

Maximum Principle, Dynamic Programming, and Their Connection in Deterministic Control^{1,2}

X. Y. ZHOU³

Communicated by L. D. Berkovitz

Abstract. Two major tools for studying optimally controlled systems are Pontryagin's maximum principle and Bellman's dynamic programming, which involve the adjoint function, the Hamiltonian function, and the value function. The relationships among these functions are investigated in this work, in the case of deterministic, finite-dimensional systems, by employing the notions of superdifferential and subdifferential introduced by Crandall and Lions. Our results are essentially non-smooth versions of the classical ones. The connection between the maximum principle and the Hamilton-Jacobi-Bellman equation (in the viscosity sense) is thereby explained by virtue of the above relationship.

Key Words. Optimal control, maximum principle, dynamic programming, viscosity solutions, superdifferential, subdifferential.

1. Introduction

Given $(s, y) \in [0, 1] \times R^d$, let us consider the following optimal control problem in R^d :

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

$$\text{subject to } dx(t)/dt = f(t, x(t), u(t)), \quad \text{a.e. } t \in [s, 1], \quad (1b)$$

$$x(s) = y, \quad (1c)$$

¹ This research was supported by the Natural Science Fund of China.

² This paper was written while the author visited Keio University, Japan. The author is indebted to Professors H. Tanaka and M. Nisio for their helpful suggestions and discussions. Thanks are also due to Professor X. J. Li for his comments and criticism.

³ PhD Student, Institute of Mathematics, Fudan University, Shanghai, People's Republic of China.

among all the admissible controls

$$u \in U_{\text{ad}}[s, 1] = \{u(\cdot) \mid u \text{ is a Lebesgue measurable function from } [s, 1] \text{ to } \Gamma\}, \quad (1d)$$

where Γ is a prescribed arbitrary set in R^m .

We denote the above problem by C_{sy} , to emphasize the importance of the initial time s and the initial state y . The value function is defined as

$$V(s, y) = \inf\{J(s, y; u) : u \in U_{\text{ad}}[s, 1]\}. \quad (2)$$

The pair (\hat{x}, \hat{u}) is called an optimal pair of the problem C_{sy} if \hat{x} is the corresponding solution of (1) for $\hat{u} \in U_{\text{ad}}[s, 1]$ and $J(s, y; \hat{u}) = V(s, y)$.

Since Pontryagin *et al.* (Ref. 1) proved the maximum principle (MP), and since Bellman (Ref. 2) put forward the method of dynamic programming (DP), most researches in optimal control theory have been carried out along one of these two lines.

The MP says that, if (\hat{x}, \hat{u}) is optimal for the problem C_{sy} , then there exists a function $\psi : [s, 1] \rightarrow R^d$, satisfying the adjoint equation

$$\begin{aligned} d\psi(t)/dt &= -f_x(t, \hat{x}(t), \hat{u}(t))\psi(t) - L_x(t, \hat{x}(t), \hat{u}(t)), \\ &\text{a.e. } t \in [s, 1], \end{aligned} \quad (3)$$

$$\psi(1) = h_x(\hat{x}(1)),$$

such that

$$\begin{aligned} H(t, \hat{x}(t), \hat{u}(t), \psi(t)) &= \max_{u \in \Gamma} H(t, \hat{x}(t), u, \psi(t)), \\ &\text{a.e. } t \in [s, 1], \end{aligned} \quad (4)$$

where the Hamiltonian is given by

$$\begin{aligned} H(t, x, u, p) &= -(p, f(t, x, u)) - L(t, x, u), \\ \forall(t, x, u, p) &\in [s, 1] \times R^d \times \Gamma \times R^d. \end{aligned} \quad (5)$$

