WEAK KAM THEORY FOR HAMILTON-JACOBI EQUATIONS DEPENDING ON UNKNOWN FUNCTIONS

XIUFENG SU
School of Mathematical Sciences
Beijing Normal University, Beijing 100875, China

LIN WANG
Yau Mathematical Sciences Center
Tsinghua University, Beijing 100084, China

JUN YAN
School of Mathematical Sciences
Fudan University, Shanghai 200433, China

(Communicated by Chong-Qing Cheng)

Abstract. We consider the evolutionary Hamilton-Jacobi equation depending on the unknown function with the continuous initial condition on a connected closed manifold. Under certain assumptions on $H(x,u,p)$ with respect to $u$ and $p$, we provide an implicit variational principle. By introducing an implicitly defined solution semigroup and an admissible value set $C_H$, we extend weak KAM theory to certain more general cases, in which $H$ depends on the unknown function $u$ explicitly. As an application, we show that for $0 \notin C_H$, as $t \to +\infty$, the viscosity solution of

\[
\begin{cases}
\partial_t u(x,t) + H(x,u(x,t),\partial_x u(x,t)) = 0, \\
u(x,0) = \varphi(x),
\end{cases}
\]

diverges, otherwise for $0 \in C_H$, it converges to a weak KAM solution of the stationary Hamilton-Jacobi equation

\[H(x,u(x),\partial_x u(x)) = 0.\]

Contents

1. Introduction and main results 6488
2. Preliminaries 6492
3. Variational principle 6494
4. Representation of the viscosity solution 6502
5. Solution semigroup 6504
6. Convergence of the solution semigroup 6510
Acknowledgements 6521
REFERENCES 6521

2010 Mathematics Subject Classification. Primary: 37J50, 35F21; Secondary: 35D40.
Key words and phrases. Weak KAM theory, Hamilton-Jacobi equations, viscosity solutions.
1. **Introduction and main results.** Let $M$ be a connected closed (compact and without boundary) $C^r(r \geq 2)$ manifold and $H : T^*M \times \mathbb{R} \to \mathbb{R}$ be a $C^r(r \geq 2)$ function called a Hamiltonian. For a given $T > 0$, we consider the following Hamilton-Jacobi equation:

$$\partial_t u(x, t) + H(x, u(x, t), \partial_x u(x, t)) = 0,$$

(1)

where $(x, t) \in M \times [0, T]$ and with the initial condition:

$$u(x, 0) = \varphi(x),$$

where $\varphi(x) \in C(M, \mathbb{R})$. The characteristics of (1) satisfy the following ordinary differential equations:

$$\begin{align*}
\dot{x} &= \frac{\partial H}{\partial p}, \\
\dot{p} &= -\frac{\partial H}{\partial x} - \frac{\partial H}{\partial u} p, \\
\dot{u} &= \frac{\partial H}{\partial p} p - H.
\end{align*}$$

(2)

The equation (2) is also referred as to the contact Hamiltonian equation on the contact manifold $(J^1(M, \mathbb{R}), du - pdx)$ (see [1]).

In 1983, M. Crandall and P. L. Lions introduced a notion of weak solution of (1) named viscosity solution for overcoming the lack of uniqueness of the solution due to the crossing of characteristics (see [1, 8, 11]). During the same period, S. Aubry and J. Mather developed a seminal work so called Aubry-Mather theory on global action minimizing orbits for area-preserving twist maps (see [3, 4, 21, 22, 23, 24] for instance). In 1991, J. Mather generalized Aubry-Mather theory to positive definite and superlinear Lagrangian systems with multi-degrees of freedom (see [25]).

There is a close connection between viscosity solutions and Aubry-Mather theory. Roughly speaking, the global minimizing orbits in Aubry-Mather theory can be embedded into the characteristic fields of PDEs. The similar ideas were reflected in pioneering papers [15] and [16] respectively. In [15], W. E was concerned with certain weak solutions of the Burgers equation, which are corresponding to area-preserving twist maps. In [16], A. Fathi considered the Hamilton-Jacobi equations under so called Tonelli conditions (see (H1)-(H2) below), which are corresponding to positive definite and superlinear Lagrangian systems. Since then, weak KAM theory has been well developed. A systematic introduction to weak KAM theory can be found in [18].

In this paper, we are devoted to exploring the dynamics of more general Hamilton-Jacobi equations, in which the Hamiltonian $H$ depends on the unknown function $u$ explicitly. Precisely speaking, we are concerned with a $C^r(r \geq 2)$ Hamiltonian $H(x, u, p)$ satisfying the following assumptions:

(H1) **Positive Definiteness.** For every $(x, u) \in M \times \mathbb{R}$, the second partial derivative $\partial^2 H/\partial p^2(x, u, p)$ is positive definite as a quadratic form;

(H2) **Superlinear Growth.** For every $(x, u) \in M \times \mathbb{R}$, $H(x, u, p)$ is superlinear with respect to $p$;

(H3) **Uniform Lipschitzity.** $H(x, u, p)$ is uniformly Lipschitzian with respect to $u$.

(H4) **Monotonicity.** $H(x, u, p)$ is non-decreasing with respect to $u$.

(H1)-(H2) are called Tonelli conditions (see [18, 25]). (H4) is referred to as “proper” condition for stationary Hamilton-Jacobi equations (see [9]). There is a broad class of Hamiltonians satisfying (H1)-(H4). Obviously, our assumptions cover Tonelli Hamiltonians independent of $u$. Besides, it also contains more general cases, for
instance, discounted Tonelli Hamiltonian $\lambda u + H(x, p)$ with $\lambda > 0$, which was focused from the view of weak KAM under weaker assumptions [12].

Based on (H1), it is easy to see that (H2) is equivalent to the superlinearity of $H(x, u, p)$ above each compact set of $M \times \mathbb{R}$. Since $M$ is compact, (H2) implies that for each $x \in M$ and $u \in I$ with a compact subset $I \subset \mathbb{R}$, $H(x, u, p)$ is superlinear with respect to $p$. (See [18, Theorem 1.3.14] for details.)

The aim of this paper is to show the main ideas of exploring the dynamics of more general Hamilton-Jacobi equations. To avoid the digression, we do not discuss whether the assumptions (H1)-(H4) are optimal, which will be focused in future works.

To state the main results, we introduce some terminology. Let us recall the Legendre transformation on the Hamiltonian independent of $u$, which is formulated as $\mathcal{L} : T^*M \to TM$ via

$$(x, \dot{x}) = \left( x, \frac{\partial H}{\partial p}(x, p) \right).$$

Let $\bar{\mathcal{L}} := (\mathcal{L}, Id)$, where $Id$ denotes the identity map from $\mathbb{R}$ to $\mathbb{R}$. Then $\bar{\mathcal{L}}$ denotes the diffeomorphism from $T^*M \times \mathbb{R}$ to $TM \times \mathbb{R}$. By $\mathcal{L}$, the Lagrangian $L(x, u, \dot{x})$ associated to $H(x, u, p)$ can be denoted by

$$L(x, u, \dot{x}) := \sup_{p \in T^*_x M} \{ (\dot{x}, p) - H(x, u, p) \}.$$ 

Since the contact vector field is of class $C^1$, by existence and uniqueness theorems in ODE, for each compact subset $K \subset T^*M \times \mathbb{R}$, there exist a neighborhood $U$ of $K$ and a $\varepsilon := \varepsilon(K) > 0$ such that one can define a local phase flow $\Psi : (-\varepsilon, \varepsilon) \times U \to T^*M \times \mathbb{R}$

$$\Psi_t(x_0, p_0, u_0) := (X(t), P(t), U(t)),$$

where we use $(X(t), U(t), P(t))$ to denote the solution of (2) with the initial data

$$(X(0), U(0), P(0)) = (x_0, u_0, p_0).$$

Without ambiguity, we change the order of variables and denote

$$\Psi_t(x_0, u_0, p_0) := (X(t), U(t), P(t))$$

for the consistency with $H(x, u, p)$. The flow generated by $L(x, u, \dot{x})$ can be denoted by $\Phi_t := \bar{\mathcal{L}} \circ \Psi_t \circ \bar{\mathcal{L}}^{-1}$. From (H1)-(H4), it follows that the Lagrangian $L(x, u, \dot{x})$ satisfies:

1. **Positive Definiteness.** For every $(x, u) \in M \times \mathbb{R}$, the second partial derivative $\frac{\partial^2 L}{\partial \dot{x}^2}(x, u, \dot{x})$ is positive definite as a quadratic form;

2. **Superlinear Growth.** For every $(x, u) \in M \times \mathbb{R}$, $L(x, u, \dot{x})$ is superlinear with respect to $\dot{x};$

3. **Uniform Lipschitzity.** $L(x, u, \dot{x})$ is uniformly Lipschitzian with respect to $u$.

4. **Monotonicity.** $L(x, u, \dot{x})$ is non-increasing with respect to $u$.

If a $C^r (r \geq 2)$ Hamiltonian $H(x, u, p)$ satisfies (H1)-(H4) (associated $L(x, u, \dot{x})$ satisfying (L1)-(L4)), then we obtain the following theorem.

**Theorem 1.1.** For any $\varphi(x) \in C(M, \mathbb{R})$, there exists a unique $u(x, t) \in C(M \times [0, T], \mathbb{R})$ satisfying $u(x, 0) = \varphi(x)$ such that

$$u(x, t) = \inf_{\gamma(t) = x} \left\{ \varphi(\gamma(0)) + \int_0^t L(\dot{\gamma}(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau \right\}. \quad (3)$$
where the infimum is taken among the continuous and piecewise $C^1$ curves $\gamma : [0, t] \to M$. The infimum is attained at a $C^1$ curve denoted by $\bar{\gamma}$. Moreover, for $\tau \in (0, t)$, $\tau \mapsto (\bar{\gamma}(\tau), \bar{\psi}(\tau), p(\tau))$ is of class $C^1$ and it satisfies the characteristic equation (2) where

$$
\bar{u}(\tau) = u(\bar{\gamma}(\tau), \tau) \quad \text{and} \quad p(\tau) = \frac{\partial L}{\partial x}(\bar{\gamma}(\tau), u(\bar{\gamma}(\tau), \tau), \dot{\bar{\gamma}}(\tau)).
$$

By analogy of the notion of weak KAM solutions of the Hamilton-Jacobi equation independent of $u$ (see [18]), we define another weak solution of (1) called variational solution under (H1)-(H4). In particular, for the discounted equation as (1), from which there exists an implicitly defined semigroup denoted by $T_t$ such that $u(x, t) = T_t \varphi(x)$, where $u(x, t)$ satisfies (3). To fix the notion, we call $T_t$ a solution semigroup. The name of solution semigroup came from [14] by A. Douglis under more restricted assumptions. In particular, for the discounted Hamiltonians $H(x, u, p) := \lambda u + H(x, p)$ with $\lambda > 0$, the solution semigroup can be reduced to the Lax-Oleinik semigroup

$$
T_t \varphi(x) = \inf_{\gamma(0) = x} \left\{ e^{-t} \varphi(\gamma(-t)) + \int_{-t}^{0} e^{s} L(\gamma(s), \dot{\gamma}(s)) ds \right\}. \quad (4)
$$

Let $a \in \mathbb{R}$ be a constant. For $u \equiv a$, we use $c(H(x, a, p))$ to denote the Mañé critical value of $H(x, a, p)$. From [10], we have

$$
c(H(x, a, p)) = \inf_{u \in C^1(M, \mathbb{R})} \sup_{x \in M} H(x, a, \partial_x u). \quad (5)
$$

Let

$$
C_H = \bigcup_{a \in \mathbb{R}} c(H(x, a, p)). \quad (6)
$$

Under the assumptions (H1)-(H4), a crucial result in weak KAM theory, the uniqueness of the Mañé’s critical value, does not hold any more. We call $C_H$ an admissible value set of $H(x, a, p)$ (and $L(x, a, \dot{x})$ by Legendre transformation). It is easy to see $C_H \neq \emptyset$. Indeed, $C_H$ is a non-empty interval (see Proposition 5.3 below). For any $c \in C_H$, there exists $a \in \mathbb{R}$ such that $c(H(x, a, p)) = c$. Let $L_c := L + c$. $H_c$ and $T_t^c$ are the Hamiltonian and solution semigroup associated to $L_c$. Let $\| \cdot \|_{\infty}$ be $C^0$-norm. We have the following theorem.

**Theorem 1.2.** For any $\varphi(x) \in C(M, \mathbb{R})$, the viscosity solution $u(x, t)$ of (1) with initial condition $u(x, 0) = \varphi(x)$ can be represented as (3).

Theorem 1.1 provides a variational principle on the evolutionary Hamilton-Jacobi equation as (1), from which there exists an explicitly defined semigroup denoted by $T_t$ such that $u(x, t) = T_t \varphi(x)$, where $u(x, t)$ satisfies (3). Following [18], we show the viscosity solution of (1) with the continuous initial condition is unique (see [11]) and it is a locally semiconcave function (see [7]). Moreover, we obtain a representation formula of the viscosity solution under (H1)-(H4). By analogy of the notion of weak KAM solutions of the Hamilton-Jacobi equation (see [18]), we define another weak solution of (1) called variational solution (see Definition 2.4 below). Based on Theorem 1.1, we construct a variational principle on the evolutionary Hamilton-Jacobi equation (2) where

$$
\bar{u}(\tau) = u(\bar{\gamma}(\tau), \tau) \quad \text{and} \quad p(\tau) = \frac{\partial L}{\partial x}(\bar{\gamma}(\tau), u(\bar{\gamma}(\tau), \tau), \dot{\bar{\gamma}}(\tau)).
$$

Theorem 1.3. For any $\varphi(x), \psi(x) \in C(M, \mathbb{R})$, $t \geq 0$ and $c \in C_H$, the solution semigroup $T_t^c$ has the following properties:

(i.) **Monotonicity:** for $\varphi \leq \psi$, $T_t^c \varphi \leq T_t^c \psi$;

(ii.) **Non-expansiveness:** $\| T_t^c \varphi - T_t^c \psi \|_{\infty} \leq \| \varphi - \psi \|_{\infty}$;

(iii.) **Uniform Boundedness:** there exists a positive constant $K$ independent of $t$ such that $\| T_t^c \varphi \|_{\infty} \leq K$;
(iv.) **Equi-Lipschitzity:** given $\delta > 0$, there exists a constant $\kappa_\delta > 0$ such that for $t \geq \delta > 0$, $x \mapsto T^c_t \varphi(x)$ is $\kappa_\delta$-Lipschitzian on $M$.

For the autonomous systems with Lagrangian $L(x, \dot{x})$, the convergence of so-called Lax-Oleinik semigroup was established in [17]. By [19], such convergence fails for the non-autonomous Lagrangian systems. A new kind of operators was found in [27] to overcome the failure of the convergence of the Lax-Oleinik semigroup for the time periodic Lagrangian systems.

Different from the previous results, the solution semigroup $T^c_t$ here is associated to $\bar{L}(t, x, \dot{x}) := L(x, u(x,t), \dot{x}) + c$, which is defined on $\mathbb{R} \times TM$. Consequently, it results in the lack of conservation of energy of the system and compactness of the underlying manifold. Either the conservation or the compactness are crucial for the previous results. Hence, it is necessary to find a completely new dynamical way for establishing the convergence of the solution semigroup $T^c_t$.

Based on Theorem 1.1 and Theorem 1.3, we obtain the convergence of the solution semigroup $T^c_t$ by considering the evolution of $H^c$ along the characteristics.

**Theorem 1.4.** For any $\varphi(x) \in C(M, \mathbb{R})$, we have the dichotomy:

(i) if $c \notin C_H$, then $T^c_t \varphi(x)$ diverges as $t \to +\infty$;

(ii) if $c \in C_H$, then $T^c_t \varphi(x)$ converges as $t \to +\infty$. Moreover, let

$$u^c_\infty(x) := \lim_{t \to +\infty} T^c_t \varphi(x),$$

then $u^c_\infty$ is a weak KAM solution of the following stationary equation:

$$H(x, u(x), \partial_x u(x)) = c. \quad (7)$$

By inspiration of [17], the large time behavior of viscosity solutions of Hamilton-Jacobi equations with Hamiltonian independent of $u$ was explored comprehensively based on both dynamical and PDE approaches (see [13, 20, 26] for instance). Theorem 1.2 implies $u(x,t) := T^c_t \varphi(x)$ is the unique viscosity solution of

$$\begin{cases}
\partial_t u(x,t) + H(x, u(x,t), \partial_x u(x,t)) = 0, \\
u(x,0) = \varphi(x).
\end{cases} \quad (8)$$

The weak KAM solution of (7) is the same as the viscosity solution. As an application of Theorem 1.4, if $0 \in C_H$, we obtain the large time behavior of the viscosity solution of (8), which appears here for the first time in this generality. In particular, for the discounted Hamiltonians $H(x,u,p) := \lambda u + \bar{H}(x,p)$ with $\lambda > 0$ and Tonelli Hamiltonian $\bar{H}(x,p)$, since $C_H = \mathbb{R}$, then the convergence in Theorem 1.4 always holds.

Note that the characteristics of (7) reads

$$\begin{cases}
\dot{x} = \frac{\partial H}{\partial p}, \\
\dot{p} = -\frac{\partial H}{\partial x} - \frac{\partial H}{\partial u} p, \\
\dot{u} = \frac{\partial H}{\partial p} p.
\end{cases} \quad (9)$$

It is easy to see that $H$ is conservative along the characteristics. Comparably, $H$ is dissipative along the characteristics of (8). Roughly speaking, Theorem 1.4 shows a conservative system can be viewed as the limit of certain dissipative system in the sense of viscosity.

