File: 409J 358101 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2869 Signs: 1820 . Length: 50 pic 3 pts, 212 mm

Journal of Mathematical Analysis and Applications – 3587

JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **176**, 000–000 (1993)

Verification Theorems within the Framework of Viscosity Solutions*

XUN YU ZHOU

Faculty of Management, University of Toronto, Toronto, Ontario, Canada M5S 1V4

Submitted by Leonard D. Berkovitz

Received October 17, 1991

In optimal control theory, the verification technique plays an important role in testing for the optimality of a given control, and in constructing optimal feedback controls. However, the existing classical verification theorem is restrictive in that it requires the associated dynamic programming equation to have smooth solutions. In this paper, some verification theorems are presented within the framework of viscosity solutions under mild assumptions. These theorems are shown to have wider applicability than the classical verification theorem., As a relevant problem, some differences and relationships between the viscosity solution and Clarke's generalized gradient are also discussed. © 1993 Academic Press, Inc.

1. INTRODUCTION

Let us consider the following optimal control problem. Given $(s, y) \in [0, 1] \times R^d$, we are to

$$\text{minimize } J(s, y; u(\cdot)) := \int_s^1 L(t, x(t), u(t))\, dt + h(x(1)), \tag{1.1}$$

$$\text{subject to } \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), & \text{a.e. } t \in [s, 1] \\ x(s) = y, \end{cases} \tag{1.2}$$

over the set of admissible controls $U_{ad}[s, 1] := \{u(\cdot) \mid u(\cdot) \text{ is a Lebesgue measurable function from } [s, 1] \text{ to } \Gamma\}$, where Γ is a prescribed arbitrary set in R^m.

We denote the above problem by $C_{s,y}$ to recall the dependence on the initial time s and the initial state y. The value function is defined as

$$V(s, y) := \inf\{J(s, y; u(\cdot)) \colon u(\cdot) \in U_{ad}[s, 1]\}. \tag{1.3}$$

* This work was partly supported by the NSERC Grant A4619 and Manufacturing Research Corporation of Ontario.

0022-247X/93 $5.00

File: 409J 358102. By:CV Date:07:04:93. Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2371 Signs: 1455. Length: 45 pic 0 pts, 190 mm

Any pair $(x(\cdot), u(\cdot))$, where $x(\cdot)$ is the solution of (1.2) corresponding to $u(\cdot) \in U_{\mathrm{ad}}[s, 1]$, is called an *admissible pair* for the problem $C_{s,y}$. An admissible pair $(x^*(\cdot), u^*(\cdot))$ is called *optimal* for $C_{s,y}$ if $u^*(\cdot)$ achieves the minimum of $J(s, y; u(\cdot))$ over $U_{\mathrm{ad}}[s, 1]$.

As a part of the dynamic programming approach, the so-called verification technique plays an important role in testing for optimality of a given admissible pair, and in constructing optimal feedback controls. The classical verification theorem is as follows (see Fleming and Rishel [6, Thm. IV.4.4 and VI.4.1]):

THEOREM 1.1. *Let $W \in C^{1,1}([0, 1] \times R^d)$ be a solution of the Hamilton–Jacobi equation (HJ for short)*

$$\begin{cases} -v_t(t, x) + \sup_{u \in \Gamma} H(t, x, u, v_x(t, x)) = 0, & (t, x) \in (0, 1) \times R^d, \\ v(1, x) = h(x), \end{cases} \tag{1.4}$$

where the Hamiltonian is defined as

$$H(t, x, u, q) := -q \cdot f(t, x, u) - L(t, x, u), \tag{1.5}$$

for $(t, x, u, q) \in [0, 1] \times R^d \times \Gamma \times R^d$. Then:

(a) *$W(s, y) \leqslant J(s, y; u(\cdot))$, for any $(s, y) \in [0, 1] \times R^d$ and any $u(\cdot) \in U_{\mathrm{ad}}[s, 1]$.*

(b) *Suppose a given admissible pair $(x^*(\cdot), u^*(\cdot))$ for the problem $C_{s,y}$ satisfies*

$$W_t(t, x^*(t)) = H(t, x^*(t), u^*(t), W_x(t, x^*(t))), \qquad \text{a.e. } t \in [s, 1]. \tag{1.6}$$

Then $(x^*(\cdot), u^*(\cdot))$ is an optimal pair for the problem $C_{s,y}$.

Remark 1.1. By the HJ equation, (1.6) is equivalent to a more familiar form:

$$\max_{u \in \Gamma} H(t, x^*(t), u, W_x(t, x^*(t))) = H(t, x^*(t), u^*(t), W_x(t, x^*(t))). \tag{1.7}$$

Then, an optimal feedback control $u^*(t, x)$ can be constructed by minimizing $H(t, x, u, W_x(t, x))$ over $u \in \Gamma$. For details, see [6].

Remark 1.2. Equality (1.6) is also equivalent to

$$W(s, y) = J(s, y; u^*(\cdot)). \tag{1.8}$$

File: 409J 358103 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2814 Signs: 2169 . Length: 45 pic 0 pts, 190 mm

Indeed, we can write

$$
\begin{aligned}
&h(x^*(1) - W(s, y)\\
&\quad = W(1, x^*(1)) - W(s, y)\\
&\quad = \int_s^1 \frac{d}{dt} W(t, x^*(t))\, dt\\
&\quad = \int_s^1 [W_t(t, x^*(t)) - H(t, x^*(t), u^*(t), W_x(t, x^*(t))) - L(t, x^*(t), u^*(t))]\, dt,
\end{aligned}
$$

which implies

$$W(s, y) = J(u^*(\cdot)) + \int_s^1 [H(t, x^*(t), u^*(t), W_x(t, x^*(t))) - W_t(t, x^*(t))]\, dt.$$

Thus (1.6) is equivalent to (1.8) by virtue of the HJ (1.4).