On the other hand, DP asserts that, if the value function $V(\cdot, \cdot)$ happens to be continuously differentiable, then it satisfies the so-called Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned} -\partial V(t, x)/\partial t + \sup_{u \in \Gamma} H(t, x, u, \partial V(t, x)/\partial x) &= 0, \\ (t, x) &\in [0, 1] \times R^d, \end{aligned} \quad (6)$$

$$V(1, x) = h(x).$$

It seems, however, that the DP fails to have a solid ground from its very beginning, since it requires the value function to be smooth, which is not true even in the simplest cases; see Example 3.1 of this paper. As for the connection between MP and DP, the known result is (Ref. 3)

$$\psi(t) = \partial V(t, \hat{x}(t)) / \partial x, \quad \forall t \in [s, 1], \quad (7)$$

which also lacks generality, due to the same reason as above. This is an important gap in the literature. Recently, this gap was partially bridged by Crandall and Lions, who invented a new concept of solutions, called viscosity solutions, of nonlinear partial differential equations, including the HJB equation (Ref. 4). Lions successfully showed that the value function V turns out to be a viscosity solution of (6), under rather mild and reasonable assumptions (Refs. 5, 6). One of the advantages of a viscosity solution is that existence, uniqueness, and stability theorems are all valid; hence, the DP can be established in a rigorous framework.

This work tries to answer the following question in the case of deterministic, finite-dimensional systems: What is the relationship between the MP and the DP, without assuming the continuous differentiability of the value function $V(\cdot, \cdot)$? We investigate the relationship among the functions ψ , H , V by employing the notions of superdifferential and subdifferential, evoked in defining the viscosity solution. A similar claim to (7) is proved, which is now interpreted as a set inclusion among superdifferential, subdifferential, and singleton $\{\psi(t)\}$. Of course, our result reduces to the classical one (7), if $V(\cdot, \cdot)$ is sufficiently smooth. Furthermore, several interesting properties of the superdifferential and the subdifferential of $V(\cdot, \cdot)$ at $(t, \hat{x}(t))$ are derived. As a byproduct, the MP can be proved directly from the DP. Hence, MP, DP, and their connection can be treated within a unified framework of viscosity solutions.

2. Maximum Principle and Dynamic Programming

In this section, we give the statement of the MP and the DP; we also give some lemmas for later use. First of all, we make some assumptions on the data appearing in problem (1), which will always remain in force:

(A1) $f(\cdot, \cdot, \cdot)$ and $L(\cdot, \cdot, \cdot)$ 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 $u \in \Gamma$;

(A2) for each $(t, u) \in [0, 1] \times \Gamma$, the functions $f(t, \cdot, u)$, $L(t, \cdot, u)$, and $h(\cdot)$ are continuously differentiable, where $h(\cdot)$ is R^1 -valued;

(A3) there exists a constant $K > 0$, which is independent of (t, u) such that

$$|f(t, x, u) - f(t, y, u)| + |L(t, x, u) - L(t, y, u)| + |h(x) - h(y)| \leq K|x - y|, \quad \forall x, y \in R^d,$$

$$|f(t, x, u)| + |L(t, x, u)| + |h(x)| \leq K(1 + |x|), \quad \forall x \in R^d.$$

Theorem 2.1. (MP) Suppose that (\hat{x}, \hat{u}) is optimal for problem C_{sv} . Then, there exists a function ψ satisfying the adjoint equation (3), such that the maximum condition (4) holds.

The proof of the above theorem is well known. One may consult Ref. 1 for the earliest rigorous proof, or Ref. 7, for the special case of distributed-parameter systems, with the use of the vector-valued theorem.

Lemma 2.1. Suppose that (\hat{x}, \hat{u}) is optimal for problem C_{sv} . Let $\Phi(\cdot, \cdot)$ be the fundamental matrix of the following ODE:

$$dy(t)/dt = f_x^T(t, \hat{x}(t), \hat{u}(t))y(t). \quad (8)$$

Then, the following results hold:

- (i) $\sup_{s \leq r \leq t \leq 1} |\Phi(t, r)| \leq \text{const}$;
- (ii) the adjoint function ψ can be represented as

$$\psi(t) = \int_t^1 \Phi^T(r, t) L_x(r, \hat{x}(r), \hat{u}(r)) dr + \Phi^T(1, t) h_x(\hat{x}(1)). \quad (9)$$

Proof. (i) It is clear.

