CENTRO DE INVESTIGACIÓN Y
DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE MATEMÁTICAS

Criterios Básicos y Avanzados de Optimalidad para Juegos Diferenciales Estocásticos de Suma–Cero

Tesis que presenta

José Daniel López Barrientos

Para obtener el grado de

Doctor en Ciencias

En la Especialidad de

Matemáticas

Directores de Tesis:

Dr. Onésimo Hernández Lerma

Dr. Héctor Jasso Fuentes

México, D.F. Junio de 2012.

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL

ZACATENCO UNIT
DEPARTMENT OF MATHEMATICS

Basic and Advanced Optimality Criteria for Zero–Sum Stochastic Differential Games

Thesis presented by

José Daniel López–Barrientos

to obtain the degree of

Doctor of Science

in the Speciality of

Mathematics

Thesis Advisors:

Onésimo Hernández–Lerma, Sc.D.

Héctor Jasso–Fuentes, Sc.D.

México, D.F. June 2012.

Agradecimientos

Quiero expresar mi sincero agradecimiento y mi admiración a mis asesores, los profesores Onésimo Hernández Lerma y Héctor Jasso Fuentes. No sólo por su sabía guía, sino por la enorme paciencia que tuvieron conmigo durante todos estos años.

Doy gracias también a mi amadísima esposa, Diana. Simplemente no puedo imaginar mi vida sin el *Oso migui* en ella. Gracias por estar a mi lado en las buenas y en las malas, gracias por cuidarme y por aventurarte conmigo. Te amo *Miu miu*.

También agradezco a mi familia, *Quiquita*, Paco y *Ximio*, por su apoyo incondicional en cada proyecto que emprendo. No habría logrado esto sin ustedes.

Vaya un agradecimiento especial a mi suegra, Asunción. Porque ella es un ejemplo de fuerza y de resistencia, pero es a la vez un modelo de ternura y comprensión.

También agradezco la beca que el CONACyT me otorgó. Es gracias al Programa Nacional de Posgrado que muchos estudiantes mexicanos pueden volver sus sueños realidad.

Gracias a mi *Compita*, a *Aru* y a Héctor. Gracias por su amistad sincera y, especialmente, por aguantarme durante este lapso de nuestras vidas. Nunca olvidaré, ni dejaré de cultivar, la amistad que han fincado conmigo.

Finalmente, dedico este trabajo a la memoria de José Meza Espítia. Él inspiró, cuidó y aconsejó a varias personas (entre las que me incluyo) con verdadera sabiduría y amor puro. Que Dios deje que nosotros, los niños del Padre Meza, podamos seguir su guía durante nuestras vidas y, como él, *pasar haciendo el bien*.

Resumen

Este trabajo lidia con juegos diferenciales estocásticos (JDEs) de dos personas y suma–cero. Estudiamos la existencia de funciones de valor y puntos silla para estos juegos con varios criterios de pago en horizonte infinito.

A lo largo de nuestra tesis usaremos un muy importante resultado que involucra intercambios de límites en una sucesión de problemas de Dirichlet de tipo elíptico. Esto nos permitirá:

- Probar la existencia de (i) funciones de valor y (ii) puntos silla para JDEs con pagos descontados en horizonte infinito.
- Invocar la técnica del descuento desvaneciente y el algoritmo de iteración de políticas (AIP) para caracterizar el valor y los equilibrios de un JDE con pago ergódico.
- Estudiar criterios avanzados de optimalidad basándonos en la búsqueda de equilibrios ergódicos.

Aquí algunas de nuestras contribuciones.

- Damos condiciones suficientes que garantizan la existencia de equilibrios de Nash para cada criterio considerado.
- Caracterizaremos la función de valor de un JDE de suma cero como la solución de cierta ecuación de Isaacs y daremos condiciones suaves bajo las cuales, tal función satisface la ecuación de programación dinámica en el sentido clásico.
- También presentamos una extensión de los resultados en [32, 33, 90, 91] al caso de JDEs con tasa de descuento aleatoria.
- Proponemos una extensión del AIP para JDEs de suma cero con pagos ergódicos.
- Damos una caracterización de los equilibrios llamados *de sesgo* y *rebasantes*.

Abstract

This work is about two–person zero–sum stochastic differential games (SDGs). We study the existence of values and saddle points for these games with several infinite–horizon payoff criteria.

Throughout our thesis we shall use an important result that involves interchanging limits in a sequence of Dirichlet problems of elliptic type. This will allow us to:

- Prove the existence of both, (i) value functions and (ii) saddle points for SDGs with discounted payoffs in infinite–horizon.
- Invoke the vanishing discount technique and the policy iteration algorithm (PIA) to find the value and saddle points of a SDG with ergodic payoff.
- Study advanced optimality criteria based on the search of ergodic equilibria.

Here are some of our contributions.

- We give conditions ensuring the existence of Nash equilibria for each criterion under consideration.
- We characterize the value function of a zero–sum SDG as the solution of certain Isaacs' equation and provide mild conditions under which, such function satisfies the dynamic programming equation in the classical sense.
- We also present an extension of the results in [32, 33, 90, 91] to the case of SDGs with random rate of discount.
- We propose an extension of the PIA for zero–sum SDGs with ergodic payoffs.
- We provide a characterization of bias and overtaking equilibria.

Contents

Chapter 1

Introduction

Dynamic games can be classiffied according to the *system* itself (for instance, we can have deterministic or stochastic systems) and/or by their *performance criteria* (for instance, total, discounted or ergodic payoff criterion). Games can also be classiffied according to their *rules*. For instance, cooperative and noncooperative games; among this last category, we can find the well–known zero and nonzero sum games. In this work we deal with several infinite–horizon zero–sum dynamic games for a general class of Markov diffusion processes, which we will refer to as stochastic differential games (SDGs). Indeed, our main objective is to give conditions for the existence, characterization, and search of saddle points for four different types of infinite–horizon criteria, which we classify as *basic* and *advanced*.

The basic criteria we shall study in this thesis are the expected discounted payoff and the long–run expected average (or ergodic) payoff in a zero–sum game. These two criteria have complementary aims; while the former focuses on early periods of the time horizon, the latter concerns only asymptotic behaviors, and it does not take into account optimality for finite intervals. To overcome these two extremal situations we consider other optimality criteria which are more selective, and can be seen as "refinements" of the average payoff criterion. They are so–named because they concern policies that optimize, for each player, the average payoff but in addition they have some other convenient features. In this work, we shall study some of these refinements, namely bias and overtaking equilibria.

Throughout our thesis we shall use an important result that involves interchanging limits in a sequence of Dirichlet problems of elliptic type. This will allow us to:

- Prove the existence of both, (i) value functions and (ii) saddle points for SDGs with discounted payoffs in infinite–horizon (for a precise definition of these concepts, see, for instance, Sections 4.1.2 and 4.1.3).
- Invoke the vanishing discount technique and the policy iteration algorithm to find the value and saddle points of a SDG with ergodic payoff (see Chapter 5).
- Study advanced optimality criteria based on the search of ergodic equilibria (see Chapter 7).

We acknowledge the great influence of professors Arapostathis, Borkar and Ghosh [3, 4, 5]; Guo, Hernández–Lerma, Jasso–Fuentes, Lasserre and Mendoza–Pérez [23, 24, 38, 39, 40, 44, 45, 46, 47, 50, 51, 52, 53, 54, 68]; and Gilbarg and Trudinger [36] for inspiring and motivating our ideas.

1.1 Related literature

There are several sources for studying the basics on stochastic games. In our present case, we have used [50]. As for SDGs, we could refer to Hamadéne and Lepeltier's works (e.g. [41]). They focus on several properties of a backward stochastic differential equation (SDE) and they use some Girsanov–like results to find value functions.

As for the literature on the basic criteria, the discounted payoff criterion for SDGs has been analyzed in Bensoussan and Frehse [9], Fujita and Morimoto [29], Swiech [92] and Kushner (see, for instance, [62]). The controlled version of this problem has been studied for many classes of systems. For non–degenerate diffusion processes we refer to Kushner [60], in which the drift of the diffusion is linear in the control variable. More recent works concerning nondegenerate diffusion processes are Arapostathis et al [4] and Borkar [15]. All these works assume boundedness from below of the cost rate. Other works include that by Guo and Hernández–Lerma [38], which considers a continuous–time controlled Markov chain with a countable state space; some more general control Markov processes are studied in Hernández–Lerma [44], and Prieto–Rumeau and Hernández–Lerma [79].

On the other hand, the controlled version of the random discount problem has been studied, for discrete–time systems with denumerable state space, in the works by González–Hernández, López–Martínez, Pérez–Hernández and Minjárez–Sosa (see [32, 33], and the references therein). Jasso–Fuentes and Yin [55] and Ghosh, Arapostathis and Marcus [4] studied the discounted payoff criterion for controlled switching diffusions and we borrow some of their ideas on the form of the corresponding Bellman equation of the problem we present. Song, Yin and Zhang [90, 91] analyzed SDGs similar to those we present here.

Regarding average payoff games, Borkar and Ghosh [18], Kushner [63], and Morimoto and Ohashi [70] have already studied the ergodic payoff criterion for SDGs. For control problems with this criterion, one of the earliest works was the paper by Kushner [61]. He used dynamic programming to study a class of diffusions with bounded coefficients with additive structure. Borkar and Ghosh [16, 17] worked in a similar context but approaching the problem by using occupation measures. The unbounded case for control problems has been studied by Arapostathis et al. [3, 4], and also by Ghosh, Arapostathis and Marcus [31] for switching diffusions, except that the cost rate is supposed to be bounded below.

Overtaking optimality, which we study in Chapter 7, was introduced in the context of economic growth problems by Frank P. Ramsey in 1928 [83]. However, its present weaker form was introduced by H. Atsumi [6] and C.C. von Weizsäcker [95] in 1965 for another class of economic problems. Later on, this sort of optimality was used in many papers on Markov decision processes and control theory. Overtaking equilibria were introduced at the same time by Brock [19] in the theory of differential games, and by Rubinstein [86] for repeated games. For discrete and continuous–time games, overtaking equilibria have been obtained for several particular classes of deterministic and stochastic games, see, for instance, the works by Carlson [20, 21] and Nowak [75]. The existence of an overtaking optimal policy is a subtle issue, and there are counterexamples showing that one has to be careful when making statements on overtaking optimality; see, for instance, Nowak and Vega–Amaya [76] and the Remark 10.9.2 in Hernández–Lerma and Lasserre [45]. The bias optimality criterion for stochastic discrete–time Markov games was implicity introduced in Nowak [73, 74]. Prieto–Rumeau and Hernández–Lerma [78] studied these criteria for a continuous–time class of Markov games, whereas Jasso–Fuentes and Hernández–Lerma [52] gave conditions for the existence of bias and overtaking optimal strategies for controlled diffusions.

The policy iteration algorithm we use for the ergodic payoff criteria is inspired in the controlled versions by Hernández–Lerma and Lasserre [46], Arapostathis [5], and by a finite game version developed by Hoffman and Karp [48] modified by Van der Wal [94]. It is important to mention that the policy iteration algorithm is due to Bellman [7], although some authors credit Howard [49] for its finding. It was later used by Fleming [26] to study some finite horizon control problems in 1963. Bismut [13] and Puterman [81, 82] studied similar problems.

1.2 Contributions and outline

Our thesis deals with two–person zero–sum stochastic differential games. We study the existence of values and saddle points for these games with several infinite–horizon payoff criteria.

Our contributions are the following.

- We give conditions ensuring the existence of Nash equilibria for each criterion under consideration. A major difference between our work and those by Elliott and Kalton (see, for instance [22]), is that they consider that each player chooses his/her action regarding what the other player did in the past history, whereas we assume that both players observe the state of the system and, independently from each other, choose their actions.

- We characterize the value function of a zero–sum SDG as the solution of certain Isaacs' equation (see [26] and [27]). A difference between our work and Fleming's or Friedman's is that our hypotheses ensure the existence of the value function in all of $\mathbb{R}^n$. An improvement with respect to, for instance, Swiech's work [92] or Hamadéne and Lepeltier's [41], is that our conditions on the coefficients of the diffusion are milder, and the dynamic programming equation has a classical solution rather than a viscosity solution.

- We give conditions for the existence of classical solutions to the so–called Poisson equation by means of a relaxation of the differentiability condition in [43] on the coefficients of the diffusion that drives the system under study. This is also sufficient to ensure the existence of a bias function in the corresponding Isaacs' equation.

- We extend the results of González–Hernández et al. [32, 33] and Song et al. (see [90, 91]) to the case of SDGs with random rate of discount. Mao and Yuan's book [67] has been a great influence on this part of our work.

- We propose an extension of Fleming's policy iteration algorithm [26] (see also [5, 94, 46, 48]) for zero–sum SDGs with ergodic payoffs.

- We provide a characterization of bias and overtaking equilibria.

Our thesis is organized as follows.

In Chapter 2 we introduce the game we are interested in. We begin by introducing general conditions on the game dynamics and the reward rate. We then present the family of admissible strategies, and a stability property of the state and the action processes. We finally impose some conditions on the payoff rates.

Chapter 3 introduces Theorem 3.4, which is a crucial tool for many of our results.

The aim of Chapter 4 is to give sufficient conditions for the existence of a value function and a saddle point for the infinite horizon game with discounted payoff. First, we work with a fixed discount factor, and then, we study the random discount payoff criterion. In both cases we give hypotheses on the model and sufficient conditions to ensure the existence of a value function and saddle points. In the case of the random discounted payoff criterion, we provide an alternative version of Theorem 3.4; namely, Theorem 4.19.

Chapter 5 is devoted to the average payoff case. To this end, we establish first some definitions associated with the average payoff context. Next, we apply the well–known vanishing discount technique to ensure the existence of both, the value of the game and average equilibria.

In Chapter 6 we continue to study the ergodic case by introducing a version of the policy iteration algorithm (PIA) that transforms a given game problem into a control problem. We will refer to the policy convergence in the sense of Schäl [87, 88] to ensure the existence of saddle points for a SDG with average payoff. To be more precise, in Section 6.1 we present the PIA, and then, in Section 6.2 we show that it converges in a suitable sense. See Lemma 6.5 and Theorem 6.7.

Chapter 7 addresses the existence of bias and overtaking equilibria. With this in mind, we characterize first the bias problem as a new average payoff problem. We attain bias equilibria by means of the techniques of Chapter 5 applied to a new average problem. Besides, we show that there is a close relation between bias and overtaking equilibria. We finish this chapter by introducing a modification of the PIA proposed in Chapter 6 to find bias optimal strategies.

We conclude our work in Chapter 8 by presenting some general remarks. The rest of the thesis presents two appendices: Appendix A contains some ancillary results that are basic for our developments and are quoted several times along the thesis. Appendix B presents the proof of Theorem 3.4 and a sketch of the proof of Theorem 4.19.

1.3 Notation

Some of our results require facts from the theory of partial differential equations. Here we will use the same type of notation given in [1] and [36].

Given $\bar{\alpha} := (\alpha_1 ..., \alpha_n)$ whose components are nonnegative integers, let $|\bar{\alpha}| := \sum_{i=1}^{n} \alpha_i$. If ϕ is a smooth function, we define the derivative of order $\bar{\alpha}$ applied to ϕ as

$$D^{\bar{\alpha}}\phi := \frac{\partial^{\bar{\alpha}}\phi}{\partial^{\alpha_1}x_1 \cdots \partial^{\alpha_n}x_n}.$$

The special case $\alpha_i = 1$ and $\alpha_j = 0$ for all $j \neq i$ reduces $D^{\bar{\alpha}}\phi$ to $\frac{\partial \phi}{\partial x_i}$, the partial derivative of ϕ with respect to x_i, in which case we write $\partial_{x_i}\phi$. If $\phi : \mathbb{R}^n \to \mathbb{R}$ then $\nabla\phi$ and $\mathbb{H}\phi$ represent the gradient vector of ϕ (i.e., the row vector $(\partial_{x_i}\phi)$, $i = 1, 2, ..., n$) and the Hessian matrix of ϕ, i.e., $\mathbb{H}\phi = \left(\partial^2_{x_i x_j}\phi\right)$, for $i, j = 1, ..., n$, respectively.

Let Ω be a subset of $\mathbb{R}^n$; κ and p positive integers; and $0 < \beta < 1$.

We will consider the following spaces.

- $\mathcal{B}_w(\Omega)$ is a normed linear space of real–valued functions on Ω with finite w–norm. See Definition 2.8.
- The space $\mathcal{C}^\kappa(\Omega)$ consists of all real–valued continuous functions ϕ on Ω such that $D^{\bar{\alpha}}\phi$, $0 \le |\bar{\alpha}| \le \kappa$, is continuous as well. A particular case can be seen in Definition 2.2.
- $\mathcal{C}^{\kappa,\beta}(\Omega)$ is the normed subspace of $\mathcal{C}^\kappa(\Omega)$ consisting of those functions f for which $D^{\bar{\alpha}}f$, $0 \le |\bar{\alpha}| \le \kappa$, satisfies a Hölder condition with exponent $\beta \in]0, 1[$ on $x_i \in \mathbb{R}$, for $i = 1, 2, ..., n$. Particular cases can be seen in Definitions 3.1 and Appendix B.
- $\mathcal{L}^p(\Omega)$ is the Banach space consisting of all measurable functions f on Ω for which

  $$\int_\Omega |f(x)|^p \, dx < \infty.$$

  See Definition 3.2.
- $\mathcal{W}^{\kappa,p}(\Omega)$ is the space of measurable functions ϕ in $\mathcal{L}^p(\Omega)$ such that $D^{\bar{\alpha}}\phi$ is in $\mathcal{L}^p(\Omega)$. Here $0 \le |\bar{\alpha}| \le \kappa$ and $D^{\bar{\alpha}}\phi$ stands for a weak (or distributional) derivative of ϕ. Definition 3.3 is a particular case of this space.

Definition 1.1. *The set Ω is said to be a domain if it is an open and connected subset of $\mathbb{R}^n$.*

Definition 1.2. *A bounded domain Ω and its boundary $\partial\Omega$ are said to be of class $\mathcal{C}^{\iota,\beta}$ for $\iota \ge 0$ and $\beta \in [0, 1]$, if for each point $x_0 \in \partial\Omega$, there exists a ball $B(x_0)$ and a one–to–one mapping ψ_{x_0} from $B(x_0)$ to $D \subset \mathbb{R}^n$ such that*

(i) $\psi_{x_0}(B(x_0) \cap \Omega) \subset \mathbb{R}^n_+$,

(ii) $\psi_{x_0}(B(x_0) \cap \partial\Omega) \subset \partial\mathbb{R}^n_+$

(iii) $\psi_{x_0} \in \mathcal{C}^{\iota,\beta}(B(x_0))$ *and* $(\psi_{x_0})^{-1} \in \mathcal{C}^{\iota,\beta}(D)$.

Definition 1.3. *Let X and Y be Banach spaces. We say that X is* continuously imbedded in Y, *which will be denoted as $X \hookrightarrow Y$, if $X \subseteq Y$ and there exists a constant C such that $\|x\|_Y \le C\|x\|_X$ for every $x \in X$. Moreover, we say that X is* compactly imbedded in Y *if $X \hookrightarrow Y$ and, in addition, the unit ball in X is precompact in Y (or equivalently, every bounded sequence in X has a subsequence that converges in Y).*

Definition 1.4. *Let X be a topological space. If there exists a complete separable metric space Y and a Borel subset $B \subset Y$ such that X is homeomorphic to B, then X is said to be a Borel space.*

For vectors x and matrices A, we use the norms

$$|x|^2 := \sum_i x_i^2 \quad \text{and} \quad |A|^2 := \mathrm{Tr}\,(AA') = \sum_{i,j} A_{ij}^2,$$

where A' and $\mathrm{Tr}(\cdot)$ denote the transpose of $A = (A_{ij})$ and the trace of a square matrix, respectively.

Chapter 2

The game model

This chapter introduces the SDG we are concerned with, as well as some important concepts.

We consider the n–dimensional process $x(\cdot)$ defined, for all $t \geq 0$, by

$$dx(t) = b\,(x(t), u_1(t), u_2(t))\,dt + \sigma\,(x(t))\,dW(t) \tag{2.1}$$

with initial condition $x(0) = x$, where $b : \mathbb{R}^n \times U^1 \times U^2 \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are given functions, and $W(\cdot)$ is an m–dimensional Wiener process. The sets U^1 and U^2 are called control (or action) spaces for the players 1 and 2, respectively. For $\ell = 1, 2$, $\{u_\ell(t) : t \geq 0\}$ is a U^ℓ–valued stochastic process representing the ℓ–player's action at each time $t \geq 0$.

Assumption 2.1. *(a) The function* b *is continuous on* $\mathbb{R}^n \times U^1 \times U^2$ *and there exists a positive constant* C_1 *such that, for each* x *and* y *in* $\mathbb{R}^n$,

$$\sup_{(u_1,u_2)\in U^1\times U^2} |b\,(x, u_1, u_2) - b\,(y, u_1, u_2)| \leq C_1|x - y|$$

(b) There exists a positive constants C_2 *such that for each* x *and* y *in* $\mathbb{R}^n$,

$$|\sigma(x) - \sigma(y)| \leq C_2|x - y|,$$

(c) There exists a positive constant γ *such that the matrix* $a := \sigma\sigma'$ *satisfies:*

$$x'a(y)x \geq \gamma|x|^2 \text{ (uniform ellipticity)}, \tag{2.2}$$

for each $x \in \mathbb{R}^n$.

(d) The control sets U^1 *and* U^2 *are compact subsets of complete and separable vector normed spaces.*

Definition 2.2. *Let* $\mathcal{C}^\kappa(\mathbb{R}^n)$ *be the space of all real–valued continuous functions on* $\mathbb{R}^n$ *with continuous* l*–th partial derivative in* $x_i \in \mathbb{R}$*, for* $i = 1, ..., N$*,* $l = 0, 1, ..., \kappa$*. In particular, when* $\kappa = 0$*,* $\mathcal{C}^0(\mathbb{R}^n)$ *stands for the space of real–valued continuous functions on* $\mathbb{R}^n$*.*

Recall the notation in Section 1.3. For (u_1, u_2) in $U^1 \times U^2$ and h in $\mathcal{C}^2(\mathbb{R}^n)$, let

$$\begin{aligned}\mathbb{L}^{u_1,u_2}h(x) &:= \langle \nabla h(x), b\,(x, u_1, u_2)\rangle + \frac{1}{2}\text{Tr}\,[[\mathbb{H}h(x)]\;a(x)] \\ &= \sum_{i=1}^{n} b_i\,(x, u_1, u_2)\,\partial_{x_i}h(x) + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\partial^2_{x_i x_j}h(x),\end{aligned} \tag{2.3}$$

with $a(\cdot)$ as in Assumption 2.1(c).

2.1 Strategies

For each $\ell = 1, 2$, we denote by V^ℓ the space of probability measures on U^ℓ endowed with the topology of weak convergence. With this topology, and in view of Assumption 2.1(d), it is well known that V^ℓ is a compact metric set (see [10, Chapter 7.4], or [11, Chapter 1] for reference).

We borrow the following from Definition C.1 in [45].

Definition 2.3. *Let* X *and* Y *be two Borel spaces (recall Definition 1.4). A* stochastic kernel *on* X *given* Y *is a function* $\mathbb{P}(\cdot|\cdot)$ *such that:*

(a) $\mathbb{P}(\cdot|y)$ *is a probability measure on* X *for each fixed* $y \in Y$*, and*

(b) $\mathbb{P}(B|\cdot)$ *is a measurable function on* Y *for each fixed Borel subset* $B \subset X$*.*

The set of all stochastic kernels on X *given* Y *is denoted as* $\mathcal{P}(X|Y)$*.*

Next, we define the set of policies we are going to deal with.

Definition 2.4. *For* $\ell = 1, 2$*, a family of functions* $\pi^\ell \equiv \left\{\pi_t^\ell : t \geq 0\right\}$ *is said to be a* randomized Markov strategy *for player* ℓ *if, for every* $t \geq 0$*,* π_t^ℓ *is a stochastic kernel in* $\mathcal{P}(U^\ell|\mathbb{R}^n)$*. We denote the family of all randomized Markov strategies for player* $\ell = 1, 2$ *as* Π_m^ℓ*. Moreover, we say that* $\pi^\ell \in \Pi_m^\ell$*,* $\ell = 1, 2$*, is a* stationary strategy *if there exists a stochastic kernel* $\varphi^\ell(\cdot|\cdot) \in \mathcal{P}(U^\ell|\mathbb{R}^n)$ *such that* $\pi_t^\ell(A|x) = \varphi^\ell(A|x)$ *for all* $t \geq 0$*,* $A \subseteq U^\ell$ *and* $x \in \mathbb{R}^n$*. As an abuse of terminology, we shall write* $\pi^\ell(\cdot|\cdot) = \pi_t^\ell(\cdot|\cdot)$ *for all* $t \geq 0$ *and* $\ell = 1, 2$*.*

The family of all stationary strategies for player $\ell = 1, 2$ *will be denoted as* Π^ℓ*. Note that* $\Pi^\ell \subseteq \Pi_m^\ell$*.*

In a general context, discounted and average equilibria can be defined in terms of randomized Markov strategies. However, we will focus on the space of stationary strategies because our hypotheses ensure the existence of saddle points for the discounted and the ergodic criteria in this set (see Theorems 4.10 and 5.5). Besides, this class of policies is typically used for defining concepts such as positive recurrence, ergodicity, w–exponential ergodicity (referred to in Assumption 2.9) and bias of a pair of strategies (see (6.6) and [23, 52, 80]). In fact, as far as we can tell, the latter objects are not even defined for nonstationary strategies.

When using randomized stationary strategies (π^1, π^2) in $\Pi^1 \times \Pi^2$, we will write, for $x \in \mathbb{R}^n$,

$$b\left(x, \pi^1, \pi^2\right) := \int_{U^2}\int_{U^1} b\left(x, u_1, u_2\right)\pi^1(du_1|x)\pi^2(du_2|x). \tag{2.4}$$

For $(\varphi, \psi) \in V^1 \times V^2$, we also introduce the notation

$$b\left(x, \varphi, \psi\right) := \int_{U^2}\int_{U^1} b\left(x, u_1, u_2\right)\varphi(du_1)\psi(du_2). \tag{2.5}$$

Moreover, recalling (2.3), for $h \in \mathcal{C}^2(\mathbb{R}^n)$, let

$$\mathbb{L}^{\pi^1, \pi^2} h(x) := \int_{U^2}\int_{U^1} \mathbb{L}^{u_1, u_2} h(x)\pi^1(du_1|x)\pi^2(du_2|x). \tag{2.6}$$

We also use

$$\mathbb{L}^{\varphi, \psi} h(x) := \int_{U^2}\int_{U^1} \mathbb{L}^{u_1, u_2} h(x)\varphi(du_1)\psi(du_2),$$

for $(\varphi, \psi) \in V^1 \times V^2$.

Remark 2.5. *A direct calculation yields that* $b(\cdot, \varphi, \psi)$ *defined in* (2.5)*, has the corresponding Lipschitz property in Assumption 2.1(a), that is, there exists a constant* C_1 *such that*

$$\sup_{(\varphi, \psi) \in V^1 \times V^2} |b\left(x, \varphi, \psi\right) - b\left(y, \varphi, \psi\right)| \leq C_1 |x - y|$$

for all $x, y \in \mathbb{R}^n$*. Moreover, the Lipschitz conditions on* b *and* σ *in Assumption 2.1(a)–(b), along with the compactness of* V^1 *and* V^2 *yield that there exists a constant* $\tilde{C} \geq C_1 + C_2$ *such that*

$$\sup_{(\varphi,\psi)\in V^1\times V^2} |b(x,\varphi,\psi)| + |\sigma(x)| \leq \tilde{C}(1+|x|)$$

for all $x \in \mathbb{R}^n$.

