STOCHASTIC DIFFERENTIAL GAMES

Theory and Applications

KANDETHODY M. RAMACHANDRAN, CHRIS P. TSOKOS

ATLANTIS STUDIES
IN PROBABILITY AND STATISTICS
SERIES EDITOR | C. P. TSOKOS

ATLANTIS PRESS

Atlantis Studies in Probability and Statistics

Volume 2

Series Editor: **Chris P. Tsokos**

Atlantis Studies in Probability and Statistics

Series Editor:

Chris P. Tsokos,

University of South Florida Tampa,

Tampa, USA

(ISSN: 1879-6893)

Aims and scope of the series

The Series 'Atlantis Studies in Probability and Statistics' publishes studies of high-quality throughout the areas of probability and statistics that have the potential to make a significant impact on the advancement in these fields. Emphasis is given to broad interdisciplinary areas at the following three levels:
(I) Advanced undergraduate textbooks, i.e., aimed at the 3rd and 4th years of undergraduate study, in probability, statistics, biostatistics, business statistics, engineering statistics, operations research, etc.;
(II) Graduate level books, and research monographs in the above areas, plus Bayesian, nonparametric, survival analysis, reliability analysis, etc.;
(III) Full Conference Proceedings, as well as Selected topics from Conference Proceedings, covering frontier areas of the field, together with invited monographs in special areas.
All proposals submitted in this series will be reviewed by the Editor-in-Chief, in consultation with Editorial Board members and other expert reviewers

For more information on this series and our other book series, please visit our website at:

www.atlantis-press.com/publications/books

ATLANTIS PRESS

PARIS – AMSTERDAM – BEIJING

Stochastic Differential Games

Theory and Applications

Kandethody M. Ramachandran, Chris P. Tsokos

University of South Florida,
Department of Mathematics and Statistics
4202 E. Fowler Avenue,
Tampa, FL 33620-5700, USA

ATLANTIS PRESS

PARIS – AMSTERDAM – BEIJING

Atlantis Press

8, square des Bouleaux
75019 Paris, France

For information on all Atlantis Press publications, visit our website at: *www.atlantis-press.com*

Atlantis Studies in Probability and Statistics

Volume 1: Bayesian Theory and Methods with Applications - Vladimir P. Savchuk, C.P. Tsokos

ISBNs
Print: 978-94-91216-46-6
E-Book: 978-94-91216-47-3
ISSN: 1879-6893

Dedications to our families:

Usha, Vikas, Vilas and Varsha Ramachandran

and

Debbie, Mathew, Jonathan, and Maria Tsokos

Preface

Conflicts in the form of wars, or competition among countries and industrial institutions are plenty in human history. The introduction of game theory in the middle of the twentieth century shed insights and enabled researchers to analyze this subject with mathematical rigor. From the ground-breaking work of VonNeumann and Morgenston, modern game theory evolved enormously. In the last few decades, Dynamic game theory framework has been deepened and generalized from the pioneering work on differential games by R. Isaacs, L.S. Pontryagin and his school, and on stochastic games by Shapley. This book will expose the reader to some of the fundamental methodology in non-cooperative game theory, and highlight some numerical methods, along with some relevant applications.

Since the early development days, differential game theory has had a significant impact in such diverse disciplines as applied mathematics, economics, systems theory, engineering, operations, research, biology, ecology, environmental sciences, among others. Modern game theory now relies on wide ranging mathematical and computational methods, and relevant applications that are rich and challenging. Game theory has been widely recognized as an important tool in many fields. Importance of game theory to economics is illustrated by the fact that numerous game theorists, such as John Forbes Nash, Jr., Robert J. Aumann and Thomas C. Schelling, have won the Nobel Memorial Prize in Economics Sciences. Simply put, game-theory has the potential to reshape the analysis of human interaction.

In Chapter 1, we will present a general introduction, survey, and background material for stochastic differential games. A brief introduction of Linear pursuit-Evation differential games will be given in Chapter 2 for a better understanding of the subject concepts. Chapter 3 will deal with two person Zero-sum stochastic differential games and various solution methods. We will also introduce games with multiple modes. Formal solutions for some classes of stochastic linear pursuit-evasion games will be given in Chapter 4. In Chapter 5, we will discuss N-person stochastic differential games. Diffusion models are in general not very good approximations for real world problems. In order to deal with those issues, we will introduce weak convergence methods for two person to the stochastic differential games in Chapter 6. In Chapter 7, will cover weak convergence methods for many player games. In Chapter 8, we will introduce some useful numerical methods for two different payoff structure; discounted payoff and ergodic payoff as well as the case of nonzero sum games. We will conclude the book in Chapter 9 by giving some real world applications of stochastic differential games to finance and competitive advertising.

We wish to express our sincere appreciation to the reviewers of the preliminary manuscript of the book for their excellent comments and suggestions.

Dr. M. Sambandham, Professor of Mathematics, Chief Editor, International Journal of Systems and applications.

Dr. G.R. Aryal, Assistant Professor of Statistics, Purdue University, Calumet, Indiana.

Dr. Rebecca Wooten, Assistant Professor of Mathematics & Statistics, University of South Florida, Tampa, Florida.

Dr. V. Laksmikatham, Distinguished Professor of Mathematics, Emeritus, Florida Institute of Technology.

Dr. Yong Xu, Assistant Professor of Mathematics at Radford University,

Dr. Kannan, Professor of Mathematics – Emeritus, University of Georgia.

Dr. Geoffrey O. Okogbaa, Professor of Industrial Engineering and Management Science, University of South Florida, Tampa, Florida.

We would also like to thank the editorial staff of Atlantis Press, in particular, the project manager Mr. Willie van Berkum.

Finally, a very special thanks to Beverly DeVine-Hoffmeyer for her excellent work in typing this book.

K.M. Ramachandran
C.P. Tsokos

Contents

Chapter 1

Introduction, Survey and Background Material

1.1 Introduction

Game theory has emerged out of the growing need for scientists and economists to have better grasp of the real world in today's technological revolution. Game theory deals with tactical interactions among multiple decision makers. These interactions can range from completely non-cooperative to completely cooperative. These decision makers are usually referred as players or agents. Each player either tries to maximize (in which case the objective function is a *utility function* or *benefit function*) or minimize (in which case the objective function is called a *cost function* or a *loss function*) using multiple alternatives (actions, or equivalently decision variable). If the players were able to enter into a cooperative agreement so that the selection of actions or decisions is done collectively and with full trust, so that all players would benefit to the extent possible, and no inefficiency would arise, then we would be in the realm of cooperative game theory. The issues of bargaining, coalition formation, excess utility distribution, etc. are of importance in cooperative game theory. However cooperative game theory will not be covered in this book. This book will only deal with non-cooperative game theory, where no cooperation is allowed among the players.

The origin of game theory and their development could be traced to the pioneering work of John Von Neumann and Oskar Morgenston [201] published in 1944. Due to the introduction of guided interceptor missiles in 1950s, the questions of pursuit and evasion took center stage. The mathematical formulation and study of the differential games was initiated by Rufus Isaacs, who was then with the Mathematics department of the RAND Corporation, in a series of RAND Corporation memoranda that appeared in 1954, [90]. This work and his further researches were incorporated into a book [91] which inspired much further work and interest in this area. After the Oscar film called "A Beautiful Mind" was released by

Universal Pictures in the year 2001, a great majority of the people started paying attention to the game theory and its usefulness. This film is about John Forbes Nash. Game theorists use the concept of Nash equilibrium to analyze outcomes of strategic interaction of two or more decision makers, Browne [33], Ho *et al.* [89], Sircar [177], and Yavin [211, 212], Yeung [214, 215, 216]. Nash's theory of non-cooperative games, [139, 140] is now recognized as one of the outstanding intellectual advances of the twentieth century, [138]. The formulation of Nash equilibrium has had a fundamental impact in economics and the social sciences.

