PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 145, Number 1, January 2017, Pages 201–214
http://dx.doi.org/10.1090/proc/13193
Article electronically published on June 30, 2016

PERTURBATION ESTIMATES OF WEAK KAM SOLUTIONS AND MINIMAL INVARIANT SETS FOR NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

QINBO CHEN AND MIN ZHOU

(Communicated by Yingfei Yi)

Abstract. For nearly integrable and Tonelli system

$$H_\epsilon = H_0(p) + \epsilon H_1(q,p,t). \quad (q,p,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T},$$

we give the perturbation estimates of weak KAM solution u_ϵ with respect to parameter ϵ and prove the stability of the Mather set $\tilde{\mathcal{M}}_\epsilon$, Aubry set $\tilde{\mathcal{A}}_\epsilon$, Mañé set $\tilde{\mathcal{N}}_\epsilon$ and even the backward (forward) calibrated curves under the perturbation.

1. Introduction

We denote by $\mathbb{T}^n \times \mathbb{R}^n$ the cotangent bundle $T^*\mathbb{T}^n$, that we endow with its usual coordinates (q,p) and its canonical symplectic form $\Omega = \sum_{i=1}^n dq_i \bigwedge dp_i$, where $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

Given $r \geq 2$ and a non-decreasing dominant function $C(x) : \mathbb{Z}^+ \to \mathbb{R}^+$, let

$$\mathcal{S} = \{f \in C^r(\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}, \mathbb{R}) : \|f(q,p,t)\|_{C^r} \leq C(K), \text{ for all } \|p\| \leq K\}.$$

We consider the following C^r nearly integrable Hamiltonian:

$$H_\epsilon(q,p,t) = H_0(p) + \epsilon H_1(q,p,t), \quad H^1 \in \mathcal{S}, \quad (q,p,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}, \tag{1.1}$$

where the integrable system $H_0(p)$ is strictly convex and superlinear, and we also assume H_ϵ satisfies the following Tonelli conditions:

(L1) Convexity: For each $(q,t) \in \mathbb{T}^n \times \mathbb{T}$, the Hamiltonian H_ϵ is strictly convex in p coordinate, i.e., the Hessian $\frac{\partial^2 H_\epsilon}{\partial p_i \partial p_j}$ is definitely positive.

(L2) Superlinearity:

$$\lim_{\|p\| \to +\infty} \frac{H_\epsilon(q,p,t)}{\|p\|} = +\infty, \quad \text{uniformly on } (q,t).$$

(L3) Completeness: All solutions of the Hamiltonian equation are well defined for the whole $t \in \mathbb{R}$.

Received by the editors December 7, 2015 and, in revised form, March 7, 2016.

2010 *Mathematics Subject Classification.* Primary 37Jxx, 70Hxx.

Key words and phrases. Weak KAM solutions, perturbation estimates, Mather theory, minimal invariant sets.

The authors were supported by the National Basic Research Program of China (973 Program) (Grant No. 2013CB834100), the National Natural Science Foundation of China (Grant No. 11171146, Grant No. 11201222) and a program PAPD of Jiangsu Province, China.

We also denote by $\mathbb{T}^n\times\mathbb{R}^n$ the tangent bundle $T\mathbb{T}^n$, and we obtain the associated C^r Lagrangian

$$L_\epsilon(q,v,t)=\langle v,\pi_p\circ\mathcal{L}^{-1}(q,v,t)\rangle-H_\epsilon\circ\mathcal{L}^{-1}(q,v,t). \tag{1.2}$$

Here $\mathcal{L}:\mathbb{T}^n\times\mathbb{R}^n\times\mathbb{T}\to\mathbb{T}^n\times\mathbb{R}^n\times\mathbb{T},\mathcal{L}(q,p,t)=(q,\frac{\partial H_\epsilon(q,p,t)}{\partial p},t)$ is the Legendre-Fenchel transformation, and π_p denotes the natural projection $\pi_p(q,p,t)=p$. Thus, L_ϵ also satisfies Tonelli conditions:

(L1) Convexity: For each $(q,t)\in\mathbb{T}^n\times\mathbb{T}$, L_ϵ is strictly convex in v coordinate, i.e., the Hessian $\frac{\partial^2 L_\epsilon}{\partial v_i\partial v_j}$ is definitely positive.

(L2) Superlinearity:

$$\lim_{\|v\|\to+\infty}\frac{L_\epsilon(q,v,t)}{\|v\|}=+\infty,\quad\text{uniformly on }(q,t).$$

(L3) Completeness: All solutions of the Euler Lagrange equation are well defined for the whole $t\in\mathbb{R}$.

Because $H^1(\mathbb{T}^n,\mathbb{R})\cong\mathbb{R}^n$, from now on, unless otherwise specified, we use the same symbol $c=(c_1,\cdots,c_n)\in\mathbb{R}^n$ to denote both the cohomology class in $H^1(\mathbb{T}^n,\mathbb{R})$ and the closed 1-form $\sum_{i=1}^n c_i dq_i$ of torus T^n.

Let's review some basic facts for nearly integrable systems. Classical KAM theory asserts that a set of nearly full measure in phase space consists of invariant tori carrying quasi-periodic motions ([2], [16], [21]). In addition, the Nekhoroshev estimates tell us that all solutions stay stable for an exponentially long time under some steepness conditions (e.g. [22]). However, for the whole time, the phenomenon of instability may occur, such as Arnold diffusion (e.g. [8], [9], [15]).

The perturbation estimates and regularity of weak KAM solutions (see Section 2) in the normally hyperbolic invariant cylinders are very important in construction of diffusion orbits (e.g. [3], [10], [24]) and propagation of singularities [7]. It was proved by [12] that if Mather's α function $\alpha(c)$ is twice differentiable at c_0, then

$$\int_{\mathbb{T}^n}\|(c+d_xu^c)-(c_0+d_xu^{c_0})\|^2d\sigma\le C\|c-c_0\|^2$$

for $\|c-c_0\|\ll 1$, where σ is the projection on $\mathbb{T}^n$ of some Mather measure μ supported on the Mather set $\tilde{\mathcal{M}}(c_0)$. Moreover, it was shown by [17] that if $\tilde{\mathcal{M}}(c_0)$ is a real analytic quasi-periodic invariant torus with a Diophantine frequency, then

$$\|(c+d_xu^c)-(c_0+d_xu^{c_0})\|\le C\|c-c_0\|$$

for $\|c-c_0\|\ll 1$. We also refer the readers to ([1], [14]) for the ϵ-regularity of weak KAM solutions. Notice that all of these results ([1], [12], [14], [17]) were established for time-independent Hamiltonians. In Theorem 1.1, we give the ϵ-regularity of weak KAM solutions for time-dependent nearly integrable Hamiltonians. Furthermore, Example 1.3 illustrates that Theorem 1.1 may not be true for general Hamiltonians, even when the Hamiltonian is time-independent.

For nearly integrable Hamiltonians, [6] provides an estimate on the speed of minimal orbits by using the globally topological trick. For quasi-integrable exact maps, [4] provides the same estimate on the speed of minimal orbits in an invariant set with homoclinic orbits for each resonant frequency. However, in Theorem 1.2, combining Mather's variational theory and Fathi's weak KAM theory, we also give similar estimates to those in [4] and [6]. Furthermore, we give the perturbation estimates of globally minimal invariant sets such as the Mather sets, Aubry sets and Mañé sets. Theorem 1.2 also gives the perturbation size of all backward (resp. forward) calibrated curves. This can be viewed as the complement to KAM theory and Nekhoroshev estimates.

