



# Global propagation of singularities for solutions of Hamilton–Jacobi equations



Paolo Albano

*Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, 40127 Bologna, Italy*

## ARTICLE INFO

### Article history:

Received 10 March 2016

Available online 25 July 2016

Submitted by H. Frankowska

### Keywords:

Hamilton–Jacobi equations

Viscosity solutions

Singular set

Generalized characteristics

## ABSTRACT

We show that the singularities of the viscosity solutions for a class of Hamilton–Jacobi equations propagate along the generalized characteristics for all the times.

© 2016 Elsevier Inc. All rights reserved.

## 1. Introduction and statement of the results

We are interested in the singularities of the viscosity solution of nonlinear first order pde's of Hamilton–Jacobi type (see [14,15] for the definition of a solution in the viscosity sense). In the present context, a singular point of a function is a point of non-differentiability for such a function.

Let us consider the viscosity solution,  $u : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$ , of the Cauchy problem

$$\begin{cases} \frac{\partial u(x, t)}{\partial t} + H(x, t, \nabla u(x, t)) = 0 & \text{in } \mathbb{R}^n \times ]0, T[, \\ u(x, 0) = u_0(x) & \text{in } \mathbb{R}^n, \end{cases} \quad (1.1)$$

where

$$H(x, t, p) = \frac{1}{2} \langle A(x, t)p, p \rangle + \mathcal{U}(x, t) \quad (1.2)$$

is the time dependent mechanical Hamiltonian.

We notice that the viscosity solution of (1.1) is locally Lipschitz continuous and solves Equation (1.1) a.e. in  $\mathbb{R}^n \times ]0, T[$ .

*E-mail address:* [paolo.albano@unibo.it](mailto:paolo.albano@unibo.it).

We denote by  $\Sigma(u)$  the *singular set* (i.e. the set of all the points of non-differentiability for  $u$ ). Furthermore, we recall that the  $C^1$  *singular support* of  $u$ ,  $\text{sing supp}_{C^1} u$ , is the complement of the largest open set on which  $u$  is of class  $C^1$ . We point out that, in general,  $\Sigma(u)$  is not a closed set and

$$\Sigma(u) \subseteq \overline{\Sigma(u)} = \text{sing supp}_{C^1} u.$$

In order to describe our result we need to introduce the notion of a generalized characteristic. For this purpose, we denote by  $D^+u(x, t)$  the superdifferential of  $u$  at  $(x, t)$ , i.e. it is the set of all the vectors  $(\xi, \tau) \in \mathbb{R}^n \times \mathbb{R}$  such that

$$u(y, s) - u(x, t) - \langle \xi, y - x \rangle - \tau(s - t) \leq o(|(y - x, t - s)|),$$

as  $(y, s) \rightarrow (x, t)$ . Let  $\Pi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$  be the projection on the “space” variables,  $\Pi(x, t) = x$  and, for every  $(x, t) \in \mathbb{R}^n \times ]0, T[$ , we define

$$\nabla^+u(x, t) = \Pi D^+u(x, t).$$

For  $(x_0, t_0) \in \mathbb{R}^n \times ]0, T[$ , we consider the differential inclusion

$$\begin{cases} \gamma'(t) \in A(\gamma(t), t) \nabla^+u(\gamma(t), t) & \text{a.e. } t \in [t_0, T[ \\ \gamma(t_0) = x_0. \end{cases} \quad (1.3)$$

We denote by  $\gamma(\cdot; x_0, t_0)$  a solution of (1.3) and we will call such a curve a generalized characteristic (it is a natural generalization of the space projection of a bicharacteristic curve).

We study the following problem: let  $(x_0, t_0) \in \mathbb{R}^n \times ]0, T[$  be a singular point and let  $\gamma(\cdot; x_0, t_0)$  be the solution of (1.3). Can we deduce that, for every  $t \in [t_0, T[$ ,  $(\gamma(t; x_0, t_0), t) \in \Sigma(u)$ ?

A positive answer to the question above was given in [16] in the special case of  $n = 1$ .

For general  $n \geq 1$ , in [5], it is shown that the singular set is invariant w.r.t. the generalized characteristics (at least for “small” times). The proof of the theorem on the propagation of the singularities given in [5] is constructive: in essence it is an iterative procedure based on the construction given in [3]. Such a proof was strongly simplified in [20] (see also [13]).

In [4], in the special case of concave solutions of the equation

$$u_t + H(Du) = 0,$$

with  $H$  of class  $C^2$ , convex and such that  $c_0 I \leq D_p^2 H(p) \leq c_1 I$  for every  $p \in \mathbb{R}^n$  and for suitable  $0 < c_0 \leq c_1$ , it is shown that the singularities propagate for all the times.

In [10] (see also [18]) the authors proved that  $\Sigma(u)$  is globally invariant w.r.t. the generalized characteristics in the constant coefficients case (i.e.  $H(p) = \langle Ap, p \rangle$ , where  $A$  is a positive definite constant matrix).

In the next example, taken from [5], it is clarified the difference between the case of constant and the one of variable coefficients. In the latter case the singularities may disappear.

**Example 1.** The function

$$u(x, t) = -a^2(t)|x|, \quad (x, t) \in \mathbb{R} \times \mathbb{R}^+,$$

where  $a(t) = \min\{0, t - 1\}$  is a solution of the equation

$$u_t(x, t) + (u_x(x, t))^2 + 2a(t)|x| - a^4(t) = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+$$

having as singular set, in  $\mathbb{R} \times \mathbb{R}^+$ , the segment

$$\Sigma(u) = \{(0, t) \mid 0 < t < 1\}.$$

We point out that, as shown in the example above, the low regularity of the Hamiltonian (the potential  $\mathcal{U}(x, t) = 2a(t)|x| - a^4(t)$  is not a smooth function) may influence the behaviour of the singularities and the singularities may disappear. In the present paper, in particular, we show that if the coefficients are at least  $C^2$  this phenomenon cannot occur (i.e. the singularities propagate for all the times).

We assume that

$$\begin{cases} 0 < A(\cdot) \leq c_0 I \text{ (as quadratic forms),} \\ u_0 \in C(\mathbb{R}^n) \text{ and } \exists c_1 \geq 0 : \quad \forall x \in \mathbb{R}^n \quad u_0(x) \geq -c_1, \\ H(\cdot, p) \in C^2(\mathbb{R}^n \times [0, T]) \quad (\forall p \in \mathbb{R}^n) \end{cases} \quad (1.4)$$

( $I$  is the identity matrix) and we prove the following

**Theorem 1.1.** *Under Assumption (1.4), let  $u$  be the viscosity solution of equation (1.1), let  $(x_0, t_0) \in \Sigma(u)$  and let  $\gamma(\cdot; x_0, t_0)$  be the solution of Equation (1.3). Then  $(\gamma(t; x_0, t_0), t) \in \Sigma(u)$  for every  $t \in [t_0, T[$ .*

A first  $n$  ( $> 1$ ) dimensional result on the global propagation of singularities in the case of variable coefficients (i.e. the Hamiltonian depends of the variables  $x$  and  $t$  and slightly more general than (1.2)) was given in [2] (see also [19]). More precisely, it is shown that the  $C^1$  singular support of  $u$  is (globally) invariant w.r.t. the generalized characteristics. We observe explicitly that such a result does not exclude that  $(x_0, t_0) \in \Sigma(u)$  and

$$(\gamma(t; x_0, t_0), t) \in \text{sing supp}_{C^1} u \setminus \Sigma(u) = \overline{\Sigma(u)} \setminus \Sigma(u),$$

for every  $t \in ]t_0, T[$ . Finally, we observe that, because of the identity  $\overline{\Sigma(u)} = \text{sing supp}_{C^1} u$ , the result given in [2] is a direct consequence of Theorem 1.1.

