HOKKAIDO UNIVERSITY

Title	A DYNAMICAL APPROACH TO LOWER GRADIENT ESTIMATES FOR VISCOSITY SOLUTIONS OF HAMILTON-JACOBI EQUATIONS
Author(s)	HAMAMUKI, NAO; HIROSE, KAZUYA
Citation	Hokkaido University Preprint Series in Mathematics, 1144, 1-34
Issue Date	2022-03-01
DOI	10.14943/101331
Doc URL	http://hdl.handle.net/2115/84215
Type	bulletin (article)
File Information	DynAppLow.pdf

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

A DYNAMICAL APPROACH TO LOWER GRADIENT ESTIMATES FOR VISCOSITY SOLUTIONS OF HAMILTON-JACOBI EQUATIONS

NAO HAMAMUKI AND KAZUYA HIROSE

Abstract. We present a new approach to deriving lower bounds for weak spatial gradients of a viscosity solution to a Hamilton-Jacobi equation when the Hamiltonian is convex. For the proof, we consider Hamiltonian systems with approximate Hamiltonians and study how the initial gradients propagate along the solutions to the Hamiltonian systems. We also apply our method to the case of homogeneous Hamiltonians, whose examples include level-set equations arising in surface evolution problems. In this case we obtain sharper gradient estimates. Moreover, we compare our results with the previous work [22]. We show that our results give better and optimal estimates in several senses.

Contents

Date: February 24, 2022.
2010 *Mathematics Subject Classification.* 35D40, 35F21, 35F25, 37J50.
Key words and phrases. Lower bound estimates for the gradient; Viscosity solutions; Hamilton-Jacobi equations; Surface evolution problems.

1. Introduction

Equation and goals. In this paper we consider a Hamilton-Jacobi equation

$$u_t(x,t) + H(x,t,D_x u(x,t)) = 0 \quad \text{in } \mathbf{R}^n \times (0,T) \tag{1.1}$$

with the initial condition

$$u(x,0) = u_0(x) \quad \text{in } \mathbf{R}^n. \tag{1.2}$$

Here $u : \mathbf{R}^n \times [0,T) \to \mathbf{R}$ is the unknown function, and $u_t = \partial_t u$, $D_x u = (\partial_{x_i} u)_{i=1}^n$ denote its derivatives. Moreover, the Hamiltonian $H : \mathbf{R}^n \times [0,T] \times \mathbf{R}^n \to \mathbf{R}$ is continuous in $\mathbf{R}^n \times [0,T] \times \mathbf{R}^n$, and the initial datum $u_0 : \mathbf{R}^n \to \mathbf{R}$ is Lipschitz continuous in $\mathbf{R}^n$. Throughout this paper, we assume these regularity conditions on H and u_0. The Lipschitz constant of u_0 over $\mathbf{R}^n$ is denoted by $\|Du_0\|_{L^\infty(\mathbf{R}^n)}$.

The goal of this paper is to present a new approach to deriving lower bounds for spatial gradients of the viscosity solution u to (1.1)–(1.2) when the Hamiltonian H is convex. More specifically, we prove that, for a fixed $(x,t) \in \mathbf{R}^n \times (0,T)$, every element $p \in D_x^- u(x,t)$ of the spatial subdifferential satisfies

$$|p| \geqq C; \tag{1.3}$$

in other words, $|D_x^- u| \geqq C$ in the viscosity solution sense. Here C depends on the initial subdifferentials over some region in $\mathbf{R}^n$. We also study the case where (1.1) is a level-set equation appearing in surface evolution problems. In this case we will derive sharper gradient estimates.

A lower bound for gradients of solutions to (1.1)–(1.2) has already been studied in [22] with a different approach. Comparison with the results in [22] is also discussed and we show that our results give better estimates in several senses.

Assumptions and methods. Our assumptions on H are as follows:

(H1) There exist $C_1 \geqq 0$ and $\beta \in \{0,1\}$ such that

$$|H(x,t,p) - H(y,t,p)| \leqq C_1(\beta + |p|)|x-y|$$

for all $(x,t,p) \in \mathbf{R}^n \times [0,T] \times \mathbf{R}^n$ and $y \in \mathbf{R}^n$.

(H2) There exist $A_2, B_2 \geqq 0$ such that

$$|H(x,t,p) - H(x,t,q)| \leqq (A_2|x| + B_2)|p-q|$$

for all $(x,t,p) \in \mathbf{R}^n \times [0,T] \times \mathbf{R}^n$ and $q \in \mathbf{R}^n$.

(H3) $p \mapsto H(x,t,p)$ is convex in $\mathbf{R}^n$ for all $(x,t) \in \mathbf{R}^n \times [0,T]$.

(H4) For every $R > 0$, H is bounded and uniformly continuous in $\mathbf{R}^n \times [0,T] \times \overline{B_R(0)}$.

Here $|\cdot|$ stands for the standard Euclidean norm, $B_r(x)$ denotes the open ball with radius r centered at x, and $\overline{B_r(x)}$ is its closure. The assumptions (H1)–(H3) are the same as the ones in [22], while we impose (H4) for uniqueness and existence of viscosity solutions to (1.1)–(1.2). Moreover, the unique solution u is Lipschitz continuous in $\mathbf{R}^n \times [0,T)$. For these results, see [22, Theorem 4.1], [21, Appendix A] and [1, Chapter 2, Sections 5 and 8] for instance. One can relax (H4) (see, e.g., [19]), but we use it just for simplification.

In order to estimate the gradients of solutions, we consider the Hamiltonian system:

$$\begin{cases} \xi'(s) = D_p H(\xi(s), s, \eta(s)), \\ \eta'(s) = -D_x H(\xi(s), s, \eta(s)). \end{cases} \tag{1.4}$$

Here we temporarily assume that H is smooth. When the solution u of (1.1)–(1.2) is smooth enough, the relation between u and the solution (ξ, η) of (1.4) is well known in the theory of classical dynamics ([11, Chapter 1], [12, Chapter 3]). The curve ξ is often called a classical characteristic. We also recall that $\eta(s)$ represents the spatial derivative of u at $(\xi(s), s)$, i.e.,

$$\eta(s) = D_x u(\xi(s), s).$$

The theory above is generalized for a possibly nonsmooth viscosity solution u ([11, Chapters 5 and 6]).

One of important ingredients for our gradient estimates is a recent result obtained in [2]. It is shown in [2, Theorem 3.2] (Theorem 2.14 below) that, for any $y \in \mathbf{R}^n$, there is a one-to-one correspondence between $p \in D^-_{pr} u_0(y)$ and a classical characteristic ξ with $\xi(0) = y$. Here $D^-_{pr} u_0(y)$ denotes the proximal subdifferential (Definition 2.1).

Let us illustrate the idea for our gradient estimates at a differentiable point (x, t) of u. We take the solution (ξ, η) of (1.4) with the terminal condition $(\xi(t), \eta(t)) = (x, D_x u(x, t))$. Then, u is differentiable at every point $(\xi(s), s)$ $(s \in (0, t))$ along ξ. By [2, Theorem 3.2] we see that

$$\eta(0) \in D^-_{pr} u_0(\xi(0)). \tag{1.5}$$

We then apply Gronwall's lemma to discover bounds for $|\eta(t) - \eta(0)|$ and $|\xi(t) - \xi(0)|$. The bound for $|\eta(t) - \eta(0)|$ yields the estimate for $|\eta(t)| = |D_x u(x, t)|$ by $|\eta(0)|$, while we find possible positions of $\xi(0)$ from the bound for $|\xi(t) - \xi(0)| = |x - \xi(0)|$. From these facts and (1.5) we deduce the gradient estimate of the form (1.3) with $p = D_x u(x, t)$. For a nonsmooth viscosity solution, the above argument is justified via approximation.

To apply the results of [2] we need more conditions on H than (H1)–(H4), which are smoothness and strict convexity as follows:

(H5) $H \in C^2(\mathbf{R}^n \times [0, T] \times \mathbf{R}^n)$.

$(\mathrm{H3})_{\mathrm{st}}$ $D_{pp} H(x, t, p)$ is positive definite for all $(x, t, p) \in \mathbf{R}^n \times [0, T] \times \mathbf{R}^n$.

See [2, page 1413, **(H)**]. In Section 3 we derive gradient estimates when H satisfies (H5) and $(\mathrm{H3})_{\mathrm{st}}$. These additional conditions are removed in Section 4, where we approximate H by H_ε satisfying (H5) and $(\mathrm{H3})_{\mathrm{st}}$ and study the equations

$$(u_\varepsilon)_t(x, t) + H_\varepsilon(x, t, D_x u_\varepsilon(x, t)) = 0 \quad \text{in } \mathbf{R}^n \times (0, T). \tag{1.6}$$

Applying the results in Section 3 to approximate solutions u_ε of (1.6), we derive the gradient estimates for (1.1).

We mention that, though our main interest lies in lower bounds for gradients of solutions, we will simultaneously present upper bounds of them because the upper bounds are obtained in an almost parallel way.

Our method provides a rather straightforward way to derive lower bounds although careful approximation is needed. Thanks to this, we are able to find the lower bounds in an explicit way and they are optimal as shown in Section 7. These facts are advantages of our method as well as contributions of this paper.

Surface evolution problem. In Section 5 we study a Hamiltonian H which is positively homogeneous of degree 1 with respect to p. More precisely, we assume

(H6) $H(x, t, \lambda p) = \lambda H(x, t, p)$ for all $(x, t, p) \in \mathbf{R}^n \times [0, T] \times \mathbf{R}^n$ and $\lambda \geqq 0$.

Such a Hamiltonian appears in application of a level-set method to surface evolution problems. When H satisfies (H6), improved gradient estimates are available.

To explain the results, let us review the level-set approach to surface evolution problems. The reader is referred to [14] for the details. For a given initial hypersurface $\Gamma(0)$ in $\mathbf{R}^n$, we consider its evolution $\{\Gamma(t)\}_{t\in[0,T)}$. To track the motion, we represent $\Gamma(t)$ as the zero level-set of an auxiliary function $u : \mathbf{R}^n \times [0,T) \to \mathbf{R}$, that is,

$$\Gamma(t) = \{x \in \mathbf{R}^n \mid u(x,t) = 0\}. \tag{1.7}$$

We now assume that the evolution law of $\Gamma(t)$ is given by

$$V = g(x,t,\mathbf{n}) \quad \text{on } \Gamma(t). \tag{1.8}$$

Here g is a given function, $\mathbf{n} = \mathbf{n}(x,t) \in \mathbf{R}^n$ is the unit normal vector to $\Gamma(t)$ at x from $\{x \in \mathbf{R}^n \mid u(x,t) > 0\}$ to $\{x \in \mathbf{R}^n \mid u(x,t) < 0\}$, and $V = V(x,t) \in \mathbf{R}$ is the normal velocity of $\Gamma(t)$ at x in the direction of $\mathbf{n}$. If u is smooth near (x,t) and $\nabla u(x,t) \neq 0$, we have

$$\mathbf{n} = -\frac{D_x u(x,t)}{|D_x u(x,t)|}, \quad V = \frac{u_t(x,t)}{|D_x u(x,t)|}.$$

Substituting these formulas for (1.8), we are led to (1.1) with the Hamiltonian

$$H(x,t,p) = -|p| g\left(x,t,-\frac{p}{|p|}\right). \tag{1.9}$$

The corresponding equation (1.1) is often called a level-set equation. Clearly, (1.9) satisfies (H6) for $p \neq 0$. When H has continuous extension to $p = 0$ and satisfies (H1)–(H4), our results can be applied to H.

Carrying out the level-set method, we first choose the initial datum u_0 so that $\Gamma(0) = \{x \in \mathbf{R}^n \mid u_0(x) = 0\}$ and next solve (1.1)–(1.2). Then the desired surfaces $\Gamma(t)$ are given by (1.7) with the solution u. The family of $\Gamma(t)$ obtained in this manner is called a level-set solution. An important feature of this method is that level-set solutions are unique. Namely, the set $\{x \in \mathbf{R}^n \mid u(x,t) = 0\}$ depends only on the initial level-set $\{x \in \mathbf{R}^n \mid u_0(x) = 0\}$ and independent of the other levels of u_0. This fact naturally raises the question whether the same holds for the gradients, i.e., whether the gradients of u on $\{x \in \mathbf{R}^n \mid u(x,t) = 0\}$ depend only on the initial gradients of u_0 on $\{x \in \mathbf{R}^n \mid u_0(x) = 0\}$.

We give a positive answer to this question. In Theorem 5.4, we demonstrate that the sub- and superdifferentials $D_x^{\pm} u(x,t)$ with $u(x,t) = 0$ depend only on $D_{pr}^{-} u_0(y)$ for y with $u_0(y) \approx 0$. (We actually prove this fact not only for the zero level-set but also for any $\gamma \in \mathbf{R}$ level-set.) A key observation for this theorem is that the solution u is constant along $(\xi(s), s)$ for the solution ξ of (1.4); the rigorous proof needs a suitable approximation of H and some error estimates. We also show that this technique applies to m-homogeneous Hamiltonian with $m > 1$. For the details, see Theorems 5.8 and 5.11.

As with Sections 3 and 4, we derive an upper bound for gradients together with the lower bound. Both bounds seem to be new for solutions to level-set equations.

One of typical surface evolution equations is the eikonal equation $V = \nu(x,t)$, where $\nu : \mathbf{R}^n \times [0,T] \to \mathbf{R}$ is called a driving force. The level-set equation is then given by

$$u_t(x,t) - \nu(x,t)|D_x u(x,t)| = 0 \quad \text{in } \mathbf{R}^n \times (0,T).$$

The transport equation $V = \langle \mathbf{X}(x,t), \mathbf{n}\rangle$ is also a surface evolution equation. Here $\mathbf{X} : \mathbf{R}^n \times [0,T] \to \mathbf{R}^n$ is a given vector field. The corresponding level-set equation is a linear one of the form

$$u_t(x,t) + \langle \mathbf{X}(x,t), D_x u(x,t)\rangle = 0 \quad \text{in } \mathbf{R}^n \times (0,T).$$

In Section 7, we show some examples of solutions to the above equations. In particular, we present equality cases of our gradient estimates, which imply optimality of the results.

Literature overview. As we have already mentioned, lower bound gradient estimates for solutions to (1.1)–(1.2) are derived in [22] with a different approach. We will describe the results in Section 2 and compare them with our results in Section 6. In [22] the author employs a notion of Barron-Jensen solutions ([8]) and derives the gradient estimates by carefully studying the inf-convolution of the solution. One of key facts in the proof is that the inf-convolution is a subsolution of (1.1) with an appropriate error. In [7] lower gradient bounds are obtained for solutions to second order geometric equations. For the proof, continuous dependence property for solutions is used in a crucial way. This gradient estimate is applied to prove short time uniqueness of solutions to nonlocal geometric equations. See also [6, 4, 5] for related results for nonlocal first order equations.

We mention other type of lower bound estimates for gradients of solutions. A lower bound for the spatially Lipschitz constant $\|D_x u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)}$ is investigated in [13, 16]. In [13] lower bounds with the optimal order of t are derived for linear parabolic equations and Hamilton-Jacobi equations. More general fully nonlinear parabolic equations are studied in [16]. For surface evolution problems, improved level-set equations are proposed in [17, 15] so that initial gradients are preserved on the zero level-sets of solutions.

Organization. This paper is organized as follows: Section 2 is devoted to preparation for some notations, definitions and useful facts. In Section 3 we give gradient estimates when the Hamiltonian H is smooth and strictly convex. These extra conditions on H are removed in Section 4, where we give our main gradient estimates via suitable approximation of H. In Section 5 we study homogeneous Hamiltonians and establish improved gradient estimates. Section 6 is concerned with comparison with the results obtained in [22]. Finally, we give some examples of solutions in Section 7.

2. Preliminaries

We prepare some definitions and known results which will be used later.

2.1. **Generalized gradients and semiconcavity.**

We introduce several notions of generalized gradients ([11, Section 3]). Throughout this subsection, we let $\Omega \subset \mathbf{R}^n$ be a nonempty open set, $f : \Omega \to \mathbf{R}$ and $x \in \Omega$. Also, $\langle \cdot,\cdot \rangle$ denotes the standard Euclidean inner product.