Theorem 1.1. *Given $c \in H^1(\mathbb{T}^n, \mathbb{R})$, there exists a small number $\epsilon_0 = \epsilon_0(H_0, c) > 0$ and a constant $D = D(H_0, c) > 0$ such that for all ϵ with $|\epsilon| < \epsilon_0$, Hamiltonian systems* (1.1) *and its associated Lagrangian* (1.2) *have the following estimates:*

for each weak KAM solution u^c_ϵ of $L_\epsilon - c$ and u^c_0 of $L_0 - c$, we have

$$\|u^c_\epsilon(x,t) - u^c_\epsilon(y,s)\| \le D\sqrt{\epsilon}(\|x-y\| + |s-t|)$$

and

$$\|du^c_\epsilon(q,t) - du^c_0(q,t)\| \le D\sqrt{\epsilon}, \quad \textit{for almost all } (q,t) \in \mathbb{T}^n \times \mathbb{T}.$$

Theorem 1.2. *Given $c \in H^1(\mathbb{T}^n, \mathbb{R})$, there exists a small number $\epsilon_0 = \epsilon_0(H_0, c) > 0$ and a constant $D = D(H_0, c) > 0$ such that for all ϵ with $|\epsilon| < \epsilon_0$, we have*

(1) *For each curve $\gamma_\epsilon(t) : (-\infty, +\infty) \to \mathbb{T}^n$ which is calibrated by a weak KAM solution u^c_ϵ of Lagrangian* (1.2)*, i.e., for all $t_1 < t_2 \in \mathbb{R}$ such that $u^c_\epsilon(\gamma_\epsilon(t_2), t_2) - u^c_\epsilon(\gamma_\epsilon(t_1), t_1) = \int_{t_1}^{t_2} (L_\epsilon - c + \alpha_\epsilon(c))(d\gamma_\epsilon(s), s)ds$, we have*

$$\|\dot\gamma_\epsilon(t) - \dot\gamma_\epsilon(0)\| \le D\sqrt{\epsilon}, \quad \forall t \in \mathbb{R}.$$

(2) *For each backward curve $\gamma_\epsilon(t) : (-\infty, t_0] \to \mathbb{T}^n$ which is calibrated by a weak KAM solution u^c_ϵ of Lagrangian* (1.2)*, i.e., for all $t_1 < t_2 \in (-\infty, t_0)$ such that $u^c_\epsilon(\gamma_\epsilon(t_2), t_2) - u^c_\epsilon(\gamma_\epsilon(t_1), t_1) = \int_{t_1}^{t_2} (L_\epsilon - c + \alpha_\epsilon(c))(d\gamma_\epsilon(s), s)ds$, we have*

$$\|\dot\gamma_\epsilon(t) - \dot\gamma_\epsilon(0)\| \le D\sqrt{\epsilon}, \quad \forall t \le t_0.$$

(3) *For each minimal orbit $\gamma_\epsilon(t)$ in the Mather set $\mathcal{M}_\epsilon(c)$ $\big($resp. Aubry set $\mathcal{A}_\epsilon(c)$, Mañé set $\mathcal{N}_\epsilon(c)\big)$, we have*

$$\|\dot\gamma_\epsilon(t) - \dot\gamma_\epsilon(0)\| \le D\sqrt{\epsilon}, \quad \forall t \in \mathbb{R}.$$

In addition, the Mather set $\tilde{\mathcal{M}}_\epsilon(c)$ $\big($resp. $\tilde{\mathcal{A}}_\epsilon(c)$, $\tilde{\mathcal{N}}_\epsilon(c)\big)$ of Lagrangian (1.2) *is contained in a $D\sqrt{\epsilon}$ neighbourhood of the Mather set $\tilde{\mathcal{M}}_0(c)$ $\big($resp. $\tilde{\mathcal{A}}_\epsilon(c)$, $\tilde{\mathcal{N}}_\epsilon(c)\big)$ of L_0, i.e., $d_H(\tilde{\mathcal{M}}_\epsilon(c), \tilde{\mathcal{M}}_0(c)) \le D\sqrt{\epsilon}$*

$$(\textit{resp.} \quad d_H(\tilde{\mathcal{A}}_\epsilon(c), \tilde{\mathcal{A}}_0(c)) \le D\sqrt{\epsilon}, \quad d_H(\tilde{\mathcal{N}}_\epsilon(c), \tilde{\mathcal{N}}_0(c)) \le D\sqrt{\epsilon}\ ),$$

where $d_H(A, B)$ denotes the Hausdorff distance, i.e.,

$$d_H(A, B) = \max\{\sup_{a\in A} \inf_{b\in B} d(a,b), \sup_{b\in B} \inf_{a\in A} d(a,b)\}.$$

The following example shows that Theorem 1.1 may not be true for general Hamiltonians.

Example 1.3. We take $H_\epsilon(q,p) = H_0(q,p) + \epsilon H_1(q,p)$, $(q,p) \in \mathbb{T} \times \mathbb{R}$, where $H_0(q,p) = \frac{1}{2}p^2 + \delta\big(\cos(2\pi q) - 1\big)$ and $H_1(q,p) = p$, we also let $\epsilon \ll \delta$. Notice that H_ϵ is not nearly integrable. It's easy to compute the associated Lagrangian of H_ϵ:

$$L_\epsilon(q,v) = L_0(q,v) + \epsilon L_1(q,v,\epsilon) = \frac{1}{2}v^2 - \delta\big(\cos(2\pi q) - 1\big) - \epsilon(v - \frac{\epsilon}{2}).$$

(1) $\epsilon = 0$, the weak KAM solution $u_0(q)$ is a 1-periodic function and satisfies:

$$d_q u_0 = \begin{cases} 2\sqrt{\delta}\sin(\pi q), & 0 \le q < \frac{1}{2}, \\ -2\sqrt{\delta}\sin(\pi q), & \frac{1}{2} < q \le 1. \end{cases}$$

(2) $0<\epsilon<\frac{4}{\pi}\sqrt{\delta}$, the weak KAM solution $u_\epsilon(q)$ is 1-periodic and satisfies:

$$d_q u_\epsilon = \begin{cases} 2\sqrt{\delta}\sin(\pi q)-\epsilon, & 0\le q<a(\epsilon), \\ -2\sqrt{\delta}\sin(\pi q)-\epsilon, & a(\epsilon)<q\le 1, \end{cases}$$

where $a(\epsilon)$ is determined by the equation $\cos(\pi a(\epsilon)) = -\epsilon/(\frac{4\sqrt{\delta}}{\pi})$ and $a(\epsilon)\ge\frac{1}{2}$.

We have,

$$\begin{aligned} \|d_q u_\epsilon - d_q u_0\| &= \max_{q\in[0,1]} |d_q u_\epsilon - d_q u_0| \ge \max_{\frac{1}{2}\le q<a(\epsilon)} |d_q u_\epsilon - d_q u_0| \\ &\ge |2\sqrt{\delta}\sin(\pi q)-\epsilon-(-2\sqrt{\delta}\sin(\pi q))| \\ &= 4\sqrt{\delta}\sqrt{1-\epsilon^2/(\frac{4\sqrt{\delta}}{\pi})^2} - \epsilon = \sqrt{16\delta-\pi^2\epsilon^2}-\epsilon. \end{aligned}$$

Then for $0<\epsilon<\frac{3\sqrt{\delta}}{\pi}$, $\|d_q u_\epsilon - d_q u_0\| \ge 7\sqrt{\delta}-\frac{3\sqrt{\delta}}{\pi} > \sqrt{\delta}$.

In view of $\epsilon \ll \delta$ and that δ is fixed, so $\|d_q u_\epsilon - d_q u_0\| \le O(\epsilon^\kappa)$, $(0<\kappa<1)$ cannot hold. Therefore, Theorem 1.1 is not true in this case.

2. Brief introduction to Mather theory and weak KAM theory

Let's review some basic results of Mather theory ([19], [20]) first. Let M be a compact connected C^∞ manifold, and TM be its tangent bundle. Let $L: TM\times\mathbb{T}\to\mathbb{R}$ be a $C^r (r\ge 2)$ Tonelli Lagrangian, where $\mathbb{T}=\mathbb{R}/\mathbb{Z}$.

