HAL open science

Value Functions and Optimality Conditions for Nonconvex Variational Problems with an Infinite Horizon in Banach Spaces

Hélène Frankowska, Nobusumi Sagara

To cite this version:

Hélène Frankowska, Nobusumi Sagara. Value Functions and Optimality Conditions for Nonconvex Variational Problems with an Infinite Horizon in Banach Spaces. Mathematics of Operations Research, In press. hal-03005905

HAL Id: hal-03005905

https://hal.science/hal-03005905v1

Submitted on 14 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Value Functions and Optimality Conditions for Nonconvex Variational Problems with an Infinite Horizon in Banach Spaces*

Hélène Frankowska†
CNRS Institut de Mathématiques de Jussieu – Paris Rive Gauche
Sorbonne Université, Campus Pierre et Marie Curie
Case 247, 4 Place Jussieu, 75252 Paris, France
e-mail: helene.frankowska@imj-prg.fr

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

February 10, 2020

*The authors are grateful to the two anonymous referees for their helpful comments and suggestions on the earlier version of this manuscript.

†The research of this author was supported by the Air Force Office of Scientific Research, USA under award number FA9550-18-1-0254. It also partially benefited from the FJMH Program PGMO and from the support to this program from EDF-THALES-ORANGE-CRITEO under grant PGMO 2018-0047H.

‡The research of this author benefited from the support of JSPS KAKENHI Grant Number JP18K01518 from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Abstract

We investigate the value function of an infinite horizon variational problem in the infinite-dimensional setting. Firstly, we provide an upper estimate of its Dini–Hadamard subdifferential in terms of the Clarke subdifferential of the Lipschitz continuous integrand and the Clarke normal cone to the graph of the set-valued mapping describing dynamics. Secondly, we derive a necessary condition for optimality in the form of an adjoint inclusion that grasps a connection between the Euler–Lagrange condition and the maximum principle.

Key Words: Infinite horizon, Dini–Hadamard subdifferential, Gelfand integral, differentiability of the value function, Euler–Lagrange condition, maximum principle, spatial Ramsey growth model.

MSC2010 Subject Classification: Primary: 34A60, 49J50, 49J52; Secondary: 49J53, 49K15, 90C39

OR/MS Subject Classification: Dynamic programming/optimal control: Applications; Deterministic; Programming: Nondifferentiable

1 Introduction

Optimal control and dynamic programming are instrumental cornerstones of modern economic growth theory originated in Ramsey [46]. In the general reduced model of capital accumulation, necessary (and sufficient) conditions for optimality are employed under the convexity assumptions on utility functions and technologies for the investigation of the existence of competitive equilibria and support prices; see Benveniste and Scheinkman [8], Magill [44], Takekuma [52, 53]. Such well-behaving properties are prominent in convex problems of optimal control explored in the classical work by Rockafeller [48] with the full power of duality theory in convex analysis. In particular, one of the advantages in convex economic models lies in the crucial observation that the differentiability of the value function is guaranteed under the smoothness assumptions on the data; see Benveniste and Scheinkman [7, 8], Bonnisseau and Le Van [9], Rincón-Zapatero and Santos [47], Takekuma [53].

On the contrary, the absence of convexity and smoothness are two major sources of complex economic dynamics in continuous time as illustrated in Askenazy and Le Van [3], Davidson and Harris [25], Hartl and Kort [37], Skiba [51], Wagener [55]. More to the point, the difficulty with the lack of convexity assumptions results in the failure of differentiability of the value function even if the underlying data are smooth. Without convexity, one can expect at best the Lipschitz continuity of the value function even for smooth problems. This causes problems with expressing optimality conditions in many nonconvex economic growth models when one attempts to apply the

Hamilton–Jacobi–Bellman (HJB) equation. Recall that the value function is its unique solution whenever it is smooth.

The well-known failure of differentiability of the value function has stimulated two alternative approaches in optimal control theory. One is the application of a "generalized" subdifferential calculus along the lines of Clarke [20], which eventually leads to the formulation of a relation between the maximum principle and dynamic programming whenever the value function is locally Lipschitz continuous; see Clarke and Vinter [21, 22]. The other independent development is the concept of "viscosity solutions" to the HJB equation initiated by Lions [43] (see also Crandall, Evans and Lions [23], Crandall and Lions [24]), which makes use of the notion of Fréchet super- and subdifferentials to claim that the value function is the unique viscosity solution of the HJB equation. For the connections between the maximum principle and the superdifferentials of the value function, see Frankowska [31, 34].

With this background in mind, we investigate the value function of an infinite horizon variational problem in the setting of an infinite-dimensional generalized control system. Our primary concern here is to go beyond convexity, smoothness, and finite dimensionality aiming the possible applications to dynamic optimization in economic theory. Since the optimal economic growth models are identified with a specific form of the general equilibrium model with single representable consumer and firm, we can deal with a rich class of commodity spaces for capital stock, which appears as a Sobolev space. In particular, spatial Ramsey growth models involve a location of each agent along the lines of Hotelling [38], in which infinite-dimensional commodity spaces naturally arise; see Boucekkine et al. [10, 11], Brito [12], Brock et al. [13]. Applying our general result, we obtain another necessary condition for optimality in spatial Ramsey growth models.

The purpose of this paper is twofold. Firstly, we provide an upper estimate of the Dini–Hadamard subdifferential of the value function in terms of the Clarke subdifferential of the Lipschitz continuous integrand and the Clarke normal cone to the set-valued mapping describing dynamics. As a result, we obtain the strict differentiability of the value function under the Fréchet differentiability of the integrand, which removes completely the convexity assumptions of the earlier works by Benveniste and Scheinkman [7, 8], Bonnisseau and Le Van [9], Rincón-Zapatero and Santos [47], Takekuma [53]. For the (sub)differentiability of the value function in the context of finite dimensional control systems with a finite horizon, see the lecture notes Frankowska [34].

Secondly, under an interiority assumption we derive a necessary condition for optimality in the form of an adjoint inclusion that grasps a connection between the Euler–Lagrange condition and the maximum principle. Our interiority assumption is weaker than those in Benveniste and Scheinkman [7, 8], Bonnisseau and Le Van [9], Takekuma [52, 53]. On the other hand, when dynamics are described by a control system, such interiority assump-

tion may be omitted. To deal with the adjoint variable in dual spaces, we introduce the Gelfand integrals of the Dini–Hadamard and Clarke subdifferential mappings, which is a new feature that does not arise in the context of finite-dimensional control systems.

For the finite-dimensional control systems, necessary conditions with or without convexity assumptions using limiting subdifferentials were obtained in Ioffe [40], Vinter and Zheng [54] in the finite horizon setting. The ones in the infinite horizon setting using Dini–Hadamard, Clarke, and limiting subdifferentials were derived in Aubin and Clarke [4], Cannarsa and Frankowska [16], Sagara [49], Ye [56]. For control systems in Hilbert spaces, a necessary condition under the convexity assumptions was obtained in Barbu [6] in the infinite horizon setting. For Banach spaces and semilinear control systems, the necessary and sufficient conditions were derived in Cannarsa and Frankowska [15] in the finite horizon setting when the set of velocities is convex.

The organization of the paper is as follows. Section 2 collects preliminary results on subdifferential calculus on Banach spaces. In Section 3 we formulate the nonconvex variational problem under investigation with the standing hypotheses and demonstrate the Lipschitz continuity and subdifferentiability of the value function. We derive in Section 4 necessary conditions for the variational and optimal control problems. Section 5 applies our main result to spatial Ramsey growth models. Appendices I and II discuss Gelfand integral of multifunctions and the Gelfand integrability of the Dini–Hadamard and Clarke subdifferential mappings, and the proofs of auxiliary results and lemmas needed to obtain the main results.

2 Preliminaries

Let $(E, \|\cdot\|)$ be a real Banach space with the dual system $\langle E^*, E\rangle$, where E^* is the norm dual of E. A real-valued function $\varphi : E \to \mathbb{R}$ is said to be *Gateaux differentiable* at $\bar{x} \in E$ if there exists an element $\nabla\varphi(\bar{x}) \in E^*$ such that

$$\lim_{\theta\to 0} \frac{\varphi(\bar{x} + \theta v) - \varphi(\bar{x})}{\theta} = \langle \nabla\varphi(\bar{x}), v\rangle \tag{2.1}$$

for every $v \in E$; $\nabla\varphi(\bar{x})$ is called the *Gateaux derivative* of φ at $\bar{x}$. If the convergence in (2.1) is uniform in $v \in C$ for every bounded subset C of E, then φ is said to be *Fréchet differentiable* at $\bar{x}$ and $\nabla\varphi(\bar{x})$ is called the *Fréchet derivative* of φ at $\bar{x}$. A function φ is said to be *strictly differentiable* at $\bar{x}$ if there exists $\nabla\varphi(\bar{x}) \in E^*$ such that

$$\lim_{\substack{x\to\bar{x}\\ \theta\to 0}} \frac{\varphi(x + \theta v) - \varphi(x)}{\theta} = \langle \nabla\varphi(\bar{x}), v\rangle \tag{2.2}$$

and the convergence in (2.2) is uniform in $v \in C$ for every compact subset C of E. Then $\nabla\varphi(\bar{x})$ is called the *strict derivative* of φ at $\bar{x}$. If φ is strictly

differentiable at $\bar{x}$, then φ is Lipschitz near $\bar{x}$; see Clarke [20, Proposition 2.2.1]. A function φ is said to be *continuously differentiable* at $\bar{x}$ if φ is Gateaux differentiable at every x in a neighborhood O of $\bar{x}$ and the mapping $x \mapsto \nabla\varphi(x)$ is continuous from O to E^*; φ is called a C^1*-function* on E if φ is continuously differentiable at any point in E. If φ is continuously differentiable at $\bar{x}$, then φ is strictly differentiable at $\bar{x}$; see Clarke [20, Corollary, p. 32]. A norm $\|\cdot\|$ on a Banach space E is said to be Gateaux (resp. Fréchet) differentiable if $\|\cdot\|$ is Gateaux (resp. Fréchet) differentiable on the open set $E \setminus \{0\}$.

The *support function* $s(\cdot, C) : E^* \to \mathbb{R} \cup \{+\infty\}$ of a nonempty subset C of E is given by $s(x^*, C) = \sup_{x \in C} \langle x^*, x \rangle$. The *polar* C^0 of C is the set $C^0 = \{x^* \in E^* \mid s(x^*, C) \leq 0\}$. The support function $s : (\cdot, K) : E \to \mathbb{R} \cup \{+\infty\}$ of a nonempty subset K of E^* is defined by $s(x, K) = \sup_{x^* \in K} \langle x^*, x \rangle$. The polar K^0 of K is the set $K^0 = \{x \in E \mid s(x, K) \leq 0\}$.

Let $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function on E. The *effective domain* of φ is the set of points where φ is finite and is denoted by $\operatorname{dom} \varphi := \{x \in E \mid \varphi(x) < +\infty\}$. If φ is Lipschitz near $\bar{x} \in \operatorname{dom} \varphi$, then its *Clarke directional derivative* at $\bar{x}$ in the direction $v \in E$ is defined by

$$\varphi^\circ(\bar{x}; v) := \limsup_{\substack{x \to \bar{x} \\ \theta \downarrow 0}} \frac{\varphi(x + \theta v) - \varphi(x)}{\theta}$$

and the *Clarke subdifferential* of φ at $\bar{x}$ is defined by

$$\partial^\circ \varphi(\bar{x}) := \{x^* \in E^* \mid \langle x^*, v \rangle \leq \varphi^\circ(\bar{x}; v) \ \forall v \in E\}.$$

Since the function $v \mapsto \varphi^\circ(\bar{x}; v)$ is positively homogeneous and subadditive, the set $\partial^\circ \varphi(\bar{x})$ is nonempty by the Hahn–Banach theorem, weakly*compact, and convex in E^*. Furthermore, the Clarke directional derivative is the support function of the Clarke subdifferential

$$\varphi^\circ(\bar{x}; v) = s(v, \partial^\circ \varphi(\bar{x}))$$

for every $v \in E$; see Clarke [20, Propositions 2.1.1 and 2.1.2]. Recall that the function φ that is Lipschitz near $\bar{x}$ is said to be *regular* at $\bar{x}$ if the classical directional derivative

$$\varphi'(\bar{x}; v) := \lim_{\theta \downarrow 0} \frac{\varphi(\bar{x} + \theta v) - \varphi(\bar{x})}{\theta}$$

exists and $\varphi'(\bar{x}; v) = \varphi^\circ(\bar{x}; v)$ for every $v \in E$.

Let $d_C : E \to \mathbb{R}$ be the *distance function* from a nonempty subset C of E defined by $d_C(x) := \inf_{\xi \in C} \|x - \xi\|$. Then d_C is nonexpansive (i.e., Lipschitz of rank one) on E. Let $\bar{x}$ be a point in C. A vector $v \in E$ is called a *tangent* to C at $\bar{x}$ if $d_C^\circ(\bar{x}; v) = 0$. The set of all tangents to C at $\bar{x}$ is called the *Clarke tangent cone* to C at $\bar{x}$ and is denoted by

$$T_C(\bar{x}) := \{v \in E \mid d_C^\circ(\bar{x}; v) = 0\}.$$

Then $T_C(\bar{x})$ is a closed convex cone because $v \mapsto d_C^\circ(\bar{x}; v)$ is nonnegative, positively homogeneous, and continuous. An intrinsic characterization of $T_C(\bar{x})$ that is independent of the use of a distance function is as follows: $v \in T_C(\bar{x})$ if and only if for every sequence $\{x_n\}_{n\in\mathbb{N}}$ in C with $x_n \to \bar{x}$ and every sequence $\{\theta_n\}_{n\in\mathbb{N}}$ of positive real numbers with $\theta_n \downarrow 0$, there is a sequence $\{v_n\}_{n\in\mathbb{N}}$ in E with $v_n \to v$ such that $x_n + \theta_n v_n \in C$ for each $n \in \mathbb{N}$; see Clarke [20, Theorem 2.4.5]. Let B be the open unit ball in E. Define the *contingent cone* $K_C(\bar{x})$ of tangents to C at $\bar{x}$ by

$$K_C(\bar{x}) := \{v \in E \mid \forall \varepsilon > 0\, \exists \theta \in (0, \varepsilon)\, \exists w \in v + \varepsilon B : \bar{x} + \theta w \in C\}\,.$$

Then $v \in K_C(\bar{x})$ if and only if there exist a sequence $\{\theta_n\}_{n\in\mathbb{N}}$ of positive real numbers with $\theta_n \downarrow 0$ and a sequence $\{v_n\}_{n\in\mathbb{N}}$ in E with $v_n \to v$ such that $\bar{x} + \theta_n v_n \in C$ for each $n \in \mathbb{N}$. It is evident that $T_C(\bar{x}) \subset K_C(\bar{x})$, but $K_C(\bar{x})$ is not necessarily convex. The set C is said to be *regular* at $\bar{x}$ if $T_C(\bar{x}) = K_C(\bar{x})$. The polar of $T_C(\bar{x})$ is called the *Clarke normal cone* to C at $\bar{x}$, which is given by

$$N_C(\bar{x}) = \{x^* \in E^* \mid \langle x^*, v\rangle \le 0\ \forall v \in T_C(\bar{x})\}.$$

The Clarke normal cone is characterized by $N_C(\bar{x}) = w^*\text{-cl}\{\bigcup_{\lambda\ge 0} \lambda \partial^\circ d_C(\bar{x})\}$ (see Clarke [20, Proposition 2.4.2]), where the right-hand side of the above equality means the weak* closure of the set. It follows from the bipolar theorem (see Aubin and Frankowska [5, Theorem 2.4.3]) that $T_C(\bar{x})$ is the polar of $N_C(\bar{x})$, i.e., $T_C(\bar{x}) = \{v \in E \mid \langle x^*, v\rangle \le 0\ \forall x^* \in N_C(\bar{x})\}$. Denote by $\operatorname{epi}\varphi = \{(x, r) \in E \times \mathbb{R} \mid \varphi(x) \le r\}$ the *epigraph* of φ. If φ is Lipschitz near $\bar{x} \in \operatorname{dom}\varphi$, then $T_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x})) = \operatorname{epi}\varphi^\circ(\bar{x}; \cdot)$ (see Aubin and Frankowska [5, Theorem 2.4.9]), and hence, $\varphi^\circ(\bar{x}; v) = \inf\{r \in \mathbb{R} \mid (v, r) \in T_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x}))\}$. Therefore, if φ is Lipschitz near $\bar{x}$, then

$$\partial^\circ\varphi(\bar{x}) = \{x^* \in E^* \mid (x^*, -1) \in N_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x}))\}.$$

The *lower directional derivative* (or *contingent*, or *Dini–Hadamard directional subderivative*) of $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ at $\bar{x} \in \operatorname{dom}\varphi$ in the direction $v \in E$ is defined by

$$\varphi^-(\bar{x}; v) := \liminf_{\substack{u\to v\\ \theta\downarrow 0}} \frac{\varphi(\bar{x} + \theta u) - \varphi(\bar{x})}{\theta} \in \mathbb{R} \cup \{\pm\infty\}$$

and the *upper directional derivative* (or *Dini–Hadamard directional superderivative*) of φ at $\bar{x}$ in the direction $v \in E$ is defined by

$$\varphi^+(\bar{x}; v) := \limsup_{\substack{u\to v\\ \theta\downarrow 0}} \frac{\varphi(\bar{x} + \theta u) - \varphi(\bar{x})}{\theta} \in \mathbb{R} \cup \{\pm\infty\}.$$