Assumption 2.1 and Remark 2.5 ensure that, for each pair (π^1, π^2) in $\Pi^1 \times \Pi^2$, the system (2.1) admits an almost surely strong solution $x(\cdot) := \{x(t) : t \geq 0\}$, which is a Markov–Feller process whose generator coincides with the operator $\mathbb{L}^{\pi^1,\pi^2}h$ in (2.6). For more details, see [35, Theorem 2.1], [30, Theorem 3.1] and [85, Chapter III.2]. To emphasize the dependence on $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, sometimes we write $x(\cdot)$ as $x^{\pi^1,\pi^2}(\cdot)$. Also, the corresponding transition probability is

$$\mathbb{P}_x^{\pi^1,\pi^2}(t, B) := \mathbb{P}\left(x^{\pi^1,\pi^2}(t) \in B | x^{\pi^1,\pi^2}(0) = x\right)$$

for every Borel set $B \subset \mathbb{R}^n$ and $t \geq 0$. The associated conditional expectation is written as $\mathbb{E}_x^{\pi^1,\pi^2}(\cdot)$.

2.2 Ergodicity assumptions

The following hypothesis is a standard Lyapunov stability condition for continuous time (controlled and uncontrolled) Markov processes.

Assumption 2.6. *There exists a function* $w \geq 1$ *in* $\mathcal{C}^2(\mathbb{R}^n)$ *and constants* $d \geq c > 0$ *such that*

(a) $\lim_{|x|\to\infty} w(x) = \infty$.

(b) $\mathbb{L}^{\pi^1,\pi^2} w(x) \leq -cw(x) + d$ *for all* (π^1, π^2) *in* $\Pi^1 \times \Pi^2$ *and* x *in* $\mathbb{R}^n$.

Assumption 2.6 gives that, for each $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, the Markov process $x^{\pi^1,\pi^2}(t)$, $t \geq 0$, is Harris positive recurrent with a unique invariant probability measure $\mu_{\pi^1,\pi^2}(\cdot)$ for which

$$\mu_{\pi^1,\pi^2}(w) := \int_{\mathbb{R}^n} w(x)\mu_{\pi^1,\pi^2}(dx) \tag{2.7}$$

is finite. (See [3, 4, 37, 42, 69].)

By Theorem 4.3 of [3], for each pair (π^1, π^2) in $\Pi^1 \times \Pi^2$ the probability measures $\mathbb{P}_x^{\pi^1,\pi^2}(t,\cdot)$ and μ_{π^1,π^2} are both equivalent to Lebesgue's measure λ on $\mathbb{R}^n$ for every $t \geq 0$ and $x \in \mathbb{R}^n$. Hence there exists a transition density function $p^{\pi^1,\pi^2}(x, t, y)$ such that

$$\mathbb{P}_x^{\pi^1,\pi^2}(t, B) = \int_B p^{\pi^1,\pi^2}(x, t, y)dy \tag{2.8}$$

for every Borel set $B \subset \mathbb{R}^n$.

Theorem A.1 (Dynkin's formula) and, again, Assumption 2.6 ensure the *boundedness* of $\mathbb{E}_x^{\pi^1,\pi^2}[w(x(t))]$ in the sense of the following result. The proof is straightforward (see, for instance, [52, Lemma 2.10] or [69, Theorem 2.1 (iii)]).

Lemma 2.7. *The condition (b) in Assumption 2.6 implies that*

$$\mathbb{E}_x^{\pi^1,\pi^2}[w(x(t))] \leq e^{-ct}w(x) + \frac{d}{c}\left(1 - e^{-ct}\right) \tag{2.9}$$

for every (π^1, π^2) *in* $\Pi^1 \times \Pi^2$*,* $t \geq 0$*, and* $x \in \mathbb{R}^n$.

We now introduce the concept of the w–weighted norm, where w is the function in Assumption 2.6.

Definition 2.8. *Let $\mathcal{B}_w(\mathbb{R}^n)$ denote the Banach space of real–valued measurable functions ν on $\mathbb{R}^n$ with finite w–norm, which is defined as*

$$\|\nu\|_w := \sup_{x\in\mathbb{R}^n} \frac{|\nu(x)|}{w(x)}.$$

Moreover, $\mathbb{M}_w(\mathbb{R}^n)$ stands for the normed linear space of finite signed measures μ on $\mathbb{R}^n$ such that

$$\|\mu\|_w := \int_{\mathbb{R}^n} w(x)d|\mu| < \infty,$$

where $|\mu| := \mu^+ + \mu^-$ denotes the total variation *of μ.*

By (2.7), μ_{π^1,π^2} belongs to $\mathbb{M}_w(\mathbb{R}^n)$ for every $(\pi^1,\pi^2) \in \Pi^1 \times \Pi^2$. In addition, for each $\nu \in \mathcal{B}_w(\mathbb{R}^n)$, letting $\mu_{\pi^1,\pi^2}(\nu) := \int \nu d\mu_{\pi^1,\pi^2}$, we get

$$|\mu_{\pi^1,\pi^2}(\nu)| \leq \|\nu\|_w \int_{\mathbb{R}^n} w d|\mu_{\pi^1,\pi^2}| = \|\nu\|_w \|\mu_{\pi^1,\pi^2}\|_w < \infty. \tag{2.10}$$

Let T be a positive constant. For (π^1,π^2) fixed, define the T–skeleton chain of $x^{\pi^1,\pi^2}(\cdot)$ by:

$$x_T^{\pi^1,\pi^2} := \left\{x^{\pi^1,\pi^2}(kT) : k = 0, 1, ...\right\}. \tag{2.11}$$

Let $Q_m^{\pi^1,\pi^2}(x,\cdot)$ be the m–step transition probability of $x_T^{\pi^1,\pi^2}$, defined as

$$Q_m^{\pi^1,\pi^2}(x,B) := \mathbb{P}_x^{\pi^1,\pi^2}(mT, B), \ B \subseteq \mathbb{R}^n,$$

with $\mathbb{P}_x^{\pi^1,\pi^2}$ as in (2.8).

Let us impose now the following condition on $x_T^{\pi^1,\pi^2}$.

Assumption 2.9. *The skeleton chain* (2.11) *is uniformly w–exponentially ergodic. That is, there exist positive constants $\rho_1 < 1$ and ρ_2 such that, for all $m \geq 1$,*

$$\sup_{(\pi^1,\pi^2)\in\Pi^1\times\Pi^2} \left\|Q_m^{\pi^1,\pi^2}(x,\cdot) - \mu_{\pi^1,\pi^2}(\cdot)\right\|_w \leq \rho_2\rho_1^m w(x). \tag{2.12}$$

Sufficient conditions for this Assumption are given, for instance, in Assumption 4.1 and Lemma 4.8 of [68].

The proof of the following result is based on those given in [51] and [52, Theorem 2.7].

Theorem 2.10. *Suppose that Assumptions 2.1, 2.6 and 2.9 hold. Then the process $x^{\pi^1,\pi^2}(\cdot)$ is uniformly w–exponentially ergodic, that is, there exist constants $C, \delta > 0$ such that*

$$\sup_{(\pi^1,\pi^2)\in\Pi^1\times\Pi^2} \left|\mathbb{E}_x^{\pi^1,\pi^2}\nu(x(t)) - \mu_{\pi^1,\pi^2}(\nu)\right| \leq Ce^{-\delta t}\|\nu\|_w w(x) \tag{2.13}$$

for all $x \in \mathbb{R}^n$, $t \geq 0$, and $\nu \in \mathcal{B}_w(\mathbb{R}^n)$. In (2.13), *$\mu_{\pi^1,\pi^2}(\nu)$ is defined as in* (2.7), *with ν in lieu of w.*

Proof. Fix $T > 0$ and note that any $t > 0$ can be expressed in terms of T as $t = mT + s$ for some $m = 0, 1, ...,$ and $s \in [0, T[$. Hence, for every $x \in \mathbb{R}^n$, $\nu \in \mathcal{B}_w(\mathbb{R}^n)$ and $t \geq 0$ we have

$$\begin{aligned}
\left|\mathbb{E}_x^{\pi^1,\pi^2}\nu(x(t)) - \mu_{\pi^1,\pi^2}(\nu)\right| &= \left|\int_{\mathbb{R}^n} \nu(y)\left[\mathbb{P}_x^{\pi^1,\pi^2}(t,dy) - \mu_{\pi^1,\pi^2}(dy)\right]\right| \\
\leq\ & \|\nu\|_w \int_{\mathbb{R}^n} w(y)\left|\mathbb{P}_x^{\pi^1,\pi^2}(t,dy) - \mu_{\pi^1,\pi^2}(dy)\right| \\
=\ & \|\nu\|_w \int_{\mathbb{R}^n} w(y)\left|\int_{\mathbb{R}^n} \mathbb{P}_z^{\pi^1,\pi^2}(mT,dy)\mathbb{P}_x^{\pi^1,\pi^2}(s,dz) - \mu_{\pi^1,\pi^2}(dy)\right|,
\end{aligned} \tag{2.14}$$

by the Chapman–Kolmogorov equation. By Fubini's Theorem, (2.14) becomes

$$\begin{aligned}
& \left|\mathbb{E}_x^{\pi^1,\pi^2}\nu\left(x(t)\right)-\mu_{\pi^1,\pi^2}(\nu)\right| \le \|\nu\|_w \int_{\mathbb{R}^n} \mathbb{P}_x^{\pi^1,\pi^2}(s,dz)\left\|\mathbb{P}_z^{\pi^1,\pi^2}(mT,\cdot)-\mu_{\pi^1,\pi^2}(\cdot)\right\|_w \\
= \; & \|\nu\|_w \int_{\mathbb{R}^n} \mathbb{P}_x^{\pi^1,\pi^2}(s,dz)\left\|Q_m^{\pi^1,\pi^2}(z,\cdot)-\mu_{\pi^1,\pi^2}(\cdot)\right\|_w \\
\le \; & \|\nu\|_w \rho_2\rho_1^m \mathbb{E}_x^{\pi^1,\pi^2} w(x(s)) \;\text{ by (2.12)} \\
\le \; & \|\nu\|_w \rho_2\rho_1^m \left[e^{-cs}w(x)+\frac{d}{c}\left(1-e^{-cs}\right)\right] \;\text{ by (2.9)} \\
\le \; & \|\nu\|_w \rho_2\rho_1^{-1}\left(\rho_1^{1/T}\right)^t\left(1+\frac{d}{c}\right)w(x).
\end{aligned}$$

Define $C := \rho_2\rho_1^{-1}(1+d/c)$ and $\delta = -(\log\rho_1)/T$, so that the result follows. □

The following result is true by virtue of Theorem A.1 and (2.13).

Lemma 2.11. *Assume that* (2.13) *holds. Let* $(\pi^1,\pi^2)\in\Pi^1\times\Pi^2$, $\nu\in\mathcal{B}_w(\mathbb{R}^n)$, *and* $x\in\mathbb{R}^n$. *Then*

$$\lim_{T\to\infty}\frac{1}{T}\mathbb{E}_x^{\pi^1,\pi^2}\nu\left(x(T)\right)=0. \tag{2.15}$$

Suppose in addition that $\nu\in\mathcal{C}^2(\mathbb{R}^n)\cap\mathcal{B}_w(\mathbb{R}^n)$ *is a harmonic function, in the sense that*

$$\mathbb{L}^{\pi^1,\pi^2}\nu(x)=0 \;\textit{ for all }\; x\in\mathbb{R}^n. \tag{2.16}$$

Then $\nu(\cdot)$ *is a constant; in fact,*

$$\nu(x)=\mu_{\pi^1,\pi^2}(\nu) \;\textit{ for all }\; x\in\mathbb{R}^n. \tag{2.17}$$

Proof. The limit (2.15) is straightforward from (2.13). Now, if $\nu\in\mathcal{C}^2(\mathbb{R}^n)\cap\mathcal{B}_w(\mathbb{R}^n)$ is a harmonic function, then for every $(\pi^1,\pi^2)\in\Pi^1\times\Pi^2$, $x\in\mathbb{R}^n$ and $t\ge 0$, Theorem A.1 yields

$$\mathbb{E}_x^{\pi^1,\pi^2}\nu\left(x(t)\right)=\nu(x)+\mathbb{E}_x^{\pi^1,\pi^2}\int_0^T\mathbb{L}^{\pi^1,\pi^2}\nu\left(x(s)\right)ds=\nu(x), \tag{2.18}$$

where the last equality follows from (2.16). Letting $T\to\infty$ in (2.18) and using (2.13), we complete the proof. □

Following the arguments of Lemma 2.7, it is easy to verify that the combination of Lemma 2.11, Assumption 2.6 and Theorem A.1 yields

$$\mu_{\pi^1,\pi^2}(w)\le\frac{d}{c} \tag{2.19}$$

for every (π^1,π^2) in $\Pi^1\times\Pi^2$.

2.3 The payoff rate

Let R be a positive real number and $\bar{B}_R$ be the closure of

$$B_R := \{x\in\mathbb{R}^n : |x|<R\}. \tag{2.20}$$

Let us now introduce the *payoff* or *reward/cost* rate function r from $\mathbb{R}^n\times U^1\times U^2$ to $\mathbb{R}$. Let us impose some conditions on r. Recall that U^1 and U^2 are compact subsets of given vector normed spaces.

Assumption 2.12. *The function* r *is*

(a) continuous on $\mathbb{R}^n\times U^1\times U^2$ *and locally Lipschitz in* x *uniformly in* $(u_1,u_2)\in U^1\times U^2$*; that is, for each* $R>0$*, there exists a constant* $C(R)$ *such that*

$$\sup_{(u_1,u_2)\in U^1\times U^2}\left|r(x,u_1,u_2)-r(y,u_1,u_2)\right|\le C(R)|x-y|$$

for all $x,y\in\bar{B}_R$*;*

(b) *in* $\mathcal{B}_w(\mathbb{R}^n)$ *uniformly in* $(u_1, u_2) \in U^1 \times U^2$*, i.e., there exists a constant* M *such that*

$$\sup_{(u_1,u_2)\in U^1\times U^2} |r(x, u_1, u_2)| \leq Mw(x)$$

for all $x \in \mathbb{R}^n$*;*

(c) *concave in* U^1*, and convex in* U^2*.*

Analogously to (2.4) and (2.5), when using randomized Markov policies (π^1, π^2) in $\Pi^1 \times \Pi^2$, we will write, for every $x \in \mathbb{R}^n$,

$$r(x, \pi^1, \pi^2) := \int_{U^2} \int_{U^1} r(x, u_1, u_2)\, \pi^1(du_1|x)\pi^2(du_2|x).$$

Similarly, for $(\varphi, \psi) \in V^1 \times V^2$, we define

$$r(x, \varphi, \psi) := \int_{U^2} \int_{U^1} r(x, u_1, u_2)\, \varphi(du_1)\psi(du_2). \tag{2.21}$$

Remark 2.13. *We can verify that, for* $(\varphi, \psi) \in V^1 \times V^2$*, the reward rate* r *satisfies Assumption 2.12(a), that is, for each* $R > 0$*, there exists a constant* $C(R)$ *such that*

$$\sup_{(\varphi,\psi)\in V^1\times V^2} |r(x, \varphi, \psi) - r(y, \varphi, \psi)| \leq C(R)|x - y|$$

for all $x, y \in \bar{B}_R$*.*

The following result provides important facts.

Lemma 2.14. *Under Assumptions 2.1 and 2.12(a), the function* $r(\cdot, \varphi, \psi)$ *is continuous in* $(\varphi, \psi) \in V^1 \times V^2$*. Moreover, for a fixed* h *in* $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$*,* $\mathbb{L}^{\varphi,\psi}h$ *is continuous in* $(\varphi, \psi) \in V^1 \times V^2$*.*

Proof. Under the given Assumptions, the functions b and r are continuous in $(u_1, u^2) \in U^1 \times U^2$, and attain their respective suprema on U^1, and infima on U^2. Hence, the definition of weak convergence yields the result. □

Remark 2.15. *[89, Theorem 4.2]. The compactness of* U^ℓ *(resp.* V^ℓ*),* $\ell = 1, 2$*, the linearity of* $h \mapsto \mathbb{L}^{\varphi,\psi}h$*, Assumption 2.12(c), and Lemma 2.14 yield Isaacs' condition:*

$$\sup_{\varphi\in V^1} \inf_{\psi\in V^2} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi,\psi}h(x) \right\} = \inf_{\psi\in V^2} \sup_{\varphi\in V^1} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi,\psi}h(x) \right\}.$$

For ease of notation we will combine expressions such as (2.4) and (2.5), that is, for $(\varphi, \psi) \in V^1 \times V^2$ and $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$,

$$r(x, \varphi, \pi^2) := r(x, \varphi, \pi^2(\cdot|x)) \quad \text{and} \quad r(x, \pi^1, \psi) := r(x, \pi^1(\cdot|x), \psi).$$

Similarly, for $h \in \mathcal{C}^2(\mathbb{R}^n)$,

$$\mathbb{L}^{\varphi,\pi^2}h(x) := \mathbb{L}^{\varphi,\pi^2(\cdot|x)}h(x) \quad \text{and} \quad \mathbb{L}^{\pi^1,\psi}h(x) := \mathbb{L}^{\pi^1(\cdot|x),\psi}h(x).$$

Chapter 3

Interchange of limits

This chapter addresses sufficient conditions to ensure that a sequence of solutions to certain Dirichlet problems converges, in some sense, to the solution of the limiting Dirichlet problem. Such convergence will be crucial to ensure the existence of the game's value (see Definition 4.4) and of saddle points. For an extensive treatment of Dirichlet problems, we refer to [36] and [64]. To begin our analysis, we introduce first some important definitions.

Definition 3.1. *The set $\mathcal{C}^{2,\beta}(\mathbb{R}^n)$ is the normed subspace of $\mathcal{C}^2(\mathbb{R}^n)$ consisting of those functions f for which f, ∇f, and $\mathbb{H}f$ satisfy a Hölder condition with exponent $\beta \in]0,1[$ on $x_i \in \mathbb{R}$, for $i = 1, 2, ..., n$. The norm $\|\cdot\|_{\mathcal{C}^{2,\beta}(\mathbb{R}^n)}$ is defined by*

$$\begin{aligned} \|f\|_{\mathcal{C}^{2,\beta}(\mathbb{R}^n)} \quad &:= \quad \max\{\sup|f(x)|, \sup|\nabla f(x)|, \sup|\mathbb{H}f(x)|\} \\ &\quad + \max\left\{\sup\frac{|f(x)-f(y)|}{|x-y|^\beta}, \sup\frac{|\nabla f(x)-\nabla f(y)|}{|x-y|^\beta}, \sup\frac{|\mathbb{H}f(x)-\mathbb{H}f(y)|}{|x-y|^\beta}\right\} \end{aligned}$$

for each f in $\mathcal{C}^{2,\beta}(\mathbb{R}^n)$. The suprema are taken over all $x, y \in \mathbb{R}^n$, with $x \neq y$.

Let Ω be a bounded *domain* in $\mathbb{R}^n$, i.e., an open and connected subset of $\mathbb{R}^n$ and denote the closure of this set by $\bar{\Omega}$.

Definition 3.2. *Fix $p \geq 1$. The normed space $\mathcal{L}^p(\Omega)$ consists of all measurable functions f on Ω for which $\|f\|_{\mathcal{L}^p(\Omega)} < \infty$, where*

$$\|f\|_{\mathcal{L}^p(\Omega)} := \left(\int_\Omega |f(x)|^p \, dx\right)^{1/p}.$$

Definition 3.3. *The set $\mathcal{W}^{2,p}(\Omega)$ is the space of measurable functions f in $\mathcal{L}^p(\Omega)$ such that f, and its first and second weak derivatives, $\partial_{x_i} f$, $\partial^2_{x_i x_j} f$, are in $\mathcal{L}^p(\Omega)$ for all $i, j = 1, ..., n$. The corresponding norm is*

$$\|f\|_{\mathcal{W}^{2,p}(\Omega)} := \left(\int_\Omega \left[|f(x)|^p + \sum_{i=1}^n |\partial_{x_i} f(x)|^p + \sum_{i,j=1}^n \left|\partial^2_{x_i x_j} f(x)\right|^p\right] dx\right)^{1/p}.$$

For every $x \in \mathbb{R}^n$, (φ, ψ) in $V^1 \times V^2$, $\alpha > 0$, and h in $\mathcal{C}^2(\mathbb{R}^n)$ let

$$\widehat{b}(x, \varphi, \psi, h, \alpha) := \langle \nabla h(x), b(x, \varphi, \psi)\rangle - \alpha h(x) + r(x, \varphi, \psi), \tag{3.1}$$

with b as in Assumption 2.1(a) and r as in Assumption 2.12. We also define

$$\widehat{\mathbb{L}}_\alpha h(x) := \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \widehat{b}(x, \varphi, \psi, h, \alpha) + \frac{1}{2}\mathrm{Tr}\,[[\mathbb{H}h(x)]a(x)], \tag{3.2}$$

with a as in Assumption 2.1, and $\widehat{b}$ and h as in (3.1). By the Remark 2.15, we are indifferent between defining $\widehat{\mathbb{L}}_\alpha h(x)$ as in (3.2), and defining it as

$$\widehat{\mathbb{L}}_\alpha h(x) := \inf_{\psi \in V^2} \sup_{\varphi \in V^1} \widehat{b}(x, \varphi, \psi, h, \alpha) + \frac{1}{2}\mathrm{Tr}\,[[\mathbb{H}h(x)]a(x)].$$

The following result is one of the main resources along the development of our work. It gives conditions ensuring that

$$\lim_{m\to\infty} \hat{\mathbb{L}}_{\alpha_m} h_m = \hat{\mathbb{L}}_{\alpha} h \tag{3.3}$$

in some sense.

Theorem 3.4. *Let Ω be a $\mathcal{C}^2$ domain and suppose that Assumptions 2.1 and 2.12 hold. In addition, assume that there exist sequences $\{h_m\} \subset \mathcal{W}^{2,p}(\Omega)$ and $\{\xi_m\} \subset \mathcal{L}^p(\Omega)$, with $p > 1$, and a sequence $\{\alpha_m\}$ of positive numbers satisfying that:*

(a) $\hat{\mathbb{L}}_{\alpha_m} h_m = \xi_m$ *in* Ω *for* $m = 1, 2, \ldots$

(b) There exists a constant M_1 *such that* $\|h_m\|_{\mathcal{W}^{2,p}(\Omega)} \leq M_1$ *for* $m = 1, 2, \ldots$

(c) ξ_m *converges in* $\mathcal{L}^p(\Omega)$ *to some function* ξ.

(d) α_m *converges to some* α.

Then:

(i) There exist a function $h \in \mathcal{W}^{2,p}(\Omega)$ and a subsequence $\{m_k\} \subset \{1, 2, \ldots\}$ such that $h_{m_k} \to h$ in the norm of $\mathcal{W}^{1,p}(\Omega)$ as $k \to \infty$. Moreover,

$$\hat{\mathbb{L}}_{\alpha} h = \xi \ \text{ in } \ \Omega. \tag{3.4}$$

(ii) If $p > n$, then $h_{m_k} \to h$ in the norm of $\mathcal{C}^{1,\eta}(\bar{\Omega})$ for $\eta < 1 - \frac{n}{p}$. If, in addition, ξ is in $\mathcal{C}^{0,\beta}(\Omega)$, with $\beta \leq \eta$, then h belongs to $\mathcal{C}^{2,\beta}(\Omega)$.

The proof is given in Appendix B.

This result can be found in similar versions in [3, Lemma 3.5], [4, Lemma 3.4.18], and [51, Proposition A.3] for optimal control problems. We extend it here to zero–sum stochastic differential games. A major difference with [3, Lemma 3.5] and [4, Lemma 3.4.18] is that our result enables a stronger type of convergence than that in $\mathcal{W}^{2,p}(\mathbb{R}^n)$. Proposition A.3 of [51] introduced the convergence of the sequence $\{\alpha_m : m = 1, 2, \ldots\}$ to a nonnegative constant α.

An important particular case of Theorem 3.4 is that where the sets V^1 and V^2 have but one element; φ and ψ, respectively. In Corollary 3.5 below, (3.2) reduces to

$$\hat{\mathbb{L}}_{\alpha}^{\varphi,\psi} h(x) := \hat{b}\left(x, \varphi, \psi, h, \alpha\right) + \frac{1}{2}\mathrm{Tr}\left[\left[\mathbb{H}h(x)\right]a(x)\right], \tag{3.5}$$

respectively.

Corollary 3.5. *Let the Assumptions of Theorem 3.4 hold. In addition, assume that (φ, ψ) is the only element in $V^1 \times V^2$ and that there exist sequences $\{h_m\} \subset \mathcal{W}^{2,p}(\Omega)$ and $\{\xi_m\} \subset \mathcal{L}^p(\Omega)$, with $p > 1$, and a sequence $\{\alpha_m\}$ of positive numbers satisfying conditions (a)–(d) of Theorem 3.4. Then the conclusions of Theorem 3.4 hold with $\hat{\mathbb{L}}_{\alpha}^{\varphi,\psi} h$ as in* (3.5).

Chapter 4

Zero–sum stochastic differential games with discounted payoffs

This chapter deals with stationary two–person zero–sum stochastic differential games with discounted payoffs. We depart from an Itô's diffusion, a payoff rate, and the dynamic programming equation associated to these. Afterwards we will borrow some techniques from the theory of elliptic partial differential equations (PDEs) to prove the existence of a solution to such PDE (see [36]). Then we will see that this solution coincides with the value of the game (see Definition 4.4).

Earlier references for SDGs with discounted payoff are Bensoussan and Frehse [9], Fujita and Morimoto [29], Swiech [92], and Kushner [62], for instance.

We borrow Song's and Mao's concepts (see [90, 91] and [67] respectively) on switching diffusions to propose what we call *SDG with random discounted payoff*. The modification we use is in the spirit of González–Hernández et al. works on controlled Markov processes [32, 33]. That is, we intend to study a *discounted payoff criterion* where the discount factor is stochastic, rather than being fixed.

4.1 The infinite–horizon discounted payoff criterion

The goal of this section is to prove the existence of saddle points as given in (4.2). To do that, we will present some connections between the discounted payoff (4.1) and the Bellman equations (4.8)–(4.10) below. The assumptions we will make are those in Chapter 2, except for Assumption 2.9.

The game we will deal with is played as follows. At each time $t > 0$, both players observe the state of the system $x(t)$, and they independently choose control actions $u_1(t)$ in U_1 and $u_2(t)$ in U_2. For every initial state $x \in \mathbb{R}^n$, the goal of player 1 (resp. player 2) is to maximize (resp. minimize) his/her reward (resp. cost) over an infinite–horizon with respect to the optimality criterion defined in (4.1).

Fix a discount factor $\alpha > 0$. For each pair of strategies (π^1, π^2) in $\Pi^1 \times \Pi^2$ and $x \in \mathbb{R}^n$, we define the infinite–horizon discounted payoff V_α as

$$V_\alpha\left(x, \pi^1, \pi^2\right) := \mathbb{E}_x^{\pi^1, \pi^2}\left[\int_0^\infty e^{-\alpha t} r\left(x(t), \pi^1, \pi^2\right) dt\right]. \tag{4.1}$$

By Lemma 2.7, Assumption 2.12(b) and Fubini's theorem, we see that the expectation and the integral in (4.1) are interchangeable.