## 2. Preliminaries

A function  $u : \mathbb{R}^n \times ]0, T[ \rightarrow \mathbb{R}$  is semiconcave if for every compact subset  $K \subset \mathbb{R}^n \times ]0, T[$  there exists  $C$  such that  $D^2 u \leq C$  in  $\mathcal{D}'(K)$ . One can show that a semiconcave function can be (locally) represented as the sum of a concave with a smooth function. We recall the following regularity result (see e.g. [12, 11]).

**Theorem 2.1.** *Under Assumptions (1.4) the viscosity solution of the Cauchy Problem (1.1) is semiconcave in  $\mathbb{R}^n \times ]0, T[$ .*

A consequence of the fact that  $u$  is a semiconcave solution of equation (1.1) is the following

**Remark 2.1.** Let  $V \subset (\mathbb{R}^n \times ]0, T[) \setminus \Sigma(u)$  be an open set. Then,  $u \in C^{1,1}(V)$ . In particular  $\text{sing supp}_{C^1} u = \text{sing supp}_{C^{1,1}} u$ .

We recall that the (generalized) characteristics flow is well defined, i.e. we have the

**Theorem 2.2.** *We assume (1.4) and let  $u$  be the viscosity solution of equation (1.1). Then, for every  $(x_0, t_0) \in \mathbb{R}^n \times ]0, T[$ , Equation (1.3) admits a unique solution.*

**Theorem 2.2** is (in essence) well-known, we omit the proof. We only remark that the existence of a generalized characteristic can be proved either arguing as in [5] (see also [20]) or using Theorem 3, page 98, in [8] (the existence of a solution of an upper semicontinuous compact-convex valued differential inclusion<sup>1</sup>). While the uniqueness can be deduced arguing as in Lemma 1 of [6].

The viscosity solution of Equation (1.1) can be represented as the value function of the classical problem of the calculus of variations

$$u(x, t) = \min_{\substack{y \in W^{1,1}([0,t]; \mathbb{R}^n) \\ y(t)=x}} \left\{ \int_0^t L(y(s), s, y'(s)) ds + u_0(y(0)) \right\}. \quad (2.1)$$

Here  $L$  is the Legendre transform of  $H$ ,

$$L(x, t, q) = \sup_{p \in \mathbb{R}^n} \{ \langle q, p \rangle - H(x, t, p) \} = \frac{1}{2} \langle A^{-1}(x, t)q, q \rangle - \mathcal{U}(x, t).$$

In the proof of **Theorem 1.1** we use the following approximation result of Yu (see [20]).

**Lemma 2.1.** *Let  $\Omega \subseteq \mathbb{R}^n \times ]0, T[$  then there exists a positive constant  $C$  such that*

$$\|Du\|_{L^\infty(\Omega)} \leq C \quad \text{and} \quad D^2u \leq CI \quad \text{in } \Omega.$$

*Let  $V$  be an open bounded subset such that  $(x_0, t_0) \in V \subseteq \Omega$ . If  $(x_0, t_0) \in \Sigma(u)$  then there exist a sequence of smooth functions  $\{u_m\}_{m \geq 1}$  in  $\Omega$  such that*

- (i)  $\lim_{m \rightarrow \infty} u_m = u$ , uniformly in  $\bar{V}$ ;
- (ii)  $\|Du_m\|_{L^\infty(\bar{V})} \leq C$  and  $D^2u_m \leq CI$  in  $V$ ;
- (iii)  $\lim_{m \rightarrow \infty} Du_m(x_0, t_0) = (\xi_0, \tau_0)$  for some  $(\xi_0, \tau_0) \in D^+u(x_0, t_0)$  satisfying  $\tau_0 + H(x_0, t_0, \xi_0) < 0$ .

**Remark 2.2.** We point out that the sequence of smooth functions given by **Lemma 2.1** is of the form

$$(\eta_m * u)(x) = \frac{1}{\varepsilon_m^{n+1}} \int_{\Omega} u(y) \eta\left(\frac{x-y}{\varepsilon_m}\right) dy =: u_m(x),$$

where  $\varepsilon_m$  is a suitable sequence converging to 0, as  $m \rightarrow \infty$ ,  $\eta \in C_0^\infty(B_1(0))$  and satisfies  $\eta > 0$  in  $B_1(0)$  and  $\int_{B_1(0)} \eta(x) dx = 1$ .

We will also use the following result which is direct consequence of a Theorem of Alexandroff [7].

**Theorem 2.3.** *Let  $u : \mathbb{R}^n \times ]0, T[ \rightarrow \mathbb{R}$  be a semiconcave function. Then there exists a set  $\text{Alex}(u) \subset \mathbb{R}^n \times ]0, T[$  such that*

- (i) *the set  $\mathbb{R}^n \times ]0, T[ \setminus \text{Alex}(u)$  has measure zero;*

<sup>1</sup> The map  $x \mapsto A(x, t) \nabla^+ u(x, t)$  is clearly compact-convex valued while its semicontinuity is a consequence of the fact that if  $u$  is semiconcave then  $x \mapsto \nabla^+ u(x, t)$  is upper semicontinuous.

(ii) for every  $(x, t) \in \text{Alex}(u)$  we have

$$\begin{aligned} u(y, s) &= u(x, t) + \langle Du(x, t), (y - x \quad s - t) \rangle \\ &\quad + \frac{1}{2} \langle D^2 u(x, t) \begin{pmatrix} y - x \\ s - t \end{pmatrix}, \begin{pmatrix} y - x \\ s - t \end{pmatrix} \rangle + o(|(y - x, s - t)|^2) \end{aligned}$$

$$\text{as } |(y - x, s - t)|^2 \rightarrow 0.$$

We observe that  $\text{Alex}(u) \cap \Sigma(u) = \emptyset$ .