The *Dini–Hadamard subdifferential* of φ at $\bar{x}$ is defined by

$$\partial^-\varphi(\bar{x}) := \{x^* \in E^* \mid \langle x^*, v\rangle \le \varphi^-(\bar{x}; v)\ \forall v \in E\}$$

and the *Dini–Hadamard superdifferential* of φ at $\bar{x}$ is defined by

$$\partial^+\varphi(\bar{x}) := \{x^* \in E^* \mid \langle x^*, v\rangle \geq \varphi^+(\bar{x}; v)\ \forall v \in E\}.$$

Because of the plus-minus symmetry with $\varphi^-(x; v) = -(-\varphi)^+(x; v)$ and $\partial^-\varphi(\bar{x}) = -\partial^+(-\varphi)(\bar{x})$, it is enough to investigate lower directional derivatives and Dini–Hadamard subdifferentials in what follows. Since

$$K_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x})) = \operatorname{epi}\varphi^-(\bar{x}; \cdot)$$

(see Aubin and Frankowska [5, Propositions 6.1.3 and 6.1.4]),

$$\varphi^-(\bar{x}; v) = \inf\{r \in \mathbb{R} \mid (v, r) \in K_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x}))\} \in \mathbb{R} \cup \{\pm\infty\}$$

with the convention that $\inf \emptyset = +\infty$. Therefore, if $\bar{x} \in \operatorname{dom}\varphi$, then

$$\partial^-\varphi(\bar{x}) = \{x^* \in E^* \mid (x^*, -1) \in K_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x}))^0\}.$$

Unlike Clarke directional derivatives, the lower directional derivative mapping $v \mapsto \varphi^-(\bar{x}; v)$ fails to be convex although it is positively homogeneous. Thus, except for a smooth or a convex function φ, it is rather typical that $\partial^-\varphi(\bar{x})$ is empty at some points for a lower semicontinuous or even a locally Lipschitz function. Note that $\partial^-\varphi(\bar{x})$ is weakly* closed and convex. If φ is locally Lipschitz, then $\varphi^-(\bar{x}, v) \leq \varphi^\circ(\bar{x}; v)$ for every $v \in E$, and hence, $\partial^-\varphi(\bar{x}) \subset \partial^\circ\varphi(\bar{x})$. In particular, if φ is also regular at $\bar{x}$, then $\varphi^-(\bar{x}, v) = \varphi^\circ(\bar{x}; v)$ for every $v \in E$ and $\partial^-\varphi(\bar{x}) = \partial^\circ\varphi(\bar{x})$. Note also that if φ has the strict derivative $\nabla\varphi(\bar{x})$ at $\bar{x} \in E$, then $\partial^-\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$.

A generic existence of Dini–Hadamard subdifferentials is assured in the following result.

Theorem 2.1 (Ioffe [39, 41])**.** *Let E be a Banach space admitting an equivalent Gateaux differentiable norm and $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. Then the set $\{x \in E \mid \partial^-\varphi(x) \neq \emptyset\}$ is dense in* $\operatorname{dom}\varphi$.

We recall that any separable Banach space has an equivalent Gateaux differentiable norm; see Fabian et al. [28, Theorem 8.2].

The next extends the well-known representation of the normal cone to the set determined by the inequality constraint.

Proposition 2.1. *Let $\varphi_i : E \to \mathbb{R}$, $i = 1, 2, \ldots, m$, be continuous real-valued functions and*

$$C := \{x \in E \mid \varphi_i(x) \leq 0,\ i = 1, 2, \ldots, m\}.$$

Define the active constraint indices at $\bar{x} \in C$ by $I(\bar{x}) := \{i \in \{1, 2, \ldots, m\} \mid \varphi_i(\bar{x}) = 0\}$. If φ_i is strictly differentiable at $\bar{x}$ for each $i \in I(\bar{x})$ and the constraint qualification $0 \notin \operatorname{co}\{\nabla\varphi_i(\bar{x}) \mid i \in I(\bar{x})\}$ is satisfied, then

$$N_C(\bar{x}) = \left\{ \sum_{i \in I(\bar{x})} \lambda_i \nabla\varphi_i(\bar{x}) \in E^* \mid \lambda_i \geq 0\ \forall i \in I(\bar{x}) \right\}.$$

The proof is provided in Subsection A.1 because we could not find this result in the literature for arbitrary Banach spaces.

3 Value Functions for an Infinite Horizon Problem

3.1 Nonconvex Variational Problems

Denote by $\mathbb{R}_+ = [0, \infty)$ the unbounded interval of the real line with the Lebesgue measure and the Lebesgue σ-algebra $\mathcal{L}$. A function $x : \mathbb{R}_+ \to E$ is said to be *simple* if there exist $x_1, x_2, \ldots, x_n \in E$ and $I_1, I_2, \ldots, I_n \in \mathcal{L}$ such that $x(\cdot) = \sum_{i=1}^n x_i \chi_{I_i}$, where $\chi_{I_i}(t) = 1$ if $t \in I_i$ and $\chi_{I_i}(t) = 0$ otherwise. A function $x(\cdot)$ is said to be *strongly measurable* if there exists a sequence of simple functions $\{x_n(\cdot)\}_{n\in\mathbb{N}}$ from $\mathbb{R}_+$ to E such that $\|x_n(t) - x(t)\| \to 0$ a.e. $t \in \mathbb{R}_+$. A strongly measurable function $x(\cdot)$ is *locally Bochner integrable* if it is Bochner integrable on every compact subset I of $\mathbb{R}_+$, that is, $\int_I \|x(t)\| dt < \infty$, where the *Bochner integral* of $x(\cdot)$ over I is defined by $\int_I x(t)dt := \lim_n \int_I x_n(t)dt$. Let $L^1_{\text{loc}}(\mathbb{R}_+, E)$ be the space of (the equivalence classes of) locally Bochner integrable functions from $\mathbb{R}_+$ to E.

A function $x(\cdot) : \mathbb{R}_+ \to E$ is said to be *strongly differentiable* at $t > 0$ if there exists $v \in E$ such that

$$\lim_{h\to 0} \frac{x(t+h) - x(t)}{h} = v.$$

The vector v is denoted by $\dot{x}(t)$ and called the *strong derivative* of x at t. Denote by $W^{1,1}_{\text{loc}}(\mathbb{R}_+, E)$ the Sobolev space, which consists of locally Bochner integrable functions $x : \mathbb{R}_+ \to E$ whose strong derivative $\dot{x}(t)$ exists a.e. $t \in \mathbb{R}_+ \setminus \{0\}$ with $\dot{x}(\cdot) \in L^1_{\text{loc}}(\mathbb{R}_+, E)$ and $x(t) = \int_0^t \dot{x}(s)ds + x(0)$ for every $t \in \mathbb{R}_+$. For each $n \in \mathbb{N}$, define the seminorm μ_n on $W^{1,1}_{\text{loc}}(\mathbb{R}_+, E)$ by $\mu_n(x(\cdot)) = \int_0^n (\|x(t)\| + \|\dot{x}(t)\|)dt$. Since $\{\mu_n\}_{n\in\mathbb{N}}$ is a countable separating family of seminorms, $W^{1,1}_{\text{loc}}(\mathbb{R}_+, E)$ is a Fréchet space under the compatible metric d given by

$$d(x(\cdot), y(\cdot)) = \max_{n\in\mathbb{N}} \frac{\mu_n(x(\cdot) - y(\cdot))}{2^n(1 + \mu_n(x(\cdot) - y(\cdot)))}, \quad x(\cdot), y(\cdot) \in W^{1,1}_{\text{loc}}(\mathbb{R}_+, E).$$

An element in $W^{1,1}_{\text{loc}}(\mathbb{R}_+, E)$ is called an *arc*. When $\mathbb{R}_+$ is replaced by a compact interval I of $\mathbb{R}_+$, the above definition simply leads to that of the Sobolev space $W^{1,1}(I, E)$ normed by $\|x(\cdot)\|_{1,1} = \int_I (\|x(t)\| + \|\dot{x}(t)\|)dt$.

Let $L : \mathbb{R}_+ \times E \times E \to \mathbb{R} \cup \{+\infty\}$ be an integrand. Given an arc $x(\cdot) \in W^{1,1}_{\text{loc}}(\mathbb{R}_+, E)$, the improper integral is defined by

$$\int_t^\infty L(s, x(s), \dot{x}(s))ds = \lim_{T\to\infty} \int_t^T L(s, x(s), \dot{x}(s))ds$$

for every $t \in \mathbb{R}_+$ provided that the above limit does exist. Let $\Gamma : \mathbb{R}_+ \times E \rightsquigarrow E$ be a multifunction. The variational problem under investigation is

to minimize the improper integral functional over the feasibility constraint governed by the differential inclusion:

$$\begin{aligned} &\inf_{x(\cdot)\in W^{1,1}_{\mathrm{loc}}([t,\infty),E)} \int_t^\infty L(s,x(s),\dot{x}(s))ds \\ &\text{s.t. } \dot{x}(s)\in\Gamma(s,x(s)) \text{ a.e. } s\in[t,\infty),\ x(t)=\xi. \end{aligned} \tag{P$_t$}$$

An arc satisfying the above differential inclusion is called an *admissible trajectory*. Define the set of admissible trajectories starting at time $t\in\mathbb{R}_+$ from a given initial condition $\xi\in E$ by

$$\mathcal{A}_{(t,\xi)} := \left\{x(\cdot)\in W^{1,1}_{\mathrm{loc}}([t,\infty),E) \mid \dot{x}(s)\in\Gamma(s,x(s)) \text{ a.e. } s\in[t,\infty),\ x(t)=\xi\right\}.$$

Then the *value function* $V:\mathbb{R}_+\times E\to\mathbb{R}\cup\{\pm\infty\}$ is defined by

$$V(t,\xi) := \inf_{x(\cdot)\in\mathcal{A}_{(t,\xi)}} \int_t^\infty L(s,x(s),\dot{x}(s))ds.$$

Here, we set $\inf\emptyset=+\infty$ if $\mathcal{A}_{(t,\xi)}$ is empty or if for every $x(\cdot)\in\mathcal{A}_{(t,\xi)}$ the integral $\int_t^\infty L(s,x(s),\dot{x}(s))ds$ is not well-defined. The *effective domain* of V is given by $\operatorname{dom}V=\{(t,x)\in\mathbb{R}_+\times E \mid V(t,x)<+\infty\}$; V is said to be *proper* if $\operatorname{dom}V$ is nonempty and $V(t,\xi)>-\infty$ for every $(t,\xi)\in\mathbb{R}_+\times E$. For every $(t,\xi)\in\operatorname{dom}V$, an admissible trajectory $x(\cdot)\in\mathcal{A}_{(t,\xi)}$ is said to be *optimal* for (P$_t$) if it satisfies $\int_t^\infty L(s,x(s),\dot{x}(s))ds=V(t,\xi)>-\infty$. For given $x\in E$, the multifunction $\Gamma(\cdot,x)\rightsquigarrow E$ is said to be *measurable* if the set $\{t\in\mathbb{R}_+ \mid \Gamma(t,x)\cap O\neq\emptyset\}$ belongs to $\mathcal{L}$ for every open subset O of E.

The standing hypothesis are described as follows.

(H$_1$) $\mathcal{A}_{(t,\xi)}$ is nonempty for every $(t,\xi)\in\mathbb{R}_+\times E$.

(H$_2$) $L(\cdot,x,y)$ is measurable for every $(x,y)\in E\times E$.

(H$_3$) There exist an integrable function $l_1:\mathbb{R}_+\to\mathbb{R}_+$ and a locally bounded, integrable function $l_2:\mathbb{R}_+\to\mathbb{R}_+$ such that

$$|L(t,0,0)|\le l_1(t)$$

and

$$|L(t,x,y)-L(t,x',y')|\le l_1(t)\|x-x'\|+l_2(t)\|y-y'\|$$

for every $t\in\mathbb{R}_+$ and every $(x,y),(x',y')\in E\times E$.

(H$_4$) Γ has nonempty closed values.

(H$_5$) $\Gamma(\cdot,x)$ is measurable for every $x\in E$.

($\mathbf{H_6}$) There exist a locally integrable function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\Gamma(t, 0) \subset \gamma(t)B$$

and

$$\Gamma(t, x) \subset \Gamma(t, x') + \gamma(t)\|x - x'\|B$$

for every $t \in \mathbb{R}_+$ and $x, x' \in E$.

($\mathbf{H_7}$) The Lipschitz modulus functions satisfy the integrability conditions:

$$\int_0^\infty \left[\exp\left(\int_0^s \gamma(\tau)d\tau\right)\left(1 + \int_0^s \gamma(\tau)d\tau\right)(l_1(s) + l_2(s)\gamma(s))\right] ds < \infty.$$

Since the integrand L is assumed to be a Carathéodory function in (H_2) and (H_3), it is jointly measurable on $\mathbb{R}_+ \times E \times E$ with respect to the product σ-algebra $\mathcal{L} \otimes \mathrm{Borel}(E, \|\cdot\|) \otimes \mathrm{Borel}(E, \|\cdot\|)$ whenever E is a separable Banach space; see Aubin and Frankowska [5, Lemma 8.2.6]. Hence, L is a normal integrand; see Appendix A.3 for the definition. Hypothesis (H_7) guarantees the integrable boundedness of $\{L(\cdot, x(\cdot), \dot{x}(\cdot)) \mid x(\cdot) \in \mathcal{A}_{(t,\xi)}\}$ over the interval $[t, \infty)$ for every $(t, \xi) \in \mathbb{R}_+ \times E$ and is needed to prove Theorem 3.1 below.

In the rest of the paper, E is assumed to be separable.

Theorem 3.1. *If* (H_1)–(H_7) *hold, then* V *is bounded and lower semicontinuous on* $\mathbb{R}_+ \times E$*, and* $V(t, \cdot)$ *is Lipschitz of rank* $k(t)$ *on* E *for every* $t \in \mathbb{R}_+$ *with a continuous decreasing function* $k : \mathbb{R}_+ \to \mathbb{R}_+$ *satisfying* $k(t) \to 0$ *as* $t \to \infty$.

The proof is deferred to Subsection B.1.

Remark 3.1. Since we impose the conditions which guarantee the integrability of the integrand on the set of admissible trajectories, the optimality criterion is unambiguous. When the integrability condition over the infinite horizon fails, there are several optimality criteria; see Carlson et al. [17], Seirstad and Sydsæter [50], Zaslavski [57]. For the derivation of the necessary condition under (weak) overtaking optimality in the finite-dimensional setting, see Halkin [36], Takekuma [52, 53].

3.2 Subdifferentials of the Value Function

In the following we always assume that optimal trajectories for (P_0) exist. To obtain an existence result in our framework, one needs standard convexity hypotheses. For the case with finite-dimensional control systems with an infinite horizon, see Cannarsa and Frankowska [16]. For the case with reflexive, separable Banach space valued semilinear control systems with a finite horizon, see Cannarsa and Frankowska [15].

Let us denote by $L_x^+(t,\bar{x},\bar{y};v)$ the upper partial directional derivative of $L(t,\cdot,\bar{y})$ at $\bar{x}\in E$ in the direction $v\in E$; $L_y^+(t,\bar{x},\bar{y};v)$ has an obvious meaning. Then $\partial_x^+ L(t,\bar{x},\bar{y})$ is the Dini–Hadamard partial superdifferential of $L(t,\cdot,\bar{y})$ at $\bar{x}$; $\partial_y^+ L(t,\bar{x},\bar{y})$ has a similar meaning. The Clarke partial directional derivatives $L_x^\circ(t,\bar{x},\bar{y};v)$ and $L_y^\circ(t,\bar{x},\bar{y};v)$, and the Clarke partial subdifferentials $\partial_x^\circ L(t,\bar{x},\bar{y})$ and $\partial_y^\circ L(t,\bar{x},\bar{y})$ are defined in a similar way. Recall that $d_{\Gamma(t,x)}:E\to\mathbb{R}$ is the distance function from the set $\Gamma(t,x)$ and denote by $N_{\Gamma(t,x)}(y)\subset E^*$ the Clarke normal cone to $\Gamma(t,x)$ at $y\in\Gamma(t,x)$.

We need another continuity assumption on Γ that replaces (H_5):

$(\mathbf{H_5'})$ $\Gamma(\cdot,x)$ is lower semicontinuous for every $x\in E$.

Our results below concern the subdifferentiability of the value function. We neither impose any convexity assumptions, nor request the interiority of the optimal trajectory. This improves results from Benveniste and Scheinkman [7, 8], Bonnisseau and Le Van [9], Rincón-Zapatero and Santos [47], Takekuma [53].

Theorem 3.2. *Let $x_0(\cdot)\in\mathcal{A}_{(0,\xi)}$ be an optimal trajectory for (P_0). If (H_1)–(H_4), (H_5'), (H_6), and (H_7) hold, then:*

(i) $V_x^-(t,x_0(t);\dot{x}_0(t)-v)\le L(t,x_0(t),v)-L(t,x_0(t),\dot{x}_0(t))$ *a.e.* $t\in\mathbb{R}_+$ *for every* $v\in\Gamma(t,x_0(t))$*;*

(ii) $-\partial_x^- V(t,x_0(t))\subset\partial_y^\circ L(t,x_0(t),\dot{x}_0(t))+N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ *a.e.* $t\in\mathbb{R}_+$.

