MASTER THESIS

Optimal Control and Vanishing Discount Limit in Continuous and Discrete Time

for obtaining the academic degree

Diplom-Ingenieur

conducted at the

Institute of
Statistics and Mathematical Methods in Economics
TU Wien

supervised by

Univ.Prof. Dr. Aris Daniilidis
Univ.Ass. Dr. Sebastián Tapia García

written by

Tobias Grammel, BSc
Student number: 11912401

Vienna, March 22, 2026

Affidavit

I, **Tobias Grammel**, hereby declare that I have written the present Master's Thesis entitled "Optimal Control and Vanishing Discount Limit in Continuous and Discrete Time", comprising 90 pages, independently and without any assistance other than the sources and resources indicated. All passages taken verbatim or in substance from other works have been clearly identified as such.

Artificial intelligence tools were used solely for the purpose of linguistic revision and, where appropriate, for improving style and punctuation. No further assistance or unauthorized tools were used.

I further declare that this Master's Thesis, or parts thereof, has not previously been submitted in any form as an examination paper, either in Austria or abroad.

Vienna, March 22, 2026

Tobias Grammel

Acknowledgements

To begin with, I would like to thank my supervisors, Aris Daniilidis and Sebastián Tapia García. They not only answered all my questions to the best of their abilities, but also provided the intuition necessary for me to understand concepts that were, at least for me, challenging.

This thesis marks the end of my journey as a student at TU Wien. I am grateful for all the experiences I have had, for everything I have learned, and for the people I have met along the way. You were all the best, and I will cherish these memories for the rest of my life.

This journey would not have come to a successful end – or might not even have begun – without the support of many people. I am grateful to each and every one who contributed to this journey in any way. I would like to thank my mother, who always believed in my abilities and supported me in every way possible; my friends and colleagues, who were always a pleasure to be around; and last but not least, my girlfriend, who always had my back, motivated me, and supported me through the not-always-easy times. Thank you!

Abstract

This thesis investigates the vanishing discount limit within infinite-horizon optimal control, exploring its asymptotic behavior across both continuous and discrete time settings. After reviewing and establishing the fundamental theory of finite- and infinite-horizon optimal control, we use the infinite-horizon framework to characterize the vanishing discount limit, namely the limit of the rescaled value function λV_λ as the discount factor λ tends to zero.

This problem has been addressed in earlier works under controllability and ergodicity assumptions ensuring that the rescaled value function converges uniformly to a constant limit. In contrast, we do not impose such conditions. When a uniform limit exists, it is in general a function that can be characterized as the supremum of the family of viscosity subsolutions to the corresponding system of Hamilton–Jacobi equations. When the subsolution is itself a viscosity solution of the corresponding system, we obtain not only the convergence of the vanishing discount limit, but also a specific rate of convergence.

An analogous analysis is carried out in discrete time via Shapley and Bellman operators. Exploiting the additional structure of Bellman operators, we show that the uniform limit of the rescaled discounted value function αv_α as α tends to zero can be characterized as the supremum of the directing vectors of sub-invariant half-lines.

Contents

1 Introduction

Optimal control theory dates back to the 1950s, with major contributions by Lev Pontryagin and Richard Bellman, building on ideas from the calculus of variations. Since then, the theory has evolved substantially and now plays a fundamental role in areas such as economics and engineering.

A major breakthrough was achieved in the 1980s by Michael Crandall and Pierre-Louis Lions, who developed the theory of viscosity solutions for Hamilton–Jacobi equations (Crandall et al., 1983). This formalism allows one to treat Hamilton–Jacobi–Bellman equations that typically do not admit classical (smooth) solutions.

Building on this theory, Chapter 2 establishes the framework of optimal control problems over a finite time horizon. We show that the associated value function V, interpreted as the optimal cost (or payoff), is the unique bounded and uniformly continuous viscosity solution of the Hamilton–Jacobi–Bellman equation

$$\begin{cases} V_t(x,t) + H(x, \nabla_x V(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,T), \\ V(x,T) = g(x) & \text{on } \mathbb{R}^n. \end{cases}$$

Chapter 3 considers the optimization problem over an infinite time horizon. Consequently, different assumptions are required to obtain results analogous to the finite-horizon case. In particular, we establish a comparison principle for bounded and uniformly continuous viscosity solutions of the corresponding Hamilton–Jacobi–Bellman equation, which yields uniqueness of the discounted value function V_λ.

Using this framework, Chapter 4 is devoted to the study of the vanishing discount problem. We analyze the asymptotic behavior of the rescaled family $\{\lambda V_\lambda\}_{\lambda>0}$ as the discount rate $\lambda \to 0^+$, with respect to the topology of uniform convergence on $\overline{\Omega}$, where $\Omega \subseteq \mathbb{R}^n$ is an open and bounded set. This limit can be interpreted as a long-run average cost criterion. We characterize the cluster points of $\{\lambda V_\lambda\}_{\lambda>0}$ and, under additional assumptions, obtain a rate of convergence. This analysis is largely based on the recent work of Cannarsa, Gaubert, Mendico, and Quincampoix in (Cannarsa et al., 2024), which serves as the main reference for this thesis.

Chapter 5 addresses the discrete-time counterpart of the vanishing discount problem. We work within the space $X := C(S)$, where S is a compact Hausdorff space. Equipped with the supremum norm and the pointwise partial order, $C(S)$ is an AM-space with unit $e \in C(S)$. We consider an operator $T : X \to X$ that is nonexpansive and commutes with the addition of the unit, reflecting the structure of Shapley operators. The discrete evolution equation

$$v^k = T(v^{k-1}), \quad k = 1, 2, \dots$$

may be viewed as a discrete-time analogue of the Hamilton–Jacobi–Bellman equation.

The discrete analogue of the discounted value function V_λ is the function $v_\alpha \in X$ satisfying

$$T((1-\alpha)v_\alpha) = v_\alpha,$$

for $\alpha \in (0,1)$. Setting $1 - \alpha = e^{-\lambda}$ establishes a link between the discrete- and continuous-time formulations. Under the additional assumption that each coordinate map $T_i : X \to \mathbb{R}$

is concave, we obtain the representation

$$T_i(x) = \inf_{a \in A_i} \{r_i^a + P_i^a x\},$$

where A_i is a nonempty index set, $r_i^a \in \mathbb{R}$, and P_i^a is a positive continuous linear functional satisfying $P_i^a e = 1$.

In this setting, we establish a discrete analogue of the convergence results obtained in the continuous case by employing the notion of sub-invariant half-lines, introduced by Kohlberg (1980). A sub-invariant half-line of T is a map of the form $s \mapsto u + s\eta$, with $s \geq 0$ and $u, \eta \in X$, such that

$$T(u + s\eta) \geq u + (s+1)\eta, \qquad \forall s \geq 0.$$

We show that the uniform limit of αv_α as $\alpha \to 0^+$, provided it exists, can be characterized as the supremum of the directing vectors η associated with such sub-invariant half-lines.

2 Finite Time Horizon Optimal Control Theory

In this chapter we develop the basic theory of deterministic optimal control on a finite time horizon. As a main result, we derive the Hamilton–Jacobi–Bellman (HJB) equation and show that the value function is its unique, bounded, and uniformly continuous viscosity solution.

In Sections 2.1–2.4 we follow the book by Evans (2010). For Section 2.5 we use (Bardi et al., 1997).

2.1 Framework

Definition 2.1 (State equation)**.** Let $X : [0,T] \to \mathbb{R}^n$, $\alpha : [0,T] \to A$, $A \subseteq \mathbb{R}^m$ compact, and $f : \mathbb{R}^n \times A \to \mathbb{R}^n$. The initial value problem

$$\begin{cases} X'(s) = f(X(s), \alpha(s)), & s \in (t,T), \\ X(t) = x, \end{cases} \tag{2.1}$$

with given initial time $t \geq 0$, initial starting point $x \in \mathbb{R}^n$, and terminal time $T \geq t$ is called *state equation*. The function α is called *control* or *control function*, and the function X is called *response*, *state function* or *state trajectory*. $X(s)$ is called the *state* of the system at time s.

Definition 2.2 (Admissible controls)**.** The set $\mathcal{A} := \{\alpha : [0,T] \to A \mid \alpha \text{ is measurable}\}$ denotes the set of all *admissible controls*.

We require that the state equation (2.1) has a unique Lipschitz continuous solution X^α for each control $\alpha \in \mathcal{A}$, existing on the time interval $[t,T]$, and satisfying the ordinary differential equation (ODE) almost everywhere in $s \in (t,T)$. To ensure this, we impose the following assumptions, which will be maintained throughout the remainder of Chapter 2. A proof that these assumptions are sufficient can be found in the book by Bardi et al. (1997).

Assumption 2.3. Let $f : \mathbb{R}^n \times A \to \mathbb{R}^n$ be a continuous function such that

$$\begin{cases} |f(x,a)| \leq C_f, \\ |f(x,a) - f(y,a)| \leq C_f |x-y|, \end{cases} \qquad \forall\, x, y \in \mathbb{R}^n,\ \forall\, a \in A$$

for some constant $C_f > 0$.

Next, we will introduce the objective function, which we wish to minimize subject to the state equation (2.1).

Definition 2.4 (Cost functional)**.** Given $x \in \mathbb{R}^n$ and $t \in [0,T]$, for each admissible control $\alpha \in \mathcal{A}$ the corresponding *cost functional* is defined by

$$C_{x,t}(\alpha(\cdot)) := \int_t^T L(X(s), \alpha(s))\, ds + g(X(T)), \tag{2.2}$$

where $X = X^\alpha$ solves the state equation (2.1). The functions $L : \mathbb{R}^n \times A \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are given.

Remark 2.5. The term $g(X(T))$ can be thought of as *terminal cost*. For infinite time horizon problems, which we will discuss in Chapter 3, this term will be excluded. Furthermore, the function L is called *Lagrangian* or *running cost per unit time*. The term Lagrangian originates from the calculus of variation which is closely related to optimal control theory. For further details, we refer the reader to the book by Liberzon (2012).

To guarantee that the infimum of the cost functional (2.2) exists and has properties we desire, we assume the following conditions, which shall hold throughout the remainder of the chapter. These desired properties will be discussed in Lemma 2.11.

Assumption 2.6. Let $L : \mathbb{R}^n \times A \to \mathbb{R}$ be a continuous function such that

$$\begin{cases} |L(x,a)| \le C_L, \\ |L(x,a) - L(y,a)| \le C_L|x-y|, \end{cases} \qquad \forall\, x,y \in \mathbb{R}^n,\ \forall\, a \in A$$

for some constant $C_L > 0$ and let $g : \mathbb{R}^n \to \mathbb{R}$ be a continuous function such that

$$\begin{cases} |g(x)| \le C_g, \\ |g(x) - g(y)| \le C_g|x-y|, \end{cases} \qquad \forall\, x,y \in \mathbb{R}^n,\ \forall\, a \in A$$

for some constant $C_g > 0$.

Given initial time $t \in [0,T]$ and initial state $x \in \mathbb{R}^n$, the goal is to find an optimal control $\alpha^* \in \mathcal{A}$ which minimizes the cost functional (2.2) among all admissible controls. This problem is called a *finite-horizon optimal control problem*.

2.2 Dynamic programming

In this section, we introduce the value function and examine its fundamental properties. The core theorem of dynamic programming is Theorem 2.9, known as principle of optimality or optimality condition of dynamic programming. Later, in Section 2.3, we will derive a partial differential equation (PDE), the Hamilton–Jacobi–Bellman equation, which can be regarded as an infinitesimal version of the optimality condition (2.4).

Definition 2.7 (Value function)**.** For each initial state $x \in \mathbb{R}^n$ and starting time $t \in [0,T]$, the function $V : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ defined by

$$V(x,t) := \inf_{\alpha \in \mathcal{A}} C_{x,t}(\alpha(\cdot)) \tag{2.3}$$

is called *value function*.

Remark 2.8. It is important to note that under our assumptions the existence of an optimal control $\alpha^* \in \mathcal{A}$, for which $V(x,t) = C_{x,t}(\alpha^*(\cdot))$ holds, is not guaranteed. However, later in this chapter we will show that there is still a way to construct an optimal control.

Theorem 2.9 (Principle of optimality)**.** *Given $x \in \mathbb{R}^n$ and $t \in [0,T)$, for each $h > 0$ so small that $t + h \le T$, we have*

$$V(x,t) = \inf_{\alpha \in \mathcal{A}} \left\{ \int_t^{t+h} L(X(s),\alpha(s))\, ds + V(X(t+h), t+h) \right\}, \tag{2.4}$$

where $X = X^\alpha$ solves the state equation (2.1) for the control α.

Proof. We proceed in two steps. In Step 1, we show that $V(x,t)$ is lower or equal than the right hand side of (2.4). In Step 2, we prove the reverse inequality.

Step 1. Choose any control $\alpha_1 \in \mathcal{A}$ and solve the ODE

$$\begin{cases} X_1'(s) = f(X_1(s), \alpha_1(s)), & s \in (t, t+h), \\ X_1(t) = x. \end{cases}$$

Fix $\varepsilon > 0$ and choose $\alpha_2 \in \mathcal{A}$ such that

$$V(X_1(t+h), t+h) + \varepsilon \geq \int_{t+h}^{T} L(X_2(s), \alpha_2(s))\, ds + g(X_2(T)), \tag{2.5}$$

where X_2 solves

$$\begin{cases} X_2'(s) = f(X_2(s), \alpha_2(s)), & s \in (t+h, T), \\ X_2(t+h) = X_1(t+h). \end{cases}$$

Therefore, with starting time $t+h$ and starting point $X_1(t+h)$, the control $\alpha_2 \in \mathcal{A}$ is almost optimal for this subproblem, meaning it achieves a value within ε of the optimal one. Let us now define the control

$$\alpha_3(s) := \begin{cases} \alpha_1(s), & s \in [t, t+h), \\ \alpha_2(s), & s \in [t+h, T], \end{cases}$$

where X_3 solves

$$\begin{cases} X_3'(s) = f(X_3(s), \alpha_3(s)), & s \in (t, T), \\ X_3(t) = x. \end{cases}$$

By uniqueness of the state equation (2.1) solutions, we have

$$X_3(s) = \begin{cases} X_1(s), & s \in [t, t+h), \\ X_2(s), & s \in [t+h, T]. \end{cases}$$

Therefore, by the definition of the value function (2.3), we get

$$\begin{aligned} V(x,t) &\leq C_{x,t}(\alpha_3(\cdot)) \\ &= \int_t^T L(X_3(s), \alpha_3(s))\, ds + g(X_3(T)) \\ &= \int_t^{t+h} L(X_1(s), \alpha_1(s))\, ds + \int_{t+h}^{T} L(X_2(s), \alpha_2(s))\, ds + g(X_2(T)) \\ &\overset{(2.5)}{\leq} \int_t^{t+h} L(X_1(s), \alpha_1(s))\, ds + V(X_1(t+h), t+h) + \varepsilon. \end{aligned}$$

Since $\alpha_1 \in \mathcal{A}$ was arbitrary, we get

$$V(x,t) \leq \inf_{\alpha \in \mathcal{A}} \left\{ \int_t^{t+h} L(X(s), \alpha(s))\, ds + V(X(t+h), t+h) \right\} + \varepsilon,$$

where $X = X^\alpha$ solves (2.1). Because $\varepsilon > 0$ was arbitrary, by passing the limit $\varepsilon \to 0$, we conclude Step 1.

Step 2. Similarly, fix $\varepsilon > 0$ and select $\alpha_4 \in \mathcal{A}$ such that

$$V(x,t) + \varepsilon \geq \int_t^T L(X_4(s), \alpha_4(s))\, ds + g(X_4(T)), \tag{2.6}$$

where X_4 solves

$$\begin{cases} X_4'(s) = f(X_4(s), \alpha_4(s)), & s \in (t, T), \\ X_4(t) = x. \end{cases}$$

By the definition of the value function (2.3),

$$V(X_4(t+h), t+h) \leq \int_{t+h}^{T} L(X_4(s), \alpha_4(s))\, ds + g(X_4(T)), \tag{2.7}$$

must hold. Therefore,

$$\begin{aligned} V(x,t) + \varepsilon &\overset{(2.6)}{\geq} \int_t^T L(X_4(s), \alpha_4(s))\, ds + g(X_4(T)) \\ &= \int_t^{t+h} L(X_4(s), \alpha_4(s))\, ds + \int_{t+h}^{T} L(X_4(s), \alpha_4(s))\, ds + g(X_4(T)) \\ &\overset{(2.7)}{\geq} \int_t^{t+h} L(X_4(s), \alpha_4(s))\, ds + V(X_4(t+h), t+h) \\ &\geq \inf_{\alpha \in \mathcal{A}} \left\{ \int_t^{t+h} L(X(s), \alpha(s))\, ds + V(X(t+h), t+h) \right\}, \end{aligned}$$

where $X = X^\alpha$ solves (2.1). Because $\varepsilon > 0$ was arbitrary again, by passing the limit $\varepsilon \to 0$, we conclude Step 2 and therefore the proof. □

Remark 2.10. Theorem 2.9 formalizes the same concepts developed by Bellman (1952) in the 1950s. Every segment of an optimal trajectory is itself optimal.

2.3 Hamilton–Jacobi–Bellman equation

Before establishing the Hamilton–Jacobi–Bellman equation we first prove the desired properties of the value function by making use of Assumption 2.6.

Lemma 2.11. *For all $x_1, x_2 \in \mathbb{R}^n$ and $t_1, t_2 \in [0, T]$, there exists $C_V > 0$ such that*

a) $|V(x_1, t_1)| \leq C_V$

b) $|V(x_1, t_1) - V(x_2, t_2)| \leq C_V(|x_1 - x_2| + |t_1 - t_2|)$.

Proof. We proceed in three steps. In Step 1, we will briefly argue why the value function is bounded. In Step 2, we will prove Lipschitz continuity in the first variable and in Step 3, in the second.

Step 1. Because of Assumption 2.6 we know that the Lagrangian and the terminal cost are bounded. Therefore, the cost functional (2.2), as a finite time horizon integral is bounded as well, for all controls. This implies the boundedness of the value function (2.3) on $\mathbb{R}^n \times [0, T]$.

Step 2. Fix $x_1, x_2 \in \mathbb{R}^n$ and $t \in [0, T)$. Let $\varepsilon > 0$ and choose $\alpha \in \mathcal{A}$ such that,

$$V(x_1, t) + \varepsilon \geq \int_t^T L(X_1(s), \alpha(s))\, ds + g(X_1(T)), \tag{2.8}$$

where X_1 solves

$$\begin{cases} X_1'(s) = f(X_1(s), \alpha(s)), & s \in (t, T), \\ X_1(t) = x_1. \end{cases}$$

This means the control α achieves a value within ε of the optimal one. Starting from x_2, the inequality

$$V(x_2,t) \leq \int_t^T L(X_2(s),\alpha(s))\,ds + g(X_2(T)), \tag{2.9}$$

where X_2 solves

$$\begin{cases} X_2'(s) = f(X_2(s),\alpha(s)), & s \in (t,T), \\ X_2(t) = x_2 \end{cases}$$

holds. Combining the inequalities (2.8) and (2.9) yields

$$\begin{aligned} V(x_2,t) - V(x_1,t) \leq \int_t^T & L(X_2(s),\alpha(s))\,ds + g(X_2(T)) \\ & - \int_t^T L(X_1(s),\alpha(s))\,ds - g(X_1(T)) + \varepsilon. \end{aligned} \tag{2.10}$$

Observe that

$$|X_1'(s) - X_2'(s)| = |f(X_1(s),\alpha(s)) - f(X_2(s),\alpha(s))| \leq C_f|X_1(s) - X_2(s)| \tag{2.11}$$

holds, because of Assumption 2.3. Define $\eta(s) := |X_1(s)-X_2(s)|$, $\phi(s) := C_f$ and $\psi(s) := 0$ for all $s \in [t,T]$. We want to apply Theorem A.1 Gronwall's inequality to these functions. The function η is nonnegative. X_1, as a state trajectory, is Lipschitz continuous and therefore absolutely continuous, which implies that η is also an absolutely continuous function. For almost every $s \in [t,T]$, the inequality

$$\eta'(s) \leq \left|X_1'(s) - X_2'(s)\right| \overset{(2.11)}{\leq} C_f|X_1(s) - X_2(s)| = \phi(s)\eta(s) + \psi(s)$$

holds, where functions ϕ and ψ are nonnegative and integrable functions. Therefore, by Gronwall's inequality

$$\begin{aligned} |X_1(s) - X_2(s)| &\leq e^{\int_t^s C_f\,ds}\left(|X_1(t) - X_2(t)| + \int_t^s 0\,ds\right) \\ &= e^{C_f(s-t)}|x_1 - x_2| \end{aligned} \tag{2.12}$$

holds for all $s \in [t,T]$. Finally, by using Assumption 2.6 and the above inequalities, we get

$$\begin{aligned} V(x_2,t) - V(x_1,t) &\overset{(2.10)}{\leq} \int_t^T L(X_2(s),\alpha(s)) - L(X_1(s),\alpha(s))\,ds + g(X_2(T)) - g(X_1(T)) + \varepsilon \\ &\leq \int_t^T C_L|X_2(s) - X_1(s)|\,ds + C_g|X_2(T) - X_1(T)| + \varepsilon \\ &\overset{(2.12)}{\leq} \int_t^T C_L e^{C_f(s-t)}|x_2 - x_1|\,ds + C_g e^{C_f(T-t)}|x_2 - x_1| + \varepsilon \\ &= \left(\frac{C_L}{C_f}\left(e^{C_f(T-t)} - 1\right) + C_g e^{C_f(T-t)}\right)|x_2 - x_1| + \varepsilon \\ &\leq \left(\frac{C_L}{C_f}\left(e^{C_f T} - 1\right) + C_g e^{C_f T}\right)|x_2 - x_1| + \varepsilon = \widehat{C}|x_2 - x_1| + \varepsilon, \end{aligned}$$

where $\widehat{C} := C_L\left(e^{C_f T} - 1\right)/C_f + C_g e^{C_f T}$. Step 2 is concluded by doing the same steps by interchanging to roles of x_1 and x_2 and passing the limit $\varepsilon \to 0$.

Step 3. Fix $x \in \mathbb{R}^n$ and assume without loss of generality $0 \le t_1 < t_2 \le T$. Take $\varepsilon > 0$ and choose a control $\alpha \in \mathcal{A}$ such that

$$V(x,t_1) + \varepsilon \ge \int_{t_1}^{T} L(X(s),\alpha(s))\,ds + g(X(T)), \tag{2.13}$$

where X solves

$$\begin{cases} X'(s) = f(X(s),\alpha(s)), & s \in (t_1,T), \\ X(t_1) = x. \end{cases}$$

Define the shifted control $\widehat{\alpha}(s) := \alpha(s + t_1 - t_2)$ for $s \in [t_2, T]$ and let $\widehat{X}$ solve

$$\begin{cases} \widehat{X}'(s) = f(\widehat{X}(s),\widehat{\alpha}(s)), & s \in (t_2,T), \\ \widehat{X}(t_2) = x. \end{cases}$$

Then because of the uniqueness of the solution $\widehat{X}(s) = X(s + t_1 - t_2)$ holds for $s \in [t_2, T]$. We also know that

$$V(x,t_2) \le \int_{t_2}^{T} L(\widehat{X}(s),\widehat{\alpha}(s))\,ds + g(\widehat{X}(T)) \tag{2.14}$$

holds. Combining the inequalities (2.13) and (2.14) and making use of the uniqueness yields

$$\begin{aligned} V(x,t_2) - V(x,t_1) &\le \int_{t_2}^{T} L(\widehat{X}(s),\widehat{\alpha}(s))\,ds + g(\widehat{X}(T)) \\ &\qquad - \int_{t_1}^{T} L(X(s),\alpha(s))\,ds - g(X(T)) + \varepsilon \\ &= \int_{t_1}^{T+t_1-t_2} L(X(s),\alpha(s))\,ds + g(X(T+t_1-t_2)) \\ &\qquad - \int_{t_1}^{T} L(X(s),\alpha(s))\,ds - g(X(T)) + \varepsilon \\ &= -\int_{T+t_1-t_2}^{T} L(X(s),\alpha(s))\,ds + g(X(T+t_1-t_2)) - g(X(T)) + \varepsilon \\ &\le \left| \int_{T+t_1-t_2}^{T} C_L\,ds \right| + C_g |X(T+t_1-t_2) - X(T)| + \varepsilon \\ &\le C_L |t_1 - t_2| + C_g C_f |t_1 - t_2| + \varepsilon = \widetilde{C} |t_1 - t_2| + \varepsilon, \end{aligned}$$

where $\widetilde{C} := C_L + C_g C_f$.

The reverse inequality works similarly. Pick $\widetilde{\alpha} \in \mathcal{A}$ such that

$$V(x,t_2) + \varepsilon \ge \int_{t_2}^{T} L(\widetilde{X}(s),\widetilde{\alpha}(s))\,ds + g(\widetilde{X}(T)), \tag{2.15}$$

where $\widetilde{X}$ solves

$$\begin{cases} \widetilde{X}'(s) = f(\widetilde{X}(s),\widetilde{\alpha}(s)), & s \in (t_2,T), \\ \widetilde{X}(t_2) = x. \end{cases}$$

Define the control

$$\bar{\alpha}(s) := \begin{cases} \widetilde{\alpha}(s + t_2 - t_1), & s \in [t_1, T + t_1 - t_2), \\ \widetilde{\alpha}(T), & s \in [T + t_1 - t_2, T], \end{cases}$$

and let $\bar{X}$ solve

$$\begin{cases} \bar{X}'(s) = f(\bar{X}(s), \bar{\alpha}(s)), & s \in (t_1, T), \\ \bar{X}(t_1) = x. \end{cases}$$

Notice that $\bar{\alpha}(s) = \widetilde{\alpha}(s + t_2 - t_1)$, $\bar{X}(s) = \widetilde{X}(s + t_2 - t_1)$ for $s \in [t_1, T + t_1 - t_2]$, and

$$V(x, t_1) \le \int_{t_1}^{T} L(\bar{X}(s), \bar{\alpha}(s))\, ds + g(\bar{X}(T)). \tag{2.16}$$

Using this information and combining the inequalities (2.15) and (2.16) yields

$$\begin{aligned} V(x,t_1) - V(x,t_2) &\le \int_{t_1}^{T} L(\bar{X}(s), \bar{\alpha}(s))\, ds + g(\bar{X}(T)) \\ &\qquad - \int_{t_2}^{T} L(\widetilde{X}(s), \widetilde{\alpha}(s))\, ds - g(\widetilde{X}(T)) + \varepsilon \\ &= \int_{t_1}^{T} L(\bar{X}(s), \bar{\alpha}(s))\, ds + g(\bar{X}(T)) \\ &\qquad - \int_{t_1}^{T+t_1-t_2} L(\widetilde{X}(s + t_2 - t_1), \widetilde{\alpha}(s + t_2 - t_1))\, ds - g(\widetilde{X}(T)) + \varepsilon \\ &= \int_{T+t_1-t_2}^{T} L(\bar{X}(s), \bar{\alpha}(s))\, ds + g(\bar{X}(T)) - g(\bar{X}(T + t_1 - t_2)) + \varepsilon \\ &\le \left| \int_{T+t_1-t_2}^{T} C_L\, ds \right| + C_g \left| \bar{X}(T) - \bar{X}(T + t_1 - t_2) \right| + \varepsilon \\ &\le C_L |t_1 - t_2| + C_g C_f |t_1 - t_2| + \varepsilon = \widetilde{C} |t_1 - t_2| + \varepsilon. \end{aligned}$$

Finally, combining the two inequalities and taking the limit $\varepsilon \to 0$ concludes Step 3 and therefore the proof. □

In the following, we formulate the Hamilton–Jacobi–Bellman equation and establish the notion of a viscosity solution in this context.

Definition 2.12 (Hamilton–Jacobi–Bellman equation)**.** Let V denote the value function associated with initial position $x \in \mathbb{R}^n$, starting time $t \in [0, T]$, and state equation (2.1). The *Hamilton–Jacobi–Bellman equation* is given by

$$\begin{cases} V_t(x,t) + H(x, \nabla_x V(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0, T), \\ V(x, T) = g(x) & \text{on } \mathbb{R}^n, \end{cases} \tag{2.17}$$

where the function $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with $H(x, p) := \min_{a \in A} \{ \langle f(x,a), p \rangle + L(x,a) \}$ is called the *Hamiltonian*.

Remark 2.13. The Hamilton–Jacobi–Bellman equation looks slightly different depending on the control problem setting. For reference, Definition 3.5 gives the corresponding formulation for the infinite-horizon problem.

Definition 2.14 (Viscosity solution of the HJB equation)**.** We say that a bounded and uniformly continuous function $V : \mathbb{R}^n \times [0, T] \to \mathbb{R}$ is a *viscosity solution* of the Hamilton–Jacobi–Bellman equation (2.17) if the following conditions hold:

1. $V(x, T) = g(x)$ on $\mathbb{R}^n$,

2. For every test function $\phi \in C^1(\mathbb{R}^n \times (0,T))$,
   a) if $V - \phi$ attains a local maximum at a point $(x_0, t_0) \in \mathbb{R}^n \times (0,T)$, then
   $$\phi_t(x_0,t_0) + H(x_0, \nabla_x \phi(x_0,t_0)) \geq 0.$$
   b) if $V - \phi$ attains a local minimum at a point $(x_0, t_0) \in \mathbb{R}^n \times (0,T)$, then
   $$\phi_t(x_0,t_0) + H(x_0, \nabla_x \phi(x_0,t_0)) \leq 0.$$

Remark 2.15. The notion of viscosity solutions for Hamilton–Jacobi equations was introduced in the 1980s by Crandall and Lions (Crandall et al., 1983), and later extended to second-order partial differential equations. For a detailed introduction to viscosity solutions, we refer the reader to Liberzon (2012).

Next, in Theorem 2.16, we prove that the value function V is a bounded and uniformly continuous viscosity solution of the Hamilton–Jacobi–Bellman equation. Afterwards, we will establish uniqueness.

Theorem 2.16 (PDE characterization of the value function)**.** *The value function V is a viscosity solution of the Hamilton–Jacobi–Bellman equation* (2.17).

Proof. We proceed in three steps. In Step 1, we briefly recall the properties of the value function and explain why the terminal condition must be satisfied. In Step 2, we prove point 2.a) of Definition 2.14 by contradiction. In Step 3, we prove point 2.b), similar to Step 2, by contradiction.

Step 1. In Lemma 2.11 we showed that the value function V is bounded, Lipschitz continuous and therefore uniformly continuous. Furthermore, from the definitions of the cost and value function we can see that

$$V(x,T) \overset{(2.3)}{=} \inf_{\alpha \in \mathcal{A}} C_{x,T}(\alpha(\cdot)) \overset{(2.2)}{=} g(x)$$

holds for all $x \in \mathbb{R}^n$. The terminal condition of (2.17) is therefore satisfied.

