Regularity of Hamilton–Jacobi Equations when Forward Is Backward

E. N. BARRON, P. CANNARSA, R. JENSEN
& C. SINISTRARI

ABSTRACT. We introduce a general principle determining in certain cases the regularity of the viscosity solutions of Hamilton–Jacobi equations. This principle says that if one can solve the equation forward in time from some initial data and then backward in time resulting in the same initial data, then the solution must be C^1 . Some cases are given when this holds as well as an example when it does not. Convexity of either the hamiltonian or the initial data plays a crucial role throughout.

1. INTRODUCTION

When one tries to solve a first order partial differential equation by the method of characteristics one sees that if the characteristics do not cross then the solution is generally smooth if the initial value is smooth. The characteristics propagate the initial data throughout the region as time progresses. This leads one to suspect that if the information of the initial data is not lost in forward time, then the characteristics do not cross and hence the solution should be smooth. The problem is how to make sense out of the phrase that the initial information is not lost. This paper looks at one approach to this. In our sense, information is not lost if solving the equation forward in time up to time, say $t = T$, and then solving the equation backward in time starting with terminal data at T , results in the recovery of the initial data at time $t = 0$.

Let us briefly explain what we mean by “forward” and “backward” solution of a Hamilton–Jacobi equation of the form $u_t + H(t, x, Du) = 0$. Suppose that $u(t, x)$ is a function defined in $[0, T] \times \mathbb{R}^n$, and let us denote by v the function obtained from u by “reversing time”, i.e. $v(t, x) = u(T - t, x)$. Clearly, if u is of class C^1 , u is a solution of $u_t + H(t, x, Du) = 0$ if and only if v solves $v_t - H(T - t, x, Dv) = 0$. It is well known, however, that the two properties are no longer equivalent when dealing with viscosity solutions of the equation. We say

therefore that u is a *forward solution* of the equation if it is a viscosity solution in the usual sense, while we say that u is a *backward solution* if $v(t, x) = u(T - t, x)$ is a viscosity solution of $v_t - H(T - t, x, Dv) = 0$.

In this paper we consider the problem $u_t + H(t, x, Du) = 0$ with some given initial data $u(0, x) = g(x)$, and we solve it forward in time up to T . Then we look at the problem $w_t + H(t, x, Dw) = 0$, solved backward in time with terminal data $w(T, x) = u(T, x)$. We then address the following two questions.

- (i) Assuming that $w \equiv u$ in $[0, T] \times \mathbb{R}^n$, can we deduce that u, w are of class C^1 ?
- (ii) Assuming only that $w(0, x) \equiv u(0, x)$, can we deduce that $w \equiv u$ in $[0, T] \times \mathbb{R}^n$ and that u, w are of class C^1 ?

In the statements above the regularity of u and w is meant in $(0, T) \times \mathbb{R}^n$, since a singularity may be present at the initial or final time. Intuitively it seems that the answer, at least to question (i), should be affirmative in general. Indeed, we prove that it is affirmative in certain cases, but we also show by examples that it is not always affirmative. In Examples 2.6 and 6.2 we provide counterexamples to property (i). The problem in the former of these examples is that the convexity of the hamiltonian changes as time progresses. At the time when the hamiltonian changes convexity, the characteristics of our solution coalesce and then fan out instantaneously. This is enough for a corner to develop but not enough to preclude solving forward and backward in time. In the latter example the hamiltonian is convex but not strictly convex. In this case, characteristics do not necessarily cross in the presence of a singularity, and so we are able to find a function which has a corner at all times and still is a forward and backward solution. Both types of behavior are excluded when the hamiltonian is strictly convex, as we will see in Theorem 2.5. However, even in this case, the answer to question (ii) can be negative, as will be shown in Example 6.3. So it seems that, for a general regularity result, a stronger condition is needed which guarantees that the characteristics never touch. Indeed the touching of the characteristics means that the characteristic system of ode's loses uniqueness and that is what leads to singularities. The problem is to determine a verifiable condition on the solution which is equivalent to no contact of characteristics. But determining such a condition is a difficult problem.

In Section 2 we give a precise definition of a forward and backward solution and consider question (i). An immediate consequence of assuming that we have a function which is both a forward and a backward solution of $u_t + F(t, x, Du) = 0$, with $F(t, x, p)$ strictly convex and superlinear in p , is that this forces u to be C^1 (see Theorem 2.5). Essentially, the strict convexity and the forward-backward property allow us to establish uniform bounds on weak second derivatives and that is why we have regularity here. Next we consider the case $u_t + H(t, x, Du) = 0$ where the hamiltonian H comes from a Mayer problem of optimal control, and so is convex in Du , but not strictly convex. Again we assume that we have in

hand a solution which is both forward and backward. It is proved in Theorem 2.8 that, if this solution is C^1 both at the initial and final time, then it is C^1 everywhere. The regularity assumption on the initial and final data is essential as shown by the above mentioned Example 6.2.

We then turn to study some cases where the stronger property (ii) can be proved. In the third section we consider $u_t + H(Du) = 0$, with $u(0, x) = g(x)$ and g convex and smooth, but no convexity assumptions on H . Now we have the Hopf formula [2] (see also [5], [6]) at our disposal and we prove that two solutions satisfying our backward and forward condition must be smooth and coincide. A necessary and sufficient condition for having $u(0, x) = w(0, x)$ under the assumptions of this section is that $p \mapsto g^*(p) + TH(p)$ is convex, where g^* is the Legendre–Fenchel conjugate of g . Using this property we can prove (see Theorem 3.1) that the answer to (ii) is affirmative. Furthermore, we see that singularities cannot develop until the first time T when $g^* + TH(p)$ becomes nonconvex.

In the fourth section we show in Theorem 4.1 that if we consider g smooth and $H = H(p)$ convex, again two solutions satisfying the backward and forward condition of (ii) must coincide and be smooth. The proof is based on the Lax formula. This result will be extended in the next section (Theorem 5.5) to hamiltonians depending also on (t, x) , but with more restrictive regularity assumptions.

In the fifth section we also address a third question, namely:

- (iii) Is it true in general (i.e. without assuming $w(0, x) \equiv u(0, x)$) that the backward solution w of $w_t + H(t, x, Dw) = 0$ with terminal data $w(T, x) = u(T, x)$ is of class C^1 ?

This conjecture is motivated by the consideration that in the special case where $H(p) = |p|^2$ the solution w is, at any fixed time $t \in (0, T)$, an inf–sup convolution of the initial value g of the forward problem, and therefore is smooth, see Lasry and Lions [13]. We also show that, in the case of a convex hamiltonian, this conjecture is closely related to the validity of (ii). We prove that the answer to (iii) is affirmative when H is strictly convex, independent of x , and the space dimension is one, see Theorem 5.6. In the cases of just convexity of H in the gradient, spatial dependence of H , or nonconvexity of H using the Hopf hypotheses, the property does not hold, as shown by counterexamples 6.1, 6.2 and 6.3. We leave as an open problem whether strict convexity of H in higher dimensions is sufficient for (iii) to hold, with $H = H(t, p)$.

2. THE CASE WITH $u_t + H(t, x, D_x u) = 0$

In this paper we consider Hamilton–Jacobi equations of the form

$$(2.2.1) \quad u_t + F(t, x, Du) = 0 \quad (t, x) \in (0, T) \times \mathbb{R}^n.$$

The assumptions on the function $F : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ will vary during the paper, but we will always suppose F at least continuous. Solutions to this equation are meant in the viscosity sense (see e.g. [3], [4], [10]).

Let us now give the definition of backward solution of our equation.

Definition 2.1. A function $u \in C([0, T] \times \mathbb{R}^n)$ is called a backward (viscosity) solution of equation (2.2.1) if $v(t, x) := u(T - t, x)$ is a viscosity solution of

$$v_t - F(T - t, x, Dv) = 0 \quad (t, x) \in [0, T] \times \mathbb{R}^n.$$

By a forward solution of the equation we mean a viscosity solution in the usual sense.