Moreover, if $L(t,x_0(t),\cdot)$ is Gateaux differentiable at $\dot{x}_0(t)$, then:

$$-\partial_x^- V(t,x_0(t))\subset\nabla_y L(t,x_0(t),\dot{x}_0(t))+N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))\ \text{ a.e. } t\in\mathbb{R}_+.$$

Furthermore, if $V(t,\cdot)$ is regular at $x_0(t)$ and $\partial_x^- V(t,x_0(t))$ is a singleton, then $V(t,\cdot)$ is strictly differentiable at $x_0(t)$ with:

$$-\nabla_x V(t,x_0(t))=\nabla_y L(t,x_0(t),\dot{x}_0(t))+q(t)\quad \text{a.e. } t\in\mathbb{R}_+,$$

where $q:\mathbb{R}_+\to E^*$ is a Borel measurable selector from $N_{\Gamma(\cdot,x_0(\cdot))}(\dot{x}_0(\cdot)):\mathbb{R}_+\rightsquigarrow E^*$ with respect to the weak topology of E^*.*

For the proof, see Subsection B.2.

Example 3.1. Let $g_i:\mathbb{R}_+\times E\times E\to\mathbb{R}$, $i=1,2,\dots,m$, be Carathéodory functions and define the velocity multifunction $\Gamma:\mathbb{R}_+\times E\rightsquigarrow E$ by

$$\Gamma(t,x):=\{y\in E\mid g_i(t,x,y)\le 0,\ i=1,2,\dots,m\}.$$

Let $x_0(\cdot)\in W^{1,1}_{\mathrm{loc}}(\mathbb{R}_+,E)$ and

$$I(t):=\{i\in\{1,2,\dots,m\}\mid g_i(t,x_0(t),\dot{x}_0(t))=0\}$$

be the active constraint indices at $(t, x_0(t), \dot{x}_0(t)) \in \mathbb{R}_+ \times E \times E$. Assume that $g_i(t, x_0(t), \cdot)$ has the strict derivative at $\dot{x}_0(t)$ for each $i \in I(t)$ and the constraint qualification $0 \notin \mathrm{co}\{\nabla_y g_i(t, x_0(t), \dot{x}_0(t)) \mid i \in I(t)\}$ holds. It follows from Proposition 2.1 that:

$$N_{\Gamma(t,x_0(t))}(\dot{x}_0(t)) = \left\{ \sum_{i \in I(t)} \lambda_i \nabla_y g_i(t, x_0(t), \dot{x}_0(t)) \in E^* \mid \lambda_i \geq 0 \ \forall i \in I(t) \right\}.$$

Under the hypotheses of Theorem 3.2, we have

$$-\nabla_x V(t, x_0(t)) = \nabla_y L(t, x_0(t), \dot{x}_0(t)) + \sum_{i \in I(t)} \lambda_i(t) \nabla_y g_i(t, x_0(t), \dot{x}_0(t))$$

for some $\lambda_i(t) \geq 0$ with $i \in I(t)$. By the measurable selection theorem, the mapping $t \mapsto \lambda_i(t)$ can be chosen in a measurable way. Under the convexity hypothesis with the constraint qualification, Rincón-Zapatero and Santos [47] provided sufficient conditions for the differentiability of the value function with the finite-dimensional state constraint without the interiority conditions (H_8) below.

4 Euler–Lagrange Conditions and the Maximum Principle

4.1 Necessary Conditions under the Interiority Assumption

A function $p : \mathbb{R}_+ \to E^*$ is said to be *locally absolutely continuous* if its restriction to the bounded closed interval $[0, \tau]$ is absolutely continuous for every $\tau > 0$, i.e., for every $\tau > 0$ and $\varepsilon > 0$ there exists $\delta > 0$ such that $0 \leq t_1 < \tau_1 \leq t_2 < \tau_2 < \cdots \leq t_n < \tau_n \leq \tau$ and $\sum_{i=1}^n |t_i - \tau_i| < \delta$ imply $\sum_{i=1}^n \|p(t_i) - p(\tau_i)\| < \varepsilon$. A function $p(\cdot)$ is said to be *weakly* differentiable* at $t > 0$ if there exists $x^* \in E^*$ such that

$$\lim_{h \to 0} \left\langle \frac{p(t+h) - p(t)}{h}, x \right\rangle = \langle x^*, x \rangle \quad \text{for every } x \in E.$$

Then vector x^* is called the *weak* derivative* of p at t and is denoted by $\dot{p}(t)$ with $d\langle p(t), x\rangle/dt = \langle \dot{p}(t), x\rangle$ for every $x \in E$.

We impose a feasibility assumption on the perturbation around a specific optimal trajectory $x_0(\cdot) \in \mathcal{A}_{(0,\xi)}$ as follows.

($\mathbf{H_8}$) For every $T > 0$ there exists $\eta > 0$ such that

$$(x_0(t) + \eta B, \dot{x}_0(t)) \subset \mathrm{gph}\, \Gamma(t, \cdot) \quad \text{a.e. } t \in [0, T].$$

This is equivalent to saying that for every $T > 0$ there exists $\eta > 0$ such that a.e. $t \in [0, T]$ we have $\dot{x}_0(t) \in \Gamma(t, x_0(t) + \eta v)$ for every $v \in B$. (H_8) is a weaker condition than the interiority condition imposed in Benveniste and Scheinkman [8]:

($\mathbf{H'_8}$) There exists $\eta > 0$ such that

$$(x_0(t) + \eta B, \dot{x}_0(t)) \subset \operatorname{gph} \Gamma(t, \cdot) \quad \text{a.e. } t \in \mathbb{R}_+.$$

Furthermore, (H_8) is a partial improvement of the interiority condition imposed in Benveniste and Scheinkman [7]:

($\mathbf{H''_8}$) There exist $T > 0$ and $\eta > 0$ such that

$$(x_0(t) + \eta B, \dot{x}_0(t) + \eta B) \subset \operatorname{gph} \Gamma(t, \cdot) \quad \text{a.e. } t \in [0, T].$$

Define the *Hamiltonian* $H : \mathbb{R}_+ \times E \times E^* \to \mathbb{R} \cup \{+\infty\}$ by

$$H(t, x, x^*) := \sup_{y \in \Gamma(t,x)} \{\langle x^*, y\rangle - L(t, x, y)\}.$$

Now we are ready to present an extension of the Euler–Lagrange necessary condition and the maximum principle with the transversality condition at infinity. In the theorem below we use the notions of weak* scalar measurability and Gelfand integrals whose definitions are recalled in Appendix A.2.

Theorem 4.1. *Suppose that* (H_1)–(H_4), (H'_5), (H_6), (H_7), *and* (H_8) *hold. If* $\partial_x^- V(0, x_0(0))$ *is nonempty and* $\partial_x^+ L(t, x_0(t), \dot{x}_0(t))$ *is nonempty a.e.* $t \in \mathbb{R}_+$, *then for every* $x^* \in \partial_x^- V(0, x_0(0))$ *and weakly* scalarly measurable selector* $f : \mathbb{R}_+ \to E^*$ *from the Dini–Hadamard superdifferential mapping* $\partial_x^+ L(\cdot, x_0(\cdot), \dot{x}_0(\cdot)) : \mathbb{R}_+ \rightsquigarrow E^*$, *the locally absolutely continuous function* $p : \mathbb{R}_+ \to E^*$ *defined by* $p(t) := \int_0^t f(s)ds - x^*$ *as a Gelfand integral satisfies:*

(i) $-p(t) \in \partial_x^- V(t, x_0(t))$ *for every* $t \in \mathbb{R}_+$;

(ii) $p(t) \in \partial_y^\circ L(t, x_0(t), \dot{x}_0(t)) + N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ *a.e.* $t \in \mathbb{R}_+$;

(iii) $\dot{p}(t) \in \partial_x^+ L(t, x_0(t), \dot{x}_0(t))$ *a.e.* $t \in \mathbb{R}_+$;

(iv) $H(t, x_0(t), p(t)) = \langle p(t), \dot{x}_0(t)\rangle - L(t, x_0(t)), \dot{x}_0(t))$ *a.e.* $t \in \mathbb{R}_+$;

(v) $\lim_{t\to\infty} p(t) = 0$,

where $\dot{p}(t)$ *denotes the weak* derivative of* $p(\cdot)$ *at* $t \in \mathbb{R}_+$. *In particular, if* $\partial_x^- V(0, x_0(0))$ *is nonempty, then* $\partial_x^- V(t, x_0(t))$ *is nonempty for every* $t \in \mathbb{R}_+$.

For the proof, see Subsection B.3.

If $\operatorname{gph}\Gamma(t,\cdot)$ is convex on which $L(t,\cdot,\cdot)$ is convex for every $t \in \mathbb{R}_+$, then $H(t,\cdot,x^*)$ is concave on E for every $x^* \in E^*$ and $H(t,x,\cdot)$ is convex on E^* for every $x \in E$. It thus follows from the Fenchel duality that the Euler–Lagrange conditions (ii) and (iii) of Theorem 4.1 are equivalent to the familiar Hamiltonian conditions $\dot{x}_0(t) \in \partial_p H(t,x_0(t),p(t))$ and $-\dot{p}(t) \in \partial_x H(t,x_0(t),p(t))$ respectively; see Rockafeller [48, Theorem 6].

Remark 4.1. Theorem 4.1 yields that "singular" points propagate forward along optimal trajectories, i.e., if the Dini–Hadamard subdifferential $\partial_x^- V(0,x_0(0)) \neq \emptyset$ is not a singleton, then so does $\partial_x^- V(t,x_0(t))$ for every $t \in \mathbb{R}_+$. In the finite-dimensional control systems, this observation is done also in Takekuma [52] for fully convex variational problems and in Cannarsa and Frankowska [16] for optimal control problems whose set of velocities are convex. Note that the nonemptiness of $\partial_x^- V(0,x_0(0))$ is an innocuous assumption because the set of points at which $\partial_x^- V(0,\cdot)$ is Dini–Hadamard subdifferentiable is dense in the separable Banach space E by Theorem 2.1.

Remark 4.2. Even if the Dini–Hadamard subdifferential $\partial_x^- V(t,x_0(t))$ is a singleton, the strict derivative $\nabla_x V(t,x_0(t))$ might not exist because of the lack of convexity of the lower directional derivative $v \mapsto V_x^-(t,x_0(t);v)$. This observation makes a sharp contrast with the case where the Clarke subdifferential $\partial_x^\circ V(t,x_0(t))$ is a singleton, in which case $\partial_x^\circ V(t,x_0(t))$ coincides with $\nabla_x V(t,x_0(t))$; see for detail the proof of Clarke [20, Proposition 2.2.4].

4.2 Necessary Conditions without the Interiority Assumption

Hypothesis (H_8) in Theorem 4.1 is stringent, mostly because the velocity multifunction Γ is too general. Hence, a "structural assumption" on the optimal trajectory $x_0(\cdot)$ compensates this generality. If instead some "structural" assumptions are imposed on Γ, then hypothesis (H_8) can be omitted. To illustrate this observation, we consider standard optimal control problems.

Let X be a complete separable metric space, $f : \mathbb{R}_+ \times E \times X \to E$ be a velocity function, and $U : \mathbb{R}_+ \rightsquigarrow X$ be a control multifunction. Denote by $\mathcal{M}(\mathbb{R}_+,X)$ the space of measurable functions on $\mathbb{R}_+$ with values in X. Define the integrand $\tilde{L} : \mathbb{R}_+ \times E \times X \to \mathbb{R}$ by $\tilde{L}(t,x,u) := L(t,x,f(t,x,u))$.

The optimal control problem under consideration is as follows:

$$\begin{aligned} &\inf_{\substack{x(\cdot)\in W^{1,1}_{\mathrm{loc}}(\mathbb{R}_+,E)\\ u(\cdot)\in\mathcal{M}(\mathbb{R}_+,X)}} \int_0^\infty \tilde{L}(t,x(t),u(t))dt \\ &\text{s.t. } u(t) \in U(t) \text{ a.e. } t \in \mathbb{R}_+, \\ &\qquad \dot{x}(t) = f(t,x(t),u(t)) \text{ a.e. } t \in \mathbb{R}_+, \quad x(0) = \xi. \end{aligned} \tag{$\tilde{\mathrm{P}}_0$}$$

The Hamiltonian for problem $(\tilde{\mathrm{P}}_0)$ is given by:

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

The velocity multifunction is defined by $\Gamma(t,x) := f(t,x,U(t))$. Now impose "usual" assumptions on f and U in order that Γ satisfies (H_1)–(H_4), (H_5'), (H_6), and (H_7). Note that $\partial_x^+ \tilde{L}(t,x,u)$ is nonempty at $(t,x,u) \in \mathbb{R}_+ \times E \times X$ whenever so is $\partial_{x,y}^+ L(t,x,f(t,x,u))$ and $f(t,\cdot,u)$ is Gateaux differentiable at x.

Denote by $\mathcal{L}(E)$ the space of bounded linear operators on E.

The following reasonable hypothesis is a "structural assumption" on f that dispenses with (H_8).

($\mathbf{H_9}$) (i) f is a Carathéodory function, i.e., $f(\cdot,x,u)$ is measurable for every $(x,u) \in E \times X$ and $f(t,\cdot,\cdot)$ is continuous for every $t \in \mathbb{R}_+$.

(ii) For every $R > 0$ and $T > 0$ there exists an integrable function $k : [0,T] \to \mathbb{R}$ such that:

(a) $\|f(t,x,u)\| \le k(t)$ for every $t \in [0,T]$, $x \in RB$, and $u \in U(t)$;

(b) $f(t,\cdot,u)$ is Lipschitz of rank $k(t)$ on RB for every $t \in [0,T]$ and $u \in U(t)$.

(iii) $f(t,\cdot,u)$ is Fréchet differentiable on E for every $(t,u) \in \mathbb{R}_+ \times X$ and the mapping $(t,x,u) \mapsto \nabla_x f(t,x,u)$ is continuous in the uniform operator topology of $\mathcal{L}(E)$.

Conditions (H_9)-(i), (ii) guarantee the existence of solutions of the integral equation

$$x(t) = \int_0^t f(s,x(s),u(s))ds + \xi \quad \text{for every } t \in \mathbb{R}_+$$

for any control $u(\cdot) \in \mathcal{M}(\mathbb{R}_+, X)$, where the locally absolutely continuous function $x(\cdot) : \mathbb{R}_+ \to E$ is a unique mild solution to the ordinary differential equation (ODE) in $(\tilde{\mathrm{P}}_0)$ (see Fattorini [29, Theorem 5.5.1]), which has the strong derivative $\dot{x}(t)$ a.e. $t \in \mathbb{R}_+$ in view of the separability of E and the Lebesgue differentiation theorem.

Let $(x_0(\cdot), u_0(\cdot)) \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}_+, E) \times \mathcal{M}(\mathbb{R}_+, X)$ be an optimal trajectory-control pair for optimal control problem $(\tilde{\mathrm{P}}_0)$. Denote by $\nabla_x f(s,x_0(s),u_0(s))^*$ in $\mathcal{L}(E^*)$ the adjoint operator of $\nabla_x f(s,x_0(s),u_0(s))$ in $\mathcal{L}(E)$.

Theorem 4.2. *Suppose that* (H_1)–(H_4), (H_5'), (H_6), (H_7), *and* (H_9) *hold with* $\Gamma(t,x) = f(t,x,U(t))$. *If* $\partial_x^- V(0,x_0(0))$ *is nonempty and the Dini–Hadamard superdifferential mapping* $\partial_x^+ \tilde{L}(\cdot,x_0(\cdot),u_0(\cdot)) : \mathbb{R}_+ \rightsquigarrow E^*$ *admits a locally Bochner integrable selector, then there exists a locally absolutely continuous function* $p : \mathbb{R}_+ \to E^*$ *such that:*

(i) $-p(t) \in \partial_x^- V(t, x_0(t))$ *for every* $t \in \mathbb{R}_+$*;*

(ii) $p(t) \in \partial_y^\circ L(t, x_0(t), f(t, x_0(t), u_0(t))) + N_{\Gamma(t,x_0(t))}(f(t, x_0(t), u_0(t)))$ *a.e.* $t \in \mathbb{R}_+$*;*

(iii) $-\dot{p}(t) \in \nabla_x f(t, x_0(t), u_0(t))^* p(t) - \partial_x^+ \tilde{L}(t, x_0(t), u_0(t))$ *a.e.* $t \in \mathbb{R}_+$*;*

(iv) $H(t, x_0(t), p(t)) = \langle p(t), f(t, x_0(t), u_0(t))\rangle - \tilde{L}(t, x_0(t), u_0(t))$ *a.e.* $t \in \mathbb{R}_+$*;*

(v) $\lim_{t\to\infty} p(t) = 0$*,*

where $\dot{p}(t)$ *denotes the strong derivative of* $p(\cdot)$ *at* $t \in \mathbb{R}_+$*. In particular, if* $\partial_x^- V(0, x_0(0))$ *is nonempty, then* $\partial_x^- V(t, x_0(t))$ *is nonempty for every* $t \in \mathbb{R}_+$*.*

The proof is provided in Subsection B.3.