Step 2. Let $\phi \in C^1(\mathbb{R}^n \times (0,T))$ and assume $V - \phi$ attains a local maximum at a point $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. We want to show $\phi_t(x_0,t_0) + \min_{a \in A}\{\langle f(x_0,a), \nabla_x \phi(x_0,t_0)\rangle + L(x_0,a)\} \geq 0$. Suppose, the contrary. Then there exist $a \in A$ and $\theta > 0$ such that

$$\phi_t(x,t) + \langle f(x,a), \nabla_x \phi(x,t)\rangle + L(x,a) \leq -\theta < 0, \tag{2.18}$$

for all points (x,t) satisfying

$$|x - x_0| + |t - t_0| < \delta, \tag{2.19}$$

where $\delta > 0$ is chosen sufficiently small. We will show that this leads to a contradiction.

Consider the constant control $\alpha(s) := a$ for $s \in [t_0, T]$ and the corresponding dynamics

$$\begin{cases} X'(s) = f(X(s), a), & s \in (t_0, T), \\ X(t_0) = x_0. \end{cases} \tag{2.20}$$

Due to the continuity of the solution and the initial condition of (2.20) we can choose $h \in [0, \delta]$ such that $|X(s) - x_0| < \delta$ for $s \in [t_0, t_0 + h]$ is satisfied. Therefore, in view of inequality (2.18),

$$\phi_t(X(s),s) + \langle f(X(s),a), \nabla_x \phi(X(s),s)\rangle + L(X(s),a) \leq -\theta \tag{2.21}$$

holds for $s \in [t_0, t_0+h]$, due to (2.19) being satisfied. Since $V-\phi$ attains a local maximum at (x_0, t_0) we can equivalently say,

$$(V-\phi)(x,t) \leq (V-\phi)(x_0,t_0) \quad \text{for all } (x,t) \text{ satisfying (2.19)}. \tag{2.22}$$

Rearranging (2.22), using the fundamental theorem of calculus, and the chain rule, we obtain

$$\begin{aligned} V(X(t_0+h), t_0+h) &- V(x_0,t_0) \\ &\leq \phi(X(t_0+h), t_0+h) - \phi(x_0,t_0) \\ &= \int_{t_0}^{t_0+h} \frac{d}{ds}\phi(X(s),s)\,ds \\ &= \int_{t_0}^{t_0+h} \phi_t(X(s),s) + \langle \nabla_x \phi(X(s),s), X'(s)\rangle\,ds \\ &\overset{(2.20)}{=} \int_{t_0}^{t_0+h} \phi_t(X(s),s) + \langle f(X(s),a), \nabla_x\phi(X(s),s)\rangle\,ds. \end{aligned} \tag{2.23}$$

Furthermore, by the principle of optimality, Theorem 2.9,

$$V(x_0,t_0) \leq \int_{t_0}^{t_0+h} L(X(s),a)\,ds + V(X(t_0+h), t_0+h) \tag{2.24}$$

holds. By adding inequalities (2.23) and (2.24), and then subtracting $V(X(t_0+h), t_0+h)$ from both sides, we obtain

$$\begin{aligned} 0 &\leq \int_{t_0}^{t_0+h} \phi_t(X(s),s) + \langle f(X(s),a), \nabla_x\phi(X(s),s)\rangle + L(X(s),a)\,ds \\ &\overset{(2.21)}{\leq} \int_{t_0}^{t_0+h} -\theta\,ds = -h\theta, \end{aligned}$$

which is a contradiction to $\theta > 0$ and therefore completes Step 2.

Step 3. Now, assume $V-\phi$ attains a local minimum at a point $(x_0,t_0) \in \mathbb{R}^n \times (0,T)$. We want to show $\phi_t(x_0,t_0) + \min_{a\in A}\{\langle f(x_0,a), \nabla_x\phi(x_0,t_0)\rangle + L(x_0,a)\} \leq 0$. Suppose the contrary. Then there exist $\theta > 0$ such that

$$\phi_t(x,t) + \langle f(x,a), \nabla_x\phi(x,t)\rangle + L(x,a) \geq \theta > 0, \tag{2.25}$$

for all $a \in A$ and all points (x,t) satisfying

$$|x-x_0| + |t-t_0| < \delta, \tag{2.26}$$

where $\delta > 0$ is chosen sufficiently small. We will show that this leads to a contradiction.

Choose $h \in (0,\delta)$ so small that $|X(s)-x_0| < \delta$ for $s \in [t_0, t_0+h]$, where X solves

$$\begin{cases} X'(s) = f(X(s), \alpha(s)), & s \in (t_0, T), \\ X(t_0) = x_0 \end{cases} \tag{2.27}$$

for any control $\alpha \in \mathcal{A}$. Since $V-\phi$ attains a local minimum at (x_0,t_0) we can equivalently say,

$$(V-\phi)(x,t) \geq (V-\phi)(x_0,t_0) \quad \text{for all } (x,t) \text{ satisfying (2.26)}. \tag{2.28}$$

By rearranging equation (2.28), using the fundamental theorem of calculus, and the chain rule, we see that for any control $\alpha \in \mathcal{A}$,

$$\begin{aligned} V(X(t_0+h), t_0+h) - V(x_0, t_0) \\ &\geq \phi(X(t_0+h), t_0+h) - \phi(x_0, t_0) \\ &= \int_{t_0}^{t_0+h} \frac{d}{ds}\phi(X(s), s)\, ds \\ &= \int_{t_0}^{t_0+h} \phi_t(X(s), s) + \langle \nabla_x \phi(X(s), s), X'(s)\rangle\, ds \\ &\overset{(2.27)}{=} \int_{t_0}^{t_0+h} \phi_t(X(s), s) + \langle f(X(s), \alpha(s)), \nabla_x \phi(X(s), s)\rangle\, ds, \end{aligned} \tag{2.29}$$

holds. Furthermore, by the principle of optimality, Theorem 2.9, we can select a control $\alpha \in \mathcal{A}$ that achieves a value within $\theta h/2$ of the optimal one. Consequently,

$$V(x_0, t_0) \geq \int_{t_0}^{t_0+h} L(X(s), \alpha(s))\, ds + V(X(t_0+h), t_0+h) - \frac{\theta h}{2}. \tag{2.30}$$

By adding inequality (2.29) and (2.30) and then subtracting $V(X(t_0+h), t_0+h)$ from both sides, we obtain

$$0 \geq \int_{t_0}^{t_0+h} \phi_t(X(s), s) + \langle f(X(s), \alpha(s)), \nabla_x \phi(X(s), s)\rangle + L(X(s), \alpha(s))\, ds - \frac{\theta h}{2}.$$

Rearranging the inequality above and using the fact that inequality (2.25) holds for all $a \in A$, we get

$$\begin{aligned} \frac{\theta h}{2} &\geq \int_{t_0}^{t_0+h} \phi_t(X(s), s) + \langle f(X(s), \alpha(s)), \nabla_x \phi(X(s), s)\rangle + L(X(s), \alpha(s))\, ds \\ &\overset{(2.25)}{\geq} \int_{t_0}^{t_0+h} \theta\, ds = h\theta. \end{aligned}$$

This contradiction shows Step 3 and concludes the proof. □

We will now prove that if V is a viscosity solution of the terminal value problem (2.17), then $\tilde{V}(x,t) := V(x, T-t)$ for $x \in \mathbb{R}^n, t \in [0,T]$ is a viscosity solution of the initial value problem

$$\begin{cases} \tilde{V}_t(x,t) - H(x, \nabla_x \tilde{V}(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,T), \\ \tilde{V}(x,0) = g(x) & \text{on } \mathbb{R}^n. \end{cases} \tag{2.31}$$

In contrast to the terminal value problem, the inequalities of Definition 2.14 are reversed when working with the initial value problem. This means a function $\tilde{V}$ is a viscosity solution to the initial value problem (2.31) if the following conditions hold:

1. $\tilde{V}(x,0) = g(x)$ on $\mathbb{R}^n$,
2. For every test function $\psi \in C^1(\mathbb{R}^n \times (0,T))$,
   a) if $\tilde{V} - \psi$ attains a local maximum at a point $(x_0, s_0) \in \mathbb{R}^n \times (0,T)$, then

$$\psi_s(x_0, s_0) - H(x_0, \nabla_x \psi(x_0, s_0)) \leq 0.$$

b) if $\tilde{V} - \psi$ attains a local minimum at a point $(x_0, s_0) \in \mathbb{R}^n \times (0, T)$, then

$$\psi_s(x_0, s_0) - H(x_0, \nabla_x \psi(x_0, s_0)) \geq 0.$$

Note that Proposition 2.17 and Lemma 2.19 are stated without proof in Evans (2010); see (Evans, 2010, Remarks, pp. 596–597).

Proposition 2.17. *If a function $V : \mathbb{R}^n \times [0, T] \to \mathbb{R}$ is a viscosity solution of the terminal value problem (2.17), then the function $\tilde{V}(x, t) := V(x, T - t)$ is a viscosity solution of the initial value problem (2.31).*

Proof. We proceed in four steps. In Step 1, we will define a time-reversed test function and establish relations we will use throughout the proof. In Step 2 and 3, we will show the points 2.a) and 2.b) from above, respectively. Finally, in Step 4 we will show the initial condition holds.

Step 1. Let $s \in (0, T)$ and set $t := T - s \in (0, T)$. Take any smooth function $\psi \in C^1(\mathbb{R}^n \times (0, T))$ and define the test function $\phi(x, t) := \psi(x, T - t)$ for V. The derivatives transform as

$$\nabla_x \phi(x, t) = \nabla_x \psi(x, T - t) \quad \text{and} \quad \phi_t(x, t) = -\psi_s(x, s)\big|_{s=T-t},$$

and in particular, if $t_0 := T - s_0$, then

$$\nabla_x \phi(x_0, t_0) = \nabla_x \psi(x_0, s_0) \quad \text{and} \quad \phi_t(x_0, t_0) = -\psi_s(x_0, s_0). \tag{2.32}$$

Also, observe that

$$\tilde{V}(x, s) - \psi(x, s) = V(x, T - s) - \psi(x, s) = V(x, t) - \phi(x, t).$$

This means:

1. If $\tilde{V} - \psi$ attains a local maximum at $(x_0, s_0) \in \mathbb{R}^n \times (0, T)$, then $V - \phi$ attains a local maximum at (x_0, t_0).
2. If $\tilde{V} - \psi$ attains a local minimum at $(x_0, s_0) \in \mathbb{R}^n \times (0, T)$, then $V - \phi$ attains a local minimum at (x_0, t_0).

Step 2. Assume that $\tilde{V} - \psi$ attains a local maximum at (x_0, s_0). Therefore, $V - \phi$ attains a local maximum at (x_0, t_0), which implies

$$\phi_t(x_0, t_0) + H(x_0, \nabla_x \phi(x_0, t_0)) \geq 0,$$

as V is assumed to be a viscosity solution of the terminal value problem. Substituting the equations (2.32) yields

$$-\psi_s(x_0, s_0) + H(x_0, \nabla_x \psi(x_0, s_0)) \geq 0,$$

or equivalently,

$$\psi_s(x_0, s_0) - H(x_0, \nabla_x \psi(x_0, s_0)) \leq 0.$$

Step 3. Analogously, assume that $\tilde{V} - \psi$ attains a local minimum at (x_0, s_0). Therefore, $V - \phi$ attains a local minimum at (x_0, t_0), which implies that

$$\phi_t(x_0, t_0) + H(x_0, \nabla_x \phi(x_0, t_0)) \leq 0.$$

Substituting the equations (2.32) yields

$$-\psi_s(x_0, s_0) + H(x_0, \nabla_x \psi(x_0, s_0)) \leq 0,$$

or equivalently,

$$\psi_s(x_0, s_0) - H(x_0, \nabla_x \psi(x_0, s_0)) \geq 0.$$

Step 4. For every $x \in \mathbb{R}^n$ the terminal condition for the function V holds. Using the definition of $\tilde{V}$ shows the initial condition

$$g(x) = V(x, T) = \tilde{V}(x, 0).$$

This concludes the proof. □

Theorem 2.18 (Uniqueness of viscosity solutions for the initial value problem)**.** *Assume that $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies for some $C_H > 0$*

$$\begin{cases} |H(x,p) - H(x,q)| \leq C_H |p - q|, \\ |H(x,p) - H(y,p)| \leq C_H |x - y|(1 + |p|), \end{cases} \qquad \forall\, x, y, p, q \in \mathbb{R}^n.$$

Then the initial value problem

$$\begin{cases} \tilde{V}_t(x,t) - H(x, \nabla_x \tilde{V}(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,T), \\ \tilde{V}(x,0) = g(x) & \text{on } \mathbb{R}^n \end{cases} \tag{2.33}$$

admits at most one bounded and uniformly continuous viscosity solution.

Proof. Assume u and v are both viscosity solutions of (2.33) with the same initial conditions, but

$$\sup_{(x,t) \in \mathbb{R}^n \times [0, T-\delta]} \{u(x,t) - v(x,t)\} =: \sigma_\delta > 0 \tag{2.34}$$

for a fixed $\delta \in (0, T)$. By constructing a contradiction, we want to show that this cannot be the case and therefore establish uniqueness.

By continuity of $u - v$ pick a point $(\bar{x}, \bar{t}) \in \mathbb{R}^n \times [0, T - \delta]$ such that

$$u(\bar{x}, \bar{t}) - v(\bar{x}, \bar{t}) \geq \sigma_\delta - \frac{\sigma_\delta}{4}. \tag{2.35}$$

Choose $\varepsilon \in (0,1)$ such that $2\varepsilon |\bar{x}|^2 \leq \sigma_\delta/8$ and $\lambda \in (0,1)$ such that $2\lambda \bar{t} \leq \sigma_\delta/8$. Consider

$$\Phi(x,y,t,s) := u(x,t) - v(y,s) - \lambda(t+s) - \frac{1}{\varepsilon^2}\big(|x-y|^2 + (t-s)^2\big) - \varepsilon\big(|x|^2 + |y|^2\big), \tag{2.36}$$

for $x, y \in \mathbb{R}^n$ and $t, s \in [0, T - \delta]$. Since u and v are bounded and the penalization

$$-\frac{1}{\varepsilon^2}\big(|x-y|^2 + (t-s)^2\big) - \varepsilon\big(|x|^2 + |y|^2\big)$$

forces $\Phi(x,y,t,s) \to -\infty$ as $|x| + |y| \to \infty$, it follows that Φ attains a maximum over $\mathbb{R}^{2n} \times [0, T-\delta]^2$. Hence, there exists a point $(x_0, y_0, t_0, s_0) \in \mathbb{R}^{2n} \times [0, T-\delta]^2$ such that

$$\Phi(x_0, y_0, t_0, s_0) = \max_{(x,y,t,s) \in \mathbb{R}^{2n} \times [0,T-\delta]^2} \Phi(x,y,t,s). \tag{2.37}$$

Because of our specific choices of ε and λ the following inequality

$$\begin{aligned}\Phi(x_0, y_0, t_0, s_0) &\geq \sup_{(x,t)\in\mathbb{R}^n\times[0,T-\delta]} \Phi(x, x, t, t) \\ &\geq \Phi(\bar{x}, \bar{x}, \bar{t}, \bar{t}) \\ &= u(\bar{x}, \bar{t}) - v(\bar{x}, \bar{t}) - 2\lambda\bar{t} - 2\varepsilon|\bar{x}|^2 \\ &\overset{(2.35)}{\geq} \sigma_\delta - \frac{\sigma_\delta}{4} - \frac{\sigma_\delta}{8} - \frac{\sigma_\delta}{8} = \frac{\sigma_\delta}{2},\end{aligned} \tag{2.38}$$

holds. Furthermore, since $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(0, 0, 0, 0) = u(0, 0) - v(0, 0)$, it follows that

$$\begin{aligned}\lambda(t_0 + s_0) + \frac{1}{\varepsilon^2}\big(|x_0 - y_0|^2 + (t_0 - s_0)^2\big) + \varepsilon\big(|x_0|^2 + |y_0|^2\big) \\ \leq u(x_0, t_0) - v(y_0, s_0) - u(0, 0) + v(0, 0).\end{aligned} \tag{2.39}$$

Dropping positive terms and using the boundedness of u and v implies

$$\frac{1}{\varepsilon^2}\big(|x_0 - y_0|^2 + (t_0 - s_0)^2\big) \leq C,$$

for some $C > 0$. We deduce that

$$|x_0 - y_0| = O(\varepsilon) \quad \text{and} \quad |t_0 - s_0| = O(\varepsilon) \quad \text{as} \quad \varepsilon \to 0, \tag{2.40}$$

where information about the big-oh and little-oh notation can be found in the Appendices. In what follows, all occurrences of these notations refer to the limit $\varepsilon \to 0$ even if not stated explicitly. Furthermore, by inequality (2.39) we get

$$\varepsilon\big(|x_0|^2 + |y_0|^2\big) \leq \tilde{C} \quad \Leftrightarrow \quad \big(|x_0|^2 + |y_0|^2\big) \leq \frac{\tilde{C}}{\varepsilon} \tag{2.41}$$

for some $\tilde{C} > 0$. Using the inequality $a + b \leq \sqrt{2(a^2 + b^2)}$ for $a := |x_0|$ and $b := |y_0|$ yields

$$|x_0| + |y_0| \leq \sqrt{2(|x_0|^2 + |y_0|^2)} \overset{(2.41)}{\leq} \sqrt{2\tilde{C}/\varepsilon} = \hat{C}\varepsilon^{-1/2}$$

for $\hat{C} := \sqrt{2\tilde{C}}$. It follows that

$$\varepsilon(|x_0| + |y_0|) = O\big(\varepsilon^{1/2}\big) \tag{2.42}$$

Similarly to before, since $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(x_0, x_0, t_0, t_0)$, we get

$$\begin{aligned}u(x_0, t_0) - v(y_0, s_0) - \lambda(t_0 + s_0) - \frac{1}{\varepsilon^2}\big(|x_0 - y_0|^2 + (t_0 - s_0)^2\big) - \varepsilon\big(|x_0|^2 + |y_0|^2\big) \\ \geq u(x_0, t_0) - v(x_0, t_0) - 2\lambda t_0 - 2\varepsilon|x_0|^2.\end{aligned}$$

Subtracting the common term $u(x_0, t_0)$ and rearranging yields

$$-v(y_0, s_0) + \lambda(t_0 - s_0) - \frac{1}{\varepsilon^2}\big(|x_0 - y_0|^2 + (t_0 - s_0)^2\big) + \varepsilon\big(|x_0|^2 - |y_0|^2\big) \geq -v(x_0, t_0).$$

Hence

$$\frac{1}{\varepsilon^2}\big(|x_0 - y_0|^2 + (t_0 - s_0)^2\big) \leq v(x_0, t_0) - v(y_0, s_0) + \lambda(t_0 - s_0) + \varepsilon\langle x_0 + y_0, x_0 - y_0\rangle. \tag{2.43}$$

From equation (2.40) we know that $\lambda(t_0 - s_0) = O(\varepsilon)$. Furthermore, by the Cauchy–Schwarz and the triangle inequality we see that

$$\varepsilon\langle x_0 + y_0, x_0 - y_0\rangle \leq \varepsilon|x_0 + y_0||x_0 - y_0| \leq \varepsilon(|x_0| + |y_0|)|x_0 - y_0|.$$

Using equation (2.40) and (2.42) yields

$$\varepsilon(|x_0| + |y_0|)|x_0 - y_0| = O\big(\varepsilon^{1/2}\big)\, O(\varepsilon) = O\big(\varepsilon^{3/2}\big).$$

Because $\varepsilon^{3/2}/\varepsilon = \varepsilon^{1/2}$ and $\lim_{\varepsilon\to 0}\varepsilon^{1/2} = 0$ it follows $O\big(\varepsilon^{3/2}\big) = o(\varepsilon)$. Using (2.40) and the uniform continuity of v, we have

$$v(x_0, t_0) - v(y_0, s_0) \to 0, \quad \lambda(t_0 - s_0) \to 0, \quad \varepsilon\langle x_0 + y_0, x_0 - y_0\rangle \to 0$$

as $\varepsilon \to 0$. Therefore the right-hand side in (2.43) tends to 0, and since the left-hand side is nonnegative, it follows that

$$\frac{1}{\varepsilon^2}\big(|x_0 - y_0|^2 + (t_0 - s_0)^2\big) \to 0 \quad \text{as } \varepsilon \to 0.$$

In particular,

$$|x_0 - y_0| = o(\varepsilon) \quad \text{and} \quad |t_0 - s_0| = o(\varepsilon). \tag{2.44}$$

Next, denote by ω the modulus of continuity of u. That is,

$$|u(x,t) - u(y,s)| \leq \omega(|x - y| + |t - s|)$$

for all $x, y \in \mathbb{R}^n$, $s, t \in [0, T - \delta]$ and $\lim_{r\to 0}\omega(r) = 0$. Analogously, we define $\tilde{\omega}$ as the modulus of continuity of v. For more information about the modulus of continuity consult for instance the book by Bishop et al. (1985). Using (2.38) and the initial condition in (2.33), which implies $u(x_0, 0) - v(x_0, 0) = 0$, together with (2.44), yields

$$\begin{aligned}
\frac{\sigma_\delta}{2} &\overset{(2.38)}{\leq} u(x_0, t_0) - v(y_0, s_0) \\
&= u(x_0, t_0) - u(x_0, 0) + u(x_0, 0) - v(x_0, 0) \\
&\qquad + v(x_0, 0) - v(x_0, t_0) + v(x_0, t_0) - v(y_0, s_0) \\
&\overset{(2.33),(2.44)}{\leq} \omega(t_0) + \tilde{\omega}(t_0) + \tilde{\omega}(o(\varepsilon)).
\end{aligned}$$

Choose $\varepsilon > 0$ small enough, to satisfy $\sigma_\delta/4 \leq \omega(t_0) + \tilde{\omega}(t_0)$ implies $t_0 \geq \mu > 0$ and $s_0 \geq \mu > 0$ for some constant $\mu > 0$. Consequently, the maximum point occurs on positive times $t_0, s_0 \in (0, T)$.

By observing (2.37) we see that the mapping $(x,t) \mapsto \Phi(x, y_0, t, s_0)$ attains a maximum at the point (x_0, t_0). By looking at the definition of Φ in (2.36) and defining

$$\psi(x,t) := v(y_0, s_0) + \lambda(t + s_0) + \frac{1}{\varepsilon^2}\big(|x - y_0|^2 + (t - s_0)^2\big) + \varepsilon\big(|x|^2 + |y_0|^2\big)$$

we notice that

$$u - \psi \text{ attains a local maximum at } (x_0, t_0).$$

Since u is a viscosity solution of (2.33), the inequality

$$\psi_t(x_0, t_0) - H(x_0, \nabla_x\psi(x_0, t_0)) \leq 0$$

holds. By differentiating we obtain,

$$\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} - H\left(x_0, \frac{2(x_0 - y_0)}{\varepsilon^2} + 2\varepsilon x_0\right) \leq 0. \tag{2.45}$$

Similarly, the mapping $(y, s) \mapsto -\Phi(x_0, y, t_0, s)$ has a minimum at the point (y_0, s_0). By defining

$$\tilde{\psi}(y, s) := u(x_0, t_0) - \lambda(t_0 + s) - \frac{1}{\varepsilon^2}\left(|x_0 - y|^2 + (t_0 - s)^2\right) - \varepsilon\left(|x_0|^2 + |y|^2\right)$$

we notice that,

$$v - \tilde{\psi} \text{ attains a local minimum at } (y_0, s_0).$$

Since v is a viscosity solution of (2.33), the inequality

$$\tilde{\psi}_s(y_0, s_0) - H(y_0, \nabla_y \tilde{\psi}(y_0, s_0)) \geq 0$$

holds. Consequently,

$$-\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} - H\left(y_0, \frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0\right) \geq 0. \tag{2.46}$$

By subtracting the inequalities (2.46) from (2.45), we obtain

$$2\lambda \leq H\left(x_0, \frac{2(x_0 - y_0)}{\varepsilon^2} + 2\varepsilon x_0\right) - H\left(y_0, \frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0\right).$$

By using the assumptions of the theorem, we get

$$\begin{aligned} 2\lambda &\leq H\Bigg(x_0, \overbrace{\frac{2(x_0 - y_0)}{\varepsilon^2} + 2\varepsilon x_0}^{=:p_1}\Bigg) - H\Bigg(y_0, \overbrace{\frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0}^{=:p_2}\Bigg) \\ &= H(x_0, p_1) - H(x_0, p_2) + H(x_0, p_2) - H(y_0, p_2) \\ &\leq C_H|p_1 - p_2| + C_H|x_0 - y_0|(1 + |p_2|). \end{aligned}$$

Together with

$$C_H|p_1 - p_2| = C_H|2\varepsilon x_0 + 2\varepsilon y_0| \leq 2\, C_H \varepsilon(|x_0| + |y_0|)$$

and

$$\begin{aligned} C_H(1 + |p_2|) &= C_H\left(1 + \left|2(x_0 - y_0)/\varepsilon^2 - 2\varepsilon y_0\right|\right) \\ &\leq 2\, C_H\left(1 + |x_0 - y_0|/\varepsilon^2 + \varepsilon|y_0|\right) \\ &\leq 2\, C_H\left(1 + |x_0 - y_0|/\varepsilon^2 + \varepsilon(|x_0| + |y_0|)\right), \end{aligned}$$

we obtain

$$\begin{aligned} \lambda &\leq C_H\varepsilon(|x_0| + |y_0|) + C_H|x_0 - y_0|\left(1 + \frac{|x_0 - y_0|}{\varepsilon^2} + \varepsilon(|x_0| + |y_0|)\right) \\ &\overset{(2.42),(2.44)}{=} O(\varepsilon^{1/2}) + o(\varepsilon)\left(1 + \frac{o(\varepsilon)}{\varepsilon^2} + O(\varepsilon^{1/2})\right). \end{aligned}$$

Consequently, by passing the limit $\varepsilon \to 0$, we discover $0 < \lambda \leq 0$. This contradicts our assumption (2.34), hence $\sigma_\delta \leq 0$. This leads to $u(x, t) \leq v(x, t)$ for all $(x, t) \in \mathbb{R}^n \times [0, T-\delta]$. Since $\delta \in (0, T)$ was arbitrary, letting $\delta \to 0^+$ yields $u \leq v$ on $\mathbb{R}^n \times [0, T)$. Interchanging the roles of u and v gives the reverse inequality, and therefore $u = v$ on $\mathbb{R}^n \times [0, T)$. □

Lemma 2.19. *The Hamiltonian* $H(x,p) = \min_{a\in A}\{\langle f(x,a),p\rangle + L(x,a)\}$ *satisfies for some* $C_H > 0$ *the conditions*

$$\begin{cases} |H(x,p) - H(x,q)| \le C_H|p-q|, \\ |H(x,p) - H(y,p)| \le C_H|x-y|(1+|p|), \end{cases} \qquad \forall\, x,y,p,q \in \mathbb{R}^n.$$

Proof. Fix $x \in \mathbb{R}^n$. Because A is assumed to be compact, see Definition 2.1, let $a_p \in A$ be the minimizer of $H(x,p)$. Due to Assumption 2.3 we get

$$\begin{aligned} H(x,p) - H(x,q) &= \langle f(x,a_p),p\rangle + L(x,a_p) - \min_{a\in A}\{\langle f(x,a),q\rangle + L(x,a)\} \\ &\le \langle f(x,a_p),p\rangle - \langle f(x,a_p),q\rangle \\ &\le |f(x,a_p)||p-q| \\ &\le C_f|p-q|. \end{aligned}$$

Interchanging p and q gives

$$|H(x,p) - H(x,q)| \le C_f|p-q|.$$

Now, fix $p \in \mathbb{R}^n$ and let $a_x \in A$ be the minimizer of $H(x,p)$. As a consequence of Assumption 2.3 and Assumption 2.6 we get

$$\begin{aligned} H(x,p) - H(y,p) &= \langle f(x,a_x),p\rangle + L(x,a_x) - \min_{a\in A}\{\langle f(y,a),p\rangle + L(y,a)\} \\ &\le \langle f(x,a_x),p\rangle + L(x,a_x) - \langle f(y,a_x),p\rangle - L(y,a_x) \\ &\le |f(x,a_x) - f(y,a_x)||p| + |L(x,a_x) - L(y,a_x)| \\ &\le C_f|x-y||p| + C_L|x-y| \\ &\le C_H|x-y|(1+|p|) \end{aligned}$$

by defining $C_H := \max\{C_f, C_L\}$. Interchanging x and y yields

$$|H(x,p) - H(y,p)| \le C_H|x-y|(1+|p|).$$

□

Corollary 2.20 (Value function as unique viscosity solution of the HJB equation)**.** *The value function* V *is the unique, bounded and uniformly continuous viscosity solution of the Hamilton–Jacobi–Bellman equation* (2.17)*.*

Proof. By Theorem 2.16 the value function V is a bounded and uniformly continuous viscosity solution of the terminal value problem (2.17).

To prove uniqueness, let $U : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ be another bounded and uniformly continuous viscosity solution of (2.17) with the same terminal condition $U(x,T) = g(x)$. Define the time-reversed functions

$$\tilde{V}(x,t) := V(x,T-t), \qquad \tilde{U}(x,t) := U(x,T-t), \qquad (x,t) \in \mathbb{R}^n \times [0,T].$$

By Proposition 2.17, both $\tilde{V}$ and $\tilde{U}$ are viscosity solutions of the initial value problem

$$\begin{cases} w_t(x,t) - H\big(x, \nabla_x w(x,t)\big) = 0 & \text{in } \mathbb{R}^n \times (0,T), \\ w(x,0) = g(x) & \text{on } \mathbb{R}^n. \end{cases}$$

Equivalently, if we introduce the Hamiltonian $\widehat{H}(x,p) := -H(x,p)$, then $\tilde{V}$ and $\tilde{U}$ solve

$$\begin{cases} w_t(x,t) + \widehat{H}\big(x, \nabla_x w(x,t)\big) = 0 & \text{in } \mathbb{R}^n \times (0,T), \\ w(x,0) = g(x) & \text{on } \mathbb{R}^n. \end{cases}$$

By Lemma 2.19 the Hamiltonian H satisfies the conditions of Theorem 2.18. It is immediate that $\widehat{H} = -H$ satisfies the same conditions, with the same constant C_H. Therefore, we may apply Theorem 2.18 to the initial value problem with Hamiltonian $\widehat{H}$ to obtain

$$\tilde{V} = \tilde{U} \quad \text{on } \mathbb{R}^n \times [0,T].$$

Undoing the time change $t \mapsto T - t$ yields

$$V(x,t) = \tilde{V}(x, T-t) = \tilde{U}(x, T-t) = U(x,t)$$

for $(x,t) \in \mathbb{R}^n \times [0,T]$, which means that any bounded and uniformly continuous viscosity solution of the HJB equation (2.17) must coincide with the value function V. This concludes the proof. □

2.4 Design of an optimal control

This section also draws on the book by Evans (2010) and the lecture notes (Evans, 2024).

We have shown that the value function is the unique viscosity solution of the Hamilton–Jacobi–Bellman equation (2.17). However, it is not immediately clear, how this characterization helps to determine the optimal control $\alpha^* \in \mathcal{A}$ that solves the original control problem. In regions where the value function V and our defined α are sufficiently regular, the following procedure can be used to construct the optimal feedback control.