This paper is outlined as follows. In Section 2, some definitions are recalled as preliminaries. In Section 3, an implicitly variational principle is established. Moreover, Theorem 1.1 can be obtained. In Section 4, a representation of the
viscosity solution is provided, which implies Theorem 1.2. In Section 5, an implicitly
defined solution semigroup is introduced and some properties are detected, from
which Theorem 1.3 is proved. In Section 6, both the divergence and convergence of
the solution semigroup are shown. Moreover, Theorem 1.4 can be verified.

2. Preliminaries. In this section, we recall the definitions of the weak KAM
solution and the viscosity solution of (1) (see [8, 11, 18]) and some aspects of Mather
type for the sake of completeness.

2.1. Weak KAM solution and viscosity solution. A function

\[ H(x, \partial_x u(x)) : TM \to \mathbb{R} \]

is called a Tonelli Hamiltonian if it satisfies (H1)-(H2). For the autonomous
Hamiltonian systems. The associated Lagrangian is denoted by \( L(x, \dot{x}) \) via the
Legendre transformation. In [16], Fathi introduced the definition of the weak KAM
solution of the following Hamilton-Jacobi equation:

\[
H(x, \partial_x u(x)) = c, \quad x \in M, \quad (10)
\]

where \( H \) is a Tonelli Hamiltonian and \( c[0] \) is the Mañé’s critical value of \( H \). We
consider the weak KAM solution of negative type here.

**Definition 2.1.** A function \( u \in C(M, \mathbb{R}) \) is called a weak KAM solution of negative
type of (10) if

(i) for each continuous piecewise \( C^1 \) curve \( \gamma : [t_1, t_2] \to M \) where \( t_2 > t_1 \), we have

\[
u(\gamma(t_2)) - u(\gamma(t_1)) \leq \int_{t_1}^{t_2} L(\gamma(\tau), \dot{\gamma}(\tau)) + c[0]d\tau; \quad (11)
\]

(ii) for any \( x \in M \), there exists a \( C^1 \) curve \( \gamma : (-\infty, 0] \to M \) with \( \gamma(0) = x \) such
that for any \( t \geq 0 \), we have

\[
u(x) - u(\gamma(-t)) = \int_{-t}^{0} L(\gamma(\tau), \dot{\gamma}(\tau)) + c[0]d\tau. \quad (12)
\]

A function \( u \) is called dominated by \( L + c[0] \) if it satisfies (i) in Definition 2.1.
It is denoted by \( u \prec L + c[0] \). For Tonelli Hamiltonians, weak KAM solutions of
negative type of (10) are equivalent to viscosity solutions of (10). By analogy of the
definition above, for \( c \in \mathcal{C}_H \), we introduce the weak KAM solution of negative type
of more general Hamilton-Jacobi equation as follows:

\[
H(x, u(x), \partial_x u(x)) = c, \quad x \in M, \quad (13)
\]

where \( \mathcal{C}_H \) is the admissible value set of \( H \) defined in (6).

**Definition 2.2.** A function \( u \in C(M, \mathbb{R}) \) is called a weak KAM solution of negative
type of (13) if

(i) for each continuous piecewise \( C^1 \) curve \( \gamma : [t_1, t_2] \to M \) where \( t_2 > t_1 \), we have

\[
u(\gamma(t_2)) - u(\gamma(t_1)) \leq \int_{t_1}^{t_2} L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) + cd\tau; \quad (14)
\]

(ii) for any \( x \in M \), there exists a \( C^1 \) curve \( \gamma : (-\infty, 0] \to M \) with \( \gamma(0) = x \) such
that for any \( t \geq 0 \), we have

\[
u(x) - u(\gamma(-t)) = \int_{-t}^{0} L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) + cd\tau. \quad (15)
\]

Following from [8, 11, 18], a viscosity solution of (1) can be defined as follows.
Definition 2.3. Let $V$ be an open subset $V \subset M$,
(i) A function $u : V \times [0, T] \to \mathbb{R}$ is called a subsolution of (1), if for every $C^1$ function $\varphi : V \times [0, T] \to \mathbb{R}$ and every point $(x_0, t_0) \in V \times [0, T]$ such that $u - \varphi$ has a maximum at $(x_0, t_0)$, we have
$$\partial_t \varphi(x_0, t_0) + H(x_0, u(x_0, t_0), \partial_x \varphi(x_0, t_0)) \leq 0;$$
(16)
(ii) A function $u : V \times [0, T] \to \mathbb{R}$ is a called supersolution of (1), if for every $C^1$ function $\psi : V \times [0, T] \to \mathbb{R}$ and every point $(x_0, t_0) \in V \times [0, T]$ such that $u - \psi$ has a minimum at $(x_0, t_0)$, we have
$$\partial_t \psi(x_0, t_0) + H(x_0, u(x_0, t_0), \partial_x \psi(x_0, t_0)) \geq 0;$$
(17)
(iii) A function $u : V \times [0, T] \to \mathbb{R}$ is called a viscosity solution of (1) if it is both a subsolution and a supersolution.

In the sequel, if not otherwise stated, solutions, subsolutions and supersolutions will be always meant in the viscosity sense, hence the adjective viscosity will be omitted in the following.

Both of Definition 2.1 and Definition 2.2 are concerned with the weak KAM solutions defined on $M$, while the solutions of (1) are defined on $M \times [0, T]$. As a bridge connecting them, we provide the definition of another weak solution of (1), called a variational solution.

Definition 2.4. For a given $T > 0$, a variational solution of (1) is a function $u : M \times [0, T] \to \mathbb{R}$ for which the following are satisfied:
(i) for each continuous piecewise $C^1$ curve $\gamma : [t_1, t_2] \to M$ where $0 \leq t_1 < t_2 \leq T$, we have
$$u(\gamma(t_2), t_2) - u(\gamma(t_1), t_1) \leq \int_{t_1}^{t_2} L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau;$$
(18)
(ii) for any $0 \leq t_1 < t_2 \leq T$ and $x \in M$, there exists a $C^1$ curve $\gamma : [t_1, t_2] \to M$ with $\gamma(t_2) = x$ such that
$$u(x, t_2) - u(\gamma(t_1), t_1) = \int_{t_1}^{t_2} L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau.$$
From the definition of $h^t(x,y)$, it follows that for each $x,y,z \in M$ and each $t,t' > 0$, we have
$$h^{t+t'}(x,z) \leq h^t(x,y) + h^{t'}(y,z).$$
(21)

In particular, we have
$$h^{t+t'}(x,y) = h^t(x,\bar{\gamma}(t)) + h^{t'}(\bar{\gamma}(t),y),$$
where $\bar{\gamma}$ is a minimal curve with $\bar{\gamma}(0) = x$ and $\bar{\gamma}(t+t') = y$.

Let $c[0]$ be the Mañé critical value of $L(x,\dot{x})$. The Peierls barrier is the function $h^\infty_{c[0]} : M \times M \to \mathbb{R}$ defined by
$$h^\infty_{c[0]}(x,y) = \lim_{t \to +\infty} h^t(x,y) + c[0]t.$$

For autonomous Lagrangians, “liminf” can be replaced by “lim”. If a function $u \in C(M,\mathbb{R})$ is dominated by $L + c[0]$, it is easy to see that
$$h^t_{c[0]}(x,y) := h^t(x,y) + c[0]t \geq u(y) - u(x) \geq -2\|u\|_{\infty}.$$
(23)

Moreover, there holds [18, Lemma 5.3.2]

**Proposition 2.5.** For each $t_0 > 0$, there exists a constant $C_{t_0}$ such that for any $t \geq t_0$ and $x,y \in M$,
$$|h^t_{c[0]}(x,y)| \leq C_{t_0}.$$  
(24)

Throughout this paper, we shall use $|\cdot|$ to denote the Euclidean norm, that is $|\alpha| = \sqrt{\alpha_1^2 + \ldots + \alpha_i^2}$ for given $\alpha = (\alpha_1, \ldots, \alpha_i) \in \mathbb{R}^i$, $i = 1$ or $i = n$.

### 3. Variational principle

For every given continuous function $\varphi$ on $M$, we define the operator $\mathbb{A}$ depending on $\varphi$ as follows:
$$\mathbb{A}[u](x,t) = \inf_{\gamma \in C^{C^{\infty}}([0,t],M)} \left\{ \varphi(\gamma(0)) + \int_0^t L(\gamma(s),u(\gamma(s),\dot{\gamma}(s)) \, ds \right\},$$
(25)

where $u(x,t) \in C(M \times [0,T],\mathbb{R})$ and $C^{C^{\infty}}([0,t],M)$ denotes the set of absolutely continuous curves $\gamma : [0,t] \to M$. It is easy to see that an operator from $C(M \times [0,T],\mathbb{R})$ to itself. For a given $u(x,t) \in C(M \times [0,T],\mathbb{R})$, let
$$\tilde{L}(t,x,\dot{x}) := L(x,u(x,t),\dot{x}).$$

Consider the action
$$\inf_{\gamma(0) = x} \int_0^t \tilde{L}(\tau,\gamma(\tau),\dot{\gamma}(\tau)) \, d\tau,$$
(26)

where the infimum is taken among the absolutely continuous curves $\gamma : [0,t] \to M$. Moreover, we have Tonelli existence theorem.

**Lemma 3.1.** For any $x,y \in M$ and $t > 0$, the infimum in (26) can be attained at an absolutely continuous curve with $\gamma(0) = x$ and $\gamma(t) = y$.

Note that $\tilde{L}$ is only continuous with respect to $x$ and $t$ rather than $C^2$. Fortunately, the loss of regularity of $\tilde{L}$ with respect to $x$ and $t$ does not cause any trouble in the proof of Lemma 3.1. We omit it for the consistency of the context, see [6, Page 114] for the details. As a corollary, the infimum in (25) can be also attained.
3.1. The fixed point of A. In the following, we will prove that the operator A has a unique fixed point.

**Lemma 3.2.** A has a unique fixed point.

**Proof.** From (L3), it follows that for every $u, v \in C(M \times [0, T], \mathbb{R})$,

$$|L(x, u, \dot{x}) - L(x, v, \dot{x})| \leq \lambda |u - v|,$$

where $\lambda$ is a positive constant independent of $x$ and $\dot{x}$. Hence, for any given $t \in [0, T]$, it follows from Lemma 3.1 that

$$\begin{align*}
(A[u] - A[v])(x, t) \\
\leq \int_0^t (L(\gamma_1(s), u(\gamma_1(s), s), \dot{\gamma}_1(s)) - L(\gamma_1(s), v(\gamma_1(s), s), \dot{\gamma}_1(s))) \, ds \\
\leq \lambda \|u - v\|_\infty t
\end{align*}$$

where $\gamma_1 \in C^{ac}([0, t], M)$ such that $\gamma_1(t) = x$ and

$$A[v](x, t) = \varphi(\gamma_1(0)) + \int_0^t L(\gamma_1(s), v(\gamma_1(s), s), \dot{\gamma}_1(s)) \, ds. \quad (27)$$

By exchanging the position of $u$ and $v$, we obtain

$$|(A[u] - A[v])(x, t)| \leq \lambda \|u - v\|_\infty t. \quad (28)$$

Let $\gamma_2 : [0, t] \to M$ be the curve such that

$$A^2[v](x, t) = \varphi(\gamma_2(0)) + \int_0^t L(\gamma_2(s), A[v](\gamma_2(s), s), \dot{\gamma}_2(s)) \, ds. \quad (29)$$

It follows from (28) that for $s \in [0, t]$, we have

$$|(A[u] - A[v])(\gamma_2(s), s)| \leq \lambda \|u - v\|_\infty s.$$

Moreover, we have the following estimates:

$$\begin{align*}
|(A^2[u] - A^2[v])(x, t)| \\
\leq \int_0^t \lambda |A[u](\gamma_2(s), s) - A[v](\gamma_2(s), s)| \, ds,
\leq \int_0^t s \lambda^2 \|u - v\|_\infty \, ds \leq \frac{(t\lambda)^2}{2} \|u - v\|_\infty.
\end{align*}$$

Moreover, continuing the above procedure, we obtain

$$\begin{align*}
|&(A^n[u] - A^n[v])(x, t)| \\
\leq &\frac{(T\lambda)^n}{n!} \|u - v\|_\infty,
\end{align*}$$

which implies

$$\|A^n[u] - A^n[v](x, t)\|_\infty \leq \frac{(T\lambda)^n}{n!} \|u - v\|_\infty. \quad (31)$$

Therefore, for any $t \in [0, T]$, there exists $N \in \mathbb{N}$ large enough such that $A^N$ is a contraction mapping and has a fixed point. That is, for any $t \in [0, T]$ and $N \in \mathbb{N}$ large enough, there exists a $u(x, t) \in C(M \times [0, T], \mathbb{R})$ such that

$$A^N[u](x, t) = u(x, t). \quad (32)$$

We now show that $u$ is a fixed point of $A$. Since

$$A[u] = A \circ A^N[u] = A^N \circ A[u],$$

we have

$$u(x, t) = A^N[u](x, t).$$

Thus, $u$ is a fixed point of $A$. 

**Remark.** The above procedure also establishes the existence of a fixed point of $A^n$ for any $n \geq 1$.
A[u] is also a fixed point of \(A^N\). By the uniqueness of fixed point of contraction mapping, we have
\[
A[u] = u.
\]
This completes the proof of Lemma 3.2.

Lemma 3.2 shows that there exists \(u(x, t) \in C(M \times [0, T], \mathbb{R})\) such that
\[
u(x, t) = \inf_{\gamma(t) = x} \left\{ \varphi(\gamma(0)) + \int_0^t L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau)) d\tau \right\}. \tag{33}
\]
To fix the notions, we call the curve \(\gamma\) achieving the infimum in (33) a minimizer of \(u\) with \(\gamma(t) = x\).

3.2. Minimizers and characteristics. In the following, we will show the relation between minimizers and characteristics of (1). More precisely, we have the following lemma.

**Lemma 3.3.** Let \(\bar{\gamma} : [0, t] \rightarrow M\) be a minimizer of \(u\), then \(\bar{\gamma}\) is a \(C^1\) curve. For \(\tau \in (0, t), \tau \rightarrow (\bar{\gamma}(\tau), u(\bar{\gamma}(\tau), p(\tau))\) is of class \(C^1\) and it satisfies the characteristic equation (2) where
\[
\bar{u}(\tau) = u(\bar{\gamma}(\tau), \tau) \quad \text{and} \quad p(\tau) = \frac{\partial L}{\partial x}(\bar{\gamma}(\tau), u(\bar{\gamma}(\tau), \tau), \dot{\gamma}(\tau)). \tag{34}
\]

In order to prove Lemma 3.3, we need the following lemma.

**Lemma 3.4.** Let \(X_{x_0, u_0}(t, p)\) be the first argument of \(\Psi_{t-t_0}(x_0, u_0, p)\), where \(\Psi\) denotes the flow generated by \(H\). For any compact set \(K \subset M \times \mathbb{R}\) and \(k > 0\), there exists \(\varepsilon := \varepsilon(K, k) > 0\) such that for \((x_0, u_0) \in K, |p| \leq k\) and \(t \in (t_0, t_0 + \varepsilon]\), \(p \mapsto X_{x_0, u_0}(t, p)\) is a \(C^1\) diffeomorphism onto its image.

**Proof.** Since \(M\) is compact, there exists \(\varepsilon := \varepsilon(K, k)\) such that for any \((x_1, u_1) \in K\) and \(0 < t \leq \varepsilon\), we can choose a domain of coordinate chart \(U\) such that the set \(\{x : x = X_{x_1, u_1}(t, p), p \in T^*_x M, |p| \leq k\}\) is contained in \(U\). Hence, we conclude that it suffices to prove the lemma for the case when \(M\) is an open subset of \(\mathbb{R}^n\).

In the sequel of the proof, we will thus suppose that \(M = U\) is an open subset of \(\mathbb{R}^n\), and thus \(T^*U = U \times \mathbb{R}^n\).