Definition 2.1. (1) The *subgradient* $D^- f(x)$ of f at x is defined as

$$D^- f(x) = \left\{ D\phi(x) \;\middle|\; \phi \in C^1(\Omega),\ f-\phi \text{ attains a local minimum at } x \right\}.$$

The *supergradient* $D^+ f(x)$ of f at x is defined by replacing "a local minimum" by "a local maximum" in the above. Moreover, for a function $u : \mathbf{R}^n \times [0,\infty) \to \mathbf{R}$ and $(z,t) \in \mathbf{R}^n \times (0,\infty)$, we define

$$D_x^\pm u(z,t) = \{p \mid (p,\tau) \in D^\pm u(z,t)\}.$$

(2) The *proximal subgradient* $D^-_{pr} f(x)$ of f at x is defined as

$$D^-_{pr} f(x) = \left\{ p \in \mathbf{R}^n \;\middle|\; \begin{array}{c} \text{there exist some } K, r > 0 \text{ such that} \\ f(y) \geqq f(x) + \langle p, y-x \rangle - \dfrac{K}{2}|y-x|^2 \text{ for all } y \in B_r(x) \subset \Omega \end{array} \right\}.$$

(3) The *reachable gradient* $D^* f(x)$ of f at x is defined as

$$D^* f(x) = \left\{ p \in \mathbf{R}^n \;\middle|\; \begin{array}{c} \text{there exists some } \{x_k\}_{k=1}^\infty \subset \Omega \setminus \{x\} \text{ such that} \\ f \text{ is differentiable at } x_k \text{ and } x_k \to x,\ Df(x_k) \to p \text{ as } k \to \infty \end{array} \right\}.$$

Remark 2.2. (1) Assume that f is Lipschitz continuous over Ω. It is then clear that, when $p \in \mathbf{R}^n$ belongs to one of $D^{\pm}f(x)$, $D^-_{pr}f(x)$ and $D^*f(x)$, we have $|p| \leqq \|Df\|_{L^\infty(\Omega)}$.

(2) By definition we easily see that $D^-_{pr}f(x) \subset D^-f(x)$, but the reverse inclusion $D^-_{pr}f(x) \supset D^-f(x)$ does not hold in general. Indeed, for $f(y) = -|y|^{1+\alpha}$ ($\alpha \in (0,1)$) we have $D^-f(0) = \{0\}$ and $D^-_{pr}f(x) = \emptyset$. However, any element of $D^-f(x)$ is reachable by elements of $D^-_{pr}f(x_\varepsilon)$ ($x_\varepsilon \approx x$) as the next proposition indicates.

Proposition 2.3. *Let $p \in D^-f(x)$. Then there exist sequences $\{x_\varepsilon\}_{\varepsilon\in(0,1]} \subset \Omega$ and $\{p_\varepsilon\}_{\varepsilon\in(0,1]} \subset \mathbf{R}^n$ such that*

$$p_\varepsilon \in D^-_{pr}f(x_\varepsilon), \quad x_\varepsilon \to x, \quad p_\varepsilon \to p \quad \textit{as } \varepsilon \to +0.$$

Proof. Take $\phi \in C^1(\Omega)$ such that $p = D\phi(x)$ and $f - \phi$ attains a local minimum at x. We may assume that $f(x) = \phi(x)$. Since ϕ is differentiable at x, there exists an increasing function $\omega : [0,\infty) \to [0,\infty)$ such that $\lim_{r\to+0}\omega(r) = \omega(0) = 0$ and

$$\phi(y) \geqq \phi(x) + \langle p, y-x\rangle - \omega(|y-x|)|y-x| \quad \text{for all } y \in \Omega.$$

Choose $\varepsilon_0 \in (0,1]$ so small that $\omega(\varepsilon_0) < 1$, $\overline{B_{\varepsilon_0}(x)} \subset \Omega$ and that x is a minimizer of $f - \phi$ over $\overline{B_{\varepsilon_0}(x)}$. Let $\varepsilon \in (0, \varepsilon_0]$ and define $r_\varepsilon := \varepsilon\omega(\varepsilon)$. Note that $0 < r_\varepsilon < \varepsilon$ by the choice of ε_0. We further define

$$\zeta_\varepsilon(y) := \phi(x) + \langle p, y-x\rangle - \frac{|y-x|^2}{\varepsilon} \quad (y \in \Omega).$$

Let x_ε be a minimum point of $f - \zeta_\varepsilon$ over $\overline{B_{r_\varepsilon}(x)}\,(\subset \Omega)$.

We now claim that $x_\varepsilon \in B_{r_\varepsilon}(x)$. First, by the definition of ζ_ε we notice that

$$f(x) - \zeta_\varepsilon(x) = f(x) - \phi(x) = 0. \tag{2.1}$$

Next, for $y \in \partial B_{r_\varepsilon}(x)$ we observe

$$\begin{aligned} f(y) - \zeta_\varepsilon(y) &\geqq \phi(y) - \zeta_\varepsilon(y) \geqq -\omega(|y-x|)|y-x| + \frac{|y-x|^2}{\varepsilon} \\ &= -\omega(r_\varepsilon)r_\varepsilon + \frac{r_\varepsilon^2}{\varepsilon} = r_\varepsilon(-\omega(r_\varepsilon) + \omega(\varepsilon)) > 0. \end{aligned}$$

This and (2.1) imply that the minimizer x_ε lies in $B_{r_\varepsilon}(x)$.

By the claim we have $p_\varepsilon := D\zeta_\varepsilon(x_\varepsilon) \in D^-_{pr}f(x_\varepsilon)$. Since $|x_\varepsilon - x| < r_\varepsilon$, we easily see that $x_\varepsilon \to x$ as $\varepsilon \to +0$. Moreover, we have

$$p_\varepsilon = p - \frac{2(x_\varepsilon - x)}{\varepsilon}, \quad \frac{|x_\varepsilon - x|}{\varepsilon} < \frac{r_\varepsilon}{\varepsilon} = \omega(\varepsilon).$$

This shows that $p_\varepsilon \to p$ as $\varepsilon \to +0$. □

From Proposition 2.3 we deduce

Corollary 2.4.

$$\begin{aligned} \sup\{|p| \mid p \in D^-f(y),\, y \in \Omega\} &= \sup\{|p| \mid p \in D^-_{pr}f(y),\, y \in \Omega\}, \\ \inf\{|p| \mid p \in D^-f(y),\, y \in \Omega\} &= \inf\{|p| \mid p \in D^-_{pr}f(y),\, y \in \Omega\}. \end{aligned}$$

The following fact is often used in the viscosity solution theory.

Lemma 2.5. *Let $\varepsilon \in (0,1]$ and $\{f_\varepsilon\}_{\varepsilon\in(0,1]} \subset C(\Omega)$. Assume that f_ε converges to f locally uniformly in Ω as $\varepsilon \to +0$. If $p \in D^+f(x)$ (resp. $p \in D^-f(x)$), there exist sequences $\{x_\varepsilon\}_{\varepsilon\in(0,1]} \subset \Omega$ and $\{p_\varepsilon\}_{\varepsilon\in(0,1]} \subset \mathbf{R}^n$ such that*

$$p_\varepsilon \in D^+f_\varepsilon(x_\varepsilon) \ (\textit{resp. } p_\varepsilon \in D^-f_\varepsilon(x_\varepsilon)), \quad x_\varepsilon \to x, \quad p_\varepsilon \to p \quad \textit{as } \varepsilon \to +0.$$

For the proof see [3, Lemma II.2.4]. Its statement is apparently different, but one can deduce Lemma 2.5 by replacing v and v_n of [3, Lemma II.2.4] by $f-\phi$ and $f_\varepsilon-\phi$, respectively, where $\phi \in C^1(\Omega)$ is a function such that $p = D\phi(x)$ and $f-\phi$ attains a strict local maximum at x.

Let $K \subset \Omega$ be a nonempty convex set. We say that f is *semiconcave* in K if $f - \frac{C}{2}|\cdot|^2$ is concave in K for some $C \geqq 0$. We list some properties of generalized gradients for a semiconcave function. For the proofs, see [3, Proposition II.4.7 (b)] and [11, Proposition 3.3.4, Theorem 3.3.6].

Proposition 2.6. *Assume that f is semiconcave in a nonempty convex set $K \subset \Omega$ and that x is an interior point of K.*

(1) *If f is not differentiable at x, then $D^-f(x) = \emptyset$.*
(2) *$D^+f(x) = \operatorname{co} D^*f(x)$, where $\operatorname{co} D^*f(x)$ stands for the convex hull of $D^*f(x)$.*
(3) *$D^*f(x) \neq \emptyset$. If f is differentiable at x, then $D^*f(x) = \{Df(x)\}$.*

2.2. **Viscosity solutions and known results on gradient estimates.** The reader is referred to [9] for the basic theory of viscosity solutions.

Definition 2.7. Let $u \in C(\mathbf{R}^n \times [0,T))$. We say that u is a *viscosity subsolution* (resp. *viscosity supersolution*) of (1.1)–(1.2) if $u(x,0) \leqq u_0(x)$ (resp. $u(x,0) \geqq u_0(x)$) for all $x \in \mathbf{R}^n$ and

$$\tau + H(x,t,p) \leqq 0 \quad (\text{resp. } \geqq 0)$$

for all $(x,t) \in \mathbf{R}^n \times (0,T)$ and $(p,\tau) \in D^+u(x,t)$ (resp. $(p,\tau) \in D^-u(x,t)$). When u is both a viscosity subsolution and a viscosity supersolution, it is called a *viscosity solution*.

We now state the results obtained in [22]. Later, we compare our results with them. For a given $x_0 \in \mathbf{R}^n$ and $r > 0$, the following assumption (U1) is imposed on u_0 in [22]:

(U1) There exists a constant $\theta > 0$ such that

$$|p| \geqq \theta \quad \text{for all } x \in B_r(x_0) \text{ and } p \in D^-u_0(x).$$

When H satisfies (H6), the next assumption (U2) is considered:

(U2) There exists a constant $\theta > 0$ such that

$$|u_0(x)| + |p| \geqq \theta \quad \text{for all } x \in \mathbf{R}^n \text{ and } p \in D^-u_0(x).$$

Let us further define

$$r(x,t) = e^{(A_2+B_2+A_2|x|)t} - 1 \tag{2.2}$$

and

$$\mathcal{D}(x_0,r) = \{(x,t) \in \mathbf{R}^n \times (0,T) \mid (r(x_0,t)+1)(|x-x_0|+1) - 1 < r\}. \tag{2.3}$$

Following [22], we call the set $\mathcal{D}(x_0,r)$ the *domain of dependence* with the base $B_r(x_0)$. The lower bound gradient estimates in [22] are stated as follows:

Theorem 2.8 ([22, Theorem 4.2 and its proof])**.** *Assume that H satisfies* (H1)–(H3). *Let u be a viscosity solution of* (1.1)–(1.2).

(1) *Let* $x_0 \in \mathbf{R}^n$, $r > 0$ *and assume* (U1). *Then, for any* $t_0 \in (0, T]$ *satisfying*

$$\theta^2 - 2\beta C_1 e^{\frac{5}{2}C_1 T} t_0 > 0,$$

we have

$$|p| \geqq \tilde{\theta} e^{-\frac{5}{4}C_1 t} \quad \text{for all } (x,t) \in \mathcal{D}(x_0, r) \cap (\mathbf{R}^n \times (0, t_0)) \text{ and } p \in D_x^- u(x,t), \tag{2.4}$$

where $\tilde{\theta} := \sqrt{\theta^2 - 2\beta C_1 e^{\frac{5}{2}C_1 T} t_0}$.

(2) *Assume furthermore that* H *satisfies* (H6) *and that* $\beta = 0$ *in* (H1). *Assume* (U2). *Then there exists a constant* $C = C(\theta) > 0$ *such that*

$$|u(x,t)| + \frac{1}{4} e^{\frac{5}{2}C_1 t} |p|^2 \geqq C \quad \text{for all } (x,t) \in \mathbf{R}^n \times (0,T) \text{ and } p \in D_x^- u(x,t).$$

Remark 2.9. Under the assumptions (H1) and (H2), a local comparison principle is proved in [22, Theorem 6.1]. See also [10, Theorem V.3], [18, Theorem 2.4], [3, Theorem III.3.12, Exercise III.3.5] and [1, Chapter 2, Theorem 5.3] for such local comparison results. This comparison principle implies that, for every $x_0 \in \mathbf{R}^n$ and $r > 0$, the values of u_0 in $B_r(x_0)$ completely determine the values of the solution u to (1.1)–(1.2) in $\mathcal{D}(x_0, r)$ ([22, Remark 6.1]). This fact and the Lipschitz continuity of the solution enable us to localize our arguments. Namely, when we study the gradients of u at a fixed $(x,t) \in \mathbf{R}^n \times (0,T)$, we may change the values of the Hamiltonian H outside $\overline{B_M(0)} \times [0,T] \times \overline{B_M(0)}$ for $M > 0$ large enough.

As an application let us consider the Hamiltonian

$$H(x,p) = |p|^m + V(x) \quad (m > 1). \tag{2.5}$$

This Hamiltonian does not satisfy (H2). However, considering a new Hamiltonian $\tilde{H}$ satisfying (H2) and $H(x,p) = \tilde{H}(x,p)$ for $|p| \leqq M$ with $M > 0$ large, one can apply the results of this paper to the solution u to (1.1)–(1.2).

We next state upper gradient estimates of solutions. Recall that $\|Du_0\|_{L^\infty(\mathbf{R}^n)}$ and $\|D_x u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)}$ denote the Lipschitz constant of u_0 and $u(\cdot,t)$ over $\mathbf{R}^n$, respectively.

Theorem 2.10 ([1, Chapter 2, Theorem 8.1]). *Assume that* H *satisfies* (H1) *and* (H4). *Let* u *be the viscosity solution of* (1.1)–(1.2). *Let* $t \in (0,T)$. *Then*

$$\|D_x u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \leqq e^{C_1 t} \|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t} - 1).$$

Though [1, Chapter 2, Theorem 8.1] assumes that u_0 is bounded, one can see that the proof works without the boundedness. We thus did not assume the boundedness of u_0 in Theorem 2.10.

2.3. **Lagrangian and the Hamiltonian system.** Let us prepare some facts about the Hamiltonian system, especially the results obtained in [2]. See also [11, Section 6 and Appendix A.2] and [20].

We first introduce the Lagrangian L associated with H. To do so, one usually needs more conditions on $H = H(x,t,p)$ such as uniform convexity and superlinear growth with respect to p ([2, page 1417, **(H1)**–**(H3)**]). Throughout this subsection we assume these conditions to define L. However, by localizing the argument (Remark 2.9), these conditions can be removed for the results below. The detailed discussion to remove the additional conditions can be found in [2, page 1421]. We thus omit the detail.

We define the *Lagrangian* $L : \mathbf{R}^n \times [0,T] \times \mathbf{R}^n \to \mathbf{R}$ associated with H by

$$L(x,t,q) := \sup_{p \in \mathbf{R}^n} \{\langle p, q\rangle - H(x,t,p)\}.$$

Then the solution u of (1.1)–(1.2) is represented as

$$u(x,t) = \inf_{\xi \in \mathcal{C}(x,t)} \left\{ u_0(\xi(0)) + \int_0^t L(\xi(s), s, \xi'(s))\, ds \right\}, \tag{2.6}$$

where

$$\mathcal{C}(x,t) := \{\xi : [0,t] \to \mathbf{R}^n \mid \xi \text{ is absolutely continuous in } [0,t],\ \xi(t) = x\}.$$

Let $\mathcal{C}_{\min}(x,t)$ be the set of $\xi \in \mathcal{C}(x,t)$ that attains the infimum of (2.6).

For the solution u, let us define

$$\mathcal{R}(u) := \{(x,t) \in \mathbf{R}^n \times (0,T) \mid u \text{ is differentiable at } (x,t)\}.$$

Theorem 2.11 ([11, page 153 and Theorem 6.4.7])**.** *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathbf{R}^n \times (0,T)$. Then $\mathcal{C}_{\min}(x,t) \neq \emptyset$ and $\mathcal{C}_{\min}(x,t) \subset C^1([0,t])$. Moreover, for every $\xi \in \mathcal{C}_{\min}(x,t)$, we have $(\xi(s), s) \in \mathcal{R}(u)$ for all $s \in (0,t)$.*

Let $\xi \in \mathcal{C}_{\min}(x,t)$. We define $\eta : [0,t] \to \mathbf{R}^n$ by

$$\eta(s) := D_q L(\xi(s), s, \xi'(s)) \quad (s \in [0,t]).$$

Then $\eta \in C^1([0,t])$. The curve η is called the *dual arc associated with* ξ.

We consider the Hamiltonian system:

$$\begin{cases} \text{(a) } \xi'(s) = D_p H(\xi(s), s, \eta(s)), \\ \text{(b) } \eta'(s) = -D_x H(\xi(s), s, \eta(s)). \end{cases} \tag{2.7}$$

Theorem 2.12 ([11, Theorems 6.3.3 and 6.4.8])**.** *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathbf{R}^n \times (0,T)$. Let $\xi \in \mathcal{C}_{\min}(x,t)$, and $\eta \in C^1([0,t])$ be the dual arc associated with ξ. Then, (ξ, η) is a solution of* (2.7)*, and*

$$\eta(s) = D_x u(\xi(s), s) \quad (s \in (0,t)).$$

Remark 2.13. By this theorem we see that

$$\xi'(s) = D_p H(\xi(s), s, D_x u(\xi(s), s)) \quad (s \in (0,t)),$$

i.e., ξ is a *classical characteristic* associated with u ([2, page 1416]).

We will apply Theorem 2.14 (2) below in a crucial way for our gradient estimates.

Theorem 2.14 ([11, Theorem 6.4.9], [2, Proposition 2.2, Theorem 3.2])**.** *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathbf{R}^n \times (0,T)$. For $(p,\tau) \in D^*u(x,t)$, let $(\xi,\eta) \in C^1([0,1])^2$ be the solution of* (2.7) *with the terminal condition*

$$\xi(t) = x, \quad \eta(t) = p. \tag{2.8}$$

Then $\xi \in \mathcal{C}_{\min}(x,t)$. Moreover,

(1) *The above map*

$$D^*u(x,t) \ni (p,\tau) \mapsto \xi \in \mathcal{C}_{\min}(x,t)$$

is bijective.