For a given initial time $t \in [0,T)$ and initial state $x \in \mathbb{R}^n$:

1. Solve the Hamilton–Jacobi–Bellman equation and obtain the value function V.

2. For each $(\tilde{x}, \tilde{t}) \in \mathbb{R}^n \times (0,T)$, define

$$\alpha(\tilde{x}, \tilde{t}) := a \in A,$$

to be the value where the minimum in the HJB equation is attained. This means select $\alpha(\tilde{x}, \tilde{t})$ such that

$$V_t(\tilde{x}, \tilde{t}) + \langle f(\tilde{x}, \alpha(\tilde{x}, \tilde{t})), \nabla_x V(\tilde{x}, \tilde{t})\rangle + L(\tilde{x}, \alpha(\tilde{x}, \tilde{t})) = 0.$$

3. Determine the corresponding state trajectory by solving

$$\begin{cases} \frac{d}{ds} X^*(s) = f(X^*(s), \alpha(X^*(s), s)), & s \in (t,T), \\ X^*(t) = x. \end{cases}$$

4. The resulting feedback control is given by

$$\alpha^*(s) := \alpha(X^*(s), s), \quad s \in (t,T).$$

To verify that this feedback control is indeed optimal, observe that along the trajectory X^* generated by α^* we have, by the HJB equation,

$$-V_t(X^*(s), s) - \langle \nabla_x V(X^*(s), s), f(X^*(s), \alpha^*(X^*(s), s))\rangle = L(X^*(s), \alpha^*(X^*(s), s)).$$

Since $\frac{d}{ds}X^*(s) = f(X^*(s), \alpha^*(X^*(s), s))$, this implies

$$-\frac{d}{ds}V(X^*(s), s) = L(X^*(s), \alpha^*(X^*(s), s)).$$

We know

$$C_{x,t}(\alpha^*) = \int_t^T L(X^*(s), \alpha^*(X^*(s), s))\, ds + g(X^*(T)).$$

By plugging in the above identity, we get

$$C_{x,t}(\alpha^*) = \int_t^T -\frac{d}{ds}V(X^*(s), s)\, ds + g(X^*(T)).$$

Using the fundamental theorem of calculus and the terminal condition $V(X^*(T), T) = g(X^*(T))$ yields

$$C_{x,t}(\alpha^*) = V(x, t).$$

Hence, the cost associated with the feedback control α^* equals the value function, confirming its optimality.

2.5 Properties of viscosity solutions of Hamilton–Jacobi equations

We briefly discuss some properties and alternative equivalent definitions of viscosity solutions, as presented in the book by Bardi et al. (1997), which will be needed in the subsequent analysis.

This section is devoted to a more general version of (2.17), namely the Hamilton–Jacobi equation

$$F(x, V(x), \nabla V(x)) = 0 \quad \text{for } x \in \Omega, \tag{2.47}$$

where $\Omega \subseteq \mathbb{R}^n$ is open and F is a real-valued function on $\Omega \times \mathbb{R} \times \mathbb{R}^n$.

Definition 2.21 (Viscosity solution)**.** A function $V \in C(\Omega)$ is a *viscosity subsolution* of (2.47) if, for any $\psi \in C^1(\Omega)$,

$$F(x_0, V(x_0), \nabla\psi(x_0)) \leq 0, \tag{2.48}$$

at any local maximum point $x_0 \in \Omega$ of $V - \psi$. Similarly, $V \in C(\Omega)$ is a *viscosity supersolution* of (2.47) if, for any $\psi \in C^1(\Omega)$,

$$F(x_1, V(x_1), \nabla\psi(x_1)) \geq 0, \tag{2.49}$$

at any local minimum point $x_1 \in \Omega$ of $V - \psi$. Finally, V is a *viscosity solution* of (2.47) if it is both a viscosity sub- and supersolution.

Remark 2.22. In the definition of subsolution we can also assume that x_0 is a local strict maximum point of $V - \psi$, by replacing $\psi(x)$ by $\psi(x) + |x - x_0|^2$. Furthermore, since the inequality (2.48) only depends on the value of $\nabla\psi(x_0)$, we may assume without loss of generality that $V(x_0) = \psi(x_0)$. Geometrically this means that the validity of the subsolution of condition (2.48) for V is tested on smooth functions "touching from above" the graph of V at x_0. Figure 2.1 tries to exemplify this. Analogously, this remark holds for supersolutions as well.

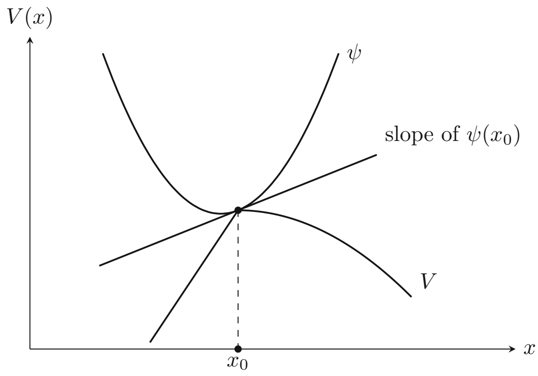

Figure 2.1: Characterization of a subsolution via a test function.

Next, we want to prove the consistency between viscosity solutions and the classical pointwise definition.

Proposition 2.23. *a) If $V \in C(\Omega)$ is a classical solution of (2.47), that is, V is differentiable at any $x \in \Omega$ and*

$$F(x, V(x), \nabla V(x)) = 0 \quad \textit{for } x \in \Omega, \tag{2.50}$$

then V is a viscosity solution of (2.47).

b) If $V \in C^1(\Omega)$ is a viscosity solution of (2.47), then V is a classical solution of (2.47).

Proof. a) Take any $\psi \in C^1(\Omega)$. By the differentiability of V, at any local maximum or minimum of $V - \psi$ we have $\nabla V(x) = \nabla \psi(x)$. Hence (2.50) yields

$$0 = F(x_0, V(x_0), \nabla V(x_0)) = F(x_0, V(x_0), \nabla \psi(x_0)) \leq 0,$$

if x_0 is a local maximum of $V - \psi$ and

$$0 = F(x_1, V(x_1), \nabla V(x_1)) = F(x_1, V(x_1), \nabla \psi(x_1)) \geq 0,$$

if x_1 is a local minimum of $V - \psi$.

b) If $V \in C^1(\Omega)$, then $\psi := V$ is a feasible choice for a test function. With this choice, any $x \in \Omega$ is simultaneously a local maximum and minimum of $V - \psi$. Since V is a viscosity solution, by (2.48) and (2.49) we have

$$0 = F(x, V(x), \nabla \psi(x)) = F(x, V(x), \nabla V(x)).$$

This concludes the proof. □

Remark 2.24. The regularity assumption $V \in C^1(\Omega)$ in part *b)* is not optimal. Proposition 2.30 strengthens this by showing that differentiability at a single point x suffices for the equation to hold at x.

The next example demonstrates that viscosity solutions are not preserved by change of sign in the equation. We show that the one dimensional function $V(x) = 1 - |x|$ is a viscosity solution to the eikonal equation $|V'(x)| - 1 = 0$ but not to $-|V'(x)| + 1 = 0$ for $x \in (-1, 1)$.

Example 2.25 (Eikonal equation)**.** The one dimensional function $V(x) = 1 - |x|$ is a viscosity solution to the eikonal equation

$$\big|V'(x)\big| - 1 = 0 \quad \text{for } x \in (-1, 1). \tag{2.51}$$

- For $x \neq 0$, the function V is differentiable with $V'(x) = 1$ for $x < 0$ and $V'(x) = -1$ for $x > 0$. Hence, V is a classical solution of equation (2.51) away from $x = 0$ and due to Proposition 2.23 a) a viscosity solution.
- For $x = 0$:
  a) To show that V is a viscosity subsolution, we assume $\psi \in C^1(-1, 1)$ and that $x = 0$ is a local maximum point of $V - \psi$. Therefore, there exists $\delta > 0$ such that for every $|y| < \delta$:

$$\begin{aligned} & V(y) - \psi(y) \leq V(0) - \psi(0) \\ \iff \quad & 1 - |y| - \psi(y) \leq 1 - \psi(0) \\ \iff \quad & -|y| \leq \psi(y) - \psi(0). \end{aligned}$$

  Hence, for

  i) $y < 0$: $1 \geq (\psi(y) - \psi(0))/y$. Passing the limit $y \to 0^-$ yields $1 \geq \psi'(0)$.
  ii) $y > 0$: $-1 \leq (\psi(y) - \psi(0))/y$. Passing the limit $y \to 0^+$ yields $-1 \leq \psi'(0)$.

  Together we get $|\psi'(0)| \leq 1 \Leftrightarrow |\psi'(0)| - 1 \leq 0$, which shows that V is a viscosity subsolution at $x = 0$.

  b) To show that V is a viscosity supersolution, we again assume $\psi \in C^1(-1, 1)$ and $x = 0$ is a local minimum point of $V - \psi$. Going through analogous steps yields $-1 \geq \psi'(0)$ and $1 \leq \psi'(0)$, a contradiction. Consequently, there exists no $\psi \in C^1(-1, 1)$ such that $V - \psi$ has a local minimum at $x = 0$. Hence V is a viscosity supersolution at $x = 0$.

On the other hand, $V(x) = 1 - |x|$ is not a viscosity solution of

$$-\big|V'(x)\big| + 1 = 0 \quad \text{for } x \in (-1, 1).$$

The subsolution condition is not fulfilled at $x_0 = 0$. Assume $\psi \in C^1(-1, 1)$ and $x_0 = 0$ is a local maximum point of $V - \psi$. For the function $\psi := 1$, we have

$$V(x) - \psi(x) = (1 - |x|) - 1 = -|x|.$$

As a result, $V - \psi$ has a local maximum at $x_0 = 0$. But $-|\psi'(0)| + 1 = 1 \nleq 0$. Therefore the subsolution condition is not satisfied. ❖

We now present an alternative definition of viscosity solutions to the Hamilton–Jacobi equation (2.47). After that, Proposition 2.27 will establish the equivalence between the two definitions.

Definition 2.26 (Super- and subdifferential)**.** Let $V \in C(\Omega)$ and $x \in \Omega$. The sets

$$\partial_\Omega^+ V(x) := \left\{ p \in \mathbb{R}^n : \limsup_{y \to x,\ y \in \Omega} \frac{V(y) - V(x) - \langle p, y - x \rangle}{|y - x|} \leq 0 \right\},$$

$$\partial_\Omega^- V(x) := \left\{ p \in \mathbb{R}^n : \liminf_{y \to x,\ y \in \Omega} \frac{V(y) - V(x) - \langle p, y - x \rangle}{|y - x|} \geq 0 \right\},$$

are called the *super-* and the *subdifferential* of V at x, respectively.

Proposition 2.27. *Let $V \in C(\Omega)$. Then,*

a) $p \in \partial^+_\Omega V(x)$ if and only if there exists $\psi \in C^1(\Omega)$ such that $\nabla\psi(x) = p$ and $V - \psi$ has a local maximum at x;

b) $p \in \partial^-_\Omega V(x)$ if and only if there exists $\psi \in C^1(\Omega)$ such that $\nabla\psi(x) = p$ and $V - \psi$ has a local minimum at x.

Proof. a) Suppose $p \in \partial^+_\Omega V(x)$. By the definition of the superdifferential, for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$V(y) \leq V(x) + \langle p, y - x\rangle + \varepsilon|y - x|, \qquad \forall\, y \in B(x, \delta).$$

Equivalently, one can introduce a continuous, increasing function $\sigma : [0, \infty) \to \mathbb{R}$ with $\sigma(0) = 0$ and find $\delta > 0$ such that

$$V(y) \leq V(x) + \langle p, y - x\rangle + \sigma(|y - x|)\,|y - x|, \qquad \forall\, y \in B(x, \delta). \tag{2.52}$$

Define $\varrho \in C^1([0, \infty))$ by

$$\varrho(r) = \int_0^r \sigma(t)\,dt.$$

We can immediately see that ϱ satisfies

$$\varrho(0) = \varrho'(0) = 0 \quad \text{and} \quad \varrho(2r) \geq r\sigma(r). \tag{2.53}$$

Define a test function ψ by

$$\psi(y) := V(x) + \langle p, y - x\rangle + \varrho(2|y - x|). \tag{2.54}$$

For $y \neq x$, by differentiation we obtain

$$\nabla\psi(y) = p + 2\varrho'(2|y - x|)\frac{y - x}{|y - x|}.$$

Since $\varrho'(0) = 0$, the function $y \mapsto \varrho(2|y - x|)$ is continuously differentiable at $y = x$ with gradient zero. Therefore $\psi \in C^1(\Omega)$ and $\nabla\psi(x) = p$. Moreover, for $y \in B(x, \delta)$,

$$\begin{aligned}(V - \psi)(y) &\overset{(2.52)}{\leq} V(x) + \langle p, y - x\rangle + \sigma(|y - x|)\,|y - x| - \psi(y)\\ &\overset{(2.54)}{=} \sigma(|y - x|)\,|y - x| - \varrho(2|y - x|)\\ &\overset{(2.53)}{\leq} 0 = (V - \psi)(x).\end{aligned}$$

Therefore, $V - \psi$ achieves a local maximum at x.

Conversely, assume there exists a test function $\psi \in C^1(\Omega)$ such that $(V - \psi)(y) \leq (V - \psi)(x)$ for all $y \in B(x, \delta)$. Rearranging and using the assumption $p = \nabla\psi(x)$ yields

$$V(y) - V(x) - \langle p, y - x\rangle \leq \psi(y) - \psi(x) - \langle\nabla\psi(x), y - x\rangle, \quad \forall\, y \in B(x, \delta).$$

Because ψ is differentiable, dividing both sides by $|y - x|$ and passing the limit $y \to x$ yields that $p \in \partial^+_\Omega V(x)$.

b) Since $\partial^-_\Omega V(x) = -(\partial^+_\Omega(-V)(x))$ the proof of part b) follows analogously when applied to $-V$. ☐

Remark 2.28. As a direct consequence of Proposition 2.27, the following definition of viscosity solution turns out to be equivalent to Definition 2.21: A function $V \in C(\Omega)$ is a viscosity subsolution of (2.47) in Ω if

$$F(x, V(x), p) \leq 0, \quad \forall x \in \Omega, \forall p \in \partial^+_\Omega V(x). \tag{2.55}$$

$V \in C(\Omega)$ is a viscosity supersolution of (2.47) in Ω if

$$F(x, V(x), p) \geq 0, \quad \forall x \in \Omega, \forall p \in \partial^-_\Omega V(x). \tag{2.56}$$

V is called a viscosity solution of (2.47) in Ω if (2.55) and (2.56) hold.

The following lemma is required to extend the results of Proposition 2.23 to the more general case presented in Proposition 2.30.

Lemma 2.29. *Let $V \in C(\Omega)$ and $x \in \Omega$. Then,*

a) if V is differentiable at x, then $\{\nabla V(x)\} = \partial^+_\Omega V(x) = \partial^-_\Omega V(x)$;

b) if for some x both $\partial^+_\Omega V(x)$ and $\partial^-_\Omega V(x)$ are nonempty, then $\partial^+_\Omega V(x) = \partial^-_\Omega V(x) = \{\nabla V(x)\}$.

Proof. Observe that for any $x, y \in \Omega$ and $p, q \in \mathbb{R}^n$ we have

$$\frac{\langle p - q, y - x\rangle}{|y - x|} = \frac{V(y) - V(x) - \langle q, y - x\rangle}{|y - x|} - \frac{V(y) - V(x) - \langle p, y - x\rangle}{|y - x|}. \tag{2.57}$$

For any $n \in \mathbb{N}$, set $y_n := x + (1/n)(p - q)$ and take $y = y_n$ in (2.57) to obtain

$$|p - q| = \frac{V(y_n) - V(x) - \langle q, y_n - x\rangle}{|y_n - x|} - \frac{V(y_n) - V(x) - \langle p, y_n - x\rangle}{|y_n - x|}.$$

Apply $\limsup_{n\to\infty}$ to both sides of the equation. By using properties of the limit inferior and superior we get

$$|p - q| \leq \limsup_{y\to x,\, y\in\Omega} \frac{V(y) - V(x) - \langle q, y - x\rangle}{|y - x|} - \liminf_{y\to x,\, y\in\Omega} \frac{V(y) - V(x) - \langle p, y - x\rangle}{|y - x|}. \tag{2.58}$$

To prove *a)*, assume V is differentiable at x. Thus, there exists a unique vector $\nabla V(x)$ such that

$$\lim_{y\to x,\, y\in\Omega} \frac{V(y) - V(x) - \langle \nabla V(x), y - x\rangle}{|y - x|} = 0.$$

This implies

$$\limsup_{y\to x,\, y\in\Omega} \frac{V(y) - V(x) - \langle \nabla V(x), y - x\rangle}{|y - x|} = 0$$

and

$$\liminf_{y\to x,\, y\in\Omega} \frac{V(y) - V(x) - \langle \nabla V(x), y - x\rangle}{|y - x|} = 0.$$

Therefore, $\partial^+_\Omega V(x) \cap \partial^-_\Omega V(x) \neq \emptyset$ since both sets contain $\nabla V(x)$. Inequality (2.58) implies that the sets $\partial^-_\Omega V(x)$ and $\partial^+_\Omega V(x)$ consist of a single element in this case.

To show *b)*, assume that both $\partial^+_\Omega V(x)$ and $\partial^-_\Omega V(x)$ are nonempty for some x. As a consequence of (2.58), $\partial^+_\Omega V(x) = \partial^-_\Omega V(x)$ is a singleton. If both the super- and subdifferentials coincide and are singletons, then V must be differentiable at x with $\{\nabla V(x)\} = \partial^+_\Omega V(x) = \partial^-_\Omega V(x)$. □

Proposition 2.30. *a) If $V \in C(\Omega)$ is a viscosity solution of (2.47), then*

$$F(x, V(x), \nabla V(x)) = 0$$

at any point $x \in \Omega$ where V is differentiable;

b) if V is locally Lipschitz continuous and is a viscosity solution of (2.47), then

$$F(x, V(x), \nabla V(x)) = 0 \quad \textit{almost everywhere in } \Omega.$$

Proof. a) If x is a point of differentiability for V, then by Lemma 2.29 a) $\{\nabla V(x)\} = \partial^+_\Omega V(x) = \partial^-_\Omega V(x)$. Hence, by (2.55) and (2.56) it follows that

$$0 \geq F(x, V(x), \nabla V(x)) \geq 0.$$

b) The second part of the proposition follows immediately from part a) and Rademacher's theorem A.2. This concludes the proof. □

Last, we establish in Proposition 2.32 a stability result in the uniform topology of $C(\Omega)$, which we for instance need in the proof of Proposition 4.4.

Lemma 2.31. *Let $u \in C(\Omega)$, $\delta > 0$, and suppose that $x_0 \in \Omega$ is a strict maximum point for u in $\overline{B(x_0, \delta)} \subseteq \Omega$. If $u_n \in C(\Omega)$ converges locally uniformly to u in Ω, then there exists a sequence $\{x_n\}$ such that*

$$x_n \to x_0 \quad \textit{and} \quad u_n(x_n) \geq u_n(x), \quad \forall\, x \in \overline{B(x_0, \delta)}.$$

Proof. For each $n \in \mathbb{N}$, u_n is continuous on the compact set $\overline{B(x_0, \delta)}$. Therefore, a maximum is attained on $\overline{B(x_0, \delta)}$. Let x_n be such a maximum point and let $\{x_{n_k}\}$, $k \in \mathbb{N}$, be any converging subsequence of $\{x_n\}$, $n \in \mathbb{N}$, with $x_{n_k} \to \tilde{x}$ as $k \to \infty$. For each $k \in \mathbb{N}$ and every $x \in \overline{B(x_0, \delta)}$ we know

$$u_{n_k}(x_{n_k}) \geq u_{n_k}(x).$$

Fix $x \in \overline{B(x_0, \delta)}$ and let $k \to \infty$. By uniform convergence,

$$u_{n_k}(x_{n_k}) \to u(\tilde{x}), \qquad u_{n_k}(x) \to u(x).$$

Therefore,

$$u(\tilde{x}) \geq u(x), \quad \forall\, x \in \overline{B(x_0, \delta)}$$

and in particular,

$$u(\tilde{x}) \geq u(x_0).$$

Since x_0 is a strict maximum point of u in $\overline{B(x_0, \delta)}$, this implies $\tilde{x} = x_0$. Consequently, every convergent subsequence of $\{x_n\}$ converges to x_0. As $\{x_n\}$ is contained in the compact set $\overline{B(x_0, \delta)}$, it follows that the whole sequence converges, $x_n \to x_0$ as $n \to \infty$. This concludes the proof. □

Proposition 2.32. *Let $V_n \in C(\Omega)$ for $n \in \mathbb{N}$ be a viscosity solution of*

$$F_n(x, V_n(x), \nabla V_n(x)) = 0 \quad \textit{in } \Omega. \tag{2.59}$$

Assume that $V_n \to V$ locally uniformly in Ω and $F_n \to F$ locally uniformly in $\Omega \times \mathbb{R} \times \mathbb{R}^n$. Then V is a viscosity solution of (2.47).

Proof. Let $\psi \in C^1(\Omega)$ and x_0 be a local maximum point of $V - \psi$. As discussed in Remark 2.22, it is not restrictive to assume that x_0 is a strict maximum of $V - \psi$ in a neighborhood of x_0. Therefore, there exists $\delta > 0$ such that

$$V(x_0) - \psi(x_0) > V(x) - \psi(x)$$

for all $x \in \overline{B(x_0, \delta)} \setminus \{x_0\}$. Since $V_n \to V$ locally uniformly in Ω apply Lemma 2.31 to $u := V - \psi$ and $u_n := V_n - \psi$. This yields points $x_n \in \overline{B(x_0, \delta)}$ so that $x_n \to x_0$ and $V_n - \psi$ attains a maximum on $\overline{B(x_0, \delta)}$ at x_n for all large n. Consequently, because V_n is a viscosity solution and in particular a viscosity subsolution of (2.59),

$$F_n(x_n, V_n(x_n), \nabla\psi(x_n)) \leq 0$$

holds. Since $x_n \to x_0$, by passing the limit $n \to \infty$, the above inequality yields

$$F(x_0, V(x_0), \nabla\psi(x_0)) \leq 0.$$

Therefore, V is a viscosity subsolution. Analogously, it can be shown that V is a viscosity supersolution. □

3 Infinite Time Horizon Optimal Control Theory

From this chapter onward our main source will be the paper by Cannarsa et al. (2024). The aim of Chapter 3 is to introduce the infinite time horizon setting and, following (Cannarsa et al., 2024), to establish a comparison principle for bounded and uniformly continuous viscosity solutions of the associated Hamilton–Jacobi–Bellman equation. As a consequence, we obtain that the value function V_λ, which is known to be a bounded and uniformly continuous viscosity solution of this equation from standard dynamic programming arguments, is unique in this class.

3.1 Framework

Definition 3.1 (Value function and state equation)**.** Let $\Omega \subseteq \mathbb{R}^n$ open. For any given $\lambda > 0$ let the function $V_\lambda : \Omega \to \mathbb{R}$ denote the *value function*

$$V_\lambda(x) := \inf_{\alpha \in \mathcal{A}} \int_0^\infty e^{-\lambda t} L\big(X_\alpha^x(t), \alpha(t)\big)\, dt, \tag{3.1}$$

where $\mathcal{A}$ stands for the set of all measurable controls $\alpha : [0, \infty) \to A$ taking values in the complete metric space A. $X_\alpha^x(\cdot)$ denotes the solution of the *state equation*

$$\begin{cases} X'(t) = f\big(X(t), \alpha(t)\big), & t > 0, \\ X(0) = x. \end{cases} \tag{3.2}$$

Remark 3.2. Comparing the definitions of the value function in the finite (2.3) and infinite (3.1) time horizon settings, we observe two main differences. First, in the infinite time horizon case, the starting time is fixed to 0 without loss of generality. Second, an exponential discount factor $e^{-\lambda t}$ is included, which ensures the convergence of the integral over an infinite time interval. In practice, this type of value function is often used in economic models.

Moreover, the discounted value function V_λ provides the starting point for the analysis carried out in Chapter 4, where the asymptotic behavior of the family $\{\lambda V_\lambda\}_{\lambda>0}$ is investigated in the vanishing discount limit $\lambda \to 0^+$.

The following assumptions are assumed to hold throughout this thesis unless stated otherwise.

Assumption 3.3. Let $f : \mathbb{R}^n \times A \to \mathbb{R}^n$ be a continuous map such that

$$\begin{cases} |f(x,a)| \le C_f, \\ |f(x,a) - f(y,a)| \le C_f |x-y|, \end{cases} \qquad \forall\, x, y \in \mathbb{R}^n,\ \forall\, a \in A \tag{3.3}$$

for some constant $C_f > 0$. Let $L : \mathbb{R}^n \times A \to \mathbb{R}$ be a continuous function satisfying

$$\begin{cases} 0 \le L(x,a) \le 1, \\ |L(x,a) - L(y,a)| \le C_L |x-y|, \end{cases} \qquad \forall\, x, y \in \mathbb{R}^n,\ \forall\, a \in A \tag{3.4}$$

for some constant $C_L > 0$. We suppose throughout the thesis that

$$\big\{\big(f(x,a),\ L(x,a)+r\big) : a \in A,\ r \geq 0\big\} \text{ is a closed and convex set for every } x \in \mathbb{R}^n.$$

Furthermore, we assume the existence of a bounded and open domain $\Omega \subseteq \mathbb{R}^n$ such that $\overline{\Omega}$ is invariant for system (3.2), that is,

$$x \in \overline{\Omega} \implies X_\alpha^x(t) \in \overline{\Omega}, \qquad \forall\, \alpha \in \mathcal{A},\ \forall\, t \geq 0. \tag{H_Ω}$$

Remark 3.4. Looking at (3.4), we can easily show that $V_\lambda(x)$ is bounded for all $x \in \overline{\Omega}$:

$$0 \overset{(3.4)}{\leq} \inf_{\alpha\in\mathcal{A}} \int_0^\infty e^{-\lambda t} L(X_\alpha^x(t), \alpha(t))\, dt \overset{(3.4)}{\leq} \int_0^\infty e^{-\lambda t}\, dt = \frac{1}{\lambda}. \tag{3.5}$$

Under the given assumptions we show that the value function V_λ is a unique viscosity solution to the following Hamilton–Jacobi–Bellman equation.

Definition 3.5 (Hamilton–Jacobi–Bellman equation)**.** Let $V : \Omega \to \mathbb{R}$ and $\lambda > 0$. The *Hamilton–Jacobi–Bellman equation* is given by

$$\lambda V(x) + H(x, -\nabla V(x)) = 0 \quad \text{on } \overline{\Omega}, \tag{3.6}$$

where the *Hamiltonian* is defined by

$$H(x,p) := \max_{a\in A}\{\langle f(x,a), p\rangle - L(x,a)\}, \qquad \forall\,(x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Finally, we introduce the notion of a viscosity solution for the Hamilton–Jacobi–Bellman equation in this setting. Notice that this version is similar to the one, we discussed in Section 2.5.

Definition 3.6. A continuous function $V : \overline{\Omega} \to \mathbb{R}$ is a *viscosity solution* of (3.6) on $\overline{\Omega}$ if and only if:

a) V is a viscosity *supersolution* on $\overline{\Omega}$, that is,

$$\lambda V(x) + H(x,-p) \geq 0, \qquad \forall\, p \in \partial_\Omega^- V(x),$$

for all $x \in \overline{\Omega}$, where

$$\partial_\Omega^- V(x) = \left\{ p \in \mathbb{R}^n :\ \liminf_{y\to x,\ y\in\overline{\Omega}} \frac{V(y) - V(x) - \langle p, y-x\rangle}{|y-x|} \geq 0 \right\}$$

is the *subdifferential* of V at x;

b) V is a viscosity *subsolution* on $\overline{\Omega}$, that is,

$$\lambda V(x) + H(x,-p) \leq 0 \qquad \forall\, p \in \partial_\Omega^+ V(x),$$

for all $x \in \overline{\Omega}$, where

$$\partial_\Omega^+ V(x) = \left\{ p \in \mathbb{R}^n :\ \limsup_{y\to x,\ y\in\overline{\Omega}} \frac{V(y) - V(x) - \langle p, y-x\rangle}{|y-x|} \leq 0 \right\}$$

is the *superdifferential* of V at x.

Remark 3.7. Under the assumptions of this chapter the value function V_λ is a bounded and uniformly continuous viscosity solution of (3.6). See for instance (Bardi et al., 1997). Uniqueness is addressed in the next section.

3.2 Uniqueness of viscosity solution

In this section we establish, within the framework introduced in Chapter 3, a comparison result for bounded and uniformly continuous viscosity sub- and supersolutions of (3.6). Combining this result with the fact that the value function V_λ is a bounded and uniformly continuous viscosity solution of (3.6), we then deduce that V_λ is the unique solution in this class.

We begin by introducing necessary definitions.

Definition 3.8. For a function $\theta : \mathbb{R}^n \to \mathbb{R}$,

a) the *epigraph* is denoted by $\mathrm{Epi}(\theta) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : \theta(x) \leq y\}$;

b) the *hypograph* is denoted by $\mathrm{Hypo}(\theta) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : \theta(x) \geq y\}$.

For a set $A \subseteq \mathbb{R}^n$,

c) the *negative polar cone* is defined as $A^- := \{x \in \mathbb{R}^n : \langle x, a\rangle \leq 0, \forall\, a \in A\}$;

d) $T_A(a)$, $a \in A$ denotes the *contingent tangent cone* to A at a. The vector u belongs to $T_A(a)$ if and only if there exist sequences $h_k \to 0^+$ and $u_k \to u$ with $a + h_k u_k \in A$ for any $k \geq 1$.

Definition 3.9 (Viability and invariance)**.** Let X be a finite-dimensional vector space and $K \subseteq X$. Consider for the set-valued function $F : X \rightrightarrows X$ the differential inclusion

$$\begin{cases} x'(t) \in F(x(t)) \\ x(0) = x_0 \end{cases} \qquad \text{for almost every } t \geq 0. \tag{3.7}$$

a) A function $x : [0, \infty) \to X$ is said to be *viable* in K on $[0, \infty)$ if and only if

$$\forall\, t \in [0, \infty) \;:\; x(t) \in K.$$

b) The set K is said to be *viable* under F, if for any initial state $x_0 \in K$, there exists a solution to the differential inclusion (3.7) starting at x_0, which is viable in K.

c) The set K is said to be *invariant* under F if for any initial state $x_0 \in K$, all solutions to the differential inclusion (3.7) are viable in K.

For Lemma 3.11 and Proposition 3.10 we define for $\tilde{\lambda} \geq 0$ and continuous $\ell : \mathbb{R}^n \times A \to \mathbb{R}$ the Hamilton–Jacobi equation as

$$\begin{cases} \tilde{\lambda} V(x) + \widetilde{H}(x, -\nabla V(x)) = 0 & \text{in } \overline{\Omega}, \\ \text{where } \widetilde{H}(x,p) := \max_{a \in A}\{\langle f(x,a), p\rangle - \ell(x,a)\}, & \forall\,(x,p) \in \mathbb{R}^n \times \mathbb{R}^n. \end{cases} \tag{3.8}$$

Notice that $\tilde{\lambda}$ can possibly be equal to zero. Moreover, (3.8) reduces to (3.6), if $\ell := L$. We assume that there exists some $C_\ell > 0$ such that the following condition,

$$\begin{cases} |\ell(x,a)| \leq C_\ell, & \forall\, x, y \in \mathbb{R}^n, \forall\, a \in A, \\ |\ell(x,a) - \ell(y,a)| \leq C_\ell |x - y|, & \\ \{(f(x,a), \ell(x,a) + r) : a \in A,\, r \geq 0\} & \text{is closed and convex for every } x \in \mathbb{R}^n \end{cases} \tag{3.9}$$

is satisfied.