First of all, we prove for given \((x_0, u_0)\), one can find \(\varepsilon > 0\) such that \(p \mapsto X_{x_0, u_0}(t, p)\) is a \(C^1\) diffeomorphism onto its image for \(|p| \leq k\) and \(t \in (t_0, t_0 + \varepsilon]\). Without ambiguity, we omit \(x_0, u_0\) of \(X_{x_0, u_0}(t, p)\). We consider the difference
\[
\Delta_{x_0, u_0}(t, p) := (X(t, p) - X(t_0, p)) - (t - t_0) \frac{\partial X}{\partial t}(t_0, p). \tag{35}
\]
It is clear that
\[
X(t_0, p) = x_0, \quad \frac{\partial X}{\partial t}(t_0, p) = \frac{\partial H}{\partial p}(x_0, u_0, p).
\]
Thus, \(\Delta_{x_0, u_0}(t, p)\) is \(C^1\) with respect to \(p\). Then we have
\[
\frac{\partial^2 X}{\partial p \partial t}(t_0, p) = \frac{\partial^2 H}{\partial p^2}(x_0, u_0, p), \tag{36}
\]
and
\[
\frac{1}{t - t_0} \frac{\partial \Delta_{x_0, u_0}}{\partial p}(t, p) = \frac{\partial X}{\partial p}(t, p) - \frac{\partial X}{\partial p}(t_0, p) - \frac{\partial^2 H}{\partial p^2}(x_0, u_0, p). \tag{37}
\]
Since \(X(t, p)\) is the first argument of \(\Psi_{t-t_0}(x_0, u_0, p)\), we denote
\[
(X(s, p), U(s, p), P(s, p)) := \Psi_{s-t_0}(x_0, u_0, p).
\]
Moreover, we have

$$X(t, p) = X(t_0, p) + \int_{t_0}^t \frac{\partial H}{\partial P}(\Psi_{s-t_0}(x_0, u_0)) ds.$$ 

By differentiation under the integral sign, we have

$$\frac{\partial X(t, p)}{\partial p} = \frac{\partial X(t_0, p)}{\partial p} + \int_{t_0}^t \frac{\partial^2 H}{\partial X \partial P} \frac{\partial X(s, p)}{\partial p} + \frac{\partial^2 H}{\partial U \partial P} \frac{\partial U(s, p)}{\partial p} + \frac{\partial^2 H}{\partial P^2} \frac{\partial P(s, p)}{\partial p} ds,$$

where

$$\frac{\partial X}{\partial p} := \left( \frac{\partial X_i}{\partial P_j} \right)_{n \times n}, \quad \frac{\partial P}{\partial p} := \left( \frac{\partial P_i}{\partial P_j} \right)_{n \times n}, \quad \frac{\partial U}{\partial p} := \left( \frac{\partial U}{\partial P_i} \right)_{1 \times n},$$

$$\frac{\partial^2 H}{\partial X \partial P} := \left( \frac{\partial^2 H}{\partial X_i \partial P_j} \right)_{n \times n}, \quad \frac{\partial^2 H}{\partial P^2} := \left( \frac{\partial^2 H}{\partial P_i \partial P_j} \right)_{n \times n}, \quad \frac{\partial^2 H}{\partial U \partial P} := \left( \frac{\partial^2 H}{\partial U \partial P_i} \right)_{n \times 1}.$$ 

Moreover, we have

$$\frac{\partial^2 X(t, p)}{\partial t \partial p} = \frac{\partial^2 H}{\partial X \partial P} \frac{\partial X(t, p)}{\partial p} + \frac{\partial^2 H}{\partial U \partial P} \frac{\partial U(t, p)}{\partial p} + \frac{\partial^2 H}{\partial P^2} \frac{\partial P(t, p)}{\partial p},$$

(38)

Since \( X(t_0, p) = x_0 \) and \( U(t_0, p) = u_0 \) are fixed, then \( \frac{\partial X(t_0, p)}{\partial p} \) and \( \frac{\partial U(t_0, p)}{\partial p} \) vanish. It follows from \( P(t_0, p) = p \) that

$$\frac{\partial^2 X(t_0, p)}{\partial P^2}(X(t_0, p), U(t_0, p), P(t_0, p)) \cdot \frac{\partial P(t_0, p)}{\partial p} = \frac{\partial^2 H}{\partial P^2}(x_0, u_0, p),$$

which together with (36) and (38) implies that \( \frac{\partial^2 X}{\partial t \partial p}(t_0, p) \) exists and

$$\frac{\partial^2 X}{\partial t \partial p}(t_0, p) = \frac{\partial^2 X}{\partial t \partial p}(t_0, p) = \frac{\partial^2 H}{\partial P^2}(x_0, u_0, p).$$

(39)

As \( t \to t_0 \), we have

$$\frac{\partial X}{\partial p}(t, p) - \frac{\partial X}{\partial p}(t_0, p) \to \frac{\partial^2 X}{\partial t \partial p}(t_0, p),$$

which together with (37) and (39) yields as \( t \to t_0 \),

$$\frac{1}{t - t_0} \frac{\partial \Delta x_0, u_0}{\partial p}(t, p) \to 0.$$ 

(40)

On the other hand, we have

$$\frac{\partial X}{\partial p}(t, p) = (t - t_0) \left( \frac{1}{t - t_0} \frac{\partial \Delta}{\partial p}(t, p) + \frac{\partial^2 H}{\partial P^2}(x_0, u_0, p) \right).$$

Combining with (H1) and the compactness of \( p \), we obtain that there exists \( \varepsilon > 0 \) small enough such that for for any \( |p| \leq k \) and \( \tau \in (t_0, t_0 + \varepsilon] \),

$$0 < \frac{1}{C} \leq \text{det} \left( \frac{\partial X}{\partial p}(t, p) \right) \leq C,$$

where \( \text{det} \) denotes the Jacobian determinant and \( C := C(k, \varepsilon) \) denotes a positive constant, thus \( p \mapsto X_{x_0, u_0}(t, p) \) is injective for any \( |p| \leq k \) and \( \tau \in (t_0, t_0 + \varepsilon] \). Moreover, it is a \( C^1 \) diffeomorphism onto its image for a given \( (x_0, u_0) \).
Second, we verify the uniform existence of \( \varepsilon \) for any \((x_0, u_0) \in \mathcal{K}\). Choosing a local coordinate chart in a neighborhood of \((x_0, u_0)\). Consider the map
\[
p \mapsto \frac{\partial H}{\partial p}(x_0, u_0, p) + \frac{1}{t - t_0} \Delta_{x_0, u_0}(t, p).
\]
By (H1), for any \((x_0, u_0) \in \mathcal{K}\), \(|p| \leq k\),
\[
p \mapsto \frac{\partial H}{\partial p}(x_0, u_0, p)
\]
is a \(C^1\) diffeomorphism onto its image.

Note that the limiting passage in (40) is uniform for any \((x_0, u_0) \in \mathcal{K}\), \(|p| \leq k\).
By (35),
\[
\frac{1}{t - t_0} \Delta_{x_0, u_0}(t, p) \to 0 \quad \text{as } t \to 0,
\]
which is also uniform for any \((x_0, u_0) \in \mathcal{K}\), \(|p| \leq k\). Moreover, for any \(\delta > 0\), one can find \(\varepsilon := \varepsilon(\mathcal{K}, k) > 0\) small enough such that for any \((x_0, u_0) \in \mathcal{K}\), \(|p| \leq k\) and \(\tau \in (t_0, t_0 + \varepsilon]\),
\[
\left\|\frac{1}{t - t_0} \Delta_{x_0, u_0}(t, p)\right\|_{C^1} < \delta.
\]
Taking \(\delta > 0\) small enough, it implies the map
\[
p \mapsto \frac{\partial H}{\partial p}(x_0, u_0, p) + \frac{1}{t - t_0} \Delta_{x_0, u_0}(t, p)
\]
a \(C^1\) diffeomorphism onto its image. Since
\[
X_{x_0, u_0}(t, p) = x_0 + (t - t_0) \left(\frac{1}{t - t_0} \Delta_{x_0, u_0}(t, p) + \frac{\partial H}{\partial p}(x_0, u_0, p)\right),
\]
thus, for any \((x_0, u_0) \in \mathcal{K}\), \(|p| \leq k\), there exists \(\varepsilon := \varepsilon(\mathcal{K}, k) > 0\) such that for \(\tau \in (t_0, t_0 + \varepsilon]\), \(p \mapsto X_{x_0, u_0}(t, p)\) is a \(C^1\) diffeomorphism onto its image. \(\square\)

Proof of Lemma 3.3. The proof is divided into three steps. Note that (2) is the characteristic equation of the Hamilton-Jacobi equation:
\[
\partial_\tau S(x, \tau) + H(x, S(x, \tau), \partial_x S(x, \tau)) = 0.
\]

(a) Construction of the classical solution
First of all, we will construct a classical solution of (41) on a cone-like region (see (43) below). Since \(\hat{\gamma} \in C^{\text{loc}}([0, t], M)\), then the derivative \(\dot{\hat{\gamma}}(\tau)\) exists almost everywhere for \(\tau \in [0, t]\). Let \(t_0 \in (0, t)\) be a differentiate point of \(\hat{\gamma}(\tau)\). It suffices to consider \(0 < t_0 < t\). Denote
\[
(x_0, u_0, v_0) := (\hat{\gamma}(t_0), u(\hat{\gamma}(t_0), t_0), \dot{\hat{\gamma}}(t_0)).
\]
Let \(k_1 := |v_0|\) and
\[
B(0, 2k_1) := \{v : |v| < 2k_1, \ v \in T_{x_0}M\}.
\]
We use \(B^*(0, 2k_1)\) to denote the image of \(B(0, 2k_1)\) via the Legendre transformation \(L^{-1} : TM \to T^*M\). That is
\[
B^*(0, 2k_1) := \left\{p : p = \frac{\partial L}{\partial v}(x_0, u_0, v), \ v \in B(0, 2k_1)\right\}.
\]
Let \(\Psi: T^*M \times \mathbb{R} \to T^*M \times \mathbb{R}\) denote the flow generated by the contact Hamiltonian equation (2). Let \(\pi\) be a projection from \(T^*M \times \mathbb{R}\) to \(M\) via \((x, u, p) \to x\) and let
\[
B^*_\pi(0, 2k_1) := \pi \circ \Psi_{t_0}(x_0, u_0, B^*(0, 2k_1)).
\]
Moreover, we denote
\[ X_{x_0, u_0}(\tau, \cdot) : B^*(0, 2k_1) \to B^*_t(0, 2k_1), \quad \text{via } p \mapsto \Psi_{\tau-t_0}(x_0, u_0, p). \]

It follows from Lemma 3.4 that for any compact set \( K \subset M \times \mathbb{R} \), there exists \( \varepsilon := \varepsilon(K, k_2) > 0 \) small enough such that for \( (x_0, u_0) \in K \) and \( \tau \in (t_0, t_0 + \varepsilon) \),
\[ X_{x_0, u_0}(\tau, \cdot) \text{ is a } C^1 \text{ diffeomorphism onto its image.} \]

We use \( \Omega_\varepsilon \) to denote the following cone-like region:
\[ \Omega_\varepsilon := \{(\tau, x) : \tau \in (t_0, t_0 + \varepsilon), x \in B^*_\tau(0, 2k_1)\}. \]

Then for any \((\tau, x) \in \Omega_\varepsilon\), there exists a unique \( p_0(x) \in B^*(0, 2k_1) \) such that
\[ X_{x_0, u_0}(\tau, p_0(x)) = x \]
where
\[ (X_{x_0, u_0}(\tau, p_0(x)), U_{x_0, u_0}(\tau, p_0(x)), P_{x_0, u_0}(\tau, p_0(x))) := \Psi_{\tau}(x_0, u_0, p_0(x)). \]

Hence, for any \((\tau, x) \in \Omega_\varepsilon\), one can define a \( C^1 \) function by \( S(x, \tau) = U_{x_0, u_0}(\tau, p_0(x)) \).

In particular, we have \( S(x, t_0) = u_0 \). Moreover, in light of the method of characteristics (see [2, Page 15, Theorem 3]) that \( S(x, \tau) \), defined on \( \Omega_\varepsilon \), is a \( C^1 \) classical solution of the Hamilton-Jacobi equation (41) with \( S(x, t_0) = u_0 \).

Without ambiguity, we denote \( X(\tau) := X_{x_0, u_0}(\tau, p_0(x)) \) for a given \( x \in M \).

(b) **Local coincidence between \( \hat{\gamma}(\tau) \) and \( X(\tau) \)**

Fix \( \tau \in (t_0, t_0 + \varepsilon] \) and let \( S_\tau(x) := S(x, \tau) \). We denote
\[ \text{grad}_L S_\tau(x) := \frac{\partial H}{\partial p}(x, S_\tau(x), p), \]
where \( p = \partial_x S_\tau(x) \). It is easy to see that \( \text{grad}_L S_\tau(x) \) gives rise to a vector field on \( M \).

**Claim A.** Let \( \gamma \) be a continuous and piecewise \( C^1 \) curve with \((\tau, \gamma(\tau)) \in \Omega_\varepsilon \) for \( \tau \in [a, b] \subset [t_0, t_0 + \varepsilon] \), we have
\[ S(\gamma(b), b) - S(\gamma(a), a) \leq \int_a^b L(\gamma(\tau), S(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau, \quad (45) \]
where the equality holds if and only if \( \gamma \) is a trajectory of the vector field \( \text{grad}_L S_\tau(x) \).

**Proof of Claim A.** From the regularity of \( S(x, \tau) \), it follows that
\[ S(\gamma(b), b) - S(\gamma(a), a) = \int_a^b \left\{ \frac{\partial S}{\partial t}(\gamma(\tau), \tau) + \left\langle \frac{\partial S}{\partial x}(\gamma(\tau), \tau), \dot{\gamma}(\tau) \right\rangle \right\} d\tau. \quad (46) \]

By virtue of Fenchel inequality, for each \( \tau \) where \( \dot{\gamma}(\tau) \) exists, we have
\[ \left\langle \frac{\partial S}{\partial x}(\gamma(\tau), \tau), \dot{\gamma}(\tau) \right\rangle \leq H(\gamma(\tau), S(\gamma(\tau), \tau), \frac{\partial S}{\partial x}(\gamma(\tau), \tau)) \]
\[ + L(\gamma(\tau), S(\gamma(\tau), \tau), \dot{\gamma}(\tau)). \]

It follows from (41) that for almost every \( \tau \in [a, b] \)
\[ \frac{\partial S}{\partial t}(\gamma(\tau), \tau) + \left\langle \frac{\partial S}{\partial x}(\gamma(\tau), \tau), \dot{\gamma}(\tau) \right\rangle \leq L(\gamma(\tau), S(\gamma(\tau), \tau), \dot{\gamma}(\tau)). \quad (47) \]

By integration, it follows from (46) that
\[ S(\gamma(b), b) - S(\gamma(a), a) \leq \int_a^b L(\gamma(\tau), S(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau. \quad (48) \]
We have equality in (48) if and only if the equality holds almost everywhere in the Fenchel inequality, i.e. 
\( \dot{\gamma}(\tau) = \text{grad}_L S_\tau(x) \). Since \( \text{grad}_L S_\tau(x) \) is continuous and defined for each \( t \in [a, b] \), it follows that \( \dot{\gamma}(t) \) can be extended by continuity to the whole interval \([a, b]\), which means that \( \gamma \) is a trajectory of the vector field \( \text{grad}_L S_\tau(x) \).

\[ \Box \]

**Claim B.** For any \( \tau \in [t_0, t_0 + \varepsilon] \), there holds
\[ S(\hat{\gamma}(\tau), \tau) = u(\hat{\gamma}(\tau), \tau), \]
where \( \hat{\gamma} \) is a minimizer of \( u \) with \( \hat{\gamma}(t_0) = x_0 \).

**Proof of Claim B.** By contradiction, we assume there exists \( \hat{t} \in [t_0, t_0 + \varepsilon] \) such that \( S(\hat{\gamma}(\hat{t}), \hat{t}) \neq u(\hat{\gamma}(\hat{t}), \hat{t}) \).