When practically applying Theorem 1.1, one usually takes the verification function W to be the value function V, since V satisfies the HJ if $V \in C^{1,1}([0, 1] \times R^d)$. Unfortunately, it is very likely that the HJ (1.4) has no smooth solution at all. Indeed, by considering the equivalent condition (1.8), Clarke [2] has given an example showing that no such function W as in Theorem 1.1 exists. This makes the applicability of the classical verification theorem very restrictive. In recent years, the viscosity solution theory of general nonlinear PDEs, which was launched by Crandall and Lions [5], has been significently developed. In this theory, all the derivatives involved are replaced by the so-called super- and sub-differentials, and the solutions in the viscosity sense can be merely continuous functions. The existence and uniqueness of viscosity solutions of the HJ can be guaranteed under very mild and reasonable assumptions, which are satisfied in the great majority of cases arising in optimal control problems. For example, the value function turns out to be the unique viscosity solution of the HJ. Now a natural question arises: Does the verification theorem still hold, with the solutions of the HJ in the classical sense replaced by the ones in the viscosity sense, and the derivatives involved replaced by the super- and/or sub-differentials?

The purpose of this paper is to answer the above question by deriving verification theorems within the framework of viscosity solutions. It should be mentioned that there is another type of nonsmooth analysis, involving the "generalized gradient" introduced by Clarke [2], within which a verification theorem has already been established [2, 3]. In that approach, the verification function W was choosen to be any "generalized solution" of the HJ, and any admissible pair $(x^*(\cdot), u^*(\cdot))$ satisfying (1.8) was proved to be an optimal pair. The significance of Clarke's theory lies in that it can treat optimal control problems with state constraints. On the other hand, Frankowska [8], after showing that any Lipschitz viscosity solution is a generalized solution of the HJ, conjectured that Clarke's verification technique applies to Lipschitz viscosity solutions of the HJ for control problems without state constraints. However, it is the following reasons that motivate us to derive verification theorems within the framework of viscosity solutions in spite of the existing Clarke's verification theorem: First, "viscosity solution" and "generalized gradient" are two *different* frameworks; see the Appendix for some of the differences. In particular, the notion of viscosity solution enjoys many merits in analysis as developed extensively in the literature, as well as some numerical advantages. So it should be of interest to have verification theorems completely within this framework. Second, Clarke's verification theorem is expressed in a form analogous to (1.8) rather than to (1.6); see [2, 3]. Hence it is not clear how to obtain an optimal feedback control, even if formally, from Clarke's verification theorem. On the other hand, if one takes the value function V to be the verification function W, then the criterion (1.8) becomes trivial. Finally, it is possible to treat optimal controls of stochastic diffusion processes within the framework of viscosity solutions, since the extension to the *second-order* super-/sub-differential is natural and straightforward (cf. [10]), whereas it is difficult even to define the corresponding "second-order generalized gradient."

The paper is organized as follows: In Section 2, some preliminary results about viscosity solutions and the associated super- and sub-differentials will be introduced. In Section 3, various verification theorems in terms of viscosity solutions and the super- and/or sub-differentials are established. In addition, an example is presented showing that the obtained theorems can test for the optimality of a given control while the classical verification cannot. Further, methods of constructing optimal feedback controls are described. Section 4 gives some concluding remarks. Finally, some differences and relationships between viscosity solutions and generalized gradients are discussed in the Appendix.

2. Super-, Sub-differentials, and Viscosity Solutions

Let Q be an open subset of R^n, and $v\colon \bar{Q} \to R$ be a continuous function.

Definition 2.1. The *super-*(resp. *sub-*) *differential* of v at $\hat{x} \in Q$, denoted by $D_x^+ v(\hat{x})$ (resp. $D_x^- vv(\hat{x})$), is a *set* defined by

$$D_x^+ v(\hat{x}) := \left\{ p \in R^n \mid \limsup_{x \in Q,\, x \to \hat{x}} \frac{v(x) - v(\hat{x}) - p \cdot (x - \hat{x})}{|x - \hat{x}|} \leqslant 0 \right\}.$$

File: 409J 358105 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2166 Signs: 1101 . Length: 45 pic 0 pts, 190 mm

(resp.

$$D_x^- v(\hat{x}) := \{ p \in R^n \mid \liminf\{\cdots\} \geqslant 0 \}.)$$

For $\hat{x} \in Q$ and $\xi \in R^n$, we denote by $v'(\hat{x}; \xi)$ the (one-sided) directional gradient (along ξ) of v at $\hat{x}$, namely,

$$v'(\hat{x}; \xi) := \lim_{h \to 0+} \frac{v(\hat{x} + h\xi) - v(\hat{x})}{h},$$

whenever the right hand side limit exists.

LEMMA 2.1. *Suppose $v'(\hat{x}; \xi)$ exists for given $\hat{x} \in Q$ and $\xi \in R^n$. Then,*

$$\sup_{p \in D_x^- v(\hat{x})} p \cdot \xi \leqslant v'(\hat{x}; \xi) \leqslant \inf_{p \in D_x^+ v(\hat{x})} p \cdot \xi, \tag{2.1}$$

where $\sup\{\varnothing\} := -\infty$, $\inf\{\varnothing\} := +\infty$.

Proof. For any $p \in D_x^+ v(\hat{x})$,

$$\limsup_{h \to 0+} \frac{v(\hat{x} + h\xi) - v(\hat{x}) - h\xi \cdot p}{h} \leqslant 0,$$

hence $v'(\hat{x}; \xi) \leqslant p \cdot \xi$. This implies the right-hand side of (2.1). Similarly for the left-hand side. ∎

Given a continuous function G: $Q \times R^1 \times R^n \to R^1$. We consider the following fully nonlinear partial differential equation:

$$G(x, v(x), \nabla v(x)) = 0. \tag{2.2}$$

DEFINITION 2.2. A function $v \in C(\bar{Q})$ is called a *viscosity solution* of (2.2), if at each $x \in Q$,

$$G(x, v(x), p) \leqslant 0, \qquad \text{for all} \quad p \in D_x^+ v(x);$$

$$G(x, v(x), p) \geqslant 0, \qquad \text{for all} \quad p \in D_x^- v(x).$$

Now let us turn to the control problem formulated in Section 1. We impose the following assumptions throughout this paper:

(A1) f and L are continuous mappings from $[0, 1] \times R^d \times \Gamma$ to R^d and R^1, respectively; moreover, f and L are continuous with respect to (t, x), uniformly in $u \in \Gamma$.