From now on, fix $\tilde{\lambda} \geq 0$. Our first goal is to establish the following proposition, which relies on Lemmas 3.11 and 3.12. Note that slight adaptations to (3.9) and (3.12) were inspired by (Plaskacz, 2003).

Proposition 3.10. *Assume that (3.9) holds true. Consider $\theta : \overline{\Omega} \to \mathbb{R}$ to be a bounded and uniformly continuous function. Then*

a) θ is a viscosity supersolution on $\overline{\Omega}$ to (3.8), if and only if

$$\forall\, x \in \overline{\Omega},\, \exists\, \alpha \in \mathcal{A},\, \forall\, t \geq 0 : \theta(x) \geq e^{-\tilde{\lambda} t}\theta(X^x_\alpha(t)) + \int_0^t e^{-\tilde{\lambda} s}\ell(X^x_\alpha(s), \alpha(s))\, ds; \qquad (3.10)$$

b) θ is a viscosity subsolution on $\overline{\Omega}$ to (3.8), if and only if

$$\forall\, x \in \overline{\Omega},\, \forall\, \alpha \in \mathcal{A},\, \forall\, t \geq 0 : \theta(x) \leq e^{-\tilde{\lambda} t}\theta(X^x_\alpha(t)) + \int_0^t e^{-\tilde{\lambda} s}\ell(X^x_\alpha(s), \alpha(s))\, ds. \qquad (3.11)$$

Lemma 3.11. *Let $\theta : \overline{\Omega} \to \mathbb{R}$ be a bounded and uniformly continuous function and assume that (3.9) holds true. Then*

a) θ satisfies (3.10) if and only if $\operatorname{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is viable for the following differential inclusion

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \in \left\{ \begin{pmatrix} f(x(t), a) \\ \tilde{\lambda} y(t) - \ell(x(t), a) - r \end{pmatrix} : a \in A,\, r \in [0, C_\ell - \ell(x(t), a)] \right\}. \qquad (3.12)$$

b) θ satisfies (3.11) if and only if $\operatorname{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is invariant for the differential inclusion (3.12).

Proof. We proceed in three steps. In Steps 1 and 2, we prove the two implications in a). In Step 3, we briefly explain why b) follows in an analogous way.

Note that, due to (3.9), the right-hand side of the differential inclusion has convex and compact values. Indeed:

- By (3.9), the set $\{(f(x, a),\, \ell(x, a) + r) : a \in A,\, r \geq 0\}$ is closed for every $x \in \mathbb{R}^n$. Since the half-space $\{(v, z) \in \mathbb{R}^n \times \mathbb{R} : z \leq C_\ell\}$ is also closed, their intersection

  $$\{(f(x, a),\, \ell(x, a) + r) : a \in A,\, r \in [0, C_\ell - \ell(x, a)]\}$$

  is closed. Since homeomorphisms map closed sets to closed sets, applying the map $\Phi(v, z) = (v, \tilde{\lambda} y(t) - z)$ ensures that the right-hand side of (3.12) is closed. Furthermore, for fixed x and y, the set is bounded because f and ℓ are bounded by assumption and r varies in a bounded interval. Since the set is both closed and bounded in $\mathbb{R}^{n+1}$, it is compact by the Heine–Borel theorem.

- Similarly, by (3.9), the set $\{(f(x, a),\, \ell(x, a) + r) : a \in A,\, r \geq 0\}$ is convex for every $x \in \mathbb{R}^n$. Intersecting with the the convex half-space $\{(v, z) \in \mathbb{R}^n \times \mathbb{R} : z \leq C_\ell\}$ yields that

  $$\{(f(x, a),\, \ell(x, a) + r) : a \in A,\, r \in [0, C_\ell - \ell(x, a)]\}$$

  is convex. Since affine maps preserve convexity, applying Φ shows that the right-hand side of (3.12) is convex.

As a consequence of assumptions (3.3) and (3.9) we also know that the right-hand side of (3.12) is a Lipschitz set-valued map. These properties will be used in the proof of Proposition 3.10.

Step 1. Suppose that $\operatorname{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is viable for (3.12). This means that for any $(x, y) \in \operatorname{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$, there exists a solution $(x(\cdot), y(\cdot))$ to (3.12) such that $(x(0), y(0)) = (x, y)$ and $(x(t), y(t)) \in \operatorname{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ for any $t \geq 0$. We want to use Theorem A.4, the Filippov measurable selection theorem.

- Define $\widehat{\Omega} := [0,\infty)$, $\widehat{\mathcal{A}}$ the Lebesgue σ-algebra on $[0,\infty)$ and μ as the Lebesgue measure. The tuple $(\widehat{\Omega}, \widehat{\mathcal{A}}, \mu)$ is a complete σ-finite measure space.
- $X := A \times [0,1]$ and $Y := \mathbb{R}^n \times \mathbb{R}$ are two complete separable metric spaces.
- The measurable set-valued map $F : \widehat{\Omega} \rightrightarrows X$ is defined as $F(t) := A \times [0,1]$ for all $t \in \widehat{\Omega}$. Hence, $F(t)$ has closed nonempty values.
- Fix a solution of (3.12), $(x(\cdot), y(\cdot))$ with $(x(0), y(0)) = (x, y)$ and $(x(t), y(t)) \in \operatorname{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ for all $t \geq 0$. Define $g : \widehat{\Omega} \times X \to Y$ by

  $$g(t,(a,s)) := \begin{pmatrix} f(x(t),a) \\ \tilde{\lambda} y(t) - \ell(x(t),a) - s\big(C_\ell - \ell(x(t),a)\big) \end{pmatrix}.$$

  The function $t \mapsto g(t,(a,s))$ is measurable for every fixed $(a,s) \in X$ and $(a,s) \mapsto g(t,(a,s))$ is continuous for every fixed $t \in \widehat{\Omega}$. Therefore, g is a Carathéodory map as in Definition A.3.
- Define the function $h : \widehat{\Omega} \to Y$ by $h(t) := (x'(t), y'(t))$, which is measurable. Since $(x(\cdot), y(\cdot))$ solves the differential inclusion (3.12), for almost every t there exist $a \in A$ and $r \in [0, C_\ell - \ell(x(t),a)]$ such that

  $$h(t) = \big(f(x(t),a),\, \tilde{\lambda} y(t) - \ell(x(t),a) - r\big).$$

  Setting $s := 0$ if $C_\ell - \ell(x(t),a) = 0$, and otherwise $s := r/(C_\ell - \ell(x(t),a)) \in [0,1]$, we obtain $h(t) \in g(t, F(t))$ for almost every $t \in \widehat{\Omega}$.

Consequently, by applying Theorem A.4, there exists $(a(t), s(t)) \in F(t) = A \times [0,1]$ such that $h(t) = g(t,(a(t),s(t)))$ for almost every $t \geq 0$. Setting $\alpha(t) := a(t)$ and $r(t) := s(t)(C_\ell - \ell(x(t),a(t)))$, we have

$$\begin{cases} x'(t) = f(x(t),\alpha(t)) \\ y'(t) = \tilde{\lambda} y(t) - \ell(x(t),\alpha(t)) - r(t) \end{cases} \quad \text{for almost every } t \geq 0.$$

Solving the differential equation with the initial condition $(x(0), y(0)) = (x, y)$ leads to

$$y(t) = e^{\tilde{\lambda} t} y - e^{\tilde{\lambda} t} \int_0^t e^{-\tilde{\lambda} s} [\ell(X_\alpha^x(s), \alpha(s)) + r(s)]\, ds.$$

Since $(X_\alpha^x(t), y(t)) \in \operatorname{Epi}(\theta)$ for all $t \geq 0$, we have

$$\theta(X_\alpha^x(t)) \leq y(t) = e^{\tilde{\lambda} t} y - e^{\tilde{\lambda} t} \int_0^t e^{-\tilde{\lambda} s} [\ell(X_\alpha^x(s), \alpha(s)) + r(s)]\, ds.$$

Since $s(t) \in [0,1]$ and $\ell \leq C_\ell$, it follows that $r(t) \geq 0$. Multiplying by $e^{-\tilde{\lambda} t}$, choosing $y = \theta(x)$, and rearranging the terms yields

$$\begin{aligned} \theta(x) &\geq e^{-\tilde{\lambda} t} \theta(X_\alpha^x(t)) + \int_0^t e^{-\tilde{\lambda} s} [\ell(X_\alpha^x(s), \alpha(s)) + r(s)]\, ds \\ &\geq e^{-\tilde{\lambda} t} \theta(X_\alpha^x(t)) + \int_0^t e^{-\tilde{\lambda} s} \ell(X_\alpha^x(s), \alpha(s))\, ds \end{aligned}$$

for all $t \geq 0$, which concludes Step 1.

Step 2. Conversely, suppose that θ satisfies (3.10). Fix $(x, y) \in \overline{\Omega} \times \mathbb{R}$ such that $y \geq \theta(x)$. By assumption, there exists $\alpha \in \mathcal{A}$ such that (3.10) holds true. Define the functions

$$x(t) := X_\alpha^x(t) \qquad \text{and} \qquad y(t) := e^{\tilde{\lambda} t} y - e^{\tilde{\lambda} t} \int_0^t e^{-\tilde{\lambda} s} \ell(X_\alpha^x(s), \alpha(s))\, ds$$

for $t \geq 0$. Then $x(0) = x$ and $y(0) = y$. Moreover,

$$x'(t) = f(x(t), \alpha(t)) \qquad \text{and} \qquad y'(t) = \tilde{\lambda} y(t) - \ell(x(t), \alpha(t))$$

for almost every $t \geq 0$. Hence, $(x(\cdot), y(\cdot))$ is a solution of the differential inclusion (3.12) with the choice $r(t) := 0$. Rearranging inequality (3.10) yields

$$\theta(x(t)) \leq e^{\tilde{\lambda} t}\theta(x) - e^{\tilde{\lambda} t} \int_0^t e^{-\tilde{\lambda} s} \ell(x(s), \alpha(s))\, ds, \qquad \forall t \geq 0.$$

Since $y \geq \theta(x)$, we obtain $\theta(x(t)) \leq y(t)$ for any $t \geq 0$. Therefore, $\text{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is viable for the differential inclusion (3.12). This concludes Step 2.

Step 3. The proof of part b) follows by an analogous argument, replacing the epigraph $\text{Epi}(\theta)$ by the hypograph $\text{Hypo}(\theta)$ and viability by invariance for the differential inclusion (3.12). We therefore omit the details. □

Lemma 3.12. *Consider $\theta : \overline{\Omega} \to \mathbb{R}$ a bounded and uniformly continuous function. Then for all $x \in \overline{\Omega}$,*

a) $p \in \partial_\Omega^+ \theta(x) \iff (-p, 1) \in \left[T_{\text{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})}(x, \theta(x))\right]^-$;

b) $p \in \partial_\Omega^- \theta(x) \iff (p, -1) \in \left[T_{\text{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})}(x, \theta(x))\right]^-$.

Proof. We proceed in two steps, proving the two implications in part a). Part b) follows by an analogous argument and is therefore omitted.

Step 1. Take $x \in \overline{\Omega}$ and assume $p \in \partial_\Omega^+ \theta(x)$. Let $(u, v) \in T_{\text{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})}(x, \theta(x))$. We have to show that

$$\langle (-p, 1), (u, v) \rangle = -\langle p, u \rangle + v \leq 0$$

holds. We do so by considering the cases $u = 0$ and $u \neq 0$ separately.

By the definition of the contingent tangent cone, there exist sequences $h_k \to 0^+$ and $(u_k, v_k) \to (u, v)$ such that for any $k \geq 1$, $(x + h_k u_k, \theta(x) + h_k v_k) \in \text{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$. This implies

$$\theta(x + h_k u_k) \geq \theta(x) + h_k v_k. \tag{3.13}$$

Since $p \in \partial_\Omega^+ \theta(x)$ we have

$$0 \geq \limsup_{y \to x,\ y \in \overline{\Omega}} \frac{\theta(y) - \theta(x) - \langle p, y - x \rangle}{|y - x|}.$$

Observe that $y_k = x + h_k u_k$ converges to x as $k \to \infty$. If $u_k = 0$ for infinitely many k, then (3.13) gives $v_k \leq 0$ along those indices and hence $v \leq 0$. Thus, without loss of generality we may assume $u_k \neq 0$ for k large enough, so that $h_k |u_k| > 0$. Therefore,

$$0 \geq \limsup_{k \to \infty} \frac{\theta(x + h_k u_k) - \theta(x) - \langle p, h_k u_k \rangle}{h_k |u_k|} \overset{(3.13)}{\geq} \limsup_{k \to \infty} \frac{v_k - \langle p, u_k \rangle}{|u_k|}. \tag{3.14}$$

If $u = 0$, we know $u_k \to 0$ and so $\langle p, u_k\rangle \to 0$ as $k \to \infty$. To satisfy (3.14), $v \leq 0$ must hold. If $u \neq 0$ by taking the limit superior in (3.14) yields

$$0 \geq \frac{v - \langle p, u\rangle}{|u|},$$

hence $v - \langle p, u\rangle \leq 0$, which concludes Step 1.

Step 2. We assume $(-p, 1) \in \left[T_{\mathrm{Hypo}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x, \theta(x))\right]^{-}$, hence

$$-\langle p, u\rangle + v \leq 0, \quad \forall\,(u, v) \in T_{\mathrm{Hypo}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x, \theta(x)) \tag{3.15}$$

and show that

$$\limsup_{y\to x,\ y\in\overline{\Omega}} \frac{\theta(y) - \theta(x) - \langle p, y - x\rangle}{|y - x|} \leq 0.$$

Fix a sequence $y_k \in \overline{\Omega}$ with $y_k \to x$ such that

$$L := \lim_{k\to\infty} \frac{\theta(y_k) - \theta(x) - \langle p, y_k - x\rangle}{|y_k - x|} = \limsup_{y\to x,\ y\in\overline{\Omega}} \frac{\theta(y) - \theta(x) - \langle p, y - x\rangle}{|y - x|}$$

and choose any $c \in \mathbb{R}$ with $c < L$. Define the sequences

$$u_k := \frac{y_k - x}{|y_k - x|} \qquad \text{and} \qquad h_k := |y_k - x|.$$

Note that $h_k \to 0$ as $k \to \infty$ and since $|u_k| = 1$ for all k, there exists a subsequence u_{k_j}, which converges to some u. By abuse of notation we may assume $u_k \to u$. Observe that $y_k = x + h_k u_k$. For any k large enough

$$\frac{\theta(y_k) - \theta(x) - \langle p, y_k - x\rangle}{|y_k - x|} = \frac{\theta(x + h_k u_k) - \theta(x) - \langle p, h_k u_k\rangle}{h_k} > c$$

is satisfied. Consequently, $\theta(x + h_k u_k) - \theta(x) - h_k\langle p, u_k\rangle > c h_k$. Rearranging the terms yields

$$\theta(x + h_k u_k) \geq \theta(x) + h_k(\langle p, u_k\rangle + c),$$

which means that $\big(x + h_k u_k, \theta(x) + h_k(\langle p, u_k\rangle + c)\big) \in \mathrm{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ for any k large enough. Consequently, $(u, \langle p, u\rangle + c) \in T_{\mathrm{Hypo}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x, \theta(x))$. By using assumption (3.15) it follows that

$$-\langle p, u\rangle + \langle p, u\rangle + c = c \leq 0. \tag{3.16}$$

Since $c < L$ was arbitrary, we must have

$$\limsup_{y\to x,\ y\in\overline{\Omega}} \frac{\theta(y) - \theta(x) - \langle p, y - x\rangle}{|y - x|} \leq 0.$$

This concludes the proof. □

Proof of Proposition 3.10. We fix $\theta : \overline{\Omega} \to \mathbb{R}$ a bounded and uniformly continuous function.

a) We proceed in two steps, showing the forward and backward directions, respectively.

Step 1. Suppose θ is a viscosity supersolution on $\overline{\Omega}$ to (3.8). We have to show that θ satisfies (3.10). In light of Lemma 3.11, it is enough to show that $\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is viable for the differential inclusion (3.12). To do this, we must prove that point iii) of Theorem A.7 is valid, use the equivalence with point i), and utilize Theorem A.9 to finish Step 1. First, we check the prerequisites for those theorems.

- $X := \mathbb{R}^n \times \mathbb{R}$ is finite-dimensional.
- $K := \mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is closed, because θ is continuous.
- The map $F(x,y) := \{(f(x,a), \tilde{\lambda} y - \ell(x,a) - r) : a \in A, r \in [0, C_\ell - \ell(x,a)]\}$ is nonempty with compact, convex values and is Lipschitz, hence upper semicontinuous. We showed these properties at the beginning of the proof of Lemma 3.11. Given the assumptions for f and ℓ, we obtain linear growth of F.

It remains to show that

$$\begin{aligned} \forall\, (x,y) \in \mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R}),\, \forall\, (p,q) \in \left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,y)\right]^- : \\ \sup_{(u,v)\in F(x,y)} \langle (u,v), -(p,q)\rangle \geq 0 \end{aligned} \tag{3.17}$$

holds. We may reduce the verification of (3.17) to the boundary points of the set $\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$. Indeed, if $(x,y) \in \mathrm{int}(\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R}))$ is an interior point of $\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$, then there exists $\varepsilon > 0$ such that $B((x,y),\varepsilon) \subseteq \mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$. Fix any vector $u \in \mathbb{R}^{n+1}$ and define $h_k := \varepsilon/(k|u|)$ and $u_k := u$ for all $k \in \mathbb{N}$. Then

$$(x,y) + h_k u_k \in B((x,y),\varepsilon) \subseteq \mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R}), \quad \forall\, k \in \mathbb{N}.$$

Since $u \in \mathbb{R}^{n+1}$ was arbitrary, it follows that

$$T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,y) = \mathbb{R}^{n+1} \quad \text{and} \quad \left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,y)\right]^- = \{0\}.$$

Therefore, (3.17) is trivially satisfied for interior points of $\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$. In particular, any point $(x,y) \in \mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ with $x \in \Omega$ and $y > \theta(x)$ is an interior point of $\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$, hence (3.17) holds automatically there. Thus it suffices to verify (3.17) at boundary points of the form $(x, \theta(x))$ with $x \in \overline{\Omega}$. Moreover, $\sup_{(u,v)\in F(x,y)} \langle (u,v), -(p,q)\rangle \geq 0$ means that there exists $(u,v) \in F(x,y)$ such that $\langle (u,v), (p,q)\rangle \leq 0$. Using these observations, it remains to verify that

$$\begin{aligned} \forall\, x \in \overline{\Omega},\, \forall\, (p,q) \in \left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\right]^-, \exists\, a \in A,\, \exists\, r \in [0, C_\ell - \ell(x,a)] : \\ \langle f(x,a), p\rangle + q\big(\tilde{\lambda}\theta(x) - \ell(x,a) - r\big) \leq 0. \end{aligned} \tag{3.18}$$

Fix $x \in \overline{\Omega}$ and let $(p,q) \in \left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\right]^-$. We first show $q \leq 0$. Since $\mathrm{Epi}(\theta) = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} : \theta(x) \leq y\}$, it follows that $\{x\} \times [\theta(x), \infty) \subseteq \mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$. Consequently,

$$T_{\{x\}\times[\theta(x),\infty)}(x,\theta(x)) = \{0\} \times [0,\infty)$$

and by the monotonicity of the tangent cone $\{0\} \times [0,\infty) \subseteq T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))$. Using the fact that $B \subseteq C \Rightarrow C^- \subseteq B^-$, yields

$$\left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\right]^- \subseteq \left[\{0\} \times [0,\infty)\right]^- = \mathbb{R}^n \times (-\infty, 0],$$

which shows that $q \leq 0$.

Consider the cases $q < 0$ and $q = 0$ separately. We begin with the case $q < 0$. Because the negative polar cone is closed under multiplication by positive scalars, every positive rescaling of (p,q) remains in the cone. Therefore, multiplying by $1/|q|$ yields

$$\left(\frac{p}{|q|}, -1\right) \in \left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\right]^-.$$

Using Lemma 3.12 yields $p/|q| \in \partial_{\Omega}^{-}\theta(x)$. Due to θ being a viscosity supersolution on $\overline{\Omega}$, we obtain

$$\tilde{\lambda}\theta(x) + \max_{a\in A}\left\{-\left\langle f(x,a), \frac{p}{|q|}\right\rangle - \ell(x,a)\right\} \geq 0.$$

Hence, by (3.9) there exists some $a \in A$ such that a maximum is attained. Multiplying both sides by $-|q|$ yields

$$\langle f(x,a), p\rangle - |q|\big(\tilde{\lambda}\theta(x) - \ell(x,a)\big) \leq 0.$$

Due to $q < 0$, this satisfies (3.18) with $r = 0$.

Consider now the case $q = 0$. Obviously, $p = 0$ satisfies (3.18). Continuing, consider $p \neq 0$ and define the function

$$\overline{\theta}(y) := \begin{cases} \theta(y), & y \in \overline{\Omega}, \\ \|\theta\|_\infty + 1, & \text{else,} \end{cases}$$

which is a bounded lower semicontinuous function. Due to $x \in \overline{\Omega}$, we have $\overline{\theta}(x) = \theta(x)$. Moreover, since $\mathrm{Epi}(\theta) \cap (\overline{\Omega}\times\mathbb{R}) \subseteq \mathrm{Epi}(\overline{\theta})$, we have $T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x)) \subseteq T_{\mathrm{Epi}(\overline{\theta})}(x,\overline{\theta}(x))$ and therefore $\big[T_{\mathrm{Epi}(\overline{\theta})}(x,\overline{\theta}(x))\big]^- \subseteq \big[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\big]^-$. Since by assumption $(p,0) \in \big[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\big]^-$, it follows that $(p,0) \in \big[T_{\mathrm{Epi}(\overline{\theta})}(x,\overline{\theta}(x))\big]^-$.

Using Rockafellar's Lemma A.10, we know that there exist converging sequences $x_k \to x$ and $(p_k, q_k) \in \big[T_{\mathrm{Epi}(\overline{\theta})}(x_k,\overline{\theta}(x_k))\big]^-$ such that

$$q_k < 0, \quad (p_k,q_k) \to (p,0), \quad \text{and} \quad \overline{\theta}(x_k) \to \overline{\theta}(x).$$

Observe that if $y \notin \overline{\Omega}$, we have $T_{\mathrm{Epi}(\overline{\theta})}(y,\overline{\theta}(y)) = T_{\mathrm{Epi}(\overline{\theta})}(y,\|\theta\|_\infty + 1)$. In light of Definition 3.8, $(y + h_k u_k, \|\theta\|_\infty + 1 + h_k v_k) \in \mathrm{Epi}(\overline{\theta})$ holds, as long as $v_k \geq 0$. Therefore,

$$T_{\mathrm{Epi}(\overline{\theta})}(y,\overline{\theta}(y)) = T_{\mathrm{Epi}(\overline{\theta})}(y,\|\theta\|_\infty + 1) = \mathbb{R}^n \times [0,\infty).$$

Consequently, we have $\big[T_{\mathrm{Epi}(\overline{\theta})}(y,\overline{\theta}(y))\big]^- = \{0\} \times (-\infty, 0]$. This means for indices k such that $x_k \notin \overline{\Omega}$, we must have $p_k = 0$. But since $(p_k,q_k) \to (p,0)$ with $p \neq 0$, for large k, p_k cannot be 0. As a result, for all indices k large enough,

$$x_k \in \overline{\Omega} \quad \text{and} \quad \overline{\theta}(x_k) = \theta(x_k)$$

holds. Hence,

$$(p_k,q_k) \in \big[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x_k,\theta(x_k))\big]^- \quad \text{for all sufficiently large } k.$$

Multiplying (p_k, q_k) by the positive scalar $1/|q_k|$ yields

$$\left(\frac{p_k}{|q_k|}, -1\right) \in \big[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x_k,\theta(x_k))\big]^-$$

and by Lemma 3.12 we get $p_k/|q_k| \in \partial_{\Omega}^{-}\theta(x_k)$. Due to θ being a viscosity supersolution on $\overline{\Omega}$, we obtain

$$\tilde{\lambda}\theta(x_k) + \max_{a\in A}\left\{-\left\langle f(x_k,a), \frac{p_k}{|q_k|}\right\rangle - \ell(x_k,a)\right\} \geq 0.$$

Multiplying both sides by $|q_k|$ and passing the limit $k \to \infty$ yields

$$\max_{a\in A}\{-\langle f(x,a), p\rangle\} \geq 0,$$

which satisfies (3.18) for $q = 0$ and $r = 0$. Consequently, (3.18) is valid and as mentioned at the beginning of the proof by Theorem A.7 and A.9 we conclude Step 1.

Step 2. We assume that θ is a bounded and uniformly continuous function satisfying (3.10) and want to show that θ is a viscosity supersolution on $\overline{\Omega}$ to (3.8). Take $x \in \overline{\Omega}$ and $p \in \partial^-_{\overline{\Omega}}\theta(x)$. By Lemma 3.12 we know that $(p, -1) \in \left[T_{\mathrm{Epi}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x, \theta(x))\right]^-$ and in view of Lemma 3.11, the set $\mathrm{Epi}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is viable for the differential inclusion (3.12). By Theorem A.9 and A.7, there exists $a \in A$ and $r \in [0, C_\ell - \ell(x, a)]$ such that

$$\langle f(x,a), p\rangle + (-1)\big(\tilde{\lambda}\theta(x) - \ell(x,a) - r\big) \leq 0.$$

Rearranging above inequality yields

$$\tilde{\lambda}\theta(x) - \langle f(x,a), p\rangle - \ell(x,a) \geq r \geq 0,$$

which means that this inequality also holds for the maximum argument $a \in A$. Consequently, θ is a viscosity supersolution on $\overline{\Omega}$.

b) The proof of part b) follows along the same lines as that of part a). Therefore, we only briefly describe it. We again proceed in two steps, showing the forward and backward directions, respectively.

Step 1. Suppose that θ is a viscosity subsolution on $\overline{\Omega}$. We wish to obtain that θ satisfies condition (3.11). By Lemma 3.11, it is enough to show that $\mathrm{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is invariant for the differential inclusion (3.12). To prove this invariance property, we use the Invariance Theorem A.5. More precisely, we verify point ii) of the theorem and utilize the equivalence with point i). For this purpose, we apply Theorem A.5 to the set-valued function $G := -F$, where F denotes the right-hand side of the differential inclusion (3.12). Applying Theorem A.6 concludes that $\mathrm{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R})$ is invariant for the differential inclusion (3.12).

To verify point ii) of Theorem A.5, we need to prove that

$$\begin{aligned}\forall\,(x,y) \in \mathrm{Hypo}(\theta) \cap (\overline{\Omega} \times \mathbb{R}),\ \forall\,(p,q) \in \left[T_{\mathrm{Hypo}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,y)\right]^- :&\\ \sup_{(u,v)\in -F(x,y)} \langle (u,v), -(p,q)\rangle \leq 0.&\end{aligned} \tag{3.19}$$

Using a similar argument as in the proof of point a), it is enough to reduce the verification of (3.19) to the boundary points $(x, \theta(x))$ of the hypograph. Therefore, it remains to verify that

$$\begin{aligned}\forall\, x \in \overline{\Omega},\ \forall\,(p,q) \in \left[T_{\mathrm{Hypo}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\right]^-,\ \forall\, a \in A,\ \forall\, r \in [0, C_\ell - \ell(x,a)] :&\\ \langle f(x,a), p\rangle + q\big(\tilde{\lambda}\theta(x) - \ell(x,a) - r\big) \leq 0.&\end{aligned} \tag{3.20}$$

We fix $x \in \overline{\Omega}$ and $(p,q) \in \left[T_{\mathrm{Hypo}(\theta)\cap(\overline{\Omega}\times\mathbb{R})}(x,\theta(x))\right]^-$ and show analogously to a) that $q \geq 0$. For the case $q > 0$, we show with the help of Lemma 3.12 that $-p/q \in \partial^+_{\overline{\Omega}}\theta(x)$ and deduce (3.20). For $q = 0$, we use an analogue version of the Rockafellar Lemma A.10 for hypographs and proceed exactly as in a).

Step 2. This step works analogously to Step 2 in the proof of a) and is therefore omitted, which concludes the proof. □

Theorem 3.13 (Comparison principle)**.** *Let θ_1 and θ_2 be two bounded and uniformly continuous functions from $\overline{\Omega}$ to $\mathbb{R}$. If θ_1 is a viscosity subsolution to (3.6) on $\overline{\Omega}$ and θ_2 is a viscosity supersolution to (3.6) on $\overline{\Omega}$, then*

$$\forall\, x \in \overline{\Omega} : \theta_1(x) \leq V_\lambda(x) \leq \theta_2(x).$$

Proof. Suppose that $\theta_1 : \overline{\Omega} \to \mathbb{R}$ and $\theta_2 : \overline{\Omega} \to \mathbb{R}$ are as in Theorem 3.13.

Using the assumption that θ_2 is a viscosity supersolution to (3.6), by Proposition 3.10 with $\ell := L$, we know that θ_2 satisfies (3.10). By passing the limit $t \to \infty$ on (3.10), due to the boundedness of θ_2, we obtain

$$\theta_2(x) \geq \int_0^\infty e^{-\lambda s} L(X_\alpha^x(s), \alpha(s))\, ds.$$

The value function minorizes the right hand side of the above inequality, hence $V_\lambda(x) \leq \theta_2(x)$.

Again, using the assumption that θ_1 is a viscosity subsolution of (3.6), by Proposition 3.10 with $\ell := L$, we know that θ_1 satisfies (3.11). By passing the limit $t \to \infty$ on (3.11), we obtain

$$\forall\, \alpha \in \mathcal{A} : \theta_1(x) \leq \int_0^\infty e^{-\lambda s} L(X_\alpha^x(s), \alpha(s))\, ds.$$

Hence, this inequality also holds for the infimum over all $\alpha \in \mathcal{A}$. Therefore,

$$\theta_1(x) \leq \inf_{\alpha \in \mathcal{A}} \int_0^\infty e^{-\lambda s} L(X_\alpha^x(s), \alpha(s))\, ds = V_\lambda(x),$$

which concludes the proof. □

Corollary 3.14. *The value function V_λ is the unique, bounded, and uniformly continuous viscosity solution on $\overline{\Omega}$ to (3.6).*

Proof. Recall from Remark 3.7 that the value function V_λ is a bounded and uniformly continuous viscosity solution on $\overline{\Omega}$ to (3.6). Hence it only remains to prove uniqueness. Let $\theta : \overline{\Omega} \to \mathbb{R}$ be another bounded and uniformly continuous viscosity solution to (3.6). Then θ is both a viscosity subsolution and a viscosity supersolution on $\overline{\Omega}$.