Remark 4.3. The existence of locally Bochner integrable selectors from the Dini–Hadamard superdifferential mapping $t \rightsquigarrow \partial_x^+ \tilde{L}(t, x_0(t), u_0(t))$ follows from (H_3) and (H_9) whenever E^* is separable in the dual norm. For the case with nonseparable E^*, the Fréchet differentiability of the integrand $L(t, \cdot, \cdot)$ on $E \times E$ and the continuity of $(t, x, y) \mapsto (\nabla_x L(t, x, y), \nabla_y L(t, x, y))$ in the dual norm of $E^* \times E^*$ guarantee the local Bochner integrability of $t \mapsto \nabla_x \tilde{L}(t, x_0(t), u_0(t))$ in E^* under (H_3) and (H_9). If $\partial_{x,y}^+ L(t, x_0(t), \dot{x}_0(t))$ is nonempty, then take any $(p, q) \in \partial_{x,y}^+ L(t, x_0(t), \dot{x}_0(t))$ and observe that for every $v \in E$ and $u \in U(t)$, we have:

$$\begin{aligned}\tilde{L}_x^+(t, x, u; v) &\le L_{x,y}^+(t, x, f(t, x, u); v, \nabla_x f(t, x, u)v) \\ &\le \langle p, v\rangle + \langle q, \nabla_x f(t, x, u)v\rangle = \langle p + \nabla_x f(t, x, u)^* q, v\rangle.\end{aligned}$$

Thus, instead of using the function $g : \mathbb{R}_+ \to E^*$ in the proof of Theorem 4.2 below, we could use as well any locally Bochner integrable selector $(\alpha(t), \beta(t)) \in \partial_{x,y}^+ L(t, x_0(t), \dot{x}_0(t))$ and write the adjoint equation involving $g(t) = \alpha(t) + \nabla_x f(t, x_0(t), u_0(t))^* \beta(t)$.

5 An Application: Spatial Ramsey Growth Models

5.1 Ramsey Meets Hotelling

Consider the spatial Ramsey growth model with a general reduced form explored in Boucekkine et al. [10, 11], Brock et al. [13], Brito [12], Camacho et al. [14] in the specific form. Let $I = [0, 1]$ be the unit interval such that the endpoints $0, 1 \in I$ are identified. Then I is homeomorphic to the unit circle in which a spatial parameter $\theta \in I$ is a location of agents along the

lines of Hotelling [38]. Let $W : \mathbb{R}_+ \times \mathbb{R}_+ \times I \to \mathbb{R}$ be a function such that $W(\cdot, \cdot, \theta)$ is an instantaneous utility function at location $\theta \in I$ satisfying $W(a, b, 0) = W(a, b, 1)$ for every $(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+$, where a denotes a capital stock and b a net investment. Let $F : \mathbb{R}_+ \times \mathbb{R}_+ \times I \to \mathbb{R}_+$ be a function such that $F(\cdot, \cdot, \theta)$ is a production function at θ satisfying $F(a, c, 0) = F(a, c, 1)$ for every $(a, c) \in \mathbb{R}_+ \times \mathbb{R}_+$, where c denotes a consumption and output $F(a, c, \theta)$ is a net investment at θ for every (a, c). Let $r > 0$ be a discount rate. For simplicity, we assume no depreciation of capital stock.

Let $x : \mathbb{R}_+ \times I \to \mathbb{R}_+$ be a capital stock trajectory in which $x(t, \theta)$ is a current capital stock, $\partial x(t, \theta)/\partial t$ is a current capital accumulation, $u : \mathbb{R}_+ \times I \to \mathbb{R}_+$ denotes a consumption trajectory in which $u(t, \theta)$ is a current consumption, and $\tau : \mathbb{R}_+ \times I \to \mathbb{R}$ denotes a net transfer trajectory in which $\tau(t, \theta)$ is a current net transfer, respectively at period $t \in \mathbb{R}_+$ and location $\theta \in I$. The capital accumulation process is described by

$$\frac{\partial x(t, \theta)}{\partial t} = F(x(t, \theta), u(t, \theta), \theta) + \tau(t, \theta)$$

for a.e. $t \in \mathbb{R}_+$ and for every $\theta \in I$. If $\tau(t, \theta) \equiv 0$, then the model describes an autarkic economy in which no capital movement occurs among locations; see Brito [12]. For the case where $\tau(t, \theta)$ is a parabolic term, see Boucekkine et al. [10, 11], Brito [12], Camacho et al. [14]. The choice of function spaces depends upon the specification of a transfer term. Following the forementioned works, we focus here on the autarkic case in which capital stock and consumption change smoothly in locations.

Let $C^2(I)$ be the space of twice continuously differentiable functions on I with their values equal at the end points $\theta = 0, 1$, endowed with the C^2-norm $\|x\|_{C^2(I)} := \sup_{\theta \in I}\{|x(\theta)| + |x'(\theta)| + |x''(\theta)|\}$, which makes $C^2(I)$ a separable Banach space. Denote by $C^2_+(I)$ the positive cone of $C^2(I)$ consisting of all nonnegative functions in $C^2(I)$. The problem under investigation is:

$$\begin{aligned}
&\max \int_0^\infty \int_I e^{-rt} W\left(x(t, \theta), \frac{\partial x(t, \theta)}{\partial t}, \theta\right) d\theta dt \\
&\text{s.t. } \frac{\partial x(t, \theta)}{\partial t} = F(x(t, \theta), u(t, \theta), \theta), \quad u(t) \in U(t) \\
&\qquad\qquad\qquad\qquad \text{a.e. } t \in \mathbb{R}_+ \text{ for every } \theta \in I, \\
&\quad x(t, 0) = x(t, 1),\ u(t, 0) = u(t, 1) \text{ a.e. } t \in \mathbb{R}_+, \\
&\quad x(0, \theta) = \xi(\theta) \text{ for every } \theta \in I.
\end{aligned} \tag{Q$_0$}$$

Here, the control set $U(t)$ is a subset of $C^2_+(I)$ for every $t \in \mathbb{R}_+$. The maximization is taken over all *nonnegative* functions $x(\cdot, \cdot)$ in the function space such that $x(\cdot, \theta)$ is a.e. differentiable on $\mathbb{R}_+$ for every $\theta \in I$ with $x(t, \cdot)$ and $\partial x(t, \cdot)/\partial t$ belonging to $C^2(I)$ a.e. $t \in \mathbb{R}_+$, and over all functions $u(\cdot, \cdot)$ such that $u(t, \cdot) \in U(t)$ and $u(\cdot, \theta)$ is measurable on $\mathbb{R}_+$ for every

$\theta \in I$ satisfying the parametrized ODE above, where the initial condition at location θ is given by $\xi(\theta)$ with $\xi \in C^2_+(I)$.

Throughout this section, we assume the following.

Assumption 5.1. (i) W has an extension to $\mathbb{R} \times \mathbb{R} \times I$ (which we do not relabel) such that $W(a, b, \cdot)$ is measurable on I for every $(a, b) \in \mathbb{R} \times \mathbb{R}$, $W(\cdot, \cdot, \theta)$ is continuously differentiable on $\mathbb{R} \times \mathbb{R}$ for every $\theta \in I$, and its partial derivatives are bounded uniformly in $(a, b, \theta) \in \mathbb{R} \times \mathbb{R} \times I$.

(ii) F has a thrice continuously differentiable extension to $\mathbb{R} \times \mathbb{R} \times I$ (which we do not relabel) such that every partial derivative of any order less or equal 3 is bounded uniformly in $(a, c, \theta) \in \mathbb{R} \times \mathbb{R} \times I$.

(iii) There exists a bounded closed subset X of $C^2_+(I)$ such that $U(t) \subset X$ for every $t \in \mathbb{R}_+$.

Define the integrands $L : \mathbb{R}_+ \times C^2(I) \times C^2(I) \to \mathbb{R}$ by

$$L(t, x(\cdot), y(\cdot)) := -e^{-rt} \int_I W(x(\theta), y(\theta), \theta) d\theta$$

and $\tilde{L} : \mathbb{R}_+ \times C^2(I) \times X \to \mathbb{R}$ by

$$\tilde{L}(t, x(\cdot), u(\cdot)) := -e^{-rt} \int_I W(x(\theta), F(x(\theta), u(\theta), \theta), \theta) d\theta$$

respectively. Consider the velocity function $f : C^2(I) \times X \to C^2(I)$ defined by

$$f(x(\cdot), u(\cdot)) := F(x(\cdot), u(\cdot), \cdot).$$

Here, $f(\cdot, u(\cdot))$ is Frèchet differentiable on $C^2(I)$ and its Frèchet derivative $\nabla_x f(x(\cdot), u(\cdot)) \in \mathcal{L}(C^2(I))$ can be calculated as

$$\nabla_x f(x(\cdot), u(\cdot)) v(\cdot) = \frac{\partial F(x(\cdot), u(\cdot), \cdot)}{\partial a} v(\cdot)$$

for every $v(\cdot) \in C^2(I)$. By construction, it is evident that

$$\tilde{L}(t, x(\cdot), u(\cdot)) = L(t, x(\cdot), f(x(\cdot), u(\cdot)))$$

for every $(x(\cdot), u(\cdot)) \in C^2(I) \times X$. Define the velocity multifunction $\Gamma : \mathbb{R}_+ \times C^2(I) \rightsquigarrow C^2(I)$ by $\Gamma(t, x(\cdot)) := f(x(\cdot), U(t))$.

We then convert the problem (Q_0) into the minimization one of the form $(\tilde{\mathrm{P}}_0)$ in the setting with $E = C^2(I)$ and $X \subset C^2_+(I)$. If $(x_0(\cdot), u_0(\cdot)) \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}_+, C^2(I)) \times \mathcal{M}(\mathbb{R}_+, X)$ is an optimal trajectory-control pair of $(\tilde{\mathrm{P}}_0)$, then $(x_0(\cdot), u_0(\cdot))$ is a solution to the associated problem (Q_0), and vice versa because any admissible trajectory of $(\tilde{\mathrm{P}}_0)$ stays in the nonnegative orthant. It is easy to see that Assumption 5.1 guarantees hypotheses in Theorem 4.2. In particular, Hypothesis (H_7) is satisfied whenever $r > 0$ is large enough.

5.2 Necessary Conditions for Optimality

Let $C(I)$ be the space of continuous functions on I endowed with the sup norm and $ca(I)$ be the space of signed Borel measures on I. Since each $x \in C^2(I)$ is represented by

$$x(\theta) = x(0) + x'(0)\theta + \int_0^\theta \int_0^\sigma x''(\omega) d\omega d\sigma \quad \text{for every } \theta \in I$$

with $x(0), x'(0) \in \mathbb{R}$ and $x'' \in C(I)$, the Banach space $C^2(I)$ is identified with the direct sum $\mathbb{R} \oplus \mathbb{R} \oplus C(I)$. Hence, $C^2(I)^* = \mathbb{R} \oplus \mathbb{R} \oplus ca(I)$ and each $x^* \in C^2(I)^*$ has the form

$$\langle x^*, x \rangle = \alpha_0 x(0) + \alpha_1 x'(0) + \int_I x''(\theta) d\mu \quad \text{for every } x \in C^2(I)$$

for some constants $\alpha_0, \alpha_1 \in \mathbb{R}$ and a signed Borel measure $\mu \in ca(I)$; see Dunford and Schwartz [27, Exercise IV.13.36]. Hence, the adjoint variables in the spatial Ramsey growth model take values in $\mathbb{R} \oplus \mathbb{R} \oplus ca(I)$.

Let $(x_0(\cdot), u_0(\cdot)) \in W^{1,1}(\mathbb{R}_+, C^2(I)) \times \mathcal{M}(\mathbb{R}_+, X)$ be an optimal trajectory-control pair of $(\tilde{\mathrm{P}}_0)$ and $V : \mathbb{R}_+ \times C^2(I) \to \mathbb{R}$ be the value function. A direct calculation shows that for every $v \in C^2(I)$ we have

$$\langle \nabla_x L(t, x_0(t), \dot{x}_0(t)), v \rangle = -e^{-rt} \int_I \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)}{\partial a} v(\theta) d\theta.$$

To evaluate the above integral, define $\alpha(\theta) := \partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)/\partial a$, $A(\theta) := -\int_\theta^1 \alpha(\omega) d\omega$, and $B(\theta) := -\int_\theta^1 A(\omega) d\omega$. The double use of integration by parts yields

$$\begin{aligned}
&\int_I \alpha(\theta) v(\theta) d\theta = [A(\theta) v(\theta)]_0^1 - \int_I A(\theta) v'(\theta) d\theta \\
= &- A(0)v(0) + B(0)v'(0) + \int_I B(\theta) v''(\theta) d\theta \\
= &\int_I \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)}{\partial a} d\theta v(0) \\
&\quad + \int_I \left[\int_\theta^1 \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\omega), \omega)}{\partial a} d\omega \right] d\theta v'(0) \\
&\quad + \int_I \left[\int_\theta^1 \int_\sigma^1 \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\omega), \omega)}{\partial a} d\omega d\sigma \right] v''(\theta) d\theta.
\end{aligned}$$

Henceforth, we obtain

$$\langle \nabla_x L(t, x_0(t), \dot{x}_0(t)), v \rangle = a_0(t) v(0) + a_1(t) v'(0) + \int_I v''(\theta) d\mu(t)$$

with

$$a_0(t) = -e^{-rt} \int_I \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)}{\partial a} d\theta,$$

$$a_1(t) = -e^{-rt} \int_I \int_\theta^1 \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\omega), \omega)}{\partial a} d\omega d\theta,$$
$$\frac{d\mu_0(t)}{d\theta}(\theta) = -e^{-rt} \int_\theta^1 \int_\sigma^1 \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\omega), \omega)}{\partial a} d\omega d\sigma,$$

where $\mu_0(t) \in ca(I)$ is given by its Radon–Nikodym derivative $d\mu_0(t)/d\theta$. Hence, $\nabla_x L(t, x_0(t), \dot{x}_0(t)) = a_0(t) \oplus a_1(t) \oplus \mu_0(t) \in \mathbb{R} \oplus \mathbb{R} \oplus ca(I)$. Similarly, replacing α by $\alpha(\theta) := \partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)/\partial b$ in the above argument with integration by parts yields

$$\begin{aligned}\langle \nabla_y L(t, x_0(t), \dot{x}_0(t)), v\rangle &= -e^{-rt} \int_I \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)}{\partial b} v(\theta) d\theta \\ &= b_0(t)v(0) + b_1(t)v'(0) + \int_I v''(\theta) d\nu_0(t)\end{aligned}$$

with

$$b_0(t) = -e^{-rt} \int_I \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)}{\partial b} d\theta,$$
$$b_1(t) = -e^{-rt} \int_I \int_\theta^1 \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\omega), \omega)}{\partial b} d\omega d\theta,$$
$$\frac{d\nu_0(t)}{d\theta}(\theta) = -e^{-rt} \int_\theta^1 \int_\sigma^1 \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\omega), \omega)}{\partial b} d\omega d\sigma.$$

Hence, $\nabla_y L(t, x_0(t), \dot{x}_0(t)) = b_0(t) \oplus b_1(t) \oplus \nu_0(t) \in \mathbb{R} \oplus \mathbb{R} \oplus ca(I)$.

If $\xi \in C^2_+(I)$ is a point such that $\partial_x^- V(0, \xi)$ is nonempty (see Remark 4.1), then in view of Theorem 4.2 there exists a locally absolutely continuous function $p : \mathbb{R}_+ \to C^2(I)^*$ such that (i) $-p(t) \in \partial_x^- V(t, x_0(t))$ for every $t \in \mathbb{R}_+$; (ii) $p(t) = \nabla_y L(t, x_0(t), \dot{x}_0(t)) + q(t)$ a.e. $t \in \mathbb{R}_+$, where $q : \mathbb{R}_+ \to C^2(I)^*$ is a Borel measurable selector from $t \rightsquigarrow N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ with respect to the weak* topology of $C^2(I)^*$; (iii) $-\dot{p}(t) = \nabla_x f(x_0(t), u_0(t))^* p(t) - \nabla_x \tilde{L}(t, x_0(t), \dot{x}_0(t))$ a.e. $t \in \mathbb{R}_+$; (iv) $H(t, x_0(t), p(t)) = \langle p(t), f(x_0(t), u_0(t))\rangle - \tilde{L}(t, x_0(t), u_0(t))$ a.e. $t \in \mathbb{R}_+$; (v) $\lim_{t\to\infty} p(t) = 0$. Since q is represented as $q(t) = \beta_0(t) \oplus \beta_1(t) \oplus \nu_1(t) \in \mathbb{R} \oplus \mathbb{R} \oplus ca(I)$, condition (ii) can be written as $p(t) = (b_0(t) + \beta_0(t)) \oplus (b_1(t) + \beta_1(t)) \oplus (\nu_0(t) + \nu_1(t))$ a.e. $t \in \mathbb{R}_+$. Since

$$\begin{aligned}&\nabla_x \tilde{L}(t, x_0(t), u_0(t)) \\ &= \nabla_x L(t, x_0(t), \dot{x}_0(t)) + \nabla_y L(t, x_0(t), \dot{x}_0(t)) \nabla_x f(x_0(t), u_0(t)) \\ &= \nabla_x L(t, x_0(t), \dot{x}_0(t)) + \nabla_x f(x_0(t), u_0(t))^* (p(t) - q(t)),\end{aligned}$$

condition (iii) yields $-\dot{p}(t) = -(a_0(t) \oplus a_1(t) \oplus \mu_0(t)) + q(t)\nabla_x f(x_0(t), u_0(t))$, and hence, $\dot{p}(t) = (a_0(t) - \alpha_0(t)) \oplus (a_1(t) - \alpha_1(t)) \oplus (\mu_0(t) - \mu_1(t))$ with setting $q(t)\nabla_x f(x_0(t), u_0(t)) = \alpha_0(t) \oplus \alpha_1(t) \oplus \mu_1(t) \in \mathbb{R} \oplus \mathbb{R} \oplus ca(I)$.

