Lecture notes on the weak KAM theorem

Hitoshi Ishii*

The following notes are based on the lectures which I delivered at Hokkaido University for the period, July 20 to July 23, 2004. Part of notes has not completed yet. They may serve as an introduction to the lecture notes [Fa2] due to A. Fathi.

1. Lagrangians and Hamiltonians: conjugate functions of convex functions

Let $L : \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{R}$ be given, where $\mathbf{T}^n$ denotes the n-dimensional torus. We assume throughout these notes:

- $L \in C^2(\mathbf{T}^n \times \mathbf{R}^n)$.
- $v \mapsto L(x, v)$ is locally *uniformly convex*. More precisely, for each $R > 0$ there is a constant $\varepsilon_R > 0$ such that

$$L_{vv}(x, v) \geq \varepsilon_R I \quad \text{ if } |v| \leq R,$$

  where I denotes the unit matrix of order n.
- L has a *superlinear growth*. That is,

$$\lim_{r\to\infty} \inf\{L(x, v)/|v| \mid |v| \geq r\} = \infty.$$

Here and henceforth we write $L_{vv}(x, v)$ for the Hessian matrix $(L_{v_i v_j}(x, v))$. Similarly we write $L_v(x, v)$ for the gradient $(L_{v_i}(x, v))$, $L_x(x, v)$ for $(L_{x_i}(x, v))$, etc.

We define the conjugate function $H : \mathbf{T}^n \times \mathbf{R}^n$ of L by

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

Here $p \cdot v$ denotes the Euclidean inner product of p and v, which may be denoted as well by pv in what follows.

---

* Supported in part by the Grant-in-Aids for Scientific Research, JSPS, No. 15340051, No. 18204009.

When we have in mind the variational problem

$$\inf_{\gamma} \int_0^T L(\gamma(t), \dot{\gamma}(t))\,\mathrm{d}t,$$

the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma(t), \dot{\gamma}(t)) = L_x(\gamma(t), \dot{\gamma}(t)),$$

or the Hamiltonian system

$$\dot{X}(t) = H_p(X(t), P(t)), \quad \dot{P}(t) = -H_x(X(t), P(t)),$$

we call L the *Lagrangian* and H the *Hamiltonian*.

A typical example of Lagrangians L is given by

$$L(x, v) = \frac{1}{2}|v|^2 + V(x),$$

where $V \in C(\mathbf{R}^n)$. The Hamiltonian H is then given by

$$H(x, p) = \frac{1}{2}|p|^2 - V(x).$$

Proposition 1.1. *H satisfies the following properties:*

(a) $H \in C^2(\mathbf{T}^n \times \mathbf{R}^n)$.

(b) $L(x, v) = \max_{p \in \mathbf{R}^n} (v \cdot p - H(x, p))$ *for all* $(x, v) \in \mathbf{T}^n \times \mathbf{R}^n$.

(c) *For each* $R > 0$ *there is a constant* $\delta_R > 0$ *such that*

$$H_{pp}(x, p) \geq \delta_R I \qquad \text{if } |p| \leq R.$$

(d) $\lim_{r \to \infty} \inf\{H(x, p)/|p| \mid |p| \geq r\} = \infty$.

Proof. 1. For fixed $(x, p) \in \mathbf{T}^n \times \mathbf{R}^n$ the function $v \mapsto p \cdot v - L(x, v)$ on $\mathbf{R}^n$ attains a maximum since it is continuous and

$$\lim_{|v| \to \infty} (p \cdot v - L(x, v)) = -\infty.$$

Let $v = V(x, p)$ be a maximum point of this function. Such a maximum point is determined uniquely by (x, v) since $v \mapsto L(x, v) - p \cdot v$ is locally uniformly convex.

2. By the elementary calculus, we find that

$$p = L_v(x, V(x,p)). \tag{1.1}$$

Since $L_{vv}(x,v) > 0$ and hence $\det L_{vv}(x, V(x,p)) \neq 0$, by the implicit function theorem we see that the function V on $\mathbf{T}^n \times \mathbf{R}^n$ is a C^1 map. Since $L_v(x, V(x,p)) = p$ for all $(x,p) \in \mathbf{T}^n \times \mathbf{R}^n$, for given $x \in \mathbf{T}^n$, the map $v \mapsto L_v(x,v)$, $\mathbf{R}^n \to \mathbf{R}^n$ is surjective. On the other hand, because of the local uniform convexity of L, for any $x \in \mathbf{T}^n$ and $v_1, v_2 \in \mathbf{R}^n$, we have

$$\begin{aligned}
&(L_v(x,v_1) - L_v(x,v_2)) \cdot (v_1 - v_2) \\
&= \int_0^1 L_{vv}(x, sv_1 + (1-s)v_2)\,\mathrm{d}s(v_1 - v_2) \cdot (v_1 - v_2) \geq \varepsilon_R |v_1 - v_2|^2,
\end{aligned}$$

where $R := \max\{|v_1|, |v_2|\}$. This shows that for each $x \in \mathbf{T}^n$, the map $v \mapsto L_v(x,v)$, $\mathbf{R}^n \to \mathbf{R}^n$ is injective. Thus we conclude that for each $x \in \mathbf{T}^n$, the map $v \mapsto L_v(x,v)$, $\mathbf{R}^n \to \mathbf{R}^n$ is a bijection.

3. Since

$$H(x,p) = p \cdot V(x,p) - L(x, V(x,p)) \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n,$$

we see that $H \in C^1(\mathbf{T}^n \times \mathbf{R}^n)$. Differentiating this relation, we have

$$\begin{aligned}
H_x &= p \cdot V_x(x,p) - L_x(x, V(x,p)) - L_v(x, V(x,p)) \cdot V_x(x,p) \\
&= p \cdot V_x(x,p) - L_x(x, V(x,p)) - p \cdot V_x(x,p) = -L_x(x, V(x,p)), \\
H_p(x,p) &= V(x,p) - p \cdot V_p(x,p) - L_v(x, V(x,p)) \cdot V_p(x,p) = V(x,p).
\end{aligned}$$

Since the functions

$$H_x(x,p) = -L_x(x, V(x,p)), \quad H_p(x,p) = V(x,p) \tag{1.2}$$

are C^1 functions on $\mathbf{T}^n \times \mathbf{R}^n$, we conclude that $H \in C^2(\mathbf{T}^n \times \mathbf{R}^n)$. Combining the latter of (1.2) with (1.1), we get

$$p = L_v(x, H_p(x,p)) \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n. \tag{1.3}$$

For fixed $(x,v) \in \mathbf{T}^n \times \mathbf{R}^n$, let $p = L_v(x,v)$. Since $w = V(x,p)$ is the unique solution of $p = L_v(x,w)$, we see that $v = V(x,p)$. Hence, we conclude that

$$v = H_p(x,p) = H_p(x, L_v(x,v)) \quad \forall (x,v) \in \mathbf{T}^n \times \mathbf{R}^n. \tag{1.4}$$

4. By the definition of H, we have

$$H(x,p) \geq p \cdot v - L(x,v) \quad \forall x \in \mathbf{T}^n,\ p, v \in \mathbf{R}^n.$$

Hence, we have

$$L(x,v) \geq p \cdot v - H(x,p) \quad \forall x \in \mathbf{T}^n,\ p, v \in \mathbf{R}^n.$$

That is,

$$L(x,v) \geq \sup_{p \in \mathbf{R}^n} (v \cdot p - H(x,p)) \quad \forall (x,v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Now fix $(x,v) \in \mathbf{T}^n \times \mathbf{R}^n$. Set $p = L_v(x,v)$. From (1.4), we have $v = H_p(x,p)$ and therefore

$$H(x,p) = p \cdot v - L(x,v).$$

That is,

$$L(x,v) = v \cdot p - H(x,p).$$

Hence

$$L(x,v) = \max_{p \in \mathbf{R}^n} (v \cdot p - H(x,p)) = v \cdot L_v(x,v) - H(x, L_v(x,p)) \quad \forall (x,v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

5. From (1.3) we have

$$p = L_v(x, V(x,p)) = L_v(x, H_p(x,v)) \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n,$$

and hence

$$I = L_{vv}(x, V(x,p)) H_{pp}(x,p) \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Hence, noting that $L_{vv}(x,v) > 0$, we have

$$H_{pp}(x,p) = L_{vv}(x, H_p(x,p))^{-1} \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Fix $R > 0$ and set

$$A_R = \max\{L_{vv}(x, H_p(x,p))\xi \cdot \xi \mid (x,p,\xi) \in \mathbf{T}^n \times B(0,R) \times S^{n-1}\}.$$

Then we have

$$L_{vv}(x, H_p(x,p)) \leq A_R I \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Consequently, we get

$$H_{pp}(x,p) = L_{vv}(x, H_p(x,p)) \geq A_R^{-1} I \quad \forall (x,p) \in \mathbf{T}^n \times B(0,R),$$

which shows (c) with $\delta_R = A_R^{-1}$.

Fix any $M > 0$. We have

$$\begin{aligned}\frac{H(x,p)}{|p|} &= \max_{v\in\mathbf{R}^n} \left(p\cdot v - \frac{L(x,v)}{|p|} \right) \geq p \cdot M\bar{p} - \frac{L(x,M\bar{p})}{|p|} \\ &= M|p| - \frac{\max_{v\in B(0,M)} L(x,v)}{|p|} \to \infty \qquad \text{as } |p| \to \infty.\end{aligned}$$

Here we have used the notation that $\bar{p}$ denotes the unit vector $p/|p|$. Thus we see that

$$\lim_{r\to\infty} \inf \left\{ \frac{H(x,p)}{|p|} \;\middle|\; x \in \mathbf{T}^n,\ |p| \geq r \right\} = \infty, \qquad \square$$

We have observed the following as well.

Proposition 1.2. *We have:*

(a) $H(x,p) = p \cdot v - L(x,v)$ *for* $v = H_p(x,p)$ *and*

$$H(x,p) > p \cdot v - L(x,v) \qquad \textit{if } v \neq H_p(x,p).$$

(b) *For each* $x \in \mathbf{T}^n$, $p \mapsto H_p(x,p),\ \mathbf{R}^n \to \mathbf{R}^n$ *is a* C^1 *diffeomorphism and its inverse map is given by*

$$v \mapsto L_v(x,v),\ \mathbf{R}^n \to \mathbf{R}^n.$$

(c) $H_x(x,p) = -L_x(x, H_p(x,p))$ *for all* $(x,p) \in \mathbf{T}^n \times \mathbf{R}^n$.

- The map $\mathcal{L} : \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{T}^n \times \mathbf{R}^n,\ (x,v) \mapsto (x, L_v(x,v))$ is called the *Legendre transform.* The Legendre transform $\mathcal{L}$ is a C^1 diffeomorphism between $\mathbf{T}^n \times \mathbf{R}^n$ and $\mathbf{T}^n \times \mathbf{R}^n$. Its inverse is given by $\mathcal{L}^{-1} : (x,p) \mapsto (x, H_p(x,p))$.

2. Euler-Lagrange equations and Hamiltonian systems

Associated with the variational problem

$$\inf_{\gamma} \int_0^T L(\gamma(t), \dot\gamma(t))\, \mathrm{d}t,$$

where the infimum is taken over all $\gamma \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ (the space of absolutely continuous functions γ on $[0,\, T]$) which satisfy $\gamma(0) = a$ and $\gamma(T) = b$, where $a,\ b \in \mathbf{T}^n$ are fixed, is the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma(t), \dot{\gamma}(t)) = L_x(\gamma(t), \dot{\gamma}(t)) \quad \forall t \in (0, T),$$

which is equivalent to

$$\ddot{\gamma}(t) = L_{vv}(\gamma(t), \dot{\gamma}(t))^{-1}(L_x(\gamma(t), \dot{\gamma}(t)) - L_{vx}(\gamma(t), \dot{\gamma}(t))\dot{\gamma}(t)).$$

Note that the function

$$(x, v) \mapsto L_{vv}(x, v)^{-1}(L_x(x, v) - L_{vx}(x, v)v),\ \ \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{R}^n$$

is a continuous function, but it is not guaranteed to be locally Lipschitz continuous.

The corresponding Hamiltonian system is given by

$$\begin{cases} \dot{X}(t) = H_p(X(t), P(t)) \\ \dot{P}(t) = -\, H_x(X(t), P(t)). \end{cases} \tag{2.1}$$

Since $(H_p, -H_x)$ is a C^1 function on $\mathbf{T}^n \times \mathbf{R}^n$, one can apply the Cauchy-Lipschitz theorem for (2.1).

Proposition 2.1. (a) *If $(X(t), P(t))$ exists for $\alpha < t < \beta$, then*

$$H(X(t), P(t)) = H(X(t_0), P(t_0)) \quad \forall t \in (\alpha,\, \beta),$$

where $t_0 \in (\alpha,\, \beta)$ is any fixed number. (b) *For any $(x_0, p_0) \in \mathbf{T}^n \times \mathbf{R}^n$ and $t_0 \in \mathbf{R}$ there is a unique solution $(X(t), P(t))$, defined on $\mathbf{R}$, of (2.1) satisfying*

$$X(t_0) = x_0, \quad P(t_0) = p_0.$$

Proof. 1. We compute that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} H(X(t), P(t)) &= H_x(X(t), P(t))\dot{X}(t) + H_p(X(t), P(t))\dot{P}(t) \\ &= H_x(X(t), P(t))H_p(X(t), P(t)) - H_p(X(t), P(t))H_x(X(t), P(t)) \\ &= 0. \end{aligned}$$

Hence we have

$$H(X(t), P(t)) = H(X(t_0), P(t_0)) \quad \forall t \in (\alpha,\, \beta),$$

which proves (a).

2. By the Cauchy-Lipschitz theorem, there is a unique solution $(X(t), P(t))$ of (2.1) satisfying $(X(t_0), P(t_0)) = (x_0, p_0)$. Let (α, β) be the maximal interval of existence for the solution $(X(t), P(t))$. There is a constant $C_1 > 0$ such that

$$H(x,p) \geq |p| - C_1 \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Then, since

$$|P(t)| - C_1 \leq H(x_0, p_0) \quad \forall t \in (\alpha,\, \beta),$$

$\{(X(t), P(t)) \mid t \in (\alpha,\, \beta)\}$ is bounded in $\mathbf{T}^n \times \mathbf{R}^n$. This implies, due to the Cauchy-Lipschitz theorem in ODE theory, that $(\alpha,\, \beta) = \mathbf{R}$, which concludes the proof of (b). □

Proposition 2.2. *Let $(X(t), P(t))$ be a solution of* (2.1) *and set $\gamma(t) := X(t)$. Then γ is a C^2 function on $\mathbf{R}$ and satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma(t), \dot\gamma(t)) = L_x(\gamma(t), \dot\gamma(t)) \quad \forall t \in \mathbf{R}. \tag{2.2}$$

Proof. Since $\dot\gamma(t) = \dot X(t) = H_p(\gamma(t), P(t))$, the function γ is a C^2 function on $\mathbf{R}$ and also, recalling that $(x,p) = (x, L_v(x,v))$ if and only if $(x,v) = (x, H_p(x,p))$, we find that

$$P(t) = L_v(\gamma(t), \dot\gamma(t)).$$

Therefore we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma(t), \dot\gamma(t)) &= \dot P(t) = -H_x(\gamma(t), P(t)) = L_x(\gamma(t), H_p(\gamma(t), P(t))) \\ &= L_x(\gamma(t), \dot\gamma(t)).\end{aligned}$$

□

Proposition 2.3. *Let $\gamma(t)$ be a C^1 function on $(\alpha,\, \beta)$ such that*

$$t \mapsto L_v(\gamma(t), \dot\gamma(t))$$

is a C^1 function on $(\alpha,\, \beta)$ and such that

$$\frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma(t), \dot\gamma(t)) = L_x(\gamma(t), \dot\gamma(t)) \quad \forall t \in (\alpha,\, \beta). \tag{2.3}$$

Then $(X(t), P(t)) := (\gamma(t), L_v(\gamma(t), \dot\gamma(t)))$ is a solution of (2.1) *on $(\alpha,\, \beta)$.*

Proof. Note first that

$$\dot\gamma(t) = H_p(\gamma(t), P(t)),$$

which, in particular, shows that $\gamma \in C^2((\alpha,\,\beta))$ and $\dot X(t) = H_p(X(t), P(t))$. By (2.3), we get

$$\dot P(t) = L_x(\gamma(t), \dot\gamma(t)) = -H_x(\gamma(t), P(t)) = -H_x(X(t), P(t)).$$

Here we have used the observation (Proposition 1.2, (c)) that

$$H_x(\gamma(t), P(t)) = -L_x(\gamma(t), H_p(\gamma(t), P(t))) = -L_x(\gamma(t), \dot\gamma(t)).$$

Thus we see that $(X(t), P(t))$ is a solution of (2.1). □

- The Legendre transform $\mathcal{L}$ maps the solutions $(\gamma, \dot\gamma)$ of the Euler-Lagrange equation (2.3) to the solutions $(X(t), P(t))$ of the Hamiltonian system (2.1).

