mathcal{E}} w(x) = v^0(x). \tag{3.11}$$

Thus, we combine (3.10) and (3.11) to get the desired result. □

4. Strictly quasi-convex Hamiltonians.

Now, we use the results in the preceding section to investigate the selection problem for strictly quasi-convex Hamilton–Jacobi equations and to prove Theorem 2.2. For convenience, we assume that $\overline{H}=0$ in (E).

Lemma 4.1. *A function, $u^\varepsilon\in C(\mathbb{T}^n)$, solves* (D$_\varepsilon$) *if and only if u^ε solves* (GE$_\varepsilon$) *for $f(r)=-e^{-\lambda_0 r}$, $r\in\mathbb{R}$, and G as in Assumption* (A5).

Proof. Clearly, u^ε is a subsolution of (D$_\varepsilon$) if and only if for any $x\in\mathbb{T}^n$ and any $p\in D^+u^\varepsilon(x)$, we have

$$\varepsilon u^\varepsilon(x) + H(x,p) \leq 0. \tag{4.1}$$

Moreover, (4.1) holds if any only if

$$-e^{-\varepsilon u^\varepsilon(x)} + G(x,p) \leq 0 \tag{4.2}$$

for any $p\in D^+u^\epsilon$ and $x\in\mathbb{T}^n$. Arguing in a similar way for the supersolution case gives the result. □

Theorem 2.2 is an immediate corollary of Theorem 2.1. The proof of Theorem 2.2 follows.

Proof of Theorem 2.2. By the preceding Lemma, u^ε solves (GE_ε). It is clear that f, G satisfy (A1)–(A3). Thus, we apply Theorem 2.1 to obtain the last statement of the theorem. □

5. One-dimensional, non-convex, double-well Hamiltonians.

For each $P \in \mathbb{R}$, we consider the discounted Hamilton–Jacobi equation

$$\varepsilon u^\varepsilon + (|P + u^\varepsilon_x|^2 - 1)^2 - V(x) = 0 \quad \text{in } \mathbb{T}. \tag{5.1}$$

As before, $\lim_{\varepsilon\to 0} \varepsilon u^\varepsilon = -\overline{H}(P) \in \mathbb{R}$, where $\overline{H}(P)$ is the unique constant for which

$$(|P + u_x|^2 - 1)^2 - V(x) = \overline{H}(P) \quad \text{in } \mathbb{T}$$

has a viscosity solution.

Assumption (2.4) means that V has a small oscillation; that is, $\mathrm{osc}\,(V) := \max_{\mathbb{T}} V - \min_{\mathbb{T}} V < 1$. Because the wells of $(|p|^2 - 1)^2$ have depth 1, which is larger than $\mathrm{osc}\,(V)$, the effect of V on $\overline{H}(P)$ is localized. Moreover, from the results in [**2**], the graph of $\overline{H}(P)$ follows and is shown in Figure 5.1. As suggested by Figure 5.1, to prove Theorem 2.3, we separately consider different cases according to the region where P lies. We have the following a priori estimates that are essential in the proof of Theorem 2.3.

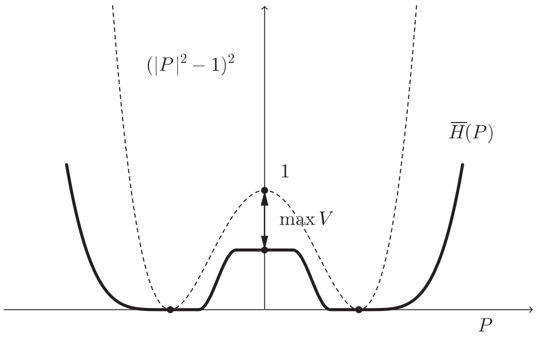


Figure 5.1. The shape of $\overline{H}(P)$.