Define the new adjoint variable by $\pi(t) := e^{rt} p(t)$ and denote it by $\pi(t) = \pi_0(t) \oplus \pi_1(t) \oplus \lambda(t) \in \mathbb{R} \oplus \mathbb{R} \oplus ca(I)$. The transversality condition

at infinity can be written as $\lim_{t\to\infty} e^{-rt}\pi(t) = 0$. It follows from condition (iii) that

$$\dot{\pi}_0(t) = r\pi_0(t) - \int_I \frac{\partial W(x_0(t,\theta), \dot{x}_0(t,\theta), \theta)}{\partial a} d\theta - e^{rt}\alpha_0(t) \tag{5.1}$$

$$\dot{\pi}_1(t) = r\pi_1(t) - \int_I \int_\theta^1 \frac{\partial W(x_0(t,\omega), \dot{x}_0(t,\omega), \omega)}{\partial a} d\omega d\theta - e^{rt}\alpha_1(t) \tag{5.2}$$

$$\begin{aligned}\dot{\lambda}(t)(A) = r\lambda(t)(A) - \int_A \int_\theta^1 \int_\sigma^1 \frac{\partial W(x_0(t,\omega), \dot{x}_0(t,\omega), \omega)}{\partial a} d\omega d\sigma d\theta \\ - e^{rt}\mu_1(t)(A)\end{aligned} \tag{5.3}$$

a.e. $t \in \mathbb{R}_+$ for every $A \in \mathcal{L}$. Hence, any stationary point $(\bar{x}, \bar{\pi}) \in C^2(I) \times C^2(I)^*$ of the dynamical system corresponding to $\dot{x}_0(t) \equiv 0$ and $\dot{\pi}(t) \equiv 0$ with the adjoint equations (5.1)–(5.3) is determined by the following conditions:

$$\begin{aligned}
&F(\bar{x}(\cdot), \bar{u}(\cdot), \cdot) = 0, \\
&\bar{\pi}_0 = r^{-1}\left(\int_I \frac{\partial W(\bar{x}(\theta), 0, \theta)}{\partial a} d\theta + \alpha_0(0)\right), \\
&\bar{\pi}_1 = r^{-1}\left(\int_I \int_\theta^1 \frac{\partial W(\bar{x}(\omega), 0, \omega)}{\partial a} d\omega d\theta + \alpha_1(0)\right), \\
&\bar{\lambda}(A) = r^{-1}\left(\int_A \int_\theta^1 \int_\sigma^1 \frac{\partial W(\bar{x}(\omega), 0, \omega)}{\partial a} d\omega d\sigma d\theta + \mu_1(0)(A)\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for every } A \in \mathcal{L}
\end{aligned}$$

along with $(\alpha_0(t), \alpha_1(t), \mu_1(t)) = (\alpha_0(0), \alpha_1(0), \mu_1(0))e^{-rt}$.

A Appendix I

A.1 Proof of Proposition 2.1

Since $T_C(\bar{x}) \subset K_C(\bar{x})$, if $v \in T_C(\bar{x})$, then there exists a sequence $\{\theta_n\}_{n\in\mathbb{N}}$ of positive real numbers with $\theta_n \downarrow 0$ and a sequence $\{v_n\}_{n\in\mathbb{N}}$ in E with $v_n \to v$ such that $\bar{x} + \theta_n v_n \in C$ for each $n \in \mathbb{N}$. Thus, $\varphi_i(\bar{x} + \theta_n v_n) \le 0$ for each $i \in I(\bar{x})$. Since φ_i is strictly differentiable (and hence Gateaux differentiable) at $\bar{x}$, we have

$$\langle \nabla\varphi_i(\bar{x}), v\rangle = \lim_{n\to\infty} \frac{\varphi_i(\bar{x} + \theta_n v) - \varphi_i(\bar{x})}{\theta_n} \le \lim_{n\to\infty} \left(\frac{\varphi_i(\bar{x} + \theta_n v_n)}{\theta_n} + \alpha_i \|v_n - v\|\right) \le 0$$

for each $i \in I(\bar{x})$, where α_i is a Lipschitz modulus of φ_i. Hence, $T_C(\bar{x}) \subset \{v \in E \mid \langle \nabla\varphi_i(\bar{x}), v\rangle \le 0 \ \forall i \in I(\bar{x})\}$. Since $0 \notin \operatorname{co}\{\nabla\varphi_i(\bar{x}) \mid i \in I(\bar{x})\}$, by the separation theorem, there exists $\bar{v} \in E$ such that $\sup_{i\in I(\bar{x})}\langle \nabla\varphi_i(\bar{x}), \bar{v}\rangle =: -\varepsilon < 0$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in C with $x_n \to \bar{x}$ and $\{\theta_n\}_{n\in\mathbb{N}}$ be a sequence of positive real numbers with $\theta_n \downarrow 0$. Then $\varphi_i(x_n) \le 0$ and

$$\limsup_{n\to\infty} \frac{\varphi_i(x_n + \theta_n \bar{v})}{\theta_n} \le \limsup_{n\to\infty} \frac{\varphi_i(x_n + \theta_n \bar{v}) - \varphi_i(x_n)}{\theta_n} = \langle \nabla\varphi_i(\bar{x}), \bar{v}\rangle < 0.$$

Thus, for every n sufficiently large, we have $\varphi_i(x_n + \theta_n \bar{v}) < 0$ for each $i \in I(\bar{x})$. Consequently, $x_n + \theta_n \bar{v} \in C$ for every n sufficiently large, and hence, $\bar{v} \in T_C(\bar{x})$. Let $v \in E$ be such that $\langle \nabla\varphi_i(\bar{x}), v\rangle \leq 0$ for each $i \in I(\bar{x})$. For $\alpha \in (0,1)$ define $v_\alpha := \alpha\bar{v} + (1-\alpha)v$. Then $\langle \nabla\varphi_i(\bar{x}), v_\alpha\rangle < 0$ for each $i \in I(\bar{x})$. Similar to the case for $\bar{v}$, we have $v_\alpha \in T_C(\bar{x})$. Since $v_\alpha \to v$ as $\alpha \to 0^+$, taking the limit yields $v \in T_C(\bar{x})$. Therefore, $T_C(\bar{x}) = \{v \in E \mid \langle \nabla\varphi_i(\bar{x}), v\rangle \leq 0 \ \forall i \in I(\bar{x})\}$.

Set $A = \sum_{i\in I(\bar{x})} \mathbb{R}_+ \nabla\varphi_i(\bar{x})$. Clearly, $\langle x^*, v\rangle \leq 0$ for every $x^* \in A$ and $v \in T_C(\bar{x})$, and hence, $A \subset N_C(\bar{x})$. We claim that $A = N_C(\bar{x})$. Toward this end, we first show that A is weakly* closed. Let $\{x^*_\nu\}$ be a net in A converging weakly* to x^* with $x^*_\nu = \sum_{i\in I(\bar{x})} \lambda^\nu_i \nabla\varphi_i(\bar{x})$ for each ν. Then $\langle x^*_\nu, \bar{v}\rangle \to \langle x^*, \bar{v}\rangle$. Since $\sum_{i\in I(\bar{x})} \lambda^\nu_i \langle \nabla\varphi_i(\bar{x}), \bar{v}\rangle \leq -\varepsilon \sum_{i\in I(\bar{x})} \lambda^\nu_i \leq 0$ for each ν, the net $\{(\lambda^\nu_{i_1}, \ldots, \lambda^\nu_{i_k})\}$ is bounded in $\mathbb{R}^k_+$, where $I(\bar{x}) = \{i_1, \ldots i_k\}$ with $k \leq m$. Thus it has a subnet converging to some $(\lambda_{i_1}, \ldots, \lambda_{i_k})$ in $\mathbb{R}^k_+$. Hence, x^*_ν converges weakly* to $x^* = \sum_{i\in I(\bar{x})} \lambda_i \nabla\varphi_i(\bar{x}) \in A$. Assume for a moment that there exists $p \in N_C(\bar{x})$ such that $p \notin A$. By the separation theorem, there exists $y \in E$ such that $0 \leq \sup_{y^*\in A}\langle y^*, y\rangle < \langle p, y\rangle$. Since A is a cone, we must have $\sup_{y^*\in A}\langle y^*, y\rangle = 0$ implying that $\langle \nabla\varphi_i(\bar{x}), y\rangle \leq 0$ for each $i \in I(\bar{x})$. Therefore, $y \in T_C(\bar{x})$ and $\langle p, y\rangle \leq 0$. The obtained contradiction yields $A = N_C(\bar{x})$. $\square$

A.2 Gelfand Integrals of Multifunctions

Let I be a nonempty closed subset of the real line $\mathbb{R}$ with the Lebesgue measure and the Lebesgue σ-algebra $\mathcal{L}$. Denote by Borel(E^*, w^*) the Borel σ-algebra of the dual space E^* generated by the weak* topology. A function $f : I \to E^*$ is said to be *weakly* scalarly measurable* if the scalar function $\langle f(\cdot), x\rangle$ is measurable for every $x \in E$. If E is a separable Banach space, then E^* is a locally convex Suslin space under the weak* topology. In this case, a function $f : I \to E^*$ is weakly* scalarly measurable if and only if it is measurable with respect to Borel(E^*, w^*); see Castaing and Valadier [19, Theorem III.36]. A weakly* scalarly measurable function f is said to be *weakly* scalarly integrable* if $\langle f(\cdot), x\rangle$ is integrable for every $x \in E$. Further, a weakly* scalarly measurable function f is said to be *Gelfand integrable* (or *weakly* integrable*) over a given set $A \in \mathcal{L}$ if there exists $x^*_A \in E^*$ such that $\langle x^*_A, x\rangle = \int_A \langle f(t), x\rangle dt$ for every $x \in E$. The element x^*_A is called the *Gelfand* (or *weak**) *integral* of f over A and is denoted by $\int_A f dt$. Note that every weakly* scalarly integrable function is Gelfand integrable over I as shown in Aliprantis and Border [1, Theorem 11.52].

A multifunction $\Phi : I \rightsquigarrow E^*$ is said to be *upper measurable* if the set $\{t \in I \mid \Phi(t) \subset O\}$ belongs to $\mathcal{L}$ for every weakly* open subset O of E^*; Φ is said to be *graph measurable* if the set $\operatorname{gph}\Phi := \{(t, x^*) \in I \times E^* \mid x^* \in \Phi(t)\}$ belongs to $\mathcal{L} \otimes$ Borel(E^*, w^*); Φ is said to be *weakly* scalarly measurable* if the scalar function $s(x, \Phi(\cdot)) : I \to \mathbb{R} \cup \{\pm\infty\}$ is measurable for every $x \in E$, where we set $s(x, \emptyset) := -\infty$. When Φ has empty values on a null set N, we can extend it to I with nonempty values at every point by setting $\Phi(t) := \{0\}$ for $t \in N$, preserving its upper, graph, and weak* scalar measurability. A function $f : I \to E^*$ is called a *selector* of Φ if $f(t) \in \Phi(t)$ a.e. $t \in I$. If E is separable, then an a.e. nonempty-valued multifunction $\Phi : I \rightsquigarrow E^*$ with measurable graph in $\mathcal{L} \otimes$ Borel(E^*, w^*) admits a Borel(E^*, w^*)-measurable selector; see Castaing and Valadier [19, Theorem III.22]. If E is separable and Φ has a.e. nonempty, weakly* compact, convex values, then Φ is weakly* scalarly measurable if and only if it is upper measurable (see Aliprantis and

Border [1, Theorem 18.31]), and in this case, Φ admits a Borel(E^*, w^*)-measurable (or equivalently, weakly* scalarly measurable) selector; see Aliprantis and Border [1, Theorem 18.33] or Cascales et al. [18, Corollary 3.1].

A multifunction $\Phi : I \rightsquigarrow E^*$ with a.e. nonempty values is *integrably bounded* if there exists an integrable function $\gamma : I \to \mathbb{R}$ such that $\sup_{x^* \in \Phi(t)} \|x^*\| \leq \gamma(t)$ a.e. $t \in I$. If Φ is integrably bounded with measurable graph, then it admits a Gelfand integrable selector whenever E is separable. Denote by $\mathcal{S}^1_\Phi$ the set of Gelfand integrable selectors of Φ. The Gelfand integral of Φ is conventionally defined as $\int \Phi dt := \{\int f dt \mid f \in \mathcal{S}^1_\Phi\}$. If Φ is an integrably bounded, weakly* closed, convex-valued multifunction with measurable graph, then $\int \Phi dt$ is nonempty, weakly* compact, and convex with $s(x, \int \Phi dt) = \int s(x, \Phi(t))dt$ for every $x \in E$ whenever E is separable; see Cascales et al. [18, Proposition 2.3 and Theorem 4.5].

A.3 Gelfand Integrals of Clarke and Dini–Hadamard Subdifferential Mappings

An $\mathcal{L} \otimes$ Borel$(E, \|\cdot\|)$-measurable function $L : I \times E \to \mathbb{R} \cup \{+\infty\}$ is called a *normal integrand* if $L(t, \cdot)$ is lower semicontinuous on E for every $t \in I$. For a given measurable function $x(\cdot) : I \to E$, let $L : I \times E \to \mathbb{R}$ be a function such that (i) $L(\cdot, x)$ is measurable for every $x \in E$; (ii) there exist $\varepsilon > 0$ and an integrable function $k : I \to \mathbb{R}$ such that $|L(t, x) - L(t, y)| \leq k(t)\|x - y\|$ for every $x, y \in x(t) + \varepsilon B$ and $t \in I$. The Clarke subdifferential mapping $t \rightsquigarrow \partial^\circ_x L(t, x(t))$ is an integrably bounded multifunction from I to E^* with weakly* compact, convex values. In view of the fact that $L^\circ_x(t, x(t); v) = s(v, \partial^\circ_x L(t, x(t)))$ for every $t \in I$ and $v \in E$, the Clarke subdifferential mapping $\partial^\circ_x L(\cdot, x(\cdot)) : I \rightsquigarrow E^*$ is weakly* scalarly measurable if and only if the Clarke directional derivative function $L^\circ_x(\cdot, x(\cdot); v) : I \to \mathbb{R}$ is measurable for every $v \in E$. This holds in particular when E is separable (see Clarke [20, Lemma, p. 78 and the proof of Theorem 2.7.8]), and hence, in this case, $\partial^\circ_x L(\cdot, x(\cdot))$ admits a Gelfand integrable selector.

We summarize the above result on the Gelfand integrability of the Clarke subdifferential mapping together with the results in Subsection A.2 as follows.

Proposition A.1 (Cascales et al. [18], Clarke [20])**.** *Let E be a separable Banach space and I be a nonempty closed subset of $\mathbb{R}$. If, for a given measurable function $x : I \to E$, the function $L : I \times E \to \mathbb{R}$ satisfies the following conditions:*

(i) *$L(\cdot, x)$ is measurable for every $x \in E$;*

(ii) *There exist $\varepsilon > 0$ and an integrable function $k : I \to \mathbb{R}$ such that $|L(t, x) - L(t, y)| \leq k(t)\|x - y\|$ for every $x, y \in x(t) + \varepsilon B$ and $t \in I$;*

then the Clarke subdifferential mapping $\partial^\circ_x L(\cdot, x(\cdot)) : I \rightsquigarrow E^*$ has a Gelfand integrable selector and the Gelfand integral $\int \partial^\circ_x L(t, x(t))dt$ is weakly* compact and convex with

$$s\left(v, \int_I \partial^\circ_x L(t, x(t))dt\right) = \int_I s\left(v, \partial^\circ_x L(t, x(t))\right) dt = \int_I L^\circ_x(t, x(t); v)dt$$

for every $v \in E$.

A similar result holds for Dini–Hadamard subdifferential mappings, but the proof is rather different from the one for Clarke subdifferential mappings since it involves a geometric aspect using the contingent cone and its polar.

Theorem A.1. *Under the hypothesis of Proposition A.1, if L is a normal integrand with $\partial_x^- L(t, x(t)) \neq \emptyset$ a.e. $t \in I$, then the Dini–Hadamard subdifferential mapping $\partial_x^- L(\cdot, x(\cdot)) : I \rightsquigarrow E^*$ has a Gelfand integrable selector and the Gelfand integral $\int \partial_x^- L(t, x(t))dt$ is weakly* compact and convex with*

$$s\left(v, \int_I \partial_x^- L(t, x(t))dt\right) = \int_I s\left(v, \partial_x^- L(t, x(t))\right) dt$$

for every $v \in E$.

Recall that a multifunction $\Gamma : I \rightsquigarrow E$ is said to be *measurable* if the set $\{t \in I \mid \Gamma(t) \cap O \neq \emptyset\}$ belongs to $\mathcal{L}$ for every open subset O of E. Denote by $\Gamma^0 : I \rightsquigarrow E^*$ the polar mapping of Γ defined by $\Gamma^0(t) := \Gamma(t)^0$.