### 3. Proof of Theorem 1.1

In order to clarify our strategy of proof we first prove Theorem 1.1 in the simpler case

$$A \equiv I \quad \text{and} \quad \mathcal{U} \equiv 0.$$

We remark that, at a point of differentiability for  $u$ ,  $D^+u = \{(\nabla u, \partial_t u)\}$  and Equation (1.1) is satisfied in the classical (pointwise) sense. Furthermore, for every  $(x, t) \in \Sigma(u)$  there exists  $(\xi, \tau) \in D^+u(x, t)$  such that

$$\tau + \frac{1}{2}|\xi|^2 < 0.$$

The inequality above gives a quantitative criterion to check whether a point  $(x, t)$  is singular or not. Hence, in order to prove Theorem 1.1, it suffices to show that if there exists  $(\xi_0, \tau_0) \in D^+u(x_0, t_0)$  such that

$$\tau_0 + \frac{1}{2}|\xi_0|^2 < 0$$

then, for every  $t \in ]t_0, T[$ , there exists  $(\xi(t), \tau(t)) \in D^+u(\gamma(t; x_0, t_0), t)$  such that

$$\tau(t) + \frac{1}{2}|\xi(t)|^2 < 0.$$

In this special case the generalized characteristic starting at  $(x_0, t_0)$  is the solution of the differential inclusion

$$\gamma'(t) \in \nabla^+ u(\gamma(t), t), \quad \text{a.e. } t \in ]t_0, T[, \quad \gamma(t_0) = x_0. \quad (3.1)$$

Define

$$\sigma_0 = \inf\{t > t_0 \mid (\gamma(t), t) \notin \Sigma(u)\}.$$

Assume, by contradiction, that  $\sigma_0 < T$ . We point out that  $(\gamma(\sigma_0), \sigma_0) \notin \Sigma(u)$  (otherwise, by the results in [5], there should be a number  $\varepsilon > 0$  such that  $(\gamma(t), t) \in \Sigma(u)$  for every  $t \in [\sigma_0, \sigma_0 + \varepsilon[$  in contradiction with the definition of  $\sigma_0$ ).

Let  $\Omega$  be a small neighbourhood of  $(\gamma(\sigma_0), \sigma_0)$  and let  $t_1 \in [t_0, \sigma_0[$  such that  $(\gamma(t_1), t_1) \in \Omega \cap \Sigma(u)$ .

Let  $u_m$  be the sequence of functions given by Lemma 2.1 (with  $(x_0, t_0) = (\gamma(t_1), t_1)$ ) and consider the solution of the ordinary differential equation

$$\gamma'_m(t) = \nabla u_m(\gamma_m(t), t), \quad t \in [t_1, \sigma_0], \quad \gamma_m(t_1) = \gamma(t_1). \quad (3.2)$$

Set

$$\varphi_m(t) = (\partial_t u_m)(\gamma_m(t), t) + \frac{1}{2} |\nabla u_m(\gamma_m(t), t)|^2.$$

We observe that because of [Lemma 2.1](#) we have that  $\varphi_m(t_1) < 0$ , for  $m$  large enough, and  $\lim_{m \rightarrow \infty} \varphi_m(t_1) < 0$ . Hence, the proof reduces to show that there exists a positive constant  $c$  (independent of  $m$ ) such that, for every  $t \in [t_1, \sigma_0]$

$$\varphi_m(t) \leq -c. \quad (3.3)$$

Indeed, by the bounds in [Lemma 2.1](#), the sequence  $\gamma_m$  is uniformly Lipschitz continuous. Hence, possibly passing to a subsequence we may assume that it converges to a limit  $\tilde{\gamma}$  uniformly on  $[t_1, \sigma_0]$ . The limit  $\tilde{\gamma}$  is the solution of the Cauchy problem

$$\tilde{\gamma}'(t) \in D^+u(\tilde{\gamma}(t), t), \quad \text{a.e. } t \in [t_1, \sigma_0], \quad \tilde{\gamma}(t_1) = \gamma(t_1). \quad (3.4)$$

(See [\[20\]](#) for a proof of this fact.) Then, because of [\(3.4\)](#) admits a unique solution (forward in time), we deduce that possibly a subsequence of  $\gamma_m$  converges to  $\gamma$  uniformly on  $[t_1, \sigma_0]$ . Using, once more, the bounds in [Lemma 2.1](#), we find that

$$\begin{aligned} u_m(y, s) &\leq u_m(\gamma_m(\sigma_0), \sigma_0) + \langle \nabla u_m(\gamma_m(\sigma_0), \sigma_0), y - \gamma_m(\sigma_0) \rangle \\ &\quad + (\partial_t u_m)(\gamma_m(\sigma_0), \sigma_0)(s - \sigma_0) + \frac{C}{2} |(y - \gamma_m(\sigma_0), s - \sigma_0)|^2 \end{aligned} \quad (3.5)$$

for  $(y, s) \in \Omega$ . Because of the estimate

$$|(\nabla u_m(\gamma_m(\sigma_0), \sigma_0), (\partial_t u_m)(\gamma_m(\sigma_0), \sigma_0))| \leq C$$

with  $C$  independent of  $m$ , possibly taking a subsequence, we may suppose that

$$\lim_{m \rightarrow \infty} (\nabla u_m(\gamma_m(\sigma_0), \sigma_0), (\partial_t u_m)(\gamma_m(\sigma_0), \sigma_0)) = (\xi, \tau) \in \mathbb{R}^n \times \mathbb{R}.$$

Hence, passing in the limit as  $m \rightarrow \infty$  in [\(3.5\)](#), we deduce that  $(\xi, \tau) \in D^+u(\gamma(\sigma_0), \sigma_0)$  and evaluating [\(3.3\)](#) at  $t = \sigma_0$  and taking the limit as  $m \rightarrow \infty$ , we would find the contradiction

$$(\gamma(\sigma_0), \sigma_0) \notin \Sigma(u) \quad \text{and} \quad \exists (\xi, \tau) \in D^+u(\gamma(\sigma_0), \sigma_0) \text{ with } \tau + \frac{1}{2} |\xi|^2 < 0.$$

Let us prove estimate [\(3.3\)](#). We have that

$$\frac{d\varphi_m(t)}{dt} = \left\langle D^2 u_m(\gamma_m(t), t) \begin{pmatrix} \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix}, \begin{pmatrix} \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix} \right\rangle, \quad (3.6)$$

hence we need to estimate the second derivatives of  $u_m$ . We point out that this estimate is the core of our proof.

We recall that, by the Hopf formula, the solution of the Cauchy problem [\(1.1\)](#) can be written as

$$u(x, t) = \inf_{y \in \mathbb{R}^n} \left( \frac{|x - y|^2}{2t} + u_0(y) \right) = \inf_{y \in \mathbb{R}^n} (S(x, t, y) + u_0(y)), \quad (3.7)$$

for every  $(x, t) \in \mathbb{R}^n \times ]0, T[$ .

We recall that we are assuming that  $(\gamma(\sigma_0), \sigma_0)$  is a point of differentiability for  $u$ . Then, there exists a unique point  $y_0 \in \mathbb{R}^n$  such that

$$u(\gamma(\sigma_0), \sigma_0) = S(\gamma(\sigma_0), \sigma_0, y_0) + u_0(y_0).$$

Possibly taking  $\Omega$  smaller we may suppose that there exists a neighbourhood of  $y_0$  in  $\mathbb{R}^n$ ,  $W$ , such that for every  $(x, t) \in \Omega$  there exists  $y \in W$  such that

$$u(x, t) = S(x, t, y) + u_0(y)$$

and, for every  $y \in W$ ,  $(x, t) \mapsto S(x, t, y)$  is in of class  $C^\infty$  in  $\Omega$ . Hence, we have

$$\partial_t S(x, t, y) + \frac{1}{2} |\nabla S(x, t, y)|^2 = 0, \quad (x, t) \in \Omega, \quad (3.8)$$

for every  $y \in W$ .