File: 409J 358106 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2479 Signs: 1420 . Length: 45 pic 0 pts, 190 mm

(A2) There exists a constant $K>0$, which is independent of (t, u), such that

$$|f(t, x, u)-f(t, y, u)|+|L(t, x, u)-L(t, y, u)|+|h(x)-h(y)|$$
$$\leqslant K|x-y|, \qquad \text{for all} \quad x, y \in R^d,$$
$$|f(t, x, u)|+|L(t, x, u)|+|h(x)| \leqslant K(1+|x|), \qquad \text{for all} \quad x \in R^d.$$

(A3) $\Gamma \subset R^m$ is compact.

The following result is well-known. One may see, e.g., Lions [9] and Crandall *et al.* [4].

LEMMA 2.2. *The value function V is globally Lipschitz continuous in (t, x), and is the unique viscosity solution of the HJ* (2.4).

LEMMA 2.3. *Let W be a locally Lipschitz viscosity solution of the HJ* (1.4). *Then, at each $(t, x) \in (0, 1) \times R^d$,*

$$-p+\sup_{u \in \Gamma} H(t, x, u, q) \leqslant 0, \qquad \text{for all} \quad (p, q) \in D_{t,x}^{+} W(t, x);$$
$$-p+\sup_{u \in \Gamma} H(t, x, u, q)=0, \qquad \text{for all} \quad (p, q) \in D_{t,x}^{-} W(t, x).$$

Proof. It suffices to prove the equality in terms of the sub-differential. But this follows from [8, Thm. 2.3] (see also Theorem 5.2 in the Appendix), since $\tilde{H}(t, x, q):=\sup_{u \in \Gamma} H(t, x, u, q)$ is convex in q. ∎

An immediate consequence of Lemma 2.3 is the following corollary.

COROLLARY 2.1. *At each $(t, x) \in (0, 1) \times R^d$,*

$$p \geqslant H(t, x, u, q), \qquad \text{for all} \quad (p, q) \in D_{t,x}^{+} W(t, x) \cup D_{t,x}^{-} W(t, x).$$

3. VERIFICATION THEOREMS

THEOREM 3.1. *Let $W \in C([0,1] \times R^d)$ be a locally Lipschitz viscosity solution of the HJ* (1.4). *Then*:

(a) *$W(s, y) \leqslant J(s, y; u(\cdot))$, for any $(s, y) \in [0,1] \times R^d$ and any $u(\cdot) \in U_{\mathrm{ad}}[s, 1]$.*

(b) *Let $(x^*(\cdot), u^*(\cdot))$ be a given admissible pair for the problem $C_{s,y}$. Suppose that for* a.e. *$t \in [s, 1]$, there exists $(p^*(t), q^*(t)) \in D_{t,x}^{+} W(t, x^*(t))$ such that*

$$p^*(t)=H(t, x^*(t), u^*(t), q^*(t)), \tag{3.1}$$

then $(x^*(\cdot), u^*(\cdot))$ is an optimal pair for the problem $C_{s,y}$.

File: 409J 358107 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2118 Signs: 1097 . Length: 45 pic 0 pts, 190 mm

Remark 3.1. Part (a) of Theorem 3.1 is trivial since $W = V$ in view of Lemma 2.2. However, we state our results in the present form purposely in order to compare with the classical verification theorem.

Proof of Theorem 3.1. We only prove part (b) of the theorem. We set $f^*(t) := f(t, x^*(t), u^*(t))$, etc., to simplify the notation. Since both W and x^* are Lipschitz, $t \mapsto W(t, x^*(t))$ is differentiable almost everywhere. Fix $r \in [s, 1]$ such that $(d/dt)\, W(t, x^*(t))|_{t=r}$ exists, that $\lim_{h \to 0+} \int_r^{r+h} f^*(t)\, dt = f^*(r)$, and that (3.1) holds. Then,

$$\begin{aligned}
\frac{d}{dt} W(t, x^*(t))|_{t=r} &= \lim_{h \to 0+} \frac{W(r+h, x^*(r+h)) - W(r, x^*(r))}{h} \\
&= \lim_{h \to 0+} \frac{W(r+h, x^*(r) + \int_r^{r+h} f^*(r)\, dr) - W(r, x^*(r))}{h} \\
&= \lim_{h \to 0+} \frac{W(r+h, x^*(r) + hf^*(r) + o(h)) - W(r, x^*(r))}{h} \\
&= \lim_{h \to 0+} \frac{W(r+h, x^*(r) + hf^*(r)) - W(r, x^*(r))}{h} \\
&\qquad \text{(by Lipschitz property of } W) \\
&= W'((r, x^*(r)); \quad (1, f^*(r))) \\
&\leqslant p^*(r) + q^*(r) \cdot f^*(r) \qquad \text{(by Lemma 2.1)} \\
&= -L^*(r) \qquad \text{(by (3.1))}.
\end{aligned}$$

Hence we conclude that

$$W(1, x^*(1)) - W(s, y) = \int_s^1 \frac{d}{dt} W(t, x^*(t))|_{t=r}\, dr \leqslant -\int_s^1 L^*(r)\, dr,$$

which implies

$$J(s, y; u^*(\cdot)) = \int_s^1 L^*(r)\, dr + h(x^*(1)) \leqslant W(s, y).$$

Therefore, it follows from a) that $u^*(\cdot)$ is an optimal control. ∎

Remark 3.2. The condition (3.1) implies that

$$\max_{u \in \Gamma} H(t, x^*(t), u, q^*(t)) = H(t, x^*(t), u^*(t), q^*(t)). \tag{3.2}$$

File: 409J 358108 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2481 Signs: 1643 . Length: 45 pic 0 pts, 190 mm

This is easily seen by recalling the fact that W is the viscosity solution of (1.4),

$$-p^*(t) + \sup_{u \in \Gamma} H(t, x^*(t), u, q^*(t)) \leqslant 0,$$

which yields (3.2) under (3.1).

