Value Functions and Transversality Conditions for Infinite-Horizon Optimal Control Problems

佐柄, 信純 / SAGARA, Nobusumi

(出版者 / Publisher)
法政大学比較経済研究所 / Institute of Comparative Economic Studies, Hosei University

(雑誌名 / Journal or Publication Title)
比較経済研究所ワーキングペーパー

(巻 / Volume)
143

(開始ページ / Start Page)
1

(終了ページ / End Page)
35

(発行年 / Year)
2008-12-18

Value Functions and Transversality Conditions for Infinite-Horizon Optimal Control Problems*

Nobusumi Sagara

Faculty of Economics, Hosei University
4342, Aihara, Machida, Tokyo, 194–0298, Japan
e-mail: nsagara@hosei.ac.jp

December 8, 2008

Abstract

This paper investigates the relationship between the maximum principle with an infinite horizon and dynamic programming and sheds new light upon the role of the transversality condition at infinity as necessary and sufficient conditions for optimality with or without convexity assumptions. We first derive the nonsmooth maximum principle and the adjoint inclusion for the value function as necessary conditions for optimality that exhibit the relationship between the maximum principle and dynamic programming. We then present sufficiency theorems that are consistent with the strengthened maximum principle, employing the adjoint inequalities for the Hamiltonian and the value function. Synthesizing these results, necessary and sufficient conditions for optimality are provided for the convex case. In particular, the role of the transversality conditions at infinity is clarified.

Key Words: Nonsmooth maximum principle; Infinite horizon; Value function; Transversality condition; Adjoint inclusion; Necessary and sufficient conditions.
MSC2000: 49K24, 49L20.

*This research is supported by a Grant-in-Aid for Scientific Research (No. 18610003) from the Ministry of Education, Culture, Sports, Science and Technology.

1 Introduction

The maximum principle in optimal control is a fundamental instrument in dynamic optimization theory. It is usually formulated in a finite horizon, but one often needs to treat the case for an infinite horizon, especially in economic growth theory. While the maximum principle with an infinite horizon was treated in a simple manner by Pontryagin et al. [28, Section 24], it was Shell [33] (later Halkin [23]) who first pointed out, by way of counterexample, that the transversality condition with a finite horizon cannot be extended in an intuitive way to that with an infinite horizon as a part of necessary conditions for optimality. Since then, the maximum principle with an infinite horizon has been elaborated by, for instance, Aseev and Kryaziimskiy [3], Aubin and Clarke [4], Feinstein and Luenberger [20], Michel [26], Seierstadt and Sydsæter [32] and Ye [38] with primal attention to the transversality condition at infinity.

On the other hand, solutions to optimal control problems can be characterized by dynamic programming, which is based on the value function as a solution to the Hamilton–Jacobi–Bellman (HJB) equation. Under some regularity conditions, the value function is a smooth solution to the HJB equation. It is well-known, however, that the regularity conditions are violated in many cases of interest and the value function fails to be continuously differentiable even if the underlying data are smooth. Indeed, one may expect the value function to be, at best, Lipschitz continuous, even in the smooth data case. (For the differentiability of the value function, see Cannarsa and Frankowska [13].)

To overcome this difficulty, there exist two lines of research. One is "nonsmooth analysis" initiated by Clarke [15, 16], which employs generalized gradients of the value function and generalized solutions to the extended HJB equation, and the linkage between the maximum principle and dynamic programming has been established by Clarke and Vinter [17] and Vinter [36]. The other, a somewhat later development, is the concept of "viscosity solutions" to the HJB equation, which makes use of the notion of super- and subdifferentials, proposed by Crandall and Lions [18] and Crandall, Evans and Lions [19]. The value function is shown to be a unique viscosity solution of the HJB equation and the connection between the adjoint equation for the Hamiltonian and that for the value function has been investigated by Barron and Jensen [7], Cannarsa and Frankowska [13], Frankowska [21], Mirică [27] and Zhou [41]. For the relation between viscosity solutions to the HJB equation and generalized solutions to the extended HJB equation, see Frankowska [22] and Zhou [42].

The purpose of this paper is to investigate the relationship between the maximum principle with an infinite horizon and dynamic programming and shed new light upon the role of the transversality condition at infinity as necessary and sufficient conditions for optimality with or without convexity assumptions.

In this paper, we mitigate the smoothness assumptions by introducing the technique of nonsmooth analysis along the line of Clarke [15, 16]. We first derive the nonsmooth maximum principle and the adjoint inclusion for the value function as necessary conditions for optimality that exhibit the relationship between the maximum principle and dynamic programming. The necessary conditions under consideration are direct extensions of those of Clarke and Vinter [17] and Vinter [36] to an infinite horizon setting. The nonsmooth maximum principle with an infinite horizon demonstrated by Ye [38] is generalized by taking into account unbounded controls and nonautonomous systems.

We then present sufficient conditions for optimality under nonsmooth nonconvex hypotheses. Two sufficiency theorems are provided. The first is an extension of the finite horizon result by Zeidan [39, 40] to the infinite horizon setting, which is stated in terms of the adjoint inequality for the Hamiltonian that is consistent with the strengthened maximum principle. The second, which exploits the adjoint inequality for the value function, is novel in the literature in that the sufficient condition is related to the adjoint inclusion of the value function as well as the adjoint inequality for the Hamiltonian.

Synthesizing these results, it is possible to characterize optimal solutions and provide necessary and sufficient conditions for optimality if one restricts attention to the convex case. In particular, the role of the transversality conditions at infinity is clarified. This characterization is analogous to the result for the finite horizon case by Rockafeller [29], who systematically developed dual problems of optimal control under convexity hypotheses. To this end, the convexity of the value function and the concavity of the Hamiltonian are established.

2 Preliminary

This section collects some preliminary results on generalized gradients for locally Lipschitz functions. When the function under investigation is a convex function, the results are reduced to the traditional subdifferential calculus. A basic reference for the results treated in this section is Clarke [15].

Denote by $\langle x, y\rangle$ the inner product of the points $x, y \in \mathbb{R}^n$. The norm of x is given by $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *Lipschitz of rank*

$K > 0$ near a given point $x \in \mathbb{R}^n$ if there exists some $\varepsilon > 0$ such that:

$$|f(y) - f(z)| \leq K\|y - z\| \quad \text{for every } y, z \in x + \varepsilon B.$$

Here, B is the open unit ball in $\mathbb{R}^n$. A function f is said to be *locally Lipschitz* on $X \subset \mathbb{R}^n$ if f is Lipschitz near x for every $x \in X$.

Let f be Lipschitz near $x \in \mathbb{R}^n$. The *generalized directional derivative* of f at x in the direction $v \in \mathbb{R}^n$, denoted by $f^\circ(x; v)$, is defined as follows:

$$f^\circ(x; v) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{f(y + \lambda v) - f(y)}{\lambda}.$$

The *generalized gradient* of f at x, denoted by $\partial f(x)$, is defined by:

$$\partial f(x) = \{\zeta \in \mathbb{R}^n \mid \langle \zeta, v \rangle \leq f^\circ(x; v) \ \forall v \in \mathbb{R}^n\}.$$

Note that $\partial f(\cdot)$ induces a set-valued mapping from $\mathbb{R}^n$ into itself and we denote it by $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

The set of points at which a given function f fails to be differentiable is denoted by Ω_f. Radenmacher's theorem states that a Lipschitz function on an open subset of $\mathbb{R}^n$ is differentiable almost everywhere on that subset. Thus, if f is Lipschitz near x, then its generalized gradient is given by:

$$\partial f(x) = \mathrm{co}\left\{ \lim_{\nu \to \infty} \nabla f(x^\nu) \mid x^\nu \to x,\ x^\nu \notin N \cup \Omega_f,\ \nu = 1, 2, \ldots \right\},$$

where $\nabla f(x^\nu)$ is the gradient of f at x^ν, N is any set of Lebesgue measure 0 in $\mathbb{R}^n$ and the convex hull is taken over all limit points $\nabla f(x^\nu)$ for which $\{x^\nu\}$ is any sequence converging to x while avoiding the set $N \cup \Omega_f$ and such that $\nabla f(x^\nu)$ converges.

Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued function, written in terms of component functions as $F(x) = (f_1(x), \ldots, f_m(x))$ such that each f_i (and hence F) is Lipschitz near a given point $x \in \mathbb{R}^n$. Denote by $JF(y)$ the $m \times n$-Jacobian matrix of partial derivatives whenever $y \in \mathbb{R}^n$ is a point at which the partial derivatives exist and by Ω_F the complement of the set of all such points. The *generalized Jacobian* of F at x, denoted by $\partial F(x)$, is defined by:

$$\partial F(x) = \mathrm{co}\left\{ \lim_{\nu \to \infty} JF(x^\nu) \mid x^\nu \to x,\ x^\nu \notin \Omega_F,\ \nu = 1, 2, \ldots \right\}.$$

The meaning of the convex hull is similar as above. It follows that:

$$\partial F(x) \subset \partial f_1(x) \times \cdots \times \partial f_m(x),$$

where the right-hand side of the inclusion denotes the set of all matrices whose i th row belongs to $\partial f_i(x)$ for each i.

The half-open interval $[0,\infty)$ of the real line is equipped with the σ-algebra $\mathscr{L}$ of Lebesgue measurable subsets of $[0,\infty)$. Denote the product of the σ-algebra of $\mathscr{L}$ and the σ-algebra $\mathscr{B}^n \times \mathscr{B}^m$ of Borel subsets of the product space $\mathbb{R}^n \times \mathbb{R}^m$ by $\mathscr{L} \times \mathscr{B}^n \times \mathscr{B}^m$.

The t-section of a subset Ω of $[0,\infty) \times \mathbb{R}^n$ is denoted by $\Omega(t)$, that is, $\Omega(t) = \{x \in \mathbb{R}^n \mid (t,x) \in \Omega\}$ for $t \in [0,\infty)$.

For later use, we present the following result.

Theorem 2.1. (i) *Let Ω be an $\mathscr{L} \times \mathscr{B}^n$-measurable subset of $[0,\infty) \times \mathbb{R}^n$. If $f : \Omega \to \mathbb{R}$ is an $\mathscr{L} \times \mathscr{B}^n$-measurable function such that $f(t,\cdot)$ is locally Lipschitz on $\Omega(t)$ for every $t \in [0,\infty)$, then $\partial_x f : \Omega \rightrightarrows \mathbb{R}^n$ is $\mathscr{L} \times \mathscr{B}^n$-measurable.*

(ii) *Let $x_0 \in \mathbb{R}^n$ and $\varepsilon > 0$ be given. If $f : (x_0 + \varepsilon B) \times \mathbb{R}^m \to \mathbb{R}$ is upper semicontinuous and $f(\,\cdot\,, y)$ is Lipschitz on $x_0 + \varepsilon B$ for every $y \in \mathbb{R}^m$, then $\partial_x f : (x_0 + \varepsilon B) \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is upper semicontinuous.*

Proof. (i) Since f is $\mathscr{L} \times \mathscr{B}^n$-measurable, it follows from the definition of the generalized directional derivative that $f_x^\circ(t,x;v)$ can be obtained as the pointwise limit of the supremum of a countable family of continuous functions, and, hence, f_x° is an $\mathscr{L} \times \mathscr{B}^n \times \mathscr{B}^n$-measurable function on $\Omega \times \mathbb{R}^n$. Note that $\partial_x f$ is *scalarly measurable*; namely, $\partial_x f$ is compact convex-valued, $f_x^\circ(t,x;\cdot\,)$ is the support function of $\partial_x f(t,x)$ in the sense that:

$$f_x^\circ(t,x;v) = \max\{\langle \zeta, v\rangle \mid \zeta \in \partial_x f(t,x)\},$$

(see Clarke [15, Proposition 2.1.2]) and f_x° is a sublinear Carathéodory function, that is, $f_x^\circ(\,\cdot\,,\cdot\,;v)$ is $\mathscr{L} \times \mathscr{B}^n$-measurable for every $v \in \mathbb{R}^n$ and $f_x^\circ(t,x;\cdot\,)$ is continuous and sublinear for every $(t,x) \in \Omega$ (see Clarke [15, Proposition 2.1.1]). Then, the measurability of $\partial_x f$ follows from its scalar measurability (see Aliprantis and Border [1, Theorem 18.32]).