Let  $(x, t) \in \Omega \cap \text{Alex}(u)$ , then we have

$$u(z, s) - u(x, t) \leq S(z, s, y) - S(x, t, y), \quad (3.9)$$

for every  $(z, s) \in \Omega$ . Here  $y \in W$  is such that

$$u(x, t) = S(x, t, y) + u_0(y).$$

Hence taking a derivative of (3.8) we find

$$\partial_t^2 S(x, t, y) + \langle \nabla \partial_t S(x, t, y), \nabla S(x, t, y) \rangle = 0$$

and

$$\partial_t \nabla S(x, t, y) + \nabla^2 S(x, t, y) \nabla S(x, t, y) = 0,$$

i.e. we get the following key identity

$$D^2 S(x, t, y) \begin{pmatrix} \nabla S(x, t, y) \\ 1 \end{pmatrix} = 0 \quad (3.10)$$

Let  $f$  be the normalized vector

$$f = \frac{1}{\sqrt{1 + |\nabla S(x, t, y)|^2}} \begin{pmatrix} \nabla S(x, t, y) \\ 1 \end{pmatrix}.$$

Then, denoting by  $\pi_f : \mathbb{R}^{n+1} \rightarrow \text{span}\{f\}$  the orthogonal projection into the linear space generated by  $f$ , because of (3.10) and the semiconcavity of  $S$ , we find

$$D^2 S(x, t, y) \leq C(I - f \otimes f),$$

in the sense of the quadratic forms, for a suitable positive constant  $C$ .<sup>2</sup>

<sup>2</sup> One can show that the constant  $C$  depends only on the set  $\Omega$  and the Hamiltonian. This fact can following the same arguments as in [1] (replacing the Dirichlet problem with the Cauchy problem).

Furthermore, using the fact that  $I - f \otimes f$  is positive semidefinite and that  $S$  is a (classical) solution of equation (3.8), we deduce that

$$\begin{aligned} D^2S(x, t, y) &\leq C \left( I(1 + |\nabla S(x, t, y)|^2) - \begin{pmatrix} \nabla S(x, t, y) \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \nabla S(x, t, y) \\ 1 \end{pmatrix} \right) \\ &= C \left( I(1 - 2\partial_t S(x, t, y)) - \begin{pmatrix} \nabla S(x, t, y) \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \nabla S(x, t, y) \\ 1 \end{pmatrix} \right). \end{aligned} \quad (3.11)$$

Hence, by (3.9), we have

$$\begin{aligned} u(z, s) - u(x, t) &\leq \langle DS(x, t, y), (z - x, s - t) \rangle \\ &\quad + \frac{1}{2} \left\langle D^2S(x, t, y) \begin{pmatrix} z - x \\ s - t \end{pmatrix}, \begin{pmatrix} z - x \\ s - t \end{pmatrix} \right\rangle + o(|(z - x, s - t)|^2), \end{aligned} \quad (3.12)$$

as  $(z, s) \rightarrow (x, t)$ , for a suitable constant  $C$ . Since  $(x, t) \in \text{Alex}(u)$ , we have, in particular, that  $u$  is differentiable at  $(x, t)$  and, by (3.12), we have that

$$DS(x, t, y) = Du(x, t),$$

and

$$D^2u(x, t) \leq C \left( I(1 - 2\partial_t u(x, t)) - \begin{pmatrix} \nabla u(x, t) \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \nabla u(x, t) \\ 1 \end{pmatrix} \right),$$

for every  $(x, t) \in \text{Alex}(u)$ . Because of  $u_m$  is defined by taking the standard convolution with a mollifier we deduce that

$$D^2u_m(x, t) \leq C \left( I(1 - 2\partial_t u_m(x, t)) - \begin{pmatrix} \nabla u_m(x, t) \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \nabla u_m(x, t) \\ 1 \end{pmatrix} \right),$$

for  $(x, t) \in V \Subset \Omega$ . Plugging the inequality above into (3.6) we find

$$\begin{aligned} \frac{d}{dt} \varphi_m(t) &\leq C(1 - 2\partial_t u_m(\gamma_m(t), t))(1 + |\nabla u_m(\gamma_m(t), t)|^2) \\ &\quad - C(1 + |\nabla u_m(\gamma_m(t), t)|^2)^2 = -2C\varphi_m(t). \end{aligned}$$

Hence we deduce that

$$\varphi_m(t) \leq e^{-2C(t-t_1)} \varphi_m(t_1) \leq e^{-2C(\sigma-t_1)} \varphi_m(t_1).$$

Recalling that, by (iii) of Lemma 2.1,  $\lim_{m \rightarrow \infty} \varphi_m(t_1) < 0$ , we conclude that there exists a positive  $c$  (independent of  $m$ ) such that (3.3) holds, for  $m$  large enough. This completes our proof in the case of constant coefficients.

Let us prove Theorem 1.1 in the general case.

Once more, we assume, by contradiction, that there exists a generalized characteristic  $(\gamma(t), t)$  such that  $(\gamma(t_0), t_0) \in \Sigma(u)$  and such that

$$\sigma_0 = \inf\{t > t_0 \mid (\gamma(t), t) \notin \Sigma(u)\} < T.$$



Furthermore, by [5], we may also assume that  $(\gamma(\sigma_0), \sigma_0) \notin \Sigma(u)$ . We will work in a small neighbourhood of such a point,  $\Omega$ .

We need a representation formula for the solution (1.1) in  $\Omega$ .

We point out that, in the general case, we have an additional technical difficulty w.r.t. the case of constant coefficients already treated (i.e.  $A \equiv I$  and  $\mathcal{U} \equiv 0$ ). Indeed, in the constant coefficients case, we have a global smooth solution of the Hamilton–Jacobi equation

$$\partial_t u + \frac{1}{2} |\nabla u|^2 = 0,$$

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

such that the solution of the Cauchy problem (1.1) can be written as the inf-convolution of  $S$  with the initial datum. In the general case, in order to have a representation formula involving a function  $S(x, t, y)$  (at least) of class  $C^2$  w.r.t. the variables  $(x, t)$  we need to consider a local (in space and in time) Cauchy problem.

We point out that it is a straightforward construction.

For this purpose, let  $\Omega \subset \mathbb{R}^{n+1}$  be a neighbourhood of  $(\gamma(\sigma_0), \sigma_0)$ , let  $(x, t), (y, s) \in \Omega$  with  $s < t$  we denote by

$$S(x, t, y, s) = \inf_s \int_s^t L(x(r), r, x'(r)) dr \quad (3.13)$$

where the infimum is taken w.r.t. all the curves  $x(\cdot) \in W^{1,1}([s, t]; \Omega)$  such that  $x(s) = y$  and  $x(t) = x$ . We introduce some subset of  $\Omega$ :

$$\Omega_s = \{(x, s) \in \Omega\}$$

and let

$$\Omega_s^+ = \{(x, t) \in \Omega \mid t > s\}.$$

We remark that, because of the  $C^2$  smoothness of the data,  $A(\cdot)$  and  $\mathcal{U}(\cdot)$ ,  $(x, t) \mapsto S(x, t, y, s)$  is a function of class  $C^2$  for every  $(y, s) \in \Omega$  with  $s < \sigma_0$  and  $(x, t) \in \Omega_s^+$ , provided that  $\Omega$  is small enough. Furthermore, we have that

$$\partial_t S(x, t, y, s) + \frac{1}{2} \langle A(x, t) \nabla_x S(x, t, y, s), \nabla_x S(x, t, y, s) \rangle + \mathcal{U}(x, t) = 0. \quad (3.14)$$

(See e.g. section 2 in [9] for a proof of the regularity of  $S$  and the fact that it is a solution of (3.14).)

