SCIENCE CHINA
Mathematics

• ARTICLES •

September 2016 Vol. 59 No. 9: 1737–1752
doi: 10.1007/s11425-016-5143-4

Lasry-Lions, Lax-Oleinik and generalized characteristics

CHEN Cui & CHENG Wei*

Department of Mathematics, Nanjing University, Nanjing 210093, *China*
Email: chenc@ujs.edu.cn, chengwei@nju.edu.cn

Received September 2, 2015; accepted December 8, 2015; published online April 1, 2016

Abstract In the recent works, an intrinsic approach of the propagation of singularities along the generalized characteristics was obtained, even in global case, by a procedure of sup-convolution with the kernel the fundamental solutions of the associated Hamilton-Jacobi equations. In the present paper, we exploit the relations among Lasry-Lions regularization, Lax-Oleinik operators (or inf/sup-convolution) and generalized characteristics, which are discussed in the context of the variational setting of Tonelli Hamiltonian dynamics, such as Mather theory and weak KAM (Kolmogorov-Arnold-Moser) theory.

Keywords Hamilton-Jacobi equations, weak KAM theory, generalized characteristics

MSC(2010) 26B25, 35A21, 49L25, 37J50, 70H20

Citation: Chen C, Cheng W. Lasry-Lions, Lax-Oleinik and generalized characteristics. Sci China Math, 2016, 59: 1737–1752, doi: 10.1007/s11425-016-5143-4

1 Introduction

Suppose $H = H(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a C^2 Tonelli Hamiltonian, i.e., H is convex in p with superlinear growth condition. Let $u : \mathbb{R}^n \to \mathbb{R}$ be a (global) viscosity solution of the Hamilton-Jacobi equation

$$H(x, Du(x)) = 0, \quad x \in \mathbb{R}^n. \tag{1.1}$$

Such a solution u is locally semiconcave (with linear modulus) on $\mathbb{R}^n$. We denote by $D^+u(x)$ the superdifferential of u at x (see, for example, [12]), which is a compact convex set in $\mathbb{R}^n$, and we call $x \in \mathbb{R}^n$ a singular point of u if $D^+u(x)$ is not a singleton. Certain "singular dynamics" was interpreted in [2] by a Hamiltonian inclusion

$$\dot{\boldsymbol{x}}(s) \in \operatorname{co} H_p(\boldsymbol{x}(s), D^+u(\boldsymbol{x}(s))), \quad \text{a.e.} \quad s \in [0, \tau],$$

and such a Lipschitz arc $\boldsymbol{x}$ is called a *generalized characteristic*. If x_0 is a singular point of u and

$$0 \notin \operatorname{co} H_p(x_0, D^+u(x_0)), \tag{1.2}$$

then the associated generalized characteristic $\boldsymbol{x}(t)$, $t \in [0, \tau]$, is composed of singular points of u. In the recent works [10], the propagation of singularities along generalized characteristics in [10] has been explained by an intrinsic variational approach (see [2–4, 11–13, 29] for the approach from Control theory or PDE), which is motivated by Mather theory (see [25]) and weak KAM theory.

*Corresponding author

Let us recall the aforementioned results in [10] at first. Let $u \in C(\mathbb{R}^n)$, for any $t > 0$, $\breve{T}_t$, the Lax-Oleinik operator of positive type, is defined as

$$\breve{T}_t u(x) := \sup_{y\in\mathbb{R}^n} \{u(y) - A_t(x, y)\} := \sup_{y\in\mathbb{R}^n} \psi_t^x(y), \quad x \in \mathbb{R}^n, \tag{1.3}$$

where

$$A_t(x, y) = \inf_{\gamma\in\Gamma^t_{x,y}} \int_0^t L(\gamma(s), \dot\gamma(s))ds$$

with

$$\Gamma^t_{x,y} = \{\gamma \in W^{1,1}([0, t], \mathbb{R}^n) : \gamma(0) = x, \gamma(t) = y\}.$$

Here L is an arbitrary Tonelli Lagrangian on $\mathbb{R}^n$ with H its Fenchel-Legendre dual, and it is well known that (1.3) is also called sup-convolution or Lax-Oleinik operators in the literature. Cannarsa and Cheng [10] have shown that the maximizers in such a procedure of sup-convolution give exactly a local or global generalized characteristic starting from a singular point of u under suitable conditions.

In the present paper, we will explain the connection between generalized characteristics and the well-known Lasry-Lions regularization at first. Throughout this paper, we suppose that L satisfies (L1) and (L2) (see Section 2).

Let M be a C^2 closed manifold, $t > 0$ and let $u : M \to \mathbb{R}$ be any semiconcave function, the following properties are already known (see, for example, [6, 22]):

(P1) $\breve{T}_t u$ belongs to class $C^{1,1}$ for $0 < t \leqslant t_0$ with t_0 is a constant dependent on the constant of semiconcavity of u.

(P2) $\breve{T}_t u$ is decreasing on $(0, +\infty)$ if u is a viscosity subsolution of Hamilton-Jacobi equation

$$H(x, Du(x)) = \alpha(0), \quad x \in M,$$

where $\alpha(\cdot)$ is Mather's α-function. Moreover, $\breve{T}_t u$ tends to u uniformly as $t \to 0^+$.

In this paper, we also have the following theorem.

Theorem 1.1. *Suppose $u : \mathbb{R}^n \to \mathbb{R}$ is a semiconcave function. Then there exists $0 < t_0 \ll 1$ such that*

(P3) *Let $x_0 \in \mathbb{R}^n$ and $L(x_0, 0) \leqslant 0$, then $\breve{T}_t u(x_0)$ is increasing on $(0, +\infty)$ and*

$$\lim_{t\to 0^+} \breve{T}_t u(x_0) = u(x_0).$$

Consequently, if

$$L(x, 0) \leqslant 0, \quad \forall\, x \in \mathbb{R}^n,$$

$\breve{T}_t u$ tends to u uniformly on any compact subset as $t \to 0^+$.

(P4) *Let $x \in \mathbb{R}^n$, suppose that the function ψ_t^x defined in* (1.3) *attains the maximizer y_t in $B(x, R(x, t))$ for all $0 < t \leqslant t_0$, where $R(x, t) > 0$ is defined in* (A.6)*. Then*

$$\lim_{t\to 0^+} D\breve{T}_t u(x) = p_0 \in D^+u(x),$$

where p_0 is the unique element with minimal energy

$$H(x, p) \geqslant H(x, p_0), \quad \forall\, p \in D^+u(x). \tag{ME}$$

For a more detailed formulation of Theorem 1.1, see Theorem 3.2.

It is worth noting that the minimal energy condition in (ME) is the same as the initial condition on the velocity of the generalized characteristic obtained by the intrinsic approach in [10], see also Proposition 3.1.

In the rest part of this paper, we try to exploit the nature of the singularities of u by the procedure of inf-convolution. As pointed out that the inf-convolution defined by

$$T_t u(x) := \inf_{y\in\mathbb{R}^n} \{u(y) + A_t(y, x)\}, \quad x \in \mathbb{R}^n \tag{1.4}$$

is not the dual procedure of sup-convolution. But, it is still meaningful to study the critical points of the local barrier function

$$\phi_t^{x_0}(x) = u(x) + A_t(x, x_0)$$

with respect to x_0. Recall that a point $x \in \mathbb{R}^n$ is a *critical point* of a locally semiconcave function u if $0 \in D^+u(x)$. Comparing with the local barrier function

$$\psi_t^{x_0}(x) = u(x) - A_t(x_0, x),$$

the function $\psi_t^{x_0}$ only admits a unique critical point (maximizer) for small time $t > 0$ since the convexity properties of the fundamental solutions $A_t(x_0, x)$ (see Appendix).

Along this line, given a singular point x_0 of u, using a nonsmooth critical point theorem by Shi [28], we obtain a critical point of the local semiconcave function $\phi_t^{x_0}$, which is not a global minimizer of $\phi_t^{x_0}$ determined by classical characteristic passing to x.