It suffices to consider the case with \( S(\hat{\gamma}(\hat{t}), \hat{t}) < u(\hat{\gamma}(\hat{t}), \hat{t}) \), the other case is similar. Let \( \hat{x} := \hat{\gamma}(\hat{t}) \). Since \( S(x, t) \) is constructed by the method of characteristics, by Lemma 3.4, there exists a \( C^1 \) curve \( \hat{\gamma} : [t_0, \hat{t}] \to M \) with \( \hat{\gamma}(t_0) = x_0 \) and \( \hat{\gamma}(\hat{t}) = \hat{x} \) such that
\[ S(\hat{x}, \hat{t}) - S(x_0, t_0) = \int_{t_0}^{\hat{t}} L(\hat{\gamma}(\tau), S(\hat{\gamma}(\tau), \tau), \dot{\hat{\gamma}}(\tau)) d\tau. \]

We denote
\[ F(\tau) = S(\hat{\gamma}(\tau), \tau) - u(\hat{\gamma}(\tau), \tau). \]

It is easy to see that \( F(\tau) \) is continuous and \( F(t_0) = 0, F(\hat{t}) < 0 \). Moreover, there exists \( t_1 \in [t_0, \hat{t}] \) such that \( F(t_1) = 0 \) and \( F(\tau) < 0 \) for any \( \tau \in (t_1, \hat{t}), \) i.e.
\[ S(\hat{\gamma}(\tau), \tau) < u(\hat{\gamma}(\tau), \tau). \]

By (L4), a simple calculation implies
\[ S(\hat{x}, \hat{t}) - u(\hat{x}, \hat{t}) \]
\[ \geq \int_{t_1}^{\hat{t}} L(\hat{\gamma}(\tau), S(\hat{\gamma}(\tau), \tau), \dot{\hat{\gamma}}(\tau)) - L(\hat{\gamma}(\tau), u(\hat{\gamma}(\tau), \tau), \dot{\hat{\gamma}}(\tau)) d\tau, \]
\[ \geq 0, \]
which contradicts the assumption \( S(\hat{x}, \hat{t}) < u(\hat{x}, \hat{t}) \). Hence, for any \( \tau \in [t_0, t_0 + \varepsilon] \), we have
\[ S(\hat{\gamma}(\tau), \tau) \geq u(\hat{\gamma}(\tau), \tau). \]

Similarly, we have the converse inequality, which verifies the claim. \( \Box \)

From the definition of \( u \) (see (33)), it follows that
\[ S(\hat{\gamma}(t_0 + \varepsilon), t_0 + \varepsilon) = S(\hat{\gamma}(t_0), t_0) + \int_{t_0}^{t_0 + \varepsilon} L(\hat{\gamma}(\tau), S(\hat{\gamma}(\tau), \tau), \dot{\hat{\gamma}}(\tau)) d\tau, \]
which implies \( \hat{\gamma}(\tau) \) is a solution of the vector field \( \text{grad}_L S_\tau(x) \), i.e.
\[ \dot{\hat{\gamma}}(\tau) = \hat{\mathcal{L}}(\partial_\tau S(\hat{\gamma}(\tau), \tau)), \]
where we use \( \hat{\mathcal{L}} : T^* M \times \mathbb{R} \to TM \times \mathbb{R} \) to denote the Legendre transformation. Let
\[ \hat{H}(\tau, x, p) := H(x, S(x, \tau), p), \quad \hat{L}(\tau, x, \hat{x}) := L(x, S(x, \tau), \hat{x}). \]

We denote
\[ x(\tau) := \hat{\gamma}(\tau), \quad u(\tau) := S(\hat{\gamma}(\tau), \tau) \quad \text{and} \quad p(\tau) := \frac{\partial \hat{L}}{\partial \hat{x}}(\tau, \hat{\gamma}(\tau), \dot{\hat{\gamma}}(\tau)). \]
Then we have
\[
\begin{align*}
\dot{x} &= \frac{\partial H}{\partial p}(t, x, p) = \frac{\partial H}{\partial p}(x, S(x, t), p), \\
\dot{p} &= -\frac{\partial H}{\partial x} = -\frac{\partial H}{\partial x} - \frac{\partial H}{\partial S} \frac{\partial S}{\partial t}, \\
\dot{u} &= L(x, S(x, t), \dot{x}) = \frac{\partial H}{\partial p} p - H(x, S(x, t), p).
\end{align*}
\] (57)

In particular, \(p(\tau)\) is class of \(C^1\). Since \(\tilde{L}\) is a diffeomorphism, it follows from (54) and (56) that
\[
p(\tau) = \partial_x S(\bar{\gamma}(\tau), \tau),
\]
which implies \(x(\tau), u(\tau), p(\tau)\) satisfy
\[
\begin{align*}
\dot{x} &= \frac{\partial H}{\partial p}(x, u, p), \\
\dot{p} &= -\frac{\partial H}{\partial x}(x, u, p) - \frac{\partial H}{\partial u}(x, u, p) p, \\
\dot{u} &= \frac{\partial H}{\partial p}(x, u, p) p - H(x, u, p).
\end{align*}
\] (58)

According to Claim B, we have
\[
\bar{u}(\tau) = u(\bar{\gamma}(\tau), \tau) \quad \text{and} \quad p(\tau) = \frac{\partial L}{\partial x}(\bar{\gamma}(\tau), u(\bar{\gamma}(\tau), \tau), \bar{\gamma}(\tau)).
\] (59)

Then \(\tau \to (\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau))\) is of class \(C^1\) and satisfies the characteristic equation (2).

(c) **Global coincidence between \(\bar{\gamma}(\tau)\) and \(X(\tau)\)**

We will show that \((\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau))\) coincides with \((X(\tau), U(\tau), P(\tau))\) at \((0, t)\).

By the density of the differentiate points of \(\bar{\gamma}\), it suffices to prove the coincidence at \([t_0, t]\).

By contradiction, we assume that \((\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau))\) coincide with \((X(\tau), U(\tau), P(\tau))\) at the maximum interval \([t_0, t_0 + \delta]\) where \(\delta < t - t_0\).

**Claim C.** There exists a constant \(k' > 0\) such that for any \(\tau \in [t_0, t_0 + \delta]\), \(|\bar{\gamma}(\tau)| < k'\).

**Proof of Claim C.** For \(\tau \in [t_0, t_0 + \delta]\), \(\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau)\) satisfy the contact Hamiltonian equation (2). It follows that
\[
\frac{dH}{d\tau}(\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau)) = \frac{\partial H}{\partial x} \bar{\gamma}(\tau) + \frac{\partial H}{\partial u} \bar{u}(\tau) + \frac{\partial H}{\partial p} \bar{p}(\tau)
\]
\[
= -\frac{\partial H}{\partial u}(\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau)) H(\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau)).
\] (60)

By (L3), we have
\[
|H(\bar{\gamma}(\tau), \bar{u}(\tau), p(\tau))| \leq e^{\lambda(\tau-t_0)}|H(\bar{\gamma}(t_0), \bar{u}(t_0), p(t_0))|.
\]

Moreover,
\[
|H(\bar{\gamma}(0), 0, p(\tau))| \leq \lambda|\bar{u}(\tau)| + e^{\lambda(\tau-t_0)}|H(\bar{\gamma}(t_0), \bar{u}(t_0), p(t_0))|.
\]

For \(\tau \in [t_0, t_0 + \delta]\), we have
\[
\bar{u}(\tau) = u(\bar{\gamma}(\tau), \tau).
\] (61)

By the continuity of \(u\), (61) still holds for \(\tau = t_0 + \delta\). Moreover, \(\bar{u}(\tau)\) is bounded for \(\tau \in [t_0, t_0 + \delta]\). Combining with (H2), it follows that \(p(\tau)\) is also bounded. Since the Legendre transformation is a diffeomorphism, there exists a constant \(k' > 0\) such that for any \(\tau \in [t_0, t_0 + \delta]\), \(|\bar{\gamma}(\tau)| < k'\). By the extension theorem of ODE, \(|\bar{\gamma}(\tau)| < k'\) still holds for \(\tau = t_0 + \delta\). \(\square\)
Hence, we have
\[ |\bar{\gamma}(t_0 + \delta) - \bar{\gamma}(t_0 + \delta/2)| \leq \int_{t_0 + \delta/2}^{t_0 + \delta} |\dot{\gamma}(\tau)|d\tau < k'\delta/2. \]

Let
\[ F(\tau) = |\bar{\gamma}(\tau) - \bar{\gamma}(t_0 + \delta/2)| - k'(\tau - (t_0 + \delta/2)). \]
Since \( F(t_0 + \delta) < 0 \), it follows from the continuity of \( F \) that for \( s > t_0 + \delta \) and close to \( t_0 + \delta \), we also have \( F(s) < 0 \). That is
\[ |\bar{\gamma}(s) - \bar{\gamma}(t_0 + \delta/2)| < k'(s - (t_0 + \delta/2)). \]

We choose a compact set \( K \subset M \times \mathbb{R} \) such that for \( \tau \in [t_0, t_0 + \delta] \),
\[ (\bar{\gamma}(\tau), u(\tau)) \in K. \]

By Lemma 3.4, there exists \( \varepsilon := \varepsilon(K, k'' \delta/2) > 0 \) small enough such that for \( \tau \in (t_0 + \delta - \varepsilon/2, t_0 + \delta + \varepsilon/2) \), \( X_{x_1, u_1}(\tau, \cdot) \) is a \( C^1 \) diffeomorphism onto its image, where
\[ x_1 := \bar{\gamma}\left(t_0 + \delta - \frac{\varepsilon}{2}\right), \quad u_1 := u\left(t_0 + \delta - \frac{\varepsilon}{2}\right). \]

Let
\[ \bar{s} := \min\left\{ s, t_0 + \delta + \frac{\varepsilon}{2} \right\}. \]
Repeat the arguments in Step (a) and Step (b), it follows that for any \( \tau \in [t_0, \bar{s}] \),
\[ (\bar{\gamma}(\tau), u(\tau), p(\tau)) = (X(\tau), U(\tau), P(\tau)), \]
which contradicts the assumption, since \( \bar{s} > \delta \). Therefore, for any \( \tau \in (0, t) \), there holds
\[ (\bar{\gamma}(\tau), u(\tau), p(\tau)) = (X(\tau), U(\tau), P(\tau)). \]
We complete the proof of Lemma 3.3. \( \square \)

So far, we complete the proof of Theorem 1.1.

**Remark 3.5.** Generally, a flow generated by \( H(x, u, p) \) may not be complete, but in this paper, we only care about the flow associated to the minimizers, which is complete necessarily from (H1)-(H4).

4. **Representation of the viscosity solution.** In this section, we will provide a representation formula of the solution of (1). By Theorem 1.1, there exists a unique \( u(x, t) \in C(M \times [0, T], \mathbb{R}) \) satisfying \( u(x, 0) = \varphi(x) \) such that
\[ u(x, t) = \inf_{\gamma(t) = x} \left\{ \varphi(\gamma(0)) + \int_0^t L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau \right\}. \] (62)
where the infimum is taken among the continuous and piecewise \( C^1 \) curves. In particular, the infimum is attained at the characteristics of (1).

**Lemma 4.1.** \( u(x, t) \) determined by (62) is a variational solution of (1) with initial condition \( u(x, 0) = \varphi(x) \).

**Proof.** Let \( \gamma : [t_1, t_2] \to M \) be a continuous piecewise \( C^1 \) curve and Let \( \bar{\gamma} : [0, t_1] \to M \) be a minimizer of \( u \) satisfying \( \bar{\gamma}(t_1) = \gamma(t_1) \). We construct a curve \( \xi : [0, t_2] \to M \) defined as follows:
\[ \xi(t) = \begin{cases} \bar{\gamma}(t), & t \in [0, t_1], \\ \gamma(t), & t \in (t_1, t_2]. \end{cases} \] (63)
From (33), it follows that
\[
u(\gamma(t_2), t_2) - u(\gamma(t_1), t_1) = \inf_{\gamma_2(t_2) = \gamma(t_2)} \left\{ \phi(\gamma_2(0)) + \int_0^{t_2} L(\gamma_2(\tau), u(\gamma_2(\tau), \tau), \dot{\gamma}_2(\tau))d\tau \right\}
\]
\[
- \inf_{\gamma_1(t_1) = \gamma(t_1)} \left\{ \phi(\gamma_1(0)) + \int_0^{t_1} L(\gamma_1(\tau), u(\gamma_1(\tau), \tau), \dot{\gamma}_1(\tau))d\tau \right\},
\]
\[
\leq \phi(\xi(0)) + \int_0^{t_2} L(\xi(\tau), u(\xi(\tau), \tau), \dot{\xi}(\tau))d\tau
\]
\[
- \phi(\bar{\xi}(0)) - \int_0^{t_1} L(\bar{\xi}(\tau), u(\bar{\xi}(\tau), \tau), \dot{\bar{\xi}}(\tau))d\tau,
\]
which together with (63) gives rise to
\[
u(\gamma(t_2), t_2) - u(\gamma(t_1), t_1) \leq \int_{t_1}^{t_2} L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau,
\]
which verifies (i) of Definition 2.4. By means of Lemma 3.1 and Lemma 3.3, there exists a C^1 minimizer \( \gamma : [t_1, t_2] \to M \) with \( \gamma(t_2) = x \) such that
\[
u(x, t_2) - u(\gamma(t_1), t_1) = \int_{t_1}^{t_2} L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau.
\]
which implies (ii) of Definition 2.4. This completes the proof of Lemma 4.1. \( \square \)

Based on Definition 2.3, it is easy to see that a variational solution of (1) is a viscosity solution.

**Lemma 4.2.** A variational solution of (1) with initial condition \( u(x, 0) = \phi(x) \) is a viscosity solution.

**Proof.** Let \( u \) be a variational solution of (1). Since \( u(x, 0) = \phi(x) \) it suffices to consider \( t \in (0, T] \). We use \( V \) to denote an open subset of \( M \). Let \( \varphi : V \times [0, T] \to \mathbb{R} \) be a C^1 test function such that \( u - \varphi \) has a maximum at \((x_0, t_0)\). This means
\[
\varphi(x_0, t_0) - \phi(x, t) \leq u(x_0, t_0) - u(x, t).
\]
Fix \( v \in T_{x_0}M \) and for a given \( \delta > 0 \), we choose a C^1 curve \( \gamma : [t_0 - \delta, t_0 + \delta] \to M \) with \( \gamma(t_0) = x_0 \) and \( \dot{\gamma}(t_0) = \xi \). For \( t \in [t_0 - \delta, t_0] \), we have
\[
\varphi(\gamma(t_0), t_0) - \varphi(\gamma(t), t) \leq u(\gamma(t_0), t_0) - u(\gamma(t), t),
\]
where the second inequality is based on (i) of Definition 2.4. Hence,
\[
\frac{\varphi(\gamma(t), t) - \varphi(\gamma(t_0), t_0)}{t - t_0} \leq \frac{1}{t - t_0} \int_{t_0}^{t} L(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau))d\tau.
\]
(66)

Let \( t \to t_0 \), we have
\[
\partial_t \varphi(x_0, t_0) + \partial_x \varphi(x_0, t_0) \cdot \xi \leq L(x_0, u(x_0, t_0), \xi),
\]
which together with Legendre transformation implies
\[
\partial_t \varphi(x_0, t_0) + H(x_0, u(x_0, t_0), \partial_x \varphi(x_0, t_0)) \leq 0,
\]
which shows that \( u \) is a subsolution.
To complete the proof of Lemma 4.2, it remains to show that $u$ is a supersolution. Let $\psi : V \times [0,T] \to \mathbb{R}$ be a $C^1$ test function and $u - \psi$ has a minimum at $(x_0, t_0)$. We have $\psi(x_0, t_0) - \psi(x, t) \geq u(x_0, t_0) - u(x, t)$. From (ii) of Definition 2.4, there exists a $C^1$ curve $\gamma : [0, t_0] \to M$ with $\gamma(t_0) = x_0$ and $\gamma'(t_0) = \eta$ such that for $0 \leq t < t_0$, we have

$$u(\gamma(t_0), t_0) - u(\gamma(t), t) = \int_t^{t_0} L(\gamma(\tau), u(\gamma(\tau), \tau), \gamma'(\tau))d\tau.$$  \hspace{1cm} (67)

Hence

$$\psi(x_0, t_0) - \psi(x, t) \geq \int_t^{t_0} L(\gamma(\tau), u(\gamma(\tau), \tau), \gamma'(\tau))d\tau.$$  

Moreover, we have

$$\frac{\psi(\gamma(t), t) - \psi(\gamma(t_0), t_0)}{t - t_0} \geq \frac{1}{t - t_0} \int_t^{t_0} L(\gamma(\tau), u(\gamma(\tau), \tau), \gamma'(\tau))d\tau.$$  

Let $t$ tend to $t_0$, it gives rise to

$$\partial_t \psi(x_0, t_0) + \partial_x \psi(x_0, t_0) \cdot \eta \geq L(x_0, u(x_0, t_0), \eta),$$

which implies

$$\partial_t \phi(x_0, t_0) + H(x_0, u(x_0, t_0), \partial_x \phi(x_0, t_0)) \geq 0.$$  

This finishes the proof of Lemma 4.2. \hfill \Box

By the comparison theorem (see [5] for instance), it yields that the solution of (1) is unique under the assumptions (H1)-(H4). So far, we have obtained that there exists a unique solution $u(x, t)$ of (1) with initial condition $u(x, 0) = \varphi(x)$ and $u(x, t)$ can be represented implicitly as

$$u(x, t) = \inf_{\gamma(t) = x} \left\{ \varphi(0) + \int_0^t L(\gamma(\tau), u(\gamma(\tau), \tau), \gamma'(\tau))d\tau \right\}. \hspace{1cm} (68)$$

This completes the proof of Theorem 1.2.

5. Solution semigroup. Based on Section 4, the solution of (1) can be represented as

$$u(x, t) = T_t \varphi(x).$$

Based on the fact that $u(x, t)$ is a viscosity solution, $T_t$ satisfies the properties of semigroup $T_{t+t'} = T_t \circ T_{t'}$. To fix the notion, we call $T_t$ solution semigroup. This notion was introduced by [14] under more strict conditions on $H$. Under the assumptions (H1)-(H4), we will detect some further properties of the solution semigroup. Moreover, we will complete the proof of Theorem 1.3.

First of all, it is easy to obtain the following proposition about the monotonicity of $T_t$.

**Proposition 5.1 (Monotonicity).** For given $\varphi, \psi \in C(M, \mathbb{R})$ and $t \geq 0$, if $\varphi \leq \psi$, then $T_t \varphi \leq T_t \psi$.