Proposition 5.1. *Let u^ε solve* (5.1). *Consider the following three cases*:

$$\text{(a) } |P| < 1,\ \overline{H}(P) > 0, \quad \text{(b) } |P| > 1,\ \overline{H}(P) > 0, \quad \text{(c) } \overline{H}(P) = 0. \tag{5.2}$$

Then, in the viscosity sense, we have,

(i) *In case* (a), *for $\varepsilon > 0$ sufficiently small, $|P + u^\varepsilon_x| \le 1$ in $\mathbb{T}$.*

(ii) *In case* (b), *if $P > 1$ then $P + u^\varepsilon_x \ge 1$ in $\mathbb{T}$ for sufficiently small $\varepsilon > 0$. If $P < -1$, then $P + u^\varepsilon_x \le -1$ in $\mathbb{T}$ for sufficiently small $\varepsilon > 0$.*

(iii) *In case* (c)*, if* $P>0$*, then* $P+u^\varepsilon_x \geq 0$ *in* $\mathbb{T}$ *for sufficiently small* $\varepsilon>0$*. If* $P<0$*, then* $P+u^\varepsilon_x \leq 0$ *in* $\mathbb{T}$ *for sufficiently small* $\varepsilon>0$*.*

PROOF. First, we extend u^ε periodically to $\mathbb{R}$.

To prove (i), we argue by contradiction. Suppose that there exists $z\in\mathbb{R}$ such that u^ε is differentiable at z and $|P+u^\varepsilon_x(z)|>1$. Define $f:\mathbb{R}\to\mathbb{R}$ by

$$f(x)=u^\varepsilon(x)+Px+|x-z|.$$

Because $|P|<1$, there exists $x_0\in\mathbb{R}$ such that $f(x_0)=\min_{x\in\mathbb{R}} f(x)$ (see Figure 5.2, case (a)). Next, we prove that $x_0\neq z$. Indeed, by setting $q=P+u^\varepsilon_x(z)$, we obtain

$$\begin{aligned} f(z-\alpha q) &= P(z-\alpha q)+u^\varepsilon(z-\alpha q)+\alpha|q| \\ &= P(z-\alpha q)+u^\varepsilon(z)-\alpha u^\varepsilon_x(z)q+o(\alpha)+\alpha|q| \\ &= f(z)-\alpha\left(|q|(|q|-1)+\frac{o(\alpha)}{\alpha}\right), \end{aligned}$$

which implies $f(z-\alpha q)<f(z)$ for a small $\alpha>0$ since $|q|>1$. Thus, $x_0\neq z$.

Because u^ε is a viscosity supersolution of (5.1) at $x=x_0$, we have

$$V(x_0)\leq \varepsilon u^\varepsilon(x_0)+\left(\left|P-\left(P+\frac{x_0-z}{|x_0-z|}\right)\right|^2-1\right)^2=\varepsilon u^\varepsilon(x_0).$$

Since $\varepsilon u^\varepsilon(x_0)\to -\overline{H}(P)<0$, the above inequality yields $V(x_0)<0$. This contradicts (2.4). Thus, (i) holds.

Next, we prove (ii). We consider only the case when $P>1$ and argue by contradiction. If $P<-1$, the argument is analogous. Suppose that there exists $z\in\mathbb{R}$ such that u^ε is differentiable at z and $P+u^\varepsilon_x(z)<1$. Define the function $f:\mathbb{R}\to\mathbb{R}$ by

$$f(x):=u^\varepsilon(x)+Px-|x-z|.$$

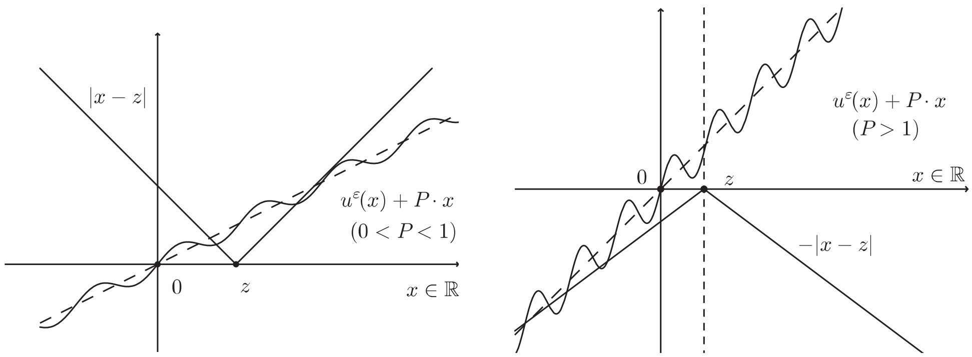
Because $P>1$, there exists $x_0\in[z,\infty)$ such that $f(x_0)=\min_{x\in[z,\infty)} f(x)$ (see Figure 5.2, case (b)). First, we prove that $x_0>z$. We start by setting $q=P+u^\varepsilon_x(z)$. For $\alpha>0$, we have