The following characterization follows immediately from the definition.

Proposition 2.2. The following properties are equivalent.

- (i) u is a backward solution of equation (2.2.1).
- (ii) u is a (forward) solution of $-u_t - F(t, x, Du) = 0$.
- (iii) If $w(t, x) = -u(T - t, x)$ and $G(t, x, p) = F(T - t, x, -p)$, then w is a solution of $w_t + G(t, x, Dw) = 0$.
- (iv) For any $(t, x) \in (0, T) \times \mathbb{R}^n$ we have

$$\begin{cases} p_t + F(t, x, p_x) \geq 0 & \text{if } (p_t, p_x) \in D^+u(t, x) \\ p_t + F(t, x, p_x) \leq 0 & \text{if } (p_t, p_x) \in D^-u(t, x) \end{cases}$$

Of course in (iv) of the proposition, if the inequalities are reversed we get the definition of a forward viscosity solution. From the above properties it is clear that the Cauchy problem for equation (2.2.1) with final data is well posed in the class of backward solutions.

There are several ways to define the sub and superdifferentials in the statement (iv) but here is the most useful:

$$D^+u(t, x) = \{p = (p_t, p_x) : p = D\varphi(t, x), \quad \exists \varphi \in C^1, u - \varphi \leq 0, (u - \varphi)(t, x) = 0\},$$

$$D^-u(t, x) = \{p = (p_t, p_x) : p = D\varphi(t, x), \quad \exists \varphi \in C^1, u - \varphi \geq 0, (u - \varphi)(t, x) = 0\}.$$

Our first aim is to prove the regularity of forward and backward solutions when F is strictly convex with respect to the third argument. Before doing this, we need to recall the definition and some basic properties of semiconcave functions.

Definition 2.3. A function $u : A \rightarrow \mathbb{R}$, with $A \subset \mathbb{R}^n$ open, is called semiconcave if, for any convex compact set $K \subset A$, there exists a nondecreasing function $\omega_K : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $\lim_{r \rightarrow 0} \omega_K(r) = 0$ and

$$(2.2.2) \quad \lambda u(x) + (1 - \lambda)u(y) - u(\lambda x + (1 - \lambda)y) \leq \lambda(1 - \lambda)|x - y|\omega_K(|x - y|)$$

for any $x, y \in K$ and $\lambda \in [0, 1]$.

The function ω_K is called a *modulus of semiconcavity*. A function u is called *semiconvex* if $-u$ is semiconcave. We remark that the previous definition is more general than the one often used in the literature, where it is required that $\omega_K(r) = c_K r$ for some $c_K > 0$. We recall the following properties ([3], [7], [1]).

- Proposition 2.4.** (i) If $u \in C^1(A)$, then u is semiconcave in A , with ω_K equal to the modulus of continuity of Du in K . Conversely, if $u : A \rightarrow \mathbb{R}$ is both semiconcave and semiconvex, then $u \in C^1(A)$. If u is both semiconcave and semiconvex with ω_K linear, then $u \in C_{loc}^{1,1}(A)$.
- (ii) If u is semiconcave then u is locally Lipschitz continuous and its superdifferential D^+u is nonempty everywhere.
- (iii) If u is semiconcave in A and $D^+u(x)$ is a singleton for all $x \in A$, then $u \in C^1(A)$.

We can now give our first regularity result.

Theorem 2.5. Let $F(t, x, \cdot)$ be strictly convex and superlinear. Assume in addition that, for any $R > 0$ there exists $C_R > 0$ such that

$$|F(t, x, p) - F(s, y, p)| \leq C_R(|t - s| + |x - y|), \quad s, t \in [0, T], \quad x, y \in \mathbb{R}^n, \quad |p| < R.$$

If u is a forward and backward locally Lipschitz solution of

$$u_t + F(t, x, Du) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n,$$

then $u \in C^1((0, T) \times \mathbb{R}^n)$.

Proof. Since u is a forward solution of the equation, u is semiconcave on $(0, T) \times \mathbb{R}^n$, by Theorem 3.2 in [8] and Proposition 2.5 in [16]. Since u is also a backward solution, it satisfies by Proposition 2.2-(iv)

$$p_t + F(t, x, p_x) = 0, \quad \forall (t, x) \in (0, T) \times \mathbb{R}^n, \quad \forall (p_t, p_x) \in D^+u(t, x).$$

By the strict convexity of F , we deduce that $D^+u(t, x)$ is a singleton for any $(t, x) \in (0, T) \times \mathbb{R}^n$. Thus, by Proposition 2.4-(iii), $u \in C^1((0, T) \times \mathbb{R}^n)$. $\square$

Remark. The local Lipschitz continuity of the solution u in the previous theorem can be obtained under general conditions on the initial data. It suffices for instance to prescribe an initial value which is lower semicontinuous and bounded from below.

Of course, the same result holds if the hamiltonian is strictly concave everywhere. On the other hand, the following example shows that, if H changes

convexity with time, the ability to solve forward and backward in time does not guarantee the smoothness of the solution.

Example 2.6. Consider the equation

$$u_t + \sin t |u_x|^2 = 0, \quad t > 0, \quad x \in \mathbb{R}.$$

We take the C^1 initial data

$$u(0, x) = \begin{cases} -|x| + 2 & \text{if } |x| \geq 4 \\ -x^2/8 & \text{if } |x| < 4. \end{cases}$$

This problem has a solution which is also a backward solution, but is not everywhere differentiable. In fact, the unique continuous viscosity solution is given by

$$(2.2.3) \quad u(t, x) = \begin{cases} -|x| + \cos t + 1, & \text{if } |x| \geq 2(1 + \cos t) \\ -\frac{x^2}{4(1 + \cos t)}, & \text{if } |x| < 2(1 + \cos t). \end{cases}$$

Observe that when $t = (2k + 1)\pi$ with $k = 0, 1, \dots$ the solution takes the value $u(t, x) = -|x|$ and hence u is not C^1 . Nevertheless, it is easily checked that u is both a forward and a backward solution of the equation.

There are interesting cases where the hamiltonian is convex but not strictly convex. One such case is the Hamilton–Jacobi equation associated with an optimal control problem of Mayer type, which we consider now. Let us first recall some basic properties. Let $H : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ be defined by

$$(2.2.4) \quad H(t, x, p) = \sup_{\alpha \in A} f(t, x, \alpha) \cdot p,$$

where A is a complete metric space and $f : [0, T] \times \mathbb{R}^n \times A \rightarrow \mathbb{R}^n$ is a continuous function with the following properties.

(P1) There exists $M > 0$ such that $|f(t, x, \alpha)| \leq M(1 + |x|)$.

(P2) There exists $L > 0$ such that

$$|f(t, x_1, \alpha) - f(t, x_2, \alpha)| \leq L|x_1 - x_2|.$$

(P3) The gradient $D_x f$ exists and is uniformly continuous on all sets of the form $[0, T] \times K \times A$, with $K \subset \mathbb{R}^n$ compact.

Then the viscosity solution of $u_t + H(t, x, Du) = 0$ can be represented as the value function of a Mayer optimal control problem, as shown by the next result (see [7]).

Theorem 2.7. *Let H be defined as in (2.2.4) for some function f which satisfies (P1)–(P3) and let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be semiconcave. Then the unique viscosity solution of*

$$(2.2.5) \quad \begin{cases} u_t + H(t, x, Du) = 0 \\ u(0, x) = g(x) \end{cases}$$

is semiconcave and is given by

$$u(x, t) = \min\{g(y(0))\}$$

over all $y(\cdot)$ solutions of

$$\begin{cases} y'(s) = f(s, y(s), \alpha(s)) \\ y(t) = x \end{cases}$$

with $\alpha : [0, t] \rightarrow A$. In addition, if f is of class $C_{loc}^{1,1}$ with respect to the x variable, and if g is semiconcave with a linear modulus, then u is also semiconcave with a linear modulus.