Now, define, for  $(x, t) \in \Omega_s^+$ ,

$$v(x, t) = \inf_{(y, s) \in \Omega_s} (S(x, t, y, s) + u(y, s)). \quad (3.15)$$

We recall that the compatibility condition for the Cauchy problem reads as

$$u(x, t) - u(y, s) \leq S(x, t, y, s) \quad \forall (x, t), (y, s) \in \Omega, \quad s < t, \quad (3.16)$$

(see e.g. [17]). Furthermore,  $v$  is a viscosity solution of the equation

$$\partial_t v(x, t) + H(x, t, \nabla v(x, t)) = 0 \text{ in } \Omega_s^+, \quad v = u \text{ on } \Omega_s. \quad (3.17)$$

(This somehow standard fact is proved in the appendix.)

Since also  $u$  is a viscosity solution of the same Cauchy problem (3.17) and the equation has a finite propagation speed (e.g. see [17]), we deduce that  $u \equiv v$  in  $\tilde{\Omega}_s^+$ , possibly taking a smaller neighbourhood  $\tilde{\Omega}_s^+ \subset \Omega_s^+$  of the point  $(\gamma(\sigma_0), \sigma_0)$ , i.e.

$$u(x, t) = \inf_{(y, s) \in \Omega_s} (S(x, t, y, s) + u(y, s)),$$

for  $(x, t) \in \tilde{\Omega}_s^+$ . From now on, for notational simplicity, we omit to write the extra tilde by writing  $\Omega_s^+$  instead of  $\tilde{\Omega}_s^+$ .

We claim that

$$(\nabla_x u(\gamma(\sigma_0), \sigma_0), \partial_t u(\gamma(\sigma_0), \sigma_0)) = (\nabla_x S(x, t, y, s), \partial_t S(x, t, y, s))|_{(x, t) = (\gamma(\sigma_0), \sigma_0)},$$

for a suitable  $y$  such that  $(y, s) \in \overline{\Omega_s}$  and

$$u(\gamma(\sigma_0), \sigma_0) = S(\gamma(\sigma_0), \sigma_0, y, s) + u(y, s).$$

Indeed, we have that, for  $(x, t) \in \Omega_s^+$ ,

$$\begin{aligned} u(x, t) - u(\gamma(\sigma_0), \sigma_0) &\leq S(x, t, y, s) - S(\gamma(\sigma_0), \sigma_0, y, s) \\ &\leq \langle (\nabla_x S)(\gamma(\sigma_0), \sigma_0, y, s), x - \gamma(\sigma_0) \rangle + (\nabla_t S)(\gamma(\sigma_0), \sigma_0, y, s)(t - \sigma_0) \\ &\quad + \frac{C}{2} |x - \gamma(\sigma_0), t - \sigma_0|^2 \end{aligned}$$

for a suitable positive constant  $C$ . Hence, we deduce that

$$((\nabla_x S)(\gamma(\sigma_0), \sigma_0, y, s), (\nabla_t S)(\gamma(\sigma_0), \sigma_0, y, s)) \in D^+ u(\gamma(\sigma_0), \sigma_0)$$

and recalling that we are assuming  $(\gamma(\sigma_0), \sigma_0) \notin \Sigma(u)$ , we conclude that  $D^+ u(\gamma(\sigma_0), \sigma_0)$  is a singleton and our claim holds. We point out that, using the argument above, one can show that if  $(x, t) \in \Omega_s^+$  is a point of differentiability of  $u$  then

$$(\nabla_x u(x, t), \partial_t u(x, t)) = (\nabla_x S(x, t, y(x, t), s), \partial_t S(x, t, y(x, t), s))$$

where  $y(x, t)$  is so that  $(y(x, t), s) \in \Omega_s$  and

$$u(x, t) = S(x, t, y(x, t), s) + u(y(x, t), s).$$

Now, let us take  $s \in ]0, \sigma_0[$ .

By (3.14), we find

$$\begin{aligned} \partial_t^2 S(x, t, y, s) + \frac{1}{2} \langle \partial_t A(x, t) \nabla S(x, t, y, s), \nabla S(x, t, y, s) \rangle \\ + \langle \partial_t \nabla S(x, t, y, s), A(x, t) \nabla S(x, t, y, s) \rangle + \partial_t \mathcal{U}(x, t) = 0 \end{aligned}$$

and

$$\begin{aligned} \partial_t \nabla S(x, t, y, s) + \frac{1}{2} \langle \nabla A(x, t) \nabla S(x, t, y, s), \nabla S(x, t, y, s) \rangle \\ + \nabla^2 S(x, t, y, s) A(x, t) \nabla S(x, t, y, s) + \nabla \mathcal{U}(x, t) = 0, \end{aligned}$$

for every  $(x, t) \in \Omega_s^+$  and for every  $y$  such that  $(y, s) \in \Omega_s$ .

We point out that the symbol

$$\langle \nabla A(x, t) \nabla S(x, t, y, s), \nabla S(x, t, y, s) \rangle$$

stands for the vector

$$\sum_{i,j=1}^n (\nabla A_{ij})(x, t) \partial_{x_i} S(x, t, y, s) \partial_{x_j} S(x, t, y, s).$$

In other words, we have

$$\begin{aligned} D^2 S(x, t, y, s) \begin{pmatrix} A(x, t) \nabla S(x, t, y, s) \\ 1 \end{pmatrix} \\ = - \left( \frac{1}{2} \langle \nabla A(x, t) \nabla S(x, t, y, s), \nabla S(x, t, y, s) \rangle + \nabla \mathcal{U}(x, t) \right) \\ - \left( \frac{1}{2} \langle \partial_t A(x, t) \nabla S(x, t, y, s), \nabla S(x, t, y, s) \rangle + \partial_t \mathcal{U}(x, t) \right), \end{aligned} \quad (3.18)$$

for every  $(x, t) \in \Omega_s^+$  and for every  $y \in \mathbb{R}^n$  with  $(y, s) \in \Omega_s$ . We argue as in the constant coefficients case.

From now on, for the sake of simplicity, we omit to write the dependence from the variables  $x, t$  and possibly  $y, s$ . We consider the vector

$$f = \frac{1}{\sqrt{1 + \langle A \nabla S, \nabla S \rangle}} \begin{pmatrix} A \nabla S \\ 1 \end{pmatrix}$$

which is normalized w.r.t. the quadratic form

$$\left\langle \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix} \cdot, \cdot \right\rangle.$$

Then, we define

$$\pi_f \begin{pmatrix} \xi \\ \tau \end{pmatrix} = \left\langle \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix} f, \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle f = \frac{\langle \nabla S, \xi \rangle + \tau}{\sqrt{1 + \langle A \nabla S, \nabla S \rangle}} f,$$

for every  $\xi \in \mathbb{R}^n$  and  $\tau \in \mathbb{R}$ . Notice that  $\pi_f f = f$  and  $\pi_f^2 = \pi_f$ .