Theorem 1.2. *Let u be a Lipschitz viscosity solution of (1.1), $t > 0$, and let $x \in \mathbb{R}^n$ be a singular point of u. Suppose there exist finite many elements in $D^*u(x)$, the set of all limiting differentials of u at x (see Definition 2.2), say $D^*u(x) = \{p_1, \dots, p_k\}$ with $k \geqslant 2$. Then there exist critical points $\{x_t^{ij}\}$ of ϕ_t^x (not global minimizers) such that, for $1 \leqslant i, j \leqslant k$, $i \neq j$, each critical point $x_t = x_t^{ij}$ has the following dichotomy*:

(a) *x_t is a differentiable point of ϕ_t^x and there exists a local minimal curve connecting x_t and x. More precisely, there exists a C^1 curve $\gamma : (-\infty, t] \to \mathbb{R}^n$ such that $\gamma(0) = x_t$, $\gamma(t) = x$ and the restriction of γ on $(-\infty, 0]$ is a $(u, L, 0)$-calibrated curve, but γ is not a $(u, L, 0)$-calibrated curve on $(-\infty, t]$*;

(b) *x_t is a singular point of u.*

For a more detailed formulation of Theorem 1.2, see Theorem 3.9.

From the theorem above, the location of singularities afford possible information to construct "local" minimal orbits for Tonelli Lagrangian systems, which is totally unknown before. In the previous works of variational approach of Hamiltonian dynamical instability problems like Arnold diffusion (see, for example, [7, 14–20]), the diffusion orbits shadow the relevant global minimizers.

The paper is organized as follows. In Section 2, we review some basic properties of viscosity solution of Hamilton-Jacobi equations. In Section 3, we discuss the relation of the generalized characteristics given by the procedure of sup-convolution and Lasry-Lions regularization, then, we also discuss what happens for the procedure of inf-convolution. Section 4 concludes the paper. We present regularity properties of fundamental solutions in Appendix.

2 Viscosity solutions and weak KAM theory

A C^2 function $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is said to be a *Tonelli Lagrangian* if the following assumptions are satisfied:

(L1) The Hessian $\frac{\partial^2 L}{\partial v^2}(x, v)$ is positive definite for all $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$.

(L2) There exists a non-decreasing function $\theta : [0, +\infty) \to [0, +\infty)$, $\theta(r)/r \to +\infty$ as $t \to +\infty$, $c_0 \geqslant 0$ and, $c_1 = c_1(x, R) \geqslant 0$ such that

$$L(x, v) \geqslant \theta(|v|) - c_0, \quad (x, v) \in \mathbb{R}^n \times \mathbb{R}^n,$$

and

$$|L_x(y, v)| + |L_v(y, v)| \leqslant c_1(x, R)\theta(|v|), \quad (y, v) \in \bar{B}(x, R) \times \mathbb{R}^n.$$

Let $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the associated *Tonelli Hamiltonian*, i.e.,

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

Throughout this paper, we suppose L is a C^2 Tonelli Lagrangian with (L1) and (L2).

2.1 Semiconcave functions

Let $\Omega \subset \mathbb{R}^n$ be a convex open set, a function $u : \Omega \to \mathbb{R}$ is *semiconcave* (with linear modulus) if there exists a constant $C > 0$ such that

$$\lambda u(x) + (1-\lambda)u(y) - u(\lambda x + (1-\lambda)y) \leqslant \frac{C}{2}\lambda(1-\lambda)|x-y|^2 \tag{2.1}$$

for any $x, y \in \Omega$ and $\lambda \in [0,1]$. Any constant C that satisfies the above inequality is called a *semiconcavity constant* for u in Ω. A function $u : \Omega \to \mathbb{R}$ is said to be *semiconvex* (with linear modulus) if $-u$ is semiconcave. A function $u : \Omega \to \mathbb{R}$ is said to be *locally semiconcave* (resp. *locally semiconvex*) if for each $x \in \Omega$, there exists an open ball $B(x,r) \subset \Omega$ such that u is a semiconcave (resp. semiconvex) function on $B(x,r)$.

Definition 2.1. Let $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a continuous function. We recall that, for any $x \in \Omega$, the closed convex sets

$$D^-u(x) = \left\{p \in \mathbb{R}^n : \liminf_{y\to x} \frac{u(y) - u(x) - \langle p, y-x\rangle}{|y-x|} \geqslant 0\right\},$$
$$D^+u(x) = \left\{p \in \mathbb{R}^n : \limsup_{y\to x} \frac{u(y) - u(x) - \langle p, y-x\rangle}{|y-x|} \leqslant 0\right\}$$

are called the (*Dini*) *subdifferential* and *superdifferential* of u at x, respectively.

Definition 2.2. Let $u : \Omega \to \mathbb{R}$ be locally Lipschitz. We recall that a vector $p \in \mathbb{R}^n$ is called a *limiting differential* of u at x if there exists a sequence $\{x_n\} \subset \Omega \setminus \{x\}$ such that u is differentiable at x_k for each $k \in \mathbb{N}$, and

$$\lim_{k\to\infty} x_k = x \quad \text{and} \quad \lim_{k\to\infty} Du(x_k) = p.$$

The set of all limiting differentials of u at x is denoted by $D^*u(x)$.

The fundamental properties of the superdifferential of a semiconcave function are listed in the following proposition. The monograph [12] is a good reference for the topic of semiconcave functions and beyond.

Proposition 2.3. *Let $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a semiconcave function and let $x \in \Omega$. Then the following properties hold*:

(a) *$D^+u(x)$ is a nonempty compact convex set in $\mathbb{R}^n$ and $D^*u(x) \subset \partial D^+u(x)$, where $\partial D^+u(x)$ denotes the topological boundary of $D^+u(x)$.*

(b) *The set-valued function $x \rightsquigarrow D^+u(x)$ is upper semicontinuous.*

(c) *If $D^+u(x)$ is a singleton, then u is differentiable at x. Moreover, if $D^+u(x)$ is a singleton for every point in Ω, then $u \in C^1(\Omega)$.*

(d) $D^+u(x) = \operatorname{co} D^*u(x)$.

(e) $D^*u(x) = \{\lim_{i\to\infty} p_i : p_i \in D^+u(x_i),\ x_i \to x,\ \operatorname{diam}(D^+u(x_i)) \to 0\}$.

From proximal analysis point of view, the following result characterizes the semiconcavity of a continuous function and its superdifferential.

Proposition 2.4. *Let $u : \Omega \to \mathbb{R}$ be a continuous function. If there exists a constant $C > 0$ such that, for any $x \in \Omega$, there exists $p \in \mathbb{R}^n$ such that*

$$u(y) \leqslant u(x) + \langle p, y-x\rangle + \frac{C}{2}|y-x|^2, \quad \forall\, y \in \Omega, \tag{2.2}$$

then u is semiconcave with constant C and $p \in D^+u(x)$. Conversely, if u is semiconcave in Ω with constant C, then (2.2) *holds for any $x \in \Omega$ and $p \in D^+u(x)$.*

Finally, we introduce the concept of singularity of a semiconcave function. A point $x \in \Omega$ is called a *singular point* of u if $D^+u(x)$ is not a singleton. The set of all singular points of u, also called the *singular set* of u, is denoted by Σ_u.

2.2 Fundamental solutions and viscosity solutions

Given $x, y \in \mathbb{R}^n$, we define

$$\Gamma^t_{x,y} = \{\gamma \in W^{1,1}([0,t], \mathbb{R}^n) : \gamma(0) = x, \gamma(t) = y\}.$$

Letting $t > 0$, we denote

$$A_t(x, y) = \inf_{\gamma \in \Gamma^t_{x,y}} \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds. \tag{2.3}$$

It is well known that the infimum can be achieved by C^2 curves. In the literature of PDEs, $A_t(x, y)$ is called a *fundamental solution* of (2.4), see, for example, [26].

