GENERALIZED CHARACTERISTICS AND LAX-OLEINIK OPERATORS: GLOBAL THEORY

PIERMARCO CANNARSA AND WEI CHENG

ABSTRACT. For autonomous Tonelli systems on $\mathbb{R}^n$, we develop an intrinsic proof of the existence of generalized characteristics using sup-convolutions. This approach, together with convexity estimates for the fundamental solution, leads to new results such as the global propagation of singularities along generalized characteristics.

1. INTRODUCTION

Let $L(x,v)$ be a Tonelli Lagrangian on $\mathbb{R}^n$ ($L:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ is a function of class C^2, strictly convex in the fibre, with superlinear growth with respect to v), and let $H(x,p)$ be the associated Hamiltonian given by the Fenchel-Legendre transform. The study of the regularity properties of the viscosity solutions of the Hamilton-Jacobi equation

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

is extremely important for several reasons. In the last two or three decades, remarkable progress in the broad area of Hamiltonian dynamical systems was achieved by Mather's theory for Tonelli systems on compact manifolds in Lagrangian formalism [31, 32], and Fathi's weak KAM theory in Hamiltonian formalism [25, 27]. Both these theories succeeded in the analysis of some hard dynamical problems such as Arnold diffusion. However, a general variational setting that applies to all the above aspects of Hamiltonian dynamics still has to be developed. As is well known, in both Mather's and Fathi's theories various global minimal sets in a variational sense—such as Mather's set, Aubry's set and Mañé's set—play a crucial role. Similarly, if we want to study the behavior of an orbit after it loses minimality, then we have to face the hard problem of dealing with the cut loci and singular sets of the associated viscosity solutions.

Although significant contributions investigating the singularities of viscosity solutions were already given in [17] and [7], the current approach to this problem goes back to [3], where the propagation of singularities was studied for general semiconcave functions. Since any viscosity solution u of (1.1) is locally semiconcave (with linear modulus), we have that singularities propagate along Lipschitz arcs starting from any singular point x of u at which the superdifferential $D^+u(x)$ satisfies the condition

$$\partial D^+u(x)\setminus D^*u(x)\neq\varnothing,$$

where $\partial D^+u(x)$ denotes the topological boundary of $D^+u(x)$ and $D^*u(x)$ the set of all reachable gradients of u at x. A more specific approach to the problem was developed in

Date: January 3, 2017.

2010 *Mathematics Subject Classification.* 35F21, 49L25, 37J50.

Key words and phrases. Hamilton-Jacobi equation, weak KAM theory, viscosity solution, generalized characteristic, singularities.

[4] by solving the generalized characteristic inclusion

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

More precisely, if the initial point x_0 belongs to the singular set of u, hereafter denoted by Sing (u), and is not a critical point of u relative to H, i.e.,

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

then it was proved in [4] that there exists a nonconstant singular arc $\mathbf{x}$ from x_0 which is a generalized characteristic. The study of the local propagation of singularities along generalized characteristics was later refined in [37] and [18]. For weak KAM solutions, local propagation results were obtained in [19] and the Lasry-Lions regularization procedure was applied in [12] to analyze the critical points of Mather's barrier functions. An interesting interpretation of the above singular curves as part of the flow of fluid particles has been recently proposed in [28] (see also [35] for related results).

Returning to our dynamical motivations, in this paper we try to give an intrinsic interpretation of generalized characteristics and study the relevant global properties of such curves. For this purpose, we use the Lax-Oleinik semigroups $T_t^{\pm}$ (see, e.g. [25]) defined as follows:

$$T_t^+u_0(x) := \sup_{y\in\mathbb{R}^n}\{u_0(y) - A_t(x,y)\},$$
$$T_t^-u_0(x) := \inf_{y\in\mathbb{R}^n}\{u_0(y) + A_t(y,x)\},$$

where $u_0 : \mathbb{R}^n \to \mathbb{R}$ is a continuous function and $A_t(x,y)$ is the fundamental solution of (1.1). These operators can be also derived from the Moreau-Yosida approximations in convex analysis ([8]) or the Lasry-Lions regularization technique based on sup- and inf-convolutions ([29], [36]).

By analyzing the maximizers y_t, for sufficiently small $t > 0$, in the sup-convolution giving $T_t^+u_0(x)$ we obtain the global propagation of singularities which represents the main result of this paper. For such a result we need the following assumptions.

(L1) *Uniform convexity*: There exists a nonincreasing function $\nu : [0,+\infty) \to (0,+\infty)$ such that

$$L_{vv}(x,v) \geqslant \nu(|v|)I$$

for all $(x,v) \in \mathbb{R}^n \times \mathbb{R}^n$.

(L2) *Growth conditions*: There exist two superlinear functions $\theta, \overline{\theta} : [0,+\infty) \to [0,+\infty)$ and a constant $c_0 > 0$ such that

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

(L3) *Uniform regularity*: There exists a nondecreasing function $K : [0,+\infty) \to [0,+\infty)$ such that, for every multindex $|\alpha| = 1, 2$,

$$|D^\alpha L(x,v)| \leqslant K(|v|) \qquad \forall (x,v) \in \mathbb{R}^n \times \mathbb{R}^n,$$

Recalling that Sing (u) stands for the singular set of u we now proceed to state our

Propagation result: Let L be a Lagrangian on $\mathbb{R}^n$ satisfying conditions (L1)-(L3), and let u be a globally Lipschitz semiconcave solution of the Hamilton-Jacobi equation (1.1), where H is the Hamiltonian associated with L. If x belongs to Sing (u), then there exists a generalized characteristic $\mathbf{x} : [0,+\infty) \to \mathbb{R}^n$ such that $\mathbf{x}(0) = x$ and $\mathbf{x}(s) \in$ Sing (u) for all $s \in [0,+\infty)$.

We observe that the study of the global propagation of singularities is much more difficult than the local one. Indeed, to this date the only known results concern geodesic systems, see [2] and [5]. More precisely, [2] studies the global propagation of the so-called C^1-singular support of solutions which—in this case—coincides with the closure of $\mathrm{Sing}(u)$, while [5] investigates the propagation of genuine singularities. For time dependent problems, global propagation was addressed in [1] for the C^1-singular support of solutions and, recently, in [15] for singularities of solutions to eikonal equations.

It is worth mentioning that, when considering geodesic systems on Riemannian manifolds, the method of generalized characteristics (or generalized gradient flow) has been successfully applied to reveal topological relations between a compact domain Ω and the cut locus enclosed in Ω ([5]). Such relations depend on global results for the propagation of singularities along the associated generalized characteristics. We will shortly apply our global propagation results to study singularities on the torus ([13]).

For the proof of the above theorem we need regularity results for the value function of the action functional (also called fundamental solution of (1.1) in [33])

$$A_t(x,y)=\inf_{\gamma\in\Gamma^t_{x,y}}\int_0^t L(\gamma(s),\dot\gamma(s))ds\quad (t>0\,,\ x,y\in\mathbb{R}^n)$$

where

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

More precisely, we need

Convexity and local $C^{1,1}$ regularity results: Suppose L is a Tonelli Lagrangian satisfying (L1)-(L3). Then the following properties hold true.

(a) For any $\lambda>0$, there exists $t_\lambda>0$ such that, for any $x\in\mathbb{R}^n$, the function $(t,y)\mapsto A_t(x,y)$ is semiconvex on the cone

$$S_\lambda(x,t_\lambda):=\left\{(t,y)\in\mathbb{R}\times\mathbb{R}^n\ :\ 0<t<t_\lambda,\ |y-x|<\lambda t\right\},$$

that is, there exists $C_\lambda>0$ such that for all $x\in\mathbb{R}^n$, all $(t,y)\in S_\lambda(x,t_\lambda)$, all $h\in[0,t/2)$, and all $z\in B(0,\lambda t)$ we have that

$$A_{t+h}(x,y+z)+A_{t-h}(x,y-z)-2A_t(x,y)\geqslant -\frac{C_\lambda}{t}(h^2+|z|^2).$$

(b) For all $t\in(0,t_\lambda]$, $A_t(x,\cdot)$ is uniformly convex on $B(x,\lambda t)$, that is, there exists $C'_\lambda>0$ such that for all $x\in\mathbb{R}^n$, all $y\in B(x,\lambda t)$, and all $z\in B(0,\lambda t)$ we have that

$$A_t(x,y+z)+A_t(x,y-z)-2A_t(x,y)\geqslant\frac{C'_\lambda}{t}|z|^2.$$

(c) For any $x\in\mathbb{R}^n$ the functions $(t,y)\mapsto A_t(x,y)$ and $(t,y)\mapsto A_t(y,x)$ are of class $C^{1,1}_{\mathrm{loc}}$ on the cone $S_\lambda(x,t_\lambda)$ defined above.

Similar regularity results were obtained in [11] by a different approach, under more restrictive structural assumptions than those we consider in this paper.

This paper is organized as follows. In section 2, we review basic properties of viscosity solution of Hamilton-Jacobi equations. In section 3, we discuss connections between sup-convolutions and generalized characteristics and we give our global result on the propagation of singularities along generalized characteristics. The paper contains four appendices that contain technical results and background material which is useful for our approach: in the first one we give a uniform bound for minimizers of the action functional following [6] and [24], in the second one we give detailed proofs of all the required regularity results for the fundamental solution, in the third one we adapt the construction of generalized characteristics from [4] to the present context, in the fourth one we provide a global semiconcavity estimate for the weak KAM solution on $\mathbb{R}^n$ constructed in [26].

Acknowledgments This work was partially supported by the Natural Scientific Foundation of China (Grant No. 11271182 and No. 11471238), the National Basic Research Program of China (Grant No. 2013CB834100), a program PAPD of Jiangsu Province, China, and the National Group for Mathematical Analysis, Probability and Applications (GNAMPA) of the Italian Istituto Nazionale di Alta Matematica "Francesco Severi". The authors thank Albert Fathi for the discussion during his visit in Nanjing. The second author is grateful to Cui Chen, Liang Jin for helpful discussions, and also Jun Yan for the communications and the hospitality during the visit to Fudan University in 2014-15.

2. Preliminaries

2.1. **Semiconcave functions.** Let $\Omega \subset \mathbb{R}^n$ be a convex set. We recall that a function $u : \Omega \to \mathbb{R}$ is said to be *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* if $-u$ is semiconcave.

When $u : \Omega \to \mathbb{R}$ is continuous, it can be proved that u is semiconcave with constant C if and only if

$$u(x) + u(y) - 2u\left(\frac{x+y}{2}\right) \leqslant \frac{C}{2}|x-y|^2$$

for any $x, y \in \Omega$.

Hereafter, we assume that Ω is a nonempty open subset of $\mathbb{R}^n$.

We recall that 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 $B(x, r)$.

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.

Let now $u : \Omega \to \mathbb{R}$ be locally Lipschitz. We recall that a vector $p \in \mathbb{R}^n$ is said to be a *reachable* (or *limiting*) *gradient* of u at x if there exists a sequence $\{x_n\} \subset \Omega \setminus \{x\}$, converging to x, such that u is differentiable at x_k for each $k \in \mathbb{N}$ and

$$\lim_{k\to\infty} Du(x_k) = p.$$

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

Now we list some well known properties of the superdifferential of a semiconcave function on $\Omega \subset \mathbb{R}^n$ (see, e.g., [16] for the proof).

Proposition 2.1. *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) *$D^+u(x) = \operatorname{co} D^*u(x)$.*

Proposition 2.2 ([16]). *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 $p \in D^+u(x)$.*

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 Sing (u).

2.2. **Tonelli Lagrangians.** In this paper, we concentrate on Lagrangians on Euclidean configuration space $\mathbb{R}^n$. We say that a function $\theta : [0, +\infty) \to [0, +\infty)$ is *superlinear* if $\theta(r)/r \to +\infty$ as $r \to +\infty$.

Definition 2.3. A function $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called a *generalized Tonelli function* if F is a function of class C^2 that satisfies the following conditions:

(T1) *Uniform convexity*: There exists a nonincreasing function $\nu : [0, +\infty) \to (0, +\infty)$ such that
$$F_{vv}(x, v) \geqslant \nu(|v|)I \qquad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$
(T2) *Growth condition*: There exist two superlinear function $\theta, \overline{\theta} : [0, +\infty) \to [0, +\infty)$ and a constant $c_0 > 0$ such that
$$\overline{\theta}(|v|) \geqslant F(x, v) \geqslant \theta(|v|) - c_0 \qquad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$
(T3) *Uniform regularity*: There exists a nondecreasing function $K : [0, +\infty) \to [0, +\infty)$ such that, for every multindex $|\alpha| = 1, 2$,
$$|D^\alpha F(x, v)| \leqslant K(|v|) \qquad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^n,$$

The convex conjugate of a superlinear function θ is defined as

$$\theta^*(s) = \sup_{r \geqslant 0}\{rs - \theta(r)\}s \qquad \forall s \geqslant 0. \tag{2.3}$$

In view of the superlinear growth of θ it is clear that θ^* is well defined and satisfies

$$\theta(r) + \theta^*(s) \geqslant rs \qquad \forall r, s \geqslant 0, \tag{2.4}$$

which in turn can be used to show that $\theta^*(s)/s \to \infty$ as $s \to \infty$.

Definition 2.4. A function $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called a *Tonelli Lagrangian* if L is a generalized Tonelli function as in Definition 2.3. If L is a Tonelli Lagrangian, the associated Hamiltonian H is the Fenchel-Legendre dual of L is defined by

$$H(x, p) = \sup_{v \in \mathbb{R}^n} \big\{\langle p, v\rangle - L(x, v)\big\} \qquad (x, p) \in \mathbb{R}^n \times \mathbb{R}^n\,. \tag{2.5}$$

The Hamiltonian H is called a *Tonelli Hamiltonian* if H is a generalized Tonelli function. We denote by (L1)-(L3) (resp. (H1)-(H3)) the corresponding conditions of a Tonelli Lagrangian L (resp. Tonelli Hamiltonian H).

Example 2.5. A typical example of a Tonelli Lagrangian L is the one of mechanical systems which has the form

$$L(x,v)=f(x)(1+|v|^2)^{\frac{q}{2}}+V(x),\quad (x,v\in\mathbb{R}^n),$$

where $q>1$, f and V are smooth functions on $\mathbb{R}^n$ with bounded derivatives up to the second order, and $\inf_{\mathbb{R}^n} f>0$.

Lemma 2.6. *If L is a Tonelli Lagrangian with H its Fenchel-Legendre dual, then H is a Tonelli Hamiltonian.*

Proof. First, let $v_{x,p}\in\mathbb{R}^n$ be such that $p=L_v(x,v_{x,p})$ or, equivalently, $v_{x,p}=H_p(x,p)$. By assumptions (L1)-(L3), we have that

$$\begin{aligned}|p||v_{x,p}|\geqslant&\langle L_v(x,v_{x,p}),v_{x,p}\rangle\geqslant L(x,v_{x,p})-L(x,0)\geqslant\theta(|v_{x,p}|)-c_0-\bar\theta(0)\\ \geqslant&(|p|+1)|v_{x,p}|-c_0-\bar\theta(0)-\theta^*(|p|+1).\end{aligned}$$

This shows that

$$|H_p(x,p)|=|v_{x,p}|\leqslant c_0+\bar\theta(0)+\theta^*(|p|+1)=C_1(|p|),$$

where $C_1(r)=c_0+\bar\theta(0)+\theta^*(r+1)$. The estimates for the other derivatives in (H3) and (H1) can be proved by using the relations

$$\begin{aligned}&H_x(x,p)=-L_x(x,H_p(x,p))\\&H_{pp}(x,p)=L_{vv}^{-1}(x,H_p(x,p))\\&H_{xp}(x,p)=-L_{xv}(x,H_p(x,p))H_{pp}(x,p)\\&H_{xx}(x,p)=-L_{xx}(x,H_p(x,p))-L_{xv}(x,H_p(x,p))H_{px}(x,p)\end{aligned}$$

and our conditions (L1)-(L3). This completes the verification of (H1) and (H3).