Notation. We define the collections $\{\phi_t^L\}_{t\in\mathbf{R}}$ and $\{\phi_t^H\}_{t\in\mathbf{R}}$ of maps of $\mathbf{T}^n \times \mathbf{R}^n$ to $\mathbf{T}^n \times \mathbf{R}^n$ by

$$\phi_t^L(x, v) = (\gamma(t), \dot\gamma(t)),$$

where γ is the solution of (2.3) which satisfies the initial condition $(\gamma(0), \dot\gamma(0)) = (x, v)$ and

$$\phi_t^H(x, p) = (X(t), P(t)),$$

where (X, P) is the solution of (2.1) satisfying the condition $(X(0), P(0)) = (x, p)$. By the uniqueness of the solution for the Cauchy problem, we see that

$$\phi_{t+s}^L = \phi_t^L \circ \phi_s^L, \qquad \phi_{t+s}^H = \phi_t^H \circ \phi_s^H \qquad \forall t, s \in \mathbf{R}.$$

As we have seen in Propositions 2.2 and 2.3,

$$\mathcal{L} \circ \phi_t^L \circ \mathcal{L}^{-1} = \phi_t^H \quad \forall t \in \mathbf{R}.$$

3. Existence of minimizers for actions

Let $L : \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{R}$ be a given Lagrangian which satisfies the assumptions described before. Let ψ be a given function on $\mathbf{T}^n$ which satisfies:

- $\psi \in C(\mathbf{T}^n)$.

Fix $T > 0$ and $x_0 \in \mathbf{T}^n$. Consider the variational problem

$$V = \inf_{\gamma} \left(\int_0^T L(\gamma(t), \dot\gamma(t))\,\mathrm{d}t + \psi(\gamma(T)) \right), \tag{3.1}$$

where γ ranges over all $\gamma \in \mathrm{AC}([0,\,T], \mathbf{T}^n)$ (the space of all absolutely continuous functions on $[0,\,T]$) such that $\gamma(0) = x_0$.

Theorem 3.1. *There exists a minimizer for* V*.*

Lemma 3.2. *Let* $x_1 \in \mathbf{T}^n$ *and define*

$$V(x_1) = \inf_{\gamma} \int_0^T L(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t,$$

where γ *ranges over all* $\gamma \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ *such that* $\gamma(0) = x_0$ *and* $\gamma(T) = x_1$*. Then there is a minimizer for* $V(x_1)$*.*

Lemma 3.3. *There is a constant* $C_0 > 0$ *such that*

$$V(x_1) \le C_0 \quad \forall x_1 \in \mathbf{T}^n.$$

Proof. Define $\gamma_0 \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ by $\gamma_0(t) := x_0 + T^{-1}t(x_1 - x_0)$. We have

$$\int_0^T L(\gamma_0(t), \dot{\gamma}_0(t))\, \mathrm{d}t = \int_0^T L(\gamma_0(t), T^{-1}(x_1 - x_0))\, \mathrm{d}t \le C_0,$$

where

$$C_0 := T \max\{L(x, v) \mid (x, v) \in \mathbf{T}^n \times \mathbf{R}^n, |v| \le T^{-1}\sqrt{n}\}. \qquad \square$$

Lemma 3.4. *Let* $\{\gamma_k\}_{k\in\mathbf{N}} \subset \mathrm{AC}([0,\, T], \mathbf{T}^n)$*. Assume that* $\gamma_k(0) = x_0$ *for all* $k \in \mathbf{N}$ *and that there is a constant* $C > 0$ *such that*

$$\int_0^T L(\gamma_k(t), \dot{\gamma}_k(t))\, \mathrm{d}t \le C \quad \forall k \in \mathbf{N}.$$

Then there exist a subsequence $\{\gamma_{k_j}\}_{j\in\mathbf{N}}$ *and* $\gamma \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ *such that*

$$\gamma_{k_j}(t) \to \gamma(t) \quad \text{uniformly on } [0,\, T]$$

as $j \to \infty$ *and*

$$\int_0^T L(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t \le \liminf_{k\to\infty} \int_0^T L(\gamma_k(t), \dot{\gamma}_k(t))\, \mathrm{d}t.$$

Using Lemma 3.4 whose proof will be given later, we first prove Lemma 3.2 and Theorem 3.1.

Proof of Lemma 3.2. 1. Fix $x_1 \in \mathbf{T}^n$. Noting that L is bounded below, we set

$$L_0 = \min_{\mathbf{T}^n\times\mathbf{R}^n} L.$$

We have

$$L_0 T \leq V(x_1) \leq C_0,$$

where C_0 is the constant from Lemma 3.3.

2. Choose a minimizing sequence $\{\gamma_k\}_{k\in\mathbf{N}} \subset \mathrm{AC}([0,\, T], \mathbf{T}^n)$ for $V(x\mathbf{1})$ so that

$$\int_0^T L(\gamma_k(t), \dot\gamma_k(t))\,\mathrm{d}t < V(x_1) + \frac{1}{k} \quad \forall k \in \mathbf{N}.$$

Here the γ_k are assumed to satisfy $\gamma_k(0) = x_0$ and $\gamma_k(T) = x_1$.

Noting that

$$\int_0^T L(\gamma_k(t), \dot\gamma_k(t))\,\mathrm{d}t \leq C_0 + 1 \quad \forall k \in \mathbf{N},$$

by virtue of Lemma 3.4, there are a subsequence $\{\gamma_{k_j}\}_{j\in\mathbf{N}}$ and $\gamma \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ such that

$$\gamma_{k_j}(t) \to \gamma(t) \quad \text{uniformly on } [0,\, T] \ \text{ as } \ j \to \infty \tag{3.2}$$

and

$$\int_0^T L(\gamma(t), \dot\gamma(t))\,\mathrm{d}t \leq \liminf_{k\to\infty} \int_0^T L(\gamma_k(t), \dot\gamma_k(t))\,\mathrm{d}t. \tag{3.3}$$

From (3.2) we have

$$\gamma(0) = x_0, \quad \gamma(T) = x_1.$$

From (3.3) we get

$$\int_0^T L(\gamma(t), \dot\gamma(t))\,\mathrm{d}t \leq V(x_1).$$

Thus we find that γ is a minimizer for $V(x_1)$. □

Lemma 3.5. *The function V is lower semicontinuous on $\mathbf{T}^n$.*

Proof. Fix $x_1 \in \mathbf{T}^n$ and a sequence $\{y_k\}_{k\in\mathbf{N}} \subset \mathbf{T}^n$ so that $y_k \to x_1$ as $k \to \infty$. According to Lemma 3.2, for each $k \in \mathbf{N}$ there is a $\gamma_k \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ satisfying $\gamma_k(0) = x_0$ and $\gamma_k(T) = y_k$ such that

$$V(y_k) = \int_0^T L(\gamma_k(t), \dot\gamma_k(t))\,\mathrm{d}t \quad \forall k \in \mathbf{N}.$$

By Lemma 3.3, there is a constant $C_1 > 0$ such that

$$\int_0^T L(\gamma_k(t), \dot\gamma_k(t))\,\mathrm{d}t \leq C_1 \quad \forall k \in \mathbf{N}.$$

Now, by Lemma 3.4, we find a $\gamma \in \mathrm{AC}([0,\, T], \mathbf{T}^n)$ satisfying $\gamma(0) = x_1$ and $\gamma(T) = x_1$ such that

$$\int_0^T L(\gamma(t), \dot\gamma(t))\, \mathrm{d}t \le \liminf_{k\to\infty} \int_0^T L(\gamma_k(t), \dot\gamma_k(t))\, \mathrm{d}t.$$

This inequality implies that

$$V(x_1) \le \liminf_{k\to\infty} V(y_k),$$

which shows that V is lower semicontinuous on $\mathbf{T}^n$.

Proof of Theorem 3.1. Note that

$$V = \inf_\gamma \int_0^T L(\gamma(t), \dot\gamma(t))\, \mathrm{d}t = \min_{x_1 \in \mathbf{T}^n} (V + \psi)(x_1).$$

Since $V + \psi$ is lower semicontinuous on $\mathbf{T}^n$, there is a point $x_1 \in \mathbf{T}^n$ where it attains a minimum. By Lemma 3.2, there is a minimizer γ_1 for $V(x_1)$. Hence, γ_1 is a minimizer for V. □

It remains to prove Lemma 3.4. We fix C and $\{\gamma_k\}$ as in Lemma 3.4. By replacing $L(x, v)$ and C by $L(x, v) + C_2$ and $C + C_2 T$, respectively, where $C_2 > 0$ is a constant such that $\min_{\mathbf{T}^n \times \mathbf{R}^n} L \ge -C_2$, if necessary, we may assume that

$$L(x, v) \ge 0 \quad \text{for all } (x, v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Lemma 3.6. *The sequence $\{\gamma_k\}$ is equi-absolutely continuous on $[0,\, T]$.*

Proof. By the superlinearity of L, for any $A > 1$ there is a constant $C_A > 0$ such that

$$L(x, v) \ge A|v| - C_A \quad \forall (x, v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Hence, for any Borel set $B \subset [0,\, T]$ we have

$$C \ge \int_0^T L(\gamma_k(t), \dot\gamma_k(t))\, \mathrm{d}t \ge \int_B L(\gamma_k(t), \dot\gamma_k(t))\, \mathrm{d}t \ge \int_B (A|\dot\gamma_k(t)| - C_A)\, \mathrm{d}t,$$

that is,

$$\int_B |\dot\gamma_k(t)|\, \mathrm{d}t \le \frac{C_0}{A} + \frac{C_A}{A}|B|.$$

Fix any $\varepsilon > 0$. Choose $A > 0$ so that $C_0/A \le \frac{\varepsilon}{2}$ and $\delta > 0$ so that

$$\frac{C_A \delta}{A} < \frac{\varepsilon}{2}.$$

Then we have

$$|B| \le \delta \implies \int_B |\dot\gamma_k(t)|\,\mathrm{d}t < \varepsilon,$$

which shows that $\{\gamma_k\}$ is equi-absolutely continuous on $[0, T]$. □

Lemma 3.7 (Selection theorem of Helly). *For $k \in \mathbf{N}$ let $f_k : [0, T] \to \mathbf{R}$ be a non-decreasing function on $[0, T]$. Assume that $\{f_k\}$ is uniformly bounded on $[0, T]$. Then there is a subsequence $\{f_{k_j}\}$ of $\{f_k\}$ such that for all $t \in [0, T]$, the sequence $\{f_{k_j}(t)\}$ is convergent.*

See [Fr] for a proof of the above lemma.

Lemma 3.8. *Fix $(x, v) \in \mathbf{T}^n \times \mathbf{R}^n$ and $\varepsilon > 0$. Then there is a constant $\delta > 0$ such that for all $(y, w) \in \mathbf{T}^n \times \mathbf{R}^n$, if $|y - x| \le \delta$, then*

$$L(y, w) \ge L(x, v) + L_v(x, v) \cdot (w - v) - \varepsilon.$$

Proof. We choose a constant $M_1 > 0$ so that

$$|L_v(x, v)| \le M_1,$$

for instance, $M_1 = |L_v(x, v)| + 1$, and a constant $M_2 > 0$ so that

$$L(y, w) \ge 2M_1|w| - M_2 \quad \forall (y, w) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Then we have

$$\begin{aligned} L(y, w) &\ge |L_v(x, v)||w| - M_2 + M_1|w| \\ &\ge L_v(x, v) \cdot w + M_1|w| - M_2 \quad \forall (y, w) \in \mathbf{T}^n \times \mathbf{R}^n. \end{aligned}$$

Noting that

$$L(x, 0) \ge L(x, v) + L_v(x, v) \cdot (0 - v) \quad \forall (x, v) \in \mathbf{T}^n \times \mathbf{R}^n,$$

we choose $M_3 > 0$ so that

$$M_1M_3 - M_2 \ge L(x, v) - L_v(x, v) \cdot v.$$

Now, for all $(y, w) \in \mathbf{T}^n \times \mathbf{R}^n$, if $|w| \ge M_3$, then we have

$$\begin{aligned} \text{(3.4)}\quad L(y, w) &\ge L_v(x, v) \cdot w + M_1|w| - M_2 \ge L_v(x, v) \cdot w + M_1M_3 - M_2 \\ &\ge L_v(x, v) \cdot w + L(x, v) - L_v(x, v) \cdot v = L(x, v) + L_v(x, v) \cdot (w - v). \end{aligned}$$

By the convexity of $w \mapsto L(x, w)$, we have

$$L(x, w) \geq L(x, v) + L_v(x, v) \cdot (w - v) \quad \forall w \in \mathbf{R}^n.$$

Since the function

$$(y, w) \mapsto L(y, w) - L(x, v) - L_v(x, v) \cdot (w - v)$$

is uniform continuous on the compact set $\mathbf{T}^n \times B(0, M_3)$, there is a constant $\delta > 0$ such that for all $(y, w) \in \mathbf{T}^n \times \mathbf{R}^n$, if $|y - x| \leq \delta$, then

$$L(y, w) - L(x, v) - L_v(x, v) \cdot (w - v) \geq -\varepsilon. \tag{3.5}$$

Combining (3.4) and (3.5), we conclude that for all $(y, w) \in \mathbf{T}^n \times \mathbf{R}^n$, if $|y - x| \leq \delta$, then

$$L(y, w) \geq L(x, v) + L_v(x, v) \cdot (w - v) - \varepsilon. \qquad \square$$

Proof of Lemma 3.4. 1. Define the functions $\beta_k : [0, T] \to \mathbf{R}$ by

$$\beta_k(t) = \int_0^t L(\gamma_k(t), \dot{\gamma}_k(t))\, \mathrm{d}t.$$

We may choose a subsequence of $\{\beta_{k_j}\}$ of $\{\beta_k\}$ so that

$$\lim_{j\to\infty} \beta_{k_j}(T) = \liminf_{k\to\infty} \int_0^T L(\gamma_k(t), \dot{\gamma}_k(t))\, \mathrm{d}t.$$

The functions β_k are non-decreasing on $[0, T]$ since $L \geq 0$, and they are uniformly bounded since

$$0 \leq \beta_k(t) \leq \int_0^T L(\gamma_k(t), \dot{\gamma}_k(t))\, \mathrm{d}t \leq C \quad \forall t \in [0, T].$$

In view of Lemma 3.7, we may assume by selecting a subsequence of $\{\beta_{k_j}\}$ if necessary that

$$\beta_{k_j}(t) \to \beta(t) \quad \text{ as } j \to \infty$$

for some non-decreasing function $\beta : [0, T] \to \mathbf{R}$.

2. In view of the compactness of $\mathbf{T}^n$, Lemma 3.6 and Ascoli-Arzela theorem, by selecting again a subsequence of $\{\gamma_{k_j}\}$ if necessary, we may assume that

$$\gamma_{k_j}(t) \to \gamma(t) \quad \text{ uniformly on } [0, T] \text{ as } j \to \infty \tag{3.6}$$

for some $\gamma \in C([0,\, T])$. By Lemma 3.6, we can see that $\gamma \in AC([0, \mathbf{T}])$. We see as well that $\gamma(0) = x_0$ and $\gamma(T) = x_1$.