Remark 3.3. Theorem 3.1 is a generalization of the classical verification theorem (Theorem 1.1). On the other hand, we do have examples showing that the classical verification theorem may not be able to verify the optimality of a given control, whereas Theorem 3.1 can.

EXAMPLE 3.1. Consider the following optimal control problem

$$\text{minimize } -x(1),$$
$$\text{subject to } \begin{cases} \dot{x}(t) = x(t)\, u(t), & \text{a.e. } t \in [s, 1], \\ x(s) = y, \end{cases}$$
$$\text{control } u(\cdot): [0, 1] \to \{r \in R^1 \mid 0 \leqslant r \leqslant 1\}.$$

The value function can be calculated as

$$V(t, x) = \begin{cases} -xe^{1-t}, & \text{if } \ x > 0, \\ -x, & \text{if } \ x \leqslant 0. \end{cases}$$

Let us consider an admissible pair $(x^*(\cdot), u^*(\cdot)) \equiv (0, 0)$ for the problem $C_{0,0}$. Theorem 1.1 cannot tell if the pair is optimal, since $V_x(t, x^*(t))$ does not exist on the *whole* trajectory $x^*(\cdot)$. On the other hand, we have $D^+_{t,x} V(t, x^*(t)) = D^+_{t,x} V(t, 0) = \{0\} \times [-e^{1-t}, -1]$. Now if we take $(p^*(t), q^*(t)) = (0, -1) \in D^+_{t,x} V(t, x^*(t))$ for each t, then (3.1) is satisfied. This implies that the pair $(x^*(\cdot), u^*(\cdot))$ is indeed optimal by virtue of Theorem 3.1.

Remark 3.4. A result similar to Part (b) of Theorem 3.1 has been proved in [1, Thm. 4.1]. It should be noted, however, that the maximum condition—the first equality of (25) in [1]—imposed there can be removed. Indeed, the condition is a *consequence* of Theorem 3.1 above and [12, Prop. 3.1].

Theorem 3.1 gives a sufficient condition for a control to be optimal. But under some extra assumptions, the condition is also necessary.

THEOREM 3.2. *Assume that f, L, h are continuously differentiable in x. Then a given admissible pair $(x^*(\cdot), u^*(\cdot))$ for the problem $C_{s,y}$ is optimal* *if and only if for* a.e. $t \in [s, t]$, *there exists* $(p^*(t), q^*(t)) \in D^+_{t,x} V(t, x^*(t))$ *such that* (3.1) *holds.*

Proof. It suffices to show the "only if" part. Let ψ be the adjoint function corresponding to the optimal pair $(x^*(\cdot), u^*(\cdot))$, namely, ψ satisfies

$$\begin{cases} \dot{\psi}(t) = H_x(t, x^*(t), u^*(t), \psi(t)), & \text{a.e. } t \in [s, 1], \\ \psi(1) = h_x(x^*(1)). \end{cases} \tag{3.3}$$

Then by Zhou [12, Thm. 3.2], for a.e. $t \in [s, 1]$,

$$(H(t, x^*), u^*(t), \psi(t)), \psi(t)) \in D^+_{t,x} V(t, x^*(t)).$$

This yields the desired result. ▮

Remark 3.5. A result analogous to Theorem 3.2 has been proved in [1, Thm. 4.2]. However, in the "if" part of our Theorem 3.2 above, $q^*(\cdot)$ is not required to satisfy the adjoint equation (3.3) and the maximum principle, as imposed in [1] (cf. Eq. (27) and (28) in [1]). In fact, the optimality can be assured as long as there is *any* $(p^*(t), q^*(t)) \in D^+_{t,x} V(t, x^*(t))$ such that (3.1) holds, even if $q^*(\cdot)$ is *not* the adjoint function. For instance, we consider the admissible pair $(x \cdot (\cdot), u^*(\cdot)) \equiv (0, 0)$ for the problem $C_{0,0}$ in Example 3.1. If we take $(p^*(t), q^*(t)) = (0, -e^{1-t}) \in D^+_{t,x} V(t, x^*(t))$ for each t, then (3.1) is satisfied. The pair $(x^*(\cdot), u^*(\cdot))$ is therefore optimal by Theorem 3.2, but $q^*(\cdot)$ does not satisfy the adjoint equation (3.3).

It should be noted that Theorem 3.1 is not adequate, since in some cases DS^+W may be empty! (E.g., when W is convex.) Therefore we need a similar result in terms of D^-W.

THEOREM 3.3. *Let* $W \in C([0, 1] \times R^d)$ *be a locally Lipschitz viscosity solution of the HJ* (1.4), *and* $(x^*(\cdot), u^*(\cdot))$ *be a given admissible pair for the problem* $C_{s,y}$. *Suppose that for* a.e. $t \in [s, 1]$, *there exists* $(p^*(t), q^*(t)) \in D^-_{t,x} W(t, x^*(t))$ *such that*

$$p^*(t) = H(t, x^*(t), u^*(t), q^*(t)), \tag{3.4}$$

then $(x^*(\cdot), u^*(\cdot))$ *is an optimal pair.*

Proof. Fix $r \in [s, 1]$ such that $(d/dt)\, W(t, x^*(t))|_{t=r}$ exists, than $\lim_{h \to 0+} \int_{r-h}^{r} f^*(t)\, dt = f^*(r)$, and that (3.4) holds. Then,