To check (H2), we have that, for all $R\geqslant 0$,

$$H(x,p)\geqslant R|p|-\bar\theta(R),\quad\forall(x,p)\in\mathbb{R}^n\times\mathbb{R}^n$$

by (2.5) and (L2). Set $\theta_1(r)=\sup_{R>0}\{Rr-\bar\theta(R)\}$, $r\in(0,+\infty)$, which is well defined by (L2). Thus, θ_1 gives the required superlinear function for (H2). □

2.3. **Hamilton-Jacobi equations.** Suppose H is the Hamiltonian associated with a Tonelli Lagrangian L and consider the Hamilton-Jacobi equation

$$H(x,Du(x))=0\qquad(x\in\mathbb{R}^n).\tag{2.6}$$

We recall that a continuous function u is called a *viscosity subsolution* of equation (2.6) if, for any $x\in\mathbb{R}^n$,

$$H(x,p)\leqslant 0,\quad\forall p\in D^+u(x)\,.\tag{2.7}$$

Similarly, u is a *viscosity supersolution* of equation (2.6) if, for any $x\in\mathbb{R}^n$,

$$H(x,p)\geqslant 0,\quad\forall p\in D^-u(x)\,.\tag{2.8}$$

Finally, u is called a *viscosity solution* of equation (2.6), if it is both a viscosity subsolution and a supersolution.

Throughout this paper we will be concerned with solutions of the above equation that are Lipschitz continuous and semiconcave on $\mathbb{R}^n$. The existence of such solution is the object of the following proposition which is essentially a consequence of the existence theorem of [26] and the semiconcavity results of this paper (see also [16] and [34]).

Proposition 2.7. *Let L be a Tonelli Lagrangian and let H be the associated Hamiltonian. Then there exists a constant $c(H) \in \mathbb{R}$ such that the Hamilton-Jacobi equation*

$$H(x, Du(x)) = c, \quad x \in \mathbb{R}^n, \tag{2.9}$$

admits a viscosity solution $u : \mathbb{R}^n \to \mathbb{R}$ for $c = c(H)$ and does not admit any such solution for $c < c(H)$. Moreover, u is globally Lipschitz continuous and semiconcave on $\mathbb{R}^n$.

The proof of Proposition 2.7 is given in Appendix D.

Let us consider the class of *admissible arcs*

$$\mathcal{A}_{t,x} = \{\xi \in W^{1,1}([0,t]; \mathbb{R}^n) : \xi(t) = x\}$$

where $W^{1,1}([a,b]; \mathbb{R}^n)$ denotes the space of all absolutely continuous $\mathbb{R}^n$-valued functions on $[a,b]$, where $-\infty < a < b < +\infty$. The functional

$$J_t(\xi) := \int_0^t L(\xi(s), \dot{\xi}(s))\, ds + u_0(\xi(0)), \quad \xi \in \mathcal{A}_{t,x}, \tag{2.10}$$

where $u_0 \in C(\mathbb{R}^n)$ is the *initial cost*, is usually called the *action functional*. A classical problem in the calculus of variations is

$$\text{to minimize } J_t \text{ over all arcs } \xi \in \mathcal{A}_{t,x}. \tag{CV$_{t,x}$}$$

We define the associated value function

$$u(t,x) = \min_{\xi \in \mathcal{A}_{t,x}} J_t(\xi). \tag{2.11}$$

It is known that $u(t,x)$ is a viscosity solution of the Cauchy problem

$$\begin{cases} u_t(t,x) + H(x, \nabla_x u(t,x)) = 0, & t > 0, x \in \mathbb{R}^n \\ u(0,x) = u_0(x), & x \in \mathbb{R}^n, \end{cases} \tag{2.12}$$

where $\nabla_x u$ denotes the spatial gradient of u. From the uniqueness of viscosity solutions of (2.12) if follows that, if the initial datum u_0 is a viscosity solution of (2.6), then the solution u of (2.12) is constant in time and coincides with u_0. In this case, because of the translation invariance of problem (CV$_{t,x}$), we have that, for all $t \geqslant 0$,

$$u_0(x) = \min_{\xi \in W^{1,1}([-t,0];\mathbb{R}^n)} \Big\{ \int_{-t}^0 L(\xi(s), \dot{\xi}(s))\, ds + u_0(\xi(-t)) \ : \ \xi(0) = x \Big\}. \tag{2.13}$$

Moreover, suppose L satisfies conditions (L1)-(L3) and let H be the associated Hamiltonian. Then we have the following result (see [16] or [34]).

Proposition 2.8. *Let $u : \mathbb{R}^n \to \mathbb{R}$ be a viscosity solution of* (2.6) *and let $x \in \mathbb{R}^n$. 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 minimizer of the problem in* (2.13) *for every $t \geqslant 0$ and $p = L_v(x, \dot{\gamma}(0))$.*

2.4. **Generalized characteristics.** The study of the structure of the singular set of a viscosity solution is a very important and hard one in many fields such as Riemannian geometry, optimal control, classical mechanics, etc. The dynamics of singularities can be described by using generalized characteristics.

Definition 2.9. A Lipschitz arc $\mathbf{x} : [0,T] \to \mathbb{R}^n$, $(T > 0)$, is said to be a *generalized characteristic* of the Hamilton-Jacobi equation (2.6) if $\mathbf{x}$ satisfies the differential inclusion

$$\dot{\mathbf{x}}(s) \in \mathrm{co}\, H_p\big(\mathbf{x}(s), D^+u(\mathbf{x}(s))\big), \quad \text{a.e. } s \in [0,T]\,. \tag{2.14}$$

A basic criterion for the propagation of singularities along generalized characteristics was given in [4] (see [18, 37] for an improved version and simplified proof of this result).

Proposition 2.10 ([4])**.** *Let u be a viscosity solution of* (2.6) *and let $x_0 \in \mathbb{R}^n$. Then there exists a generalized characteristic $\mathbf{x} : [0,T] \to \mathbb{R}^n$ with initial point $\mathbf{x}(0) = x_0$. Moreover, if $x_0 \in \mathrm{Sing}\,(u)$, then $\tau \in (0,T)$ exists such that $\mathbf{x}(s) \in \mathrm{Sing}\,(u)$ for all $s \in [0,\tau]$. Furthermore, if*

$$0 \notin \mathrm{co}\, H_p(x_0, D^+u(x_0))\,, \tag{2.15}$$

then $\mathbf{x}(s) \neq x_0$ for every $s \in [0,\tau]$.

Condition (2.15) is the key point to guarantee the propagation of singularities along generalized characteristics. For the cell problem with L in the form

$$L(x,v) = \frac{1}{2}\langle A(x)v, v\rangle - V(x) + E,$$

a local propagation result can be obtained replacing assumption (2.15) by the energy condition $E > \max_{x\in\mathbb{R}^n} V(x)$ (see [14]).

3. Generalized characteristics and Lax-Oleinik operators

For any $t > 0$, given $x, y \in \mathbb{R}^n$, we set

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

and define

$$A_t(x,y) = \min_{\xi\in\Gamma^t_{x,y}} \int_0^t L(\xi(s),\dot{\xi}(s))ds \qquad (x,y\in\mathbb{R}^n). \tag{3.1}$$

The existence of the above minimum is a well-known result in Tonelli's theory (see, for instance, [25]). Any $\xi \in \Gamma^t_{x,y}$ at which the minimum in (3.1) will be called a *minimizer* for $A_t(x,y)$ and such a minimizer ξ is of class C^2 by classical results. In the PDE literature, $A_t(x,y)$ is also called the *fundamental solution* of (2.6), see, for instance, [33].

Let L be a Tonelli Lagrangian satisfying (L1)-(L3) and let H be the associated Hamiltonian. In this section we study the singularities of a Lipschitz continuous semiconcave solution u of the Hamilton-Jacobi equation

$$H(x, Du(x)) = 0, \quad x \in \mathbb{R}^n. \tag{3.2}$$

The existence of such a solution is guaranteed by Proposition 2.7.

3.1. Lax-Oleinik operators.

For any Lipschitz continuous function $u : \mathbb{R}^n \to \mathbb{R}$ we set

$$\mathrm{Lip}(u) = \sup_{y\neq x} \frac{|u(y)-u(x)|}{|y-x|}\,. \tag{3.3}$$

For all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^n$ let

$$\phi^x_t(y) = u(y) - A_t(x,y) \quad \text{and} \quad \psi^x_t(y) = u(y) + A_t(y,x) \qquad (y\in\mathbb{R}^n) \tag{3.4}$$

where A_t is the fundamental solution of (3.2). The Lax-Oleinik operators T^-_t and T^+_t are defined as follows

$$T^+_t u(x) = \sup_{y\in\mathbb{R}^n} \phi^x_t(y), \quad x\in\mathbb{R}^n, \tag{3.5}$$

$$T^-_t u(x) = \inf_{y\in\mathbb{R}^n} \psi^x_t(y), \quad x\in\mathbb{R}^n. \tag{3.6}$$

The functions ϕ^x_t and ψ^x_t are also called *local barrier functions*.

Lemma 3.1. *Suppose L is a Tonelli Lagrangian and let u be a Lipschitz function on $\mathbb{R}^n$. Then the supremum in (3.5) is attained for every $(t,x)\in\mathbb{R}_+\times\mathbb{R}^n$. Moreover, there exists a constant $\lambda_0>0$, depending only on $\mathrm{Lip}(u)$, such that, for any $(t,x)\in\mathbb{R}_+\times\mathbb{R}^n$ and any maximum point $y_{t,x}$ of ϕ^x_t, we have*

$$|y_{t,x}-x|\leqslant\lambda_0 t\,. \tag{3.7}$$

Proof. Let $k_u=\mathrm{Lip}(u)+1$. Then, for any $(t,x)\in\mathbb{R}_+\times\mathbb{R}^n$ and $y\in\mathbb{R}^n$, (2.4) yields

$$\begin{aligned}A_t(x,y)\geqslant&\inf_{\xi\in\Gamma^t_{x,y}}\int_0^t\theta_1(|\dot\xi(s)|)\,ds-c_0t\geqslant\inf_{\xi\in\Gamma^t_{x,y}}k_u\int_0^t|\dot\xi(s)|\,ds-\big(\theta_1^*(k_u)+c_0\big)t\\ \geqslant&k_u|y-x|-\big(\theta_1^*(k_u)+c_0\big)t,\end{aligned}$$

where c_0 is the constant in assumption (L2). Therefore

$$\begin{aligned}&\phi^x_t(y)-\phi^x_t(x)=u(y)-u(x)-A_t(x,y)+A_t(x,x)\\ \leqslant&\,\mathrm{Lip}(u)|y-x|-k_u|y-x|+t\big(\theta_1^*(k_u)+c_0\big)+tL(x,0)\\ \leqslant&-|y-x|+t\big(\theta_1^*(k_u)+c_0+\theta_2(0)\big)\end{aligned}$$

where θ_2 is given by condition (L2). Now, taking $\lambda_0=\theta_1^*(k_u)+c_0+\theta_2(0)$ it follows that

$$\Lambda^x_t:=\{y:\phi^x_t(y)\geqslant\phi^x_t(x)\}\subset\overline{B}(x,\lambda_0t). \tag{3.8}$$

Therefore Λ^x_t is compact and the supremum in (3.5) is indeed a maximum. Moreover, (3.7) is a consequence of (3.8). □

A similar result holds for the inf-convolution defined in (3.6). A more detailed study of the properties of inf/sup-convolutions can be found in [8] with respect to the quadratic Hamiltonian $H(p)=|p|^2/2$ and, consequently, the kernel $A_t(x,y)=\frac{1}{2t}|x-y|^2$. This type of regularization, also called Moreau-Yosida regularization in convex analysis, was developed into a well-known procedure by Lasry and Lions [29]. In our context, we recover more information from the dynamical systems point of view by replacing quadratic kernels with the fundamental solutions.

3.2. **Propagation of singularities.** In this section, we will discuss the connection between sup-convolutions, singularities, and generalized characteristics. We begin our analysis with the local propagation of singularities of viscosity solutions along generalized characteristics. For Tonelli systems under rather general conditions, a local propagation result was obtained in [3] by a different method, without relating singular arcs to generalized characteristics. In the following lemma, we construct a singular arc starting from any singular point of the solution. A crucial point of this result is the fact that the interval $[0,t_0]$ on which the singular arc is defined turns out to be independent of the starting point x.

Lemma 3.2. *Let L be a Tonelli Lagrangian and let H be the associated Hamiltonian. Suppose $u:\mathbb{R}^n\to\mathbb{R}$ is a Lipschitz continuous semiconcave viscosity solution of (3.2). Then there exists t_0 in $(0,1]$ such that, for all $(t,x)\in(0,t_0]\times\mathbb{R}^n$, there is a unique maximum point $y_{t,x}$ of ϕ^x_t and the curve*

$$\mathbf{y}(t):=\begin{cases}x & \text{if}\quad t=0\\ y_{t,x} & \text{if}\quad t\in(0,t_0]\end{cases} \tag{3.9}$$

satisfies $\lim_{t\to0}\mathbf{y}(t)=x$.

Moreover, if $x\in\mathrm{Sing}\,(u)$, then $\mathbf{y}(t)\in\mathrm{Sing}\,(u)$ for all $t\in(0,t_0]$.

Proof. Let $C_1 > 0$ be a semiconcavity constant for u on $\mathbb{R}^n$ and let λ_0 be the positive constant in Lemma 3.1. By Proposition B.8 with $\lambda = 1 + \lambda_0$, we deduce that there exists $t_0 \in (0,1]$ and a constant $C_2 > 0$ such that for every $(t,x) \in (0,t_0] \times \mathbb{R}^n$, every $y \in B(x,\lambda t)$, and every $z \in B(0,\lambda t)$ we have that

$$A_t(x,y+z) + A_t(x,y-z) - 2A_t(x,y) \geqslant \frac{C_2}{t}|z|^2.$$

Thus, $\phi_t^x(y) = u(y) - A_t(x,y)$ is strictly concave on $\overline{B}(x,\lambda t)$ for all $t \in (0,t_0]$ provided that we further restrict t_0 in order to have

$$t_0 < \frac{C_2}{C_1}. \tag{3.10}$$

Then, for all such numbers t, there exists a unique maximum point $y_{t,x}$ of ϕ_t^x in $\overline{B}(x,\lambda t)$. In fact, $y_{t,x}$ is an interior point of $B(x,\lambda t)$ since, by Lemma 3.1, we have that $|y_t - x| \leqslant \lambda_0 t$.