We can now prove a regularity result for this class of equations. Observe that the hamiltonian H defined above is homogeneous of degree one in the third argument and so it does not satisfy the hypotheses of Theorem 2.5.

Theorem 2.8. *Let H be the function defined in (2.2.4) for some given f which satisfies (P1)–(P3). Let $u \in C([0, T] \times \mathbb{R}^n)$ be a forward and backward viscosity solution of*

$$u_t + H(t, x, Du) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n.$$

- (i) *If $u(0, x)$ and $u(T, x)$ are both of class C^1 , then $u \in C^1([0, T] \times \mathbb{R}^n)$.*
- (ii) *If f is of class $C_{loc}^{1,1}$ with respect to the x variable, if $u(0, \cdot)$ is semiconcave and $u(T, \cdot)$ is semiconvex, both with a linear modulus, then $u \in C^{1,1}([0, T] \times \mathbb{R}^n)$.*

Proof. The regularity assumption in (i) implies in particular, by Proposition 2.4-(i), that $u(0, \cdot)$ is semiconcave and that $u(T, \cdot)$ is semiconvex. Since u is a forward solution of the equation with semiconcave initial data, u is semiconcave on $[0, T] \times \mathbb{R}^n$, by Theorem 2.7. Moreover, by Proposition 2.2-(iii), the function $w(t, x) := -u(T - t, x)$ satisfies $w_t + G(t, x, Dw) = 0$, where G is the hamiltonian of the Mayer problem associated with the function $g(t, x, \alpha) := -f(T - t, x, \alpha)$. Since g also satisfies (P1)–(P3) and $w(0, x)$ is semiconcave, it follows that w is also semiconcave. This implies that u is semiconvex and therefore C^1 , by Proposition 2.4-(i). The proof of statement (ii) is analogous. $\square$

Remark. Observe that, in contrast with Theorem 2.5, we had to assume here some regularity of the initial and final data. Such a requirement is essential, as will be shown in Example 6.2.

3. THE CASE $u_t + H(Du) = 0$, $u(0, x) = g(x)$ AND g CONVEX.

We now consider the Hamilton–Jacobi equation

$$(3.3.1) \quad u_t + H(Du) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n$$

with

$$(3.3.2) \quad u(0, x) = g(x), \quad x \in \mathbb{R}^n.$$

In this section we will assume the following:

- (A) $H : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable and convex.

Under assumption (A), the Hopf formula ([2, Theorem 3.1]) gives us the unique uniformly continuous explicit solution

$$(3.3.3) \quad u(t, x) = (g^*(p) + t H(p))^* = \sup_{p \in \mathbb{R}^n} p \cdot x - g^*(p) - t H(p).$$

Recall that the Legendre–Fenchel conjugates of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ are defined by $f^*(p) = \sup\{p \cdot x - f(x) : x \in \mathbb{R}^n\}$, $p \in \mathbb{R}^n$ and $f^{**}(x) = \sup\{p \cdot x - f^*(p) : p \in \mathbb{R}^n\}$.

The following theorem says that if we have a forward solution and a backward solution equal at the initial and final times, then the two solutions must be equal and smooth.

Theorem 3.1. *Let $T > 0$ denote a fixed positive time. Assume (A) and let $w : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ be the backward solution of*

$$(3.3.4) \quad w_t + H(Dw) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n$$

with terminal condition $w(T, x) = u(T, x)$, $x \in \mathbb{R}^n$. Then, if $w(0, x) = g(x)$, we have $w(t, x) = u(t, x)$ for $(t, x) \in [0, T] \times \mathbb{R}^n$ and $u \in C^1([0, T] \times \mathbb{R}^n)$. Furthermore $w(0, x) = g(x)$ if and only if $p \mapsto g^(p) + TH(p)$ is convex.*

Proof. The proof will proceed in two parts and will use Hopf’s formula in an essential way.

The function u is given by (3.3.3). Backward in time, we have the function w given by the Hopf formula

$$(3.3.5) \quad w(t, x) = (u^*(T, p) - (T - t)H(p))^*.$$

Now,

$$(3.3.6) \quad u^*(t, p) = (g^*(p) + t H(p))^{**} = g^*(p) + t H(p) - f(t, p)$$

where

$$f(t, p) = g^*(p) + t H(p) - (g^*(p) + t H(p))^{**},$$

is the smallest nonnegative function which makes $g^* + tH - f$ convex. Now we make the following claim.

Lemma 3.2. *Assume that $w(0, x) = g(x)$. The function $p \mapsto g^*(p) + TH(p)$ is convex.*

Proof. By (3.3.5) we have

$$g(x) = w(0, x) = (u^*(T, p) - TH(p))^*.$$

Let us set

$$k(t, p) = u^*(t, p) - t H(p) - (u^*(t, p) - t H(p))^{**}.$$

Then we have, using (3.3.6),

$$\begin{aligned} g^*(p) &= u^*(T, p) - TH(p) - k(T, p) \\ &= g^*(p) + TH(p) - f(T, p) - TH(p) - k(T, p) \\ &= g^*(p) - f(T, p) - k(T, p). \end{aligned}$$

and so $f(T, p) = k(T, p) = 0$ since they are both nonnegative. We conclude that $g^* + TH$ is convex. $\square$

Conversely, if $g^* + TH$ is convex, we have

$$\begin{aligned} w(0, x) &= (u^*(T, p) - TH(p))^* = ((g^*(p) + TH(p))^{**} - TH(p))^* \\ &= (g^*(p) + TH(p) - TH(p))^* = g^{**}(x) = g(x). \end{aligned}$$

Consequently, we have proved the last statement of the theorem.

Assume now that $w(0, x) = g(x)$. It follows from Rockafellar [14], Theorem 26.3, p. 253, that an everywhere finite convex function is strictly convex iff its conjugate is smooth. Since $g = g^{**}$ is C^1 and convex, g^* must be strictly convex. Now we come to our second lemma.

Lemma 3.3. *$p \mapsto g^*(p) + tH(p)$ is strictly convex for all $0 \leq t < T$.*

Proof. We have that $p \mapsto g^*(p) + TH(p)$ is convex. Then, writing

$$g^*(p) + tH(p) = \frac{t}{T}(g^*(p) + TH(p)) + \frac{T-t}{T}g^*(p)$$

for any $t \in [0, T)$, we see that $g^* + tH$ is the sum of a convex function and a strictly convex function and so it must be strictly convex for all $t \in [0, T)$. $\square$

We conclude that $u^*(t, p) = g^*(p) + tH(p)$ is strictly convex for each $t \in [0, T]$ and hence

$$u(t, x) = (u^*(t, p))^*$$

is C^1 in $x \in \mathbb{R}^n$, for $0 \leq t < T$.

Now the directional derivative of u at the point (t, x) in the direction (s, y) is given by

$$(3.3.7) \quad Du(t, x; s, y) = \max_{p \in S(t, x)} -H(p) s + p \cdot y = -H(D_x u(t, x)) s + D_x u(t, x) \cdot y$$

where $p_0 \in S(t, x)$ iff $p_0 \cdot x - g^*(p_0) - tH(p_0) = \sup_p p \cdot x - g^*(p) - tH(p)$. But we have proved that $S(t, x) = \{D_x u(t, x)\}$. Hence, the last part of (3.3.7) follows. Consequently, $(s, y) \mapsto Du(t, x; s, y)$ is linear, and we conclude that ([14, Theorem 25.2]) u is differentiable, and hence ([14, Theorem 25.5]) continuously differentiable at any $(t, x) \in [0, T] \times \mathbb{R}^n$. This implies that u is also a backward solution of the equation, and therefore $u = w$ in $[0, T] \times \mathbb{R}^n$. That is, once the solution is known to be C^1 it must be the only solution and it is generated by the method of characteristics. Hence it is immediately a forward and backward solution. $\square$

Remark 3.4. The sufficient condition in the theorem gives us a way of determining the length of time that the solution remains smooth, viz., as long as $g^*(p) + TH(p)$ remains convex.