Throughout this section, we suppose the C^2 Tonelli Lagrangian L satisfies (L1)–(L2). We discuss the associated Nagumo type conditions and the essential regularity results of the fundamental solutions in Appendix. For the main regularity results we will use, see, Propositions A.3 and A.4.

Suppose H is a Tonelli Hamiltonian, throughout this paper we will be concerned with the Hamilton-Jacobi equation

$$H(x, Du(x)) = 0, \quad x \in \mathbb{R}^n. \tag{2.4}$$

We recall that a continuous function u is called a *viscosity subsolution* of equation (2.4) if, for any $x \in \mathbb{R}^n$,

$$H(x, p) \leqslant 0, \quad \forall\, p \in D^+u(x). \tag{2.5}$$

Similarly, u is a *viscosity supersolution* of (2.4) if, for any $x \in \mathbb{R}^n$,

$$H(x, p) \geqslant 0, \quad \forall\, p \in D^-u(x). \tag{2.6}$$

Finally, u is called a *viscosity solution* of (2.4), if it is both a viscosity subsolution and a supersolution.

Proposition 2.5. *Any viscosity solution of the Hamilton-Jacobi equation* (2.4) *is locally semiconcave with linear modulus.*

Proposition 2.6. *There holds* $\operatorname{Ext} D^+u(x) = D^*u(x)$[1)] *for any viscosity solution u of* (2.4) *and any $x \in \mathbb{R}^n$.*

Proposition 2.7 (See [27])**.** *Let $x \in \mathbb{R}^n$ and $u : \mathbb{R}^n \to \mathbb{R}$ be a viscosity solution of the Hamilton-Jacobi equation* (2.4)*. Then $p \in D^*u(x)$ if and only if there exists a unique C^2 curve $\gamma : (-\infty, 0] \to \mathbb{R}^n$ with $\gamma(0) = x$ which is a $(u, L, 0)$-calibrated curve*[2)]*, and $p = L_v(x, \dot{\gamma}(0))$.*

2.3 Generalized characteristics

The construction of the singular set or cut loci of viscosity solutions is a very important and hard problem in many fields such as Riemannian geometry, optimal control, classical mechanics, etc. It is known that the study of propagation of singularities can go back to [1] for general semiconcave functions by the method from nonsmooth analysis. Some dynamical nature of the singularity was found by the concept of generalized characteristic.

Definition 2.8. A Lipschitz arc $\boldsymbol{x} : [0, \tau] \to \mathbb{R}^n$ is said to be a *generalized characteristic* of the Hamilton-Jacobi equation (2.4) if $\boldsymbol{x}$ satisfies the differential inclusion

$$\dot{\boldsymbol{x}}(s) \in \operatorname{co} H_p(\boldsymbol{x}(s), D^+u(\boldsymbol{x}(s))), \quad \text{a.e. } s \in [0, \tau]. \tag{2.7}$$

A basic criterion for the propagation of singularities along generalized characteristic was given in [2] (see [13, 29] for an improved version and simplified proof).

---

1) For any convex closed subset of $\mathbb{R}^n$, we denote by $\operatorname{Ext} C$ the set of extremal points of C.

2) For the concept of dominated functions and calibrated curves, see, for example, [22]

Proposition 2.9 (See [2]). *Let u be a viscosity solution of Hamilton-Jacobi equation* (2.4) *and let $x_0 \in \mathbb{R}^n$. Then there exists a generalized characteristic $\boldsymbol{x} : [0,\tau] \to \mathbb{R}^n$ with initial point $\boldsymbol{x}(0) = x_0$. Moreover, if $x_0 \in \Sigma_u$, then $\boldsymbol{x}(s) \in \Sigma_u$ for all $s \in [0,\tau]$. Furthermore, if*

$$0 \notin \operatorname{co} H_p(x_0, D^+u(x_0)), \tag{2.8}$$

then $\boldsymbol{x}(\cdot)$ is injective for every $s \in [0,\tau]$.

3 Procedure of sup-convolution and generalized characteristics

Let H be a Tonelli Hamiltonian on $\mathbb{R}^n$. Recall the Lax-Oleinik operators T_t and $\breve{T}_t$, i.e., for any $u \in C(\mathbb{R}^n)$,

$$\breve{T}_t u(x) := \sup_{y\in\mathbb{R}^n} \{u(y) - A_t(x,y)\}, \tag{3.1}$$

$$T_t u(x) := \inf_{y\in\mathbb{R}^n} \{u(y) + A_t(y,x)\}. \tag{3.2}$$

When taking $H(p) = |p|^2/2$ and the kernel

$$A_t(x,y) = \frac{1}{2t}|x-y|^2,$$

the two operators above are closely linked to the so-called Lasry-Lions regularization procedure (see [24]) which is written in the form of *sup-convolution* and *inf-convolution*, respectively. This type of regularization is also called Moreau-Yosida regularization in convex analysis. A more detailed formulation can be found in [5] with respect to the aforementioned quadratic kernel.

3.1 Procedure of sup-convolution and generalized characteristics

Recently, Cannarsa and Cheng [10] studied the intrinsic relation of propagation of singularities along the generalized characteristics and the following procedure of sup-convolution.

Fixing $x \in \mathbb{R}^n$, $0 < t \leqslant t_0 \ll 1$, then there exists $R(x,t) > 0$ such that, the function

$$\psi_t^x(y) := u(y) - A_t(x,y), \quad y \in \bar{B}(x, R(x,t)) \tag{3.3}$$

has a unique maximizer for each $t \in (0, t_0]$, where $A_t(x,y)$ is a fundamental solution with respect to the associated Tonelli Lagrangian L.

Suppose that $u(\cdot)$ is semiconcave while $A_t(x,\cdot)$ is both locally semiconcave, and convex when $t \in (0,t_0]$ (see Proposition A.3), say $C_1 > 0$ (resp. $-C_2(t)$) is the semiconcavity (resp. convexity) constant of $u(\cdot)$ (resp. $A_t(x,\cdot)$). Note that, by Proposition A.4, the constant $C_2(t) = \frac{C}{t}$, thus $\psi_t^x(\cdot)$ is strictly concave in $\bar{B}(x, R(x,t))$ and consequently we have a unique maximizer for each $t \in (0,t_0]$, which is also a unique critical point of ψ_t^x if $y_t \in B(x, R(x,t))$.

Let us define the arc $\boldsymbol{y} : [0,t_0] \to \mathbb{R}^n$ by

$$\boldsymbol{y}(t) = \begin{cases} x, & t = 0, \\ y_t, & t \in (0, t_0]. \end{cases} \tag{3.4}$$

If $\xi_t : [0,t] \to \mathbb{R}^n$ is the unique minimizer in the definition of $A_t(x,y)$, we define

$$p_t(s) := L_v(\xi_t(s), \dot{\xi}_t(s)), \quad s \in [0,t], \tag{3.5}$$

the associated dual arc with respect to $\xi_t(s)$.

Proposition 3.1 (See [10]). *Let u be a locally semiconcave function and $x \in \Sigma_u$, the singular set of u. If ϕ_t^x attains the unique maximizer $y_t \in B(x, R(x,t))$ for all $t \in [0,t_0]$, then the arc $\boldsymbol{y} : [0,t_0] \to \mathbb{R}^n$*

defined in (3.4) *is a generalized characteristic composed of singular points of* u, *i.e.*, $\boldsymbol{y} : [0, t_0] \to \mathbb{R}^n$ *is Lipschitz continuous*, $\boldsymbol{y}(t) \in \Sigma_u$ *for all* $t \in [0, t_0]$, *and satisfies*

$$\dot{\boldsymbol{y}}(\tau) \in \operatorname{co} H_p(\boldsymbol{y}(\tau), D^+u(\boldsymbol{y}(\tau))), \qquad a.e. \ \ \tau \in [0, t_0]. \tag{3.6}$$

Moreover,

$$\dot{\boldsymbol{y}}^+(0) = H_p(x, p_0), \tag{3.7}$$

where $p_0 \in D^+u(x)$ *is the unique element of minimal energy*

$$H(x, p) \geqslant H(x, p_0), \quad \forall\, p \in D^+u(x).$$

3.2 Lasry-Lions regularization

In this section, we will explain the connection between Lasry-Lions regularization (see [24]) and generalized characteristics first found in [2]. We only concentrate on the case of sup-convolution $\breve{T}_t u$ with u a locally semiconcave function.