The Tonelli conditions imply that the Legendre-Fenchel transformation $\mathcal{L}$ is a C^{r-1} diffeomorphism of $TM\times\mathbb{T}$ onto $T^*M\times\mathbb{T}$,

$$\mathcal{L}(q,v,t) = (q, \frac{\partial L}{\partial v}(q,v,t), t).$$

Therefore, we obtain the associated Hamiltonian $H(q,p,t) = \langle p,v\rangle - L(q,v,t)$, where $v=v(q,p,t)$ is implicitly determined by $p=\frac{\partial L}{\partial v}(q,v,t)$.

Let $I=[a,b]$ be an interval, and $\gamma: I\to M$ be an absolutely continuous curve. We denote by

$$A(\gamma) = \int_a^b L(d\gamma(t),t)dt$$

the action of γ. An absolute curve $\gamma: I\to M$ is called a minimizer or action minimizing curve if

$$A(\gamma) = \min_{\substack{\xi(a)=\gamma(a),\xi(b)=\gamma(b)\\ \xi\in C^{ac}(I,M)}} \int_a^b L(d\xi(t),t)dt.$$

We call $\gamma: (-\infty,+\infty)\to M$ a globally minimizing curve if for all $a<b, \gamma$ is a minimizer on $[a,b]$. Notice that the minimizer satisfies the Euler Lagrange equation.

Let $\mathcal{M}_L$ be the space of Euler Lagrangian flow invariant probability measures on $TM\times\mathbb{T}$. To each $\mu\in\mathcal{M}_L$, note that $\int\lambda d\mu$=0 for each exact 1-form λ. Therefore, given $c\in H^1(M,\mathbb{R})$ and a closed 1-form $\eta_c\in c=[\eta_c]$, we can define Mather's α function

$$\alpha(c) = -\inf_{\mu\in\mathcal{M}_L} A_c(\mu) = -\inf_{\mu\in\mathcal{M}_L}\int_{TM\times\mathbb{T}} L-\eta_c d\mu.$$

It's easy to check that $\alpha(c)$ is finite everywhere, convex and superlinear.

We associate to $\mu \in \mathcal{M}_L$ its rotation vector $\rho(\mu) \in H_1(M,\mathbb{R})$ in the following sense:

$$\langle \rho(\mu), [\eta_c] \rangle = \int_{TM\times\mathbb{T}} \eta_c d\mu, \quad \forall c \in H^1(M,\mathbb{R}).$$

So we can define Mather's β function:

$$\beta(h) = \inf_{\mu\in\mathcal{M}_L, \rho(\mu)=h} \int L d\mu.$$

β is finite, convex, and superlinear and β is the Legendre-Fenchel dual of α .

Let $\mathcal{M}^c = \{\mu \in \mathcal{M}_L | A_c(\mu) = -\alpha(c)\}$, $\mathcal{M}_h = \{\mu \in \mathcal{M}_L | \rho(\mu) = h, A(\mu) = \beta(h)\}$. $\mu \in \mathcal{M}_L$ is called a $c-$minimal measure if $\mu \in \mathcal{M}^c$ and we can define the Mather set:

$$\tilde{\mathcal{M}}(c) = \bigcup_{\mu\in\mathcal{M}^c} \text{supp}\mu.$$

To study more properties of dynamic systems, we need to find "larger" invariant sets and study their topology structure. First, we define a function Φ_c,

$$\Phi_c : (M \times \mathbb{T}) \times (M \times \mathbb{T}) \to \mathbb{R},$$

$$((x,\tau),(x',\tau')) \mapsto \inf_{\substack{t'>t,\ t\equiv\tau \bmod 1\\ t'\equiv\tau' \bmod 1,\ \gamma\in\Gamma}} \int_t^{t'} (L - \eta_c + \alpha(c))(d\gamma(s), s)ds,$$

where Γ is a set of absolutely continuous curves γ satisfying $\gamma(t) = x, \gamma(t') = x'$, and η_c is a closed 1-form such that $[\eta_c] = c \in H^1(M,\mathbb{R})$. A curve $\gamma : \mathbb{R} \to M$ is called c-semi-static if

$$A_c(\gamma|[t,t']) = \Phi_c\big((\gamma(t), t \bmod 1),\ (\gamma(t'), t' \bmod 1)\big).$$

A curve $\gamma : \mathbb{R} \to M$ is called c-static if

$$A_c(\gamma|[t,t']) = -\Phi_c\big((\gamma(t'), t' \bmod 1),\ (\gamma(t), t \bmod 1)\big).$$

Thus, we define the Aubry set $\tilde{\mathcal{A}}(c)$ and the Mañé set $\tilde{\mathcal{N}}(c)$ in $TM \times \mathbb{T}$ as

$$\tilde{\mathcal{A}}(c) = \bigcup\{(d\gamma(t), t \bmod 1) \mid \gamma \text{ is } c\text{-static}\},$$
$$\tilde{\mathcal{N}}(c) = \bigcup\{(d\gamma(t), t \bmod 1) \mid \gamma \text{ is } c\text{-semi-static}\}.$$

Then, we have the following relation:

$$\tilde{\mathcal{M}}(c) \subseteq \tilde{\mathcal{A}}(c) \subseteq \tilde{\mathcal{N}}(c).$$

Let $\pi : TM \times \mathbb{T} \to M \times \mathbb{T}$ be the natural projection. We denote

$$\mathcal{M}(c) \triangleq \pi \circ \tilde{\mathcal{M}}(c),\ \mathcal{A}(c) \triangleq \pi \circ \tilde{\mathcal{A}}(c),\ \mathcal{N}(c) \triangleq \pi \circ \tilde{\mathcal{N}}(c).$$

Now, we will review some basic results of weak KAM theory. For the autonomous case, see [13]. For the non-autonomous case, see e.g. [5], [11], [23].

Definition 2.1. We say $u^- : M \times \mathbb{T} \to \mathbb{R}$ is a backward weak KAM solution if

(1) u^- is dominated by $L + \alpha(0)$, i.e.,

$$u^-(x,s) - u^-(y,t) \le \Phi_0((y,t),(x,s)).$$

(2) For each $(x,s) \in M \times \mathbb{T}$, there exists a calibrated curve $\gamma : (-\infty, s] \to M$ such that

$$u^-(x,s) - u^-(\gamma(t),t) = A(\gamma|_{[t,s]}) + \alpha(0)(s-t), \quad \forall t \in (-\infty, s].$$

Similarly, we can also define the forward weak KAM solution u^+.

The Lax-Oleinik semigroup is well known in PDE and in Calculus of Variations. Now we will introduce the associated Lax-Oleinik operator on $C^0(M \times \mathbb{T}, \mathbb{R})$ for non-autonomous and time 1-periodic Lagrangian:

$$\begin{aligned} T^-_{\eta_c,n} &: C^0(M \times \mathbb{T}, \mathbb{R}) \to C^0(M \times \mathbb{T}, \mathbb{R}), \\ T^-_{\eta_c,n} u(x,t) &= \inf_{\substack{\gamma \in C^{ac} \\ \gamma(t)=x}} \big(u(\gamma(t-n), t-n) + \int_{t-n}^{t} (L - \eta_c + \alpha(c))(d\gamma(s), s) ds \big), \end{aligned}$$

where $n \in \mathbb{N}$. The sequence $\{T^-_{\eta_c,n}\}_{n\in\mathbb{N}}$ constitutes a semigroup.

