Viscosity Solutions of Hamilton-Jacobi Equations and Optimal Control Problems

(a short tutorial)

Alberto Bressan

S.I.S.S.A., Trieste 34014 Italy
and
Department of Mathematical Sciences, NTNU,
Trondheim N-7491, Norway

These notes are intended to provide a brief introduction to the theory of viscosity solutions to Hamilton-Jacobi equations, and applications to optimal control.

In the first section we introduce the notions of super- and sub-differential for a continuous scalar function, and derive some elementary properties. The second section contains the basic definition of *viscosity solution* to a Hamilton-Jacobi equation, while in section 3 we prove a standard comparison and uniqueness result for such solutions. In section 4 we consider a class of optimal control problems and prove a version of Bellman's *Dynamic Programming Principle*. The last section contains a derivation of the first order P.D.E. satisfied by the value function, in the viscosity sense. In view of the uniqueness of the viscosity solution, this yields a sufficient condition for the optimality of a control function.

1 - One-sided Differentials

Let $u : \Omega \mapsto I\!R$ be a scalar function, defined on an open set $\Omega \subseteq I\!R^n$. The set of *super-differentials* of u at a point x is defined as

$$D^+u(x) \doteq \left\{ p \in I\!R^n\,; \quad \limsup_{y\to x} \frac{u(y)-u(x)-p\cdot(y-x)}{|y-x|} \leq 0 \right\}. \tag{1.1}$$

In other words, a vector $p \in I\!R^n$ is a super-differential iff the plane $y \mapsto u(x)+p\cdot(y-x)$ is tangent from above to the graph of u at the point x (fig. 1a). Similarly, the set of *sub-differentials* of u at a point x is defined as

$$D^-u(x) \doteq \left\{ p \in I\!R^n\,; \quad \liminf_{y\to x} \frac{u(y)-u(x)-p\cdot(y-x)}{|y-x|} \geq 0 \right\}, \tag{1.2}$$

so that a vector $p \in I\!R^n$ is a sub-differential iff the plane $y \mapsto u(x)+p\cdot(y-x)$ is tangent from below to the graph of u at the point x (fig. 1b).

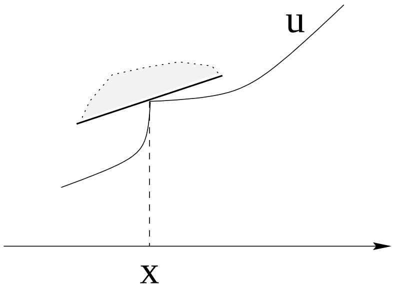

figure 1a

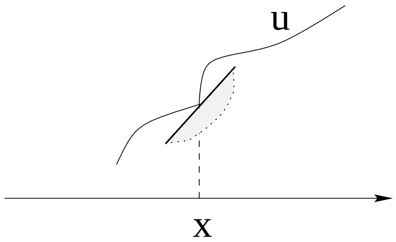

figure 1b

Example 1. Consider the function (fig. 2)

$$u(x) \doteq \begin{cases} 0 & \text{if} \quad x < 0, \\ \sqrt{x} & \text{if} \quad x \in [0,1], \\ 1 & \text{if} \quad x > 1. \end{cases}$$

In this case we have

$$D^+u(0) = \emptyset, \qquad\qquad D^-u(0) = [0, \infty[\,,$$

$$D^+u(x) = D^-u(x) = \big\{1/2\sqrt{x}\big\} \qquad\qquad x \in \,]0,1[\,,$$

$$D^+u(1) = \big[0,\ 1/2\big], \qquad\qquad D^-u(1) = \emptyset.$$

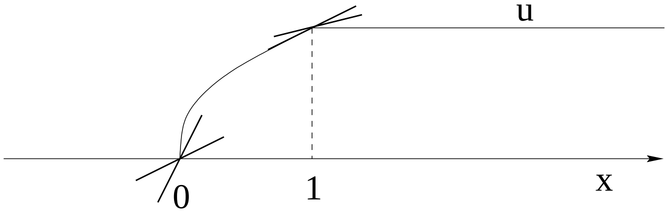

figure 2

If $\varphi \in \mathcal{C}^1$, its differential at a point x is written as $\nabla\varphi(x)$. The following characterization of super- and sub-differential is very useful.

Lemma 1. *Let $u \in \mathcal{C}(\Omega)$. Then*

(i) $p \in D^+u(x)$ if and only if there exists a function $\varphi \in \mathcal{C}^1(\Omega)$ such that $\nabla\varphi(x) = p$ and $u - \varphi$ has a local maximum at x.

(ii) $p \in D^-u(x)$ if and only if there exists a function $\varphi \in \mathcal{C}^1(\Omega)$ such that $\nabla\varphi(x) = p$ and $u - \varphi$ has a local minimum at x.

By adding a constant, it is not restrictive to assume that $\varphi(x) = u(x)$. In this case, we are saying that $p \in D^+u(x)$ iff there exists a smooth function $\varphi \geq u$ with $\nabla\varphi(x) = p$, $\varphi(x) = u(x)$. In other words, the graph of φ touches the graph of u from above at the point x (fig. 3a). A similar

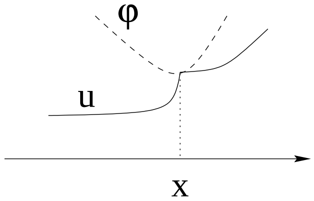

figure 3a

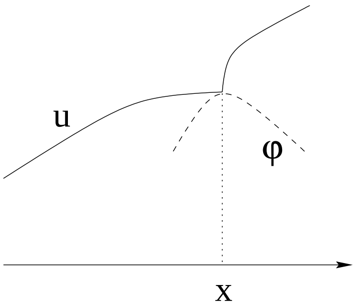

figure 3b

property holds for subdifferentials: $p \in D^- u(x)$ iff there exists a smooth function $\varphi \leq u$ whose graph touches from below the graph of u at the point x. (fig. 3b).

Proof of Lemma 1. Assume that $p \in D^+ u(x)$. Then we can find $\delta > 0$ and a continuous, increasing function $\sigma : [0, \infty[\mapsto I\!R$, with $\sigma(0) = 0$, such that

$$u(y) \leq u(x) + p \cdot (y - x) + \sigma\big(|y - x|\big)|y - x|$$

for $|y - x| < \delta$. Define

$$\rho(r) \doteq \int_0^r \sigma(t)\, dt$$

and observe that

$$\rho(0) = \rho'(0) = 0, \qquad \rho(2r) \geq \sigma(r)\, r\,.$$

By the above properties, the function

$$\varphi(y) \doteq u(x) + p \cdot (y - x) + \rho\big(2|y - x|\big)$$

is in $\mathcal{C}^1(\Omega)$ and satisfies

$$\varphi(x) = u(x), \qquad \nabla\varphi(x) = p.$$

Moreover, for $|y - x| < \delta$ we have

$$u(y) - \varphi(y) \leq \sigma\big(|y - x|\big)|y - x| - \rho\big(2|y - x|\big) \leq 0.$$

Hence, the difference $u - \varphi$ attains a local maximum at the point x.