Definition 4.1. *A pair $(\pi^1_*, \pi^2_*) \in \Pi^1 \times \Pi^2$ is said to be a saddle point (also known as a* Nash equilibrium *or a* noncooperative equilibrium*) if*

$$V_\alpha\left(x, \pi^1, \pi^2_*\right) \leq V_\alpha\left(x, \pi^1_*, \pi^2_*\right) \leq V_\alpha\left(x, \pi^1_*, \pi^2\right) \tag{4.2}$$

for all $x \in \mathbb{R}^n$ *and* $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$.

Remark 4.2. *An economic intepretation of this definition is that when both players are in equilibrium, if one of them wishes to change his strategy, he would not earn more, on the contrary, he might actually loose value with respect to what he would earn if he sticked to the strategies of the saddle point.*

The following result establishes that the infinite–horizon discounted payoff $V_\alpha(\cdot,\cdot,\cdot)$ is dominated by the Lyapunov function w in Assumption 2.6, in a certain sense.

Proposition 4.3. *Assumptions 2.6 and 2.12(b) imply that the infinite–horizon discounted payoff* $V_\alpha(\cdot, \pi^1, \pi^2)$ *belongs to the space* $\mathcal{B}_w(\mathbb{R}^n)$ *for each* (π^1, π^2) *in* $\Pi^1 \times \Pi^2$. *Actually, for each* $x \in \mathbb{R}^n$ *we have*

$$\sup_{(\pi^1,\pi^2)\in\Pi^1\times\Pi^2} \left|V_\alpha\left(x, \pi^1, \pi^2\right)\right| \leq M(\alpha)w(x) \text{ with } M(\alpha) = M\frac{\alpha + d}{\alpha c}. \tag{4.3}$$

Here, c *and* d *are as in Assumption 2.6 and* M *is the constant in Assumption 2.12.*

The proof of Proposition 4.3 is based on Lemma 2.7. We omit it because it follows the same arguments of [51, Proposition 2.2.3], or [53, Proposition 3.6].

4.1.1 Value of the game

The functions L and U on $\mathbb{R}^n$ defined by

$$L(x) := \sup_{\pi^1\in\Pi^1} \inf_{\pi^2\in\Pi^2} V_\alpha\left(x, \pi^1, \pi^2\right) \text{ and} \tag{4.4}$$

$$U(x) := \inf_{\pi^2\in\Pi^2} \sup_{\pi^1\in\Pi^1} V_\alpha\left(x, \pi^1, \pi^2\right) \tag{4.5}$$

are called the *lower value* and the *upper value*, respectively, of the discounted payoff game. It is clear that

$$L(x) \leq U(x) \text{ for all } x \in \mathbb{R}^n. \tag{4.6}$$

When the equality holds, we obtain the following.

Definition 4.4. *If* $L(x) = U(x)$ *for all* $x \in \mathbb{R}^n$, *then the common function is called the* value of the infinite–horizon game *and it is denoted by* V.

Observe that if $(\pi^1_*, \pi^2_*) \in \Pi^1 \times \Pi^2$ satisfies the saddle point condition (4.2), a trivial calculation yields

$$U(x) \leq V_\alpha\left(x, \pi^1_*, \pi^2_*\right) \leq L(x) \text{ for all } x \in \mathbb{R}^n.$$

This fact, along with (4.6) gives that, if a saddle point (π^1_*, π^2_*) exists, then the infinite–horizon game has the value

$$V(x) = V_\alpha\left(x, \pi^1_*, \pi^2_*\right) \text{ for all } x \in \mathbb{R}^n. \tag{4.7}$$

The converse is not necessarily true.

Observe that, by (4.3), the lower and the upper values of the game, and therefore the value of the game, are in $\mathcal{B}_w(\mathbb{R}^n)$.

Let us introduce now the *α–discount Bellman equations*.

Definition 4.5. *We say that a function* v *and a pair of strategies* $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ *verify the* _α–discount Bellman equations_ *if*

$$\alpha v(x) = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} v(x) \tag{4.8}$$

$$= \sup_{\varphi\in V^1} \left\{ r\left(x, \varphi, \pi^2\right) + \mathbb{L}^{\varphi,\pi^2} v(x) \right\} \tag{4.9}$$

$$= \inf_{\psi\in V^2} \left\{ r\left(x, \pi^1, \psi\right) + \mathbb{L}^{\pi^1,\psi} v(x) \right\} \tag{4.10}$$

for all $x \in \mathbb{R}^n$.

The following result establishes a well–known relation between the infinite–horizon α–discounted payoff V_α and the solution of equation (4.11) below. It can be obtained by seeing $V_\alpha\left(\cdot,\pi^1,\pi^2\right)$ in (4.1) as the resolvent of a Markov semigroup (see, for instance, [25, p. 11] or [47, Lemma 2.2]), or by invoking Theorem A.1 for $e^{-\alpha T}v(x(T))$ and then letting T tend to ∞.

Proposition 4.6. *Fix $\alpha > 0$ and $\left(\pi^1,\pi^2\right)$ in $\Pi^1\times\Pi^2$. If a function $v\in\mathcal{C}^2(\mathbb{R}^n)\cap\mathcal{B}_w(\mathbb{R}^n)$ satisfies*

$$\alpha v(x) = r\left(x,\pi^1,\pi^2\right) + \mathbb{L}^{\pi^1,\pi^2}v(x) \tag{4.11}$$

for all $x\in\mathbb{R}^n$, then

$$v(x) = V_\alpha\left(x,\pi^1,\pi^2\right). \tag{4.12}$$

Moreover, if the equality in (4.11) *is replaced with "$\leq$" or "$\geq$", then* (4.12) *holds with the corresponding inequality.*

4.1.2 Existence of the value function

For $\left(\pi^1,\pi^2\right)$ in $\Pi^1\times\Pi^2$ and $R>0$, define the exit time

$$\tau_R^{\pi^1,\pi^2} := \inf\left\{t\geq 0 : x^{\pi^1,\pi^2}(t)\notin B_R\right\}, \tag{4.13}$$

with B_R as in (2.20). Since B_R is a bounded set, $\mathbb{E}_x^{\pi^1,\pi^2}\left[\tau_R^{\pi^1,\pi^2}\right]$ is finite. See [77, p. 119].

Define now

$$h_{\alpha,R}^{\pi^1,\pi^2}(x) := \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^{\tau_R^{\pi^1,\pi^2}} e^{-\alpha t} r\left(x(t),\pi^1,\pi^2\right)dt\right] \tag{4.14}$$

for $x\in B_R$. As in Proposition 4.3,

$$\begin{aligned}\left|h_{\alpha,R}^{\pi^1,\pi^2}(x)\right| &\leq M\int_0^\infty e^{-\alpha t}\mathbb{E}_x^{\pi^1,\pi^2}w\left(x(t)\right)dt\\ &\leq M(\alpha)w(x),\end{aligned} \tag{4.15}$$

with $M(\alpha)$ as in (4.3). This implies, in particular, that $h_{\alpha,R}^{\pi^1,\pi^2}$ belongs to $\mathcal{B}_w(\mathbb{R}^n)$.

Our next result establishes the existence of a solution to equation (4.11). Its proof is inspired in the results of Section 3.5 in [4] (see also [9, Chapter 3], [51] and [92]). We include it here for the sake of completeness.

Proposition 4.7. *Fix $\alpha>0$, $p>n$, and $\left(\pi^1,\pi^2\right)$ in $\Pi^1\times\Pi^2$. Then there exists a function v in $\mathcal{C}^2(\mathbb{R}^n)\cap\mathcal{B}_w(\mathbb{R}^n)$ that satisfies* (4.11) *for all $x\in\mathbb{R}^n$.*

Proof. We use Corollary 3.5 to prove that (4.11) admits a solution v, which is a member of $\mathcal{C}^2(\mathbb{R}^n)$.

Fix $R>0$ and consider the following *linear* Dirichlet problem:

$$\begin{aligned}\alpha v_R(x) &= r\left(x,\pi^1,\pi^2\right) + \mathbb{L}^{\pi^1,\pi^2}v_R(x) \ \text{ for all } x\in B_R, &(4.16)\\ v_R(x) &= 0 \ \text{ for all } x\in\partial B_R, &(4.17)\end{aligned}$$

with B_R as in (2.20). Observe that v_R depends on the selection of $\left(\pi^1,\pi^2\right)\in\Pi^1\times\Pi^2$. However, we will use the symbol v_R for ease of notation.

Theorem 9.15 of [36] ensures that (4.16)–(4.17) has a unique solution v_R in $\mathcal{W}^{2,p}(B_R)$.

A direct calculation yields that this solution is in fact $v_R\equiv h_{\alpha,R}^{\pi^1,\pi^2}$, with $h_{\alpha,R}^{\pi^1,\pi^2}$ as in (4.14).

Now, fix $p > n$, where n is the dimension of the system (2.1). Let $R_m \uparrow \infty$ be an increasing sequence of positive numbers such that $R_1 > 2R$. Then, for each $m = 1, 2, \ldots$ we invoke Theorem A.3 to assert the existence of a constant C_0 (independent of the sequence $\{R_m\}$) such that

$$\begin{aligned}
\|v_{R_m}\|_{\mathcal{W}^{2,p}(B_R)} &\leq C_0 \left(\|v_{R_m}\|_{\mathcal{L}^p(B_{2R})} + \left\| r\left(\cdot, \pi^1, \pi^2\right) \right\|_{\mathcal{L}^p(B_{2R})} \right) && (4.18)\\
&\leq C_0 \left(M(\alpha) \|w\|_{\mathcal{L}^p(B_{2R})} + M \|w\|_{\mathcal{L}^p(B_{2R})} \right) && (4.19)\\
&\leq C_0 (M(\alpha) + M) |\bar{B}_{2R}|^{1/p} \max_{x \in \bar{B}_R} w(x), && (4.20)
\end{aligned}$$

where $M(\alpha)$ and M are the constants in (4.3) and in Assumption 2.12, respectively. In (4.20), $|\bar{B}_{2R}|$ denotes the volume of the closed ball with radius 2R. Thus, applying Corollary 3.5, we ensure the existence of a function $v \in \mathcal{C}^{2,\beta}(B_R)$, with $\beta \in]0,1[$, such that $v_R \to v$ uniformly on B_R, and also

$$\alpha v(x) = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1, \pi^2} v(x) \text{ for all } x \in B_R.$$

Since $R > 0$ was arbitrary, we can extend the previous convergence $v_{R_m} \to v$ to all of $\mathbb{R}^n$. Finally, using the fact that $v_{R_m} \equiv h^{\pi^1, \pi^2}_{\alpha, R_m}$ and combining it with inequality (4.15) we get that v_{R_m} is in $\mathcal{B}_w(\mathbb{R}^n)$. Hence, by the uniform convergence of $v_{R_m} \to v$ and by Lemma A.4, we conclude that v is in $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$. This completes the proof. □

Remark 4.8. *From Proposition 4.7, it should be noted that the function $v \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ depends implicitly on the choice of $\left(\pi^1, \pi^2\right)$ in $\Pi^1 \times \Pi^2$.*

4.1.3 Existence of a saddle point

In this subsection we establish a result on the existence of a solution to the α–discount Bellman equations (4.8)–(4.10).

Theorem 4.9. *Recall that n is the dimension of the diffusion* (2.1). *Let $p > n$. Fix an arbitrary $\alpha \in]0,1[$. If Assumptions 2.1, 2.6 and 2.12 hold, then there exist a function v in $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ and a pair of strategies $\left(\pi^1, \pi^2\right)$ in $\Pi^1 \times \Pi^2$ that verify the α–discount Bellman equations* (4.8)–(4.10).

Proof. Fix $\alpha > 0$, $p > n$, $R > 0$, and consider the Dirichlet problem

$$\begin{aligned}
\alpha v_R(x) &= \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi, \psi} v_R(x) \right\} && (4.21)\\
&= \inf_{\psi \in V^2} \sup_{\varphi \in V^1} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi, \psi} v_R(x) \right\} && (4.22)
\end{aligned}$$

for all $x \in B_R$, and the boundary condition

$$v_R(x) = 0 \text{ for all } x \in \partial B_R, \qquad (4.23)$$

with B_R as in (2.20).

By Theorem 15.2 of [36] (or Theorem 3.4.17 of [4] in the context of controlled diffusions), the problem (4.21)–(4.23) has a solution $v_R \in \mathcal{C}^{2,\beta}(B_R)$, with $0 < \beta < 1$.

By Theorem A.2, we can assert the existence of a pair $\left(\pi^1_*, \pi^2_*\right) \in \Pi^1 \times \Pi^2$ such that (4.16) holds. With this in mind, it is easy to verify that $v_R \equiv h^{\pi^1, \pi^2}_{\alpha, R}$ with $h^{\pi^1, \pi^2}_{\alpha, R}$ as in (4.14). Thus, by (4.15), we see that v_R is in $\mathcal{B}_w(B_R)$.

Now let $R_m \uparrow \infty$ be an increasing sequence with $R_1 > 2R$, and let $\left(\pi^1_m, \pi^2_m\right)$ be such that (4.16) holds. By Theorem A.3 there exists a constant C_0 (independent of R_m) such that

$$\|v_{R_m}\|_{\mathcal{W}^{2,p}(B_R)} \leq C_0 \left(\|v_{R_m}\|_{\mathcal{L}^p(B_{2R})} + \left\| r\left(\cdot, \pi^1_m, \pi^2_m\right) \right\|_{\mathcal{L}^p(B_{2R})} \right)$$

$$\leq \quad C_0(M(\alpha)+M)|\bar{B}_{2R}|^{1/p} \max_{x\in\bar{B}_R} w(x) < \infty.$$

Thus, Theorem 3.4 yields the existence of a function $v \in \mathcal{C}^{2,\beta}(B_R)$, with $\beta \in]0,1[$, such that $v_{R_m} \to v$ and

$$\begin{aligned} \alpha v(x) &= \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \left\{ r(x,\varphi,\psi) + \mathbb{L}^{\varphi,\psi} v(x) \right\} \\ &= \inf_{\psi\in V^2} \sup_{\varphi\in V^1} \left\{ r(x,\varphi,\psi) + \mathbb{L}^{\varphi,\psi} v(x) \right\} \end{aligned}$$

with $x \in B_R$. Now, since $R > 0$ was arbitrary, we can extend the convergence $v_{R_m} \to v$ to all of $\mathbb{R}^n$.

Finally, since v_R is in $\mathcal{B}_w(\mathbb{R}^n)$, and due to the uniform convergence of v_R to v, we can use Lemma A.4 to conclude that v is a member of $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$. □

The following result establishes the equivalence between a saddle point of the α–discount game and the strategies that verify equations (4.8)–(4.10).

Theorem 4.10. *Assume the hypotheses of Theorem 4.9. Then the function v in $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ and the pair (π^1_*, π^2_*) of strategies in $\Pi^1 \times \Pi^2$ that satisfy the α–discount Bellman equations* (4.8)–(4.10) *are such that:*

(a) The function $v(x)$ equals the value function $V(x)$ in Definition 4.4 for all $x \in \mathbb{R}^n$, and

(b) The pair (π^1_*, π^2_*) is a saddle point, and therefore, from (4.7)*, $V_\alpha(x, \pi^1_*, \pi^2_*) = V(x)$ for all $x \in \mathbb{R}^n$.*

Proof. (a) Since the existence of the pair $(\pi^1_*, \pi^2_*) \in \Pi^1 \times \Pi^2$ is ensured by Theorem A.2, a comparison of (4.8) with (4.11) yields that part *(a)* follows from Proposition 4.6.

(b) Let $V^*(x, \pi^1, \pi^2)$ be defined by

$$V^*(x, \pi^1, \pi^2) := r(x, \pi^1, \pi^2) + \mathbb{L}^{\pi^1,\pi^2} v(x). \tag{4.24}$$

Interpreting this function as the payoff of a certain game, it follows from (4.8)–(4.10) and [50, Proposition 4.3] that the pair (π^1_*, π^2_*) is a saddle point, that is, $V^*(x, \pi^1_*, \pi^2_*) = \alpha v(x)$ satisfies (4.2). More explicitly, from (4.24) and the equality $V^*(x, \pi^1_*, \pi^2_*) = \alpha v(x)$, (4.2) becomes:

$$r(x, \pi^1, \pi^2_*) + \mathbb{L}^{\pi^1,\pi^2_*} v(x) \leq \alpha v(x) \leq r(x, \pi^1_*, \pi^2) + \mathbb{L}^{\pi^1_*,\pi^2} v(x)$$

for all $x \in \mathbb{R}^n$, and $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. These two inequalities, along with the second part of Proposition 4.6, give (4.2).

□

4.2 The infinite–horizon random discounted payoff criterion

In this section we consider a class of SDGs with two special features: (i) the system evolves according to a *Markov–modulated diffusion* and (ii) the *sum* of the switching parameters, up to time $t \geq 0$, serves as *discount rate* for studying an extension of the discounted payoff criterion of Section 4.1.

To do this, we shall replace the constant $\alpha > 0$ by a time–homogeneous continuous–time irreducible Markov chain, namely $\alpha(\cdot)$. We assume the state space of $\alpha(\cdot)$ is a finite set $E = \{\alpha_1, \alpha_2, ..., \alpha_N\}$ of positive real numbers. Let $q_{ij} \geq 0$ be the transition rate from state i to state j and observe that the transition probabilities are given by

$$\mathbb{P}(\alpha(s+t) = \alpha_j | \alpha(s) = \alpha_i) = q_{ij} t + o(t), \tag{4.25}$$

for states $\alpha_i \neq \alpha_j$, and $q_{ii} = -\sum_{j\neq i} q_{ij}$. The matrix $Q = [q_{ij}]$ is known as the infinitesimal matrix of the process $\alpha(\cdot)$.

Now let the evolution of the game be driven by a diffusion of the form

$$dx(t) = b(x(t), \alpha(t), u_1(t), u_2(t))dt + \sigma(x(t), \alpha(t))dW(t), \tag{4.26}$$

with initial conditions $x(0) = x$, and $\alpha(0) = \alpha_i$. Here, $W(\cdot)$ is an m–dimensional Brownian motion as in (2.1) and it is assumed to be independent of $\alpha(\cdot)$. The expressions (4.25)–(4.26) are known as a Markov–modulated diffusion with switching parameter $\alpha(\cdot)$.

It is well known that, even though $x(t)$ itself may not satisfy the Markov property, the joint process $(x(t), \alpha(t))$ is Markov. See, for instance, [67, p. 104–106].

We will denote the transition probability of the process $(x(t), \alpha(t))$ by

$$\mathbb{P}^{u_1,u_2}_{x,\alpha_i}(t, B \times J) := \mathbb{P}\left((x^{u_1,u_2}(t), \alpha(t)) \in B \times J \mid x(0) = x, \alpha(0) = \alpha_i\right)$$

for every Borel set $B \subset \mathbb{R}^n$ and $J \subset E$. The associated conditional expectation is written $\mathbb{E}^{u_1,u_2}_{x,\alpha_i}(\cdot)$.

The game is played as before in Section 4.1, except that the state of the game is now $(x(t), \alpha(t))$, rather than just $x(\cdot)$. For every initial state $(x, \alpha_i) \in \mathbb{R}^n \times E$, the goal of player 1 (resp. player 2) is to choose a strategy π^1 (resp. π^2) –see Definition 4.12 below– that maximizes (resp. minimizes) his/her random discounted payoff over an infinite–horizon with respect to the optimality criterion defined by

$$V\left(x, \alpha_i, \pi^1, \pi^2\right) := \mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left[\int_0^\infty \exp(-S_t) r\left(x(t), \alpha(t), \pi^1, \pi^2\right) dt\right], \tag{4.27}$$

where, for $t \geq 0$,

$$S_t := \int_0^t \alpha(s)ds, \ S_0 := 0. \tag{4.28}$$

We will refer to (4.27) as the *infinite horizon random discounted payoff*. The fact that $\alpha_j > 0$ for $j = 1, ..., N$, along with Assumption 4.15, below, ensures that (4.27) is finite.

Let us impose some conditions on the model (4.25)–(4.26). These conditions are much alike Assumption 2.1, except they use $\mathbb{R}^n \times E$ in lieu of just $\mathbb{R}^n$.

Assumption 4.11. *(a) The function* b *is continuous on* $\mathbb{R}^n \times E \times U^1 \times U^2$ *and there exists a positive constant* C_1 *such that, for each* x *and* y *in* $\mathbb{R}^n$,

$$\sup_{(\alpha,u_1,u_2)\in E\times U^1\times U^2} |b(x, \alpha, u_1, u_2) - b(y, \alpha, u_1, u_2)| \leq C_1|x - y|$$

(b) There exists a positive constant C_2 *such that for each* x *and* y *in* $\mathbb{R}^n$,

$$\sup_{\alpha\in E} |\sigma(x, \alpha) - \sigma(y, \alpha)| \leq C_2|x - y|.$$

(c) There exists a constant and $\gamma > 0$ *such that, for each* x *in* $\mathbb{R}^n$, *the matrix* $a(\cdot, \cdot) := \sigma(\cdot, \cdot)\sigma'(\cdot, \cdot)$ *satisfies*

$$\inf_{\alpha\in E} x'a(y, \alpha)x \geq \gamma|x|^2 \text{ (uniform ellipticity).}$$

(d) The control sets U^1 *and* U^2 *are compact subsets of complete and separable vector normed spaces.*

Using the notation in Section 1.3, let $\mathcal{C}^2(\mathbb{R}^n \times E)$ be the space of real–valued continuous functions h on $\mathbb{R}^n \times E$ such that $h(x, \alpha)$ is continuously differentiable in $x \in \mathbb{R}^n$ for each $\alpha_i \in E$. For $h \in \mathcal{C}^2(\mathbb{R}^n \times E)$, let

$$\mathcal{Q}h(x, \alpha_i) := \sum_{j=1}^N q_{ij} h(x, \alpha_j).$$

Analogously to (2.3), for $(u_1, u_2) \in U^1 \times U^2$ and $h \in \mathcal{C}^2(\mathbb{R}^n \times E)$, let

$$\mathbb{L}^{u_1,u_2} h(x, \alpha_i) := \langle \nabla h(x, \alpha_i), b(x, \alpha_i, u_1, u_2)\rangle + \frac{1}{2}\text{Tr}\left[[\mathbb{H}h(x, \alpha_i)] \cdot a(x, \alpha_i)\right] + \mathcal{Q}h(x, \alpha), \tag{4.29}$$

with $a(\cdot, \cdot)$ as in Assumption 4.11(c).

Strategies

As in Section 2.1, for each $\ell = 1, 2$, we denote by V^ℓ the space of probability measures on U^ℓ endowed with the topology of weak convergence. We define now the control policies we are going to use for the present variation of the game model. The following definition matches Definition 2.4 (except for the fact that, here, we put $\mathbb{R}^n \times E$ instead of just $\mathbb{R}^n$).

Definition 4.12. *For $\ell = 1, 2$, a family of functions $\pi^\ell \equiv \{\pi^\ell_t : t \geq 0\}$ is said to be a randomized Markov strategy for player ℓ if, for every $t \geq 0$, π^ℓ_t is a stochastic kernel in $\mathcal{P}(U^\ell|\mathbb{R}^n \times E)$. We denote the family of all randomized Markov strategies for player $\ell = 1, 2$ as Π^ℓ_m. Moreover, we say that $\pi^\ell \in \Pi^\ell_m$, $\ell = 1, 2$, is a stationary strategy if there exists a stochastic kernel $\varphi^\ell(\cdot|\cdot,\cdot) \in \mathcal{P}(U^\ell|\mathbb{R}^n \times E)$ such that $\pi^\ell_t(A|x,\alpha) = \varphi^\ell(A|x,\alpha)$ for all $t \geq 0$, $A \subseteq U^\ell$ and $(x,\alpha) \in \mathbb{R}^n \times E$. In this case we write $\pi^\ell(\cdot|\cdot,\cdot)$ rather than $\pi^\ell_t(\cdot|\cdot,\cdot)$.*

The family of all stationary strategies for player $\ell = 1, 2$ will be denoted as Π^ℓ.

When using randomized stationary strategies (π^1, π^2) in $\Pi^1 \times \Pi^2$, we will write, for $(x,\alpha) \in \mathbb{R}^n \times E$,

$$b\left(x, \alpha, \pi^1, \pi^2\right) := \int_{U^2} \int_{U^1} b\left(x, \alpha, u_1, u_2\right) \pi^1(du_1|x,\alpha)\pi^2(du_2|x,\alpha).$$

For $(\varphi, \psi) \in V^1 \times V^2$, we also introduce the notation

$$b\left(x, \alpha, \varphi, \psi\right) := \int_{U^2} \int_{U^1} b\left(x, \alpha, u_1, u_2\right) \varphi(du_1)\psi(du_2).$$

Moreover, for $h \in \mathcal{C}^2(\mathbb{R}^n \times E)$, let

$$\mathbb{L}^{\pi^1,\pi^2} h(x,\alpha) := \int_{U^2} \int_{U^1} \mathbb{L}^{u_1,u_2} h(x,\alpha)\pi^1(du_1|x,\alpha)\pi^2(du_2|x,\alpha). \tag{4.30}$$

We also use

$$\mathbb{L}^{\varphi,\psi} h(x,\alpha) := \int_{U^2} \int_{U^1} \mathbb{L}^{u_1,u_2} h(x,\alpha)\varphi(du_1)\psi(du_2),$$

for $(\varphi, \psi) \in V^1 \times V^2$.

Assumption 4.11 ensures that, for each pair (π^1, π^2) in $\Pi^1 \times \Pi^2$, the system (4.25)–(4.26) admits an almost surely unique strong solution $x(\cdot) := \{x(t) : t \geq 0\}$, such that $((x\cdot), \alpha(\cdot))$ is a Markov–Feller process whose generator coincides with the operator $\mathbb{L}^{\pi^1,\pi^2} h$ in (4.30). For details, see, for instance, [67, pp. 88–90]. Moreover, the operator $\mathbb{L}^{\pi^1,\pi^2}$ coincides with the infinitesimal generator associated to the pair $(x(\cdot), \alpha(\cdot))$ in (4.25)–(4.26). See [67, p. 48], for instance.

Some assumptions and definitions

The following hypothesis is a Lyapunov–like condition analogous to Assumption 2.6.

Assumption 4.13. *There exists a function $w \geq 1$ in $\mathcal{C}^2(\mathbb{R}^n \times E)$ and constants $d \geq c > 0$ such that*

(a) $\lim_{|x|\to\infty} w(x,\alpha) = \infty$ *for all* $\alpha \in E$.

(b) $\mathbb{L}^{\pi^1,\pi^2} w(x,\alpha) \leq -cw(x,\alpha) + d$ *for all* (π^1, π^2) *in* $\Pi^1 \times \Pi^2$ *and* $(x,\alpha) \in \mathbb{R}^n \times E$.

Here, as in Lemma 2.7, the condition (b) in Assumption 4.13 implies that

$$\mathbb{E}^{\pi^1,\pi^2}_{x,\alpha}\left[w\left(x(t), \alpha(t)\right)\right] \leq e^{-ct} w(x,\alpha) + \frac{d}{c}\left(1 - e^{-ct}\right) \tag{4.31}$$

for every (π^1, π^2) in $\Pi^1 \times \Pi^2$, $t \geq 0$, and $(x,\alpha) \in \mathbb{R}^n \times E$.