Proposition 2.2 ([13], [23]). *There exist backward weak KAM solutions corresponding to the Lagrangian $L - \eta_c$. Let v_c^- be any backward weak KAM solution; then*

(1) $T^-_{\eta_c,n} v_c^- = v_c^-$, *for all n. In addition, v_c^- is Lipschitz and it is a viscosity solution of the following Hamilton-Jacobi equation:*

$$\partial_t f + H(q, d_q f + \eta_c, t) = \alpha(c).$$

(2) *If the curve $\gamma : (-\infty, s] \to M$ is calibrated by v_c^-, then γ is c-semi-static and v_c^- is differentiable at $(\gamma(t), t)$ for all $t \in (-\infty, s)$, i.e.,*

$$d_q v_c^-(\gamma(t), t) = \frac{\partial(L - \eta_c)}{\partial v}(d\gamma(t), t),$$

$$d_t v_c^-(\gamma(t), t) = -H(q, \frac{\partial(L - \eta_c)}{\partial v}(d\gamma(t), t) + \eta_c, t) + \alpha(c).$$

3. Proof of the main theorems

In this section, we turn to the proof of Theorem 1.1 and Theorem 1.2. We need the following lemmas. In this section, for all $K \in \mathbb{Z}^+$, we define the following set by using Legendre-Fenchel transformation $\mathcal{L}$:

$$\begin{aligned} \mathcal{D}_\epsilon(K) &= \mathcal{L}\{(q,p,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T} : \|p\| \le K\} \\ &= \{(q,v,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T} : \exists\, p, \|p\| \le K \text{ such that } v = \partial_p H_\epsilon(q,p,t)\}. \end{aligned}$$

Lemma 3.1. *The associated Lagrangian L_ϵ (with respect to H_ϵ) has the form $L_\epsilon = L_0(v) + \epsilon L_1(q,v,t,\epsilon)$ and*

$$\|L_1(q,v,t,\epsilon)\|_{C^0} \le C(K), \quad \forall (q,v,t) \in \mathcal{D}_\epsilon(K),$$

where $L_0(v)$ is the associated Lagrangian of $H_0(p)$, and $C(x) : \mathbb{Z}^+ \to \mathbb{R}^+$ is the non-decreasing dominant function in (1.1).

Proof. Because $H_0(p)$ is convex in p, we have

$$\begin{aligned} H_0(p) &\ge H_0(\partial_v L_0(v)) - \langle \partial_p H_0(\partial_v L_0(v)), p - \partial_v L_0(v) \rangle \\ &= H_0(\partial_v L_0(v)) - \langle v, p - \partial_v L_0(v) \rangle. \end{aligned}$$

Then for $(q, v, t) \in \mathcal{D}_\epsilon(K)$,

$$\begin{aligned}
L_\epsilon(q, v, t) &= \sup_{\|p\| \leq K} \{\langle p, v\rangle - H_0(p) - \epsilon H_1(q, p, t)\} \\
&\leq \sup_{\|p\| \leq K} \{\langle p, v\rangle - H_0(\partial_v L_0(v)) - \langle v, p - \partial_v L_0(v)\rangle - \epsilon H_1\} \\
&= \sup_{\|p\| \leq K} \{\langle \partial_v L_0(v), v\rangle - H_0(\partial_v L_0(v)) - \epsilon H_1\} \\
&= \sup_{\|p\| \leq K} \{L_0(v) - \epsilon H_1\} \leq L_0(v) + \epsilon C(K).
\end{aligned}$$

In addition,

$$\begin{aligned}
L_\epsilon(q, v, t) &= \sup_p \{\langle p, v\rangle - H_0(p) - \epsilon H_1\} \\
&\geq \langle \partial_v L_0(v), v\rangle - H_0(\partial_v L_0(v)) - \epsilon H_1(q, \partial_v L_0(v), t) \\
&= L_0(v) - \epsilon H_1(q, \partial_v L_0(v), t) \geq L_0(v) - \epsilon C(K).
\end{aligned}$$

This completes the proof. □

Lemma 3.2. *There exist two constants $R_0 = R_0(H_0) > 0$, $K_0 = K_0(H_0) > 0$ large enough and a small constant $\epsilon_1 = \epsilon_1(H_0) > 0$, such that $\{(q, v, t) : \|v\| \leq R_0\} \subseteq \mathcal{D}_\epsilon(K_0)$ and*

(1) *For all $v \in \mathbb{R}^n$ with $\|v\| \geq \frac{R_0}{2}$, we have $L_0(v) \geq A+2$, where $A = \max_{\|v\| \leq 2\sqrt{n}} L_0(v)$.*

(2) *For all $\epsilon, |\epsilon| < \epsilon_1$, each minimizing curve $\gamma(t)$ of $L_\epsilon : [t_1, t_2] \to \mathbb{T}^n, t_2 - t_1 \geq 1$ satisfies*

$$\|\dot{\gamma}(t)\| \leq R_0 - 1, \quad \forall t \in [t_1, t_2].$$

(3) *For all $\epsilon, |\epsilon| < \epsilon_1$, the weak KAM solution u_ϵ of L_ϵ satisfies:*

$$|u_\epsilon(x, t) - u_\epsilon(y, t)| \leq K_0 \|x - y\|.$$

Proof. (1) is a straight consequence of the property of superlinearity.

For the proof of (2), it's enough to show this lemma for $t_2 - t_1 = 1$. Indeed, for general $[t_1, t_2]$ and $t \in [t_1, t_2]$, we can find an interval of the form $[c, c+1]$, with $t \in [c, c+1] \subseteq [t_1, t_2]$.

For simplicity, we assume $[t_1, t_2] = [0, 1]$. Obviously, we can find a geodesic segment connecting $\gamma(0)$ to $\gamma(1)$ in T^n, and parameterize it by the time interval [0,1] with speed of constant norm. We denote by $\eta(t) : [0, 1] \to \mathbb{T}^n, \eta(0) = \gamma(0), \eta(1) = \gamma(1), \|\dot{\eta}(t)\| = d(\gamma(0), \gamma(1)) \leq diam(\mathbb{T}^n) = diam(\mathbb{R}^n/\mathbb{Z}^n) = \sqrt{n}$.

Take a constant $R_0 > 0$ large enough, so $(\eta(s), \dot{\eta}(s), s) \in \{(q, v, t) : \|v\| \leq \sqrt{n}\} \subseteq \{(q, v, t) : \|v\| \leq R_0\}$. There exists a constant K_0 such that $\{(q, v, t) : \|v\| \leq R_0\} \subseteq \mathcal{D}_\epsilon(K_0)$; notice that R_0, K_0 only depend on H_0.

Take a small number $\epsilon_1 > 0$ satisfying $\epsilon_1 C(K_0) < 1$. By Lemma 3.1, for all $|\epsilon| < \epsilon_1$, we have

$$\int_0^1 L_\epsilon(\gamma(s), \dot{\gamma}(s), s) ds \leq \int_0^1 L_\epsilon(\eta(s), \dot{\eta}(s), s) ds \leq \int_0^1 A + \epsilon C(K_0) ds < A + 1,$$

where $A = \max_{\|v\| \leq 2\sqrt{n}} L_0(v)$. Hence, there exists $\xi \in [0, 1]$ such that

$$L_\epsilon(\gamma(\xi), \dot{\gamma}(\xi), \xi) < A + 1. \tag{3.1}$$

Next, we claim that

$$L_\epsilon(\gamma(\xi), v, \xi) \geq A + 1 \quad \text{for all} \quad v, \|v\| \geq R_0. \tag{3.2}$$

We prove by contradiction. Indeed, suppose there exists $(\gamma(\xi), v_1, \xi), \|v_1\| \geq R_0$ such that

$$L_\epsilon(\gamma(\xi), v_1, \xi) < A + 1. \tag{3.3}$$

Since we already know that for all $(\gamma(\xi), v_2, \xi), \|v_2\| \leq \sqrt{n} < \frac{R_0}{2}$,

$$L_\epsilon(\gamma(\xi), v_2, \xi) \leq A + \epsilon C(K_0) < A + 1. \tag{3.4}$$

Moreover, Lemma 3.2 (1) implies that for all $(\gamma(\xi), v_3, \xi), \frac{R_0}{2} \leq \|v_3\| < R_0$,

$$L_\epsilon(\gamma(\xi), v_3, \xi) \geq A + 2 - \epsilon C(K_0) > A + 1. \tag{3.5}$$

Therefore, (3.3), (3.4) and (3.5) lead to a contradiction to the strict convexity of $L_\epsilon(\gamma(\xi), \dots, \xi)$.