To prove the opposite implication, assume that $D\varphi(x) = p$ and $u - \varphi$ has a local maximum at x. Then

$$\limsup_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} \leq \limsup_{y \to x} \frac{\varphi(y) - \varphi(x) - p \cdot (y - x)}{|y - x|} = 0\,. \tag{1.3}$$

This completes the proof of (i). The proof of (ii) is entirely similar.

Remark 1. By possibly replacing the function φ with $\tilde{\varphi}(y) = \varphi(y) \pm |y - x|^2$, it is clear that in the above lemma we can require that $u - \varphi$ attains a *strict* local maximum or local minimum at the point x.

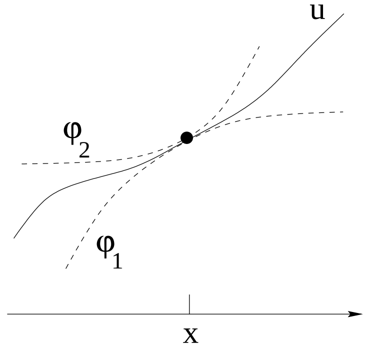

figure 4

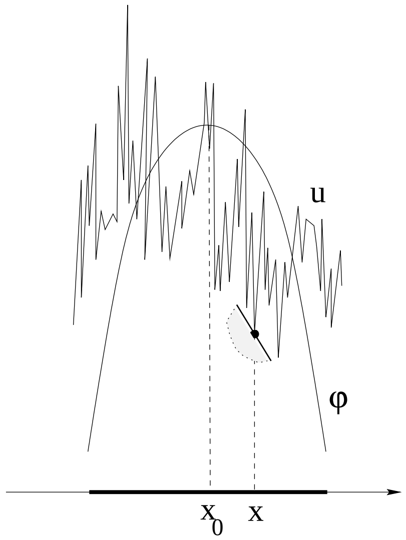

figure 5

Some simple properties of super- and sub-differential are collected in the next lemma.

Lemma 2. *Let $u \in \mathcal{C}(\Omega)$. Then*

(i) If u is differentiable at x, then

$$D^+u(x) = D^-u(x) = \{\nabla u(x)\}. \tag{1.4}$$

(ii) If the sets $D^+u(x)$ and $D^-u(x)$ are both non-empty, then u is differentiable at x, hence (1.4) holds.

(iii) The sets of points where a one-sided differential exists:

$$\Omega^+ \doteq \{x \in \Omega;\ \ D^+u(x) \neq \emptyset\}, \qquad \Omega^- \doteq \{x \in \Omega;\ \ D^-u(x) \neq \emptyset\} \tag{1.5}$$

are both non-empty. Indeed, they are dense in Ω.

Proof. Concerning (i), assume u is differentiable at x. Trivially, $\nabla u(x) \in D^\pm u(x)$. On the other hand, if $\varphi \in \mathcal{C}^1(\Omega)$ is such that $u - \varphi$ has a local maximum at x, then $\nabla\varphi(x) = \nabla u(x)$. Hence $D^+u(x)$ cannot contain any vector other than $\nabla u(x)$.

To prove (ii), assume that the sets $D^+u(x)$ and $D^-u(x)$ are both non-empty. Then there we can find $\delta > 0$ and $\varphi_1, \varphi_2 \in \mathcal{C}^1(\Omega)$ such that (fig. 4)

$$\varphi_1(x) = u(x) = \varphi_2(x), \qquad \varphi_1(y) \leq u(y) \leq \varphi_2(y) \qquad |y - x| < \delta.$$

By a standard comparison argument, this implies that u is differentiable at x and $\nabla u(x) = \nabla\varphi_1(x) = \nabla\varphi_2(x)$.

Concerning (iii), fix any ball $B(x_0, \rho) \subseteq \Omega$. By choosing $\varepsilon > 0$ sufficiently small, the smooth function (fig. 5)

$$\varphi(x) \doteq u(x_0) - \frac{|x - x_0|^2}{2\varepsilon}$$

is strictly negative on the boundary of the ball, where $|x - x_0| = \rho$. Since $u(x_0) = \varphi(x_0)$, the function $u - \varphi$ attains a local minimum at an interior point $x \in B(x_0, \rho)$. By Lemma 1, the sub-differential of u at x is non-empty. Indeed, $\nabla\varphi(x) = (x - x_0)/\varepsilon \in D^- u(x)$. The previous argument shows that, for every $x_0 \in \Omega$ and $\rho > 0$, the set Ω^- has non-empty intersection with the ball $B(x_0, \rho)$. Therefore Ω^- is dense in Ω. The case of super-differentials is entirely similar.

2 - Viscosity Solutions

In the following, we consider the first order, partial differential equation

$$F\big(x,\, u(x),\, \nabla u(x)\big) = 0 \tag{2.1}$$

defined on an open set $\Omega \in I\!R^n$. Here $F : \Omega \times I\!R \times I\!R^n \mapsto I\!R$ is a continuous (nonlinear) function.

Definition 1. A function $u \in \mathcal{C}(\Omega)$ is a *viscosity subsolution* of (2.1) if

$$F\big(x, u(x), p\big) \leq 0 \qquad \text{for every } x \in \Omega,\ p \in D^+u(x). \tag{2.2}$$

Similarly, $u \in \mathcal{C}(\Omega)$ is a *viscosity supersolution* of (2.1) if

$$F\big(x, u(x), p\big) \geq 0 \qquad \text{for every } x \in \Omega,\ p \in D^-u(x). \tag{2.3}$$

We say that u is a *viscosity solution* of (2.1) if it is both a supersolution and a subsolution in the viscosity sense.

Similar definitions also apply to evolution equations of the form

$$u_t + H\big(t, x, u, \nabla u\big) = 0, \tag{2.4}$$

where ∇u denotes the gradient of u w.r.t. x. Recalling Lemma 1, we can reformulate these definitions in an equivalent form:

Definition 2. A function $u \in \mathcal{C}(\Omega)$ is a *viscosity subsolution* of (2.4) if, for every $\mathcal{C}^1$ function $\varphi = \varphi(t, x)$ such that $u - \varphi$ has a local maximum at (t, x), there holds

$$\varphi_t(t, x) + H(t, x, u, \nabla\varphi) \leq 0. \tag{2.5}$$

Similarly, $u \in \mathcal{C}(\Omega)$ is a *viscosity supersolution* of (2.4) if, for every $\mathcal{C}^1$ function $\varphi = \varphi(t, x)$ such that $u - \varphi$ has a local minimum at (t, x), there holds

$$\varphi_t(t, x) + H(t, x, u, \nabla\varphi) \geq 0. \tag{2.6}$$

Remark 2. In the definition of subsolution, we are imposing conditions on u only at points x where the super-differential is non-empty. Even if u is merely continuous, say nowhere differentiable, there are many of these points. Indeed, by Lemma 2, the set of points x where $D^+u(x) \neq \emptyset$ is dense on Ω. Similarly, for supersolutions we impose conditions only at points where $D^-u(x) \neq \emptyset$.