Applying Theorem 3.13 with $\theta_1 := \theta$ and $\theta_2 := V_\lambda$ yields

$$\theta(x) \leq V_\lambda(x), \qquad \forall\, x \in \overline{\Omega}.$$

Applying Theorem 3.13 again with $\theta_1 := V_\lambda$ and $\theta_2 := \theta$ gives

$$V_\lambda(x) \leq \theta(x), \qquad \forall\, x \in \overline{\Omega}.$$

Thus $\theta(x) = V_\lambda(x)$ for all $x \in \overline{\Omega}$, which shows that V_λ is the unique, bounded, and uniformly continuous viscosity solution to (3.6). □

4 Analysis of Vanishing Discount Limits in Continuous Time

This chapter is devoted to analyzing the behavior of the rescaled family of value functions $\{\lambda V_\lambda\}_{\lambda>0} \subseteq C(\overline{\Omega})$ as the discount rate $\lambda \to 0^+$ in the topology of uniform convergence on $\overline{\Omega}$. The limit $\lim_{\lambda\to 0^+} \lambda V_\lambda$ can be thought of as a long-run average cost criterion.

The main results are Theorem 4.10, which characterizes the Lipschitz cluster points of $\{\lambda V_\lambda\}_{\lambda>0}$, if they exist, and Theorem 4.23, which estimates the rate of convergence of λV_λ.

The framework established in Section 3.1 is utilized throughout this chapter. In particular, Assumption 3.3 remains in effect.

Remark 4.1. The present chapter investigates structural properties under the assumption that cluster points of $\{\lambda V_\lambda\}_{\lambda>0}$ exist. The existence of cluster points of such a family, in the uniform topology on $\overline{\Omega}$ can be deduced by assuming equicontinuity on $\overline{\Omega}$ of $\{\lambda V_\lambda\}_{\lambda>0}$. Under this additional hypothesis, the Arzelà–Ascoli Theorem A.11 yields the desired compactness.

Remark 4.2. Numerous works have proposed conditions ensuring the existence of cluster points of the family $\{\lambda V_\lambda\}_{\lambda>0}$. A typical assumption is to impose coercivity of $H(x,\cdot)$, which for instance is discussed in (Bardi et al., 1997):

$$H(x,p) \to \infty, \qquad \text{as } |p| \to \infty,$$

uniformly with respect to x. This assumption yields that every cluster point of $\{\lambda V_\lambda\}_{\lambda>0}$, in the uniform convergence topology, is constant.

Another common assumption is dissipativity, see Remark 4.12. This condition likewise guarantees that λV_λ converges uniformly to a unique constant.

4.1 Cluster points of $\{\lambda V_\lambda\}_{\lambda>0}$

Definition 4.3 (Reduced Hamiltonian)**.** The *reduced Hamiltonian* $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$h(x,p) := \max_{a\in A} \langle f(x,a), p\rangle, \qquad \forall\, (x,p) \in \mathbb{R}^n \times \mathbb{R}^n. \tag{4.1}$$

The next result shows that a cluster point of the family $\{\lambda V_\lambda\}_{\lambda>0}$ is a viscosity solution to the following Hamilton–Jacobi equation.

Proposition 4.4. *If v^* is a cluster point of the family $\{\lambda V_\lambda\}_{\lambda>0}$ as $\lambda \to 0^+$ in the uniform topology on $\overline{\Omega}$, then v^* is a viscosity solution of the equation*

$$h(x, -\nabla v^*(x)) = 0 \quad \textit{in } \overline{\Omega}.$$

Proof. Let $\lambda_j \overset{j\to\infty}{\longrightarrow} 0^+$ be a subsequence such that $\lambda_j V_{\lambda_j}$ converges uniformly to v^*. We know that for each $j \in \mathbb{N}$, the value function V_{λ_j} is a viscosity solution to the HJB equation

$$\lambda_j V_{\lambda_j} + H(x, -\nabla V_{\lambda_j}) = 0.$$

Multiplying by $\lambda_j > 0$ yields

$$
\begin{aligned}
0 &= \lambda_j\big(\lambda_j V_{\lambda_j} + H(x, -\nabla V_{\lambda_j})\big) \\
&= \lambda_j^2 V_{\lambda_j} + \max_{a\in A}\{\langle f(x,a), -\nabla(\lambda_j V_{\lambda_j})\rangle - \lambda_j L(x,a)\}
\end{aligned}
$$

in the viscosity sense. By the stability result of Proposition 2.32, as $\lambda_j \to 0^+$, we conclude

$$0 = \max_{a\in A}\langle f(x,a), -\nabla v^*(x)\rangle = h(x, -\nabla v^*(x)) \quad \text{in } \overline{\Omega}.$$

□

We now introduce the set of coupled viscosity subsolutions $\mathcal{S}(\overline{\Omega})$.

Definition 4.5. We define $\mathcal{S}(\overline{\Omega})$ to be the set of all pairs $(u,v) \in C(\overline{\Omega}) \times C(\overline{\Omega})$ which satisfy

$$\begin{cases} h(x, -\nabla v(x)) \le 0 \\ v(x) + H(x, -\nabla u(x)) \le 0 \end{cases} \quad \text{on } \overline{\Omega} \tag{H}$$

in the viscosity sense, meaning in the sense of Definition 3.6 b).

Proposition 4.6. *Let* $\lambda_j \overset{j\to\infty}{\longrightarrow} 0^+$ *be such that* $\lambda_j V_{\lambda_j} \overset{j\to\infty}{\longrightarrow} v^*$ *uniformly on* $\overline{\Omega}$. *Then*

$$\big(V_{\lambda_j}, v^* - \varepsilon_j\big) \in \mathcal{S}(\overline{\Omega}), \qquad \forall j \in \mathbb{N},$$

where $\varepsilon_j := \|\lambda_j V_{\lambda_j} - v^*\|_\infty$.

Proof. Due to v^* being a cluster point of the family $\{\lambda V_\lambda\}_{\lambda>0}$ as $\lambda \to 0^+$, thanks to Proposition 4.4 we have that

$$h(x, -\nabla v^*) = h(x, -\nabla(v^* - \varepsilon_j)) = 0 \quad \text{in } \overline{\Omega}$$

in the viscosity sense. Consequently, $h(x, -\nabla(v^* - \varepsilon_j)) \le 0$ holds for all $j \in \mathbb{N}$ in the viscosity sense. This shows the first inequality of system (H).

Since V_{λ_j} is a viscosity solution of the HJB equation $\lambda_j V_{\lambda_j} + H(x, -\nabla V_{\lambda_j}) = 0$ we have

$$H(x, -\nabla V_{\lambda_j}) = -\lambda_j V_{\lambda_j} \quad \text{on } \overline{\Omega}$$

in the viscosity sense. We know that $\lambda_j V_{\lambda_j} \to v^*$ uniformly. Therefore, $|\lambda_j V_{\lambda_j}(x) - v^*(x)| \le \varepsilon_j$ for all $x \in \overline{\Omega}$, which yields

$$H(x, -\nabla V_{\lambda_j}) = -\lambda_j V_{\lambda_j} \le -v^* + \varepsilon_j \quad \text{on } \overline{\Omega}$$

in the viscosity sense. Consequently, by rearranging above inequality we have

$$(v^* - \varepsilon_j) + H(x, -\nabla V_{\lambda_j}) \le 0 \quad \text{on } \overline{\Omega},$$

in the viscosity sense, which shows the second inequality of system (H). □

4.2 Characterization of the vanishing discount limit

The goal of this section is to establish a precise characterization of the vanishing discount limit for the rescaled value functions λV_λ as $\lambda \to 0^+$. Furthermore, we show that this limit coincides with the infimum over the set of reachable points of the minimal Lagrangian.

Definition 4.7. We define the set $\mathcal{S}_{\text{Lip}}(\overline{\Omega})$ as the set of all pairs of Lipschitz continuous functions in $\mathcal{S}(\overline{\Omega})$:

$$\mathcal{S}_{\text{Lip}}(\overline{\Omega}) := \{(u,v) \in \mathcal{S}(\overline{\Omega}) : u, v \in \text{Lip}(\overline{\Omega})\},$$

where

$$\text{Lip}(\overline{\Omega}) := \left\{ f \in C(\overline{\Omega}) : \sup_{x \neq y} \frac{|f(y) - f(x)|}{|y - x|} < \infty \right\}.$$

Proposition 4.8. *Let $(\bar{u}, \bar{v}) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega})$. Then the following hold true for any $\lambda > 0$:*

a) for any $c \in \mathbb{R}$, the function

$$u_c(x) := \frac{\bar{v}(x)}{\lambda} + \bar{u}(x) + c \qquad \text{for } x \in \overline{\Omega}, \tag{4.2}$$

satisfies

$$\lambda u_c + H(x, -\nabla u_c) \leq \lambda(\bar{u} + c) \quad \text{on } \overline{\Omega} \tag{4.3}$$

in the viscosity sense;

b) V_λ satisfies the lower bound

$$\bar{u}(x) - \|\bar{u}\|_\infty \leq V_\lambda(x) - \frac{\bar{v}(x)}{\lambda} \qquad \forall x \in \overline{\Omega}. \tag{4.4}$$

Proof. We apply Proposition 3.10 b) for $\theta := \bar{u}$, $\ell := L - \bar{v}$ and $\tilde{\lambda} := 0$. We first verify that (3.9) holds true. By Assumption 3.3, the function L is uniformly bounded and Lipschitz in its first argument. Moreover, $\bar{v}$ is a Lipschitz continuous function on the bounded set $\overline{\Omega}$. Hence the first and the second condition in (3.9) are satisfied. Assumption 3.3 also ensures the convexity and closedness of the set

$$\{(f(x,a),\, L(x,a) + r) : a \in A,\ r \geq 0\} \tag{4.5}$$

for every $x \in \overline{\Omega}$. Subtracting $\bar{v}(x)$ from the second component of (4.5) is an affine translation, therefore convexity and closedness are preserved. Hence the third condition of (3.9) is satisfied.

Recall the definition of $\widetilde{H}$ in (3.8). Because $(\bar{u}, \bar{v}) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega}) \subseteq \mathcal{S}(\overline{\Omega})$, the pair $(\bar{u}, \bar{v})$ satisfies (H). We rearrange the second line of (H) to

$$\bar{v}(x) + H(x, -\nabla \bar{u}) \leq 0 \quad \Longleftrightarrow \quad \widetilde{H}(x, -\nabla \bar{u}) \leq 0 \quad \text{on } \overline{\Omega}.$$

Therefore $\bar{u}$ is a viscosity subsolution to (3.8) and by Proposition 3.10 b) for all $x \in \overline{\Omega}$ and $\alpha \in \mathcal{A}$

$$\bar{u}(x) \leq \bar{u}(X^x_\alpha(t)) + \int_0^t L(X^x_\alpha(s), \alpha(s)) - \bar{v}(X^x_\alpha(s))\, ds, \quad \forall t \geq 0. \tag{4.6}$$

The condition (3.9) is also satisfied for $\theta := \bar{v}$, $\ell := 0$ and $\tilde{\lambda} := 0$. Rearranging the first line of (H) yields that $\bar{v}$ is a viscosity subsolution to (3.8), therefore by applying Proposition 3.10 b), for all $x \in \overline{\Omega}$ and $\alpha \in \mathcal{A}$

$$\bar{v}(x) \leq \bar{v}(X^x_\alpha(t)), \qquad \forall t \geq 0. \tag{4.7}$$

Multiplying (4.7) by $1/\lambda$, adding any $c \in \mathbb{R}$, and combining it with (4.6) yields

$$u_c(x) \le u_c(X^x_\alpha(t)) + \int_0^t L(X^x_\alpha(s), \alpha(s)) - \bar{v}(X^x_\alpha(s))\, ds, \quad \forall t \ge 0.$$

By using Proposition 3.10 b) one last time with $\theta := u_c$, $\ell := L - \bar{v}$ and $\tilde{\lambda} := 0$ we obtain

$$\widetilde{H}(x, -\nabla u_c) \le 0 \quad \text{on } \overline{\Omega} \tag{4.8}$$

in the viscosity sense. Rearranging (4.8) yields

$$\begin{aligned} \widetilde{H}(x, -\nabla u_c) \le 0 \quad &\iff \quad \sup_{a \in A}\{\langle f(x,a), -\nabla u_c\rangle - L(x,a) + \bar{v}(x)\} \le 0 \\ &\iff \quad \bar{v}(x) + H(x, -\nabla u_c) \le 0 \quad \text{on } \overline{\Omega} \end{aligned} \tag{4.9}$$

in the viscosity sense. By using the identity $\bar{v}(x) = \lambda(u_c(x) - \bar{u}(x) - c)$ we obtain (4.3).

Defining $c := -\|\bar{u}\|_\infty$ on (4.3), we have

$$\lambda u_c + H(x, -\nabla u_c) \le \lambda(\bar{u} - \|\bar{u}\|_\infty) \le 0 \quad \text{on } \overline{\Omega}$$

in the viscosity sense. This shows that u_c is a viscosity subsolution of (3.6). By the comparison principle Theorem 3.13 we obtain $u_c \le V_\lambda$ on $\overline{\Omega}$. Inserting the definition of u_c from (4.2) and rearranging the resulting inequality yields (4.4). □

Remark 4.9. Note that in general the existence of a cluster point v^* of the family $\{\lambda V_\lambda\}_{\lambda>0}$ is insufficient to guarantee that $\mathcal{S}_{\text{Lip}}(\overline{\Omega}) \ne \emptyset$. Although for each fixed λ_j the corresponding value function V_{λ_j} is Lipschitz continuous, see for instance (Bardi et al., 1997), this property alone does not yield the regularity needed. However, by Proposition 4.6, if $v^* \in \text{Lip}(\overline{\Omega})$, then $\mathcal{S}_{\text{Lip}}(\overline{\Omega})$ is indeed nonempty.

If we assume that a Lipschitz continuous cluster point of $\{\lambda V_\lambda\}_{\lambda>0}$ exists, then it is unique and can be characterized as follows:

Theorem 4.10 (Characterization)**.** *Let v^* be a cluster point of $\{\lambda V_\lambda\}_{\lambda>0}$ as $\lambda \to 0^+$, in the uniform topology on $\overline{\Omega}$ such that $v^* \in \text{Lip}(\overline{\Omega})$. Then, we have that*

$$v^*(x) = \sup\{v(x) : (u,v) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega})\} \qquad \text{for } x \in \overline{\Omega}. \tag{4.10}$$

Proof. Define v_0 as the right-hand side of (4.10), that is

$$v_0(x) := \sup\{v(x) : (u,v) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega})\} \qquad \text{for } x \in \overline{\Omega}. \tag{4.11}$$

We prove that the cluster point v^* is equal to v_0.

Let $\lambda_j \to 0^+$ be such that $\lambda_j V_{\lambda_j}$ converges uniformly to v^* as $j \to \infty$. As mentioned in Remark 4.9, Proposition 4.6 yields $(V_{\lambda_j}, v^* - \varepsilon_j) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega})$, where $\varepsilon_j := \|\lambda_j V_{\lambda_j} - v^*\|_\infty$ for all $j \in \mathbb{N}$. Consequently, by the definition of v_0 in (4.11), we have $v^* - \varepsilon_j \le v_0$ for all $j \in \mathbb{N}$, which implies $v^* \le v_0$.

Let $(\bar{u}, \bar{v}) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega})$ be arbitrary. Then by Proposition 4.8 b), V_{λ_j} satisfies the lower bound

$$\bar{u}(x) - \|\bar{u}\|_\infty \le V_{\lambda_j}(x) - \frac{\bar{v}(x)}{\lambda_j}, \qquad \forall x \in \overline{\Omega}, \forall j \in \mathbb{N}.$$

Multiplying both sides by λ_j and rearranging yields

$$\bar{v}(x) \le \lambda_j V_{\lambda_j}(x) + \lambda_j(\|\bar{u}\|_\infty - \bar{u}(x)), \qquad \forall x \in \overline{\Omega}, \forall j \in \mathbb{N}.$$

By passing the limit $j \to \infty$, we have $\lambda_j \to 0^+$ and by assumption, $\lambda_j V_{\lambda_j} \to v^*$. Consequently, $\bar{v} \le v^*$. As $(\bar{u}, \bar{v}) \in \mathcal{S}_{\text{Lip}}(\overline{\Omega})$ was arbitrary, $v_0 \le v^*$ follows. □

Corollary 4.11. *Suppose $\{\lambda V_\lambda\}_{\lambda>0}$ is equi-Lipschitz. Then λV_λ converges uniformly as $\lambda \to 0^+$ to the function v^* defined in (4.10).*

Proof. In Remark 3.4 we discussed the boundedness of the value function. Multiplying (3.5) by λ yields the uniform boundedness of $\{\lambda V_\lambda\}_{\lambda>0}$. Consequently, by the Arzelà–Ascoli Theorem A.11, we obtain that there exists a sequence $\lambda_j \to 0^+$ such that $\lambda_j V_{\lambda_j} \to v_0$ uniformly on $\overline{\Omega}$ for some $v_0 \in C(\overline{\Omega})$. Since $\{\lambda V_\lambda\}_{\lambda>0}$ is equi-Lipschitz, $v_0 \in \mathrm{Lip}(\overline{\Omega})$. Hence, by Theorem 4.10, v_0 must coincide with the function v^* defined in (4.10).

In particular, every cluster point of $\{\lambda V_\lambda\}_{\lambda>0}$ must coincide with v^*. So, recalling that $\{\lambda V_\lambda\}_\lambda$ is relatively compact with respect to the uniform distance, we deduce that $\{\lambda V_\lambda\}_{\lambda>0}$ converges uniformly to v^* as $\lambda \to 0^+$. □

Remark 4.12. Any one of the following additional assumptions independently ensures that the family $\{\lambda V_\lambda\}_{\lambda>0}$ is equi-Lipschitz.

a) Controllability, see (Bardi et al., 1997):

$$\exists\, r > 0,\ \forall\, x \in \overline{\Omega} : B(0, r) \subseteq \overline{\mathrm{co}}(f(x, A)).$$

b) Dissipativity, see (Bardi et al., 1997):

$$\exists\, K > 0,\ \forall\, x, y \in \mathbb{R}^n : \sup_{a\in A} \inf_{b\in A} \langle f(x, a) - f(y, b), x - y\rangle \le -K|x - y|^2.$$

c) Nonexpansivity, see (Cannarsa et al., 2015): Suppose for all $x, y \in \overline{\Omega}$ and $a \in A$, there exists $b \in A$ such that

$$\begin{cases} i) & \langle f(x, a) - f(y, b), x - y\rangle \le 0, \\ ii) & L(x, a) - L(y, b) \le K|x - y|. \end{cases}$$

The example below shows that under the assumptions of Corollary 4.11, it is not guaranteed to find $u_0 \in \mathrm{Lip}(\overline{\Omega})$ such that $(u_0, v^*) \in \mathcal{S}_{\mathrm{Lip}}(\overline{\Omega})$.

Example 4.13. Let $A = [0, 1] \subseteq \mathbb{R}$. Consider $\overline{\Omega} = [0, 1]$ and $X'(t) = -\alpha(t)X(t)$ for $t > 0$, $X(0) = x_0 \in [0, 1]$. Therefore,

$$X(t) = x_0 e^{-\int_0^t \alpha(s)\,ds} \qquad \text{for } t \ge 0.$$

As $\alpha(t) \in [0, 1]$, the trajectory $X(t)$ remains in $[0, 1]$ for all $t \ge 0$, which means that $\overline{\Omega}$ is invariant for X'. Furthermore, the dynamic $f(x, a) = -ax$ is bounded and Lipschitz with respect to x for $x, a \in [0, 1]$.

Consider the Lagrangian

$$L(x, a) = 1 - x\sqrt{a} \qquad \text{for } x, a \in [0, 1].$$

This function is Lipschitz with respect to x and bounded between 0 and 1. For the example to satisfy Assumption 3.3, it remains to show that the set

$$\big\{\big(f(x, a),\ L(x, a) + r\big) : a \in [0, 1],\ r \ge 0\big\} = \big\{\big(-ax,\ 1 - x\sqrt{a} + r\big) : a \in [0, 1],\ r \ge 0\big\}$$

is closed and convex for $x \in [0, 1]$. After some calculations, one can show that the above set is the epigraph of the convex and continuous function $u \mapsto 1 - \sqrt{-ux}$ on $[-x, 0]$. Consequently, it is convex and closed.

Moreover,

$$V_\lambda(x) = \frac{1}{\lambda} - \frac{x}{2\sqrt{\lambda}} \qquad \text{for } x, a \in [0,1]$$

is the viscosity solution of the HJB equation

$$\lambda V_\lambda(x) + \max_{a\in[0,1]} \left\{ a x V_\lambda'(x) + x\sqrt{a} - 1 \right\} = 0.$$

Indeed, for $x \in (0,1)$, the function $V_\lambda(x)$ satisfies the equation pointwise in the classical sense. Furthermore, direct computation confirms that the sub- and supersolution inequalities are satisfied at the boundaries $x = 0$ and $x = 1$.

If we consider λ in a bounded interval, for instance $\lambda \in (0,1]$, the set $\{\lambda V_\lambda\}_{\lambda\in(0,1]}$ is equi-Lipschitz with constant $1/2$.

Observe that, although λV_λ converges uniformly to the constant function $v^* = 1$ on $[0,1]$ as $\lambda \to 0^+$, the family

$$V_\lambda(x) - \frac{\bar{v}(x)}{\lambda} = -\frac{x}{2\sqrt{\lambda}} \qquad \text{for } x \in [0,1]$$

fails to be bounded from below uniformly in λ. Therefore, by Proposition 4.8 b) there exists no $u_0 \in \mathrm{Lip}(\overline{\Omega})$ such that $(u_0, v^*) \in \mathcal{S}_{\mathrm{Lip}}(\overline{\Omega})$. Note that this does not contradict Proposition 4.6, as $(V_{\lambda_j}, v^* - \varepsilon_j) \in \mathcal{S}_{\mathrm{Lip}}(\overline{\Omega})$, where by definition ε_j tends to 0 as $j \to \infty$.❖

In Proposition 4.15 below, we show that the limit of $\{\lambda V_\lambda\}_{\lambda>0}$ coincides with the infimum over the set of reachable points of the minimal Lagrangian.

Definition 4.14 (Reachable points)**.** By $\mathcal{R}(x)$ we denote the set of *reachable points* with initial position $x \in \overline{\Omega}$. That is

$$\mathcal{R}(x) := \{X_\alpha^x(t) : \alpha \in \mathcal{A},\, t \geq 0\}.$$

We define the function I by

$$I(x) := \inf_{y\in\mathcal{R}(x)} \min_{a\in A} L(y,a), \qquad \forall\, x \in \overline{\Omega}. \tag{4.12}$$

Proposition 4.15. *Assume $\{\lambda V_\lambda\}_{\lambda>0}$ converges to the function v^* defined in (4.10). Moreover, suppose that for any $x \in \overline{\Omega}$ and any $\varepsilon > 0$ there exists $\alpha_\varepsilon \in \mathcal{A}$, $T_\varepsilon > 0$, and $a_\varepsilon \in A$ such that*

$$L(X_{\alpha_\varepsilon}^x(T_\varepsilon), a_\varepsilon) \leq I(x) + \varepsilon \quad \text{and} \quad f(X_{\alpha_\varepsilon}^x(T_\varepsilon), a_\varepsilon) = 0. \tag{4.13}$$

Then, $v^*(x) = I(x)$ for all $x \in \overline{\Omega}$.

Proof. We prove that $\lambda V_\lambda(x) \to I(x)$ as $\lambda \to 0^+$, which implies that $v^*(x) = I(x)$ for any $x \in \overline{\Omega}$. We proceed in two steps. In Step 1, we show that $\lambda V_\lambda(x) \geq I(x)$ for $x \in \overline{\Omega}$ and in Step 2 that $\limsup_{\lambda\to 0^+} \lambda V_\lambda(x) \leq I(x)$ for $x \in \overline{\Omega}$.

Step 1. By the definition of the value function in (3.1), we have

$$\lambda V_\lambda(x) = \lambda \inf_{\alpha\in\mathcal{A}} \int_0^\infty e^{-\lambda t} L(X_\alpha^x(t), \alpha(t))\, dt \qquad \text{for } x \in \overline{\Omega}.$$

Also, the definition of $I(x)$ yields $L(X_\alpha^x(t), \alpha(t)) \geq I(x)$ for all $x \in \overline{\Omega}$. Consequently,

$$\lambda V_\lambda(x) \geq \lambda \int_0^\infty e^{-\lambda t} I(x)\, dt = I(x) \qquad \text{for } x \in \overline{\Omega}. \tag{4.14}$$

Step 2. Fix $x \in \overline{\Omega}$. For any $\varepsilon > 0$ take $\alpha_\varepsilon \in \mathcal{A}$, $T_\varepsilon > 0$ and $a_\varepsilon \in A$ such that (4.13) holds. Moreover, define the control $\tilde{\alpha}$ as

$$\tilde{\alpha}(t) := \begin{cases} \alpha_\varepsilon(t), & t \in [0, T_\varepsilon], \\ a_\varepsilon, & t > T_\varepsilon. \end{cases}$$

As $\tilde{\alpha}$ might not be the optimal control, we obtain

$$\lambda V_\lambda(x) \leq \lambda \int_0^{T_\varepsilon} e^{-\lambda t} L(X^x_{\alpha_\varepsilon}(t), \alpha_\varepsilon(t))\, dt + \lambda \int_{T_\varepsilon}^{\infty} e^{-\lambda t} L(X^x_{\tilde{\alpha}}(t), a_\varepsilon)\, dt. \tag{4.15}$$

Due to $f(X^x_{\alpha_\varepsilon}(T_\varepsilon), a_\varepsilon) = 0$ and the definition of $\tilde{\alpha}$, the solution of the state equation stays constant after T_ε, which means

$$L(X^x_{\tilde{\alpha}}(t), a_\varepsilon) = L(X^x_{\alpha_\varepsilon}(T_\varepsilon), a_\varepsilon) \overset{(4.13)}{\leq} I(x) + \varepsilon, \quad \forall\, t \geq T_\varepsilon. \tag{4.16}$$

As the Lagrangian is bounded from above by 1, see (3.4), combining (4.16) with the above inequality (4.15) yields

$$\begin{aligned} \lambda V_\lambda(x) &\leq \lambda \int_0^{T_\varepsilon} e^{-\lambda t}\, dt + \lambda \int_{T_\varepsilon}^{\infty} e^{-\lambda t}(I(x) + \varepsilon)\, dt. \\ &= (1 - e^{-\lambda T_\varepsilon}) + (I(x) + \varepsilon) e^{-\lambda T_\varepsilon}. \end{aligned}$$

Passing the limit $\lambda \to 0^+$ yields

$$\limsup_{\lambda \to 0^+} \lambda V_\lambda(x) \leq I(x) + \varepsilon.$$

By the arbitrariness of $\varepsilon > 0$ we conclude the proof. □

Our next goal is to establish Proposition 4.17 as a dual counterpart to Proposition 4.8.

Definition 4.16 (Joint Hamiltonian)**.** For any $x \in \overline{\Omega}$, $u \in \mathbb{R}$, $p, q \in \mathbb{R}^n$ let us define the *joint Hamiltonian* J as

$$J(x, u, p, q) := \max_{a \in A} \min\{\langle f(x, a), p\rangle,\ \langle f(x, a), q\rangle - L(x, a) + u\}. \tag{4.17}$$

We say a pair $(v, w) \in C(\overline{\Omega}) \times C(\overline{\Omega})$ is a viscosity supersolution of

$$J(x, v, -\nabla v, -\nabla w) = 0 \quad \text{on } \overline{\Omega}, \tag{4.18}$$

if for all $x \in \overline{\Omega}$ we have that

$$J(x, v(x), -p, -q) \geq 0, \qquad \forall\, p \in \partial^-_{\overline{\Omega}} v(x),\ \forall\, q \in \partial^-_{\overline{\Omega}} w(x).$$

Proposition 4.17. *Let $v, w \in \mathrm{Lip}(\overline{\Omega})$ be such that (v, w) is a viscosity supersolution of (4.18). Then the following holds true for any $\lambda > 0$:*

a) for any $c \in \mathbb{R}$, the function

$$w_c(x) := \frac{v(x)}{\lambda} + w(x) + c \qquad \text{for } x \in \overline{\Omega}, \tag{4.19}$$

satisfies

$$\lambda w_c + H(x, -\nabla w_c) \geq \lambda(w + c) \quad \text{on } \overline{\Omega} \tag{4.20}$$

in the viscosity sense;

b) V_λ satisfies the upper bound

$$V_\lambda(x) - \frac{v(x)}{\lambda} \leq w(x) + \|w\|_\infty, \qquad \forall\, x \in \overline{\Omega}. \tag{4.21}$$

Proof. We want to apply Proposition 3.10 a) to the functions v and w. For w we take $\ell := L - v$ and $\tilde{\lambda} := 0$, whereas for v we use $\ell := 0$ and $\tilde{\lambda} := 0$. Due to $v \in \mathrm{Lip}(\overline{\Omega})$, analogously to the proof of Proposition 4.8, condition (3.9) is satisfied for this choice of functions. To obtain a single control $\alpha \in \mathcal{A}$ for which the corresponding dynamic inequalities for v and w hold simultaneously, one formally applies the viability arguments of Proposition 3.10 and Lemma 3.11 to a suitable extended state space. We omit the details here.

Since (v, w) is a viscosity supersolution of (4.18), for all $x \in \overline{\Omega}$,

$$\max_{a\in A} \min\{\langle f(x,a), -p\rangle, \langle f(x,a), -q\rangle - L(x,a) + v(x)\} \geq 0, \quad \forall\, p \in \partial^-_{\overline{\Omega}} v(x), \forall\, q \in \partial^-_{\overline{\Omega}} w(x).$$

This condition implies that for any $x \in \overline{\Omega}$, any $p \in \partial^-_{\overline{\Omega}} v(x)$ and any $q \in \partial^-_{\overline{\Omega}} w(x)$, there exists a common $a \in A$ satisfying both arguments of the minimum are nonnegative. Therefore, by the viability arguments mentioned above, applied in combination with Proposition 3.10 a), for every $x \in \overline{\Omega}$ there exists $\alpha \in \mathcal{A}$ such that

$$\begin{cases} w(x) \geq w(X^x_\alpha(t)) + \displaystyle\int_0^t L(X^x_\alpha(s), \alpha(s)) - v(X^x_\alpha(s))\, ds, \\ v(x) \geq v(X^x_\alpha(t)), \end{cases} \qquad \forall\, t \geq 0.$$

Hence, combining the two inequalities yields

$$w_c(x) \geq w_c(X^x_\alpha(t)) + \int_0^t L(X^x_\alpha(s), \alpha(s)) - v(X^x_\alpha(s))\, ds, \quad \forall\, t \geq 0.$$

By using Proposition 3.10 a) with $\theta := w_c$, $\ell := L - v$ and $\tilde{\lambda} := 0$ we obtain

$$\widetilde{H}(x, -\nabla w_c) \geq 0 \quad \text{on } \overline{\Omega} \tag{4.22}$$

in the viscosity sense, where $\widetilde{H}$ is defined in (3.8). Rearranging (4.22) yields

$$\begin{aligned} \widetilde{H}(x, -\nabla w_c) \geq 0 \quad &\Longleftrightarrow \quad \sup_{a\in A}\{\langle f(x,a), -\nabla w_c\rangle - L(x,a) + v(x)\} \geq 0 \\ &\Longleftrightarrow \quad v(x) + H(x, -\nabla w_c) \geq 0 \quad \text{on } \overline{\Omega} \end{aligned} \tag{4.23}$$

in the viscosity sense. By using the identity $v(x) = \lambda(w_c(x) - w(x) - c)$ we obtain (4.20).