The relationship between differential games and optimal control theory and the publication of Isaacs [91] at a time when interest in optimal control theory was very great served to further stimulate interest in differential games, Berkovitz [25]. For a good coverage on the connection between control theory and game theory, readers are referred to Krasovskii and Subbotin [100]. Earlier works on differential games and optimal control theory appeared almost simultaneously, independently of each other. At first, it seems natural to view a differential game as a control process where the controls are divided among various players who are willing to use them for objectives which possibly conflict with each other. However a more deeper study will reveal that the development of the two fields followed different paths. Both have the evolutionary aspect in common, but differential games have in addition a game-theoretic aspect. As a result, the techniques developed for the optimal control theory cannot be simply reused.

In the 1960s researchers started working on what have been called stochastic differential games. These games are stochastic in the sense that noise is added to the players' observations of the state of the system or to the transition equation itself. A stochastic differential game problem was solved in Ho [87] using variational techniques where one player controlled the state and attempted to minimize the error and confuse the other player who could only make noisy measurements of the state and attempted to minimize his/her error estimate. Later in Basar and Haurie [15], a problem of pursuit-evasion is considered where the pursuer has perfect knowledge whereas the evader can only make noisy measurements of the state of the game. In Bafico [5], Roxin and Tsokos [170], a definition of stochastic differential game is given. A connection between stochastic differential games and control theory is discussed in Nocholas [141]. In the 1970s rigorous discussion of existence and uniqueness results for stochastic differential games using martingale problem techniques and variational inequality techniques ensued, Elliot [47, 48, 49, 50], Bensoussan and Lions [22], Bensoussan and Friedman [23, 24], among many others. There are many aspects of differential games such as pursuit evasion games, zero-sum games, cooperative and non-cooperative games and other types of dynamic games. For some survey papers on such diverse topics as pursuit-evasion games, viscosity solutions, discounted stochastic games, numerical methods, and others, we refer to Bardi and Raghavan [7], which serves as a rich source of information on these topics. In this article we will restrict ourselves to mostly strictly non-cooperative stochastic differential games.

The early works on differential games are based on the dynamic programming method now called as Hamiltonian-Jacobi-Isaacs (HJI). Many authors worked on making the concept of value of a differential game precise and providing a rigorous derivation of HJI equation, which does not have a classical solution in most cases. For HJI equations smooth solutions do not exist in general and nonsmooth solutions are highly nonunique. Some of the works in this direction include, Berkovitz [25], Fleming [61], Elliott [47, 49], Firedman [67], Kalton, Krasovskii, and Subbotin [95], Roxin and Tsokos [182], Uchida [197], Varaiya [198, 199]. In the 1980s a new notion of generalized solutions for Hamilton-Jacobi equations, namely, viscosity solutions, Crandall and Lions [43], Fleming and Soner [63], Lions and Souganidis [124], [125], [126], Souganidis [180], Nisio [143], provided a means of characterizing the value function as the unique solution of HJI equation satisfying suitable boundary conditions. This method also provided the tools to show the convergence of the algorithms based on Dynamic Programming to the correct solution of the differential game and to establish the rate of convergence. A rigorous analysis of the viscosity solution of the Hamilton-Jacobi-Bellman-Isaacs equations in infinite dimensions is given in Swiech [190]. In the 1990s a method based on an occupation measure approach is introduced for stochastic differential games in a relaxed control setting in which the differential game problem reduces to a static game problem on the set of occupation measures, the dynamics of the game being captured in these measures, Borkar and Ghosh [31]. The major advantage of this method is that it enabled one to consider the dynamic game problems in much more physically appropriate wideband noise settings and use the powerful weak convergence methods, Ramachandran [158, 159, 163]. As a result, discrete games and differential games could be considered in a single setting.

The information structure plays an important role in the stochastic differential games. All the above referenced works assumes that all the players of the game have full information of the state. This need not be the case in many applications. The interplay of information structure in the differential games is described in Friedman [68], Ho [88], Olsder [145], Ramachandran [160], Sun and Ho [184]. The stochastic differential game problems with incomplete information are not as much developed as the stochastic control problems with partial observations.

One of the earlier works on obtaining computational method for stochastic differential games is given in Kushner and Chamberlain [111]. Following the work on numerical solutions for stochastic control Kushner and Dupuis [112] and many references in there, currently there are some efforts in deriving numerical schemes for stochastic differential games, Kushner [107, 108]. For a numerical scheme for the viscosity solution of the Isaacs' equation, we refer to Basar and Haurie [16]. Also, as a result of weak convergence analysis Ramachandran [158], Ramachandran and Rao [163], it is easier to obtain numerical methods for stochastic differential games similar to that of Kushner and Dupuis [112] and to develop new computational methods.

The key step to a general formulation from control theory to game theory was the distinction between state and control variables. The nature of a strategy is then clear; make the control variables functions of the state variables. This is an immediate generalization of the strategies of discrete games and is general enough for a far wider range of applications than just combat problems. In his book *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*, Isaacs gives examples in athletics and steel production as well as several pursuit and pursuit-evasion examples.

The Mathematical Theory of Optimal Processes published in 1962 by Pontryagin, *et al.* [152], deals with minimizing problems which could be called one-player differential games. This work was extended to two players in Kelendzeridze [97]. At the same time other research was going on in the United States. Control theory can be considered as equivalent to one-player differential games and is thus a special case.

When the connection between differential games and optimal control theory was realized in the early 1960s a flurry of research followed. Much of this work was by scientists working in control theory. Thus, there was a natural tendency to view differential games as an extension of optimal control theory. It gradually became evident that this view is not quite satisfactory.

Simply stated, differential games are a class of two-sided optimal control problems. More precisely, optimal control theory can be considered as a special case of differential games; but differential games are not easily viewed as an extension of optimal control theory. It is important to note certain differences. First, although feedback control is desirable in the one-sided problem it becomes almost mandatory in the game. Second, in more general games it is not at all certain that the game will terminate.

It is argued that both optimal control and differential games should be viewed as special cases of a much larger framework of dynamic optimization such as *Generalized Control Theory* as discussed in Ho, Bryson and Baron [89].

Whether it is deterministic or stochastic, there are three basic parts to an optimization problem:

i) the criterion (payoff) function;
ii) the controller(s) or players; and
iii) the information available to the players.

In optimal control theory there is just one controller who tries to minimize the criterion function on the basis of one set of information. Although this model can account for some real life situations, we can quickly think of situations in which there are more than one measure of performance and more than one intelligent controller operating with or without cooperation from the other controllers. It is also easy to think of situations where all the controllers do not receive the same information. We summarize these ideas in Table 1.1.1. The last column lists some typical references.

In the next section, we will give a brief survey of the literature on deterministic and stochastic differential games respectively. In Section 1.3, we will present a brief survey of stochastic differential games in the sixties and seventies. In Section 1.4 basic formulation of the problem will be presented. We will conclude this chapter with some basic definitions.

1.2 Deterministic Differential Games: A Brief Survey

The object of this section is to give a brief survey of the literature on deterministic differential games as they were introduced and as they have been generalized by other authors, Nicholas [141].

1.2.1 *Two-person, zero-sum differential games state and control variables*

As implied by the title of this subsection, a two-person, zero-sum differential game involves two players with opposing aims. The first two notions, state and control variables, are borrowed from control theory. In the theory of differential games as presented by Isaacs [91], both players know the values of the state variables at all times (games with complete information); and they are precisely the values on which they both make their decisions.

Table 1.1.1 Summary of Generalized Control Theory Problems

	Criterion			Number of Controllers			Information			Typical References
	One	Two	N 2	One J	Two $J_1 - J_2$	Multiple	Perfect	Imperfect	Multiple or Incomplete	
Deterministic Optimal Control	√			√			√			
Stochastic Optimal Control	√			√				√		
Vector-valued optimization Problem	√					√	√			Zadeh DaChuna & Polak
Zero-Sum Differential Game		√			√		√			Ho, Bryson, and Baron
Stochastic zero-sum Differential Game		√			√				√	Behn and Ho Rhodes & Luenberger Willman
Nonzero-sum Differential Game			√			√	√			Case Starr and Ho
Stochastic nonzero-sum Differential Game			√			√			√	

The control variables, as the name implies, are those variables which the players can manipulate.
The initial development of game theory was inspired by the problems in social science and economics. However, the main motivation of differential games was the study of military problems such as Pursuit-Evasion games.

Pursuer and Evader The terms pursuer, P and evader, E are carry overs from the early applications of differential game theory strictly to pursuit problems. As a convention, we assume that P controls the variables u_i and that E controls v_i.