Lemma A.1. *Let E be a separable Banach space. If $\Gamma : I \rightsquigarrow E$ is a measurable multifunction with nonempty closed values, then its polar mapping $\Gamma^0 : I \rightsquigarrow E^*$ has the graph in $\mathcal{L} \otimes \mathrm{Borel}(E^*, w^*)$ and Γ^0 admits a weakly* scalarly measurable selector.*

Proof. Let $\{g_n\}_{n\in\mathbb{N}}$ be a *Castaing representation* of Γ, that is, each $g_n : I \to E$ is a measurable selector of Γ such that $\mathrm{cl}\{g_n(t) \mid n \in \mathbb{N}\} = \Gamma(t)$ for every $t \in I$. Since $(t, x^*) \mapsto \langle x^*, g_n(t)\rangle$ is $\mathcal{L} \otimes \mathrm{Borel}(E^*, w^*)$-measurable for each $n \in \mathbb{N}$ (see Castaing and Valadier [19, Theorem III.36]) and $s(x^*, \Gamma(t)) = \sup_n \langle x^*, g_n(t)\rangle$ for every $x^* \in E^*$ and $t \in I$, $\mathrm{gph}\,\Gamma^0$ is $\mathcal{L} \otimes \mathrm{Borel}(E^*, w^*)$-measurable. Therefore, Γ^0 admits a $\mathrm{Borel}(E^*, w^*)$-measurable, and hence, weakly* scalarly measurable selector. □

Proof of Theorem A.1. Define the multifunction $\Gamma : I \rightsquigarrow E \times \mathbb{R}$ by

$$\Gamma(t) := K_{\mathrm{epi}\, L(t,\cdot)}(x(t), L(t, x(t))).$$

Since L is a normal integrand, the epigraph mapping $t \rightsquigarrow \mathrm{epi}\, L(t, \cdot)$ is a nonempty, closed-valued multifunction with its graph in $\mathcal{L} \otimes \mathrm{Borel}(E \times \mathbb{R}, \|\cdot\|)$; see Castaing and Valadier [19, Lemma VII.1]. It follows from Aubin and Frankowska [5, Theorem 8.5.1] that Γ is a measurable multifunction with nonempty closed values. Then by Lemma A.1, the polar mapping $\Gamma^0 : I \rightsquigarrow E^* \times \mathbb{R}$ of Γ has the graph in $\mathcal{L} \otimes \mathrm{Borel}(E^* \times \mathbb{R}, w^*)$. Define the multifunction $\Phi : I \rightsquigarrow E^* \times \mathbb{R}$ by

$$\Phi(t) := \Gamma^0(t) \cap (E^* \times \{-1\}) = \left\{(x^*, -1) \in K_{\mathrm{epi}\, L(t,\cdot)}(x(t), L(t, x(t)))^0\right\}.$$

Then $\Phi(t) \neq \emptyset$ a.e. $t \in I$ and $\mathrm{gph}\,\Phi$ belongs to $\mathcal{L} \otimes \mathrm{Borel}(E^* \times \mathbb{R}, w^*)$. Therefore, Φ admits a weakly* scalarly measurable selector, and hence, there exists a weakly* scalarly measurable function $f : I \to E^*$ such that $(f(t), -1) \in \Phi(t)$ a.e. $t \in I$. Since $f(t) \in \partial_x^- L(t, x(t))$ and the Dini–Hadamard subdifferential mapping $\partial_x^- L(\cdot, x(\cdot))$ is integrably bounded, f is Gelfand integrable. Under the assumptions of the theorem, $\partial_x^- L(t, x(t))$ is nonempty, weakly* compact, and convex a.e. $t \in I$. Therefore, the Gelfand integral of $\partial_x^- L(\cdot, x(\cdot))$ is nonempty, weakly* compact and convex, and the desired equality holds as noted in Subsection A.2. □

Remark A.1. Note that unlike Clarke directional derivatives, the lack of convexity of the function $v \mapsto L_x^-(t, x(t); v)$ leads to the failure of the equality $L_x^-(t, x(t); v) = s(v, \partial_x^- L(t, x(t)))$ even if $\partial_x^- L(t, x(t))$ is nonempty. This is a disadvantage of the

use of Dini–Hadamard subdifferentials because $\partial_x^- L(t, x(t))$ may be empty even if $L(t, \cdot)$ is Lipschitz on $x(t)+\varepsilon B$. On the other hand, if $L(t, \cdot)$ is strictly differentiable at $x(t)$, then its Dini–Hadamard subdifferential is the singleton $\{\nabla_x L(t, x(t))\}$. See also Remark 4.2 for a further discussion.

B Appendix II

B.1 Lipschitz Continuity of the Value Function

The following result is a special case of Frankowska [33, Theorem 1.2], which is an infinite-dimensional analogue of the celebrated Filippov theorem; see Filippov [30]. Note that the solution concept adopted in Frankowska [33] is a mild solution to an evolution differential inclusion involving semigroups of unbounded linear operators. As in our case the semigroup is given by the identity operator, it follows from the Lebesgue differentiation theorem that the mild solution has a strong derivative that is Bochner integrable whenever E is separable.

Lemma B.1 (Frankowska [33])**.** *Let E be a separable Banach space and $[t_0, t_1]$ be any closed interval in $\mathbb{R}_+$. If* (H_4), (H_5), *and* (H_6) *hold, and $y(\cdot) \in W^{1,1}([t_0, t_1], E)$ is such that $t \mapsto d_{\Gamma(t,y(t))}(\dot{y}(t))$ is integrable with $y(t_0) = \xi \in E$, then for every $\xi' \in E$ and $\varepsilon > 0$ there exists $x(\cdot) \in W^{1,1}([t_0, t_1], E)$ such that:*

(i) $\dot{x}(t) \in \Gamma(t, x(t))$ *a.e.* $t \in [t_0, t_1]$ *with* $x(t_0) = \xi'$*;*

(ii) $$\|x(t) - y(t)\| \le \exp\left(\int_{t_0}^{t} \gamma(s)ds\right) \times \left(\|\xi - \xi'\| + \int_{t_0}^{t} d_{\Gamma(s,y(s))}(\dot{y}(s))ds + \varepsilon(t - t_0)\right)$$
for every $t \in [t_0, t_1]$*;*

(iii) $$\|\dot{x}(t) - \dot{y}(t)\| \le \exp\left(\int_{t_0}^{t} \gamma(s)ds\right) \gamma(t)(\|\xi - \xi'\| + \varepsilon(t - t_0)) + d_{\Gamma(t,y(t))}(\dot{y}(t)) + \varepsilon$$
a.e. $t \in [t_0, t_1]$.

Proof of Theorem 3.1. Take any $(t, \xi) \in \mathbb{R}_+ \times E$. Since, by ($\mathrm{H}_6$), every admissible trajectory $x(\cdot) \in \mathcal{A}_{(t,\xi)}$ satisfies the inequality $\|\dot{x}(s)\| \le \gamma(s) + \gamma(s)\|x(s)\|$ a.e. $s \in [t, \infty)$ by (H_6), the Gronwall's inequality yields

$$\|x(s)\| \le \exp\left(\int_0^s \gamma(\tau)d\tau\right)\left(\|\xi\| + \int_0^s \gamma(\tau)d\tau\right) =: \gamma_{\|\xi\|}(s) < \infty$$

for every $s \in [t, \infty)$. It follows from (H_3) and (H_7) that

$$|L(s, x(s), \dot{x}(s))| \le l_1(s) + l_1(s)\gamma_{\|\xi\|}(s) + l_2(s)(\gamma(s) + \gamma(s)\gamma_{\|\xi\|}(s)) =: k_{\|\xi\|}(s)$$

and $k_{\|\xi\|}(\cdot)$ is integrable over $[t, \infty)$. Therefore, V is bounded. Furthermore, it follows from $|\int_t^\infty L(s, x(s), \dot{x}(s))ds| \le \int_t^\infty k_{\|\xi\|}(s)ds$ that for every $\varepsilon > 0$ there exists $t_0 \in \mathbb{R}_+$ such that $|\int_T^\infty L(s, x(s), \dot{x}(s))ds| < \varepsilon$ for every $T > t_0$ and $x(\cdot) \in \mathcal{A}_{(t,\xi)}$. This implies that $\sup_{x(\cdot)\in\mathcal{A}_{(t,\xi)}} |V(T, x(T))| \to 0$ as $T \to \infty$ for every $(t, \xi) \in \mathbb{R}_+ \times E$.

Next, we demonstrate the Lipschitz continuity of $V(t, \cdot)$. Let $\xi, \xi' \in E$ be arbitrary. Take any $\varepsilon > 0$ and $T \in [t, \infty)$. Then by (H_1) and the Bellman principle of

optimality, there exists $x(\cdot) \in \mathcal{A}_{(t,\xi)}$ such that $\int_t^T L(s, x(s), \dot{x}(s))ds + V(T, x(T)) < V(t, \xi) + \varepsilon$. It follows from Lemma B.1 that there exists $x^T(\cdot) \in W^{1,1}([t, T], E)$ such that:

(i) $\dot{x}^T(s) \in \Gamma(s, x^T(s))$ a.e. $s \in [t, T]$ with $x^T(t) = \xi'$;

(ii) $\|x^T(s) - x(s)\| \leq \exp\left(\int_t^s \gamma(\tau)d\tau\right)(\|\xi' - \xi\| + \varepsilon(s - t))$ for every $s \in [t, T]$;

(iii) $\|\dot{x}^T(s) - \dot{x}(s)\| \leq \exp\left(\int_t^s \gamma(\tau)d\tau\right)\gamma(s)(\|\xi' - \xi\| + \varepsilon(s - t)) + \varepsilon$ a.e. $s \in [t, T]$.

Take any $x_T(\cdot) \in \mathcal{A}_{(T,x^T(T))}$ and define $y_T(\cdot) \in \mathcal{A}_{(t,\xi')}$ by $y_T(\cdot) = x^T(\cdot)$ on $[t, T]$ and $y_T(\cdot) = x_T(\cdot)$ on (T, ∞). As observed in the above, we obtain

$$\lim_{T\to\infty} |V(T, x^T(T))| = \lim_{T\to\infty} |V(T, y_T(T))| \leq \lim_{T\to\infty} \sup_{z(\cdot)\in\mathcal{A}_{(t,\xi')}} |V(T, z(T))| = 0.$$

Similarly, $\lim_{T\to\infty} V(T, x(T)) = 0$. By the Bellman principle of optimality, we have

$$\begin{aligned}
V(t, \xi') - V(t, \xi) &\leq \int_t^T L(s, x^T(s), \dot{x}^T(s))ds + V(T, x^T(T)) \\
&\quad - \int_t^T L(s, x(s), \dot{x}(s))ds - V(T, x(T)) + \varepsilon \\
&\leq \int_t^T \left[l_1(s)\|x^T(s) - x(s)\| + l_2(s)\|\dot{x}^T(s) - \dot{x}(s)\|\right] ds \\
&\quad + V(T, x^T(T)) - V(T, x(T)) + \varepsilon \\
&\leq k_1(t)(\|\xi' - \xi\| + \varepsilon(T - t)) + k_2(t)(\|\xi' - \xi\| + \varepsilon(T - t)) \\
&\quad + V(T, x^T(T)) - V(T, x(T)) + \varepsilon,
\end{aligned}$$

where we set in the last inequality $k_1(t) := \int_t^\infty \exp(\int_t^s \gamma(\tau)d\tau)l_1(s)ds$ and $k_2(t) := \int_t^\infty \exp(\int_t^s \gamma(\tau)d\tau)l_2(s)\gamma(s)ds$. Since ε is arbitrary, we obtain

$$V(t, \xi') - V(t, \xi) \leq k(t)\|\xi' - \xi\| + V(T, x^T(T)) - V(T, x(T))$$

for every $T \in [t, \infty)$ with $k(t) := k_1(t) + k_2(t)$. Then $k : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous decreasing function with $k(t) \to 0$ as $t \to \infty$. Letting $T \to \infty$ in this inequality yields $V(t, \xi') - V(t, \xi) \leq k(t)\|\xi' - \xi\|$. Since the role of ξ and ξ' is interchangeable in the above argument, we have demonstrated that $V(t, \cdot)$ is Lipschitz of rank $k(t)$ on E for every $t \in \mathbb{R}_+$.

Finally, we show the lower semicontinuity of V. Toward this end, fix $t \in \mathbb{R}_+$. It suffices to show that $V(\cdot, \xi)$ is lower semicontinuous on $\mathbb{R}_+$ for every $\xi \in E$. Indeed, since we have

$$V(t', \xi) - k(t')\|\xi' - \xi\| \leq V(t', \xi')$$

for every $(t', \xi') \in \mathbb{R}_+ \times E$, taking the limit inferior in the both sides of the above inequality yields

$$\liminf_{t'\to t} V(t', \xi) \leq \liminf_{(t',\xi')\to(t,\xi)} V(t', \xi').$$

Take any $t' \in \mathbb{R}_+$. If $t' \in [t, \infty)$, then

$$V(t, \xi) \leq \int_t^{t'} L(s, x(s), \dot{x}(s))ds + V(t', x(t')) \tag{B.1}$$

for every $x(\cdot) \in \mathcal{A}_{(t,\xi)}$. Since $V(t',\cdot)$ is Lipschitz of rank $k(t') \le k(t)$, we obtain $|V(t',x(t')) - V(t',\xi)| \le k(t)\|x(t') - \xi\| \to 0$ as $t' \downarrow t$. Taking the limit inferior in the both sides of (B.1) yields

$$V(t,\xi) \le \liminf_{t' \downarrow t} V(t',\xi).$$

Similarly, if $t' \in [0,t)$ with $t > 0$, then for every $\varepsilon > 0$ there exists $y(\cdot) \in \mathcal{A}(t',\xi)$ such that

$$\int_{t'}^{t} L(s,y(s),\dot{y}(s))ds + V(t,y(t)) \le V(t',\xi) + \varepsilon \tag{B.2}$$

Since $V(t,\cdot)$ is Lipschitz of rank $k(t)$, we have

$$\begin{aligned}
&|V(t,y(t)) - V(t,\xi)| \\
&\le k(t)\|y(t) - \xi\| = k(t)\left\|\int_{t'}^{t} \dot{y}(s)ds\right\| \\
&\le k(t)\int_{t'}^{t} \left[\gamma(s) + \gamma(s)\|y(s)\|\right] ds \le k(t)\int_{t'}^{t} \left[\gamma(s) + \gamma(s)\gamma_{\|\xi\|}(s)\right] ds \to 0
\end{aligned}$$

as $t' \uparrow t$ and

$$\left|\int_{t'}^{t} L(s,y(s),\dot{y}(s))ds\right| \le \int_{t'}^{t} k_{\|\xi\|}(s)ds \to 0$$

as $t' \uparrow t$. Hence, taking the limit inferior in (B.2) yields

$$V(t,\xi) \le \liminf_{t' \uparrow t} V(t',\xi) + \varepsilon.$$

Since ε is arbitrary, we obtain

$$V(t,\xi) \le \liminf_{t' \to t} V(t',\xi).$$

Therefore, $V(\cdot,\xi)$ is lower semicontinuous at every $t \in \mathbb{R}_+$. □

B.2 Subdifferentiability of the Value Function

Denote by $o(h) > 0$ the Landau symbol with $\lim_{h\downarrow 0} h^{-1}o(h) = 0$.

Lemma B.2. *Suppose that* (H_4), (H_5'), *and* (H_6) *hold. Let $t \in \mathbb{R}_+$ be such that the strong derivative $\dot{x}_0(t)$ exists and $v \in \Gamma(t,x_0(t))$ be arbitrarily fixed. Then for every $h > 0$ there exists $x_h(\cdot) \in W^{1,1}([t,t+h],E)$ such that:*

(i) $\dot{x}_h(s) \in \Gamma(s,x_h(s))$ *a.e.* $s \in [t,t+h]$ *with* $x_h(t+h) = x_0(t+h)$*;*

(ii) $\|x_h(t) - x_0(t) - h(\dot{x}_0(t) - v)\| = o(h)$*;*

(iii) $\|\dot{x}_h(\cdot) - v\|_{L^1([t,t+h])} = o(h)$.