(2) $\eta(0) \in D^-_{pr} u_0(\xi(0))$.

When the Hamiltonian H is smooth and strictly convex, the solution u is locally semiconcave. See, e.g., [11, Theorem 6.4.1, Corollary 6.4.4].

Theorem 2.15. *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$*. Then the viscosity solution u of* (1.1)–(1.2) *is locally semiconcave in $\mathbf{R}^n \times (0, T)$.*

2.4. **Approximation of Hamiltonian.** For a Hamiltonian H satisfying only (H1)–(H4), we approximate it so that it becomes smooth (H5) and strictly convex $(\mathrm{H3})_{\mathrm{st}}$. We approximate H as follows:

First we extend H as

$$H(x,t,p) = \begin{cases} H(x,0,p) & (t<0), \\ H(x,T,p) & (t>T). \end{cases}$$

Let $\varepsilon \in (0,1]$. We define $H_\varepsilon : \mathbf{R}^{2n+1} \to \mathbf{R}$ by

$$H_\varepsilon(x,t,p) = (H * \rho_\varepsilon)(x,t,p) + h_\varepsilon(p) \quad ((x,t,p) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n). \tag{2.9}$$

Here $h_\varepsilon(p) := \varepsilon\sqrt{|p|^2+1}$ ($p \in \mathbf{R}^n$) and

$$(H * \rho_\varepsilon)(x,t,p) = (H * \rho_\varepsilon)(z) := \int_{B_\varepsilon(0)} H(z-w)\rho_\varepsilon(w)\, dw,$$

where $\rho_\varepsilon : \mathbf{R}^{2n+1} \to \mathbf{R}$ is the standard Friedrichs mollifier such that

$$\operatorname{supp} \rho_\varepsilon := \overline{\{z \in \mathbf{R}^{2n+1} \mid \rho_\varepsilon(z) \neq 0\}} = \overline{B_\varepsilon(0)}$$

and $\int_{B_\varepsilon(0)} \rho_\varepsilon(z)\, dz = 1$.

Remark 2.16. By properties of mollification and h_ε, we see that H_ε satisfies (H5) and $(\mathrm{H3})_{\mathrm{st}}$. Furthermore, (H1) holds with the same C_1. The condition (H2) is also fulfilled by H_ε, but the Lipschitz constant $A_2|x| + B_2$ must be replaced by $A_2|x| + B_2 + \varepsilon$ since $|Dh_\varepsilon(p)| \leqq \varepsilon$.

The definition of H_ε implies that H_ε converges to H locally uniformly in $\mathbf{R}^{2n+1}$. Therefore standard stability results for viscosity solutions ([9, Section 6]) yield locally uniform convergence of solutions.

Proposition 2.17. *Assume that H satisfies* (H1)–(H4)*, and let H_ε be defined as in* (2.9)*. Let u and u_ε be respectively the viscosity solutions of* (1.1)–(1.2) *and* (1.6)–(1.2)*. Then u_ε converges to u locally uniformly in $\mathbf{R}^n \times [0, T)$ as $\varepsilon \to +0$.*

3. Smooth and strictly convex Hamiltonian

In this section, we derive gradient estimates when the Hamiltonian H satisfies (H1)–(H5) and $(\mathrm{H3})_{\mathrm{st}}$. The results will be used in the next section to study general Hamiltonians.

3.1. **Solutions of the Hamiltonian system.** We derive some estimates for the solution (ξ, η) of (2.7)–(2.8) when $(x,t) \in \mathcal{R}(u)$. Namely, we solve (2.7) with the terminal condition

$$\xi(t) = x, \quad \eta(t) = D_x u(x,t). \tag{3.1}$$

Proposition 3.1. *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$*. Let u be the viscosity solution of* (1.1)–(1.2)*. Let $(x,t) \in \mathcal{R}(u)$ and $(\xi,\eta) \in C^1([0,t])^2$ be the solution of* (2.7)–(3.1)*. Then*

$$\begin{cases} |D_x u(x,t) - \eta(0)| \leqq (\beta + |D_x u(x,t)|)(e^{C_1 t} - 1), \\ |D_x u(x,t) - \eta(0)| \leqq (\beta + |\eta(0)|)(e^{C_1 t} - 1), \end{cases} \tag{3.2}$$

where C_1 and β are the constants in (H1). *Moreover,*

$$|\eta(0)|e^{-C_1 t} - \beta(1 - e^{-C_1 t}) \leqq |D_x u(x,t)| \leqq |\eta(0)|e^{C_1 t} + \beta(e^{C_1 t} - 1), \tag{3.3}$$

and

$$|\eta(0)|e^{-C_1 t} \leqq |D_x u(x,t)| \leqq |\eta(0)|e^{C_1 t} \quad \text{if } \beta = 0, \tag{3.4}$$

$$D_x u(x,t) = \eta(0) \quad \text{if } C_1 = 0. \tag{3.5}$$

Proof. (3.4) and (3.5) are immediate from (3.3) and (3.2), respectively. Moreover, applying $|\eta(0)| - |D_x u(x,t)| \leqq |D_x u(x,t) - \eta(0)|$ to the first inequality in (3.2), we obtain the lower estimate for $|D_x u(x,t)|$ in (3.3). The upper estimate in (3.3) is derived in a similar manner from the second inequality in (3.2).

Let us prove (3.2). Let $\tau \in [0,t)$. Integrating both the sides of (2.7)(b) over $[\tau, t]$, we have

$$\eta(t) - \eta(\tau) = -\int_\tau^t D_x H(\xi(s), s, \eta(s))\, ds.$$

By (H1),

$$\begin{aligned} |\eta(t) - \eta(\tau)| &\leqq \int_\tau^t |D_x H(\xi(s), s, \eta(s))|\, ds \leqq \int_\tau^t C_1(\beta + |\eta(s)|)\, ds \\ &\leqq \int_\tau^t C_1(\beta + |\eta(t)| + |\eta(t) - \eta(s)|)\, ds \\ &= C_1(\beta + |\eta(t)|)(t - \tau) + C_1 \int_\tau^t |\eta(t) - \eta(s)|\, ds. \end{aligned} \tag{3.6}$$

Applying Gronwall's lemma, we have

$$|\eta(t) - \eta(\tau)| \leqq C_1(\beta + |\eta(t)|)(t - \tau) + e^{C_1(t-\tau)} \int_\tau^t e^{-C_1(t-s)} C_1^2 (\beta + |\eta(t)|)(t - s)\, ds.$$

In particular, when $\tau = 0$,

$$\begin{aligned} |\eta(t) - \eta(0)| &\leqq C_1(\beta + |\eta(t)|)t + e^{C_1 t} \int_0^t e^{-C_1(t-s)} C_1^2 (\beta + |\eta(t)|)(t - s)\, ds \\ &= C_1(\beta + |\eta(t)|) \left\{ t + C_1 \int_0^t (t - s) e^{C_1 s}\, ds \right\}. \end{aligned}$$

Here

$$\begin{aligned} C_1 \int_0^t (t - s) e^{C_1 s}\, ds &= \int_0^t (t - s)(e^{C_1 s})'\, ds = \left[(t - s)e^{C_1 s}\right]_0^t + \int_0^t e^{C_1 s}\, ds \\ &= -t + \left[\frac{e^{C_1 s}}{C_1}\right]_0^t = -t + \frac{e^{C_1 t} - 1}{C_1}, \end{aligned}$$

and therefore we obtain

$$|\eta(t) - \eta(0)| \leqq (\beta + |\eta(t)|)(e^{C_1 t} - 1), \tag{3.7}$$

which is the first estimate in (3.2) since $\eta(t) = D_x u(x,t)$.

If one integrates (2.7)(b) over $[0, \tau]$ ($\tau \in (0, t]$), then the second estimate in (3.2) is derived in a similar way. $\square$

Proposition 3.2. *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathcal{R}(u)$ and $(\xi,\eta) \in C^1([0,t])^2$ be the solution of* (2.7)–(3.1). *Then*

$$\begin{cases} |x-\xi(0)| \leqq \left(\dfrac{B_2}{A_2}+|x|\right)(e^{A_2 t}-1) & \text{if } A_2>0, \\ |x-\xi(0)| \leqq B_2 t & \text{if } A_2=0, \end{cases} \tag{3.8}$$

where A_2 and B_2 are the constants in (H2).

Proof. The proof of (3.8) is similar to that of the first inequality in (3.2). In fact, integrating both the sides of (2.7)(a) over $[\tau,t]$ ($\tau \in [0,t)$), one has

$$\begin{aligned} |\xi(t)-\xi(\tau)| &= \left|\int_\tau^t D_p H(\xi(s),s,\eta(s))\,ds\right| \leqq \int_\tau^t |D_p H(\xi(s),s,\eta(s))|\,ds \\ &\leqq \int_\tau^t (A_2|\xi(s)|+B_2)\,ds. \end{aligned}$$

If $A_2=0$, (3.8) is immediate from this. Assume that $A_2>0$. Comparing the above estimate with (3.6), we see that the resulting estimate for $|\xi(t)-\xi(0)|$ is obtained by replacing C_1 and β by A_2 and $\frac{B_2}{A_2}$ in (3.7), respectively. Hence,

$$|\xi(t)-\xi(0)| \leqq \left(\frac{B_2}{A_2}+|\xi(t)|\right)(e^{A_2 t}-1).$$

This is what we have to prove since $\xi(t)=x$. □

3.2. **Gradient estimates.** By (3.8) we are led to the following definitions.

Definition 3.3. For $(x,t) \in \mathbf{R}^n \times (0,T)$ we define

$$R(x,t) := \begin{cases} \left(\dfrac{B_2}{A_2}+|x|\right)(e^{A_2 t}-1) & \text{if } A_2>0, \\ B_2 t & \text{if } A_2=0. \end{cases} \tag{3.9}$$

Moreover,

$$\begin{aligned} S(x,t;u_0) &:= \sup\left\{|p| \;\middle|\; p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)}(x)}\right\}, \\ I(x,t;u_0) &:= \inf\left\{|p| \;\middle|\; p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)}(x)}\right\}. \end{aligned}$$

In Section 6 we will compare $R(x,t)$ with $r(x,t)$ in (2.2). By Theorem 2.14 (2), Propositions 3.1, 3.2 and Definition 3.3, we have

Proposition 3.4. *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathcal{R}(u)$. Then*

$$I(x,t;u_0)e^{-C_1 t} - \beta(1-e^{-C_1 t}) \leqq |D_x u(x,t)| \leqq S(x,t;u_0)e^{C_1 t} + \beta(e^{C_1 t}-1),$$

where C_1 and β are the constants in (H1). *In particular,*

$$I(x,t;u_0)e^{-C_1 t} \leqq |D_x u(x,t)| \leqq S(x,t;u_0)e^{C_1 t} \quad \text{if } \beta=0.$$

Hereafter we omit to state estimates in the case where $\beta=0$.

We next study the generalized gradient at a point $(x,t) \notin \mathcal{R}(u)$. We prepare

Lemma 3.5. *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let* $(x,t) \in \mathbf{R}^n \times (0,T)$ *and* $\xi \in \mathcal{C}_{\min}(x,t)$. *Then,*

$$B_{R(\xi(s),s)}(\xi(s)) \subset B_{R(x,t)}(x) \quad (s \in (0,t)). \tag{3.10}$$

Proof. Recall that a pair (ξ,η), where η is the dual arc associated with ξ, solves (2.7). Let $s \in (0,t)$. We have to prove

$$|x-\xi(s)| + R(\xi(s),s) \leqq R(x,t).$$

Assume that $A_2 > 0$. By a similar argument to the proof of (3.8), we have

$$|x-\xi(s)| \leqq \left(\frac{B_2}{A_2} + |x|\right)(e^{A_2(t-s)}-1).$$

Using this, we compute

$$\begin{aligned}
|x-\xi(s)| + R(\xi(s),s) &= |x-\xi(s)| + \left(\frac{B_2}{A_2} + |\xi(s)|\right)(e^{A_2 s}-1)\\
&\leqq |x-\xi(s)| + \left(\frac{B_2}{A_2} + |x-\xi(s)| + |x|\right)(e^{A_2 s}-1)\\
&= |x-\xi(s)|e^{A_2 s} + \left(\frac{B_2}{A_2} + |x|\right)(e^{A_2 s}-1)\\
&\leqq \left(\frac{B_2}{A_2} + |x|\right)(e^{A_2(t-s)}-1)e^{A_2 s} + \left(\frac{B_2}{A_2} + |x|\right)(e^{A_2 s}-1)\\
&= \left(\frac{B_2}{A_2} + |x|\right)(e^{A_2 t}-1) = R(x,t).
\end{aligned}$$

When $A_2 = 0$, the proof of (3.10) is more straightforward. □

Theorem 3.6 (Gradient estimates I)**.** *Assume that H satisfies* (H1)–(H5) *and* $(\mathrm{H3})_{\mathrm{st}}$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let* $(x,t) \in \mathbf{R}^n \times (0,T)$.

(1) *If $p \in D_x^- u(x,t)$, then*

$$I(x,t;u_0)e^{-C_1 t} - \beta(1-e^{-C_1 t}) \leqq |p| \leqq S(x,t;u_0)e^{C_1 t} + \beta(e^{C_1 t}-1).$$

(2) *If $p \in D_x^+ u(x,t)$, then $|p| \leqq S(x,t;u_0)e^{C_1 t} + \beta(e^{C_1 t}-1)$.*

Proof. Recall that u is locally semi-concave in $\mathbf{R}^n \times (0,T)$ (Theorem 2.15). If $(x,t) \in \mathcal{R}(u)$, then the conclusions follow from Proposition 3.4 since $D_x^{\pm} u(x,t) = \{D_x u(x,t)\}$. In what follows, we assume that $(x,t) \notin \mathcal{R}(u)$.

(1) By Proposition 2.6 (1), we have $D^- u(x,t) = \emptyset$, which concludes the proof.

(2) We first prove the following fact:

$$\text{If } (p,\tau) \in D^* u(x,t), \text{ then } |p| \leqq S(x,t;u_0)e^{C_1 t} + \beta(e^{C_1 t}-1). \tag{3.11}$$

Fix $(p,\tau) \in D^* u(x,t)$. Let $(\xi,\eta) \in C^1([0,t])^2$ be the solution of (2.7)–(2.8). Then, by Theorem 2.11 we see that $(\xi(s),s) \in \mathcal{R}(u)$ $(s \in (0,t))$ and

$$p = \eta(t) = \lim_{s\to t-0} \eta(s) = \lim_{s \to t-0} D_x u(\xi(s),s).$$

Proposition 3.4 implies that, for any $s \in (0,t)$,

$$|D_x u(\xi(s),s)| \leqq S(\xi(s),s;u_0)e^{C_1 s} + \beta(e^{C_1 s}-1).$$

By (3.10) we have $S(\xi(s), s; u_0) \leqq S(x, t; u_0)$, and therefore

$$|D_x u(\xi(s), s)| \leqq S(x, t; u_0) e^{C_1 s} + \beta(e^{C_1 s} - 1).$$

Sending $s \to t - 0$ yields the inequality in (3.11).

Let us take $(p, \tau) \in D^+ u(x, t)$. By Proposition 2.6 (2), we see that $(p, \tau) \in \operatorname{co} D^* u(x, t)$. From this and (3.11) we easily deduce the desired inequality of (2). Indeed, Carathéodory's theorem implies that there exist some $\{(q_k, \sigma_k)\}_{k=1}^{n+1} \subset D^* u(x, t)$ and $\{\lambda_k\}_{k=1}^{n+1} \subset [0, 1]$ such that $\sum_{k=1}^{n+1} \lambda_k = 1$ and $(p, \tau) = \sum_{k=1}^{n+1} \lambda_k (q_k, \sigma_k)$. In particular, $p = \sum_{k=1}^{n+1} \lambda_k q_k$. Thus, for $K \in \{1, \ldots, n+1\}$ such that $|q_K| = \max\{|q_1|, \ldots, |q_{n+1}|\}$, we have

$$|p| \leqq \sum_{k=1}^{n+1} \lambda_k |q_k| \leqq |q_K| \sum_{k=1}^{n+1} \lambda_k = |q_K| \leqq S(x, t; u_0) e^{C_1 t} + \beta(e^{C_1 t} - 1),$$

which completes the proof. □

Remark 3.7. Theorem 3.6 gives better estimates than the ones in Theorem 4.3, which applies to a more general Hamiltonian H. See Definition 4.1 and Remark 4.2.

Remark 3.8. It is clear that the upper bounds appearing in Theorem 3.6 are dominated by

$$\|Du_0\|_{L^\infty(\mathbf{R}^n)} \cdot e^{C_1 t} + \beta(e^{C_1 t} - 1).$$

This agrees with the upper bound in Theorem 2.10.