Definition 4.14. *Let $\mathcal{B}_w(\mathbb{R}^n \times E)$ denote the Banach space of real–valued measurable functions v on $\mathbb{R}^n \times E$ with finite w–norm, which is defined as*

$$\|v\|_w := \sup_{(x,\alpha)\in\mathbb{R}^n\times E} \frac{|v(x,\alpha)|}{w(x,\alpha)}.$$

Let $r : \mathbb{R}^n \times E \times U^1 \times U^2 \to \mathbb{R}$ be a measurable function, which we call the payoff rate. The following conditions are analogous to those in Assumption 2.12.

Assumption 4.15. *The function r is*

(a) continuous on $\mathbb{R}^n \times E \times U^1 \times U^2$ and locally Lipschitz in x uniformly in $(\alpha, u_1, u_2) \in E \times U^1 \times U^2$; that is, for each $R > 0$, there exists a constant $C(R)$ such that

$$\sup_{(\alpha,u_1,u_2)\in E\times U^1\times U^2} |r(x,\alpha,u_1,u_2) - r(y,\alpha,u_1,u_2)| \leq C(R)|x-y|$$

for all $|x|, |y| \leq R$;

(b) in $\mathcal{B}_w(\mathbb{R}^n \times E)$ uniformly in $(u_1, u_2) \in U^1 \times U^2$, i.e., there exists a constant M such that

$$\sup_{(u_1,u_2)\in U^1\times U^2} |r(x,\alpha,u_1,u_2)| \leq Mw(x,\alpha)$$

for all $(x, \alpha) \in \mathbb{R}^n \times E$;

(c) concave in U^1 and convex in U^2 for every $(x, \alpha) \in \mathbb{R}^n \times E$.

When using randomized Markov strategies (π^1, π^2) in $\Pi^1 \times \Pi^2$, we will write, for every $(x, \alpha) \in \mathbb{R}^n \times E$,

$$r\left(x,\alpha,\pi^1,\pi^2\right) := \int_{U^2}\int_{U^1} r(x,\alpha,u_1,u_2)\,\pi^1(du_1|x,\alpha)\pi^2(du_2|x,\alpha); \tag{4.32}$$

and, for $(\varphi, \psi) \in V^1 \times V^2$,

$$r(x,\alpha,\varphi,\psi) := \int_{U^2}\int_{U^1} r(x,\alpha,u_1,u_2)\,\varphi(du_1)\psi(du_2).$$

Similarly, for $(\varphi, \psi) \in V^1 \times V^2$ and $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$,

$$r\left(x,\alpha,\varphi,\pi^2\right) := r\left(x,\alpha,\varphi,\pi^2(\cdot|x,\alpha)\right),$$

and

$$r\left(x,\alpha,\pi^1,\psi\right) := r\left(x,\alpha,\pi^1(\cdot|x,\alpha),\psi\right).$$

Finally, for $h \in \mathcal{C}^2(\mathbb{R}^n \times E)$, we also write

$$\mathbb{L}^{\varphi,\pi^2}h(x,\alpha) := \mathbb{L}^{\varphi,\pi^2(\cdot|x,\alpha)}h(x,\alpha)$$

and

$$\mathbb{L}^{\pi^1,\psi}h(x,\alpha) := \mathbb{L}^{\pi^1(\cdot|x,\alpha),\psi}h(x,\alpha).$$

The following result is analogous to Lemma 2.14.

Lemma 4.16. *Fix h in $\mathcal{C}^2(\mathbb{R}^n \times E) \cap \mathcal{B}_w(\mathbb{R}^n \times E)$. Under Assumptions 4.11 and 4.15(c), the functions $r(x, \alpha, \varphi, \psi)$ and $\mathbb{L}^{\varphi,\psi}$ are continuous in $(\varphi, \psi) \in V^1 \times V^2$ for every $(x, \alpha) \in \mathbb{R}^n \times E$.*

Remark 4.17. *Analogously to Remark 2.15, the compactness of U^ℓ ($\ell = 1, 2$), the linearity of $h \mapsto \mathbb{L}^{\varphi,\psi}h$, Assumption 4.15(c), and Lemma 4.16 yield Isaacs' condition:*

$$\begin{aligned}
&\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\left\{r(x,\alpha,\varphi,\psi) + \mathbb{L}^{\varphi,\psi}h(x,\alpha)\right\} \\
=\ &\inf_{\psi\in V^2}\sup_{\varphi\in V^1}\left\{r(x,\alpha,\varphi,\psi) + \mathbb{L}^{\varphi,\psi}h(x,\alpha)\right\}.
\end{aligned}$$

Recalling (4.27)–(4.28), we will say that a pair $(\pi^1_*, \pi^2_*) \in \Pi^1 \times \Pi^2$ is a saddle point of the infinite horizon random discounted game if

$$V\left(x,\alpha,\pi^1,\pi^2_*\right) \le V\left(x,\alpha,\pi^1_*,\pi^2_*\right) \le V\left(x,\alpha,\pi^1_*,\pi^2\right) \tag{4.33}$$

for all $(x,\alpha) \in \mathbb{R}^n \times E$ and $(\pi^1,\pi^2) \in \Pi^1 \times \Pi^2$.

Following the arguments of [51, Proposition 2.2.3] or [53, Proposition 3.6] we can use (4.31) to see that $V(\cdot,\cdot,\pi^1,\pi^2)$ is in $\mathcal{B}_w(\mathbb{R}^n \times E)$ for each (π^1,π^2) in $\Pi^1 \times \Pi^2$.

The lower and upper value functions L and U, respectively, in $\mathbb{R}^n \times E$ are defined similarly to (4.4)–(4.5), with (x,α) instead of x. If these function are equal, we denote such equality as $V(x,\alpha)$ for all (x,α) in $\mathbb{R}^n \times E$.

Definition 4.18. *We say that a function v and a pair of strategies $(\pi^1,\pi^2) \in \Pi^1 \times \Pi^2$ verify the random discount Bellman equations if*

$$\begin{aligned}
\alpha_i v(x,\alpha_i) &= r\left(x,\alpha_i,\pi^1,\pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} v(x,\alpha_i) && (4.34)\\
&= \sup_{\varphi\in V^1}\left\{ r\left(x,\alpha_i,\varphi,\pi^2\right) + \mathbb{L}^{\varphi,\pi^2} v(x,\alpha_i)\right\} && (4.35)\\
&= \inf_{\psi\in V^2}\left\{ r\left(x,\alpha_i,\pi^1,\psi\right) + \mathbb{L}^{\pi^1,\psi} v(x,\alpha_i)\right\} && (4.36)
\end{aligned}$$

for all $(x,\alpha_i) \in \mathbb{R}^n \times E$ and $t \ge 0$.

Interchange of limits

For every $(x,\alpha_i) \in \mathbb{R}^n \times E$, $(\varphi,\psi) \in V^1 \times V^2$, $h^j(\cdot) \equiv h(\cdot,\alpha_j) \in \mathcal{C}^2(\mathbb{R}^n)$ for $j \ne i$, define

$$\hat{b}\left(x,\alpha_i,\varphi,\psi,h^1,...,h^N\right) := \left\langle \nabla h^i(x), b\left(x,\alpha_i,\varphi,\psi\right)\right\rangle - \alpha_i h^i(x) + r\left(x,\alpha_i,\varphi,\psi\right) + \sum_{j=1}^{N} q_{ij} h^j(x), \tag{4.37}$$

with b as in Assumption 4.11(a), r as in Assumption 4.15, and i such that $h^i \equiv h(\cdot,\alpha_i)$. We also define

$$\hat{\mathbb{L}}\left(x,\alpha_i,h^1,...,h^N\right) := \sup_{\varphi\in V^1}\inf_{\psi\in V^2} \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1,...,h^N\right) + \frac{1}{2}\mathrm{Tr}\left[[\mathbb{H}h^i(x)]\cdot a(x,\alpha_i)\right].$$

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, as in Chapter 3. Consider also its closure $\bar{\Omega}$.

The following is an extension of Theorem 3.4.

Theorem 4.19. *Let Assumptions 4.11 and 4.15 hold. In addition, assume that there exist sequences $\left\{h^j_m\right\} \subset \mathcal{W}^{2,p}(\Omega)$, $j = 1,...,N$ and $\{\xi_m\} \subset \mathcal{L}^p(\Omega)$, with $p > 1$, satisfying that:*

- *(a) $\hat{\mathbb{L}}\left(x,\alpha_i,h^1_m,...,h^N_m\right) = \xi_m$ in Ω for $m = 1,2,...$, and $\alpha_i \in E$.*
- *(b) There exists a constant M^j_1 such that $\left\|h^j_m\right\|_{\mathcal{W}^{2,p}(\Omega)} \le M^j_1$ for $m = 1,2,...$ and $j = 1,...,N$.*
- *(c) ξ_m converges in $\mathcal{L}^p(\Omega)$ to some function ξ.*

Then:

- *(i) For each $j = 1,...,N$, there exist a function $h^j \in \mathcal{W}^{2,p}(\Omega)$ and a subsequence $\left\{m^j_k\right\} \subset \{1,2,...\}$ such that $h^j_{m^j_k} \to h^j$ as $k \to \infty$ strongly in $\mathcal{W}^{1,p}(\Omega)$, and weakly in $\mathcal{W}^{2,p}(\Omega)$. Moreover,*

$$\hat{\mathbb{L}}\left(\cdot,\cdot,h^1,...,h^N\right) = \xi \;\; in \;\; \Omega \times E. \tag{4.38}$$

- *(ii) If $p > n$, then $h^j_{m^j_k} \to h^j$ in the norm of $\mathcal{C}^{0,\eta}(\bar{\Omega})$ for $\eta < 1 - \frac{n}{p}$ and $j = 1,...,N$. If, in addition, ξ is in $\mathcal{C}^{0,\beta}(\Omega)$, with $\beta \le \eta$, then h^j belongs to $\mathcal{C}^{2,\beta}(\Omega)$.*

A sketch of the proof is provided in Section B.3.

Existence of value and saddle points

We establish now extensions of the main results of Sections 4.1.2 and 4.1.3.

First we give an analogue of Proposition 4.7.

Proposition 4.20. *Fix $p > n$, and (π^1, π^2) in $\Pi^1 \times \Pi^2$. Then there exists a function v in $\mathcal{C}^2(\mathbb{R}^n \times E) \cap \mathcal{B}_w(\mathbb{R}^n \times E)$ that satisfies*

$$\alpha_i v(x, \alpha_i) = r\left(x, \alpha_i, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} v(x, \alpha_i) \tag{4.39}$$

for all $(x, \alpha_i) \in \mathbb{R}^n$.

The proof of Proposition 4.20 resembles that of Proposition 4.7. The difference lies in the fact that one should replace x by (x, α_i), and $h^{\pi^1,\pi^2}_{\alpha,R}(x)$ by

$$h^{\pi^1,\pi^2}_R(x, \alpha_i) := \mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left[\int_0^{\tau^{\pi^1,\pi^2}_R} \exp(-S_t) r\left(x(t), \alpha(t), \pi^1, \pi^2\right) dt\right],$$

with S_t as in (4.28).

Theorem 4.21. *If Assumptions 4.11, 4.13, and 4.15 hold, then there exist a function $v \in \mathcal{C}^2(\mathbb{R}^n \times E) \cap \mathcal{B}_w(\mathbb{R}^n \times E)$ and a pair of strategies (π^1, π^2) in $\Pi^1 \times \Pi^2$ that satisfy the random discount HJB equations (4.34)–(4.36). Moreover, such a function v coincides with the value function V of the random discounted game, and (π^1, π^2) is a corresponding saddle point.*

The proof of this result is very much alike those of Theorems 4.9 and 4.10 (except one should replace x by (x, α_i) and put the space $\mathcal{C}^2(\mathbb{R}^n \times E) \cap \mathcal{B}_w(\mathbb{R}^n \times E)$ in lieu of $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$). We should note that the proof of Theorem 4.9 quotes Theorem 3.4, and that Theorem 4.10 uses Proposition 4.6. We should replace these with proper invokations of Theorem 4.19 and the following result.

Proposition 4.22. *Fix (π^1, π^2) in $\Pi^1 \times \Pi^2$. If a function $v \in \mathcal{C}^2(\mathbb{R}^n \times E) \cap \mathcal{B}_w(\mathbb{R}^n \times E)$ satisfies (4.39) for all $(x, \alpha_i) \in \mathbb{R}^n \times E$, then*

$$v(x, \alpha_i) = V\left(x, \alpha_i, \pi^1, \pi^2\right), \tag{4.40}$$

where V is the random discounted payoff criterion defined in (4.27).

Moreover, if the equality in (4.39) is replaced with "$\leq$" or "$\geq$", then (4.40) holds with the corresponding inequality.

Proof. Fix $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. Observe that Dynkin's formula for Markov modulated diffusions (see [67, Theorem 1.45 and Lemma 1.9]) yields

$$\begin{aligned}
\mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left(\exp\{-S_T\} v(x(T), \alpha(T))\right) &= v(x, \alpha_i) + \mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left[\int_0^T \exp\{-S_t\}\left(-\alpha(t) v(x(t), \alpha(t)) + \mathbb{L}^{\pi^1,\pi^2} v(x(t), \alpha(t))\right) dt\right] \\
&= v(x, \alpha_i) - \mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left[\int_0^T \exp\{-S_t\} r\left(x(t), \alpha(t), \pi^1, \pi^2\right) dt\right] \quad \text{(by (4.39))}.
\end{aligned}$$

Define

$$\alpha_* := \min_{j=1,\ldots,N} \alpha_j.$$

Since v is in $\mathcal{B}_w(\mathbb{R}^n \times E)$,

$$\begin{aligned}
\left|\mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left(\exp\{-S_T\} v(x(T), \alpha(T))\right)\right| &\leq \mathbb{E}^{\pi^1,\pi^2}_{x,\alpha_i}\left(e^{-\alpha_* T}\|v\|_w\, w(x(T), \alpha(T))\right) \\
&\leq e^{-\alpha_* T}\|v\|_w\left[e^{-cT} w(x, \alpha) + \frac{d}{c}\left(1 - e^{-cT}\right)\right] \quad \text{(by (4.31))} \\
&\to 0 \text{ as } T \to \infty.
\end{aligned}$$

This implies that

$$\begin{aligned} v(x,\alpha_i) &= \mathbb{E}_{x,\alpha_i}^{\pi^1,\pi^2}\left[\int_0^\infty \exp\{-S_t\} r\left(x(t),\alpha_i,\pi^1,\pi^2\right) dt\right] \\ &= V\left(x,\alpha_i,\pi^1,\pi^2\right). \end{aligned}$$

Similarly, the proof of the second statement uses the same arguments by replacing the equality in (4.39) by either "$\leq$" or "$\geq$". □

4.3 Concluding remarks

This chapter was intended to characterize saddle points for a general class of SDGs with discounted payoffs. We studied two models of zero–sum games in infinite–horizon: with a fixed discount rate and with a rate of discount driven by a Markov chain. We gave sufficient conditions for the existence of value functions and equilibria in both contexts.

Theorems 4.10 and 4.21 are typical verification results, so they yield the existence of saddle points in the infinite–horizon games we studied.

As for the value of the infinite–horizon game, the uniform ellipticity condition in Assumption 2.1(c) on the diffusion (2.1), along with Theorem 3.4 provided us with a powerful tool. Proposition 4.7 was proved thanks to Theorem 3.4 without making explicit use of semigroup theory. Moreover, this theorem has a broad range of applications in, for instance, the vanishing discount technique for proving the existence of equilibria in SDGs with ergodic payoff and the policy iteration algorithm for finding those equilibria. The following chapters are devoted to develop some of these applications.

We provided an analogous version of Theorem 3.4 for the context of the random discounted game in Section 4.2. The key of Theorem 4.19 is the assumption that the diffusion (2.1) is replaced by the Markov–modulated dynamic (4.25)–(4.26). The main changes we had to include in our model for this problem (with respect to Section 4.1) were the substitution of $x \in \mathbb{R}^n$ by $(x,\alpha) \in \mathbb{R}^n \times E$, and of the generator displayed in (2.3), by that of (4.29). Moreover, Proposition 4.22 gave us that the Bellman equation associated with this game is of the form (4.5). However, this fact was to be expected, since the controlled discrete–time version of the problem studied in [32] and [33] presents the same feature.

Chapter 5

The vanishing discount technique for zero–sum SDGs with ergodic payoff

This chapter is devoted to the study of ergodic zero–sum SDGs. This type of games has been studied, for instance, in [9, 18, 63]. Our main goal is to look for saddle points in the sense of (5.10) below, when the payoff function for each player is given by (5.2). To ensure the validity of several results here, we will use the uniform *w*–exponential ergodicity referred to in Theorem 2.10.

We will use the *vanishing discount technique* to study the connection between some Bellman–type equations arising of a discounted game and the saddle points of an ergodic game. This approach is one of the most common methods to deal with the average payoff criterion. It is so–named because it is based on the convergence of certain sequence of discounted problems as in Chapter 4 (indexed by a discount rate α) as the discount rate *vanishes*, i.e. $\alpha \downarrow 0$.

5.1 Average optimality

Recall that U^1 and U^2 are compact subsets of given vector normed spaces (see Assumption 2.1(d) and Section 2.1). By the results in the beginning of Section 2.1, V^1 and V^2 are also compact spaces. Furthermore, consider the family $\Pi^1 \times \Pi^2$ of all pairs of stationary randomized Markov strategies for players 1 and 2 (see Definition 2.4). Finally, let R be a positive real number and let B_R and $\bar{B}_R$ be as in (2.20).

For each $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$ and $T \geq 0$, let

$$J_T\left(x, \pi^1, \pi^2\right) := \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^T r\left(x(t), \pi^1, \pi^2\right) dt\right] \tag{5.1}$$

be the *total expected payoff* of $\left(\pi^1, \pi^2\right)$ over the time interval $[0, T]$, when the initial state is $x \in \mathbb{R}^n$. The *ergodic payoff* (also known as long–run average payoff) given the initial state x is given by

$$J\left(x, \pi^1, \pi^2\right) := \limsup_{T \to \infty} \frac{1}{T} J_T(x, \pi^1, \pi^2). \tag{5.2}$$

Proposition 5.1. *Let Assumptions 2.1, 2.6, 2.9 and 2.12 hold. Then the payoff rate* r *is* μ_{π^1,π^2}*–integrable for every pair* $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$

Proof. Given $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$, define

$$J(\pi^1, \pi^2) := \mu_{\pi^1,\pi^2}\left(r\left(\cdot, \pi^1, \pi^2\right)\right) = \int_{\mathbb{R}^n} r\left(x, \pi^1, \pi^2\right) \mu_{\pi^1,\pi^2}(dx). \tag{5.3}$$

with μ_{π^1,π^2} as in (2.7).

By the definition of $J(\pi^1,\pi^2)$ in (5.3), Assumption 2.12(b), (2.7) and (2.10) yield

$$\left|J\left(\pi^1,\pi^2\right)\right| \le \int_{\mathbb{R}^n} \left|r\left(x,\pi^1,\pi^2\right)\right| \mu_{\pi^1,\pi^2}(dx) \le M\cdot\mu_{\pi^1,\pi^2}(w) < \infty \tag{5.4}$$

for all $\left(\pi^1,\pi^2\right)\in\Pi^1\times\Pi^2$. In fact, by (2.19),

$$\left|J\left(\pi^1,\pi^2\right)\right| \le M\cdot\mu_{\pi^1,\pi^2}(w) \le M\cdot\frac{d}{c}, \tag{5.5}$$

so that $J\left(\pi^1,\pi^2\right)$ is uniformly bounded on $\Pi^1\times\Pi^2$. This yields the desired result. □

It follows from (2.13) that the average payoff (5.2) coincides with the constant $J\left(\pi^1,\pi^2\right)$ in (5.3) for every $\left(\pi^1,\pi^2\right)\in\Pi^1\times\Pi^2$. Indeed, note that J_T defined in (5.1) can be expressed as

$$J_T\left(x,\pi^1,\pi^2\right) = T\cdot J\left(\pi^1,\pi^2\right) + \int_0^T \left[\mathbb{E}_x^{\pi^1,\pi^2} r\left(x(t),\pi^1,\pi^2\right) - J\left(\pi^1,\pi^2\right)\right] dt.$$

Hence, multiplying the latter equality by $\frac{1}{T}$ and letting $T\to\infty$, by (2.13), we obtain,

$$J\left(x,\pi^1,\pi^2\right) = \limsup_{T\to\infty}\frac{1}{T}J_T\left(x,\pi^1,\pi^2\right) = J\left(\pi^1,\pi^2\right) \quad \text{for all } x\in\mathbb{R}^n. \tag{5.6}$$

By virtue of this last expression, we can write (5.2) simply as $J\left(\pi^1,\pi^2\right)$.

We now define the constant values

$$\mathcal{U} := \inf_{\pi^2\in\Pi^2}\sup_{\pi^1\in\Pi^1} J\left(\pi^1,\pi^2\right) \tag{5.7}$$

and

$$\mathcal{L} := \sup_{\pi^1\in\Pi^1}\inf_{\pi^2\in\Pi^2} J\left(\pi^1,\pi^2\right). \tag{5.8}$$

The function $\mathcal{L}$ is called the game's *lower value*, and $\mathcal{U}$ is the game's *upper value*. Clearly, we have $\mathcal{L}\le\mathcal{U}$. If these two numbers coincide, then the game is said to have a *value*, say $\mathcal{V}$. This number is the common value of $\mathcal{L}$ and $\mathcal{U}$, i.e.,

$$\mathcal{V} := \mathcal{L} = \mathcal{U}. \tag{5.9}$$

As a consequence of (5.6) and (5.5), $\mathcal{L}$ and $\mathcal{U}$ are finite. This implies that $\mathcal{V}$ is also finite if the second equality in (5.9) holds.

The basic problem we are concerned with is to find *average payoff equilibria* or *saddle points* of the average payoff SDG. Namely, we are interested in pairs $\left(\pi^1_*,\pi^2_*\right)\in\Pi^1\times\Pi^2$ for which

$$J\left(\pi^1,\pi^2_*\right) \le J\left(\pi^1_*,\pi^2_*\right) \le J\left(\pi^1_*,\pi^2\right) \tag{5.10}$$

for every $\left(\pi^1,\pi^2\right)\in\Pi^1\times\Pi^2$. The set of pairs of average payoff equlibria is denoted by $\left(\Pi^1\times\Pi^2\right)_{ae}$.

Remark 5.2. *Observe that if $\left(\pi^1_*,\pi^2_*\right)$ is an average payoff equilibrium, then the game has a value $J\left(\pi^1_*,\pi^2_*\right) =: \mathcal{V}$. As in the discounted payoff case, the converse is not necessarily true.*

Definition 5.3. *We say that a constant $J\in\mathbb{R}$, a function $h\in\mathcal{C}^2(\mathbb{R}^n)\cap\mathcal{B}_w(\mathbb{R}^n)$, and a pair of strategies $\left(\pi^1,\pi^2\right)\in\Pi^1\times\Pi^2$ verify the average payoff optimality equations if, for every $x\in\mathbb{R}^n$,*

$$\begin{aligned} J &= r\left(x,\pi^1,\pi^2\right) + \mathbb{L}^{\pi^1,\pi^2}h(x) && (5.11)\\ &= \sup_{\varphi\in V^1}\left\{r\left(x,\varphi,\pi^2\right) + \mathbb{L}^{\varphi,\pi^2}h(x)\right\} && (5.12)\\ &= \inf_{\psi\in V^2}\left\{r\left(x,\pi^1,\psi\right) + \mathbb{L}^{\pi^1,\psi}h(x)\right\}. && (5.13)\end{aligned}$$

In this case, the pair of strategies $\left(\pi^1,\pi^2\right)\in\Pi^1\times\Pi^2$ satisfying (5.11)–(5.13) is called a canonical equilibrium.

The following result holds by virtue of Theorem A.1. It is the ergodic version of Proposition 4.6 in the discounted payoff case. We omit the proof because it is immediate.

Proposition 5.4. *If there is a constant* J*, a function* h *in* $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ *and a pair* (π^1, π^2) *in* $\Pi^1 \times \Pi^2$ *such that*

$$J \geq r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} h(x) \text{ for all } x \in \mathbb{R}^n, \tag{5.14}$$

then

$$J \geq J\left(\pi^1, \pi^2\right). \tag{5.15}$$

Similarly, if the inequality (5.14) *is replaced by "*$\leq$*", then* (5.15) *should be replaced by the same inequality, i.e., if*

$$J \leq r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} h(x), \text{ then } J \leq J\left(\pi^1, \pi^2\right).$$

Therefore, if the equality holds in (5.14)*, then we have* $J = J\left(\pi^1, \pi^2\right)$*.*

Our next result ensures the existence of solutions to equations (5.11)–(5.13). Furthermore, it relates some components of these equations with the properties of the game in (5.9) and (5.10).

Theorem 5.5. *If Assumptions 2.1, 2.6, 2.9, and 2.12 hold, then:*

(i) There exist a solution $(J, h) \in \mathbb{R} \times \left(\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)\right)$*, and a pair* $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ *such that the average payoff optimality equations* (5.11)–(5.13) *are satisfied. Moreover, the constant* J *equals* $\mathcal{V}$*, the value of the game, and the function* h *is unique up to additive constants, under the extra condition that* $h(0) = 0$*.*

(ii) A pair of strategies is an average payoff equilibrium if, and only if, it is canonical.

The proof we will offer is based on an extension of the vanishing discount technique for control problems (cf. [8, Chapter II], [17, Corollary 6.2], and [71]). We present such extension in the following section.

5.2 The vanishing discount technique

We will prove the existence of solutions to the average payoff optimality equations (5.11)–(5.13) using the so–called vanishing discount approach. The idea is to impose conditions on an associated α–discounted payoff game in such a way that, when $\alpha \downarrow 0$, we obtain equations (5.11)–(5.13).

To this end, recall the payoff rate r given in Assumption 2.12, and let V_α be the *expected* α*–discounted payoff* defined in (4.1), that is

$$V_\alpha\left(x, \pi^1, \pi^2\right) := \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^\infty e^{-\alpha t} r\left(x(t), \pi^1, \pi^2\right) dt\right]. \tag{5.16}$$

In Theorem 4.9 we showed that, under Assumptions 2.1, 2.6 and 2.12, there exist a function v_α in $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$, and a pair (π^1, π^2) in $\Pi^1 \times \Pi^2$ that satisfy (4.8)–(4.10), i.e.,

$$\begin{aligned}
\alpha v_\alpha(x) &= r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} v_\alpha(x) && (5.17)\\
&= \sup_{\varphi \in V^1} \left\{ r\left(x, \varphi, \pi^2\right) + \mathbb{L}^{\varphi,\pi^2} v_\alpha(x) \right\} && (5.18)\\
&= \inf_{\psi \in V^2} \left\{ r\left(x, \pi^1, \psi\right) + \mathbb{L}^{\pi^1,\psi} v_\alpha(x) \right\}, && (5.19)
\end{aligned}$$

for all $x \in \mathbb{R}^n$. These results and Isaacs' condition of Remark 2.15 give that

$$\begin{aligned}
v_\alpha(x) &:= \inf_{\psi \in V^2} \sup_{\varphi \in V^1} V_\alpha(x, \varphi, \psi) && (5.20)\\
&= \sup_{\varphi \in V^1} \inf_{\psi \in V^2} V_\alpha(x, \varphi, \psi). && (5.21)
\end{aligned}$$

Passage to the limit as $\alpha \downarrow 0$:

We characterize the classical solution $(J, h) \in \mathbb{R} \times \left(\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)\right)$ of (5.11)–(5.13) as the limit, as $\alpha \downarrow 0$, of v_α in (5.17)–(5.19).