File: 409J 358110 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2393 Signs: 1276 . Length: 45 pic 0 pts. 190 mm

$$
\begin{aligned}
\frac{d}{dt} W(t, x^*(t))|_{t=r} &= -\lim_{h\to 0+} \frac{W(r-h, x^*(r-h)) - W(r, x^*(r))}{h} \\
&= -\lim_{h\to 0+} \frac{W(r-h, x^*(r) - \int_{r-h}^{r} f^*(r)\,dr) - W(r, x^*(r))}{h} \\
&= -\lim_{h\to 0+} \frac{W(r-h, x^*(r) - hf^*(r)) - W(r, x^*(r))}{h} \\
&= -W'((r, x^*(r)); (-1, -f^*(r))) \\
&\leqslant -\max_{(p,q)\in D^-_{t,x} W(r, x^*(r))} [-p - q\cdot f^*(r)] \\
&= \min_{(p,q)\in D^-_{t,x} W(r, x^*(r))} [p + q\cdot f^*(r)] \\
&\leqslant p^*(r) + q^*(r)\cdot f^*(r) \\
&= -L^*(r).
\end{aligned}
$$

Hence the desired result follows similarly as in the proof of Theorem 3.1. ∎

Remark 3.6. Due to Lemma 2.3, the condition (3.4) is *equivalent* to

$$\max_{u\in\Gamma} H(t, x^*(t), u, q^*(t)) = H(t, x^*(t), u^*(t), q^*(t)).$$

Set

$$D^{\pm}_{t,x} v(t, x) := D^+_{t,x} v(t, x) \cup D^-_{t,x} v(t, x),$$

for any $v \in C([0, 1] \times R^d)$.

THEOREM 3.4. *Let $(x^*(\cdot), u^*(\cdot))$ be a given admissible pair for the problem $C_{s,y}$. We have the following conclusions*:

(a) *If for a.e. $t\in[s, 1]$, there exists $(p^*(t), q^*(t)) \in D^{\pm}_{t,x} V(t, x^*(t))$ such that $p^*(t) = H(t, x^*(t), u^*(t), q(t))$, then $(x^*(\cdot), u^*(\cdot))$ is optimal.*

(b) *If there exists a non-zero Lebesgue measurable set $T_0 \subseteq [s, 1]$ such that for any $t \in T_0$, there is $(p^*(t), q^*(t)) \in D^{\pm}_{t,x} V(t, x^*(t))$ and $p^*(t) > H(t, x^*(t), u^*(t), q^*(t))$, then $(x^*(\cdot), u^*(\cdot))$ is* not *optimal.*

Proof. (a) follows by combining the proofs of Theorems 3.1 and 3.3; (b) follows from [12, Prop. 3.1]. ∎

Let us conclude this section by describing how to construct optimal feedback controls by the verification theorems obtained. First, we recall the definition of admissible feedback controls, following [6].

File: 409J 358111 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2692 Signs: 1889 . Length: 45 pic 0 pts, 190 mm

DEFINITION 3.1. A measurable function $\mathbf{u}$ from $[0, 1] \times R^d$ to Γ is called an *admissible feedback control* if for any $(s, y) \in [0, 1] \times R^d$ there is a unique solution $x(\cdot\,; s, y)$ of the following equation

$$\begin{cases} \dot{x}(t) = f(t, x(t), \mathbf{u}(t, x(t))), & \text{a.e. } t \in [s, 1], \\ x(s) = y. \end{cases} \tag{3.5}$$

An admissible feedback control $\mathbf{u}^*$ is called optimal if $(x^*(\cdot\,; s, y), \mathbf{u}^*(\cdot, x^*(\cdot\,; s, y)))$ is optimal for the problem $C_{s,y}$ for *each* (s, y), where $x^*(\cdot\,; s, y)$ is the solution of (3.5) corresponding to $\mathbf{u}^*$.

THEOREM 3.5. *Let $\mathbf{u}^*$ be an admissible feedback control, and $\mathbf{p}^*$ and $\mathbf{q}^*$ be two measurable functions satisfying $(\mathbf{p}^*(t, x), \mathbf{q}^*(t, x)) \in D^{\pm}_{t,x} V(t, x)$ for all (t, x). If*

$$\begin{aligned} \mathbf{p}^*(t, x) - H(t, x, \mathbf{u}^*(t, x), \mathbf{q}^*(t, x)) &= \inf_{(p, q, u) \in D^{\pm}_{t,x} V(t,x) \times \Gamma} [p - H(t, x, u, q)] \\ &= 0 \end{aligned} \tag{3.6}$$

for all $(t, x) \in [0, 1] \times R^d$. Then $\mathbf{u}^*$ is optimal.

Proof. The result follows readily from (a) of Theorem 3.4. ∎

Remark 3.7. By Theorem 3.5, we can formally obtain an optimal feedback control by minimizing $p - H(t, x, u, q)$ over $D^{\pm}_{t,x} V(t, x) \times \Gamma$ for each (t, x). We said "formally" because there are some points which are not clear. First, although the infimum in (3.6) can be achieved (note that each $D^{\pm}_{t,x} V(t, x)$ is compact due to the Lipschitz property of V), we do not know in general if the infimum is zero and if there is a *measurable selector* of $(\mathbf{p}^*(t, x), \mathbf{q}^*(t, x), \mathbf{u}^*(t, x))$ (The answers are positive if V is convex or semiconcave). Second, even if there exists a measurable selector such that (3.6) holds, it is difficult to verify whether the equation (3.5) under $\mathbf{u}^*$ has a unique solution. This is not clear even when V is smooth (cf. [6, p. 99 and p. 170]). All these remain challenging open problems.