By (3.1), (3.2) and (3.5), we obtain

$$\|\dot\gamma(\xi)\| \leq \frac{R_0}{2}.$$

Denote by Φ_ϵ^T the time T map of Lagrangian flow. Since R_0 is very large and ϵ_1 is very small, it's not hard to check that $\|\Phi_\epsilon^T(\gamma(\xi), \dot\gamma(\xi), \xi)\| \leq \frac{2}{3}R_0 \leq R_0 - 1,\ \forall\ |T| \leq 1$, which completes (2).

Finally, we turn to the proof of (3). Taking a differentiable point (x, t) of u_ϵ, there exists a backward minimizing curve $\gamma(s) : (-\infty, t] \to \mathbb{T}^n$ such that $\gamma(t) = x$ and $d_q u_\epsilon(x, t) = \partial_v L_\epsilon(\gamma(t), \dot\gamma(t), t)$ by Proposition 2.2. By Lemma 3.2 (2), we obtain

$$\|\dot\gamma(t)\| \leq R_0, \quad \|d_q u_\epsilon(x, t)\| = \|\partial_v L_\epsilon(\gamma(t), \dot\gamma(t), t)\| \leq K_0.$$

Because u_ϵ is Lipschitz and differentiable almost everywhere, we have

$$|u_\epsilon(x, t) - u_\epsilon(y, t)| \leq K_0\|x - y\|.$$

□

Lemma 3.3. *Let ϵ_1, K_0 be the constants as in Lemma 3.2. Then, for all ϵ such that $|\epsilon| < \epsilon_1$, we have*

$$|\alpha_\epsilon(0) - \alpha_0(0)| \leq \epsilon C(K_0),$$

where $\alpha_\epsilon(\cdot)$ (resp. $\alpha_0(\cdot)$) is the α-function with respect to L_ϵ (resp. L_0).

Proof. It's not hard to obtain that $\alpha_0(c) = H_0(c)$. By the Legendre-Fenchel inequality, $L_\epsilon(q, v, t) + H_\epsilon(q, 0, t) \geq 0$, so

$$L_\epsilon(q, v, t) \geq -H_\epsilon(q, 0, t) \geq -\max_{q\in\mathbb{T}^n, t\in\mathbb{T}} H_\epsilon(q, 0, t).$$

It follows

$$\begin{aligned} -\alpha_\epsilon(0) &= \inf_\mu \int_{\mathbb{T}^n\times\mathbb{R}^n\times\mathbb{T}} L_\epsilon d\mu \geq \inf_\mu \int_{\mathbb{T}^n\times\mathbb{R}^n\times\mathbb{T}} -\max_{q\in\mathbb{T}^n, t\in\mathbb{T}} H_\epsilon(q, 0, t) d\mu \\ &= -\max_{q\in\mathbb{T}^n, t\in\mathbb{T}} H_\epsilon(q, 0, t), \\ \alpha_\epsilon(0) &\leq \max_{q\in\mathbb{T}^n, t\in\mathbb{T}} H_\epsilon(q, 0, t) \leq H_0(0) + \epsilon C(K_0) = \alpha_0(0) + \epsilon C(K_0). \end{aligned}$$

Let $v_0 = \partial_p H_0(0)$ and take a closed curve $\gamma(t) = (t, 0, 0, \dots, 0) \subseteq \mathbb{T}^1 \times \mathbb{T}^{n-1}$; then $\tilde\gamma(t) = (\gamma(t), v_0, t)$ gives a closed curve on $\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}$. This introduces a probability measure $\mu_{\tilde\gamma}$ on $\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}$ defined by,

$$\int_{\mathbb{T}^n\times\mathbb{R}^n\times\mathbb{T}} \Psi \mathrm{d}\mu_{\tilde\gamma} = \int_0^1 \Psi(\gamma(t), v_0, t)\mathrm{d}t,$$

for all continuous function Ψ with compact support. Thus, by an equivalent definition of α function (see Section 1 of [18]), we have

$$-\alpha_\epsilon(0) \le \int L_\epsilon \mathrm{d}\mu_{\tilde{\gamma}} = L_0(v_0) + \int_0^1 \epsilon L_1(\gamma(t), v_0, t, \epsilon)\mathrm{d}t$$
$$= -H_0(0) + \epsilon \int_0^1 L_1(\gamma(t), v_0, t, \epsilon)\mathrm{d}t \le -\alpha_0(0) + \epsilon C(K_0).$$

The last inequality follows from Lemma 3.1. Thus, $\alpha_\epsilon(0) \ge \alpha_0(0) - \epsilon C(K_0)$. This leads to our conclusion. □

For brevity, given $f \in C^r(\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T})$, we denote by $\frac{\partial^2 f}{\partial p^2}$ the Hessian matrix $\big(\frac{\partial^2 f}{\partial p_i \partial p_j}\big)_{n\times n}$, by $\frac{\partial^2 f}{\partial p \partial q}$ the Hessian matrix $\big(\frac{\partial^2 f}{\partial p_i \partial q_j}\big)_{n\times n}$ and by $\frac{\partial^2 f}{\partial q^2}$ the Hessian matrix $\big(\frac{\partial^2 f}{\partial q_i \partial q_j}\big)_{n\times n}$.

Lemma 3.4. *There exists a constant $\epsilon_2 = \epsilon_2(H_0) > 0$ such that for all $\epsilon, |\epsilon| < \epsilon_2$, we have the following estimates: There exists a constant $\lambda_0 = \lambda_0(H_0) \in (0,1)$ such that*

$$\lambda_0 Id \le \frac{\partial^2 H_\epsilon(p)}{\partial p^2} \le \frac{1}{\lambda_0} Id,\ \|\frac{\partial^2 H_\epsilon}{\partial p \partial q}\|_{C^0} \le \epsilon C(K_0),\ \|\frac{\partial^2 H_\epsilon}{\partial q^2}\|_{C^0} \le \epsilon C(K_0), \tag{3.6}$$

for all $(q,p,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}, \|p\| \le K_0$.

$$\lambda_0 Id \le \frac{\partial^2 L_\epsilon(v)}{\partial v^2} \le \frac{1}{\lambda_0} Id,\ \|\frac{\partial^2 L_\epsilon}{\partial v \partial q}\|_{C^0} \le \frac{\epsilon C(K_0)}{\lambda_0},\ \|\frac{\partial^2 L_\epsilon}{\partial q^2}\|_{C^0} \le \frac{2\epsilon C(K_0)}{\lambda_0}, \tag{3.7}$$

for all $(q,v,t) \in \mathcal{D}_\epsilon(K_0)$.

Proof. Obviously, there exists a constant $\lambda \in (0,1)$ such that

$$\lambda Id \le \frac{\partial^2 H_0(p)}{\partial p^2} \le \frac{1}{\lambda} Id, \quad \forall (q,p,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T},\ \|p\| \le K_0.$$

Moreover, we have

$$(\lambda - \epsilon C(K_0)) Id \le \frac{\partial^2 H_\epsilon(p)}{\partial p^2} \le (\frac{1}{\lambda} + \epsilon C(K_0)) Id,$$
$$\|\frac{\partial^2 H_\epsilon}{\partial p \partial q}\|_{C^0} \le \epsilon C(K_0), \quad \|\frac{\partial^2 H_\epsilon}{\partial q^2}\|_{C^0} \le \epsilon C(K_0),$$

for all $(q,p,t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}, \|p\| \le K_0$. Since $C(K_0)$ is fixed, we choose ϵ_2 sufficiently small and $\lambda_0 \in (0,1)$, such that for all $\epsilon, |\epsilon| < \epsilon_2$,

$$\lambda_0 \le \lambda - \epsilon C(K_0), \quad \frac{1}{\lambda} + \epsilon C(K_0) \le \frac{1}{\lambda_0}.$$

Notice that λ_0 and ϵ_2 only depend on the function H_0. Therefore, (3.6) holds.