Theorem 5.6. *For each $\alpha > 0$, let $h_\alpha(x) := v_\alpha(x) - v_\alpha(0)$, where $v_\alpha(\cdot)$ satisfies* (5.17)–(5.19). *If Assumptions 2.1, 2.6, 2.9 and 2.12 hold, then there exists a constant J, a function $h \in \mathcal{C}^2(\mathbb{R}^n) \times \mathcal{B}_w(\mathbb{R}^n)$, and a sequence $\alpha_m \downarrow 0$ such that*

$$\alpha_m v_{\alpha_m}(0) \to J \tag{5.22}$$

and, for all $x \in \mathbb{R}^n$,

$$h_{\alpha_m}(x) \quad \to \quad h(x), \tag{5.23}$$

$$\alpha_m h_{\alpha_m}(x) \quad \to \quad 0. \tag{5.24}$$

In addition, the limit (J, h) satisfies (5.11)–(5.13).

The proof of Theorem 5.6 is based on Theorem 3.4. Hence we devote the following lines to the verification of the hypotheses of such result.

Recall from Chapter 3 that Ω is a bounded, open and connected subset of $\mathbb{R}^n$. To invoke Theorem 3.4, we need to ensure the existence of $\{h_m\} \subset \mathcal{W}^{2,p}(\Omega)$ and $\{\xi_m\} \subset \mathcal{L}^p(\Omega)$, with $p > 1$, and a sequence $\{\alpha_m\}$ of positive numbers satisfying that:

(a) For $m = 1, 2, ...,$

$$\begin{aligned} \xi_m(x) &= \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi,\psi} h_m(x) \right\} - \alpha_m h_m(x) \\ &= \inf_{\psi \in V^2} \sup_{\varphi \in V^1} \left\{ r(x, \varphi, \psi,) + \mathbb{L}^{\varphi,\psi} h_m(x) \right\} - \alpha_m h_m(x) \end{aligned}$$

for all x in Ω.

(b) There exists a constant M_1 such that $\|h_m\|_{\mathcal{W}^{2,p}(\Omega)} \leq M_1$ for $m = 1, 2, ...$

(c) ξ_m converges in $\mathcal{L}^p(\Omega)$ to some function ξ

(d) α_m converges to some α.

To this end, let $\alpha_m > 0$ be a sequence of positive numbers such that $\alpha_m \downarrow 0$ as $m \to \infty$. Define $h_{\alpha_m}(x) := v_{\alpha_m}(x) - v_{\alpha_m}(0)$ for each $m = 1, 2, ...$ as in Theorem 5.6. A direct calculation yields that $v_{\alpha_m}(x) = h_{\alpha_m}(x) + v_{\alpha_m}(0)$ satisfies (5.17)–(5.19), i.e.,

$$\alpha_m v_{\alpha_m}(0) + \alpha_m h_{\alpha_m}(x) = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} h_{\alpha_m}(x) \tag{5.25}$$

$$= \sup_{\varphi \in V^1} \left\{ r\left(x, \varphi, \pi^2\right) + \mathbb{L}^{\varphi,\pi^2} h_{\alpha_m}(x) \right\} \tag{5.26}$$

$$= \inf_{\psi \in V^2} \left\{ r\left(x, \pi^1, \psi\right) + \mathbb{L}^{\pi^1,\psi} h_{\alpha_m}(x) \right\} \tag{5.27}$$

for all $x \in \mathbb{R}^n$ and all $\alpha_m > 0$.

Verification of the hypotheses of Theorem 3.4:

(a) Define the constant functions

$$\xi_m(x) := \alpha_m v_{\alpha_m}(0) \ \text{ for all } \ x \in \mathbb{R}^n. \tag{5.28}$$

Replacing these in (5.25), we can see that hypothesis (a) of Theorem 3.4 holds.

(b) Fix an arbitrary $R > 0$, and let $\overline{B}_R$ be as in (2.20). Then, by Theorem A.3, there exists a constant C_0 such that, for fixed $p > n$,

$$\|h_{\alpha_m}\|_{\mathcal{W}^{2,p}(B_R)} \leq C_0 \left(\|h_{\alpha_m}\|_{L^p(B_{2R})} + \left\| r\left(\cdot, \pi^1, \pi^2\right) \right\|_{L^p(B_{2R})} + |\alpha_m v_{\alpha_m}(0)| \right). \tag{5.29}$$

Note now that, by (5.16), for every $\left(\pi^1, \pi^2\right)$ in $\Pi^1 \times \Pi^2$ and $\alpha > 0$, we get

$$\begin{aligned}
\left|V_\alpha\left(x, \pi^1, \pi^2\right) - V_\alpha\left(0, \pi^1, \pi^2\right)\right| &\leq \int_0^\infty e^{-\alpha t} \left| \mathbb{E}_x^{\pi^1,\pi^2} r\left(x(t), \pi^1, \pi^2\right) - \mathbb{E}_0^{\pi^1,\pi^2} r\left(x(t), \pi^1, \pi^2\right) \right| dt \\
&\leq \int_0^\infty e^{-\alpha t} \left| \mathbb{E}_x^{\pi^1,\pi^2} r\left(x(t), \pi^1, \pi^2\right) - J\left(\pi^1, \pi^2\right) \right| dt \\
&\quad + \int_0^\infty e^{-\alpha t} \left| \mathbb{E}_0^{\pi^1,\pi^2} r\left(x(t), \pi^1, \pi^2\right) - J\left(\pi^1, \pi^2\right) \right| dt \\
&\leq \int_0^\infty CMe^{-(\alpha+\delta)t} \left(w(x) + w(0)\right) dt \quad \text{(by (2.13))}
\end{aligned}$$

with M as in Assumption 2.12(b). Let $\hat{M} := \frac{CM}{\delta}(1 + w(0))$. Since $w \geq 1$, we see that

$$\left|V_\alpha\left(x, \pi^1, \pi^2\right) - V_\alpha\left(0, \pi^1, \pi^2\right)\right| \leq \hat{M} w(x).$$

Hence, since $\hat{M}$ is independent of $\left(\pi^1, \pi^2\right)$, it follows from (5.20)–(5.21) that

$$|h_{\alpha_m}(x)| \leq \sup_{(\pi^1,\pi^2)\in\Pi^1\times\Pi^2} \left|V_{\alpha_m}\left(x, \pi^1, \pi^2\right) - V_{\alpha_m}\left(0, \pi^1, \pi^2\right)\right|,$$

and so

$$|h_{\alpha_m}(x)| \leq \hat{M} w(x) \text{ for all } x \in \mathbb{R}^n. \tag{5.30}$$

Furthermore, the relation (4.3) implies that the sequence in (5.28) is bounded by a positive number, say ρ. Combine (5.29), (5.30) and Assumption 2.12(b) to obtain that, independently of our choice for α_m,

$$\begin{aligned}
\|h_{\alpha_m}\|_{\mathcal{W}^{2,p}(B_R)} &\leq C_0 \left(\hat{M} \|w\|_{L^p(B_{2R})} + M \|w\|_{L^p(B_{2R})} + \rho \right) \\
&\leq C_0 \left(\hat{M} + M\right) |\bar{B}_{2R}|^{1/p} \max_{x\in\bar{B}_{2R}} w(x) + \rho C_0,
\end{aligned} \tag{5.31}$$

where $\left|\bar{B}_{2R}\right|$ denotes the volume of the closed ball $\bar{B}_{2R}$ with radius $2R$. Hence, by (5.31), hypothesis (b) of Theorem 3.4 holds.

(c) Since $\alpha_m v_m(0)$ in (5.28) is bounded, there exist a number J and a subsequence of $\{\alpha_m\}$ (again denoted as $\{\alpha_m\}$), such that (5.22) holds. Hence hypothesis (c) of Theorem 3.4 follows.

(d) Since $\alpha \downarrow 0$, hypothesis (d) trivially holds.

Proof of Theorem 5.6. Since hypotheses (a)–(d) of Theorem 3.4 hold, we invoke that result to assert the existence of a function h in the class $\mathcal{W}^{2,p}(B_R)$ such that h_{α_m} (or a subsequence thereof) converges to h in B_R. In fact, we can use (5.31) again, along with the compactness of the embedding $\mathcal{W}^{2,p}(B_R) \hookrightarrow \mathcal{C}^{0,\eta}(\bar{B}_R)$ for $\eta < 1$ and $p > n$, as well as Arzelà–Ascoli's Theorem, to ensure that the convergence $h_{\alpha_m} \to h$ is uniform on any bounded, open and connected subset $B_R \subset B_{2R}$, and that h actually belongs to $\mathcal{C}^{2,\beta}(B_R)$ for all $0 < \beta < 1$.

Also observe that, by (5.31), $\alpha_m h_{\alpha_m}(x) \to 0$ in $\mathcal{W}^{2,p}(B_R)$. Indeed, let

$$M_1 := C_0 \left[(\hat{M} + M)|\bar{B}_{2R}|^{1/p} \max_{x\in\bar{B}_{2R}} w(x) + \rho \right]$$

to see that:

$$\|\alpha_m h_{\alpha_m}\|_{\mathcal{W}^{2,p}(B_R)} = \alpha_m \|h_{\alpha_m}\|_{\mathcal{W}^{2,p}(B_R)}$$

$$\begin{aligned} &\leq \alpha_m M_1 \\ &\to 0 \text{ as } \alpha \downarrow 0. \end{aligned}$$

This proves (5.24).

To prove the last part of Theorem 5.6, we apply Theorem 3.4 to obtain

$$J = \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi, \psi} h(x) \right\} \tag{5.32}$$

$$= \inf_{\psi \in V^2} \sup_{\varphi \in V^1} \left\{ r(x, \varphi, \psi) + \mathbb{L}^{\varphi, \psi} h(x) \right\} \tag{5.33}$$

for all $x \in B_R$.

Since the choice of $R > 0$ was arbitrary, we can extend the convergence $h_{\alpha_m} \to h$ to all of $\mathbb{R}^n$ with h satisfying (5.32)–(5.33). Actually, by (5.30), we can ensure that h_{α_m} is in $\mathcal{B}_w(\mathbb{R}^n)$. Now, the uniform convergence of h_{α_m} to h on bounded, open and connected subsets of $\mathbb{R}^n$ and the use of Lemma A.4, yield that h belongs to $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$.

Finally, the existence of a pair $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ such that equations (5.11)–(5.13) are satisfied, is given by (5.32)–(5.33) and by Theorem A.2. □

5.3 Proof of Theorem 5.5.

(*i*) The proof of the existence of a constant J and a function h such that (5.11)–(5.13) hold was given in Theorem 5.6.

On the other hand, Propositions 4.2 and 4.3 in [50] combined with (5.11)–(5.13) and Proposition 5.4 yield that $J = \mathcal{V}$. Further, the proof that h is unique up to additive constants requires us to note that if h satisfies (5.11)–(5.13), then so does $h + k$, with k constant, because $\mathbb{L}^{u_1, u_2}$ is a differential operator (see (2.3)). To prove the uniqueness of solutions to equations (5.11)–(5.13), let us suppose that (J, h_1) and (J, h_2) are two solutions in $\mathbb{R} \times \left(\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)\right)$ of (5.11)–(5.13), that is

$$\begin{aligned} J &= r(x, \pi^1, \pi^2) + \mathbb{L}^{\pi^1, \pi^2} h_1(x), \\ J &= r(x, \pi^1, \pi^2) + \mathbb{L}^{\pi^1, \pi^2} h_2(x). \end{aligned}$$

The substraction of these two equalities yields that $\mathbb{L}^{\pi^1, \pi^2} \eta(x) = 0$, with $\eta(\cdot) := h_1(\cdot) - h_2(\cdot)$. Hence, by Lemma 2.11,

$$\begin{aligned} \eta(x) &= \mu_{\pi^1, \pi^2}(\eta) \text{ for all } x \in \mathbb{R}^n \\ &= \mu_{\pi^1, \pi^2}(h_1 - h_2). \end{aligned}$$

But $\mu_{\pi^1, \pi^2}(h_1 - h_2)$ must be zero, since $\eta(0) = h_1(0) - h_2(0) = 0$. This gives $h_1 \equiv h_2$.

(*ii*) The *only if part*. We use the same arguments in the proof of [17, Corollary 6.2]. Suppose that (π^1_*, π^2_*) is an average equilibrium that is not canonical. Then, either (5.12) or (5.13) does not hold. Assume that, say, (5.12) is not satisfied. Then, by the continuity of $x \to r(x, \pi^1, \pi^2) + \mathbb{L}^{\pi^1, \pi^2} h(x)$, there exists a constant $\epsilon > 0$ and a Borel set $B \subset \mathbb{R}^n$, with $\lambda(B) > 0$ (recall that λ stands for the Lebesgue measure on $\mathbb{R}^n$) such that

$$J \geq r(x, \pi^1_*, \pi^2_*) + \mathbb{L}^{\pi^1_*, \pi^2_*} h(x) + \epsilon \chi_B(x) \text{ for } x \in \mathbb{R}^n, \tag{5.34}$$

where $\chi_B(\cdot)$ stands for the indicator function of B. Combining Theorem A.1 and (5.34) we obtain, for all $t \geq 0$,

$$\mathbb{E}_x^{\pi^1_*, \pi^2_*} h(x(t)) - h(x) \leq Jt - \mathbb{E}_x^{\pi^1_*, \pi^2_*} \left(\int_0^t r(x(s), \pi^1_*, \pi^2_*) \, ds \right) - \epsilon \mathbb{E}_x^{\pi^1_*, \pi^2_*} \left(\int_0^t \chi_B(x(s)) \, ds \right). \tag{5.35}$$

Multiplying by t^{-1} and letting $t \to \infty$ yields

$$J(\pi^1_*, \pi^2_*) + \epsilon \mu_{\pi^1_*, \pi^2_*}(B) \leq J \text{ (by (2.13) and (2.15))}. \tag{5.36}$$

Moreover, by [3, Theorem 4.3], $\mu_{\pi^1_*,\pi^2_*}$ is equivalent to the Lebesgue measure λ. Hence $\lambda(B) > 0$ yields $\mu_{\pi^1_*,\pi^2_*}(B) > 0$, and thus $J\left(\pi^1_*, \pi^2_*\right) < J = \mathcal{V}$, which contradicts the equilibrium property of $\left(\pi^1_*, \pi^2_*\right)$.

The if part. Suppose that $\left(\pi^1_*, \pi^2_*\right)$ satisfies the average optimality equations. Then, by (5.11) and Proposition 5.4 we obtain that $\left(\pi^1_*, \pi^2_*\right)$ is average optimal. □

5.4 Concluding remarks

This chapter introduces the average payoff criterion for SDGs. This criterion is the basis for the developments to come in our work, such as the so–named policy iteration algorithm and the obtention of bias and overtaking equilibria. A central hypothesis for our developments is the uniform w–exponential ergodicity condition (2.13). The main result is Theorem 5.5, because it gives us elements to study ergodic payoff games as a limit of discounted payoff problems. A key to this fact is Theorem 3.4.

Chapter 6

Policy iteration for zero–sum SDGs with ergodic payoff

The results in Chapter 5 (specially Theorem 5.5) ensure the existence of the pairs (J, h) in $\mathbb{R} \times \left(\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)\right)$ and $\left(\pi^1, \pi^2\right)$ in $\Pi^1 \times \Pi^2$ to the average optimality equations (5.11)–(5.13). However, the question now is: how can we find (or at least approximate) values of $J(= \mathcal{V})$, h, and $\left(\pi^1, \pi^2\right)$? Providing an answer to this and other related questions is the goal of this chapter. Our aim is to give conditions under which a certain algorithm, in our case, the *policy iteration algorithm* (PIA) produces convergent sequences of values and policies for a SDG with ergodic payoff.

The PIA was used by Fleming [26] to study some finite horizon controlled diffusion problems in 1963. Fleming himself attributed the PIA to Bellman. Now, since Howard [49] determined optimal policies for processes in infinite–horizon by proposing a solution based upon successive approximation in a policy space, some authors know the PIA as *Howard's algorithm*. It was studied later by Bismut [13] and Puterman [81, 82], who found its convergence rate for controlled diffusions in compact regions of $\mathbb{R}^n$. Arapostathis [5] studied a version of the PIA also for controlled diffusions. For discrete–time zero–sum games, Van der Wal [94] presented a convergent version of the PIA under the assumption that the state space and the action space are both finite. The goal of the PIA for a SDG is to generate sequences of strategies and value functions that converge to the equilibrium and value function of the SDG.

The algorithm we present resembles that introduced in [46] for controlled Markov decision processes in Borel spaces and is inspired in the Hoffman–Karp [48] version presented in [94]. In our algorithm, we propose to fix the action of one of the players to find the other player's *best* action, thus reducing the game in that stage to a Markov control process. Then, we find the *current* value of the game and we move on to the next iteration, where we fix the other player's *best* action.

The set of assumptions we used in Chapter 5 ensures the convergence of the PIA to a saddle point of the zero–sum SDG with ergodic payoff. To prove this, we will use again Theorem 3.4 and, for a given pair of strategies, we will use the concept of its bias from the game's value (see equation (6.6) below).

Throughout this chapter we will consider that Assumptions 2.1, 2.6, 2.9 and 2.12 hold.

6.1 The policy iteration algorithm

We now introduce the PIA, also known as policy improvement algorithm. The version we present in this section was inspired by the results in [46, 94].

The PIA:

Step 1. Set $m = 0$. Select a strategy $\pi_0^2 \in \Pi^2$, and define $J\left(\pi_{-1}^1, \pi_{-1}^2\right) := -\infty$.

Step 2. Find a policy $\pi_m^1 \in \Pi^1$, a constant $J\left(\pi_m^1, \pi_m^2\right)$, and a function $h_m : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ such that $\left(J\left(\pi_m^1, \pi_m^2\right), h_m\right)$ is a solution of (6.1)–(6.2):

$$J\left(\pi_m^1, \pi_m^2\right) = \sup_{\varphi \in V^1} \left[r\left(x, \varphi, \pi_m^2\right) + \mathbb{L}^{\varphi, \pi_m^2} h_m(x)\right] \tag{6.1}$$

$$= r\left(x, \pi_m^1, \pi_m^2\right) + \mathbb{L}^{\pi_m^1, \pi_m^2} h_m(x) \text{ for all } x \in \mathbb{R}^n. \tag{6.2}$$

Observe that

$$J\left(\pi_m^1, \pi_m^2\right) \geq \inf_{\psi \in V^2} \left[r\left(x, \pi_m^1, \psi\right) + \mathbb{L}^{\pi_m^1, \psi} h_m(x)\right] \text{ for all } x \in \mathbb{R}^n. \tag{6.3}$$

Step 3. If $J\left(\pi_m^1, \pi_m^2\right) = J\left(\pi_{m-1}^1, \pi_{m-1}^2\right)$, then $J\left(\pi_m^1, \pi_m^2\right)$ is the value of the game and $\left(\pi_m^1, \pi_m^2\right)$ is a saddle point. Terminate PIA. Otherwise, go to step 4.

Step 4. Determine a strategy $\pi_{m+1}^2 \in \Pi^2$ that attains the minimum on the right hand side of (6.3), i.e., for all $x \in \mathbb{R}^n$

$$r\left(x, \pi_m^1, \pi_{m+1}^2\right) + \mathbb{L}^{\pi_m^1, \pi_{m+1}^2} h_m(x) = \inf_{\psi \in V^2} \left[r\left(x, \pi_m^1, \psi\right) + \mathbb{L}^{\pi_m^1, \psi} h_m(x)\right]. \tag{6.4}$$

Increase m in 1 and go back to step 2.

Remark 6.1. *Observe that Remark 2.15 makes us indifferent between using the PIA version we have proposed, and using a modification that minimizes in* (6.2) *in step 2, and maximizes in* (6.4) *in step 4.*

Definition 6.2. *The PIA is said to converge if the sequence* $J\left(\pi_m^1, \pi_m^2\right)$ *converges to the value of the game defined in* (5.10). *That is,*

$$J\left(\pi_m^1, \pi_m^2\right) \to \mathcal{V}.$$

To ensure the convergence of the PIA, we need to guarantee it is well–defined. To do this, it is necessary to satisfy the following conditions.

1. For every pair $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$, there exists an invariant probability measure μ_{π^1, π^2}. This is the first consequence of Assumption 2.6.

2. For every pair $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$, the payoff rate $r\left(\cdot, \pi^1, \pi^2\right)$ is μ_{π^1, π^2}–integrable, so that (5.3) holds, that is,

$$J\left(\pi^1, \pi^2\right) := \mu_{\pi^1, \pi^2}\left(r\left(\cdot, \pi^1, \pi^2\right)\right) = \int_{\mathbb{R}^n} r\left(x, \pi^1, \pi^2\right) \mu_{\pi^1, \pi^2}(dx).$$

   This follows from Proposition 5.1.

3. For every pair $\left(\pi^1, \pi^2\right)$ there is a unique solution $\left(J\left(\pi^1, \pi^2\right), h_{\pi^1, \pi^2}\right)$ to the *Poisson equation*

$$J\left(\pi^1, \pi^2\right) = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1, \pi^2} h_{\pi^1, \pi^2}(x) \text{ for all } x \in \mathbb{R}^n, \tag{6.5}$$

   which is guaranteed by Proposition 6.4 below.

4. For each $\pi_m^2 \in \Pi^2$, there exists a strategy $\pi_m^1 \in \Pi^1$ such that (6.2) holds. This is indeed the case by virtue of Assumption 2.12, the compactness of V^1, and Theorem A.2.

5. For every function h_m in a suitable set, there exists a strategy $\pi_{m+1}^2 \in \Pi^2$ such that (6.4) holds. This statement is true by Assumption 2.12, the compactness of V^2, and again Theorem A.2.

As already noted above, a necessary condition for the algorithm to be well–defined is the existence of a solution $\left(J\left(\pi^1, \pi^2\right), h_{\pi^1, \pi^2}\right)$ to the Poisson equation (6.5). To prove this, we introduce the concept of *bias* of $\left(\pi^1, \pi^2\right)$.

Definition 6.3. *Let* $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. *The* <u>*bias*</u> *of* (π^1, π^2) *is the function given by*

$$h_{\pi^1,\pi^2}(x) := \int_0^\infty \left[\mathbb{E}_x^{\pi^1,\pi^2} r\left(x(t), \pi^1, \pi^2\right) - J\left(\pi^1, \pi^2\right)\right] dt. \tag{6.6}$$

Observe that this function is finite–valued because (2.13) and the Assumption 2.12(b) give, for all $t \geq 0$,

$$\left|\mathbb{E}_x^{\pi^1,\pi^2} r\left(x(t), \pi^1, \pi^2\right) - J\left(\pi^1, \pi^2\right)\right| \leq Ce^{-\delta t} M w(x). \tag{6.7}$$

Hence, by (6.6) and (6.7), the bias of (π^1, π^2) is such that

$$|h_{\pi^1,\pi^2}(x)| \leq \delta^{-1} CMw(x), \tag{6.8}$$

and so

$$\|h_{\pi^1,\pi^2}\|_w \leq \delta^{-1} CM.$$

This means that the bias h_{π^1,π^2} is a finite–valued function and, in fact, is in $\mathcal{B}_w(\mathbb{R}^n)$. Actually, its w–norm is uniformly bounded on $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. The following result is necessary to ensure that the PIA is well–defined.

Proposition 6.4. *For each* $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, *the pair* $\left(J\left(\pi^1, \pi^2\right), h_{\pi^1,\pi^2}\right)$ *is the unique solution of the Poisson equation* (6.5) *for which the* μ_{π^1,π^2}*–expectation of* h_{π^1,π^2} *is zero:*

$$\mu_{\pi^1,\pi^2}\left(h_{\pi^1,\pi^2}\right) = \int_{\mathbb{R}^n} h_{\pi^1,\pi^2}(x) \mu_{\pi^1,\pi^2}(dx) = 0. \tag{6.9}$$

Moreover, h_{π^1,π^2} *is in* $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$.

Proof. A slight variation of the vanishing discount technique of Section 5.2 gives us that, for fixed $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, the Poisson equation (6.5) has a solution $\tilde{h}_{\pi^1,\pi^2}$, which is a member of $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$, i.e.,

$$\tilde{J}\left(\pi^1, \pi^2\right) = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1,\pi^2} \tilde{h}_{\pi^1,\pi^2}(x) \text{ for all } x \in \mathbb{R}^n. \tag{6.10}$$

The difference between the technique of Section 5.2 and the one we use here, is that, instead of invoking Theorem 3.4, we invoke Corollary 3.5.

To obtain (6.9) first note that, by (2.7) and (6.8), h_{π^1,π^2} is indeed μ_{π^1,π^2}–integrable for every (π^1, π^2) in $\Pi^1 \times \Pi^2$. Then, in (6.9) choose the distribution of the initial state to be μ_{π^1,π^2} and so (6.9) follows from Fubini's theorem and the invariance of μ_{π^1,π^2}. Moreover, the fact that h_{π^1,π^2} is in $\mathcal{B}_w(\mathbb{R}^n)$ follows from (6.8).

On the other hand, the fact that $\tilde{J}\left(\pi^1, \pi^2\right)$ coincides with the ergodic payoff $J\left(\pi^1, \pi^2\right)$ in (5.3) is a direct consequence of the proof of Proposition 5.4 and the part that adresses uniqueness in Theorem 5.5(i).

Next, to ensure that $\tilde{h}_{\pi^1,\pi^2}$ equals the bias h_{π^1,π^2} in (6.6) for all $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, we can use Theorem A.1 on $\tilde{h}_{\pi^1,\pi^2}(x(t))$ to obtain

$$\mathbb{E}_x^{\pi^1,\pi^2}\left[\tilde{h}_{\pi^1,\pi^2}(x(t))\right] = \tilde{h}_{\pi^1,\pi^2}(x) + \tilde{J}\left(\pi^1, \pi^2\right) t - \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^t r\left(x(s), \pi^1, \pi^2\right) ds\right].$$

This implies

$$\tilde{h}_{\pi^1,\pi^2}(x) = \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^t \left(r\left(x(s), \pi^1, \pi^2\right) - J\left(\pi^1, \pi^2\right)\right) ds\right] + \mathbb{E}_x^{\pi^1,\pi^2}\left[\tilde{h}_{\pi^1,\pi^2}(x(t))\right]. \tag{6.11}$$

Since h_{π^1,π^2} is in $\mathcal{B}_w(\mathbb{R}^n)$ for all $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, we see that the uniform w–exponential ergodicity condition (2.12) yields that the second term of the right hand side of (6.11) converges to $\mu_{\pi^1,\pi^2}\left(\tilde{h}_{\pi^1,\pi^2}\right)$ as t goes to infinity; but, by (6.9), this last limit becomes zero. Therefore, letting $t \to \infty$ in both sides of (6.11), we obtain

$$\tilde{h}_{\pi^1,\pi^2}(x) = \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^\infty \left(r\left(x(s), \pi^1, \pi^2\right) - J\left(\pi^1, \pi^2\right)\right) ds\right],$$

which coincides with the bias h_{π^1,π^2} defined in (6.6). These facts yield also uniqueness of solutions to equation (6.5), and Proposition 6.4 follows. □

6.2 Convergence

This section is intended to prove that the PIA converges. But first, we present the following extension of [46, Lemma 4.5]. Part (b) of Lemma 6.5, along with (6.3) gives that if $J\left(\pi_m^1, \pi_m^2\right) = J\left(\pi_{m+1}^1, \pi_{m+1}^2\right)$ in the PIA, then $\left(\pi_m^1, \pi_m^2\right)$ is a saddle point of the average SDG.