The kinematic equations The motion of a point $x = (x_1, \ldots, x_n) \in E$, where E is the playing space (usually $\mathbb{R}^n$), is governed by the kinematic equations,

$$\dot{x} = f_j(t, x_1, \ldots, x_n, u_1, \ldots, u_p, v_1, \ldots, v_q)$$

for $j = 1, \ldots, n$ or briefly, $\dot{x} = (t, x, u, v)$ where $x_1, \ldots, x_n$ are the state variables and $u_1, \ldots, u_p$ and $v_1, \ldots, v_q$ are the control variables. We shall use the notation $\cdot \equiv \frac{d}{dt}$.

Terminal surface A game is terminated when x reaches the terminal surface C which is part of the boundary of E, or after a prescribed time T has elapsed. Since much of differential games are devoted to pursuit games, the surface C can be thought of as the set of all points where capture can occur. For this, P and E are also used as reference points on the two players. Clearly, we need not require that P and E coincide but just that they are "near" each other. It is obvious that bodies with large masses such as a plane and a rocket will collide before $d(P,E) = 0$, where $d(P,E)$ denotes the distance between the reference points P and E; so we require only that $d(P,E) < \ell$ where ℓ is some positive number. Thus, we can usually think of the capture region as "circular".

The Payoff The payoff is a numerical quantity which the players strive to minimize or maximize. For a game of degree (one which has a continuum of outcomes) the payoff is of the form

$$P(u,v) = H(t_f) \div \int G(x,u,v)dt,$$

where the integral is over the path in E and H is a smooth function on C which is the terminal value of the game. If $H = 0$, the game is said to have an integral payoff and if $G = 0$, a terminal payoff. Pursuit games with time to capture as payoff have an integral payoff with $G = 1$.

The Value Since P, controlling u, tries to minimize the payoff, while E, controlling v, tries to maximize the payoff, and the value of a differential game is defined as the minimax of the payoff,

$$v(x) = \min_u \max_v \ \text{(payoff)}.$$

Solution. The solution of a differential game is not a very rigorous concept. A game is considered solved when one or more of the following have been found:

i) The value function $v(x)$;

ii) The optimal paths;

and

iii) The optimal strategies (functions)

$$u^{\circ}(x) \text{ and } v^{\circ}(x) \text{ defined over } E.$$

Isaacs' approach was basically formal and did not make extensive use of classical variational techniques. His approach closely resembled the dynamic programming approach to optimization problems. In 1957 Berkowitz and Fleming [27] applied rigorous calculus of variation techniques to simple differential games. In a later definitive paper Berkowitz [26] expanded the applicable class of problems.

1.2.2 *Pursuit-Evasion Differential Games*

A two-person, zero-sum differential game problem may be stated crudely as follows. Determine a saddle point for

$$J = H(x(t_f), t_f) + \int_{t_0}^{t_f} G(t, x, u, v)dt \tag{1.2.1}$$

subject to the constraints

$$\dot{x} = f(t, x, u, v); \quad x(t_0) = x_0 \tag{1.2.2}$$

and

$$u \in U(t), \quad v \in U(t) \tag{1.2.3}$$

where J is the payoff, x is the state of the game, u and v are piecewise continuous functions, called strategies, which are restricted to certain sets U and V of admissible strategies, and a saddle point is defined as a pair of strategies (u°, v°) satisfying

$$J(u^{\circ}, v) \leqslant J(u^{\circ}, v^{\circ}) \leqslant (u, v^{\circ}) \tag{1.2.4}$$

for arbitrary $u \in U$ and $v \in V$. If (1.2.4) can be realized u° and v° are called optimal pure strategies, and $J(u^{\circ}, v^{\circ})$ is called the *Value* of the game.

Many control theorists have investigated the problem of controlling a dynamic system so as to hit a moving target. Most of these only allowed the pursuer to control his motion. Ho, Bryson, and Baron [89], allowing both players to control their motions, derived conditions for capture and optimality. Under the usual simplifying approximations to the equations of motion of the missile and the target, they showed that the proportional navigation law used in many missile guidance systems is actually an optimal pursuit strategy.

Ho *et al.*, considered the following game. Determine a saddle point for

$$J = \frac{a^2}{2} \|x_P(t_f) - x_e\left(t_f\right)\|^2 A'A + \frac{1}{2}\int_{t_0}^{t_f} \left[\|u(t)\|^2 R_P(t) - \|v(t)\|^2 R_e(t)\right] dt \tag{1.2.5}$$

subject to the constraints

$$\dot{x}_P = F_P(t)x_P + \overline{G}_P(t)u; \quad x_P(t_0) = x_{P_0} \tag{1.2.6}$$

$$\dot{x}_e = F_e(t)x_e + \overline{G}_e(t)u; \qquad x_e(t_0) = x_{e_0} \tag{1.2.7}$$

and

$$u(t),\ v(t) \in \mathbb{R}^m$$

where x_p is an n-dimensional vector describing the pursuer's state, $u(t)$ is the m-dimensional pursuer's control, $F_P(t)$ and $\overline{G}_P(t)$ are $(n \times n)$ and $(n \times m)$ matrices continuous in t; x_e, $v(t)$, $F_e(t)$ and $\overline{G}_e(t)$ are defined similarly. $R_P(t)$ and $R_e(t)$ are $(m \times m)$ positive definite matrices and $A = [I_K : 0]$ is a $(k \times n)$, $1 \leqslant k \leqslant n$, matrix. The quantity a^2 was introduced to allow for weighting terminal miss against energy. They considered a game of finite duration and perfect information. That is, t_f is a fixed terminal time and both players know the dynamics of both systems, (1.2.6) and (1.2.7), and at any time t, they know the state of each system.

A considerable and meaningful simplification is possible by reformulating the problem in terms of the k-dimensional vector.

$$z(t) = A\left[\Phi_P(t_f,t)x_P(t) - \Phi_e(t_f,t)x_e(t)\right].$$

In terms of $z(t)$, a completely equivalent problem is, determine a saddle point of

$$J = \frac{a^2}{2}\|z(t_f)\|^2 + \frac{1}{2}\int_{t_0}^{t_f}\left[\|u(t)\|^2 R_P(t) - \|v(t)\|^2 R_e(t)\right] dt \tag{1.2.8}$$

subject to the constraints

$$\dot{z} = G_P(t_f,t)u - G_e(t_f,t)v; \quad z(t_0) = z_0 \tag{1.2.9}$$

where

$$G_P = A\Phi_P(t_f,t)\overline{G}_P(t)$$

and

$$G_e = A\Phi_e(t_f,t)\overline{G}_e(t).$$

It is this approach which we will use throughout this study. The problem is essentially reduced from $2n$ dimensions to $k \leqslant n$ dimensions.

The problem presented in (1.2.8) and (1.2.9) is classified as a linear-quadratic differential game. That is the state equation (1.2.9) is linear in the controls and the payoff (1.2.8) is quadratic.

1.2.3 *The Problem of Two Cars*

The Problem of Two Cars is a good example of a two-person zero-sum pursuit-evasion game which is not based on warfare strategies. It is just like the classical Homicidal Chauffeur game, Isaacs [91], except that the evader's radius of curvature is also constrained. Here we have two cars traveling on an infinite parking lot at constant but (possibly) different speeds. Cockayne [41] found that necessary and sufficient conditions for the capture region to be the entire state space are (1) the pursuer must be faster than the evader; and (2) the pursuer must have greater lateral acceleration capability, as embodied in the minimum radius of curvature, than the evader. Meier [132] studied the problem when the pursuer is slower than the evader. Although the capture region could be found analytically using Isaacs' theory, the geometric methods presented by Meier are simpler and give more insight. The technique appears to be applicable to a general class of pursuit-evasion problems in which the dynamics of the players are independent of their positions and in which termination depends only on their relative positions.

1.2.4 *The Lanchester Combat Model*

Some research which can be classified under the broad heading of differential games was carried on at Virginia Tech a few years ago. Springall and Conolly [42] obtained some theoretical results for the probability of victory in the Lanchester combat model described by the deterministic differential equations

$$\dot{m} = -\mu mn - \delta n$$

and

$$\dot{n} = -\lambda mn - \gamma n$$

where m denotes the first player's forces and n denotes the second player's forces. Let us call the two sides P and E.