On the other hand, it's not hard to obtain the following:

$$\frac{\partial^2 L_\epsilon}{\partial v^2} = (\frac{\partial^2 H_\epsilon}{\partial p^2})^{-1},\ \frac{\partial^2 L_\epsilon}{\partial q \partial v} = -\frac{\partial^2 H_\epsilon}{\partial p \partial q}\frac{\partial^2 L_\epsilon}{\partial v^2},\ \frac{\partial^2 L_\epsilon}{\partial q^2} = -\frac{\partial^2 H_\epsilon}{\partial p \partial q}\frac{\partial^2 L_\epsilon}{\partial v \partial q} - \frac{\partial^2 H_\epsilon}{\partial q^2}.$$

Thus we have

$$\lambda_0 Id \le \frac{\partial^2 L_\epsilon}{\partial v^2} \le \frac{1}{\lambda_0} Id, \quad \| \frac{\partial^2 L_\epsilon}{\partial q \partial v} \|_{C^{r-2}} \le \frac{\epsilon C(K_0)}{\lambda_0},$$
$$\| \frac{\partial^2 L_\epsilon}{\partial q^2} \|_{C^{r-2}} \le \frac{(\epsilon C(K_0))^2}{\lambda_0} + \frac{\epsilon C(K_0)}{\lambda_0} \le \frac{2\epsilon C(K_0)}{\lambda_0},$$

for all $(q,v,t)\in\mathcal{D}_\epsilon(K_0)$. □

Now, we begin to prove Theorem 1.1.

Proof of Theorem 1.1. For simplicity, we just prove our theorem for $c=0\in H^1(\mathbb{T}^n,\mathbb{R})$. Take $\epsilon_0=\min\{\epsilon_1,\epsilon_2\}$, where ϵ_1,ϵ_2 are the constants in Lemma 3.2 and Lemma 3.4.

Let u_ϵ denote the weak KAM solution of L_ϵ.

Step 1. First we claim that there exists $C_0=C_0(H_0)>0$ such that

$$|u_\epsilon(x+\Delta x,t)-u_\epsilon(x,t)|\le C_0\sqrt{\epsilon}\|\Delta x\|. \tag{3.8}$$

Since $u_\epsilon(x,t)$ can be viewed as a $\mathbb{Z}^{n+1}$-periodic function in $\mathbb{R}^n\times\mathbb{R}$, it's enough to prove

$$|u_\epsilon(x+\Delta x,0)-u_\epsilon(x,0)|\le C_0\sqrt{\epsilon}\|\Delta x\| \quad \text{for all} \quad x\in\mathbb{R}^n,\ \|\Delta x\|\le 1,$$

where $\Delta x=(\Delta x_1,\cdots,\Delta x_n)$. For general t, it can be proved similarly. From now on, we use the symbol $[x]$ to represent the integer part of x and take

$$N=[\frac{1}{\sqrt{\epsilon}}].$$

By Proposition 2.2 (1), there exists a C^2 minimizer $\gamma_0(s):[0,N]\to\mathbb{R}^n$ with $\gamma_0(N)=x$, such that

$$u_\epsilon(x,0)=u_\epsilon(x,N)=u_\epsilon(\gamma_0(0),0)+\int_0^N L_\epsilon(\gamma_0(s),\dot\gamma_0(s),s)+\alpha_\epsilon(0)ds.$$

Let $\eta(s)\triangleq\gamma_0(s)+s\frac{\Delta x}{N},\eta(N)=x+\Delta x$. So

$$u_\epsilon(x+\Delta x,0)\le u_\epsilon(\eta(0),0)+\int_0^N L_\epsilon(\eta(s),\dot\eta(s),s)+\alpha_\epsilon(0)ds$$

and

$$u_\epsilon(x+\Delta x,0)-u_\epsilon(x,0)\le\int_0^N L_\epsilon(\eta(s),\dot\eta(s),s)-L_\epsilon(\gamma_0(s),\dot\gamma_0(s),s)ds. \tag{3.9}$$

Fix s and use the integral form of the Taylor formula. We have

$$\begin{aligned}L_\epsilon(\eta(s),\dot\eta(s),s)=L_\epsilon(\gamma_0(s),\dot\gamma_0(s),s)&+\langle\frac{\partial L_\epsilon}{\partial q}(\gamma_0(s),\dot\gamma_0(s),s),\frac{\Delta x}{N}s\rangle\\&+\langle\frac{\partial L_\epsilon}{\partial v}(\gamma_0(s),\dot\gamma_0(s),s),\frac{\Delta x}{N}\rangle+\mathcal{R}(s),\end{aligned}$$

where

$$\mathcal{R}(s)=\int_0^1(1-t)(\frac{\Delta x}{N}s,\frac{\Delta x}{N})M(t)(\frac{\Delta x}{N}s,\frac{\Delta x}{N})^T dt$$

$$\text{and } M(t)=\frac{\partial^2 L_\epsilon}{\partial q\partial v}(t\eta(s)+(1-t)\gamma_0(s),t\dot\eta(s)+(1-t)\dot\gamma_0(s),s).$$

Because γ_0 is a minimizer, by Lemma 3.2 and $\|\Delta x\|\le 1$, we have

$$\|\dot\gamma_0(s)\|\le R_0-1,\quad \|\dot\eta(s)\|\le R_0,\quad \forall s\in[0,N].$$

Setting $v(t)=(t\eta(s)+(1-t)\gamma_0(s),\ t\dot\eta(s)+(1-t)\dot\gamma_0(s),\ s)$, we have

$$v(t)\in\{(q,v,t)\in\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}:\|v\|\le R_0\}. \tag{3.10}$$

Moreover,

$$\begin{aligned}(\frac{\Delta x}{N}s,\frac{\Delta x}{N})M(t)(\frac{\Delta x}{N}s,\frac{\Delta x}{N})^T =& \sum_{i,j=1}^{n}\frac{\partial^2 L_\epsilon}{\partial q_i\partial q_j}(v(t))(\frac{\Delta x_i}{N}s)(\frac{\Delta x_j}{N}s)\\ &+2\sum_{i,j=1}^{n}\frac{\partial^2 L_\epsilon}{\partial q_i\partial v_j}(v(t))(\frac{\Delta x_i}{N}s)(\frac{\Delta x_j}{N})\\ &+\sum_{i,j=1}^{n}\frac{\partial^2 L_\epsilon}{\partial v_i\partial v_j}(v(t))(\frac{\Delta x_i}{N})(\frac{\Delta x_j}{N}).\end{aligned}$$

Thus by (3.10), Lemma 3.2 and Lemma 3.4, we obtain

$$|\mathcal{R}(s)|\le n^2\big(\frac{2C(K_0)\epsilon}{\lambda_0 N^2}s^2+\frac{1}{\lambda_0 N^2}+\frac{2C(K_0)\epsilon}{\lambda_0 N^2}s\big)\|\Delta x\|^2.$$