Lemma 6.5. *Let $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$ be an arbitrary pair of randomized stationary strategies. Let $\pi_*^1 \in \Pi^1$ be such that*

$$\begin{aligned} J\left(\pi_*^1, \pi^2\right) &= \sup_{\varphi \in V^1} \left[r\left(x, \varphi, \pi^2\right) + \mathbb{L}^{\varphi, \pi^2} h(x) \right] && (6.12) \\ &= r\left(x, \pi_*^1, \pi^2\right) + \mathbb{L}^{\pi_*^1, \pi^2} h_{\pi_*^1, \pi^2}(x) \ \textit{for all} \ x \in \mathbb{R}^n, && (6.13) \end{aligned}$$

and let $\pi_*^2 \in \Pi^2$ be such that

$$\inf_{\psi \in V^2} \left[r\left(x, \pi_*^1, \psi\right) + \mathbb{L}^{\pi_*^1, \psi} h_{\pi_*^1, \pi^2}(x) \right] = r\left(x, \pi_*^1, \pi_*^2\right) + \mathbb{L}^{\pi_*^1, \pi_*^2} h_{\pi_*^1, \pi^2}(x) \tag{6.14}$$

for all $x \in \mathbb{R}^n$. Then

(a) $J\left(\pi_*^1, \pi_*^2\right) \leq J\left(\pi_*^1, \pi^2\right)$, *and*

(b) if $J\left(\pi^1, \pi_*^2\right) \leq J\left(\pi_*^1, \pi_*^2\right)$, then $\left(\pi_*^1, \pi_*^2\right)$ is a saddle point of the SDG with average payoff.

Proof. The relations (6.12)–(6.14) imply

$$\begin{aligned} & r\left(x, \pi_*^1, \pi_*^2\right) + \mathbb{L}^{\pi_*^1, \pi_*^2} h_{\pi_*^1, \pi^2}(x) \\ \leq \ & r\left(x, \pi_*^1, \pi^2\right) + \mathbb{L}^{\pi_*^1, \pi^2} h_{\pi_*^1, \pi^2}(x) \\ = \ & J\left(\pi_*^1, \pi^2\right). \end{aligned}$$

An application of Proposition 5.4 yields (a). Part (b) of the result is immediate from (a) and (5.10). □

Proposition 6.6 guarantees the existence of a pair of policies $\left(\pi_*^1, \pi_*^2\right)$ in $\Pi^1 \times \Pi^2$ that satisfies that, for every fixed $x \in \mathbb{R}^n$, there exists a subsequence $m_k \equiv m_k(x)$ of $\{m\}$ such that

$$\left(\pi_{m_k}^1(\cdot|x), \pi_{m_k}^2(\cdot|x)\right) \to \left(\pi_*^1(\cdot|x), \pi_*^2(\cdot|x)\right) \ \text{ as } \ k \to \infty \tag{6.15}$$

in the topology of weak convergence of $V^1 \times V^2$. This type of policy convergence was first introduced in [87, Lemma 4] for the case of nonstationary, deterministic, discrete–time policies. It can also be found in [88, Proposition 12.2] and [45, Theorem D.7]. In this case we say that the sequence $\left\{\left(\pi_m^1, \pi_m^2\right) : m = 1, 2, ...\right\}$ converges *in the sense of Schäl* to $\left(\pi_*^1, \pi_*^2\right)$.

Proposition 6.6. *Let $\left\{\left(\pi_m^1, \pi_m^2\right) : m = 1, 2, ...\right\} \subset \Pi^1 \times \Pi^2$ be the sequence generated by the PIA. If Assumptions 2.1, 2.6 and 2.12 hold, then, there exists $\left(\pi_*^1, \pi_*^2\right) \in \Pi^1 \times \Pi^2$ such that $\left(\pi_*^1, \pi_*^2\right)$ is the limit in the sense of Schäl of $\left\{\left(\pi_m^1, \pi_m^2\right) : m = 1, 2, ...\right\}$.*

Proof. Fix $x \in \mathbb{R}^n$. By the compactness of $V^1 \times V^2$, we can ensure the existence of a subsequence $m_k \equiv m_k^\ell(x)$, $\ell = 1, 2$, such that $\pi_{m_k}^\ell(\cdot|x) \to \pi_*^\ell(\cdot|x)$. Using again the compactness of V^ℓ, $\ell = 1, 2$, we easily see that $\pi_*^\ell(\cdot|x)$ is a probability measure. Furthermore, for all $B \subseteq U^\ell$, by [87, Lemma 4], $\pi_*^\ell(B|\cdot)$ is measurable on $\mathbb{R}^n$. Hence, π_*^ℓ is in Π^ℓ. This proves the result. □

Theorem 6.7. *Let $p > n$. Let Assumptions 2.1, 2.6, 2.9 and 2.12 hold. In addition, let $\left(\pi_m^1, \pi_m^2\right)$ be a pair of randomized stationary policies generated by the PIA. Then $\left\{\left(\pi_m^1, \pi_m^2\right) : m = 1, 2, ...\right\}$ converges in the sense of Schäl to a saddle point $\left(\pi_*^1, \pi_*^2\right)$ of the average SDG. Therefore the PIA converges.*

Proof. For each pair $\left(\pi_m^1, \pi_m^2\right)$ generated by the PIA, Proposition 6.4 ensures the existence of a function $h_m \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ such that (6.2) holds. Now, an analogous argument to those presented in Sections 4.1.2 and 5.2 allows us to invoke Corollary 3.5, thus establishing the existence of a function $h \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ such that

$$\lim_{m \to \infty} h_m = h \text{ for all } x \in \mathbb{R}^n.$$

On the other hand, observe that Proposition 6.6 asserts the existence of the limit $\left(\pi_*^1, \pi_*^2\right)$ (in the sense of Schäl) of the sequence of policies $\left\{\left(\pi_m^1, \pi_m^2\right)\right\}$ generated by the PIA.

Now, fix an arbitrary state $x \in \mathbb{R}^n$ and let m_k be as in (6.15). Next, in (6.2), replace m by m_k and let $k \to \infty$ to obtain

$$J\left(\pi_*^1, \pi_*^2\right) = r\left(x, \pi_*^1, \pi_*^2\right) + \mathbb{L}^{\pi_*^1, \pi_*^2} h(x).$$

We shall use Lemma 6.5 to conclude the proof. Namely, observe that (6.1) in step 2 of the PIA ensures that (6.12) holds. In addition, (6.4) in step 4 yields (6.14). Hence, Lemma 6.5(b) asserts that $\left(\pi_*^1, \pi_*^2\right)$ is a saddle point of the ergodic game and the result is thus proved. □

6.3 Concluding remarks

This chapter gives sufficient conditions under which the PIA converges in a certain class of games. The version of the algorithm under study is an extension to the continuous–time scheme of that presented in [46] and of Hoffman–Karp [94] algorithm. Besides, our state space and action sets are nondenumerable. Our two results, Lemma 6.5 and Theorem 6.7 are inspired in [46, Lemma 4.5] and [46, Theorem 4.3], respectively.

Chapter 7

Bias and overtaking equilibria for zero–sum SDGs

The aim of the present chapter is to study conditions ensuring the existence of bias and overtaking equilibria in a zero–sum SDG. We will introduce these criteria by means of the classic average optimality criterion studied in Chapters 5 and 6.

The chapter represents an extension to SDGs of some bias and overtaking optimality results for controlled diffusions [52] and for continuous–time Markov games with a denumerable state–space [78]. In some stochastic control problems, the concepts of bias and overtaking optimality are equivalent (see, for instance, [52]). However, the evidence found in [78] indicates that in continuous–time games this equivalence does not hold, although this remains an open problem for games with a nondenumerable state space.

7.1 Bias optimality

Throughout the following we will suppose that Assumptions 2.1, 2.6, 2.9 and 2.12 hold.

We recall that the set of strategies that satisfy (5.10) is denoted by $\left(\Pi^1 \times \Pi^2\right)_{ae}$, that is, $\left(\pi^1_*, \pi^2_*\right)$ is in $\left(\Pi^1 \times \Pi^2\right)_{ae}$ if and only if

$$J\left(\pi^1, \pi^2_*\right) \leq J\left(\pi^1_*, \pi^2_*\right) \leq J\left(\pi^1_*, \pi^2\right) \quad \text{for every} \quad \left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2.$$

Recall as well Definition 6.3 of the bias h_{π^1,π^2} and its characterization as solution of the Poisson equation given in Proposition 6.4.

The following definition uses the concept of average payoff equilibira introduced above.

Definition 7.1. *We say that an average payoff equilibrium* $\left(\pi^1_*, \pi^2_*\right) \in \left(\Pi^1 \times \Pi^2\right)_{ae}$ *is a* *bias equilibrium* *if*

$$h_{\pi^1,\pi^2_*}(x) \leq h_{\pi^1_*,\pi^2_*}(x) \leq h_{\pi^1_*,\pi^2}(x) \tag{7.1}$$

for all $x \in \mathbb{R}^n$ *and every pair of average payoff equilibria* $\left(\pi^1, \pi^2\right) \in \left(\Pi^1 \times \Pi^2\right)_{ae}$. *The function* $h_{\pi^1_*,\pi^2_*}$ *is called the* *optimal bias function.*

The next result is an extension to SDGs of [52, Proposition 4.2]. It gives an expression for the bias function of $\left(\pi^1, \pi^2\right)$ by using any solution h of the average optimality equations (5.11)–(5.13).

Proposition 7.2. *If* $\left(\pi^1, \pi^2\right) \in \left(\Pi^1 \times \Pi^2\right)_{ae}$, *then its bias* h_{π^1,π^2} *and any solution* h *of the average optimality equations* (5.11)–(5.13) *coincide up to an additive constant; in fact,*

$$h_{\pi^1,\pi^2}(x) = h(x) - \mu_{\pi^1,\pi^2}(h) \quad \textit{for all } x \in \mathbb{R}^n. \tag{7.2}$$

Proof. Let $(\pi^1, \pi^2) \in (\Pi^1 \times \Pi^2)_{ae}$ be an arbitrary average payoff equilibrium. Then (π^1, π^2) satisfies Theorem 5.5*(ii)* with $J = \mathcal{V}$, i.e.,

$$\mathcal{V} = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1, \pi^2} h(x) \text{ for all } x \in \mathbb{R}^n. \tag{7.3}$$

In addition, the Poisson equation for (π^1, π^2) is

$$\mathcal{V} = r\left(x, \pi^1, \pi^2\right) + \mathbb{L}^{\pi^1, \pi^2} h_{\pi^1, \pi^2}(x) \text{ for all } x \in \mathbb{R}^n. \tag{7.4}$$

The subtraction of (7.3) from (7.4) yields that $h - h_{\pi^1, \pi^2}$ is a harmonic function. Consequently, (7.2) follows from Theorem A.1, Lemma 2.11 and (6.9). □

If the optimal bias function $h_{\pi^1_*, \pi^2_*}$ exists, then, by Proposition 6.4, it is in $\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ for any bias equilibrium (π^1_*, π^2_*).

Let $(\Pi^1 \times \Pi^2)_{bias}$ be the family of bias equilibria. By Definition 7.1,

$$(\Pi^1 \times \Pi^2)_{bias} \subset (\Pi^1 \times \Pi^2)_{ae}.$$

Let $(J, h) \in \mathbb{R} \times \left(\mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)\right)$ be a solution of the average optimality equations (5.11)–(5.13). We define for every $x \in \mathbb{R}^n$ the sets

$$\Pi^1(x) := \left\{ \varphi \in V^1 : J = \inf_{\psi \in V^2} \left[r(x, \varphi, \psi) + \mathbb{L}^{\varphi, \psi} h(x) \right] \right\}, \tag{7.5}$$

$$\Pi^2(x) := \left\{ \psi \in V^2 : J = \sup_{\varphi \in V^1} \left[r(x, \varphi, \psi) + \mathbb{L}^{\varphi, \psi} h(x) \right] \right\}. \tag{7.6}$$

We now present an extension of [78, Lemma 4.6].

Lemma 7.3. *For every $x \in \mathbb{R}^n$, $\Pi^1(x)$ and $\Pi^2(x)$ are convex compact sets. Moreover, the multifunctions $x \mapsto \Pi^1(x)$ are such that $\Pi^1(\cdot)$ is upper semicontinuous, and $\Pi^2(\cdot)$ is lower semicontinuous.*

Proof. Recall from Section 2.1 that the sets V^1 and V^2 (endowed with the topology of weak convergence) are compact. Thus, we only need to show that $\Pi^\ell(x)$ is a closed set ($\ell = 1, 2$). But this is a consequence of Lemma 4.4 in [78] and Lemma 2.14. The proof that $\Pi^1(x)$ and $\Pi^2(x)$ are convex mimics that of Lemma 4.6 in [78]. The upper semicontinuity of $\Pi^1(\cdot)$ was given, for the case of controlled diffusions, in [52, Lemma 5.2]. However, it is not difficult to see why it holds in the present case (the same goes for the lower semicontinuity of $\Pi^2(\cdot)$). □

Remark 7.4. *By Theorem 5.5*(ii)*, (π^1, π^2) is in $(\Pi^1 \times \Pi^2)_{ae}$ if and only if $\pi^1(\cdot|x)$ is in $\Pi^1(x)$ and $\pi^2(\cdot|x)$ is in $\Pi^2(x)$ for all $x \in \mathbb{R}^n$.*

Theorem 7.5. *The set $(\Pi^1 \times \Pi^2)_{bias}$ is nonempty.*

Proof. Let $(\pi^1, \pi^2) \in (\Pi^1 \times \Pi^2)_{ae}$ be an average payoff equilibrium. Using the expression (7.2) for the bias function h_{π^2, π^2}, we obtain that finding bias equilibria is equivalent to solving a new SDG with ergodic payoff. Let us call this problem *bias game*. The components of this game are:

- The dynamic system (2.1),
- The action sets $\Pi^1(x)$ and $\Pi^2(x)$ for each $x \in \mathbb{R}^n$, and
- The reward rate

$$r'\left(x, \pi^1, \pi^2\right) := -h(x).$$

Observe that the bias game satisfies Assumptions 2.1, 2.6, 2.9 and 2.12. Hence, Theorem 5.5 ensures the existence of a constant $\tilde{J}$; a function $\tilde{h} \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$; and a pair (π^1, π^2) such that $(\pi^1(\cdot|x), \pi^2(\cdot|x))$ is in $\Pi^1(x) \times \Pi^2(x)$ for every $x \in \mathbb{R}^n$. Moreover, $\tilde{h}$ and (π^1, π^2) satisfy the average payoff optimality equations (5.11)–(5.13); that is,

$$\begin{aligned}
\tilde{J} &= -h(x) + \mathbb{L}^{\pi^1,\pi^2}\tilde{h}(x) && (7.7)\\
&= \sup_{\varphi\in\Pi^1(x)} \left[-h(x) + \mathbb{L}^{\varphi,\pi^2}\tilde{h}(x)\right] && (7.8)\\
&= \inf_{\psi\in\Pi^2(x)} \left[-h(x) + \mathbb{L}^{\pi^1,\psi^2}\tilde{h}(x)\right]. && (7.9)
\end{aligned}$$

Hence, by virtue of Theorem 5.5(ii) and (7.2), (π^1, π^2) is a bias equilibrium. i.e., $(\pi^1, \pi^2) \in (\Pi^1 \times \Pi^2)_{bias}$. By (7.7)–(7.9) the value of the bias game is

$$\tilde{J} = \mu_{\pi^1,\pi^2}(-h) := \mathcal{V}_h^*.$$

□

Let $(\pi^1, \pi^2) \in (\Pi^1 \times \Pi^2)_{bias}$. Using (7.2), we define

$$\mathcal{H}(x) := h_{\pi^1,\pi^2}(x) = h(x) + \mathcal{V}_h^*, \qquad (7.10)$$

where $\mathcal{V}_h^*$ is the value of the bias game.

Bias optimality equations

We give a characterization of bias equilibria by means of the bias optimality equations defined as follows.

Definition 7.6. *We say that the constant $J \in \mathbb{R}$, the functions $\mathcal{H}, \tilde{h} \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ and the pair $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ verify the bias optimality equations if J and $\mathcal{H}$ satisfy the average optimality equations* (5.11)–(5.13) *and, in addition for every $x \in \mathbb{R}^n$, $\tilde{h}$ satisfies*

$$\begin{aligned}
\mathcal{H}(x) &= \mathbb{L}^{\pi^1,\pi^2}\tilde{h}(x) && (7.11)\\
&= \sup_{\varphi\in\Pi^1(x)} \mathbb{L}^{\varphi,\pi^2}\tilde{h}(x) && (7.12)\\
&= \inf_{\psi\in\Pi^2(x)} \mathbb{L}^{\pi^1,\psi}\tilde{h}(x). && (7.13)
\end{aligned}$$

Theorem 7.7. *Under our hypotheses, the following assertions are true.*

(i) A solution of the bias optimality equations (5.11)–(5.13) *and* (7.11)–(7.13), *with $\mathcal{H}(0) = \mathcal{V}_h^*$, exists, is unique, and, further, $J = \mathcal{V}$, with $\mathcal{V}$ as in* (5.9).

(ii) A pair of stationary strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ is a bias equilibrium if and only if it verifies the bias optimality equations.

Proof. By Theorem 5.5 we know that the equations (5.11)–(5.13) have a unique solution (V, h). Now, since $\mathbb{L}^{\pi^1,\pi^2}$ is a differential operator, it follows that, if h satisfies (5.11)–(5.13), then, so does $\mathcal{H}$ in (7.10). On the other hand, the same arguments in the proof of Theorem 7.5 ensure the existence of a function, say $\tilde{h} \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ such that $(\mathcal{V}_h^*, \tilde{h})$ is the unique solution to the average optimal equations for the bias game with reward rate $-h(\cdot)$, i.e., $(\mathcal{V}_h^*, \tilde{h})$ satisfies (7.7)–(7.9). Hence, from (7.10) we can see that $\mathcal{H}(x)$ satisfies (7.11)–(7.13).

Part *(ii)* follows from Theorem 5.5*(ii)* applied to the bias game. □

7.1.1 The PIA for the bias game

By the proof of Theorem 7.5, the bias game can be expressed as a SDG with a particular average payoff. We will use this and a modification of the PIA presented in Chapter 6 to find another characterization of bias equilibria.

We assume that the original SDG with average payoff of Chapter 5 has been solved, i.e., J is the game value, $\left(\pi_0^1, \pi_0^2\right)$ belongs to $\left(\Pi^1 \times \Pi^2\right)_{ae}$, and $h(x) = h_{\pi_0^1,\pi_0^2}(x) + \mu_{\pi_0^1,\pi_0^2}(h)$ for all $x \in \mathbb{R}^n$.

Step 1. Set $m = 0$. Fix $\pi_0^2 \in \Pi^2(x)$ and define $\tilde{J}_0 := -\infty$.

Step 2. Find a policy $\pi_m^1(\cdot|x) \in \Pi^1(x)$, a constant $\tilde{J}_m$, and a function $\tilde{h}_m : \mathbb{R}^n \to \mathbb{R}$ such that $\left(\tilde{J}_m, \tilde{h}_m\right)$ is a solution of (7.8).

Step 3. If $\tilde{J}_m = \tilde{J}_{m-1}$, then $\left(\pi_m^1, \pi_m^2\right) \in \left(\Pi^1 \times \Pi^2\right)_{bias}$. Terminate PIA. Otherwise, go to step 4.

Step 4. Determine an average optimal strategy $\pi_{m+1}^2(\cdot|x) \in \Pi^2(x)$ that attains the minimum on (7.9).

Increase m in 1 and go to step 2.

Analogously to the end of Section 6.1, there are some critical parts we must verify to ensure that this version of the PIA for the bias game is well–defined and yields a pair of bias equilibria.

1. In step 2, $\pi_m^1(\cdot|x)$ is such that

$$\begin{aligned}\tilde{J}_m &= \sup_{\varphi\in\Pi^1(x)} \left[-h(x) + \mathbb{L}^{\varphi,\pi_m^2}\tilde{h}_m(x)\right] \\ &= -h(x) + \mathbb{L}^{\pi_m^1,\pi_m^2}\tilde{h}_m(x),\end{aligned}$$

   which is consistent with Theorem 7.7. Similarly, the strategy $\pi_{m+1}^2(\cdot|x)$ of step 3 is such that

$$\begin{aligned}\tilde{J}_m &= \inf_{\psi\in\Pi^2(x)} \left[-h(x) + \mathbb{L}^{\pi_m^1,\psi}\tilde{h}_m(x)\right] \\ &= -h(x) + \mathbb{L}^{\pi_m^1,\pi_{m+1}^2}\tilde{h}_m(x).\end{aligned}$$

2. Proposition 5.1 gives that $-h$ is μ_{π^1,π^2}–integrable.

3. Lemma 7.3 can be invoked to ensure the compactness of $\Pi^1(x)$. Thus Theorem A.2 (with $V^1 \times V^2$ replaced by $\Pi^1(x) \times \Pi^2(x)$) allows us to extend [52, Theorem 3.2] to the context of randomized strategies. These steps enable us to guarantee the existence of $\tilde{J}_m$, a function $\tilde{h}_m$ in $\mathcal{C}^2(\mathbb{R}^n)\cap\mathcal{B}_w(\mathbb{R}^n)$, and φ in $\Pi^1(x)$, that maximizes (7.8).

4. Assumption 2.12, Lemma 7.3 and Theorem A.2 (with $\Pi^2(x)$ in lieu of V^2) allow the extension of [52, Theorem 3.2] that ensures that, for h_m given, there exists $\pi^2(\cdot|x) \in \Pi^2(x)$ that minimizes (7.9).

These remarks, together with Lemma 6.5 and Theorem 6.7 give that the PIA for the bias game is well–defined.

7.2 Overtaking optimality

In this section we introduce the overtaking optimality criterion and show some relations between this criterion and bias optimality.

Definition 7.8. *A pair of strategies $\left(\pi_*^1, \pi_*^2\right) \in \Pi^1 \times \Pi^2$ is said to be an* overtaking equilibrium *in the class $\Pi^1 \times \Pi^2$ if for each $\left(\pi^1, \pi^2\right) \in \Pi^1 \times \Pi^2$ and $x \in \mathbb{R}^n$ we have*

$$\liminf_{T\to\infty} \left[J_T\left(x, \pi_*^1, \pi_*^2\right) - J_T\left(x, \pi^1, \pi_*^2\right)\right] \geq 0 \tag{7.14}$$

and

$$\limsup_{T\to\infty} \left[J_T\left(x,\pi_*^1,\pi_*^2\right) - J_T\left(x,\pi_*^1,\pi^2\right)\right] \leq 0. \tag{7.15}$$

The set of pairs of overtaking equilibria is denoted by $\left(\Pi^1 \times \Pi^2\right)_{oe}$.

Remark 7.9. *If* (π_1, π_2) *is an overtaking equilibrium in* $\Pi^1 \times \Pi^2$*, then it is an average payoff equilibrium. To see this, it suffices to compare the definitions of lim inf, in* (7.14)*, and lim sup, in* (7.15)*, with expressions* (5.2)*,* (5.6)*, and* (5.10)*.*

By Theorem A.1 and the definition in (5.1), we can write $J_T\left(x,\pi^1,\pi^2\right)$ as follows

$$J_T\left(x,\pi^1,\pi^2\right) = T\cdot J\left(\pi^1,\pi^2\right) + h_{\pi^1,\pi^2}(x) - \mathbb{E}_x^{\pi^1,\pi^2} h_{\pi^1,\pi^2}\left(x(T)\right) \tag{7.16}$$

for every $\left(\pi^1,\pi^2\right)$ in $\left(\Pi^1 \times \Pi^2\right)_{ae}$.

Theorem 7.10. *If a pair of strategies* $\left(\pi_*^1,\pi_*^2\right)$ *is an overtaking equilibrium in* $\Pi^1 \times \Pi^2$*, then it is a bias equilibrium.*

Proof. Let $\left(\pi_*^1,\pi_*^2\right) \in \Pi^1 \times \Pi^2$ be a pair of overtaking optimal strategies. Then by the Remark 7.9, we have that $\left(\pi_*^1,\pi_*^2\right) \in \left(\Pi^1 \times \Pi^2\right)_{ae}$. Then, by using (7.16), we obtain

$$\begin{aligned} & J_T\left(x,\pi_*^1,\pi_*^2\right) - J_T\left(x,\pi^1,\pi_*^2\right) \\ = \; & h_{\pi_*^1,\pi_*^2}(x) - h_{\pi^1,\pi_*^2}(x) + \mathbb{E}_x^{\pi_*^1,\pi_*^2} h_{\pi_*^1,\pi_*^2}(x(T)) - \mathbb{E}_x^{\pi^1,\pi_*^2} h_{\pi^1,\pi_*^2}(x(T)), \end{aligned} \tag{7.17}$$

for each $\pi^1 \in \Pi^1(x)$ (recall the definition of $\Pi^1(x)$ and $\Pi^2(x)$ in (7.5)–(7.6)). Equation (7.17), along with (2.13), (6.9) and (7.14), yields

$$\liminf_{T\to\infty} \left[J_T\left(x,\pi_*^1,\pi_*^2\right) - J_T\left(x,\pi^1,\pi_*^2\right)\right] = h_{\pi_*^1,\pi_*^2}(x) - h_{\pi^1,\pi_*^2}(x) \geq 0. \tag{7.18}$$

Similar arguments show that for all $\pi^2 \in \Pi^2(x)$,

$$\limsup_{T\to\infty} \left[J_T\left(x,\pi_*^1,\pi_*^2\right) - J_T\left(x,\pi_*^1,\pi^2\right)\right] = h_{\pi_*^1,\pi_*^2}(x) - h_{\pi_*^1,\pi^2}(x) \leq 0. \tag{7.19}$$

Inequalities (7.18) and (7.19) yield condition (7.1) in Definition 7.1 for all $\left(\pi_*^1,\pi_*^2\right) \in \left(\Pi^1 \times \Pi^2\right)_{ae}$. Hence the pair $\left(\pi_*^1,\pi_*^2\right)$ is a bias equilibrium. □

Theorem 7.11. *Suppose that a pair of strategies* $\left(\pi_*^1,\pi_*^2\right) \in \Pi^1 \times \Pi^2$ *is a bias equilibrium, then it is an overtaking equilibrium in the class* $\left(\Pi^1 \times \Pi^2\right)_{ae}$.