(ii) Choose any $0 < \varepsilon' < \varepsilon$ and put $\eta = \varepsilon - \varepsilon'$. Then, $\{x + \eta B \mid x \in x_0 + \varepsilon' B\}$ is an open covering of $x_0 + \varepsilon' B$ whose union is $x_0 + \varepsilon B$. Let S be the set of unit vectors in $\mathbb{R}^n$. We claim that the function $g_\eta : (x_0 + (\varepsilon - \eta)B) \times \mathbb{R}^m \times S \to \mathbb{R}$ defined by

$$g_\eta(x,y,v) = \sup_{\substack{x' \in x + \eta B \\ 0 < \lambda < \eta}} \frac{f(x' + \lambda v, y) - f(x', y)}{\lambda}$$

is upper semicontinuous. To this end, note first that the supremum is finite because of the Lipschitz continuity of $f(\,\cdot\,, y)$. Define the set-valued mapping

$\Gamma_\eta : (x_0 + (\varepsilon - \eta)B) \times \mathbb{R}^m \times S \rightrightarrows \mathbb{R}^n \times \mathbb{R}$ by $\Gamma_\eta(x, y, v) = (x + \eta B) \times (0, \eta)$. Then, g_η is the marginal function of the upper semicontinuous function:

$$(x', y, v, \lambda) \mapsto \frac{f(x' + \lambda v, y) - f(x', y)}{\lambda},$$

on $(x_0 + (\varepsilon - \eta)B) \times \mathbb{R}^m \times S \times (0, \eta)$ maximized over $(x', \lambda) \in \Gamma_\eta(x, y, v)$. Note that Γ_η is upper semicontinuous and relatively compact-valued. Then, applying the maximum theorem and noticing that its proof is valid even if Γ_η is not compact-valued but relatively compact-valued, g_η is upper semicontinuous (see Aubin and Frankowska [5, Theorem 1.4.16]). Since this is true for every $0 < \eta < \varepsilon$ and g_η is nondecreasing in η, the pointwise infimum $\inf_{0<\eta<\varepsilon} g_\eta(x, y, v)$ of $\{g_\eta\}$ is upper semicontinuous on $x_0 + \varepsilon B$, but the infimum coincides with $f_x^\circ(x, y; v)$ by definition. Therefore, f_x° is upper semicontinuous on $(x_0 + \varepsilon B) \times \mathbb{R}^m \times S$. From the fact that the function $v \mapsto f_x^\circ(x, y; v)$ is finite and positively homogeneous (see Clarke [15, Proposition 2.1.1]), it follows that f_x° is upper semicontinuous on $(x_0 + \varepsilon B) \times \mathbb{R}^m \times \mathbb{R}^n$. Henceforth, the upper semicontinuity of $f_x^\circ(\,\cdot\,, \cdot\,; v)$ is equivalent to that of the set-valued mapping $\partial_x f$ because $\partial_x f$ is compact convex-valued and $f_x^\circ(x, y; \cdot\,)$ is the support function of $\partial_x f(x, y)$ (see Aliprantis and Border [1, Theorem 17.41]). □

3 Necessary Condition for Optimality

We are given $\mathscr{L} \times \mathscr{B}^n \times \mathscr{B}^m$-measurable functions $L : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $f : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, an $\mathscr{L} \times \mathscr{B}^n$-measurable subset Ω of $[0, \infty) \times \mathbb{R}^n$ and a set-valued mapping $U : [0, \infty) \rightrightarrows \mathbb{R}^m$ with the $\mathscr{L} \times \mathscr{B}^m$-measurable graph. An *$\varepsilon$-tube* about the continuous function $x : [0, \infty) \to \mathbb{R}^n$ is a set of the form:

$$T(x(\cdot); \varepsilon) = \{(t, x) \in [0, \infty) \times \mathbb{R}^n \mid x \in x(t) + \varepsilon B\},$$

with $\varepsilon > 0$.

The optimal control problem under investigation is the following:

$$\begin{aligned} \min\ & J(x(\cdot), u(\cdot)) := \int_0^\infty L(t, x(t), u(t))dt \\ \text{s.t. } & \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [0, \infty), \\ & x(0) = x_0, \\ & x(t) \in \Omega(t) \quad \text{for every } t \in [0, \infty), \\ & u(t) \in U(t) \quad \text{a.e. } t \in [0, \infty). \end{aligned} \tag{P}$$

Here, the minimization is taken over all locally absolutely continuous functions (arcs) $x : [0, \infty) \to \mathbb{R}^n$ and $\mathscr{L}$-measurable functions $u : [0, \infty) \to \mathbb{R}^m$ satisfying the control system for the problem (P).

Because the objective integral functional with an infinite horizon admits its values to be infinite, there are several criteria for optimality (see, for example, Feinstein and Luenberger [20], Halkin [23], Kamihigashi [24], Seierstadt and Sydsæter [32], Takekuma [34]). For simplicity, we restrict ourselves to the class of pairs $(x(\cdot), u(\cdot))$ of functions for which the improper integral converges, as in Aseev and Kryaziimskiy [3], Aubin and Clarke [4], Michel [26], Pontryagin et al. [28] and Ye [38].

A *process* on a given subinterval I of $[0, \infty)$ is a pair $(x(\cdot), u(\cdot))$ of functions on I of which $x : I \to \mathbb{R}^n$ is a locally absolutely continuous function and $u : I \to \mathbb{R}^m$ is a measurable function such that the control system for (P) with I in place of $[0, \infty)$ and the initial condition $x(t) = x_0$, where t is the left endpoint of I, is satisfied. A process $(x(\cdot), u(\cdot))$ on I is *admissible* if the integrand $L(\,\cdot\,, x(\cdot), u(\cdot))$ is integrable on I. A process on I is *minimizing* if it minimizes the value of the integral functional $\int_I L dt$ over all admissible processes on I. When $I = [0, \infty)$, we shall abbreviate the domain on which processes are defined. In this section, $(x_0(\cdot), u_0(\cdot))$ is taken to be a fixed minimizing process on $[0, \infty)$ for (P).

We define the value function $V : \Omega \to \mathbb{R} \cup \{\pm\infty\}$ by:

$$V(t, x) = \inf \left\{ \int_t^\infty L(s, x(s), u(s)) ds \right\},$$

where the infimum is taken over all admissible processes $(x(\cdot), u(\cdot))$ on $[t, \infty)$ for which $x(t) = x \in \Omega(t)$. When no such admissible processes exist, the value is supposed to be $+\infty$, as usual.

3.1 Maximum Principle with an Infinite Horizon

The basic hypotheses to derive necessary conditions for optimality are as follows.

Hypothesis 3.1. (i) $L(\,\cdot\,, x, \cdot\,)$ is measurable for every $x \in \mathbb{R}^n$ and $L(t, \cdot\,, u)$ is Lipschitz of rank $k_L(t)$ on $\Omega(t)$ for every $(t, u) \in \operatorname{graph}(U)$ with k_L an integrable function.

(ii) There exists an integrable function φ on $[0, \infty)$ such that $|L(t, x_0(t), u)| \leq \varphi(t)$ for every $(t, u) \in \operatorname{graph}(U)$.

(iii) $f(\,\cdot\,, x, \cdot\,)$ is measurable for every $x \in \mathbb{R}^n$ and $f(t, \cdot\,, u)$ is Lipschitz of rank $k_f(t)$ on $\Omega(t)$ for every $(t, u) \in \operatorname{graph}(U)$ with k_f a locally integrable function.

(iv) The function k on $[0, \infty)$ given by $k(t) := k_L(t) \exp(\int_0^t k_f(s)ds)$ is integrable.

(v) There exists an ε-tube about $x_0(\cdot)$ contained in Ω such that $V(t, \cdot)$ is Lipschitz of rank K on $x_0(t) + \varepsilon B$ for every $t \in [0, \infty)$.

The Lipschitz continuity of the value function in the condition (v) of the hypothesis is nonstringent because, as seen in Appendix A, the condition is implied from the hypothesis guaranteeing the existence of minimizing processes for every initial condition. In particular, when $\Omega = [0, \infty) \times \mathbb{R}^n$, it is redundant because it is obtained from other conditions (i) to (iv) of the hypothesis.

The Pontryagin (or pseudo) Hamiltonian H_P and the (true) Hamiltonian H for (P) are given respectively by:

$$H_P(t, x, u, p) = \langle p, f(t, x, u) \rangle - L(t, x, u),$$

and

$$H(t, x, p) = \sup_{u \in U(t)} \{\langle p, f(t, x, u) \rangle - L(t, x, u)\}.$$

Theorem 3.1. *Suppose that Hypothesis 3.1 is satisfied. Then, there exists a locally absolutely continuous function $p : [0, \infty) \to \mathbb{R}^n$ with the following properties.*

(i) $-\dot{p}(t) \in \partial_x H_P(t, x_0(t), u_0(t), p(t))$ *a.e.* $t \in [0, \infty)$.

(ii) $H_P(t, x_0(t), u_0(t), p(t)) = H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$.

(iii) $-p(t) \in \partial_x V(t, x_0(t))$ *a.e.* $t \in [0, \infty)$.

(iv) $-p(0) \in \partial_x V(0, x_0(0))$.

It should be noted that the adjoint inclusion in the maximum principle is sometimes relevant to the (true) Hamiltonian. (See Corollary 4.1 and Theorems 5.1 and 5.2 in the sequel or Clarke [16].) The following result is a weaker form of Theorem 3.1.

Corollary 3.1. *The condition (i) of Theorem 3.1 implies that:*

$$-\dot{p}(t) \in \partial_x H(t, x_0(t), p(t)) \quad a.e.\ t \in [0, \infty).$$

Proof. It suffices to show that:

$$\partial_x H_P(t, x_0(t), u_0(t), p(t)) \subset \partial_x H(t, x_0(t), p(t)) \quad \text{a.e. } t \in [0, \infty).$$

Let $t \in [0, \infty)$ be any point at which the condition (ii) of Theorem 3.1 is true. Define the Lipschitz functions on $\Omega(t)$ by $h_P(x) = H_P(t, x, u_0(t), p(t))$ and $h(x) = H(t, x, p(t))$. Then, $h_P(x) \le h(x)$ for every $x \in \mathbb{R}^n$ with equality at $x = x_0(t)$. It follows from the definition of the generalized directional derivative that:

$$\begin{aligned} h_P^\circ(x_0(t); v) &= \limsup_{\substack{x \to x_0(t) \\ \lambda \downarrow 0}} \frac{h_P(x + \lambda v) - h_P(x)}{\lambda} \\ &\le \limsup_{\substack{x \to x_0(t) \\ \lambda \downarrow 0}} \left[\frac{h(x + \lambda v) - h(x)}{\lambda} + \frac{h(x) - h_P(x)}{\lambda} \right] = h^\circ(x_0(t); v) \end{aligned}$$

for every $v \in \mathbb{R}^n$. Since the generalized directional derivatives are support functions of the generalized gradients (see Clarke [15, Proposition 2.1.2]), the above inequality is equivalent to the inclusion $\partial h_P(x_0(t)) \subset \partial h(x_0(t))$. (See Clarke [15, Proposition 2.1.4].) □

Theorem 3.1 does not exclude the possibility that $-p(t) \notin \partial_x V(t, x_0(t))$ for every t in the null set of $[0, \infty)$. The question naturally arises whether this null set can be eliminated in special circumstances. The proof of the following result is the same as that of Clarke and Vinter [17].

Corollary 3.2. *The condition (iii) of Theorem 3.1 can be strengthened to:*

$$-p(t) \in \partial_x V(t, x_0(t)) \quad \textit{for every } t \in [0, \infty),$$

if (i) $\partial_x V(\,\cdot\,, x_0(\cdot)) : [0, \infty) \rightrightarrows \mathbb{R}^n$ is upper semicontinuous; or (ii) $\Omega(t)$ is convex for every $t \in [0, \infty)$ and $V(t, \cdot\,)$ is a convex function on $\Omega(t)$ for every $t \in [0, \infty)$.