4. THE CASE $w_t + H(Dw) = 0$, $w(T, x) = g(x)$ AND H CONVEX.

In this section we will assume the following conditions on H and g .

- (B) $H : \mathbb{R}^n \rightarrow \mathbb{R}$ is in $C^1(\mathbb{R}^n)$, H is convex and $\lim_{|p| \rightarrow \infty} H(p)/|p| = +\infty$. The function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is in $C^1(\mathbb{R}^n)$ and Lipschitz.

At the risk of confusing the reader we consider in this section the forward and the backward problem in the opposite order. That is, we first define w to be the backward solution of $w_t + H(Dw) = 0$ with $w(T, x) = g(x)$ and then let u be the forward solution of $u_t + H(Du) = 0$ with $u(0, x) = w(0, x)$. The forward and backward condition in this section is that $u(T, x) = g(x)$. We will prove the following theorem.

Theorem 4.1. Assume (B). Let w solve $w_t + H(Dw) = 0$ backward in time, with terminal condition $w(T, x) = g(x)$ and let u solve $u_t + H(Du) = 0$ with initial condition $u(0, x) = w(0, x)$.

(i) A necessary and sufficient condition for $u(T, x) = g(x)$ is

$$(4.4.1) \quad g(x) = \min_z \max_y \left\{ g(y) - T H^* \left(\frac{y - z}{T} \right) + T H^* \left(\frac{x - z}{T} \right) \right\}.$$

(ii) If (4.4.1) holds, then $u(t, x) = w(t, x)$, $(t, x) \in [0, T] \times \mathbb{R}^n$ and u is continuously differentiable on $(0, T] \times \mathbb{R}^n$.

Proof. To simplify the notation we will take $T = 1$ in the proof. In case (B) holds the unique solution for $(t, x) \in [0, 1] \times \mathbb{R}^n$ is given by the Lax formula (see [2, Theorem 2.1] and Strömberg [15] for the uniqueness):

$$u(t, x) = \min_{z \in \mathbb{R}^n} \left\{ \sigma(z) + t H^* \left(\frac{x - z}{t} \right) \right\} = \min_{z \in \mathbb{R}^n} \{ \sigma(x - tz) + t H^*(z) \}$$

in forward time. This function u is the unique viscosity solution of $u_t + H(Du) = 0$ with $u(0, x) = \sigma(x)$.

For the backward problem we have, by the Lax formula,

$$(4.4.2) \quad w(t, x) = \max_{y \in \mathbb{R}^n} \left\{ g(y) - (1 - t) H^* \left(\frac{y - x}{1 - t} \right) \right\},$$

is the unique viscosity solution of $w_t + H(Dw) = 0$ with $w(1, x) = g(x)$.

Solve w back from g at time 1. Take the initial data for u , $\sigma(x) = w(0, x)$. Then the solution of $u_t + H(Du) = 0$ with $u(0, x) = w(0, x)$ is

$$u(t, x) = \min_{z \in \mathbb{R}^n} \left\{ w(0, z) + t H^* \left(\frac{x - z}{t} \right) \right\}.$$

Now we make use of the forward and backward assumption. We have

$$\begin{aligned} u(1, x) &= \min_{z \in \mathbb{R}^n} \{ w(0, z) + H^*(x - z) \} \\ &= \min_z \max_y \{ g(y) - H^*(y - z) + H^*(x - z) \}. \end{aligned}$$

Part (i) follows trivially.

Remark. Under the assumption (B), the minima and maxima in the Lax formulas for u and w are achieved for each (t, x) . The easiest way to see this is as a result of the domain of dependence property of the solution. That is, for any fixed (t_0, x_0) , $u(t_0, x_0)$ depends only on the values of the initial data in the region $\{x : |x - x_0| \leq L|t - t_0|\}$ where L is a Lipschitz constant for H .

Now we suppose that (4.4.1) holds and we prove part(ii). First, the fact that $u = w$ is immediate from the uniqueness of viscosity solutions since under (4.4.1) both functions satisfy the same initial and terminal data. Now we begin the proof of continuous differentiability.

We will use the notation that for each fixed $x \in \mathbb{R}^n$, the infimum in (4.4.1) is achieved at z_x , i.e.,

$$g(x) = \max_y g(y) - H^*(y - z_x) + H^*(x - z_x).$$

But then, we have

$$g(x) - H^*(x - z_x) = \max_y g(y) - H^*(y - z_x).$$

If $\hat{y}$ is a point where the maximum of the right side is achieved, we conclude that

$$(4.4.3) \quad g(x) - H^*(x - z_x) = g(\hat{y}) - H^*(\hat{y} - z_x)$$

and

$$(4.4.4) \quad Dg(\hat{y}) - DH^*(\hat{y} - z_x) = 0.$$

Note that under (B), $H \in C^1$ and convex implies that $H = H^{**}$, H^* is strictly convex, everywhere finite, and therefore smooth.

Next, from (4.4.1) we have for any fixed $x_0 \in \mathbb{R}^n$

$$g(x) - H^*(x - z_{x_0}) \leq \max_y \{g(y) - H^*(y - z_{x_0})\},$$

and

$$g(x_0) - H^*(x_0 - z_{x_0}) = \max_y \{g(y) - H^*(y - z_{x_0})\}.$$

This says that x_0 maximizes $x \mapsto g(x) - H^*(x - z_{x_0})$ and consequently,

$$(4.4.5) \quad Dg(x_0) - DH^*(x_0 - z_{x_0}) = 0.$$

Lemma 4.2. *If $x_1 \in \mathbb{R}^n$ is any point so that $Dg(x_1) - DH^*(x_1 - z_{x_0}) = 0$, then $z_{x_1} = z_{x_0}$ and x_1 is a point achieving the maximum in $\max_y g(y) - H^*(y - z_{x_0})$.*

Proof. Indeed, applying (4.4.5) with x_0 replaced by x_1 , we have $Dg(x_1) - DH^*(x_1 - z_{x_1}) = 0$. Using the assumption of the lemma gives us $DH^*(x_1 - z_{x_0}) = DH^*(x_1 - z_{x_1})$. Since H^* is strictly convex, DH^* is a strictly monotone map, and this implies that $z_{x_1} = z_{x_0}$. Therefore, from (4.4.3),

$$\begin{aligned} g(x_1) - H^*(x_1 - z_{x_1}) &= g(\hat{y}) - H^*(\hat{y} - z_{x_1}) \\ &= g(\hat{y}) - H^*(\hat{y} - z_{x_0}) = \max_y g(y) - H^*(y - z_{x_0}). \end{aligned}$$

The lemma is proved. □

Now fix a point z_0 and choose a point x_{z_0} which satisfies

$$g(x_{z_0}) - H^*(x_{z_0} - z_0) = \max_{y \in \mathbb{R}^n} g(y) - H^*(y - z_0).$$

Then

$$\begin{aligned} g(x_{z_0}) &= \max_y \{g(y) - H^*(y - z_0) + H^*(x_{z_0} - z_0)\} \\ &\geq \min_z \max_y \{g(y) - H^*(y - z) + H^*(x_{z_0} - z)\} = g(x_{z_0}). \end{aligned}$$

Therefore, z_0 achieves the minimum for the point x_{z_0} .

Now let y' be any point for which $Dg(y') - DH^*(y' - z_0) = 0$. For this z_0 there is a y_0 so that

$$g(y_0) - H^*(y_0 - z_0) = \max_y g(y) - H^*(y - z_0),$$

and z_0 provides the minimum in

$$g(y_0) = \min_z \max_y g(y) - H^*(y - z) + H^*(y_0 - z).$$

Furthermore,

$$Dg(y') - DH^*(y' - z_0) = 0 = Dg(y_0) - DH^*(y_0 - z_0).$$

Hence, using Lemma 4.2, z_0 also provides the minimum in

$$g(y') = \min_z \max_y g(y) - H^*(y - z) + H^*(y' - z)$$

and not just $g(y_0)$.