EXAMPLE 3.2. Consider a deterministic manufacturing system. Let $U(t)$ be the control variable (rates of production, advertising expenditures, etc.) at time t, $X(t)$ the state (inventories, sales, etc.), and $Z(t)$ some given input to the system (demands, etc.). The dynamics is

$$\dot{X}(t) = b(X(t), Z(t)) + BU(t), \qquad X(0) = X_0, \quad Z(0) = Z_0.$$

File: 409J 358112 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2671 Signs: 1966 . Length: 45 pic 0 pts, 190 mm

The cost functional is

$$J(X_0, Z_0; U(\cdot)) = \int_0^{\infty} e^{-\rho t} G(X(t), U(t))\, dt.$$

The corresponding HJ equation is as follows

$$\rho v(x, z) = \inf_{u \in \Gamma} [(b(x, z) + Bu)\, v_x(x, z) + G(x, u)]. \tag{3.7}$$

Sethi and Zhang [11] considered a slightly different version of the above model, and obtained an optimal feedback control by minimizing the right hand side of (3.7) with the value function v. However, this is not quite rigorous since $v_x(x, z)$ may not exist. (For a very simple case, $v_x(x, z)$ does exist, and an optimal feedback control can be obtained explicitly; see Fleming *et al.* [7].) By Remark 3.8, we can formally obtain an optimal feedback by minimizing $(b(x, z) + Bu)q + G(x, u)$ over $(q, u) \in D_x^{\pm} v(x, z) \times \Gamma$ for all (x, z) (Note that all the results in this paper adapt readily to the problems with discounted cost functions over infinite horizons).

4. Concluding Remarks

In this paper we have given some verification theorems in the language of viscosity solutions and the associated super- and sub-differentials. The conditions under which these theorems valid are quite mild and reasonable, compared with the restrictive classical verification theorem. We have also conjectured that optimal feedback controls may be constructed by virtue of the verification theorems obtained in this paper.

It should be noted that the results of this paper were derived when there is no state constraint in the optimal control problem. We do not know how to treat the state constraint problems. Indeed, the presence of state constraints causes great difficulty to the analysis: they bring some particular boundary conditions (depending on the particular features of the state constraints imposed) to the associated HJ equations, while the existing viscosity solutions theory on nonlinear PDEs with boundary conditions is far from satisfactory and complete.

5. Appendix

In this Appendix, we discuss some differences and relationships between the two frameworks of nonsmooth analysis: viscosity solution and generalized gradient.

File: 409J 358113 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2640 Signs: 1739 . Length: 45 pic 0 pts, 190 mm

Let Q be an open subset of R^n, and $v: \bar{Q} \to R$ be a locally Lipschitz function. We recall the following definition [2].

DEFINITION 5.1. The *generalized gradient* of v at $\hat{x} \in Q$, denoted by $\partial v(\hat{x})$, is a set defined by

$$\partial v(\hat{x}) := \{ p \in R^n \mid p \cdot \xi \leqslant v^0(\hat{x}; \xi), \quad \text{for any } \xi \in R^n \},$$

where $v^0(\hat{x}; \xi) := \limsup_{x \to \hat{x}, h \to 0+} (v(x + h\xi) - v(x))/h$.

Remark 5.1. $\partial v(\hat{x})$ is a *nonempty* convex set satisfying $\partial(-v)(\hat{x}) = -\partial v(\hat{x})$ and

$$v^0(\hat{x}; \xi) = \max_{p \in \partial v(\hat{x})} p \cdot \xi. \tag{5.1}$$

See [2] for a thorough treatment of the generalized gradients.

Remark 5.2. One of the relationships between the super-/sub-differential and the generalized gradient can be written as

$$D_x^{\pm} v(x) \subseteq \partial v(x), \qquad \text{for any} \quad x \in Q,$$

provided that v is locally Lipschitz [8, Thm. 1.4]. It should be noted that the above inclusion may be strict. To see this, take

$$v(x) := \begin{cases} x^2 \sin(1/x), & \text{if} \quad x \in R^1, x \neq 0, \\ 0, & \text{if} \quad x = 0. \end{cases}$$

v is differentiable at 0, hence $D_x^+ v(0) = D_x^- v(0) = \{0\}$. But $\partial v(0)$ equals the convex hull of the set of limits of the form $\lim v_x(y)$, where $y \to 0$ (cf. [2, Thm. 2.5.1]). So $\partial v(0) = [0, 1]$. Through this example, we may capture some sense about the difference between the super-/sub-differential and the generalized gradient: if the former is a nonsmooth notion of "differentiability," then the latter may be regarded as a nonsmooth notion of "continuous differentiability."

Now let us introduce the generalized solution of the HJ (1.4) in the frame-work of Clarke's generalized gradient.

DEFINITION 5.2. A locally Lipschitz function W is called a *generalized solution* of the HJ (1.4), if $W(1, x) = h(x)$, and, at each $(t, x) \in (0, 1) \times R^d$,

$$\max_{(p,q) \in \partial W(t,x)} \left[-p + \sup_{u \in \Gamma} H(t, x, u, q) \right] = 0.$$

It should be noted that, unlike the viscosity solution, generalized solutions of the associated HJ may *not* be unique under the assumptions (Aw1)–(A3). This is seen from the following example.

File: 409J 358114 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2311 Signs: 1346 . Length: 45 pic 0 pts, 190 mm

EXAMPLE 5.1. Consider the following optimal control problem

$$\text{minimize } \tfrac{1}{4}\int_s^1 u^2(t)\,dt,$$

$$\text{subject to } \begin{cases} \dot{x}(t)=u(t), & \text{a.e. } t\in[s,1], \\ x(s)=y, \end{cases}$$

$$\text{control } u(\cdot)\colon [s,1]\to\{r\in R^1 \mid -2\leqslant r\leqslant 2\}.$$

The associated HJ (1.4) in this problem reads

$$\begin{cases} -v_t(t,x)+\sup_{-2\leqslant u\leqslant 2}(-v_x(t,x)u-\tfrac{1}{4}u^2)=0, \\ v(1,x)=0. \end{cases} \tag{5.2}$$

$V\equiv 0$ is obviously a generalized solution of the above equation, which is also the value function of the optimal control problem (and therefore the unique viscosity solution of (5.2)). Now we are to show that the function given by

$$\hat{W}(t,x):=\begin{cases} 0, & \text{if } |x|\geqslant 1-t, \\ t+|x|-1, & \text{if } |x|<1-t, \end{cases} \tag{5.3}$$

is also a generalized solution of the HJ (5.2). Note $\hat{W}$ is globally Lipschitz continuous in (t,x), and is not continuously differentiable only at the lines $x=\pm(1-t)$ and $x=0$. Appealing to [2, Thm. 2.5.1], it is not difficult to verify that for any $t\in(0,1)$,

$$\begin{aligned} \partial\hat{W}(t,1-t)&=\{(p,q)\in R^2 \mid 0\leqslant p=q\leqslant 1\}, \\ \partial\hat{W}(t,t-1)&=\{(p,q)\in R^2 \mid 0\leqslant p<-q\leqslant 1\}, \\ \partial\hat{W}(t,0)&=\{1\}\times[-1,1]. \end{aligned}$$