Using the Euler-Lagrange equation $\frac{d}{dt}(\frac{\partial L_\epsilon}{\partial v})(d\gamma_0(s),s)=\frac{\partial L_\epsilon}{\partial q}(d\gamma_0(s),s)$,

$$\begin{aligned}(3.9)\le&\int_0^N\langle\frac{\partial L_\epsilon}{\partial q}(d\gamma_0(s),s),s\frac{\Delta x}{N}\rangle+\langle\frac{\partial L_\epsilon}{\partial v}(d\gamma_0(s),s),\frac{\Delta x}{N}\rangle ds+\int_0^N\mathcal{R}(s)\ ds\\ \le&\langle\frac{\partial L_\epsilon}{\partial v}(d\gamma_0(s),s),s\frac{\Delta x}{N}\rangle|_0^N+\int_0^N n^2(\frac{2C(K_0)\epsilon s^2+1+2C(K_0)\epsilon s}{\lambda_0 N^2})\|\Delta x\|^2 ds\\ \le&\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\Delta x\rangle+\big(\frac{4n^2C(K_0)N}{3\lambda_0}\epsilon+\frac{n^2}{\lambda_0 N}+\frac{C(K_0)\epsilon}{\lambda_0}\big)\|\Delta x\|^2\\ \le&\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\Delta x\rangle+C_1\sqrt{\epsilon}\|\Delta x\|^2.\end{aligned}$$

The last inequality follows from $N=[\frac{1}{\sqrt{\epsilon}}]$. Notice that C_1 only depends on H_0. Therefore,

$$u_\epsilon(x+\Delta x,0)-u_\epsilon(x,0)\le\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\Delta x\rangle+C_1\sqrt{\epsilon}\|\Delta x\|^2. \tag{3.11}$$

We set $\vec{e_1}=(1,0,\dots,0)^T,\vec{e_i}=(0,\dots,1,\dots,0)^T\in\mathbb{R}^n$, so by (3.11) we get

$$0=u_\epsilon(x\pm\vec{e_i},0)-u_\epsilon(x,0)\le\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\pm\vec{e_i}\rangle+C_1\sqrt{\epsilon}\|\pm\vec{e_i}\|,$$

$$|\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\vec{e_i}\rangle|\le C_1\sqrt{\epsilon}\|\pm\vec{e_i}\|=C_1\sqrt{\epsilon}.$$

From the Cauchy inequality,

$$|\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\Delta x\rangle|=|\sum_{i=1}^{n}\langle\frac{\partial L_\epsilon}{\partial v}(x,\dot\gamma_0(N),N),\Delta x_i\vec{e_i}\rangle|\le C_1\sqrt{n\epsilon}\|\Delta x\|.$$

Therefore, use (3.11) and we obtain

$$u_\epsilon(x+\Delta x,0)-u_\epsilon(x,0)\le C_1(\sqrt{n}+1)\sqrt{\epsilon}\|\Delta x\|,\quad\forall x\in\mathbb{R}^n,\ \|\Delta x\|\le 1.$$

Similarly, we can prove

$$u_\epsilon(x,0)-u_\epsilon(x+\Delta x,0)\le C_1(\sqrt{n}+1)\sqrt{\epsilon}\|-\Delta x\|,\quad\forall x\in\mathbb{R}^n,\ \|\Delta x\|\le 1.$$

Since $u_\epsilon(x,0)$ is periodic in x,

$$|u_\epsilon(x+\Delta x,0)-u_\epsilon(x,0)|\le C_1(\sqrt{n}+1)\sqrt{\epsilon}\|\Delta x\|,\quad\forall x\in\mathbb{R}^n.$$

For general t, we can prove similarly. Hence (3.8) holds.

Step 2. u_ϵ is a viscosity solution of the following Hamilton-Jacobi equation:

$$\partial_t u_\epsilon + H_\epsilon(x, d_x u_\epsilon, t) = \alpha_\epsilon(0).$$

Suppose u_ϵ is differentiable at (x, t). By (3.8), we have

$$\|d_x u_\epsilon(x,t)\| \le C_1\sqrt{n\epsilon}.$$

Then, there exists a constant $C_2 = C_2(H_0) > 0$ such that

$$|H_0(d_x u_\epsilon(x,t)) - H_0(0)| \le C_2\sqrt{\epsilon}. \tag{3.12}$$

On the other hand, by Lemma 3.3 and (3.12), we obtain

$$\begin{aligned}|\partial_t u_\epsilon(x,t)| &= |\alpha_\epsilon(0) - H_0(d_x u_\epsilon(x,t)) - \epsilon H_1(x, d_x u_\epsilon(x,t), t)| \\ &= |\alpha_\epsilon(0) - H_0(0) + H_0(0) - H_0(d_x u_\epsilon(x,t)) - \epsilon H_1(x, d_x u_\epsilon(x,t), t)| \\ &\le |\alpha_\epsilon(0) - \alpha_0(0)| + |H_0(0) - H_0(d_x u_\epsilon(x,t))| + |\epsilon H_1| \\ &\le \epsilon C(K_0) + C_2\sqrt{\epsilon} + \epsilon C(K_0).\end{aligned}$$

Combining Step 1 with Step 2, we obtain that, for almost every $(x,t) \in \mathbb{R}^n \times \mathbb{R}$,

$$\|d_x u_\epsilon(x,t)\| \le C_1(\sqrt{n}+1)\sqrt{\epsilon}, \quad |\partial_t u_\epsilon(x,t)| \le \epsilon C(K_0) + C_2\sqrt{\epsilon} + \epsilon C(K_0).$$

Since $u_\epsilon(x,t)$ is Lipschitz and differentiable almost everywhere, it's easy to know that there exists a constant $D = D(H_0)$ such that

$$|u_\epsilon(x,t) - u_\epsilon(y,s)| \le D\sqrt{\epsilon}(\|x-y\| + \|t-s\|).$$

This completes our proof. □

Proof of Theorem 1.2. For simplicity, we just prove our theorem for $c = 0 \in H^1(\mathbb{T}^n, \mathbb{R})$.

(1). Because $\gamma_\epsilon(t)$ is calibrated by some weak KAM solution u_ϵ, by Proposition 2.2, u_ϵ must be differentiable at $(\gamma_\epsilon(t), t)$, $\forall\, t \in (-\infty, +\infty)$. Thus,

$$\dot\gamma_\epsilon(t) = \frac{\partial H_\epsilon}{\partial p}(\gamma_\epsilon(t), d_{\gamma_\epsilon(t)}u_\epsilon, t) = \frac{\partial H_0}{\partial p}(d_{\gamma_\epsilon(t)}u_\epsilon) + \epsilon\frac{\partial H_1}{\partial p}(\gamma_\epsilon(t), d_{\gamma_\epsilon(t)}u_\epsilon, t),$$

$$\dot\gamma_\epsilon(0) = \frac{\partial H_\epsilon}{\partial p}(\gamma_\epsilon(0), d_{\gamma_\epsilon(0)}u_\epsilon, 0) = \frac{\partial H_0}{\partial p}(d_{\gamma_\epsilon(0)}u_\epsilon) + \epsilon\frac{\partial H_1}{\partial p}(\gamma_\epsilon(0), d_{\gamma_\epsilon(0)}u_\epsilon, 0).$$

Invoking the Taylor formula, there exists $\theta \in (0,1)$ such that

$$\begin{aligned}\|\dot\gamma_\epsilon(t) - \dot\gamma_\epsilon(0)\| =& \|\frac{\partial^2 H_0}{\partial p^2}\big(\theta d_{\gamma_\epsilon(t)}u_\epsilon + (1-\theta)d_{\gamma_\epsilon(0)}u_\epsilon\big)\big(d_{\gamma_\epsilon(t)}u_\epsilon - d_{\gamma_\epsilon(0)}u_\epsilon\big) \\ &+ \epsilon\frac{\partial H_1}{\partial p}(\gamma_\epsilon(t), d_{\gamma_\epsilon(t)}u_\epsilon, t) - \epsilon\frac{\partial H_1}{\partial p}(\gamma_\epsilon(0), d_{\gamma_\epsilon(0)}u_\epsilon, 0)\|.\end{aligned}$$

In view of Theorem 1.1 and Lemma 3.4, we conclude that

$$\begin{aligned}\|\dot\gamma_\epsilon(t) - \dot\gamma_\epsilon(0)\| &\le \frac{1}{\lambda_0}\|d_{\gamma_\epsilon(t)}u_\epsilon - d_{\gamma_\epsilon(0)}u_\epsilon\| + 2\epsilon C(K_0) \\ &\le \frac{2}{\lambda_0}D\sqrt{\epsilon} + 2\epsilon C(K_0) \le 2(\frac{D}{\lambda_0} + C(K_0))\sqrt{\epsilon}.\end{aligned}$$

This completes the proof of (1).