4. General Hamiltonian

Let us study a Hamiltonian H satisfying only (H1)–(H4). Following the way in Section 2.4, we approximate H by smooth Hamiltonians H_ε satisfying (H5) and $(\mathrm{H3})_{\mathrm{st}}$. We also recall that H_ε fulfills (H1) with the Lipschitz constant $A_2|x| + B_2 + \varepsilon$ (Remark 2.16). For this reason, we prepare the following:

$$R_\varepsilon(x, t) := \begin{cases} \left(\dfrac{B_2 + \varepsilon}{A_2} + |x| \right) (e^{A_2 t} - 1) & \text{if } A_2 > 0, \\ (B_2 + \varepsilon) t & \text{if } A_2 = 0, \end{cases}$$

and

$$S_\varepsilon(x, t; u_0) := \sup \left\{ |p| \;\middle|\; p \in D^-_{pr} u_0(y),\, y \in \overline{B_{R_\varepsilon(x,t)}(x)} \right\},$$
$$I_\varepsilon(x, t; u_0) := \inf \left\{ |p| \;\middle|\; p \in D^-_{pr} u_0(y),\, y \in \overline{B_{R_\varepsilon(x,t)}(x)} \right\}.$$

These will appear in gradient estimates for the viscosity solution u_ε of (1.6)–(1.2).

Besides, since the limiting process is carried out, we need the following bounds instead of $S(x, t; u_0)$ and $I(x, t; u_0)$ in Definition 3.3.

Definition 4.1. Let $(x, t) \in \mathbf{R}^n \times (0, T)$. We define

$$\overline{S}(x, t; u_0) := \lim_{\delta \to +0} \sup \left\{ |p| \;\middle|\; p \in D^-_{pr} u_0(y),\, y \in \overline{B_{R(x,t)+\delta}(x)} \right\},$$
$$\underline{I}(x, t; u_0) := \lim_{\delta \to +0} \inf \left\{ |p| \;\middle|\; p \in D^-_{pr} u_0(y),\, y \in \overline{B_{R(x,t)+\delta}(x)} \right\}.$$

Remark 4.2. It is clear that

$$\underline{I}(x, t; u_0) \leqq I(x, t; u_0) \leqq S(x, t; u_0) \leqq \overline{S}(x, t; u_0).$$

Moreover, Corollary 2.4 implies that we may replace "$D^{\pm}_{pr}u_0(y)$" by "$D^{\pm}u_0(y)$". Namely, we have

$$\overline{S}(x,t;u_0) = \lim_{\delta\to+0} \sup\left\{|p| \,\middle|\, p\in D^{-}u_0(y),\, y\in\overline{B_{R(x,t)+\delta}(x)}\right\},$$
$$\underline{I}(x,t;u_0) = \lim_{\delta\to+0} \inf\left\{|p| \,\middle|\, p\in D^{-}u_0(y),\, y\in\overline{B_{R(x,t)+\delta}(x)}\right\}.$$

The same replacement is possible for Definitions 5.3 and 5.10.

Theorem 4.3 (Gradient estimates II)**.** *Assume that H satisfies* (H1)–(H4)*. Let u be the viscosity solution of* (1.1)–(1.2)*. Let $(x,t)\in\mathbf{R}^n\times(0,T)$.*

(1) *If $p\in D^-_x u(x,t)$, then*
$$\underline{I}(x,t;u_0)e^{-C_1t}-\beta(1-e^{-C_1t}) \leqq |p| \leqq \overline{S}(x,t;u_0)e^{C_1t}+\beta(e^{C_1t}-1).$$
(2) *If $p\in D^+_x u(x,t)$, then $|p| \leqq \overline{S}(x,t;u_0)e^{C_1t}+\beta(e^{C_1t}-1)$.*

Proof. Let $\varepsilon\in(0,1]$ and H_ε be the approximate Hamiltonian defined in (2.9). Let u_ε be the viscosity solution of (1.6)–(1.2). By Proposition 2.17, u_ε converges to u locally uniformly in $\mathbf{R}^n\times[0,T)$ as $\varepsilon\to+0$.

Take an arbitrary $\delta>0$. Then, since $R_\varepsilon(y,s)\to R(x,t)$ as $(y,s,\varepsilon)\to(x,t,+0)$, there exists some $\theta\in(0,1]$ such that

$$R_\varepsilon(y,s) < R(x,t)+\delta \quad \text{for all } (y,s)\in B_\theta(x,t) \text{ and } \varepsilon\in(0,\theta). \tag{4.1}$$

(1) Suppose that $(p,\tau)\in D^-u(x,t)$. By Lemma 2.5 there exist sequences $\{(x_\varepsilon,t_\varepsilon)\}_{\varepsilon\in(0,1]}\subset\mathbf{R}^n\times(0,T)$ and $\{(p_\varepsilon,\tau_\varepsilon)\}_{\varepsilon\in(0,1]}\subset\mathbf{R}^n\times\mathbf{R}$ such that

$$(p_\varepsilon,\tau_\varepsilon)\in D^-u_\varepsilon(x_\varepsilon,t_\varepsilon),\quad (x_\varepsilon,t_\varepsilon)\to(x,t),\quad (p_\varepsilon,\tau_\varepsilon)\to(p,\tau)\quad \text{as } \varepsilon\to+0. \tag{4.2}$$

We now apply Theorem 3.6 to obtain

$$I_\varepsilon(x_\varepsilon,t_\varepsilon;u_0)e^{-C_1t_\varepsilon}-\beta(1-e^{-C_1t_\varepsilon}) \leqq |p_\varepsilon| \leqq S_\varepsilon(x_\varepsilon,t_\varepsilon;u_0)e^{C_1t_\varepsilon}+\beta(e^{C_1t_\varepsilon}-1).$$

For ε small enough we have $(x_\varepsilon,t_\varepsilon)\in B_\theta(x,t)$ and $\varepsilon\in(0,\theta)$, and so $R_\varepsilon(x_\varepsilon,t_\varepsilon)<R(x,t)+\delta$ by (4.1). This implies that

$$\begin{aligned}
&\inf\left\{|q| \,\middle|\, q\in D^-_{pr}u_0(y),\, y\in\overline{B_{R(x,t)+\delta}(x)}\right\}\cdot e^{-C_1t_\varepsilon}-\beta(1-e^{-C_1t_\varepsilon})\\
&\leqq |p_\varepsilon|\\
&\leqq \sup\left\{|q| \,\middle|\, q\in D^-_{pr}u_0(y),\, y\in\overline{B_{R(x,t)+\delta}(x)}\right\}\cdot e^{C_1t_\varepsilon}+\beta(e^{C_1t_\varepsilon}-1).
\end{aligned}$$

Sending $\varepsilon\to+0$ and $\delta\to+0$, we obtain the desired inequalities.

(2) The proof is the same as (1). ☐

For a given $x_0\in\mathbf{R}^n$ and $r>0$, let us define

$$\mathcal{E}(x_0,r) := \{(x,t)\in\mathbf{R}^n\times(0,T) \mid R(x,t)+|x-x_0|<r\}, \tag{4.3}$$

where $R(x,t)$ is the constant defined in (3.9). The set $\mathcal{E}(x_0,r)$ is the domain of dependence with the base $B_r(x_0)$ obtained in this paper. Indeed, we easily see that Theorem 4.3 is rephrased as follows under the assumption (U1):

Theorem 4.4 (Gradient estimates II′)**.** *Assume that H satisfies* (H1)–(H4)*. Let u be the viscosity solution of* (1.1)–(1.2)*. Let $x_0\in\mathbf{R}^n$, $r>0$ and assume* (U1)*. Then,*

$$|p| \geqq \theta e^{-C_1t}-\beta(1-e^{-C_1t}) \quad \textit{for all } (x,t)\in\mathcal{E}(x_0,r) \textit{ and } p\in D^-_x u(x,t), \tag{4.4}$$

where β and C_1 are the constants in (H1).

Theorem 4.4 does not need restriction of time by some t_0 as in Theorem 2.8 (1). Instead, when $\beta = 1$, the lower bound in (4.4) becomes negative after some time t_L. See Remark 6.2 and Theorem 6.3 for comparison between t_0 and t_L.

5. Homogeneous Hamiltonian

In this section, we study a Hamiltonian H satisfying the homogeneity (H6) in addition to (H1)–(H4). We further investigate m-homogeneous Hamiltonians; the precise assumption will be given as $(\mathrm{H6})_m$ in Section 5.3.

5.1. **Heuristic discussion.**

Suppose that H is a smooth Hamiltonian satisfying (H6). Then we have

$$\langle D_pH(x,t,p), p\rangle = H(x,t,p) \tag{5.1}$$

for all $(x,t,p) \in \mathbf{R}^n \times [0,T] \times \mathbf{R}^n$. In fact, differentiating both the sides of $H(x,t,\lambda p) = \lambda H(x,t,p)$ with respect to $\lambda > 0$, we find that $\langle D_pH(x,t,\lambda p), p\rangle = H(x,t,p)$. Letting $\lambda = 1$, we obtain (5.1).

Making a use of (5.1), one can observe that the viscosity solution u of (1.1)–(1.2) keeps its value along $\xi \in \mathcal{C}_{\min}(x,t)$ for each $(x,t) \in \mathbf{R}^n \times (0,T)$. To see this, take any $(p,\tau) \in D^*u(x,t)$ and let $(\xi,\eta) \in C^1([0,1])^2$ be the solution of (2.7)–(2.8). Let us show that

$$u(x,t) = u_0(\xi(0)). \tag{5.2}$$

Since $\xi \in \mathcal{C}_{\min}(x,t)$ by Theorem 2.14, it follows from Theorem 2.11 that $(\xi(s),s) \in \mathcal{R}(u)$ for all $s \in (0,t)$. Using (5.1), we compute

$$\begin{aligned}\frac{d}{ds}u(\xi(s),s) &= \langle D_xu(\xi(s),s), \xi'(s)\rangle + u_t(\xi(s),s)\\ &= \langle \eta(s), D_pH(\xi(s),s,\eta(s))\rangle + u_t(\xi(s),s)\\ &= H(\xi(s),s,\eta(s)) + u_t(\xi(s),s)\\ &= H(\xi(s),s,D_xu(\xi(s),s)) + u_t(\xi(s),s).\end{aligned} \tag{5.3}$$

Since u solves (1.1) at $(\xi(s),s)$ in a classical sense, the right-hand side (5.3) is equal to 0. This shows that

$$u(\xi(s),s) = u(\xi(0),0) = u_0(\xi(0)) \quad (s \in [0,t]).$$

Letting $s = t$, we obtain (5.2).

Since $\eta(0) \in D^-_{pr}u_0(\xi(0))$ by Theorem 2.14 (2), this observation suggests that the generalized gradients of u at (x,t) depend on $D^-_{pr}u_0(y)$ only for y such that $u(x,t) = u_0(y)$. In other words, the gradients depend on the initial gradients on the same level-set. In the next subsection, we prove that this is true via approximation of H by smooth H_ε given as (2.9).

We note that a smooth H satisfying (H6) should be linear in p. Accordingly, the approximation is crucial to study a general homogeneous Hamiltonian.

Remark 5.1. Let L be the Lagrangian associated with H. Then we have

$$H(x,t,p) + L(x,t,D_pH(x,t,p)) = \langle D_pH(x,t,p), p\rangle.$$

See, e.g., [11, Theorem A.2.5, (A.21)]. In the same manner as (5.3), we then deduce that

$$\frac{d}{ds}u(\xi(s),s) = L(\xi(s),s,\xi'(s)).$$

Thus, if there exists a bound $C > 0$ for $|L(\xi(s),s,\xi'(s))|$, one obtains the estimate $|u(x,t) - u_0(\xi(0))| \leqq Ct$, and this inequality gives gradient bounds for u depending on $D^-_{pr}u_0(y)$ for y such that $|u(x,t) - u_0(y)| \leqq Ct$. The gradient estimate in this direction, without (H6), is also interesting, but we do not pursue it anymore in this paper. Instead, we will carry out a similar calculation for m-homogeneous Hamiltonians; see Section 5.3.

5.2. **Gradient estimates.** By (5.1) it is naturally expected that

$$H_\varepsilon(x,t,p) \approx \langle D_p H_\varepsilon(x,t,p), p\rangle$$

for an approximate Hamiltonian H_ε. Let us give an error estimate for this.

Since Lipschitz regularity with respect to t is not required in our assumptions (H1)–(H4) and (H6), we hereafter assume that

$$H \text{ is locally Lipschitz continuous in } \mathbf{R}^n \times [0,T] \times \mathbf{R}^n \tag{5.4}$$

to guarantee that $D_p H(x,t,p)$ exists almost everywhere in $\mathbf{R}^n \times [0,T] \times \mathbf{R}^n$.

Lemma 5.2. *Assume that H satisfies* (H2), (H6) *and* (5.4). *Let $\varepsilon \in (0,1]$ and H_ε be the approximate Hamiltonian defined in* (2.9). *For $(x,t,p) \in \mathbf{R}^n \times [0,T] \times \mathbf{R}^n$ let us define*

$$E_\varepsilon(x,t,p) := H_\varepsilon(x,t,p) - \langle D_p H_\varepsilon(x,t,p), p\rangle. \tag{5.5}$$

Then,

$$|E_\varepsilon(x,t,p)| \leqq (A_2|x| + A_2\varepsilon + B_2 + 1)\varepsilon \tag{5.6}$$

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

Proof. For $h_\varepsilon(p) = \varepsilon\sqrt{|p|^2+1}$ we compute

$$h_\varepsilon(p) - \langle Dh_\varepsilon(p), p\rangle = \varepsilon\sqrt{|p|^2+1} - \left\langle \frac{\varepsilon p}{\sqrt{|p|^2+1}}, p \right\rangle = \frac{\varepsilon}{\sqrt{|p|^2+1}}.$$

We next study the function $(H * \rho_\varepsilon)(x,t,p)$. Let us write $z = (x,t,p) \in \mathbf{R}^{2n+1}$ and $w = (y,s,q) \in \mathbf{R}^{2n+1}$. As H is locally Lipschitz continuous (5.4), we have

$$D_p(H*\rho_\varepsilon)(z) = D_p\left(\int_{B_\varepsilon(0)} H(z-w)\rho_\varepsilon(w)\,dw\right) = \int_{B_\varepsilon(0)} D_p H(z-w)\rho_\varepsilon(w)\,dw,$$

and then

$$\begin{aligned}\langle D_p(H*\rho_\varepsilon)(z), p\rangle &= \int_{B_\varepsilon(0)} \langle D_p H(z-w), p\rangle \rho_\varepsilon(w)\,dw \\ &= \int_{B_\varepsilon(0)} \langle D_p H(z-w), p-q\rangle \rho_\varepsilon(w)\,dw + \int_{B_\varepsilon(0)} \langle D_p H(z-w), q\rangle \rho_\varepsilon(w)\,dw \\ &=: J_1 + J_2.\end{aligned}$$

Since (5.1) holds almost everywhere in $\mathbf{R}^{2n+1}$, we have

$$J_1 = \int_{B_\varepsilon(0)} H(z-w)\rho_\varepsilon(w)\,dw = (H*\rho_\varepsilon)(z).$$

Next, for any $w = (y,s,q) \in B_\varepsilon(0)$, the inner product in J_2 is dominated by

$$|D_p H(z-w)| \cdot |q| \leqq (A_2|x-y| + B_2)\varepsilon \leqq (A_2|x| + A_2\varepsilon + B_2)\varepsilon.$$

Thus,

$$\begin{aligned}|J_2| &\leqq \int_{B_\varepsilon(0)} |\langle D_p H(z-w), q\rangle| \rho_\varepsilon(w)\,dw \leqq (A_2|x| + A_2\varepsilon + B_2)\varepsilon \int_{B_\varepsilon(0)} \rho_\varepsilon(w)\,dw \\ &= (A_2|x| + A_2\varepsilon + B_2)\varepsilon. \end{aligned} \tag{5.7}$$

Summarizing the above calculations, we are led to

$$\begin{aligned}E_\varepsilon(x,t,p) &= (H*\rho_\varepsilon)(x,t,p) + h_\varepsilon(p) - \langle D_p(H*\rho_\varepsilon)(x,t,p), p\rangle - \langle Dh_\varepsilon(p), p\rangle \\ &= \frac{\varepsilon}{\sqrt{|p|^2+1}} - J_2,\end{aligned}$$

and

$$|E_\varepsilon(x,t,p)| \leqq \frac{\varepsilon}{\sqrt{|p|^2+1}} + |J_2| \leqq \varepsilon + (A_2|x| + A_2\varepsilon + B_2)\varepsilon = (A_2|x| + A_2\varepsilon + B_2 + 1)\varepsilon,$$

which is (5.6). □

The following bounds will appear in gradient estimates in this subsection.