Proof. Let $\left(\pi^1,\pi^2\right)$ in $\left(\Pi^1 \times \Pi^2\right)_{ae}$ be a pair of average optimal strategies and let $\left(\pi_*^1,\pi_*^2\right) \in \left(\Pi^1 \times \Pi^2\right)_{bias}$. Then $\left(\pi_*^1,\pi_*^2\right)$ is in $\left(\Pi^1 \times \Pi^2\right)_{ae}$. Using (7.16) yields (7.17) again. Hence, (2.13), (6.9) and (7.1) give

$$\liminf_{T\to\infty} \left[J_T\left(x,\pi_*^1,\pi_*^2\right) - J_T\left(x,\pi^1,\pi_*^2\right)\right] \geq 0.$$

Similarly

$$\limsup_{T\to\infty} \left[J_T\left(x,\pi_*^1,\pi_*^2\right) - J_T\left(x,\pi_*^1,\pi^2\right)\right] \leq 0$$

Hence $\left(\pi_*^1,\pi_*^2\right)$ is an overtaking equilibrium in the class $\left(\Pi^1 \times \Pi^2\right)_{ae}$ of average payoff equilibria. □

Remark 7.12. *Theorem 7.10 shows that overtaking optimality implies bias optimality in the class* $\Pi^1 \times \Pi^2$ *of stationary strategies. On the other hand, Theorem 7.11 gives a partial converse; the results in [78] lead us to think that the full converse holds only in the class* $\left(\Pi^1 \times \Pi^2\right)_{ae}$*, but this remains an open problem. In contrast, overtaking and bias optimality are equivalent in control (or single–player) problems in the class* $\Pi^1 \times \Pi^2$ *of stationary strategies; see, for instance, [52, Theorem 5.5].*

7.3 Concluding remarks

This chapter presents a unified analysis of bias and overtaking optimality for a general class of zero–sum SDGs. Under suitable hypotheses such as uniform ellipticity in Assumption 2.1(b) and the uniform w–exponential ergodicity in (2.13), we have shown the existence and give characterizations of bias equilibria and their connection with overtaking equilibria. Moreover, we provide an algorithm to find bias equilibria in terms of the so–called *bias game*.

Our characterizations follow a lexicographical type in the sense that, first, we identify the set of average payoff equilibria, and then, within this set, we look for some special strategies. Finally, we show that overtaking equilibrium implies bias equilibrium (Theorem 7.10). However, the results in [78] lead us to believe that a partial converse holds: bias equilibria are overtaking equilibria in the class of average payoff equilibria. This is an open issue.

Chapter 8

Final remarks

In this thesis we have studied several infinite–horizon zero–sum games for a general class of Markov diffusion processes. Our work begins with the study of two basic optimality criteria, namely, discounted payoff (in Chapter 4) and average payoff (in Chapters 5 and 6). We used an alternative method —with Theorem 3.4 in its core— to prove the existence of value functions and saddle points under these two criteria, and we used Chapter 6 to propose an algorithm to characterize average payoff equilibria and game's value for the corresponding SDG. Moreover, we also gave conditions milder than those given in [43] to prove that the Poisson equation (6.5) has a solution $h_{\pi^1,\pi^2} \in \mathcal{C}^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ (see Proposition 6.4).

In Chapter 4.2 we studied a special class of zero–sum SDGs under what we called *random discounted payoff criterion*. We started by replacing the fixed parameter $\alpha > 0$ by the continuous–time Markov chain $\alpha(\cdot)$ and we used this process to index the SDE (4.26) so that we had a Markov–modulated diffusion. We proved the existence of saddle points and value functions as we did in Section 4.1. A major difference between our approach and that of [90, 91] is that we consider that the *switching parameter* is present in both: the diffusion itself, and the discount factor. Actually, our Theorem 4.19 is suitable for the problem we present here, as well as for the examples presented by Song, Yin and Zhang in their works.

We also provided a proof for the existence of overtaking optimal policies and gave two characterizations of these policies. In particular, we related the concept of overtaking optimality with the concept of bias optimality. We also presented a modified version of the PIA that looks for bias optimal strategies in $\left(\Pi^1 \times \Pi^2\right)_{ao}$ the set of average optimal strategies.

Nevertheless, there are several research lines in the theory of SDGs that our work leaves open. For instance:

1. Solving a non–stationary version of the discounted payoff criterion and proving an analogue of Theorem 3.4 for this problem. This would imply to work with Cauchy problems of parabolic type, thus applying the results in Kolokoltsov [58, 59], Reed and Simon [84], and Ladyzhenskaya and Uraltseva [65] to the context of our interest.

2. Finding a suitable PIA for the discounted payoff criterion.

3. Propose appropriate applications, in economics, environmental or actuarial sciences, of our results on (*i*) the random discounted payoff criterion (we think that the examples of [90, 91] make a fair point of departure); and (*ii*) the bias and overtaking equilibria (an extension to the work of Kawaguchi and Morimoto [56] seems appealing enough in this instance).

Our work is also a collection of applications of the powerful Theorem 3.4 and within that approach lies the key to a possible extension of our results: discounted and ergodic payoff criteria for nonzero–sum SDGs.

We will devote our future efforts to finding answers in these and other related topics.

Appendix A

Frequently used results

This Appendix presents four crucial results that are repeatedly quoted along our work.

For our first two results, Ω is a bounded subset of $\mathbb{R}^n$, $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, $h \in \mathcal{W}^{2,p}(\Omega)$, and $x(\cdot) = \{x(t) : t \geq 0\}$ is an almost surely strong solution to

$$dx(t) = b\left(x(t), \pi^1, \pi^2\right) dt + \sigma\left(x(t)\right) dW(t)$$

with initial condition $x(0) = x$, where $b : \mathbb{R}^n \times \Pi^1 \times \Pi^2 \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are given functions, and $W(\cdot)$ is an m–dimensional Wiener process. We also assume that $\mathbb{L}^{\pi^1,\pi^2}h(\cdot)$ is the infinitesimal generator of $x(\cdot)$; and is given by

$$\begin{aligned}\mathbb{L}^{\pi^1,\pi^2}h(x) &:= \left\langle \nabla h(x), b\left(x, \pi^1, \pi^2\right)\right\rangle + \frac{1}{2}\mathrm{Tr}\left[\left[\mathbb{H}h(x)\right]\cdot a(x)\right] \\ &= \sum_{i=1}^{n} b_i\left(x, \pi^1, \pi^2\right)\partial_{x_i}h(x) + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\partial^2_{x_i x_j}h(x),\end{aligned}$$

with $a(\cdot)$ as in Assumption 2.1(c).

For the sake of completeness, we list first Dynkin's formula. See [57, Corollary 6.5].

Theorem A.1. *Let τ be the first exit time of Ω (cf.* (4.13)*). If Assumption 2.1 and Remark 2.5 hold, then*

$$\mathbb{E}_x^{\pi^1,\pi^2}\left[h\left(x(\tau)\right)\right] = h(x) + \mathbb{E}_x^{\pi^1,\pi^2}\left[\int_0^{\tau}\mathbb{L}^{\pi^1,\pi^2}h(x(t))dt\right].$$

The next result is a tuned version of [72, Theorem 2.1]. Let V^ℓ be the space of probability measures on U^ℓ endowed with the topology of weak convergence ($\ell = 1, 2$).

Theorem A.2. *If Assumption 2.1, Remark 2.5, Assumptions 2.6 and 2.12, and Remark 2.13 hold, then, there exists a pair $\left(\pi^1_*, \pi^2_*\right) \in \Pi^1 \times \Pi^2$ such that*

$$\begin{aligned} r\left(x, \pi^1_*, \pi^2\right) + \mathbb{L}^{\pi^1_*,\pi^2}h(x) &= \sup_{\varphi\in V^1}\left\{r\left(x, \varphi, \pi^2\right) + \mathbb{L}^{\varphi,\pi^2}h(x)\right\}, \ \textit{and} \\ r\left(x, \pi^1, \pi^2_*\right) + \mathbb{L}^{\pi^1,\pi^2_*}h(x) &= \inf_{\psi\in V^2}\left\{r\left(x, \pi^1, \psi\right) + \mathbb{L}^{\pi^1,\psi}h(x)\right\}.\end{aligned}$$

Now let Ω be as in Chapter 3, that is, a bounded, open and connected subset of $\mathbb{R}^n$. Assume that the space $V^1 \times V^2$ has a single element, namely (φ, ψ). Recall as well definitions (3.1) and (3.5).

For every $x \in \mathbb{R}^n$, $\alpha > 0$, and h in $\mathcal{C}^2(\mathbb{R}^n)$ let

$$\hat{b}\left(x, \varphi, \psi, h, \alpha\right) := \left\langle \nabla h(x), b\left(x, \varphi, \psi\right)\right\rangle - \alpha h(x) + r\left(x, \varphi, \psi\right), \tag{A.1}$$

with b as in (2.5) and r as in (2.21). We also recall

$$\hat{\mathbb{L}}_{\alpha}^{\varphi,\psi}h(x) := \hat{b}\left(x,\varphi,\psi,h,\alpha\right) + \frac{1}{2}\mathrm{Tr}\left[[\mathbb{H}h(x)]a(x)\right], \tag{A.2}$$

with a as in Assumption 2.1(b).

We now borrow a version of Theorem 9.11 of [36] that is appropriate to our context.

Theorem A.3. *Let* $h \in \mathcal{W}^{2,p}(\Omega)$ *be a solution of* $\hat{\mathbb{L}}_{\alpha}^{\varphi,\psi}h = \xi$*, with* $\xi \in \mathcal{L}^p(\Omega)$*, and* $\hat{\mathbb{L}}_{\alpha}^{\varphi,\psi}$ *as in* (A.2)*. Suppose that* a *satisfies Assumption 2.1, and that* b *and* r *are as in Remarks 2.5 and 2.13, respectively. Then, for any open connected subset* Ω' *of* Ω*, there exists a constant* C_0 *that depends on* n, p, α, Ω *and* Ω' *such that:*

$$\|h\|_{\mathcal{W}^{2,p}(\Omega')} \leq C_0\left(\|h\|_{\mathcal{L}^p(\Omega)} + \|\xi - r(\cdot,\varphi,\psi)\|_{\mathcal{L}^p(\Omega)}\right).$$

Observe that $r(\cdot,\varphi,\psi)$ is in $\mathcal{L}^p(\Omega)$ whenever Ω is a bounded set. Indeed, by Assumption 2.12(b), we have that

$$\left(\int_{\Omega}|r(x,\varphi,\psi)|^p\,dx\right)^{1/p} \leq M\left(\int_{\Omega}(w(x))^p\,dx\right)^{1/p} \leq M\sup_{x\in\bar{\Omega}} w(x)|\bar{\Omega}| < \infty,$$

where $|\bar{\Omega}|$ denotes the volume of the closure of Ω.

Our final result ensures the convergence in the space of weighted functions $\mathcal{B}_w(\Omega)$ of certain sequences.

Lemma A.4. *Let* $\mathcal{B}_w(\Omega)$ *be the space in Definition 2.8 (with* Ω *rather than* $\mathbb{R}^n$*). In addition, consider a sequence* $\{v_m\}$ *of functions in* $\mathcal{B}_w(\Omega)$*, and suppose that there exists a real–valued function* v *on* Ω *such that* $v_m \to v$ *uniformly. Then* v *is in* $\mathcal{B}_w(\Omega)$*.*

Proof. By the triangle inequality and the uniform convergence of $\{v_m\}$ to v, for each $\epsilon > 0$, there exists a natural number N_ϵ such that, for all $m \geq N_\epsilon$,

$$|v(x)| - |v_m(x)| \leq |v(x) - v_m(x)| < \epsilon \ \text{ for all } \ x \in \Omega.$$

This yields

$$\begin{aligned} |v(x)| &< \epsilon + |v_m(x)| \\ &\leq \epsilon + \|v_m\|_w\, w(x) \ \text{ for all } \ x \in \Omega. \end{aligned}$$

Since $w \geq 1$, the claim follows. $\square$

Appendix B

Proof of Theorem 3.4

This Appendix requires definitions (3.1) and (3.2) (or (A.1) and (A.2)). We will also require the definition of Ω in Section 3.

B.1 Proof of Theorem 3.4(*i*)

We first show that there exist a function h in $\mathcal{W}^{2,p}(\Omega)$ and a subsequence $\{m_k\} \subset \{1, 2, ...\}$ such that, as $k \to \infty$, $h_{m_k} \to h$ weakly in $\mathcal{W}^{2,p}(\Omega)$ and strongly in $\mathcal{W}^{1,p}(\Omega)$. Namely, since $\mathcal{W}^{2,p}(\Omega)$ is reflexive [1, Theorem 3.5], then, by the Banach–Alaouglu Theorem [1, Theorem 1.17], the ball

$$\mathcal{H} := \left\{ h \in \mathcal{W}^{2,p}(\Omega) : \|h\|_{\mathcal{W}^{2,p}(\Omega)} \leq M_1 \right\} \tag{B.1}$$

is weakly sequentially compact. Hence, the compactness of the imbedding $\mathcal{W}^{2,p}(\Omega) \hookrightarrow \mathcal{W}^{1,p}(\Omega)$ [1, Theorem 6.2 part II] implies that $\mathcal{H}$ is precompact in $\mathcal{W}^{1,p}(\Omega)$, that is, there exist a function $h \in \mathcal{W}^{2,p}(\Omega)$ and a subsequence $\{h_{m_k}\} \equiv \{h_m\} \subset \mathcal{H}$ such that

$$h_m \to h \;\text{ weakly in }\; \mathcal{W}^{2,p}(\Omega) \;\text{ and strongly in }\; \mathcal{W}^{1,p}(\Omega). \tag{B.2}$$

The second step is to show that, as $m \to \infty$,

$$\sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}(\cdot, \varphi, \psi, h_m, \alpha_m) \to \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}(\cdot, \varphi, \psi, h, \alpha) \;\text{ in }\; \mathcal{L}^p(\Omega). \tag{B.3}$$

To this end, recall (A.1) and note that, given $x \in \Omega$, two functions $h \in \mathcal{W}^{2,p}(\Omega)$ and $h_m \in \mathcal{H}$, and a pair of positive numbers α and α_m, the following holds.

$$\begin{aligned}
& \left| \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\,(x, \varphi, \psi, h, \alpha) - \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\,(x, \varphi, \psi, h_m, \alpha_m) \right|^p \\
\leq \; & \sup_{(\varphi,\psi) \in V^1 \times V^2} \left| \hat{b}\,(x, \varphi, \psi, h, \alpha) - \hat{b}\,(x, \varphi, \psi, h_m, \alpha_m) \right|^p \\
= \; & \sup_{(\varphi,\psi) \in V^1 \times V^2} \left| \langle (\nabla h - \nabla h_m)\,(x), b(x, \varphi, \psi) \rangle - (\alpha h - \alpha_m h_m)\,(x) \right|^p \\
\leq \; & \left[\left| (\nabla h - \nabla h_m)\,(x) \right| \sup_{(\varphi,\psi) \in V^1 \times V^2} |b(x, \varphi, \psi)| + |(\alpha h - \alpha_m h_m)\,(x)| \right]^p \\
\leq \; & \left[2 \max \left\{ |(\nabla h - \nabla h_m)\,(x)| \sup_{(\varphi,\psi) \in V^1 \times V^2} |b(x, \varphi, \psi)|\,, |(\alpha h - \alpha_m h_m)\,(x)| \right\} \right]^p \\
\leq \; & 2^p \left[\left(|(\nabla h - \nabla h_m)\,(x)|\, C(\Omega) \right)^p + |(\alpha h - \alpha_m h_m)\,(x)|^p \right].
\end{aligned} \tag{B.4}$$

The existence of the constant $C(\Omega)$ follows from the boundedness of the set Ω and from Remark 2.5. Hence

$$\begin{aligned}
& \left| \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(x, \varphi, \psi, h, \alpha\right) - \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(x, \varphi, \psi, h_m, \alpha_m\right) \right|^p \\
\leq \;& 2^p \left[\left| \left(\nabla h - \nabla h_m\right)(x) \right| C(\Omega) \right]^p + 2^p \left[|h(x)| \left|\alpha - \alpha_m\right| + |\alpha_m| \left| \left(h - h_m\right)(x) \right| \right]^p \\
\leq \;& 2^p \left[\left| \left(\nabla h - \nabla h_m\right)(x) \right| C(\Omega) \right]^p + 2^p \left[2 \max\left\{ |h(x)| \left|\alpha - \alpha_m\right|, |\alpha_m| \left| \left(h - h_m\right)(x) \right| \right\} \right]^p \\
\leq \;& 2^p \left[\left| \left(\nabla h - \nabla h_m\right)(x) \right| C(\Omega) \right]^p + 4^p \left(|h(x)|^p \left|\alpha - \alpha_m\right|^p + |\alpha_m|^p \left| \left(h - h_m\right)(x) \right|^p \right) \\
\leq \;& [2C(\Omega)]^p n^p \left[\max_{1 \leq i \leq n} \left| \left(\partial_{x_i} h - \partial_{x_i} h_m \right)(x) \right| \right]^p + 4^p \left(|h(x)|^p \left|\alpha - \alpha_m\right|^p + |\alpha_m|^p \left| \left(h - h_m\right)(x) \right|^p \right) \\
\leq \;& [2C(\Omega) n]^p \left[\max_{1 \leq i \leq n} \left| \left(\partial_{x_i} h - \partial_{x_i} h_m \right)(x) \right| \right]^p + 4^p \left(|h(x)|^p \left|\alpha - \alpha_m\right|^p + |\alpha_m|^p \left| \left(h - h_m\right)(x) \right|^p \right). \qquad \text{(B.5)}
\end{aligned}$$

It follows from (B.5) that, for h_m and h as in (B.2) and $x \in \Omega$,

$$\begin{aligned}
& \left\| \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \varphi, \psi, h, \alpha\right) - \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \varphi, \psi, h_m, \alpha_m\right) \right\|^p_{\mathcal{L}^p(\Omega_T)} \\
\leq \;& [2C(\Omega) n]^p \, n \max_{1 \leq i \leq n} \left\| \partial_{x_i} h_m - \partial_{x_i} h \right\|^p_{\mathcal{L}^p(\Omega_T)} \\
& + \left(4 \|h\|_{\mathcal{L}^p(\Omega_T)} \right)^p \left| \alpha_m - \alpha \right|^p + (4\alpha_m)^p \left\| h_m - h \right\|^p_{\mathcal{L}^p(\Omega_T)}. \qquad \text{(B.6)}
\end{aligned}$$

Hence, as $m \to \infty$, it follows from (B.2) and hypothesis (d) that the right–hand side of (B.6) tends to zero; thus proving (B.3).

The existence of the constant $C(\Omega)$ in (B.4) also gives that for each l in $\mathcal{L}^{\frac{p}{p-1}}(\Omega)$,

$$\frac{1}{2} \left| \int_\Omega \left(l \cdot \operatorname{Tr}\left[a \mathbb{H}\left(h_m - h\right) \right] \right)(x) dx \right| \leq n^2 \frac{[C(\Omega)]^2}{2} \sum_{i,j=1}^{n} \left| \int_\Omega \left(l \cdot \partial^2_{x_i x_j} \left(h_m - h\right) \right)(x) dx \right|. \qquad \text{(B.7)}$$

Thus the weak convergence of $\{h_m\}$ to h in $\mathcal{W}^{2,p}(\Omega)$ yields that the right–hand side of (B.7) converges to zero as $m \to \infty$. Combining (B.3), (B.7), and hypothesis (c) to see that for every l in $\mathcal{L}^{\frac{p}{p-1}}(\Omega)$,

$$\int_\Omega \left(l \cdot \left[\hat{\mathbb{L}}_\alpha h - \xi \right] \right)(x) dx = \lim_{m \to \infty} \int_\Omega \left(l \cdot \left[\hat{\mathbb{L}}_{\alpha_m} h_m - \xi_m \right] \right)(x) dx = 0.$$

This fact, along with Theorem 2.10 in [66], implies (3.4), i.e.

$$\hat{\mathbb{L}}_\alpha h = \xi \text{ in } \Omega,$$

which completes the proof of part (*i*).

B.2 Proof of Theorem 3.4 (*ii*)

Let us introduce the following auxiliary results.

Lemma B.1. *[36, Theorem 9.19] Let $h \in \mathcal{W}^{2,p}(\Omega)$ be a solution of the equation $\hat{\mathbb{L}}_\alpha h = \xi$ in Ω. If the coefficients of $\hat{\mathbb{L}}_\alpha$ and ξ belong to $\mathcal{C}^{0,\beta}(\Omega)$, with $\beta \in]0,1[$, then h is in $\mathcal{C}^{2,\beta}(\Omega)$.*

Lemma B.2. *Consider a sequence of functions $\{f_m\}$ in $\mathcal{C}^{0,\eta}(\Omega)$, with $0 < \eta < 1$. Suppose the existence of*

- *a uniform bound H^* for the sequence $\{f_m\}$ in $\mathcal{C}^{0,\eta}(\Omega)$, i.e., for $m = 1, 2, ...$*

$$\|f_m\|_{\mathcal{C}^{0,\eta}(\Omega_T)} \leq H^*, \qquad \text{(B.8)}$$

- *a real valued function f on Ω, such that f_m converges uniformly to f on Ω; i.e., for every $\epsilon > 0$ there exists $M(\epsilon) \in \{1, 2, ...\}$ such that, for all $m \geq M(\epsilon)$ and all $x \in \Omega$,*

$$|f_m(x) - f(x)| < \epsilon. \qquad \text{(B.9)}$$

Then f *belongs to* $\mathcal{C}^{0,\eta}(\Omega)$.

Proof. Consider $f_m \in \mathcal{C}^{0,\eta}(\Omega)$ and observe that, by (B.8) for all $x, y \in \Omega$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_m(x)| + |f_m(x) - f_m(y)| + |f_m(y) - f(y)| \\ &\leq \sup_{x\in\Omega} |f(x) - f_m(x)| + H^*|x-y|^\eta + \sup_{y\in\Omega} |f_m(y) - f(y)|. \end{aligned}$$

Therefore, by (B.9), letting $m \to \infty$, $|f(x) - f(y)| \leq H^*|x-y|^\eta$ for all $x, y \in \Omega$. Hence f is in $\mathcal{C}^{0,\eta}(\Omega)$. ☐

To prove part (*ii*) of Theorem 3.4, we need to verify first that, if $p > n$, then $h_{m_k} \to h$ in $\mathcal{C}^{1,\eta}(\bar{\Omega})$ for all $\eta < 1 - \frac{n}{p}$.

By the Rellich–Kondrachov Theorem [1, Theorem 6.2, Part III], the imbedding $\mathcal{W}^{2,p}(\Omega) \hookrightarrow \mathcal{C}^{1,\eta}(\bar{\Omega})$, for $0 \leq \eta < 1 - \frac{n}{p}$ is compact; hence, it is also continuous. This implies that the set $\mathcal{H}$ in (B.1) is relatively compact in $\mathcal{C}^{1,\eta}(\bar{\Omega})$. Recall now that, from the proof of part (*i*), $\mathcal{H}$ is weakly sequentially compact. Hence, there exist $h \in \mathcal{W}^{2,p}(\Omega)$ and a subsequence $\{h_{m_k}\} \equiv \{h_m\} \subset \mathcal{H}$ such that h_m converges weakly to h in $\mathcal{W}^{2,p}(\Omega)$ and strongly in $\mathcal{C}^{1,\eta}(\bar{\Omega})$.

To complete the proof of part (*ii*), suppose that ξ is in $\mathcal{C}^{0,\beta}(\Omega)$ with $\beta \leq \eta < 1 - \frac{n}{p}$. We wish to show that the limit function h is in $\mathcal{C}^{2,\beta}(\Omega)$. To do this, we will proceed in several steps.

First we will show that, for each $m \geq 1$

$$\sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot, \varphi, \psi, h_m, \alpha_m\right) \text{ is in } \mathcal{C}^{0,\eta}(\Omega), \tag{B.10}$$

and that the sequence of functions

$$\sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot, \varphi, \psi, h_m, \alpha_m\right) \text{ is uniformly bounded on } \mathcal{C}^{0,\eta}(\Omega). \tag{B.11}$$

Afterwards we will show that

$$\sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot, \varphi, \psi, h_m, \alpha_m\right) \to \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot, \varphi, \psi, h, \alpha\right) \text{ uniformly on } \Omega. \tag{B.12}$$

Then we will invoke Lemma B.2 to conclude that

$$\sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot, \varphi, \psi, h, \alpha\right) \text{ is in } \mathcal{C}^{0,\eta}(\Omega). \tag{B.13}$$

Since we assumed that ξ is in $\mathcal{C}^{0,\beta}(\Omega)$, we will see that

$$\frac{1}{2}\mathrm{Tr}\left[a\mathbb{H}h\right] = \xi - \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot, \varphi, \psi, h, \alpha\right) \text{ is in } \mathcal{C}^{0,\beta}(\Omega). \tag{B.14}$$

Hence part (ii) will follow from Lemma B.1.

Let us proceed with the completion of the proof.