We now prove that $y_{t,x}$ is a singular point of u for every $t \in (0,t_0]$. Let $\xi_{t,x} \in \Gamma^t_{x,y_{t,x}}$ be the unique minimizer for $A_t(x,y_{t,x})$ and let

$$p_{t,x}(s) := L_v(\xi_{t,x}(s),\dot{\xi}_{t,x}(s)), \quad s \in [0,t_0],$$

be the associated dual arc. We claim that

$$p_{t,x}(t) \in D^+u(y_{t,x}) \setminus D^*u(y_{t,x}), \tag{3.11}$$

which in turn yields $y_{t,x} \in \operatorname{Sing}(u)$. Indeed, if $p_{t,x}(t) \in D^*u(y_{t,x})$, then by Proposition 2.8 there would exist a C^2 curve $\gamma_{t,x} : (-\infty,t] \to \mathbb{R}^n$ solving the minimum problem

$$\min_{\gamma \in W^{1,1}([\tau,t];\mathbb{R}^n)} \left\{ \int_\tau^t L(\gamma(s),\dot{\gamma}(s))\,ds + u(\gamma(\tau)) \ : \ \gamma(t) = y_{t,x} \right\}$$

for all $\tau \leqslant t$. It is easily checked that $\gamma_{t,x}$ and $\xi_{t,x}$ coincide on $[0,t]$ since both of them are extremal curves for L and satisfy the same endpoint condition at $y_{t,x}$, i.e.,

$$L_v(\xi_{t,x}(t),\dot{\xi}_{t,x}(t)) = p_{t,x}(t) = L_v(\gamma_{t,x}(t),\dot{\gamma}_{t,x}(t)).$$

This leads to a contradiction since $x \in \operatorname{Sing}(u)$ while u should be smooth at $\gamma_{t,x}(0)$. Thus, (3.11) holds true and $y_t \in \operatorname{Sing}(u)$. □

Next, we proceed to show that the singular arc in Lemma 3.2 is a generalized characteristic.

Lemma 3.3. *Let t_0 and $\mathbf{y}$ be given by Lemma 3.2 for a given $x \in \mathbb{R}^n$. For any $t \in (0,t_0]$ let $\xi_{t,x} \in \Gamma^t_{x,\mathbf{y}(t)}$ be a minimizer for $A_t(x,\mathbf{y}(t))$. Then*

$$\{\dot{\xi}_{t,x}(\cdot)\}_{t\in(0,t_0]} \text{ is an equi-Lipschitz family.} \tag{3.12}$$

Proof. Since $\mathbf{y}(t) \in \overline{B}(x,\lambda t)$ by Lemma 3.1, we have that $(\xi_{t,x}(s),p_{t,x}(s)) \in \mathbf{K}^*_{x,\lambda_0}$ for all $s \in [0,t]$, where the compact set $\mathbf{K}^*_{x,\lambda_0}$ is defined in (B.2). Therefore, being solutions of the Hamiltonian system

$$\begin{cases} \dot{\xi}_{t,x}(s) = H_p(\xi_{t,x}(s),p_{t,x}(s)) \\ \dot{p}_{t,x}(s) = -H_x(\xi_{t,x}(s),p_{t,x}(s)) \end{cases} \qquad s \in [0,t],$$

both $\{\dot{\xi}_{t,x}(\cdot)\}_{t\in(0,t_0]}$ and $\{\dot{p}_{t,x}(\cdot)\}_{t\in(0,t_0]}$ are uniformly bounded. Consequently,

$$\ddot{\xi}_{t,x}(s) = H_{px}(\xi_{t,x}(s),p_{t,x}(s))\dot{\xi}_{t,x}(s) + H_{pp}(\xi_{t,x}(s),p_{t,x}(s))\dot{p}_{t,x}(s) \qquad (s \in [0,t])$$

is also bounded, uniformly for $t \in (0,t_0]$. □

Proposition 3.4. *Let L be a Tonelli Lagrangian. Let $t_0 \in (0,1]$ be given by Lemma 3.2. For any fixed $x \in \mathbb{R}^n$, let $\mathbf{y} : [0,t_0] \to \mathbb{R}^n$ be the curve constructed in Lemma 3.2. Then*

(a) *$\mathbf{y}$ is Lipschitz on $[0,t_0]$.*

Moreover, for any $t \in (0,t_0]$, let $\xi_{t,x} \in \Gamma^t_{x,\mathbf{y}(t)}$ be a minimizer for $A_t(x,\mathbf{y}(t))$. Then the following properties hold true:

(b) *The right derivative $\dot{\mathbf{y}}^+(0)$ exists and*

$$\dot{\mathbf{y}}^+(0) = \lim_{t\to 0^+} \dot{\xi}_{t,x}(t) = H_p(x,p_x) \tag{3.13}$$

where p_x is the unique element of $D^+u(x)$ such that

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

(c) *The arc $\mathbf{p}(t) := L_v(\xi_{t,x}(t), \dot{\xi}_{t,x}(t))$ is continuos on $(0,t_0]$ and $\lim_{t\to 0^+} \mathbf{p}(t) = p_x$.*

(d) *There exist $0 < \rho \leqslant t_0$ and constants $C_1, C_2 > 0$ such that*

$$H(\mathbf{y}(t),\mathbf{p}(t)) \leqslant H(x,p_x) + C_1 t - C_2|\mathbf{p}(t) - p_x|^2, \quad \forall t \in (0,\rho]. \tag{3.14}$$

Proof. Having fixed $x \in \mathbb{R}^n$, we shall abbreviate $\xi_{t,x} = \xi_t$. Let $0 < t,s \leqslant t_0$ and let $\xi_t \in \Gamma^t_{x,\mathbf{y}(t)}$, $\xi_s \in \Gamma^s_{x,\mathbf{y}(s)}$ and $\eta \in \Gamma^t_{x,\mathbf{y}(s)}$ be minimizers for $A_t(x,\mathbf{y}(t))$, $A_s(x,\mathbf{y}(s))$, and $A_t(x,\mathbf{y}(s))$ respectively. Setting $p_t = L_v(\xi_t(t), \dot{\xi}_t(t))$, $p_s = L_v(\xi_s(s), \dot{\xi}_s(s))$, and $p = L_v(\eta(t), \dot{\eta}(t))$ we have that

$$\begin{aligned}
&\frac{C_2}{t}|\mathbf{y}(t) - \mathbf{y}(s)|^2\\
&\qquad\leqslant\langle p_t - p, \mathbf{y}(t) - \mathbf{y}(s)\rangle = \langle p_t - p_s, \mathbf{y}(t) - \mathbf{y}(s)\rangle + \langle p_s - p, \mathbf{y}(t) - \mathbf{y}(s)\rangle\\
&\qquad\leqslant C_1|\mathbf{y}(t) - \mathbf{y}(s)|^2 + \langle p_s - p, \mathbf{y}(t) - \mathbf{y}(s)\rangle,
\end{aligned}$$

where we have used the notation of the proof of Lemma 3.2. By Proposition B.9, the function $(t,y) \mapsto A_t(x,y)$ is locally $C^{1,1}$ in the set $\{(t,y) \in \mathbb{R}\times\mathbb{R}^n : 0 < t < t_0, |y - x| < \lambda t\}$. Moreover, Proposition B.3 together with Proposition B.8 ensures that

$$|p_s - p| \leqslant \frac{C_3}{t}|s - t|$$

for some constant $C_3 > 0$. Therefore

$$\left(\frac{C_2}{t} - C_1\right)|\mathbf{y}(t) - \mathbf{y}(s)|^2 \leqslant \frac{C_3}{t}|s-t||\mathbf{y}(t) - \mathbf{y}(s)|.$$

Recalling (3.10) we have that $C_2/t - C_1 > 0$ for all $0 < t \leqslant t_0$. Thus

$$|\mathbf{y}(t) - \mathbf{y}(s)| \leqslant \frac{C_3}{C_2 - C_1 t_0}|t - s|,$$

and this proves (a).

Now we turn to the proof of (b). Since $\{\dot{\xi}_t(\cdot)\}_{t\in(0,t_0]}$ are equi-Lipschitz by Lemma 3.3, for any sequence $t_k \to 0^+$ such that $v_k := (\xi_{t_k}(t_k) - x)/t_k$ converges, we obtain

$$\begin{aligned}
\left|\frac{\xi_{t_k}(t_k) - x}{t_k} - \dot{\xi}_{t_k}(t_k)\right| &\leqslant \frac{1}{t_k}\int_0^{t_k} |\dot{\xi}_{t_k}(s) - \dot{\xi}_{t_k}(t_k)|\, ds\\
&\leqslant \frac{C}{t_k}\int_0^{t_k} (t_k - s)\, ds = \frac{C}{2}t_k.
\end{aligned} \tag{3.15}$$

This implies that

$$v_0 := \lim_{k\to\infty} v_k = \lim_{k\to\infty} \dot{\xi}_{t_k}(t_k).$$

By the semiconcavity of u, for any $p\in D^+u(x)$, we have

$$
\begin{aligned}
u(x)&\leqslant u(\mathbf{y}(t_k))+\langle L_v(\mathbf{y}(t_k),\dot{\xi}_{t_k}(t_k)),x-\mathbf{y}(t_k)\rangle+\frac{C}{2}|x-\mathbf{y}(t_k)|^2\\
&\leqslant u(x)+\langle p,\mathbf{y}(t_k)-x\rangle+\langle L_v(\mathbf{y}(t_k),\dot{\xi}_{t_k}(t_k)),x-\mathbf{y}(t_k)\rangle+C|x-\mathbf{y}(t_k)|^2.
\end{aligned}
$$

Then, recalling that $\xi_{t_k}(t_k)=\mathbf{y}(t_k)$ we have

$$
\langle p-L_v(\mathbf{y}(t_k),\dot{\xi}_{t_k}(t_k)),v_k\rangle+t_kC|v_k|^2\geqslant 0,\quad \forall p\in D^+u(x). \tag{3.16}
$$

Taking the limit in (3.16) as $k\to\infty$ we obtain

$$
\langle p,v_0\rangle\geqslant\langle L_v(x,v_0),v_0\rangle=\langle p_x,v_0\rangle,\quad \forall p\in D^+u(x), \tag{3.17}
$$

where $p_x:=L_v(x,v_0)\in D^+u(x)$ by the upper semicontinuity of $x\rightsquigarrow D^+u(x)$. So,

$$
H(x,p)\geqslant\langle L_v(x,v_0),v_0\rangle-L(x,v_0)=H(x,p_x),\quad \forall p\in D^+u(x), \tag{3.18}
$$

and p_x is the unique minimum point of $H(x,\cdot)$ on $D^+u(x)$. The uniqueness of p_x implies the uniqueness of v_0 since $L_v(x,\cdot)$ is injective. This leads to the assertion that

$$
v_0=\lim_{t\to0^+}\frac{\xi_t(t)-x}{t}=\lim_{t\to0^+}\dot{\xi}_t(t)
$$

and, together with (3.18), implies (3.13). This completes the proof of (b).

The conclusion (c) is a straight consequence of (a), (b) and the locally $C^{1,1}$ regularity property of the function $(t,y)\mapsto A_t(x,y)$.

Finally, we turn to prove (d). First, using Tailor's expansion, we have that

$$
\begin{aligned}
&H(x,p_x)-H(\mathbf{y}(s),\mathbf{p}(s))\\
=&H_x(\mathbf{y}(s),\mathbf{p}(s))(x-\mathbf{y}(s))+H_p(\mathbf{y}(s),\mathbf{p}(s))(p_x-\mathbf{p}(s))\\
&+\frac12\langle(H_{xp}(\mathbf{y}(s),\mathbf{p}(s))+H_{px}(\mathbf{y}(s),\mathbf{p}(s)))(p_x-\mathbf{p}(s)),(x-\mathbf{y}(s))\rangle\\
&+\frac12\langle H_{xx}(\mathbf{y}(s),\mathbf{p}(s))(x-\mathbf{y}(s)),(x-\mathbf{y}(s))\rangle\\
&+\frac12\langle H_{pp}(\mathbf{y}(s),\mathbf{p}(s))(p_x-\mathbf{p}(s)),(p_x-\mathbf{p}(s))\rangle+o(|\mathbf{y}(s)-x|^2+|\mathbf{p}(s)-p_x|^2).
\end{aligned}
$$

Thus, by (a), (b), (c) and our assumptions on H, there exist $\rho>0$ such that, for $s\in(0,\rho]$, we have

$$
\begin{aligned}
&H(x,p_x)-H(\mathbf{y}(s),\mathbf{p}(s))\\
\geqslant&-C_1s+\langle\dot{\xi}_s(s),p_x-\mathbf{p}(s)\rangle-C_\varepsilon s^2-\varepsilon|\mathbf{p}(s)-p_x|^2+C_2|\mathbf{p}(s)-p_x|^2
\end{aligned}
$$

Taking $\varepsilon>0$ small enough, we have

$$
H(x,p_x)-H(\mathbf{y}(s),\mathbf{p}(s))\geqslant-C_3s+\langle\dot{\xi}_s(s),p_x-\mathbf{p}(s)\rangle+C_4|\mathbf{p}(s)-p_x|^2.
$$

In view of (3.15), we have

$$
H(x,p_x)-H(\mathbf{y}(s),\mathbf{p}(s))\geqslant-C_5s+\left\langle\frac{\mathbf{y}(s)-x}{s},p_x-\mathbf{p}(s)\right\rangle+C_4|\mathbf{p}(s)-p_x|^2.
$$

Therefore, by the semiconcavity of u, we obtain

$$
H(x,p_x)-H(\mathbf{y}(s),\mathbf{p}(s))\geqslant-C_6s+C_4|\mathbf{p}(s)-p_x|^2,
$$

which completes the proof of (d). □

Remark 3.5. Observe that (3.17), that is,

$$\langle p - p_x, v_0\rangle \geqslant 0, \quad \forall p \in D^+u(x),$$

is exactly the key condition for propagation of singularities in [4] and [18].