3.2 Auxiliary Result

Theorem 3.1 can be proven by extending the necessary condition for the finite horizon case provided by Clarke and Vinter [17] to the infinite horizon case. To this end, we introduce a perturbed infinite-horizon optimal control problem with free left endpoints and deduce the maximum principle for it. The adjoint variable of the finite horizon problem restricted to the arbitrarily fixed finite interval $[0, T]$ is extended to $[0, \infty)$ as $T \to \infty$ by making use of the diagonalization method based on the equicontinuity of the relevant sequence of adjoint variables.

3.2.1 Perturbed Problem

Fix $\varepsilon > 0$ such that the ε-tube about $x_0(\cdot)$ is contained in Ω given in Hypothesis 3.1(v). A triplet $(x(\cdot), u(\cdot), v(\cdot))$ of functions on $[0, \infty)$ is called a *perturbed process* if it satisfies the perturbed control system:

$$\begin{aligned}
&\dot{x}(t) = f(t, x(t), u(t)) + v(t) \quad \text{a.e. } t \in [0, \infty), \\
&x(t) \in x_0(t) + \varepsilon B \quad \text{for every } t \in [0, \infty), \\
&u(t) \in U(t) \quad \text{a.e. } t \in [0, \infty), \\
&v(t) \in B \quad \text{a.e. } t \in [0, \infty).
\end{aligned}$$

Here, an $\mathscr{L}$-measurable function $v : [0, \infty) \to \mathbb{R}^n$ is viewed as a new control function.

Define the function $\sigma_\varepsilon : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ by:

$$\sigma_\varepsilon(t, v) = \max\{\langle p, v \rangle \mid p \in \partial_x V(t, x_0(t) + \varepsilon \bar{B})\}.$$

Here, $\bar{B}$ is the closure of B. Since $\partial_x V(t, \cdot\,)$ is compact-valued and upper semicontinuous (see Clarke [15, Proposition 2.1.1]), $\partial_x V(t, x_0(t) + \varepsilon \bar{B})$ is compact for every $t \in [0, \infty)$. Therefore, the maximum in the above is indeed attained.

Lemma 3.1. (i) *σ_ε is $\mathscr{L} \times \mathscr{B}^n$-measurable and $\sigma_\varepsilon(t, \cdot\,)$ is continuous for every $t \in [0, \infty)$;*

(ii) *$\sigma_\varepsilon(\,\cdot\,, v(\cdot))$ is locally integrable on $[0, \infty)$ if $v(\cdot)$ is locally integrable on $[0, \infty)$.*

Proof. (i) Since V is continuous on $T(x_0(\cdot); \varepsilon)$ by Theorem A.1 and $V(t, \cdot\,)$ is Lipschitz on $x_0(t) + \varepsilon B$ for every $t \in [0, \infty)$ by Hypothesis 3.1(v), we can apply Theorem 2.1(i) and noting that the t-section of $T(x_0(\cdot); \varepsilon)$ is $x_0(t) + \varepsilon B$. Thus, $\partial_x V$ is $\mathscr{L} \times \mathscr{B}^n$-measurable on $T(x_0(\cdot); \varepsilon)$, and, hence, the set-valued mapping Σ defined by $\Sigma(t) = \partial_x V(t, x_0(t) + \varepsilon \bar{B})$ is $\mathscr{L}$-measurable. Since $\sigma_\varepsilon(t, \cdot\,)$ is the support function of $\Sigma(t)$, the measurability of σ_ε follows from that of Σ. The continuity of $\sigma_\varepsilon(t, \,\cdot\,)$ follows from the fact that $\Sigma(t)$ is compact in $\mathbb{R}^n$ for every $t \in [0, \infty)$.

(ii) Since $V(t, \cdot\,)$ is Lipschitz of rank K on $x_0(t) + \varepsilon B$ for every $t \in [0, \infty)$, we have $\max\{\|p\| \mid p \in \Sigma(t)\} \leq K$ (see Clarke [15, Proposition 2.1.2]). Thus, if $v(\cdot)$ is locally integrable on $[0, \infty)$, then $|\sigma_\varepsilon(t, v(t))| \leq K\|v(t)\|$ for a.e. $t \in [0, \infty)$. Therefore, $\sigma_\varepsilon(\,\cdot\,, v(\cdot))$ is locally integrable. □

The following result is an obvious extension of Clarke and Vinter [17, Lemma 8.4].

Lemma 3.2. *If $(x(\cdot), u(\cdot), v(\cdot))$ is a perturbed process, then:*

$$\int_0^t L(s, x(s), u(s))ds + \int_0^t \sigma_\varepsilon(s, -v(s))ds - V(0, x(0)) \geq 0,$$

for every $t \in [0, \infty)$ with the equality at $(x_0(\cdot), u_0(\cdot), v(\cdot) \equiv 0)$.

Consider the following perturbed infinite-horizon optimal control problem with free left endpoints:

$$\begin{aligned}
\min \int_0^\infty & L(t, x(t), u(t))dt + \int_0^\infty \sigma_\varepsilon(t, -v(t))dt - V(0, x(0)) \\
\text{s.t. } & \dot{x}(t) = f(t, x(t), u(t)) + v(t) \quad \text{a.e. } t \in [0, \infty), \\
& x(t) \in x_0(t) + \varepsilon B \quad \text{for every } t \in [0, \infty), \\
& u(t) \in U(t) \quad \text{a.e. } t \in [0, \infty), \\
& v(t) \in B \quad \text{a.e. } t \in [0, \infty).
\end{aligned} \tag{P$_\varepsilon$}$$

Here, $u(\cdot)$ and $v(\cdot)$ are control functions and $x(\cdot)$ is a state function. Note that, by Hypothesis 3.1, for every perturbed process $(x(\cdot), u(\cdot), v(\cdot))$, we have:

$$\begin{aligned}
& |L(t, x(t), u(t)) - L(t, x_0(t), u_0(t))| \\
\leq\ & |L(t, x(t), u(t)) - L(t, x_0(t), u(t))| + |L(t, x_0(t), u(t)) - L(t, x_0(t), u_0(t))| \\
\leq\ & k_L(t)\|x(t) - x_0(t)\| + 2\varphi(t) \\
\leq\ & \varepsilon k_L(t) + 2\varphi(t),
\end{aligned}$$

a.e. $t \in [0, \infty)$. Thus, the improper integral $\int_0^\infty L dt$ converges over all perturbed process. A perturbed process is *admissible* for the problem (P$_\varepsilon$) if the improper integral $\int_0^\infty \sigma_\varepsilon dt$ converges. A minimizing process for (P$_\varepsilon$) is a perturbed process that minimizes the objective integral functional of (P$_\varepsilon$) over all admissible process. By Lemma 3.2, $(x_0(\cdot), u_0(\cdot), v(\cdot) \equiv 0)$ is a minimizing process for (P$_\varepsilon$).

3.2.2 Necessary Condition for the Perturbed Problem

Let $l : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz. Consider the following free left and right endpoint infinite-horizon problem:

$$\begin{aligned}
\min\ & l(x(0)) + \int_0^\infty L(t, x(t), u(t))dt \\
\text{s.t. } & \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [0, \infty), \\
& x(t) \in \Omega(t) \quad \text{for every } t \in [0, \infty), \\
& u(t) \in U(t) \quad \text{a.e. } t \in [0, \infty).
\end{aligned} \tag{Q$^\infty$}$$

We say that a process is admissible for the problem (Q^∞) if the improper integral $\int_0^\infty L dt$ converges.

A necessary condition for (P_ε) is obtained from that for the more general problem (Q^∞). While the following result was exploited by Ye [38] with a sketchy outline of the proof, the suggested proof requires an adequate diagonalization method. For completeness, we render an alternative proof. (The compactness argument in Step 3 in the sequel is where we depart from the argument by Ye [38].)

Theorem 3.2. *Let $(x_0(\cdot), u_0(\cdot))$ be a minimizing process for* (Q^∞) *with Hypothesis 3.1. Then, there exists a locally absolutely continuous function $p : [0, \infty) \to \mathbb{R}^n$ such that*

(i) $-\dot{p}(t) \in \partial_x H_P(t, x_0(t), u_0(t), p(t))$ *a.e.* $t \in [0, \infty)$,

(ii) $H_P(t, x_0(t), u_0(t), p(t)) = H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$,

(iii) $p(0) \in \partial l(x_0(0))$.

Proof. [Step 1]: Let $T \in [0, \infty)$ be given arbitrarily. Consider the truncated problem:

$$\begin{aligned}
&\min \ l(x(0)) + \int_0^T L(t, x(t), u(t)) dt \\
&\text{s.t. } \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [0, T], \\
&\qquad x(T) = x_0(T), \\
&\qquad x(t) \in \Omega(t) \quad \text{for every } t \in [0, T], \\
&\qquad u(t) \in U(t) \quad \text{a.e. } t \in [0, T].
\end{aligned} \tag{Q^T}$$

It follows from the Bellman principle of optimality that the minimizing process $(x_0(\cdot), u_0(\cdot))$ for (Q^∞) restricted to the interval $[0, T]$ is a minimizing process for the problem (Q^T) as well. To show this, suppose to the contrary that $(x_0(\cdot), u_0(\cdot))$ is not a minimizing process for (Q^T). Then, there exists an admissible process $(x(\cdot), u(\cdot))$ for (Q^T) such that:

$$l(x(0)) + \int_0^T L(t, x(t), u(t)) dt < l(x_0(0)) + \int_0^T L(t, x_0(t), u_0(t)) dt.$$

Let $\hat{x} : [0, \infty) \to \mathbb{R}^n$ be a locally absolutely continuous function such that $\hat{x}(t) = x(t)$ for $t \in [0, T)$ and $\hat{x}(t) = x_0(t)$ for $t \in [T, \infty)$ and let $\hat{u} : [0, \infty) \to \mathbb{R}^m$ be a measurable function defined by $\hat{u}(t) = u(t)$ for $t \in [0, T)$ and $\hat{u}(t) = u_0(t)$ for $t \in [T, \infty)$. By construction, $(\hat{x}(\cdot), \hat{u}(\cdot))$ is an admissible process for (Q^∞) satisfying:

$$l(\hat{x}(0)) + \int_0^\infty L(t, \hat{x}(t), \hat{u}(t)) dt < l(x_0(0)) + \int_0^\infty L(t, x_0(t), u_0(t)) dt,$$

contradicting the minimality of $(x_0(\cdot), u_0(\cdot))$ for (Q^∞).

While the truncated problem (Q^T) is a free left endpoint problem with fixed right endpoint $x_0(T)$, by reversing the elapse of time, the problem can be readily transformed into a free right endpoint problem with fixed left endpoint $x_0(T)$ (simply replace t with $T-t$). Therefore, the multiplier in the Hamiltonian attached to L in (Q^T) can be normalized to 1. By the nonsmooth maximum principle (see Clarke [15, Theorem 5.2.1]; Vinter [37, Theorem 6.2.1]), there exists an absolutely continuous function $p^T : [0,T] \to \mathbb{R}^n$ such that

(a) $-\dot{p}^T(t) \in \partial_x H_P(t, x_0(t), u_0(t), p^T(t))$ a.e. $t \in [0,T]$,

(b) $H_P(t, x_0(t), u_0(t), p^T(t)) = H(t, x_0(t), p^T(t))$ a.e. $t \in [0,T]$,

(c) $p^T(0) \in \partial l(x_0(0))$.