Defining $c := \|w\|_\infty$ on (4.20), we satisfy

$$\lambda w_c + H(x, -\nabla w_c) \geq \lambda(w + \|w\|_\infty) \geq 0 \quad \text{on } \overline{\Omega}$$

in the viscosity sense, which shows that w_c is a viscosity supersolution of (3.6). By the comparison principle, Theorem 3.13, we obtain $w_c \geq V_\lambda$ on $\overline{\Omega}$. Inserting the definition of w_c from (4.19) and rearranging this inequality yields (4.21). □

4.3 Hamilton–Jacobi system

This section is devoted to the system of Hamilton–Jacobi equations

$$\begin{cases} h(x,-\nabla u(x)) = 0 \\ u(x) + H(x,-\nabla w(x)) = 0 \\ J(x,u(x),-\nabla u(x),-\nabla w(x)) = 0 \end{cases} \quad \text{on } \overline{\Omega}, \tag{S}$$

where h is the reduced Hamiltonian in (4.1) and J the joint Hamiltonian in (4.17). We consider the pair $(u,w) \in C(\overline{\Omega}) \times C(\overline{\Omega})$ to be a solution of (S) in the viscosity sense. This system plays a central role in establishing the rate of convergence of the discounted value function λV_λ to its limit v^*, presented in Theorem 4.23.

Remark 4.18. Note that if (u,w) is a viscosity solution of system (S), it follows that the swapped pair (w,u) belongs to the set $\mathcal{S}(\overline{\Omega})$.

Remark 4.19. If the dynamics are independent of the control variable, meaning that $X' = f(X(t))$ in (3.2), any pair (u,w) satisfying the first two equations in system (S) in the viscosity sense necessarily satisfies the third equation as well.

Indeed, the independence of f from a implies $\langle f(x), -\nabla u(x)\rangle = 0$ for all $x \in \overline{\Omega}$ in the first equation. The second equation $u(x) + H(x,-\nabla w(x)) = 0$ becomes

$$\max_{a\in A}\{\langle f(x), -\nabla w(x)\rangle + u(x) - L(x,a)\} = 0 \qquad \text{on } \overline{\Omega}.$$

Consequently, for all $a \in A$, the term inside the maximum is non-positive. Given the independence of the dynamics from the control variable, the third equation takes the form

$$\max_{a\in A}\min\{\langle f(x), -\nabla u(x)\rangle, \langle f(x), -\nabla w(x)\rangle + u(x) - L(x,a)\} = 0 \qquad \text{on } \overline{\Omega}.$$

Since the first term inside the minimum is exactly 0 and the second term is non-positive, the minimum simply equals the second term. Thus, the entire expression reduces to the left-hand side of the second equation, which evaluates to 0, confirming that the third equation is automatically satisfied.

We start by providing a uniqueness property of solutions of system (S) in Proposition 4.21.

Lemma 4.20. *Let (u,w) be a solution of (S) with $u \in \mathrm{Lip}(\overline{\Omega})$. Then for all $x \in \overline{\Omega}$ and $\alpha \in \mathcal{A}$*

$$\begin{cases} a) \quad u(x) \le u(X^x_\alpha(t)), \\ b) \quad w(x) \le \int_0^t L\big(X^x_\alpha(s),\alpha(s)) - u(X^x_\alpha(s)\big)\,ds + w(X^x_\alpha(t)), \end{cases} \qquad \forall\, t \ge 0. \tag{4.24}$$

Proof. a) Since (u,w) is a solution of (S), we have

$$h(x,-\nabla u(x)) = 0 \quad \text{on } \overline{\Omega}$$

in the viscosity sense. Hence, u is a viscosity subsolution of (3.8) with $\tilde{\lambda} := 0$ and $\ell := 0$. As condition (3.9) holds, by Proposition 3.10 b) the first inequality in (4.24) is satisfied.

b) Again, since (u,w) is a solution of (S), we have

$$u(x) + H(x,-\nabla w(x)) = 0 \quad \text{on } \overline{\Omega}$$

in the viscosity sense. Comparing with (3.8), we see that for the choices $\tilde{\lambda} := 0$ and $\ell := L - u$,

$$\begin{aligned} 0 = \widetilde{H}(x, -\nabla w(x)) &:= \max_{a \in A}\{\langle f(x,a), -\nabla w(x)\rangle - L(x,a) + u(x)\} \\ &= u(x) + H(x, -\nabla w(x)) \quad \text{on } \overline{\Omega}, \end{aligned}$$

holds in the viscosity sense. Hence, w is a viscosity solution and in particular a viscosity subsolution of (3.8). As $u \in \mathrm{Lip}(\overline{\Omega})$, the condition (3.9) is satisfied, hence by Proposition 3.10 b), the second inequality in (4.24) is satisfied. □

Proposition 4.21. *Let (u_1, w_1) and (u_2, w_2) with $u_1, u_2 \in \mathrm{Lip}(\overline{\Omega})$ be solutions of (S). Then $u_1 = u_2$ on $\overline{\Omega}$.*

Proof. Fix an arbitrary $x \in \overline{\Omega}$. Since (u_1, w_1) is a solution of (S), it is in particular a viscosity supersolution to the joint Hamiltonian equation $J(x, u_1, -\nabla u_1, -\nabla w_1) = 0$. Using the same viability arguments as in the proof of Proposition 4.17, there exists a control $\alpha_x \in \mathcal{A}$ such that for all $t \geq 0$,

$$\begin{cases} w_1(x) \geq w_1(X^x_{\alpha_x}(t)) + \displaystyle\int_0^t L(X^x_{\alpha_x}(s), \alpha_x(s)) - u_1(X^x_{\alpha_x}(s))\, ds, \\ u_1(x) \geq u_1(X^x_{\alpha_x}(t)). \end{cases}$$

On the other hand, Lemma 4.20 applied to (u_1, w_1) with this specific control α_x gives the reverse inequalities. Consequently, the equalities

$$\begin{cases} w_1(x) = w_1(X^x_{\alpha_x}(t)) + \displaystyle\int_0^t L(X^x_{\alpha_x}(s), \alpha_x(s)) - u_1(X^x_{\alpha_x}(s))\, ds, \\ u_1(x) = u_1(X^x_{\alpha_x}(t)) \end{cases} \tag{4.25}$$

are satisfied for all $t \geq 0$.

Now consider the second solution (u_2, w_2). By Lemma 4.20 b), using the same control α_x, we have

$$w_2(x) \leq \int_0^t L(X^x_{\alpha_x}(s), \alpha_x(s)) - u_2(X^x_{\alpha_x}(s))\, ds + w_2(X^x_{\alpha_x}(t)) \tag{4.26}$$

for all $t \geq 0$. Combining the equality for w_1, in the first line of (4.25), with the inequality (4.26), yields

$$\begin{aligned} w_2(x) - w_1(x) \leq &\int_0^t L(X^x_{\alpha_x}(s), \alpha_x(s)) - u_2(X^x_{\alpha_x}(s))\, ds + w_2(X^x_{\alpha_x}(t)) \\ &- \int_0^t L(X^x_{\alpha_x}(s), \alpha_x(s)) - u_1(X^x_{\alpha_x}(s))\, ds - w_1(X^x_{\alpha_x}(t)) \\ = &\int_0^t u_1(X^x_{\alpha_x}(s)) - u_2(X^x_{\alpha_x}(s))\, ds + w_2(X^x_{\alpha_x}(t)) - w_1(X^x_{\alpha_x}(t)) \end{aligned} \tag{4.27}$$

for all $t \geq 0$. Using Lemma 4.20 a) with (u_2, w_2) and the control α_x we have $u_2(X^x_{\alpha_x}(s)) \geq u_2(x)$, which implies $-u_2(X^x_{\alpha_x}(s)) \leq -u_2(x)$ for all $s \geq 0$. Adding this inequality to the second line of (4.25) yields

$$u_1(X^x_{\alpha_x}(s)) - u_2(X^x_{\alpha_x}(s)) \leq u_1(x) - u_2(x), \qquad \forall\, s \geq 0. \tag{4.28}$$

Furthermore, using (4.28) to estimate the integral in (4.27) leads to

$$w_2(x) - w_1(x) \le t[u_1(x) - u_2(x)] + w_2(X^x_{\alpha_x}(t)) - w_1(X^x_{\alpha_x}(t)), \qquad \forall\, t \ge 0.$$

Since w_1 and w_2 are bounded on $\overline{\Omega}$, we may consider $t > 0$, divide by t, and pass to the limit $t \to \infty$. This yields

$$u_1(x) - u_2(x) \ge 0 \quad \Longrightarrow \quad u_1(x) \ge u_2(x).$$

Exchanging the roles of (u_1, w_1) and (u_2, w_2) symmetrically yields $u_1(x) \le u_2(x)$. Therefore, $u_1 = u_2$ on $\overline{\Omega}$. □

Next, we want to introduce the function w_λ as a tool for quantifying the rate of convergence of the discounted value function λV_λ to its limit v^*.

Definition 4.22. Let V_λ be the value function and let v^* be as in (4.10). Then the function $w_\lambda : \overline{\Omega} \to \mathbb{R}$ is defined as

$$w_\lambda(x) := \frac{\lambda V_\lambda(x) - v^*(x)}{\lambda} \quad \text{for } x \in \overline{\Omega}.$$

Theorem 4.23. *Let (u, w) be a solution of (S) with $u, w \in \mathrm{Lip}(\overline{\Omega})$. Then for any $\lambda > 0$,*

$$w(x) - \|w\|_\infty \le w_\lambda(x) \le w(x) + \|w\|_\infty, \qquad \forall\, x \in \overline{\Omega}. \tag{4.29}$$

Proof. Since (u, w) is a solution of (S) and u, w are both Lipschitz continuous, by Remark 4.18, we have $(w, u) \in \mathcal{S}_{\mathrm{Lip}}(\overline{\Omega})$. Hence the assumption for Proposition 4.8 is satisfied. Furthermore, (u, w) being a solution for (S) implies that (u, w) is a viscosity supersolution of (4.18). Therefore, also the assumptions for Proposition 4.17 are satisfied. Hence,

$$w(x) - \|w\|_\infty \overset{(4.4)}{\le} V_\lambda(x) - \frac{u(x)}{\lambda} \overset{(4.21)}{\le} w(x) + \|w\|_\infty \tag{4.30}$$

is satisfied for all $x \in \overline{\Omega}$. Multiplying by λ yields

$$\lambda(w(x) - \|w\|_\infty) \le \lambda V_\lambda(x) - u(x) \le \lambda(w(x) + \|w\|_\infty), \qquad \forall\, x \in \overline{\Omega}$$

and so, due to the boundedness of w,

$$|\lambda V_\lambda(x) - u(x)| \le 2\lambda \|w\|_\infty, \qquad \forall\, x \in \overline{\Omega}.$$

This implies that the uniform limit of λV_λ as $\lambda \to 0^+$ is u, hence $v^* = u$ on $\overline{\Omega}$. Finally, substituting $v^* = u$ into to definition of w_λ and using (4.30), we conclude

$$w(x) - \|w\|_\infty \le w_\lambda(x) \le w(x) + \|w\|_\infty, \qquad \forall\, x \in \overline{\Omega}.$$

□

Remark 4.24. Theorem 4.23 links the Hamilton–Jacobi system (S) directly to the asymptotic behavior of the discounted value function. In particular, any Lipschitz continuous solution (u, w) of (S) uniquely determines the vanishing discount limit as $v^* = u$.

On the other hand, the next example shows that system (S) may fail to have a solution even when λV_λ converges uniformly as $\lambda \to 0^+$.

Example 4.25. We return to the setting of Example 4.13 where we have obtained

$$v^*(x) = 1 \quad \text{and} \quad V_\lambda(x) = \frac{1}{\lambda} - \frac{x}{2\sqrt{\lambda}} \quad \text{for } x \in [0,1].$$

Thus the family

$$w_\lambda(x) = \frac{\lambda V_\lambda(x) - v^*(x)}{\lambda} = -\frac{x}{2\sqrt{\lambda}} \quad \text{for } x \in [0,1],$$

fails to be bounded uniformly in λ. Hence, Theorem 4.23 implies that system (S) has no solution. ❖

In this next example we want to deduce the convergence of V_λ and especially the speed of convergence in the form of (4.29).

Example 4.26. Consider the uncontrolled two dimensional linear system

$$\begin{pmatrix} X'(t) \\ Y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}, \; t > 0 \quad \text{and} \quad (X(0), Y(0)) = (x_0, y_0).$$

Consider also a Lipschitz continuous function $L : \mathbb{R}^2 \to [0,1]$ which does not depend on controls. Hence,

$$V_\lambda(x_0, y_0) := \int_0^\infty e^{-\lambda t} L(X^{x_0,y_0}(t), Y^{x_0,y_0}(t))\, dt.$$

Observe that any ball $\overline{\Omega} = \overline{B(0,R)}$ is invariant under the above differential equation. Define two functions $u, w : \mathbb{R}^2 \to \mathbb{R}$ in polar coordinates $(x,y) = (r\cos(\theta), r\sin(\theta))$ as

$$u(x,y) := \frac{1}{2\pi} \int_0^{2\pi} L(r\cos(\sigma), r\sin(\sigma))\, d\sigma \quad \text{and}$$

$$w(x,y) := \int_0^\theta L(r\cos(\sigma), r\sin(\sigma)) - u(r\cos(\sigma), r\sin(\sigma)\, d\sigma.$$

We now show that (u, w) satisfies system (S). Observe that

$$f(x,y) = (y, -x) = (r\sin(\theta), -r\cos(\theta)) = -re_\theta,$$

where $e_\theta = (-\sin(\theta), \cos(\theta))$ denotes the angular unit vector.

i) The polar gradient of the radial function u is $\nabla u = u'(r)e_r$, where $e_r = (\cos(\theta), \sin(\theta))$ denotes the radial unit vector. Hence

$$h(x, -\nabla u) = \langle f, -\nabla u\rangle = -ru'(r)\overbrace{\langle e_\theta, e_r\rangle}^{=0} = 0,$$

which satisfies the first equation of (S).

ii) The polar gradient of w is $\nabla w = w_r e_r + 1/r\,(w_\theta e_\theta)$. Calculate

$$\langle f, -\nabla w\rangle = \langle -re_\theta, -w_r e_r - \frac{1}{r} w_\theta e_\theta\rangle = rw_r\langle e_\theta, e_r\rangle + w_\theta \overbrace{\langle e_\theta, e_\theta\rangle}^{=1} = w_\theta.$$

Rearranging the second equation of (S) yields

$$u + H(x, -\nabla w) = 0 \iff u + \big(\langle f, -\nabla w\rangle - L\big) = 0 \iff w_\theta = L - u.$$

By the definition of w, the equation is satisfied.

As mentioned in Remark 4.19, the third equation of (S) is automatically satisfied. By Theorem 4.23, we have that λV_λ converges uniformly to u on $\overline{B(0,R)}$ as $\lambda \to 0^+$. Moreover, the theorem provides the bounds

$$|\lambda V_\lambda(x,y) - u(x,y)| \leq \lambda |w(x,y) + \|w\|_\infty| \leq 2\lambda \|w\|_\infty, \quad \forall\,(x,y) \in \overline{B(0,R)}^2, \forall\,\lambda > 0.$$

Since $\|L\|_\infty \leq 1$, hence $\|u\|_\infty \leq 1$ and so $\|w\|_\infty \leq 4\pi$. Thus,

$$|\lambda V_\lambda(x,y) - u(x,y)| \leq 8\lambda\pi, \quad \forall\,(x,y) \in \overline{B(0,R)}^2, \forall\,\lambda > 0.$$

❖

Remark 4.27. In this chapter, we focused on the analysis of Abel means (λV_λ as $\lambda \to 0^+$). An alternative method for studying long-run average costs is the analysis of Cesàro means ($\frac{1}{T} V_T$ as $T \to \infty$). In (Buckdahn et al., 2015), the authors investigate the relationship between these two approaches. Their main result establishes that in the topology of uniform convergence, both families share the same cluster points, provided they exist.

5 Analysis of Vanishing Discount Limits in Discrete Time

We now turn to the discrete-time counterpart of the vanishing discount analysis developed in the previous chapters. This setting naturally encompasses Markov decision processes, stochastic games, and more general dynamic programming operators.

In the classical finite state and finite action framework, the long-run mean payoff has been characterized using two main approaches. The first relies on the notion of invariant half-lines of dynamic programming operators, introduced by Kohlberg (1980) for polyhedral operators and further studied in the one-player case. The second approach is based on multichain linear programming formulations introduced in (Denardo et al., 1968), which provide an alternative characterization of the average cost.

The purpose of this chapter is to show that the structural results obtained earlier extend beyond the finite-dimensional setting. In particular, we develop an abstract framework for discrete-time control problems that allows for infinite state and action spaces. Within this framework, sub-invariant half-lines emerge as the natural analogue of subsolutions in the continuous-time Hamilton–Jacobi theory, leading to a characterization of vanishing discount limits that parallels the continuous-time case.

5.1 Mean payoff and sub-invariant half-lines of Shapley operators

We begin by introducing the abstract framework for the discrete-time problem. We recall the following definition, which will be used throughout the discrete-time analysis. For more details, we refer to the book by Schaefer (1974).

Definition 5.1 (AM-space with unit)**.** A Banach lattice $(X, \|\cdot\|, \leq)$ in which there is an element $e > 0$, the *unit*, such that $\|x\| = \inf\{t > 0 : -te \leq x \leq te\}$ is called *abstract M-space with unit* or briefly *AM-space with unit*.

A typical example of an AM-space with unit is $C(S)$, the space of real-valued continuous functions on a compact Hausdorff topological space S, equipped with the sup-norm and the pointwise partial order. The constant function $f = 1 \in C(S)$ serves as the unit e.

Moreover, by the Kakutani-Krein theorem (Schaefer, 1974, Chapter 2, Theorem 7.4), every AM-space with unit is lattice isometrically isomorphic to $C(K)$ for a suitable compact Hausdorff space K. That is, for any AM-space with unit there exists a compact space K and a linear bijection to $C(K)$ which preserves the norm and the lattice operations.

Therefore, throughout the remainder of this section, we may identify the AM-space with unit with

$$X := C(S),$$

where the explicit knowledge of the underlying compact space S is not required.

In discrete-time control problems, the evolution of value functions is described by operators acting on X. We now introduce the properties of the operators relevant for our analysis, which lead to the definition of a Shapley operator.

Definition 5.2. Let $T : X \to X$. Denote by $\leq$ the pointwise partial order on X and by e the unit of X. We say that:

a) T is *order preserving*, if for all $x, y \in X : x \leq y \implies T(x) \leq T(y)$.

b) T *commutes with the addition of the unit*, if for all $x \in X$ and $\lambda \in \mathbb{R} : T(\lambda e + x) = \lambda e + T(x)$.

Proposition 5.3. *If $T : X \to X$ is order preserving and commutes with the addition of the unit, then T is nonexpansive, that is,*

$$\|T(x) - T(y)\| \leq \|x - y\|, \quad \forall\, x, y \in X.$$

Proof. Let $x, y \in X$ be arbitrary. Set $\delta := \|x - y\|$. By the definition of the norm in an AM-space with unit, we have

$$-\delta e \leq x - y \leq \delta e.$$

Rearranging the terms and using the fact that T is order preserving yields

$$T(y - \delta e) \leq T(x) \leq T(y + \delta e).$$

Since T commutes with the addition of the unit, we obtain

$$T(y) - \delta e \leq T(x) \leq T(y) + \delta e,$$

and hence

$$-\delta e \leq T(x) - T(y) \leq \delta e.$$

Applying again the definition of the norm implies

$$\|T(x) - T(y)\| \leq \delta = \|x - y\|.$$

□

Definition 5.4 (Shapley operator)**.** A map $T : X \to X$ is called a *Shapley operator* if it is order preserving and commutes with the addition of the unit.

In Example 5.5 we present the original operator introduced by Shapley (1953).

Example 5.5. We consider the finite-state case $S = \{1, \dots, n\} =: [n]$ with two players, denoted by A and B. We identify $C(S)$ with $\mathbb{R}^n$, endowed with the coordinatewise order, the supremum norm, and unit $e = (1, \dots, 1)$, so that $\mathbb{R}^n$ is an AM-space with unit.

For each state $i \in S$ let A_i and B_i denote the finite sets of actions available to Player 1 and Player 2, respectively, and let $g_i : A_i \times B_i \to \mathbb{R}$ be the corresponding payoff function. Furthermore, let $P_i : A_i \times B_i \to \Delta([n])$, where for any finite set K, $\Delta(K)$ denotes the set of probability measures on K, which can be identified with nonnegative vectors $p = (p_k)_{k \in K} \in \mathbb{R}^K$ such that $\sum_{k \in K} p_k = 1$. The quantity $(P_i(a, b))_j$ represents the transition probability from the current state i to the next state j, given the actions $a \in A_i$ and $b \in B_i$.

For a given value vector $x \in \mathbb{R}^n$, the quantity $T_i(x)$ denotes the i-th component of the vector $T(x)$ and represents the optimal expected payoff when the current state is i. It is defined by

$$T_i(x) = \min_{\mu \in \Delta(A_i)} \max_{\nu \in \Delta(B_i)} \sum_{a \in A_i} \sum_{b \in B_i} \Big(g_i(a, b) + \sum_{j \in [n]} (P_i(a, b))_j \, x_j \Big) \mu_a \nu_b.$$

This operator fits the above definition of a Shapley operator. Indeed, since $P_i(a,b)$ is a probability distribution on $[n]$, we have

$$\sum_{j\in[n]} (P_i(a,b))_j\,(x_j+\lambda) = \sum_{j\in[n]} (P_i(a,b))_j\,x_j + \lambda,$$

which shows that T commutes with the addition of the unit. Order preservation holds since increasing x can only increase the value $T_i(x)$. ❖

For a Shapley operator T we consider the discrete evolution equation

$$v^k = T(v^{k-1}), \qquad k = 1, 2, \dots$$

for $v^k \in X$ and initial condition $v^0 \in X$. We denote by T^k the k-fold iterate of T, so that $v^k = T^k(v^0)$. This evolution equation may be regarded as a discrete-time analogue of the Hamilton–Jacobi–Bellman equation

$$V_t + H(x, -\nabla V) = 0.$$

In analogy with the discounted Hamilton–Jacobi–Bellman equation

$$\lambda V_\lambda + H(x, -\nabla V_\lambda) = 0, \quad \lambda > 0,$$

we now introduce the discrete-time analogue of the discounted value function V_λ.

Definition 5.6 (Discounted value function)**.** Let $T : X \to X$ be nonexpansive. For $\alpha \in (0,1)$, the *discounted value function* $v_\alpha \in X$ is defined as the unique fixed point satisfying

$$T((1-\alpha)v_\alpha) = v_\alpha. \tag{5.1}$$

The existence and uniqueness of v_α follow from the fact that the map $u \mapsto T((1-\alpha)u)$ is a contraction on X. We shall sometimes write $v_\alpha = v_\alpha(T)$ to emphasize its dependence on the operator. The correspondence between the continuous- and discrete-time discount parameters is obtained by setting

$$1-\alpha = e^{-\lambda},$$

so that the vanishing discount limit $\lambda \to 0^+$ in continuous time corresponds to $\alpha \to 0^+$ in discrete time.

Remark 5.7. The equation

$$T((1-\alpha)v_\alpha) = v_\alpha, \qquad \alpha \in (0,1),$$

may be interpreted as a discrete-time discounted dynamic programming principle, in which future costs are discounted by the factor $1-\alpha$. The rescaled family $\{\alpha v_\alpha\}_{\alpha\in(0,1)}$ therefore plays the role of a long-run average cost in the vanishing discount limit, in direct analogy with the family $\{\lambda V_\lambda\}_{\lambda>0}$ in continuous time.

We now introduce the concept of invariant half-lines first proposed by Kohlberg (1980). In his work he used an operator $T : \mathbb{R}^n \to \mathbb{R}^n$, which was only required to be nonexpansive. In the present setting, the operator $T : X \to X$ is not only nonexpansive but also order preserving. Therefore, it is convenient to consider a one-sided variant of Kohlberg's notion.

Definition 5.8 (Sub-, super-, and invariant half-lines)**.** Suppose $T : X \to X$. A *sub-invariant half-line* is a map of the form

$$s \mapsto u + s\eta, \qquad s \geq 0,$$

with $u, \eta \in X$, such that

$$T(u + s\eta) \geq u + (s+1)\eta, \qquad \forall\, s \geq 0.$$

We say that the vector η *directs* the half-line. A *super-invariant half-line* is defined by reversing the preceding inequality. An *invariant half-line* is both sub- and super-invariant.

Remark 5.9. If T is order preserving with a sub-invariant half-line $s \mapsto u + s\eta$, then, by induction and order preservation, one obtains

$$T^k(u) \geq u + k\eta, \qquad \forall\, k \geq 0. \tag{5.2}$$

If the half-line is invariant, then (5.2) holds with equality.

The following proposition establishes the existence of a sub-invariant half-line for any Shapley operator.

Proposition 5.10. *If $T : X \to X$ is a Shapley operator, then it has a sub-invariant half-line. More precisely, the half-line $s \mapsto -s\|T(0)\|e$ is sub-invariant.*

Proof. We need to show that for all $s \geq 0$,

$$T(-s\|T(0)\|e) \geq -(s+1)\|T(0)\|e,$$

holds. Since T commutes with the addition of the unit, we have

$$T(-s\|T(0)\|e) = T(0 + (-s\|T(0)\|)e) = T(0) - s\|T(0)\|e.$$

Hence, the desired inequality is equivalent to

$$T(0) - s\|T(0)\|e \geq -(s+1)\|T(0)\|e \quad \Longleftrightarrow \quad T(0) \geq -\|T(0)\|e.$$

By the definition of the AM-space, we have

$$-\|T(0)\|e \leq T(0) \leq \|T(0)\|e,$$

and in particular $T(0) \geq -\|T(0)\|e$, which proves the claim. □

The statement of Proposition 5.11 can be compared with Corollary 2.2 in (Kohlberg, 1980). The result shows that the existence of an invariant half-line yields not only the existence of the limit $\lim_{k\to\infty} T^k(x)/k$, which coincides with the directing vector η, but also the property that the deviation $T^k(x) - k\eta$ remains bounded as $k \to \infty$. An analogous phenomenon holds for the vanishing discount limit.

Note that Proposition 5.11 is not proven explicitly in the work by Cannarsa et al. (2024).

Proposition 5.11. *Suppose that $T : X \to X$ is nonexpansive and admits an invariant half-line $s \mapsto u + s\eta$. Then the following assertions hold.*

a) The sequence $(T^k(x) - k\eta)_{k\geq 0}$ remains bounded for every $x \in X$.

b) For every $x \in X$, the limit

$$\lim_{k\to\infty} \frac{T^k(x)}{k}$$

exists and coincides with the directing vector η.

c) The family $(v_\alpha - \alpha^{-1}\eta)_{\alpha\in(0,1)}$ remains bounded as $\alpha \to 0^+$, and in particular

$$\lim_{\alpha\to 0^+} \alpha v_\alpha = \eta.$$

Proof. a) Since $s \mapsto u + s\eta$ is an invariant half-line, as discussed in Remark 5.9, we have

$$T^k(u) = u + k\eta, \qquad \forall\, k \geq 0.$$

By nonexpansiveness of T,

$$\|T^k(x) - (u + k\eta)\| = \|T^k(x) - T^k(u)\| \leq \|x - u\|, \qquad \forall\, k \geq 0. \tag{5.3}$$

Using $T^k(x) - k\eta = \big(T^k(x) - (u + k\eta)\big) + u$ and the triangle inequality, we get

$$\|T^k(x) - k\eta\| \leq \|T^k(x) - (u + k\eta)\| + \|u\| \overset{(5.3)}{\leq} \|x - u\| + \|u\|, \qquad \forall\, k \geq 0.$$

Therefore $\sup_{k\geq 0} \|T^k(x) - k\eta\| \leq \|x - u\| + \|u\| < \infty$, and the sequence $(T^k(x) - k\eta)_{k\geq 0}$ is bounded.

b) We write

$$\frac{T^k(x)}{k} = \frac{T^k(x) - k\eta}{k} + \eta.$$

By point a), there exists $M > 0$ such that $\|T^k(x) - k\eta\| \leq M$ for all $k \geq 0$. Hence,

$$\left\| \frac{T^k(x) - k\eta}{k} \right\| \leq \frac{M}{k} \xrightarrow[k\to\infty]{} 0$$

and therefore

$$\lim_{k\to\infty} \frac{T^k(x)}{k} = \eta.$$

c) Define

$$y_\alpha := u + \alpha^{-1}\eta, \qquad \alpha \in (0, 1).$$

Since $s \mapsto u + s\eta$ is an invariant half-line of T, we have

$$T(u + s\eta) = u + (s + 1)\eta, \qquad \forall\, s \geq 0.$$

Taking $s = \alpha^{-1} - 1$, we obtain

$$T(u + (\alpha^{-1} - 1)\eta) = u + \alpha^{-1}\eta = y_\alpha \tag{5.4}$$

Moreover,

$$(1 - \alpha)y_\alpha = (1 - \alpha)u + (\alpha^{-1} - 1)\eta. \tag{5.5}$$

Combining (5.4) and (5.5), we get

$$\|T((1 - \alpha)y_\alpha) - y_\alpha\| = \|T((1 - \alpha)u + (\alpha^{-1} - 1)\eta) - T(u + (\alpha^{-1} - 1)\eta)\|.$$

By nonexpansiveness of T,

$$\begin{aligned}\|T((1-\alpha)y_\alpha)-y_\alpha\| &\le \|(1-\alpha)u+(\alpha^{-1}-1)\eta-\big(u+(\alpha^{-1}-1)\eta\big)\|\\ &= \alpha\|u\|.\end{aligned}$$

Since v_α satisfies $v_\alpha = T((1-\alpha)v_\alpha)$, we may write

$$\begin{aligned}\|v_\alpha - y_\alpha\| &= \|T((1-\alpha)v_\alpha)-y_\alpha\|\\ &\le \|T((1-\alpha)v_\alpha)-T((1-\alpha)y_\alpha)\| + \|T((1-\alpha)y_\alpha)-y_\alpha\|\\ &\le (1-\alpha)\|(v_\alpha-y_\alpha)\| + \alpha\|u\|.\end{aligned}$$

Rearranging yields

$$\alpha\|v_\alpha - y_\alpha\| \le \alpha\|u\| \quad\implies\quad \|v_\alpha-y_\alpha\|\le\|u\|.$$

Finally,

$$\|v_\alpha-\alpha^{-1}\eta\| = \|v_\alpha-\alpha^{-1}\eta-u+u\| \le \|v_\alpha-y_\alpha\|+\|u\|\le 2\|u\|.$$

This shows that the family $(v_\alpha-\alpha^{-1}\eta)_{\alpha\in(0,1)}$ is bounded as $\alpha\to 0^+$. Moreover,

$$\alpha v_\alpha = \alpha(v_\alpha-\alpha^{-1}\eta)+\eta \xrightarrow[\alpha\to 0^+]{} \eta,$$

which concludes the proof. □

Theorem 5.14 below provides a one-sided extension of Kohlberg's result based on sub-invariant half-lines under the assumption that T is order preserving. Corollary 5.15 shows that if both sub- and super-invariant half-lines with the same direction exist, then statements a), b) and c) from Proposition 5.11 hold also true.