Proof. Define $y_h(s) := x_0(t+h) - sv$ for $s \in [0,h]$ and the multifunction $\Gamma_h : \mathbb{R} \times E \rightsquigarrow E$ by

$$\Gamma_h(s,x) := \begin{cases} -\Gamma(t+h-s,x) & \text{if } s \in [0,h], \\ -\Gamma(t,x) & \text{if } s \in (h,\infty). \end{cases}$$

By (H_6), we have $\Gamma_h(s, y_h(s)) \subset \Gamma_h(s, x_0(t)) + \gamma(t+h-s)(\|x_0(t+h) - x_0(t)\| + h\|v\|)B$ for every $s \in [0, h]$, which yields the inequality

$$d_{\Gamma_h(s,y_h(s))}(\dot{y}_h(s)) \le d_{\Gamma_h(s,x_0(t))}(-v) + \gamma(t+h-s)(\|x_0(t+h) - x_0(t)\| + h\|v\|).$$

Since the multifunction $(s, h) \rightsquigarrow \Gamma_h(s, x_0(t))$ is lower semicontinuous, the distance function $(s, h) \mapsto d_{\Gamma_h(s,x_0(t))}(-v)$ is upper semicontinuous; see Aubin and Frankowska [5, Corollary 1.4.17]. Let $\varphi(s, h) := d_{\Gamma_h(s,x_0(t))}(-v)$ and $\hat{\varphi}(h) := \sup_{s\in[0,h]} \varphi(s, h)$. Then for every $\varepsilon > 0$ and $h \ge 0$ there exists $s_h \in [0, h]$ such that $\hat{\varphi}(h) < \varphi(s_h, h) + \varepsilon$. Since $s_h \to 0$ as $h \downarrow 0$ and φ is upper semicontinuous at the origin with $\varphi(0, 0) = d_{\Gamma(t,x_0(t))}(v) = 0$, taking the limit superior of the above inequality yields $\limsup_{h\downarrow 0} \hat{\varphi}(h) \le \limsup_{h\downarrow 0} \varphi(s_h, h) + \varepsilon \le \varepsilon$. Since ε is arbitrary, we have $\lim_{h\downarrow 0} \hat{\varphi}(h) = 0$. Consequently, it follows from the inequality $\|x_0(t+h) - x_0(t)\| \le h\|\dot{x}_0(t)\| + o(h)$ that

$$d_{\Gamma_h(s,y_h(s))}(\dot{y}_h(s)) \le \hat{\varphi}(h) + \gamma(t+h-s)(h(\|\dot{x}_0(t)\| + \|v\|) + o(h)).$$

By Lemma B.1 applied with $\varepsilon = h$, there exists $z(\cdot) \in W^{1,1}([0, h], E)$ such that $\dot{z}(s) \in \Gamma_h(s, z(s))$ a.e. $s \in [0, h]$ with $z(0) = y_h(0) = x_0(t+h)$ satisfying

$$\begin{aligned}
\|z(s) - y_h(s)\| &\le \exp\left(\int_0^h \gamma(t+h-\tau)d\tau\right)\left(\int_0^h d_{\Gamma_h(\tau,y_h(\tau))}(\dot{y}_h(\tau))d\tau + h^2\right) \\
&\le \exp\left(\int_t^{t+h} \gamma(\tau)d\tau\right) \\
&\quad \times \left(h\hat{\varphi}(h) + (h(\|\dot{x}_0(t)\| + \|v\|) + o(h))\int_t^{t+h} \gamma(\tau)d\tau + h^2\right) = o(h)
\end{aligned}$$

and

$$\begin{aligned}
\|\dot{z}(s) - \dot{y}_h(s)\| &\le h^2 \exp\left(\int_0^h \gamma(t+h-\tau)d\tau\right)\gamma(t+h-s) + d_{\Gamma_h(s,y_h(s))}(\dot{y}_h(s)) + h \\
&\le h^2 \exp\left(\int_t^{t+h} \gamma(\tau)d\tau\right)\gamma(t+h-s) + \hat{\varphi}(h) \\
&\quad + \gamma(t+h-s)(h(\|\dot{x}_0(t)\| + \|v\|) + o(h)) + h
\end{aligned}$$

for a.e. $s \in [0, h]$. Integrating the both sides of the above inequality over $[0, h]$ yields

$$\begin{aligned}
&\|\dot{z}(\cdot) - \dot{y}_h(\cdot)\|_{L^1([0,h])} \\
&= \int_t^{t+h} \gamma(\tau)d\tau \left(h^2 \exp\left(\int_t^{t+h} \gamma(\tau)d\tau\right) + h(\|\dot{x}_0(t)\| + \|v\|) + o(h)\right) \\
&\quad + h\hat{\varphi}(h) + h^2 = o(h).
\end{aligned}$$

Set $x_h(\tau) := z(t+h-\tau)$ for $\tau \in [t, t+h]$. Then $x_h(t+h) = z(0) = x_0(t+h)$ and $\dot{x}_h(\tau) = -\dot{z}(t+h-\tau) \in -\Gamma_h(t+h-\tau, x_h(\tau)) = \Gamma(\tau, x_h(\tau))$ a.e. $\tau \in [t, t+h]$. Thus, condition (i) is verified. Since $\|x_h(t) - y_h(h)\| = \|z(h) - y_h(h)\| = o(h)$, we have $\|x_h(t) - x_0(t+h) + hv\| = o(h)$. Hence, $\|x_h(t) - x_0(t) - h(\dot{x}_0(t) - v)\| \le \|x_h(t) - x_0(t+h) + hv\| + \|x_0(t+h) - x_0(t) - h\dot{x}_0(t)\| = o(h)$ and we obtain

condition (ii). In view of $\dot{x}_h(\tau) = -\dot{z}(t+h-\tau)$ and $\dot{y}_h(t+h-\tau) = -v$, we obtain $\|\dot{x}_h(\cdot) - v\|_{L^1([t,t+h])} = \|\dot{y}_h(\cdot) - \dot{z}(\cdot)\|_{L^1([0,h])} = o(h)$, which implies condition (iii). $\square$

Proof of Theorem 3.2. (i): Let $t > 0$, $v \in \Gamma(t, x_0(t))$, and $x_h(\cdot) \in W^{1,1}([t, t+h], E)$ be as in the claim of Lemma B.2. By condition (ii) of Lemma B.2, for every $s \in [t, t+h]$ we have

$$\begin{aligned}
\|x_h(s) - x_0(t)\| &\le \|x_h(s) - x_h(t)\| + \|x_h(t) - x_0(t)\| \\
&\le h\|\dot{x}_h(t)\| + o(h) + h\|\dot{x}_0(t) - v\| \\
&\le h(\gamma(t) + \gamma(t)\|x_h(t)\|) + h\|\dot{x}_0(t) - v\| + o(h) \\
&\le h(\gamma(t) + \gamma(t)(\|x_0(t)\| + h\|\dot{x}_0(t) - v\| + o(h)) + h\|\dot{x}_0(t) - v\| \\
&\qquad + o(h) \\
&= h(\gamma(t) + \gamma(t)\|x_0(t)\|) + h\|\dot{x}_0(t) - v\| + o(h),
\end{aligned}$$

which yields the following estimates:

$$\begin{aligned}
&\left| \int_t^{t+h} L(s, x_h(s), \dot{x}_h(s))ds - \int_t^{t+h} L(s, x_0(t), v)ds \right| \\
&\le \int_t^{t+h} [l_1(s)\|x_h(s) - x_0(t)\| + l_2(s)\|\dot{x}_h(s) - v\|]\, ds \\
&\le (h(\gamma(t) + \gamma(t)\|x_0(t)\|) + h\|\dot{x}_0(t) - v\| + o(h)) \int_t^{t+h} l_1(s)ds \\
&\qquad + \sup_{s \in [t,t+h]} l_2(s)\|\dot{x}_h(\cdot) - v\|_{L^1([t,t+h])} = o(h).
\end{aligned}$$

By the separability of E and Frankowska, Plaskacz and Rzeżuchowski [35, Theorem 2.5], there exists a subset I of $\mathbb{R}_+$ such that the Lebesgue measure of its complement $\mathbb{R}_+ \setminus I$ is zero with $\lim_{h \downarrow 0} h^{-1} \int_t^{t+h} L(s, x, v)ds = L(t, x, v)$ for every $(t, x, v) \in I \times E \times E$. We thus obtain

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} L(s, x_h(s), \dot{x}_h(s))ds = \lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} L(s, x_0(t), v)ds = L(t, x_0(t), v)$$

for every $t \in I$. Let $t \in \mathbb{R}_+$ be a Lebesgue point of $L(\cdot, x_0(\cdot), \dot{x}_0(\cdot))$. By the Bellman principle of optimality, we have

$$V(t, x_h(t)) \le \int_t^{t+h} L(s, x_h(s), \dot{x}_h(s))ds + V(t+h, x_0(t+h)).$$

Subtracting $V(t, x_0(t)) = \int_t^{t+h} L(s, x_0(s), \dot{x}_0(s))ds + V(t+h, x_0(t+h))$ from the both sides of the above inequality yields

$$\begin{aligned}
&\int_t^{t+h} L(s, x_h(s), \dot{x}_h(s))ds - \int_t^{t+h} L(s, x_0(s), \dot{x}_0(s))ds \\
&\ge V(t, x_h(t)) - V(t, x_0(t)) \ge V(t, x_0(t) + h(\dot{x}_0(t) - v)) - V(t, x_0(t)) - o(h)
\end{aligned}$$

because of the Lipschitz continuity of $V(t, \cdot)$. Dividing the both sides of the above inequality by $h > 0$ and taking the limit inferior as $h \to 0$ yield the inequality

$$V_x^-(t, x_0(t); \dot{x}_0(t) - v) \le L(t, x_0(t), v) - L(t, x_0(t), \dot{x}_0(t)).$$

Since $v \in \Gamma(t, x_0(t))$ is arbitrary, the above holds true for any such v.

(ii): Take any $x^* \in \partial_x^- V(t, x_0(t))$. If $u \in K_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$, then there exist a sequence $\{\theta_n\}_{n\in\mathbb{N}}$ of positive real numbers with $\theta_n \to 0$ and a sequence $\{u_n\}_{n\in\mathbb{N}}$ in E with $u_n \to u$ such that $\dot{x}_0(t) + \theta_n u_n \in \Gamma(t, x_0(t))$ for each $n \in \mathbb{N}$. Since it follows from condition (i) that $V_x^-(t, x_0(t); -\theta_n u_n) \leq L(t, x_0(t), \dot{x}_0(t) + \theta_n u_n) - L(t, x_0(t), \dot{x}_0(t))$, we have

$$\langle x^*, -u_n \rangle \leq \frac{L(t, x_0(t), \dot{x}_0(t) + \theta_n u_n) - L(t, x_0(t), \dot{x}_0(t))}{\theta_n}.$$

Letting $n \to \infty$ in the both sides of the above inequality yields

$$\langle -x^*, u \rangle \leq L_y^+(t, x_0(t), \dot{x}_0(t); u) \leq L_y^\circ(t, x_0(t), \dot{x}_0(t); u) \tag{B.3}$$

for every $u \in K_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$. Suppose, by way of contradiction, that $-x^* \notin \partial_y^\circ L(t, x_0(t), \dot{x}_0(t)) + N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$. Since $\partial_y^\circ L(t, x_0(t), \dot{x}_0(t))$ is weakly* compact and convex and $N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ is weakly* closed and convex, $\partial_y^\circ L(t, x_0(t), \dot{x}_0(t)) + N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ is weakly* closed and convex. Then by the separation theorem, there exists $v \in E$ such that

$$\langle -x^*, v \rangle > \sup_{y^* \in \partial_y^\circ L(t,x_0(t),\dot{x}_0(t))} \langle y^*, v \rangle + \sup_{z^* \in N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))} \langle z^*, v \rangle.$$

Since $N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ is a cone in E^*, we must have $\langle z^*, v \rangle \leq 0$ for every $z^* \in N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$. This means that $v \in T_{\Gamma(t,x_0(t))}(\dot{x}_0(t)) \subset K_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$ by the bipolar theorem; see Aubin and Frankowska [5, Theorem 2.4.3]. Since the support function of the Clarke subdifferential $\partial_y^\circ L(t, x_0(t), \dot{x}_0(t))$ coincides with the Clarke directional derivative $L_y^\circ(t, x_0(t), \dot{x}_0(t); v)$, the inequality above finally implies that $\langle -x^*, v \rangle > L_y^\circ(t, x_0(t), \dot{x}_0(t); v)$ with $v \in K_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$, in contradiction with inequality (B.3). Consequently, we have $-x^* \in \partial_y^\circ L(t, x_0(t), \dot{x}_0(t)) + N_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$. If $L(t, x_0(t), \cdot)$ is Gateaux differentiable at $\dot{x}_0(t)$, then (B.3) can be replaced by the inequality

$$\langle -x^*, u \rangle \leq \nabla_y L(t, x_0(t), \dot{x}_0(t); u)$$

for every $u \in K_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$. This implies that the above argument is also valid when we replace $\partial_y^\circ L(t, x_0(t), \dot{x}_0(t))$ and $L_y^\circ(t, x_0(t), \dot{x}_0(t); v)$ respectively by $\nabla_y L(t, x_0(t), \dot{x}_0(t))$ and $\langle \nabla_y L(t, x_0(t), \dot{x}_0(t)), v \rangle$. □

B.3 Proof of Necessary Conditions for Optimality

It should be underlined that unlike the real-valued case, locally absolutely continuous functions with values in Banach spaces fail to be strongly differentiable almost everywhere; see Petrakis and Uhl [45, Examples 1 and 2] or Deimling [26, Example 4.2] for such examples. The failure of the strong differentiability of locally absolutely continuous functions disappears under the reflexivity assumption. Specifically, every locally absolutely function $p : \mathbb{R}_+ \to E^*$ has the Bochner integrable strong derivative $\dot{p}(t)$ a.e. $t \in \mathbb{R}_+ \setminus \{0\}$ with $p(t) = \int_0^t \dot{p}(s)ds + p(0)$ for every $t \in \mathbb{R}_+$ whenever E is reflexive; see Kōmura [42, Lemma, p. 505].

We construct an adjoint variable $p : \mathbb{R}_+ \to E^*$ as a locally absolutely continuous function to express optimality conditions. However, we dispense with the reflexivity of E. The weak* differentiability of locally absolutely continuous functions is fundamental in the sequel and is virtually contained in the argument of the

proof of Kōmura [42, Lemma]. We provide a proof for the sake of completeness to make clear why the reflexivity of E is irrelevant to weak* differentiability. See also Ambrosio and Kirchheim [2, Theorem 3.5] for a strengthened version of the weak* differentiability of Lipschitz functions.

Lemma B.3 (Kōmura [42])**.** *Let E be a separable Banach space. Then every locally absolutely continuous function $p : \mathbb{R}_+ \to E^*$ possesses the weak* derivative $\dot{p}(t)$ a.e. $t \in \mathbb{R}_+$.*

Proof. Define the variation of $p : \mathbb{R}_+ \to E^*$ over the compact interval $[0, \tau]$ by $\mathrm{var}(p, [0, \tau]) := \sup \sum_{i=1}^n \|p(t_i) - p(\tau_i)\|$, where the supremum is taken over all finite sets of points $t_i, \tau_i \in [0, \tau]$ with $0 \le t_1 < \tau_1 \le t_2 < \tau_2 \le \cdots \le t_n < \tau_n \le \tau$. Since p is locally absolutely continuous, $\mathrm{var}(p, [0, \tau]) < \infty$ for every $\tau > 0$. Define $p_h(t) := h^{-1}(p(t+h) - p(t))$ for $t \in [0, \tau]$ and $h \neq 0$, and $P^+(t) := \limsup_{h \downarrow 0} \|p_h(t)\|$ and $P^-(t) := \limsup_{h \uparrow 0} \|p_h(t)\|$. Since p is continuous, so is p_h. Hence, P^+ and P^- are measurable on $[0, \tau]$. We claim that $P^+(t)$ and $P^-(t)$ are finite a.e. $t \in [0, \tau]$. Suppose to the contrary that the Lebesgue measure λ of the set $\{t \in [0, \tau] \mid P^+(t) = \infty\}$ is positive. Let

$$A_n := \left\{ t \in [0, \tau] \;\middle|\; \begin{array}{l} \sup \left[\left\| \dfrac{p(t+h) - p(t)}{h} \right\| \;\middle|\; h \ge \dfrac{1}{n},\ 0 \le t < t + h \le \tau \right] \\ \ge \dfrac{2}{\lambda} \mathrm{var}(p, [0, \tau]) \end{array} \right\}.$$

Then each A_n is a closed set and $\{t \in [0, \tau] \mid P^+(t) = \infty\} \subset \bigcup_{n \in \mathbb{N}} A_n$. Since $\{A_n\}_{n \in \mathbb{N}}$ is an increasing sequence, $|A_n| > \lambda/2$ for some $n \in \mathbb{N}$, where $|A_n|$ denotes the Lebesgue measure of A_n. Let $\{t_i\}_{i \in \mathbb{N}}$ and $\{h_i\}_{i \in \mathbb{N}}$ be defined inductively by

$$\begin{aligned} &t_1 := \inf A_n, \\ &t_{i+1} := \inf\{t \in A_n \mid t \ge t_i + h_i\}, \text{ and} \\ &h_i := \sup \left\{ h > 0 \mid t_i + h \le \tau,\ \left\| \frac{p(t_i + h) - p(t_i)}{h} \right\| \ge \frac{2}{\lambda} \mathrm{var}(p, [0, \tau]) \right\}. \end{aligned}$$

Then by construction, we have $A_n \subset \bigcup_{i \in \mathbb{N}} [t_i, t_i + h_i]$, and hence, $|A_n| \le \sum_{i \in \mathbb{N}} h_i$. Consequently,

$$\sum_{i \in \mathbb{N}} \|p(t_i + h_i) - p(t_i)\| \ge \frac{2}{\lambda} \mathrm{var}(p, [0, \tau]) \sum_{i \in \mathbb{N}} h_i \ge \frac{2}{\lambda} \mathrm{var}(p, [0, \tau]) |A_n| > \mathrm{var}(p, [0, \tau]),$$

a contradiction. In the same way we show that the set $\{t \in [0, \tau] \mid P^-(t) = \infty\}$ is of Lebesgue measure zero. Therefore, there exists a null set $N_0 \subset [0, \tau]$ such that for every $t \in [0, \tau] \setminus N_0$ the set $\{p_h(t) \mid h \neq 0\}$ is bounded in E^*.