Remark 3. If u is a $\mathcal{C}^1$ function that satisfies (2.1) at every $x \in \Omega$, then u is also a solution in the viscosity sense. Viceversa, if u is a viscosity solution, then the equality (2.1) must hold at every point x where u is differentiable. In particular, if u is Lipschitz continuous, then by Rademacher's theorem it is a.e. differentiable. Hence (2.1) holds a.e. in Ω.

Example 2. Set $F(x,u,u_x) \doteq 1 - |u_x|$. Then the function $u(x) = |x|$ is a viscosity solution of

$$1 - |u_x| = 0 \tag{2.7}$$

defined on the whole real line. Indeed, u is differentiable and satisfies the equation (2.7) at all points $x \neq 0$. Moreover, we have

$$D^+u(0) = \emptyset, \qquad\qquad D^-u(0) = [-1,\ 1]. \tag{2.8}$$

To show that u is a subsolution, there is nothing else to check. To show that u is a supersolution, take any $p \in [-1, 1]$. Then $1 - |p| \geq 0$, as required.

It is interesting to observe that the same function $u(x) = |x|$ is NOT a viscosity solution of the equation

$$|u_x| - 1 = 0 \tag{2.9}$$

Indeed, at $x = 0$, taking $p = 0 \in D^-u(0)$ we find $|0| - 1 < 0$. In conclusion, the function $u(x) = |x|$ is a viscosity subsolution of (2.9), but not a supersolution.

The definition of viscosity solution is naturally motivated by the study of vanishing viscosity limits.

Lemma 3. *Let u_ε be a sequence of smooth solutions to the viscous equation*

$$F\big(x,\ u_\varepsilon(x),\ \nabla u_\varepsilon(x)\big) = \varepsilon\, \Delta u_\varepsilon\,. \tag{2.10}$$

Assume that, as $\varepsilon \to 0+$, we have the convergence $u_\varepsilon \to u$ uniformly on an open set $\Omega \subseteq I\!\!R^n$. Then u is a viscosity solution of (2.1).

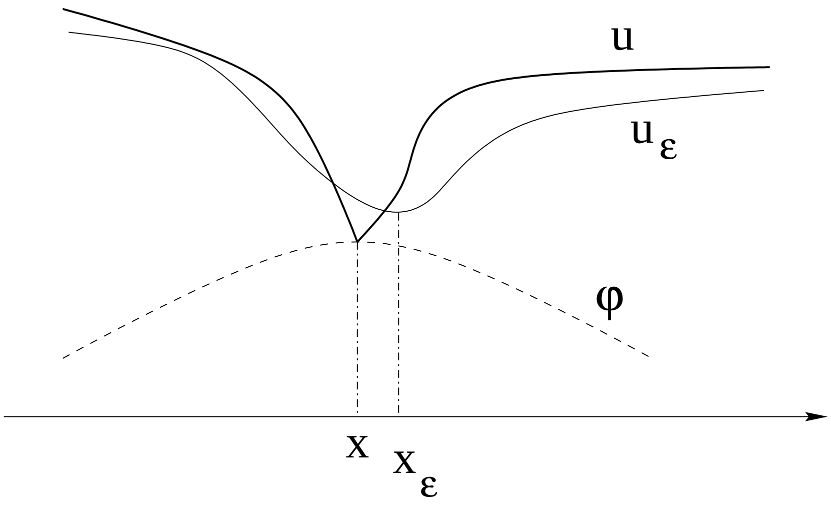

figure 6

Proof. Fix $x \in \Omega$ and assume $p \in D^-u(x)$. To prove that u is a supersolution we need to show that $F(x,\, u(x),\, p) \geq 0$.

1. By Lemma 1 and Remark 1, there exists $\varphi \in \mathcal{C}^1$ with $\nabla\varphi(x) = p$, $\varphi(x) = u(x)$ and $\varphi(y) < u(y)$ for all $y \neq x$. For any $\delta > 0$ we can then find $0 < \rho \leq \delta$ and a function $\psi \in \mathcal{C}^2$ such that

$$\big|\nabla\varphi(y) - \nabla\varphi(x)\big| \leq \delta \qquad\qquad \text{if} \quad |y-x| \leq \rho\,, \tag{2.11}$$

$$\|\psi - \varphi\|_{\mathcal{C}^1} \leq \delta \tag{2.12}$$

and such that each function $u_\varepsilon - \psi$ has a local minimum inside the ball $B(x;\,\rho)$, for $\varepsilon > 0$ small enough (fig. 6).

2. Let x_ε be the location of this local minimum of $u_\varepsilon - \psi$. Since u_ε is smooth, this implies

$$\nabla\psi(x_\varepsilon) = \nabla u(x_\varepsilon), \qquad\qquad \Delta u(x_\varepsilon) \geq \Delta\psi(x_\varepsilon). \tag{2.13}$$

Hence from (2.10) it follows

$$F\big(x,\, u_\varepsilon(x_\varepsilon),\;\; \nabla\psi(x_\varepsilon)\big) \geq \varepsilon\,\Delta\psi(x_\varepsilon) \tag{2.14}$$

3. Extract a convergent subsequence $x_\varepsilon \to \tilde{x}$. Clearly $|\tilde{x} - x| \leq \rho$. Since $\psi \in \mathcal{C}^2$, we can pass to the limit in (2.14) and conclude

$$F\big(x,\, u(\tilde{x}),\;\; \nabla\psi(\tilde{x})\big) \geq 0 \tag{2.15}$$

By (2.11)-(2.12) we have

$$\begin{aligned} \big|\nabla\psi(\tilde{x}) - p\big| &\leq \big|\nabla\psi(\tilde{x}) - \nabla\varphi(\tilde{x})\big| + \big|\nabla\varphi(\tilde{x}) - \nabla\varphi(x)\big| \\ &\leq \delta + \delta\,. \end{aligned} \tag{2.16}$$

Since $\delta > 0$ can be taken arbitrarily small, (2.16) and the continuity of F imply (2.3), showing that u is a supersolution. The other half of the proof is entirely similar.

3 - Comparison Theorems

A most remarkable feature of the notion of viscosity solutions is that on one hand it requires a minimum amount of regularity (just continuity), and on the other hand it is stringent enough to yield general comparison and uniqueness theorems.