Because of the semiconcavity of  $S$ , we have

$$\begin{aligned} \left\langle D^2 S \begin{pmatrix} \xi \\ \tau \end{pmatrix}, \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle \\ \leq C \left\langle (I - \pi_f) \begin{pmatrix} \xi \\ \tau \end{pmatrix}, \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle + 2 \left\langle D^2 S \pi_f \begin{pmatrix} \xi \\ \tau \end{pmatrix}, (I - \pi_f) \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle + \left\langle D^2 S \pi_f \begin{pmatrix} \xi \\ \tau \end{pmatrix}, \pi_f \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle, \end{aligned}$$

for every  $\xi \in \mathbb{R}^n$ ,  $\tau \in \mathbb{R}$ . Recalling (3.18) we obtain

$$\begin{aligned} & \left\langle D^2 S \left( \frac{\xi}{\tau} \right), \left( \frac{\xi}{\tau} \right) \right\rangle \\ & \leq \frac{C}{1 + \langle A \nabla S, \nabla S \rangle} \left\{ (1 - 2\partial_t S - 2\mathcal{U})(|\xi|^2 + \tau^2) - \left\langle (\langle \nabla S, \xi \rangle + \tau) \begin{pmatrix} A \nabla S \\ 1 \end{pmatrix}, \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle \right\} \\ & \quad - 2 \frac{\langle \nabla S, \xi \rangle + \tau}{1 + \langle A \nabla S, \nabla S \rangle} \left\langle \begin{pmatrix} \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla S, \nabla S \rangle \\ \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla S, \nabla S \rangle \end{pmatrix}, (I - \pi_f) \begin{pmatrix} \xi \\ \tau \end{pmatrix} \right\rangle \\ & \quad - \left( \frac{\langle \nabla S, \xi \rangle + \tau}{1 + \langle A \nabla S, \nabla S \rangle} \right)^2 \left\langle \begin{pmatrix} \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla S, \nabla S \rangle \\ \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla S, \nabla S \rangle \end{pmatrix}, \begin{pmatrix} A \nabla S \\ 1 \end{pmatrix} \right\rangle, \end{aligned}$$

for a suitable  $C > 0$ . Hence, we find

$$\begin{aligned} \left\langle D^2 S \left( \frac{\xi}{\tau} \right), \left( \frac{\xi}{\tau} \right) \right\rangle & \leq C \{ (1 - 2\partial_t S - 2\mathcal{U})(|\xi|^2 + \tau^2) - (\langle \nabla S, \xi \rangle + \tau)(\langle A \nabla S, \xi \rangle + \tau) \} \\ & \quad - 2 \frac{\langle \nabla S, \xi \rangle + \tau}{1 + \langle A \nabla S, \nabla S \rangle} \left\langle \begin{pmatrix} \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla S, \nabla S \rangle \\ \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla S, \nabla S \rangle \end{pmatrix}, \xi \right\rangle \\ & \quad + \left( \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla S, \nabla S \rangle \right) \tau \\ & \quad + \left( \frac{\langle \nabla S, \xi \rangle + \tau}{1 + \langle A \nabla S, \nabla S \rangle} \right)^2 \left\langle \begin{pmatrix} \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla S, \nabla S \rangle \\ \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla S, \nabla S \rangle \end{pmatrix}, A \nabla S \right\rangle \\ & \quad + \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla S, \nabla S \rangle, \end{aligned}$$

for a suitable  $C > 0$ .

Arguing as in the constant coefficients case (i.e. using the Alexandroff theorem) we deduce that

$$\begin{aligned} & \left\langle D^2 u \left( \frac{\xi}{\tau} \right), \left( \frac{\xi}{\tau} \right) \right\rangle \\ & \leq C \{ (1 - 2\partial_t u - 2\mathcal{U})(|\xi|^2 + \tau^2) - (\langle \nabla u, \xi \rangle + \tau)(\langle A \nabla u, \xi \rangle + \tau) \} \\ & \quad - 2 \frac{\langle \nabla u, \xi \rangle + \tau}{1 + \langle A \nabla u, \nabla u \rangle} \left\langle \begin{pmatrix} \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla u, \nabla u \rangle \\ \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u, \nabla u \rangle \end{pmatrix}, \xi \right\rangle + \left( \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u, \nabla u \rangle \right) \tau \\ & \quad + \left( \frac{\langle \nabla u, \xi \rangle + \tau}{1 + \langle A \nabla u, \nabla u \rangle} \right)^2 \left\langle \begin{pmatrix} \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla u, \nabla u \rangle \\ \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u, \nabla u \rangle \end{pmatrix}, A \nabla u \right\rangle + \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u, \nabla u \rangle, \quad (3.19) \end{aligned}$$

on  $\Omega_s^+ \cap \text{Alex}(u)$ , for a suitable  $C > 0$ . Notice that the R.H.S. of the inequality above depends only on the gradient of  $u$  which is in  $L^\infty(\Omega)$ .

Hence, taking the convolution with the mollifier  $\eta_m$ , we conclude that (possibly taking a smaller  $\Omega$ )

$$D^2 u_m(x, t) \leq (\eta_m * B)(x, t), \quad \text{for every } (x, t) \in \Omega. \quad (3.20)$$

Here,  $B(x, t)$  is the R.H.S. of the inequality (3.19). Let us consider the ordinary differential equation

$$\gamma'_m(t) = A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \quad t \in [t_1, \sigma_0], \quad \gamma_m(t_1) = \gamma(t_1),$$

and set

$$\varphi_m(t) = \partial_t u_m(\gamma_m(t), t) + \frac{1}{2} \langle A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle + \mathcal{U}(\gamma_m(t), t),$$

for  $t \in [t_1, \sigma_0]$ . In order to complete the proof, it suffices to show that there exists  $c > 0$  such that

$$\varphi_m(t) \leq -c \quad (3.21)$$

for every  $t \in [t_1, \sigma_0]$  and for  $m$  large enough.

Indeed, once (3.21) is established a verbatim repetition of the arguments used in the case  $A \equiv I$  and  $\mathcal{U} \equiv 0$  yields the conclusion.

In order to estimate  $\varphi_m(t)$  let us compute and estimate the time derivative of  $\varphi_m$ :

$$\begin{aligned} \frac{d}{dt} \varphi_m(t) &= \left\langle D^2 u_m(\gamma_m(t), t) \begin{pmatrix} A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix}, \begin{pmatrix} A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix} \right\rangle \\ &\quad + \frac{1}{2} \langle (\partial_t A)(\gamma_m(t), t) \nabla u_m(\gamma_m(t)), \nabla u_m(\gamma_m(t)) \rangle \\ &\quad + \frac{1}{2} \langle \langle \nabla A(\gamma_m(t), t) \nabla u_m(\gamma_m(t)), \nabla u_m(\gamma_m(t)) \rangle, A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\ &\quad + (\partial_t \mathcal{U})(\gamma_m(t), t) + \langle \nabla \mathcal{U}(\gamma_m(t), t), A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\ &\leq \frac{1}{2} \langle (\partial_t A)(\gamma_m(t), t) \nabla u_m(\gamma_m(t)), \nabla u_m(\gamma_m(t)) \rangle \\ &\quad + \left\langle (\eta_m * B)(\gamma_m(t), t) \begin{pmatrix} A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix}, \begin{pmatrix} A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix} \right\rangle \\ &\quad + \frac{1}{2} \langle \langle \nabla A(\gamma_m(t), t) \nabla u_m(\gamma_m(t)), \nabla u_m(\gamma_m(t)) \rangle, A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\ &\quad + (\partial_t \mathcal{U})(\gamma_m(t), t) + \langle \nabla \mathcal{U}(\gamma_m(t), t), A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle. \end{aligned}$$