The model studied by Conolly and Springall [42] is unusual in that they assume that both sides deploy only a constant fraction of their initial strengths in the field, holding the remainder in reserve to replace casualties. Due to the formulation of the model, although the results of a combat do depend on the initial strengths, it was found that neither the probability of P's victory, nor the probability of E's victory, depends on how side E partitions his troops. Both probabilities are, however, dependent on how P partitions his forces. Conditions are given in Conolly and Springall [42] on how side P can divide his forces to maximize his probability of winning. Using data on Civil War battles it was found that the actual outcomes agreed favorably with the outcomes which would be predicted based on the initial percentages of forces sent into the field.

For realistic applications to other fields, such as biology or economics, it is usually necessary to study games which are not zero-sum and which involve more than two players.

1.2.5 *Nonzero-sum N-person Differential Games*

The theory of differential games has been extended to the situation where there are N players $(N > 2)$ and the players try to minimize different performance criteria.

In the general nonzero-sum, N-player differential game, the following situation arises. For $i = 1, 2, \ldots, N$, player i wants to choose his control u_i to minimize

$$J_i = K_i(x(t_f), t_f) + \int_{t_0}^{t_f} L_i(t, x, u_1, \ldots u_N)\, dt$$

subject to the constraint

$$\dot{x} = f(t, x, u_1, \ldots u_N); \quad x(t_0) = x_0.$$

There may also be some inequality constraints on the state and/or control variables as well as restrictions on the terminal state. The terminal time t_f may be fixed or variable.

Case [38] was concerned only with pure strategies and with games which he expected to have pure strategy solutions. This dictated that all the players had perfect information throughout the course of the game.

When we have N players the definition of a solution is no longer obvious. Many new concepts arise which force one to sharpen his definition of optimality. In a pair of papers Starr and Ho [185] discussed three types of solutions: *Nash equilibrium, non-inferior set of strategies, and minimax.*

Nash equilibrium A Nash solution u_i^*, $i = 1, 2, \ldots, N$, is defined by

$$J_i\left(u_1^*, u_2^*, \ldots, u_i^*, \ldots, u_N^*\right) \leqslant J_i\left(u_1^*, \ldots, u_{i-1}^*, u_i, u_{i+1}^*, \ldots, u_N^*\right)$$

for all u_i, $i = 1, 2, \ldots, N$, where J_i is the criterion which the ith player is attempting to minimize.

Noninferior solutions Let Λ denote the set of non-inferior solutions. Then, the strategy N-tuple $s^* = \{s_1^*, \ldots, s_N^*\}$, belongs to Λ if, and only if, for any other strategy N-tuple $s = \{s_1, \ldots, s_N\}$, the following is satisfied.

$$J_i(s) \leqslant J_i\left(s^*\right), \quad i = 1, \ldots, N$$

The set Λ contains those solutions which are worthy of consideration for cooperation or negotiation. It is called the *Pareto-optimal set* for the problem.

Minimax Consider the other players to be irrational in that they try to maximize our criterion instead of minimizing their own criterion. That is, each solves a zero-sum game with respect to his own criterion with all the other players opposing it. The minimax value of the criterion is then security level of a controller when operating against other irrational controllers, while the Nash value is the level against rational ones.

For a linear-quadratic deterministic differential game, that is, a game with linear dynamics and quadratic payoff, all three of these solutions can be obtained by solving sets of ordinary matrix differential equations.

Applying their theory to a nonzero-sum version of a simple pursuit-evasion game considered by Ho, Bryson, and Baron [89], Starr and Ho [186] found *negotiated solutions* which give both players better results than the usual saddle-point solution. They also outlined an application to economic analysis involving the dividend policies of firms operating in an imperfectly competitive market.

In a recent paper Case [38] casts the problem of profit maximization for two firms manufacturing the same commodity (coal in his example) as a two-person differential game. That is, he supposes that all of the coal deposits in some small country, isolated from the rest of the world by high mountains, are owned by two competing firms. Because the demand for coal in this country is highly elastic, the two firms could overcharge.

The assumptions are similar to those made by Isaacs for his steel production example [91]. It takes coal to mine coal and to open new mines. Thus, each firm must decide how much coal to allocate to the production of coal for the present market; how much to invest in new mines; and how much to stockpile for future demands. The qualitative aspects of the Nash equilibrium point of the game are discussed.

Case's model is applicable to the theory of protective tariffs. The high mountain range offers full tariff protection for a given period. Assigning values to the physical constants, one could actually calculate the prices which would evolve and a tariff rate which should be sufficient to guarantee the desired protection. If such calculations could be made for realistic and complicated models, many people would want to use them.

1.2.6 *Friedman's approach to differential games*

No survey of differential games would be complete without the mention of Avner Friedman's works [67, 68]. We have not referenced him previously because his research publications transcend our section headings.

Friedman [67] defines a differential game in terms of a sequence of approximating discrete games. He assumes that the controls appear *separated* in the kinematic equation and in the integral part of the payoff and gives an example to show that the game may not have Value if the controls are not separated. His work generalizes that of Varaiya and Lin [200]. In Friedman [67] he proves, under suitable conditions, that a pursuit-evasion game with a general payoff which is usually not a continuous functional, has Value and saddle points and that the Value is Lipschitz continuous. These results are extended to differential games of survival. Another paper by Friedman [67] extends the results to the case where the state x is restricted to a given *phase set* X which is a subset of Euclidean n-dimensional space $\mathbb{R}^n$. Friedman also computes the Value for a class of games with fixed duration; and gives a general method for computing saddle points for games of fixed duration as well as games of pursuit and evasion.

Friedman's research is not limited to two-person differential games. In addition he considers linear-quadratic differential games with non-zero sum and N players. His approach to N-person differential games is similar to his approach to two-person differential games in that he defines the game through the concept of δ-games, i.e. discrete approximating games of fixed duration δ. In this paper he derives bounds on the optimal strategy for a δ-game and proves a theorem which asserts that the differential game has Value under certain conditions on the controls.

Earlier a similar approach to differential games was investigated by Fleming [62, 63]. He introduced the idea of *a majorant* and *minorant* game in which the information is biased to favor one player or the other. He then gave conditions for the majorant value and minorant value to converge to the Value of the game.

Consider the following two differential games.

$$\dot{x}_1 = f_1(t, x_1, u_1, v_1); \quad x_1(t_0) = x_{10}$$

and

$$\dot{x}_2 = f_2(t, x_2, u_2, v_2); \quad x_2(t_0) = x_{20}.$$

Using the definition of a differential game given by Friedman [67] and from differential inequalities, it can be shown that if the functions f_1 and f_2 are close in some sense then so are their Values. Such a comparison is of use in approximating a differential game by a simpler one.

One source of differential games is the study of optimal control problems in which the system to be controlled is subject to unknown random disturbances. We now go on to a discussion of stochastic differential games.

1.3 Stochastic Differential Games: Definition and Brief Discussion

In recent years a number of articles have appeared in the journals on what have been called stochastic differential games. These games are stochastic in the sense that noise (zero mean, Gaussian, white noise) is added to the players' observations of the state of the system or to the transition equation itself.

In 1966, Ho [87] solved a stochastic differential game using variational techniques. One player controlled the state and attempted to minimize the terminal error and confuse the other player who could only make noisy measurements of the state and attempted to minimize the error of his estimate. Since only one player actually controlled the state, the game was not of the pursuit-evasion type, and could be solved subsequently by first determining the form of the first player's controller and using this to determine the form of the second player's estimator. The solution indicated that a certain time the first player should change strategies from trying to confuse the other player to trying to minimize the terminal criterion. A logical extension is an investigation of a pursuit-evasion problem in which both players have imperfect knowledge of the states involved.

1.3.1 *Stochastic Linear Pursuit-Evasion Games*

Behn and Ho [19] made some progress in this direction when they studied the problem where the pursuer has perfect knowledge, but the evader can make only noisy measurements of the state of the game. They showed that the evader can use the noisy measurements to obtain an optimal estimate of the state and then use this estimate in the feedback strategy for the deterministic problem.