$$\begin{aligned} f(z+\alpha) &= P(z+\alpha)+u^\varepsilon(z+\alpha)-\alpha \\ &= P(z+\alpha)+u^\varepsilon(z)+\alpha u^\varepsilon_x(z)+o(\alpha)-\alpha \\ &= f(z)+\alpha\left(q-1+\frac{o(\alpha)}{\alpha}\right). \end{aligned}$$

Because $q<1$, the preceding identity implies that $f(z+\alpha)<f(z)$ for a small $\alpha>0$. Hence, $x_0>z$. Consequently, by the argument of the last part of the proof of (i), we get a contradiction.

Finally, we prove (iii). We consider only the case when $P>0$ and prove that $P+u^\varepsilon_x\geq 0$ in $\mathbb{T}$. The case when $P<0$ is analogous.

The proof proceeds by contradiction. First, if $P+u^\varepsilon_x\leq 0$ for almost everywhere



Case (a). Case (b).

Figure 5.2.

$x \in \mathbb{T}$, then

$$0 = u^\varepsilon(1) - u^\varepsilon(0) = \int_0^1 u_x^\varepsilon(x)\,dx \le \int_0^1 (-P)\,dx = -P < 0,$$

which is a contradiction. Therefore, we need to consider only the case when there exists $x_1, x_2 \in \mathbb{R}$ with $x_1 \neq x_2$ such that u^ε is differentiable at x_1, x_2 and

$$P + u_x^\varepsilon(x_1) > 0 > P + u_x^\varepsilon(x_2).$$

We can assume that $x_1 < x_2$ without loss of generality. Otherwise, by periodicity, we replace x_2 by $x_2 + k$ for some large enough $k \in \mathbb{N}$.

In view of [**3**, Lemma 2.6], there exists $x_3 \in (x_1, x_2)$ such that $0 \in P + D^+u^\varepsilon(x_3)$. By the definition of the viscosity subsolution, we have

$$\varepsilon u^\varepsilon(x_3) + |0^2 - 1|^2 - V(x_3) = \varepsilon u^\varepsilon(x_3) + 1 - V(x_3) \le 0,$$

which is a contradiction for sufficiently small $\varepsilon > 0$ as $\lim_{\varepsilon\to 0} \varepsilon u^\varepsilon(x_3) = 0$ and $\max V < 1$. $\square$

Remark 3. By inspecting the proof of case (a) of the preceding proposition, we see that the argument extends to arbitrary dimensions. In contrast, the proofs of the other two cases are one dimensional in nature, and we do not know how to generalize them for higher dimensions.

Finally, we present the proof of Theorem 2.3.

Proof of Theorem 2.3. First, we use Proposition 5.1 to transform (5.1) into (GE_ε). Then, we proceed as follows. We set $v^\varepsilon := u^\varepsilon + \overline{H}(P)/\varepsilon$. Thus, the ergodic constant becomes 0.

In case (a), we use (i) of Proposition 5.1 to rewrite (5.1) as

$$-\sqrt{V(x)+\overline{H}(P)-\varepsilon v^\varepsilon}-|P+v^\varepsilon_x|^2+1=0 \quad \text{in } \mathbb{T}.$$

Next, we set $G(x,p):=-|P+p|^2+1$. Then, G is concave (not convex) in p. We set $w^\varepsilon=-v^\varepsilon$ to have that w^ε is the viscosity solution to

$$\sqrt{V(x)+\overline{H}(P)+\varepsilon w^\varepsilon}+|P-w^\varepsilon_x|^2-1=0 \quad \text{in } \mathbb{T}.$$

This new equation is convex. Thus, the convergence of w^ε (hence v^ε) is straightforward by Theorem 2.1.