The uniqueness proofs are based on a technique of doubling of variables, which reminds of Kruzhkov's uniqueness theorem for conservation laws [K]. We now illustrate this basic technique in a simple setting.

Theorem 1. *Let $\Omega \subset I\!R^n$ be a bounded open set. Assume that $u_1, u_2 \in \mathcal{C}(\overline{\Omega})$ are, respectively, viscosity sub- and supersolutions of*

$$u + H(x, \nabla u) = 0 \qquad\qquad x \in \Omega, \tag{3.1}$$

and

$$u_1 \le u_2 \qquad\qquad on \;\; \partial\Omega. \tag{3.2}$$

Moreover, assume that the function $H : \Omega \times I\!R^n \mapsto I\!R$ *satisfies*

$$\big|H(x,p) - H(y,p)\big| \le \omega\Big(|x-y|\big(1+|p|\big)\Big), \tag{3.3}$$

where $\omega : [0, \infty[\mapsto [0, \infty[$ *is continuous and non-decreasing, with* $\omega(0) = 0$*. Then*

$$u_1 \le u_2 \qquad\qquad on \;\; \overline{\Omega}. \tag{3.4}$$

Proof. To appreciate the main idea of the proof, consider first the case where u_1, u_2 are smooth. If the conclusion (3.4) fails, then the difference $u_1 - u_2$ attains a positive maximum at a point $x_0 \in \Omega$. This implies $p \doteq \nabla u_1(x_0) = \nabla u_2(x_0)$. By definition of sub- and supersolution, we now have

$$\begin{aligned} u_1(x_0) + H(x_0, p) &\le 0, \\ u_2(x_0) + H(x_0, p) &\ge 0. \end{aligned} \tag{3.5}$$

Subtracting the second from the first inequality in (3.5) we conclude $u_1(x_0) - u_2(x_0) \le 0$, reaching a contradiction.

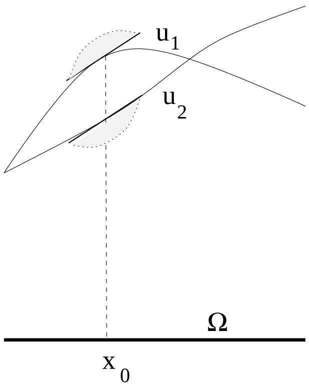

figure 7a

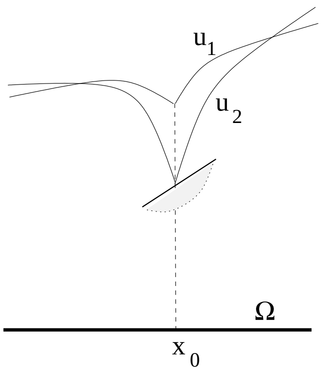

figure 7b

Next, consider the non-smooth case. We can repeat the above argument and reach again a contradiction provided that we can find a point x_0 such that (fig. 7a)

(i) $u_1(x_0) > u_2(x_0)$,

(ii) some vector p lies at the same time in the upper differential $D^+u_1(x_0)$ and in the lower differential $D^-u_2(x_0)$.

A natural candidate for x_0 is a point where $u_1 - u_2$ attains a global maximum. Unfortunately, at such point one of the sets $D^+u_1(x_0)$ or $D^-u_2(x_0)$ may be empty, and the argument breaks down (fig. 7b). To proceed further, the key observation is that we don't need to compare values of u_1 and u_2 at exactly the same point. Indeed, to reach a contradiction, it suffices to find nearby points x_ε and y_ε such that (fig. 8) (i') $u_1(x_\varepsilon) > u_2(y_\varepsilon)$,

(ii') some vector p lies at the same time in the upper differential $D^+u_1(x_\varepsilon)$ and in the lower differential $D^-u_2(y_\varepsilon)$.

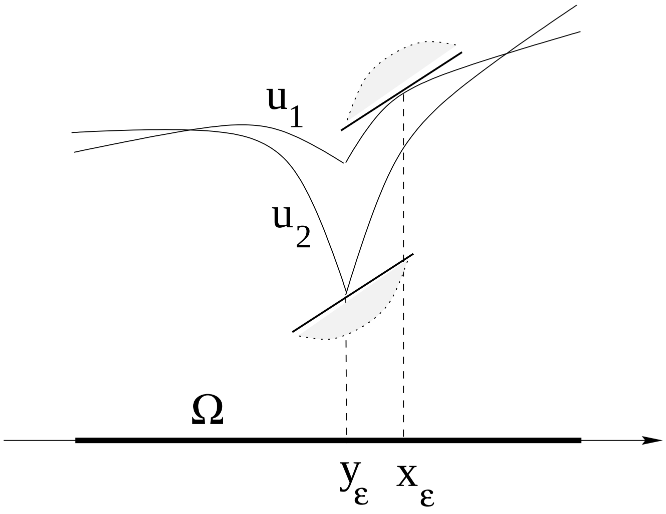

figure 8

Can we always find such points? It is here that the variable-doubling technique comes in. The trick is to look at the function of two variables

$$\Phi_\varepsilon(x,y) \doteq u_1(x) - u_2(y) - \frac{|x-y|^2}{2\varepsilon}. \tag{3.6}$$

This clearly admits a global maximum over the compact set $\overline{\Omega} \times \overline{\Omega}$. If $u_1 > u_2$ at some point x_0, this maximum will be strictly positive. Moreover, taking $\varepsilon > 0$ sufficiently small, the boundary conditions imply that the maximum is attained at some interior point $(x_\varepsilon, y_\varepsilon) \in \Omega \times \Omega$. Notice that the points $x_\varepsilon, y_\varepsilon$ must be close to each other, otherwise the penalization term in (3.6) will be very large and negative.

We now observe that the function of just one variable

$$x \mapsto u_1(x) - \left(u_2(y_\varepsilon) + \frac{|x-y_\varepsilon|^2}{2\varepsilon}\right) = u_1(x) - \varphi_1(x) \tag{3.7}$$

attains its maximum at the point x_ε. Hence by Lemma 1

$$\frac{x_\varepsilon - y_\varepsilon}{\varepsilon} = \nabla\varphi_1(x_\varepsilon) \in D^+u_1(x_\varepsilon).$$

Moreover, the function of just one variable

$$y \mapsto u_2(y) - \left(u_1(x_\varepsilon) - \frac{|x_\varepsilon - y|^2}{2\varepsilon}\right) = u_2(y) - \varphi_2(y) \tag{3.8}$$

attains its minimum at the point y_ε. Hence

$$\frac{x_\varepsilon - y_\varepsilon}{\varepsilon} = \nabla\varphi_2(y_\varepsilon) \in D^- u_2(y_\varepsilon).$$

We have thus discovered two points x_ε, y_ε and a vector $p = (x_\varepsilon - y_\varepsilon)/\varepsilon$ which satisfy the conditions (i')-(ii').