For $t > 0$, recalling that

$$\breve{T}_t u(x) := \sup_{y \in \mathbb{R}^n} \{u(y) - A_t(x, y)\}, \tag{3.8}$$

where $u : \mathbb{R}^n \to \mathbb{R}$ is any locally semiconcave function, and $A_t(x, y)$ is the fundamental solution with respect to any Tonelli Lagrangian L.

Theorem 3.2. *Suppose* $u : \mathbb{R}^n \to \mathbb{R}$ *be a semiconcave function with constant* C. *Then there exists* $0 < t_0 \ll 1$ *such that if* $\breve{T}_t u$ *is defined as in* (3.8), *we have the following*:

(P3) *Fix* $x \in \mathbb{R}^n$, *then* $\breve{T}_t u(x)$ *is increasing on* $(0, +\infty)$ *and* $\lim_{t \to 0^+} \breve{T}_t u(x) = u(x)$ *if* $L(x, 0) \leqslant 0$. *Consequently, if*

$$L(x, 0) \leqslant 0, \quad \forall\, x \in \mathbb{R}^n,$$

then $\breve{T}_t u$ *tends to* u *uniformly on any compact subset as* $t \to 0^+$.

(P4) *Let* $x \in \mathbb{R}^n$, *suppose that the function* ψ_t^x *defined in* (3.3) *attains the maximizer* y_t *in* $B(x, R(x, t))$ *for all* $t \in (0, t_0]$. *Then*

$$\lim_{t \to 0^+} D\breve{T}_t u(x) = p_0,$$

where p_0 *is the unique element with minimal energy, i.e.*,

$$H(x, p_0) = \min_{p \in D^+u(x)} H(x, p). \tag{3.9}$$

(P5) *In particular, when* L *has the form*

$$L(x, v) = \frac{1}{2}\langle Av, v\rangle, \quad x \in \mathbb{R}^n, \quad v \in \mathbb{R}^n,$$

where A *is an* $n \times n$ *symmetric and positive definite matrix. If* $t \leqslant \kappa C^{-1}$, *then, the functions* u *and* $\breve{T}_t u$ *have the same critical points and critical values where* $\kappa > 0$ *is the smallest eigenvalue of* A.

Remark 3.3. The properties (P1) and (P2) (see the introduction) is already known (in the case of compact manifolds), see, for example, [6] or [22]. Since this is a local result, it is not hard to generalize to the manifolds using local charts. We collect the known results here just for the comparison interests like (P2) (in the introduction) and (P3). The property (P5) is a slight generalization of a known result (see [5]).

Remark 3.4. It is worth noting that the assumption in (P4) that the function ψ_t^x defined in (3.3) attains the maximizer y_t in $B(x, R(x, t))$ for all $t \in (0, t_0]$, is not easy to be checked in general. Fortunately, if we consider a certain type of nearly integrable systems or mechanical systems, this condition holds. The readers can refer to [10] for more general discusssion.

Proof. Let $x, y \in \mathbb{R}^n$ and $t > 0$, for any $0 < s < t$, it is easily checked that

$$A_t(x, y) \leqslant A_s(x, y) + A_{t-s}(x, x).$$

Taking the constant curve $\gamma(\tau) \equiv x$, $\tau \in [0, t-s]$, we have

$$A_{t-s}(x, x) \leqslant \int_0^{t-s} L(\gamma(\tau), \dot{\gamma}(\tau)) d\tau = (t-s)L(x, 0).$$

Therefore, for any fixed $x \in \mathbb{R}^n$, we have $A_t(x, \cdot) \leqslant A_s(x, \cdot)$ since $L(x, 0) \leqslant 0$, and thus, $\psi_t^x(y) \geqslant \psi_s^x(y)$ for all $y \in \mathbb{R}^n$. This leads to the conclusion that $\breve{T}_s u(x) \leqslant \breve{T}_t u(x)$ if $0 < s < t$. The uniform convergence result is a direct consequence of Dini's lemma on monotone sequence of continuous functions. This completes the proof of (P3).

Now, we turn to the proof of (P4). Fix $x \in \mathbb{R}^n$ and $t \in (0, t_0]$. Adopting the same terminologies as before, since $\psi_t^x(\cdot)$ attains the maximum at $y = y_t \in B(x, R(x, t))$ and $\xi_t \in \Gamma_{x, y_t}^t$ is the minimal curve in the definition of $A_t(x, y_t)$, we have

$$L_v(\xi_t(t), \dot{\xi}_t(t)) = D_y A_t(x, y_t) \in D^+ u(y_t),$$

since the results in Proposition A.4 and $0 \in D^+ \psi_t^x(y_t)$. Moreover, we have that the family $\{\dot{\xi}_t\}_{t \in (0, t_0]}$ is equi-Lipschitz, by Proposition A.2.

Therefore, we have

$$\begin{aligned} \left| \frac{\xi_t(t) - x}{t} - \dot{\xi}_t(0) \right| &\leqslant \frac{1}{t} \int_0^t |\dot{\xi}_t(s) - \dot{\xi}_t(0)| ds \\ &\leqslant \frac{C}{t} \int_0^t s ds = \frac{C}{2} t. \end{aligned}$$

Thus, we obtain

$$v_0 = \lim_{t \to 0^+} v_t = \lim_{t \to 0^+} \dot{\xi}_t(0) = \lim_{t \to 0^+} \dot{\xi}_t(t),$$

where $v_t = (y_t - x)/t$. Since u is a locally semiconcave function, by the monotone property of semiconcave functions (see, e.g., [12]), we have

$$\langle p - L_v(\xi_t(t), \dot{\xi}_t(t)), v_t \rangle + tC|v_t|^2 \geqslant 0, \quad \forall\, p \in D^+ u(x). \tag{3.10}$$

Taking limit in (3.10), then

$$\langle p, v_0 \rangle \geqslant \langle L_v(x, v_0), v_0 \rangle, \quad \forall\, p \in D^+ u(x). \tag{3.11}$$

In other words,

$$H(x, p) \geqslant \langle L_v(x, v_0), v_0 \rangle - L(x, v_0) = H(x, p_0), \quad \forall\, p \in D^+ u(x), \tag{3.12}$$

where

$$p_0 = L_v(x, v_0) \in D^+ u(x),$$

by the upper semicontinuity of the set valued function $x \rightsquigarrow D^+ u(x)$, is the unique element solve the associated optimization problem (3.12), and

$$\lim_{t \to 0^+} D\breve{T}_t u(x) = \lim_{t \to 0^+} L_v(\xi_t(0), \dot{\xi}_t(0)) = p_0.$$

This completes the proof of (P4).