To simplify the proof of Theorem 5.14 it will be convenient to define for all $u\in X$ the conjugate map

$$T_u(x) := -u + T(u+x).$$

We will use the following properties of T_u.

Lemma 5.12. *Let $T : X\to X$ be nonexpansive and let $u\in X$. Then T_u is nonexpansive.*

Proof. Since T is nonexpansive, for all $x,y\in X$ we have $\|T(x)-T(y)\|\le\|x-y\|$. Therefore, for any $x,y\in X$,

$$\begin{aligned}\|T_u(x)-T_u(y)\| &= \|(-u+T(u+x))-(-u+T(u+y))\|\\ &= \|T(u+x)-T(u+y)\|\\ &\le \|(u+x)-(u+y)\|\\ &= \|x-y\|.\end{aligned}$$

Hence, T_u is nonexpansive. □

Lemma 5.13. *Let $T : X\to X$ be nonexpansive and let $u\in X$.*

a) If $s\mapsto u+s\eta$ is a sub-invariant half-line of T, then $s\mapsto s\eta$ is a sub-invariant half-line of T_u.

b) For all $x\in X$ and $k\ge 0$: $T_u^k(x) = -u+T^k(u+x)$.

c) For all $\alpha\in(0,1)$: $\|v_\alpha(T)-v_\alpha(T_u)\|\le 2\|u\|$.

Proof. a) Since $s \mapsto u + s\eta$ is a sub-invariant half-line for T, we have

$$T(u + s\eta) \geq u + (s+1)\eta, \qquad \forall\, s \geq 0.$$

Subtracting u from both sides and using the definition of T_u yields

$$T_u(s\eta) = -u + T(u + s\eta) \geq (s+1)\eta,$$

so $s \mapsto s\eta$ is sub-invariant for T_u.

b) We prove the identity by induction on k. For $k = 0$,

$$T_u^0(x) = x = -u + (u + x) = -u + T^0(u + x),$$

holds. Assume for some $k \geq 0$ that $T_u^k(x) = -u + T^k(u + x)$ is satisfied. Then

$$\begin{aligned} T_u^{k+1}(x) &= T_u(T_u^k(x)) = -u + T(u + T_u^k(x)) \\ &= -u + T(u + (-u + T^k(u + x))) = -u + T(T^k(u + x)) \\ &= -u + T^{k+1}(u + x), \end{aligned}$$

which completes the induction.

c) By Lemma 5.12, T_u is nonexpansive, hence $v_\alpha(T_u)$ exists and is unique. By Definition 5.6 we have

$$v_\alpha(T) = T((1-\alpha)v_\alpha(T)) \qquad \text{and} \qquad \begin{aligned} v_\alpha(T_u) &= T_u((1-\alpha)v_\alpha(T_u)) \\ &= -u + T(u + (1-\alpha)v_\alpha(T_u)). \end{aligned}$$

Therefore,

$$v_\alpha(T) - v_\alpha(T_u) - u = T((1-\alpha)v_\alpha(T)) - T(u + (1-\alpha)v_\alpha(T_u)).$$

Using nonexpansiveness of T, we get

$$\|v_\alpha(T) - v_\alpha(T_u) - u\| \leq \big\|(1-\alpha)v_\alpha(T) - \big(u + (1-\alpha)v_\alpha(T_u)\big)\big\|.$$

Rewrite the right-hand side as

$$(1-\alpha)\big(v_\alpha(T) - v_\alpha(T_u) - u\big) - \alpha u,$$

so by the triangle inequality,

$$\|v_\alpha(T) - v_\alpha(T_u) - u\| \leq (1-\alpha)\|v_\alpha(T) - v_\alpha(T_u) - u\| + \alpha\|u\|.$$

Rearranging yields

$$\alpha\|v_\alpha(T) - v_\alpha(T_u) - u\| \leq \alpha\|u\| \quad \Longrightarrow \quad \|v_\alpha(T) - v_\alpha(T_u) - u\| \leq \|u\|.$$

Consequently,

$$\begin{aligned} \|v_\alpha(T) - v_\alpha(T_u)\| &\leq \|v_\alpha(T) - v_\alpha(T_u) - u\| + \|u\| \\ &\leq 2\|u\|, \end{aligned}$$

which completes the proof. □

Note that the limit inferior and limit superior appearing in the next theorem are well defined in the order sense in Banach lattices. In the case $X = C(S)$, they coincide with the usual pointwise definitions for functions. For further details, see for example the book by Aliprantis et al. (2005).

Theorem 5.14 (Comparison principle)**.** *Suppose that $T : X \to X$ is a Shapley operator and that $s \mapsto u + s\eta$ is a sub-invariant half-line of T. Then the following assertions hold.*

a) For every $x \in X$, the sequence

$$\left(T^k(x) - k\eta\right)_{k\geq 0}$$

is bounded from below.

b) The family

$$\left(v_\alpha - \alpha^{-1}\eta\right)_{\alpha\in(0,1)}$$

is bounded from below as $\alpha \to 0^+$.

c) Consequently,

$$\liminf_{\alpha\to 0^+} \alpha v_\alpha \geq \eta \qquad \textit{and} \qquad \liminf_{k\to\infty} \frac{T^k(x)}{k} \geq \eta, \quad \forall\, x \in X.$$

Proof. a) Since $s \mapsto u + s\eta$ is a sub-invariant half-line, we have

$$T(u + s\eta) \geq u + (s+1)\eta, \qquad \forall\, s \geq 0.$$

As discussed in Remark 5.9, order preservation yields

$$T^k(u) \geq u + k\eta, \qquad \forall\, k \geq 0. \tag{5.6}$$

Fix $x \in X$. Since T is a Shapley operator, it is nonexpansive, hence so is T^k. Therefore

$$\|T^k(x) - T^k(u)\| \leq \|x - u\|, \qquad \forall\, k \geq 0.$$

By the definition of the AM-space norm, this implies the inequality

$$-\|x - u\|e \leq T^k(x) - T^k(u) \leq \|x - u\|e,$$

and in particular

$$T^k(x) \geq T^k(u) - \|x - u\|e, \qquad \forall\, k \geq 0. \tag{5.7}$$

Combining the two inequalities (5.6) and (5.7) yields

$$T^k(x) - k\eta \overset{(5.7)}{\geq} T^k(u) - \|x-u\|e - k\eta \overset{(5.6)}{\geq} u - \|x-u\|e, \qquad \forall\, k \geq 0,$$

which shows that $(T^k(x) - k\eta)_{k\geq 0}$ is bounded from below.

b) By Lemma 5.13 a) we know that $s \mapsto s\eta$ is a sub-invariant half-line of T_u. This means

$$T_u(s\eta) \geq (s+1)\eta, \qquad \forall\, s \geq 0. \tag{5.8}$$

Denote by $v_\alpha(T)$ and $v_\alpha(T_u)$ the discounted value functions, as in Definition 5.6, to the respective operator. By Lemma 5.13 c),

$$\|v_\alpha(T) - v_\alpha(T_u)\| \leq 2\|u\|, \qquad \forall\, \alpha \in (0,1).$$

By the definition of the AM-space norm, this implies

$$-2\|u\|e \leq v_\alpha(T) - v_\alpha(T_u), \qquad \forall\, \alpha \in (0,1),$$

hence

$$v_\alpha(T) \geq v_\alpha(T_u) - 2\|u\|e, \qquad \forall\, \alpha \in (0,1). \tag{5.9}$$

Thus, it suffices to prove that $(v_\alpha(T_u) - \alpha^{-1}\eta)_{\alpha\in(0,1)}$ is bounded from below.

Fix $\alpha \in (0,1)$ and set $z := \alpha^{-1}\eta$ and $s := \alpha^{-1} - 1 \geq 0$. Then $(1-\alpha)z = s\eta$, and by sub-invariance of $s \mapsto s\eta$ for T_u,

$$T_u((1-\alpha)z) = T_u(s\eta) \overset{(5.8)}{\geq} (s+1)\eta = \alpha^{-1}\eta = z. \tag{5.10}$$

Define $F_\alpha(x) := T_u((1-\alpha)x)$. Since T_u is nonexpansive (Lemma 5.12), F_α is a contraction. Furthermore, since T is order preserving, so is T_u, hence F_α is order preserving. Moreover, $v_\alpha(T_u)$ is its unique fixed point,

$$F_\alpha(v_\alpha(T_u)) = v_\alpha(T_u).$$

From $F_\alpha(z) \geq z$ in (5.10) and order preservation, we obtain $F_\alpha^k(z) \geq z$ for all $k \geq 0$. Since $F_\alpha^k(z) \to v_\alpha(T_u)$ as $k \to \infty$, this yields

$$v_\alpha(T_u) \geq z = \alpha^{-1}\eta.$$

Hence, $v_\alpha(T_u) - \alpha^{-1}\eta \geq 0$ for all $\alpha \in (0,1)$, which means that $(v_\alpha(T_u) - \alpha^{-1}\eta)_{\alpha\in(0,1)}$ is bounded from below. Together with (5.9), this shows that $(v_\alpha(T) - \alpha^{-1}\eta)_{\alpha\in(0,1)}$ is bounded from below as $\alpha \to 0^+$.

c) From point b) we know that there exists $w_0 \in X$ and $\alpha_0 \in (0,1)$ such that

$$v_\alpha - \alpha^{-1}\eta \geq w_0, \qquad \forall\, \alpha \in (0, \alpha_0).$$

Multiplying by $\alpha > 0$ and letting $\alpha \to 0^+$, we obtain

$$\liminf_{\alpha\to 0^+} \alpha v_\alpha \geq \eta,$$

since $\alpha w_0 \to 0$.

From point a) we know that there exists $w_1 \in X$ such that

$$T^k(x) - k\eta \geq w_1 \qquad \forall\, k \geq 1.$$

Dividing by k yields

$$\frac{T^k(x)}{k} - \eta \geq \frac{w_1}{k}.$$

Since $w_1/k \to 0$ as $k \to \infty$, we conclude that

$$\liminf_{k\to\infty} \frac{T^k(x)}{k} \geq \eta,$$

for every $x \in X$. This completes the proof. □

Corollary 5.15. *Suppose that $T : X \to X$ is a Shapley operator and that it admits both a sub-invariant and a super-invariant half-line with the same directing vector η. Then the families*

$$\left(v_\alpha - \alpha^{-1}\eta\right)_{\alpha\in(0,1)} \qquad \textit{and} \qquad \left(T^k(x) - k\eta\right)_{k\geq 0}$$

are bounded for every $x \in X$. In particular,

$$\lim_{\alpha\to 0^+} \alpha v_\alpha = \lim_{k\to\infty} \frac{T^k(x)}{k} = \eta, \qquad \forall\, x \in X.$$

Proof. Let $T : X \to X$ be a Shapley operator. By Theorem 5.14 applied to a sub-invariant half-line with direction η, we have, for all $x \in X$,

$$\liminf_{\alpha\to 0^+} \alpha v_\alpha \geq \eta \qquad \text{and} \qquad \liminf_{k\to\infty} \frac{T^k(x)}{k} \geq \eta.$$

We now obtain the corresponding upper bounds from the super-invariant half-line. Define for $x \in X$ the dual operator

$$\widehat{T}(x) := -T(-x).$$

Since T is order preserving and commutes with the addition of the unit, the same holds for $\widehat{T}$, so $\widehat{T}$ is a Shapley operator. If $s \mapsto \bar{u} + s\eta$ is a super-invariant half-line of T, that is

$$T(\bar{u} + s\eta) \leq \bar{u} + (s+1)\eta, \qquad \forall\, s \geq 0,$$

then $s \mapsto -\bar{u} + s(-\eta)$ is a sub-invariant half-line of $\widehat{T}$. Indeed, for all $s \geq 0$,

$$\widehat{T}(-\bar{u} + s(-\eta)) = -T(\bar{u} + s\eta) \geq -\bar{u} + (s+1)(-\eta).$$

Applying Theorem 5.14 to $\widehat{T}$ and the direction $-\eta$ yields

$$\liminf_{\alpha\to 0^+} \alpha v_\alpha(\widehat{T}) \geq -\eta \qquad \text{and} \qquad \liminf_{k\to\infty} \frac{\widehat{T}^k(x)}{k} \geq -\eta, \quad \forall\, x \in X.$$

Since $\widehat{T}^k(x) = -T^k(-x)$ for all $k \geq 0$, we obtain

$$\limsup_{k\to\infty} \frac{T^k(x)}{k} \leq \eta, \qquad \forall\, x \in X.$$

By noticing that u solves $T((1-\alpha)u) = u$ if and only if $-u$ satisfies $\widehat{T}((1-\alpha)(-u)) = -u$, we have

$$v_\alpha(\widehat{T}) = -v_\alpha(T), \qquad \alpha \in (0,1).$$

Hence

$$\limsup_{\alpha\to 0^+} \alpha v_\alpha(T) \leq \eta.$$

Combining the lower and upper bounds gives

$$\lim_{\alpha\to 0^+} \alpha v_\alpha = \eta \qquad \text{and} \qquad \lim_{k\to\infty} \frac{T^k(x)}{k} = \eta, \quad \forall\, x \in X.$$

□

5.2 Sub-invariant half-lines of Bellman operators

Building on the abstract framework of Shapley operators introduced in the previous section, we now focus on the one-player minimization case. Exploiting the representation of the state space X as $C(S)$, we write the operator coordinatewise on S, which allows us to impose concavity on each component. This leads to the notion of a Bellman operator.

Let S denote the representation space of the AM-space $X = C(S)$. For $i \in S$ and $x \in X$ we write $x_i := x(i)$ for the value of x at i. Similarly, for an operator $T : X \to X$, we denote the i-th coordinate map of T by

$$T_i(x) := T(x)(i).$$

We say that an operator $T : X \to X$ is *concave* if, for every $i \in S$, the map $T_i : X \to \mathbb{R}$ is concave.

Definition 5.16 (Bellman operator)**.** A map $T : X \to X$ is called a *Bellman operator* if it is both a Shapley operator and concave.

We will discuss how this relates to a one-player minimization problem after a few preliminary considerations.

It is convenient to work in $B(S) \supseteq C(S) = X$, the space of bounded functions from S to $\mathbb{R}$. This space can be equipped with the topology of uniform convergence and with the topology of pointwise convergence. On order bounded families, the latter coincides with order convergence, see Lemma 8.17 in (Aliprantis et al., 2005). To avoid any ambiguity in this section, the symbol lim will refer to the uniform convergence topology, whereas the symbol o-lim will refer to the pointwise/order topology. It should be noted that convergence in the uniform topology implies convergence in the order topology, but not conversely.

Definition 5.17 (Recession function)**.** Let $T : X \to X$ be concave. The *recession function* of T is the map $\hat{T}$ defined by

$$\hat{T}(y) := \operatorname*{o-lim}_{s\to\infty} s^{-1}T(sy), \qquad y \in X,$$

whenever the limit exists.

The following lemma is a result in convex analysis.

Lemma 5.18. *Suppose that $T : X \to X$ is concave and nonexpansive. Then the recession function $\hat{T}$ is well defined, takes values in $B(S)$, and satisfies, for all $x, y \in X$,*

$$\hat{T}(y) = \inf_{s>0} s^{-1}(T(x + sy) - T(x)).$$

Proof. Fix $y \in X$. For each $i \in S$, the coordinate map $T_i : X \to \mathbb{R}$ is concave by assumption, hence the function $f_i := -T_i$ is convex. By Theorem A.12, for every $x \in X$, the limit

$$\lim_{s\to\infty} s^{-1}(T_i(x + sy) - T_i(x))$$

exists in $\overline{\mathbb{R}}$. Taking $x = 0$ shows that $\hat{T}_i(y) = \lim_{s\to\infty} s^{-1}T_i(sy)$ exists in $\overline{\mathbb{R}}$.

Since T is nonexpansive we have for all $s > 0$,

$$\|T(sy)\| \leq \|T(sy) - T(0)\| + \|T(0)\| \leq s\|y\| + \|T(0)\|,$$

and therefore, for every $i \in S$, $|T_i(sy)| \leq s\|y\| + \|T(0)\|$. Dividing by s yields

$$|s^{-1}T_i(sy)| \leq \|y\| + \frac{\|T(0)\|}{s} \quad \text{for all } i \in S.$$

Passing the limit $s \to \infty$, we see that $|\hat{T}_i(y)| \leq \|y\|$. Thus, $\hat{T}_i(y)$ is finite for every $i \in S$ and $\hat{T}(y)$ takes values in $B(S)$.

Finally, applying Theorem A.12 to $-T_i$ yields the representation

$$\hat{T}_i(y) = \inf_{s>0} s^{-1}(T_i(x + sy) - T_i(x))$$

for any $x \in X$. As this holds for every $i \in S$, the claim follows. □

$\hat{T}$ inherits the property of commutation with the addition of the unit from T.

Lemma 5.19. *Suppose that $T : X \to X$ is a Bellman operator. Then its recession function $\hat{T}$ commutes with the addition of the unit, that is,*

$$\hat{T}(y + ce) = \hat{T}(y) + ce, \qquad \forall\, y \in X,\ \forall\, c \in \mathbb{R}.$$

Proof. Fix $y \in X$ and $c \in \mathbb{R}$ and an arbitrary $x \in X$. By Lemma 5.18,

$$\hat{T}(y+ce) = \inf_{s>0} \frac{T(x+s(y+ce)) - T(x)}{s}.$$

Since $x + s(y+ce) = (x+sy) + sce$ and T commutes with the addition of the unit, we have for every $s > 0$,

$$T((x+sy)+sce) = T(x+sy) + sce.$$

Hence

$$\frac{T(x+s(y+ce)) - T(x)}{s} = \frac{T(x+sy) - T(x)}{s} + ce,$$

and therefore

$$\hat{T}(y+ce) = \inf_{s>0} \left\{ \frac{T(x+sy) - T(x)}{s} + ce \right\} = \inf_{s>0} \frac{T(x+sy) - T(x)}{s} + ce = \hat{T}(y) + ce.$$

□

The following Proposition provides a characterization of sub-invariant half-lines.

Proposition 5.20. *Assume that $T : X \to X$ is concave and nonexpansive. Then, $s \mapsto u + s\eta$ is a sub-invariant half-line of T if and only if*

$$T(u) \geq u + \eta \qquad \text{and} \qquad \hat{T}(\eta) \geq \eta. \tag{5.11}$$

Proof. We structure this proof in two steps proving both directions of the equivalence.

Step 1. Let $s \mapsto u + s\eta$ be a sub-invariant half-line of T, that is $T(u+s\eta) \geq u + (s+1)\eta$ for $s \geq 0$. Taking $s = 0$ gives $T(u) \geq u + \eta$. For $s > 0$, subtracting $T(u)$ and dividing by s yields

$$\frac{T(u+s\eta) - T(u)}{s} \geq \eta + \frac{u + \eta - T(u)}{s}.$$

Letting $s \to \infty$ and using $(u + \eta - T(u))/s \to 0$ gives

$$\hat{T}(\eta) = \operatorname*{o-lim}_{s\to\infty} s^{-1}(T(u+s\eta) - T(u)) \geq \eta.$$

Step 2. Conversely, assume $T(u) \geq u + \eta$ and $\hat{T}(\eta) \geq \eta$. By Lemma 5.18 we have for $s > 0$

$$\hat{T}(\eta) \leq s^{-1}(T(u+s\eta) - T(u)),$$

hence

$$T(u+s\eta) \geq T(u) + s\hat{T}(\eta) \overset{(5.11)}{\geq} u + \eta + s\eta = u + (s+1)\eta.$$

Thus $s \mapsto u + s\eta$ is a sub-invariant half-line of T. □

We now justify the terminology *Bellman operator*. Let $T : X \to X$ be an order preserving, concave operator that commutes with the addition of the unit. We show in Remark 5.21 that for each $i \in S$, the coordinate map T_i admits the explicit representation

$$T_i(x) = \inf_{a \in A_i} \{ r_i^a + P_i^a x \}, \tag{5.12}$$

where A_i is a nonempty index set, $r_i^a \in \mathbb{R}$, and $P_i^a : X \to \mathbb{R}$ is a positive continuous linear functional satisfying $P_i^a e = 1$.

Recall that $X = C(S)$, where S is a compact Hausdorff space. By the Riesz–Markov–Kakutani representation theorem, the dual space $C(S)^*$ can be identified with the space

$\mathrm{ca_r}(S)$ of regular signed Borel measures with bounded total variation on S. (Folland, 1999, Theorem 7.17). Fix $a \in A_i$. Since $P_i^a : X \to \mathbb{R}$ is a continuous linear functional, there exists a unique $\mu_i^a \in \mathrm{ca_r}(S)$ such that

$$P_i^a x = \int_S x \, d\mu_i^a, \qquad \forall\, x \in C(S).$$

Moreover, since P_i^a is positive, the representing measure μ_i^a is positive, $\mu_i^a \geq 0$. Finally, using $P_i^a e = 1$, where e is the unit, we obtain

$$1 = P_i^a e = \int_S 1 \, d\mu_i^a = \mu_i^a(S).$$

Hence, μ_i^a is a probability measure on S. Then T coincides with the one-step dynamic programming operator of a Markov decision process with state space S and action sets A_i, $i \in S$. The real number r_i^a represents the instantaneous cost associated with state i and action $a \in A_i$. μ_i^a is the probability law according to which the next state is selected.

Remark 5.21. To justify the representation of T_i in the form (5.12), we make use of the Legendre–Fenchel conjugate of the convex function $-T_i$. Recall that the conjugate of $-T_i : X \to \mathbb{R}$ is defined by

$$(-T_i)^* : X^* \to \overline{\mathbb{R}}, \qquad (-T_i)^*(x^*) := \sup_{x \in X} \big\{ \langle x^*, x \rangle - (-T_i(x)) \big\} = \sup_{x \in X} \big\{ \langle x^*, x \rangle + T_i(x) \big\},$$

where $\langle \cdot, \cdot \rangle$ denotes the canonical pairing between X^* and X. For further information, see for instance Chapter 3.3 of (Borwein et al., 2006). Since $-T_i$ is convex and continuous, and hence lower semicontinuous, the Fenchel–Moreau theorem (Borwein et al., 2006, Theorem 4.2.1) yields

$$-T_i = (-T_i)^{**} \quad \text{on } X.$$

In particular, for every $x \in X$,

$$-T_i(x) = \sup_{\nu \in X^*} \big\{ \langle x, \nu \rangle - (-T_i)^*(\nu) \big\} = \sup_{\nu \in \mathrm{ca_r}(S)} \big\{ \langle x, \nu \rangle - (-T_i)^*(\nu) \big\}. \tag{5.13}$$

Fix $\mu \in \mathrm{ca_r}(S)$. Observing that

$$(-T_i)^*(-\mu) = \sup_{x \in X} \big\{ \langle -\mu, x \rangle - (-T_i(x)) \big\},$$

we may restrict the supremum to functions of the form $x = se$, $s \in \mathbb{R}$, to obtain

$$(-T_i)^*(-\mu) \geq \sup_{s \in \mathbb{R}} \big\{ \langle -\mu, se \rangle - (-T_i(se)) \big\}.$$

Since T_i commutes with the addition of the unit, we have $T_i(se) = s + T_i(0)$, and hence

$$\sup_{s \in \mathbb{R}} \big\{ \langle -\mu, se \rangle - (-T_i(se)) \big\} = \sup_{s \in \mathbb{R}} \big\{ s(1 - \langle \mu, e \rangle) + T_i(0) \big\}.$$

It follows that $(-T_i)^*(-\mu) < \infty$ implies $\langle \mu, e \rangle = 1$. We therefore introduce the set

$$\mathcal{P}_i := \{ \mu \in \mathrm{ca_r}(S) : (-T_i)^*(-\mu) < \infty \},$$

for which the previous implication shows that $\mu \in \mathcal{P}_i$ implies $\langle \mu, e \rangle = 1$. Next, we use the order-preserving property of T_i. If $x \geq 0$, then $T_i(x) \geq T_i(0)$, and hence

$$(-T_i)^*(-\mu) = \sup_{x \in X} \big\{ \langle -\mu, x \rangle + T_i(x) \big\} \geq \sup_{x \in X_+} \big\{ \langle -\mu, x \rangle + T_i(x) \big\} \geq \sup_{x \in X_+} \big\{ \langle -\mu, x \rangle \big\} + T_i(0),$$

where $X_+ := \{x \in C(S) : x \geq 0\}$. If μ is not positive, then there exists $x_0 \in X_+$ such that $\langle \mu, x_0 \rangle < 0$, hence $\langle -\mu, x_0 \rangle > 0$. Scaling tx_0 with $t \to \infty$ yields

$$\sup_{x \in X_+} \{\langle -\mu, x \rangle\} = \infty,$$

and therefore $(-T_i)^*(-\mu) = \infty$. Consequently, $\mu \in \mathcal{P}_i$ implies $\mu \geq 0$.

Returning to (5.13) and changing the variables $\nu = -\mu$, we obtain

$$-T_i(x) = \sup_{\mu \in \mathrm{ca}_r(S)} \{\langle x, -\mu \rangle - (-T_i)^*(-\mu)\}.$$

Multiplying by -1 yields

$$T_i(x) = \inf_{\mu \in \mathrm{ca}_r(S)} \{\langle x, \mu \rangle + (-T_i)^*(-\mu)\} = \inf_{\mu \in \mathcal{P}_i} \{\langle x, \mu \rangle + (-T_i)^*(-\mu)\},$$

where the restriction to $\mathcal{P}_i$ is justified since $(-T_i)^*(-\mu) = \infty$ for $\mu \notin \mathcal{P}_i$. This is precisely the form (5.12), in which in state $i \in S$, $P_i^a x := \langle x, \mu \rangle$ representing the expected future value and $r_i^a := (-T_i)^*(-\mu)$ the instantaneous cost associated with the choice of the transition law $\mu \in \mathcal{P}_i$.

Remark 5.22. When T is of the form (5.12), the conditions characterizing sub-invariant half-lines (5.11) are equivalent to an infinite system of linear inequalities indexed by $a \in A_i$. Indeed, if $s \mapsto u + s\eta$ is a sub-invariant half-line of T, the condition $T(u) \geq u + \eta$ takes the following form for every $i \in S$:

$$\inf_{a \in A_i} \{r_i^a + P_i^a u\} \geq u_i + \eta_i,$$

which is equivalent to

$$r_i^a + P_i^a u \geq u_i + \eta_i, \qquad \forall a \in A_i.$$

Moreover, for T of the form (5.12), the recession function is $\hat{T}_i(x) = \inf_{a \in A_i} P_i^a x$. Consequently, the second inequality in (5.11), namely $\hat{T}(\eta) \geq \eta$, is equivalent to

$$P_i^a \eta \geq \eta_i, \qquad \forall a \in A_i.$$

For each $i \in S$ and $a \in A_i$, we introduce the affine map

$$T_i^a(x) := r_i^a + P_i^a x, \tag{5.14}$$

whose recession function is given by

$$\hat{T}_i^a(x) = \operatorname*{o-lim}_{s \to \infty} \frac{T_i^a(sx)}{s} = P_i^a x.$$

In order to prove Theorem 5.23, which provides a discrete-time analogue of Theorem 4.23, we introduce a discrete counterpart of the third equation in system (S).

$$\inf_{a \in A_i} \max\{\hat{T}_i^a(\eta) - \eta_i, T_i^a(u) - u_i - \eta_i\} \leq 0, \qquad \forall\, i \in S. \tag{5.15}$$

Theorem 5.23. *Suppose that $T : X \to X$ is a Bellman operator. Assume in addition that there exist $u, \eta \in X$ such that*

$$T(u) \geq u + \eta \qquad \text{and} \qquad \hat{T}(\eta) \geq \eta,$$

and that the complementarity condition (5.15) holds, with the infimum being attained for each $i \in S$. Then the map $s \mapsto u + s\eta$ is an invariant half-line of T. Moreover, the family $(v_\alpha - \alpha^{-1}\eta)_{\alpha \in (0,1)}$ remains bounded as $\alpha \to 0^+$.

Proof. By Proposition 5.20, the conditions $T(u) \geq u+\eta$ and $\hat{T}(\eta) \geq \eta$ imply that $s \mapsto u+s\eta$ is a sub-invariant half-line of T.

We now show that it is also super-invariant. By the assumption (5.15), there exists $a \in A_i$ such that

$$\max\{\hat{T}_i^a(\eta) - \eta_i,\, T_i^a(u) - u_i - \eta_i\} \leq 0.$$

Hence,

$$\hat{T}_i^a(\eta) \leq \eta_i \qquad \text{and} \qquad T_i^a(u) \leq u_i + \eta_i. \tag{5.16}$$

Using $T_i(x) = \inf_{a\in A_i} T_i^a(x)$ and the linearity of P_i^a, we have for all $s \geq 0$,

$$T_i(u+s\eta) \leq T_i^a(u+s\eta) \overset{(5.14)}{=} r_i^a + P_i^a(u+s\eta) = T_i^a(u) + sP_i^a\eta.$$

Since $\hat{T}_i^a(\eta) = P_i^a\eta$, (5.16) yields

$$T_i(u+s\eta) \leq (u_i+\eta_i) + s\eta_i = u_i + (s+1)\eta_i.$$

As this holds for every $i \in S$, we obtain $T(u+s\eta) \leq u+(s+1)\eta$ for all $s \geq 0$, so $s \mapsto u+s\eta$ is a super-invariant half-line. Together with sub-invariance, it is an invariant half-line.

Finally, the boundedness of $(v_\alpha - \alpha^{-1}\eta)_{\alpha\in(0,1)}$ as $\alpha \to 0^+$ follows from Proposition 5.11. □

The next lemma shows that any accumulation points of αv_α as $\alpha \to 0^+$ and $T^k(x)/k$ as $k \to \infty$ must satisfy a fixed point relation for the recession operator. Combined with the characterization of sub-invariant half-lines, this will allow us to identify these limits as the supremum of the corresponding director vectors.

Lemma 5.24. *Let $T : X \to X$ be a Bellman operator. Any accumulation point $\eta \in X$ of the family αv_α as $\alpha \to 0^+$ with respect to the topology of uniform convergence satisfies*

$$\hat{T}(\eta) = \eta. \tag{5.17}$$

Moreover, for any $x \in X$, any accumulation point of the sequence $T^k(x)/k$ as $k \to \infty$ with respect to the uniform convergence topology satisfies (5.17).

Proof. We structure our proof into two steps. In the first step, we prove the statement regarding the family αv_α and in the second step regarding $T^k(x)/k$.