Theorem 3.6. *Let L be a Tonelli Lagrangian and let H be the associated Hamiltonian. Suppose $u : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz continuous semiconcave viscosity solution of* (3.2) *and $x \in \mathrm{Sing}\,(u)$. Then the singular arc $\mathbf{y} : [0, t_0] \to \mathbb{R}^n$ defined in Lemma 3.2 is a generalized characteristic and satisfies*

$$\dot{\mathbf{y}}(\tau) \in \mathrm{co}\, H_p(\mathbf{y}(\tau), D^+u(\mathbf{y}(\tau))), \quad \textit{a.e. } \tau \in [0, t_0]. \tag{3.19}$$

Moreover,

$$\dot{\mathbf{y}}^+(0) = H_p(x, p_0), \tag{3.20}$$

where p_0 is the unique element of minimal energy:

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

Proof. The conclusion can be derived directly from Lemma 3.2 and Proposition 3.4 except for (3.19). For the proof of (3.19), see Appendix C. □

To study the genuine propagation of singularities along generalized characteristics, we have to check that the singular arc $\mathbf{y}(t)$ in Lemma 3.2 does not keep constant locally. As we show below, the following condition can be useful for this purpose:

$$D_yA_t(x, x) \notin D^+u(x), \quad \text{for all } t \in (0, t_0]. \tag{3.21}$$

Proposition 3.7. *Let $\mathbf{y} : [0, t_0] \to \mathbb{R}^n$ be the singular generalized characteristic in Theorem 3.6, and let $t \in (0, t_0]$. Then $\mathbf{y}(t) = x$ if and only if $D_yA_t(x, x) \in D^+u(x)$. Consequently, if* (3.21) *holds, then $\mathbf{y}(t) \neq x$ for every $t \in (0, t_0]$.*

Proof. Let $p \in D^+u(x)$, $p' = D_yA_t(x, x)$, and let $t \in (0, t_0]$. Recalling that $\mathbf{y}(t)$ is the unique maximizer of ϕ_t^x we have

$$\begin{aligned}
0 \leqslant \phi_t^x(\mathbf{y}(t)) - \phi_t^x(x) =& [u(\mathbf{y}(t)) - u(x) - \langle p, \mathbf{y}(t) - x\rangle - \frac{C_1}{2}|\mathbf{y}(t) - x|^2] \\
&- [A_t(x, \mathbf{y}(t)) - A_t(x, x) - \langle p', \mathbf{y}(t) - x\rangle - \frac{C_2}{2t}|\mathbf{y}(t) - x|^2] \\
&+ \langle p - p', \mathbf{y}(t) - x\rangle + \frac{1}{2}\Big(C_1 - \frac{C_2}{t}\Big)|\mathbf{y}(t) - x|^2, \\
\leqslant& \langle p - p', \mathbf{y}(t) - x\rangle + \frac{1}{2}\Big(C_1 - \frac{C_2}{t}\Big)|\mathbf{y}(t) - x|^2,
\end{aligned}$$

where—like in the proof of Lemma 3.2—$C_1 > 0$ is a semiconcavity constant for u on $\mathbb{R}^n$ and $C_2 > 0$ a convexity constant for $A_t(x, \cdot)$ on $B(x, (1 + \lambda_0)t)$. So,

$$0 \leqslant |\mathbf{y}(t) - x| \leqslant \frac{2|p - p'|}{C_2/t - C_1}.$$

If $D_yA_t(x, x) \in D^+u(x)$, then taking $p = p'$ in the above inequality yields $\mathbf{y}(t) = x$. Conversely, if $\mathbf{y}(t) = x$, then the nonsmooth Fermat rule yields $0 \in D^+u(x) - D_yA_t(x, x)$ which completes the proof. □

Another condition that ensures the genuine propagation of singularities is related to the notion of critical point.

Definition 3.8. We say that $x \in \mathbb{R}^n$ is a *critical point* of a viscosity solution u of (3.2) if $0 \in \operatorname{co} H_p(x, D^+u(x))$, and a *strong critical point* of u if $0 \in H_p(x, D^+u(x))$.

Remark 3.9. For a mechanical Lagrangian of the form

$$L(x,v) = \frac{1}{2}\langle A(x)v, v\rangle - V(x), \tag{3.22}$$

with $\langle A(x)\cdot,\cdot\rangle$ the matrix associated with a Riemannian metric in $\mathbb{R}^n$ and V a smooth potential, x is a critical point of a semiconcave solution u of the corresponding Hamilton-Jacobi equation

$$\frac{1}{2}\langle A(x)^{-1}Du, Du\rangle + V(x) = 0$$

if and only if $0 \in D^+u(x)$, i.e., x is a critical point of u in the sense of nonsmooth analysis.

It is already known the condition that x is not a critical point is a key point to guarantee the genuine propagation of singularities along generalized characteristics (see, for instance, [4]).

Corollary 3.10. *Let* $\mathbf{y} : [0, t_0] \to \mathbb{R}^n$ *be the singular generalized characteristic in Theorem 3.6. If x is not a strong critical point of u then there exists* $t \in (0, t_0]$ *such that* $\mathbf{y}(s) \neq x$ *for all* $s \in (0, t]$.

Proof. It suffices to show that $0 \in H_p(x, D^+u(x))$ whenever a sequence $t_k \to 0$ exists such that $\mathbf{y}(t_k) = x$ for all $k \in \mathbb{N}$. Indeed, denoting by $\xi_k \in \Gamma^{t_k}_{x,\mathbf{y}(t_k)}$ the unique minimizer of $A_{t_k}(x, \mathbf{y}(t_k))$, as in the proof of Proposition 3.4 we have that

$$\lim_{k\to\infty} \dot{\xi}_k(t_k) = \lim_{k\to\infty} \frac{\xi_k(t_k) - x}{t_k} = 0$$

because $\xi_k(t_k) = \mathbf{y}(t_k)$. Therefore the dual arc $p_k(s) = L_v(\xi_k(s), \dot{\xi}_k(s))$ satisfies

$$\lim_{k\to\infty} H_p(\xi_k(t_k), p_k(t_k)) = \lim_{k\to\infty} \dot{\xi}_k(t_k) = 0.$$

Now, since $p_k(t_k) = D_y A_{t_k}(x, \mathbf{y}(t_k)) \in D^+u(\mathbf{y}(t_k))$ by the nonsmooth Fermat rule, the upper semicontinuity of $z \rightsquigarrow H_p(x, D^+u(z))$ yields $0 \in H_p(x, D^+u(x))$. □

The above results on the propagation of singularities along generalized characteristics leads to the following global propagation property.

Theorem 3.11. *Let L be a Tonelli Lagrangian and let H be the associated Hamiltonian. Suppose* $u : \mathbb{R}^n \to \mathbb{R}$ *is a Lipschitz continuous semiconcave viscosity solution of* (3.2)*. If* $x \in \operatorname{Sing}(u)$*, then there exists a generalized characteristic* $\mathbf{x} : [0, +\infty) \to \mathbb{R}^n$ *such that* $\mathbf{x}(0) = x$ *and* $\mathbf{x}(s) \in \operatorname{Sing}(u)$ *for all* $s \in [0, +\infty)$.

For geodesic systems, global propagation results were obtained in [1], [2], and [5] even on Riemannian manifolds. Theorem 3.11 above applies to mechanical systems with a Lagrangian L of the form (3.22).

Corollary 3.12. *Let L be the Tonelli Lagrangian in* (3.22) *and let H be the associated Hamiltonian. Suppose that A and V are bounded together with and all their derivatives up to the second order and let u be a viscosity solution of the Hamilton-Jacobi equation* (3.2)*. If* $x \in \operatorname{Sing}(u)$*, then there exists a unique generalized characteristic* $\mathbf{x} : [0, +\infty) \to \mathbb{R}^n$ *such that* $\mathbf{x}(0) = x$ *and* $\mathbf{x}(s) \in \operatorname{Sing}(u)$ *for all* $s \in [0, +\infty)$.

Proof. Observing that all the conditions (L1)-(L3) are satisfied, the main part of the conclusion is an immediate consequence of Theorem 3.11. The uniqueness of the generalized characteristic is a well-known consequence of the semiconcavity of u (see, e.g., [16]). □

Appendix A. Uniform Lipschitz bound for minimizers

In this appendix we adapt to the present context a Lipschitz estimate for minimizers of the action functional that was obtained in [24] (see also [6]). We give a detailed proof of this result for the readers' convenience. We assume that the Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a function of class C^2 that satisfies the following conditions:

(L1') *Convexity*: $L_{vv}(x, v) > 0$ for all $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$.

(L2') *Growth condition*: There exists a superlinear function $\theta : [0, +\infty) \to [0, +\infty)$ and a constant $c_0 > 0$ such that

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

(L3') *Uniform bound*: There exists a nondecreasing function $K : [0, +\infty) \to [0, +\infty)$ such that

$$L(x, v) \leqslant K(|v|) \qquad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Observe that (L1')-(L3') are weaker than assumptions (L1)-(L3).

We define the *energy function*

$$E(x, v) = \langle v, L_v(x, v)\rangle - L(x, v), \quad (x, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Proposition A.1. *Let $t, R > 0$ and suppose L satisfies condition* (L1')-(L3'). *Given any $x \in \mathbb{R}^n$ and $y \in \overline{B}(x, R)$, let $\xi \in \Gamma^t_{x,y}$ be a minimizer for $A_t(x, y)$. Then we have that*

$$\sup_{s\in[0,t]} |\dot\xi(s)| \leqslant \kappa(R/t), \tag{A.1}$$

where $\kappa : (0, \infty) \to (0, \infty)$ is nondecreasing. Moreover, if $t \leqslant 1$, then

$$\sup_{s\in[0,t]} |\xi(s) - x| \leqslant \kappa(R/t). \tag{A.2}$$

Proof. Fix $t > 0$, $R > 0$, $x \in \mathbb{R}^n$, let $y \in \overline{B}(x, R)$, and let $\xi \in \Gamma^t_{x,y}$ 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 $\sigma(s) = x + \frac{s}{t}(y - x)$, $s \in [0, t]$, in view of (L2') and (L3') we have that

$$\int_0^t \theta(|\dot\xi(s)|)\, ds - c_0 t \leqslant \int_0^t L(\xi(s), \dot\xi(s))\, ds \leqslant \int_0^t L(\sigma(s), \dot\sigma(s))\, ds$$
$$= \int_0^t L\Big(x + \frac{s}{t}(y - x), \frac{y - x}{t}\Big)\, ds \leqslant tK(R/t).$$

Therefore

$$\int_0^t \theta(|\dot\xi(s)|)\, ds \leqslant c_0 t + tK(R/t) = tC_1(R/t)$$

with $C_1(r) = K(r) + c_0$. Since $|\dot\xi(s)| \leqslant \theta(|\dot\xi(s)|) + \theta^*(1)$, where θ^* is the convex conjugate of θ defined in (2.3), we have that

$$\int_0^t |\dot\xi(s)| \leqslant tC_2(R/t),$$

with $C_2(r) = C_1(r) + \theta^*(1)$. Hence

$$|\xi(s) - x| \leqslant \int_0^s |\dot\xi(s)|\, ds \leqslant tC_2(R/t), \quad \forall s \in [0, t], \tag{A.3}$$

and

$$\inf_{s\in[0,t]}|\dot{\xi}(s)|\leqslant\frac{1}{t}\int_0^t|\dot{\xi}(s)|\,ds\leqslant C_2(R/t). \tag{A.4}$$

Now, define $l_\xi(s,\lambda)=L(\xi(s),\dot{\xi}(s)/\lambda)\lambda$ for all $s\in[0,t]$ and $\lambda>0$. Then we have

$$\frac{d}{d\lambda}l_\xi(s,\lambda)|_{\lambda=1}=L(\xi(s),\dot{\xi}(s))-\langle\dot{\xi}(s),L_v(\xi(s),\dot{\xi}(s))\rangle=-E(\xi(s),\dot{\xi}(s)).$$

Since the energy is constant along a minimizer, there exists a constant c_ξ such that

$$\frac{d}{d\lambda}l_\xi(s,\lambda)|_{\lambda=1}=c_\xi,\quad\forall s\in[0,t].$$

Moreover, a simple computation shows that $l_\xi(s,\lambda)$ is convex in λ. So, we have

$$c_\xi\geqslant\sup_{\lambda<1}\frac{l_\xi(s,\lambda)-l_\xi(s,1)}{\lambda-1},\quad\forall s\in[0,t].$$

Let us now take, in the above inequality, $\lambda=3/4$ and $s_0\in[0,t]$ such that $|\dot{\xi}(s_0)|=\inf_{s\in[0,t]}|\dot{\xi}(s)|$. Then, by (L2'), (L3'), and (A.4) we conclude that

$$\begin{aligned}c_\xi\geqslant&4(l_\xi(s_0,1)-l_\xi(s_0,3/4))\geqslant4(-c_0-l_\xi(s_0,3/4))\\=&-4c_0-3L\Big(\xi(s_0),\frac{4}{3}\dot{\xi}(s_0)\Big)\geqslant-4c_0-3K\Big(\frac{4}{3}|\dot{\xi}(s_0)|\Big)\\\geqslant&-4c_0-3K\Big(\frac{4}{3}C_2(R/t)\Big)=-C_3(R/t),\end{aligned} \tag{A.5}$$

where $C_3(r)=4c_0+3K(4C_2(r)/3)$.

By the convexity of $l_\xi(s,\cdot)$ we also have, for any $\varepsilon\in(0,1)$,

$$c_\xi\leqslant\frac{l_\xi(s,2-\varepsilon)-l_\xi(s,1)}{1-\varepsilon}.$$

In other words,

$$(1-\varepsilon)c_\xi+\varepsilon l_\xi(s,1)\leqslant l_\xi(s,2-\varepsilon)-(1-\varepsilon)l_\xi(s,1).$$

Moreover, again by convexity, we have that

$$l_\xi(s,2-\varepsilon)=l_\xi\Big(s,\varepsilon\cdot\frac{1}{\varepsilon}+(1-\varepsilon)\cdot1\Big)\leqslant\varepsilon l_\xi\Big(s,\frac{1}{\varepsilon}\Big)+(1-\varepsilon)l_\xi(s,1).$$

Therefore

$$(1-\varepsilon)c_\xi+\varepsilon l_\xi(s,1)\leqslant\varepsilon l_\xi\Big(s,\frac{1}{\varepsilon}\Big),$$

that is,

$$(1-\varepsilon)c_\xi+\varepsilon L(\xi(s),\dot{\xi}(s))\leqslant L(\xi(s),\varepsilon\dot{\xi}(s)).$$

Hence, combining (A.5) and condition (L2'), we obtain

$$-(1-\varepsilon)C_3(R/t)+\varepsilon(\theta(|\dot{\xi}(s)|)-c_0)\leqslant L(\xi(s),\varepsilon\dot{\xi}(s)).$$

Set $S_\xi=\{s\in[0,t]:|\dot{\xi}(s)|\geqslant2\}$ and $\varepsilon=\varepsilon(s)=1/|\dot{\xi}(s)|$ for $s\in S_\xi$. Then

$$-C_3(R/t)+\frac{1}{|\dot{\xi}(s)|}C_3(R/t)+\frac{\theta(|\dot{\xi}(s)|)-c_0}{|\dot{\xi}(s)|}\leqslant L\Big(\xi(s),\frac{\dot{\xi}(s)}{|\dot{\xi}(s)|}\Big)\leqslant K(1),\quad\forall s\in S_\xi.$$

Thus

$$\theta(|\dot{\xi}(s)|)\leqslant(K(1)+C_3(R/t))|\dot{\xi}(s)|+(c_0-C_3(R/t)),\quad\forall s\in S_\xi.$$

Therefore, by the Young-Fenchel inequality we deduce that

$$|\dot{\xi}(s)| \leqslant (c_0 - C_3(R/t)) + \theta^*(K(1) + C_3(R/t) + 1) := C_4(R/t), \quad \forall s \in S_\xi.$$

Consequently,

$$\sup_{s\in[0,t]} |\dot{\xi}(s)| \leqslant \max\{2, C_4(R/t)\} := C_5(R/t). \tag{A.6}$$

The conclusion follows from (A.6) and (A.3) taking $\kappa(r) = \max\{C_5(r), C_2(r)\}$. □

Corollary A.2. *In Proposition A.1, assume the additional condition:*

(L3”) *There exists a nondecreasing function $K_1 : [0, +\infty) \to [0, +\infty)$ such that*

$$|L_v(x, v)| \leqslant K_1(|v|) \qquad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Then the dual arc $p(\cdot)$ associated with $\xi(\cdot)$ satisfies

$$\sup_{s\in[0,t]} |p(s)| \leqslant \kappa_1(R/t), \tag{A.7}$$

where $\kappa_1 : (0, \infty) \to (0, \infty)$ is nondecreasing.