**Proof.** For given $\varphi, \psi \in C(M, \mathbb{R})$ with $\varphi \leq \psi$, by contradiction, we assume that there exist $t_1 > 0$ and $x_1 \in M$ such that $T_{t_1} \varphi(x_1) > T_{t_1} \psi(x_1)$. Let $\gamma_0 : [0, t_1] \to M$ be a minimizer of $T_{t_1} \psi$ with $\gamma_0(t_1) = x_1$. We denote

$$F(\tau) = T_\tau \varphi(\gamma_0(\tau)) - T_\tau \psi(\gamma_0(\tau)).$$

It is easy to see that $F(\tau)$ is continuous and $F(t_1) > 0$. Since $F(0) = T_0 \varphi(\gamma_0(0)) - T_0 \psi(\gamma_0(0)) \leq 0$,
there exists $t_0 \in [0,t_1)$ such that $F(t_0) = 0$ and for any $\tau \in [t_0,t_1]$, $F(\tau) \geq 0$, i.e.

$$T_\tau \varphi(\gamma_\psi(\tau)) \geq T_\tau \psi(\gamma_\psi(\tau)).$$

(69)

Moreover, it follows that

$$T_{t_1} \varphi(x_1) - T_{t_1} \psi(x_1)$$

$$= \inf_{\gamma(t_1) = x_1} \left\{ T_{t_0} \varphi(\gamma(t_0)) + \int_{t_0}^{t_1} L(\gamma(\tau), T_\tau \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) \, d\tau \right\}$$

$$- \inf_{\gamma(t_1) = x_1} \left\{ T_{t_0} \psi(\gamma(t_0)) + \int_{t_0}^{t_1} L(\gamma(\tau), T_\tau \psi(\gamma(\tau)), \dot{\gamma}(\tau)) \, d\tau \right\}$$

$$\leq T_{t_0} \varphi(\gamma_\psi(t_0)) - T_{t_0} \psi(\gamma_\psi(t_0)) + \int_{t_0}^{t_1} (L(\gamma_\psi(\tau), T_\tau \varphi(\gamma_\psi(\tau)), \dot{\gamma}_\psi(\tau)) - L(\gamma_\psi(\tau), T_\tau \psi(\gamma_\psi(\tau)), \dot{\gamma}_\psi(\tau))) \, d\tau,$$

$$\leq \int_{t_0}^{t_1} (L(\gamma(\tau), T_\tau \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) - L(\gamma(\tau), T_\tau \psi(\gamma(\tau)), \dot{\gamma}(\tau))) \, d\tau.$$

(70)

Combining with (69) and (L4), we have

$$T_{t_1} \varphi(x_1) \leq T_{t_1} \psi(x_1),$$

which is a contradiction. This finishes the proof of Proposition 5.1.

By a similar argument as the one in Proposition 5.1, one can obtain the non-expansiveness of $T_\tau$. For $\varphi \in C(M, \mathbb{R})$, we use $\|\varphi\|_\infty$ to denote $C^0$-norm of $\varphi$. We have the following proposition.

**Proposition 5.2 (Non-expansiveness).** For given $\varphi, \psi \in C(M, \mathbb{R})$ and $t \geq 0$, we have $\|T_t \varphi - T_t \psi\|_\infty \leq \|\varphi - \psi\|_\infty$.

**Proof.** By contradiction, we assume that there exist $t_1 > 0$ and $x_1 \in M$ such that

$$T_{t_1} \varphi(x_1) - T_{t_1} \psi(x_1) > \|\varphi - \psi\|_\infty.$$  

Let $\gamma_\psi : [0,t_1] \to M$ be a minimizer of $T_t \psi$ with $\gamma_\psi(t_1) = x_1$. We denote

$$G(\tau) = T_\tau \varphi(\gamma_\psi(\tau)) - T_\tau \psi(\gamma_\psi(\tau)) - \|\varphi - \psi\|_\infty.$$  

It is easy to see that $G(\tau)$ is continuous and $G(t_1) > 0$. Since

$$G(0) = T_0 \varphi(\gamma_\psi(0)) - T_0 \psi(\gamma_\psi(0)) - \|\varphi - \psi\|_\infty \leq 0,$$

there exists $t_0 \in [0,t_1)$ such that $G(t_0) = 0$ and for any $\tau \in [t_0,t_1]$, $G(\tau) \geq 0$, i.e.

$$T_\tau \varphi(\gamma_\psi(\tau)) - T_\tau \psi(\gamma_\psi(\tau)) \geq \|\varphi - \psi\|_\infty \geq 0.$$  

(71)
A similar calculation as (70) implies
\[ T_{t_1} \varphi(x_1) - T_{t_1} \psi(x_1) - \| \varphi - \psi \|_{\infty} = \inf_{\gamma(t_1) = x_1} \left\{ T_{t_0} \varphi(\gamma(t_0)) + \int_{t_0}^{t_1} L(\gamma(\tau), T_{\tau} \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) \, d\tau \right\} \]
\[ \quad - \inf_{\gamma(t_1) = x_1} \left\{ T_{t_0} \psi(\gamma(t_0)) + \int_{t_0}^{t_1} L(\gamma(\tau), T_{\tau} \psi(\gamma(\tau)), \dot{\gamma}(\tau)) \, d\tau \right\} - \| \varphi - \psi \|_{\infty}, \]
\[ \leq T_{t_0} \varphi(\psi(t_0)) - T_{t_0} \psi(\gamma(t_0)) - \| \varphi - \psi \|_{\infty} + \int_{t_0}^{t_1} \left( L(\gamma(\tau), T_{\tau} \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) - L(\gamma(\tau), T_{\tau} \psi(\gamma(\tau)), \dot{\gamma}(\tau)) \right) \, d\tau, \]
\[ \leq \int_{t_0}^{t_1} \left( L(\gamma(\tau), T_{\tau} \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) - L(\gamma(\tau), T_{\tau} \psi(\gamma(\tau)), \dot{\gamma}(\tau)) \right) \, d\tau. \]

By (L4), we have
\[ T_{t_1} \varphi(x_1) - T_{t_1} \psi(x_1) \leq \| \varphi - \psi \|_{\infty} \] (73)
which is a contradiction. Hence, we have
\[ T_{t_1} \varphi(x_1) - T_{t_1} \psi(x_1) \leq \| \varphi - \psi \|_{\infty}. \]

Similarly, we have
\[ T_{t_1} \varphi(x_1) - T_{t_1} \psi(x_1) > -\| \varphi - \psi \|_{\infty}. \]
This finishes the proof of Proposition 5.1.

We use \( c(H(x, a, p)) \) to denote the Mañé critical value of \( H(x, a, p) \). It is easy to see that \( c(H(x, a, p)) \) is continuous with respect to \( a \). Let
\[ C_H = \cup_{a \in \mathbb{R}} c(H(x, a, p)). \] (74)

**Proposition 5.3.** \( C_H \) is a non-empty interval.

**Proof.** It is clear that \( C_H \) is non-empty. It remains to show \( C_H \) is connected. That is, if \( c_1, c_2 \in C_H \), then for any \( c \in [c_1, c_2], c \in C_H \).

For \( c_1, c_2 \in C_H \) and \( c_1 \neq c_2 \), one can find \( a_1 \neq a_2 \in \mathbb{R} \) such that \( c_1 = c(H(x, a_1, p)) \) and \( c_2 = c(H(x, a_2, p)) \). Since \( c(H(x, a, p)) \) is continuous with respect to \( a \), then for any \( c \in [c_1, c_2] \), there exists at least one \( \tilde{a} \in [a_1, a_2] \) such that \( c = c(H(x, \tilde{a}, p)) \), which is contained in \( C_H \).

Let \( L_c = L + c \). For the sake of simplicity, we will prove Proposition 5.4 and Proposition 5.5 by taking \( c_0 := c(H(x, 0, p)) = c(L(x, 0, \dot{x})) \). It is similar to prove the cases with other elements in \( C_H \). In the following context, we consider \( L_{c_0} \) instead of \( L \). Without ambiguity, we still denote \( T_{t} := T_{t}^{c_0} \), i.e.
\[ T_{t} \varphi(x) = \inf_{\gamma(t) = x} \left\{ \varphi(\gamma(0)) + \int_{0}^{t} L_{c_0}(\gamma(\tau), T_{\tau} \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) \, d\tau \right\}. \] (75)

It is easy to see that the following Proposition 5.4 and 5.5 still hold for other elements in \( C_H \).

**Proposition 5.4 (Uniform boundedness).** For every \( \varphi \in C(M, \mathbb{R}) \), there exists a positive constant \( K \) such that for any \( t \geq 0 \)
\[ \| T_{t} \varphi \|_{\infty} \leq K. \] (76)
Proof. For $t = 0$, $T_0 \phi = \phi$, which is bounded. Let $u(x, t) := T_t \phi(x)$.

On the one hand, we show that $u(x, t)$ is bounded from below. We assume there exists $(x, t) \in M \times (0, +\infty)$ such that $u(x, t) < 0$. Otherwise $u(x, t) \geq 0$ for any $(x, t) \in M \times (0, +\infty)$, which gives the lower bound of $u(x, t)$. Moreover, there exists a minimizer $\gamma$ of $u$ with $\gamma(t) = x$ such that

$$u(x, t) = \phi(\gamma(0)) + \int_0^t L_{\text{co}}(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau)) d\tau. \quad (77)$$

Then, we have the following two cases:

(I) there exists a $\tau_0 \in [0, t]$ such that $u(\gamma(\tau_0), \tau_0) = 0$ and $u(\gamma(\tau), \tau) < 0$ for $\tau \in [\tau_0, t]$;

(II) for every $\tau \in [0, t]$, $u(\gamma(\tau), \tau) < 0$.

For Case (I), it follows from (L4) that

$$u(x, t) = u(\gamma(\tau_0), \tau_0) + \int_{\tau_0}^t L_{\text{co}}(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau)) d\tau,$$

$$\geq 0 + \int_{\tau_0}^t L_{\text{co}}(\gamma(\tau), 0, \dot{\gamma}(\tau)) d\tau$$

$$\geq h_{t_{\text{co}}}^{\tau_0}(\gamma(\tau_0), x).$$

where $h_{t_{\text{co}}}^{\tau_0}(\gamma(\tau_0), x)$ denotes the minimal action from $\gamma(\tau_0)$ to $x$ for the Lagrangian

$L_{\text{co}}(x, 0, \dot{x}) := L(x, 0, \dot{x}) + c_0$.

It follows from (23) that $h_{t_{\text{co}}}^{\tau_0}(\gamma(\tau_0), x)$ is bounded from below. Hence, $u(x, t)$ is bounded from below.

For Case (II), a similar calculation yields

$$u(x, t) \geq \min_{x \in M} \phi(x) + h_{t_{\text{co}}}^{\tau_0}(\gamma(0), x).$$

It follows from the compactness of $M$ and (23) that there exists a constat $K_1$ independent of $(x, t)$ such that $u(x, t) \geq K_1$ for any $(x, t) \in M \times (0, +\infty)$.

On the other hand, we show that $u(x, t)$ is bounded from above. We assume that there exists $(x, t) \in M \times (0, +\infty)$ such that $u(x, t) > 0$. Otherwise, we have $u(x, t) \leq 0$ for any $(x, t) \in M \times (0, +\infty)$, which gives the upper bound of $u(x, t)$.

Let $\gamma : [0, t] \to M$ be a minimal curve with $\gamma(0) = \gamma(t) = x$. Let $\tilde{\gamma} : [0, t] \to M$ be a minimal curve with $\tilde{\gamma}(0) = \gamma(0)$ and $\tilde{\gamma}(t) = x$ such that

$$\int_0^t L_{\text{co}}(\tilde{\gamma}(\tau), 0, \dot{\tilde{\gamma}}(\tau)) d\tau = \inf_{\tilde{\gamma}(0) = \gamma(0)} \int_0^t L_{\text{co}}(\tilde{\gamma}(\tau), 0, \dot{\tilde{\gamma}}(\tau)) d\tau,$$

where the infimum is taken among the continuous and piecewise $C^1$ curves. Then, we have the following two cases:

(I) there exists a $\tau_0 \in [0, t]$ such that $u(\tilde{\gamma}(\tau_0), \tau_0) = 0$ and $u(\tilde{\gamma}(\tau), \tau) > 0$ for $\tau \in [\tau_0, t]$;

(II) for every $\tau \in [0, t]$, $u(\tilde{\gamma}(\tau), \tau) > 0$.  

For Case (I), we have
\[ u(x, t) \leq u(\bar{\gamma}(\tau_0), \tau_0) + \int_{\tau_0}^{t} L_{c_0}(\bar{\gamma}(\tau), u(\bar{\gamma}(\tau), \tau), \dot{\bar{\gamma}}(\tau)) \, d\tau \]
\[ \leq 0 + \int_{\tau_0}^{t} L_{c_0}(\bar{\gamma}(\tau), 0, \dot{\bar{\gamma}}(\tau)) \, d\tau \]
\[ = h^{t-\tau_0}_c(\bar{\gamma}(\tau_0), x). \]

To verify that \( u(x, t) \) is bounded from above, it suffices to prove that \( h^{t-\tau_0}_c(\bar{\gamma}(\tau_0), x) \) is bounded from above. Without loss of generality, we assume \( t > 1 \). Hence, among \( \tau_0 \) and \( t - \tau_0 \), there exists at least one not less than \( \frac{t}{2} \). If \( t - \tau_0 \geq \frac{1}{2} \), then it follows from Proposition 2.5 that \( h^{t-\tau_0}_c(\bar{\gamma}(\tau_0), x) \) is bounded from above. If \( \tau_0 \geq \frac{1}{2} \), it follows from (22) that
\[ h^{t-\tau_0}_c(\bar{\gamma}(\tau_0), x) = h^{\tau_0}_c(x_0, x) - h^{\tau_0}_c(x_0, \bar{\gamma}(\tau_0)). \]

which also implies \( h^{t-\tau_0}_c(\bar{\gamma}(\tau_0), x) \) is bounded from above. Hence, \( u(x, t) \) is bounded from above.

For Case (II), a similar calculation yields
\[ u(x, t) \leq \max_{x \in M} \varphi(x) + h^{t}_c(\gamma(0), x). \]

By the compactness of \( M \), there exists a constant \( K_2 \) independent of \( (x, t) \) such that \( u(x, t) \leq K_2 \) for any \( (x, t) \in M \times [0, +\infty) \). This completes the proof of Proposition 5.4.

Based on Proposition 5.4, one can obtain the equi-Lipschitzity of the family of functions \( T_t \varphi(x) \).

**Proposition 5.5 (Equi-Lipschitzity).** Given \( \delta > 0 \), there exists a constant \( \kappa_\delta > 0 \) such that for \( t \geq \delta > 0 \), \( x \mapsto T_t \varphi(x) \) is \( \kappa_\delta \)-Lipschitzian on \( M \).

The key point to prove Proposition 5.5 is a priori compactness, from which it is easy to verify Proposition 5.5 following from a similar argument as [18]. Let \( u(x, t) := T_t \varphi(x) \).

**Lemma 5.6 (a priori Compactness).** Given \( \delta > 0 \), there exists a compact subset \( \mathcal{K}_\delta \) such that for every minimizer \( \gamma \) of \( u \) and any \( t \geq \delta \), we have
\[ (\gamma(t), u(\gamma(t), t), \dot{\gamma}(t)) \in \mathcal{K}_\delta. \]

**Proof.** It suffices to prove that for a given \( t_0 \geq \delta \),
\[ (\gamma(s), u(\gamma(s), s), \dot{\gamma}(s)) \in \mathcal{K}_\delta \quad \forall s \in [0, t_0]. \]

Indeed, for \( s > t_0 \), one can find \( \bar{t} > 0 \) such that \( \bar{t} \in [\bar{t}, \bar{t} + t_0] \). The curve \( \gamma_{\bar{t}} : [0, t_0] \rightarrow M \) via \( s \mapsto \gamma(\bar{t} + s) \) satisfies the assumption of Lemma 5.6 with \([0, t_0]\) in place of \([\delta, +\infty)\).

By Proposition 5.4, we have \( |u(x, t)| \leq K \). It follows from (L2) that there exists \( C > 0 \) such that
\[ L_{c_0}(x, K, \dot{x}) \geq |\dot{x}| - C. \quad (78) \]

It follows that
\[ u(\gamma(t_0), t_0) - u(\gamma(0), 0) = \int_{0}^{t_0} L_{c_0}(\gamma(\tau), u(\gamma(\tau), \tau), \dot{\gamma}(\tau)) \, d\tau \]
\[ \geq \int_{0}^{t_0} L_{c_0}(\gamma(\tau), K, \dot{\gamma}(\tau)) \, d\tau \geq \int_{0}^{t_0} |\dot{\gamma}(\tau)| \, d\tau - Ct_0, \]
where the second inequality is owing to the monotonicity assumption. On the other hand, we have
\[ |u(\gamma(t_0), t_0) - u(\gamma(0), 0)| \leq 2K. \] 
(79)
Hence, one can find \( s_0 \in [0, t_0] \) such that
\[ |\dot{\gamma}(s_0)| \leq \frac{2K}{t_0} + C \leq \frac{2K}{\delta} + C. \]

Let \( \Phi_s \) be the flow generated by \( L_{c_0}(x, u, \dot{x}) \). Theorem 1.1 implies
\[ (\gamma(s), u(\gamma(s), s), \dot{\gamma}(s)) \]
is an orbit generated by \( L_{c_0}(x, u, \dot{x}) \). It follows that for \( s \in [0, t_0] \),
\[ \Phi_{s-s_0}(\gamma(s_0), u(\gamma(s_0), s_0), \dot{\gamma}(s_0)) = (\gamma(s), u(\gamma(s), s), \dot{\gamma}(s)). \]
Then, Lemma 5.6 follows from the continuity of \( \Phi_s \).

Proof of Proposition 5.5. Taking \( x, y \in M \), let \( \eta := \text{dist}(x, y) \), where “dist” denotes the distance induced by the Riemannian metric on \( M \). Let \( u(x, t) := T_t \varphi(x) \).
Without loss of generality, we assume \( M = U \subset \mathbb{R}^n \).