For the proof of (P5), note that, in our case, the minimal curve

$$\xi_t(s) = \frac{y_t - x}{t} \cdot s,$$

thus, by (3.10), we have

$$\langle p - Av_t, v_t\rangle + tC|v_t|^2 \geqslant 0, \quad \forall\, p \in D^+u(x).$$

If $0 \in D^+u(x)$, take $p = 0$ in the inequality above, then it follows

$$\langle -\kappa v_t, v_t\rangle + tC|v_t|^2 \geqslant \langle -Av_t, v_t\rangle + tC|v_t|^2 \geqslant 0, \quad \forall\, p \in D^+u(x),$$

where $\kappa > 0$ is the smallest eigenvalue of A. Therefore,

$$(tC - \kappa)|v_t|^2 \geqslant 0, \quad \forall\, p \in D^+u(x).$$

If $t \leqslant \kappa C^{-1}$, then $v_t \equiv 0$, $y_t \equiv x$ and $u(x) = \breve{T}_t u(x)$. Conversely, if $0 = D\breve{T}_t u(x)$, then $v_t = 0$ and $y_t \equiv x$. It follows $0 \in D^+u(x)$ which proves (P5). □

3.3 What happens for the inf-convolution

In this subsection, we discuss the procedure of inf-convolution. Let u be a locally semiconcave function on $\mathbb{R}^n$, and let L be a C^2 Tonelli Lagrangian, for any fixed $x \in \mathbb{R}^n$, define

$$\phi_t^x(y) := u(y) + A_t(y, x), \quad y \in \mathbb{R}^n.$$

It is worth noting that ϕ_t^x is the sum of two locally semiconcave functions, and it is also locally semiconcave consequently.

For the convenience of our discussion, we suppose that u is a global viscosity solution of the Hamilton-Jacobi equation

$$H(x, Du(x)) = 0, \quad x \in \mathbb{R}^n, \tag{3.13}$$

where H is the associated Hamiltonian with respect to L.

At this stage, we have

$$u(x) = T_t u(x) = \inf_{y \in \mathbb{R}^n} \phi_t^x(y)$$

for all $t > 0$ by the well-known facts from weak KAM theory.

Lemma 3.5. *Let u be a viscosity solution of (3.13), and ϕ_t^x be defined as above for $t > 0$. Then for $t > 0$, there exists z_t such that*

$$\phi_t^x(z_t) = \inf_{y \in \mathbb{R}^n} \phi_t^x(y).$$

Proof. This is actually obvious. Indeed, by Proposition 2.7, for any $t > 0$, and $p \in D^*u(x)$, there exists a C^2 curve $\gamma : (-\infty, t] \to \mathbb{R}^n$ such that $\gamma(t) = x$, $p = L_v(\gamma(t), \dot\gamma(t))$ and

$$u(\gamma(t)) - u(\gamma(s)) = \int_s^t L(\gamma(\tau), \dot\gamma(\tau))\, d\tau, \quad \forall\, s < t.$$

Taking $z_t = \gamma(0)$, then we have the expected result. □

Now, we can impose such a question: Is the aforementioned procedure of inf-convolution efficient for tracking the information of the propagation of singularities along generalized characteristics?

We will try to answer this question using the technique from nonsmooth critical point theory, see also [9] for the applications by standard using Lasry-Lions regularization.

Lemma 3.6. *Let u be a viscosity solution of (3.13) and the function ϕ_t^x be defined as above for any fixed $x \in \mathbb{R}^n$ and $t > 0$. Then there exists a one-to-one correspondence between $p \in D^*u(x)$ and the global minimizers z_t of ϕ_t^x for all $t > 0$.*

Proof. Let $z_t \in \mathbb{R}^n$ be a minimizer of ϕ_t^x, $t > 0$. Then ϕ_t^x is differentiable at z_t since ϕ_t^x is locally semiconcave. Thus, z_t is a differentiable point for both u and $A_t(\cdot, x)$. Consequently, there exist two C^1 curves $\gamma_1 : (-\infty, 0] \to \mathbb{R}^n$ and $\gamma_2 \in \Gamma_{z_t,x}^t$ such that

$$\gamma_1(0) = \gamma_2(0) = z_t,$$

$$p = Du(z_t) = L_v(\gamma_1(0), \dot\gamma_1(0)),$$
$$p' = D_x A_t(z_t, x) = -L_v(\gamma_2(0), \dot\gamma_2(0)),$$

by Propositions 2.7 and A.4, and $p + p' = 0$ since z_t is a critical point of ϕ_t^x. Moreover, γ_1 is a $(u, L, 0)$-calibrated curve, i.e., for any $s > 0$,

$$u(\gamma_1(0)) - u(\gamma_1(-s)) = \int_{-s}^{0} L(\gamma_1(\tau), \dot\gamma_1(\tau))d\tau,$$

and, similarly,

$$u(x) - u(\gamma_2(0)) = A_t(\gamma_2(0), x) = \int_0^t L(\gamma_2(\tau), \dot\gamma_2(\tau))d\tau.$$

By the juxtaposition of γ_1 and γ_2, we define

$$\eta_t(\tau) = \begin{cases} \gamma_1(\tau), & \tau \leqslant 0, \\ \gamma_2(\tau), & 0 < \tau \leqslant t. \end{cases}$$

It is clear that η_t is a C^1 curve on $(-\infty, t]$ with $\eta_t(t) = x$, and

$$u(x) - u(\eta_t(-s)) = \int_{-s}^{t} L(\eta_t(\tau), \dot\eta_t(\tau))d\tau, \quad s > 0,$$

which follows that η_t is also a $(u, L, 0)$-calibrated curve, and such a $(u, L, 0)$-calibrated curve passing through z_t with x the terminal datum is unique. Therefore, the correspondence between z_t and η_t is one-to-one.

The rest of the proof is a direct consequence of Proposition 2.7. □

Now, we fix a point $x \in \mathbb{R}^n$.

(1) If x is a differentiable (or regular) point of u, then $D^*u(x) = \{Du(x)\}$, and ϕ_t^x has a unique global minimizer z_t which determines a unique $(u, L, 0)$-calibrated curve passing though z_t with x the terminal endpoint.

(2) If x is singular point of u, it become relatively complicated. Let

$$Z_{x,E} = \{p \in \mathbb{R}^n : H(x, p) \leqslant E\},$$

which is a non-empty compact and convex set when the energy E, say $E = 0$, is suitably chosen. It is known that $D^*u(x) = \operatorname{Ext} D^+u(x)$, the set of extremal points of $D^+u(x)$, by Proposition 2.6. This means the elements of $D^*u(x)$ is exactly the set $\operatorname{Ext} D^+u(x)$ which is located in the energy hypersurface $\partial Z_{x,E}$ since $H(x, \cdot)$ is strictly convex.

In the spirit of Lemma 3.6, we want to look for the critical points of ϕ_t^x. A point $x \in \mathbb{R}^n$ is a *critical point* of a locally semiconcave function u if $0 \in D^+u(x)$. To find the critical points of ϕ_t^x besides the global minimizers as in Lemma 3.6, we cannot apply the standard Lasry-Lions regularization directly since such a function ϕ_t^x is only locally semiconcave. Fortunately, recall a well known nonsmooth critical point theorem, see, for example, [28]. We only need the result in the following finite dimension setting.

Proposition 3.7. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function. Suppose that $x_1, x_2 \in \mathbb{R}^n$, $x_2 \notin \bar B(x_1, r)$ with $r > 0$ such that*

$$\max\{f(x_1), f(x_2)\} < b_0 < \inf_{\partial B(x_1,r)} f,$$

and define

$$b = \inf_{\gamma\in\Gamma} \max_{t\in[0,1]} f(\gamma(t)),$$

where

$$\Gamma = \{\gamma \in C([0, 1], \mathbb{R}^n) : \gamma(0) = x_1, \gamma(1) = x_2\}.$$

If f is coercive, then there exists x_3 such that $f(x_3) = b$ and $0 \in \partial f(x_3)$, where $\partial f(x_3)$ is the Clarke's generalized gradient of f at x_3.

The readers can refer to [21] for the definition and properties of Clarke's generalized gradients. Applying Proposition 3.7 to $f = \phi_t^x$ above, we obtain the following lemma.