We now work out the details of the proof, in several steps.

1. If the conclusion fails, then there exists $x_0 \in \Omega$ such that

$$u_1(x_0) - u_2(x_0) = \max_{x\in\overline{\Omega}} \big\{u_1(x) - u_2(x)\big\} \doteq \delta > 0. \tag{3.9}$$

For $\varepsilon > 0$, call $(x_\varepsilon, y_\varepsilon)$ a point where the function Φ_ε in (3.6) attains its global maximum on the compact set $\overline{\Omega}\times\overline{\Omega}$. By (3.9) one has

$$\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \geq \delta > 0. \tag{3.10}$$

2. Call M an upper bound for all values $\big|u_1(x)\big|$, $\big|u_2(x)\big|$, as $x \in \overline{\Omega}$. Then

$$\Phi_\varepsilon(x,y) \leq 2M - \frac{|x-y|^2}{2\varepsilon},$$

$$\Phi_\varepsilon(x,y) \leq 0 \qquad \text{if} \quad |x-y|^2 \geq M\varepsilon.$$

Hence (22) implies

$$|x_\varepsilon - y_\varepsilon| \leq \sqrt{M\varepsilon}\,. \tag{3.11}$$

3. By the uniform continuity of the functions u_2 on the compact set $\overline{\Omega}$, for $\varepsilon' > 0$ sufficiently small we have

$$\big|u_2(x) - u_2(y)\big| < \frac{\delta}{2} \qquad\qquad \text{whenever} \quad |x-y| \leq \sqrt{M\varepsilon'}\,. \tag{3.12}$$

We now show that, choosing $\varepsilon < \varepsilon'$, the points $x_\varepsilon, y_\varepsilon$ cannot lie on the boundary of Ω. Indeed, if $x_\varepsilon \in \partial\Omega$, then by (3.11) and (3.12)

$$\begin{aligned}\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) &\leq \big(u_1(x_\varepsilon) - u_2(x_\varepsilon)\big) + \big|u_2(x_\varepsilon) - u_2(y_\varepsilon)\big| - \frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon}\\ &\leq 0 + \delta/2 + 0,\end{aligned}$$

against (3.10).

4. Having shown that $x_\varepsilon, y_\varepsilon$ are interior points, we consider the functions of one single variable φ_1, φ_2 defined at (3.7)-(3.8). Since x_ε provides a local maximum for $u_1 - \varphi_1$ and y_ε provides a local minimum for $u_2 - \varphi_2$, we conclude that

$$p \doteq \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} \in D^+ u_1(x_\varepsilon) \cap D^- u_2(y_\varepsilon). \tag{3.13}$$

From the definition of viscosity sub- and supersolution we now obtain

$$\begin{aligned} u_1(x_\varepsilon) + H(x_\varepsilon, p) &\leq 0, \\ u_2(y_\varepsilon) + H(y_\varepsilon, p) &\geq 0. \end{aligned} \tag{3.14}$$

5. Observing that

$$\delta \leq \Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \leq u_1(x_\varepsilon) - u_2(x_\varepsilon) + \big|u_2(x_\varepsilon) - u_2(y_\varepsilon)\big| - \frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon},$$

by (3.9) we see that

$$\big|u_2(x_\varepsilon) - u_2(y_\varepsilon)\big| - \frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \geq 0$$

and hence, by the uniform continuity of u_2,

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \to 0 \qquad \text{as } \varepsilon \to 0. \tag{3.15}$$

6. Recalling (3.10) and subtracting the second from the first inequality in (3.14) we obtain

$$\begin{aligned} \delta &\leq \Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \\ &\leq u_1(x_\varepsilon) - u_2(y_\varepsilon) \\ &\leq \big|H(x_\varepsilon, p) - H(y_\varepsilon, p)\big| \\ &\leq \omega\Big((|x_\varepsilon - y_\varepsilon| \cdot \big(1 + |x_\varepsilon - y_\varepsilon|\varepsilon^{-1}\big)\Big). \end{aligned} \tag{3.16}$$

This yields a contradiction, Indeed, by (3.11) and (3.15) the right hand side of (3.16) can be made arbitrarily small by letting $\varepsilon \to 0$.

An easy consequence of the above result is the following uniqueness theorem for the boundary value problem

$$u + H(x, \nabla u) = 0 \qquad x \in \Omega, \tag{3.17}$$

$$u = \psi \qquad x \in \partial\Omega. \tag{3.18}$$

Corollary 1. *Let $\Omega \subset I\!R^n$ be a bounded open set. Let the Hamiltonian function H satisfy the equicontinuity assumption (3.3). Then the boundary value problem (3.17)-(3.18) admits at most one viscosity solution.*

Proof. Let u_1, u_2 be viscosity solutions. Since u_1 is a subsolution and u_2 is a supersolution, and $u_1 = u_2$ on $\partial\Omega$, by Theorem 1 we conclude $u_1 \leq u_2$ on $\overline{\Omega}$. Reversing the roles of u_1 and u_2, we deduce $u_2 \leq u_1$, completing the proof.

By similar techniques, comparison and uniqueness results can be proved also for Hamilton-Jacobi equations of evolutionary type. Consider the Cauchy problem

$$u_t + H(t, x, \nabla u) = 0 \qquad (t, x) \in \,]0, T[\, \times I\!R^n, \tag{3.19}$$

$$u(0,x) = g(x) \qquad\qquad x \in I\!R^n. \tag{3.20}$$

Here and in the sequel, it is understood that $\nabla u \doteq (u_{x_1}, \ldots, u_{x_n})$ always refers to the gradient of u w.r.t. the space variables. We shall assume that the function H satisfies the following equicontinuity conditions:

(H1) H is uniformly continuous on $[0,T] \times I\!R^n \times K$, for every compact set $K \subset I\!R^n$.

(H2) There exists a continuous, non-decreasing function $\omega : [0,\infty[\, \times [0,\infty[$ with $\omega(0) = 0$ such that

$$\big|H(t,x,p) - H(t,y,p)\big| \leq \omega\Big(|x-y|\big(1+|p|\big)\Big). \tag{3.21}$$

Theorem 2. *Let the function $H : [0,T] \times I\!R^n \times I\!R^n$ satisfy the equicontinuity assumptions (H1)-(H2). Let u_1, u_2 be uniformly continuous sub- and super-solutions of (3.19) respectively. If $u_1(0,x) \leq u_2(0,x)$ for all $x \in I\!R^n$, then*

$$u_1(t,x) \leq u_2(t,x) \qquad\qquad \text{for all } \ (t,x) \in [0,T] \times I\!R^n. \tag{3.22}$$

For a proof, see [B-C] or [E].