In case (b), we consider only the case when $P>1$ as the case when $P<-1$ is similar. In light of (ii) in Proposition 5.1, we rewrite (5.1) as

$$-\sqrt{V(x)+\overline{H}(P)-\varepsilon v^\varepsilon}+|P+v^\varepsilon_x|^2-1=0 \quad \text{in } \mathbb{T}.$$

Here, a direct application of Theorem 2.1 implies the convergence of v^ε as $\varepsilon\to 0$.

Finally, in case (c), we consider only the case when $P>0$. Because $P+u^\varepsilon_x\geq 0$ in $\mathbb{T}$, only the positive branch ($p\geq 0$) of the graph of $(|p|^2-1)^2$ plays a role here. Note that this branch is quasi-convex and satisfies (A4). Therefore, Theorem 2.2 gives the convergence of u^ε as $\varepsilon\to 0$. □

5.1. A further generalization in one dimension.

The argument in the proof of Theorem 2.3 can be adapted to handle the case when the oscillation of the potential energy, V, is smaller than the depth of any well of the kinetic energy, H. Thus, we have the convergence of the discounted approximation. We can generalize this idea as follows. Consider a Hamiltonian of the form

$$H(x,p)=F(p)-V(x),$$

where $F(p)$ is the kinetic energy and $V(x)$ is the potential energy. Assume that $-\infty=p_0<p_1<p_2<\cdots<p_{2L+1}<p_{2L+2}=+\infty$ exists for some $L\in\mathbb{N}$ such that

- $\lim_{|p|\to\infty}F(p)=+\infty$,
- $F'(p_i)=0$ and $F''(p_i)\neq 0$ for $1\leq i\leq 2L+1$,
- $F'(p)>0$ for $p\in(p_{2i+1},p_{2i+2})$ for $0\leq 1\leq L$,
- $F'(p)<0$ for $p\in(p_{2i},p_{2i+1})$ for $0\leq 1\leq L$.

Set

$$m=\min_{1\leq i\leq 2L}|F(p_i)-F(p_{i+1})|.$$

Assume that

$$\operatorname{osc}(V)<m.$$

Under these assumptions, we can prove that the solution u^ε of $(\mathrm{D})_\varepsilon$ converges to a solution of (E) for any $P\in\mathbb{R}$, which generalizes Theorem 2.3.

REMARK 4. If the oscillation of V is larger than m, we cannot localize the convergence argument. The qualitative behavior of $\overline{H}$ was examined in [**3**]. However, the characterization of the convergence of the discounted approximation remains an open problem. In this setting, for some values of P, we see that due to the non-convex nature of the gradient jumps, the problem cannot be transformed into an equation of the form (GE$_\varepsilon$).

6. An example: A quasi-convex Hamiltonian with flat parts.

In this last section, we study a selection problem for which the results in Sections 3 and 4 do not apply. We consider a continuous, piecewise C^1, quasi-convex Hamiltonian that has a level set with a flat part. Thus, (A4) does not hold and, therefore, we need an alternative approach.

Assume that $n = 1$. For $p \geq 0$, let

$$F(p) = \begin{cases} p & \text{for } 0 \leq p \leq 1, \\ 1 & \text{for } 1 \leq p \leq 2, \\ p - 1 & \text{for } p \geq 2. \end{cases}$$

Consider the Hamiltonian

$$H(x, p) = F(|p|) - V(x), \tag{6.1}$$

where V is as follows. First, we select a sufficiently small $s > 0$. Then, we set

$$V(x) = \begin{cases} x & \text{for } 0 \leq x \leq s, \\ 2s - x & \text{for } s \leq x \leq 2s, \\ 0 & \text{for } 2s \leq x \leq 1. \end{cases} \tag{6.2}$$

To study the effect of the flat part of F, we fix $P = 3/2$. For this value of P, we examine the maximal subsolutions of (2.5) and the discounted problem.

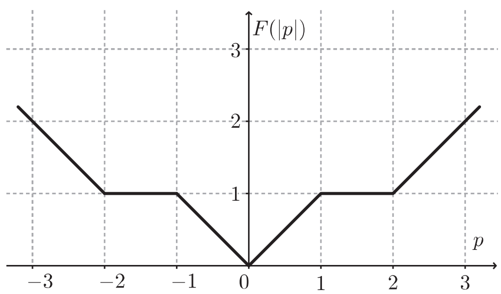


Image of $F(|p|)$.