With these preliminaries out of the way, we begin the argument to show that the maximum in (4.4.2) is achieved at a unique point for each fixed $(t, x) \in (0, 1) \times \mathbb{R}^n$. Fix $t \in (0, 1)$. Let $y_0 \in \mathbb{R}^n$ be a point for which

$$w(t, x) = \max_{y \in \mathbb{R}^n} g(y) - (1-t)H^*\left(\frac{y-x}{1-t}\right) = g(y_0) - (1-t)H^*\left(\frac{y_0-x}{1-t}\right).$$

At this point, $Dg(y_0) - DH^*((y_0 - x)/(1-t)) = 0$. We rewrite this as

$$Dg(y_0) - DH^*\left(y_0 - \left(y_0 - \frac{y_0-x}{1-t}\right)\right) = 0,$$

and conclude (from the preliminaries) that y_0 also satisfies

$$\begin{aligned} (4.4.6) \quad g(y_0) - H^*\left(y_0 - \left(y_0 - \frac{y_0-x}{1-t}\right)\right) \\ = \max_{y \in \mathbb{R}^n} g(y) - H^*\left(y - \left(y_0 - \frac{y_0-x}{1-t}\right)\right). \end{aligned}$$

Next,

$$\begin{aligned} g(y) - (1-t)H^*\left(\frac{y-x}{1-t}\right) \\ = g(y) - H^*\left(y - y_0 + \frac{y_0-x}{1-t}\right) + H^*\left(y - y_0 + \frac{y_0-x}{1-t}\right) \\ \quad - (1-t)H^*\left(\frac{y-y_0}{1-t} + \frac{y_0-x}{1-t}\right) \\ = \left\{g(y) - H^*\left(y - y_0 + \frac{y_0-x}{1-t}\right)\right\} + \left\{G(y-y_0) - (1-t)G\left(\frac{y-y_0}{1-t}\right)\right\} \\ \equiv I_1(y) + I_2(y), \end{aligned}$$

where

$$\begin{aligned} G(p) &:= H^*\left(p + \frac{y_0-x}{1-t}\right) - p \cdot DH^*\left(\frac{y_0-x}{1-t}\right), \\ I_1(y) &= \left\{g(y) - H^*\left(y - y_0 + \frac{y_0-x}{1-t}\right)\right\}, \quad \text{and} \\ I_2(y) &= \left\{G(y-y_0) - (1-t)G\left(\frac{y-y_0}{1-t}\right)\right\}. \end{aligned}$$

Observe that

$$DG(p) = DH^* \left(p + \frac{y_0 - x}{1 - t} \right) - DH^* \left(\frac{y_0 - x}{1 - t} \right),$$

and so $DG(0) = 0$. In addition G is a strictly convex function minus a linear function, and hence G is strictly convex.

Recall that $I_1(y)$ is maximized when $y = y_0$. As for $I_2(y)$, we have $DI_2(y_0) = 0$. Set $\delta = 1 - t$, $\xi = y - y_0$, and

$$\gamma(\beta) = G(\beta\xi) - \delta G(\beta\xi/\delta), \quad \beta \in \mathbb{R}.$$

We assume that $t \in (0, 1)$, so $0 < \delta < 1$. Then, $I_2(y) = \gamma(1)$, $I_2(y_0) = \gamma(0)$, $\gamma'(0) = 0$, and

$$(4.4.7) \quad d\gamma/d\beta = (\beta\xi - \beta\xi/\delta) \cdot (DG(\beta\xi) - DG(\beta\xi/\delta)) \frac{\delta}{\beta(\delta - 1)}.$$

Since G is strictly convex, $(\beta\xi - \beta\xi/\delta) \cdot (DG(\beta\xi) - DG(\beta\xi/\delta)) > 0$. If $\beta > 0$ the right hand side of (4.4.7) is therefore negative, and if $\beta < 0$ the right hand side of (4.4.7) is positive. Therefore, γ has a strict global maximum at $\beta = 0$. This implies that $I_2(y) < I_2(y_0)$. We conclude from

$$g(y) - (1 - t)H^* \left(\frac{y - x}{1 - t} \right) = I_1(y) + I_2(y)$$

that y_0 provides the unique maximum of $I_1 + I_2$ and it provides the unique maximum for $w(t, x)$. We conclude ([3, Prop. 2.13, chap. 2]) that w is continuously differentiable at every point of $(0, 1] \times \mathbb{R}^n$ and that u coincides with w . $\square$

Remark 4.3. The condition (4.4.1) tells us that the solution will be smooth for a length of time T as long as the data g satisfies the condition.

5. BACKWARD REGULARITY

The purpose of this section is twofold. First, we want to generalize the result of the previous two sections to some equations of the form

$$(5.5.1) \quad u_t + H(t, x, Du) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n$$

where there is dependence on t, x , and so we can no longer use the Hopf or Lax formula. For the reader's convenience, let us recall again our notation. Given some initial value $g \in \text{Lip}(\mathbb{R}^n)$, we call u the forward solution of equation (5.5.1) with initial value $u(0, x) = g(x)$. Then we denote by w the backward solution of

the same equation with final data $w(T, x) = u(T, x)$. As in the previous sections, we seek conditions on H , g which ensure the validity of the following implication:

$$(5.5.2) \quad w(0, x) = u(0, x) \text{ for all } x \in \mathbb{R}^n \Rightarrow u \equiv w \text{ and } u, w \in C^1((0, T) \times \mathbb{R}^n).$$

Our second aim in this section is to check, given u , w as above, the validity of the property

$$(5.5.3) \quad w \in C^1((0, T) \times \mathbb{R}^n).$$

In other words, we investigate whether the special form of the final data prescribed for w implies the regularity of w even without requiring that $u = w$ at time $t = 0$. Clearly property (5.5.3) implies the validity of (5.5.2); actually we will see that, in the case when H is convex, the two properties are in some sense equivalent (see Corollary 5.3 and the subsequent remark) and so they can be studied in a parallel way.

The conjecture that a property like (5.5.3) may hold is motivated by the following observation. Suppose that H is convex and depends only on Du . Then the Lax formula tells us that

$$w(t, x) = \max_{y \in \mathbb{R}^n} \min_{z \in \mathbb{R}^n} g(z) + TH^* \left(\frac{y - z}{T} \right) - (T - t)H^* \left(\frac{y - x}{T - t} \right).$$

Now, if $H(p) = H^*(p) = |p|^2/2$ then $w(t, \cdot)$ is an sup-inf convolution of g (see Lasry and Lions [13]), i.e.,

$$w(t, x) = \max_{y \in \mathbb{R}^n} \min_{z \in \mathbb{R}^n} g(z) + \frac{1}{T}|y - z|^2 - \frac{1}{(T - t)}|y - x|^2.$$

Hence w is of class $C^{1,1}$ for any $t \in (0, T)$. It is natural to conjecture that w should be smooth in $(0, T) \times \mathbb{R}^n$ for more general choices of H . That is, we ask ourselves how general is the principle that letting a function g evolve twice by the semigroup generated by H (once forward and once backward) yields a smooth function. By contrast we know that a single application of the semigroup in general does not regularize and in fact creates singularities.

Let us also remark that there is another simple case for which property (5.5.3) can be proved immediately: when g is concave and $H = H(Du)$ is strictly convex. Indeed, in this case, u stays concave and $-w(x, T - t)$ turns out to be the forward solution of an equation with convex initial data and strictly convex hamiltonian. So it is differentiable.