Corollary 2. *Let the function H satisfy the assumptions (H1)-(H2). Then the Cauchy problem (3.19)-(3.20) admits at most one uniformly continuous viscosity solution $u : [0,T] \times I\!R^n \mapsto I\!R$.*

4 - An Optimal Control Problem

We now consider a control system of the form

$$\dot x(t) = f\big(x(t),\, \alpha(t)\big) \qquad\qquad s < t < T, \tag{4.1}$$

$$x(s) = y. \tag{4.2}$$

Here $x \in I\!R^n$, while the *control function* $\alpha : [s,T] \mapsto A$ is required to take values inside a compact set $A \subset I\!R^m$. Moreover, we assume that $f : I\!R^n \times A \mapsto I\!R^n$ is a continuous function such that

$$\big|f(x,a)\big| \leq C, \qquad\qquad \big|f(x,a) - f(y,a)\big| \leq C\,|x-y| \qquad \text{for all } x,y \in I\!R^n,\ a \in A \tag{4.3}$$

for some constant C. Under these assumptions, for every choice of the measurable control function $\alpha(\cdot)$ the Cauchy problem (4.1)-(4.2) has a unique solution, which we denote as

$$t \mapsto x(t; s, y, \alpha) \tag{4.4}$$

or sometimes simply as $t \mapsto x(t)$. We call $x(\cdot\,; \alpha)$ the *response* of the system to the control $\alpha(\cdot)$, and $x(t)$ the *state* of the system at time t.

For a given initial condition $x(s) = y$, different choices of the control function α thus yield different trajectories. In an optimal control problem, we wish to single out one particular control $\alpha^* : [s,T] \mapsto A$, which is optimal w.r.t. a specified *cost criterion*. In the following, we consider a cost of the form

$$J(s,y,\alpha) \doteq \int_s^T h\big(x(t),\, \alpha(t)\big)\, dt + g\big(x(T)\big). \tag{4.5}$$

Here it is understood that $x(t) = x(t; s, y, \alpha)$, while

$$h : I\!R^n \times A \mapsto I\!R, \qquad\qquad g : I\!R^n \mapsto I\!R$$

are continuous functions. We call h the *running cost* and g the *terminal cost*. We shall assume that the functions h, g satisfy the bounds

$$\big|h(x,a)\big| \leq C, \qquad\qquad \big|g(x)\big| \leq C, \tag{4.6}$$

$$\big|h(x,a) - h(y,a)\big| \leq C\,|x-y|, \qquad\qquad \big|g(x) - g(y)\big| \leq C\,|x-y|, \tag{4.7}$$

for all $x, y \in I\!R^n$, $a \in A$.

According to the method of *dynamic programming*, an optimal control problem can be studied by looking at the *value function*:

$$u(s,y) \doteq \inf_{\alpha(\cdot)} J(s,y,\alpha). \tag{4.8}$$

We consider here a whole family of optimal control problem, all with the same dynamics (4.1) and cost functional (4.5). We are interested in how the minimum cost varies, as a function of the initial conditions (4.2). It turns out that the value function u can be characterized as the unique viscosity solution to a Hamilton-Jacobi equation. Towards this goal, a basic step is provided by Bellman's *Dynamic Programming Principle*:

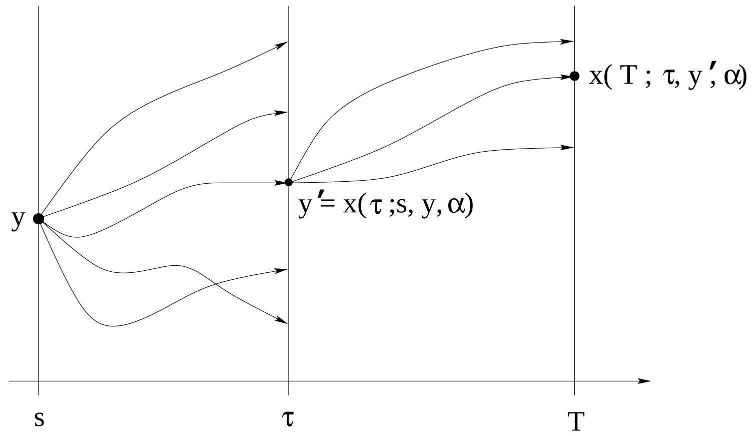

figure 9

Theorem 3. *For every $s < \tau < T$, $y \in I\!R^n$, one has*

$$u(s,y) = \inf_{\alpha(\cdot)} \left\{ \int_s^\tau h\big(x(t;s,y,\alpha),\ \alpha(t)\big)\,dt \ +\ u\big(\tau,\ x(\tau;s,y,\alpha)\big)\right\}. \tag{4.9}$$

In other words (fig. 9), the optimization problem on the time interval $[s,T]$ can be split into two separate problems:

- As a first step, we solve the optimization problem on the sub-interval $[\tau, T]$, with running cost h and terminal cost g. In this way, we determine the value function $u(\tau,\cdot)$, at time τ.

- As a second step, we solve the optimization problem on the sub-interval $[s,\tau]$, with running cost h and terminal cost $u(\tau,\cdot)$, determined by the first step.

At the initial time s, by (4.9) we are saying that the value function $u(s,\cdot)$ obtained in step 2 is the same as the value function corresponding to the global optimization problem over the whole interval $[s,T]$.

Proof. Call J_τ the right hand side of (4.9).

1. To prove that $J_\tau \leq u(s,y)$, fix $\varepsilon > 0$ and choose a control $\alpha : [s,T] \mapsto A$ such that

$$J(s,y,\alpha) \leq u(s,y) + \varepsilon.$$

Observing that

$$u\big(\tau, x(\tau;s,y,\alpha)\big) \leq \int_\tau^T h\big(t,\ x(t;s,y,\alpha)\big)\,dt + g\big(x(T;s,y,\alpha)\big),$$

we conclude

$$\begin{aligned} J_\tau &\leq \int_s^\tau h\big(t,\ x(t;s,y,\alpha)\big)\,dt + u\big(\tau, x(\tau;s,y,\alpha)\big) \\ &\leq J(s,y,\alpha) \\ &\leq u(s,y) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this first inequality is proved.

2. To prove that $u(s,y) \leq J_\tau$, fix $\varepsilon > 0$. Then there exists a control $\alpha' : [s,\tau] \mapsto A$ such that

$$\int_s^\tau h\big(t,\ x(t;s,y,\alpha')\big)\,dt + u\big(\tau,\ x(\tau;s,y,\alpha')\big) \leq J_\tau + \varepsilon. \tag{4.10}$$

Moreover, there exists a control $\alpha'' : [\tau,T] \mapsto A$ such that

$$J\big(\tau,\ x(\tau;s,y,\alpha'),\ \alpha''\big) \leq u\big(\tau,\ x(\tau;s,y,\alpha')\big) + \varepsilon. \tag{4.11}$$

One can now define a new control $\alpha : [s,T] \mapsto A$ as the concatenation of α', α'':

$$\alpha(t) \doteq \begin{cases} \alpha'(t) & \text{if} \quad t \in [s,\tau], \\ \alpha''(t) & \text{if} \quad t \in \,]\tau,T]. \end{cases}$$

By (4.10) and (4.11) it is now easy to check that

$$u(s,y) \leq J(s,y,\alpha) \leq J_\tau + 2\varepsilon.$$

Since $\varepsilon > 0$ can be arbitrarily small, this second inequality is also proved.