Recall the definition of $\mathcal{H}$ in (B.1). To prove that (B.10) and (B.11) hold, observe first that for all $x, y \in \Omega$, $\alpha_m > 0$ and $h_m \in \mathcal{H}$,

$$\begin{aligned} &\left| \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}(x, \varphi, \psi, h_m, \alpha_m) - \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}(y, \varphi, \psi, h_m, \alpha_m) \right| \\ \leq\ & \sup_{(\varphi,\psi)\in V^1\times V^2} \left| \hat{b}(x, \varphi, \psi, h_m, \alpha_m) - \hat{b}(y, \varphi, \psi, h_m, \alpha_m) \right| \\ =\ & \sup_{(\varphi,\psi)\in V^1\times V^2} \left| \langle \nabla h_m(x) - \nabla h_m(y), b\left(x, \varphi, \psi\right) \rangle + \langle \nabla h_m(y), b\left(x, \varphi, \psi\right) - b\left(y, \varphi, \psi\right) \rangle \right. \end{aligned}$$

$$-\alpha_m\left[h_m(x)-h_m(y)\right]+\left[r(x,\varphi,\psi)-r(y,\varphi,\psi)\right]\Big|$$

$$\begin{aligned}
\leq\ & \sup_{(\varphi,\psi)\in V^1\times V^2}|b(x,\varphi,\psi)|\,|\nabla h_m(x)-\nabla h_m(y)|+\sup_{(\varphi,\psi)\in V^1\times V^2}|b(x,\varphi,\psi)-b(y,\varphi,\psi)|\,|\nabla h_m(y)|\\
& +\alpha_m|h_m(x)-h_m(y)|+\sup_{(\varphi,\psi)\in V^1\times V^2}|r(x,\varphi,\psi)-r(y,\varphi,\psi)|\\
\leq\ & n\max_{1\leq i\leq n}|\partial_{x_i}h_m(x)-\partial_{x_i}h_m(y)|\,\tilde{C}(\bar{\Omega})+|\nabla h_m(y)|\,C_1|x-y|+\alpha_m|h_m(x)-h_m(y)|+C(R)|x-y|,
\end{aligned}$$

where $\tilde{C}(\bar{\Omega})$, C_1 and $C(R)$ are the constants in the Remarks 2.5 and 2.13. In this case, R stands for the radius of a ball B_R such that $\Omega\subset B_R$. Hence

$$\begin{aligned}
&\left|\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(x,\varphi,\psi,h_m,\alpha_m)-\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(y,\varphi,\psi,h_m,\alpha_m)\right|\\
\leq\ & n\|h_m\|_{\mathcal{C}^{1,\eta}(\Omega)}|x-y|^\eta\tilde{C}(\bar{\Omega})+|\nabla h_m(y)|\,C_1|x-y|+\alpha_m\|h_m\|_{\mathcal{C}^{1,\eta}(\Omega)}|x-y|^\eta+C(R)|x-y|. \qquad \text{(B.15)}
\end{aligned}$$

Now note that, by the continuity of the imbedding $\mathcal{W}^{2,p}(\Omega)\hookrightarrow\mathcal{C}^{1,\eta}(\bar{\Omega})$, there exists a so–called *imbedding constant* M_2 such that

$$\begin{aligned}
\max\left\{\sup_{x\in\bar{\Omega}}|h_m(x)|,\max_{1\leq i\leq n}\sup_{x\in\bar{\Omega}}|\partial_{x_i}h_m(x)|\right\} &\leq \|h_m\|_{\mathcal{C}^{1,\eta}(\bar{\Omega})}\\
&\leq M_2\|h_m\|_{\mathcal{W}^{2,p}(\Omega)}\\
&\leq M_2M_1 \qquad \text{(B.16)}
\end{aligned}$$

with M_1 as in the hypothesis (b). Therefore, combining (B.15) with (B.16) yields

$$\begin{aligned}
&\left|\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(x,\varphi,\psi,h_m,\alpha_m)-\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(y,\varphi,\psi,h_m,\alpha_m)\right|\\
\leq\ & M_2M_1|x-y|^\eta n\tilde{C}(\bar{\Omega})+C_1|x-y|nM_2M_1+\alpha_mM_2M_1|x-y|^\eta+C(R)|x-y|. \qquad \text{(B.17)}
\end{aligned}$$

Now, if $|x-y|<1$, then (B.17) yields

$$\left|\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(x,\varphi,\psi,h_m,\alpha_m)-\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(y,\varphi,\psi,h_m,\alpha_m)\right|\leq H_m^{(1)}|x-y|^\eta,$$

where $H_m^{(1)}:=M_1M_2n\tilde{C}(\bar{\Omega})+C_1nM_1M_2+\alpha_mM_1M_2+C(R)$. Observe that, since the sequence $\{\alpha_m\}$ converges, there exist a constant H_1^* such that $H_1^*\geq H_m^{(1)}$ for all $m\in\{1,2,...\}$, and so

$$\left|\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(x,\varphi,\psi,h_m,\alpha_m)-\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(y,\varphi,\psi,h_m,\alpha_m)\right|\leq H_1^*|x-y|^\eta. \qquad \text{(B.18)}$$

Otherwise, if $|x-y|\geq 1$, let $K^*:=\max_{x,y\in\bar{\Omega}}|x-y|$. Hence, from (B.17) again,

$$\left|\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(x,\varphi,\psi,h_m,\alpha_m)-\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(y,\varphi,\psi,h_m,\alpha_m)\right|\leq H_m^{(2)}|x-y|^\eta,$$

where $H_m^{(2)}:=M_1M_2n\tilde{C}(\bar{\Omega})+C_1K^*nM_1M_2+\alpha_mM_1M_2+C(R)K^*$. Yet, by the boundedness of the convergent sequence $\{\alpha_m\}$, there exists a constant H_2^* such that $H_2^*\geq H_m^{(2)}$ for all $m\in\{1,2,...\}$, and

$$\left|\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(x,\varphi,\psi,h_m,\alpha_m)-\sup_{\varphi\in V^1}\inf_{\psi\in V^2}\hat{b}(y,\varphi,\psi,h_m,\alpha_m)\right|\leq H_2^*|x-y|^\eta. \qquad \text{(B.19)}$$

From (B.18) and (B.19) we obtain (B.10) and (B.11).

We will now see that (B.12) holds. Let h_m and h be as in (B.2) and $\alpha_m \to \alpha$ as in hypothesis (d). Then, by (A.1),

$$
\begin{aligned}
& \sup_{x\in\bar{\Omega}} \left| \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\,(x,\varphi,\psi,h_m,\alpha_m) - \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\,(x,\varphi,\psi,h,\alpha) \right| \\
\leq\; & \sup_{x\in\bar{\Omega}} \sup_{(\varphi,\psi)\in V^1\times V^2} \left| \hat{b}\,(x,\varphi,\psi,h_m,\alpha_m) - \hat{b}\,(x,\varphi,\psi,h,\alpha) \right| \\
\leq\; & \tilde{C}(\bar{\Omega}) \sum_{i=1}^{n} \sup_{x\in\bar{\Omega}} \left| (\partial_{x_i} h_m - \partial h_{x_i})\,(x) \right| + \sup_{x\in\bar{\Omega}} |h_m(x)| \cdot |\alpha_m - \alpha| + \alpha \sup_{x\in\bar{\Omega}} |(h_m - h)\,(x)|\,.
\end{aligned}
\tag{B.20}
$$

Moreover, by (B.16) we have $\sup_{x\in\bar{\Omega}} |h_m(x)| \leq M_1 M_2$. Consequently, since h_m converges to h in $\mathcal{C}^{1,\beta}(\bar{\Omega})$, the right hand side of (B.20) tends to zero as $m \to \infty$. This gives (B.12).

Since (B.10) and (B.12) hold, we may use Lemma B.2 to assert that (B.13) holds.

To conclude the proof of (*ii*), observe that a direct calculation yields (B.14), and so Lemma B.1 completes the proof of part (*ii*). Hence, the proof of Theorem 3.4 is now complete. □

B.3 Notes on the proof of Theorem 4.19

The proof of Theorem 4.19 resembles that of Theorem 3.4. We will state a few remarks on the details. Recall the definition of $\hat{b}$ given in (4.37), that is

$$
\hat{b}\left(x,\alpha,\varphi,\psi,h^1,...,h^N\right) := \left\langle \nabla h^i(x), b\,(x,\alpha,\varphi,\psi) \right\rangle - \alpha h^i(x) + r\,(x,\alpha,\varphi,\psi) + \sum_{j=1}^{N} q_{ij} h^j(x).
$$

Our first step is to show the existence of a subsequence $\left\{h^j_{m_k}\right\}$ of $\left\{h^j_m\right\}_m$, whose convergence to h^j is weak in $\mathcal{W}^{2,p}(\Omega)$ and strong in $\mathcal{W}^{1,p}(\Omega)$ for $j = 1, ..., N$.

To do this, we repeat, for $j = 1, ..., N$, the argument that led us to (B.2). Then, it is necessary to prove that, as $m \to \infty$,

$$
\sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot,\cdot,\varphi,\psi,h^1_m,...,h^N_m\right) \to \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(\cdot,\cdot,\varphi,\psi,h^1,...,h^N\right)
\tag{B.21}
$$

in $\mathcal{L}^p(\Omega)$. Observe that (B.4) changes to

$$
\begin{aligned}
& \left| \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1_m,...,h^N_m\right) - \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1,...,h^N\right) \right|^p \\
\leq\; & \sup_{(\varphi,\psi)\in V^1\times V^2} \left| \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1,...,h^N\right) - \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1_m,...,h^N_m\right) \right|^p \\
\leq\; & \sup_{(\varphi,\psi)\in V^1\times V^2} \left| \left\langle \left(\nabla h^i_m - \nabla h^i_m\right)(x), b(x,\alpha_i,\varphi,\psi) \right\rangle + \sum_{j=1}^{N} \left(h^j_m(x) - h^j(x)\right) q_{ij} \right|^p \\
\leq\; & \left(\left| \left(\nabla h^i_m - \nabla h^i_m\right)(x) \right| \tilde{C}(\bar{\Omega}) + \left| \sum_{j=1}^{N} \left(h^j_m(x) - h^j(x)\right) q_{ij} \right| \right)^p \\
\leq\; & \left(\left| \left(\nabla h^i_m - \nabla h^i_m\right)(x) \right| \tilde{C}(\bar{\Omega}) + \left| \max_{1\leq j\leq N} \left(h^j_m(x) - h^j(x)\right) \sum_{j=1}^{N} q_{ij} \right| \right)^p
\end{aligned}
$$

But, by the conservativeness of the chain $\alpha(\cdot)$, we have that $\sum_{j=1}^{N} q_{ij} = 0$. Hence,

$$
\left| \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1_m,...,h^N_m\right) - \sup_{\varphi\in V^1} \inf_{\psi\in V^2} \hat{b}\left(x,\alpha_i,\varphi,\psi,h^1,...,h^N\right) \right|^p
$$

$$\leq \quad \left(\left|\left(\nabla h_m^i - \nabla h_m^i\right)(x)\right| \tilde{C}(\bar{\Omega})\right)^p \to 0 \text{ as } m \to \infty.$$

Thus, using the same steps as (B.6), we prove (B.21).

The relation (B.21), an analogous inequality to (B.7), hypothesis (c) and [66, Theorem 2.10] yields (4.38). (See the last paragraph of Section B.1.)

To prove part (*ii*), we need to verify first that, if $p > n$, then $h^j_{m_k} \to h^j$ in $\mathcal{C}^{1,\eta}(\Omega)$ for $0 \leq \eta < 1 - \frac{n}{p}$ and $j = 1, ..., N$. This is essentially the same we did in the first part of Section B.2.

Afterwards, we assume that $\xi \in \mathcal{C}^{0,\beta}(\Omega)$ with $\beta \leq \eta < 1 - \frac{n}{p}$, and show that the corresponding relations to (B.10) and (B.11) (in the present context) hold, that is:

$$\sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \alpha_i, \varphi, \psi, h_m^1, ..., h_m^N\right) \text{ is in } \mathcal{C}^{0,\eta}(\Omega), \tag{B.22}$$

and that the sequence of functions

$$\sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \alpha_i, \varphi, \psi, h_m^1, ...h_m^N\right) \text{ is uniformly bounded on } \mathcal{C}^{0,\eta}(\Omega). \tag{B.23}$$

To this end, it is necessary to obtain the following analogous of (B.15).

$$\begin{aligned} &\left| \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(x, \alpha_i, \varphi, \psi, h_m^1, ..., h_m^N\right) - \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(y, \alpha_i, \varphi, \psi, h_m^1, ..., h_m^N\right) \right| \\ \leq \quad & n \left\|h_m^i\right\|_{\mathcal{C}^{1,\eta}(\bar{\Omega})} |x-y|^\eta \tilde{C}(\bar{\Omega}) + \left|\nabla h_m^i(x)\right| C_1 |x-y| + \alpha_i \left\|h_m^i\right\|_{\mathcal{C}^{1,\eta}(\bar{\Omega})} |x-y|^\eta + C(R)|x-y|. \end{aligned}$$

This relation can be taken as the point of departure to argue as in (B.16)–(B.19), and therefore, obtain (B.22) and (B.23).

The convergence of $h^j_{m^i_k}$ to h^j, for $j = 1, ..., N$, yields

$$\sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \alpha_i, \varphi, \psi, h_m^1, ..., h_m^N\right) \to \sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \alpha_i, \varphi, \psi, h^1, ..., h^N\right) \text{ uniformly on } \Omega,$$

and by Lemma B.2 we deduce that

$$\sup_{\varphi \in V^1} \inf_{\psi \in V^2} \hat{b}\left(\cdot, \alpha_i, \varphi, \psi, h^1, ..., h^N\right) \text{ is in } \mathcal{C}^{0,\eta}(\Omega).$$

This last assertion, combined with an analogous relation to (B.14), and Lemma B.1 yield part (*ii*).

Bibliography

[1] Adams, R.A. (1975) *Sobolev Spaces*. Academic Press. New York.

[2] Akella, R. and Kumar, P.R. (1985) *Optimal control of production rate in a failure prone manufacturing system*. IEEE Trans. Automatic Control. **31**, 116–126.

[3] Arapostathis, A., Borkar, V.S. (2010) *Uniform recurrence properties of controlled diffusions and applications to optimal control*. SIAM J. Control Optim. **48**, no. 7, 4181–4223.

[4] Arapostathis, A., Ghosh, M.K., Borkar, V.S. (2011) *Ergodic Control of Diffusion Processes*. Cambridge University Press.

[5] Arapostathis, A. (2012) *On the policy iteration algorithm for non–degenerate controlled diffusions under the ergodic criterion*. In Optimization, Control and Applications of Stochastic Systems, edited by Daniel Hernández–Hernández and Adolfo Minjárez–Sosa. Birkhäuser, New York.

[6] Atsumi, H. (1965) *Neoclassical growth and the efficient program of capital accumulation*. Rev. Econ. Stud. **32**, 127–136.

[7] Bellman, R. (1957) *Dynamic Programming*. Princeton University Press, Princeton, New Jersey.

[8] Bensoussan, A. (1998) *Perturbation Methods in Optimal Control.* Wiley, New York.

[9] Bensoussan, A. and Frehse, J. (2000) *Regularity Results for Nonlinear Elliptic Systems and Applications.* Springer. Berlin.

[10] Bertsekas, D.P. and Shreve, S.E. (1996) *Stochastic Optimal Control: The Discrete–Time Case*. Athena Scientific, Belmont, Massachusetts.

[11] Billingsley, P. (1999) *Convergence of Probability Measures*. Wiley Series in Probability and Statistics, New York. 2nd edition.

[12] Bhatt, A.G. and Karandikar, R.L. (1993) *Invariant measures and evolution equations for Markov processes characterized via martingale problems.* Ann. Probab. **21**, no. 4, 2246–2268.

[13] Bismut, J. (1978) *An aproximation method in optimal stochastic control.* SIAM Journal on Control and Optimization. **16**, 122–130.

[14] Borkar, V. S. (1989) *A topology for Markov Controls*. Appl. Math. Optim. **20**, 55–62.

[15] Borkar, V.S. (2005) *Controlled diffusion processes.* Probab. Surveys **2**, 213–244.

[16] Borkar, V.S., Ghosh, M.K. (1988) *Ergodic control of multidimensional diffusions I: The existence of results*. SIAM J. Control Optim. **26**, 112–126.

[17] Borkar, V.S. and Ghosh, M.K. (1990) *Ergodic control of multidimensional diffusions II: Adaptive control.* Appl. Math. Opt. **21**, 191–220.

[18] Borkar, V.S. and Ghosh, M.K. (1992) *Stochastic differential games: an occupation measure based approach.* J. Optimization Theory Appl. **73**, 359–385. Correction: **88** (1996), 251–252.

[19] Brock, W.A. (1977) *Differential games with active and passive variables*. Mathematical Economics and Game Theory; Essays in honor of Oscar Morgenstern. Edited by Henn R. and Moesclin O., Springer–Verlag, Berlin. 34–52.

[20] Carlson, D. (2005) *Normalized Overtaking Nash Equilibrium for a Class of Distributed Parameter Dynamic Games.* Advances in Dynamic Games (A.S. Nowak and K. Swajowski, eds.), Birkhauser, Boston.

[21] Carlson, D., Haurie, A. (1995) *A turnpike theory for infinite horizon open–loop differential games with decoupled controls*. New Trends in Dynamic Games and Applications (G.J. Olsder, editor) Annals of the ISDG **3**, Birkhauser, Boston, 353–376.

[22] Elliott, R.J., Kalton, N.J. (1972) *Values in differential games* Bull. Amer. Math. Soc. **78** 427–431

[23] Escobedo–Trujillo, B.A., Hernández–Lerma, O. (2011) *Overtaking optimality for controlled Markov–modulated diffusions*. Submitted.

[24] Escobedo–Trujillo, B.A., López–Barrientos, J.D., Hernández–Lerma, O. (2012) *Bias and overtaking equilibria for stochastic differential games*. J. Optimization Theory and Applications. DOI 10.1007/s10957-011-9974-4.

[25] Ethier, S.N., Kurtz, T.G. (1986) *Markov Processes: Characterization and Convergence.* Wiley, New York.

[26] Fleming, W.H. (1963) *Some Markovian optimization problems*. J. Math. Mech. **12**, 131–140.

[27] Friedman, A. (1972) *Stochastic differential games.* J. Differential Equations. **11**, 79–108.

[28] Fort, G. and Roberts, G.O. (2005) *Subgeometric ergodicity of strong Markov processes.* Ann. Appl. Prob. **15**, 1565–1589.

[29] Fujita, Y., Morimoto, H. (1987) *On bang–bang solutions of stochastic differential games*. IEEE Trans. Autom. Control. **32**, 535–537.

[30] Ghosh, M.K., Arapostathis, A. and Marcus, S.I. (1993) *Ergodic control of switching diffusions to flexible manufacturing systems*. SIAM J. Control Optim. **31**, 1183–1204.

[31] Ghosh, M.K., Arapostathis, A. and Marcus, S.I. (1997) *Ergodic control of switching diffusions*. SIAM J. Control Optim. **35**, 1962–1988.

[32] González–Hernández, J., López–Martínez, R. R. and Pérez–Hernández, J. R. (2007) *Markov control processes with randomized discounted cost*. Math. Meth. Oper.Res. **65**, 27–44.

[33] González–Hernández, J., López–Martínez, R. R. and Minjárez–Sosa, J. A. (2008) *Adaptive policies for stochastic systems under a randomized discounted cost criterion.* Boletín de la Sociedad Matemática Mexicana **14**, 149–163.

[34] Ghosh, M.K., Marcus, S.I. (1991) *Infinite horizon controlled diffusions problems with nonstandard criteria.* J. Math. Systems Estim Control **1**, 45–69.

[35] Ghosh, M.K., Arapostathis, A., Marcus, S.I. (1992) *Optimal control of switching diffusions with application to flexible manufacturing systems.* J. Control and Optim. **30**, 1–23.

[36] Gilbarg, D., Trudinger, N.S. (1998) *Elliptic Partial Differential Equations of Second Order.* Reprinted version, Springer. Heidelberg.

[37] Glynn, P.W. and Meyn, S.P. (1996) *A Lyapunov bound for solutions of the Poisson equation*. Ann. Prob. **24**, 916–931.

[38] Guo, X.P., Hernández–Lerma, O. (2003) *Continuous–time Markov chains with discounted rewards*. Acta Appl. Math. **79**, 195–216.

[39] Guo, X.P., Hernández–Lerma, O. (2003) *Drift and monotonicity conditions for continuous–time controlled Markov chains with average criterion*. IEEE Trans. Automatic Control. **48**, 236–244.

[40] Guo, X.P., Hernández–Lerma, O. (2003) *Zero–sum games for continuous–time Markov chains with unbounded transition and average payoff rates*. J. Appl. Prob. **40**, 327-345

[41] Hamadéne, S., Lepeltier, J. P. (1995) *Backward equations, stochastic control and zero-sum stochastic differential games*. Stochastics An International Journal of Probability and Stochastic Processes. **54**, 221–231.

[42] Has'minskii, R.Z. (1980) *Stochastic Stability of Differential Equations*. Sijthoff and Noordhoff, Germantown, Md.

[43] Hashemi, S.N., Heunis, A.J. (2005) *On the Poisson equation for singular diffusions*. Stochastics **77**, 155–189.

[44] Hernández–Lerma, O. (1994) *Lectures on Continuous–Time Markov Control Processes.* Sociedad Matemática Mexicana, Mexico City.

[45] Hernández–Lerma, O., Lasserre, J.B. (1996) *Discrete–Time Markov Control Processes: Basic Optimality Criteria.* Springer, New York.

[46] Hernández–Lerma, O., Lasserre, J.B. (1997) *Policy iteration for average cost Markov control processes on Borel spaces*. Acta Appl. Math. **47**, 125–154.

[47] Hernández–Lerma, O., Govindan, T.E. (2001) *Nonstationary continuous–time Markov control processes with discounted costs on infinite horizon.* Acta Appl. Math. **67**, 277–293.

[48] Hoffman, A.K., Karp, R.M. (1966) *On nonterminating stochastic games*. Management Science, **12**, 359–370.

[49] Howard, R.A. (1960) *Dynamic Programming and Markov Processes*. MIT Press, Cambridge, Massachusetts.

[50] Jasso–Fuentes, H. (2004) *Non–cooperative Markov games.* Morfismos **9**, 39–54.
Available at `http://www.morfismos.cinvestav.mx/Artículos/Volumen9Número1.aspx`

[51] Jasso–Fuentes, H. (2007) *Infinite–horizon optimal control problems for Markov diffusion processes.* Sc.D. dissertation. Mathematics Department, CINVESTAV–IPN, México.

[52] Jasso–Fuentes H., Hernández–Lerma O. (2008) *Characterizations of overtaking optimality for controlled diffusion processes*. Appl. Math. Optim. **57**, 349–369.

[53] Jasso–Fuentes H., Hernández–Lerma O. (2009) *Ergodic control, bias, and sensitive discount optimality for Markov diffusion processes*. Stochastic Analysis and Applications, **27**, 363–385.

[54] Jasso–Fuentes, H., Hernández–Lerma, O. (2009) *Blackwell optimality for controlled diffusion processes*. J. Appl. Prob. **46**, 372–391.

[55] Jasso–Fuentes, H., Yin, G. (2012) *Advanced criteria for Markov–modulated diffusions in infinite horizon: Overtaking, bias and Blackwell optimality*. Submitted.

[56] Kawaguchi, K. and Morimoto, H. (2007) *Long–run average welfare in a pollution accumulation model*. Journal of Economic Dynamics and Control. **31**, 703-720.

[57] Klebaner, F. C. (2005) *Introduction to Stochastic Calculus with Applications.* Imperial College Press, 2nd. ed. London.

[58] Kolokoltsov, V. N. (2010) *Nonlinear Markov Processes and Kinetic Equations*. Cambridge University Press.

[59] Kolokoltsov, V. N. (2011) *Markov Processes, Semigroups and Generators*. DeGruyter tracts in Mathematics. DeGruyter.

[60] Kushner, H.J. (1967) *Optimal discounted stochastic control for diffusion processes.* SIAM J. Control **5**, 520–531.

[61] Kushner, H.J. (1978) *Optimality conditions for the average cost per unit time problem with a diffusion model*. SIAM J. Control Optim. **16**, 330–346.

[62] Kushner, H.J. (2002) *Numerical approximations for stochastic differential games*. SIAM J. Control Optim. **41**, 457–486.

[63] Kushner, H.J. (2004) *Numerical approximations for stochastic differential games: the ergodic case*. SIAM J. Control Optim. **42**, 1911–1933.

[64] Ladyzhenskaya O.A., Uraltseva, N.N. (1968) *Linear and Quasilinear Elliptic Equations*. Academic Press. New York.

[65] Ladyzhenskaya O.A., Uraltseva, N.N. (1968) *Linear and Quasilinear Equations of Parabolic Type*. Translations of Mathematical Monographs, 23. Providence, RI. AMS.

[66] Lieb, E.H., Loss, M. (2001) *Analysis*. Second Edition. American Mathematical Society. Providence, R.I.

[67] Mao, X., Yuan, C. (2006) *Stochastic Differential Equations with Markovian Switching.* Imperial College Press. London.

[68] Mendoza–Pérez, A. F., Hernández–Lerma, O. (2010) Markov control processes with pathwise constraints. Math. Meth. Oper. Res. **71**, p. 477–502.

[69] Meyn, S.P., Tweedie, R.L. (1993) *Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous–time processes.* Adv. Appl. Prob. **25**, 518–548.

[70] Morimoto, H., Ohashi, M. (1987) *On bang–bang solutions of stochastic differential games*. IEEE Trans. Autom. Control **32**, 535–537.

[71] Morimoto, H., Okada, M. (1999) *Some results on the Bellman equation of ergodic control*. SIAM J. Control Optim. **38**, 159–174.

[72] Nowak, A.S. (1985) *Measurable selection theorems for minimax stochastic optimization problems*. SIAM J. Control Optimization **23**, 466–476.

[73] Nowak, A.S. (1999) *Sensitive equilibria for ergodic stochastic games with countable state spaces*. Math. Meth. Oper. Res. **50**, 65–76.

[74] Nowak, A.S. (1999) *Optimal strategies in a class of zero–sum ergodic stochastic games*. Math. Meth. Oper. Res. **50**, 399–419.

[75] Nowak, A.S. (2008) *Equilibrium in a dynamic game of capital accumulation with the overtaking criterion*. Econom. Lett., **99**, 233–237.

[76] Nowak, A.S., Vega–Amaya, O. (1998) *A counterexample on overtaking optimality*. Math. Meth. Oper. Res. **49**, 435–439.

[77] Øksendal, B. (2005) *Stochastic Differential Equations. An Introduction with Applications*. Springer, 5th ed. Berlin.

[78] Prieto–Rumeau, T., Hernández–Lerma, O. (2005) *Bias and overtaking equilibria for zero-sum continuous–time Markov games*. Math. Meth. Oper. Res. **61**, 437–454.

[79] Prieto–Rumeau, T., Hernández-Lerma, O. (2006) *A unified approach to continuous–time discounted Markov control processes*. Morfismos **10**, 1–40.
Available at `http://chucha.math.cinvestav.mx/morfismos/v10n1/uni.pdf`

[80] Prieto–Rumeau, T., Hernández–Lerma, O. (2006) *Bias optimality for continuous–time controlled Markov chains.* SIAM J. Control Optim. **45**, 51–73.

[81] Puterman, M.L. (1977) *Optimal control of diffusion processes with reflection.* J. Optim. Theory and Appl. **22**, 103–116.

[82] Puterman, M.L. (1981) *On the convergence of policy iteration for controlled diffusions.* J. Optim. Theory and Appl. **33**, 137–144.

[83] Ramsey, F.P. (1928) *A mathematical theory of savings*. Econ. J. **38**, 543–559.

[84] Reed, M., Simon, B. (1975) *Methods of Modern Mathematical Physics*, v.2. Harmonic analysis. Academic Press, N.Y.

[85] Rogers, L.C.G., Williams, D. (2000) *Diffusions, Markov Processes and Martingales, Vol.1, Foundations.* Cambridge University Press.

[86] Rubinstein, A. (1979) *Equilibrium in supergames with the overtaking criterion*. J. Econom. Theory. **21**, 1–9.

[87] Schäl, M. (1974) *A selection theorem for optimization problems*. Arch. Math. **XXV**, 219–224.

[88] Schäl, M. (1975) *Conditions for optimality in dynamic programming of n–stage optimal policies to be optimal.* Z. Wahrs. verw. Geb **32**, 179–196.

[89] Sion, M. (1958) *On general minimax theorems.* Pacific J. Math., **8**, 171–176.

[90] Song, Q.S., Yin, G., Zhang, Z. (2006) *Numerical methods for controlled regime–switching diffusions and regime–switching jump diffusions.* Automatica, **42**, 1147-1157.

[91] Song, Q.S., Yin, G., Zhang, Z. (2008) *Numerical solutions for stochastic differential games with regime switching*. IEEE Trans. Automat. Control **53** 509–521.

[92] Swiech, A. (1996) *Another approach to the existence of value functions of stochastic differential games.* J. Mathematical Analysis and Applications. **204**, 884-897.

[93] Taylor, H. M. (1965) *Markovian sequential replacement processes*, Ann. Math. Stat. **36**, 1677–1694.

[94] Van der Wal, J. (1978) *Discounted Markov games: Generalized policy iteration method*. J. Optim. Theory and Applications. **25**, 125–138.

[95] Von Weizsäcker, C.C. (1965) *Existence of optimal programs of accumulation for an infinite horizon*. Rev. Econ. Stud. **32**, 85–104.

[96] Veretennikov, A.Y. and Klokov, S.A. (2005) *On the subexponencial rate of mixing Markov processes.* Theory Prob. Appl. **49**, 110–122.

Aprobación

El jurado designado por el Departamento de Matemáticas del CINVESTAV aprobó esta tesis el día 13 de junio de 2012.

Dr. Onésimo Hernández Lerma	**Dr. Héctor Jasso Fuentes**
Dr. Carlos Pacheco González	**Dr. Eduardo Santillán Zerón**
Dr. Alin Andrei Carsteanu	**Dr. Carlos Ibarra Valdéz**