When we restrict the problem to linear dynamics and quadratic criterion with Gaussian noises as the sources of randomness, then specific results are available. Liu [129] considered this problem and converted it into a stochastic differential game under the assumption that one of the players fixes his strategy in a linear form with a linear filter. The other player must then use a linear strategy for optimality. Liu obtained optimal pairs of linear strategies when one player has corrupted information and when both have corrupted information.

Consider the zero-sum, two-person stochastic differential game with the linear transition equation

$$\dot{x} = Fx + G_P u + G_e v; \quad x(t_0) = x_0$$

and quadratic criterion given by,

$$J = \frac{1}{2}E\left\{x'(t_f)S_f x\left(t_f\right) + \int_{t_0}^{t_f} \begin{bmatrix} x' & u' & v' \end{bmatrix} \begin{bmatrix} Q & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & -C \end{bmatrix} \begin{bmatrix} x \\ u \\ v \end{bmatrix} dt \right\},$$

where B and C are symmetric, positive definite matrices and $S_f = S(t_f)$ is a symmetric, positive semi definite matrix which is the solution of a Riccati-like equation. Let the observations (measurements) be given by

$$z_P = h_1 x + w_1, \tag{1.3.1}$$

and

$$z_e = h_2 x + w_2. \tag{1.3.2}$$

It is assumed that $x_0 \sim N(0, P_0)$ and is independent of w_1 and w_2 which are white Gaussian processes such that

$$E(w_1) = 0, \quad E(w_1 w_1') = R_1,$$

and

$$E(w_2) = 0, \quad E(w_2 w_2') = R_2.$$

We shall summarize the work done on this particular type of problem in Table 1.3.1.
Perfect measurements for player i, $i = 1, 2$, is denoted by $R_i = 0$, $H_i = I$ (the identity matrix) where $R_i = 0$ denotes the degenerate case $w_1 \equiv 0 \equiv w_2$. Similarly, no measurements are denoted by $R_i = \infty$, $i = 1, 2$. Thus, there are nine cases to be considered. Either player's measurements may be perfect, noisy, or omitted.
The case where both players can make perfect measurements is referred to as the closed-loop game and the case where neither player has any measurement as the open loop game. Borh are treated by Bryson and Ho in *Applied Optimal Control* [34]. The solution to cases 1, 3, 6, 7, 8, and 9 requires only the solution of Riccati-like equations because the measurements involved are degenerate. Cases 6 and 8 are extensions of stochastic control theory since one player operates open loop. The other three cases, 2, 4, and 5, give rise to complicated equations of the two point boundary value problem type.
A further stochastization can be achieved by making the transition equation (1.2.2) itself stochastic. Willman [207] did this by considering a random version of (1.2.9), namely:

$$\dot{x} = G_P u - G_e v + q$$

Table 1.3.1 Summary of Research Publications on Stochastic Pursuit-Evasion Games

		PURSUER		
		Perfect Measurements $R_1 = 0$, $H_1 = 1$	Noisy Measurements $0 < R_1 < \infty$	No Measurements $R_1 = \infty$
EVADER	Perfect Measurements $R_2 = 0$, $H_2 = 1$	1 Ho, Bryson, and Baron	2 Behn and Ho Rhodes & Luenberger	3 Bryson and Ho
	Noisy Measurements $0 < R_2 < \infty$	4 Behn and Ho Rhodes & Luenberger	5 Willman Rhodes & Luenberger	6 Rhodes & Luenberger
	No Measurements $R_2 = \infty$	7 Bryson and Ho	8 Rhodes & Luenberger	9 Bryson and Ho

with criterion

$$J = \frac{1}{2} E \left\{ x'(t_f) S_f x(t_f) + \int_{t_0}^{t_f} \begin{bmatrix} u' & v' \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} dt \right\}$$

and measurements given by (1.3.1) and (1.3.2). It was assumed that

$$\begin{bmatrix} q \\ w_1 \\ w_2 \end{bmatrix}$$

is Gaussian white noise process with mean vector and covariance matrix given by the pair

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} Q & 0 & 0 \\ 0 & R_1 & 0 \\ 0 & 0 & R_2 \end{bmatrix}$$

which is assumed to be statistically independent of the prior. As before B and C are positive definite time-dependent matrices and S_f is positive semi-definite.

Willman [207] was able to get formal solutions for games of this type which depend on the solutions of certain sets of implicit equations of the two point boundary value problem type.

Although each of the authors referenced above has tacitly admitted that the real world is not static, or deterministic, they have made their games stochastic by adding independent, zero mean, Gaussian white noise to the observations and/or the transition equations. That is, they have perturbed the games. Can this normality assumption be relaxed? Roxin and Tsokos [170] recently introduced a more general and realistic definition of a stochastic differential game.

1.3.2 *The Definition of a Stochastic Differential Game*

Consider a two-person, zero-sum differential game defined by the differential equation

$$\dot{x} = f(t,x,u,v,w) \tag{1.3.3}$$

where

i) $t \in \mathbb{R}_+$ is the time;
ii) $x \in \mathbb{R}^n$ is the state variable:
iii) $u = u(t) \in E^P$ and $v = v(t) \in E^q$ are control variables;
and
iv) $w = w(t,\omega) \in E^r$ is some stochastic process defined over the complete probability space $(\Omega, \mathrm{A}, \mu)$ where Ω is a nonempty abstract set, A is a σ-algebra of subsets of Ω, and μ is a probability measure on A.

The initial state of the differential game is given by

$$x_0 = x(t_0),$$

and the terminal condition by

$$\Phi(t_f, x(t_f)) = 0, \tag{1.3.4}$$

where Φ is a given vector-valued function which defines a manifold in E^{n+1} space. In this game, one player, called u, controls the functions $u(t)$ and the other, called v, controls the function $v(t)$. We take as admissible controls $u(t)$ and $v(t)$ all measurable functions whose values belong (almost everywhere) to some given compact sets $U \subset E^P$ and $V \subset E^q$.
The payoff which player u must pay to player v at the end of the game is given by the functional

$$J(t_0,x,u,v) = H(t_f,x(t_f)) + \int_{t_0}^{t_f} G(x,u,v)dt. \tag{1.3.5}$$

Thus, players u and v, want to choose their actions so as to minimize and maximize the expected value of the payoff functional J respectively.

The definition given by Roxin and Tsokos [170] assumes that both players know the dynamics of the game (1.3.3), the terminal condition (1.3.4), the admissibility sets U and V, and the payoff (1.3.5). They must choose their controls on the basis of the observations

$$y(t,w) = \psi_1\{t, x(s), w(s,\omega),\ 0 \leqslant s \leqslant t\}$$

and

$$z(t,w) = \psi_2\{t, x(s), w(s,\omega),\ 0 \leqslant s \leqslant t\}$$

respectively. A strategy $\sigma_u[\sigma_v]$ for player $u[v]$ is a rule for determining the control $u(t)$ $[v(t)]$ as a functional.
That is,

$$u(t) = \sigma_u\{y(t)\}$$

and

$$v(t) = \sigma_v\{z(t)\}.$$

The differential game described above is called a two-person, zero-sum, stochastic differential game. Unfortunately even for the discrete case, there is no general way to ascertain the existence or the uniqueness of the solution of the game for a given pair of strategies. An approach which goes back to the works of Fleming [61] is to consider the continuous differential game as the limit of a discrete game obtained by dividing the time interval into N short subintervals. For deterministic differential games this approach was investigated by Varaiya and Lin [200], and Friedman [67]. It was applied to stochastic differential games by Roxin and Tsokos [170].