Proof. By (L3”) together with (A.1) and (A.7) follows from

$$\sup_{s\in[0,t]} |p(s)| = \sup_{s\in[0,t]} |L_v(\xi(s), \dot{\xi}(s))| \leqslant K_1(\kappa(R/t)) = \kappa_1(R/t),$$

where $\kappa_1(r) = K_1 \circ \kappa(r)$. □

Appendix B. Convexity and $C^{1,1}$ estimate of fundamental solutions

Let L be a Tonelli Lagrangian (which implies conditions (L1’)-(L3’) and (L3”) in Appendix A). Then we have the following fundamental bounds for the velocity of minimizers.

Fix $x \in \mathbb{R}^n$ and suppose $R > 0$ and L is a Tonelli Lagrangian. For any $0 < t \leqslant 1$ and $y \in \overline{B}(x, R)$, let $\xi \in \Gamma^t_{x,y}$ be a minimizer for $A_t(x, y)$ and let p be its dual arc. Then there exists a nondecreasing function $\kappa : (0, \infty) \to (0, \infty)$ such that

$$\sup_{s\in[0,t]} |\dot{\xi}(s)| \leqslant \kappa(R/t), \qquad \sup_{s\in[0,t]} |p(s)| \leqslant \kappa(R/t), \tag{B.1}$$

by Proposition A.1 and Corollary A.2. Now, $x \in \mathbb{R}^n$ and $\lambda > 0$ define compact sets

$$\begin{aligned} \mathbf{K}_{x,\lambda} &:= \overline{B}(x, \kappa(4\lambda)) \times \overline{B}(0, \kappa(4\lambda)) \subset \mathbb{R}^n \times \mathbb{R}^n, \\ \mathbf{K}^*_{x,\lambda} &:= \overline{B}(x, \kappa(4\lambda)) \times \overline{B}(0, \kappa(4\lambda)) \subset \mathbb{R}^n \times (\mathbb{R}^n)^*. \end{aligned} \tag{B.2}$$

The following is one of the key technical points of this paper.

Proposition B.1. *Suppose L is a Tonelli Lagrangian. Fix $x \in \mathbb{R}^n$, $\lambda > 0$, $t \in (0, 1)$, and $y \in B(x, \lambda t)$. Let $z \in \mathbb{R}^n$ and $h \in \mathbb{R}$ be such that*

$$|z| < \lambda t \qquad \text{and} \qquad -\frac{t}{2} < h < 1 - t. \tag{B.3}$$

Then any minimizer $\xi \in \Gamma^{t+h}_{x,y+z}$ for $A_{t+h}(x, y + z)$ and corresponding dual arc p satisfy the following inclusions

$$\begin{aligned} &\{(\xi(s), \dot{\xi}(s)) : s \in [0, t + h]\} \subset \mathbf{K}_{x,\lambda}, \\ &\{(\xi(s), p(s)) : s \in [0, t + h]\} \subset \mathbf{K}^*_{x,\lambda}. \end{aligned}$$

Proof. Since $t/2 < t+h < 1$ and $y+z \in B(x, 2\lambda t)$ by (B.3), we can use (A.2) and (B.1) to obtain

$$\sup_{s\in[0,t+h]} |\xi(s) - x| \leqslant \kappa\Big(\frac{2\lambda t}{t+h}\Big) \leqslant \kappa(4\lambda)$$

and

$$\sup_{s\in[0,t]} |\dot{\xi}(s)| \leqslant \kappa\Big(\frac{2\lambda t}{t+h}\Big) \leqslant \kappa(4\lambda).$$

Since a similar bound holds true for $\sup_{s\in[0,t]}|p(s)|$, the conclusion follows. □

Remark B.2. For any $x \in \mathbb{R}^n$ and $y \in B(x, \lambda t)$, condition (B.3) is satisfied when

$$|h| < t/2 \qquad \text{and} \qquad |z| < \lambda t \tag{B.4}$$

provided that $0 < t < 2/3$.

B.1. **Semiconcavity of the fundamental solution.** The role of semiconcavity in optimal control problems has been widely investigated, see [16]. For the minimization problem in (3.1), the local semiconcavity of $A_t(x,y)$ with respect to y was proved in [10]. In this paper, we give a local semiconcavity result of the map $(t,y) \mapsto A_t(x,y)$.

The following result is essentially known.

Proposition B.3 (Semiconcavity of the fundamental solution)**.** *Suppose L is a Tonelli Lagrangian. Then for any $\lambda > 0$ there exists a constant $C_\lambda > 0$ such that for any $x \in \mathbb{R}^n$, $t \in (0, 2/3)$, $y \in B(x, \lambda t)$, and $(h,z) \in \mathbb{R} \times \mathbb{R}^n$ satisfying $|h| < t/2$ and $|z| < \lambda t$ we have*

$$A_{t+h}(x, y+z) + A_{t-h}(x, y-z) - 2A_t(x,y) \leqslant \frac{C_\lambda}{t}\big(|h|^2 + |z|^2\big). \tag{B.5}$$

Consequently, $(t,y) \mapsto A_t(x,y)$ is locally semiconcave in $(0,1) \times \mathbb{R}^n$, uniformly with respect to x.

Remark B.4. For the purposes of this paper, it suffices to assume $0 < t < 2/3$. In the general case $t > 0$, a local semiconcavity result holds true for $A_t(x,y)$ in the same form as (B.5) with C_λ depending on t.

B.2. **Main Regularity Lemma.** We begin with the following known properties.

Lemma B.5. *Suppose L is a Tonelli Lagrangian on $\mathbb{R}^n$. For any $x, y \in \mathbb{R}^n$, $t > 0$, let $\xi \in \Gamma^t_{x,y}$ be a minimizer for $A_t(x,y)$. Then $\xi \in C^2([0,t])$ is an extremal curve*[1]*, and the dual arc $p(s) := L_v(\xi(s), \dot{\xi}(s))$ satisfies the sensitivity relation*

$$p(s) \in D^+_y A_t(x, \xi(s)), \quad s \in [0,t]. \tag{B.6}$$

Moreover, $A_t(x,\cdot)$ is differentiable at y if and only if there is a unique minimizer $\xi \in \Gamma^t_{x,y}$. In this case, we have

$$D_y A_t(x,y) = L_v(\xi(t), \dot{\xi}(t)). \tag{B.7}$$

Proof. The sensitivity relation (B.6) is obtained in, for instance, [16, Theorem 6.4.8], for a problem with initial cost. Here the proof is similar. The uniqueness of the minimizer and regularity are classical results. □

[1] An arc $\xi(s)$ is called an extremal curve if it satisfies the associated Euler-Lagrange equation.

Lemma B.6 (Main Regularity Lemma)**.** *Suppose L is a Tonelli Lagrangian. Then for any $\lambda > 0$ there exists $t_\lambda \in (0,1]$ and constants $C_\lambda, C'_\lambda, C''_\lambda > 0$ such that, for all $t \in (0,t_\lambda)$, $x \in \mathbb{R}^n$, $y_1, y_2 \in B(x,\lambda t)$, and any minimizer ξ_i $(i=1,2)$ for $A_t(x,y_i)$, we have*

$$\|\xi_2-\xi_1\|^2_{L^\infty(0,t)} \leqslant \frac{C_\lambda}{t}|y_2-y_1|^2 \tag{B.8}$$

$$\int_0^t |p_2-p_1|^2 ds \leqslant \frac{C'_\lambda}{t}|y_2-y_1|^2 \tag{B.9}$$

$$\int_0^t |\dot\xi_2-\dot\xi_1|^2 ds \leqslant \frac{C''_\lambda}{t}|y_2-y_1|^2, \tag{B.10}$$

where p_i denotes the dual arc of ξ_i.

Proof. Since L is a Tonelli Lagrangian, we have that $\xi_i(s)$ $(i=1,2)$ of class C^2 and, by Proposition B.1 with $h=0=z$, it follows that

$$\sup_{s\in[0,t]}|\dot\xi_i(s)| \leqslant \kappa(4\lambda), \qquad \sup_{s\in[0,t]}|p_i(s)| \leqslant \kappa(4\lambda),$$

where $p_i(s)=L_v(\xi_i(s),\dot\xi_i(s))$. Moreover, the pair $(\xi_i(\cdot),p_i(\cdot))$ satisfies the Hamiltonian system

$$\begin{cases}\dot\xi_i = H_p(\xi_i,p_i)\\ \dot{p_i} = -H_x(\xi_i,p_i)\end{cases} \qquad \text{on } [0,t]$$

with

$$\xi_i(t)=y_i, \quad \xi_i(0)=x.$$

Furthermore, owing to Lemma B.5,

$$p_i(s)\in D^+_y A_t(x,\xi_i(s)), \quad \forall s\in[0,t].$$

Therefore

$$\frac{1}{2}\frac{d}{ds}|\xi_2-\xi_1|^2 = \langle H_p(\xi_2,p_2)-H_p(\xi_1,p_1),\xi_2-\xi_1\rangle.$$

Integrating over $[s,t]$, we conclude that

$$\begin{aligned}|\xi_2(t)-\xi_1(t)|^2-|\xi_2(s)-\xi_1(s)|^2 \geqslant& -C_1\int_s^t \big(|\xi_2-\xi_1|^2+|p_2-p_1|\cdot|\xi_2-\xi_1|\big)d\tau\\ \geqslant& -C_1\int_s^t |p_2-p_1|^2 d\tau - 2C_1\int_s^t |\xi_2-\xi_1|^2 d\tau,\end{aligned}$$

where $C_1=C_1(\lambda)>0$ is an upper bound for $D^2H(x,p)$ on $\{(x,p)\ :\ |p|\leqslant\kappa(4\lambda)\}$. So,

$$\|\xi_2-\xi_1\|^2_{L^\infty(0,t)} \leqslant |y_2-y_1|^2 + C_1\int_s^t |p_2-p_1|^2 d\tau + 2C_1 t\|\xi_2-\xi_1\|^2_{L^\infty(0,t)}. \tag{B.11}$$

Now,

$$\begin{aligned}&\frac{d}{ds}\langle p_2-p_1,\xi_2-\xi_1\rangle\\ =&\langle p_2-p_1, H_p(\xi_2,p_2)-H_p(\xi_1,p_1)\rangle - \langle H_x(\xi_2,p_2)-H_x(\xi_1,p_1),\xi_2-\xi_1\rangle\\ =&\langle p_2-p_1, \widehat H_{px}(\xi_2-\xi_1)+\widehat H_{pp}(p_2-p_1)\rangle - \langle \widehat H_{xp}(p_2-p_0)+\widehat H_{xx}(\xi_2-\xi_1),\xi_2-\xi_1\rangle,\end{aligned}$$

where

$$\widehat H_{px}(s) = \int_0^1 H_{px}\big(\lambda\xi_2(s)+(1-\lambda)\xi_1(s),\lambda p_2(s)+(1-\lambda)p_1(s)\big)\,d\lambda,$$

and $\widehat{H}_{pp}$, $\widehat{H}_{xp}$, $\widehat{H}_{xx}$ are defined in a similar way. Thus, owing to (L1)-(L3), and since $(\widehat{H}_{px}(s))^* = \widehat{H}_{xp}(s)$ where $(\widehat{H}_{px})^*$ stands for the adjoint matrix, we have

$$\frac{d}{ds}\langle p_2 - p_1, \xi_2 - \xi_1\rangle \geqslant \nu|p_2 - p_1|^2 - C_2|\xi_2 - \xi_1|^2$$

for some positive constants $\nu = \nu(\lambda)$ and $C_2 = C_2(\lambda)$. Now, by (B.5),

$$\begin{aligned}\nu\int_0^t |p_2 - p_1|^2 ds &\leqslant C_2\int_0^t |\xi_2 - \xi_1|^2 ds + \langle p_2(t) - p_1(t), \xi_2(t) - \xi_1(t)\rangle\\ &\leqslant C_2\int_0^t |\xi_2 - \xi_1|^2 ds + \frac{C_3}{t}|y_2 - y_1|^2.\end{aligned}$$

Therefore,

$$\int_0^t |p_2 - p_1|^2 ds \leqslant \frac{C_2 t}{\nu}\|\xi_2 - \xi_1\|^2_{L^\infty(0,t)} + \frac{C_3}{\nu t}|y_2 - y_1|^2. \tag{B.12}$$

Combining (B.11) and (B.12), we obtain

$$\begin{aligned}\|\xi_2 - \xi_1\|^2_{L^\infty(0,t)} \leqslant{}& |y_2 - y_1|^2 + \frac{C_1C_2t}{\nu}\|\xi_2 - \xi_1\|^2_{L^\infty(0,t)} + \frac{C_1C_3}{\nu t}|y_2 - y_1|^2\\ &+ 2C_1t\|\xi_2 - \xi_1\|^2_{L^\infty(0,t)},\\ ={}& \Big(\frac{C_2}{\nu} + 2\Big)C_1t\|\xi_2 - \xi_1\|^2_{L^\infty(0,t)} + \Big(1 + \frac{C_1C_3}{\nu t}\Big)|y_2 - y_1|^2\end{aligned}$$

Then, taking

$$t_\lambda = \min\Big\{1\,,\ \frac{\nu}{2C_1(C_2 + 2\nu)}\Big\},$$

for all $t \in (0, t_\lambda)$ we conclude that

$$\|\xi_2 - \xi_1\|^2_{L^\infty(0,t)} \leqslant 2\Big(1 + \frac{C_1C_3}{\nu t}\Big)|y_2 - y_1|^2 \leqslant \frac{C}{t}|y_2 - y_1|^2.$$

This proves (B.8) and also (B.9) owing to (B.12). Finally, observing that

$$\begin{aligned}\int_0^t |\dot\xi_2 - \dot\xi_1|^2 ds &= \int_0^t |H_p(\xi_2, p_2) - H_p(\xi_1, p_1)|^2 ds\\ &\leqslant 2\int_0^t |H_p(\xi_2, p_2) - H_p(\xi_2, p_1)|^2 ds + \int_0^t |H_p(\xi_2, p_1) - H_p(\xi_1, p_1)|^2 ds\\ &\leqslant 2C_4\Big(\int_0^t |p_2 - p_1|^2 ds + \int_0^t |\xi_2 - \xi_1|^2 ds\Big)\end{aligned}$$

we obtain (B.10) by appealing to (B.8) and (B.9). $\square$

B.3. **Convexity of the fundamental solution for small time.** For any $t > 0$, $x, y, z \in \mathbb{R}^n$, and any $h \in [0, t)$, let $\xi_+ \in \Gamma^{t+h}_{x,y+z}$ and $\xi_- \in \Gamma^{t-h}_{x,y-z}$ be given. Define $\tilde\xi_\pm \in \Gamma^t_{x,y\pm z}$ by

$$\tilde\xi_+(\tau) = \xi_+\Big(\frac{t+h}{t}\tau\Big), \quad \tilde\xi_-(\tau) = \xi_-\Big(\frac{t-h}{t}\tau\Big), \quad \tau \in [0, t]. \tag{B.13}$$

Obviously,

$$\tilde{\xi}_+(0) = \tilde{\xi}_-(0) = x, \quad \tilde{\xi}_+(t) = y + z, \quad \tilde{\xi}_-(t) = y - z,$$

$$\frac{\tilde{\xi}_+ + \tilde{\xi}_-}{2}(0) = x, \quad \frac{\tilde{\xi}_+ + \tilde{\xi}_-}{2}(t) = y, \tag{B.14}$$

$$\frac{\tilde{\xi}_+ - \tilde{\xi}_-}{2}(0) = 0, \quad \frac{\tilde{\xi}_+ - \tilde{\xi}_-}{2}(t) = z. \tag{B.15}$$

Lemma B.7. *Suppose L is a Tonelli Lagrangian. For any $\lambda > 0$ let $t_\lambda > 0$ be given by Lemma B.6 and define $t'_\lambda = \min\{t_\lambda, 2/3\}$. Then there exist constants $C_\lambda, C'_\lambda > 0$ such that, for any $t \in (0, t'_\lambda)$, any $x \in \mathbb{R}^n$, any $y \in B(x, \lambda t)$, any $(h, z) \in [0, t/2) \times B(0, \lambda t)$, and any pair of minimizers, $\xi_\pm \in \Gamma^{t\pm h}_{x,y\pm z}$, for $A_{t\pm h}(x, y \pm z)$ we have*

$$\|\tilde{\xi}_+ - \tilde{\xi}_-\|^2_{L^\infty(0,t)} \leqslant \frac{C_\lambda}{t}(h^2 + |z|^2), \tag{B.16}$$

$$\int_0^t |\dot{\tilde{\xi}}_+ - \dot{\tilde{\xi}}_-|^2 \, d\tau \leqslant \frac{C'_\lambda}{t}(h^2 + |z|^2), \tag{B.17}$$

where $\tilde{\xi}_\pm \in \Gamma^t_{x,y\pm z}$ are defined in (B.13).