[Step 2]: By the finite sum formula for generalized gradients (see Clarke [15, Proposition 2.3.3]), we have from the condition (a) that:

$$-\dot{p}^T(t) \in \partial_x f(t, x_0(t), u_0(t))^* p^T(t) - \partial_x L(t, x_0(t), u_0(t)) \quad \text{a.e. } t \in [0,T],$$

where $\partial_x f(t, x_0(t), u_0(t))^*$ denotes the transpose of the generalized Jacobian $\partial_x f(t, x_0(t), u_0(t))$. Denote f in terms of the component functions such that $f = (f_1, \dots, f_n)$. Since $(t,x) \mapsto f_i(t, x, u_0(t))$ and $(t,x) \mapsto L(t, x, u_0(t))$ are measurable functions on Ω satisfying the conditions of Theorem 2.1(i), there exist measurable selections $A^T(\cdot)$ and $a^T(\cdot)$ satisfying:

$$A^T(t) \in (\partial_x f_1(t, x_0(t), u_0(t)) \times \cdots \times \partial_x f_n(t, x_0(t), u_0(t)))^*$$

for every $t \in [0,T]$ with $A^T(t)$ an $n \times n$-matrix and

$$a^T(t) \in \partial_x L(t, x_0(t), u_0(t)) \quad \text{for every } t \in [0,T]$$

with $a^T(t)$ an vector in $\mathbb{R}^n$ such that:

$$-\dot{p}^T(t) = A^T(t) p^T(t) - a^T(t) \quad \text{a.e. } t \in [0,T].$$

(See Clarke [15, Theorem 3.1.1].) Then, the solution $p^T(\cdot)$ to the linear differential equation system is given by:

$$p^T(t) = S^T(t,0) p^T(0) - \int_0^t S^T(t,s) a^T(s) ds \quad \text{for every } t \in [0,T], \tag{3.1}$$

where S^T is the fundamental matrix of the system $\dot{z}(t) = A^T(t) z(t)$ defined by

$$\frac{d}{dt} S^T(t,s) = A^T(t) S^T(t,s), \ S^T(s,s) = I \quad \text{for every } s, t \in [0,T], \tag{3.2}$$

and I is the $n \times n$-identity matrix. Because $\|A^T(t)\| \leq k_f(t)$ for every $t \in [0, T]$ by Hypothesis 3.1(iii), it follows from Gronwall's inequality that:

$$\|S^T(t,s)\| \leq \exp\left(\int_s^t k_f(\tau)d\tau\right) \quad \text{for } 0 \leq s \leq t \leq T.$$

Thus, by combining (3.1) and (3.2), we have:

$$\begin{aligned}
\|\dot{p}^T(t)\| &= \left\|\tfrac{d}{dt}S^T(t,0)p^T(0) + a^T(t)\right\| \\
&\leq \|p^T(0)\|\|A^T(t)\|\|S^T(t,0)\| + \|a^T(t)\| \\
&\leq K_l k_f(t) \exp\left(\int_0^t k_f(s)ds\right) + k_L(t) =: \psi(t),
\end{aligned} \tag{3.3}$$

a.e. $t \in [0, T]$, where K_l is a Lipschitz bound of l near $x_0(0)$. Since ψ is locally integrable on $[0, \infty)$ by Hypothesis 3.1(i), (iii) and (iv), we have:

$$\|p^T(t)\| \leq \|p^T(0)\| + \int_0^t \|\dot{p}^T(s)\|ds \leq K_l + \int_0^t \psi(s)ds \tag{3.4}$$

for every $t \in [0, T]$.

[Step 3]: Note that, for each $k = 1, 2, \ldots$, the sequence $\{p^\nu(\cdot)\}_{\nu \geq k}$ is bounded in the sup norm and equicontinuous on $[0, k]$ by (3.3) and (3.4). By the Ascoli–Arzela theorem, there exists a subsequence $\{p^{\nu_1}(\cdot)\}$ of $\{p^\nu(\cdot)\}_{\nu \geq 1}$ that converges uniformly on $[0, 1]$ to $p_1(\cdot)$. Since the sequence $\{\dot{p}^{\nu_1}(\cdot)\}$ is bounded in $L^1([0, 1]; \mathbb{R}^n)$ and uniformly integrable on $[0, 1]$ by (3.3), we then invoke the Dunford–Pettis criterion to extract a subsequence of $\{\dot{p}^{\nu_1}(\cdot)\}$ that converges weakly to $\alpha(\cdot)$ in $L^1([0, 1]; \mathbb{R}^n)$. Thus:

$$p^{\nu_1}(t) = p^{\nu_1}(0) + \int_0^t \dot{p}^{\nu_1}(s)ds \to p_1(0) + \int_0^t \alpha(s)ds = p_1(t),$$

for every $t \in [0, 1]$, and, hence, $p_1 : [0, 1] \to \mathbb{R}^n$ is absolutely continuous. Similarly, there exists a subsequence $\{p^{\nu_2}(\cdot)\}$ of $\{p^{\nu_1}(\cdot)\}$ that converges uniformly on $[0, 2]$ to an absolutely continuous function $p_2 : [0, 2] \to \mathbb{R}^n$. Continuing in this way, for each $k = 1, 2, \ldots$, we can extract a subsequence $\{p^{\nu_k}(\cdot)\}$ of $\{p^{\nu_{k-1}}(\cdot)\}$ converging uniformly on $[0, k]$ to an absolutely continuous function $p_k : [0, k] \to \mathbb{R}^n$. Note that, by construction, $p_k(t) = p_{k+1}(t)$ for $t \in [0, k]$.

Define the locally absolutely continuous function $p : [0, \infty) \to \mathbb{R}^n$ by:

$$p(t) = p_k(t) \quad \text{for } t \in [0, k].$$

Then, $p(\cdot)$ is well-defined and for every compact interval I of $[0, \infty)$, the restriction $p(\cdot)$ to I is the uniform limit of a subsequence (which possibly depends on I) of $\{p^\nu(\cdot)\}$ restricted to I. By construction of $p(\cdot)$, we have $\dot{p}(t) \leq \psi(t)$ a.e. $t \in [0,\infty)$ and $\|p(t)\| \leq K_l + \int_0^t \psi(s)ds$ for every $t \in [0,\infty)$.

[Step 4]: Define the set-valued mapping $\Sigma : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ by:

$$\Sigma(t,p) = \partial_x H_P(t, x_0(t), u_0(t), p).$$

By Hypothesis 3.1(i) and (iii), $(x,p) \mapsto H_P(t,x,u_0(t),p)$ is a continuous function on $(x_0(t)+\varepsilon B) \times \mathbb{R}^n$ for every $t \in [0,\infty)$ and satisfies the condition of Theorem 2.1(ii). Thus, $\Sigma(t,\cdot)$ is upper semicontinuous on $\mathbb{R}^n$ for every $t \in [0,\infty)$. Note also that Σ is $\mathscr{L} \times \mathscr{B}^n$-measurable by Theorem 2.1(i) and compact convex-valued. Therefore, $\Sigma(t,\cdot)$ satisfies property (Q) and, hence, property (K) a.e. $t \in [0,\infty)$ (see Cesari [14, Theorem 8.5.4]). As demonstrated above, there exists a locally absolutely continuous function $p : [0,\infty) \to \mathbb{R}^n$ such that for each $k = 1, 2, \ldots$, there exists a subsequence $\{p^\nu(\cdot)\}$ (which we do not relabel) such that $\dot{p}^\nu(t) \in \Sigma(t, p^\nu(t))$ a.e. $t \in [0,k]$ for each ν, $p^\nu(\cdot)$ converges uniformly to $p(\cdot)$ on $[0,k]$ and $\dot{p}^\nu(\cdot)$ converges weakly to $\dot{p}(\cdot)$ in $L^1([0,k];\mathbb{R}^n)$. Therefore, by the closure theorem of Cesari (see Cesari [14, Theorem 10.6.i]; or one can apply Clarke [15, Theorem 3.1.7]), we have $\dot{p}(t) \in \Sigma(t,p(t))$ a.e. $t \in [0,k]$ for each k. Henceforth, the condition (i) of the theorem is satisfied.

Since H_P is linear with respect to p, taking the limit along a suitable subsequence (we do not relabel it) in the inequality that follows from the condition (b):

$$H_P(t, x_0(t), u, p^\nu(t)) \leq H_P(t, x_0(t), u_0(t), p^\nu(t))$$

for every $u \in U(t)$ a.e. $t \in [0,\nu]$, yields:

$$H_P(t, x_0(t), u, p(t)) \leq H_P(t, x_0(t), u_0(t), p(t)),$$

for every $u \in U(t)$ a.e. $t \in [0,\infty)$, which is the condition (ii) of the theorem.

The condition (iii) of the theorem follows from the condition (c) and the compactness of $\partial l(x_0(0))$ with $p^\nu(0) \to p(0)$ in $\partial l(x_0(0))$ along a subsequence.

□

3.3 Proof of Theorem 3.1

Now, back to the necessary condition for (P_ε). Since $(x_0(\cdot), u_0(\cdot), v(\cdot) \equiv 0)$ is a minimizing process for (P_ε) by Lemma 3.2, it follows from Theorem 3.2 that there exists a locally absolutely continuous function $p_\varepsilon : [0,\infty) \to \mathbb{R}^n$ such that

(1) $-\dot{p}_\varepsilon(t) \in \partial_x H_P(t, x_0(t), u_0(t), p_\varepsilon(t))$ a.e. $t \in [0,\infty)$,

(2) $H_P(t, x_0(t), u_0(t), p_\varepsilon(t)) = H(t, x_0(t), p_\varepsilon(t))$ a.e. $t \in [0, \infty)$,

(3) $\max_{v \in B}\{\langle p_\varepsilon(t), v\rangle - \sigma_\varepsilon(t, -v)\} = 0$ a.e. $t \in [0, \infty)$,

(4) $-p_\varepsilon(0) \in \partial_x V(0, x_0(0))$.

Since $\|\dot{p}_\varepsilon(t)\| \le \psi(t)$ a.e. $t \in [0, \infty)$ and $\|p_\varepsilon(t)\| \le K + \int_0^t \psi(s)ds$ for every $t \in [0, \infty)$ with $\psi(t) = Kk_f(t)\exp(\int_0^t k_f(s)ds) + k_L(t)$, where K is the Lipschitz bound of $V(0, \cdot\,)$ given in Hypothesis 3.1(v). Thus, the net $\{p_\varepsilon(\cdot)\}$ is an equicontinuous family of locally absolutely continuous functions on $[0, \infty)$ and, hence, the similar diagonalization process as in Step 3 of the proof of Theorem 3.2 yields: there exists a locally absolutely continuous function $p : [0, \infty) \to \mathbb{R}^n$ such that, for every compact subset I of $[0, \infty)$, the net $\{p_\varepsilon(\cdot)\}$ contains a subnet (which we do not relabel) such that $p_\varepsilon(\cdot)$ converges uniformly to $p(\cdot)$ on I and $\dot{p}_\varepsilon(\cdot)$ converges weakly to $\dot{p}(\cdot)$ in $L^1(I; \mathbb{R}^n)$ as $\varepsilon \to 0$. Therefore, by taking the limits in the conditions (1), (2) and (4) along a suitable subnet as in Step 4 of the proof of Theorem 3.2, at the limit, we obtain the conditions (i), (ii) and (iv) of the theorem.

Finally, we investigate the implication of the condition (3) according to the argument by Clarke and Vinter [17]. Take a point $t \in [0, \infty)$ at which (3) is true. Then:

$$-p_\varepsilon(t) \in \overline{\mathrm{co}}\, \partial_x V(t, x_0(t) + \varepsilon \bar{B}) =: \Pi_\varepsilon(t),$$

for otherwise $-p_\varepsilon(t)$ and the closed convex set $\Pi_\varepsilon(t)$ can be strictly separated, i.e., there exists a vector v in B such that:

$$\langle p_\varepsilon(t), v\rangle > \max\{-\langle p, v\rangle \mid p \in \Pi_\varepsilon(t)\} = \sigma_\varepsilon(t, -v)$$

in contradiction of (3). Thus, $-p_\varepsilon(t) \in \Pi_\varepsilon(t)$ a.e. $t \in [0, \infty)$ and passing to the limit along a subnet yields:

$$-p(t) \in \bigcap_{\varepsilon > 0} \overline{\mathrm{co}}\, \partial_x V(t, x_0(t) + \varepsilon \bar{B}) \quad \text{a.e. } t \in [0, \infty). \tag{3.5}$$

We claim that the condition (iii) of the theorem:

$$-p(t) \in \partial_x V(t, x_0(t)) \quad \text{a.e. } t \in [0, \infty),$$

holds. Otherwise, we can strictly separate the point $-p(t)$ and the closed convex set $\partial_x V(t, x_0(t))$, i.e., there exists $v \in \mathbb{R}^n$ and $\delta > 0$ such that:

$$-\langle p(t), v\rangle - \delta > \max\{\langle p, v\rangle \mid p \in \partial_x V(t, x_0(t))\} = V^\circ(t, x_0(t); v).$$

Since the generalized partial derivative $V^{\circ}(t, \cdot\,; \cdot\,)$ is upper semicontinuous (see Clarke [15, Proposition 2.1.1]):

$$-\langle p(t), v\rangle - \tfrac{1}{2}\delta > V^{\circ}(t, x; v),$$

whenever $x \in x_0(t) + \varepsilon B \subset \Omega$ for some $\varepsilon > 0$. Then:

$$\begin{aligned} -\langle p(t), v\rangle - \tfrac{1}{2}\delta &> \sup\{\langle p, v\rangle \mid p \in \partial_x V(t, x_0(t) + \varepsilon \bar{B})\} \\ &= \max\{\langle p, v\rangle \mid p \in \overline{\mathrm{co}}\, \partial_x V(t, x_0(t) + \varepsilon \bar{B})\}. \end{aligned}$$

But this implies that:

$$-p(t) \notin \overline{\mathrm{co}}\, \partial_x V(t, x_0(t) + \varepsilon \bar{B}),$$

in contradiction of (3.5). Therefore, the condition (iii) of the theorem is true.