Conclusion (2) can be proved in the same way.

The first part of (3) is similar to (1). Let's prove the second part. For integrable systems, one has $\tilde{\mathcal{M}}_0 = \tilde{\mathcal{A}}_0 = \tilde{\mathcal{N}}_0 = \mathbb{T}^n \times \{\frac{\partial H_0}{\partial p}(0)\} \times \mathbb{T}$. Each minimal orbit $\gamma_\epsilon(t)$ in the Mather set $\mathcal{M}_\epsilon$ is calibrated by some weak KAM solution u_ϵ (see, for instance, [5])

$$\begin{aligned}\|\dot{\gamma}_\epsilon(t)-\frac{\partial H_0}{\partial p}(0)\| &= \|\frac{\partial H_0}{\partial p}(d_{\gamma_\epsilon(t)}u_\epsilon)+\epsilon\frac{\partial H_1}{\partial p}(\gamma_\epsilon(t),d_{\gamma_\epsilon(t)}u_\epsilon,t)-\frac{\partial H_0}{\partial p}(0)\| \\ &\leq \frac{D\sqrt{\epsilon}}{\lambda_0}+\epsilon C(K_0)\sim O(\sqrt{\epsilon}),\quad \forall t\in(-\infty,+\infty).\end{aligned}$$

Then,

$$d_H(\tilde{\mathcal{M}}_\epsilon,\tilde{\mathcal{M}}_0)\sim O(\sqrt{\epsilon}),$$

where d_H is the Hausdorff distance. Similarly,

$$d_H(\tilde{\mathcal{A}}_\epsilon,\tilde{\mathcal{A}}_0)\sim O(\sqrt{\epsilon}),\quad d_H(\tilde{\mathcal{N}}_\epsilon,\tilde{\mathcal{N}}_0)\sim O(\sqrt{\epsilon}).$$

This completes our proof. □

Acknowledgements

The authors greatly appreciate the discussions with Professor Chong-Qing Cheng and Professor Wei Cheng.

References

[1] Diogo Aguiar Gomes, *Regularity theory for Hamilton-Jacobi equations*, J. Differential Equations **187** (2003), no. 2, 359–374, DOI 10.1016/S0022-0396(02)00013-X. MR1949445

[2] V. I. Arnol′d, *Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian*, Uspehi Mat. Nauk **18** (1963), no. 5 (113), 13–40. MR0163025 (29 #328)

[3] P. Bernard, V. Kaloshin, and K. Zhang, *Arnold diffusion in arbitrary degrees of freedom and crumpled 3-dimensional normally hyperbolic invariant cylinders*, preprint, arXiv:1112.2773 [math.DS] (2011).

[4] Patrick Bernard, *Homoclinic orbits to invariant sets of quasi-integrable exact maps*, Ergodic Theory Dynam. Systems **20** (2000), no. 6, 1583–1601, DOI 10.1017/S0143385700000870. MR1804946

[5] Patrick Bernard, *The dynamics of pseudographs in convex Hamiltonian systems*, J. Amer. Math. Soc. **21** (2008), no. 3, 615–669, DOI 10.1090/S0894-0347-08-00591-2. MR2393423

[6] David Bernstein and Anatole Katok, *Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians*, Invent. Math. **88** (1987), no. 2, 225–241, DOI 10.1007/BF01388907. MR880950

[7] Piermarco Cannarsa, Wei Cheng, and Qi Zhang, *Propagation of singularities for weak KAM solutions and barrier functions*, Comm. Math. Phys. **331** (2014), no. 1, 1–20, DOI 10.1007/s00220-014-2106-x. MR3231994

[8] Chong-Qing Cheng, *Arnold diffusion in nearly integrable Hamiltonian systems*, preprint, arXiv:1207.4016 [math.DS] (2012).

[9] Chong-Qing Cheng and Jinxin Xue, *Arnold diffusion in nearly integrable Hamiltonian systems of arbitrary degrees of freedom*, preprint, arXiv:1503.04153 [math.DS] (2015).

[10] Chong-Qing Cheng and Min Zhou, *Global normally hyperbolic cylinders in Lagrangian systems*, to appear in Math. Res. Lett.

[11] G. Contreras, R. Iturriaga, and H. Sánchez-Morgado, *Weak solutions of the Hamilton-Jacobi equation for time periodic Lagrangians*, preprint, arXiv:1307.0287 [math.DS] (2000).

[12] L. C. Evans and D. Gomes, *Effective Hamiltonians and averaging for Hamiltonian dynamics. I*, Arch. Ration. Mech. Anal. **157** (2001), no. 1, 1–33, DOI 10.1007/PL00004236. MR1822413

[13] Albert Fathi, *Weak KAM theorem in Lagrangian dynamics*, to be published by Cambridge University Press.

[14] Diogo Aguiar Gomes, *Perturbation theory for viscosity solutions of Hamilton-Jacobi equations and stability of Aubry-Mather sets*, SIAM J. Math. Anal. **35** (2003), no. 1, 135–147, DOI 10.1137/S0036141002405960. MR2001468

[15] V. Kaloshin and K. Zhang, *A strong form of Arnold diffusion for two and a half degrees of freedom*, preprint, arXiv:1212.1150 [math.DS] (2012).

[16] A. N. Kolmogorov, *On conservation of conditionally periodic motions for a small change in Hamilton's function* (Russian), Dokl. Akad. Nauk SSSR (N.S.) **98** (1954), 527–530. MR0068687

[17] Zhenguo Liang, Jun Yan, and Yingfei Yi, *Viscous stability of quasi-periodic tori*, Ergodic Theory Dynam. Systems **34** (2014), no. 1, 185–210, DOI 10.1017/etds.2012.120. MR3163030

[18] Ricardo Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996), no. 2, 273–310, DOI 10.1088/0951-7715/9/2/002. MR1384478

[19] John N. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. **207** (1991), no. 2, 169–207, DOI 10.1007/BF02571383. MR1109661

[20] John N. Mather and Giovanni Forni, *Action minimizing orbits in Hamiltonian systems*, Transition to chaos in classical and quantum mechanics (Montecatini Terme, 1991), Lecture Notes in Math., vol. 1589, Springer, Berlin, 1994, pp. 92–186, DOI 10.1007/BFb0074076. MR1323222

[21] J. Moser, *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1962** (1962), 1–20. MR0147741

[22] N. N. Nehorošev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Uspehi Mat. Nauk **32** (1977), no. 6(198), 5–66, 287. MR0501140 (58 #18570)

[23] Kaizhi Wang and Jun Yan, *A new kind of Lax-Oleinik type operator with parameters for time-periodic positive definite Lagrangian systems*, Comm. Math. Phys. **309** (2012), no. 3, 663–691, DOI 10.1007/s00220-011-1375-x. MR2885604

[24] Min Zhou, *Hölder regularity of weak KAM solutions in a priori unstable systems*, Math. Res. Lett. **18** (2011), no. 1, 75–92, DOI 10.4310/MRL.2011.v18.n1.a6. MR2770583

Department of Mathematics, Nanjing University, Nanjing, Jiangsu, People's Republic of China 210093

E-mail address: `qinboChen1990@gmail.com`

School of Information Management, Nanjing University, Nanjing, Jiangsu, People's Republic of China 210093

E-mail address: `minzhou@nju.edu.cn`