3. Note that non-decreasing functions and absolutely continuous functions are a.e. differentiable. Accordingly, β and γ are a.e. differentiable on $[0,\, T]$. Fix any of differentiability points in $[0,\, T)$ of (β, γ) and denote it by c. Fix any $\varepsilon > 0$ and, in view of Lemma 3.8, select $\delta > 0$ so that for all $(y, w) \in \mathbf{T}^n \times \mathbf{R}^n$, if $|y - \gamma(c)| \le \delta$, then

$$L(y, w) \ge L(\gamma(c), \dot\gamma(c)) + L_v(\gamma(c), \dot\gamma(c)) \cdot (w - \dot\gamma(c)) - \varepsilon.$$

4. In view of Lemma 3.6 and that as $j \to \infty$,

$$\gamma_{k_j}(c) \to \gamma(c),$$

we may choose $J \in \mathbf{N}$ so that for all $t \in [c,\, c + J^{-1}]$ and $j \ge J$,

$$|\gamma_{k_j}(t) - \gamma(c)| \le \delta.$$

Fix $j, m \in \mathbf{N}$ so that $j \ge J$ and $m \ge J$. We have

$$L(\gamma_{k_j}(t), \dot\gamma_{k_j}(t)) \ge L(\gamma(c), \dot\gamma(c)) + L_v(\gamma(c), \dot\gamma(c)) \cdot (\dot\gamma_{k_j}(t) - \dot\gamma(c)) - \varepsilon$$

for a.e. $t \in [c,\, c + m^{-1}]$. Integrating this over $[c,\, c + m^{-1}]$ and multiplying the resulting inequality by m, we get

$$\begin{aligned} m \int_c^{c+m^{-1}} & L(\gamma_{k_j}(t), \dot\gamma_{k_j}(t))\, \mathrm{d}t \ge L(\gamma(c), \dot\gamma(c)) \\ & + L_v(\gamma(c), \dot\gamma(c)) \cdot (m \int_c^{c+m^{-1}} \dot\gamma_{k_j}(t)\, \mathrm{d}t - \dot\gamma(c)) - \varepsilon \\ = {} & L(\gamma(c), \dot\gamma(c)) + L_v(\gamma(c), \dot\gamma(c)) \cdot [m(\gamma_k(c + m^{-1}) - \gamma_{k_j}(c)) - \dot\gamma(c)] - \varepsilon. \end{aligned}$$

By the definition of β_k, we get

$$\begin{aligned} & m(\beta_{k_j}(c + m^{-1}) - \beta_{k_j}(c)) \\ & \ge L(\gamma(c), \dot\gamma(c)) + L_v(\gamma(c), \dot\gamma(c)) \cdot [m(\gamma_{k_j}(c + m^{-1}) - \gamma_{k_j}(c)) - \dot\gamma(c)] - \varepsilon. \end{aligned}$$

Sending $j \to \infty$, we have

$$\begin{aligned} & m(\beta(c + m^{-1}) - \beta(c)) \\ & \ge L(\gamma(c), \dot\gamma(c)) + L_v(\gamma(c), \dot\gamma(c)) \cdot [m(\gamma(c + m^{-1}) - \gamma(c)) - \dot\gamma(c)] - \varepsilon. \end{aligned}$$

Next, sending $m \to \infty$ yields

$$\dot\beta(c) \geq L(\gamma(c), \dot\gamma(c)) + L_v(\gamma(c), \dot\gamma(c)) \cdot [\dot\gamma(c) - \dot\gamma(c)] - \varepsilon = L(\gamma(c), \dot\gamma(c)) - \varepsilon.$$

From this we conclude that for every point $t \in [0, T)$ of differentiability of (β, γ), we have

$$\dot\beta(t) \geq L(\gamma(t), \dot\gamma(t)). \tag{3.7}$$

5. Integrating both sides of (3.7), we get

$$\int_0^T L(\gamma(t), \dot\gamma(t))\, \mathrm{d}t \leq \int_0^T \dot\beta(t)\, \mathrm{d}t \leq \beta(T) - \beta(0) = \beta(T).$$

Notice that for any non-decreasing function g on $[0, T]$, we have

$$\int_0^T \dot g(t)\, \mathrm{d}t \leq g(T) - g(0).$$

Since

$$\beta(T) = \lim_{j\to\infty} \beta_{k_j}(T) = \liminf_{k\to\infty} \int_0^T L(\gamma_k(t), \dot\gamma_k(t))\, \mathrm{d}t,$$

we have

$$\int_0^T L(\gamma(t), \dot\gamma(t))\, \mathrm{d}t \leq \liminf_{k\to\infty} \int_0^T L(\gamma_k(t), \dot\gamma_k(t))\, \mathrm{d}t.$$

This and (3.6) together complete the proof. □

The following lemma will be useful later.

Lemma 3.8. *There is a constant $C_2 > 0$, depending only on T and L, such that for any $x_1, x_2 \in \mathbf{T}^n$ and any minimizer $\gamma \in \mathrm{AC}([0,T], \mathbf{T}^n)$ for $V(x_1)$,*

$$\operatorname*{ess\,inf}_{t\in[0,\,T]} |\dot\gamma(t)| \leq C_2.$$

Proof. As before there are constants $C_0 > 0$ and $C_1 > 0$, which depend only on T and L, such that

$$V(x_1) \leq C_0,$$
$$L(x, v) \geq |v| - C_1 \quad \forall (x, v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Since γ is a minimizer for $V(x_1)$, we have

$$\int_0^T |\dot{\gamma}(t)|\,\mathrm{d}t \le C_0 + C_1 T.$$

Hence,

$$\operatorname*{ess\,inf}_{t\in[0,\,T]} |\dot{\gamma}(t)| \le C_0 T^{-1} + C_1. \qquad \square$$

4. Regularity of minimizers

Let $\gamma \in \mathrm{AC}([0,T],\mathbf{T}^n)$ be a minimizer for V given by (3.1).

1. The minimizer $\gamma \in \mathrm{AC}([0,\,T],\mathbf{T}^n)$ is a.e. differentiable. Fix $t_0 \in (0,\,T)$, where γ is differentiable. Choose a constant $C > 0$ so that $|\dot{\gamma}(t_0)| < C$ and a constant $\delta > 0$ so that $[t_0-\delta,\,t_0+\delta] \subset [0,\,T]$ and

$$|\gamma(t)-\gamma(t_0)| \le C|t-t_0| \quad \forall t \in [t_0-\delta,\,t_0+\delta].$$

2. Due to ODE theory and the implicit function theorem, there exists a constant $\delta_1 \in (0,\,\delta]$ such that

$$\pi \circ \phi_t^L(\{\gamma(t_0)\}\times B(0,2C)) \supset B(\gamma(t_0),C|t|) \quad \forall t \in [-\delta_1,\,\delta_1].$$

3. For each $v \in B(0,2C)$ let $p = L_v(\gamma(t_0),v)$ and choose $\psi_v \in C^2(\mathbf{T}^n)$ so that $D\psi_v(\gamma(t_0)) = p$. We can choose the family $\{\psi_v\}_{v\in B(0,2C)}$ so that it is bounded in $C^2(\mathbf{T}^n)$. According to the method of characteristics (see, e.g., [L]), there exists a constant $\delta_2 \in (0,\,\delta_1]$ and for each $v \in B(0,2C)$ a function $S^v \in C(\mathbf{T}^n \times [t_0-\delta_2,\,t_0+\delta_2])$ such that

$$\begin{aligned} S^v(x,t_0) &= \psi_v(x) \quad \forall x \in \mathbf{T}^n,\\ S^v_t(x,t) + H(x,S^v_x(x,t)) &= 0 \quad \forall (x,t) \in \mathbf{T}^n \times [t_0-\delta_2,\,t_0+\delta_2]. \end{aligned}$$

4. Fix any $\tau \in (0,\,\delta_2]$ and set $t_1 = t_0+\tau$, $y_0 = \gamma(t_0)$, and $y_1 = \gamma(t_1)$. Choose $v \in B(0,2C)$ so that if $\mu(t) = \pi\circ\phi^L_{t-t_0}(\gamma(t_0),v)$, then $\mu(t_1) = \gamma(t_1)$. Observe that for any $\nu \in \mathrm{AC}([t_0,\,t_1],\mathbf{T}^n)$ such that $\nu(t_0) = y_0$ and $\nu(t_1) = y_1$, since

$$S^v_x(\nu(t),t)\dot{\nu}(t) \le H(\nu(t),S^v_x(\nu(t),t)) + L(\nu(t),\dot{\nu}(t)),$$

we have

$$
\begin{aligned}
S^v(y_1,t_1)-S^v(y_0,t_0) &= S^v(\nu(t_1),t_1)-S^v(\nu(t_0),t_0)\\
&= \int_{t_0}^{t_1} (S^v_t(\nu(t),t)+S^v_x(\nu(t),t)\dot\nu(t))\ \mathrm{d}t\\
&\le \int_{t_0}^{t_1} (S^v_t(\nu(t),t)+H(\nu(t),S^v_x(\nu(t),t))+L(\nu(t),\dot\nu(t)))\ \mathrm{d}t\\
&= \int_{t_0}^{t_1} L(\nu(t),\dot\nu(t))\,\mathrm{d}t.
\end{aligned}
$$

Observe as well that, since

$$
\begin{aligned}
S^v_x(\mu(t),t)\dot\mu(t) &= S^v_x(\mu(t),t)H_p(\mu(t),S^v_x(\mu(t),t))\\
&= H(\mu(t),S^v_x(\mu(t),t))+L(\mu(t),\dot\mu(t))
\end{aligned}
$$

by Proposition 1.2, (a), we have

$$
\begin{aligned}
S^v(y_1,t_1)-S^v(y_0,t_0) &= S^v(\mu(t_1),t_1)-S^v(\mu(t_0),t_0)\\
&= \int_{t_0}^{t_1} (S^v_t(\mu(t),t)+S^v_x(\mu(t),t)\dot\mu(t))\ \mathrm{d}t\\
&= \int_{t_0}^{t_1} (S^v_t(\mu(t),t)+H(\mu(t),S^v_x(\mu(t),t))+L(\mu(t),\dot\mu(t)))\ \mathrm{d}t\\
&= \int_{t_0}^{t_1} L(\mu(t),\dot\mu(t))\,\mathrm{d}t.
\end{aligned}
$$

These observations show that for any $\nu\in \mathrm{AC}([t_0,\,t_1],\mathbf{T}^n)$ such that $\nu(t_0)=y_0$ and $\nu(t_1)=y_1$, if $\nu\neq\mu$, then

$$
\int_{t_0}^{t_1} L(\mu(t),\dot\mu(t))\,\mathrm{d}t < \int_{t_0}^{t_1} L(\nu(t),\dot\nu(t))\,\mathrm{d}t.
$$

Consequently, we find that $\mu(t)=\gamma(t)$ for all $t\in[t_0,t_1]$ and hence $\mu(t)=\gamma(t)$ for all $t\in[0,T]$. Since $\mu\in C^2(\mathbf{R})$, we conclude that $\gamma\in C^2([0,T])$. Thus we have

Theorem 4.1. *Let $\gamma\in \mathrm{AC}([0,T],\mathbf{T}^n)$ be a minimizer for V defined by* (3.1). *Then $\gamma\in C^2([0,T])$.*

5. Weak KAM theorem

The weak KAM theorem [Fa1] due to A. Fathi is now stated as

Theorem 5.1 (weak KAM theorem). *There are functions $u_-, u_+ \in \mathrm{Lip}(\mathbf{T}^n)$ and a constant $c_0 \in \mathbf{R}$ having the properties:*

(a) *For any $\gamma \in \mathrm{AC}([a,\, b], \mathbf{T}^n)$, where $a < b$,*

$$u_\pm(\gamma(b)) - u_\pm(\gamma(a)) \le \int_0^T L(\gamma(t), \dot\gamma(t))\, \mathrm{d}t + c_0(b-a).$$

(b) *For each $x \in \mathbf{T}^n$, there are functions $\gamma_- \in \mathrm{AC}((-\infty,\, 0], \mathbf{T}^n)$, $\gamma_+ \in \mathrm{AC}([0,\, \infty), \mathbf{T}^n)$ such that $\gamma_\pm(0) = x$ and*

$$u_-(\gamma_-(0)) - u_-(\gamma_-(-t)) = \int_{-t}^0 L(\gamma_-(t), \dot\gamma_-(t))\, \mathrm{d}s + c_0 t \quad \forall t > 0,$$

and

$$u_+(\gamma_+(t)) - u_+(\gamma_+(0)) = \int_0^t L(\gamma_+(s), \dot\gamma_+(s))\, \mathrm{d}s + c_0 t \quad \forall t > 0.$$

For each $t > 0$ and $\phi \in C(\mathbf{T}^n)$ we introduce $T_t^- \phi : \mathbf{T}^n \to \mathbf{R}$ by

$$T_t^-\phi(x) = \inf_{\gamma(t)=x} \left[\int_0^t L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + \phi(\gamma(0)) \right],$$

where the infimum is taken over all $\gamma \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$ such that $\gamma(t) = x$. Similarly we define $T_t^+\phi : \mathbf{T}^n \to \mathbf{R}$ by

$$T_t^+\phi(x) = \sup_{\gamma(0)=x} \left[-\int_0^t L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + \phi(\gamma(t)) \right],$$

where the infimum is taken over all $\gamma \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$ such that $\gamma(0) = x$.

With this notation, as we will see later, Theorem 5.1 is equivalent to the following theorem.

Theorem 5.2. *There are functions $u_-, u_+ \in \mathrm{Lip}(\mathbf{T}^n)$ and a constant $c_0 \in \mathbf{R}$ such that*

$$u_-(x) = T_t^- u_-(x) + c_0 t \quad \forall t > 0,\ x \in \mathbf{T}^n, \tag{5.1}$$

and

$$u_+(x) = T_t^+ u_+(x) - c_0 t \quad \forall t > 0,\ x \in \mathbf{T}^n. \tag{5.2}$$

In the rest of this section we are mostly devoted to proving a weaker form of Theorem 5.2. That is, we prove the following proposition.

Theorem 5.3. *There are functions u_-, $u_+ \in \mathrm{Lip}(\mathbf{T}^n)$ and constants c_0, $d_0 \in \mathbf{R}$ such that*

$$u_-(x) = T_t^- u_-(x) + c_0 t \quad \forall t > 0,\ x \in \mathbf{T}^n,$$

and

$$u_+(x) = T_t^+ u_+(x) - d_0 t \quad \forall t > 0,\ x \in \mathbf{T}^n. \tag{5.3}$$

We postpone until next section to prove that $d_0 = c_0$, and in this section, assuming that $d_0 = c_0$, we prove that Theorem 5.3 is equivalent to Theorem 5.1.

Define the $\hat{L} : \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{R}$ by $\hat{L}(x, v) = L(x, -v)$. Fix $\lambda \in (0, 1)$. Define v^λ on $\mathbf{T}^n$ by

$$v^\lambda(x) = \inf_{\gamma(0)=x} \int_0^\infty e^{-\lambda t} \hat{L}(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t.$$

Lemma 5.4. $\min_{\mathbf{T}^n \times \mathbf{R}^n} \hat{L} \leq \lambda v^\lambda(x) \leq \hat{L}(x, 0)$ *for $x \in \mathbf{T}^n$.*

Proof. Set $C = \min_{\mathbf{T}^n \times \mathbf{R}^n} L$. Fix $x \in \mathbf{T}^n$. For any $\gamma \in \mathrm{AC}([0, \infty), \mathbf{T}^n)$ satisfying $\gamma(0) = x$, we have

$$\int_0^\infty e^{-\lambda t} \hat{L}(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t \geq \int_0^\infty e^{-\lambda t} C\, \mathrm{d}t = C\lambda^{-1}.$$

Hence,

$$\lambda v^\lambda(x) \geq C.$$

Also, we have

$$v^\lambda(x) \leq \int_0^\infty e^{-\lambda t} \hat{L}(x, 0)\, \mathrm{d}t = \hat{L}(x, 0)\lambda^{-1}.$$

Thus we get

$$C \leq \lambda v^\lambda(x) \leq \hat{L}(x, 0). \qquad \square$$

Lemma 5.5 (Dynamic programming principle). *For any $T > 0$ and $x \in \mathbf{T}^n$, we have*

$$v^\lambda(x) = \inf_{\gamma(0)=x} \left[\int_0^T e^{-\lambda t} \hat{L}(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t + e^{-\lambda T} v^\lambda(\gamma(T)) \right].$$

Proof. We denote by $w(x)$ the right hand side of the above formula. Fix $x \in \mathbf{T}^n$. Fix any $\gamma \in \mathrm{AC}([0, \infty), \mathbf{T}^n)$ such that $\gamma(0) = x$. Note that

$$\begin{aligned}
&\int_0^\infty e^{-\lambda t} \hat{L}(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t \\
&= \int_0^T e^{-\lambda t} \hat{L}(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t + e^{-\lambda T} \int_0^\infty e^{-\lambda t} \hat{L}(\gamma(t+T), \dot{\gamma}(t+T))\, \mathrm{d}t \\
&\geq \int_0^T e^{-\lambda t} \hat{L}(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t + e^{-\lambda T} v^\lambda(\gamma(T)) \geq w(x).
\end{aligned}$$

Hence we have $v^\lambda(x) \geq w(x)$.