Proof. In view of Proposition B.1 and Remark B.2 we have that $\{(\xi_\pm(s), \dot{\xi}_\pm(s))\}_{s\in[0,t\pm h]}$ is contained in the convex compact set $\mathbf{K}_{x,\lambda}$. So, for any $\tau \in [0, t]$,

$$\begin{aligned}
&|\tilde{\xi}_+(\tau) - \tilde{\xi}_-(\tau)| = \Big|\xi_+\Big(\frac{t+h}{t}\tau\Big) - \xi_-\Big(\frac{t-h}{t}\tau\Big)\Big| \\
\leqslant&\Big|\xi_+\Big(\frac{t+h}{t}\tau\Big) - \xi_+\Big(\frac{t-h}{t}\tau\Big)\Big| + \Big|\xi_+\Big(\frac{t-h}{t}\tau\Big) - \xi_-\Big(\frac{t-h}{t}\tau\Big)\Big| \\
\leqslant&C_1 h + \Big|\xi_+\Big(\frac{t-h}{t}\tau\Big) - \xi_-\Big(\frac{t-h}{t}\tau\Big)\Big| \leqslant \kappa(4\lambda)h + \max_{s\in[0,t-h]} |\xi_+(s) - \xi_-(s)|.
\end{aligned}$$

Since

$$\begin{aligned}
|\xi_+(t-h) - \xi_-(t-h)| \leqslant&|\xi_+(t-h) - \xi_+(t+h)| + |\xi_+(t+h) - \xi_-(t-h)| \\
\leqslant&2\big(\kappa(4\lambda)h + |z|\big),
\end{aligned}$$

by (B.8) applied to ξ_+, ξ_- on $[0, t-h]$ we obtain

$$\max_{s\in[0,t-h]} |\xi_+(s) - \xi_-(s)|^2 \leqslant \frac{C_1}{t}(h^2 + |z|^2),$$

for some constant $C_1 = C_1(\lambda) > 0$. Similarly, we have

$$\begin{aligned}
&\big|\dot{\tilde{\xi}}_+(\tau) - \dot{\tilde{\xi}}_-(\tau)\big| = \Big|\frac{t+h}{t}\dot{\xi}_+\Big(\frac{t+h}{t}\tau\Big) - \frac{t-h}{t}\dot{\xi}_-\Big(\frac{t-h}{t}\tau\Big)\Big| \\
\leqslant&\Big|\frac{t+h}{t}\dot{\xi}_+\Big(\frac{t+h}{t}\tau\Big) - \frac{t+h}{t}\dot{\xi}_-\Big(\frac{t-h}{t}\tau\Big)\Big| + \Big|\frac{t+h}{t}\dot{\xi}_-\Big(\frac{t-h}{t}\tau\Big) - \frac{t-h}{t}\dot{\xi}_-\Big(\frac{t-h}{t}\tau\Big)\Big| \\
\leqslant&\frac{t+h}{t}\Big|\dot{\xi}_+\Big(\frac{t+h}{t}\tau\Big) - \dot{\xi}_-\Big(\frac{t-h}{t}\tau\Big)\Big| + \kappa(4\lambda)\frac{h}{t}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\Big|\dot\xi_+\Big(\frac{t+h}{t}\tau\Big)-\dot\xi_-\Big(\frac{t-h}{t}\tau\Big)\Big|\\
\leqslant&\Big|\dot\xi_+\Big(\frac{t+h}{t}\tau\Big)-\dot\xi_+\Big(\frac{t-h}{t}\tau\Big)\Big|+\Big|\dot\xi_+\Big(\frac{t-h}{t}\tau\Big)-\dot\xi_-\Big(\frac{t-h}{t}\tau\Big)\Big|\\
\leqslant&C_2\frac{h}{t}+\Big|\dot\xi_+\Big(\frac{t-h}{t}\tau\Big)-\dot\xi_-\Big(\frac{t-h}{t}\tau\Big)\Big|.
\end{aligned}$$

Thus, for $t\in(0,t'_\lambda)$ the bound in (B.10) yields

$$\begin{aligned}
\int_0^t|\dot{\tilde\xi}_+(\tau)-\dot{\tilde\xi}_-(\tau)|^2\,d\tau\leqslant&C_3\frac{h^2}{t}+\frac{(t+h)^2}{t^2}\int_0^t\Big|\dot\xi_+\Big(\frac{t-h}{t}\tau\Big)-\dot\xi_-\Big(\frac{t-h}{t}\tau\Big)\Big|^2\,d\tau\\
=&C_3\frac{h^2}{t}+\frac{(t+h)^2}{t^2}\cdot\frac{t}{t-h}\int_0^{t-h}|\dot\xi_+(s)-\dot\xi_-(s)|^2\,ds\\
\leqslant&C_3\frac{h^2}{t}+\frac{(t+h)^2}{t^2}\cdot\frac{t}{t-h}\cdot\frac{C_4}{t-h}(h^2+|z|^2).
\end{aligned}$$

This leads to our conclusion. □

Proposition B.8. *Suppose L is a Tonelli Lagrangian and, for any $\lambda>0$, let $t'_\lambda>0$ be the number given by Lemma B.7. Then, for any $x\in\mathbb{R}^n$, the function $(t,y)\mapsto A_t(x,y)$ is semiconvex on the cone*

$$S_\lambda(x,t'_\lambda):=\big\{(t,y)\in\mathbb{R}\times\mathbb{R}^n\ :\ 0<t<t'_\lambda,\ |y-x|<\lambda t\big\}\,,\tag{B.18}$$

and there exists a constant $C''_\lambda>0$ such that for all $(t,y)\in S_\lambda(x,t'_\lambda)$, all $h\in[0,t/2)$, and all $z\in B(0,\lambda t)$ we have that

$$A_{t+h}(x,y+z)+A_{t-h}(x,y-z)-2A_t(x,y)\geqslant-\frac{C''_\lambda}{t}(h^2+|z|^2).\tag{B.19}$$

Moreover, there exists $t''_\lambda\in(0,t'_\lambda]$ and $C'''_\lambda>0$ such that for all $t\in(0,t''_\lambda]$ the function $A_t(x,\cdot)$ is uniformly convex on $B(x,\lambda t)$ and for all $y\in B(x,\lambda t)$ and $z\in B(0,\lambda t)$ we have that

$$A_t(x,y+z)+A_t(x,y-z)-2A_t(x,y)\geqslant\frac{C'''_\lambda}{t}|z|^2.\tag{B.20}$$

Proof. Let $x\in\mathbb{R}^n$ and fix $(t,y)\in S_\lambda(x,t'_\lambda)$, $h\in[0,t/2)$, and $z\in B(0,\lambda t)$. Let $\xi_+\in\Gamma^{t+h}_{x,y+z}$ and $\xi_-\in\Gamma^{t-h}_{x,y-z}$ be minimizers for $A_{t+h}(x,y+z)$ and $A_{t-h}(x,y-z)$ respectively, and define $\tilde\xi_\pm$ as in (B.13). In view of Proposition B.1 and Remark B.2 we have that $\{(\xi_\pm(s),\dot\xi_\pm(s))\}_{s\in[0,t\pm h]}$ and $\{(\tilde\xi_\pm(s),\dot{\tilde\xi}_\pm(s))\}_{s\in[0,t]}$ are all contained in the convex compact set $\mathbf{K}_{x,\lambda}$. Moreover

$$\begin{aligned}
&A_{t+h}(x,y+z)+A_{t-h}(x,y-z)\\
=&\int_0^{t+h}L(\xi_+(s),\dot{\xi}_+(s))\,ds+\int_0^{t-h}L(\xi_-(s),\dot{\xi}_-(s))\,ds\\
=&\frac{t+h}{t}\int_0^t L\Big(\xi_+\Big(\frac{t+h}{t}\tau\Big),\dot{\xi}_+\Big(\frac{t+h}{t}\tau\Big)\Big)\,d\tau\\
&+\frac{t-h}{t}\int_0^t L\Big(\xi_-\Big(\frac{t-h}{t}\tau\Big),\dot{\xi}_-\Big(\frac{t-h}{t}\tau\Big)\Big)\,d\tau\\
=&\frac{t+h}{t}\int_0^t L\Big(\tilde{\xi}_+(\tau),\frac{t}{t+h}\dot{\tilde{\xi}}_+(\tau)\Big)\,d\tau+\frac{t-h}{t}\int_0^t L\Big(\tilde{\xi}_-(\tau),\frac{t}{t-h}\dot{\tilde{\xi}}_-(\tau)\Big)\,d\tau\\
=&\int_0^t L\Big(\tilde{\xi}_+(\tau),\frac{t}{t+h}\dot{\tilde{\xi}}_+(\tau)\Big)\,d\tau+\int_0^t L\Big(\tilde{\xi}_-(\tau),\frac{t}{t-h}\dot{\tilde{\xi}}_-(\tau)\Big)\,d\tau+I_1,
\end{aligned}$$

where

$$I_1=\frac{h}{t}\int_0^t\Big\{L\Big(\tilde{\xi}_+(\tau),\frac{t}{t+h}\dot{\tilde{\xi}}_+(\tau)\Big)-L\Big(\tilde{\xi}_-(\tau),\frac{t}{t-h}\dot{\tilde{\xi}}_-(\tau)\Big)\Big\}\,d\tau.$$

Set

$$\begin{aligned}
I_2=&\int_0^t L(\tilde{\xi}_+(\tau),\dot{\tilde{\xi}}_+(\tau))\,d\tau+\int_0^t L(\tilde{\xi}_-(\tau),\dot{\tilde{\xi}}_-(\tau))\,d\tau-2A_t(x,y),\\
I_3=&\int_0^t L\Big(\tilde{\xi}_+(\tau),\frac{t}{t+h}\dot{\tilde{\xi}}_+(\tau)\Big)\,d\tau-\int_0^t L(\tilde{\xi}_+(\tau),\dot{\tilde{\xi}}_+(\tau))\,d\tau\\
&+\int_0^t L\Big(\tilde{\xi}_-(\tau),\frac{t}{t-h}\dot{\tilde{\xi}}_-(\tau)\Big)\,d\tau-\int_0^t L(\tilde{\xi}_-(\tau),\dot{\tilde{\xi}}_-(\tau))\,d\tau.
\end{aligned}$$

Then

$$A_{t+h}(x,y+z)+A_{t-h}(x,y-z)-2A_t(x,y)=I_1+I_2+I_3.$$

Now we turn to the estimates of I_1, I_2 and I_3.

Estimate of I_1: Let $C_0=C_0(\lambda)>0$ be an upper bound for $|L_v|$ on $\mathbf{K}_{x,\lambda}$. Then

$$\begin{aligned}
I_1=&\frac{h}{t}\int_0^t\Big\{L\Big(\tilde{\xi}_+,\frac{t}{t+h}\dot{\tilde{\xi}}_+\Big)-L\Big(\tilde{\xi}_+,\frac{t}{t-h}\dot{\tilde{\xi}}_+\Big)\Big\}\,d\tau\\
&+\frac{h}{t}\int_0^t\Big\{L\Big(\tilde{\xi}_+,\frac{t}{t-h}\dot{\tilde{\xi}}_+\Big)-L\Big(\tilde{\xi}_+,\frac{t}{t-h}\dot{\tilde{\xi}}_-\Big)\Big\}\,d\tau\\
&+\frac{h}{t}\int_0^t\Big\{L\Big(\tilde{\xi}_+,\frac{t}{t-h}\dot{\tilde{\xi}}_-\Big)-L\Big(\tilde{\xi}_-,\frac{t}{t-h}\dot{\tilde{\xi}}_-\Big)\Big\}\,d\tau\\
\geqslant&-C_0\frac{th^2}{t^2-h^2}-C_0\frac{h}{t-h}\int_0^t|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-|d\tau-C_0\frac{h}{t}\int_0^t|\tilde{\xi}_+-\tilde{\xi}_-|d\tau
\end{aligned}$$

By (B.17) it is easy to check that

$$\int_0^t|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-|d\tau\leqslant\Big(\int_0^t|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-|^2d\tau\Big)^{\frac12}t^{\frac12}\leqslant\sqrt{C'_\lambda(h^2+|z|^2)}.$$

Similarly, by (B.16),

$$\int_0^t|\tilde{\xi}_+-\tilde{\xi}_-|d\tau\leqslant\sqrt{C_\lambda t(h^2+|z|^2)}.$$

Thus, recalling that $h \in [0, t/2)$ and $z \in B(0, \lambda t)$, we conclude that

$$\begin{aligned} I_1 \;&\geqslant\; -\frac{2C_0}{t}h^2 - \frac{2C_0}{t}h\sqrt{C'_\lambda(h^2+|z|^2)} - \frac{C_0}{t}h\sqrt{C_\lambda t(h^2+|z|^2)} \qquad \text{(B.21)}\\ &\geqslant\; -\frac{C_1}{t}(h^2+|z|^2), \qquad \text{(B.22)} \end{aligned}$$

for some constant $C_1 = C_1(\lambda) > 0$.