Lemma 3.8. *Let u be a Lipschitz viscosity solution of* (3.13), $t > 0$, *and let $x \in \mathbb{R}^n$ be a singular point of u. Suppose there exist finitely many elements in $D^*u(x)$, say $D^*u(x) = \{p_1, \ldots, p_k\}$ with $k \geqslant 2$. Then there exist k distinct global minimizers $z_t^1, \ldots, z_t^k$ of ϕ_t^x.*

Moreover, if

$$b_{ij} = \inf_{\gamma_{ij} \in \Gamma_{ij}} \max_{s \in [0,1]} \phi_t^x(\gamma_{ij}(s)), \quad 1 \leqslant i, j \leqslant k, \quad i \neq j, \tag{3.14}$$

where

$$\Gamma_{ij} = \{\gamma \in C([0,1], \mathbb{R}^n) : \gamma(0) = z_t^i, \gamma(1) = z_t^j\},$$

then, for each pair of (i, j) with $i \neq j$, there exists a third critical point x_t^{ij} of ϕ_t^x such that $\phi_t^x(x_t^{ij}) = b_{ij} > \min_{y \in \mathbb{R}^n} \phi_t^x(y)$.

Proof. We suppose $k = 2$ and the proof in the general case is definitely similar. Suppose $\{p_1, p_2\} = D^*u(x)$, and $t > 0$. Then there exists two $(u, L, 0)$-calibrated C^1 curves η_t^1 and η_t^2, and two global minimizers z_t^1 and z_t^2 of ϕ_t^x such that

$$\eta_t^1(t) = \eta_t^2(t) = x, \quad \eta_t^1(0) = z_t^1, \quad \eta_t^2(0) = z_t^2$$

by Lemma 3.6. Since z_t^1 and z_t^2 are two isolated (global) minimizers of local Lipschitz function ϕ_t^x which is coercive by the superlinear growth condition on L, a third critical point is obtained in the context of mountain pass method as in Proposition 3.7. The rest part of the proof is a direct consequence of Proposition 3.7 and the fact that the Clarke's generalized gradient $\partial\phi_t^x(\cdot)$ coincides with the (proximal) superdifferential $D^+\phi_t^x(\cdot)$ (see [12]) since ϕ_t^x is locally semiconcave. □

Therefore, we claim that if $x \in \mathbb{R}^n$ is a singular point of viscosity solution u, there exists a third critical point $x_t = x_t^{12}$ of ϕ_t^x determined by two isolated global minimizers z_t^1 and z_t^2 as Lemma 3.8.

Theorem 3.9. *Let u be a Lipschitz viscosity solution of* (3.13), $t > 0$, *and let $x \in \mathbb{R}^n$ be a singular point of u. Suppose there exist finite many elements in $D^*u(x)$, say $D^*u(x) = \{p_1, \ldots, p_k\}$ with $k \geqslant 2$. Then there exist critical points $\{x_t^{ij}\}$ of ϕ_t^x (not global minimizers) such that, for $1 \leqslant i, j \leqslant k$, $i \neq j$, each critical point $x_t = x^{ij}$ has the following dichotomy*:

(a) *x_t is a differentiable point of ϕ_t^x and there exists a local minimal curve connecting x_t and x. More precisely, there exists a C^1 curve $\gamma : (-\infty, t] \to \mathbb{R}^n$ such that $\gamma(0) = x_t$, $\gamma(t) = x$ and the restriction of γ on $(-\infty, 0]$ is a $(u, L, 0)$-calibrated curve, but γ is not a $(u, L, 0)$-calibrated curve on $(-\infty, t]$*;

(b) *x_t is a singular point of u.*

Proof. The existence of such critical points $\{x_t^{ij}\}_{1 \leqslant i,j \leqslant k}$ of ϕ_t^x is a direct consequence of Lemma 3.8. The critical points $\{x_t^{ij}\}$ are not global minimizers of ϕ_t^x since each global minimizer z_t^i, $1 \leqslant i \leqslant k$, is isolated.

Suppose x_t is not a singular point of ϕ_t^x, thus both $u(\cdot)$ and $A_t(\cdot, x)$ is differentiable at x_t and

$$Du(x_t) + D_x A_t(x_t, x) = 0$$

since x_t is a critical point of ϕ_t^x. Let $p = Du(x_t)$ and $p' = -D_x A_t(x_t, x)$, then there exist two C^1 curves $\gamma_1 : (-\infty, 0] \to \mathbb{R}^n$ and $\gamma_2 : [0, t] \to \mathbb{R}^n$ such that

$$\gamma_1(0) = \gamma_2(0) = x_t, \quad \gamma_2(t) = x,$$

and

$$p = L_v(\gamma_1(0), \dot{\gamma}_1(0)) = -p' = L_v(\gamma_2(0), \dot{\gamma}_2(0)).$$

It follows that $\dot{\gamma}_1(0) = \dot{\gamma}_2(0)$ and γ, the juxtaposition of γ_1 and γ_2, is a C^1 curve which is an extremal. But, $\gamma : (-\infty, t] \to \mathbb{R}$ is not a $(u, L, 0)$-calibrated curve, otherwise, $\gamma(0) = x_t$ is a global minimizer of ϕ_t^x by Lemma 3.6. □

Remark 3.10. It is not clear, even when $t > 0$ is sufficient small, whether the critical point x_t^{ij} of ϕ_t^x found by Lemma 3.8 is close to the singular point x of u. It is not hard to prove that when the positive time t tends to 0, the global minimizers z_t^i and z_t^j tend to x along the direction determined by the associated limiting differentials in $D^*u(x)$, respectively. We hope to dig out more information from this approach in the future.

4 Concluding remarks

Recently, an intrinsic approach of the study of propagation of singularities along the generalized characteristics is obtained, in both global and local cases. One of the essential parts is the regularity properties of the fundamental solutions $A_t(x.y)$. Then, it is natural to relate to some important facts in the literature:

• In the spirit of Lasry-Lions regularization (see [5] or [24]), the standard kernel of such inf/sup-convolutions is defined by an integrable Lagrangian $L(x,v) = \frac{|v|^2}{2}$, and the associated fundamental solution has the implicit form $A_t(x,y) = \frac{|y-x|^2}{2t}$.

• The Lax-Oleinik operators introduced in the theory of viscosity solutions of Hamilton-Jacobi equations play an essential role in the study of both Hamilton-Jacobi equations (see, for example, [6–8, 12, 22, 23]) and Hamiltonian dynamical systems (see, for example, [14–19]).

In view of the regularity results of $A_t(x,y)$ in [10] for small $t > 0$, together with some rescaling technique, it is worth imposing the following open questions:

• Is there a precise formulation on the correspondence between the classic characteristics determined by the inf-convolution procedure (see Subsection 3.3) and the generalized characteristics determined by the sup-convolution procedure (see Subsection 3.1)? What is the ω-limit sets of the generalized characteristics?

• What is a more general setting for the problem of the long-time behavior of Lax-Oleinik operators, even if the Aubry sets are empty?

• What is the relation between the Aubry sets and the cut loci in topological sense?

• How can we find a local minimizer in the procedure of inf-convolution as in Subsection 3.3?

• What can we obtained in analogy to the property (P4) in Theorem 3.2 for the t-dependent systems?

We will try to answer these questions in the future.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 11271182 and 11471238) and the National Basic Research Program of China (Grant No. 2013CB834100). The authors are grateful to Liang Jin for helpful discussions on the results of this paper.