Fix any $\gamma \in \mathrm{AC}([0,\infty),\mathbf{T}^n)$ such that $\gamma(0)=x$, and then any $\mu \in \mathrm{AC}([0,\infty),\mathbf{T}^n)$ such that $\mu(0)=\gamma(T)$. Define $\nu \in \mathrm{AC}([0,\infty),\mathbf{T}^n)$ by

$$\nu(t) = \begin{cases} \gamma(t) & (0 \leq t < T), \\ \mu(t-T) & (T \leq t). \end{cases}$$

Then we have

$$\begin{aligned} v^\lambda(x) &\leq \int_0^T e^{-\lambda t}\hat{L}(\nu(t),\dot{\nu}(t))\,\mathrm{d}t + e^{-\lambda T}\int_0^\infty e^{-\lambda t}\hat{L}(\nu(t+T),\dot{\nu}(t+T))\,\mathrm{d}t \\ &\leq \int_0^T e^{-\lambda t}\hat{L}(\gamma(t),\dot{\gamma}(t))\,\mathrm{d}t + e^{-\lambda T}\int_0^\infty e^{-\lambda t}\hat{L}(\mu(t),\dot{\mu}(t))\,\mathrm{d}t. \end{aligned}$$

Consequently, we have

$$v^\lambda(x) \leq \int_0^T e^{-\lambda t}\hat{L}(\gamma(t),\dot{\gamma}(t))\,\mathrm{d}t + e^{-\lambda T}v^\lambda(\gamma(T)).$$

From this we find that

$$v^\lambda(x) \leq w(x). \qquad \square$$

Lemma 5.6. *The functions v^λ, with $\lambda \in (0,\,1)$, are equi-Lipschitz continuous on $\mathbf{T}^n$.*

Proof. Set

$$C = \min_{\mathbf{T}^n\times\mathbf{R}^n} L,$$

and note that

$$v^\lambda(x) - \lambda^{-1}C = v^\lambda(x) - C\int_0^\infty e^{-\lambda t}\,\mathrm{d}t = \inf_\gamma \int_0^\infty e^{-\lambda t}(\hat{L}(\gamma(t),\dot{\gamma}(t)) - C)\,\mathrm{d}t.$$

Thus, by replacing $v^\lambda(x)$ and $\hat{L}$, respectively, by $v^\lambda(x)-\lambda^{-1}C$ and $\hat{L}-C$ if necessary, we may assume that $\hat{L}\geq 0$ on $\mathbf{T}^n\times\mathbf{R}^n$, so that $v^\lambda \geq 0$ on $\mathbf{T}^n$.

Fix $x,\ y \in \mathbf{T}^n$. Assume that $x\neq y$. By Lemma 5.5, for any $\gamma \in \mathrm{AC}([0,\,|x-y|],\mathbf{T}^n)$ with $\gamma(0)=y$, we have

$$v^\lambda(y) \leq \int_0^{|x-y|} e^{-\lambda t}L(\gamma(t),\dot{\gamma}(t))\,\mathrm{d}t + e^{-\lambda|x-y|}v^\lambda(\gamma(|x-y|)).$$

Define $\mu \in \mathrm{AC}([0,\,|x-y|],\mathbf{T}^n)$ by

$$\mu(t) = y + t|x-y|^{-1}(x-y).$$

Then we get

$$\begin{aligned}
v^\lambda(y) &\le \int_0^{|x-y|} e^{-\lambda t}\hat{L}(\mu(t),\dot{\mu}(t))\,\mathrm{d}t + e^{-|x-y|}v^\lambda(x) \\
&\le \int_0^{|x-y|} \hat{L}(\mu(t),|x-y|^{-1}(x-y))\,\mathrm{d}t + v^\lambda(x) \\
&\le C_1\int_0^{|x-y|} \mathrm{d}t + v^\lambda(x) = v^\lambda(x) + C_1|x-y|.
\end{aligned}$$

Here C_1 is a positive constant such that

$$\max_{\mathbf{T}^n\times B(0,1)} L \le C_1.$$

Thus we get

$$v^\lambda(y) - v^\lambda(x) \le C_1|x-y| \quad \forall x,y\in\mathbf{T}^n,$$

and conclude that

$$|v^\lambda(x) - v^\lambda(y)| \le C_1|x-y| \quad \forall x,y\in\mathbf{T}^n. \qquad \square$$

Proof of Theorem 5.3. 1. For $T>0$ and $\phi\in C(\mathbf{T}^n)$ we define $Q_T^-\phi:\mathbf{T}^n\to\mathbf{R}$ by

$$Q_T^-\phi(x) = \inf_{\gamma(0)=x}\left[\int_0^T \hat{L}(\gamma(t),\dot{\gamma}(t))\,\mathrm{d}t + \phi(\gamma(T))\right],$$

where the infimum is taken over all $\gamma\in\mathrm{AC}([0,\,T],\mathbf{T}^n)$ such that $\gamma(0)=x$. We show that there exist a function $u_-\in\mathrm{Lip}(\mathbf{T}^n)$ and a constant $c_0\in\mathbf{R}$ such that

$$u_-(x) = Q_T^-u_-(x) + c_0T \quad \forall T>0. \tag{5.4}$$

2. By Lemma 5.4, the collection $\{\lambda v^\lambda(0)\mid \lambda\in(0,\,1)\}$ is bounded. Therefore we can select a sequence $\{\lambda_j\}_{j\in\mathbf{N}}\subset(0,\,1)$ so that, as $j\to\infty$,

$$\lambda_j\to 0 \quad\text{and}\quad \lambda_j v^{\lambda_j}(0)\to -c_0$$

for some $c_0\in\mathbf{R}$.

3. Set $w^\lambda(x) = v^\lambda(x) - v^\lambda(0)$ for $x\in\mathbf{T}^n$. The collection $\{w^\lambda\mid\lambda\in(0,\,1)\}\subset\mathrm{Lip}(\mathbf{T}^n)$ is uniformly bounded and equi-Lipschitz on $\mathbf{T}^n$. Hence, we may assume that, as $j\to\infty$,

$$w^{\lambda_j}(x)\to w(x) \quad \text{uniformly on } \mathbf{T}^n$$

for some $w\in\mathrm{Lip}(\mathbf{T}^n)$.

4. Fix $x \in \mathbf{T}^n$ and $T > 0$. Using Lemma 5.5, we get

$$w^\lambda(x) = \inf_{\gamma(0)=x} \left[\int_0^T e^{-\lambda t} \hat{L}(\gamma(t), \dot\gamma(t))\, \mathrm{d}t + e^{-\lambda T} w^\lambda(\gamma(T)) \right] + \left(e^{-\lambda T} - 1\right) v^\lambda(0). \tag{5.5}$$

Fix any $\gamma \in \mathrm{AC}([0,\infty), \mathbf{T}^n)$ so that $\gamma(0) = x$. From the above, we have

$$w^\lambda(x) \le \int_0^T e^{-\lambda t} \hat{L}(\gamma(t), \dot\gamma(t))\, \mathrm{d}t + e^{-\lambda T} w^\lambda(\gamma(T)) - T \frac{e^{-\lambda T} - 1}{-\lambda T} \lambda v^\lambda(0)$$

Passing to the limit along the sequence $\lambda = \lambda_j$ as $j \to \infty$, we get

$$\begin{aligned} w(x) &\le \int_0^T \hat{L}(\gamma(t), \dot\gamma(t))\, \mathrm{d}t + w(\gamma(T)) - T \cdot 1 \cdot (-c_0) \\ &= \int_0^T \hat{L}(\gamma(t), \dot\gamma(t))\, \mathrm{d}t + w(\gamma(T)) + c_0 T. \end{aligned}$$

Thus we have

$$w(x) \le Q_T^- w(x) + c_0 T \quad \forall x \in \mathbf{T}^n,\ T > 0. \tag{5.6}$$

5. In view of (5.5), we choose $\gamma_\lambda \in \mathrm{AC}([0,\infty), \mathbf{T}^n)$, with $\gamma_\lambda(0) = x$, so that

$$w^\lambda(x) + \lambda > \int_0^T e^{-\lambda t} \hat{L}(\gamma_\lambda(t), \dot\gamma_\lambda(t))\, \mathrm{d}t + e^{-\lambda T} w^\lambda(\gamma_\lambda(T)) + (e^{-\lambda T} - 1) v^\lambda(0). \tag{5.7}$$

We rewrite this as

$$\begin{aligned} w^\lambda(x) + \lambda &> e^{-\lambda T} \Big(\int_0^T \hat{L}(\gamma_\lambda(t), \dot\gamma_\lambda(t))\, \mathrm{d}t + w(\gamma_\lambda(T)) \Big) \\ &\quad + (e^{-\lambda T} - 1) v^\lambda(0) + e_\lambda \\ &\ge e^{-\lambda T} Q_T^- w(x) + (e^{-\lambda T} - 1) v^\lambda(0) + e_\lambda, \end{aligned} \tag{5.8}$$

where

$$e_\lambda = \int_0^T (e^{-\lambda t} - e^{-\lambda T}) \hat{L}(\gamma_\lambda(t), \dot\gamma_\lambda(t))\, \mathrm{d}t + e^{-\lambda T} \left[w^\lambda(\gamma_\lambda(T)) - w(\gamma_\lambda(T)) \right].$$

6. Noting that there is a constant $C_2 > 0$ such that

$$\hat{L}(y, v) \ge -C_2 \quad \forall (y, v) \in \mathbf{T}^n \times \mathbf{R}^n,$$

we have

$$\int_0^T (e^{-\lambda t} - e^{-\lambda T})\hat{L}(\gamma_\lambda(t), \dot\gamma_\lambda(t))\,\mathrm{d}t \ge -C_2 \int_0^T (e^{-\lambda t} - e^{-\lambda T})\,\mathrm{d}t \ge -C_2 T(1 - e^{-\lambda T}).$$

Consequently, we have

$$e_\lambda \ge -C_2 T(1 - e^{-\lambda T}) - e^{-\lambda T} \max_{\mathbf{T}^n} |w^\lambda - w|. \tag{5.9}$$

7. From (5.8) and (5.9), we get

$$w^\lambda(x) + \lambda > e^{-\lambda T} Q_T^- w(x) + (e^{-\lambda T} - 1)v^\lambda(0) - C_2 T(1 - e^{-\lambda T}) - \max_{\mathbf{T}^n} |w^\lambda - w|.$$

Sending $\lambda \to 0$ along the sequence $\lambda = \lambda_j$, we now find that

$$w(x) \ge Q_T^- w(x) + c_0 T.$$

This together with (5.6) yields

$$w(x) = Q_T^- w(x) + c_0 T \quad \forall x \in \mathbf{T}^n,\ T > 0.$$

8. Let $(x, t) \in \mathbf{T}^n \times (0, \infty)$. From the above identity we get

$$\begin{aligned} w(x) &= Q_t^- w(x) + c_0 t = \inf_{\gamma(0)=x} \left[\int_0^t \hat{L}(\gamma(s), \dot\gamma(s))\,\mathrm{d}s + w(\gamma(t))\right] + c_0 t \\ &= \inf_{\mu(t)=x} \left[\int_0^t \hat{L}(\mu(s), -\dot\mu(s))\,\mathrm{d}s + w(\mu(0))\right] + c_0 t \\ &= \inf_{\mu(t)=x} \left[\int_0^t L(\mu(s), \dot\mu(s))\,\mathrm{d}s + w(\mu(0))\right] + c_0 t \\ &= T_t^- w(x) + c_0 t. \end{aligned}$$

Here we used the observation that for $\gamma \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$, with $\gamma(t) = x$, if we set $\mu(s) = \gamma(t - s)$ for $s \in [0,\, t]$, then $\mu \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$ and $\mu(0) = x$.

Thus we find that the pair $(w, c_0) \in \mathrm{Lip}(\mathbf{T}^n) \times \mathbf{R}$ has the required properties for (u_-, c_0) in Theorem 5.3.

9. We repeat the arguments in the paragraphs 1 to 7 above with L in place of $\hat{L}$, to conclude that there is a function $v \in \mathrm{Lip}(\mathbf{T}^n)$ and a constant $d_0 \in \mathbf{R}$ such that

$$v(x) = \inf_{\gamma(0)=x} \left[\int_0^t L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s + v(\gamma(t))\right] + d_0 t \quad \forall t > 0,$$

where the infimum is taken over all $\gamma \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$ such that $\gamma(0) = x$. Multiplying this by -1 and writing $u = -v$, we find that

$$u(x) = \sup_{\gamma(0)=x} \left[-\int_0^t L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + u(\gamma(t)) \right] - d_0 t \quad \forall t > 0,$$

which shows that the pair (u, d_0) has the properties required for (u_+, d_0) in Theorem 5.3. $\square$

Now, we turn to the proof of the equivalence of Theorems 5.1 and 5.2.

Proof of Theorem 5.1 from Theorem 5.2. Let u_-, u_+, c_0 be those from Theorem 5.2.

1. Fix any $\gamma \in \mathrm{AC}([a,\, b], \mathbf{T}^n)$, with $a < b$. Define $\mu \in \mathrm{AC}([0,\, b-a], \mathbf{T}^n)$ by $\mu(s) = \gamma(s + a)$. Since

$$u_-(x) = T_{b-a} u_-(x) + c_0(b - a) \quad \forall x \in \mathbf{T}^n,$$

we get

$$u_-(\mu(b-a)) \le \int_0^{b-a} L(\mu(s), \dot\mu(s))\, \mathrm{d}s + u_-(\mu(0)) + c_0(b-a),$$

and hence

$$u(\gamma(b)) - u_-(\gamma(a)) \le \int_a^b L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + c_0(b-a).$$

Similarly, we get

$$u_+(\mu(0)) \ge -\int_0^{b-a} L(\mu(s), \dot\mu(s))\, \mathrm{d}s + u_+(\mu(b-a)) - c_0(b-a)$$

and

$$u_+(\gamma(a)) = -\int_a^b L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + u_+(\gamma(b)) - c_0(b-a).$$

from which we find that

$$u_+(\gamma(b)) - u_+(\gamma(a)) \le \int_a^b L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + c_0(b-a).$$

Thus assertion (a) has been shown.

2. To prove (b), fix $x \in \mathbf{T}^n$. We construct $\gamma_- \in \mathrm{AC}((-\infty,\, 0], \mathbf{T}^n)$ as follows. First note that

$$u_-(y) = \inf_{\gamma(t)=y} \left[\int_{t-1}^t L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + u_-(\gamma(t-1)) \right] + c_0 \quad \forall y \in \mathbf{T}^n.$$

Define the sequence $\gamma_k \in C^2([-k,\, -k+1], \mathbf{T}^n)$, $k \in \mathbf{N}$ by selecting γ_k inductively. We first select γ_1 so that

$$\begin{aligned} &\gamma_1(0) = x, \\ &u_-(x) = \int_{-1}^{0} L(\gamma_1(s), \dot\gamma_1(s))\,\mathrm{d}s + u_-(\gamma_1(-1)) + c_0. \end{aligned}$$

For $k > 1$ we select γ_k so that

$$\begin{aligned} &\gamma_k(-k+1) = \gamma_{k-1}(-k+1), \\ &u_-(\gamma_k(-k+1)) = \int_{-k}^{-k+1} L(\gamma_k(s), \dot\gamma_k(s))\,\mathrm{d}s + u_-(\gamma_k(-k)) + c_0. \end{aligned}$$

Indeed, according to Theorem 3.1, such γ_k, with $k \in \mathbf{N}$, exist. Define $\gamma_- \in \mathrm{AC}((-\infty,\, 0], \mathbf{T}^n)$ by setting

$$\gamma_-(s) = \gamma_k(s) \quad \text{ for } s \in [-k, -k+1],\ k \in \mathbf{N}.$$

3. We have

$$u_-(x) = \int_{-k}^{0} L(\gamma_-(s), \dot\gamma_-(s))\,\mathrm{d}s + u_-(\gamma_-(-k)).$$

This and (a) guarantee that γ_- is a minimizer for

$$\inf_\gamma \int_a^b L(\gamma(s), \dot\gamma(s)),$$

with any $-\infty < a < b \le 0$, where the infimum is taken over all $\gamma \in \mathrm{AC}([a,\, b], \mathbf{T}^n)$ such that $\gamma(a) = \gamma_-(a)$ and $\gamma(b) = \gamma_-(b)$. This shows that

$$u_-(x) = \int_{-t}^{0} L(\gamma_-(s), \dot\gamma_-(s))\,\mathrm{d}s + u_-(\gamma_-(-t)) \quad \forall t > 0$$

and that $\gamma_- \in C^2((-\infty,\, 0])$ by Theorem 4.1.

4. Fix $x \in \mathbf{T}^n$. We can select $\gamma_1 \in C^2([0,\, 1])$ so that

$$\begin{aligned} &\gamma_1(0) = x, \\ &u_+(x) = -\int_{0}^{1} L(\gamma_1(s), \dot\gamma_1(s))\,\mathrm{d}s + u_+(\gamma_1(1)) - c_0. \end{aligned}$$

Next, we can choose $\gamma_k \in C^2([k-1,\, k])$ inductively for $k > 1$ so that

$$\begin{aligned} &\gamma_k(k-1) = \gamma_{k-1}(k-1), \\ &u_+(\gamma_k(k-1)) = -\int_{k-1}^{k} L(\gamma_k(s), \dot\gamma_k(s))\,\mathrm{d}s + u_+(\gamma_k(k)) - c_0. \end{aligned}$$

Setting

$$\gamma_+(s) = \gamma_k(s) \quad \text{for } s \in [k-1,\, k],\ k \in \mathbf{N},$$

we find a $\gamma_+ \in C^2([0,\, \infty))$ such that $\gamma_+(0) = x$ and

$$u_+(\gamma_+(t)) = u_+(x) + \int_0^t L(\gamma_+(s), \dot\gamma_+(s))\,\mathrm{d}s + c_0 t \quad \forall t > 0.$$

The proof is now complete. □

Proof of Theorem 5.2 from Theorem 5.1. 1. Let u_-, u_+ c_0 be those from Theorem 5.1. We show that

$$u_-(x) = T_t^- u_-(x) + c_0 t \quad \forall t > 0, \tag{5.10}$$

$$u_+(x) = T_t^+ u_+(x) - c_0 t \quad \forall t > 0. \tag{5.11}$$

We only prove (5.10). The proof of (5.11) can be done in a parallel way.