Step 1. Let $\eta \in X$ be an accumulation point of αv_α as $\alpha \to 0^+$ in the topology of uniform convergence. Then there exists a sequence α_k with $\alpha_k \to 0^+$ such that

$$\alpha_k v_{\alpha_k} \to \eta \quad \text{uniformly as } k \to \infty,$$

that is,

$$\alpha_k v_{\alpha_k} = \eta + o(1). \tag{5.18}$$

$o(\cdot)$ denotes the little-oh notation, see Appendices. Recall that v_{α_k} is the fixed point of the discounted operator $x \mapsto T((1-\alpha_k)x)$, which means

$$v_{\alpha_k} = T((1-\alpha_k)v_{\alpha_k}). \tag{5.19}$$

Multiplying (5.19) by α_k yields

$$\alpha_k v_{\alpha_k} = \alpha_k T((1-\alpha_k)v_{\alpha_k}) = \alpha_k T((\alpha_k^{-1}-1)\alpha_k v_{\alpha_k}). \tag{5.20}$$

We claim that $(\alpha_k v_{\alpha_k})_{k\geq 1}$ is bounded. Indeed, by the fixed point identity (5.19) and nonexpansiveness of T,

$$\|v_{\alpha_k}\| = \|T((1-\alpha_k)v_{\alpha_k})\| \leq \|T((1-\alpha_k)v_{\alpha_k}) - T(0)\| + \|T(0)\| \leq (1-\alpha_k)\|v_{\alpha_k}\| + \|T(0)\|.$$

Hence $\alpha_k\|v_{\alpha_k}\| \leq \|T(0)\|$, so $\alpha_k v_{\alpha_k} = O(1)$, where $O(\cdot)$ denotes the big-oh notation. Now expand the argument of T in (5.20) to

$$(\alpha_k^{-1} - 1)\alpha_k v_{\alpha_k} = \alpha_k^{-1}(\alpha_k v_{\alpha_k}) - (\alpha_k v_{\alpha_k}) \overset{(5.18)}{=} \alpha_k^{-1}(\eta + o(1)) + O(1), \tag{5.21}$$

By nonexpansiveness of T,

$$\begin{aligned}\|\alpha_k T((\alpha_k^{-1} - 1)\alpha_k v_{\alpha_k}) - \alpha_k T(\alpha_k^{-1}\eta)\| &\overset{(5.21)}{=} \|\alpha_k T(\alpha_k^{-1}(\eta + o(1)) + O(1)) - \alpha_k T(\alpha_k^{-1}\eta)\| \\ &\leq \alpha_k\|\alpha_k^{-1}o(1) + O(1)\| \\ &\leq \|o(1)\| + \alpha_k\|O(1)\| = o(1),\end{aligned}$$

so that

$$\alpha_k T((\alpha_k^{-1} - 1)\alpha_k v_{\alpha_k}) = \alpha_k T(\alpha_k^{-1}\eta) + o(1). \tag{5.22}$$

Combining (5.18), (5.20) and (5.22), shows that

$$\eta = \alpha_k T(\alpha_k^{-1}\eta) + o(1).$$

Since uniform convergence implies order convergence, by passing the limit $k \to \infty$, we obtain

$$\eta = \underset{k\to\infty}{\text{o-lim}}\, \alpha_k T(\alpha_k^{-1}\eta) = \hat{T}(\eta).$$

Step 2. Now, let $\eta \in X$ be an accumulation point of $T^k(x)/k$ in the topology of uniform convergence. Then there exists a subsequence n_k with $n_k \to \infty$ such that

$$\frac{T^{n_k}(x)}{n_k} \to \eta \quad \text{uniformly as } k \to \infty.$$

We may assume that $x = 0$. Indeed, by nonexpansiveness we have

$$\frac{\|T^{n_k}(x) - T^{n_k}(0)\|}{n_k} \leq \frac{\|x - 0\|}{n_k} \xrightarrow[k\to\infty]{} 0,$$

hence $T^{n_k}(0)/n_k = \eta + o(1)$. We claim that $\|T^{m+1}(0) - T^m(0)\| \leq \|T(0)\|$ for all $m \geq 0$. Indeed, for $m = 0$ this statement is trivial as $\|T^1(0) - T^0(0)\| = \|T(0)\|$. Assume that for some fixed m, the statement holds, then by nonexpansiveness,

$$\|T^{m+2}(0) - T^{m+1}(0)\| = \|T(T^{m+1}(0)) - T(T^m(0))\| \leq \|T^{m+1}(0) - T^m(0)\| \leq \|T(0)\|.$$

Thus,

$$\frac{\|T^{n_k+1}(0) - T^{n_k}(0)\|}{n_k} \leq \frac{\|T(0)\|}{n_k} \xrightarrow[k\to\infty]{} 0,$$

which means that $\lim_{k\to\infty} T^{n_k+1}(0)/n_k = \eta$. Moreover,

$$\frac{T^{n_k+1}(0)}{n_k} = \frac{T(n_k n_k^{-1} T^{n_k}(0))}{n_k} = \frac{T(n_k(\eta + o(1)))}{n_k}.$$

Due to nonexpansiveness,

$$\frac{\|T(n_k(\eta + o(1))) - T(n_k\eta)\|}{n_k} \leq \frac{\|n_k\, o(1)\|}{n_k} = \|o(1)\|,$$

and therefore,

$$\frac{T^{n_k+1}(0)}{n_k} = \frac{T(n_k\eta)}{n_k} + o(1).$$

Passing the limit $k \to \infty$, yields

$$\eta = \operatorname*{o-lim}_{k\to\infty} \frac{T(n_k\eta)}{n_k} = \hat{T}(\eta),$$

which concludes the proof. □

Theorem 5.25. *Suppose that $T : X \to X$ is a Bellman operator. Then any accumulation point in the uniform convergence topology of αv_α as $\alpha \to 0^+$, coincides with the supremum, taken in X, of the director vectors of sub-invariant half-lines of T. Moreover, for any $x \in X$ the same conclusion holds for any accumulation point of the sequence $T^k(x)/k$ as $k \to \infty$ with respect to the same topology.*

Proof. We proceed in two steps. In Step 1, we prove the statement regarding αv_α and in Step 2 the one regarding $T^k(x)/k$.

Step 1. As T is a Bellman operator and in particular a Shapley operator, Proposition 5.10 ensures the existence of at least one sub-invariant half-line of T. Let $\overline{\eta}$ denote the supremum of the director vectors of all sub-invariant half-lines of T.

By Theorem 5.14 c), for every $x \in X$,

$$\liminf_{\alpha\to 0^+} \alpha v_\alpha \geq \overline{\eta} \qquad \text{and} \qquad \liminf_{k\to\infty} \frac{T^k(x)}{k} \geq \overline{\eta}. \tag{5.23}$$

Let η be the uniform limit of $\alpha_k v_{\alpha_k}$ along some sequence $\alpha_k \to 0^+$. Since uniform convergence implies order convergence, (5.23) yields

$$\eta \geq \overline{\eta}. \tag{5.24}$$

Fix $\varepsilon > 0$. Since $\alpha_k v_{\alpha_k}$ converges uniformly to η, for k large enough,

$$\alpha_k v_{\alpha_k} \geq \eta - \varepsilon e. \tag{5.25}$$

Define $u := v_{\alpha_k} - \alpha_k v_{\alpha_k}$. Using the fixed point relation $v_{\alpha_k} = T((1-\alpha_k)v_{\alpha_k})$, we obtain

$$T(u) = T(v_{\alpha_k} - \alpha_k v_{\alpha_k}) = v_{\alpha_k} = u + \alpha_k v_{\alpha_k} \overset{(5.25)}{\geq} u + \eta - \varepsilon e. \tag{5.26}$$

By Lemma 5.24, $\hat{T}(\eta) = \eta$ and by Lemma 5.19, $\hat{T}$ commutes with the addition of the unit, therefore

$$\hat{T}(\eta - \varepsilon e) = \hat{T}(\eta) - \varepsilon e = \eta - \varepsilon e.$$

Hence, by Proposition 5.20, the map $s \mapsto u + s(\eta - \varepsilon e)$ is a sub-invariant half-line of T, so $\overline{\eta} \geq \eta - \varepsilon e$, as $\overline{\eta}$ is the supremum of all director vectors. Since $\varepsilon > 0$ is arbitrary, we conclude

$$\overline{\eta} \geq \eta.$$

Together with (5.24), this yields $\overline{\eta} = \eta \in X$.

Step 2. Let η be any accumulation point of the sequence $T^k(x)/k$ for $x \in X$. As in the proof of Lemma 5.24, we may assume $x = 0$. Fix $\varepsilon > 0$. Then there exists $p \geq 1$, such that $\|T^p(0)/p - \eta\| \leq \varepsilon$. By the definition of the norm in the AM-space this implies

$$\eta - \frac{T^p(0)}{p} \leq \varepsilon e,$$

and so $T^p(0) \geq p(\eta - \varepsilon e)$. Set $\tilde{\eta} := \eta - \varepsilon e$. Therefore,

$$T^p(0) \geq p\tilde{\eta}. \tag{5.27}$$

By Lemma 5.24, we have $\hat{T}(\eta) = \eta$. Moreover, Lemma 5.19 shows that $\hat{T}$ commutes with the addition of the unit. Hence,

$$\eta = \hat{T}(\eta) = \hat{T}(\tilde{\eta} + \varepsilon e) = \hat{T}(\tilde{\eta}) + \varepsilon e.$$

Rearranging yields

$$\tilde{\eta} = \hat{T}(\tilde{\eta}). \tag{5.28}$$

By Lemma 5.18, for all $x \in X$, the recession function satisfies

$$\hat{T}(\tilde{\eta}) = \inf_{s>0} \frac{T(x + s\tilde{\eta}) - T(x)}{s}$$

and so for all $t > 0$,

$$t\hat{T}(\tilde{\eta}) \leq T(x + t\tilde{\eta}) - T(x).$$

Consequently,

$$T(x + t\tilde{\eta}) \geq T(x) + t\hat{T}(\tilde{\eta}) \overset{(5.28)}{=} T(x) + t\tilde{\eta}$$

and in particular for $x = 0$,

$$T(t\tilde{\eta}) \geq T(0) + t\tilde{\eta} \quad \text{for all } t > 0. \tag{5.29}$$

We define the vector u by

$$\begin{aligned} u &:= \sup\{(p-1)\tilde{\eta}, T(0) + (p-2)\tilde{\eta}, T^2(0) + (p-3)\tilde{\eta}, \dots, T^{p-2}(0) + \tilde{\eta}, T^{p-1}(0)\} \\ &= \sup_{j=1,\dots,p} \{T^{p-j}(0) + (j-1)\tilde{\eta}\}. \end{aligned}$$

Therefore,

$$u + s\tilde{\eta} = \sup_{j=1,\dots,p} \{T^{p-j}(0) + (s+j-1)\tilde{\eta}\} \quad \text{for all } s \geq 0.$$

For all $j = 1, \dots, p$,

$$u + s\tilde{\eta} \geq T^{p-j}(0) + (s+j-1)\tilde{\eta}$$

holds and so by the order preservation of T,

$$\begin{aligned} T(u + s\tilde{\eta}) &= T\big(\sup\{(s+p-1)\tilde{\eta}, T(0) + (s+p-2)\tilde{\eta}, \dots, T^{p-1}(0) + s\tilde{\eta}\}\big) \\ &\geq T\big(T^{p-j}(0) + (s+j-1)\tilde{\eta}\big), \qquad \forall\, j = 1, \dots, p. \end{aligned}$$

Since the left-hand side does not depend on j, taking the supremum over $j = 1, \dots, p$ yields

$$T(u + s\tilde{\eta}) \geq \sup\big\{T\big((s+p-1)\tilde{\eta}\big), T\big(T(0) + (s+p-2)\tilde{\eta}\big), \dots, T\big(T^{p-1}(0) + s\tilde{\eta}\big)\big\}.$$

Consequently, by inequality (5.29),

$$
\begin{aligned}
T(u+s\tilde{\eta}) &\geq \sup\{T(0)+(s+p-1)\tilde{\eta}, T(T(0))+(s+p-2)\tilde{\eta},\ldots,T(T^{p-1}(0))+s\tilde{\eta}\} \\
&= \sup\{T(0)+(s+p-1)\tilde{\eta}, T^2(0)+(s+p-2)\tilde{\eta},\ldots,T^p(0)+s\tilde{\eta}\}.
\end{aligned}
$$

By (5.27), $T^p(0)+s\tilde{\eta}\geq(s+p)\tilde{\eta}$, therefore

$$
T(u+s\tilde{\eta}) \geq \sup\{T(0)+(s+p-1)\tilde{\eta},\ldots,T^{p-1}(0)+(s+1)\tilde{\eta},(s+p)\tilde{\eta}\}.
$$

Rewriting the above supremum yields

$$
\begin{aligned}
T(u+s\tilde{\eta}) &\geq (s+1)\tilde{\eta}+\sup\{T(0)+(p-2)\tilde{\eta},\ldots,T^{p-1}(0),(p-1)\tilde{\eta}\} \\
&= u+(s+1)\tilde{\eta}.
\end{aligned}
$$

Thus $s\mapsto u+s\tilde{\eta}$ is a sub-invariant half-line of T with direction $\tilde{\eta}$. Therefore, by the definition of $\overline{\eta}$, we have $\overline{\eta}\geq\tilde{\eta}=\eta-\varepsilon e$. As $\varepsilon>0$ was arbitrary, we have

$$
\overline{\eta}\geq\eta. \tag{5.30}
$$

Together with (5.23), which implies $\eta\geq\overline{\eta}$, we obtain $\overline{\eta}=\eta$. This concludes the proof. □

Corollary 5.26. *Let $T:\mathbb{R}^n\to\mathbb{R}^n$ be a Bellman operator. Then the two limits $\lim_{\alpha\to 0^+}\alpha v_\alpha$ and $\lim_{k\to\infty}T^k(x)/k$ for all $x\in\mathbb{R}^n$ coincide with the supremum of the director vectors of sub-invariant half-lines of T.*

Proof. Let $\overline{\eta}$ denote the supremum of the director vectors of sub-invariant half-lines of T. By Theorem 5.25, every accumulation point of αv_α as $\alpha\to 0^+$ and of $T^k(x)/k$ as $k\to\infty$ coincides with $\overline{\eta}$.

Since $\mathbb{R}^n$ is finite-dimensional, every bounded subset of $\mathbb{R}^n$ is relatively compact. Hence, the bounded families αv_α and $T^k(x)/k$ admit accumulation points. As these accumulation points are unique, both families converge, and their limits equal $\overline{\eta}$. □

Remark 5.27. As shown in Remark 5.22, the characterization of sub-invariant half-lines associated with a Bellman operator leads to an infinite-dimensional linear programming problem. In view of the previous theorem and corollary, the mean payoff is obtained by maximizing the directing vector η subject to the corresponding system of linear inequalities.

In the special case where the state space S and the action sets A_i are finite, this system reduces to a finite collection of constraints and coincides with the finite linear program originally introduced by Denardo et al. (1968).

Remark 5.28. In the case of Markov decision processes with finite state space, several authors, see for instance (Renault, 2011), have shown that the limits

$$
\lim_{\alpha\to 0^+}\alpha v_\alpha=\lim_{k\to\infty}T^k(x)/k
$$

exist and coincide. By comparison, the novelty of the result proved in (Cannarsa et al., 2015), and recovered here in Theorem 5.25, lies in the characterization of this limit in terms of sub-invariant half-lines. This approach is analogous to Perron's method in potential theory for the Dirichlet problem, where a harmonic solution is constructed as the supremum of an appropriate class of subsolutions.

6 Conclusion

In this thesis, we have established a framework for the finite and the infinite time horizon optimal control theory. In both settings, we showed that the value function is the unique viscosity solution to the associated Hamilton–Jacobi–Bellman equation.

The central focus of our analysis was the infinite time horizon framework. We investigated the asymptotic behavior of the rescaled family $\{\lambda V_\lambda\}_{\lambda>0}$ as the discount rate $\lambda \to 0^+$. Following the approach of Cannarsa et al. (2024), we then characterized the possible uniform limit as the supremum of the family of viscosity subsolutions to the corresponding system of Hamilton–Jacobi equations. In contrast to classical ergodic results, where the limit is a constant under strong controllability assumptions, the limit in the general case is a function. When the subsolution is itself a viscosity solution of the corresponding system, we not only obtained convergence of λV_λ but also a specific rate of convergence.

This continuous-time analysis is structurally mirrored by the discrete-time part of the thesis. Working within the abstract framework of AM-spaces with unit, we analyzed Shapley and Bellman operators and the asymptotics of their discounted fixed points. We showed that the vanishing discount limit αv_α as $\alpha \to 0^+$ corresponds to the supremum of the director vectors of sub-invariant half-lines of the dynamic programming operator.

Looking forward, several open problems and directions for future research emerge from these considerations.

First, our study was primarily concerned with the Abel mean, corresponding to the vanishing discount limit. However, other notions of mean value are of interest too, such as Cesàro means or general weighted averages, see for instance the work by Li et al. (2016). A natural but challenging direction for future research would be to develop a Hamilton–Jacobi approach for such general means.

Second, a remaining challenge for the discrete framework is to identify general conditions ensuring the existence of invariant half-lines. Currently, only specific cases are resolved. Kohlberg (1980) proved the existence of invariant half-lines for nonexpansive polyhedral operators in finite dimensions. Additionally, invariant half-lines arise when the operator $T : X \to X$ admits an ergodic eigenvector, meaning there exist $u \in X$ and $\lambda \in \mathbb{R}$ such that $T(u) = u + \lambda e$, where e denotes the unit element of X. In this case, the map $s \mapsto u + s\lambda e$ defines an invariant half-line of T. Establishing relaxed conditions that guarantee existence outside of these special cases remains a compelling open problem.

Appendix

Theorems

In this section, we present classical theorems that are used throughout this thesis.

Gronwall's inequality

This version of the Gronwall's inequality can be found in the appendix of the book by Evans (2010).

Theorem A.1 (Gronwall). *Let $\eta(\cdot)$ be a nonnegative, absolutely continuous function on $[0,T]$, which satisfies for almost every t the differential inequality*

$$\eta'(t) \leq \phi(t)\eta(t) + \psi(t),$$

where $\phi(t)$ and $\psi(t)$ are nonnegative, integrable functions on $[0,T]$. Then

$$\eta(t) \leq e^{\int_0^t \phi(s)\,ds}\left(\eta(0) + \int_0^t \psi(s)\,ds\right)$$

for all $0 \leq t \leq T$.

Rademacher's theorem

Rademacher's theorem can for instance be found in Chapter 3 of (Evans et al., 2015).

Theorem A.2 (Rademacher). *Assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ is a locally Lipschitz continuous function. Then f is differentiable almost everywhere in $\mathbb{R}^n$.*

Filippov measurable selection theorem

This theorem can be found in Chapter 8 of the book by Aubin et al. (2009).

Definition A.3 (Carathéodory map). We call a map $\varphi : \widehat{\Omega} \times X \to Y$, where Y is a metric space, a *Carathéodory map*, if for every $x \in X$, $\varphi(\cdot, x)$ is measurable and for every $\omega \in \widehat{\Omega}$, $\varphi(\omega, \cdot)$ is continuous.

Theorem A.4 (Filippov). *Consider a complete σ-finite measure space $(\widehat{\Omega}, \widehat{\mathcal{A}}, \mu)$, complete separable metric spaces X and Y, and a measurable set-valued map $F : \widehat{\Omega} \rightrightarrows X$ with closed nonempty images. Let $g : \widehat{\Omega} \times X \to Y$ be a Carathéodory map. Then for every measurable map $h : \widehat{\Omega} \to Y$ satisfying*

$$h(\omega) \in g(\omega, F(\omega)) \quad \textit{for almost all } \omega \in \widehat{\Omega}$$

there exists a measurable selection $f(\omega) \in F(\omega)$ such that

$$h(\omega) = g(\omega, f(\omega)) \quad \textit{for almost all } \omega \in \widehat{\Omega}.$$

Invariance theorems

The following theorems about the invariance of differential inclusions can be found in the paper by Frankowska et al. (2000) and Chapter 5 of (Aubin, 1991).

Theorem A.5 (Invariance Theorem 1). *Assume that $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is locally Lipschitz continuous with nonempty compact values and K is a locally compact subset of $\mathbb{R}^n$. Then the following conditions are equivalent:*

$$\begin{cases} i) & K \text{ is a backward invariance domain of } G, \\ & \quad \text{which means that for every } x \in K : -G(x) \subseteq T_K(x) \text{ holds.} \\ ii) & \forall\, x \in K, \forall\, n \in \left[T_K(x)\right]^- : \sup_{g \in G(x)} \langle g, -n \rangle \leq 0 \\ iii) & \forall\, x_0 \in K, \exists\, T > 0 \text{ such that for every solution } x : [0,T) \to K \text{ to} \\ & \quad x'(t) \in G(x(t)),\ x(0) = x_0 \text{ we have } x(t) \in K \text{ for all } t \in [-T, 0]. \end{cases}$$

Theorem A.6 (Invariance Theorem 2). *Let F be a nontrivial set-valued map and $K \subseteq \mathrm{Dom}(F)$. If F is locally bounded and $\forall\, x \in \mathrm{Dom}(F) : F(x) \subseteq T_K(x)$, then K is invariant under F.*

Viability theorems

We need the following characterizations of viability. The theorems can be found in Chapter 3 of the book (Aubin, 1991) and Chapter 10 of (Aubin et al., 2009).

Theorem A.7 (Viability Theorem 1). *Let X be a finite-dimensional vector space and $K \subseteq X$ closed. Assume that the set-valued map $F : K \rightrightarrows X$ is upper semicontinuous with convex compact values. Then the three following properties are equivalent:*

$$\begin{cases} i) & \forall\, x \in K : F(x) \cap T_K(x) \neq \emptyset, \\ ii) & \forall\, x \in K : F(x) \cap \overline{\mathrm{co}}(T_K(x)) \neq \emptyset, \\ iii) & \forall\, x \in K, \forall\, p \in \left[T_K(x)\right]^- : \sup_{v \in F(x)} \langle v, -p \rangle \geq 0. \end{cases}$$

Definition A.8 (Marchaud map). We say that a set-valued map F is a *Marchaud map* if it is nontrivial, upper semicontinuous, has compact convex images and linear growth.

Theorem A.9 (Viability Theorem 2). *Let X be a finite-dimensional vector space. Let us consider a Marchaud map $F : X \rightrightarrows X$ and a closed subset $K \subseteq \mathrm{Dom}(F)$. Then $\forall\, x \in K : F(x) \cap T_K(x) \neq \emptyset$ if and only if K is viable to the differential inclusion*

$$\begin{cases} x'(t) \in F(x(t)) \\ x(0) = x_0 \in K \end{cases} \qquad \text{for almost every } t \geq 0.$$

Rockafellar lemma

The following lemma can be found in the paper by Frankowska (1993) or Plaskacz et al. (2002). It is due to a result of Rockafellar (1981).

Lemma A.10 (Rockafellar). *Let $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a lower semicontinuous function and $x \in \mathrm{Dom}(\phi)$. Let $(p, 0) \in \left[T_{\mathrm{Epi}(\phi)}(x, \phi(x))\right]^-$ be such that $p \neq 0$. Then there exist sequences x_k, $p_k \in \mathbb{R}^n$ and $q_k < 0$ such that*

$$\phi(x_k) \to \phi(x), \quad x_k \to x, \quad p_k \to p, \quad q_k \to 0, \quad \text{and} \quad (p_k, q_k) \in \left[T_{\mathrm{Epi}(\phi)}(x_k, \phi(x_k))\right]^-.$$

Arzelà–Ascoli theorem

This theorem can for instance be found in Chapter 10 of the the book by Royden et al. (2010).

Theorem A.11 (Arzelà–Ascoli). *Let X be a compact metric space and $\{f_n\}$ a uniformly bounded equicontinuous sequence of real-valued functions on X. Then $\{f_n\}$ has a subsequence that converges uniformly on X to a continuous function f on X.*

Recession function theorem

This classic result in convex analysis can for instance be found in Chapter 3 of (Rockafellar et al., 2009) for the finite dimensional case $X = \mathbb{R}^n$. For the case, where X is an arbitrary topological vector space, we refer to (Baiocchi et al., 1988).

Theorem A.12 (Recession function theorem). *Let (X, σ) be a Hausdorff topological vector space. Let $f : X \to \overline{\mathbb{R}}$ be a proper, σ-lower semicontinuous convex function. The recession function of f, denoted by $f^\infty : X \to \overline{\mathbb{R}}$, is defined by*

$$f^\infty(x) := \lim_{\lambda \searrow 0} \lambda f(\lambda^{-1} x),$$

where the limit exists for all $x \in X$. Then f^∞ is lower semicontinuous, convex, and positively homogeneous. Moreover, for any $\bar{x} \in \operatorname{dom} f$ and any $x \in X$, one has

$$f^\infty(x) = \lim_{\lambda \to \infty} \frac{f(\bar{x} + \lambda x) - f(\bar{x})}{\lambda} = \sup_{\lambda > 0} \frac{f(\bar{x} + \lambda x) - f(\bar{x})}{\lambda}.$$

Notation

i) $|x|$ denotes the Euclidean norm of $x \in \mathbb{R}^n$ and $\langle \cdot, \cdot \rangle$ the associated scalar product.

ii) We write $a := b$ to mean that a is defined to be b.

iii) $[n] := \{1, \ldots, n\}$.

iv) $\|f\|_\infty := \sup\{|f(x)| : x \in X\}$ denotes the uniform norm or sup-norm for a real-valued function f on X.

v) $B(x_0, r) := \{x \in \mathbb{R}^n : |x - x_0| < r\}$ denotes the open ball.

vi) $\operatorname{int}(X)$ denotes the interior of a set X.

vii) $\overline{X}$ denotes the closure of a set X.

viii) $\operatorname{Dom}(F)$ denotes the domain of a map F.

ix) $\operatorname{co}(X) := \{\sum_{i=1}^n \lambda_i a_i : n \in \mathbb{N},\ a_i \in X,\ \lambda_i \geq 0,\ \sum_{i=1}^n \lambda_i = 1\}$ denotes the convex hull of $X \subseteq \mathbb{R}^n$.

x) $B(\Omega)$ denotes the space of bounded functions on the set Ω.

xi) $C(\Omega)$ denotes the space of continuous functions on Ω.

xii) $C^k(\Omega)$ denotes the space of k-times continuously differentiable functions on Ω.

xiii) $\mathrm{ca_r}(S)$ denotes the space of regular signed Borel measures with bounded total variation on S.

Citing from the book of Evans (2010) we state what we mean by the big-oh and little-oh notation.

xiv) We write $f(x) = O(g(x))$ as $x \to x_0$, provided there exists a constant $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all x sufficiently close to x_0.

xv) We write $f(x) = o(g(x))$ as $x \to x_0$, provided $\lim_{x \to x_0} |f(x)|/|g(x)| = 0$.

Bibliography

ALIPRANTIS, C.D., BORDER, K.C. (2005). Infinite Dimensional Analysis. Springer, 3rd edition.

AUBIN, J.P. (1991). Viability Theory. Birkhäuser.

AUBIN, J.P., FRANKOWSKA, H. (2009). Set-Valued Analysis. Birkhäuser.

BAIOCCHI, C., BUTTAZZO, G., GASTALDI, F., TOMARELLI, F., General Existence Theorems for Unilateral Problems in Continuum Mechanics. *Archive for Rational Mechanics and Analysis*, **100 (1988)**, 149–189.

BARDI, M., CAPUZZO-DOLCETTA, I. (1997). Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Birkhäuser.

BELLMAN, R., On the Theory of Dynamic Programming. *Proceedings of the National Academy of Sciences*, **38 (1952)**, 716–719.

BISHOP, E., BRIDGES, D. (1985). Constructive Analysis. Springer.

BORWEIN, J.M., LEWIS, A.S. (2006). Convex Analysis and Nonlinear Optimization: Theory and Examples. Springer, 2nd edition.

BUCKDAHN, R., QUINCAMPOIX, M., RENAULT, J., On Representation Formulas for Long Run Averaging Optimal Control Problem. *Journal of Differential Equations*, **259 (2015)**, 5554–5581.

CANNARSA, P., GAUBERT, S., MENDICO, C., QUINCAMPOIX, M., Analysis of the Vanishing Discount Limit for Optimal Control Problems in Continuous and Discrete Time. *Mathematical Control and Related Fields*, **14 (2024)**, 1275–1305.

CANNARSA, P., QUINCAMPOIX, M., Vanishing Discount Limit and Nonexpansive Optimal Control and Differential Games. *SIAM Journal on Control and Optimization*, **53 (2015)**, 1789–1814.

CRANDALL, M.G., LIONS, P.L., Viscosity Solutions of Hamilton–Jacobi Equations. *Transactions of the American Mathematical Society*, **277 (1983)**, 1–42.

DENARDO, E.V., FOX, B.L., Multichain Markov Renewal Programs. *SIAM Journal on Applied Mathematics*, **16 (1968)**, 468–487.

EVANS, L.C. (2010). Partial Differential Equations. American Mathematical Society, 2nd edition.

EVANS, L.C. (2024). An Introduction to Mathematical Optimal Control Theory. URL https://math.berkeley.edu/~evans/control.course.pdf. Lecture notes, Department of Mathematics, University of California, Berkeley. URL accessed: 2025-11-10.

Evans, L.C., Gariepy, R.F. (2015). Measure Theory and Fine Properties of Functions, Revised Edition. CRC Press.

Folland, G.B. (1999). Real Analysis: Modern Techniques and Their Applications. John Wiley & Sons, 2nd edition.

Frankowska, H., Lower Semicontinuous Solutions of Hamilton–Jacobi–Bellman Equations. *SIAM Journal on Control and Optimization*, **31 (1993)**, 257–272.

Frankowska, H., Plaskacz, S., Semicontinuous Solutions of Hamilton–Jacobi–Bellman Equations with Degenerate State Constraints. *Journal of Mathematical Analysis and Applications*, **251 (2000)**, 818–838.

Kohlberg, E., Invariant Half-Lines of Nonexpansive Piecewise-Linear Transformations. *Mathematics of Operations Research*, **5 (1980)**, 366–372.

Li, X., Quincampoix, M., Renault, J., Limit Value for Optimal Control with General Means. *Discrete and Continuous Dynamical Systems*, **36 (2016)**, 2113–2132.

Liberzon, D. (2012). Calculus of Variations and Optimal Control Theory: A Concise Introduction. Princeton University Press.

Plaskacz, S. (2003). Value Functions in Control Systems and Differential Games: A Viability Method. Nicolaus Copernicus University. Habilitation dissertation.

Plaskacz, S., Quincampoix, M., On Representation Formulas for Hamilton–Jacobi's Equations Related to Calculus of Variations Problems. *Topological Methods in Nonlinear Analysis*, **20 (2002)**, 85–118.

Renault, J., Uniform Value in Dynamic Programming. *Journal of the European Mathematical Society*, **13 (2011)**, 309–330.

Rockafellar, R.T., Proximal Subgradients, Marginal Values, and Augmented Lagrangians in Nonconvex Optimization. *Mathematics of Operations Research*, **6 (1981)**, 424–436.

Rockafellar, R.T., Wets, R.J.B. (2009). Variational Analysis. Springer.

Royden, H.L., Fitzpatrick, P.M. (2010). Real Analysis. Pearson, 4th edition.

Schaefer, H.H. (1974). Banach Lattices and Positive Operators. Springer.

Shapley, L.S., Stochastic Games. *Proceedings of the National Academy of Sciences*, **39 (1953)**, 1095–1100.