In the above estimate we used (3.20). We need some more information on the matrix  $\eta_m * B$ . By the definition of  $B$ , we obtain the identity

$$\begin{aligned} &\left\langle (\eta_m * B)(\gamma_m(t), t) \begin{pmatrix} A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix}, \begin{pmatrix} A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \\ 1 \end{pmatrix} \right\rangle \\ &= C(1 - 2\partial_t u_m(\gamma_m(t), t) - 2\mathcal{U}_m(\gamma_m(t), t))(|A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|^2 + 1) \\ &\quad - C \left\{ \eta_m * [(\langle \nabla u, \xi \rangle + \tau)(\langle A \nabla u, \xi \rangle + \tau)] \right\}(\gamma_m(t), t) \\ &\quad - 2 \left\{ \eta_m * \left[ \frac{\langle \nabla u, \xi \rangle + \tau}{1 + \langle A \nabla u, \nabla u \rangle} \left( \left\langle \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla u, \nabla u \rangle, \xi \right\rangle \right. \right. \right. \\ &\quad \left. \left. \left. + \left( \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u, \nabla u \rangle \right) \tau \right) \right] \right\}(\gamma_m(t), t) \\ &\quad + \left\{ \eta_m * \left[ \left( \frac{\langle \nabla u, \xi \rangle + \tau}{1 + \langle A \nabla u, \nabla u \rangle} \right)^2 \left( \left\langle \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla u, \nabla u \rangle, A \nabla u \right\rangle \right. \right. \right. \right. \\ &\quad \left. \left. \left. + \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u, \nabla u \rangle \right) \right] \right\}(\gamma_m(t), t), \end{aligned} \quad (3.22)$$

where  $(\xi, \tau) = (A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), 1)$ . Then we find

$$\begin{aligned}
 & \left\langle (\eta_m * B)(\gamma_m(t), t) \left( A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \right)_1, \left( A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \right)_1 \right\rangle \\
 &= C(1 - 2\partial_t u_m(\gamma_m(t), t) - 2U(\gamma_m(t), t))(|A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|^2 + 1) \\
 &\quad - C(\langle \nabla u_m(\gamma_m(t), t), \xi \rangle + \tau)(\langle A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \xi \rangle + \tau) \\
 &\quad - 2 \left\{ \frac{\langle \nabla u_m, \xi \rangle + \tau}{1 + \langle A \nabla u_m, \nabla u_m \rangle} \left( \left\langle \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla u_m, \nabla u_m \rangle, \xi \right\rangle \right. \right. \\
 &\quad \left. \left. + \left( \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u_m, \nabla u_m \rangle \right) \tau \right) \right\} (\gamma_m(t), t) \\
 &\quad + \left\{ \left( \frac{\langle \nabla u_m, \xi \rangle + \tau}{1 + \langle A \nabla u_m, \nabla u_m \rangle} \right)^2 \left( \left\langle \left( \nabla \mathcal{U} + \frac{1}{2} \langle \nabla A \nabla u_m, \nabla u_m \rangle, A \nabla u_m \right) \right. \right. \right. \\
 &\quad \left. \left. + \partial_t \mathcal{U} + \frac{1}{2} \langle \partial_t A \nabla u_m, \nabla u_m \rangle \right) \right\} (\gamma_m(t), t) + E_m(t) \\
 &= C(1 - 2\partial_t u_m(\gamma_m(t), t) - 2\mathcal{U}(\gamma_m(t), t))(|A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|^2 + 1) \tag{3.23} \\
 &\quad - C(\langle A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle + 1)(|A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|^2 + 1) \\
 &\quad - 2\langle \nabla \mathcal{U}(\gamma_m(t), t), A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad - \langle \langle \nabla A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle, A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad - 2\partial_t \mathcal{U}(\gamma_m(t), t) - \langle \partial_t A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad + \langle \nabla \mathcal{U}(\gamma_m(t), t), A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad + \frac{1}{2} \langle \langle \nabla A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle, A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad + \frac{1}{2} \langle \partial_t A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad + \partial_t \mathcal{U}(\gamma_m(t), t) + E_m(t) \\
 &= -2C(|A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|^2 + 1)\varphi_m(t) \\
 &\quad - \langle \nabla \mathcal{U}(\gamma_m(t), t) A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle \\
 &\quad - \frac{1}{2} \langle \langle \nabla A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle, A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t) \rangle - \partial_t \mathcal{U}(\gamma_m(t), t) \\
 &\quad - \frac{1}{2} \langle \partial_t A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t), \nabla u_m(\gamma_m(t), t) \rangle + E_m(t)
 \end{aligned}$$

where  $E_m$  is implicitly defined by (3.22) and (3.23).

Hence, we have

$$\frac{d}{dt} \varphi_m(t) \leq -2C(|A(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|^2 + 1)\varphi_m(t) + E_m(t)$$

and we deduce the (rough) estimate

$$\varphi_m(t) \leq \varphi_m(t_1) + C \int_{t_1}^{\sigma_0} |E_m(s)| ds$$

for every  $t \in [t_1, \sigma_0]$  and for a suitable positive constant  $C$ .

In order to complete the proof it suffices to show that

$$\lim_{m \rightarrow \infty} \int_{t_1}^{\sigma_0} |E_m(s)| ds = 0.$$

We observe that a typical term in  $E_m$  is of the form

$$\eta_m * \left( a \frac{q(|B\nabla u|)}{r(|B\nabla u|)} \right) - a \frac{q(|B\nabla u_m|)}{r(|B\nabla u_m|)}$$

where  $a$  is a smooth function,  $B$  a matrix with smooth entries and  $q/r$  is a rational function with  $r$  a positive polynomial. In order to complete the proof it suffices to show that

$$\begin{aligned} \lim_{m \rightarrow \infty} \int_{t_1}^{\sigma_0} & \left| \left( \eta_m * a \frac{q(|B\nabla u|)}{r(|B\nabla u|)} \right) (\gamma_m(t), t) \right. \\ & \left. - a(\gamma_m(t), t) \frac{q(|B(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|)}{r(|B(\gamma_m(t), t) \nabla u_m(\gamma_m(t), t)|)} \right| dt = 0. \end{aligned}$$

We observe that the integral above can be estimated by

$$\left\| \left( \eta_m * a \frac{q(|B\nabla u|)}{r(|B\nabla u|)} \right) - a \frac{q(|B\nabla u_m|)}{r(|B\nabla u_m|)} \right\|_{L^1(V)}$$

( $V$  is the set given by [Lemma 2.1](#)). Since the function

$$a \frac{q(|B\nabla u|)}{r(|B\nabla u|)}$$

is in  $L^\infty(V)$  (and the mollified function  $L^1$ -converges) it suffices to show that

$$\lim_{m \rightarrow \infty} \left\| \frac{q(|B\nabla u|)}{r(|B\nabla u|)} - \frac{q(|B\nabla u_m|)}{r(|B\nabla u_m|)} \right\|_{L^1(V)} = 0.$$

This fact is a consequence of the identity  $\nabla u_m = \eta_m * \nabla u$  because of  $\|\nabla u - \nabla u_m\|_{L^p(V)} \rightarrow 0$ , as  $m \rightarrow \infty$ , for every  $p \in [1, \infty[$ . This completes our proof of [Theorem 1.1](#).