This completes the proof of Theorem 3.2. □

4 Sufficient Conditions for Optimality

We now turn for the important issue of *sufficient conditions*; that is, conditions that assure that a given admissible process is in fact an optimal solution of the problem.

4.1 Sufficiency Theorems

Definition 4.1. An admissible process $(x_0(\cdot), u_0(\cdot))$ for (P) is *locally minimizing* in $T(x_0(\cdot); \varepsilon)$ if there exists some $\varepsilon > 0$ such that $(x_0(\cdot), u_0(\cdot))$ minimizes the functional $J(x(\cdot), u(\cdot))$ over all admissible processes $(x(\cdot), u(\cdot))$ satisfying $x(t) \in x_0(t) + \varepsilon B$ for every $t \in [0, \infty)$.

Note that, if $\varepsilon = +\infty$, then $(x_0(\cdot), u_0(\cdot))$ is a minimizing process for (P).

Hypothesis 4.1. (i) $L(t, \cdot\,, \cdot\,)$ is lower semicontinuous on $\Omega(t) \times U(t)$ for every $t \in [0, \infty)$.

(ii) $f(t, \cdot\,, \cdot\,)$ is continuous on $\Omega(t) \times U(t)$ for every $t \in [0, \infty)$.

(iii) $U(t)$ is closed for every $t \in [0, \infty)$ and graph (U) is $\mathscr{L} \times \mathscr{B}^m$-measurable.

(iv) For every $t \in [0, \infty)$ and for every bounded subset Z of $\mathbb{R}^n \times \mathbb{R}^n$, the set:

$$\{u \in U(t) \mid \exists (x, v) \in Z : f(t, x, u) = v\},$$

is bounded.

The following result is an extension of Zeidan [40] to the infinite horizon case.

Theorem 4.1. *Suppose that Hypothesis 4.1 is satisfied. Let* $(x_0(\cdot), u_0(\cdot))$ *be an admissible process for* (P) *such that there exist a locally absolutely continuous function* $p : [0, \infty) \to \mathbb{R}^n$*, a locally absolutely continuous* $n \times n$*-symmetric matrix-valued function* P *on* $[0, \infty)$ *and some* $\varepsilon > 0$ *with the following properties.*

(i) *For a.e.* $t \in [0, \infty)$ *and for every* $v \in \varepsilon B$ *and* $u \in U(t)$*:*

$$\begin{aligned} &H_P(t, x_0(t) + v, u, p(t) - P(t)v) \\ &\leq H_P(t, x_0(t), u_0(t), p(t)) - \langle \dot{p}(t) + P(t)\dot{x}_0(t), v\rangle + \tfrac{1}{2}\langle v, \dot{P}(t)v\rangle. \end{aligned}$$

(ii) *For every* $\eta > 0$*, there exists some* $t_0 \in [0, \infty)$ *such that:*

$$\tfrac{1}{2}\langle v, P(t)v\rangle < \langle p(t), v\rangle + \eta \quad \textit{for every } v \in \varepsilon B \textit{ and } t \in [t_0, \infty).$$

Then, $(x_0(\cdot), u_0(\cdot))$ *is a locally minimizing process in* $T(x_0(\cdot); \varepsilon)$ *for* (P).

Note that the condition (i) of the theorem implies the condition (ii) of Theorem 3.1. When $\varepsilon = +\infty$ and the matrix-valued function P in the theorem happens to be identically the zero matrix, the condition (i) of the theorem reduces to the supergradient inequality for H:

$$H(t, x_0(t) + v, p(t)) - H(t, x_0(t), p(t)) \leq -\langle \dot{p}(t), v\rangle, \tag{4.1}$$

for every $v \in \mathbb{R}^n$. The condition (4.1) is imposed by Feinstein and Luenberger [20] to obtain the sufficiency result. This is, of course, satisfied if $H(t, x, p(t))$ is concave in x for every $t \in [0, \infty)$. Thus, the condition (i) of the theorem can be viewed as a strengthening of the necessary conditions (i) and (ii) of Theorem 3.1 under the convexity hypothesis.

If $P(t)$ is negative semidefinite for every $t \in [0, \infty)$ and $\lim_{t\to\infty} p(t) = 0$, then the condition (ii) of the theorem is satisfied. On the other hand, if $P = 0$, then the condition (ii) of the theorem is equivalent to the transversality condition at infinity:

$$\lim_{t\to\infty} p(t) = 0. \tag{4.2}$$

For the finite horizon case, sufficient conditions for optimality were given by Mangasarian [25] under the hypothesis that the Hamiltonian H_P is concave and differentiable in (x, u), whose result was extended by Seierstadt and Sydsæter [32] to the infinite horizon case. Thus, the above observation leads to an extension of the Mangasarian sufficiency theorem with an infinite horizon as follows.

Corollary 4.1. *Suppose that Hypothesis 4.1 is satisfied. Let* $(x_0(\cdot), u_0(\cdot))$ *be an admissible process for* (P) *and* $p : [0, \infty) \to \mathbb{R}^n$ *be a locally absolutely continuous function with the following properties.*

(i) $H(t, \cdot, p(t))$ *is concave on* $\mathbb{R}^n$ *for every* $t \in [0, \infty)$.

(ii) $-\dot{p}(t) \in \partial_x H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$.

(iii) $H_P(t, x_0(t), u_0(t), p(t)) = H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$.

(iv) $\lim_{t\to\infty} p(t) = 0$.

Then, $(x_0(\cdot), u_0(\cdot))$ *is a minimizing process for* (P).

For the derivation of the transversality condition (4.2) as a necessary condition for optimality, see Aseev and Kryaziimskiy [3] and Michel [26] for the smooth case and Ye [38] for the nonsmooth case.

Consider the following transversality condition at infinity:

$$\lim_{t\to\infty} \langle p(t), x(t) - x_0(t)\rangle \geq 0, \tag{4.3}$$

for every admissible arc for (P). To obtain the sufficiency result, Seierstadt and Sydsæter [32] imposed the condition (4.3) in addition to the conditions (i) and (ii) of the corollary as well as the differentiability assumption on (L, f) and Feinstein and Luenberger [20] assumed (4.3) for the nonsmooth nonconcave Hamiltonians along with the condition (4.1).

Note that the condition (4.3) is implied by the condition (4.2) if every admissible arc is bounded. However, (4.3) is difficult to check in practice when admissible arcs are unbounded because it involves possible information on the limit behavior of all admissible arcs. The condition (4.2) on its own right needs no such information and improves upon (4.3). Its derivation as a sufficient condition is novel in the literature.

Let V be an extension of the value function on Ω (which we do not relabel) to $[0, \infty) \times \mathbb{R}^n$ given by $V(t, x) = +\infty$ for $(t, x) \notin \Omega$. We now provide a new sufficient condition in terms of the adjoint inequality for the value function.

Theorem 4.2. *Suppose that Hypothesis 4.1 is satisfied. Let* $(x_0(\cdot), u_0(\cdot))$ *be an admissible process for* (P) *such that there exist a locally absolutely continuous function* $p : [0, \infty) \to \mathbb{R}^n$ *and a locally absolutely continuous* $n \times n$*-symmetric matrix-valued function* P *on* $[0, \infty)$ *with the following properties.*

(i) *For a.e.* $t \in [0, \infty)$ *and for every* $v \in \mathbb{R}^n$ *and* $u \in U(t)$*:*

$$\begin{aligned}&H_P(t, x_0(t) + v, u, p(t) - P(t)v)\\ &\leq H_P(t, x_0(t), u_0(t), p(t)) - \langle \dot{p}(t) + P(t)\dot{x}_0(t), v\rangle + \tfrac{1}{2}\langle v, \dot{P}(t)v\rangle.\end{aligned}$$

(ii) *For every* $v \in \mathbb{R}^n$ *and* $t \in [0, \infty)$*:*

$$V(t, x_0(t)) - \langle p(t) + P(t)x(t), v \rangle + \tfrac{1}{2}\langle v, P(t)v \rangle \leq V(t, x_0(t) + v).$$

(iii) $\lim_{t \to \infty} V(t, x_0(t)) = 0$.

Then, $(x_0(\cdot), u_0(\cdot))$ *is a minimizing process for* (P).

For the case in which $P = 0$ in the theorem, the condition (ii) of the theorem reduces to the subgradient inequality for $V(t, \cdot\,)$:

$$V(t, x_0(t) + v) - V(t, x_0(t)) \geq -\langle p(t), v \rangle,$$

for every $v \in \mathbb{R}^n$. This is, indeed, satisfied if $V(t, x)$ is convex in x for every $t \in [0, \infty)$. Thus, the condition (ii) of the theorem can be viewed as a strengthening of the adjoint inclusions (iii) and (iv) of Theorem 3.1.

While the role of the limit behavior of the value function at infinity in the condition (iii) of the theorem is novel in optimal control theory, it is clarified in the derivation of the sufficiency result for convex problems of calculus of variations with an infinite horizon by Benveniste and Scheinkman [12] and Takekuma [35].

4.2 Proof of Sufficiency Theorems

Let $F : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be an $\mathscr{L} \times \mathscr{B}^n \times \mathscr{B}^n$-measurable function. Consider the problem of Lagrange in calculus of variations:

$$\min \ \mathscr{J}(x(\cdot)) := \int_0^\infty F(t, x(t), \dot{x}(t))dt, \tag{L}$$

where the minimum is taken over all locally absolutely continuous functions (arcs) $x : [0, \infty) \to \mathbb{R}^n$ satisfying the initial condition $x(0) = x_0$. We say that $x(\cdot)$ is an *admissible arc* if $\mathscr{J}(x(\cdot))$ is finite and the initial condition is satisfied and that $x_0(\cdot)$ is *locally minimizing* in $T(x_0(\cdot); \varepsilon)$ for the problem (L) if there exists some $\varepsilon > 0$ such that $x_0(\cdot)$ minimizes $\mathscr{J}(x(\cdot))$ over all admissible arcs $x(\cdot)$ satisfying $x(t) \in x_0(t) + \varepsilon B$ for every $t \in [0, \infty)$. The Hamiltonian for (L) is given by:

$$\mathscr{H}(t, x, p) = \sup_{v \in \mathbb{R}^n} \{\langle p, v \rangle - F(t, x, v)\}.$$

The sufficiency theorem for problems of Bolza due to Zeidan [39] is adapted to the infinite horizon setting here.