Estimate of I_2: Let $\nu_\lambda := \nu\big(\kappa(4\lambda)\big) > 0$ be the lower bound for L_{vv} on $\mathbf{K}_{x,\lambda}$ provided by assumption (L1). Then

$$\begin{aligned} I_2 =& \int_0^t L(\tilde\xi_+(\tau),\dot{\tilde\xi}_+(\tau))\,d\tau + \int_0^t L(\tilde\xi_-(\tau),\dot{\tilde\xi}_-(\tau))\,d\tau - 2A_t(x,y),\\ \geqslant& \int_0^t \Big\{L(\tilde\xi_+,\dot{\tilde\xi}_+) + L(\tilde\xi_-,\dot{\tilde\xi}_-) - 2L\Big(\frac{\tilde\xi_++\tilde\xi_-}{2},\frac{\dot{\tilde\xi}_++\dot{\tilde\xi}_-}{2}\Big)\Big\}\,d\tau\\ =& \int_0^t \Big\{L\Big(\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_+\Big) + L\Big(\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_-\Big) - 2L\Big(\frac{\tilde\xi_++\tilde\xi_-}{2},\frac{\dot{\tilde\xi}_++\dot{\tilde\xi}_-}{2}\Big)\Big\}\,d\tau\\ &+ \int_0^t \Big\{L(\tilde\xi_+,\dot{\tilde\xi}_+) - L\Big(\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_+\Big)\Big\}d\tau + \int_0^t \Big\{L(\tilde\xi_-,\dot{\tilde\xi}_-) - L\Big(\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_-\Big)\Big\}d\tau\\ \geqslant& \nu_\lambda\int_0^t \Big|\frac{\dot{\tilde\xi}_+-\dot{\tilde\xi}_-}{2}\Big|^2\,d\tau + \int_0^t\int_0^1 \Big\langle L_x\Big(\lambda\tilde\xi_+ + (1-\lambda)\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_+\Big),\frac{\tilde\xi_+-\tilde\xi_-}{2}\Big\rangle\,d\lambda d\tau\\ &+ \int_0^t\int_0^1 \Big\langle L_x\Big(\lambda\tilde\xi_- + (1-\lambda)\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_-\Big),-\frac{\tilde\xi_+-\tilde\xi_-}{2}\Big\rangle\,d\lambda d\tau \end{aligned}$$

Setting

$$\widehat{L_x}(\lambda,\tau) = L_x\Big(\lambda\tilde\xi_+ + (1-\lambda)\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_+\Big) - L_x\Big(\lambda\tilde\xi_- + (1-\lambda)\frac{\tilde\xi_++\tilde\xi_-}{2},\dot{\tilde\xi}_-\Big),$$

we have that

$$\begin{aligned} I_2 \geqslant& \nu_\lambda\int_0^t \Big|\frac{\dot{\tilde\xi}_+-\dot{\tilde\xi}_-}{2}\Big|^2\,d\tau + \int_0^t\int_0^1 \big\langle\widehat{L_x},\frac{\tilde\xi_+-\tilde\xi_-}{2}\big\rangle\,d\lambda d\tau\\ \geqslant& \nu_\lambda\int_0^t \Big|\frac{\dot{\tilde\xi}_+-\dot{\tilde\xi}_-}{2}\Big|^2\,d\tau - C_2\int_0^t \big(|\tilde\xi_+-\tilde\xi_-|^2 + |\dot{\tilde\xi}_+-\dot{\tilde\xi}_-|\cdot|\tilde\xi_+-\tilde\xi_-|\big)\,d\tau, \qquad \text{(B.23)} \end{aligned}$$

where $C_2 = C_2(\lambda) > 0$ is such that

$$|L_v|,|L_{xx}|,|L_{xv}|,|L_{vv}| \leqslant C_2 \qquad \text{on} \quad \mathbf{K}_{x,\lambda}. \qquad \text{(B.24)}$$

Estimate of I_3: As above, let $\nu_\lambda = \nu\big(\kappa(4\lambda)\big) > 0$. Then

$$\begin{aligned} I_3 =& \int_0^t L\Big(\tilde\xi_+(\tau),\frac{t}{t+h}\dot{\tilde\xi}_+(\tau)\Big)\,d\tau - \int_0^t L(\tilde\xi_+(\tau),\dot{\tilde\xi}_+(\tau))\,d\tau\\ &+ \int_0^t L\Big(\tilde\xi_-(\tau),\frac{t}{t-h}\dot{\tilde\xi}_-(\tau)\Big)\,d\tau - \int_0^t L(\tilde\xi_-(\tau),\dot{\tilde\xi}_-(\tau))\,d\tau\\ \geqslant& \int_0^t \Big\{\big\langle L_v(\tilde\xi_+,\dot{\tilde\xi}_+),-\frac{h}{t+h}\dot{\tilde\xi}_+\big\rangle + \frac{\nu_\lambda h^2}{|t+h|^2}|\dot{\tilde\xi}_+|^2\Big\}\,d\tau\\ &+ \int_0^t \Big\{\big\langle L_v(\tilde\xi_-,\dot{\tilde\xi}_-),\frac{h}{t-h}\dot{\tilde\xi}_-\big\rangle + \frac{\nu_\lambda h^2}{|t-h|^2}|\dot{\tilde\xi}_-|^2\Big\}d\tau. \end{aligned} \qquad \text{(B.25)}$$

Since $2\big(|\dot{\tilde{\xi}}_+|^2+|\dot{\tilde{\xi}}_-|^2\big)\geqslant\big|\dot{\tilde{\xi}}_+\pm\dot{\tilde{\xi}}_-|^2$ and

$$\frac{1}{|t\pm h|^2}\geqslant\Big(\frac{2}{3t}\Big)^2,$$

(B.25) yields

$$\begin{aligned}I_3\geqslant{}&\nu_\lambda h^2\Big(\frac{2}{3t}\Big)^2\Big\{\int_0^t\Big|\frac{\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-}{2}\Big|^2\,d\tau+\int_0^t\Big|\frac{\dot{\tilde{\xi}}_++\dot{\tilde{\xi}}_-}{2}\Big|^2\,d\tau\Big\}\\&+\int_0^t\Big\{\big\langle L_v(\tilde{\xi}_-,\dot{\tilde{\xi}}_-),\frac{h}{t-h}\dot{\tilde{\xi}}_-\big\rangle-\big\langle L_v(\tilde{\xi}_+,\dot{\tilde{\xi}}_+),\frac{h}{t+h}\dot{\tilde{\xi}}_+\big\rangle\Big\}d\tau:=I_4+I_5\end{aligned}\tag{B.26}$$

Now, since $\tilde{\xi}_+-\tilde{\xi}_-$ is an arc connecting 0 to z, comparison with $s\mapsto\frac{s}{t}z$ yields

$$\int_0^t\big|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-\big|^2\,d\tau\geqslant\frac{|z|^2}{t}.\tag{B.27}$$

Similarly,

$$\int_0^t\big|\dot{\tilde{\xi}}_++\dot{\tilde{\xi}}_-\big|^2\,d\tau\geqslant\frac{|y-x|^2}{t}.$$

So,

$$I_4\geqslant\frac{\nu_\lambda h^2}{9t^3}(|z|^2+|y-x|^2).\tag{B.28}$$

As for I_5, we have

$$\begin{aligned}I_5={}&\frac{h}{t-h}\int_0^t\big\langle L_v(\tilde{\xi}_-,\dot{\tilde{\xi}}_-),\dot{\tilde{\xi}}_--\dot{\tilde{\xi}}_+\big\rangle d\tau+\Big(\frac{h}{t-h}-\frac{h}{t+h}\Big)\int_0^t\big\langle L_v(\tilde{\xi}_-,\dot{\tilde{\xi}}_-),\dot{\tilde{\xi}}_+\big\rangle d\tau\\&+\frac{h}{t+h}\int_0^t\big\langle L_v(\tilde{\xi}_-,\dot{\tilde{\xi}}_-)-L_v(\tilde{\xi}_-,\dot{\tilde{\xi}}_+),\dot{\tilde{\xi}}_+\big\rangle\,d\tau\\&+\frac{h}{t+h}\int_0^t\big\langle L_v(\tilde{\xi}_-,\dot{\tilde{\xi}}_+)-L_v(\tilde{\xi}_+,\dot{\tilde{\xi}}_+),\dot{\tilde{\xi}}_+\big\rangle\,d\tau.\end{aligned}$$

Thus, by (B.24) and Lemma B.7 we have

$$\begin{aligned}I_5\geqslant{}&-C_2\frac{2h}{t}\int_0^t|\dot{\tilde{\xi}}_--\dot{\tilde{\xi}}_+|\,d\tau-C_2\kappa(4\lambda)\frac{2h^2t}{t^2-h^2}\\&-C_2\kappa(4\lambda)\frac{h}{t+h}\Big\{\int_0^t|\dot{\tilde{\xi}}_--\dot{\tilde{\xi}}_+|\,d\tau+\int_0^t|\tilde{\xi}_--\tilde{\xi}_+|\,d\tau\Big\}\\\geqslant{}&-4C_2\kappa(4\lambda)\frac{h^2}{t}-C_2\big[2+\kappa(4\lambda)\big]\frac{h}{\sqrt{t}}\Big(\int_0^t|\dot{\tilde{\xi}}_--\dot{\tilde{\xi}}_+|^2\,d\tau\Big)^{\frac12}-C_2\kappa(4\lambda)h\sqrt{\frac{C_\lambda}{t}(h^2+|z|^2)}\\\geqslant{}&-C_3\frac{h}{t}(h+|z|)\end{aligned}$$

for some constant $C_3=C_3(\lambda)>0$. By the last inequality, (B.26), and (B.28) we get

$$I_3\geqslant\frac{\nu_\lambda h^2}{9t^3}(|z|^2+|y-x|^2)-C_3\frac{h}{t}(h+|z|).\tag{B.29}$$

We are now ready to prove (B.19). Indeed, by combining (B.22), (B.23), and (B.29) we obtain, thanks to Lemma B.7,

$$\begin{aligned}
&A_{t+h}(x,y+z)+A_{t-h}(x,y-z)-2A_t(x,y)\geqslant -\frac{C_1}{t}(h^2+|z|^2)\\
&\qquad -C_2\int_0^t\left(\left|\tilde{\xi}_+-\tilde{\xi}_-\right|^2+\left|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-\right|\cdot\left|\tilde{\xi}_+-\tilde{\xi}_-\right|\right)d\tau-C_3\frac{h}{t}(h+|z|)\\
&\geqslant\ -\frac{C_1}{t}(h^2+|z|^2)-\frac{C_2}{2}\Big(3C_\lambda+\frac{C'_\lambda}{t}\Big)(h^2+|z|^2)-\frac{C_3}{t}\Big(\frac{3}{2}h^2+|z|^2\Big).
\end{aligned}$$

Finally, in order to prove (B.20) observe that taking $h=0$ in (B.21) and (B.29) we conclude that $I_1, I_3\geqslant 0$. Therefore, for any $\varepsilon>0$, (B.23) yields the lower bound

$$\begin{aligned}
&A_t(x,y+z)+A_t(x,y-z)-2A_t(x,y)\\
\geqslant\ &\frac{\nu_\lambda}{4}\int_0^t\left|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-\right|^2d\tau-C_2\int_0^t\left(\left|\tilde{\xi}_+-\tilde{\xi}_-\right|^2+\left|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-\right|\cdot\left|\tilde{\xi}_+-\tilde{\xi}_-\right|\right)d\tau\\
\geqslant\ &\Big(\frac{\nu_\lambda}{4}-\frac{\varepsilon}{2}\Big)\int_0^t\left|\dot{\tilde{\xi}}_+-\dot{\tilde{\xi}}_-\right|^2d\tau-\Big(C_2+\frac{C_2^2}{2\varepsilon}\Big)\int_0^t\left|\tilde{\xi}_+-\tilde{\xi}_-\right|^2d\tau.
\end{aligned}$$

Hence, taking $\varepsilon=\nu_\lambda/4$, recalling (B.27), and appealing to Lemma B.7 it follows that

$$A_t(x,y+z)+A_t(x,y-z)-2A_t(x,y)\geqslant\Big\{\frac{\nu_\lambda}{8t}-\Big(C_2+\frac{2C_2^2}{\nu_\lambda}\Big)C_\lambda\Big\}|z|^2.$$

Now, choosing $t''_\lambda\in(0,t'_\lambda]$ such that

$$\frac{\nu_\lambda}{8t''_\lambda}>\Big(C_2+\frac{2C_2^2}{\nu_\lambda}\Big)C_\lambda$$

one completes the proof. □

B.4. $C^{1,1}_{loc}$ **regularity of the fundamental solution.**

Proposition B.9. *Suppose L is a Tonelli Lagrangian and, for any $\lambda>0$, let $t'_\lambda>0$ be the number given by Lemma B.7.*

Then, for any $x\in\mathbb{R}^n$ the functions $(t,y)\mapsto A_t(x,y)$ and $(t,y)\mapsto A_t(y,x)$ are of class $C^{1,1}_{loc}$ on the cone $S_\lambda(x,t'_\lambda)$ defined in (B.18)*. Moreover, for all $(t,y)\in S(x,t'_\lambda)$*

$$D_yA_t(x,y)=L_v(\xi(t),\dot{\xi}(t)),\tag{B.30}$$

$$D_xA_t(x,y)=-L_v(\xi(0),\dot{\xi}(0)),\tag{B.31}$$

$$D_tA_t(x,y)=-E_{t,x,y},\tag{B.32}$$

where $\xi\in\Gamma^t_{x,y}$ is the unique minimizer for $A_t(x,y)$ and

$$E_{t,x,y}:=H(\xi(s),p(s))\qquad\forall\, s\in[0,t]$$

is the energy of the Hamiltonian trajectory (ξ,p) with

$$p(s)=L_v(\xi(s),\dot{\xi}(s)).\tag{B.33}$$

Proof. $C^{1,1}$-regularity on $S(x,t'_\lambda)$ is a corollary of propositions B.3 and B.8. (B.30) follows from Lemma B.5 and (B.31) can be proved by a similar argument.

For the proof of (B.32), first, we define for $h>0$ small enough

$$\xi_+(\tau)=\xi\Big(\frac{t}{t+h}\tau\Big),\quad \tau\in[0,t+h].$$

Then

$$
\begin{aligned}
A_{t+h}(x,y)-A_t(x,y)\leqslant&\int_0^{t+h}L(\xi_+(\tau),\dot\xi_+(\tau))\,d\tau-\int_0^t L(\xi(s),\dot\xi(s))\,ds\\
=&\frac{t+h}{t}\int_0^t L\Big(\xi(s),\frac{t}{t+h}\dot\xi(s)\Big)-\int_0^t L(\xi(s),\dot\xi(s))\,ds.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\limsup_{h\to 0}\frac{A_{t+h}(x,y)-A_t(x,y)}{h}\\
\leqslant&\lim_{h\to 0}\frac{1}{t}\int_0^t L\Big(\xi(s),\frac{t}{t+h}\dot\xi(s)\Big)ds\\
&+\lim_{h\to 0}\frac{1}{h}\int_0^t\Big\{L\Big(\xi(s),\frac{t}{t+h}\dot\xi(s)\Big)-L(\xi(s),\dot\xi(s))\Big\}ds\\
=&\frac{1}{t}\int_0^t\Big\{L(\xi(s),\dot\xi(s))-\big\langle L_v(\xi(s),\dot\xi(s)),\dot\xi(s)\big\rangle\Big\}ds=-\frac{1}{t}\int_0^t H(\xi(s),p(s))\,ds,
\end{aligned}
$$

where $p(\cdot)$ is given by (B.33). The study of the function $(t,y)\mapsto A_t(y,x)$ is similar. □

APPENDIX C. $\mathbf{y}$ IS A GENERALIZED CHARACTERISTIC

The existence of singular generalized characteristics was first proved in [4] (see also [16]) in a constructive way, and then a simplified proof was given in [37] using an approximation method. Here, in order to prove that the curve $\mathbf{y}(t)$, $t\in[0,t_0]$, in Lemma 3.2 is a generalized characteristic, we follow the idea of the original proof from [4]. We sketch the proof for completeness. The following result is an analogy to Lemma 5.5.6 in [16].