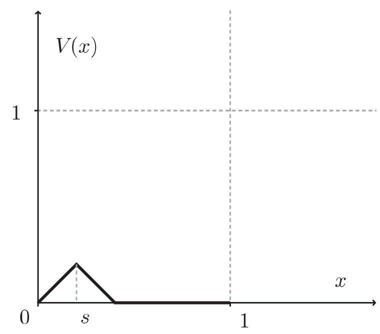


Image of $V(x)$.

6.1. Maximal subsolutions and the Aubry set at $P = 3/2$.

First, we compute the effective Hamiltonian at 3/2; that is, the unique value $\overline{H}(P)$ for which (2.5) has a viscosity solution.

Lemma 6.1. *Assume that* (6.1) *and* (6.2) *hold. Then,* $\overline{H}(3/2) = 1$.

Proof. Choose a function $v \in C(\mathbb{T})$ such that

$$\begin{cases} v_x = \dfrac{1}{2} + V & \text{in } (0, 2s), \\ v_x \in \left[-\dfrac{1}{2}, \dfrac{1}{2}\right] & \text{in } (2s, 1), \\ \displaystyle\int_0^1 v_x \, dx = 0. \end{cases}$$

Clearly, v is a viscosity solution to

$$F\left(\left|\frac{3}{2} + v_x\right|\right) - V(x) = 1 \quad \text{in } \mathbb{T}. \tag{6.3}$$

Thus, $\overline{H}(3/2) = 1$. □

Next, we define the corresponding maximal subsolutions. First, we fix a vertex $y \in \mathbb{T}$ and set

$$S(x, y) = \sup\{w(x) - w(y) \,:\, w \text{ is a subsolution of (6.3)}\}. \tag{6.4}$$

Clearly, $S(y, y) = 0$, $x \mapsto S(x, y)$ is a subsolution of (6.3) in the whole torus, $\mathbb{T}$, and it is a solution of (6.3) in $\mathbb{T} \setminus \{y\}$. Because $S(\cdot, y)$ is the largest subsolution w of (6.3) satisfying $w(y) = 0$, we call it the maximal subsolution with vertex y.

In the conference "New connections between dynamical systems and PDEs" at the American Institute of Mathematics in 2003, Sergey Bolotin posed the following question (see [**6**], question 12 in the list of open problems):

Question 1. Does there exist $y \in \mathbb{T}$ such that $x \mapsto S(x, y)$ is a solution of (6.3) in $\mathbb{T}$?

The answer to the preceding question was found to be yes if H is strictly quasiconvex (see [**14**]). For the general nonconvex case, this question has remained open. Here, we answer no to this question (see also [**13**, Example 12.7]). More precisely, we offer the following proposition.

Proposition 6.2. *For all* $y \in \mathbb{T}$, $S(\,\cdot\,, y)$ *is not a solution of* (6.3) *in* $\mathbb{T}$.

Proof. Fix $y \in \mathbb{T}$. Let $w : \mathbb{T} \to \mathbb{R}$ be a function such that $w(y) = 0$ and

$$w_x(x) = \begin{cases} -\dfrac{7}{2} & \text{for } x \in \left(y - \dfrac{1}{8}, y\right), \\ \dfrac{1}{2} & \text{for } x \in \left(y, y + \dfrac{7}{8}\right). \end{cases} \tag{6.5}$$

It is straightforward that w is a subsolution of (6.3) in the almost everywhere sense. Hence, it is a viscosity subsolution. Therefore, $S(x, y) \geq w(y)$ for all $x \in \mathbb{T}$. In particular, this implies that

$$\left[-\frac{7}{2}, \frac{1}{2}\right] \subset D^- S(y, y). \tag{6.6}$$

Next, we select $q = -3/2 \in D^- S(y, y)$ and notice that

$$F\left(\left|\frac{3}{2} + q\right|\right) - V(y) \leq F(0) = 0 < 1.$$

Consequently, $S(\,\cdot\,, y)$ is not a supersolution of (6.3) at y. □

We observe that the maximal subsolution, $S(x, y)$, can be computed explicitly. However, we do not need this computation here and, thus, omit it.