Hence,

$$\max_{(p,q)\in\partial\hat{W}(t,1-t)}\Big[-p+\sup_{-2\leqslant u\leqslant 2}(-qu-\tfrac{1}{4}u^2)\Big]=\max_{0\leqslant p\leqslant 1}(-p+p^2)=0.$$

We can check the similar equality in terms of $\partial\hat{W}(t,t-1)$ and $\partial\hat{W}(t,0)$. Therefore, according to Definition 5.2, $\hat{W}$ is a generalized solution of the HJ (5.2).

While there may be a lot of generalized solutions of the HJ, it would be interesting to compare the generalized solutions and the unique viscosity solution (i.e., the value function). The following Theorems 5.1 and 5.2 are concerned with the problem.

File: 409J 358115 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2150 Signs: 987 . Length: 45 pic 0 pts, 190 mm

THEOREM 5.1. *Let W be any generalized solution of the HJ* (1.4). *Then*

$$W(s, y) \leqslant V(s, y)$$

for any $(s, y) \in [0, 1] \times R^d$.

Proof. Let $x(\cdot)$, $u(\cdot))$ be any admissible pair of the problem $C_{s,y}$. Set $f(t) := f(t, x(t), u(t))$. A similar argument as in proof of Theorem 3.1 yields that for a.e. $t \in [s, 1]$,

$$\begin{aligned}
\frac{d}{dt} W(t, x(t)) &= \lim_{h \to 0+} \frac{W(t+h, x(t)+hf(t)) - W(t, x(t))}{h} \\
&\geqslant -(-W)^0((t, x(t)); \quad (1, f(t))) \\
&= -\max_{(p,q) \in \partial(-W)(t,x))} [p + q \cdot f(t)] \\
&= \min_{(p,q) \in \partial W(t,x(t))} [p + q \cdot f(t)] \\
&\geqslant \min_{(p,q) \in \partial W(t,x(t))} [p - \sup_{u \in \Gamma} H(t, x(t), u, q) - L(t, x(t), u(t))] \\
&= -L(t, x(t), u(t)).
\end{aligned}$$

(by the definition of the generalized solutions)

Hence we conclude

$$W(1, x(1)) - W(s, y) = \int_s^1 \frac{d}{dt} W(t, x(t))\, dt \geqslant -\int_s^1 L(t, x(t), u(t))\, dt,$$

which implies

$$W(s, y) \leqslant \int_s^1 L(t, x(t), u(t))\, dt + h(x(1)) = J(s, y; u(\cdot)).$$

The desired result thus follows since $u(\cdot)$ is arbitrary. ∎

DEFINITION 5.3. The *closed super*-(resp. *sub*-)*differential* of v at $\hat{x} \in Q$, denoted by $\bar{D}^+ v_x(\hat{x})$(resp. $\bar{D}_x^- v(\hat{x})$), is a set defined by

$$\bar{D}_x^+ v(\hat{x}) := \{ p \in R^n \mid \text{there exist } x_n \to \hat{x},\ p_n \to p \text{ such that } p_n \in D_x^+ v(x_n) \}.$$

(resp.

$$\bar{D}_x^- v(\hat{x}) := \{ p \in R^n \mid \text{there exist } x_n \to \hat{x},\ p_n \to p \text{ such that } p_n \in D_x^- v(x_n) \}.)$$

File: 409J 358116 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2871 Signs: 1826 . Length: 45 pic 0 pts, 190 mm

Since $\{x \in Q \mid D_x^+ v(x) \neq \varnothing\}$ is dense in Q (cf. [5, 9]), $\bar{D}_x^+ v(x)$ is a *nonempty* set at each $x \in Q$ due to the locally Lipschitz property of v. Similarly, $\bar{D}_x^- v(x)$ is also *nonempty*.

LEMMA 5.1. *Let $v \in C(\bar{Q})$ be locally Lipschitz continuous. Then at each $x \in Q$:*

(a) $\text{Co } \bar{D}_z^+ v(z) = \text{Co } \bar{D}_x^- v(x) = \partial v(x)$, *where* Co *denotes the convex hull of a set.*

(b) $D_x^+ v(x) \subseteq \bar{D}_x^+ v(x) \subseteq \partial v(x)$, $D_x^- v(x) \subseteq \bar{D}_x^- v(x) \subseteq \partial v(x)$.

(c) $\bar{D}_x^+ v(x)$ *and* $\bar{D}_x^- v(x)$ *are compact sets.*

Proof. (a) Let $p \in \bar{D}_x^+ v(x)$. Then there exist $x_n \to x$, $p_n \to p$ such that $p_n \in D_x^+ v(x_n) \subseteq \partial v(x_n)$. By the upper semicontinuity of the multi-value function $\partial v(\cdot)$ [2, Prop. 2.1.5], it follows that $p \in \partial v(x)$. Note $\partial v(x)$ is closed and convex, so we have $\text{Co } \bar{D}_x^+ v(x) \subseteq \partial v(x)$. Conversely, since $\{p \mid$ there is $x_n \to x$ such that $p = \lim_{x_n \to x} v_x(x_n)\} \subseteq \bar{D}_x^+ v(x)$, hence $\partial v(x) \subseteq \text{Co } \bar{D}_x^+ v(x)$. Therefore we conclude that $\partial v(x) = \text{Co } \bar{D}_x^+ v(x)$. Similarly for the sub-case.