Lemma C.1. *Let L be a Tonelli Lagrangian, and $t_0\in(0,1]$ be given by Lemma 3.2. For any fixed $x\in\mathbb{R}^n$, let $\mathbf{y}:[0,t_0]\to\mathbb{R}^n$ be the curve constructed in Lemma 3.2, and let $\mathbf{p}:[0,t_0]\to\mathbb{R}^n$ be the arc defined in Proposition 3.4. Then, for any $\varepsilon>0$, there exist arcs $\mathbf{x}_\varepsilon:[0,t_0]\to\mathbb{R}^n$, $\mathbf{p}_\varepsilon:[0,t_0]\to\mathbb{R}^n$, with $\mathbf{x}_\varepsilon(0)=x$ and $\mathbf{p}_\varepsilon(0)=p_x$ where p_x is the unique element in $\arg\min_{p\in D^+u(x)}H(x,p)$, and a partition $0=s_0<s_1<\cdot<s_{k-1}<s_k=t_0$ with the following properties:*

(i) $\max\{|s_{j+1}-s_j|:0\leqslant j\leqslant k-1\}<\varepsilon$;
(ii) $\mathbf{p}_\varepsilon(s)\in D^+u(\mathbf{x}_\varepsilon(s))$ *for all* $s\in[0,t_0]$;
(iii) *For each* $j\in\{0,\dots,k-1\}$, $\mathbf{p}_\varepsilon(\cdot)$ *is continuous on* $[s_j,s_{j+1})$;
(iv) $|\mathbf{x}_\varepsilon(t)-\mathbf{x}_\varepsilon(s)|\leqslant C_1|t-s|$ *for all* $t,s\in[0,t_0]$;
(v) $|\mathbf{p}_\varepsilon(s)-\mathbf{p}_\varepsilon(s')|\leqslant C_2\sqrt{\varepsilon}$ *for all* $s,s'\in[s_j,s_{j+1})$, $j\in\{0,\dots,k-1\}$.
(vi) *The number of the nodes satisfies* $k\leqslant[\frac{C_3}{\varepsilon}]+1$.

Proof. The construction of the arcs $\mathbf{x}_\varepsilon$ and $\mathbf{p}_\varepsilon$ follows the reasoning of Lemma 5.5.6 in [16]. The proof of properties (i)-(vi) also goes as in [16], except for the use of the essential inequality (5.64) in [16], which can be replaced by property (d) in Proposition 3.4. □

Lemma C.2. *Let the arcs $\mathbf{x}_\varepsilon$ and $\mathbf{p}_\varepsilon$ be defined as in Lemma C.1. Then there exists a constant $C>0$, independent of ε, such that*

$$
\left|\mathbf{x}_\varepsilon(r)-\mathbf{x}_\varepsilon(s)-\int_s^r H_p(\mathbf{x}_\varepsilon(\tau),\mathbf{p}_\varepsilon(\tau))\,d\tau\right|\leqslant C\sqrt{\varepsilon}. \tag{C.1}
$$

Proof. In view of property (vi) in Lemma C.1, it is sufficient to prove (C.1) under the extra assumption that $s = 0$ and $r < s_1$. Thus, we have

$$\begin{aligned} &\left|\mathbf{x}_\varepsilon(r) - \mathbf{x}_\varepsilon(s) - \int_s^r H_p(\mathbf{x}_\varepsilon(\tau), \mathbf{p}_\varepsilon(\tau))\, d\tau\right| = \left|\mathbf{y}(r) - \mathbf{y}(0) - \int_0^r H_p(\mathbf{y}(\tau), \mathbf{p}(\tau))\, d\tau\right| \\ =&\left|\xi_r(r) - x - \int_0^r H_p(\xi_\tau(\tau), p_\tau(\tau))\, d\tau\right| = \left|\int_0^r \dot{\xi}_r(\tau) - H_p(\xi_\tau(\tau), p_\tau(\tau))\, d\tau\right| \\ =&\left|\int_0^r H_p(\xi_r(\tau), p_r(\tau)) - H_p(\xi_\tau(\tau), p_\tau(\tau))\, d\tau\right|, \end{aligned}$$

where $\xi_t \in \Gamma^t_{x,\mathbf{y}(t)}$ is a minimizer for $A_t(x, \mathbf{y}(t))$, $t \in (0, s_1)$ and $p_t(\tau) := L_v(\xi_t(\tau), \dot{\xi}_t(\tau))$, $\tau \in [0, t]$. Then, by Lemma 3.3, Proposition 3.4 (a) and Lemma C.1, we have

$$\begin{aligned} &|H_p(\xi_r(\tau), p_r(\tau)) - H_p(\xi_\tau(\tau), p_\tau(\tau))| \\ \leqslant&|H_p(\xi_r(\tau), p_r(\tau)) - H_p(\xi_r(\tau), p_\tau(\tau))| + |H_p(\xi_r(\tau), p_\tau(\tau)) - H_p(\xi_\tau(\tau), p_\tau(\tau))| \\ \leqslant&C_1|p_r(\tau)) - p_\tau(\tau))| + C_2|\xi_r(\tau) - \xi_\tau(\tau)| \\ \leqslant&C_1(|p_r(\tau)) - p_r(r))| + |p_r(r)) - p_\tau(\tau))|) + C_2(|\xi_r(\tau) - \xi_r(r)| + |\xi_r(r) - \xi_\tau(\tau)|) \\ \leqslant&C_1(C_3\varepsilon + |\mathbf{p}(r) - \mathbf{p}(\tau)|) + C_2(C_4\varepsilon + |\mathbf{y}(r') - \mathbf{y}(\tau)|) \\ \leqslant&C_5\varepsilon + C_6\sqrt{\varepsilon} \leqslant C_7\sqrt{\varepsilon}, \end{aligned}$$

which leads to (C.1). □

The rest of the proof is standard, see, e.g., [4] or [16]. As $\varepsilon \to 0$ in (C.1), we obtain

$$\dot{\mathbf{y}}(s) \in \operatorname{co} H_p(\mathbf{y}(s), D^+u(\mathbf{y}(s))).$$

We omit the rest of the proof. The reader can refer to, for instance, [16, Page 133-135].

Appendix D. Global viscosity solutions on $\mathbb{R}^n$

In this section we prove Proposition 2.7.

Proof. The first part of the conclusion, that is, the fact that there exists a constant $c(H) \in \mathbb{R}$ such that the Hamilton-Jacobi equation (2.9) admits a viscosity solution $u : \mathbb{R}^n \to \mathbb{R}$ for $c = c(H)$ and does not admit any such solution for $c < c(H)$ is guaranteed by Theorem 1.1 in [26]. Moreover, in view of Proposition 4.1 in [26], we have that

$$u = T_t^- u + c(H)t \qquad \forall\, t \geqslant 0,$$

where T_t^- is defined in (3.6). Therefore, u is Lipschitz continuous on $\mathbb{R}^n$ on account of Proposition 3.2 in [26].

We proceed to show that u is also semiconcave. Let $A_t(x, y)$ be the fundamental solution of (2.9) with $c = c(H)$ and fix $t_0 \in (0, 2/3)$. For every $x \in \mathbb{R}^n$ we have that

$$u(x) = \min_{y \in \mathbb{R}^n} \left\{u(y) + A_{t_0}(y, x)\right\} + c(H)t_0.$$

Let $y_x \in \mathbb{R}^n$ be a point at which the above minimum is attained. Then taking $\lambda = 1$ in Proposition B.3 we conclude that, for all $y \in B(y_x, t_0)$ and all $z \in B(0, t_0)$,

$$A_{t_0}(y_x, y + z) + A_{t_0}(y_x, y - z) - 2A_{t_0}(y_x, y) \leqslant \frac{C_1}{t_0}|z|^2$$

for some constant $C_1 > 0$ independent of y_x. Therefore, taking $y = x$ in the above inequality we obtain

$$\begin{aligned} & u(x+z)+u(x-z)-2u(x) \\ \leqslant \quad & A_{t_0}(y_x,x+z)+A_{t_0}(y_x,x-z)-2A_{t_0}(y_x,x) \leqslant \frac{C_1}{t_0}|z|^2 \end{aligned}$$

all $z \in B(0,t_0)$. So, being Lipschitz, u is semiconcave on $\mathbb{R}^n$. □

REFERENCES

[1] Albano, P., *Propagation of singularities for solutions of Hamilton-Jacobi equations*. J. Math. Anal. Appl. **411**(2014), no. 2, 684–687.

[2] Albano, P., *The regularity of the distance function propagates along minimizing geodesics*. Nonlinear Anal. **95**(2014), 308–312.

[3] Albano, P.; Cannarsa, P., *Structural properties of singularities of semiconcave functions.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28**(1999), no. 4, 719–740.

[4] Albano, P., Cannarsa, P., *Propagation of singularities for solutions of nonlinear first order partial differential equations*, Arch. Ration. Mech. Anal., **162**(2002), 1–23.

[5] Albano, P., Cannarsa, P., Nguyen, K.T. and Sinestrari, C., *Singular gradient flow of the distance function and homotopy equivalence*, Math. Ann., **356**(2013), 23-43.

[6] Ambrosio, L.; Ascenzi, O.; Buttazzo, G., *Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands*. J. Math. Anal. Appl. **142**(1989), no. 2, 301–316.

[7] Ambrosio, L., Cannarsa, P. and Soner, H. M., *On the propagation of singularities of semi-convex functions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **20**(1993), 597–616.

[8] Attouch, H. Variational convergence for functions and operators. *Applicable Mathematics Series*. Pitman, 1984.

[9] Bernard, P., *Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds*. Ann. Sci. École Norm. Sup. (4) 40 (2007), no. 3, 445–452.

[10] Bernard, P., *The dynamics of pseudographs in convex Hamiltonian systems*. J. Amer. Math. Soc. 21 (2008), no. 3, 615–669.

[11] Bernard, P., The Lax-Oleinik semi-group: a Hamiltonian point of view. Proc. Roy. Soc. Edinburgh Sect. A **142**(2012), no. 6, 1131–1177.

[12] Cannarsa, P.; Cheng, W., *Homoclinic orbits and critical points of barrier functions*. Nonlinearity **28** (2015), no. 6, 1823–1840.

[13] Cannarsa, P., Cheng, W., *Dynamic and asymptotic behavior of singularities of weak KAM solutions on the torus*, preprint, 2016.

[14] Cannarsa, P.; Cheng, W.; Zhang, Q., *Propagation of singularities for weak KAM solutions and barrier functions*. Comm. Math. Phys. **331**(2014), no. 1, 1–20.

[15] Cannarsa, P.; Mazzola, M.; Sinestrari, C., *Global Propagation of Singularities for Time Dependent Hamilton-Jacobi Equation*. Discrete Contin. Dyn. Syst. Ser. A **38**(2015), no. 9, 441–469

[16] Cannarsa, P.; Sinestrari, C., Semiconcave functions, Hamilton-Jacobi equations, and optimal control. *Progress in Nonlinear Differential Equations and their Applications, 58*, Birkhäuser Boston, Inc., Boston, MA, 2004.

[17] Cannarsa, P.; Soner, H. M. *On the singularities of the viscosity solutions to Hamilton-Jacobi-Bellman equations*. Indiana Univ. Math. J. **36**(1987), no. 3, 501–524.

[18] Cannarsa, P.; Yu, Y. *Singular dynamics for semiconcave functions*. J. Eur. Math. Soc. **11**(2009), no. 5, 999–1024.

[19] Chen, C.; Cheng, W., *Lasry-Lions, Lax-Oleinik and generalized characteristics*, Sci. China Math., to appear.

[20] Cheng, C.-Q., *Uniform hyperbolicity of invariant cylinder*. arXiv: 1509.03160, preprint, 2015.

[21] Cheng, C.-Q., *A way to cross double resonance*. arXiv: 1510.08754, preprint, 2015.

[22] Clarke, F. H., Optimization and nonsmooth analysis. Second edition. *Classics in Applied Mathematics, 5*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[23] Dafermos, C. M. *Generalized characteristics and the structure of solutions of hyperbolic conservation laws*. Indiana Univ. Math. J. 26 (1977), no. 6, 1097–1119.

[24] Dal Maso, G.; Frankowska, H., *Autonomous integral functionals with discontinuous nonconvex integrands: Lipschitz regularity of minimizers, DuBois-Reymond necessary conditions, and Hamilton-Jacobi equations*. Appl. Math. Optim. **48**(2003), no. 1, 39–66.

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

[26] Fathi, A.; Maderna, E. *Weak KAM theorem on non compact manifolds*. NoDEA Nonlinear Differential Equations Appl. **14**(2007), no. 1-2, 1–27.

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

[28] Khanin K; Sobolevski A., *On dynamics of Lagrangian trajectories for Hamilton-Jacobi equations*. Arch. Ration. Mech. Anal., **219**(2016), no. 2, 861–885.

[29] Lasry, J.-M.; Lions, P.-L., *A remark on regularization in Hilbert spaces*. Israel J. Math. 55 (1986), no. 3, 257–266.

[30] Lions, P.-L., *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics 69, Pitman, Boston, 1982.

[31] Mather, J. N., *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. **207**(1991), no. 2, 169–207.

[32] Mather, J. N., *Variational construction of connecting orbits*, Ann. Inst. Fourier (Grenoble) **43**(1993), no. 5, 1349–1386.

[33] 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. **53**(2015), no. 5, 2898–2933.

[34] Rifford, L., *On viscosity solutions of certain Hamilton-Jacobi equations: regularity results and generalized Sard's theorems*. Comm. Partial Differential Equations 33 (2008), no. 1-3, 517–559.

[35] Strömberg T; Ahmadzadeh F. *Excess action and broken characteristics for HamiltonJacobi equations*. Nonlinear Analysis, **110**(2014), 113–129.

[36] Villani, C. Optimal transport: old and new. *Grundlehren der Mathematischen Wissenschaften*, 338. Springer-Verlag, Berlin, 2009.

[37] Yu, Y., *A simple proof of the propagation of singularities for solutions of Hamilton–Jacobi equations*. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) **5**(2006), 439–444.

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI ROMA "TOR VERGATA", VIA DELLA RICERCA SCIENTIFICA 1, 00133 ROMA, ITALY

E-mail address: cannarsa@mat.uniroma2.it

DEPARTMENT OF MATHEMATICS, NANJING UNIVERSITY, NANJING 210093, CHINA

E-mail address: chengwei@nju.edu.cn