First of all, we consider \( \eta \leq 1 \). Let \( \gamma : [0, t] \to M \) be a minimizer of \( u \) with \( \gamma(t) = x \), then
\[ u(x, t) = \varphi(\gamma(0)) + \int_0^t L_{c_0}(\gamma(s), u(\gamma(s), s), \dot{\gamma}(s))ds. \]
By Lemma 5.6, there exists a constant \( A_\delta \) independent of \( t \) such that for \( t \geq \delta > 0 \), \( |\dot{\gamma}(t)| \leq A_\delta \). For \( t \geq \delta > 0 \), Taking
\[ \varepsilon := \min \left\{ t, \frac{\eta}{A_\delta} \right\}, \]
we consider another curve \( \tilde{\gamma} : [t - \varepsilon, t] \to M \) defined by
\[ \tilde{\gamma}(s) := \frac{s - (t - \varepsilon)}{\varepsilon}(y - x) + \gamma(s). \]
It is clear that \( \tilde{\gamma} \) is of class \( C^1 \) and connects \( \gamma(t - \varepsilon) \) and \( y \). Then we have
\[ u(y, t) \leq \varphi(\gamma(0)) + \int_0^{t-\varepsilon} L_{c_0}(\gamma(s), u(\gamma(s), s), \dot{\gamma}(s))ds \]
\[ + \int_{t-\varepsilon}^t L_{c_0}(\tilde{\gamma}(s), u(\tilde{\gamma}(s), s), \dot{\gamma}(s))ds. \]
Moreover,
\[ u(y, t) - u(x, t) \]
\[ \leq \int_{t-\varepsilon}^t L_{c_0}(\tilde{\gamma}(s), u(\tilde{\gamma}(s), s), \dot{\gamma}(s))ds - \int_{t-\varepsilon}^t L_{c_0}(\gamma(s), u(\gamma(s), s), \dot{\gamma}(s))ds, \]
\[ \leq \int_{t-\varepsilon}^t |L_{c_0}(\tilde{\gamma}(s), u(\tilde{\gamma}(s), s), \dot{\gamma}(s)) - L_{c_0}(\gamma(s), u(\gamma(s), s), \dot{\gamma}(s))|ds. \]
Based on the regularity of \( L_{c_0} \), we have
\[ |L_{c_0}(x, u, v) - L_{c_0}(x', u', v')| \leq C \max\{\text{dist}(x, x'), |u - u'|, |v - v'|\}. \]
By the definition of \( \tilde{\gamma} \), it follows that for \( s \in [t - \varepsilon, t] \),
\[ \text{dist}(\tilde{\gamma}(s), \gamma(s)) \leq \eta, \quad |\dot{\gamma}(s) - \dot{\gamma}(s)| = \frac{1}{\varepsilon}\eta. \]
which together with the uniform boundedness of \( u(x,t) \) implies

\[
u(y,t) - u(x,t) \leq C\varepsilon \max \left\{ \eta, 2K, \frac{1}{\varepsilon} \eta \right\} \leq \frac{C}{A_\delta} \max \left\{ \eta^2, 2K\eta, A_\delta \eta \right\},
\]

where \( K \) denote the bound of \( u(x,t) \). Since \( \eta \leq 1 \), there exists a constant \( \kappa \) such that

\[
u(y,t) - u(x,t) \leq \kappa \eta.
\]

For the case with \( \eta > 1 \), let \( \bar{\gamma} : [0,\eta] \to M \) be a geodesic of length \( \eta \), parameterized by arclength and connecting \( x \) and \( y \). One can find a finite sequence

\[
t_0 = 0 \leq t_1 \leq \ldots \leq t_n = \eta \quad \text{with} \quad t_{i+1} - t_i \leq 1.
\]

Since

\[
dist(\bar{\gamma}(t_{i+1}), \bar{\gamma}(t_i)) = t_{i+1} - t_i \leq 1,
\]

it follows that for any \( i \in \{0,\ldots,n-1\} \),

\[
u(\bar{\gamma}(t_{i+1}), t) - u(\bar{\gamma}(t_i), t) \leq \kappa(t_{i+1} - t_i).
\]

Adding these inequalities, we obtain

\[
u(y,t) - u(x,t) \leq \kappa \eta.
\]

By exchanging the roles of \( x \) and \( y \), we have

\[
\|u(x,t) - u(y,t)\| \leq \kappa \text{dist}(x,y).
\]

Note that \( \kappa \) is independent of \( t \) for \( t \geq \delta > 0 \). This finishes the proof of Proposition 5.5. \( \square \)

So far, we complete the proof of Theorem 1.3.

6. **Convergence of the solution semigroup.** In this section, we will prove Theorem 1.4. First of all, we are concerned with the divergence of the solution semigroup generated by the Lagrangian \( L_c := L + c \).

**Lemma 6.1.** For \( \varphi(x) \in C(M,\mathbb{R}) \), if \( c \notin C_H \), then there exists \( x_1 \in M \) and a sequence \( \{t_n\}_{n \in \mathbb{N}} \) such that either \( T^\varphi_{t_n} \varphi(x_1) \to +\infty \) or \( T^\varphi_{t_n} \varphi(x_1) \to -\infty \) as \( t_n \to +\infty \).

**Proof.** Fix \( c \notin C_H \). Without loss of generality, we assume \( c = 0 \notin C_H \). Let \( u(x,t) := T^\varphi_{t} \varphi(x) \). It is similar to prove the cases with other elements which are not contained in \( C_H \).

For \( 0 \notin C_H \), it follows from Proposition 5.3 that there are two cases:

(I) for any \( a \in \mathbb{R} \), \( 0 < c(H(x,a,p)) \); 

(II) for any \( a \in \mathbb{R} \), \( 0 > c(H(x,a,p)) \).

For Case (I), we will show that there exists \( x_1 \in M \) and \( \{t_n\}_{n \in \mathbb{N}} \) such that \( u(x_1,t_n) \to -\infty \) as \( t_n \to +\infty \). By contradiction, we assume that there exists a finite \( k \in \mathbb{R} \) such that for any \( x \in M \) and \( t \to +\infty \), we have

\[
u(x,t) \geq k.
\]

That is, there exists \( t_1 > 0 \) such that for any \( t \geq t_1 \) and \( x \in M \), (80) holds. Let \( c_k := c(L(x,k,x)) \), then \( c_k > 0 \). Let \( \gamma_k : [0,t] \to M \) be a minimizer of \( u \) with \( \gamma_k(t) = x \). Let \( \bar{\gamma}_k : [0,t-t_1] \to M \) be a minimal curve with \( \bar{\gamma}_k(0) = \gamma_k(t_1) \) and \( \bar{\gamma}_k(t-t_1) = x \) such that

\[
\int_0^{t-t_1} L_{c_k}(\bar{\gamma}_k(\tau), k, \dot{\bar{\gamma}}_k(\tau)) d\tau = \inf_{\gamma(0)=\gamma_k(t_1)} \int_0^{t-t_1} L_{c_k}(\gamma(\tau), k, \dot{\gamma}(\tau)) d\tau,
\]
where the infimum is taken among the continuous and piecewise $C^1$ curves. Moreover, we have
\[
  u(x,t) \leq u(\gamma_t(t_1), t_1) + \int_0^{t-t_1} L(\gamma_t(\tau), u(\gamma_t(\tau), \tau + t_1), \dot{\gamma}_t(\tau)) \, d\tau \\
  \leq u(\gamma_t(t_1), t_1) + \int_0^{t-t_1} L(\bar{\gamma}_t(\tau), k, \dot{\gamma}_t(\tau)) \, d\tau \\
  = u(\gamma_t(t_1), t_1) + \int_0^{t-t_1} L(\bar{\gamma}_t(\tau), k, \dot{\gamma}_t(\tau)) + c_k - c_k \, d\tau \\
  = u(\gamma_t(t_1), t_1) + h_{c_k}^{t-t_1}(\gamma_t(t_1), x) - c_k(t-t_1).
\]
It follows from Proposition 2.5 and $c_k > 0$ that $u(x,t)$ tends to $-\infty$ as $t \to +\infty$, which contradicts (80).

For Case (II), we will show that there exists $t_2 > 0$ such that for any $t \geq t_2$ and $x \in M$, (81) holds. Let $c_k := c(L(x,k,x))$, then $c_k < 0$. Let $\gamma_t : [0,t] \to M$ be a minimizer of $u$ with $\gamma_t(t) = x$. Moreover, we have
\[
  u(x,t) = u(\gamma_t(t_2), t_2) + \int_0^{t-t_2} L(\gamma_t(\tau), u(\gamma_t(\tau), \tau + t_2), \dot{\gamma}_t(\tau)) \, d\tau \\
  \geq u(\gamma_t(t_2), t_2) + \int_0^{t-t_2} L(\gamma_t(\tau), k, \dot{\gamma}_t(\tau)) \, d\tau \\
  = u(\gamma_t(t_2), t_2) + \int_0^{t-t_2} L(\gamma_t(\tau), k, \dot{\gamma}_t(\tau)) + c_k - c_k \, d\tau \\
  \geq u(\gamma_t(t_2), t_2) + h_{c_k}^{t-t_2}(\gamma_t(t_2), x) - c_k(t-t_2).
\]
It follows from Proposition 2.5 and $c_k < 0$ that $u(x,t)$ tends to $+\infty$ as $t \to +\infty$, which contradicts (81).

This verifies (i) of Theorem 1.4. In the following, we are concerned with the convergence of the solution semigroup generated by the Lagrangian $L_c := L + c$ for a given $c \in C_H$. $H_c$ and $T_t^c$ are the associated Hamiltonian and solution semigroup. Without ambiguity, we still use $L$, $H$ and $T_t$ instead of $L_c$, $H_c$ and $T_t^c$ for the simplicity.

By the monotonicity assumption (H4), we have $\frac{\partial H}{\partial u} \geq 0$. If $\frac{\partial H}{\partial u} \equiv 0$, it was proved by Fathi [17, 18], which is based on the conservation of energy and some properties of Mather sets for the corresponding Hamiltonian systems. Note that $\frac{\partial H}{\partial s} = -\frac{\partial H}{\partial u} H$, the energy is not conservative generally. Besides, the Mather theory for contact Hamiltonian systems has not been established yet. In order to overcome these difficulties, we have to establish a completely new and unified dynamical method to handle the general case with $\frac{\partial H}{\partial u} \geq 0$.

We will show that for any $\varphi(x) \in C(M, \mathbb{R})$, $T_t \varphi(x)$ converges as $t \to +\infty$ to a weak KAM solution of
\[
  H(x, u, \partial_x u) = 0,
\]
which will verify (ii) of Theorem 1.4. The proof will be divided into four steps.
6.1. **Step 1: weak KAM solutions of the stationary equation.** In this step, we will prove the existence of weak KAM solutions of (82).

**Lemma 6.2.** $u$ is a weak KAM solution of (82) if and only if $T_t u = u$ for each $t \geq 0$.

**Proof.** First of all, we suppose $u$ is a weak KAM solution of (82). By Definition 2.2, we have

$$u(x) = \inf_{\gamma(t) = x} \left\{ u(\gamma(0)) + \int_0^t L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau \right\},$$

where the infimum is taken among the continuous and piecewise $C^1$ curves. We will prove the existence of weak KAM solutions of (82). Step 1: weak KAM solutions of the stationary equation.

6.1. Let $F(\tau) = u(\gamma(\tau)) - T_\tau u(\gamma(\tau))$. Since $F(t) > 0$ and $F(0) = 0$, then one can find $s_0 \in [0, t]$ such that $F(s_0) = 0$ and $F(s) > 0$ for $s \in (s_0, t]$. A direct calculation shows

$$F(s) \leq \lambda \int_{s_0}^s F(\tau) d\tau,$$

which implies $F(s) \leq 0$ for $s \in (s_0, t]$ from Gronwall inequality. It contradicts $F(t) > 0$.

Next, we suppose $T_t u = u$ for each $t \geq 0$. Let $\gamma : [t_1, t_2] \rightarrow M$ be a continuous piecewise $C^1$ curve and let $\bar{\gamma} : [0, t_1] \rightarrow M$ be a minimizer of $u$ satisfying $\bar{\gamma}(t_1) = \gamma(t_1)$. We construct a curve $\xi : [0, t_2] \rightarrow M$ defined as follows:

$$\xi(t) = \begin{cases} \bar{\gamma}(t), & t \in [0, t_1], \\ \gamma(t), & t \in (t_1, t_2]. \end{cases}$$

(85)

It follows that

$$u(\gamma(t_2)) - u(\gamma(t_1)) = T_{t_2} u(\gamma(t_2)) - T_{t_1} u(\gamma(t_1)),$$

$$= \inf_{\gamma_2(t_2) = \gamma(t_2)} \left\{ u(\gamma_2(0)) + \int_0^{t_2} L(\gamma_2(\tau), T_\tau u(\gamma_2(\tau)), \dot{\gamma}_2(\tau)) d\tau \right\}$$

$$- \inf_{\gamma_1(t_1) = \gamma(t_1)} \left\{ u(\gamma_1(0)) + \int_0^{t_1} L(\gamma_1(\tau), T_\tau u(\gamma_1(\tau)), \dot{\gamma}_1(\tau)) d\tau \right\},$$

$$\leq u(\xi(0)) + \int_0^{t_2} L(\xi(\tau), T_\tau u(\xi(\tau)), \dot{\xi}(\tau)) d\tau$$

$$- u(\bar{\gamma}(0)) - \int_0^{t_1} L(\bar{\gamma}(\tau), T_\tau u(\bar{\gamma}(\tau)), \dot{\gamma}(\tau)) d\tau,$$

which together with (85) gives rise to

$$u(\gamma(t_2)) - u(\gamma(t_1)) \leq \int_{t_1}^{t_2} L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau,$$

(86)

which verifies (i) of Definition 2.2.
For each $t > 0$, there exists a $C^1$ minimizer $\gamma_t : [−t, 0] \to M$ with $\gamma_t(0) = x$ such that for any $t' \in [−t, 0]$, we have

$$u(x) − u(\gamma_t(t')) = \int_{t'}^0 L(\gamma_t(\tau), u(\gamma_t(\tau)), \dot{\gamma}_t(\tau))d\tau.$$  \hfill (87)

Based on the a priori compactness, for a given $\delta > 0$, there exists a compact subset $K_\delta$ such that for any $t > \delta$ and $s \in [−t, 0]$, we have

$$(\gamma_t(s), u(\gamma_t(s)), \dot{\gamma}_t(s)) \in K_\delta.$$  

Since $\gamma_t$ is a minimizer, it follows from the implicit variational principle that

$$(\gamma_t(s), u(\gamma_t(s)), \dot{\gamma}_t(s)) = \Phi_s(\gamma_t(0), u(\gamma_t(0)), \dot{\gamma}_t(0)) = \Phi_s(x, u(x), \dot{\gamma}_t(0)).$$

The points $(\gamma_t(0), u(\gamma_t(0)), \dot{\gamma}_t(0))$ are contained in a compact subset, then one can find a sequence $t_n$ such that $(\gamma_{t_n}(0), u(\gamma_{t_n}(0)), \dot{\gamma}_{t_n}(0))$ tends to $(x, u(x), v_\infty)$ as $n \to \infty$. Fixing $t' \in (−\infty, 0]$, the function $s \mapsto \Phi_s(x, u(x), \dot{\gamma}_{t_n}(0))$ is defined on $[t', 0]$ for $n$ large enough. By the continuity of $\Phi_s$, the sequence converges uniformly on the compact interval $[t', 0]$ to the map $s \mapsto \Phi_s(x, u(x), v_\infty)$. Moreover, we have

$$(\gamma_\infty(s), u(\gamma_\infty(s)), \dot{\gamma}_\infty(s)) = \Phi_s(x, u(x), v_\infty),$$

then for any $t' \in (−\infty, 0]$, we have

$$u(x) − u(\gamma_\infty(t')) = \int_{t'}^0 L(\gamma_\infty(\tau), u(\gamma_\infty(\tau)), \dot{\gamma}_\infty(\tau))d\tau,$$  \hfill (88)

which implies (ii) of Definition 2.2. Hence, $u$ is a weak KAM solution of (82).

\[\square\]

**Remark 6.3.** By Lemma 4.1 and Lemma 4.2, the weak KAM solution of the evolutionary Hamilton-Jacobi equation (1) is the viscosity solution. Using a similar argument, one can obtain that the weak KAM solution of (82) is a viscosity solution. Let $u(x)$ be a viscosity solution of (82). Note that for each $t \geq 0$, both $T_t u(x)$ and $u(x)$ are the viscosity solutions of

\[
\begin{aligned}
\partial_t v(x, t) + H(x, v(x, t), \partial_x v(x, t)) &= 0, \\
v(x, 0) &= u(x).
\end{aligned}
\]  \hfill (89)

The uniqueness of the solution of (89) implies $T_t u(x) = u(x)$. Hence, the weak KAM solution of (82) is the same as the viscosity solution.

By virtue of Proposition 5.4, we have $T_t \varphi$ is bounded for any $\varphi \in C(M, \mathbb{R})$. Hence, $\limsup_{t \to +\infty} T_t \varphi$ does exist, which is denoted by $\bar{u}(x)$. We have the following lemma.

**Lemma 6.4.** The limit $\lim_{t \to +\infty} T_t \bar{u}$ exists. Moreover, let

$$u_\infty(x) := \lim_{t \to +\infty} T_t \bar{u}(x),$$

then $u_\infty(x)$ is a weak KAM solution of (82).