Let us now turn to the analysis of the more general equation (5.5.1). We assume in the following that the equation comes from a problem in calculus of variations. More precisely, we assume that H satisfies the conditions (10.13) in chapter I of [12]:

- (i) H is smooth and $H_{pp} > 0$
- (ii) $\lim_{|p| \rightarrow \infty} H(t, x, p)/|p| = +\infty$
- (iii) $p \cdot H_p - H \geq |H_p| \gamma(H_p)$, $\gamma(r) \rightarrow \infty$ as $|r| \rightarrow \infty$
- (iv) $H(t, x, 0) \leq c_1$, $H(t, x, p) \geq -c_1$ for some c_1
- (v) $|H_x| \leq c_2(p \cdot H_p - H) + c_3$, for some c_2, c_3
- (vi) $|H_p| \leq R \Rightarrow |p| \leq C(R)$, for some $C(R)$.

Then, the forward solution of (5.5.1) with initial data $u(0, x) = g(x)$ is given by

$$(5.5.4) \quad u(t, x) = \min_{\substack{\xi \in C^1([0, t], \mathbb{R}^n) \\ \xi(t) = x}} \left\{ g(\xi(0)) + \int_0^t L(s, \xi(s), \xi'(s)) ds \right\},$$

where $L = L(t, x, q) = H^*(t, x, \cdot)$ in the third argument.

Similarly, the backward solution w with terminal data $w(T, x) = u(T, x)$ is given by

$$(5.5.5) \quad w(t, x) = \max_{\substack{\zeta \in C^1([0, t], \mathbb{R}^n) \\ \zeta(t) = x}} \left\{ u(T, \zeta(T)) - \int_t^T L(s, \zeta(s), \zeta'(s)) ds \right\}.$$

By standard arguments it follows that the minimum (resp. maximum) in the above problems exists. Moreover, u is semiconcave in $(0, T] \times \mathbb{R}^n$ and w is semiconvex in $[0, T] \times \mathbb{R}^n$ (see e.g. [8]).

We collect some results about the above problems, which are in part classical, in part due originally to Kuznetsov and Šiškin (see [11] or [12] for a proof).

Theorem 5.1. *Let H satisfy assumptions (i)–(vi), let $L = H^*$ and let $g \in C^2(\mathbb{R}^n)$. Let $\xi : [0, t] \rightarrow \mathbb{R}^n$ be a minimizer for problem (5.5.4). Then the forward solution u to equation (5.5.1) with initial data g is differentiable at all points of the form $(s, \xi(s))$ with $s \in (0, t)$. In addition, there exists a unique $p : [0, t] \rightarrow \mathbb{R}^n$ (called the dual arc to ξ) such that the pair (ξ, p) satisfies the hamiltonian system*

$$(5.5.6) \quad \begin{cases} \xi'(s) = H_p(s, \xi(s), p(s)) \\ p'(s) = -H_x(s, \xi(s), p(s)) \end{cases}$$

and such that $p(s) = Du(s, \xi(s))$ for any $s \in [0, t)$. Moreover, the optimal trajectory starting from (t, x) is unique if and only if u is differentiable at (t, x) , and in that case $p(t) = Du(t, x)$.

An analogous result holds for the maximizers of the backward problem. The statement has to be changed in an obvious way; we only point out that the hamiltonian ordinary differential system is the same for both problems.

Remark. In the following we will need a version of the above theorem in the case when the initial value g is not in C^2 , but is merely semiconvex. The statement remains unchanged, except for the property $p(0) = Du(0, \xi(0))$, which has to be replaced by $p(0) \in D^-g(\xi(0))$. This last fact follows by the nonsmooth version of the maximum principle (see e.g. [9, Theor. IV.9.1]). The proof of the other properties remains the same, since it is based on the smoothness of the extremals and so no differentiability is required for g .

We begin our analysis of our pair of solutions u, w defined at the beginning of the section by looking at the way the equality $w(0, x) = u(0, x)$ can possibly fail.

Lemma 5.2. *We have $w(0, x) \leq g(x)$ for all $x \in \mathbb{R}^n$. In addition, the two following properties are equivalent, for a given x_0 :*

- (i) $w(0, x_0) = g(x_0)$;
- (ii) *if $\zeta : [0, T] \rightarrow \mathbb{R}^n$ with $\zeta(0) = x_0$ is a maximizer for the backward problem (5.5.5) with endpoint $(t, x) = (0, x_0)$, then ζ is also a minimizer for the forward problem (5.5.4) with endpoint $(t, x) = (T, \zeta(T))$.*

Proof. Given $x_0 \in \mathbb{R}^n$, let $\zeta : [0, T] \rightarrow \mathbb{R}^n$ with $\zeta(0) = x_0$ be an optimal arc for the backward problem with endpoint $(0, x_0)$. Then we have

$$w(0, x_0) = u(T, \zeta(T)) - \int_0^T L(s, \zeta(s), \zeta'(s)) ds.$$

On the other hand, ζ is an admissible arc for the forward problem with endpoint $(T, \zeta(T))$ and so we have, by (5.5.4),

$$(5.5.7) \quad u(T, \zeta(T)) \leq g(x_0) + \int_0^T L(s, \zeta(s), \zeta'(s)) ds.$$

It follows that $w(0, x_0) \leq g(x_0)$, with equality if and only if there is equality in (5.5.7), that is, if ζ is optimal for the forward problem at $(T, \zeta(T))$. $\square$

Remark. Conversely, if we first consider the solution v_1 to the backward problem with some final data $v_1(T, x) = f(x)$, and call v_2 the forward solution with initial data $v_2(0, x) = v_1(0, x)$, we obtain by the same argument of the lemma that $v_2(T, x) \geq f(x)$.

Corollary 5.3. *If $\tilde{u}$ is the forward solution of equation (5.5.1) with initial data $\tilde{u}(0, x) = w(0, x)$, then we have $\tilde{u}(T, x) \equiv u(T, x)$.*

Proof. By the previous remark we obtain that $\tilde{u}(T, x) \geq w(T, x) = u(T, x)$. On the other hand, we have $\tilde{u}(0, x) \leq u(0, x)$ by Lemma 5.2, and this implies $\tilde{u}(T, x) \leq u(T, x) = w(T, x)$ by (5.5.4) (or by the comparison principle for viscosity solutions). $\square$

In other words, even if in general the inequality $w(0, x) \leq g$ is strict, we always have $u(T, x) = \tilde{u}(T, x)$ and so the process of taking alternately forward and backward solutions stabilizes after the first step, a property which is formally analogous to the behavior of the Legendre–Fenchel convex conjugate, i.e., after two conjugations we get nothing new. This result depends in an essential way on the convexity of H .

Remark. From the previous corollary it follows that, when we want to prove property (5.5.3), it is not restrictive to assume that $u(0, \cdot) = w(0, \cdot)$, since we can always replace g with $w(0, x)$ and u with $\tilde{u}$, and this leaves w unchanged. Thus, properties (5.5.2) and (5.5.3) are in fact equivalent, and from now on we will focus our attention on the proof of the former one. One point however should be noted if we want to deduce property (5.5.3) from (5.5.2) by the above argument. Since all the regularity we can expect from $w(0, x)$ in general is semiconvexity, we have to prove property (5.5.2) without requiring g to be more regular than that. Any result where property (5.5.2) is proved assuming the differentiability of g , like Theorem 4.1 in the previous section, cannot be used to prove property (5.5.3).

An important consequence of the forward and backward assumption is the following.