Remark 5. We recall that we can define the Aubry set for strictly quasiconvex Hamilton–Jacobi equations as the set of all points, y, such that $S(\,\cdot\,, y)$ is a solution on $\mathbb{T}^n$ (see [**14**] for the details). Proposition 6.2 implies that if we define the Aubry set in the same way, it is empty. However, this does not contradict the results in [**14**] as the Hamiltonian of the example in this section violates an assumption of the strictly quasiconvexity. This fact indeed highlights a significant difference between convex and non-convex cases. Therefore, if an analog of the Aubry set exists, it has to be defined in a different way. In the general non-convex case, we can construct Mather measures [**7**] using the nonlinear adjoint method. When the Hamiltonian is strictly quasiconvex, these measures are invariant under the Hamiltonian flow. Moreover, the Mather measures are supported in a subset of the Aubry set called the Mather set. This, of course, cannot hold if the Aubry set is empty. Besides, in the general non-convex case, Mather measures may not be invariant under the Hamiltonian flow, and the loss of invariance is encoded in dissipation measures that record the gradient jump structure [**7**].

6.2. Discounted approximation at $P = 3/2$.

Finally, we consider the discounted approximation problem for $P = 3/2$.

$$\varepsilon u^\varepsilon + F\left(\left|\frac{3}{2} + u_x^\varepsilon\right|\right) - V(x) = 0 \quad \text{in } \mathbb{T}. \tag{6.7}$$

Proposition 6.3. *There exists a solution of* (6.3), $u^0 \in C(\mathbb{T})$, *such that*

$$\lim_{\varepsilon \to 0}\left(u^\varepsilon + \frac{1}{\varepsilon}\right) = u^0 \quad \textit{in } C(\mathbb{T}).$$

Proof. Let $v^\varepsilon = u^\varepsilon + 1/\varepsilon$. Then, v^ε solves

$$\varepsilon v^\varepsilon + F\left(\left|\frac{3}{2} + v_x^\varepsilon\right|\right) = 1 + V(x) \quad \text{in } \mathbb{T}. \tag{6.8}$$

Next, we give an explicit construction for v^ε.

Step 1. Set

$$v^\varepsilon(x) = e^{-\varepsilon x} \int_0^x e^{\varepsilon r}\left(\frac{1}{2} + V(r)\right) dr \quad \text{for } x \in (0, a^\varepsilon),$$

where a^ε is a number to be chosen such that $a^\varepsilon \in (s, 2s)$ and $v_x^\varepsilon(a^\varepsilon-) = 1/2$.

It is clear that

– $v^\varepsilon(0) = 0$ and $v_x^\varepsilon(0+) = 1/2$.

– $v^\varepsilon(x) \le x$ and

$$v_x^\varepsilon(x) = \frac{1}{2} + V(x) - \varepsilon v^\varepsilon(x).$$

– In particular, for $0 < x < s$, we have $v_x^\varepsilon(x) \ge 1/2$ and thus

$$\varepsilon v^\varepsilon(x) + F\left(\left|\frac{3}{2} + v_x^\varepsilon\right|\right) = \varepsilon v^\varepsilon(x) + \left(\frac{1}{2} + v_x^\varepsilon\right) = 1 + V(x). \tag{6.9}$$

– v^ε is increasing and always $\varepsilon v^\varepsilon = O(\varepsilon)$. We choose $a^\varepsilon \in (s, 2s)$ such that $\varepsilon v^\varepsilon(a^\varepsilon) = V(a^\varepsilon)$. Then, $\lim_{\varepsilon\to 0} a^\varepsilon = 2s$ and $v_x^\varepsilon(a^\varepsilon-) = 1/2$. Clearly, (6.9) holds for all $x \in (0, a^\varepsilon)$.

Step 2. Define

$$v^\varepsilon(x) = e^{\varepsilon(a^\varepsilon - x)} v^\varepsilon(a^\varepsilon) + e^{-\varepsilon x} \int_{a^\varepsilon}^x e^{\varepsilon r}\left(-\frac{1}{2} + V(r)\right) dr \quad \text{for } x \in (a^\varepsilon, b^\varepsilon),$$

where $b^\varepsilon > 2s$ is a number to be chosen later. We have that

– $v_x^\varepsilon(a^\varepsilon+) = -1/2$.