(ii) We can check that the function ψ given by (9) satisfies Eq. (3), using the known properties of the fundamental matrix. Hence, the result follows from the uniqueness of the solution of (3). \square

Now, let us recall briefly the definition of viscosity solution (Refs. 4, 5, 6).

Definition 2.1. Let Q be an open set in R^n , and let $v \in C(Q)$. The superdifferential [resp., subdifferential] of v at $x \in Q$, denoted by $D^+v(x)$ [resp., $D^-v(x)$], is a set defined as

$$D^+v(x) = \{p \in R^n \mid \overline{\lim}_{y \rightarrow x} [v(y) - v(x) - (p, y - x)] / |y - x| \leq 0\}$$

$$[\text{resp., } D^-v(x) = \{p \in R^n \mid \underline{\lim}_{y \rightarrow x} \{\dots\} \geq 0].$$

Definition 2.2. Let $H \in C(Q \times R \times R^n)$ and $v \in C(Q)$. The solution v is said to be a viscosity solution of the nonlinear PDE

$$H(x, v, Dv) = 0,$$

if

$$H(x, v(x), p) \leq 0, \quad \forall x \in Q, \forall p \in D^+ v(x),$$

$$H(x, v(x), p) \geq 0, \quad \forall x \in Q, \forall p \in D^- v(x).$$

From now on, we denote by $D_{ix}^+ v(\cdot, \cdot)$ the superdifferential in the (t, x) -variables and by $D_x^+ v(t, \cdot)$ the superdifferential in the x -variable for each fixed t , etc.

Theorem 2.2. (DP) The value function $V(\cdot, \cdot)$ is a viscosity solution of the HJB equation (6).

Remark 2.1. Our definition of the value function (2) follows that in Ref. 3, which is somewhat different from that in Lions (Refs. 5, 6). The difference lies in that our value function satisfies a backward HJB equation, while Lions' value function satisfies a forward HJB equation. But the proof of Theorem 2.2 is similar to that in Ref. 5.

3. Main Results

The adjoint function ψ , the Hamiltonian function H , and the value function V are the most important data appearing in the MP and DP. In this section, we show how they relate to each other within the framework of the superdifferential and the subdifferential.

Theorem 3.1. Suppose that (\hat{x}, \hat{u}) is optimal for problem C_{sv} . Then, for any $t \in [s, 1]$, $D_x^- V(t, \hat{x}(t)) \subset \{\psi(t)\} \subset D_x^+ V(t, \hat{x}(t))$.

Proof. For $z \in R^d$, let $x(\cdot; z)$ be the solution of (1) with control \hat{u} , initial time t , and initial state z . Then,

$$\begin{aligned} & x(r; z) - \hat{x}(r) \\ &= z - \hat{x}(t) + \int_t^r \{f(\theta, x(\theta; z), \hat{u}(\theta)) - f(\theta, \hat{x}(\theta), \hat{u}(\theta))\} d\theta \\ &= z - \hat{x}(t) + \int_t^r f_x^T(\theta, \hat{x}(\theta), \hat{u}(\theta))(x(\theta; z) - \hat{x}(\theta)) d\theta \\ &+ \int_t^r \epsilon(\theta; z)(x(\theta; z) - \hat{x}(\theta)) d\theta, \end{aligned} \tag{10}$$

where

$$\epsilon(\theta; z) \equiv \int_0^1 \{f_x^T(\theta, \hat{x}(\theta) + \alpha(x(\theta; z) - \hat{x}(\theta)), \hat{u}(\theta)) - f_x^T(\theta, \hat{x}(\theta), \hat{u}(\theta))\} d\alpha,$$

with the following properties:

$$\lim_{z \rightarrow \hat{x}(t)} \epsilon(\theta; z) = 0, \quad \forall \theta \in [t, 1], \quad (11a)$$

$$\sup_{\theta, z} |\epsilon(\theta; z)| \leq \text{const.} \quad (11b)$$

By (10) and the variation-of-constants formula, we get

$$\begin{aligned} x(r; z) - \hat{x}(r) &= \Phi(r, t)(z - \hat{x}(t)) + \int_t^r \Phi(r, \theta) \epsilon(\theta; z)(x(\theta; z) - \hat{x}(\theta)) d\theta, \end{aligned} \quad (12)$$

where Φ is defined in Lemma 2.1. Therefore,

$$\begin{aligned} &V(t, z) - V(t, \hat{x}(t)) \\ &\leq \int_t^1 \{L(r, x(r; z), \hat{u}(r)) - L(r, \hat{x}(r), \hat{u}(r))\} dr \\ &+ h(x(1; z)) - h(\hat{x}(1)) \\ &= \int_t^1 L_x^T(r, \hat{x}(r), \hat{u}(r))(x(r; z) - \hat{x}(r)) dr \\ &+ \int_t^1 \epsilon_0(r; z)(x(r; z) - \hat{x}(r)) dr \\ &+ h_x^T(\hat{x}(1))(x(1; z) - \hat{x}(1)) + o(|x(1; z) - \hat{x}(1)|), \end{aligned} \quad (13)$$

where $\epsilon_0(\cdot; z)$ is defined in a similar way to $\epsilon(\cdot; z)$, with f_x replaced by L_x . By virtue of (12), we can rewrite (13) as follows:

$$\begin{aligned} &V(t, z) - V(t, \hat{x}(t)) \\ &\leq \int_t^1 L_x^T(r, \hat{x}(r), \hat{u}(r)) \Phi(r, t) dr \cdot (z - \hat{x}(t)) \\ &+ h_x^T(\hat{x}(1)) \Phi(1, t)(z - \hat{x}(t)) + o(|z - \hat{x}(t)|) \\ &= \psi^T(t)(z - \hat{x}(t)) + o(|z - \hat{x}(t)|). \end{aligned}$$

The above inequality leads immediately to the conclusion that

$$\psi(t) \in D_x^+ V(t, \hat{x}(t)).$$

On the other hand, suppose that $p \in D_x^- V(t, \hat{x}(t))$. Then,

$$0 \leq \lim_{z \rightarrow \hat{x}(t)} \{V(t, z) - V(t, \hat{x}(t) - (p, z - \hat{x}(t)))\} / |z - \hat{x}(t)|$$

$$\leq \lim_{z \rightarrow \hat{x}(t)} (\psi(t) - p, z - \hat{x}(t)) / |z - \hat{x}(t)|.$$

Hence, $p = \psi(t)$. The proof is now complete. □

Remark 3.1. Theorem 3.1 is a nonsmooth version of (7). Indeed, if $V(t, \cdot)$ is differentiable, then the theorem reduces to (7). But in general, both the inclusions $D_x^- V(t, \hat{x}(t)) \subset \{\psi(t)\}$ and $\{\psi(t)\} \subset D_x^+ V(t, \hat{x}(t))$ may be strict, which will be shown in the following example.

Example 3.1. Consider the following optimal control problem C_{00} :

$$\begin{aligned} &\text{minimize} && -x(1), \\ &\text{subject to} && dx(t)/dt = x(t)u(t), \quad \text{a.e. } t \in [0, 1], \\ &&& x(0) = 0, \\ &&& u(\cdot): [0, 1] \rightarrow \{r \in R^1 \mid 0 \leq r \leq 1\}. \end{aligned}$$

The optimal pair of this problem is $(\hat{x}(\cdot), \hat{u}(\cdot)) = (0, 0)$. Moreover, one can easily compute that

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

Then,

$$D_x^- V(0, 0) = \phi, \quad D_x^+ V(0, 0) = [-e, -1], \quad \psi(0) = -1.$$

Theorem 3.1 is lacking, since it says nothing about the Hamiltonian. In what follows, we will show that the Hamiltonian is related to the superdifferential and the subdifferential of $V(\cdot, \cdot)$ at $(t, \hat{x}(t))$.