Definition 5.3. Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and $\gamma \in \mathbf{R}$. We define

$$\begin{aligned}\overline{S}(x,t;u_0,\gamma) &= \lim_{\delta\to+0} \sup \left\{ |p| \,\middle|\, p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)+\delta}(x)},\ |u_0(y)-\gamma| \leqq \delta \right\}, \\ \underline{I}(x,t;u_0,\gamma) &= \lim_{\delta\to+0} \inf \left\{ |p| \,\middle|\, p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)+\delta}(x)},\ |u_0(y)-\gamma| \leqq \delta \right\}.\end{aligned}$$

Also, in the proof of the gradient estimates we solve the approximate Hamiltonian system:

$$\begin{cases} \text{(a) } \xi_\varepsilon'(s) = D_pH_\varepsilon(\xi_\varepsilon(s), s, \eta_\varepsilon(s)), \\ \text{(b) } \eta_\varepsilon'(s) = -D_xH_\varepsilon(\xi_\varepsilon(s), s, \eta_\varepsilon(s)). \end{cases} \tag{5.8}$$

Theorem 5.4 (Gradient estimates III)**.** *Assume that H satisfies* (H1)–(H4), (H6) *and* (5.4). *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and set $\gamma := u(x,t)$.*

(1) *If $p \in D^-_x u(x,t)$, then*

$$\underline{I}(x,t;u_0,\gamma)e^{-C_1t} - \beta(1-e^{-C_1t}) \leqq |p| \leqq \overline{S}(x,t;u_0,\gamma)e^{C_1t} + \beta(e^{C_1t}-1).$$

(2) *If $p \in D^+_x u(x,t)$, then $|p| \leqq \overline{S}(x,t;u_0,\gamma)e^{C_1t} + \beta(e^{C_1t}-1)$.*

Proof. The idea of the proof is similar to that of Theorem 4.3, but we need an additional error estimate for $|u_\varepsilon(x_\varepsilon,t_\varepsilon) - u(\xi_\varepsilon(0))|$, where ξ_ε is given below. We only prove (1) since (2) is shown in a similar manner.

(1) Let $\varepsilon \in (0,1]$ and H_ε be the approximate Hamiltonian defined in (2.9). Let u_ε be the viscosity solution of (1.6)–(1.2). Take $(p,\tau) \in D^-u(x,t)$ arbitrarily. Then (4.2) holds for some $\{(x_\varepsilon,t_\varepsilon)\}_{\varepsilon\in(0,1]} \subset \mathbf{R}^n \times (0,T)$ and $\{(p_\varepsilon,\tau_\varepsilon)\}_{\varepsilon\in(0,1]} \subset \mathbf{R}^n \times \mathbf{R}$. Since $(p_\varepsilon,\tau_\varepsilon) \in D^-u_\varepsilon(x_\varepsilon,t_\varepsilon)$ and u_ε is semiconcave, Proposition 2.6 (1) implies that u_ε is differentiable at $(x_\varepsilon,t_\varepsilon)$, i.e., $(x_\varepsilon,t_\varepsilon) \in \mathcal{R}(u_\varepsilon)$. We let $(\xi_\varepsilon,\eta_\varepsilon) \in C^1([0,1])^2$ be the solution of (5.8) with the terminal condition

$$\xi_\varepsilon(t_\varepsilon) = x_\varepsilon, \quad \eta_\varepsilon(t_\varepsilon) = p_\varepsilon = D_xu_\varepsilon(x_\varepsilon,t_\varepsilon).$$

Note that $(\xi_\varepsilon(s), s) \in \mathcal{R}(u_\varepsilon)$ $(s \in (0,t_\varepsilon))$ by Theorem 2.11. In particular,

$$(u_\varepsilon)_t(\xi_\varepsilon(s),s) + H_\varepsilon(\xi_\varepsilon(s), s, D_xu_\varepsilon(\xi_\varepsilon(s),s)) = 0.$$

Moreover, $\eta_\varepsilon(0) \in D^-_{pr}u_0(\xi_\varepsilon(0))$ by Theorem 2.14 (2).

Let E_ε be the function defined in (5.5). Then, by a similar calculation to (5.3), we have

$$\begin{aligned}\frac{d}{ds}u_\varepsilon(\xi_\varepsilon(s),s) &= \langle D_x u_\varepsilon(\xi_\varepsilon(s),s), \xi_\varepsilon'(s)\rangle + (u_\varepsilon)_t(\xi_\varepsilon(s),s)\\ &= \langle \eta_\varepsilon(s), D_p H_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s))\rangle + (u_\varepsilon)_t(\xi_\varepsilon(s),s)\\ &= -E_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s)) + H_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s)) + (u_\varepsilon)_t(\xi_\varepsilon(s),s)\\ &= -E_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s)) + H_\varepsilon(\xi_\varepsilon(s),s,D_x u_\varepsilon(\xi_\varepsilon(s),s)) + (u_\varepsilon)_t(\xi_\varepsilon(s),s)\\ &= -E_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s)). \end{aligned} \tag{5.9}$$

Integrating both the sides over $[0,t_\varepsilon]$, we find that

$$u_\varepsilon(x_\varepsilon,t_\varepsilon) = u_0(\xi_\varepsilon(0)) - \int_0^{t_\varepsilon} E_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s))\,ds.$$

Now, we see by (3.10) that

$$\xi_\varepsilon(s) \in \overline{B_{R_\varepsilon(x_\varepsilon,t_\varepsilon)}(x_\varepsilon)} \quad (s\in[0,t_\varepsilon]). \tag{5.10}$$

This implies that there exists some $K>0$ independent of ε and s such that $|\xi_\varepsilon(s)| \leqq K$ for all $\varepsilon\in(0,1]$ and $s\in[0,t_\varepsilon]$. Then, by Lemma 5.2

$$|E_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s))| \leqq (A_2|\xi_\varepsilon(s)| + A_2\varepsilon + B_2 + 1)\varepsilon \leqq (A_2K + A_2\varepsilon + B_2 + 1)\varepsilon,$$

and therefore

$$|u_\varepsilon(x_\varepsilon,t_\varepsilon) - u_0(\xi_\varepsilon(0))| \leqq \int_0^{t_\varepsilon} |E_\varepsilon(\xi_\varepsilon(s),s,\eta_\varepsilon(s))|\,ds \leqq (A_2K + A_2\varepsilon + B_2 + 1)\varepsilon t_\varepsilon. \tag{5.11}$$

Take any $\delta>0$ and choose ε small enough that

$$|x - x_\varepsilon| < \frac{\delta}{2}, \quad R_\varepsilon(x_\varepsilon,t_\varepsilon) < R(x,t) + \frac{\delta}{2} \tag{5.12}$$

(see (4.1)) and

$$|u_\varepsilon(x_\varepsilon,t_\varepsilon) - u(x,t)| < \frac{\delta}{2}, \quad (A_2K + A_2\varepsilon + B_2 + 1)\varepsilon t_\varepsilon < \frac{\delta}{2}. \tag{5.13}$$

Using (5.10) and (5.12), we have

$$|x - \xi_\varepsilon(0)| \leqq |x - x_\varepsilon| + |x_\varepsilon - \xi_\varepsilon(0)| < \frac{\delta}{2} + R_\varepsilon(x_\varepsilon,t_\varepsilon) < R(x,t) + \delta. \tag{5.14}$$

In addition, (5.11) and (5.13) imply that

$$|u_0(\xi_\varepsilon(0)) - \gamma| \leqq |u_0(\xi_\varepsilon(0)) - u_\varepsilon(x_\varepsilon,t_\varepsilon)| + |u_\varepsilon(x_\varepsilon,t_\varepsilon) - u(x,t)| < \delta. \tag{5.15}$$

In view of (3.3) we have

$$|\eta_\varepsilon(0)|e^{-C_1t_\varepsilon} - \beta(1 - e^{-C_1t_\varepsilon}) \leqq |p_\varepsilon| \leqq |\eta_\varepsilon(0)|e^{C_1t_\varepsilon} + \beta(e^{C_1t_\varepsilon} - 1).$$

Recalling that $\eta_\varepsilon(0) \in D^-_{pr}u_0(\xi_\varepsilon(0))$, we deduce from (5.14) and (5.15) that

$$\begin{aligned}&\inf\left\{|q| \;\middle|\; q\in D^-_{pr}u_0(y),\, y\in\overline{B_{R(x,t)+\delta}(x)},\, |u_0(y)-\gamma|\leqq\delta\right\}\cdot e^{-C_1t_\varepsilon} - \beta(1-e^{-C_1t_\varepsilon})\\ &\leqq |p_\varepsilon|\\ &\leqq \sup\left\{|q| \;\middle|\; q\in D^-_{pr}u_0(y),\, y\in\overline{B_{R(x,t)+\delta}(x)},\, |u_0(y)-\gamma|\leqq\delta\right\}\cdot e^{C_1t_\varepsilon} + \beta(e^{C_1t_\varepsilon}-1).\end{aligned}$$

Finally, we send $\varepsilon\to+0$ and $\delta\to+0$ to conclude the proof. $\square$

Under the assumption (U2) we have

Theorem 5.5 (Gradient estimates III′). *Assume that H satisfies* (H1)–(H4), (H6) *and* (5.4). *Let u be the viscosity solution of* (1.1)–(1.2). *Assume* (U2). *Then,*

$$|u(x,t)| + e^{C_1 t}|p| \geqq \theta - \beta(e^{C_1 t} - 1) \quad \textit{for all } (x,t) \in \mathbf{R}^n \times (0,T) \textit{ and } p \in D_x^- u(x,t).$$

In particular, if $\beta = 0$ in (H1), *then*

$$|u(x,t)| + e^{C_1 t}|p| \geqq \theta \quad \textit{for all } (x,t) \in \mathbf{R}^n \times (0,T) \textit{ and } p \in D_x^- u(x,t). \tag{5.16}$$

Proof. For any $(x,t) \in \mathbf{R}^n \times (0,T)$ we deduce from (U2) that

$$\underline{I}(x,t;u_0,u(x,t)) \geqq \theta - |u(x,t)|.$$

Applying this to the result in Theorem 5.4 (1), we obtain the desired inequality. □

Remark 5.6. The exponent of $|p|$ in Theorem 5.5 is different from the one in Theorem 2.8 (2), where $|p|^2$ appears. We also notice that, when $\beta = 0$, the initial lower bound θ is preserved globally-in-time in the sense of (5.16).

5.3. **m-homogeneous Hamiltonian.** The method in the previous subsection also applies to a Hamiltonian which is positively homogeneous of degree $m > 1$ with respect to p. We assume

(H6)$_m$ H is independent of t, and there exists some $m > 1$ such that $H(x,\lambda p) = \lambda^m H(x,p)$ for all $(x,p) \in \mathbf{R}^n \times \mathbf{R}^n$ and $\lambda \geqq 0$.

A typical Hamiltonian is the one given by (2.5).

We note that there is no Hamiltonian H satisfying both (H2) and (H6)$_m$. For this reason, we localize the assumption (H2) as follows:

(H2)$_L$ There exist $L > 0$ and $A_2, B_2 \geqq 0$ such that

$$|H(x,p) - H(x,q)| \leqq (A_2|x| + B_2)|p - q|$$

for all $(x,p) \in \mathbf{R}^n \times \overline{B_L(0)}$ and $q \in \overline{B_L(0)}$.

In the rest of this subsection, we define $R(x,t)$ for the constants A_2 and B_2 appearing in (H2)$_L$.

Assume that H is smooth and satisfies (H6)$_m$. In a similar way to (5.1), we then have

$$\langle D_p H(x,p), p\rangle = mH(x,p) \tag{5.17}$$

for all $(x,p) \in \mathbf{R}^n \times \mathbf{R}^n$. By a similar calculation to (5.3), we deduce from (5.17) that

$$\begin{aligned}
\frac{d}{ds}u(\xi(s),s) &= \langle D_x u(\xi(s),s), \xi'(s)\rangle + u_t(\xi(s),s) \\
&= \langle \eta(s), D_p H(\xi(s),\eta(s))\rangle + u_t(\xi(s),s) \\
&= mH(\xi(s),\eta(s)) + u_t(\xi(s),s) \\
&= (m-1)H(\xi(s),\eta(s)) + \{H(\xi(s), D_x u(\xi(s),s)) + u_t(\xi(s),s)\} \\
&= (m-1)H(\xi(s),\eta(s)).
\end{aligned} \tag{5.18}$$

The last term is not zero but it does not depend on s. In fact, since the Hamiltonian H is now independent of t, it is constant along $(\xi(s),\eta(s))$, and therefore we have $H(\xi(s),\eta(s)) = H(\xi(0),\eta(0))$ and

$$u(x,t) = u_0(\xi(0)) + (m-1)tH(\xi(0),\eta(0)).$$

This observation implies that the gradients of u depend on the initial gradients *near* the same level-set. When the Hamiltonian H is smooth and strictly convex, the above argument works without any approximation. Even for a general Hamiltonian H, we are able to make the above discussion rigorous via an approximation of H and some error estimates, but we omit most part of the proofs since they are parallel to the proof in the previous subsection.

We first state the result for a smooth and strictly convex Hamiltonian.

Definition 5.7. Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and $\gamma \in \mathbf{R}$. We define

$$S^m(x,t;u_0,\gamma) = \sup\left\{ |p| \;\middle|\; \begin{array}{l} p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)}(x)}, \\ u_0(y) + (m-1)tH(y,p) = \gamma \end{array} \right\},$$

$$I^m(x,t;u_0,\gamma) = \inf\left\{ |p| \;\middle|\; \begin{array}{l} p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)}(x)}, \\ u_0(y) + (m-1)tH(y,p) = \gamma \end{array} \right\}.$$

Theorem 5.8 (Gradient estimates IV)**.** *Assume that H satisfies* (H1)*,* (H2)$_L$*,* (H3)$_{\mathrm{st}}$*,* (H4)*,* (H5) *and* (H6)$_m$*. Let u be the viscosity solution of* (1.1)–(1.2)*. Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and set $\gamma := u(x,t)$. Assume that L in* (H2)$_L$ *satisfies*

$$L \geqq e^{C_1 t}\|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t} - 1). \tag{5.19}$$

(1) *If $p \in D^-_x u(x,t)$, then*
$$I^m(x,t;u_0,\gamma)e^{-C_1 t} - \beta(1 - e^{-C_1 t}) \leqq |p| \leqq S^m(x,t;u_0,\gamma)e^{C_1 t} + \beta(e^{C_1 t} - 1).$$

(2) *If $p \in D^+_x u(x,t)$, then $|p| \leqq S^m(x,t;u_0,\gamma)e^{C_1 t} + \beta(e^{C_1 t} - 1)$.*

Proof. It suffices to check that the estimate (3.8) in Proposition 3.2 still holds in this case. Let (ξ,η) be the same one as Proposition 3.2. Then, by Theorem 2.10 and (5.19) we see that

$$|\eta(s)| = |D_x u(\xi(s),s)| \leqq e^{C_1 t}\|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t} - 1) \leqq L$$

for all $s \in (0,t)$. Thus (H2)$_L$ yields

$$|D_p H(\xi(s),\eta(s))| \leqq A_2|\xi(s)| + B_2.$$

Using this inequality, we again obtain (3.8). $\square$

We next investigate a Hamiltonian which is not necessarily smooth. The error estimate corresponding to Lemma 5.2 is

Lemma 5.9. *Assume that H satisfies* (H1)*,* (H2)$_L$ *and* (H6)$_m$*. Let $\varepsilon \in (0,1]$ and H_ε be the approximate Hamiltonian defined in* (2.9)*. For $(x,p) \in \mathbf{R}^n \times \mathbf{R}^n$ let us define*

$$E^m_\varepsilon(x,p) := mH_\varepsilon(x,p) - \langle D_p H_\varepsilon(x,p), p\rangle.$$

Let $L' \in (0,L)$. Then,

$$|E^m_\varepsilon(x,p)| \leqq (A_2|x| + A_2\varepsilon + B_2 + m + (m-1)|p|)\varepsilon \tag{5.20}$$

for all $\varepsilon \in (0, L - L')$ and $(x,p) \in \mathbf{R}^n \times B_{L'}(0)$.