Theorem 4.3. *Let $x_0(\cdot)$ be an admissible arc for* (L)*. Suppose that there exist a locally absolutely continuous function $p : [0, \infty) \to \mathbb{R}^n$, a locally absolutely continuous $n \times n$-symmetric matrix-valued function P on $[0, \infty)$ and some $\varepsilon > 0$ with the following properties.*

(i) *For every $v \in \mathbb{R}^n$ and a.e. $t \in [0, \infty)$:*

$$F(t, x_0(t), \dot{x}_0(t) + v) - F(t, x_0(t), \dot{x}_0(t)) \geq \langle p(t), v \rangle.$$

(ii) *For every $v \in \varepsilon B$ and a.e. $t \in [0, \infty)$:*

$$\begin{aligned} &\mathscr{H}(t, x_0(t) + v, p(t) - P(t)v) - \mathscr{H}(t, x_0(t), p(t)) \\ \leq\ &- \langle \dot{p}(t) + P(t)\dot{x}_0(t), v \rangle + \tfrac{1}{2}\langle v, \dot{P}(t)v \rangle. \end{aligned}$$

(iii) *For every $\eta > 0$, there exists some $t_0 \in [0, \infty)$ such that:*

$$\tfrac{1}{2}\langle v, P(t)v \rangle < \langle p(t), v \rangle + \eta \quad \textit{for every } v \in \varepsilon B \textit{ and } t \in [t_0, \infty).$$

Then, $x_0(\cdot)$ is a locally minimizing arc in $T(x_0(\cdot); \varepsilon)$ for (L)*.*

Proof. Under the conditions (i) and (ii) of the theorem, for every admissible arc $x(\cdot)$ with $x(t) \in x_0(t) + \varepsilon B$ for every $t \in [0, \infty)$:

$$\begin{aligned} &\langle p(t), x(t) - x_0(t) \rangle - \tfrac{1}{2}\langle x(t) - x_0(t), P(t)(x(t) - x_0(t)) \rangle \\ \leq\ &\int_0^t [F(t, x(s), \dot{x}(s)) - F(s, x_0(s), \dot{x}_0(s))] ds. \end{aligned} \tag{4.4}$$

(See Zeidan [39] or Clarke [15, Theorem 4.3.1].) Let $\eta > 0$ be given arbitrarily. By the condition (iii) of the theorem and (4.4), we have:

$$\int_0^t [F(t, x(s), \dot{x}(s)) - F(s, x_0(s), \dot{x}_0(s))] dt > -\eta \quad \text{for every } t \in [t_0, \infty).$$

Letting $t \to \infty$ in this inequality gives $\mathscr{J}(x_0(\cdot)) \leq \mathscr{J}(x(\cdot)) + \eta$ for every admissible arc $x(\cdot)$ with $x(s) \in x_0(s) + \varepsilon B$ for every $s \in [0, \infty)$. Hence, $x_0(\cdot)$ is locally minimizing in $T(x_0(\cdot); \varepsilon)$ for (L). □

Define the function $F : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by:

$$F(t, x, v) = \inf\{L(t, x, u) \mid u \in U(t) : f(t, x, u) = v\}. \tag{4.5}$$

(Note that the infimum over the empty set is taken to be $+\infty$.) An established technique for transforming the problem of optimal control (P) into that of calculus of variations (L) is available here (see Rockafeller [30, 31]). It is based on the observation that the Hamiltonian H for (P) coincides with the Hamiltonian $\mathscr{H}$ for (L) on Ω. Indeed:

$$\begin{aligned}
&\sup_{v\in\mathbb{R}^n}\{\langle p,v\rangle - F(t,x,v)\}\\
&= \sup_{v\in\mathbb{R}^n}\{\langle p,v\rangle - \inf\{L(t,x,u)\mid u\in U(t) : f(t,x,u)=v\}\}\\
&= \sup_{u\in U(t)}\{\langle p, f(t,x,u)\rangle - L(t,x,u)\},
\end{aligned}$$

and, hence, for every $(t,x,p)\in[0,\infty)\times\mathbb{R}^n\times\mathbb{R}^n$:

$$\mathscr{H}(t,x,p) = H(t,x,p). \tag{4.6}$$

The following result is a special case of the equivalence theorem due to Rockafeller [31]. (See also Clarke [15, Theorem 5.4.1].)

Equivalence Theorem. *Suppose that Hypothesis 4.1 is satisfied. Let F be given in* (4.5). *Then, $x_0(\cdot)$ is a minimizing arc for* (L) *if and only if there is a control function $u_0 : [0,\infty)\to\mathbb{R}^m$ corresponding to $x_0(\cdot)$ such that $(x_0(\cdot), u_0(\cdot))$ is a minimizing process for* (P).

Proof of Theorem 4.1. The argument is based on Zeidan [40] and Clarke [15, Theorem 5.4.2]. Hypothesis 4.1 assures that F is $\mathscr{L}\times\mathscr{B}^n\times\mathscr{B}^n$-measurable and $F(t,\cdot,\cdot)$ is lower semicontinuous for every $t\in[0,\infty)$. (See Clarke [15, Theorem 5.4.1] and Rockafeller [31].) The condition (i) of the theorem and (4.6) imply that:

$$F(t,x_0(t),\dot{x}_0(t)) = L(t,x_0(t),u_0(t)) \quad \text{a.e. } t\in[0,\infty). \tag{4.7}$$

On the other hand, (4.5) implies that $\mathscr{J}(x(\cdot))\le J(x(\cdot),u(\cdot))$ for every admissible process $(x(\cdot),u(\cdot))$ for (P) with $x(t)\in x_0(t)+\varepsilon B$ for every $t\in[0,\infty)$. Therefore, to show that $(x_0(\cdot),u_0(\cdot))$ is a locally minimizing process in $T(x_0(\cdot);\varepsilon)$ for (P), it suffices to demonstrate that $x_0(\cdot)$ is a locally minimizing arc in $T(x_0(\cdot);\varepsilon)$ for (L), which is guaranteed if the conditions (i) and (ii) of Theorem 4.3 are shown to be met. It is easy to verify that the condition (i) of Theorem 4.1 and (4.6) imply that:

$$\mathscr{H}(t,x_0(t),p(t)) = \langle p(t),\dot{x}_0(t)\rangle - F(t,x_0(t),\dot{x}_0(t)) \quad \textit{a.e. } t\in[0,\infty).$$

Thus, the condition (i) of Theorem 4.3 is satisfied. The condition (i) of Theorem 4.1 and (4.6) again yield the condition (ii) of Theorem 4.3. $\square$

Proof of Theorem 4.2. Let $(x_0(\cdot), u_0(\cdot))$ be an admissible process for (P) satisfying the conditions of the theorem. It suffices to show that:

$$V(0, x_0(0)) = \int_0^t L(s, x_0(s), u_0(s))ds + V(t, x_0(t)), \tag{4.8}$$

for every $t \in [0, \infty)$, because taking the limit as $t \to \infty$ in (4.8) yields:

$$V(0, x_0(0)) = \int_0^\infty L(s, x_0(s), u_0(s))ds,$$

from which the optimality of $(x_0(\cdot), u_0(\cdot))$ follows.

Suppose to the contrary that (4.8) is not true. By the definition of V, there exists some $\eta > 0$ such that:

$$V(0, x_0(0)) + \eta < \int_0^T L(t, x_0(t), u_0(t))dt + V(T, x_0(T)), \tag{4.9}$$

for some $T \in [0, \infty)$. Again by the definition of V, there exists an admissible process $(x(\cdot), u(\cdot))$ for (P) such that:

$$\int_0^\infty L(t, x(t), u(t))dt < V(0, x_0(0)) + \eta.$$

Thus, the inequality (4.9) implies the existence of an admissible process $(x(\cdot), u(\cdot))$ for (P) such that:

$$\int_0^T L(t, x(t), u(t))dt + V(T, x(T)) < \int_0^T L(t, x_0(t), u_0(t))dt + V(T, x_0(T)). \tag{4.10}$$

It follows from (4.5) that:

$$L(t, x(t), u(t)) - L(t, x_0(t), u_0(t)) \geq F(t, x(t), \dot{x}(t)) - F(t, x_0(t), \dot{x}_0(t)), \tag{4.11}$$

a.e. $t \in [0, \infty)$. As noted in the proof of Theorem 4.1, the conditions (i) and (ii) of Theorem 4.3 are satisfied for $\varepsilon = +\infty$. Thus, integrating the inequality (4.11) together with (4.4) and the condition (ii) of the theorem yield:

$$\int_0^T [L(t, x(t), u(t)) - L(t, x_0(t), u_0(t))]dt \geq -(V(T, x(T)) - V(T, x_0(T))),$$

which contradicts (4.10). $\square$

5 Necessary and Sufficient Conditions for Optimality

In this section, we derive the necessary and sufficient conditions for optimality under convexity hypotheses. Convex problems of optimal control examined here clarify the role of the limit behavior of the value function for a complete characterization of optimality. Furthermore, we investigate the role of transversality conditions at infinity and derive them as necessary and sufficient conditions for optimality under some additional assumptions.

5.1 Limit Behavior of the Value Function at Infinity

As demonstrated in the Appendix, the hypothesis that follows is derived from the convexity hypothesis on the primitive (L, f, Ω, U).

Hypothesis 5.1. (i) $\Omega(t) \times U(t)$ is convex for every $t \in [0, \infty)$.

(ii) $H(t, \cdot\,, p)$ is concave on $\mathbb{R}^n$ for every $(t, p) \in [0, \infty) \times \mathbb{R}^n$.

(iii) $V(t, \cdot\,)$ is convex on $\Omega(t)$ for every $t \in [0, \infty)$.

Theorem 5.1. *Suppose that Hypotheses 3.1, 4.1 and 5.1 are satisfied. An admissible process $(x_0(\cdot), u_0(\cdot))$ is a minimizing process for* (P) *if and only if the following conditions are satisfied.*

(i) *There exists a locally absolutely continuous function $p : [0, \infty) \to \mathbb{R}^n$ such that*

(a) $-\dot{p}(t) \in \partial_x H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$,

(b) $H_P(t, x_0(t), u_0(t), p(t)) = H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$,

(c) $-p(t) \in \partial_x V(t, x_0(t))$ *for every* $t \in [0, \infty)$,

(ii) $\lim_{t\to\infty} V(t, x_0(t)) = 0$.

Proof. [Necessity]: Let $(x_0(\cdot), u_0(\cdot))$ be a minimizing process for (P). The conditions (i-a) to (i-c) of the theorem follow from Theorem 3.1 and Corollary 3.2. Since:

$$-\infty < V(t, x_0(t)) = \int_t^\infty L(s, x_0(s), u_0(s))ds < \infty,$$

by the Bellman principle of optimality, taking the limit as $t \to \infty$ in this equality yields the condition (ii) of the theorem.

[Sufficiency]: Since $H(t, \cdot\,, p)$ is concave on $\mathbb{R}^n$ and $V(t, \cdot\,)$ is convex on $\Omega(t)$ by Hypothesis 5.1, Theorem 4.2 applies with $P = 0$. □

5.2 Transversality Condition at Infinity

To derive a sharper result on the transversality condition at infinity, one must specify the problem in more detail. The following hypothesis is in accordance with the standard conditions in economic growth theory such as Benveniste and Scheinkman [12] and Takekuma [35].

Hypothesis 5.2. (i) $\Omega(t) \subset \mathbb{R}^n_+$ for every $t \in [0, \infty)$.

(ii) $0 \in U(t)$ a.e. $t \in [0, \infty)$.

(iii) $f(t, 0, 0) = 0$ a.e. $t \in [0, \infty)$.

(iv) $L(t, 0, 0) \leq 0$ a.e. $t \in [0, \infty)$.

(v) $L(t, \cdot\,, u)$ is nondecreasing on $\Omega(t)$ for every $u \in U(t)$ a.e. $t \in [0, \infty)$.

Theorem 5.2. *Suppose that Hypotheses 3.1, 4.1, 5.1 and 5.2 are satisfied. An admissible process* $(x_0(\cdot), u_0(\cdot))$ *is a minimizing process for* (P) *if and only if there exists a locally absolutely continuous function* $p : [0, \infty) \to \mathbb{R}^n$ *such that*

(i) $-\dot{p}(t) \in \partial_x H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$*;*

(ii) $H_P(t, x_0(t), u_0(t), p(t)) = H(t, x_0(t), p(t))$ *a.e.* $t \in [0, \infty)$*;*

(iii) $-p(t) \in \partial_x V(t, x_0(t))$ *for every* $t \in [0, \infty)$*;*

(iv) $\lim_{t\to\infty} \langle p(t), x_0(t)\rangle = 0$.