Lemma 5.4. *Let u, w be defined as at the beginning of this section, for some semiconvex initial value g . Suppose that $u(0, x) = w(0, x)$. Let x_0 be a point where g is differentiable. Denote by (ξ, p) the solution of the hamiltonian system (5.5.6) with initial data $\xi(0) = x_0, p(0) = Dg(x_0)$. Then the following properties hold.*

- (i) ξ is an optimal trajectory both for the backward maximization problem (5.5.5) with endpoint $(0, x_0)$ and for the forward minimization problem (5.5.4) with endpoint $(T, \xi(T))$.
- (ii) u and w coincide along ξ and satisfy

$$u(t, \xi(t)) = w(t, \xi(t)) = g(x_0) + \int_0^t L(s, \xi(s), \xi'(s)) ds, \quad \forall t \in [0, T].$$

- (iii) u and w are differentiable along ξ (except possibly at the terminal point) and satisfy

$$Du(t, \xi(t)) = Dw(t, \xi(t)) = p(t), \quad \forall t \in [0, T].$$

Proof. We apply Theorem 5.1 to the maximization problem (5.5.5) with endpoint $(0, x_0)$. Since g is differentiable at x_0 , there exists a unique optimal trajectory ξ , which is obtained by solving system (5.5.6) with initial data $\xi(0) = x_0$, $p(0) = Dg(x_0)$. Then we obtain from Theorem 5.1 the parts of properties (i)–(iii) concerning w . On the other hand, from Lemma 5.2–(ii) we deduce that ξ is optimal also for the forward problem with endpoint $(T, \xi(T))$. We can now apply Theorem 5.1 to the forward problem. Let us call $\tilde{p}$ the dual arc associated with ξ as in Theorem 5.1, this time considered as a minimizer for the forward problem. Since g is differentiable at x_0 , we have $\tilde{p}(0) = Dg(x_0)$. The two pairs $(\xi, \tilde{p})$ and (ξ, p) satisfy the same differential system and coincide at $t = 0$, and so they coincide everywhere. The remaining assertions of the lemma are then also a consequence of Theorem 5.1. $\square$

As a consequence, we can prove property (5.5.2) holds in the case where g is smooth. It is similar to Theorem 4.1; the smoothness and growth assumptions on H are here more restrictive, but dependence on t, x is allowed.

Theorem 5.5. *Let H satisfy assumptions (1)–(6) and let $g \in C^1(\mathbb{R}^n) \cap \text{Lip}(\mathbb{R}^n)$. Let u and w be defined as in the beginning of this section and suppose that $w(0, x) = u(0, x)$ for all $x \in \mathbb{R}^n$. Then u and w coincide in $[0, T] \times \mathbb{R}^n$ and are C^1 in $[0, T) \times \mathbb{R}^n$.*

Proof. It suffices to prove that u is differentiable in $[0, T) \times \mathbb{R}^n$. Consider a point $(\bar{t}, \bar{x}) \in (0, T) \times \mathbb{R}^n$. Let ξ be a minimizer for the forward problem (5.5.4) at $(\bar{t}, \bar{x})$ and let us set $x_0 = \xi(0)$, $p_0 = Dg(x_0)$. By Theorem 5.1, ξ is the first component of the solution to system (5.5.6) with initial data $\xi(0) = x_0$, $p(0) = p_0$. Then, by part (iii) of the previous lemma, u is differentiable at the point $(\bar{t}, \bar{x})$, because it lies along ξ . $\square$

As already remarked, the above result cannot be used to prove property (5.5.3), since it contains the assumption that g is differentiable. It is therefore desirable to remove this assumption. This is done in the next theorem, at the price of requiring that the hamiltonian does not depend on x and that the space dimension is one. The first restriction is unavoidable, since Example 6.3 below shows that if H depends on x the theorem is false. On the other hand, we do not know whether the assumption on the space dimension is essential.

Theorem 5.6. *Suppose the space dimension n is one. Let the function H in (5.5.1) be independent of x and satisfy assumptions (1)–(6). Given $g \in \text{Lip}(\mathbb{R})$ and semiconvex, let u, w be defined as at the beginning of this section. Suppose that $w(0, x) = u(0, x)$ for all $x \in \mathbb{R}$. Then u and w coincide in $[0, T] \times \mathbb{R}$ and are C^1 in $(0, T) \times \mathbb{R}$.*

Proof. Let us assume that u is not differentiable at a point $(\bar{t}, \bar{x}) \in (0, T) \times \mathbb{R}$. Then, by Theorem 5.1, there exist two distinct minimizers ξ_1, ξ_2 for the forward problem (5.5.4) with endpoint $(\bar{t}, \bar{x})$. Let us denote by p_1, p_2 the dual arcs associated with ξ_1, ξ_2 . Since H does not depend on x , we see from (5.5.6) that these arcs are constant, and that the trajectories are solutions of

$$\xi'_i(s) = H_p(s, p_i), \quad s \in [0, \bar{t}].$$

Since H_p is strictly monotone in p , and since $\xi_1(\bar{t}) = \xi_2(\bar{t}) = \bar{x}$ we easily obtain that $\xi_1(0) \neq \xi_2(0)$. Let us assume for instance that $\xi_1(0) < \xi_2(0)$. Since g is Lipschitz, we can find $x_0 \in (\xi_1(0), \xi_2(0))$ where g is differentiable. By Lemma 5.4-(i), x_0 is the starting point of a curve $\xi : [0, T] \rightarrow \mathbb{R}$ which is a minimizer for the forward problem at $(T, \xi(T))$. But then ξ must cross either ξ_1 or ξ_2 somewhere in $(0, \bar{t}]$. Since $x_0 \neq \xi_i(0)$ and $\bar{t} < T$ the intersection point is not an endpoint of ξ , and so u is differentiable at that point. Then ξ and the intersected trajectory ξ_i , together with their dual arcs, solve the same system with the same conditions at the intersection point. This implies that $\xi \equiv \xi_i$, which is impossible because $\xi(0) \neq \xi_i(0)$. $\square$

By the arguments of the remark after Corollary 5.3, we can use the above result to prove, under the same hypotheses on H , property (5.5.3).

Corollary 5.7. *Let $n = 1$ and let $H = H(t, p)$ satisfy assumptions (1)–(6). Given $g \in \text{Lip}(\mathbb{R}^n)$, let u and w be defined as at the beginning of this section. Then $w \in C^1((0, T) \times \mathbb{R}^n)$.*

6. EXAMPLES

We present a few examples to show that our regularity results are sharp.

Example 6.1. In this example we will show that at least convexity of the hamiltonian is essential for backward regularity. Let $H(p) = (1 - p^2)^2$. For fixed $\varepsilon > 0$ to be determined later, set $g^*(p) = 2\varepsilon p^2$. This comes from the function $g(x) = x^2/8\varepsilon$. Taking $T = 1$ we have, by the Hopf formula,

$$\begin{aligned} u(t, x) &= (g^*(p) + tH(p))^*, \\ w(t, x) &= (u^*(1, p) - (1 - t)H(p))^*. \end{aligned}$$

The function u is the forward solution and w is the backward solution. Then

$$u^*(1, p) = g^*(p) + H(p) - k(p) = [p^2 - (1 - \varepsilon)]^2 + 2\varepsilon - \varepsilon^2 - k(p),$$

where

$$k(p) = g^*(p) + H(p) - (g^*(p) + H(p))^{**} = \begin{cases} 0, & \text{if } p^2 > 1 - \varepsilon \\ (p^2 - (1 - \varepsilon))^2, & \text{if } p^2 \leq 1 - \varepsilon. \end{cases}$$

In order for w to be smooth we must have that

$$w^*(t, p) = (u^*(1, p) - (1 - t)H(p))^{**}$$

is strictly convex with respect to p . We will show that this is not the case. Indeed,

$$f(t, p) := u^*(1, p) - (1 - t)H(p) = tp^4 - 2(t - \varepsilon)p^2 + t - k(p).$$

Then, for $p^2 < 1 - \varepsilon$ we have

$$f_{pp}(t, p) = (t - 1)(12p^2 - 4)$$

which is negative provided $0 < \varepsilon < 2/3$, and therefore $f(t, p)$ is not convex. Hence, $w^*(t, p)$ must have flat spots, i.e., it is not strictly convex, and so $w(t, x)$ cannot be smooth.