2. Fix any $x \in \mathbf{T}^n$ and $t > 0$. Let $\gamma \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$ be such that $\gamma(t) = x$. By Theorem 5.1, (a), we have

$$u_-(x) \le \int_0^t L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s + u_-(\gamma(0)) + c_0 t.$$

Hence we have

$$u_-(x) \le T_t^- u_-(x) + c_0 t. \tag{5.12}$$

Let $\gamma_- \in C^2((-\infty,\, 0], \mathbf{T}^n)$ be the one from Theorem 5.1, (b). Setting $\mu(s) = \gamma_-(s-t)$ for $s \in [0,\, t]$ and noting that $\mu(t) = x$, we have

$$\begin{aligned} u_-(x) &= \int_{-t}^0 L(\gamma_-(s), \dot\gamma_-(s))\,\mathrm{d}s + u_-(\gamma_-(-t)) + c_0 t \\ &= \int_0^t L(\mu(s), \dot\mu(s))\,\mathrm{d}s + u_-(\mu(0)) + c_0 t \ge T_t^- u_-(x) + c_0 t. \end{aligned}$$

This together with (5.12) proves (5.10). □

6. A PDE approach

We consider a general scalar first order partial differential equation

$$F(x, u(x), Du(x)) = 0 \quad \text{in } \Omega, \tag{6.1}$$

where Ω is an open subset of $\mathbf{R}^n$ and Du denotes the gradient of $u : \Omega \to \mathbf{R}$. We assume that F is continuous on $\Omega \times \mathbf{R} \times \mathbf{R}^n$.

A lower semicontinuous function $u : \Omega \to \mathbf{R}$ is called a *viscosity supersolution* of (6.1) if for any $(\psi, x) \in C^1(\Omega) \times \Omega$ such that $(u - \psi)(x) = \min_\Omega (u - \psi)$,

$$F(x, u(x), D\psi(x)) \geq 0.$$

An upper semicontinuous function $u : \Omega \to \mathbf{R}$ is called a *viscosity subsolution* of (6.1) if for any $(\psi, x) \in C^1(\Omega) \times \Omega$ such that $(u - \psi)(x) = \max_\Omega (u - \psi)$,

$$F(x, u(x), D\psi(x)) \leq 0.$$

A continuous function $u : \Omega \to \mathbf{R}$ is called a *viscosity solution* of (6.1) if it is both a viscosity supersolution and a viscosity subsolution of (6.1).

Note that u is a viscosity supersolution (resp., subsolution) of (6.1) if and only if $v := -u$ is a viscosity subsolution (resp., supersolution) of

$$-F(x, -v(x), -Dv(x)) = 0 \quad \text{ in } \Omega.$$

We refer the reader to [CL, CEL, BC, B, L] for general references on viscosity solutions of first order PDE.

A first remark based on the PDE approach on the weak KAM theorem is the following.

Proposition 6.1. *Let $\phi \in C(\mathbf{T}^n)$ and define $u : \mathbf{T}^n \times [0, \infty) \to \mathbf{R}$ by $u(x, t) = T_t^- \phi(x)$. Then*

(a) *u is continuous on $\mathbf{T}^n \times [0, \infty)$;*

(b) *for each $t > 0$ there is a constant $C_t > 0$ such that*

$$|u(x, s) - u(y, s)| \leq C_t |x - y| \quad \forall x, y \in \mathbf{T}^n,\ s \geq t;$$

(c) *u is a viscosity solution of*

$$u_t(x, t) + H(x, u_x(x, t)) = 0 \quad \textit{in } \mathbf{T}^n \times (0, \infty). \tag{6.2}$$

Remark. To be precise, the definition of u for $t = 0$ should be understood as

$$u(x, 0) = T_0^- \phi(x) = \phi(x).$$

Lemma 6.2 (Dynamic programming principle). *For any $t \geq 0$, $s \geq 0$, $\phi \in C(\mathbf{T}^n)$, and $x \in \mathbf{T}^n$ we have*

$$T^-_{t+s}\phi(x) = T^-_t \circ T^-_s \phi(x).$$

The arguments in the proof of Lemma 5.5 apply to the proof of this lemma, which we omit to reproduce here.

Proof of Proposition 6.1. 1. We first show the continuity of u at $t = 0$. Since

$$\begin{aligned} u(x,t) &= \inf_{\gamma(t)=x} \left[\int_0^t L(\gamma(s), \dot\gamma(s))\, \mathrm{d}s + \phi(\gamma(0)) \right] \\ &\leq \int_0^t L(x, 0)\, \mathrm{d}s + \phi(x) \quad \forall (x,t) \in \mathbf{T}^n \times [0,\, \infty), \end{aligned}$$

we have

$$u(x,t) - \phi(x) \leq t \max_{x \in \mathbf{T}^n} L(x, 0) \quad \forall (x,t) \in \mathbf{T}^n \times [0,\, \infty). \tag{6.3}$$

2. Let ω_ϕ be the modulus of continuity of ϕ. That is,

$$\omega_\phi(r) = \sup\{|\phi(x) - \phi(y)| \mid x, y \in \mathbf{T}^n,\ |x-y| \leq r\} \quad \text{ for } r \geq 0.$$

Fix $(x,t) \in \mathbf{T}^n \times (0,\, t)$. Let $\gamma \in C^2([0,\, t])$ be a minimizer for

$$\inf_{\mu(t)=x} \left[\int_0^t L(\mu(s), \dot\mu(s))\, \mathrm{d}s + \phi(\mu(0)) \right].$$

For each $A \geq 1$ we choose a constant $C_A > 0$ so that

$$L(x,v) \geq A|v| - C_A \quad \forall (x,v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

We compute that

$$\begin{aligned} T^-_t \phi(x) - \phi(x) &\geq A \int_0^t |\dot\gamma(s)|\, \mathrm{d}s - tC_A + \phi(\gamma(0)) - \phi(x) \\ &\geq A|\gamma(t) - \gamma(0)| - tC_A - \omega_\phi(|\gamma(t) - \gamma(0)|). \end{aligned}$$

Define the function $\nu : [1,\, \infty) \to \mathbf{R}$ by

$$\nu(A) = \sup\{\omega_\phi(r) - Ar \mid r \geq 0\}.$$

Observe that ν is a non-increasing function on $[1, \infty)$ and $\nu(A) \geq \omega_\phi(0) = 0$ for all $A \geq 1$. Also, since $\omega_\phi(r)$ is bounded on $[0, \infty)$, we have

$$\omega_\phi(r) \leq C \quad \forall r \geq 0$$

for some constant $C > 0$, and hence

$$\nu(A) = \sup\{\omega_\phi(r) - Ar \mid 0 \leq r \leq A^{-1}C\} \leq \omega_\phi(A^{-1}C).$$

Note that

$$T_t^- \phi(x) - \phi(x) \geq -\nu(A) - tC_A \quad \forall A \geq 1.$$

Setting

$$\rho(s) = \inf_{A \geq 1} (\nu(A) + sC_A) \qquad \text{for } s \geq 0,$$

we get

$$T_t^- \phi(x) - \phi(x) \geq -\rho(t).$$

Observe that ρ is upper semicontinuous on $[0, \infty)$, $\rho(s) \geq 0$ for all $s \geq 0$, $\rho(s) \leq \nu(1) + sC_A$ for all $s \geq 0$, and

$$\rho(0) \leq \inf_{A \geq 1} \nu(A) \leq \inf_{A \geq 1} \omega_\phi(A^{-1}C) = 0.$$

This and (6.3) together show that there is a continuous function σ on $[0, \infty)$, with $\sigma(0) = 0$, such that

$$|u(x,t) - \phi(x)| \leq \sigma(t) \quad \forall (x,t) \in \mathbf{T}^n \times [0, \infty). \tag{6.4}$$

We may assume that $\sigma(t) \leq C_0(t+1)$ for all $t \geq 0$ and for some constant $C_0 > 0$. Finally note that ρ depends only on ω_ϕ and the family $\{C_A \mid A > 1\}$ and hence σ depends only on ω_ϕ, $\{C_A\}_{A>1}$, and $\max_{x \in \mathbf{T}^n} L(x, 0)$.

3. Next we prove (b). Let $C_0 > 0$ be a constant for which $\sigma(t) \leq C_0(t+1)$ for all $t \geq 0$. Choose $C_1 > 0$ so that

$$L(x, v) \geq |v| - C_1 \quad \forall (x, v) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Fix any $x, y \in \mathbf{T}^n$. Choose a minimizer $\gamma \in C^2([0, t])$ for

$$T_t^- \phi(x) = \inf_{\mu(t)=x} \left[\int_0^t L(\mu(s), \dot{\mu}(s))\, \mathrm{d}s + \phi(\mu(0)) \right].$$

Observing that

$$\int_0^t |\dot\gamma(s)|\,\mathrm{d}s \le tC_1 + \sigma(t) - \phi(\gamma(0)) \le tC_1 + C_0(t+1) + \max_{\mathbf{T}^n}|\phi|,$$

we find a $\tau \in [0,\, t]$ such that

$$t|\dot\gamma(\tau)| \le tC_1 + C_0(t+1) + \max_{\mathbf{T}^n}|\phi|.$$

Setting $C_1(r) = C_1 + C_0 + r^{-1}(C_0 + \max_{\mathbf{T}^n}|\phi|)$, we have $|\dot\gamma(\tau)| \le C_1(r)$.

4. In view of Proposition 2.1, (a), we have

$$H(\gamma(s), L_v(\gamma(s), \dot\gamma(s))) = H(\gamma(\tau), L_v(\gamma(\tau), \dot\gamma(\tau))) \quad \forall s \in [0,\, t].$$

Consequently,

$$H(\gamma(s), L_v(\gamma(s), \dot\gamma(s))) \le \max_{(x,v)\in\mathbf{T}^n\times B(0,C_1(r))} H(x, L_v(x,v)) \quad \forall s \in [0,\, t].$$

By the superlinearity of H, there exists a constant $C_2(r) > 0$ such that

$$|L_v(\gamma(s), \dot\gamma(s))| \le C_2(r) \quad \forall s \in [0,\, t].$$

Since $\dot\gamma(s) = H_p(\gamma(s), L_v(\gamma(s), \dot\gamma(s)))$ for all $s \in [0,\, t]$, we find a constant $C_3(r) > 0$ such that

$$|\dot\gamma(s)| \le C_t^3 \quad \forall s \in [0,\, t].$$

5. We define $\mu \in \mathrm{AC}([0,\, t], \mathbf{T}^n)$ by

$$\mu(s) = \begin{cases} \gamma(s) & \text{for } 0 \le s \le t-r, \\ \gamma(s) + \dfrac{s-t+r}{r}(y-x) & \text{for } t-r \le s \le t. \end{cases}$$

Noting that $\mu(0) = \gamma(0)$, $\mu(t) = \gamma(t) + y - x = y$, and $|\dot\mu(s)| \le |\dot\gamma(s)| + \frac{1}{r}|x-y| \le C_3(r) + \frac{\sqrt{n}}{r}$ and writing $C_4(r) = C_3(r) + \frac{\sqrt{n}}{r}$, we have

$$\begin{aligned} T_t^- \phi(y) &\le \int_0^t L(\mu(s), \dot\mu(s))\,\mathrm{d}s + \phi(\mu(0)) \\ &\le \int_0^t L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s + \phi(\gamma(0)) \\ &\quad + \left(\max_{\mathbf{T}^n\times B(0,C_4(r))} |L_x| r + \max_{\mathbf{T}^n\times B(0,C_4(r))} |L_v| \right) |x-y|. \end{aligned}$$

Hence we get

$$T_t^- \phi(y) - T_t^- \phi(x) \le \left(\max_{\mathbf{T}^n \times B(0,C_4(r))} |L_x| r + \max_{\mathbf{T}^n \times B(0,C_4(r))} |L_v| \right) |x-y|.$$

From this, setting

$$C_5(r) = \left(\max_{\mathbf{T}^n \times B(0,C_4(r))} |L_x| r + \max_{\mathbf{T}^n \times B(0,C_4(r))} |L_v| \right),$$

we conclude that

$$|u(x,t) - u(y,t)| \le C_5(r)|x-y| \quad \forall x, y \in \mathbf{T}^n,\ t \ge r, \tag{6.5}$$

and hence assertion (b).

6. From (6.4) we get

$$|u(x,t) - u(y,t)| \le |\phi(x) - \phi(y)| + 2\sigma(t) \le \omega_\phi(|x-y|) + 2\sigma(t) \quad \forall x, y \in \mathbf{T}^n,\ t \ge 0.$$

Fix any $\varepsilon > 0$, $x, y \in \mathbf{T}^n,\ t \ge 0$. From the above, if $|x-y| + \varepsilon \ge t$, then

$$|u(x,t) - u(y,t)| \le |\phi(x) - \phi(y)| + 2\sigma(t) \le \omega_\phi(|x-y|) + 2\sigma(|x-y| + \varepsilon).$$

On the other hand, by (6.5), if $|x-y| + \varepsilon < t$, then

$$|u(x,t) - u(y,t)| \le C_5(\varepsilon)|x-y|.$$

Combining these yields

$$|u(x,t) - u(y,t)| \le \omega_\phi(|x-y|) + 2\sigma(|x-y| + \varepsilon) + C_5(\varepsilon)|x-y|.$$

Define $\bar{\omega} : [0, \infty) \to \mathbf{R}$ by

$$\bar{\omega}(r) = \inf_{s>0} \left(\omega_\phi(r) + 2\sigma(r+s) + C_5(s) r \right).$$

We have then

$$|u(x,t) - u(y,t)| \le \bar{\omega}(|x-y|) \quad \forall x, y \in \mathbf{T}^n,\ t \ge 0. \tag{6.6}$$

Observe that $\bar{\omega}(r) \ge 0$ for all $r \ge 0$ and $\bar{\omega}(r) = \omega_\phi(r) + 2\sigma(r+s) + C_5(s)r$ for all $r \ge 0$ and $s > 0$ and hence that $\lim_{r \searrow 0} \bar{\omega}(r) = 0$. Therefore, (6.6) guarantees that the collection $\{u(\cdot, t) \mid t \ge 0\} \subset C(\mathbf{T}^n)$ is equi-continuous.

7. The above arguments 1 and 2 applied to $T_t^- \psi$, with $\psi = u(\cdot, s)$, where $t \geq 0$ and $s \geq 0$, yields a modulus $\bar{\sigma} \in C([0, \infty))$ such that

$$|T_t^- \circ T_s^- \phi(x) - T_s^- \phi(x)| \leq \bar{\sigma}(t) \quad \forall t \geq 0, s \geq 0, \ x \in \mathbf{T}^n.$$

Here the function $\bar{\sigma}$ depends only on $\bar{\omega}$ and the Lagrangian L. The above inequality can be rewritten as

$$|u(x,t) - u(x,s)| \leq \bar{\sigma}(|t-s|) \quad \forall x \in \mathbf{T}^n, \ t, s \in [0, \infty).$$

This and (6.6) show that u is indeed uniformly continuous on $\mathbf{T}^n \times [0, \infty)$, thus proving (a).

8. Next we show that u is a viscosity subsolution of (6.2). Let $\psi \in C^1(\mathbf{T}^n \times (0, \infty))$ and $(x_0, t_0) \in \mathbf{T}^n \times (0, \infty)$. Assume that $u - \psi$ attains a maximum at (x_0, t_0). By adding a constant to ψ, we may assume that $u(x_0, t_0) = \psi(x_0, t_0)$ and $u \leq \psi$ on $\mathbf{T}^n \times (0, \infty)$.