Proof. First, we compute

$$mh_\varepsilon(p) - \langle Dh_\varepsilon(p), p\rangle = m\varepsilon\sqrt{|p|^2+1} - \left\langle \frac{\varepsilon p}{\sqrt{|p|^2+1}}, p \right\rangle = \frac{m\varepsilon}{\sqrt{|p|^2+1}} + \frac{(m-1)\varepsilon|p|^2}{\sqrt{|p|^2+1}}.$$

Let $\varepsilon \in (0, L - L')$, $z = (x,p) \in \mathbf{R}^n \times B_{L'}(0)$ and $w = (y,q) \in B_\varepsilon(0)$. Then $z - w \in \mathbf{R}^n \times B_L(0)$, and thus we deduce from (H2)$_L$ that

$$|D_p H(z - w)| \leqq A_2|x - y| + B_2$$

almost everywhere. Following a similar calculation to the proof of Lemma 5.2, we have

$$\begin{aligned}E_\varepsilon^m(x,p) &= m(H*\rho_\varepsilon)(x,p) + mh_\varepsilon(p) - \langle D_p(H*\rho_\varepsilon)(x,p),p\rangle - \langle Dh_\varepsilon(p),p\rangle \\ &= \frac{m\varepsilon}{\sqrt{|p|^2+1}} + \frac{(m-1)\varepsilon|p|^2}{\sqrt{|p|^2+1}} - J_2,\end{aligned}$$

where J_2 is the same one as in the proof of Lemma 5.2. Hence, by (5.7)

$$\begin{aligned}|E_\varepsilon^m(x,p)| &\leqq \frac{m\varepsilon}{\sqrt{|p|^2+1}} + \frac{(m-1)\varepsilon|p|^2}{\sqrt{|p|^2+1}} + |J_2| \leqq m\varepsilon + (m-1)\varepsilon|p| + (A_2|x| + A_2\varepsilon + B_2)\varepsilon \\ &= \{m + (m-1)|p| + A_2|x| + A_2\varepsilon + B_2\}\varepsilon,\end{aligned}$$

which completes the proof. □

Definition 5.10. Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and $\gamma \in \mathbf{R}$. We define

$$\begin{aligned}\overline{S}^m(x,t;u_0,\gamma) &= \lim_{\delta\to+0} \sup\left\{ |p| \;\middle|\; \begin{array}{l} p \in D_{pr}^- u_0(y),\ y \in \overline{B_{R(x,t)+\delta}(x)}, \\ |u_0(y) + (m-1)tH(y,p) - \gamma| \leqq \delta \end{array}\right\}, \\ \underline{I}^m(x,t;u_0,\gamma) &= \lim_{\delta\to+0} \inf\left\{ |p| \;\middle|\; \begin{array}{l} p \in D_{pr}^- u_0(y),\ y \in \overline{B_{R(x,t)+\delta}(x)}, \\ |u_0(y) + (m-1)tH(y,p) - \gamma| \leqq \delta \end{array}\right\}.\end{aligned}$$

Theorem 5.11 (Gradient estimates V). *Assume that H satisfies* (H1), $(\mathrm{H2})_L$, (H3), (H4) *and* $(\mathrm{H6})_m$. *Let u be the viscosity solution of* (1.1)–(1.2). *Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and set $\gamma := u(x,t)$. Assume that L in* $(\mathrm{H2})_L$ *satisfies*

$$L > e^{C_1 t}\|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t} - 1).$$

(1) *If $p \in D_x^- u(x,t)$, then*
$$\underline{I}^m(x,t;u_0,\gamma)e^{-C_1 t} - \beta(1 - e^{-C_1 t}) \leqq |p| \leqq \overline{S}^m(x,t;u_0,\gamma)e^{C_1 t} + \beta(e^{C_1 t} - 1).$$

(2) *If $p \in D_x^+ u(x,t)$, then $|p| \leqq \overline{S}^m(x,t;u_0,\gamma)e^{C_1 t} + \beta(e^{C_1 t} - 1)$.*

Proof. We use the same notations as the proof of Theorem 5.4. As in (5.9) and (5.18), we have

$$\begin{aligned}\frac{d}{ds}u_\varepsilon(\xi_\varepsilon(s),s) &= -E_\varepsilon^m(\xi_\varepsilon(s),\eta_\varepsilon(s)) + (m-1)H_\varepsilon(\xi_\varepsilon(s),\eta_\varepsilon(s)) \\ &= -E_\varepsilon^m(\xi_\varepsilon(s),\eta_\varepsilon(s)) + (m-1)H_\varepsilon(\xi_\varepsilon(0),\eta_\varepsilon(0)),\end{aligned}$$

where we have applied the fact that H_ε is constant along $(\xi_\varepsilon(s),\eta_\varepsilon(s))$. This shows that

$$u_\varepsilon(x_\varepsilon,t_\varepsilon) = u_0(\xi_\varepsilon(0)) + (m-1)t_\varepsilon H_\varepsilon(\xi_\varepsilon(0),\eta_\varepsilon(0)) - \int_0^{t_\varepsilon} E_\varepsilon^m(\xi_\varepsilon(s),\eta_\varepsilon(s))\,ds. \tag{5.21}$$

We now take a constant $L' > 0$ such that

$$L > L' > e^{C_1 t}\|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t} - 1).$$

Applying Theorem 2.10 to the solution u_ε of (1.6)–(1.2), we deduce that

$$|\eta_\varepsilon(s)| = |D_x u_\varepsilon(\xi_\varepsilon(s),s)| \leqq e^{C_1 t_\varepsilon}\|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t_\varepsilon} - 1)$$

for all $s \in (0,t_\varepsilon)$. Since $t_\varepsilon \to t$ as $\varepsilon \to +0$, we have

$$|\eta_\varepsilon(s)| < L'$$

for all $s \in (0, t_\varepsilon)$ provided that ε is sufficiently small. Thus the estimate (5.20) in Lemma 5.9 is applicable at $(x,p) = (\xi_\varepsilon(s), \eta_\varepsilon(s))$, and then

$$\begin{aligned}|E_\varepsilon^m(\xi_\varepsilon(s), \eta_\varepsilon(s))| &\leqq (A_2|\xi_\varepsilon(s)| + A_2\varepsilon + B_2 + m + (m-1)|\eta_\varepsilon(s)|)\varepsilon \\ &\leqq (A_2K + A_2\varepsilon + B_2 + m + (m-1)L')\varepsilon.\end{aligned}$$

The rest of the proof runs as before by combining the above estimate and (5.21). □

6. Comparison with the previous results in [22]

We compare our results with the ones in [22]. We prove that both the lower bound and the domain of dependence obtained in this paper are larger. Let us recall an elemental inequality

$$e^X > X + 1 \ (X \neq 0),$$

which will be used below.

6.1. **Lower bounds.** We recall the lower bound in (2.4). Since $t < t_0 \leqq T$, the lower bound satisfies

$$\tilde{\theta} e^{-\frac{5}{4}C_1 t} = e^{-\frac{5}{4}C_1 t}\sqrt{\theta^2 - 2\beta C_1 e^{\frac{5}{2}C_1 T} t_0} \leqq e^{-\frac{5}{4}C_1 t}\sqrt{\theta^2 - 2\beta C_1 e^{\frac{5}{2}C_1 t} t}.$$

We compare the right-hand side with our lower bound obtained in (4.4).

Let us prepare notations.

Definition 6.1. For $\theta > 0$ we define

$$l(t) = e^{-\frac{5}{4}C_1 t}\sqrt{\theta^2 - 2\beta C_1 e^{\frac{5}{2}C_1 t} t}, \quad L(t) = \theta e^{-C_1 t} - \beta(1 - e^{-C_1 t}),$$

where β and C_1 are the constants in (H1). When $\beta = 1$, we define $t_l, t_L > 0$ as the unique numbers such that $l(t_l) = 0$ and $L(t_L) = 0$.

Remark 6.2. We have $l(0) = L(0) = \theta$, and both l and L are decreasing with respect to t. Moreover, for the constant t_0 in Theorem 2.8 (1), we have $l(t_0) > 0$. Therefore, $t_0 < t_l$ when $\beta = 1$.

The next theorem shows that our result gives a sharper lower bound and applies for a longer time.

Theorem 6.3. *Let $\theta > 0$. Assume that $C_1 > 0$.*

(1) *If $\beta = 0$, then $l(t) < L(t)$ for all $t \in (0, \infty)$.*
(2) *If $\beta = 1$, then $t_l < t_L$ and $l(t) < L(t)$ for all $t \in (0, t_l]$.*

Proof. (1) This is obvious since $l(t) = \theta e^{-\frac{5}{4}C_1 t}$ and $L(t) = \theta e^{-C_1 t}$.

(2) We have

$$l(t) = e^{-\frac{5}{4}C_1 t}\sqrt{\theta^2 - 2C_1 e^{\frac{5}{2}C_1 t} t}, \quad L(t) = \theta e^{-C_1 t} - (1 - e^{-C_1 t}).$$

Since $l(t) \geqq 0$ for all $t \in (0, t_l]$, the assertions of (2) follow if we prove that $\{l(t)\}^2 < \{L(t)\}^2$ for all $t \in (0, t_l]$. Observe

$$\begin{aligned}\{l(t)\}^2 &= e^{-\frac{5}{2}C_1 t}\left(\theta^2 - 2C_1 e^{\frac{5}{2}C_1 t} t\right) = \theta^2 e^{-\frac{5}{2}C_1 t} - 2C_1 t, \\ \{L(t)\}^2 &= \theta^2 e^{-2C_1 t} - 2\theta e^{-C_1 t}(1 - e^{-C_1 t}) + (1 - e^{-C_1 t})^2\end{aligned}$$

and

$$\{L(t)\}^2 - \{l(t)\}^2 = \theta^2 e^{-2C_1 t}(1 - e^{-\frac{1}{2}C_1 t}) - 2\theta e^{-C_1 t}(1 - e^{-C_1 t}) + (1 - e^{-C_1 t})^2 + 2C_1 t.$$

Let us denote by Δ the discriminant of the above quadratic polynomial of θ. It suffices to prove that $\Delta < 0$. We compute

$$\begin{aligned}\frac{\Delta}{4} &= \{-e^{-C_1t}(1-e^{-C_1t})\}^2 - e^{-2C_1t}(1-e^{-\frac{1}{2}C_1t})\cdot\{(1-e^{-C_1t})^2+2C_1t\}\\ &= e^{-2C_1t}[(1-e^{-C_1t})^2-(1-e^{-\frac{1}{2}C_1t})\cdot\{(1-e^{-C_1t})^2+2C_1t\}]\\ &= e^{-2C_1t}\{e^{-\frac{1}{2}C_1t}(1-e^{-C_1t})^2-2C_1t(1-e^{-\frac{1}{2}C_1t})\}.\end{aligned}$$

Applying $e^{-C_1t} > -C_1t+1$, we see that

$$\begin{aligned}\frac{\Delta}{4} &< e^{-2C_1t}\{e^{-\frac{1}{2}C_1t}(C_1t)^2-2C_1t(1-e^{-\frac{1}{2}C_1t})\}\\ &= e^{-2C_1t}\cdot 2C_1t\left\{\left(\frac{1}{2}C_1t+1\right)e^{-\frac{1}{2}C_1t}-1\right\}.\end{aligned}$$

Since $\frac{1}{2}C_1t+1 < e^{\frac{1}{2}C_1t}$, we conclude that $\Delta < 0$. $\square$

6.2. **Domains of dependence.** We next investigate the domains of dependence $\mathcal{D}(x_0,r)$ in (2.3) and $\mathcal{E}(x_0,r)$ in (4.3), where gradient estimates are available. Let us begin with comparison between $r(x,t)$ and $R(x,t)$.

Proposition 6.4. *Let $(x,t)\in\mathbf{R}^n\times(0,T)$. Then*

$$R(x,t)=r(x,t)=0 \quad \text{if } (A_2,B_2)=(0,0), \qquad R(x,t)<r(x,t) \quad \text{if } (A_2,B_2)\neq(0,0).$$

Proof. By definition we easily see that $R(x,t)=r(x,t)=0$ when $(A_2,B_2)=(0,0)$. Assume that $(A_2,B_2)\neq(0,0)$. First, if $A_2=0$ and $B_2\neq 0$, we have

$$r(x,t)=e^{B_2t}-1>B_2t=R(x,t).$$

We next assume that $A_2\neq 0$. Observe

$$e^{A_2t}=\frac{A_2t}{A_2te^{-A_2t}}>\frac{1-e^{-A_2t}}{A_2te^{-A_2t}}=\frac{e^{A_2t}-1}{A_2t}.$$

Then,

$$\begin{aligned}r(x,t)&=e^{A_2t}\cdot e^{(B_2+A_2|x|)t}-1>\frac{e^{A_2t}-1}{A_2t}\cdot\{(B_2+A_2|x|)t+1\}-1\\ &=(e^{A_2t}-1)\left\{\left(\frac{B_2}{A_2}+|x|\right)+\frac{1}{A_2t}\right\}-1=R(x,t)+\frac{e^{A_2t}-1-A_2t}{A_2t}>R(x,t).\end{aligned}$$

The proof is complete. $\square$

The following theorem shows that our gradient estimates are obtained in a larger set.

Theorem 6.5. *Let $x_0\in\mathbf{R}^n$ and $r>0$. Then*

$$\mathcal{D}(x_0,r)=\mathcal{E}(x_0,r) \quad \text{if } (A_2,B_2)=(0,0), \qquad \mathcal{D}(x_0,r)\subsetneq\mathcal{E}(x_0,r) \quad \text{if } (A_2,B_2)\neq(0,0).$$

Proof. If $(A_2,B_2)=(0,0)$, we easily see that $\mathcal{D}(x_0,r)=\mathcal{E}(x_0,r)=B_r(x_0)\times(0,T)$. Assume that $(A_2,B_2)\neq(0,0)$. It suffices to prove that

$$\{(x,t)\in\mathbf{R}^n\times(0,T)\mid (r(x_0,t)+1)(|x-x_0|+1)-1\leqq r\}\subset\mathcal{E}(x_0,r),$$

where "$<r$" in (2.3) was replaced by "$\leqq r$". Take $(x,t)\in\mathbf{R}^n\times(0,T)$ satisfying

$$(r(x_0,t)+1)(|x-x_0|+1)-1\leqq r. \tag{6.1}$$

We have to demonstrate that $R(x,t)+|x-x_0|<r$. By (6.1) we have

$$r-(R(x,t)+|x-x_0|)\geqq r(x_0,t)(|x-x_0|+1)-R(x,t)=:J.$$

Let us prove that $J>0$. If $A_2=0$, we have

$$J=(e^{B_2t}-1)(|x-x_0|+1)-B_2t>B_2t(|x-x_0|+1)-B_2t\geqq 0.$$

Assume next that $A_2\neq 0$. When $|x|\leqq|x_0|$, we have $r(x,t)\leqq r(x_0,t)$ by the definition (2.2). Thus, Proposition 6.4 implies that

$$J\geqq r(x,t)(|x-x_0|+1)-R(x,t)>R(x,t)(|x-x_0|+1)-R(x,t)\geqq 0.$$

Suppose that $|x|>|x_0|$. Observe

$$\begin{aligned}J&=(e^{(A_2+B_2+A_2|x_0|)t}-1)(|x-x_0|+1)-\left(\frac{B_2}{A_2}+|x|\right)(e^{A_2t}-1)\\&\geqq(e^{(A_2+B_2+A_2|x_0|)t}-1)(|x|-|x_0|+1)-\left(\frac{B_2}{A_2}+|x|\right)(e^{A_2t}-1).\end{aligned}$$

Let us define $j(\rho)$ as the right-hand side above with $\rho=|x|$. Then $j(\rho)$ is an affine function of ρ, and the coefficient of ρ is

$$e^{(A_2+B_2+A_2|x_0|)t}-e^{A_2t}=e^{A_2t}(e^{(B_2+A_2|x_0|)t}-1)\geqq 0.$$

Hence $j(\rho)$ is nondecreasing. Moreover,

$$j(|x_0|)=(e^{(A_2+B_2+A_2|x_0|)t}-1)-\left(\frac{B_2}{A_2}+|x_0|\right)(e^{A_2t}-1)=r(x_0,t)-R(x_0,t)>0.$$

Summarizing the above, we see that $J\geqq j(|x|)\geqq j(|x_0|)>0$, which completes the proof. $\square$

Remark 6.6. Let us study some geometrical properties of the domains $\mathcal{D}(x_0,r)$ and $\mathcal{E}(x_0,r)$. To do so, we fix $x_0\in\mathbf{R}^n$, $r>0$ and assume that $(A_2,B_2)\neq(0,0)$. Figures 1, 2 and 3 show $\mathcal{D}(x_0,r)$ and $\mathcal{E}(x_0,r)$ for several A_2 and B_2 in the case where $|x_0|<r$. Let us explain why they are given as those.

We first note that the base of each domain, which means the intersection with the plane $t=0$, is $B_r(x_0)$ in common. However, they have different vertexes. Here a point $(\bar{x},\bar{t})\in\partial\mathcal{D}(x_0,r)$ is called a vertex of $\mathcal{D}(x_0,r)$ if $\bar{t}\geqq t$ for any $(x,t)\in\mathcal{D}(x_0,r)$. We define a vertex of $\mathcal{E}(x_0,r)$ in a similar manner.

Let us consider $\mathcal{D}(x_0,r)$, for which see also [22, Remark 6.1]. It is a cone-like domain. In fact, the condition $(r(x_0,t)+1)(|x-x_0|+1)-1<r$ in (2.3) is equivalent to

$$|x-x_0|<(r+1)e^{-(A_2+B_2+A_2|x_0|)t}-1.$$

This implies that $\mathcal{D}(x_0,r)$ is axially symmetric around the line $x=x_0$ and the vertex is

$$\left(x_0,\frac{1}{A_2+B_2+A_2|x_0|}\log(r+1)\right).$$

Contrary to $\mathcal{D}(x_0,r)$, our domain $\mathcal{E}(x_0,r)$ is not necessarily axially symmetric. Also, the x-coordinate of the vertex depends on the values of A_2 and B_2; in particular, it may not be x_0. Let us investigate these properties in more details.

If $A_2=0$, then $R(x,t)=B_2t$, and so $(x,t)\in\mathcal{E}(x_0,r)$ if and only if $|x-x_0|<r-B_2t$. Accordingly, $\mathcal{E}(x_0,r)$ is a cone with the vertex $(x_0,\frac{r}{B_2})$. See Figure 1.