Let us explore some properties of $D_{ix}^+ V(t, \hat{x}(t))$ beforehand.

Proposition 3.1. Suppose that (\hat{x}, \hat{u}) is optimal for problem C_{xy} . Then, a.e. $t \in (s, 1)$, for any $(p, q) \in D_{ix}^\pm V(t, \hat{x}(t))$, we have

$$p = H(t, \hat{x}(t), \hat{u}(t), q) = \max_{u \in I} H(t, \hat{x}(t), u, q). \tag{14}$$

Proof. We note that, for a.e. $t \in (s, 1)$,

$$\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} \begin{bmatrix} f(\theta, \hat{x}(\theta), \hat{u}(\theta)) \\ L(\theta, \hat{x}(\theta), \hat{u}(\theta)) \end{bmatrix} d\theta = \begin{bmatrix} f(t, \hat{x}(t), \hat{u}(t)) \\ L(t, \hat{x}(t), \hat{u}(t)) \end{bmatrix}. \quad (15)$$

Fix $t \in (s, 1)$ such that (15) holds. If $(p, q) \in D_{ix}^- V(t, \hat{x}(t))$, then

$$\lim_{r \downarrow t} \frac{V(r, \hat{x}(r)) - V(t, \hat{x}(t)) - p(r-t) - (q, \hat{x}(r) - \hat{x}(t))}{|r-t| + |\hat{x}(r) - \hat{x}(t)|} \geq 0.$$

Hence,

$$\begin{aligned} 0 &\leq \lim_{r \downarrow t} \left[- \int_t^r L(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta - p(r-t) \right. \\ &\quad \left. - \int_t^r (q, f(\theta, \hat{x}(\theta), \hat{u}(\theta))) d\theta \right] / |r-t| \\ &= -L(t, \hat{x}(t), \hat{u}(t)) - p - (q, f(t, \hat{x}(t), \hat{u}(t))). \end{aligned}$$

Thus,

$$p \leq H(t, \hat{x}(t), \hat{u}(t), q).$$

Similarly, letting $r \uparrow t$, we may deduce that

$$p \geq H(t, \hat{x}(t), \hat{u}(t), q).$$

Therefore, the left-half part of (14) holds. Since $q \in D_x^- V(t, \hat{x}(t))$, hence $q = \psi(t)$ by Theorem 3.1. Now, the right-half part of (14) follows from the maximum principle.

On the other hand, suppose that $(p, q) \in D_{ix}^+ V(t, \hat{x}(t))$. Then, we can check the left-half part of (14) by an argument similar to the above. Furthermore, since $V(\cdot, \cdot)$ is a viscosity solution of the HJB equation (6),

$$-p + \sup_{u \in \Gamma} H(t, \hat{x}(t), \hat{u}(t), q) \leq 0,$$

which is precisely the right-half part of (14). The proof is completed. \square

Remark 3.2. We can see from Proposition 3.1 that, if we choose $q(t) \in (D_{ix}^+ V(t, \hat{x}(t)))_x$, where $(D_{ix}^+ V(t, \hat{x}(t)))_x$ means the x -section of the set $D_{ix}^+ V(t, \hat{x}(t))$, which is nonempty by the forthcoming Theorem 3.2, then $q(\cdot)$ satisfies the maximum condition almost everywhere. Together with Theorem 2.1 and Theorem 3.1, we know that $\psi(\cdot)$ is a continuous selection of the set-valued function $t \rightarrow (D_{ix}^+ V(t, \hat{x}(t)))_x$ that satisfies the MP.