(b) and (c) are obvious from the definitions of the closed super-/sub-differentials. ∎

Remark 5.3. It is possible that all the inclusions in (b) of Lemma 5.1 are strict. For example, take $v(x) = |x|$. Then $D_x^+ v(0) = \varnothing$, $\bar{D}_x^+ v(0) = \{-1, 1\}$, and $\partial v(0) = [-1, 1]$.

The following lemma is readily seen on account of the continuity of the Hamiltonian H and the compactness of the control region Γ.

LEMMA 5.2. *Let W be a locally Lipschitz viscosity solution of the HJ (1.4), then at each $(t, x) \in (0, 1) \times R^d$,*

$$-p + \sup_{u \in \Gamma} H(t, x, u, q) \leqslant 0, \qquad \textit{for all} \quad (p, q) \in \bar{D}_{t,x}^+ W(t, x);$$

$$-p + \sup_{u \in \Gamma} H(t, x, u, q) \geqslant 0, \qquad \textit{for all} \quad (p, q) \in \bar{D}_{t,x}^- W(t, x).$$

THEOREM 5.2. *A locally Lipschitz function W is a viscosity solution of the HJ (1.4) if and only if W is a generalized solution and*

$$-p + \sup_{u \in \Gamma} H(t, x, u, q) = 0, \qquad \textit{for all} \quad (p, q) \in \bar{D}_{t,x}^- W(t, x). \tag{5.4}$$

Proof. It suffices to prove the "only if" part. By virtue of Lemma 5.2 together with the facts that $\partial W(t, x) = \text{Co } \bar{D}_{t,x}^+ W(t, x)$ and that $-p + \sup_{u \in \Gamma} H(t, x, u, q)$ is convex in (p, q), we conclude that

$$\max_{(p,q) \in \partial W(t,x)} \left[-p + \sup_{u \in \Gamma} H(t, x, u, q)\right] \leqslant 0. \tag{5.5}$$

File: 409J 358117 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 2801 Signs: 1918 . Length: 45 pic 0 pts, 190 mm

Therefore (5.4) follows from (b) of Lemma 5.1 and Lemma 5.2. Moreover, since $\bar{D}^-_{t,x} W(t, x)$ is *never* empty, the maximum in the hand side of (5.5) is precisely zero. This implies that W is a generalized solution of the HJ (1.4). ∎

Remark 5.4. The above theorem is originally due to Frankowska [8, Thm. 2.3]. But there is an argument in [8, p. 25] that seems to be inadequate: (5.5) together with the fact that

$$-p + \sup_{u \in \Gamma} H(t, x, u, q) = 0, \qquad \text{for all} \quad (p, q) \in D^-_{t,x} W(t, x)$$

cannot verify that the maximum in (5.5) is zero, since $D^-_{t,x} W(t, x)$ may be *empty*! Here, we modified Frankowska's proof by introducing the closed super- and sub-differential that are always nonempty.

Remark 5.5. That a generalized solution may not be the viscosity solution can also be seen from Example 5.1. The function $\hat{W}$ given by (5.2) is a generalized solution as shown in Example 5.1. On the other hand, for any $t \in (0, 1)$,

$$D^+_{t,x} \hat{W}(t, 0) = \varnothing, \qquad D^-_{t,x} \hat{W}(t, 0) = \{1\} \times [-1, 1].$$

Then for any $(p, q) \in D^-_{t,x} \hat{W}(t, 0)$,

$$-p + \sup_{-2 \leqslant u \leqslant 2} (-qu - \tfrac{1}{4}u^2) = -1 + q^2 \leqslant 0,$$

which violates the definition of the viscosity solutions. We see also that (5.4) is not satisfied by $\hat{W}$.

Acknowledgments

This research was carried out while the author was visiting the Department of Mathematics, Kobe University, and the Faculty of Management, University of Toronto. The support and the hospitality of the host institutions are gratefully acknowledged. Thanks are especially due to Professor Leonard D. Berkovitz for pointing out an error in the earlier version of the paper, and to the referee for helpful suggestions.

References

1. P. Cannarsa and H. Frankowska, Some characterizations of optimal trajectories in control theory, *SIAM J. Control Optim.* **29** (1991), 1322–1347.
2. F. H. Clarke, "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
3. F. H. Clarke and R. B. Vinter, Local optimality conditions and Lipschitzian solutions to the Hamilton–Jacobi equations, *SIAM J. Control Optim.* **21** (1983), 856–870.

File: 409J 358118 . By:CV . Date:07:04:93 . Time:12:51 LOP8M. V8.0. Page 01:01
Codes: 1812 Signs: 1369 . Length: 45 pic 0 pts, 190 mm

4. M. G. Crandall, L. C. Evans, and P. L. Lions, Some properties of viscosity solutions of Hamilton–Jacobi equations, *Trans. Amer. Math. Soc.* **282** (1984), 487–502.
5. M. G. Crandall and P. L. Lions, Viscosity solutions of Hamilton–Jacobi equations, *Trans. Amer. Math. Soc.* **277** (1983), 1–42.
6. W. H. Fleming and R. W. Rishel, "Deterministic and Stochastic Optimal Control," Springer-Verlag, New York, 1975.
7. W. H. Fleming, S. P. Sethi, and H. M. Soner, An optimal stochastic production planning problem with randomly fluctuating demand, *SIAM J. Control Optim.* **25** (1987), 1494–1502.
8. H. Frankowska, Hamilton–Jacobi equations: Viscosity solutions and generalized gradients, *J. Math. Anal. Appl.* **141** (1989), 21–26.
9. P. L. Lions, "Generalized Solutions of Hamilton–Jacobi Equations," Pitman, London, 1982.
10. P. L. Lions, Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations, part 2: Viscosity solutions and uniqueness, *Comm. Partial Differential Equations* **11** (1983), 1229–1276.
11. S. P. Sethi and Q. Zhang, "Hierarchical Production Planning in Dynamic Stochastic Manufacturing Systems: Asymptotic Optimality and Error Bounds, "Working Paper, University of Toronto, 1991.
12. X. Y. Zhou, Maximum principle, dynamic programming, and their connection in deterministic control, *J. Optim. Theory Appl.* **65** (1990), 363–373.