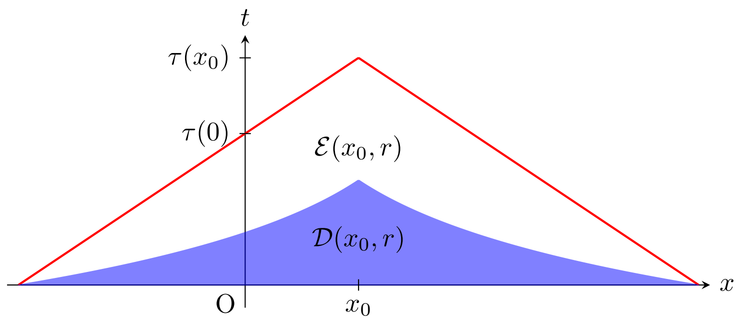


FIGURE 1. $A_2 = 0$.

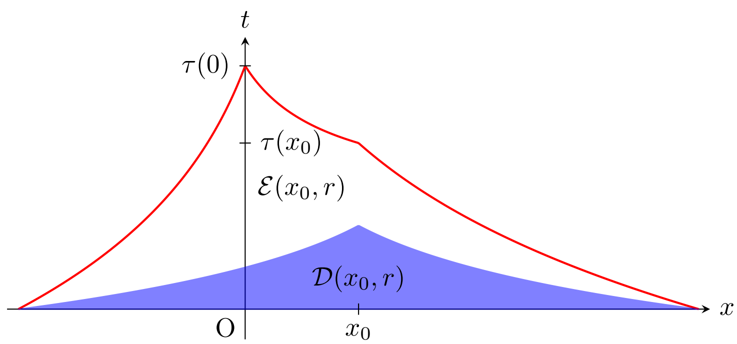


FIGURE 2. $A_2(r - |x_0|) > B_2$.

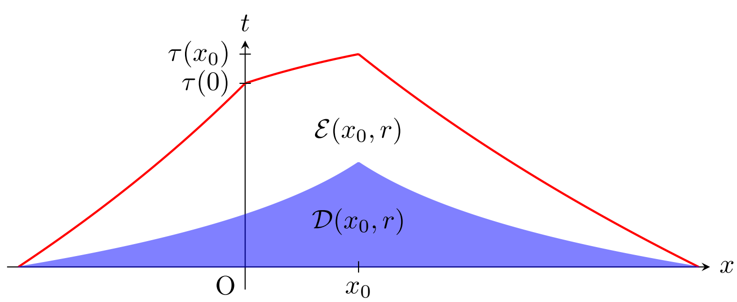


FIGURE 3. $A_2(r - |x_0|) < B_2$.

Assume that $A_2 > 0$. To find the vertex of $\mathcal{E}(x_0, r)$, let us define $\tau(x) > 0$ for $x \in B_r(x_0)$ as the unique number such that $R(x, \tau(x)) + |x - x_0| = r$, that is,

$$\tau(x) = \frac{1}{A_2} \log \left(\frac{A_2(r - |x - x_0|)}{B_2 + A_2|x|} + 1 \right). \tag{6.2}$$

It suffices to look for the vertex in points $(x, \tau(x)) \in \partial\mathcal{E}(x_0, r)$ for $x \in B_r(x_0)$. Our purpose is thus to find a maximizer of

$$\max_{x \in B_r(x_0)} \tau(x). \tag{6.3}$$

When $x_0 = 0$, the definition (4.3) implies that $\mathcal{E}(x_0, r)$ is axially symmetric around the line $x = 0$. Moreover, by (6.2) we easily see that the maximum of (6.3) is attained only at 0. In consequence, the vertex of $\mathcal{E}(x_0, r)$ is $(0, \tau(0))$.

Let $x_0 \neq 0$ in what follows. Then $\mathcal{E}(x_0, r)$ is not axially symmetric by (4.3). We note that, for every $\rho > 0$, the minimum of $\min_{x \in \partial B_\rho(0)} |x - x_0|$ is attained only at $y_\rho := \frac{\rho x_0}{|x_0|}$. This fact and (6.2) imply that the maximizer of $\max_{x \in \partial B_\rho(0)} \tau(x)$ is y_ρ. Therefore the maximizing problem (6.3) is reduced to the one dimensional one over $I := [(|x_0| - r)_+, |x_0| + r]$, that is,

$$\max_{\rho \in I} \tau(y_\rho), \tag{6.4}$$

where we set $y_0 := 0$ and $a_+ = \max\{a, 0\}$ denotes the plus part of $a \in \mathbf{R}$.

Let $\rho \in I$. Since $|y_\rho| = \rho$ and $|y_\rho - x_0| = |\rho - |x_0||$, we have

$$\begin{aligned}
\tau(y_\rho) &= \frac{1}{A_2} \log\left(\frac{A_2(r - |\rho - |x_0||)}{B_2 + A_2\rho} + 1\right) \\
&= \begin{cases} \dfrac{1}{A_2} \log\left(\dfrac{-B_2 + A_2(r - |x_0|)}{B_2 + A_2\rho} + 2\right) & \text{if } \rho \in [(|x_0| - r)_+, |x_0|], \\ \dfrac{1}{A_2} \log \dfrac{B_2 + A_2(r + |x_0|)}{B_2 + A_2\rho} & \text{if } \rho \in [|x_0|, |x_0| + r]. \end{cases}
\end{aligned}$$

Let us study monotonicity of $\tau(y_\rho)$ as a function of ρ.

Case 1: $|x_0| \geqq r$, i.e., $0 \notin B_r(x_0)$. Then $\tau(y_\rho)$ is increasing in $[(|x_0| - r)_+, |x_0|]$ and decreasing in $[|x_0|, |x_0| + r]$. Thus the maximum of (6.4) is attained at $\rho = |x_0|$, and hence the vertex of $\mathcal{E}(x_0, r)$ is $(x_0, \tau(x_0))$.

Case 2: $|x_0| < r$, i.e., $0 \in B_r(x_0)$. Then $\tau(y_\rho)$ is decreasing in $[|x_0|, |x_0| + r]$, while the monotonicity in $[(|x_0| - r)_+, |x_0|]$ depends on the sign of $-B_2 + A_2(r - |x_0|)$. Namely, in $[(|x_0| - r)_+, |x_0|]$,

$$\tau(y_\rho) \text{ is } \begin{cases} \text{decreasing} & \text{if } A_2(r - |x_0|) > B_2, \\ \text{constant} & \text{if } A_2(r - |x_0|) = B_2, \\ \text{increasing} & \text{if } A_2(r - |x_0|) < B_2. \end{cases}$$

Thus it follows that the maximum of (6.4) is attained at $\rho = 0$, $\rho \in [0, |x_0|]$ and $\rho = |x_0|$ if $A_2(r - |x_0|) > B_2$, $A_2(r - |x_0|) = B_2$ and $A_2(r - |x_0|) < B_2$, respectively. We therefore conclude that the vertex of $\mathcal{E}(x_0, r)$ is

$$\begin{cases} (0, \tau(0)) & \text{if } A_2(r - |x_0|) > B_2, \\ (1 - \mu)(0, \tau(0)) + \mu(x_0, \tau(x_0)) \text{ for any } \mu \in [0, 1] & \text{if } A_2(r - |x_0|) = B_2, \\ (x_0, \tau(x_0)) & \text{if } A_2(r - |x_0|) < B_2. \end{cases}$$

See Figures 2 and 3 for the case $A_2(r - |x_0|) > B_2$ and $A_2(r - |x_0|) < B_2$, respectively.

7. Examples

We show some examples of viscosity solutions to illustrate our results.

Example 7.1. We begin with a linear Hamiltonian $H(x, p) = -\langle x, p\rangle$. Then H satisfies (H1) with $C_1 = 1$, $\beta = 0$ and (H2) with $A_2 = 1$, $B_2 = 0$. We thus have $R(x, t) = |x|(e^t - 1)$. Also, H satisfies (H3), (H5), (H6) and (5.4). Though H does not satisfy (H4), one can apply Theorems 4.3 and 5.4 by modifying H outside $\overline{B_M(0)} \times \mathbf{R}^n$ for a large $M > 0$ so that (H4) is fulfilled. For example, let $\mathbf{X} : \mathbf{R}^n \to \mathbf{R}^n$ be a bounded and Lipschitz continuous function in $\mathbf{R}^n$ such that $\|D\mathbf{X}\|_{L^\infty(\mathbf{R}^n)} = 1$ and

$$\mathbf{X}(x) = x \quad \text{for all } x \in \overline{B_M(0)}.$$

Then the new Hamiltonian $\tilde{H}(x,p) = -\langle \mathbf{X}(x), p\rangle$ satisfies (H4). In what follows, however, we consider the original H in order to simplify our presentation.

The equation is

$$u_t(x,t) - \langle x, D_x u(x,t)\rangle = 0 \quad \text{in } \mathbf{R}^n \times (0,T), \tag{7.1}$$

which is the transport equation. It is easily seen that the viscosity solution of (7.1)–(1.2) is given by

$$u(x,t) = u_0(xe^t). \tag{7.2}$$

We therefore have

$$D_x^\pm u(x,t) = e^t D^\pm u_0(xe^t) = \{e^t p \mid p \in D^\pm u_0(xe^t)\} \quad \text{for all } (x,t) \in \mathbf{R}^n \times (0,T).$$

This means that $D_x^\pm u(x,t)$ depend only on the gradients of u_0 at xe^t. Moreover, (7.2) implies that the gradients $D_x^\pm u(x,t)$ depend only on the gradients of u_0 on the same level-set.

(i) Assume that $u_0 \in C^1(\mathbf{R}^n)$ and $D_{pr}^- u_0(x) = \{Du_0(x)\}$ for all $x \in \mathbf{R}^n$. Let $(x,t) \in \mathbf{R}^n \times (0,T)$. Then

$$D_x^\pm u(x,t) = \{e^t Du_0(xe^t)\}.$$

In addition, since

$$|xe^t - x| = |x|(e^t - 1) = R(x,t), \tag{7.3}$$

we have $xe^t \in \overline{B_{R(x,t)}(x)}$. Accordingly, it follows that

$$e^t|Du_0(xe^t)| \leqq \overline{S}(x,t;u_0,u(x,t))e^t \leqq S(x,t;u_0)e^t,$$

which correspond to the upper estimates in Theorems 4.3 and 5.4. Moreover, we see by (7.3) that the radius $R(x,t)$ is optimal to get the gradient estimates. If $S(x,t;u_0) = |Du_0(xe^t)|$, the inequalities above become equalities, and hence the upper estimates are optimal.

(ii) One may wonder if it is possible to replace $\overline{S}(x,t;u_0,\gamma)$ and $\underline{I}(x,t;u_0,\gamma)$ in Theorem 5.4 by

$$S(x,t;u_0,\gamma) = \sup\left\{|p| \;\middle|\; p \in D_{pr}^- u_0(y),\, y \in \overline{B_{R(x,t)}(x)},\, u_0(y) = \gamma\right\},$$
$$I(x,t;u_0,\gamma) = \inf\left\{|p| \;\middle|\; p \in D_{pr}^- u_0(y),\, y \in \overline{B_{R(x,t)}(x)},\, u_0(y) = \gamma\right\},$$

respectively, but it is impossible. Let $\alpha \in (0,1)$, $K > 1$ and

$$u_0(x) = \max\{-|x|^{1+\alpha},\, -K\}.$$

Let u be the solution of (7.1)–(1.2), which is given by (7.2). Fix $t \in (0,T)$, and we consider the gradients $D_x^\pm u(0,t)$. By (7.2) we have $u(0,t) = u_0(0) = 0$. From the definition of u_0 we deduce that, if $y \in \overline{B_{R(0,t)}(0)}$ and $u_0(y) = 0$, then $y = 0$. However, as we noted in Remark 2.2, we have $D_{pr}^- u_0(0) = \emptyset$. Therefore the set

$$\left\{|p| \;\middle|\; p \in D_{pr}^- u_0(y),\, y \in \overline{B_{R(x,t)}(x)},\, u_0(y) = 0\right\}$$

is empty. Thus the definitions of $S(x,t;u_0,0)$ and $I(x,t;u_0,0)$ do not make sense.

(iii) If the equation is

$$u_t(x,t) + \langle x, D_x u(x,t)\rangle = 0 \quad \text{in } \mathbf{R}^n \times (0,T), \tag{7.4}$$

which has the opposite sign to (7.1), then the solution u of (7.4)–(1.2) is given by

$$u(x,t) = u_0(xe^{-t}).$$

Fix $(x,t) \in \mathbf{R}^n \times (0,T)$. Under the same regularity conditions on u_0 as **(i)**, one has $D_x^{\pm} u(x,t) = \{e^{-t} Du_0(xe^{-t})\}$ and

$$I(x,t;u_0)e^{-t} \leqq \underline{I}(x,t;u_0,u(x,t))e^{-t} \leqq e^{-t}|Du_0(xe^{-t})|.$$

These are lower estimates of the same forms as Theorems 4.3 and 5.4. If $I(x,t;u_0) = |Du_0(xe^{-t})|$, they are optimal estimates.

Example 7.2. This example reveals a difference among Theorems 3.6, 4.3 and 5.4 when the Hamiltonian H is not smooth or strictly convex. Let $H(p) = c|p|$, where $c > 0$ is a constant. Then H satisfies (H1) with $C_1 = 0$, $\beta = 0$, and (H2) with $A_2 = 0$, $B_2 = c$. Moreover, we have $R(x,t) = ct$. The corresponding equation is the eikonal equation:

$$u_t(x,t) + c|D_x u(x,t)| = 0 \quad \text{in } \mathbf{R}^n \times (0,T). \tag{7.5}$$

The viscosity solution of (7.5)–(1.2) is given by

$$u(x,t) = \min_{y \in \overline{B_{ct}(x)}} u_0(y). \tag{7.6}$$

This formula can be derived by regarding (7.5) as a Bellman equation and applying the theory of optimal controls ([12, Chapter 10]). Since H satisfies (H1)–(H4), (H6) and (5.4), Theorems 4.3 and 5.4 are applicable.

(i) Let us consider the initial datum

$$u_0(x) = \min\{(2 - |x|)_+,\, 1\}, \tag{7.7}$$

where we recall that $(\cdot)_+$ stands for the plus part. The proximal subgradients of u_0 are given as follows:

$$D^-_{pr} u_0(x) = \begin{cases} \{0\} & \text{if } |x| < 1 \text{ or } |x| > 2, \\ \emptyset & \text{if } |x| = 1. \\ \left\{ -\dfrac{x}{|x|} \right\} & \text{if } 1 < |x| < 2, \\ \left\{ -\dfrac{sx}{|x|} \;\middle|\; s \in [0,1] \right\} & \text{if } |x| = 2. \end{cases} \tag{7.8}$$

By (7.6) we find that

$$u(x,t) = \min\{(2 - ct - |x|)_+,\, 1\}.$$

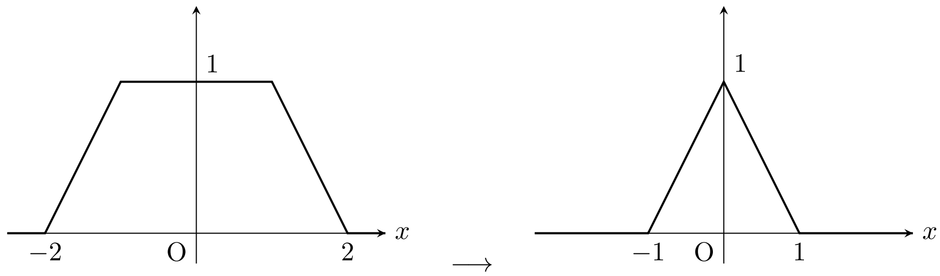
In particular, $u(x, \frac{1}{c}) = (1 - |x|)_+$. The graphs of $u_0(x)$ and $u(x, \frac{1}{c})$ are shown in Figure 4. Note that $t = \frac{1}{c}$ is the first time that $u(x,t)$ is non-differentiable at $x = 0$.

Let us consider the supergradient $D_x^+ u(0, \frac{1}{c})$. It is clear that $D_x^+ u(0, \frac{1}{c}) = \overline{B_1(0)}$. Furthermore, we notice that $R(0, \frac{1}{c}) = 1$ and

$$\overline{S}\left(0, \frac{1}{c}; u_0\right) = \lim_{\delta \to +0} \sup \left\{ |p| \;\middle|\; p \in D^-_{pr} u_0(y),\, y \in \overline{B_{1+\delta}(0)} \right\} = 1.$$

Therefore we deduce that

$$|p| \leqq \overline{S}\left(0, \frac{1}{c}; u_0\right) \quad \text{for any } p \in D_x^+ u\left(0, \frac{1}{c}\right), \tag{7.9}$$



(A) $u_0(x) = \min\{(2-|x|)_+,\, 1\}$. (B) $u(x, \frac{1}{c}) = (1-|x|)_+$.

Figure 4. Solution of (7.5).

which is the assertion of Theorem 4.3 (2). In addition, (7.9) is an optimal estimate in the sense that there is $p \in D_x^+ u(0, \frac{1}{c})$ such that $|p| = 1$. Also, we easily see that

$$\overline{S}\left(0, \frac{1}{c}; u_0, u\left(0, \frac{1}{c}\right)\right) = \overline{S}\left(0, \frac{1}{c}; u_0, 1\right) = 1,$$

and hence $\overline{S}(0, \frac{1}{c}; u_0)$ in (7.9) can be replaced by $\overline{S}(0, \frac{1}{c}; u_0, 1)$. This is what Theorem 5.4 (2) asserts.