Theorem 3.2. Suppose that (\hat{x}, \hat{u}) is optimal for problem C_{sv} . Then, for a.e. $t \in [s, 1]$,

$$D_{ix}^- V(t, \hat{x}(t)) \subset \{(H(t, \hat{x}(t), \hat{u}(t), \psi(t)), \psi(t))\} \subset D_{ix}^+ V(t, \hat{x}(t)).$$

Proof. The left inclusion of the theorem is a direct consequence of Proposition 3.1 and Theorem 3.1. Now, let us turn to the right one.

Fix $t \in (s, 1)$ such that (15) holds. For $(\tau, z) \in [s, 1] \times R^d$, with $\tau > t$, let $x(\cdot; \tau, z)$ be the solution of (1) with control \hat{u} , initial time τ , and initial state z . Then, for any $r \in [\tau, 1]$,

$$\begin{aligned} x(r; \tau, z) - \hat{x}(r) &= \left\{ z - \hat{x}(t) - \int_t^\tau f(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta \right\} \\ &\quad + \int_\tau^r f_x^T(\theta, \hat{x}(\theta), \hat{u}(\theta))(x(\theta; \tau, z) - \hat{x}(\theta)) d\theta \\ &\quad + \int_\tau^r \epsilon(\theta; \tau, z)(x(\theta; \tau, z) - \hat{x}(\theta)) d\theta, \end{aligned}$$

where $\epsilon(\cdot; \tau, z) \rightarrow 0$, as $\tau \rightarrow t$ and $z \rightarrow \hat{x}(t)$, and $\epsilon(\theta; \tau, z)$ is uniformly bounded. By the variation-of-constants formula, we can write

$$\begin{aligned} x(r; \tau, z) - \hat{x}(r) &= \Phi(r, \tau) \left\{ z - \hat{x}(t) - \int_t^\tau f(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta \right\} \\ &\quad + \int_\tau^r \Phi(r, \theta) \epsilon(\theta; \tau, z)(x(\theta; \tau, z) - \hat{x}(\theta)) d\theta. \end{aligned}$$

By a similar computation as in the proof of Theorem 3.1, we have

$$\begin{aligned} &V(\tau, z) - V(t, \hat{x}(t)) \\ &\leq \left(\psi(\tau), z - \hat{x}(t) - \int_t^\tau f(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta \right) \\ &\quad - \int_t^\tau L(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta + o(|\tau - t| + |z - \hat{x}(t)|) \\ &= \left(\psi(t) + \dot{\psi}(t)(\tau - t) + o(|\tau - t|), z - \hat{x}(t) - \int_t^\tau f(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta \right) \\ &\quad - \int_t^\tau L(\theta, \hat{x}(\theta), \hat{u}(\theta)) d\theta + o(|\tau - t| + |z - \hat{x}(t)|) \\ &= (\psi(t), z - \hat{x}(t)) + (\tau - t)H(t, \hat{x}(t), \hat{u}(t), \psi(t)) \\ &\quad + o(|\tau - t| + |z - \hat{x}(t)|). \tag{16} \end{aligned}$$

Moreover, when $\tau < t$, a similar calculation also leads to (16). Hence, the desired result follows. \square

Remark 3.3. Let us see what the smooth version of Theorem 3.2. In fact, if $V(\cdot, \cdot)$ is differentiable, then Theorem 3.1 yields

$$\begin{aligned}\partial V(t, \hat{x}(t))/\partial t &= H(t, \hat{x}(t), \hat{u}(t), \psi(t)) \\ &= H(t, \hat{x}(t), \hat{u}(t), \partial V(t, \hat{x}(t))/\partial x),\end{aligned}$$

which is well known in classical dynamic programming theory (Refs. 3, 8)

Remark 3.4. In Ref. 9, Barron and Jensen proved the MP via the DP. Here, by virtue of Theorem 3.2, we can present another simple proof of the MP by using the DP. In fact, since $(H(t, \hat{x}(t), \hat{u}(t), \psi(t)), \psi(t)) \in D_{\alpha}^+ V(t, \hat{x}(t))$, by the DP, we have

$$-H(t, \hat{x}(t), \hat{u}(t); \psi(t)) + \sup_{u \in \Gamma} H(t, \hat{x}(t), u, \psi(t)) \leq 0,$$

which yields the MP.