## Appendix A

In this appendix we just include for the sake of completeness the proof of the following

**Lemma A.1.** *The function  $v(x, t) = \inf_{(y,s) \in \Omega_s} (S(x, t, y, s) + u(y, s))$  defined in [\(3.15\)](#) is the viscosity solution of the equation*

$$\begin{cases} \partial_t v(x, t) + \frac{1}{2} \langle A(x, t) \nabla v(x, t), \nabla v(x, t) \rangle + \mathcal{U}(x, t) = 0 & \text{in } \Omega_s^+ \\ v(x, t) = u(x, t) & (x, t) \in \Omega_s. \end{cases} \quad (\text{A.1})$$

**Proof.** Due to the convexity of the Hamiltonian w.r.t. the gradient it suffices to show that if  $\varphi = \varphi(x, t)$  is a function of class  $C^1$  and  $v - \varphi$  has a local minimum at  $(x_0, t_0) \in \Omega_s^+$  then

$$\partial_t \varphi(x_0, t_0) + \langle A(x_0, t_0) \nabla \varphi(x_0, t_0), \nabla \varphi(x_0, t_0) \rangle + \mathcal{U}(x_0, t_0) = 0. \quad (\text{A.2})$$

We have that

$$v(x, t) \geq v(x_0, t_0) + \langle \nabla \varphi(x_0, t_0), x - x_0 \rangle + \partial_t \varphi(x_0, t_0)(t - t_0) + o(|(x - x_0, t - t_0)|)$$

as  $\Omega_s^+ \ni (x, t) \rightarrow (x_0, t_0)$ . By the definition of  $v$  we deduce that

$$\begin{aligned} S(x, t, y, s) &\geq S(x_0, t_0, y, s) + \langle \nabla \varphi(x_0, t_0), x - x_0 \rangle \\ &\quad + \partial_t \varphi(x_0, t_0)(t - t_0) + o(|(x - x_0, t - t_0)|) \end{aligned}$$

as  $\Omega_s^+ \ni (x, t) \rightarrow (x_0, t_0)$ . Here  $y \in \mathbb{R}^n$  is so that

$$v(x_0, t_0) = S(x_0, t_0, y, s) + u(y, s).$$

Hence, the vector  $(\nabla \varphi(x_0, t_0), \partial_t \varphi(x_0, t_0))$  is in the subdifferential of  $S$  (w.r.t.  $(x, t)$ ) at  $(x_0, t_0, y, s)$  and, in light of the differentiability of  $S$  and (3.14), we deduce that (A.2) holds, i.e.  $v$  is a viscosity solution of Equation (A.1). It remains to show that  $v$  attains the initial datum at time  $t = s$ . This follows, by observing that

$$v(x, s) = \inf_{(y, s) \in \Omega_s} (S(x, s, y, s) + u(y, s)) \leq u(x, s). \quad (\text{A.3})$$

(Here we used the fact that, by the definition of  $S$ ,  $S(x, s, y, s) = 0$ .) Furthermore, by the compatibility condition (3.16), we deduce that

$$u(x, s) \leq v(x, s). \quad \square \quad (\text{A.4})$$

Hence, (A.3) together with (A.4) yield the conclusion: the Cauchy datum at  $t = s$  is attained. This completes our proof.

## References

- [1] P. Albano, On the local semiconcavity of the solutions of the eikonal equation, *Nonlinear Anal.* 73 (2) (2010) 458–464.
- [2] P. Albano, Propagation of singularities for solutions of Hamilton–Jacobi equations, *J. Math. Anal. Appl.* 411 (2) (2014) 684–687.
- [3] P. Albano, P. Cannarsa, Structural properties of singularities of semiconcave functions, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* 28 (1999) 719–740.
- [4] P. Albano, P. Cannarsa, Propagation of singularities for concave solutions of Hamilton–Jacobi equations, in: *International Conference on Differential Equations, Vol. 1, 2*, Berlin, 1999, World Sci. Publ., River Edge, NJ, 2000, pp. 583–588.
- [5] P. Albano, P. Cannarsa, Propagation of singularities for solutions of nonlinear first order partial differential equations, *Arch. Ration. Mech. Anal.* 162 (2002) 1–23.
- [6] P. Albano, P. Cannarsa, K.T. Nguyen, C. Sinestrari, Singular gradient flow of the distance function and homotopy equivalence, *Math. Ann.* 356 (1) (2013) 23–43.
- [7] A.D. Alexandroff, Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it, *Leningrad State Univ. Ann. [Uchenye Zapiski] Math. Ser.* 6 (1939) 3–35 (in Russian).
- [8] J.P. Aubin, A. Cellina, *Differential Inclusions. Set-Valued Maps and Viability Theory*, Grundlehren der Mathematischen Wissenschaften, vol. 264, Springer-Verlag, Berlin, 1984.
- [9] P. Bernard, The Lax–Oleinik semi-group: a Hamiltonian point of view, *Proc. Roy. Soc. Edinburgh Sect. A* 142 (6) (2012) 1131–1177.
- [10] P. Cannarsa, M. Mazzola, C. Sinestrari, Global propagation of singularities for time dependent Hamilton–Jacobi equations, *Discrete Contin. Dyn. Syst.* 35 (9) (2015) 4225–4239.
- [11] P. Cannarsa, C. Sinestrari, *Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control*, Progress in Nonlinear Differential Equations and Their Applications, vol. 58, Birkhäuser, Boston, 2004.
- [12] P. Cannarsa, H.M. Soner, On the singularities of the viscosity solutions to Hamilton–Jacobi–Bellman equations, *Indiana Univ. Math. J.* 36 (1987) 501–524.
- [13] P. Cannarsa, Y. Yu, Singular dynamics for semiconcave functions, *J. Eur. Math. Soc. (JEMS)* 11 (2009) 999–1024.
- [14] M.G. Crandall, L.C. Evans, P.L. Lions, Some properties of viscosity solutions of Hamilton–Jacobi equations, *Trans. Amer. Math. Soc.* 282 (1984) 487–502.



- [15] M.G. Crandall, P.L. Lions, Viscosity solutions of Hamilton–Jacobi equations, *Trans. Amer. Math. Soc.* 277 (1983) 1–42.
- [16] C.M. Dafermos, Generalized characteristics and the structure of solutions of hyperbolic conservation laws, *Indiana Univ. Math. J.* 26 (1977) 1097–1119.
- [17] P.L. Lions, Generalized Solutions of Hamilton–Jacobi Equations, *Research Notes in Mathematics*, vol. 69, Pitman, 1982.
- [18] T. Strömberg, A new proof of indefinite propagation of singularities for a Hamilton–Jacobi equations, preprint, 2016.
- [19] T. Strömberg, F. Ahmadzadeh, Excess action and broken characteristics for Hamilton–Jacobi equations, *Nonlinear Anal.* 110 (2014) 113–129.
- [20] Y. Yu, A simple proof of the propagation of singularities for solutions of Hamilton–Jacobi equations, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (2006) 439–444.