We note that H satisfies neither (H5) nor $(\mathrm{H3})_{\mathrm{st}}$, and thus Theorem 3.6 cannot be applied. In fact, it is not possible to replace $\overline{S}(0, \frac{1}{c}; u_0)$ in (7.9) by $S(0, \frac{1}{c}; u_0)$ since we have $S(0, \frac{1}{c}; u_0) = 0$. This means that the assertion of Theorem 3.6 (2) fails in the present case.

Finally, we mention the optimality of $R(0, \frac{1}{c}) = 1$. If $R \in (0, 1)$, then

$$\overline{S}_R := \lim_{\delta \to +0} \sup \left\{ |p| \;\middle|\; p \in D^-_{pr} u_0(y),\, y \in \overline{B_{R+\delta}(0)} \right\} = 0,$$

which shows that (7.9) does not hold if we replace $\overline{S}(0, \frac{1}{c}; u_0)$ by $\overline{S}_R$ above.

(ii) We change the initial datum to

$$u_0(x) = \min\{(2-|x|)_+,\, 1\} + f(x),$$

where $f \in C^1(\mathbf{R}^n)$ is a nonnegative function such that $\operatorname{supp} f \subset B_1(0)$ and $\sup_{B_1(0)} |Df| > 1$. Let u be the solution of (7.5)–(1.2). Then we have $u(x, \frac{1}{c}) = (1-|x|)_+$ again, but the estimate (7.9), which is the result of Theorem 4.3 (2), becomes worse. Indeed, due to the additional term f, we have

$$\overline{S}\left(0, \frac{1}{c}; u_0\right) = \sup_{B_1(0)} |Df| > 1.$$

In contrast, the estimate in Theorem 5.4 (2) still keeps the optimality. In fact, since we only need to see the level-sets of u_0 close to 1-level, we find that $\overline{S}(0, \frac{1}{c}; u_0, 1) = 1$.

Example 7.3. Let us next study the case where the Hamiltonian H is smooth and strictly convex. Let $H(p) = \frac{1}{2}|p|^2$. Then H satisfies (H1) with $C_1 = 0$, $\beta = 0$, $(\mathrm{H3})_{\mathrm{st}}$, (H4), (H5) and $(\mathrm{H6})_m$ with $m = 2$. Since (H2) is not fulfilled, we will modify H or localize the assumption later.

The equation associated with H is

$$u_t(x,t) + \frac{1}{2}|D_x u(x,t)|^2 = 0 \quad \text{in } \mathbf{R}^n \times (0, T). \tag{7.10}$$

The viscosity solution of (7.10)–(1.2) is given by the *inf-convolution*:

$$u(x,t) = \inf_{y\in\mathbf{R}^n}\left\{u_0(y) + \frac{1}{2t}|x-y|^2\right\}. \tag{7.11}$$

This is a special case of the Hopf-Lax formula ([11], [12, Chapters 3 and 10]).

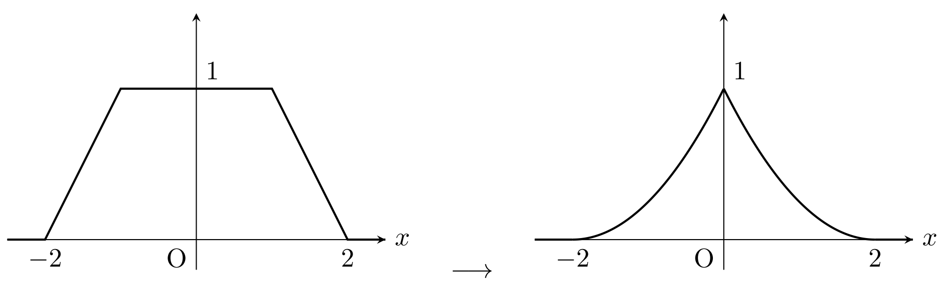
Let us consider the initial datum u_0 given by (7.7). By a straightforward calculation we find that, when $t \geqq 2$, all $y \in B_2(0)$ can be removed from the range of the infimum in (7.11). Namely,

$$u(x,t) = \inf_{y\in\mathbf{R}^n,\,|y|\geqq 2}\left\{u_0(y) + \frac{1}{2t}|x-y|^2\right\} = \begin{cases} \frac{1}{2t}(|x|-2)^2 & \text{if } |x| \leqq 2, \\ 0 & \text{if } |x| \geqq 2 \end{cases} \quad \text{for } t \geqq 2.$$

In particular,

$$u(x,2) = \begin{cases} \frac{1}{4}(|x|-2)^2 & \text{if } |x| \leqq 2, \\ 0 & \text{if } |x| \geqq 2 \end{cases} = \frac{1}{4}\{(2-|x|)_+\}^2. \tag{7.12}$$

See Figure 5 for the graphs of $u_0(x)$ and $u(x,2)$. The graph of $u(x,2)$ has a cusp at $x=0$, while, for every $t \in (0,2)$, the graph of $u(x,t)$ is flat near $x=0$.



(A) $u_0(x) = \min\{(2-|x|)_+,\, 1\}$.

(B) $u(x,2) = \frac{1}{4}\{(2-|x|)_+\}^2$.

Figure 5. Solution of (7.10).

(i) We now have $\|Du_0\|_{L^\infty(\mathbf{R}^n)} = 1$. By a property of the inf-convolution, the Lipschitz constant of $u(\cdot,t)$ does not increase. Namely, we have $\|D_x u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \leqq 1$ for all $t \in (0,T)$. This Lipschitz bound 1 is also deduced from Theorem 2.10 since

$$e^{C_1 t}\|Du_0\|_{L^\infty(\mathbf{R}^n)} + \beta(e^{C_1 t}-1) = 1. \tag{7.13}$$

The above observation implies that we may change the value of $H(p)$ for $|p| > 1$. For example, for any $\varepsilon > 0$ small, we take a convex Hamiltonian $\tilde{H} \in C^2(\mathbf{R}^n)$ such that

$$\tilde{H} = H \quad \text{in } \overline{B_1(0)}, \qquad |D_p\tilde{H}| \leqq 1+\varepsilon \quad \text{in } \mathbf{R}^n.$$

Then $\tilde{H}$ satisfies (H2) with $A_2 = 0$, $B_2 = 1+\varepsilon$, and we have $R(x,t) = (1+\varepsilon)t$.

We investigate the supergradient $D_x^+u(0,2)$. By (7.12) we see that $u(0,2) = 1$ and $D_x^+u(0,2) = \overline{B_1(0)}$. Moreover, since $R(0,2) = 2+2\varepsilon$, we deduce from (7.8) that

$$S(0,2;u_0) = \sup\left\{|p| \;\middle|\; p \in D_{pr}^- u_0(y),\ y \in \overline{B_{2+2\varepsilon}(0)}\right\} = 1.$$

In consequence, we find that the assertion of Theorem 3.6 (2):

$$|p| \leqq S(0,2;u_0) \quad \text{for any } p \in D_x^+u(0,2) \tag{7.14}$$

holds. Unlike (7.9) in Example 7.2, we do not need $\overline{S}(0,2;u_0)$ for the estimate. Furthermore, (7.14) is an optimal estimate since there is $p \in D_x^+ u(0,2)$ such that $|p| = 1$.

(ii) For any $L > 0$ the present Hamiltonian H satisfies $(\mathrm{H2})_L$. Since (7.13) holds, let us now take $L = 1$ and apply Theorem 5.8. When $L = 1$, the assumption $(\mathrm{H2})_L$ is fulfilled for $A_2 = 0$, $B_2 = 1$. This yields $R(x,t) = t$, which is smaller than the previous one in **(i)**.

Recall that

$$S^2(x,t;u_0,u(x,t)) = \sup\left\{ |p| \;\middle|\; p \in D^-_{pr}u_0(y),\ y \in \overline{B_{R(x,t)}(x)},\ u_0(y) + \frac{1}{2}t|p|^2 = u(x,t) \right\}$$

if $H(p) = \frac{1}{2}|p|^2$ and $m = 2$ (Definition 5.7). At a point $(x,t) = (0,2)$ we have

$$S^2(0,2;u_0,1) = \sup\left\{ |p| \;\middle|\; p \in D^-_{pr}u_0(y),\ y \in \overline{B_2(0)},\ u_0(y) + |p|^2 = 1 \right\}.$$

When $|y| = 2$, there exists $p \in D^-_{pr}u_0(y)$ such that $|p| = 1 \Longleftrightarrow u_0(y) + |p|^2 = 1$. This shows that $S^2(0,2;u_0,1) = 1$. We thus see that the assertion of Theorem 5.8 (2) holds.

Remark 7.4. For the Hamiltonian $H(p) = \frac{1}{2}|p|^2$, gradient estimates similar to Theorem 5.8 (2) are derived from several properties of the inf-convolution. Let $(x,t) \in \mathbf{R}^n \times (0,T)$ and u be the viscosity solution of (7.10)–(1.2), which is given by (7.11).

Assume that $p \in D_x^- u(x,t)$. Then it is known that

$$p \in D^- u_0(x+tp), \quad u_0(x+tp) + \frac{1}{2}t|p|^2 = u(x,t).$$

See, e.g., [9, Lemma A.5] and [3, Lemmas II.4.11, II.4.12 and their proofs]. Let us write $L = \|Du_0\|_{L^\infty(\mathbf{R}^n)}$. As already mentioned we have $\|D_x u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \leqq L$, and therefore $|p| \leqq L$. This implies that $x + tp \in \overline{B_{Lt}(x)}$. Hence

$$\begin{aligned}
&\inf\left\{ |q| \;\middle|\; q \in D^- u_0(y),\, y \in \overline{B_{Lt}(x)},\, u_0(y) + \frac{1}{2}t|q|^2 = u(x,t) \right\} \\
&\leqq |p| \\
&\leqq \sup\left\{ |q| \;\middle|\; q \in D^- u_0(y),\, y \in \overline{B_{Lt}(x)},\, u_0(y) + \frac{1}{2}t|q|^2 = u(x,t) \right\}.
\end{aligned}$$

Example 7.5. This last example shows that the lower bound gradient estimate is not available when the Hamiltonian is not convex. Let $H(p) = -c|p|$, where $c > 0$ is a constant. This is not a convex Hamiltonian. The corresponding equation is

$$u_t(x,t) - c|D_x u(x,t)| = 0 \quad \text{in } \mathbf{R}^n \times (0,T), \tag{7.15}$$

which is the eikonal equation with the opposite sign to (7.5). The viscosity solution of (7.15)–(1.2) is

$$u(x,t) = \max_{y \in \overline{B_{ct}(x)}} u_0(y).$$

This formula is obtained in a similar way to (7.6).

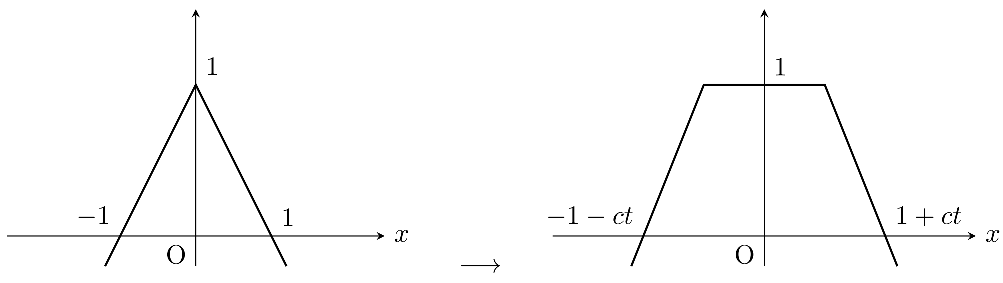
Let $u_0(x) = 1 - |x|$. Then the proximal subgradients of u_0 are

$$D^-_{pr}u_0(x) = \begin{cases} \emptyset & \text{if } x = 0, \\ \left\{ -\dfrac{x}{|x|} \right\} & \text{if } x \neq 0. \end{cases}$$

The solution u is of the form

$$u(x,t) = \min\{1 + ct - |x|,\, 1\}.$$

The graphs of $u_0(x)$ and $u(x,t)$ are shown in Figure 6. We can observe that the solution u instantly develops a flat part near $x=0$, i.e., $D_x u(0,t)=0$ for every $t>0$.



(A) $u_0(x) = 1 - |x|$.   (B) $u(x,t) = \min\{1 + ct - |x|,\, 1\}$.

Figure 6. Solution of (7.15).

Fix $t>0$. If Theorem 4.3 (1) were applicable at $(0,t)$, we would have

$$|p| \geqq \underline{I}(0,t;u_0) \quad \text{for any } p \in D_x^- u(0,t).$$

Since $D_x^- u(0,t) = \{0\}$, this is equivalent to

$$0 \geqq \underline{I}(0,t;u_0). \tag{7.16}$$

However, this inequality does not hold. In fact, for any $R>0$ we have

$$I_R := \inf\left\{ |p| \;\middle|\; p \in D_{pr}^- u_0(y),\, y \in \overline{B_R(0)} \right\} = 1.$$

In particular, $\underline{I}(0,t;u_0) = 1$ and thus (7.16) fails. Even if we replace $\underline{I}(0,t;u_0)$ by I_R, (7.16) does not hold whatever $R>0$ is chosen.

Acknowledgements

The work of the first author was partly supported by No. 19K14564 (Early-Career Scientists).

References

[1] Y. Achdou, G. Barles, H. Ishii, G. L. Litvinov, *Hamilton–Jacobi equations: approximations, numerical analysis and applications*, Lecture Notes in Mathematics **2074**, Springer, Heidelberg; Fondazione C.I.M.E., Florence, 2011.

[2] P. Albano, P. Cannarsa, C. Sinestrari, *Generation of singularities from the initial datum for Hamilton-Jacobi equations*, Journal of Differential Equations **268** (2020), 1412–1426.

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

[4] G. Barles, P. Cardaliaguet, O. Ley, R. Monneau, *Global existence results and uniqueness for dislocation equations*, SIAM J. Math. Anal. **40** (2008), 44–69.

[5] G. Barles, P. Cardaliaguet, O. Ley, A. Monteillet, *Uniqueness results for nonlocal Hamilton-Jacobi equations*, J. Funct. Anal. **257** (2009), 1261–1287.

[6] G. Barles, O. Ley, *Nonlocal first-order Hamilton-Jacobi equations modelling dislocations dynamics*, Comm. Partial Differential Equations **31** (2006), 1191–1208.

[7] G. Barles, O. Ley, H. Mitake, *Short time uniqueness results for solutions of nonlocal and non-monotone geometric equations*, Math. Ann. **352** (2012), 409–451.

[8] E. N. Barron, R. Jensen, *Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex Hamiltonians*, Comm. Partial Differential Equations **15** (1990), 1713–1742.

[9] M. G. Crandall, H. Ishii, P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), 1–67.

[10] M. G. Crandall, P.-L. Lions, *Viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), 1–42.

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

[12] L. C. Evans, *Partial differential equations. The second edition*, Graduate Studies in Mathematics **19**, American Mathematical Society, Providence, RI, 2010.

[13] Y. Fujita, *Lower estimates of L^∞-norm of gradients for Cauchy problems*, J. Math. Anal. Appl. **458** (2018), 910–924.

[14] Y. Giga, *Surface evolution equations: A level set approach*, Monographs in Mathematics **99**, Birkhäuser Verlag, Basel, 2006.

[15] N. Hamamuki, *An improvement of level set equations via approximation of a distance function*, Appl. Anal. **98** (2019), 1901–1915.

[16] N. Hamamuki, S. Kikkawa, *A lower spatially Lipschitz bound for solutions to fully nonlinear parabolic equations and its optimality*, to appear in Indiana Univ. Math. J.

[17] N. Hamamuki, E. Ntovoris, *A rigorous setting for the reinitialization of first order level set equations*, Interfaces Free Bound. **18** (2016), 579–621.

[18] H. Ishii, *Uniqueness of unbounded viscosity solution of Hamilton–Jacobi equations*, Indiana Univ. Math. J. **33** (1984), 721–748.

[19] H. Ishii, *Existence and uniqueness of solutions of Hamilton-Jacobi equations*, Funkcial. Ekvac. **29** (1986), 167–188.

[20] H. Ishii, *Lecture notes on the weak KAM theorem*, `http://www.f.waseda.jp/hitoshi.ishii/`

[21] N. Ichihara, H. Ishii, *Asymptotic solutions of Hamilton-Jacobi equations with semi-periodic Hamiltonians*, Comm. Partial Differential Equations **33** (2008), 784–807.

[22] O. Ley, *Lower-bound gradient estimates for first-order Hamilton–Jacobi equations and applications to the regularity of propagating fronts*, Adv. Differential Equations **6** (2001), 547–576.

(N. Hamamuki) Department of Mathematics, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-0810, Japan

Email address: `hnao@math.sci.hokudai.ac.jp`

(K. Hirose) Department of Mathematics, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-0810, Japan

Email address: `hirose.kazuya.w2@elms.hokudai.ac.jp`