– v^ε is decreasing in $(a^\varepsilon, b^\varepsilon)$, and

$$v_x^\varepsilon(x) = -\frac{1}{2} + V(x) - \varepsilon v^\varepsilon(x).$$

– We argue that, for $x \in (a^\varepsilon, 2s)$, we have $\varepsilon v^\varepsilon(x) \ge V(x)$. This is correct as $\varepsilon v^\varepsilon(a^\varepsilon) = V(a^\varepsilon)$ and $\varepsilon v_x^\varepsilon(x) \ge -1 = V'(x)$ in $(a^\varepsilon, 2s)$. Thus, $v_x^\varepsilon(x) \le -1/2$ in $(a^\varepsilon, 2s)$ and also $v^\varepsilon(2s) > 0$.

– Pick $b^\varepsilon > 2s$ to be the smallest number such that $v^\varepsilon(b^\varepsilon) = 0$. Then, $v_x^\varepsilon(b^\varepsilon-) = -1/2$, and, for $x \in (a^\varepsilon, b^\varepsilon)$, we always have $v_x^\varepsilon(x) \le -1/2$ and

$$\varepsilon v^\varepsilon(x) + F\left(\left|\frac{3}{2} + v^\varepsilon_x\right|\right) = \varepsilon v^\varepsilon(x) + \left(\frac{3}{2} + v^\varepsilon_x\right) = 1 + V(x).$$

STEP 3. For $x \in (b^\varepsilon, 1)$, we set $v^\varepsilon(x) = 0$. As $V = 0$ in $(b^\varepsilon, 1)$, we have that, for $x \in (b^\varepsilon, 1)$,

$$\varepsilon v^\varepsilon(x) + F\left(\left|\frac{3}{2} + v^\varepsilon_x\right|\right) = F\left(\frac{3}{2}\right) = 1 = 1 + V(x).$$

From the preceding three steps of the construction, v^ε is 1-periodic. To check that v^ε solves (6.8), we only need to check the definition of viscosity solutions at the points where there are gradient jumps. These points are $x = 0, a^\varepsilon, b^\varepsilon$. As we have $F(p) = 1$ for $p \in [1, 2]$, the verification at these three points is obvious.

Now, we are concerned with the convergence of v^ε as $\varepsilon \to 0$. As discussed in Step 1, we have that $\lim_{\varepsilon\to 0} a^\varepsilon = 2s$. Set $b = \lim_{\varepsilon\to 0} b^\varepsilon$. We use the explicit formula of v^ε to get that $v^\varepsilon \to u^0$, uniformly in $\mathbb{T}$, where u^0 satisfies

$$\begin{cases} u^0(0) = 0, & \\ (u^0)'(x) = \dfrac{1}{2} + V(x) & \text{in } (0, 2s), \\ (u^0)'(x) = -\dfrac{1}{2} & \text{in } (2s, b), \\ u^0 \equiv 0 & \text{on } [b, 1]. \end{cases}$$

Finally, we see that u^0 solves (6.3). □

References

[1] E. S. Al-Aidarous, E. O. Alzahrani, H. Ishii and A. M. M. Younas, A convergence result for the ergodic problem for Hamilton–Jacobi equations with Neumann type boundary conditions, Proc. Royal Soc. Edinburgh Sect. A, **146** (2016), 225–242.

[2] S. Armstrong, H. V. Tran and Y. Yu, Stochastic homogenization of a nonconvex Hamilton–Jacobi equation, Calc. Var. Partial Differential Equations, **54** (2015), 1507–1524.

[3] S. Armstrong, H. V. Tran and Y. Yu, Stochastic homogenization of nonconvex Hamilton–Jacobi equations in one space dimension, J. Differential Equations, **261** (2016), 2702–2737.

[4] M. Bardi and I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 1997. With appendices by M. Falcone and P. Soravia.

[5] U. Bessi, Aubry–Mather theory and Hamilton–Jacobi equations, Comm. Math. Phys., **235** (2003), 495–511.

[6] S. Bolotin, List of open problems, http://www.aimath.org/WWN/dynpde/articles/html/20a/.