In view of the separability of E, there is a countable dense subset $\{v_i\}_{i \in \mathbb{N}}$ of E. Since each scalar function $\varphi_i(t) := \langle p(t), v_i \rangle$ is absolutely continuous on $[0, \tau]$, its derivative $\dot{\varphi}_i(t)$ exists except at a point of a null set $N_i \subset [0, \tau]$. This means that $\dot{\varphi}_i(t) = \lim_{h \to 0} \langle p_h(t), v_i \rangle$ for every $i \in \mathbb{N}$ and $t \in [0, \tau] \setminus \bigcup_{j \in \mathbb{N}} N_j$. Recalling that $\{p_h(t) \mid h \neq 0\}$ is relatively weakly* compact for every $t \in [0, \tau] \setminus N_0$, it has a subnet (which we do not relabel) that converges weakly* to an element in E^*. Therefore, $\dot{p}(t) = w^*\text{-}\lim_{h \to 0} p_h(t)$ exists for every $t \in [0, \tau] \setminus \bigcup_{i=0}^\infty N_i$ because $\{v_i\}_{i \in \mathbb{N}}$ is a total family of E. $\square$

Proof of Theorem 4.1. Let $t \in \mathbb{R}_+$ be arbitrarily given and $\eta > 0$ be as in (H_8). Take any $x^* \in \partial_x^- V(0, x_0(0))$ and let $f : \mathbb{R}_+ \to E^*$ be a Gelfand integrable selector of the Dini–Hadamard superdifferential mapping $s \rightsquigarrow \partial_x^+ L(s, x_0(s), \dot{x}_0(s))$, whose existence is guaranteed in Theorem A.1. Define $p(t) = \int_0^t f(s)ds - x^*$ as a Gelfand integral. We claim that $-p(t) \in \partial_x^- V(t, x_0(t))$. To this end, fix any $v \in E$ and consider the local perturbation of $x_0(\cdot)$ over $[0, t]$ given by $x_\theta(s) := x_0(s) + \theta v$ for $s \in [0, t]$. By construction, $\dot{x}_\theta(s) = \dot{x}_0(s)$ a.e. $s \in [0, t]$ and $x_\theta(s) \in x_0(s) + \eta B$ whenever $0 < \theta \leq (1+\|v\|)^{-1}\eta$, and hence, $(x_\theta(s), \dot{x}_\theta(s)) \in \operatorname{gph} \Gamma(s, \cdot)$ a.e. $s \in [0, t]$. By the Bellman principle of optimality, we have

$$V(0, x_\theta(0)) \leq \int_0^t L(s, x_\theta(s), \dot{x}_\theta(s))ds + V(t, x_\theta(t)).$$

Subtracting $V(0, x_0(0)) = \int_0^t L(s, x_0(s), \dot{x}_0(s))ds + V(t, x_0(t))$ from the both sides of the above inequality yields

$$\begin{aligned} V(0, x_\theta(0)) - V(0, x_0(0)) \leq \int_0^t & \left[L(s, x_\theta(s), \dot{x}_\theta(s)) - L(s, x_0(s), \dot{x}_0(s))\right] ds \\ & + V(t, x_\theta(t)) - V(t, x_0(t)). \end{aligned}$$

Let $\{\theta_n\}_{n\in\mathbb{N}}$ be a sequence of positive real numbers with $\theta_n \to 0$ such that

$$V_x^-(t, x_0(t); v) = \lim_{n\to\infty} \frac{V(t, x_0(t) + \theta_n v) - V(t, x_0(t))}{\theta_n}.$$

Dividing the both sides of the above inequality by θ_n and taking the limit as $n \to \infty$ yields

$$\begin{aligned} V_x^-(0, x_0(0); v) &\leq \int_0^t L_x^+(s, x_0(s), \dot{x}_0(s); v)ds + V_x^-(t, x_0(t); v) \\ &\leq \int_0^t \langle f(s), v\rangle ds + V_x^-(t, x_0(t); v) \\ &= \langle p(t), v\rangle + \langle x^*, v\rangle + V_x^-(t, x_0(t); v) \end{aligned}$$

for every $v \in E$, where we employ the Lebesgue dominated convergence theorem and Fatou's lemma to derive that

$$\begin{aligned} & \limsup_{n\to\infty} \int_0^t \frac{L(s, x_0(s) + \theta_n v, \dot{x}_0(s)) - L(s, x_0(s), \dot{x}_0(s))}{\theta_n} ds \\ \leq & \int_0^t L_x^+(s, x_0(s), \dot{x}_0(s); v)ds. \end{aligned}$$

On the other hand, $\langle x^*, v\rangle \leq V_x^-(0, x_0(0); v)$. Hence, $\langle -p(t), v\rangle \leq V_x^-(t, x_0(t); v)$ for every $v \in E$ and thus our claim is true.

Since $\langle p(t), y\rangle = \int_0^t \langle f(s), y\rangle ds - \langle x^*, y\rangle$ for every $t \in \mathbb{R}_+$ and $y \in E$ with $|\langle f(s), y\rangle| \leq \|f(s)\|\|y\| \leq l_1(s)\|y\|$, we get $|\langle p(t+h) - p(t), y\rangle| \leq \|y\| \int_t^{t+h} l_1(s)ds$, and therefore, $\|p(t+h) - p(t)\| \leq \int_t^{t+h} l_1(s)ds$ for every $h > 0$. This means that the function $p : \mathbb{R}_+ \to E^*$ constructed above is locally absolutely continuous. In view of Lemma B.3, the weak* derivative $\dot{p}(t) = f(t)$ exists a.e. $t \in \mathbb{R}_+$. This demonstrates that the adjoint inclusions (i) and (iii) hold. Since Theorem 3.2

and condition (i) yield $\langle p(t), \dot{x}_0(t)\rangle - L(t, x_0(t), \dot{x}_0(t)) \geq \langle p(t), y\rangle - L(t, x_0(t), y)$ for every $y \in \Gamma(t, x_0(t))$, the maximum principle (iv) holds. Thus, for a.e. $t \in \mathbb{R}_+$ and every $v \in T_{\Gamma(t,x_0(t))}(\dot{x}_0(t))$, we have $\langle p(t), v\rangle \leq L_y^-(t, x_0(t), \dot{x}_0(t); v) \leq L_y^\circ(t, x_0(t), \dot{x}_0(t); v)$ and condition (ii) follows from the separation argument as in the proof of Theorem 3.2(ii). To verify the transversality condition (v) at infinity, recall that by Theorem 3.1, $V(t, \cdot)$ is Lipschitz of rank $k(t)$ with $k(t) \to 0$ as $t \to \infty$. Therefore, $\|p(t)\| \leq k(t) \to 0$. □

Proof of Theorem 4.2. Let $t \in \mathbb{R}_+$ and $v \in E$. Consider the variational equation:

$$\dot{w}(s) = \nabla_x f(s, x_0(s), u_0(s))w(s) \text{ a.e. } s \in [0, t], \quad w(t) = v. \tag{B.4}$$

In view of the separability of E and (H_9), a unique mild solution $w(\cdot) \in W^{1,1}([0, t], E)$ to (B.4) satisfies

$$w(s) = v - \int_s^t \nabla_x f(\tau, x_0(\tau), u_0(\tau))w(\tau)d\tau \quad \text{for every } s \in [0, t]$$

and its strong derivative $\dot{w}(s)$ exists a.e. $s \in [0, t]$ and satisfies (B.4) by the Lebesgue differentiation theorem. It follows from Frankowska [33, Theorem 4.2] (applied to $F(s, x) \equiv \{f(s, x, u_0(s))\}$ and $A \equiv 0$) that for every $\theta > 0$ there exists a mild solution $x_\theta(\cdot)$ to

$$\dot{x}(s) = f(s, x(s), u_0(s)) \text{ a.e. } s \in [0, t], \quad x(t) = x_0(t) + \theta v$$

such that $(x_\theta(s) - x_0(s))/\theta \to w(s)$ uniformly in $s \in [0, t]$ as $\theta \to 0$.

Take any $x^* \in \partial_x^- V(0, x_0(0))$ and let $g : \mathbb{R}_+ \to E^*$ be a locally Bochner integrable selector from $s \rightsquigarrow \partial_x^+ \tilde{L}(s, x_0(s), u_0(s))$. Consider the adjoint system:

$$-\dot{p}(s) = \nabla_x f(s, x_0(s), u_0(s))^* p(s) - g(s) \text{ a.e. } s \in [0, t], \quad p(0) = -x^*. \tag{B.5}$$

Since the mapping $(s, x, u) \mapsto \nabla_x f(s, x, u)^*$ has separable values in E^* in view of (H_9), a unique mild solution to (B.5) does exist. As in the proof of Theorem 4.1 (via the Bellman principle of optimality), we obtain the inequality

$$\begin{aligned} & V(0, x_\theta(0)) - V(0, x_0(0)) \\ \leq & \int_0^t [L(s, x_\theta(s), f(s, x_\theta(s), u_0(s)) - L(s, x_0(s), f(s, x_0(s), u_0(s))]\, ds \\ & + V(t, x_0(t) + \theta v) - V(t, x_0(t)). \end{aligned}$$

Divide the both sides of the above inequality by θ and let $\theta \to 0$ to get

$$\langle x^*, w(0)\rangle \leq \int_0^t \langle g(s), w(s)\rangle ds + V_x^-(t, x_0(t); v).$$

It follows from the a.e. strong differentiability of $p(\cdot)$ and $w(\cdot)$ that

$$\begin{aligned} \int_0^t \langle g(s), w(s)\rangle ds &= \int_0^t \langle \dot{p}(s) + \nabla_x f(s, x_0(s), u_0(s))^* p(s), w(s)\rangle ds \\ &= \int_0^t [\langle \dot{p}(s), w(s)\rangle + \langle p(s), \nabla_x f(s, x_0(s), u_0(s))w(s)\rangle]\, ds \end{aligned}$$

$$= \int_0^t [\langle \dot{p}(s), w(s)\rangle + \langle p(s), \dot{w}(s)\rangle]\, ds$$
$$= \int_0^t \frac{d}{ds}\langle p(s), w(s)\rangle ds = \langle p(t), v\rangle - \langle p(0), w(0)\rangle.$$

Hence, $\langle -p(t), v\rangle \leq V_x^-(t, x_0(t); v)$ for every $v \in E$. This means that $-p(t) \in \partial_x^- V(t, x_0(t))$. This being true for every $t \in \mathbb{R}_+$, we deduce the adjoint inclusions (i) and (iii). The rest of the conditions follows as in the proof of Theorem 4.1. □

References

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

[2] Ambrosio, L. and B. Kirchheim, "Rectifiable sets in metric and Banach spaces", *Math. Ann.* **318** (2000), 527–555.

[3] Askenazy, P. and C. Le Van, "A model of optimal growth strategy", *J. Econom. Theory* **85** (1999), 24–51.

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

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

[6] Barbu, V., "Convex control problems and Hamiltonian systems on infinite intervals", *SIAM J. Control Optim.* **16** (1978), 895–911.

[7] Benveniste, L. M. and J. A. Scheinkman, "On the differentiability of the value function in dynamic models of economics", *Econometrica* **47** (1979), 727–732.

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

[9] Bonnisseau, J. M. and C. Le Van, "On the subdifferential of the value function in economic optimization problems", *J. Math. Econom.* **25** (1996), 55–73.

[10] Boucekkine, R., C. Camacho and G. Fabbric, "Spatial dynamics and convergence: The spatial AK model", *J. Econom. Theory* **148** (2013), 2719–2736.

[11] Boucekkine, R., C. Camacho and B. Zou, "Bridging the gap between growth theory and the new economic geography: The Spatial Ramsey model", *Macroeconom. Dynamics* **13** (2009), 20–45.

[12] Brito, P., "The dynamics of growth and distribution in a spatially heterogeneous world", Working Papers Department of Economics 2004/14, ISEG, Lisbon School of Economics and Management, Department of Economics, Universidade de Lisboa, `https://econpapers.repec.org/paper/iseisegwp/wp142004.htm`

[13] Brock, W. A., A. Xepapadeas and A. N. Yannacopoulos, "Optimal agglomerations in dynamic economics", *J. Math. Econom.* **53** (2014), 1–15.

[14] Camacho, C., B. Zou and M. Briani, "On the dynamics of capital accumulation across space", *European J. Oper. Res.* **186** (2008), 451–465.

[15] Cannarsa, P. and H. Frankowska, "Value function and optimality conditions for semilinear control problems", *Appl. Math. Optim.* **26** (1992), 139–169.

[16] Cannarsa, P. and H. Frankowska, "Value functions, relaxation, and transversality conditions in infinite horizon optimal control", *J. Math. Anal. Appl.* **457** (2018), 1188–1217.

[17] Carlson, D. A., A. B. Haurie and A. Leizarowitz, *Infinite Horizon Optimal Control*, 2nd edn., Springer, Berlin, 1991.

[18] Cascales, B., V. Kadets and J. Rodríguez, "The Gelfand integral for multivalued functions", *J. Convex Anal.* **18** (2011), 873–895.

[19] Castaing, C. and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes Math. **580** (1977), Springer, Berlin.

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

[21] Clarke, F. H. and R. B. Vinter, "Local optimality conditions and Lipschitzian solutions to the Hamilton–Jacobi equation", *SIAM J. Control Optim.* **21** (1983), 856–870.

[22] Clarke, F. H. and R. B. Vinter, "The relationship between the maximum principle and dynamic programming", *SIAM J. Control Optim.* **25** (1987), 1291–1311.

[23] Crandall, M. G., L. C. Evans and P. L. Lions, "Some properties of viscosity solutions of Hamilton–Jacobi equations", *Trans. Amer. Math. Soc.* **282** (1984), 487–502.

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

[25] Davidson, R. and R. Harris, "Non-convexities in continuous-time investment theory", *Rev. Econ. Stud.* **48** (1981), 235–253.

[26] Deimling, K., *Multivalued Differential Equations*, Walter de Gruyter, Berlin, 1992.

[27] Dunford, N. and J. T. Schwartz, *Linear Operators, Part I: General Theory*, Wiley, New York, 1958.

[28] Fabian, M. J., P. Habala, P. Hájek, V. Montesinos, and V. Zizler, *Banach Space Theory: The Basic for Linear and Nonlinear Analysis*, Springer, Berlin, 2011.

[29] Fattorini, H. O., *Infinite Dimensional Optimization and Control Theory*, Cambridge Univ. Press, Cambridge, 1999.

[30] Filippov, A. F., "Classical solutions of differential equations with multivalued right-hand side", *SIAM J. Control* **5** (1967), 609–621.

[31] Frankowska, H., "Optimal trajectories associated with a solution of the contingent Hamilton–Jacobi equation", *Appl. Math. Optim.* **19** (1989), 291–311.

[32] Frankowska, H., "Hamilton–Jacobi equations: Viscosity solutions and generalized gradients", *J. Math. Anal. Appl.* **141** (1989), 21–26.

[33] Frankowska, H., "A priori estimates for operational differential inclusions", *J. Differential Equations* **84** (1990), 100–128.

[34] Frankowska, H., *Value Function in Optimal Control*, ICTP Lecture Note Series **8** (2002), `http://users.ictp.it/~pub_off/lectures/lns008/Frankowska/Frankowska02.pdf`.

[35] Frankowska, H., S. Plaskacz and T. Rzeżuchowski, "Measurable viability theorems and the Hamilton–Jacobi–Bellman equations", *J. Differential Equations* **116** (1995), 263–305.

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

[37] Hartl, R. F. and P. M. Kort, "History dependence without unstable steady state: A non-differentiable framework", *J. Math. Econom.* **39** (2003), 891–900.

[38] Hotelling, H., "Stability in competition", *Econom. J.* **39** (1929), 41–57.

[39] Ioffe, A. D., "Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps", *Nonlinear Anal.* **8** (1984), 517–539.

[40] Ioffe, A. D., "Euler–Lagrange and Hamiltonian formalism in dynamic optimization", *Trans. Amer. Math. Soc.* **349** (1997), 2871–2900.

[41] Ioffe, A. D., *Variational Analysis of Regular Mappings: Theory and Applications*, Springer, Cham, Switzerland, 2017.

[42] Kōmura, Y., "Nonlinear semi-groups in Hilbert space", *J. Math. Soc. Japan* **19** (1967), 493–507.

[43] Lions, P. L., *Generalized solutions of Hamilton–Jacobi equations*, Pitman, London, 1982.

[44] Magill, M. J. P., "Pricing infinite horizon programs", *J. Math. Anal. Appl.* **88** (1982), 398–421.

[45] Petrakis, M. and J. J. Uhl, Jr., "Differentiation in Banach spaces", in: S. T. L. Choy, J. P. Jesudason, P. Y. Lee (eds.), *Proceedings of the Analysis Conference, Singapore 1986*, North-Holland, Amsterdam, 1988, pp. 219–241.

[46] Ramsey, F. P., "A mathematical theory of saving", *Economic J.* **38** (1928), 543–559.

[47] Rincón-Zapatero, J. P and M. S. Santos, "Differentiability of the value function in continuous-time economic models", *J. Math. Anal. Appl.* **394** (2012), 305–323.

[48] Rockafellar, R. T., "Conjugate convex functions in optimal control and the calculus of variations", *J. Math. Anal. Appl.* **32** (1970), 174–222.

[49] Sagara, N., "Value functions and transversality conditions for infinite-horizon optimal control problems", *Set-Valued Var. Anal.* **18** (2010), 1–28.

[50] Seirstad, A. and K. Sydsæter, *Optimal Control Theory with Economic Applications*, Elsevier, Amsterdam, 1987.

[51] Skiba, A. K., "Optimal growth with a convex-concave production function", *Econometrica* **48** (1978), 527–539.

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

[53] Takekuma, S.-I., "On duality theory for the continuous time model of capital accumulation", *Hitotsubashi J. Econom.* **25** (1984), 145–154.

[54] Vinter, R. B. and H. Zheng, "The extended Euler–Lagrange condition for nonconvex variational problems", *SIAM J. Control Optim.* **35** (1997), 56–77.

[55] Wagener, F. O. O., "Skiba points and heteroclinic bifurcations, with applications to the shallow lake system", *J. Econom. Dynam. Control* **27** (2003), 1533–1561.

[56] Ye, J. J., "Nonsmooth maximum principle for infinite-horizon problems", *J. Optim. Theory Appl.* **76** (1993), 485–500.

[57] Zaslavski, A., *Turnpike Properties in the Calculus of Variations and Optimal Control*, Springer, New York, 2006.