The discretized game Without loss of generality we can assume that the game starts at $t_0 = 0$ and ends at a fixed time $t_f = T$. For any positive integer k let $N = 2^k$ and $\delta = \frac{T}{N}$. Define the subintervals

$$I_j = \{t : (j-1)\delta \leqslant t < j\delta\}, \quad j = 1, 2, \ldots, N.$$

We can now define a discrete interpretation of the continuous stochastic differential game (1.2.3) as follows. At each instant $t_j = j\delta$, $j = 0, 1, \ldots N-1$, players u and v choose their controls $u(t)$ and $v(t)$ for the succeeding subinterval $j\delta \leqslant t < (j+1)\delta$. They have at their disposal the observations $y(t_j)$ and $z(t_j)$ respectively. Since neither player knows the control chosen by his opponent, it is well known that each should use a randomized decision function. To avoid randomized decisions we give player v, the maximizing player, a slight advantage. Player u must choose $u(t)$ for $t_j \leqslant t < t_{j+1}$ based only on his observations $y(t_j)$

but player v chooses his control $v(t)$ based on $z(t_j)$ and $u(t)$ for $t_j \leqslant t < t_{j+1}$. However, player v is not allowed to store this information, that is, he cannot use $u(t)$for $t < t_j$ when choosing $v(t)$ for $t_j \leqslant t < t_{j+1}$. This is called an upper δ-game. A lower δ-game is similarly defined.

The Expected Payoff It is clear that, even when the initial conditions (t_0, x_0) are given and the players have chosen strategies σ_u and σ_v, the resulting payoff function is still a random variable. This is because the payoff also depends on the stochastic process $w(t,\omega)$. The expected value of the payoff is therefore defined to be

$$J_0\left(t_0, x_0, \sigma_u, \sigma_v\right) = E\left\{J\left(t_0, x_0, \sigma_u, \sigma_v\right)\right\}.$$

This expected value is unknown to the players but player u tries to minimize it based on his information and the most unfavorable strategies of player v. Similarly, player v tries to maximize it on the basis of his information and the most unfavorable u-strategies.

The Value of The Game Define V_1 and V_2 as follows:

$$V_1 = \underset{\sigma_u}{\text{glb}}\ \underset{\sigma_v}{\text{lub}} \quad E\left\{J_0\left(t_0, x_0, \sigma_u, \sigma_v\right) | y(t_0)\right\} \tag{1.3.6}$$

and

$$V_2 = \underset{\sigma_v}{\text{lub}}\ \underset{\sigma_u}{\text{glb}} \quad E\left\{J_0\left(t_0, x_0, \sigma_u, \sigma_v\right) | z(t_0)\right\}. \tag{1.3.7}$$

Let the optimal strategies be denoted by σ_u^* and σ_v^*. For the discrete upper δ-game define

$$V^{\delta}\left(t_0, x_0\right) = J_0^{\delta}\left(t_0, x_0, \sigma_u^{*\delta}, \sigma_v^{*\delta}\right)$$

and for the lower δ-game

$$V_{\delta}\left(t_0, x_0\right) = J_{0\,\delta}\left(t_0, x_0, \sigma_{u\,\delta}^{*}, \sigma_{v\,\delta}^{*}\right)$$

where $\sigma_u^{*\,\delta}$ and $\sigma_v^{*\delta}$ are optimal in (1.3.6) and (1.3.7) for the upper δ-game and $\sigma_{u\,\delta}^{*}$ and $\sigma_{v\delta}^{*}$ are optimal for the lower δ-game. If $\delta \to 0$

$$\lim_{\delta \to 0} V^{\delta}(t_0, x_0) = \lim_{\delta \to 0} V_{\delta}(t_0, x_0) = V,$$

V is called the *value* of the continuous stochastic differential game.

The definition of a stochastic differential game used in this study will be slightly different from that outlined above. Inspired by the definition of Roxin and Tsokos, we will apply an idea similar to that recently used by Morazon [136] and Tsokos [196] in the study of the stability of linear systems. We will adopt the idea of letting the functions which constitute the game be random functions themselves. This interpretation is more realistic and consistent with the terminology. Games in which white noise has been added to the observations and/or the transition equation itself would better be called *perturbed differential games.*

1.4 Formulation of the Problem

Let the triple $(\Omega, \mathrm{A}, \mu)$ denote a *probability measure space*. That is, Ω is a nonempty set, A is a σ-algebra of subsets of Ω and μ is a complete probability measure on A. Let $x(t;\omega)$, $t \in \mathbb{R}_+$, denote a *stochastic process* (or *random function*) whose index set is the set of non-negative real numbers $\mathbb{R}_+ = \{t : t \geqslant 0\}$, $\omega \in \Omega$. That is, for each $t \in \mathbb{R}_+$, $x(t;\omega)$ is a random variable defined on Ω,

Perhaps a picture is in order here; the reader is referred to Figure 1.4.1 and 1.4.2 for a Graphical Explanation of $x(t;\omega)$. Of course in a deterministic game we simply have a single trajectory. However, when we consider a solution to be a stochastic process, we have an ensemble of trajectories. In Figure 1.4.1 each line represents a possible realization of $x(t;\omega)$ for a given $\omega \in \Omega$. If we let Ω vary also, then we get an ensemble of paths for the motion of the point x. On the other hand in Figure 1.4.2 the time t is held fixed and Ω is varying. We see that we have a distribution of points for each fixed $t \in \mathbb{R}_+$. If we connect the points for each $\omega \in \Omega$ we will get continuous curves as in Figure 1.4.1.

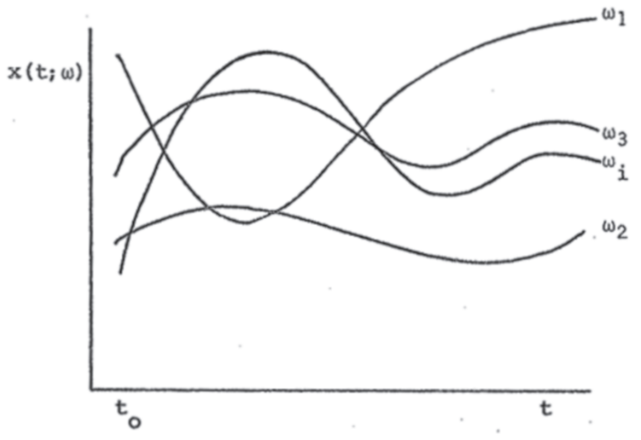

Fig. 1.4.1 t varying, ω fixed

Definition 1.4.1. A stochastic process (or random function) $x(t;\omega)$, $t \in \mathbb{R}_+$, is said to be a *second order* (or *regular*) *process* or to belong to the space $L_2(\Omega, \mathrm{A}, \mu)$ if for each $t \in \mathbb{R}_+$, the second absolute moment exists. That is,

$$E\left\{|x(;\omega)|^2\right\} = \int_\Omega |x(;\omega)|^2 \, d\mu(\omega) < \infty.$$

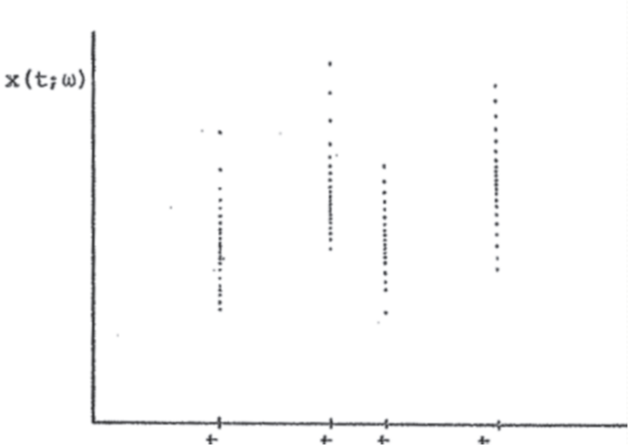

Fig. 1.4.2 t fixed, ω varying

In other words, $x(t;\omega)$, $t \in \mathbb{R}_+$ is *square-summable* with respect to μ-measure.

The norm of $x(t;\omega) \in L_2\,(\Omega, \mathrm{A}, \mu)$ is defined by

$$||x(t;\omega)||_{L_2(\Omega,\mathrm{A},\mu)} = \left\{ \left[|x(t;\omega)|^2 \right] \right\}^{1/2} \tag{1.4.1}$$

for each $t \in \mathbb{R}_+$, $L_2\,(\Omega, \mathrm{A}, \mu)$ is a Hilbert space with inner product defined for each pair of random variables $x(t;\omega)$ and $y(t;\omega)$ by

$$\begin{aligned}(x(t;\omega), \quad y(t;\omega))_{L_2(\Omega,\mathrm{A},\mu)} &= \int_\Omega x(t;\omega)\overline{y(t;\omega)}\,d\mu(\omega) \\ &= E\left[x(t;\omega)\overline{y(t;\omega)}\right],\end{aligned} \tag{1.4.2}$$

where the bar denotes the complex conjugate in case we are talking about complex-valued random variables. Combining equations (1.4.1) and (1.4.2), the norm in $L_2(\Omega, \mathrm{A}, \mu)$ is defined in terms of the inner product. Thus, for a second order process the covariance function always exists and is finite.