Proof. [Necessity]: Let $(x_0(\cdot), u_0(\cdot))$ be a minimizing process for (P). Then, by Theorem 5.1, the conditions (i) to (iii) of the theorem are satisfied. By the convexity of $V(t, \cdot\,)$ in Hypothesis 5.1(iii), the condition (iii) of the theorem implies that

$$V(t, 0) - V(t, x_0(t)) \geq -\langle p(t), x_0(t)\rangle \quad \text{for every } t \in [0, \infty).$$

Since $V(t, x)$ is nondecreasing in x by Hypothesis 5.2(v), $-p(t)$ is in the nonnegative orthant. We thus have $-\langle p(t), x_0(t)\rangle \geq 0$ for every $t \in [0, \infty)$ because $x_0(t)$ is in the nonnegative orthant by Hypothesis 5.2(i). Note that $(x(\cdot) \equiv 0, u(\cdot) \equiv 0)$ is an admissible process on $[t, \infty)$ with initial condition $0 \in \Omega(t)$ by Hypothesis 5.2(i) to (iii). Thus, we have:

$$V(t, 0) \leq \int_t^\infty L(s, 0, 0)ds \leq 0,$$

by Hypothesis 5.2(iv). Therefore, we have:

$$-V(t, x_0(t)) \geq -\langle p(t), x_0(t) \rangle \geq 0 \quad \text{for every } t \in [0, \infty).$$

Taking the limit in this inequality yields:

$$0 = -\lim_{t \to \infty} V(t, x_0(t)) \geq \lim_{t \to \infty} -\langle p(t), x_0(t) \rangle \geq 0.$$

Therefore, the condition (iv) of the theorem is true.

[Sufficiency]: Let $(x_0(\cdot), u_0(\cdot))$ be an admissible process satisfying the conditions (i) to (iv) of the theorem. Since $H(t, \cdot, p(t))$ is concave by Hypothesis 5.1, the adjoint inclusion in the condition (i) of the theorem implies:

$$H(t, x(t), p(t)) - H(t, x_0(t), p(t)) \leq \langle -\dot{p}(t), x(t) - x_0(t) \rangle, \tag{5.1}$$

a.e. $t \in [0, \infty)$ for every admissible process $(x(\cdot), u(\cdot))$ for (P). It follows from the condition (ii) of the theorem that:

$$H(t, x_0(t), p(t)) = \langle p(t), \dot{x}_0(t) \rangle - L(t, x_0(t), u_0(t)),$$

and

$$H(t, x(t), p(t)) \geq \langle p(t), \dot{x}(t) \rangle - L(t, x(t), u(t)),$$

and substituting these into (5.1) yields:

$$L(t, x(t), u(t)) - L(t, x_0(t), u_0(t)) \geq \frac{d}{dt} \langle p(t), x(t) - x_0(t) \rangle \quad \text{a.e. } t \in [0, \infty).$$

Integrating this inequality gives:

$$\begin{aligned} \int_0^T [L(t, x(t), u(t)) - L(t, x_0(t), u_0(t))] dt &\geq \langle p(T), x(T) - x_0(T) \rangle \\ &\geq \langle p(T), x_0(T) \rangle, \end{aligned}$$

for every $T \in [0, \infty)$. Here, the fact that $p(T)$ and $x_0(T)$ are in the nonnegative orthant is employed. Letting $T \to \infty$ in this inequality along with the transversality condition (iv) of the theorem yields:

$$\int_0^\infty L(t, x(t), u(t)) dt \geq \int_0^\infty L(t, x_0(t), u_0(t)) dt.$$

This completes the proof. □

While the transversality condition at infinity:

$$\lim_{t\to\infty} \langle p(t), x_0(t)\rangle = 0,$$

is familiar in economic growth theory, the derivation of this condition as a necessary and sufficient condition for optimality in optimal control is novel in the literature. Aseev and Kryaziimskiy [3] obtained this as a necessary condition for optimality under somewhat restrictive smoothness assumptions with quasi-linear control systems.

For convex problems of Lagrange in calculus of variations, Araujo and Scheinkman [2], Benveniste and Scheinkman [12] and Takekuma [35] obtained this condition as a necessary and sufficient condition for optimality for the nonsmooth case and Becker and Boyd [10] did so for the smooth case. For the derivation of the variant of this condition as a necessary condition in nonconvex smooth problems of Lagrange in calculus of variations with unbounded integrands, see Kamihigashi [24].

A Properties of the Value Function and the Hamiltonian

We have assumed in Hypothesis 3.1(v) that $V(t,\cdot\,)$ is Lipschitz of rank K on $x_0(t)+\varepsilon B$ for every $t\in[0,\infty)$. In Appendix A.1, we demonstrate the continuity of V on the ε-tube about $x_0(\cdot)$ and the Lipschitz continuity of $V(t,\cdot\,)$ under the existence of a minimizing process for any initial condition. For the finite horizon case, the result is well-known (see, for instance, Vinter [37, Proposition 12.3.5]), but some intricate arguments are involved for the infinite horizon case concerning the integrability of the integrand and the interiority of the minimizing arcs.

The convexity of the value function is proven in Appendix A.2 under some additional assumptions. The concavity of the Hamiltonian is demonstrated in Appendix A.3.

A.1 Lipschitz Continuity of the Value Function

Theorem A.1. *Suppose that Hypothesis 3.1 is satisfied. Then, V is continuous on the ε-tube about $x_0(\cdot)$.*

Proof. Let $(t,x)\in T(x_0(\cdot);\varepsilon)\subset\Omega$ be given arbitrarily. Since $V(t,x)$ is finite by Hypothesis 3.1(v), the Bellman principle of optimality implies that, for every admissible process $(x(\cdot), u(\cdot))$ on $[t, \infty)$ with the initial condition $x(t) = x$, we have:

$$V(t, x) \le \int_t^\tau L(s, x(s), u(s))ds + V(\tau, x(\tau)),$$

whenever $\tau \in [t, \infty)$. The continuity of $x(\cdot)$ implies that there exists some $\delta > 0$ such that $(\tau, x) \in T(x_0(\cdot); \varepsilon)$ for every $\tau \in [t, t+\delta)$. Thus, by employing the Lipschitz continuity of $V(\tau, \cdot\,)$ on $x_0(\tau)+\varepsilon B$, for every $(\tau, y) \in T(x_0(\cdot); \varepsilon)$ with $\tau \in [t, t + \delta)$, we have:

$$\begin{aligned} V(t, x) - V(\tau, y) &\le \int_t^\tau L(s, x(s), u(s))ds + |V(\tau, x(\tau)) - V(\tau, x)| \\ &\qquad\qquad + |V(\tau, x) - V(\tau, y)| \\ &\le \int_t^\tau |L(s, x(s), u(s))|ds + K(\|x(\tau) - x\| + K\|x - y\|). \end{aligned}$$

The similar argument applied for the case $0 \le \tau < t$ yields:

$$|V(t, x) - V(\tau, y)| \le \int_{t\wedge\tau}^{t\vee\tau} |L(s, x(s), u(s))|ds + K(\|x(\tau) - x\| + K\|x - y\|),$$

for every $\tau \in [t - \delta, t + \delta] \cap [0, \infty)$ with some $\delta > 0$. Letting $(\tau, y) \to (t, x)$ yields $V(\tau, y) \to V(t, x)$. Therefore, V is continuous at every $(t, x) \in T(x_0(\cdot); \varepsilon)$. □

We extend the notion of an ε-tube. Let $\theta_\varepsilon : [0, \infty) \to \mathbb{R}$ be a positive measurable function given by $\theta_\varepsilon(s) = \varepsilon \exp(\int_0^s k_f(\tau)d\tau)$ for $s \in [0, \infty)$ with $\varepsilon > 0$. An *extended ε-tube* about continuous function $x : [t, \infty) \to \mathbb{R}^n$ is of the form:

$$T(x(\cdot); \theta_\varepsilon) := \{(s, x) \in [t, \infty) \times \mathbb{R}^n \mid x \in x(s) + \theta_\varepsilon(s)B\}.$$

Hypothesis A.1. There exists some $\varepsilon > 0$ such that, for every $(t, x) \in \Omega$, there exists a minimizing process $(x(\,\cdot \mid t, x), u(\,\cdot \mid t, x))$ on $[t, \infty)$ with the initial condition $x(t \mid t, x) = x$ such that the extended ε-tube about $x(\cdot \mid t, x)$ is contained in Ω.

Without loss of generality, we may assume that $x_0(\cdot) = x(\,\cdot \mid 0, x_0)$.

Theorem A.2. *Suppose that the conditions (i) to (iv) of Hypothesis 3.1, and Hypothesis A.1, are satisfied. Then, $V(t, \cdot\,)$ is Lipschitz of rank K on $x_0(t) + \frac{\varepsilon}{2}B$ for every $t \in [0, \infty)$.*

Proof. Let $T \in [0, \infty)$, $t \in [0, T]$ and $x, y \in x_0(t) + \frac{\varepsilon}{2} B \subset \Omega(t)$ be given arbitrarily. Consider the ordinary differential equation (ODE):

$$\dot{x}(s) = f(s, x(s), u(s \mid t, y)) \quad \text{for } s \in [t, \infty),\ x(t) = x.$$

Then, Hypothesis 3.1(iii) guarantees the existence and uniqueness of a solution $x(\cdot)$ of the ODE satisfying:

$$\|x(T) - x(T \mid t, y)\| \leq \exp\left(\int_t^T k_f(s) ds\right) \|x - y\| \leq \theta_\varepsilon(T).$$

(See Vinter [37, Corollary 2.4.5].) This shows that $(T, x(T))$ is in the extended ε-tube about $x(\cdot \mid t, y)$ for every $T \in [0, \infty)$. Therefore, $x(s) \in \Omega(s)$ for every $s \in [t, \infty)$ by Hypothesis A.1 and $(x(\cdot), u(\cdot \mid t, y))$ is an admissible process on $[t, \infty)$ with initial condition $x(t) = x \in \Omega(t)$. We thus have:

$$V(t, x) \leq \int_t^\infty L(s, x(s), u(s \mid t, y)) ds.$$

In view of:

$$\|x(s) - x(s \mid t, y)\| \leq \exp\left(\int_t^T k_f(s) ds\right) \|x - y\|,$$

for every $s \in [t, T]$ and $T \in [0, \infty)$, we obtain:

$$\begin{aligned} V(t, x) - V(t, y) &\leq \int_t^\infty [L(s, x(s), u(s \mid t, y)) - L(s, x(s \mid t, y), u(s \mid t, y))] ds \\ &\leq \int_t^\infty \left[k_L(s) \exp\left(\int_t^s k_f(\tau) d\tau\right)\right] ds \|x - y\| \leq K \|x - y\|, \end{aligned}$$

with $K = \int_0^\infty k(s) ds < \infty$ by Hypothesis 3.1(iv). Interchanging the role of x with y in the above, we arrive at $|V(t, x) - V(t, y)| \leq K \|x - y\|$. Therefore, $V(t, \cdot)$ is Lipschitz of rank K on $x_0(t) + \frac{\varepsilon}{2} B$ for every $t \in [0, \infty)$ and the Lipschitz constant K is independent of t. $\square$

A.2 Convexity of the Value Function

Define the set-valued mapping $\Gamma : \Omega \rightrightarrows \mathbb{R} \times \mathbb{R}^n$ by:

$$\Gamma(t, x) = \{(v, w) \in \mathbb{R}^n \times \mathbb{R} \mid \exists u \in U(t) : w \geq L(t, x, u),\ v = f(t, x, u)\},$$

and the set M by:

$$M = \{(t, x, u) \in [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \mid (x, u) \in \Omega(t) \times U(t)\}.$$

Hypothesis A.2. (i) L and f are continuous on M.

(ii) $-\infty < V(t, x)$ for every $(t, x) \in \Omega$.

(iii) Ω and graph (U) are closed.

(iv) $\Omega(t)$ is convex for every $t \in [0, \infty)$.

(v) $\Gamma(t, \cdot\,) : \Omega(t) \rightrightarrows \mathbb{R}^n \times \mathbb{R}$ has the convex graph for every $t \in [0, \infty)$.

The condition (ii) of the hypothesis is automatically satisfied if Hypothesis A.1 is imposed. The conditions (iv) and (v) of the hypothesis are somewhat stronger than the standard convexity hypothesis guaranteeing the existence of a minimizing process that $\Gamma(t, \cdot\,)$ is convex-valued for every $t \in [0, \infty)$. (See Balder [6], Bates [8], Baum [9], Bell et al. [11], Feinstein and Luenberger [20].)