References

1 Albano P, Cannarsa P. Structural properties of singularities of semiconcave functions. Ann Scuola Norm Sup Pisa Cl Sci, 1999, 28: 719–740

2 Albano P, Cannarsa P. Propagation of singularities for solutions of nonlinear first order partial differential equations. Arch Ration Mech Anal, 2002, 162: 1–23

3 Albano P, Cannarsa P, Nguyen K T, et al. Singular gradient flow of the distance function and homotopy equivalence. Math Ann, 2013, 356: 23–43

4 Ambrosio L, Cannarsa P, Soner H M. On the propagation of singularities of semi-convex functions. Ann Scuola Norm Sup Pisa Cl Sci, 1993, 20: 597–616

5 Attouch H, Azé D. Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry-Lions method. Ann Inst H Poincaré Anal Non Linéaire, 1993, 10: 289–312

6 Bernard P. Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds. Ann Sci École Norm Sup, 2007, 40: 445–452

7 Bernard P. The dynamics of pseudographs in convex Hamiltonian systems. J Amer Math Soc, 2008, 21: 615–669

8 Bernard P. The Lax-Oleinik semi-group: A Hamiltonian point of view. Proc Roy Soc Edinburgh Sect A, 2012, 142: 1131–1177

9 Cannarsa P, Cheng W. Homoclinic orbits and critical points of barrier functions. Nonlinearity, 2015, 28: 1823–1840

10 Cannarsa P, Cheng W. Generalized characteristics and Lax-Oleinik operators: Global theory. ArXiv:1605.07581, 2016
11 Cannarsa P, Cheng W, Zhang Q. Propagation of singularities for weak KAM solutions and barrier functions. Comm Math Phys, 2014, 331: 1–20
12 Cannarsa P, Sinestrari C. Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control. In: Progress in Nonlinear Differential Equations and Their Applications, vol. 58. Boston: Birkhäuser Boston, 2004
13 Cannarsa P, Yu Y. Singular dynamics for semiconcave functions. J Eur Math Soc, 2009, 11: 999–1024
14 Cheng C Q. Arnold diffusion in nearly integrable Hamiltonian systems. ArXiv:1207.4016v2, 2013
15 Cheng C Q. Uniform hyperbolicity of invariant cylinder. ArXiv:1509.03160, 2015
16 Cheng C Q. A way to cross double resonance. ArXiv:1510.08754, 2015
17 Cheng C Q, Xue J. Arnold diffusion in nearly integrable Hamiltonian systems of arbitrary degrees of freedom. ArXiv:1503.04153, 2015
18 Cheng C Q, Yan J. Existence of diffusion orbits in a priori unstable Hamiltonian systems. J Differential Geom, 2004, 67: 457–517
19 Cheng C Q, Yan J. Arnold diffusion in Hamiltonian systems: A priori unstable case. J Differential Geom, 2009, 82: 229–277
20 Cheng C Q, Zhou M. Non-degeneracy of extremal points in multi-dimensional space. Sci China Math, 2015, 58: 2255–2260
21 Clarke F H. Optimization and Nonsmooth Analysis, 2nd ed. Philadelphia: SIAM, 1990
22 Fathi A. Weak KAM Theorem in Lagragian Dynamics. Cambridge: Cambridge University Press, 2005
23 Fathi A, Siconolfi A. Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation. Invent Math, 2004, 155: 363–388
24 Lasry J M, Lions P L. A remark on regularization in Hilbert spaces. Israel J Math, 1986, 55: 257–266
25 Mather J N. Variational construction of connecting orbits. Ann Inst Fourier (Grenoble), 1993, 43: 1349–1386
26 McEneaney W M, Dower P M. The principle of least action and fundamental solutions of mass-spring and N-body two-point boundary value problems. SIAM J Control Optim, 2015, 53: 2898–2933
27 Rifford L. On viscosity solutions of certain Hamilton-Jacobi equations: Regularity results and generalized Sard's theorems. Comm Partial Differential Equations, 2008, 33: 517–559
28 Shi S. Ekeland's variational principle and the mountain pass lemma. Acta Math Sin Engl Ser, 1985, 1: 348–355
29 Yu Y. A simple proof of the propagation of singularities for solutions of Hamilton-Jacobi equations. Ann Scuola Norm Sup Pisa Cl Sci, 2006, 5: 439–444

Appendix: Regularity properties of fundamental solutions

For the details of the proofs of the results in this appendix, the readers can refer to [10] or [8] under certain special conditions.

Proposition A.1. *Let $0 < t \leqslant 1$, $R > 0$ and suppose L satisfies* (L1) *and* (L2). *Let $\xi \in \Gamma^t_{x,y}$ be a minimizer for $A_t(x,y)$, $x \in \mathbb{R}^n$, $y \in \bar{B}(x,R)$, and let $p(s)$ be the dual arc of $\xi(s)$. Then we have*

$$\sup_{s\in[0,t]} |\dot{\xi}(s)| \leqslant \Delta(x, R/t), \qquad \sup_{s\in[0,t]} |p(s)| \leqslant \Delta(x, R/t), \qquad \sup_{s\in[0,t]} |\xi(s)| \leqslant \Delta(x, R/t),$$

where $\Delta(x,\cdot)$ is non-decreasing.

Proof. For any $t > 0$, $R > 0$, let $x \in \mathbb{R}^n$, $y \in \bar{B}(x,R)$ and $\xi \in \Gamma^t_{x,y}$ be a minimizer for $A_t(x,y)$, i.e.,

$$A_t(x,y) = \int_0^t L(\xi(s), \dot{\xi}(s))ds.$$

Denoting by $\sigma \in \Gamma^t_{x,y}$ the straight line segment defined by

$$\sigma(s) = x + \frac{s}{t}(y-x), \quad s \in [0,t],$$

then by the Nagumo conditions in (L2) we have

$$\int_0^t \theta(|\dot{\xi}(s)|)ds - c_0 t \leqslant \int_0^t L(\xi(s), \dot{\xi}(s))ds \leqslant \int_0^t L(\sigma(s), \dot{\sigma}(s))ds$$

$$
\begin{aligned}
&= \int_0^t L\left(x+\frac{s}{t}(y-x), \frac{y-x}{t}\right) - L\left(\frac{s}{t}(y-x), \frac{y-x}{t}\right) + L\left(\frac{s}{t}(y-x), \frac{y-x}{t}\right) ds \\
&\leqslant c_1 t|x|\theta\left(\left|\frac{y-x}{t}\right|\right) + t \max_{y\in\bar{B}(x,R), s\in[0,t]} \left|L\left(\frac{s}{t}(y-x), \frac{y-x}{t}\right)\right| \\
&\leqslant c_1 t|x|\theta\left(\left|\frac{y-x}{t}\right|\right) + tM(t,R) \\
&\leqslant c_1 t|x|\theta(R/t) + t \max_{|x|,|v|\leqslant R/t} |L(x,v)| =: C_1(t,R).
\end{aligned}
$$

By Condition (L2), we have

$$
\begin{aligned}
|L(x,v) - L(0,0)| &\leqslant |L(x,v) - L(x,0)| + |L(x,0) - L(0,0)| \\
&\leqslant c_1\theta(|v|)|v| + c_2(x).
\end{aligned}
$$

Thus,

$$
C_1(t,R) \leqslant c_3(x,R)t\kappa_1(R/t) \tag{A.1}
$$

with

$$
\kappa_1(s) = \theta(s)(1+s) + 1.
$$

By the superlinear growth condition of θ, we have that

$$
\int_0^t |\dot{\xi}(s)| ds \leqslant c_4(x,R)t\kappa_2(R/t),
$$

where $\kappa_2(s) = 1 + \kappa_1(s)$. Hence

$$
|\xi(s) - x| \leqslant \int_0^s |\dot{\xi}(s)| ds \leqslant c_4(x,R)t\kappa_2(R/t), \quad s \in [0,t], \tag{A.2}
$$

and

$$
\inf_{s\in[0,t]} |\dot{\xi}(s)| \leqslant \frac{1}{t}\int_0^t |\dot{\xi}(s)| ds \leqslant c_4(x,R)\kappa_2(R/t). \tag{A.3}
$$

Now, we turn to estimating $\sup_{s\in[0,t]} |\dot{\xi}(s)|$. By Condition (L2) and the convexity of L, we have

$$
\begin{aligned}
\theta(|\dot{\xi}(s)|) - c_0 \leqslant L(\xi(s), \dot{\xi}(s)) &\leqslant L(\xi(s), 0) + \langle L_v(\xi(s), \dot{\xi}(s)), \dot{\xi}(s)\rangle \\
&= L(\xi(s), 0) + \left\langle \int_0^s L_x(\xi(\tau), \dot{\xi}(\tau)) d\tau + L_v(\xi(0), \dot{\xi}(0)), \dot{\xi}(s)\right\rangle.
\end{aligned} \tag{A.4}
$$