Fix $\varepsilon \in (0, t_0)$ and observe by Lemma 6.2 that

$$\begin{aligned}\psi(x_0, t_0) &= (T_\varepsilon^- u(\cdot, t_0 - \varepsilon))(x_0) = \inf_{\gamma(\varepsilon) = x_0} \left[\int_0^\varepsilon L(\gamma(s), \dot{\gamma}(s))\, \mathrm{d}s + u(\gamma(0), t_0 - \varepsilon) \right] \\ &\leq \inf_{\gamma(\varepsilon) = x_0} \left[\int_0^\varepsilon L(\gamma(s), \dot{\gamma}(s))\, \mathrm{d}s + \psi(\gamma(0), t_0 - \varepsilon) \right].\end{aligned}$$

Fix any $v \in \mathbf{R}^n$ and consider the function (or curve) γ defined by $\gamma(s) = x_0 + (\varepsilon - s)v$ for $s \in [0, \varepsilon]$, to find

$$\psi(x_0, t_0) \leq \int_0^\varepsilon L(x_0 + (\varepsilon - s)v, -v)\, \mathrm{d}s + \psi(x_0 + \varepsilon v, t_0 - \varepsilon),$$

from which we get

$$\begin{aligned} 0 &\geq \int_0^\varepsilon \left[-L(x_0 + (\varepsilon - s)v, -v) + \frac{\mathrm{d}}{\mathrm{d}s} \psi(x_0 + (\varepsilon - s)v, t_0 + s - \varepsilon) \right] \mathrm{d}s \\ &= \int_0^\varepsilon \big[-L(x_0 + (\varepsilon - s)v, -v) + \psi_t(x_0 + (\varepsilon - s)v, t_0 + s - \varepsilon) \\ &\quad - v\psi_x(x_0 + (\varepsilon - s)v, t_0 + s - \varepsilon) \big]\, \mathrm{d}s. \end{aligned}$$

Dividing this by ε and sending $\varepsilon \to 0$, we get

$$-L(x_0, -v) - v\psi_x(x_0, t_0) + \psi_t(x_0, t_0) \leq 0 \quad \forall v \in \mathbf{R}^n.$$

Taking the supremum over $v \in \mathbf{R}^n$ yields

$$\psi_t(x_0, t_0) + H(x_0, \psi_x(x_0, t_0)) \le 0,$$

which was to be shown.

9. What remains is to show that u is a viscosity supersolution of (6.2). Let $\psi \in C^1(\mathbf{T}^n \times (0, \infty))$ and $(x_0, t_0) \in \mathbf{T}^n \times (0, \infty)$. Assume that $u - \psi$ attains a minimum at (x_0, t_0). We may assume that $u(x_0, t_0) = \psi(x_0, t_0)$ and $u \ge \psi$ on $\mathbf{T}^n \times (0, \infty)$.

Fix $\varepsilon \in (0, t_0)$ and observe that

$$\begin{aligned}\psi(x_0, t_0) &= (T_\varepsilon^- u(\cdot, t_0 - \varepsilon))(x_0) = \inf_{\gamma(\varepsilon)=x_0} \left[\int_0^\varepsilon L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s + u(\gamma(0), t_0 - \varepsilon) \right] \\ &\ge \inf_{\gamma(\varepsilon)=x_0} \left[\int_0^\varepsilon L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s + \psi(\gamma(0), t_0 - \varepsilon) \right].\end{aligned}$$

Choose a minimizer $\gamma_\varepsilon \in \mathrm{AC}([0, \varepsilon], \mathbf{T}^n)$ for the last variational problem. Compute that

$$\begin{aligned}(6.7)\quad 0 &\le \int_0^\varepsilon \left[L(\gamma_\varepsilon(s), \dot\gamma_\varepsilon(s)) + \frac{\mathrm{d}}{\mathrm{d}s}\psi(\gamma_\varepsilon(s), t_0 + s - \varepsilon) \right] \mathrm{d}s \\ &= \int_0^\varepsilon [-L(\gamma_\varepsilon(s), \dot\gamma_\varepsilon(s)) + \dot\gamma_\varepsilon(s)\psi_x(\gamma_\varepsilon(s), t_0 + s - \varepsilon) + \psi_t(\gamma_\varepsilon(s), t_0 + s - \varepsilon)]\,\mathrm{d}s \\ &\le \int_0^\varepsilon [H(\gamma_\varepsilon(s), \psi_x(\gamma_\varepsilon(s), t_0 + s - \varepsilon)) + \psi_t(\gamma_\varepsilon(s), t_0 + s - \varepsilon)]\,\mathrm{d}s.\end{aligned}$$

Now observe as in the proof of Lemma 3.6 that for any $A > 1$ there exists a $C_A > 0$ such that $L(x, v) \ge A|v| - C_A$ for all $(x, v) \in \mathbf{T}^n \times \mathbf{R}^n$ and hence

$$A \int_0^\varepsilon |\dot\gamma_\varepsilon(t)|\,\mathrm{d}t \le C_A \varepsilon + 2 \max_{\mathbf{T}^n \times [t_0/2, 2t_0]} |\psi| \quad \text{if } \varepsilon \in (0, t_0/2),$$

and moreover

$$A|x_0 - \gamma_\varepsilon(0)| \le C_A \varepsilon + 2 \max_{\mathbf{T}^n \times [t_0/2, 2t_0]} |\psi| \quad \text{if } \varepsilon \in (0, t_0/2).$$

Consequently we have

$$\gamma_\varepsilon(0) \to x_0 \quad \text{as } \varepsilon \to 0.$$

Dividing (6.7) by ε and sending $\varepsilon \to 0$, we get

$$\psi_t(x_0, t_0) + H(x_0, \psi_x(x_0, t_0)) \ge 0.$$

This shows that u is a viscosity supersolution of (6.2). The proof is now complete. □

A remark on T_t^+ similar to Proposition 6.1 is stated as follows.

Proposition 6.3. *Let $\phi \in C(\mathbf{T}^n)$ and define $u : \mathbf{T}^n \times [0, \infty) \to \mathbf{R}$ by*

$$u(x,t) = T_t^+ \phi(x).$$

Then

(a) *u is continuous on $\mathbf{T}^n \times [0, \infty)$;*

(b) *for each $t > 0$ there is a constant $C_t > 0$ such that*

$$|u(x,s) - u(y,s)| \leq C_t |x - y| \quad \forall x, y \in \mathbf{T}^n,\ s \geq t;$$

(c) *u is a viscosity solution of*

$$u_t(x,t) - H(x, u_x(x,t)) = 0 \quad \text{in } \mathbf{T}^n \times (0, \infty). \tag{6.8}$$

Proof. Fix $\phi \in C(\mathbf{T}^n)$. Observe that

$$\begin{aligned} T_t^+ \phi(x) &= \sup_{\gamma(0)=x} \left[-\int_0^t L(\gamma(s), \dot{\gamma}(s))\, \mathrm{d}s + \phi(\gamma(t)) \right] \\ &= \sup_{\mu(t)=x} \left[-\int_0^t \hat{L}(\mu(s), \dot{\mu}(s))\, \mathrm{d}s + \phi(\mu(0)) \right] \\ &= -\inf_{\mu(t)=x} \left[\int_0^t \hat{L}(\mu(s), \dot{\mu}(s))\, \mathrm{d}s - \phi(\mu(0)) \right] \\ &= -\hat{T}_t^-(-\phi)(x) \quad \forall (x,t) \in \mathbf{T}^n \times [0,\, \infty), \end{aligned}$$

where $\hat{L}(x,v) := L(x, -v)$ and

$$\hat{T}_t^- \psi(x) := \inf_{\gamma(t)=x} \left[\int_0^t \hat{L}(\gamma(s), \dot{\gamma}(s))\, \mathrm{d}s + \psi(\gamma(0)) \right] \quad \text{for } \psi \in C(\mathbf{T}^n).$$

By Proposition 6.1, setting

$$v(x,t) = \hat{T}_t^-(-\phi)(x) \quad \text{for } (x,t) \in \mathbf{T}^n \times [0,\, \infty),$$

we see that v has properties (a) and (b) and so does $u = -v$. Also, v is a viscosity solution of

$$v_t + \hat{H}(x, v_x) = 0 \quad \text{in } \mathbf{T}^n \times (0,\, \infty),$$

where $\hat{H}(x,p) := \sup\{vp - \hat{L}(x,v) \mid v \in \mathbf{R}^n\}$. Note that

$$\hat{H}(x,p) = \sup_{v\in\mathbf{R}^n}(-vp - L(x,v)) = H(x,-p) \quad \forall (x,p) \in \mathbf{T}^n \times \mathbf{R}^n.$$

Hence we find that v is a viscosity solution of

$$v_t + H(x, -v_x) = 0 \quad \text{ in } \mathbf{T}^n \times (0, \infty).$$

As we remarked before, the function $u = -v$ is a viscosity solution of

$$-[(-u)_t + H(x, -(-u)_x)] = 0 \quad \text{ in } \mathbf{T}^n \times (0, \infty).$$

That is, u is a viscosity solution of

$$u_t - H(x, u_x) = 0 \quad \text{ in } \mathbf{T}^n \times (0, \infty).$$

The proof is now complete. □

Lemma 6.4. *Let $G \in C(\mathbf{T}^n \times \mathbf{R} \times \mathbf{R}^n)$ have the properties:* (a) *for each $(x,p) \in \mathbf{T}^n \times \mathbf{R}^n$, the function $r \mapsto G(x,r,p)$ is non-decreasing on $\mathbf{R}$;* (b) *for each $r \in \mathbf{R}$,*

$$\lim_{R\to\infty} \inf\{G(x,r,p) \mid (x,p) \in \mathbf{T}^n \times \mathbf{R}^n,\ |p| \geq R\} > 0.$$

Let $c, d \in \mathbf{R}$ satisfy $c < d$. Let $u \in C(\mathbf{T}^n)$ and $v \in C(\mathbf{T}^n)$ be a viscosity subsolution of

$$G(x, u, u_x) = c \quad \text{in } \mathbf{T}^n, \tag{6.9}$$

and a viscosity supersolution of

$$G(x, v, v_x) = d \quad \text{in } \mathbf{T}^n, \tag{6.10}$$

respectively. Then $u \leq v$ on $\mathbf{T}^n$.

Proof. We argue by contradiction. Thus we assume that $\max_{\mathbf{T}^n}(u - v) > 0$ and will get a contradiction. We work on $\mathbf{R}^n$. That is, we regard $u,\ v,\ G(\cdot, r, p)$ as periodic functions on $\mathbf{R}^n$.

Note first that u is a Lipschitz continuous function. Indeed, we choose a constant $C > 0$ so that

$$G(x, \min_{\mathbf{T}^n} u, p) > c \quad \forall (x,p) \in \mathbf{T}^n \times (\mathbf{R}^n \setminus B(0, C)).$$

Fix any $y \in \mathbf{R}^n$ and consider the function $\phi \in C^1(\mathbf{R}^n \times \setminus\{y\})$ defined by

$$\phi(x) = u(y) + C|x - y|.$$

Choosing $R > 0$ large enough, we observe that

$$u(x) < \phi(x) \quad \forall x \in \partial B(y, R),$$

and that

$$G(x, u(x), \phi_x(x)) = G\Big(x, u(x), C\frac{x - y}{|x - y|}\Big) > c. \tag{6.11}$$

We compare u with ϕ on the set $B(y, R)$: if $u(\bar{x}) > \phi(\bar{x})$ at a point $\bar{x} \in B(y, R)$, then $\bar{x} \in \operatorname{int} B(y, R) \setminus \{y\}$ and, since u is a viscosity subsolution of (6.9), we must have

$$G\Big(\bar{x}, u(\bar{x}), C\frac{\bar{x} - y}{|\bar{x} - y|}\Big) \le c.$$

This contradicts (6.11), which shows that $u(x) \le \phi(x)$ in $B(y, R)$. Here R can be chosen independently of y. Accordingly we get

$$u(x) \le u(y) + C|x - y| \quad \text{ if } |x - y| \le R,$$

which implies that

$$|u(x) - u(y)| \le C|x - y| \quad \forall x, y \in \mathbf{R}^n.$$

Now we consider the function

$$\Phi(x, y) = u(x) - v(y) - \alpha|x - y|^2 - \varepsilon(|y|^2 + 1)^{1/2}$$

on $\mathbf{R}^n \times \mathbf{R}^n$, where $\alpha > 1$ and $\varepsilon > 0$ are constants to be sent to ∞ and 0, respectively. Let $(\bar{x}, \bar{y})$ be a maximum point of Φ. Note that

$$\Phi(\bar{x}, \bar{y}) \ge \Phi(\bar{y}, \bar{y}),$$

which yields

$$\alpha|\bar{x} - \bar{y}|^2 \le u(\bar{x}) - u(\bar{y}) \le C|\bar{x} - \bar{y}|,$$

and hence

$$\alpha|\bar{x} - \bar{y}| \le C.$$

Since u and v are a viscosity subsolution of (6.9) and a viscosity supersolution of (6.10), respectively, we get

$$G(\bar{x}, u(\bar{x}), 2\alpha(\bar{x}-\bar{y})) \leq c,$$
$$G\left(\bar{y}, v(\bar{y}), 2\alpha(\bar{x}-\bar{y}) - \varepsilon(|\bar{y}|^2+1)^{-1/2}\bar{y}\right) \geq d.$$

Sending $\varepsilon \to 0$ and $\alpha \to \infty$ together, we find that for some $\hat{x} \in \mathbf{R}^n$ and $\hat{p} \in B(0, C)$,

$$G(\hat{x}, u(\hat{x}), \hat{p}) \leq c < d \leq G(\hat{x}, v(\hat{x}), \hat{p}),$$

which is a contradiction. □

Remark. The above proposition is valid under the weaker assumption that $u \in \mathrm{USC}(\mathbf{T}^n)$ and $v \in \mathrm{LSC}(\mathbf{T}^n)$. The same proof as above yields this result.

Proposition 6.5. (a) *There is a pair of a constant $c_0 \in \mathbf{R}$ and a function $u \in \mathrm{Lip}(\mathbf{T}^n)$ such that u is a viscosity solution of*

$$H(x, u_x) = c_0 \quad \text{in } \mathbf{T}^n. \tag{6.12}$$

(b) *If $(d, v) \in \mathbf{R} \times C(\mathbf{T}^n)$ is another pair for which v is a viscosity solution of*

$$H(x, v_x) = d \quad \text{in } \mathbf{T}^n,$$

then $d = c_0$.

Proof. 1. The underlining idea of the arguments here parallels the proof of Theorem 5.3. We consider the Hamilton-Jacobi equation

$$\lambda u^\lambda(x) + H(x, u_x^\lambda) = 0 \quad \text{in } \mathbf{T}^n, \tag{6.13}$$

where $\lambda \in (0, 1)$ is a parameter to be sent to zero later. This equation has a unique viscosity solution. Indeed, to see the uniqueness, let $u, v \in C(\mathbf{T}^n)$ be a viscosity subsolution and a viscosity supersolution of (6.13). Fix any $\varepsilon > 0$ and set $u_\varepsilon(x) = u(x) - \varepsilon$ for $x \in \mathbf{R}^n$. Then u_ε is a viscosity subsolution of

$$\lambda u_\varepsilon + H(x, Du_\varepsilon) = -\lambda\varepsilon \quad \text{in } \mathbf{T}^n.$$

By Lemma 6.4, we see that $u_\varepsilon \leq v$ on $\mathbf{T}^n$. Since $\varepsilon > 0$ is arbitrary, we get $u \leq v$ on $\mathbf{T}^n$, which implies the uniqueness of viscosity solutions of (6.13). The existence of a viscosity solution of (6.13) can be deduced by Perron's method. In fact, it is easily seen that the function $f(x) := -\lambda^{-1}\max_{x\in\mathbf{T}^n} H(x,0)$ and $g(x) := -\lambda^{-1}\min_{x\in\mathbf{T}^n} H(x,0)$ are classical (and hence viscosity) subsolution and superslution of (6.13), respectively. Note also that $f \le g$ on $\mathbf{R}^n$. Therefore, by Perron's method, we find a function u^λ such that the upper semicontinuous envelope $(u^\lambda)^*$ of u^λ is a viscosity subsolution of (6.13) and the lower semicontinous envelope u^λ_* of u^λ is a viscosity supersolution of (6.13). As above, we may apply Lemma 6.4 to $(u^\lambda)^* - \varepsilon$, with any $\varepsilon > 0$, and u^λ_*, to deduce that $(u^\lambda)^* - \varepsilon \le u^\lambda_*$ on $\mathbf{R}^n$, which yields that $(u^\lambda)^* \le u^\lambda_*$ on $\mathbf{R}^n$. This last inequality implies that $u^\lambda = (u^\lambda)^* = u^\lambda_* \in C(\mathbf{T}^n)$, proving the existence of a viscosity solution of (6.15).