In this study we will be dealing with stochastic differential games with transition equations of the form

$$\frac{d}{dt}x(t;\omega) = f\left(x(t;\omega), u, v, t\right); \qquad x\left(t_0;\omega\right) = x_0(\omega)$$

where $x(t;\omega) \in L_2(\Omega, \mathrm{A}, \mu)$ for each $t \in \mathbb{R}_+$. The control variables u and v may be random, i.e. belong to $L_2(\Omega, \mathrm{A}, \mu)$, or deterministic. Further assumptions concerning their behavior will be given at the appropriate points in the study.

We will now list some definitions which will be necessary for the presentation of this study.

1.5 Basic Definitions

It will be assumed that the reader is familiar with the fundamentals of measure theory and integration, functional analysis, and topology. Therefore such definitions are linear (vector) space, norm, semi-norm, normed linear space, and complete normed linear space or Banach space will not be given. We refer the reader to such texts as Yosida [218]. However, some definitions from these and related fields will be repeated here for the convenience of the reader.

Definition 1.5.1. By a *random solution* of the stochastic differential equation (1.3.4) we shall mean a function $x(t;\omega)$ which satisfies equation (1.3.4) μ-a.e.

We have already defined what we mean by $x(t;\omega)$ in $L_2(\Omega, \mathrm{A}, \mu)$. For fixed $t \in \mathbb{R}_+$ we shall denote $x(t;\omega)$ by $x(\omega)$ and call it a *random variable*. Recall Figure 1.4.2.

Definition 1.5.2. A random variable $x(\omega)$, $\omega \in \Omega$, is said to be μ-*essentially bounded* or to belong to the space $L_\infty(\Omega, \mathrm{A}, \mu)$ if it is measureable with respect to μ and there is a constant $a > 0$ such that

$$\mu\{\omega : |x(\omega)| > a\} = 0. \tag{1.5.1}$$

That is, $x(\omega)$ is bounded in the usual sense except maybe on a set of probability measure zero.

The greatest lower bound (glb) of the set of all values for which (1.5.1) holds is called the *essential supremum* of $|x(\omega)|$ with respect to μ and is denoted by

$$\begin{aligned}\mu\text{-ess sup}x(\omega) &= \text{glb}\{a : \mu[\omega : |x(\omega)| > a] = 0\} \\ &= \inf\left\{\sup_{\Omega\text{-}\Omega_0} |x(\omega)|\right\},\end{aligned}$$

where Ω_0 is a set of probability measure zero, $\mu(\Omega_0) = 0$.

The norm of $x(\omega) \in L_\infty(\Omega, \mathrm{A}, \mu)$ is defined by

$$\|x(\omega)\|_{L_\infty(\Omega,\mathrm{A},\mu)} = \mu\text{-ess}_{\omega\in\Omega}\text{sup}x(\omega).$$

Definition 1.5.3. Consider a mapping $f : X \to X$. f is said to be a *contraction mapping* if there exists a number $a \in (0,1)$ such that $d(f(x), f(y)) \leqslant ad(x,y)$ for any $x, y \in X$.

Definition 1.5.4. Let $x(s)$ is a finite function defined on the closed interval $[a,b]$. Suppose that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\left|\sum_{k=1}^{n}\{x(b_k) - x(a_k)\}\right| < \varepsilon,$$

for all $a_1, b_1, \ldots, a_n, b_n$ such that $a_1 < b_1 \leqslant a_2 < b_2 \leqslant \ldots \leqslant a_n < b_n$ and $\sum_{k=1}^{n}(b_k - a_k) < \delta$. Then the function $x(s)$ is said to be *absolutely continuous.*

We note that if a function $x(s)$ is absolutely continuous, then its derivative exists almost everywhere.

We now state an important inequality known as the *generalized Gronwall's Inequality*.

Definition 1.5.5. Let $x(t)$ be a continuous non-negative function on $[t_0, t_f]$ and assume that

$$x(t) \leqslant M + \int_{t_0}^{t} x(s) d\sigma(s), \quad t \in [t_0, t_f]$$

where M is a positive constant and $\sigma(t)$ is a nondecreasing function on $[t_0, t_f]$ such that $\sigma(t) = \sigma(t+0)$. Then $x(t)$ satisfies

$$x(t) \leqslant M e^{\sigma(t)} - \sigma(t_0).$$

Definition 1.5.6. Consider the stochastic system

$$\frac{d}{dt} x(t;\omega) = f(t, x, u, v), \quad t \geqslant 0$$

with initial condition. The system is called *stochastically asymptotically stable* if the following two conditions are satisfied:

(i) for each $\varepsilon > 0$, $t_0 \in \mathbb{R}_+$, there exists a $\delta = \delta(t_0, \varepsilon) > 0$ such that $\|x(\omega)\|_{L_2(\Omega, \mathrm{A}, \mu)} \leqslant \delta$ implies

$$\|x(t, t_0, x_0(\omega), u, v,)\|_{L_2(\Omega, \mathrm{A}, \mu)} < \varepsilon, \quad t \geqslant t_0$$

for every admissible pair of controls u, v;

and

(ii) for each $\varepsilon > 0$, $t_0 \in \mathbb{R}_+$, there exists numbers $\delta_0 = \delta_0(t_0)$ and $T = T(t_0, \varepsilon)$ such that whenever

$$\|x_0(\omega)\|_{L_2(\Omega, \mathrm{A}, \mu)} \leqslant \delta_0,$$

and

$$\|x(t, t_0, x_0)(\omega), u, v\| < \varepsilon, \quad t \geqslant t_0 + T$$

for every admissible pair of controls u, v.

Definition 1.5.7. Consider an $(n \times n)$ matrix $A(\omega) = (a_{ij}(\omega))$, $\omega \in \Omega$. $A(\omega)$ is called a *random matrix* if $A(\omega) \times (t, \omega)$ is a random n-vector with values in $L_2(\Omega, \mathrm{A}, \mu)$ for every n-vector $x(t;\omega) \in L_2(\Omega, \mathrm{A}, \mu)$, for $t \in \mathbb{R}_+$. That is, $A(\omega)$ is a matrix whose n^2 elements $a_{ij}(\omega)$ are random variables.

Definition 1.5.8. Let $x(\omega)$ be a random vector. The matrix norm $|A(\omega)|$ will be defined by

$$|A(\omega)| = \sup_{\|x(\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant 1} \|A(\omega)x(\omega)\|_{L_2(\Omega,\mathrm{A},\mu)}.$$

With respect to the completion of a differential game we give the following criteria.

Definition 1.5.9. A pair of strategies $\left(U^0, V^0\right)$ is called *optimal*, or *minimax*, if it satisfies the *saddle-point criterion*. That is, the strategy pair $\left(U^0, V^0\right)$ is minimax if for any other strategy pair (U, V) is based on the available information set

$$J\left(U^0, V\right) \leqslant J\left(U^0, V^0\right) \leqslant J\left(U, V^0\right).$$

Definition 1.5.10. Consider an n-person deterministic differential game. If $J_1(s_1, \ldots, s_n)$, $\ldots$, $J_N(s_1, \ldots, s_n)$ are the cost functions for players $1, \ldots, N$, then the strategy set $(s_1^*, \ldots, s_N^*)$ is called a *Nash equilibrium* strategy set if, for $i = 1, 2, \ldots, N$,

$$J_i\left(s_1^*, \ldots, s_{i-1}^*, \quad s_i, s_{i+1}^*, \ldots, s_N^*\right) \geqslant J_i\left(s_1^*, \ldots, s_N^*\right)$$

where s_i is any admissible strategy for player i.

We will use these definitions in subsequent chapters.