Note that we use the Euler-Lagrange equation in the last equality. By (L2) and the estimates above, we have

$$
\int_0^s |L_x(\xi(\tau), \dot{\xi}(\tau))| d\tau \leqslant \int_0^s c_1\theta(|\dot{\xi}(\tau)|) d\tau \leqslant c_1c_3(x,R)t\kappa_1(R/t) + c_1c_0t.
$$

For any $s \in [0,t]$, we also have

$$
|L_v(\xi(0), \dot{\xi}(0))| \leqslant |L_v(\xi(s), \dot{\xi}(s))| + \int_0^s |L_x(\xi(\tau), \dot{\xi}(\tau))| d\tau.
$$

Then, by (L2) and (A.3), it follows

$$
|L_v(\xi(0), \dot{\xi}(0))| \leqslant c_1\theta(|\dot{\xi}(s)|) + c_1c_3(x,R)t\kappa_1(R/t) + c_1c_0t,
$$

and this implies

$$
|L_v(\xi(0), \dot{\xi}(0))| \leqslant c_1\theta(\inf_{s\in[0,t]} |\dot{\xi}(s)|) + c_1c_3(x,R)t\kappa_1(R/t) + c_1c_0t
$$

$$\leqslant c_1\theta(c_4(x,R)\kappa_2(R/t)) + c_1c_3(x,R)t\kappa_1(R/t) + c_1c_0t.$$

It follows there exists $M>0$ sufficient large and $\mu>0$ such that

$$\begin{aligned}M|\dot\xi(s)| &\leqslant [c_1\theta(c_4(x,R)\kappa_2(R/t)) + c_0t] + 2[c_1c_3(x,R)t\kappa_1(R/t) + c_1c_0t]\cdot|\dot\xi(s)| + \mu\\ &\leqslant c_5(x,R/t)\cdot|\dot\xi(s)| + \mu,\end{aligned}$$

since $c_3(x,\cdot)$ is non-decreasing and $0<t\leqslant 1$. This implies

$$|\dot\xi(s)| \leqslant \frac{\mu}{M - c_5(x,R/t)} =: c_6(x,R/t).$$

So, if $t\leqslant 1$, we have

$$\sup_{s\in[0,t]}|\dot\xi(s)| \leqslant c_6(x,R/t).$$

As for the dual arc $p(\cdot)$, by (L2), we have

$$\sup_{s\in[0,t]}|p(s)| = \sup_{s\in[0,t]}|L_v(\xi(s),\dot\xi(s))| \leqslant c_7(x,R/t).$$

We complete the proof by defining $\Delta(x,R/t)=\max\{c_6(x,R/t),c_7(x,R/t)\}$. □

Fix $x\in\mathbb{R}^n$ and suppose $R>0$ and L satisfies (L1)–(L2). In this case, the following observation is one of the key points of the results on the local regularity properties of $A_t(x,y)$. For any $t>0$ and $y\in\bar B(x,R)$, let $\xi_{t,y}\in\Gamma^t_{x,y}$ be a minimizer for $A_t(x,y)$, and $p_{t,y}$ be its dual arc. Then we have

$$\sup_{s\in[0,t]}|\dot\xi_{t,y}(s)| \leqslant \Delta(x,R/t), \quad \sup_{s\in[0,t]}|p_{t,y}(s)| \leqslant \Delta(x,R/t)$$

by Proposition A.1. Now, define

$$\begin{aligned}\boldsymbol{K}_x &:= \bar B(x,\Delta(x,1))\times\bar B(0,\Delta(x,1))\subset\mathbb{R}^n\times\mathbb{R}^n,\\ \boldsymbol{K}^*_x &:= \bar B(x,\Delta(x,1))\times\bar B(0,\Delta(x,1))\subset\mathbb{R}^n\times(\mathbb{R}^n)^*.\end{aligned} \tag{A.5}$$

Then, by defining a function $R(x,\cdot):\mathbb{R}^n\times(0,1]\to(0,\infty)$, $R(x,t)=\frac{t}{2}$, we have

$$\Delta(x,1/2)\leqslant\Delta(x,1). \tag{A.6}$$

Because of the monotonicity properties of $\Delta(x,\cdot)$. So, if $y\in\bar B(x,R(x,t))$, and $\xi_t\in\Gamma_{x,y}$ is a minimizer in the definition of $A_t(x,y)$, then

$$\{\xi(s),p(s)\}_{s\in[0,t],t\in(0,1]}\subset\boldsymbol{K}^*_x, \quad \{\xi(s),\dot\xi(s)\}_{s\in[0,t],t\in(0,1]}\subset\boldsymbol{K}_x.$$

Proposition A.2 (See [10]). *Fix any $x\in\mathbb{R}^n$ and $t>0$ with $R(x,t)$ defined as in* (A.6). *If y_t is the unique maximizer of ψ^x_t in $\bar B(x,R(x,t))$ for all $t\in(0,t_0]$, and $\xi_t\in\Gamma^t_{x,y_t}$ is a minimal curve in the definition of $A_t(x,y_t)$, $t\in(0,t_0]$, then the family $\{\dot\xi_t\}$ is equi-Lipschitz.*

The proof of the following result is similar to those in [10] since the estimates involving certain first and second order partial derivatives of L and H which are bounded on the *a priori* compact sets $\boldsymbol{K}^*_x$ or $\boldsymbol{K}_x$. The difference between the cases here and what in [10] is that the bound for the minimal curves and the dual arc is independent of x in the latter.

Proposition A.3. *Suppose L is a Tonelli Lagrangian satisfying* (L1)–(L2). *Fix any $x\in\mathbb{R}^n$. Then there exists $t_0>0$, such that for $0<t\leqslant t_0$, $(t,y)\mapsto A_t(x,y)$ is locally convex in*

$$S(x,t_0)=\{(t,y)\in\mathbb{R}\times\mathbb{R}^n: 0<t\leqslant t_0, |y-x|\leqslant R(x,t)\}$$

with $R(x,t)$ defined in (A.6).

More precisely, there exist constants $C_1,C_2>0$ such that, if $y\in B(x,R(x,t))$, then, for $|h|\ll 1$ and $|z|\ll 1$, we have

$$A_{t+h}(x,y+z)+A_{t-h}(x,y-z)-2A_t(x,y)\geqslant\frac{C_1}{t^3}|h|^2+\frac{C_2}{t}|z|^2. \tag{A.7}$$

Proposition A.4. *Suppose L is a Tonelli Lagrangian satisfying* (L1)–(L2). *For any $x \in \mathbb{R}^n$, there exists $t_0 > 0$, such that the functions $w : (t, y) \mapsto A_t(x, y)$ and $(t, y) \mapsto A_t(y, x)$ are both of class $C^{1,1}_{\rm loc}$ in*

$$S(x, t_0) = \{(t, y) \in \mathbb{R} \times \mathbb{R}^n : 0 < t \leqslant t_0, |y - x| \leqslant R(x, t)\},$$

with $R(x, t)$ defined in (A.6), *for $0 < t \leqslant t_0$. In Particular, for any $t \in (0, t_0]$,*

$$D_y A_t(x, y) = L_v(\xi(t), \dot{\xi}(t)), \tag{A.8}$$

$$D_x A_t(x, y) = - L_v(\xi(0), \dot{\xi}(0)), \tag{A.9}$$

$$D_t A_t(x, y) = - E_{t,x,y}, \tag{A.10}$$

where $\xi \in \Gamma^t_{x,y}$ is the unique minimizer for $A_t(x, y)$ and $E_{t,x,y}$ is the energy of the Hamiltonian trajectory $(\xi(s), p(s))$ with $p(s) = L_v(\xi(s), \dot{\xi}(s))$.