5 - The Hamilton-Jacobi-Bellman Equation

The main goal of this section is to characterize the value function as the unique solution of a first order P.D.E., in the viscosity sense. In turn, this will provide a sufficient condition for the global optimality of a control function $\alpha(\cdot)$. As in the previous section, we assume here that the function f satisfies the bounds (4.3), and that the cost functions h, g satisfy (4.6)-(4.7). We first derive some regularity properties of the value function.

Lemma 4. *There exists a constant C' such that the value function u in (4.8) satisfies*

$$\big|u(s,y)\big| \leq C', \tag{5.1}$$

$$\big|u(s,y) - u(s',y')\big| \leq C'\big(|s-s'| + |y-y'|\big) \tag{5.2}$$

Proof. By (4.6), for any control function α one has

$$\big|J(s,y,\alpha)\big| \leq \int_s^T \Big|h\big(x(t),\, \alpha(t)\big)\Big|\, dt + \Big|g\big(x(T)\big)\Big| \leq CT + C.$$

Hence (5.1) holds with $C' \doteq C(1+T)$.

To prove (5.2), let s, y, y' be given. For any $\varepsilon > 0$, choose a control $\alpha : [s,T] \mapsto A$ such that

$$J(s,y,\alpha) \leq u(s,y) + \varepsilon.$$

If we now use the same control α starting from y' instead of y, by the Lipschitz continuity of f, the corresponding trajectory will satisfy

$$\big|x(t;\, s,y',\alpha) - x(t;\, s,y,\alpha)\big| \leq e^{C(t-s)}|y'-y| \leq C_1|y'-y| \tag{5.3}$$

with $C_1 \doteq e^{CT}$. Therefore

$$\begin{aligned} J(s,y',\alpha) \leq J(s,y,\alpha) &+ \int_s^T \Big|h\big(x(t;s,y',\alpha),\, \alpha(t)\big) - h\big(x(t;s,y,\alpha),\, \alpha(t)\big)\Big|\, dt \\ &+ \Big|g\big(x(T;s,y',\alpha)\big) - g\big(x(T;s,y,\alpha)\big)\Big| \\ &\leq J(s,y,\alpha) + T\cdot CC_1|y'-y| + CC_1|y'-y|, \\ &\leq u(s,y) + \varepsilon + (T+1)CC_1|y'-y|. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this implies

$$u(s,y') \leq u(s,y) + C_2|y'-y| \qquad \text{for all } \ s \in [0,T], \quad y, y' \in I\!R^n,$$

with $C_2 \doteq (T+1)CC_1$. Reversing the role of y, y' we conclude

$$\big|u(s,y') - u(s,y)\big| \leq C_2|y'-y|, \tag{5.4}$$

proving the Lipschitz continuity of the value function w.r.t. y.

To prove the Lipschitz continuity w.r.t. the time variable s, assume $0 \leq s < s' \leq T$. Fix any control function $\alpha(\cdot)$. Recalling that $|\dot{x}| = \big|f(x,\alpha)\big| \leq C$, using (4.9) we obtain the bound

$$
\begin{aligned}
u(s,y) &\leq \int_s^{s'} h\big(x(t;s,y,\alpha),\ \alpha(t)\big)\,dt + u\big(s',\ x(s',s,y,\alpha)\big) \\
&\leq C|s'-s| + u(s',y) + C_2\big|y - x(s',s,y,\alpha)\big| \\
&\leq u(s',y) + (C + C_2C)|s'-s|.
\end{aligned} \tag{5.5}
$$

Viceversa, for any $\varepsilon > 0$, there exists a nearly optimal control $\alpha : [s,T] \mapsto A$ such that

$$
\begin{aligned}
u(s,y) &\geq J(s,y,\alpha) - \varepsilon \\
&\geq \int_s^{s'} h\big(x(t;s,y,\alpha)\,\alpha(t)\big)\,dt + u\big(s',\ x(s',s,y,\alpha)\big) - \varepsilon \\
&\geq -C|s'-s| + u(s',y) - C_2\big|y - x(s',s,y,\alpha)\big| - \varepsilon \\
&\geq u(s',y) - (C + C_2C)|s'-s| - \varepsilon,
\end{aligned} \tag{5.6}
$$

because of (5.4). Together, (5.5) and (5.6) prove the Lipschitz continuity of u also w.r.t. the time variable s.

We can now state the main result of this section, which characterizes of the value function as the viscosity solution of a suitable first order P.D.E.

Theorem 4. *In connection with the control system (4.1), consider the value function $u = u(s,y)$ defined by (4.8) and (4.5). Then u is a viscosity solution of the Hamilton-Jacobi-Bellman equation*

$$
-\Big[u_t + H(x,\nabla u)\Big] = 0 \qquad (t,x) \in]0,T[\,\times I\!R^n, \tag{5.7}
$$

with terminal condition

$$
u(T,x) = g(x) \qquad x \in I\!R^n, \tag{5.8}
$$

and Hamiltonian function

$$
H(x,p) \doteq \min_{a\in A} \big\{f(x,a)\cdot\nabla u + h(x,a)\big\}. \tag{5.9}
$$

Proof. By Lemma 3, the value function is uniformly Lipschitz continuous on $[0,T] \times I\!R^n$. The terminal conditions (5.8) is obvious. To show that u is a viscosity solution, let $\varphi \in \mathcal{C}^1\big(]0,T[\,\times I\!R^n\big)$. Two separate statements need to be proved:

(P1) If $u - \varphi$ attains a local maximum at a point $(t_0,x_0) \in]0,T[\,\times I\!R^n$, then

$$
\varphi_t(t_0,x_0) + \min_{a\in A} \big\{f(x_0,a)\cdot\nabla\varphi(t_0,x_0) + h(x_0,a)\big\} \geq 0. \tag{5.10}
$$

(P2) If $u - \varphi$ attains a local minimum at a point $(t_0,x_0) \in]0,T[\,\times I\!R^n$, then

$$
\varphi_t(t_0,x_0) + \min_{a\in A} \big\{f(x_0,a)\cdot\nabla\varphi(t_0,x_0) + h(x_0,a)\big\} \leq 0. \tag{5.11}
$$

1. To prove (P1), we can assume that

$$u(t_0, x_0) = \varphi(t_0, x_0), \qquad\qquad u(t,x) \le \varphi(t,x) \quad \text{for all } t, x. \tag{5.12}$$