2. Perron's method has yielded a solution u^λ which is given by

$$u^\lambda(x) = \sup\{v(x) \mid v \in C(\mathbf{T}^n) \text{ is a viscosity subsolution of (6.15)},\ f \le v \le g \text{ on } \mathbf{R}^n\} \quad \forall x \in \mathbf{R}^n.$$

From this we see that

$$-\max_{x\in\mathbf{T}^n} H(x,0) \le \lambda u^\lambda(x) \le -\min_{x\in\mathbf{T}^n} H(x,0) \quad \forall x \in \mathbf{R}^n. \tag{6.14}$$

Hence we find that u^λ is a viscosity subsolution of

$$H(x, u^\lambda_x) = \max_{x\in\mathbf{T}^n} H(x,0) \quad \text{in } \mathbf{R}^n.$$

As in the proof of Lemma 6.4, we see that there is a constant $C > 0$, independent of $\lambda \in (0,\,1)$, such that

$$|u^\lambda(x) - u^\lambda(y)| \le C|x-y| \quad \forall x, y \in \mathbf{R}^n,\ \lambda \in (0,\,1). \tag{6.15}$$

We set $w^\lambda(x) = u^\lambda(x) - u^\lambda(0)$ for $x \in \mathbf{R}^n$ and $\lambda \in (0,\,1)$ and $c^\lambda = -\lambda u^\lambda(0)$ for $\lambda \in (0,\,1)$. Then (6.14) and (6.15) guarantee that $\{c^\lambda\}_{0<\lambda<1} \subset \mathbf{R}$ is bounded and $\{w^\lambda\}_{0<\lambda<1} \subset C(\mathbf{T}^n)$ is a uniformly bounded and equi-continuous on $\mathbf{R}^n$. Therefore we can select a sequence $\{\lambda_j\} \in (0,\,1)$ so that as $j \to \infty$,

$$\lambda_j \to 0,$$
$$c^{\lambda_j} \to c_0,$$
$$w^{\lambda_j}(x) \to u(x) \quad \text{uniformly for } x \in \mathbf{T}^n,$$

for some constant c_0 and some function $u \in \mathrm{Lip}(\mathbf{T}^n)$. By the stability of the viscosity property, noting that

$$\lambda w^\lambda(x) + H(x, w^\lambda_x) = c^\lambda \quad \text{in } \mathbf{R}^n$$

in the viscosity sense, we see that u is a viscosity solution of

$$H(x, u_x) = c_0 \quad \text{in } \mathbf{R}^n.$$

Thus we have proved (a).

3. By assumption, we have

$$H(x, u_x) = c_0 \quad \text{in } \mathbf{T}^n,$$
$$H(x, v_x) = d \quad \text{in } \mathbf{T}^n$$

in the viscosity sense. By adding a constant to u, we may assume that $u > v$ on $\mathbf{R}^n$. By Lemma 6.4, we may deduce that $c_0 \geq d$. By adding another constant to u, we may assume in turn that $u < v$ on $\mathbf{R}^n$ and we may deduce as above that $c_0 \leq d$. Thus we see that $c_0 = d$, completing the proof. □

Proposition 6.6. *Let $c_0 \in \mathbf{R}$ be such that there is a viscosity solution $u \in \mathrm{Lip}(\mathbf{T}^n)$ of*

$$H(x, u_x) = c_0 \quad \textit{in } \mathbf{T}^n. \tag{6.16}$$

Then

$$c_0 = \inf_{\phi \in C^1(\mathbf{T}^n)} \sup_{x \in \mathbf{T}^n} H(x, \phi_x(x)). \tag{6.17}$$

Proof. 1. Let $u \in \mathrm{Lip}(\mathbf{T}^n)$ be a viscosity solution of (6.16). Let $\rho \in C_0^\infty(\mathbf{R}^n)$ be a standard mollification kernel such that $\operatorname{spt} \rho \subset B(0, 1)$. Fix $\varepsilon \in (0, 1)$, and set $\rho_\varepsilon(x) = \varepsilon^{-n} \rho(x/\varepsilon)$ and $u_\varepsilon = u * \rho_\varepsilon$. Let $C_0 > 0$ be a Lipschitz constant of the function u. Since u is differentiable a.e. and the a.e. derivatives are identical with the distributional derivatives, using the Jensen inequality, we have

$$H(x, Du_\varepsilon(x)) \leq \rho_\varepsilon * H(x, Du(\cdot)) \leq \rho_\varepsilon * H(\cdot, Du(\cdot)) + \omega(\varepsilon) \leq c_0 + \omega(\varepsilon) \quad \forall x \in \mathbf{R}^n,$$

where ω is the modulus of the function H on $\mathbf{R}^n \times B(0, C_0)$. Now, letting c_1 denote the right hand side of (6.17), we have

$$c_1 \leq \sup_{x \in \mathbf{R}^n} H(x, Du_\varepsilon(x)) \leq c_0 + \omega(\varepsilon) \quad \forall \varepsilon \in (0, 1).$$

Because of the arbitrariness of ε, we find that

$$c_1 \leq c_0. \tag{6.18}$$

2. To prove that $c_0 \le c_1$, we argue by contradiction, and so suppose that $c_0 > c_1$. Let $u \in \mathrm{Lip}(\mathbf{T}^n)$ be a viscosity solution of (6.16) as before. Set $c = (c_0 + c_1)/2$. Then u is a viscosity supersolution of By the definition of c_1, there is a function $\phi \in C^1(\mathbf{T}^n)$ which is a subsolution of

$$H(x, u_x) = c \quad \text{in } \mathbf{R}^n$$

in the classical sense (and hence in the viscosity sense). We may assume by adding a constant to u if necessary that $\phi > u$ on $\mathbf{R}^n$. By Lemma 6.4, since $c < c_0$, we have $\phi \le u$ on $\mathbf{R}^n$, which is a contradiction. Thus we see that $c_0 \le c_1$, completing the proof. □

Now, we turn to the PDE

$$-H(x, u_x) = -d_0 \quad \text{in } \mathbf{T}^n, \tag{6.19}$$

where $d_0 \in \mathbf{R}$ is a constant.

We remark that $u \in C(\mathbf{T}^n)$ is a viscosity solution of (6.19) if and only if $v := -u$ is a viscosity solution of

$$H(x, -v_x) = d_0 \quad \text{in } \mathbf{T}^n.$$

The Hamiltonian $(x, p) \mapsto H(x, -p)$ has the properties required in Propositions 6.5 and 6.6. Therefore, we have the following proposition.

Proposition 6.7. (a) *There is a pair of a constant $d_0 \in \mathbf{R}$ and a function $v \in \mathrm{Lip}(\mathbf{T}^n)$ such that v is a viscosity solution of*

$$H(x, -v_x) = d_0 \quad \textit{in } \mathbf{T}^n.$$

(b) *If $(e, w) \in \mathbf{R} \times C(\mathbf{T}^n)$ is another pair for which w is a viscosity solution of*

$$H(x, -w_x) = e \quad \textit{in } \mathbf{T}^n,$$

then $e = d_0$. (c) *The formula*

$$d_0 = \inf_{\phi \in C^1(\mathbf{T}^n)} \sup_{x \in \mathbf{T}^n} H(x, -\phi_x(x)). \tag{6.20}$$

holds.

Corollary 6.8. *Let c_0 and d_0 be constants from Propositions 6.5 and 6.7, respectively. Then we have $c_0 = d_0$.*

Proof. From (6.17) and (6.20), we have

$$c_0 = \inf_{\phi \in C^1(\mathbf{T}^n)} \sup_{x \in \mathbf{T}^n} H(x, \phi_x(x)) = \inf_{\phi \in C^1(\mathbf{T}^n)} \sup_{x \in \mathbf{T}^n} H(x, -\phi_x(x)) = d_0. \qquad \square$$

7. Some consequences of the main theorem

Define $\mathrm{P}_{\mathrm{inv}}$ as the set of Borel probability measures μ on $\mathbf{T}^n \times \mathbf{R}^n$ which are invariant under the flow $\{\phi_t^L\}_{t\in\mathbf{R}}$. Here, by definition, μ is *invariant* under the flow $\{\phi_t^L\}$ if for all $\theta \in C_b(\mathbf{T}^n \times \mathbf{R}^n)$,

$$\int_{\mathbf{T}^n\times\mathbf{R}^n} \theta \circ \phi_t^L \;\mathrm{d}\mu = \int_{\mathbf{T}^n\times\mathbf{R}^n} \theta \;\mathrm{d}\mu \quad \forall t \in \mathbf{R}.$$

Theorem 7.1. *We have*

$$-c_0 = \inf_{\mu\in\mathrm{P}_{\mathrm{inv}}} \int_{\mathbf{T}^n\times\mathbf{R}^n} L \;\mathrm{d}\mu.$$

Proof. Let (u_-, γ_-, c_0) be from the weak KAM theorem, where we not not specify the value $\gamma_-(0)$.

1. By property (a) of u_-, for all $(x,v) \in \mathbf{T}^n \times \mathbf{R}^n$, we have

$$u_-(\pi \circ \phi_t^L(x,v)) - u_-(\pi \circ \phi_0^L(x,v)) \le \int_0^t L(\phi_s^L(x,v)) \;\mathrm{d}s + c_0 t \quad \forall t > 0.$$

Let $\mu \in \mathrm{P}_{\mathrm{inv}}$. We integrate the above by μ over $\mathbf{T}^n \times \mathbf{R}^n$, to get

$$\int_{\mathbf{T}^n\times\mathbf{R}^n} u_- \circ \pi \circ \phi_t^L \;\mathrm{d}\mu - \int_{\mathbf{T}^n\times\mathbf{R}^n} u_- \circ \pi \;\mathrm{d}\mu \\ \le \int_0^t \left(\int_{\mathbf{T}^n\times\mathbf{R}^n} L \circ \phi_s^L \;\mathrm{d}\mu\right) \mathrm{d}s + c_0 t \quad \forall t > 0.$$

Hence, using the invariance of μ under $\{\phi_t^L\}$, we find that

$$0 \le \int_{\mathbf{T}^n\times\mathbf{R}^n} L \;\mathrm{d}\mu + c_0.$$

Thus we have

$$-c_0 \le \inf_{\mu\in\mathrm{P}_{\mathrm{inv}}} \int_{\mathbf{T}^n\times\mathbf{R}^n} L \;\mathrm{d}\mu. \tag{7.1}$$

2. Define the Borel probability measures μ_k, with $k \in \mathbf{N}$, on $\mathbf{T}^n \times \mathbf{R}^n$ by

$$\int_{\mathbf{T}^n\times\mathbf{R}^n} \theta \;\mathrm{d}\mu_k = \frac{1}{k}\int_{-k}^0 \theta(\gamma_-(s), \dot{\gamma}_-(s)) \;\mathrm{d}s \quad \forall \theta \in C_b(\mathbf{T}^n \times \mathbf{R}^n).$$

Since

$$(\gamma_-(t), \dot{\gamma}_-(t)) = \phi_t^L(\gamma_-(0), \dot{\gamma}_-(0)) \quad \forall t \le 0,$$

there is a constant $R > 0$ such that

$$|\dot\gamma_-(t)| \le R \quad \forall t \le 0.$$

Therefore we have

$$\operatorname{spt} \mu_k \subset \mathbf{T}^n \times B(0,R),$$

where the set on the right hand side is a compact set. Thus, we find a subsequnece $\{\mu_{k_j}\}_{j\in\mathbf{N}}$ and a Borel probability measure μ_- such that as $j \to \infty$,

$$\mu_{k_j} \to \mu_- \qquad \text{weakly in the sense of measures.}$$

3. We now show that μ_- is invariant under the flow $\{\phi_t^L\}$. Fix any $t \in \mathbf{R}$ and $\theta \in C_b(\mathbf{T}^n \times \mathbf{R}^n)$. We have

$$\begin{aligned}\int_{\mathbf{T}^n\times\mathbf{R}^n} \theta\circ\phi_t^L \ \mathrm{d}\mu_- &= \lim_{j\to\infty} \frac{1}{k_j}\int_{-k_j}^0 \theta\circ\phi_t^L\circ\phi_s^L(x_0,v_0)\ \mathrm{d}s \\ &= \lim_{j\to\infty} \frac{1}{k_j}\int_{-k_j}^0 \theta\circ\phi_{t+s}^L(x_0,v_0)\ \mathrm{d}s,\end{aligned}$$

where $(x_0,v_0) = (\gamma_-(0),\dot\gamma_-(0))$, and

$$\begin{aligned}&\int_{-k}^0 \theta\circ\phi_{t+s}(x_0,v_0)\ \mathrm{d}s = \int_{t-k}^t \theta\circ\phi_s^L(x_0,v_0)\ \mathrm{d}s \\ &= \int_{-k}^0 \theta\circ\phi_s^L(x_0,v_0)\ \mathrm{d}s + \int_0^t \theta\circ\phi_s^L(x_0,v_0)\ \mathrm{d}s + \int_{t-k}^{-k} \theta\circ\phi_s^L(x_0,v_0)\ \mathrm{d}s.\end{aligned}$$

Hence, dividing this by k and sending $k = k_j \to \infty$, we get

$$\int_{\mathbf{T}^n\times\mathbf{R}^n} \theta\circ\phi_t^L\ \mathrm{d}\mu_- = \lim_{j\to\infty}\frac{1}{k_j}\int_{k_j}^0 \theta\circ\phi_s^L(x_0,v_0)\ \mathrm{d}s = \int_{\mathbf{T}^n\times\mathbf{R}^n} \theta\ \mathrm{d}\mu_-,$$

and conclude that $\mu_- \in \mathrm{P}_{\mathrm{inv}}$. Therefore we have

$$\inf_{\mu\in\mathrm{P}_{\mathrm{inv}}} \int_{\mathbf{T}^n\times\mathbf{R}^n} L\ \mathrm{d}\mu \le \int_{\mathbf{T}^n\times\mathbf{R}^n} L\ \mathrm{d}\mu_-. \tag{7.2}$$

4. We observe that

$$\begin{aligned}\int_{\mathbf{T}^n\times\mathbf{R}^n} L\ \mathrm{d}\mu_- &= \lim_{j\to\infty}\frac{1}{k_j}\int_{-k_j}^0 L(\gamma(t),\dot\gamma_-(t))\ \mathrm{d}t \\ &= \lim_{j\to\infty} \frac{u_-(\gamma(0)) - u_-(\gamma(-k_j)) - c_0k_j}{k_j} = -c_0.\end{aligned}$$

Combine this with (7.1) and (7.2), to conclude that

$$-c_0 = \inf_{\mu \in \mathrm{P_{inv}}} \int_{\mathbf{T}^n \times \mathbf{R}^n} L \, \mathrm{d}\mu = \int_{\mathbf{T}^n \times \mathbf{R}^n} L \, \mathrm{d}\mu_-. \qquad \square$$

Remark. The variational problem

$$\inf_{\mu \in \mathrm{P_{inv}}} \int_{\mathbf{T}^n \times \mathbf{R}^n} L \, \mathrm{d}\mu$$

has a minimizer as the above proof shows. In what follows we write

$$\mathrm{P_{min}} = \{\mu \in \mathrm{P_{inv}} \mid \int_{\mathbf{T}^n \times \mathbf{R}^n} L \, \mathrm{d}\mu = \inf_{\nu \in \mathrm{P_{inv}}} \int_{\mathbf{T}^n \times \mathbf{R}^n} L \, \mathrm{d}\nu\}.$$

We introduce the *Aubry set* A_ε^-, with parameter $\varepsilon > 0$, as the set of points $x \in \mathbf{T}^n$ such that there exists $\gamma_x \in \mathrm{AC}([-\varepsilon, \varepsilon], \mathbf{T}^n)$ satisfying $\gamma_x(0) = x$ for which

$$u_-(\gamma_x(\varepsilon)) - u_-(\gamma_x(-\varepsilon)) = \int_{-\varepsilon}^{\varepsilon} L(\gamma_x(s), \dot{\gamma}_x(s)) \, \mathrm{d}s + 2\varepsilon c_0. \tag{7.3}$$

Remark. Note that A_ε^- depends also on the choice of u^-. We refer to [Fa2, FS1, FS2] for recent developments related to Aubry sets.

Theorem 7.2. *We have:*

(a) *u^- is differentiable at every $x \in A_\varepsilon^-$.*

(b) *$Du_-(x) = L_v(x, \dot{\gamma}_x(0))$ for all $x \in A_\varepsilon^-$.*

(c) *The map $x \mapsto Du_-(x)$, $A_\varepsilon^- \to \mathbf{R}^n$ is Lipschitz continuous.*

Proof. We write u for u_-.