**Proof.** Due to the definition of limsup, for every $\epsilon > 0$, there exists $s_0 > 0$ such that for any $s \geq s_0$, we have

$$T_s \varphi \leq \bar{u} + \epsilon,$$  \hfill (90)

which the non-expansiveness and monotonicity of $T_t$ implies

$$T_t \circ T_s \varphi \leq T_t(\bar{u} + \epsilon) \leq T_t \bar{u} + \epsilon.$$  \hfill (91)
Fixing \( t \geq 0 \), we take \( \limsup \) for the above inequality as \( s \to +\infty \). Since \[
limsup_{s \to +\infty} T_t \circ T_s \varphi = \limsup_{t+s \to +\infty} T_{t+s} \varphi = \bar{u},
\] then we obtain \[
\bar{u} \leq T_t \bar{u} + \epsilon.
\] Since \( \epsilon \) is arbitrary, we have \[
\bar{u} \leq T_t \bar{u}.
\] By the monotonicity of \( T_t \), it follows from the semigroup property that \( T_t \bar{u} \) is non-decreasing with respect to \( t \). Combining with boundedness of \( T_t \bar{u} \), it follows that the limit \( \lim_{t \to +\infty} T_t \bar{u} \) does exist, which is denoted by \( u_\infty \). Then, we have \[
T_t u_\infty = u_\infty.
\] Based on Proposition 5.4 and Proposition 5.5, it follows from Arzela-Ascoli theorem that \( u_\infty(x) \in C(M, \mathbb{R}) \). By Lemma 6.2, \( u_\infty \) is a weak KAM solution of (82).

This completes the proof of Lemma 6.4.

6.2. Step 2: zero level set of the modified Lagrangian. Since \( u_\infty(x) \) is a weak KAM solution, then it is easy to see that \( u_\infty(x) \) is Lipschitzian. Moreover, it follows from (H1) and (H2), \( u_\infty \) is locally semiconcave (see [7]). Let \( D \) be the set of all differentiable points of \( u_\infty \) on \( M \). Due to the Lipschitzian property of \( u_\infty \), it follows that \( D \) has full Lebesgue measure. For \( x \in D \), we have \[
H(x, u_\infty(x), \partial_x u_\infty(x)) = 0.
\] We define \[
\tilde{L}(x, \dot{x}) = L(x, u_\infty(x), \dot{x}) - \langle \partial_x u_\infty(x), \dot{x} \rangle, \quad x \in \mathcal{D}.
\] Denote \[
\Gamma := \left\{ \left(x, \frac{\partial H}{\partial p}(x, u_\infty(x), \partial_x u_\infty(x)) \right) : x \in \mathcal{D} \right\},
\] where \( \frac{\partial H}{\partial p} \) denotes the partial derivative of \( H \) with respect to the third argument. We have the following lemma.

**Lemma 6.5.** For any \( x \in \mathcal{D} \), \( \tilde{L}(x, \dot{x}) \geq 0 \). In particular, \( \tilde{L}(x, \dot{x}) = 0 \) if and only if \( (x, \dot{x}) \in \Gamma \).

**Proof.** By (96) and (97), we have \[
\left. \frac{\partial L}{\partial p} \right|_{\Gamma} = -H(x, u_\infty(x), \partial_x u_\infty(x)) = 0.
\] In addition, we have \[
\left. \frac{\partial \tilde{L}}{\partial \dot{x}} \right|_{\Gamma} = \frac{\partial L}{\partial \dot{x}}(x, u_\infty(x), \dot{x}) - \partial_x u_\infty(x) = 0.
\] By (L2), it follows from (98) that there exists \( K_1 > 0 \) large enough such that for \( |\dot{x}| > K_1 \), \[
\tilde{L}(x, \dot{x}) \geq d > 0,
\] where \( d \) is a constant independent of \( (x, \dot{x}) \).

For \( x \in \mathcal{D} \), \( u_\infty(x) \) satisfies the equation (95). Since \( u_\infty(x) \) is Lipschitzian, then \( \partial_x u_\infty(x) \) is bounded. Let \[
\dot{x}_0 := \frac{\partial H}{\partial p}(x, u_\infty(x), \partial_x u_\infty(x)),
\]
then there exists $K_2 > 0$ independent of $x$ such that $|\dot{x}_0| \leq K_2$. Take $K_3 := \max\{K_1, K_2\}$. From the assumption (L1), it follows that $\frac{\partial^2L}{\partial x^2}(x, u_\infty(x), \dot{x})$ is positive definite. Hence, for $|\dot{x}| \leq K_3$, it follows from (98) and (99) that there exists $\Lambda > 0$ independent of $(x, \dot{x})$ such that

$$
\bar{L}(x, \dot{x}) \geq \Lambda \left| \dot{x} - \frac{\partial H}{\partial p}(x, u_\infty(x), \partial_x u_\infty(x)) \right|^2.
$$

(100)

Consequently, it is easy to see that

$$
\bar{L}(x, \dot{x}) \begin{cases} 
0, & (x, \dot{x}) \in \Gamma, \\
> 0, & (x, \dot{x}) \notin \Gamma.
\end{cases}
$$

(101)

This completes the proof of Lemma 6.5.

6.3. **Step 3: energy evolution along the characteristics.** In this step, we focus on the evolution of $H$ along the characteristics. First of all, we recall the definition of $T_t$:

$$
T_t \varphi(x) = \inf_{\gamma(t) = x} \left\{ \varphi(\gamma(0)) + \int_0^t L(\gamma(\tau), T_\tau \varphi(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau \right\},
$$

(102)

where the infimum is taken among the continuous and piecewise $C^1$ curves. In particular, the infimum is attained at the characteristics of (1) based on Theorem 1.1. More precisely, let $\gamma : [0, t] \to M$ be a minimizer of $T_t \varphi$, then $(\gamma(s), u(s), p(s))$ defined as

$$
(\gamma(s), u(s) := T_s \varphi(\gamma(s)), p(s) = \frac{\partial L}{\partial x}(\gamma(s), T_s \varphi(\gamma(s)), \dot{\gamma}(s))
$$

(103)

is $C^1$ and satisfies the characteristic equation (2). To avoid the ambiguity, we denote the characteristics by $(X(t), U(t), P(t))$. Along the characteristics, there holds for any $s \in [0, t]$,

$$
\frac{dH}{ds}(X(s), U(s), P(s)) = \frac{\partial H}{\partial x} \dot{X}(s) + \frac{\partial H}{\partial u} \dot{U}(s) + \frac{\partial H}{\partial p} \dot{P}(s)
$$

$$
= -\frac{\partial H}{\partial u}(X(s), U(s), P(s)) H(X(s), U(s), P(s)).
$$

(104)

Let $\tilde{H}(s) := H(X(s), U(s), P(s))$. It follows from (H4) that

- $\tilde{H}(s)$ is a non-increasing function with respect to $s$ if $\tilde{H}(0) > 0$;
- $\tilde{H}(s)$ is a non-decreasing function with respect to $s$ if $\tilde{H}(0) < 0$;
- $\tilde{H}(s) = 0$ if $\tilde{H}(0) = 0$.

By virtue of Proposition 5.4 and Proposition 5.5, it follows from Arzela-Ascoli theorem that there exist a sequence $t_n \to +\infty$ such that $\lim_{n \to +\infty} T_{t_n} \varphi$ exists. We denote

$$
\tilde{u}(x) := \lim_{n \to +\infty} T_{t_n} \varphi.
$$

(105)

It is easy to see that $\tilde{u}(x)$ is Lipschitz. Based on Lemma 6.4, we have

$$
\tilde{u} \leq u_\infty,
$$

(106)

where $u_\infty$ denotes the weak KAM solution given by Lemma 6.4.

For a given $s > 0$, extracting a subsequence if necessary, we denote

$$
u_s(x) := \lim_{n \to +\infty} T_{t_{n-s}} \varphi(x).
$$

(107)
$u_s(x)$ is also Lipschitzian. In addition, we have $u_s \leq u_{\infty}$. Let $w_s(x, t) := T_t u_s(x)$. By the monotonicity of $T_t$, we have

$$w_s(x, t) \leq u_{\infty}(x). \quad (108)$$

It follows from Theorem 1.2 that $w(x, t) := w_s(x, t)$ is a solution of the following equation:

$$\begin{align*}
\frac{\partial}{\partial t} w(x, t) + H(x, w(x, t), \frac{\partial}{\partial x} w(x, t)) &= 0, \\
w(x, 0) &= u_s(x).
\end{align*} \quad (109)$$

A standard argument shows that $w_s(x, t)$ is locally semiconcave ([7, Theorem 5.3.8]). In particular, we have $w_s(x, s) = \tilde{u}(x)$, where $\tilde{u}$ is given by (105). Indeed, from non-expansiveness, we have

$$w_s(x, s) = T_s u_s(x) = T_s \left( \lim_{n \to +\infty} T_{t_n - s} \varphi \right) = \lim_{n \to +\infty} T_s \circ T_{t_n - s} \varphi = \lim_{n \to +\infty} T_{t_n} \varphi = \tilde{u}(x).$$

Let $\mathcal{D}'$ be the set of all differentiable points of $\tilde{u}$ on $M$. Due to the Lipschitzian property of $\tilde{u}$, it follows that $\mathcal{D}'$ has full Lebesgue measure. Let $\mathcal{D}'' := \mathcal{D} \cap \mathcal{D}'$, then $\mathcal{D}''$ also has full Lebesgue measure. For the simplicity of notations, we still use $\mathcal{D}$ instead of $\mathcal{D}''$. We have the following lemma.

**Lemma 6.6.** For each $x \in \mathcal{D}$, we have

$$H(x, \tilde{u}(x), \frac{\partial}{\partial x} \tilde{u}(x)) \leq 0. \quad (110)$$

**Proof.** By contradiction, we assume that there exists $x_0 \in \mathcal{D}$ such that

$$H(x_0, \tilde{u}(x_0), \frac{\partial}{\partial x} \tilde{u}(x_0)) = \delta > 0. \quad (111)$$

From the Legendre transformation,

$$H(x_0, \tilde{u}(x_0), \frac{\partial}{\partial x} \tilde{u}(x_0)) = \sup_{v \in T_{x_0} M} \{ \langle v, \frac{\partial}{\partial x} \tilde{u}(x_0) \rangle - L(x_0, \tilde{u}(x_0), v) \},$$

$$= \left\langle \frac{\partial L}{\partial \dot{x}}(x_0, \tilde{u}(x_0), \dot{x}_0), \dot{x}_0 \right\rangle - L(x_0, \tilde{u}(x_0), \dot{x}_0),$$

where

$$\dot{x}_0 = \frac{\partial H}{\partial \dot{p}}(x_0, \tilde{u}(x_0), \frac{\partial}{\partial x} \tilde{u}(x_0)). \quad (112)$$

By (111), it yields that

$$L(x_0, \tilde{u}(x_0), \dot{x}_0) - \left\langle \frac{\partial L}{\partial \dot{x}}(x_0, \tilde{u}(x_0), \dot{x}_0), \dot{x}_0 \right\rangle = -H(x_0, \tilde{u}(x_0), \frac{\partial}{\partial x} \tilde{u}(x_0)) = -\delta. \quad (113)$$

We denote

$$\bar{L}(x, w_s(x, t), \dot{x}) := L(x, w_s(x, t), \dot{x}) - \langle \frac{\partial}{\partial x} u_{\infty}(x), \dot{x} \rangle,$$

where $w_s(x, t)$ is a solution of (109). In particular, $w_s(x_0, s) = \tilde{u}(x_0)$. In terms of (96), we have $\bar{L}(x, u_{\infty}(x), \dot{x}) = \bar{L}(x, \dot{x})$. By (108), we have $w_s(x, t) \leq u_{\infty}(x)$. Then it follows from (L4) that

$$\bar{L}(x, w_s(x, t), \dot{x}) \geq \bar{L}(x, \dot{x}). \quad (114)$$

By Lemma 6.5, we have

$$\bar{L}(x, w_s(x, t), \dot{x}) \geq 0. \quad (115)$$

We denote

$$\bar{L}(x, w_s(x, t), \dot{x}) := L(x, w_s(x, t), \dot{x}) - \left\langle \frac{\partial L}{\partial \dot{x}}(x, w_s(x, t), \dot{x}), \dot{x} \right\rangle. \quad (116)$$
We consider the following two sets:

\[ \Delta_s := \left\{ (x, w_s(x, t), \dot{x}) : \dot{L}(x, w_s(x, t), \dot{x}) = 0, \ (x, t) \in M \times [0, s], \ \dot{x} \in T_x M \right\}, \]

\[ \Sigma_s(\delta) := \left\{ (x, w_s(x, t), \dot{x}) : \dot{L}(x, w_s(x, t), \dot{x}) \leq -\delta, \ (x, t) \in M \times [0, s], \ \dot{x} \in T_x M \right\}, \]

where \( \Delta_s \) and \( \Sigma_s(\delta) \) denote the zero level set and the sublevel set of \( \dot{L} \) respectively.

**Claim D.** \( \Delta_s \) is compact.

**Proof of Claim D.** By the Legendre transformation, we have

\[ \dot{L}(x, w_s(x, t), \dot{x}) = -H(x, w_s(x, t), p), \]

where \( \dot{x} = \frac{\partial H}{\partial p}(x, w_s(x, t), p) \). If \( (x, w_s(x, t), \dot{x}) \in \Delta_s \), then \( H(x, w_s(x, t), p) = 0 \). According to Proposition 5.4, \( w_s(x, t) \) is bounded, which together with \( (H2) \) yields there exists \( C \) independent of \( (x, t) \) such that \( |p| \leq C \). Moreover, \( \dot{x} \) is contained in a compact set. Hence, the claim follows from the compactness of \( M \). \( \square \)

Next, we show that \( \Delta_s \) and \( \Sigma_s(\delta) \) are separated by \( \theta(\delta) > 0 \) for \( \delta > 0 \). More precisely, we have the following claim.

**Claim E.** For each \( (x, w_s(x, t), \dot{x}) \in \Sigma_s(\delta) \), there exists \( \theta(\delta) > 0 \) such that

\[ \text{dist}((x, w_s(x, t), \dot{x}), \Delta_s) \geq \theta(\delta), \]

where “dist” denotes a distance induced by the Riemannian metric.

**Proof of Claim E.** By contradiction, we assume that for any \( \epsilon > 0 \),

\[ \text{dist}((x, w_s(x, t), \dot{x}), \Delta_s) < \epsilon. \]

Hence, there exists a sequence \( (x_n, w_s(x_n, t_n), v_n) \) contained in \( \Sigma_s(\delta) \) such that for \( n \) large enough, extracting a subsequence if necessary,

\[ \text{dist}((x_n, w_s(x_n, t_n), v_n), \Delta_s) < \frac{1}{n}. \]

Up to a local trivialization on \( TM \), extracting a subsequence if necessary, one obtain

\[ (x_n, w_s(x_n, t_n), v_n) \rightarrow (\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}). \]

Based on the compactness of \( \Delta_s \), we have \( (\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) \in \Delta_s \). Then, it follows from the definition of \( \Delta_s \) that

\[ \dot{L}(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) = 0. \]  

(117)

Since \( \dot{L}(x_n, w_s(x_n, t_n), v_n) \leq -\delta < 0 \), then it follows from the continuity of \( \dot{L} \) with respect to \( x \) that

\[ \dot{L}(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) \leq -\delta < 0, \]  

(118)

which contradicts (117). Hence, for \( (x, w_s(x, t), \dot{x}) \in \Sigma_s(\delta) \), there exists \( \theta > 0 \) such that

\[ \text{dist}((x, w_s(x, t), \dot{x}), \Delta_s) \geq \theta(\delta), \]

which verifies the claim. \( \square \)

**Claim F.** For each \( x \in D \) and \( (x, w_s(x, t), \dot{x}) \in \Sigma_s(\delta) \), there exists \( \delta' > 0 \) independent of \( (x, w_s(x, t), \dot{x}) \) such that

\[ \dot{L}(x, w_s(x, t), \dot{x}) \geq \delta' > 0, \]  

(119)

**Proof of Claim F.** By contradiction, we assume that for any \( \epsilon > 0 \),

\[ \dot{L}(x, w_s(x, t), \dot{x}) < \epsilon. \]
Hence, there exists a sequence \((x_n, w_s(x_n, t_n), v_n)\) satisfying \(x_n \in \mathcal{D}\) and
\[
(x_n, w_s(x_n, t_n), v_n) \in \Sigma_s(\delta)
\]
such that for \(n\) large enough, extracting a subsequence if necessary,
\[
\bar{L}(x_n, w_s(x_n, t_n), v_n) < \frac{1}{n}.
\]
By the definition of \(\bar{L}\), we have
\[
L(x_n, w_s(x_n, t_n), v_n) - \langle \partial_x \bar{u}_\infty(x_n), v_n \rangle < \frac{1}{n}.
\] (120)
Since \(u_\infty\) is Lipschitzian, then there exists a positive constant \(C\) independent of \(x\) such that \(|\partial_x u_\infty(x)| \leq C\) for \(x \in \mathcal{D}\). Let \(y_n := \partial_x u_\infty(x_n)\). Up to a local trivialization on \(TM\), extracting a subsequence if necessary, one obtains \((x_n, v_n, t_n, y_n) \rightarrow (\bar{x}, \bar{v}, \bar{v}, \bar{y})\) as \(n \rightarrow +\infty\). From (115), we have
\[
L(x_n, w_s(x_n, t_n), v_n) - \langle y_n, v_n \rangle \geq 0,
\]
which together with (120) implies
\[
L(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) - \langle \bar{y}, \bar{v} \rangle = 0.
\] (121)
Based on Lemma 6.5, it follows from (114) that for any \(\xi \in T_{\bar{x}}\mathcal{M}\), we have
\[
L(\bar{x}, w_s(\bar{x}, \bar{t}), \xi) - \langle \bar{y}, \xi \rangle \geq 0,
\]
which together with (121) yields
\[
\bar{y} = \frac{\partial L}{\partial \bar{x}}(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}).
\]
Hence, we have
\[
\bar{L}(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) = L(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) - \left\langle \frac{\partial L}{\partial \bar{x}}(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}), \bar{v} \right\rangle = 0,
\]
which implies
\[
(\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) \in \Delta_s.
\] (122)
On the other hand, since \((x_n, w_s(x_n, t_n), v_n) \in \Sigma_s(\delta)\), then \((\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}) \in \Sigma_s(\frac{\delta}{2})\).
Moreover, there exists \(\theta(\frac{\delta}{2}) > 0\) such that
\[
\text{dist} ((\bar{x}, w_s(\bar{x}, \bar{t}), \bar{v}), \Delta_s) \geq \theta(\frac{\delta}{2}),
\]
which contradicts (122). This verifies the claim. \(\square\)