Theorem A.3. *Suppose that Hypothesis A.2 is satisfied. Then, $V(t, \cdot\,)$ is convex on $\Omega(t)$ for every $t \in [0, \infty)$.*

Proof. Let $x, y \in \Omega(t)$ and $\lambda \in (0, 1)$ be arbitrary. We must show the inequality:

$$V(t, \lambda x + (1 - \lambda)y) \leq \lambda V(t, x) + (1 - \lambda)V(t, y).$$

If $V(t, x) = +\infty$ or $V(t, y) = +\infty$, then the inequality is trivially true. Suppose that $V(t, x) < +\infty$ and $V(t, y) < +\infty$. Denote by $(x(\cdot), u(\cdot))$ an admissible process on $[t, \infty)$ with the initial condition $x(t) = x$ and by $(x'(\cdot), u'(\cdot))$ an admissible process on $[t, \infty)$ with the initial condition $x'(t) = y$. Define $z(s) = L(s, x(s), u(s))$ and $z'(s) = L(s, x'(s), u'(s))$ for $s \in [t, \infty)$. We then have $(z(s), \dot{x}(s)) \in \Gamma(s, x(s))$ and $(z'(s), \dot{x}'(s)) \in \Gamma(s, x'(s))$ a.e. $s \in [t, \infty)$. Let $(z_\lambda(\cdot), x_\lambda(\cdot)) = \lambda(z(\cdot), x(\cdot)) + (1 - \lambda)(z'(\cdot), x'(\cdot))$. From Hypothesis A.2(iv) and (v), it follows that $(z_\lambda(s), \dot{x}_\lambda(s)) \in \Gamma(s, x_\lambda(s))$ a.e. $s \in [t, \infty)$ with the initial condition $x_\lambda(t) = \lambda x + (1 - \lambda)y$. Then, by the implicit function theorem for orientor fields (see Cesari [14, Section 8.2.C.2]), there exists an $\mathscr{L}$-measurable function $u_\lambda : [t, \infty) \to \mathbb{R}^m$ such that $z_\lambda(s) \geq L(s, x_\lambda(s), u_\lambda(s))$, $\dot{x}_\lambda(s) = f(s, x_\lambda(s), u_\lambda(s))$ and $u_\lambda(s) \in U(s)$ a.e. $s \in [t, \infty)$. We thus have:

$$V(t, \lambda x + (1 - \lambda)y) \leq \int_t^\infty L(s, x_\lambda(s), u_\lambda(s))ds \leq \int_t^\infty z_\lambda(s)ds$$
$$= \lambda \int_t^\infty L(s, x(s), u(s))ds + (1 - \lambda) \int_t^\infty L(s, x'(s), u'(s))ds.$$

Taking the infimum of the admissible processes $(x(\cdot), u(\cdot))$ on $[t, \infty)$ with $x(t) = x$ and $(x'(\cdot), u'(\cdot))$ with $x'(t) = y$ in the right-hand side of the inequality yields:

$$V(t, \lambda x + (1-\lambda)y) \leq \lambda V(t, x) + (1-\lambda)V(t, y),$$

from which the convexity of $V(t, \cdot\,)$ follows. □

A.3 Concavity of the Hamiltonian

The concavity of the Hamiltonian is subtler than the convexity of the value function. Specifically, Hypothesis A.2, guaranteeing the convexity of the value function $V(t, x)$ in x, is insufficient to establish the concavity of the Hamiltonian $H(t, x, p)$ in x.

Note that, by (4.6), for every $(t, x, p) \in [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$:

$$H(t, x, p) = \sup_{v \in \mathbb{R}^n} \{\langle p, v\rangle - F(t, x, v)\}.$$

Thus, $H(t, x, p)$ is concave in x if $F(t, x, v)$ is convex in (x, v). As shown by Feinstein and Luenberger [20], the following hypothesis is sufficient for $F(t, \cdot\,, \cdot\,)$ to be a convex function on $\Omega(t) \times \mathbb{R}^n$ for every $t \in [0, \infty)$, from which the concavity of the Hamiltonian follows.

Hypothesis A.3. (i) $\Omega(t) \times U(t)$ is convex for every $t \in [0, \infty)$.

(ii) $L(t, \cdot\,, \cdot\,)$ is convex on $\Omega(t) \times U(t)$ for every $t \in [0, \infty)$ and $L(t, x, \cdot\,)$ is nondecreasing on $U(t)$ for every $(t, x) \in \Omega$.

(iii) $f(t, \cdot\,, \cdot) : \Omega(t) \times U(t) \to \mathbb{R}^n$ is concave for every $t \in [0, \infty)$.

(iv) $f(t, \cdot\,, U(t)) : \Omega(t) \rightrightarrows \mathbb{R}^n$ has the convex graph for every $t \in [0, \infty)$.

(v) For every $v \in f(t, x, U(t))$ and $u \in U(t)$ with $v \leq f(t, x, u)$ and $x \in \Omega(t)$, there exists some $u' \in U(t)$ such that $u' \leq u$ and $v = f(t, x, u')$.

Theorem A.4. *$H(t, \cdot\,, p)$ is concave on $\mathbb{R}^n$ for every $(t, p) \in [0, \infty) \times \mathbb{R}^n$ if Hypothesis A.3 is satisfied.*

Note also that the conditions (i) to (iii) and (v) of the hypothesis imply Hypothesis A.2 and, thus, the convexity of the value function.

References

[1] Aliprantis, C. D. and K. C. Border, (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd edn., Berlin, Springer-Verlag.

[2] Araujo, A. and J. A. Scheinkman, (1983). "Maximum principle and transversality condition for concave infinite horizon economic models", *Journal of Economic Theory*, vol. 30, pp. 1–16.

[3] Aseev, S. M. and A. Kryaziimskiy, (2004). "The Pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons", *SIAM Journal on Control and Optimization*, vol. 43, pp. 1094–1119.

[4] Aubin, J.-P. and F. H. Clarke, (1979). "Shadow prices and duality for a class of optimal control problems", *SIAM Journal on Control and Optimization*, vol. 17, pp. 567–586.

[5] Aubin, J.-P. and H. Frankowska, (1990). *Set-Valued Analysis*, Boston, Birkhäuser.

[6] Balder, E. J., (1983). "An existence result for optimal economic growth problems", *Journal of Mathematical Analysis and Applications*, vol. 95, pp. 195–213.

[7] Barron, E. N. and R. Jensen, (1986). "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.

[8] Bates, G. R., (1978). "Lower closure existence theorems for optimal control problems with infinite horizon", *Journal of Optimization Theory and Applications*, vol. 24, pp. 639–649.

[9] Baum, R. F., (1976). "Existence theorems for Lagrange control problems with unbounded time domain", *Journal of Optimization Theory and Applications*, vol. 19, pp. 89–116.

[10] Becker, R. A. and J. H. Boyd, III, (1992). "Recursive utility and optimal capital accumulation. II. Sensitivity and duality theory", *Economic Theory*, vol. 2, pp. 547–563.

[11] Bell, M. L., R. W. H. Sargent and R. B. Vinter, (1997). "Existence of optimal controls for continuous time infinite horizon problems", *International Journal of Control*, vol. 68, pp. 887–896.

[12] Benveniste, L. M. and J. A. Scheinkman, (1982). "Duality theory for dynamic optimization models of economics: The continuous time case", *Journal of Economic Theory*, vol. 27, pp. 1–19.

[13] Cannarsa, P. and H. Frankowska, (1991). "Some characterization of optimal trajectories in control theory", *SIAM Journal on Control and Optimization*, vol. 29, pp. 1322–1347.

[14] Cesari, L., (1983). *Optimization–Theory and Applications: Problems with Ordinary Differential Equations*, Berlin, Springer-Verlag.

[15] Clarke, F. H., (1983). *Optimization and Nonsmooth Analysis*, New York, John Wiley & Sons.

[16] Clarke, F. H., (1989). *Methods of Dynamic and Nonsmooth Optimization*, Philadelphia, SIAM.

[17] Clarke, F. H. and R. B. Vinter, (1987). "The relationship between the maximum principle and dynamic programming", *SIAM Journal on Control and Optimization*, vol. 25, pp. 1291–1311.

[18] Crandall, M. G. and P. L. Lions, (1983). "Viscosity solutions of Hamilton–Jacobi equations", *Transactions of the American Mathematical Society*, vol. 277, pp. 1–42.

[19] Crandall, M. G., L. C. Evans and P. L. Lions, (1983). "Some properties of viscosity solutions of Hamilton–Jacobi equations", *Transactions of the American Mathematical Society*, vol. 282, pp. 487–502.

[20] Feinstein, C. D. and D. G. Luenberger, (1981). "Analysis of the asymptotic behavior of optimal control trajectories: The implicit programming problem", *SIAM Journal on Control and Optimization*, vol. 19, pp. 561–585.

[21] Frankowska, H., (1989). "Optimal trajectories associated with a solution of the contingent Hamilton–Jacobi equation", *Applied Mathematics and Optimization*, vol. 19, pp. 291–311.

[22] Frankowska, H., (1989). "Hamilton–Jacobi equations: Viscosity solutions and generalized gradients", *Journal of Mathematical Analysis and Applications*, vol. 141, pp. 21–26.

[23] Halkin, H., (1974). "Necessary conditions for optimal control problems with infinite horizon", *Econometrica*, vol. 42, pp. 267–272.

[24] Kamihigashi, T., (2001). “Necessity of transversality conditions for infinite horizon problems”, *Econometrica*, vol. 69, pp. 995–1012.

[25] Mangasarian, O. L., (1966). “Sufficient conditions for the optimal control of nonlinear systems”, *SIAM Journal on Control and Optimization*, vol. 4, pp. 139–151.

[26] Michel, P., (1982). “On the transversality condition in infinite horizon optimal problems”, *Econometrica*, vol. 50, pp. 975–984.

[27] Mirică, S., (1992). “A proof of Pontryagin’s minimum principle using dynamic programming”, *Journal of Mathematical Analysis and Applications*, vol. 170, pp. 501–512.

[28] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mischenko, (1962). *The Mathematical Theory of Optimal Processes*, New York, John Wiley & Sons.

[29] Rockafellar, R. T., (1970). “Conjugate convex functions in optimal control and the calculus of variations”, *Journal of Mathematical Analysis and Applications*, vol. 32, pp. 174–222.

[30] Rockafellar, R. T., (1973). “Optimal arcs and the minimum value function in problems of Lagrange”, *Transaction of the American Mathematical Society*, vol. 180, pp. 53–83.

[31] Rockafellar, R. T., (1975). “Existence theorems for generalized control problems of Bolza and Lagrange”, *Advances in Mathematics*, vol. 15, pp. 312–333.

[32] Seierstadt, A. and K. Sydsæter, (1977). “Sufficient conditions in optimal control theory”, *International Economic Review*, vol. 18, pp. 367–391.

[33] Shell, K., (1969). “Applications of Pontryagin’s maximum principle to economics”, in H. Kuhn and G.P. Szegö, (eds.), *Mathematical Systems Theory and Economics I*, Springer–Verlag, Berlin, pp. 241–292.

[34] Takekuma, S.-I., (1982). “Support price theorem for the continuous time model of capital accumulation”, *Econometrica*, vol. 50, pp. 427–442.

[35] Takekuma, S.-I., (1984). “On duality theory for the continuous time model of capital accumulation”, *Hitotsubashi Journal of Economics*, vol. 25, pp. 145–154.

[36] Vinter, R. B., (1988). “New results on the relationship between dynamic programming and the maximum principle”, *Mathematics of Control, Signals, and Systems*, vol. 1, pp. 97–105.

[37] Vinter, R. B., (2000). *Optimal Control*, Boston, Birkhäuser.

[38] Ye, J. J., (1993). “Nonsmooth maximum principle for infinite-horizon problems”, *Journal of Optimization Theory and Applications*, vol. 76, pp. 485–500.

[39] Zeidan, V., (1984). “A modified Hamilton–Jacobi approach in the generalized problem of Bolza”, *Applied Mathematics and Optimization*, vol. 11, pp. 97–109.

[40] Zeidan, V., (1984). “First and second order sufficient conditions for optimal control and the calculus of variations”, *Applied Mathematics and Optimization*, vol. 11, pp. 209–226.

[41] Zhou, X. Y., (1990). “Maximum principle, dynamic programming, and their connection in deterministic control”, *Journal of Optimization Theory and Applications*, vol. 65, pp. 363–373.

[42] Zhou, X. Y., (1993). “Verification theorems within the framework of viscosity solutions”, *Journal of Mathematical Analysis and Applications*, vol. 177, pp. 208–225.