Example 6.2. Here we show that if the hamiltonian is convex, but not strictly convex, we cannot expect regularity of the backward solution, even if the initial data of the forward problem are smooth and the space dimension is one. Consider $u_t + |u_x| = 0$ on $[0, \pi/2] \times \mathbb{R}$, with $u(0, x) = \cos(x)$. Then using the Lax formula, we easily find that the forward solution is 2π -periodic in x and is given by

$$(6.6.1) \quad u(t, x) = \begin{cases} -1, & \text{if } \pi - t \leq |x| \leq \pi \\ \cos(|x| + t), & \text{if } |x| < \pi - t, \end{cases}$$

for $|x| \leq \pi$, while the backward solution w satisfying $w(\pi/2, x) = u(\pi/2, x)$ is given by

$$(6.6.2) \quad w(t, x) = \begin{cases} -1, & \text{if } \pi - t \leq |x| \leq \pi \\ \cos(|x| + t), & \text{if } \pi/2 - t \leq |x| < \pi - t \\ 0, & \text{if } |x| < \pi/2 - t. \end{cases}$$

Thus, w is nondifferentiable along the lines $x = \pm(\pi/2 - t)$. Observe in addition that w is also a forward solution of the equation. Therefore w satisfies all the assumptions of Theorem 2.8-(ii) except the semiconvexity of the final data. Since w is not C^1 , this shows that such an assumption on the final data is essential for that theorem.

Example 6.3. This example shows that the conjecture of backward regularity is false when we have x dependence even with a strictly convex and coercive hamiltonian. Consider the equation

$$(6.6.3) \quad u_t + \frac{1}{16}u_x^2 + 4 + 3K^2x^2 = 0,$$

with initial data $u(0, x) = g(x)$. Here $K > 0$ is to be determined later and g is any smooth even function, with $g(0) = 0$. The viscosity solution can be represented as the value function of the following problem of calculus of variations:

$$u(t, x) = \min_{\substack{\xi \in C^1([0, t], \mathbb{R}^n) \\ \xi(t) = x}} g(\xi(t)) - \int_0^t 3K^2\xi^2(s) - 4(\xi'^2(s) - 1) \, ds.$$

Given some $T > 0$, let $w(t, x)$ be the solution of (6.6.3) with $w(T, x) = u(T, x)$. We claim that, if K is chosen large enough, then w cannot be smooth at all points $(t, 0)$ for $0 < t < T$. To see this, we use the property that w is given by

$$w(t, x) = \max_{\substack{\xi \in C^1([t, T], \mathbb{R}^n) \\ \xi(t) = x}} u(T, \xi(T)) + \int_t^T 3K^2\xi^2(s) - 4(\xi'^2(s) - 1) \, ds.$$

Given a point of the form $(t_0, 0)$, with $0 < t_0 < T$, let us suppose that the trajectory $\xi \equiv 0$ is a maximizer for the above problem. Then we have $w(t_0, 0) = u(T, 0) + 4(T - t_0)$. On the other hand, if we choose $\xi'(s) = +1$ on $t_0 \leq s \leq (T + t_0)/2$ and then $\xi'(s) = -1$ on $(T + t_0)/2 < s \leq T$, we get

$$w(t_0, 0) \geq u(T, 0) + \int_{t_0}^T 3K^2\xi^2(s) \, ds = u(T, 0) + 2K^2 \left(\frac{T - t_0}{2} \right)^3,$$

which implies that $K(T - t_0) \leq 4$. Thus we see that, if K is chosen greater than $4/T$ and if t_0 is small enough, then $\xi \equiv 0$ cannot be optimal for the point $(t_0, 0)$. In this case there exists a maximizer ξ^* for $(t_0, 0)$ which is not identically zero. But then, by symmetry, $-\xi^*$ is also a maximizer. Hence, from $(t_0, 0)$ there are at least two optimal trajectories and this is impossible if w is smooth (see Theorem 5.1). Let us also remark that, if we denote by $\tilde{u}$ the solution of (6.6.3) with initial data $\tilde{u}(0, x) = w(0, x)$, then, by Corollary 5.3, $\tilde{u}$ and w are a forward and a backward solution which coincide at initial and final times but not in $(0, T) \times \mathbb{R}^n$. This shows that property (5.5.2) is in general false when the hamiltonian depends on x .

REFERENCES

- [1] G. ALBERTI, L. AMBROSIO & P. CANNARSA, *On the singularities of convex functions*, Manuscripta Math. **76** (1992), pp. 421–435.
- [2] M. BARDI & L. C. EVANS, *On Hopf formula for solutions of Hamilton–Jacobi equations*, Nonlinear Anal. TMA **8** (1984), pp. 1373–1381.
- [3] M. BARDI & I. CAPUZZO-DOLCETTA, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*, Birkhäuser, Boston, 1997.
- [4] G. BARLES, *Solutions de Viscosité des Équations de Hamilton–Jacobi*, Mathématiques & Applications **17**, Springer Verlag, New York, 1994.
- [5] E. N. BARRON, R. JENSEN & W. LIU, *Hopf–Lax formula for $u_t + H(u, Du) = 0$* , J. Diff. Eqs. **126** (1996), pp. 48–61.
- [6] E. N. BARRON, R. JENSEN & W. LIU, *Hopf–Lax formula for $u_t + H(u, Du) = 0$: II*, Comm. in PDE **22** (1997), pp. 1141–1160.
- [7] P. CANNARSA & H. FRANKOWSKA, *Some characterizations of the optimal trajectories in control theory*, SIAM J. Control Optim. **29** (1991), pp. 1322–1347.
- [8] P. CANNARSA & H. M. SONER, *Generalized one–sided estimates for solutions of Hamilton–Jacobi equations and applications*, Nonlinear Anal. TMA **13** (1989), pp. 305–323.
- [9] F. H. CLARKE, Y. S. LEDYAEV, R. J. STERN, P. R. WOLENSKI, *Nonsmooth analysis and control theory*, Springer–Verlag, New York, 1998.
- [10] M. G. CRANDALL, H. ISHII, P. L. LIONS, *User’s guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. **27** (1992), pp. 1–67.
- [11] W. H. FLEMING, *The Cauchy problem for a nonlinear first order partial differential equation*, J. Differential Equations, **5** (1969), pp. 515–530.
- [12] W. H. FLEMING & H. M. SONER, *Controlled Markov Processes and Viscosity Solutions*, Springer–Verlag, New York, 1993.
- [13] J. M. LASRY & P. L. LIONS, *A remark on regularization in Hilbert spaces*, Israel J. Math. **55** (1986), pp. 257–266.
- [14] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
- [15] T. STRÖMBERG, *Hopf’s formula gives the unique viscosity solution*, to appear.
- [16] C. SINISTRARI, *Semiconcavity of solutions of stationary Hamilton–Jacobi equations*, Nonlinear Anal. TMA **24** (1995), pp. 1321–1326.
- [17] A. I. SUBBOTIN, *Generalized Solutions of First Order PDEs*, Birkhäuser, Boston, 1995.

Research by E. N. Barron and R. Jensen was supported in part by a grant DMS-9532030 from the National Science Foundation, and a grant from Loyola University–Chicago.

E. N. BARRON & R. JENSEN:
Department of Mathematical and Computer Sciences
Loyola University Chicago
Chicago, Illinois 60626, U. S. A
EMAIL: enb@math.luc.edu (E. N. Barron)
EMAIL: rrj@math.luc.edu (R. Jensen)

P. CANNARSA & C. SINISTRARI
Dipartimento di Matematica
Università di Roma “Tor Vergata”
Via della Ricerca Scientifica
I-00133 Roma, Italia
EMAIL: cannarsa@axp.mat.uniroma2.it (P. Cannarsa)
EMAIL: sinistra@axp.mat.uniroma2.it (C. Sinestrari)

Submitted: July 2, 1998; revised: February 6, 1999.