[7] F. Cagnetti, D. Gomes and H. V. Tran, Aubry–Mather measures in the non convex setting, SIAM J. Math. Anal., **43** (2011), 2601–2629.

[8] F. Cagnetti, D. Gomes, H. Mitake and H. V. Tran, A new method for large time behavior of degenerate viscous Hamilton–Jacobi equations with convex Hamiltonians, Ann. Inst. H. Poincaré Anal. Non Linéaire, **32** (2015), 183–200.

[9] A. Davini, A. Fathi, R. Iturriaga and M. Zavidovique, Convergence of the solutions of the discounted Hamilton–Jacobi equation, Invent. Math., **206** (2016), 29–55.

[10] L. C. Evans, Adjoint and compensated compactness methods for Hamilton–Jacobi PDE, Arch. Rat. Mech. Anal., **197** (2010), 1053–1088.

[11] A. Fathi, Weak KAM Theorem in Lagrangian Dynamics, Cambridge Univ. PR(Txp), 2016.

[12] A. Fathi, Sur la convergence du semi-groupe de Lax-Oleinik, C. R. Acad. Sci. Paris Sér. I Math., **327** (1998), 267–270.

[13] A. Fathi, Weak KAM from a PDE point of view: viscosity solutions of the Hamilton–Jacobi equation and Aubry set, Proc. Roy. Soc. Edinburgh Sect. A, **142** (2012), 1193–1236.

[14] A. Fathi and A. Siconolfi, PDE aspects of Aubry–Mather theory for quasiconvex Hamiltonians, Calc. Var. Partial Differential Equations, **22** (2005), 185–228.

[15] H. Gao, Random homogenization of coercive Hamilton–Jacobi equations in 1d, Calc. Var. Partial Differential Equations, **55** (2016), Art. 30, 39pp.

[16] D. A. Gomes, Generalized Mather problem and selection principles for viscosity solutions and Mather measures, Adv. Calc. Var., **1** (2008), 291–307.

[17] H. Ishii, H. Mitake and H. V. Tran, The vanishing discount problem and viscosity Mather measures, Part 1: the problem on a torus, J. Math. Pures Appl. (9), **108** (2017), 125–149.

[18] H. Ishii, H. Mitake and H. V. Tran, The vanishing discount problem and viscosity Mather measures, Part 2: boundary value problems, J. Math. Pures Appl. (9), **108** (2017), 261–305.

[19] H. R. Jauslin, H. O. Kreiss and J. Moser, On the forced Burgers equation with periodic boundary conditions, Proc. Sympos. Pure Math., **65**, Amer. Math. Soc., Province, RI, 1999.

[20] P.-L. Lions, G. Papanicolaou and S. R. S. Varadhan, Homogenization of Hamilton–Jacobi equations, unpublished work (1987).

[21] R. Mañé, Generic properties and problems of minimizing measures of Lagrangian systems, Nonlinearity, **9** (1996), 273–310.

[22] J. N. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, Math. Z., **207** (1991), 169–207.

[23] H. Mitake and H. V. Tran, Selection problems for a discount degenerate viscous Hamilton–Jacobi equation, Adv. Math., **306** (2017), 684–703.

[24] K. Soga, Selection problems of $\mathbb{Z}^2$-periodic entropy solutions and viscosity solutions, Calc. Var. Partial Differential Equations, **56** (2017), Art 119.

[25] H. V. Tran, Adjoint methods for static Hamilton–Jacobi equations, Calc. Var. Partial Differential Equations, **41** (2011), 301–319.

Diogo A. Gomes
King Abdullah University of Science
and Technology (KAUST)
CEMSE Division
Thuwal 23955-6900, Saudi Arabia
E-mail: diogo.gomes@kaust.edu.sa

Hiroyoshi Mitake
Institute of Engineering
Division of Electrical, Systems
and Mathematical Engineering
Hiroshima University
1-4-1 Kagamiyama
Higashi-Hiroshima-shi
Hiroshima 739-8527, Japan
E-mail: hiroyoshi-mitake@hiroshima-u.ac.jp

Hung V. Tran
Department of Mathematics
University of Wisconsin-Madison
Van Vleck hall
480 Lincoln drive, Madison
WI 53706, USA
E-mail: hung@math.wisc.edu