If (5.10) does not hold, then there exists $a \in A$ and $\theta > 0$ such that

$$\varphi_t(t_0, x_0) + f(x_0, a) \cdot \nabla\varphi(t_0, x_0) + h(x_0, a) < -\theta. \tag{5.13}$$

We shall derive a contradiction by showing that this control value a is "too good to be true". Namely, by choosing a control function $\alpha(\cdot)$ with $\alpha(t) \equiv a$ for $t \in [t_0,\ t_0 + \delta]$ and such that α is nearly optimal on the remaining interval $[t_0 + \delta,\ T]$, we obtain a total cost $J(t_0, x_0, \alpha)$ strictly smaller than $u(t_0, x_0)$. Here is the proof. By continuity, (5.13) implies

$$\varphi_t(t,x) + f(x,a) \cdot \nabla\varphi(t,x) < -h(x,a) - \theta. \tag{5.14}$$

whenever

$$|t - t_0| < \delta, \qquad\qquad |x - x_0| \le C\delta, \tag{5.15}$$

for some $\delta > 0$ small enough and C the constant in (4.3). Let $x(t) \doteq x(t; t_0, x_0, a)$ be the solution of

$$\dot x(t) = f\big(x(t), a\big), \qquad\qquad x(t_0) = x_0,$$

i.e. the trajectory corresponding to the constant control $\alpha(t) \equiv a$. We then have

$$\begin{aligned} u\big(t_0 + \delta,\ x(t_0 + \delta)\big) - u(t_0, x_0) &\le \varphi\big(t_0 + \delta,\ x(t_0 + \delta)\big) - \varphi(t_0, x_0) \\ &= \int_{t_0}^{t_0+\delta} \frac{d}{dt}\varphi\big(t,\ x(t)\big)\, dt \\ &= \int_{t_0}^{t_0+\delta} \Big\{\varphi_t\big(t,\ x(t)\big) + f\big(x(t), a\big) \cdot \nabla\varphi\big(t,\ x(t)\big)\Big\}\, dt \\ &\le -\int_{t_0}^{t_0+\delta} h\big(x(t),\ a\big)\, dt - \delta\theta, \end{aligned} \tag{5.16}$$

because of (5.13). On the other hand, the Dynamic Programming Principle (4.9) yields

$$u\big(t_0 + \delta,\ x(t_0 + \delta)\big) - u(t_0, x_0) \ge -\int_{t_0}^{t_0+\delta} h\big(t,\ x(t)\big)\, dt. \tag{5.17}$$

Together, (5.16) and (5.17) yield a contradiction, hence (P1) must hold.

2. To prove (P2), we can assume that

$$u(t_0, x_0) = \varphi(t_0, x_0), \qquad\qquad u(t,x) \ge \varphi(t,x) \quad \text{for all } t, x. \tag{5.18}$$

If (P2) fails, then there exists $\theta > 0$ such that

$$\varphi_t(t_0, x_0) + f(x_0, a) \cdot \nabla\varphi(t_0, x_0) + h(x_0, a) > \theta \qquad \text{for all } a \in A. \tag{5.19}$$

In this case, we shall reach a contradiction by showing that no control function $\alpha(\cdot)$ is good enough. Namely, whatever control function $\alpha(\cdot)$ we choose on the initial interval $[t_0,\ t_0 + \delta]$, even if during the remaining time $[t_0+\delta,\, T]$ our control is optimal, the total cost will still be considerably larger than $u(t_0, x_0)$. Indeed, by continuity, (5.9) implies

$$\varphi_t(t,x) + f(x,a)\cdot\nabla\varphi(t,x) > \theta - h(x,a) \qquad \text{for all } a\in A, \tag{5.20}$$

for all t, x close to t_0, x_0, i.e. such that (5.15) holds. Choose an arbitrary control function $\alpha : [t_0,\, t_0+\delta] \mapsto A$, and call $t\mapsto x(t) = x(t; t_0, x_0, \alpha)$ the corresponding trajectory. We now have

$$\begin{aligned} u\big(t_0+\delta,\ x(t_0+\delta)\big) - u(t_0,x_0) &\geq \varphi\big(t_0+\delta,\ x(t_0+\delta)\big) - \varphi(t_0,x_0) \\ &= \int_{t_0}^{t_0+\delta} \frac{d}{dt}\varphi\big(t,\ x(t)\big)\,dt \\ &= \int_{t_0}^{t_0+\delta} \varphi_t\big(t,\ x(t)\big) + f\big(x(t),a\big)\cdot\nabla\varphi\big(t,\ x(t)\big)\,dt \\ &\geq \int_{t_0}^{t_0+\delta} \theta - h\big(x(t),\alpha(t)\big)\,dt, \end{aligned} \tag{5.21}$$

according to (5.20). Therefore, for every control function $\alpha(\cdot)$ we have

$$u\big(t_0+\delta,\ x(t_0+\delta)\big) + \int_{t_0}^{t_0+\delta} h\big(x(t),\alpha(t)\big)\,dt \geq u(t_0,x_0) + \delta\theta. \tag{5.22}$$

Taking the infimum of the left hand side of (5.22) over all control functions α, we see that this infimum is still $\geq u(t_0,x_0)+\delta\theta$, in contradiction with (4.8). This establishes (P2), completing the proof.

Theorems 2 and 4 can be effectively combined, to obtain sufficient conditions for the optimality of a control function. The usual setting is the following.

Consider the problem of minimizing the cost functional (4.5). Assume that, for each initial condition (s,y), we have guessed a "candidate" optimal control $\alpha^{s,y} : [s,T]\mapsto A$. We then call

$$v(s,y) \doteq J(s,y,\alpha^{s,y}) \tag{5.23}$$

the corresponding cost. Typically, these control functions $\alpha^{s,y}$ are found by applying the Pontryagin Maximum Principle [F-R], which provides a necessary condition for optimality. On the other hand, consider the true value function u, defined at (4.8) as the infimum over all control functions $\alpha(\cdot)$. By Theorem 4, this function u provides a viscosity solution to the Hamilton-Jacobi-Bellman equation (5.7) with terminal condition $u(T,y) = g(y)$. We can now check if our function v at (5.23) also provides a viscosity solution to (5.7)-(5.8). If this is indeed the case, by the uniqueness of the viscosity solution we can conclude that $u\equiv v$, and hence all controls $\alpha^{s,y}$ are optimal.

References

[B-C] M. Bardi and I. Capuzzo Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, 1997.

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

[E] L. C. Evans, *Partial Differential Equations*, Amer. Math. Soc. Graduate Studies in Mathematics, Vol. 19, Providence 1998.

[F-R] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer, 1975.

[K] S. Kruzhkov, First order quasilinear equations with several space variables, *Math. USSR Sbornik* **10** (1970), 217-243.