1. We prove first (a) and (b). Fix $x \in A_\varepsilon^-$ and let $\gamma_x \in C^2([-\varepsilon, \varepsilon], \mathbf{T}^n)$ satisfy (7.3) and $\gamma_x(0) = x$. We have

$$u(\gamma_x(0)) - u(\gamma_x(-\varepsilon)) = \int_{-\varepsilon}^{0} L(\gamma_x(s), \dot{\gamma}_x(s)) \, \mathrm{d}s + c_0 \varepsilon, \tag{7.4}$$

$$u(\gamma_x(\varepsilon)) - u(\gamma_x(0)) = \int_{0}^{\varepsilon} L(\gamma_x(s), \dot{\gamma}_x(s)) \, \mathrm{d}s + c_0 \varepsilon. \tag{7.5}$$

2. Fix $y \in \mathbf{T}^n$. Define $\mu_- \in C^2([-\varepsilon, 0], \mathbf{T}^n)$ by

$$\mu_-(t) = \gamma_x(t) + \frac{\varepsilon + t}{\varepsilon}(y - x).$$

Note that $\mu_-(0) = y$, $\dot{\mu}_-(t) = \dot{\gamma}_x(t) + \frac{1}{\varepsilon}(y-x)$ for all $t \in [-\varepsilon, 0]$, and $\mu_-(-\varepsilon) = \gamma_x(-\varepsilon)$. By the property (a) of u_- in the weak KAM theorem, we have

$$u(\mu_-(0)) - u(\mu_-(-\varepsilon)) \le \int_{-\varepsilon}^{0} L(\mu_-(s), \dot{\mu}_-(s))\,\mathrm{d}s + c_0\varepsilon. \tag{7.6}$$

3. Combining (7.4) and (7.6), we get

$$\begin{aligned} u(y) - u(x) &\le u(\mu_-(-\varepsilon)) - u(\gamma_x(-\varepsilon)) + \int_{-\varepsilon}^{0} [L(\mu_-, \dot{\mu}_-) - L(\gamma_x, \dot{\gamma}_x)]\ \mathrm{d}s \\ &= \int_{-\varepsilon}^{0} [L(\mu_-, \dot{\mu}_-) - L(\gamma_x, \dot{\gamma}_x)]\ \mathrm{d}s. \end{aligned}$$

We choose a constant $C > 0$ so that

$$|\dot{\gamma}_x(t)| \le C \quad \forall t \in [-\varepsilon,\, \varepsilon].$$

Noting that

$$\max\{|\dot{\gamma}_x(t)|,\ |\dot{\mu}_-(t)|\} \le C + \frac{|y-x|}{\varepsilon} \le C_\varepsilon \quad \forall t \in [-\varepsilon, 0],$$

where we may assume that $|x - y| \le \sqrt{n}$ and consequently we may choose $C_\varepsilon = C + \varepsilon^{-1}\sqrt{n}$, and applying the Taylor theorem, we get

$$\begin{aligned} &u(y) - u(x) \\ &\le \int_{-\varepsilon}^{0} \left(L_x(\gamma_x(s), \dot{\gamma}_x(s)) \cdot \frac{\varepsilon + s}{\varepsilon}(y-x) + L_v(\gamma_x(s), \dot{\gamma}_x(s)) \cdot \frac{1}{\varepsilon}(y-x) \right) \mathrm{d}s \\ &+ K_\varepsilon |y-x|^2 \end{aligned} \tag{7.7}$$

for some constant $K_\varepsilon > 0$, for instance,

$$K_\varepsilon = \frac{1}{2}\left(1 + \frac{1}{\varepsilon}\right) \max_{(x,v)\in\mathbf{T}^n\times B(0,C_\varepsilon)} \left\| \begin{pmatrix} L_{xx} & L_{xv} \\ L_{vx} & L_{vv} \end{pmatrix} \right\|.$$

Since γ_x satisfies the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma_x(t), \dot{\gamma}_x(t)) = L_x(\gamma_x(t), \dot{\gamma}_x(t)) \quad \forall t \in [-\varepsilon, \varepsilon],$$

by integration by parts, we find that

$$\begin{aligned} &\int_{-\varepsilon}^{0} \frac{\varepsilon + t}{\varepsilon} L_x(\gamma_x(t), \dot{\gamma}_x(t)) \cdot (y-x)\,\mathrm{d}t = \int_{-\varepsilon}^{0} \frac{\varepsilon + t}{\varepsilon} \frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma_x(t), \dot{\gamma}_-(t)) \cdot (y-x)\,\mathrm{d}t \\ &= \left[\frac{\varepsilon + t}{\varepsilon} L_v(\gamma_x(t), \dot{\gamma}_x(t)) \cdot (y-x) \right]_{t=-\varepsilon}^{t=0} - \int_{-\varepsilon}^{0} \frac{1}{\varepsilon} L_v(\gamma_x(t), \dot{\gamma}_x(t)) \cdot (y-x)\,\mathrm{d}t \\ &= L_v(\gamma_x(0), \dot{\gamma}_x(0)) \cdot (y-x) - \int_{-\varepsilon}^{0} \frac{1}{\varepsilon} L_v(\gamma_x(t), \dot{\gamma}_x(t)) \cdot (y-x)\,\mathrm{d}t. \end{aligned}$$

This together with (7.7) yields

$$u(y) - u(x) \le L_v(x, \dot\gamma_x(0)) \cdot (y - x) + K_\varepsilon |y - x|^2. \tag{7.8}$$

4. Next define $\mu_+ \in C^2([0,\varepsilon], \mathbf{T}^n)$ by

$$\mu_+(t) = \gamma_x(t) + \frac{\varepsilon - t}{\varepsilon}(y - x).$$

Note that $\mu_+(0) = y$, $\dot\mu_+(t) = \dot\gamma_x(t) + \frac{1}{\varepsilon}(x - y)$ for all $t \in [0, \varepsilon]$, and $\mu_+(\varepsilon) = \gamma_x(\varepsilon)$.

By property (a) of u_- in the weak KAM theorem, we have

$$u(\mu_+(\varepsilon)) - u(\mu_+(0)) \le \int_0^\varepsilon L(\mu_+(t), \dot\mu_+(t))\,\mathrm{d}t + c_0\varepsilon.$$

Combine this with

$$u(\gamma_x(\varepsilon)) - u(\gamma_x(0)) = \int_0^\varepsilon L(\gamma_x(t), \dot\gamma_x(t))\,\mathrm{d}t + c_0\varepsilon,$$

to get

$$\begin{aligned} u(x) - u(y) &\le u(\gamma_x(\varepsilon)) - u(\mu_+(\varepsilon)) + \int_0^\varepsilon [L(\mu_+(t), \dot\mu_+(t)) - L(\gamma_x(t), \dot\gamma_x(t))]\ \mathrm{d}t \\ &\le \int_0^\varepsilon [L(\mu_+(t), \dot\mu_+(t)) - L(\gamma_x(t), \dot\gamma_x(t))]\ \mathrm{d}t. \end{aligned}$$

Using the Taylor theorem, the Euler-Lagrange equation, and integration by parts, we get

$$\begin{aligned} &u(x) - u(y) \\ &\le \int_0^\varepsilon (L_x(\gamma_x(t), \dot\gamma_x(t)) \cdot (\mu_+(t) - \gamma_x(t)) + L_v(\gamma_x(t), \dot\gamma_x(t)) \cdot (\dot\mu_+(t) - \dot\gamma_x(t)))\ \mathrm{d}t \\ &+ K_\varepsilon |x - y|^2 \\ &= \int_0^\varepsilon \left(\frac{\varepsilon - t}{\varepsilon} \frac{\mathrm{d}}{\mathrm{d}t} L_v(\gamma_x(t), \dot\gamma_x(t)) \cdot (y - x) - \frac{1}{\varepsilon} L_v(\gamma_x(t), \dot\gamma_x(t)) \cdot (y - x) \right) \mathrm{d}t \\ &+ K_\varepsilon |x - y|^2 \\ &= -L_v(\gamma_x(0), \dot\gamma_x(0)) \cdot (y - x) + K_\varepsilon |y - x|^2. \end{aligned}$$

This and (7.8) yield

$$|u(y) - u(x) - L_v(x, \dot\gamma_x(0)) \cdot (y - x)| \le K_\varepsilon |x - y|^2 \quad \forall x \in A_\varepsilon^-,\ y \in \mathbf{T}^n.$$

In particular, we see that

$$Du(x) = L_v(x, \dot\gamma_x(0)) \quad \forall x \in A_\varepsilon^-,$$

which proves (7.1) and (7.2). In order to complete the proof, we just need to apply the following lemma. $\square$

Lemma 7.3. *Let $A \subset \mathbf{R}^n$ and $u : \mathbf{R}^n \to \mathbf{R}$. Assume that u is differentiable at every point $x \in A$ and that there is a constant $K > 0$ for which*

$$|u(y) - u(x) - Du(x)\cdot(y-x)| \le K|y-x|^2 \quad \forall y \in \mathbf{R}^n. \tag{7.9}$$

Then

$$|Du(y) - Du(x)| \le 6K|y-x|.$$

Proof. Let $x_1, x_2 \in \mathbf{R}^n$. Let $h \in \mathbf{R}^n$ be a vector to be fixed later on. We assume that $|h| = |x_1 - x_2|$. By (7.9), we find that

$$\begin{aligned}
|u(x_1+h) - u(x_1) - Du(x_1)\cdot h| &\le K|h|^2,\\
|u(x_1) - u(x_2) - Du(x_2)\cdot(x_1 - x_2)| &\le K|h|^2,\\
|u(x_1+h) - u(x_2) - Du(x_2)\cdot(h + x_1 - x_2)| &\le K|x_1 - x_2 + h|^2 \le 4K|h|^2.
\end{aligned}$$

Noting that

$$\begin{aligned}
&u(x_1+h) - u(x_1) + Du(x_1)\cdot h\\
&- u(x_1) + u(x_2) + Du(x_2)\cdot(x_1 - x_2)\\
&+ u(x_1+h) - u(x_2) - Du(x_2)\cdot(h + x_1 - x_2)\\
&= (Du(x_1) - Du(x_2))\cdot h,
\end{aligned}$$

we get

$$\begin{aligned}
|(Du(x_1) - Du(x_2))\cdot h| \le{}& |u(x_1+h) - u(x_1) - Du(x_1)\cdot h|\\
&+ |u(x_1) - u(x_2) - Du(x_2)\cdot(x_1 - x_2)|\\
&+ |u(x_1+h) - u(x_2) - Du(x_2)\cdot(h + x_1 - x_2)| \le 6K|h|^2.
\end{aligned}$$

We amy assume that $x_1 \ne x_2$. Setting

$$h = \frac{Du(x_1) - Du(x_2)}{|Du(x_1) - Du(x_2)|}|x_1 - x_2|,$$

we find from the above inequality that

$$|Du(x_1) - Du(x_2)| \le 6K|x_1 - x_2|,$$

which completes the proof. $\square$

Define the *Mather set* $\widetilde{\mathcal{M}}_0$ and the *projected Mather set* $\mathcal{M}_0$ by

$$\widetilde{\mathcal{M}}_0 = \text{ closure of } \bigcup\{\operatorname{spt}\mu \mid \mu \in \mathrm{P}_{\min}\},$$
$$\mathcal{M}_0 = \pi(\widetilde{\mathcal{M}}_0).$$

Remark. By definition, for a Borel probability measure μ on $\mathbf{T}^n \times \mathbf{R}^n$,

$$\operatorname{spt}\mu = \{(x,v) \in \mathbf{T}^n \times \mathbf{R}^n \mid \mu(U) > 0 \text{ for all neighborhood } U \text{ of } (x,v)\}.$$

Proposition 7.4. *$\widetilde{\mathcal{M}}_0$ is invariant under the flow $\{\phi_t^L\}$.*

Proof. We argue by contradiction. Suppose that there were a point $(x_0, v_0) \in \widetilde{\mathcal{M}}_0$ and $t \in \mathbf{R}$ such that

$$\phi_t^L(x_0, v_0) \notin \widetilde{\mathcal{M}}_0.$$

Choose a neighborhood $U \subset \mathbf{T}^n \times \mathbf{R}^n$ of $\phi_t^L(x_0, v_0)$ such that

$$U \cap \widetilde{\mathcal{M}}_0 = \emptyset.$$

Set $V = \phi_{-t}^L(U)$. Since $V \ni \phi_{-t}^L(\phi_t^L(x_0, v_0)) = (x_0, v_0)$ and $\phi_{-t}^L : \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{T}^n \times \mathbf{R}^n$ is a homeomorphism, V is a neighborhood of (x_0, v_0). By the definition of $\widetilde{\mathcal{M}}_0$, there is a $\mu \in \mathrm{P}_{\min}$ such that

$$\int \mathbf{1}_V \,\mathrm{d}\mu > 0.$$

Note that

$$\int \mathbf{1}_U \,\mathrm{d}\mu = 0$$

and that $\mathbf{1}_{\phi_{-t}^L(U)} = \mathbf{1}_U \circ \phi_t^L$. Using the invariance of μ under $\{\phi_s^L\}$, we find that

$$0 < \int \mathbf{1}_V \,\mathrm{d}\mu = \int \mathbf{1}_{\phi_{-t}^L(U)} \,\mathrm{d}\mu = \int \mathbf{1}_U \circ \phi_t^L \,\mathrm{d}\mu = \int \mathbf{1}_U \,\mathrm{d}\mu = 0.$$

This is a contradiction, which completes the proof. $\square$

Theorem 7.5. *$\mathcal{M}_0 \subset A_\varepsilon^-$ for all $\varepsilon > 0$.*

Proof. 1. Define the function $\psi : \mathbf{T}^n \times \mathbf{R}^n \to \mathbf{R}$ by

$$\psi(x,v) = \int_{-\varepsilon}^{\varepsilon} L \circ \phi_s^L(x,v)\,\mathrm{d}s + 2c_0\varepsilon - u_- \circ \pi \circ \phi_\varepsilon^L(x,v) + u_- \circ \pi \circ \phi_{-\varepsilon}^L(x,v).$$

Note that $\psi \in C(\mathbf{T}^n \times \mathbf{R}^n)$ and $\psi \geq 0$ on $\mathbf{T}^n \times \mathbf{R}^n$ by property (a) of u_- in the weak KAM theorem.

2. We show that

$$\psi(x,v) = 0 \quad \forall (x,v) \in \widetilde{\mathcal{M}}_0. \tag{7.10}$$

To this end, we argue by contradiction. We suppose that there were a point $(x_0, v_0) \in\in \widetilde{\mathcal{M}}_0$ such that $\psi(x_0,v_0) > 0$. There is an open neighborhood $U \subset \mathbf{T}^n \times \mathbf{R}^n$ of (x_0,v_0) such that $\psi(x,v) > 0$ for all $(x,v) \in U$. Since $(x_0,v_0) \in \widetilde{\mathcal{M}}_0$, there is a $\mu \in \mathrm{P}_{\min}$ such that $\operatorname{spt}\mu \cap U \neq \emptyset$. Then we have

$$\int_U \psi\,\mathrm{d}\mu > 0. \tag{7.11}$$

On the other hand, we have

$$\int_U \psi\,\mathrm{d}\mu \leq \int_{\mathbf{T}^n \times \mathbf{R}^n} \psi\,\mathrm{d}\mu$$

and

$$\begin{aligned}\int_{\mathbf{T}^n \times \mathbf{R}^n} \psi\,\mathrm{d}\mu &= \int_{-\varepsilon}^{\varepsilon} \mathrm{d}s \int_{\mathbf{T}^n \times \mathbf{R}^n} L \circ \phi_s\,\mathrm{d}\mu \\ &\quad + 2c_0\varepsilon - \int_{\mathbf{T}^n \times \mathbf{R}^n} u_- \circ \pi \circ \phi_s^L\,\mathrm{d}\mu + \int_{\mathbf{T}^n \times \mathbf{R}^n} u_- \circ \pi\,\mathrm{d}\mu \\ &= 2(-c_0)\varepsilon + 2c_0\varepsilon = 0.\end{aligned}$$

Hence,

$$\int_U \psi\,\mathrm{d}\mu \leq 0.$$

This contradicts with (7.11), which completes the proof. □

References

[BC] M. Bardi, I. Capuzzo Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, With appendices by Maurizio Falcone and Pierpaolo Soravia, Systems & Control: Foundations & Applications. Birkhauser Boston, Inc., Boston, MA, 1997.

[B] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Mathematiques & Applications (Berlin) [Mathematics & Applications], 17, Springer-Verlag, Paris, 1994.

[CL] M. G. Crandall, P. L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. **277** (1983), no. 1, 1–42.

[CEL] M. G. Crandall, L. C. Evans, P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. **282** (1984), no. 2, 487–502.

[Fa1] A. Fathi, Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens, C. R. Acad. Sci. Paris Ser. I Math. **324** (1997), no. 9, 1043–1046.

[Fa2] A. Fathi, *Weak KAM theorem in Lagrangian dynamics*, fourth preliminary version.

[FS1] A. Fathi, A. Siconolfi, Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation, Invent. Math. **155** (2004), no. 2, 363–388.

[FS2] A. Fathi, A. Siconolfi, PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians, Calc. Var. Partial Differential Equations **22** (2005), no. 2, 185–228.

[Fr] A. Friedman, *Foundations of modern analysis*, Dover, New York, 1982.

[L] P. L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics, 69. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982.