4. Conclusions

The idea of viscosity solution, including the superdifferential and the subdifferential, can be regarded as a type of nonsmooth analysis. Dynamic programming and its relationship to the maximum principle is interpreted in the language of viscosity solutions, while the illusory assumption that $V(\cdot, \cdot)$ is smooth is removed.

Another framework of nonsmooth analysis, called generalized gradient, was proposed by Clarke; see Ref. 10 for a thorough treatment of the generalized gradient. The MP and the DP have been stated within this framework (Refs. 11, 12); more recently, Clarke and Vinter related the MP and DP by

$$\psi(t) \in \partial_x V(t, \hat{x}(t)), \tag{17}$$

where $\partial_x V(t, \cdot)$ denotes the generalized gradient of $V(t, \cdot)$; see Ref. 8.

Obviously, Theorem 3.1 in this work is an analogous result to (17), which suggests that the MP, the DP, and their relationship can be treated uniformly within the framework of viscosity solution.

References

1. PONTRYAGIN, M. L., BOLTYANSKII, V. G., GAMKRELIDZE, R. V., and MISCHENKO, E. F., *The Mathematics Theory of Optimal Processes*, Interscience, New York, New York, 1962.

2. BELLMAN, R., *Dynamic Programming*, Princeton University Press, Princeton, New Jersey, 1957.
3. FLEMING, W. H., and RISHEL, R. W., *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, New York, 1975.
4. CRANDALL, M. G., and LIONS, P. L., *Viscosity Solutions of Hamilton-Jacobi Equations*, Transactions of the American Mathematical Society, Vol. 277, pp. 1-42, 1983.
5. LIONS, P. L., *Generalized Solutions of Hamilton-Jacobi Equations*, Pitman, London, England, 1982.
6. LIONS, P. L., *Optimal Control and Viscosity Solutions*, Proceedings of the Conference on Recent Mathematical Methods in Dynamic Programming, Rome, Italy, 1984; Edited by I. Capuzzo-Dolcetta, W. H. Fleming, and T. Zolezzi, Springer-Verlag, New York, New York, pp. 94-112, 1985.
7. LI, X. J., and YAO, Y. L., *Maximum Principle of Distributed-Parameter Systems with Time Lags*, Proceedings of the Conference on Control Theory of Distributed-Parameter Systems and Applications, Vorau, Austria, 1984; Edited by F. Kappel, K. Kunisch, and W. Schappacher, Springer-Verlag, New York, New York, pp. 410-427, 1985.
8. CLARKE, F. H., and VINTER, R. B., *The Relationship between the Maximum Principle and Dynamic Programming*, SIAM Journal on Control and Optimization, Vol. 25, pp. 1291-1311, 1987.
9. BARRON, E. N., and JENSEN, R., *The Pontryagin Maximum Principle from Dynamic Programming and Viscosity Solutions to First-Order Partial Differential Equations*, Transactions of the American Mathematical Society, Vol. 298, pp. 635-641, 1986.
10. CLARKE, F. H., *Optimization and Nonsmooth Analysis*, John Wiley, New York, New York, 1983.
11. CLARKE, F. H., *The Maximum Principle under Minimal Hypotheses*, SIAM Journal on Control and Optimization, Vol. 14, pp. 1078-1091, 1976.
12. CLARKE, F. H., and VINTER, R. B., *Local Optimality Conditions and Lipschitzian Solutions to the Hamilton-Jacobi Equations*, SIAM Journal on Control and Optimization, Vol. 21, pp. 856-870, 1983.