Since \(w_s(x, t)\) is semiconcave, it follows that
\[
\partial_x w_s(x_0, s) = \partial_x \tilde{u}(x_0) \in \Pi_x D^* w_s(x_0, s),
\]
where \(\Pi_x D^* u(x, t)\) means the projection of the reachable gradient of \(u\) at \((x, t)\) to the \(x\)-argument. According to [7, Theorem 6.4.9], combining with Theorem 1.1, there exists a minimizer \(\gamma_{w_s} : [0, s] \rightarrow M\) of \(w_s(x, t)\) with \(\gamma_{w_s}(s) = x_0\) and
\[
\frac{\partial L}{\partial \bar{x}}(\gamma_{w_s}(s), w_s(\gamma_{w_s}(s), s), \gamma_{w_s}(s)) = \partial_x \tilde{u}(x_0).
\]
It follows from (111) that
\[
H \left( \gamma_{w_s}(s), w_s(\gamma_{w_s}(s), s), \frac{\partial L}{\partial \bar{x}}(\gamma_{w_s}(s), w_s(\gamma_{w_s}(s), s), \gamma_{w_s}(s)) \right) = \delta.
\]
In terms of (104), for any \( \tau \in [0, s] \),
\[
H \left( \gamma_{w_s}(\tau), w_s(\gamma_{w_s}(\tau), \tau), \frac{\partial L}{\partial x} (\gamma_{w_s}(\tau), w_s(\gamma_{w_s}(\tau), \tau), \dot{\gamma}_{w_s}(\tau)) \right) \geq \delta,
\]
where \( \frac{\partial L}{\partial x} \) denotes the partial derivative of \( L \) with respect to the third argument. By the Legendre transformation, for any \( \tau \in [0, s] \),
\[
(\gamma_{w_s}(\tau), w_s(\gamma(\tau), \tau), \dot{\gamma}_{w_s}(\tau)) \in \Sigma_s(\delta).
\]
It follows that for any \( \tau \in [0, s] \),
\[
\text{dist} \left( (\gamma_{w_s}(\tau), w_s(\gamma(\tau), \tau), \dot{\gamma}_{w_s}(\tau)) \right) \geq \theta(\delta).
\]
Let \( \Theta \) be the set of \( \gamma(\tau) \) along which the directional derivative \( \partial_{\gamma_{w_s}}(\tau)u_\infty(\dot{\gamma}_{w_s}(\tau)) \) exists. For \( \gamma_{w_s}(\tau) \in \Theta \), we denote
\[
\hat{L}(\gamma_{w_s}(\tau)) := L(\gamma_{w_s}(\tau), w_s(\gamma_{w_s}(\tau), \tau), \dot{\gamma}_{w_s}(\tau)) - \partial_{\gamma_{w_s}}(\tau)u_\infty(\dot{\gamma}_{w_s}(\tau)).
\]
Since \( u_\infty \) is a weak KAM solution (also a viscosity solution), then it is locally semiconcave. By [7, Proposition 3.3.4 and Theorem 3.3.6], one can find a sequence \( x_n^\tau \in D \) with \( x_n^\tau \to \gamma_{w_s}(\tau) \) as \( n \to +\infty \) for a given \( \tau \in [0, s] \) such that for \( n \) large enough, extracting a subsequence if necessary,
\[
\partial_{\gamma_{w_s}}(\tau)u_\infty(\dot{\gamma}_{w_s}(\tau)) \leq \langle \partial_x u_\infty(x_n^\tau), \dot{\gamma}_{w_s}(\tau) \rangle + \frac{1}{n}.
\]
It follows from (123) that
\[
H \left( x_n^\tau, w_s(x_n^\tau, \tau), \frac{\partial L}{\partial x} (x_n^\tau, w_s(x_n^\tau, \tau), \gamma_{w_s}(\tau)) \right) \geq \frac{\delta}{2},
\]
which implies
\[
\text{dist} \left( (x_n^\tau, w_s(x_n^\tau, \tau), \dot{\gamma}_{w_s}(\tau)) \right) \geq \theta(\frac{\delta}{2}).
\]
Since \( x_n^\tau \in D \) and \( (x_n^\tau, w_s(x_n^\tau, \tau), \dot{\gamma}_{w_s}(\tau)) \in \Sigma_s(\frac{\delta}{2}) \), then there exists \( \delta'' > 0 \) independent of \( \tau \) and \( n \) such that for any \( \tau \in [0, s] \),
\[
\hat{L}(x_n^\tau, w_s(x_n^\tau, \tau), \dot{\gamma}_{w_s}(\tau)) = L(x_n^\tau, w_s(x_n^\tau, \tau), \dot{\gamma}_{w_s}(\tau)) - \langle \partial_x u_\infty(x_n^\tau), \dot{\gamma}_{w_s}(\tau) \rangle \geq \delta''
\]
which together with (124) and (125) implies
\[
\hat{L}(\gamma_{w_s}(\tau)) \geq \frac{\delta''}{2}.
\]
Moreover, we have
\[
\int_0^s \hat{L}(\gamma_{w_s}(\tau)) d\tau \geq \frac{\delta''}{2} s.
\]
On the other hand, since \( \gamma_{w_s} \) is a minimizer of \( w_s(x, t) \), then we have
\[
\int_0^s \hat{L}(\gamma_{w_s}(\tau)) d\tau
\[
= \int_0^s L(\gamma_{w_s}(\tau), w_s(\gamma_{w_s}(\tau), \tau), \dot{\gamma}_{w_s}(\tau)) - \partial_{\gamma_{w_s}}(\tau)u_\infty(\dot{\gamma}_{w_s}(\tau)) d\tau,
\]
\[
= w_s(x_n^s, s) - w_s(\gamma_{w_s}(0), 0) - (u_\infty(\gamma_{w_s}(s)) - u_\infty(\gamma_{w_s}(0))).
\]
By Proposition 5.4, there exists a positive constant \( C \) independent of \( s \) such that
\[
\int_0^s \hat{L}(\gamma_{w_s}(\tau)) d\tau \leq C,
\]
which contradicts (129) for $s$ large enough. Therefore, it yields that for $x \in \mathcal{D}$,
\[
H(x, \tilde{u}(x), \partial_x \tilde{u}(x)) \leq 0.
\] (130)
This completes the proof of Lemma 6.6.

\textbf{Lemma 6.7.} For each continuous and piecewise $C^1$ curve $\gamma : [0, t] \rightarrow M$, we have
\[
\tilde{u}(\gamma(t)) - \tilde{u}(\gamma(0)) \leq \int_0^t L(\gamma(\tau), \dot{u}(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau.
\] (131)

\textbf{Proof.} It follows from Lemma 6.6 that $H(x, u(x), \partial_x u(x)) \leq 0$ for almost all $x \in M$. Using a covering of a curve by coordinates charts, one can assume $M = U$ is an open convex set in $\mathbb{R}^n$. Note that a $C^1$ curve can be approximated by piecewise affine curves in the topology of uniform convergence. The following proof is similar to [18, Proposition 4.2.3]. We omit the details.

\textbf{6.4. Step 4: proof of Theorem 1.4.}

\textbf{Lemma 6.8.} Given $u(x) \in C(M, \mathbb{R})$, for each $t \geq 0$, there holds $u(x) \leq T_t u(x)$ if and only if for each continuous and piecewise $C^1$ curve $\gamma : [0, t] \rightarrow M$, we have
\[
u(\gamma(t)) - u(\gamma(0)) \leq \int_0^t L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau.
\] (132)

\textbf{Proof.} We first suppose $u(x) \leq T_t u(x)$ for each $t \geq 0$. For each continuous and piecewise $C^1$ curve $\gamma : [0, t] \rightarrow M$, we have
\[
u(\gamma(t)) \leq T_t u(\gamma(t)) \leq u(\gamma(0)) + \int_0^t L(\gamma(\tau), T_t u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau,
\]
\[
\leq u(\gamma(0)) + \int_0^t L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau,
\]
where the last inequality is from the monotonicity assumption (L4). Then
\[
u(\gamma(t)) - u(\gamma(0)) \leq \int_0^t L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau.
\]

Conversely, we suppose that for each continuous and piecewise $C^1$ curve $\gamma : [0, t] \rightarrow M$, we have
\[
u(\gamma(t)) - u(\gamma(0)) \leq \int_0^t L(\gamma(\tau), u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau.
\]
By contradiction, we assume $u(x) > T_t u(x)$. Let $\gamma : [0, t] \rightarrow M$ be a minimizer of $T_t u$ with $\gamma(t) = x$, i.e.
\[
T_t u(x) = u(\gamma(0)) + \int_0^t L(\gamma(\tau), T_t u(\gamma(\tau)), \dot{\gamma}(\tau)) d\tau.
\] (133)
Let $F(\tau) = u(\gamma(\tau)) - T_t u(\gamma(\tau))$. Since $F(t) > 0$ and $F(0) = 0$, then one can find $s_0 \in [0, t]$ such that $F(s_0) = 0$ and $F(s) > 0$ for $s \in (s_0, t]$. A direct calculation shows
\[
F(s) \leq \lambda \int_{s_0}^{s} F(\tau) d\tau,
\]
which implies $F(s) \leq 0$ for $s \in (s_0, t]$ from Gronwall inequality. It contradicts $F(t) > 0$. This completes the proof.
Based on the preparations above, by Proposition 5.1, we conclude that for each \( t' \geq t \geq 0 \),

\[ \tilde{u} \leq T_t \tilde{u} \leq T_t u. \]

Let us recall \( \lim_{t_n \to +\infty} T_{t_n} \varphi = \tilde{u} \). Up to a subsequence, we choose \( t_{n+1} - t_n \to +\infty \).

Let \( s_n := t_{n+1} - t_n \). Note that \( T_{s_n} \circ T_{t_n} u = T_{t_{n+1}} u \), it follows from Proposition 5.2 that

\[
\|T_{s_n} \tilde{u} - \hat{u}\|_{\infty} \leq \|T_{s_n} \tilde{u} - T_{s_n} \circ T_{t_n} \varphi\|_{\infty} + \|T_{t_{n+1}} \varphi - \hat{u}\|_{\infty},
\]

which together with \( \lim_{t_n \to +\infty} T_{t_n} \varphi = \tilde{u} \) shows \( T_{s_n} \tilde{u} \to \hat{u} \) as \( s_n \to +\infty \). Since \( T_t \tilde{u} \) is non-decreasing with respect to \( t \), then \( \hat{u} \) is a fixed point of \( T_t \) for \( t \geq 0 \). By virtue of Lemma 4.1, we have \( \hat{u} \) is a weak KAM solution of (82). Moreover, using Proposition 5.2 again, it follows that for \( t > t_n \), we have

\[
\|T_t \varphi - \hat{u}\|_{\infty} = \|T_{t-t_n} \circ T_{t_n} \varphi - T_{t-t_n} \tilde{u}\|_{\infty} \leq \|T_{t_n} \varphi - \hat{u}\|_{\infty}.
\]

Since \( T_{t_n} \varphi \to \tilde{u} \) as \( t_n \to +\infty \), we obtain

\[
\lim_{t \to +\infty} T_t \varphi = \tilde{u},
\]

where \( \tilde{u} \) is a weak KAM solution of (82). This finishes the proof of Theorem 1.4. \( \square \)

**Remark 6.9.** Based on the uniqueness of the limit of \( T_t \varphi(x) \) as \( t \to +\infty \), we know that \( u_\infty \) given by Lemma 6.4 is the same as \( \tilde{u} \) given by (105).

**Acknowledgements.** The authors sincerely thank the referees for their careful reading of the manuscript and invaluable comments which were very helpful in improving this paper. X. Su was partially supported by both National Natural Science Foundation of China (Grant No. 11301513) and “the Fundamental Research Funds for the Central Universities”. L. Wang was partially under the support of National Natural Science Foundation of China (Grant No. 11401107). J. Yan was partially under the support of National Natural Science Foundation of China (Grant No. 11325103) and National Basic Research Program of China (Grant No. 2013CB834100).

**REFERENCES**

1. V. I. Arnold, *Geometric Methods in The Theory of Ordinary Differential Equations*, Springer-Verlag, New York, 1983.
2. V. I. Arnold, *Lectures on Partial Differential Equations*, Springer, Berlin Heidelberg, 2004.
3. S. Aubry, The twist map, the extended Frenkel-Kontorova model and the devil’s staircase, *Phys. D.*, 7 (1983), 240–258.
4. S. Aubry and P. Y. Le Daeron, The discrete Frenkel-Kontorova model and its extensions I: Exact results for the ground states, *Phys. D.*, 8 (1983), 381–422.
5. G. Barles, *Solutions de Viscosité des Équations de Hamilton-Jacobi Mathématiques & Applications*, (Berlin) 17, Springer, Paris, 1994.
6. G. Buttazzo, M. Giaquinta and S. Hildebrandt, *One-dimensional Variational Problems, An Introduction*, (Oxford Lecture Series in Mathematics and Its Applications), Clarendon Press Oxford, 1998.
7. P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Vol. 58, Springer, 2004.
8. M. G. Crandall, L. C. Evans and P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.*, 282 (1984), 487–502.
9. M. G. Crandall, H. Ishii and P.-L. Lions, User’s guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (N.S.)*, 27 (1992), 1–67.
[10] G. Contreras, R. Iturriaga, G. P. Paternain and M. Paternain, Lagrangian graphs, minimizing measures and Mañé’s critical values, *Geom. Funct. Anal.*, 8 (1998), 788–809.
[11] M. G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.*, 277 (1983), 1–42.
[12] A. Davini, A. Fathi, R. Iturriaga and M. Zavidovique, Convergence of the solutions of the discounted Hamilton-Jacobi equation, *Invent. Math.*, 105 (2016), 1–27.
[13] A. Davini and A. Siconolfi, A generalized dynamical approach to the large time behavior of solutions of Hamilton-Jacobi equations, *SIAM J. Math. Anal.*, 38 (2006), 478–502.
[14] A. Douglis, Solutions in the large for multi-dimensional, non-linear partial differential equations of first order, *Ann. Inst. Fourier (Grenoble)*, 15 (1965), 1–35.
[15] W. E, Aubry-Mather theory and periodic solutions of the forced Burgers equation, *Comm. Pure Appl. Math.*, 52 (1999), 811–828.
[16] A. Fathi, Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens, *C. R. Acad. Sci. Paris Sér. I Math.*, 324 (1997), 1043–1046.
[17] A. Fathi, Sur la convergence du semi-groupe de Lax-Oleinik, *C. R. Acad. Sci. Paris Sér. I Math.*, 327 (1998), 267–270.
[18] A. Fathi, *Weak KAM Theorem in Lagrangian Dynamics*, Preliminary Version Number 10, 2008.
[19] A. Fathi and J. N. Mather, Failure of convergence of the Lax-Oleinik semi-group in the time-periodic case, *Bull. Soc. Math. France.*, 128 (2000), 473–483.
[20] N. Ichihara and H. Ishii, Long-time behavior of solutions of Hamilton-Jacobi equations with convex and coercive Hamiltonians, *Arch. Ration. Mech. Anal.*, 194 (2009), 383–419.
[21] J. N. Mather, Existence of quasi periodic orbits for twist homeomorphisms of the annulus, *Topology*, 21 (1982), 457–467.
[22] J. N. Mather, More Denjoy minimal sets for area preserving diffeomorphisms, *Comment. Math. Helv.*, 60 (1985), 508–557.
[23] J. N. Mather, A criterion for the non-existence of invariant circle, *Publ. Math. IHES*, 63 (1986), 301–309.
[24] J. N. Mather, Modulus of continuity for Peierls’s barrier, *Periodic Solutions of Hamiltonian Systems and Related Topics*, ed. P.H.Rabinowitz et al NATO ASI Series C 209, Reidel: Dordrecht, (1987), 177–202.
[25] J. N. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, *Math. Z.*, 207 (1991), 169–207.
[26] G. Namah and J.-M. Roquejoffre, Remarks on the long time behavior of the solutions of Hamilton-Jacobi equations, *Commun. Partial Differ. Equ.*, 24 (1999), 883–893.
[27] K. Wang and J. Yan, A new kind of Lax-Oleinik type operator with parameters for time-periodic positive definite Lagrangian systems, *Comm. Math. Phys.*, 309 (2012), 663–691.

Received September 2015; revised July 2016.

E-mail address: xfsu@bnu.edu.cn
E-mail address: lwang@math.tsinghua.edu.cn
E-mail address: yanjun@fudan.edu.cn