Chapter 2

Stochastic Linear Pursuit-Evasion Game

2.1 Introduction

The 1950's saw the introduction of guided interceptor missiles and the launching of Sputnik I. Questions on pursuit and evasion were suddenly in everyone's mind. What is the best strategy to intercept a moving target? How can friendly planes best avert midair collisions? Thus the theory of Differential Games is permeated with the theory of military pursuit games. Dr. Rufus Isaacs, who was then with the Mathematics Department of the RAND Corporation realized that no one guidance scheme can be optimal against all types of evasion. An intelligent evader can deliberately maneuver to confuse the pursuer's predictions. Thus optimal pursuit and evasion must be considered equally.

Consider a stochastic linear pursuit-evasion game described by a linear stochastic differential equation of the form

$$\frac{d}{dt}x(t;\omega) = A(\omega)x(t;\omega) + B(\omega)u(t;\omega) - C(\omega)v(t;\omega) \tag{2.1.1}$$

for $t \geqslant 0$ and $\omega \in \Omega$, where

i) Ω is the supporting set of a complete probability measure space $(\Omega, \mathrm{A}, \mu)$;
ii) $x(t;\omega)$ is the unknown random n-dimensional state variable;
iii) $u(t;\omega)$ and $v(t;\omega)$ are the random control vectors;
and
iv) $A(\omega)$, $B(\omega)$, and $C(\omega)$ are random matrices of appropriate dimensions.

The problem is to choose a control $u_v(t;\omega)$, depending on the evader's control $v(t;\omega)$ such that

$$x\left(t_{u,v};\omega\right) \in M_\varepsilon \text{ for some } t_{u,v} \in \mathbb{R}_+$$

where M_ε is the terminal set to be defined in Section 2.3.

The object of this chapter is to prove the existence and uniqueness of a random solution, that is, a second order stochastic process, which satisfies equation (2.1.1) with probability one. In order to do this we integrate equation (2.1.1) with respect to t obtaining a vector stochastic integral equation of the form

$$\begin{aligned} x(t;\omega) &= x_0(\omega)e^{A(\omega)t} \\ &+ \int_0^t e^{A(\omega)(t-\tau)}\left[B(\omega)u(\tau;\omega) - C(\omega)v(\tau;\omega)\right] d\tau \end{aligned} \tag{2.1.2}$$

for $t \geqslant 0$ and $\omega \in \Omega$, with initial condition $x(0;\omega) = x_0(\omega)$. In the theory of stochastic integral equations the term $x_0(\omega)e^{A(\omega)t}$ is referred to as the *free stochastic term* or *free random vector* and $e^{A(\omega)(t-\tau)}$ as the *stochastic kernel*.

We will approach the question of existence and uniqueness of a random solution of equation (2.1.2) using the technique of admissibility theory introduced into the study of random integral equations by Tsokos [184]. To do this we must first define some topological spaces and state some results which are essential to this presentation.

2.2 Preliminaries and an Existence Theorem

We will be concerned with the space of random vectors in $L_2(\Omega, \mathrm{A}, \mu)$ where $L_2(\Omega, \mathrm{A}, \mu)$ denotes the set of all μ-equivalence classes of random vectors of the form $(x_1(\omega), \ldots, x_n(\omega)) = x(\omega)$ where for each $i = 1, 2, \ldots, n$, $x_i(\omega)$ is an element of $L_\infty(\Omega, \mathrm{A}, \mu)$. It is well known that $L_2(\Omega, \mathrm{A}, \mu)$ is a normed linear space over the real numbers with the usual definitions of component-wise addition and scalar multiplication with norm given by

$$\|x(\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} = \left\{ \int_\Omega \left[x_1(\omega)^2 + x_2(\omega)^2 + \cdots + x_n(\omega)^2\right] d\mu(\omega) \right\}^{\frac{1}{2}}.$$

Definition 2.2.1. Let $C_c = C_c\left(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu)\right)$ denote the space for all continuous vector valued functions from $\mathbb{R}_+$ into $L_2(\Omega, \mathrm{A}, \mu)$, or second order stochastic processes on $\mathbb{R}_+$, with the topology of uniform convergence on every compact interval $[0, T]$, $T > 0$. That is, the sequence $x(t;\omega)_k$ converges to $x(t;\omega)$ in C_c if and only if

$$\lim_{k\to\infty} \left\{ E\left|x(t;\omega)_k - x(t;\omega)\right|^2 \right\}^{\frac{1}{2}} = \lim_{k\to\infty} \left\{ \int_\Omega |x(t;\omega)_k - x(t;\omega)|^2 d\mu(\omega) \right\}^{\frac{1}{2}} = 0$$

uniformly on every interval $[0, T]$, $T > 0$.

Definition 2.2.1 simply says that the map $t \to x(t;\omega) = (x_1(t;\omega), x_2(t;\omega), \ldots, x_n(t;\omega))$ is continuous and that for each $t \in \mathbb{R}_+$ and each $i = 1, 2, \ldots, n$, $x_i(t;\omega) \in L_\infty(\Omega, \mathrm{A}, \mu)$. Thus for fixed $t \in \mathbb{R}_+$

$$\|x(t;\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} = \left\{ \int_\Omega \left[x_1(t;\omega)^2 + \cdots + x_n(t;\omega)^2 \right] d\mu(\omega) \right\}^{\frac{1}{2}}.$$

$C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu))$ is a linear space over the nonnegative real numbers with the usual definitions of addition and scalar multiplication for continuous functions. It should also be noted that C_c is locally convex with topology defined by the following family of semi-norms, Yoshida [207]

$$\left\{ \|x(t;\omega)\|_n : \|x(t;\omega)\|_n = \sup_{0 \leqslant t \leqslant n} \left[\int_\Omega |x(t;\omega)|^2 d\mu(\omega) \right]^{\frac{1}{2}}, \quad n = 1,2,\ldots \right\}.$$

Let T denote a linear operator from the space $C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu))$ into itself; and let B and D denote Banach spaces contained C_c.

Definition 2.2.2. The pair of Banach spaces (B,D) is called *admissible* with respect to the operator T if and only if $TB \subseteq D$.

Definition 2.2.3. The operator T is called *closed* if

$$x(t;\omega)_k \xrightarrow{B} x(t;\omega)$$

and

$$(Tx_k)(t;\omega) \xrightarrow{D} y(t;\omega)$$

imply that

$$(Tx)(t;\omega) = y(t;\omega).$$

Definition 2.2.4. The Banach space B is called *stronger* than the space $C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu))$ if every sequence which converges in B with respect to its norm also converges in C_c. The converse need not be true.

The following lemmas due to Tsokos [184] and Banach's fixed point theorem are the basic tools used in the following results.

Lemma 2.2.1. *Let T be a continuous operator from $C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu))$ into itself. If B and D are Banach spaces stronger than C_c; and if the pair (B,D) is admissible with respect to T, then T is a continuous operator from B to D.*

Theorem 2.2.1 (Banach's Fixed Point Theorem). *If T is a contraction operator from a Banach space B into itself, then there exists a unique point $x^* \in B$ such that $T(x^*) = x^*$. That is, $x^* \in B$ is the unique fixed point of the operator T.*

Since T is a continuous linear operator from B to D, it is bounded in the sense that there exists a constant $M > 0$ such that

$$\|(Tx)(t;\omega)\|_D \leqslant M\|x(t;\omega)\|_B$$

for $x(t;\omega) \in B$. Thus we can define a norm for the operator T by

$$\|T\|_0 = \sup\left[\frac{\|(Tx)(t;\omega)\|_D}{\|x(t;\omega)\|_B} : \;\; x(t;\omega) \in B, \;\; \|x(t;\omega)\|_B \neq 0\right].$$

We are also guaranteed that

$$\|(Tx)(t;\omega)\|_D \leqslant \|T\|_0\|x(t;\omega)\|_B.$$

We can now state and prove a theorem on the existence and uniqueness of a random solution of a stochastic integral equation of which equation (2.2.1) is a special case.

2.2.1 *An Existence Theorem*

Consider a stochastic integral equation of the general form

$$x(t;\omega) = h(t,x(t;\omega)) + \int_0^t k(t,x(t;\omega);\omega)d\tau, \quad t \geqslant 0 \tag{2.2.1}$$

where

i) as usual $\Omega = \{\text{all } \omega\}$ is the supporting set of the complete probability measure space $(\Omega, \mathrm{A}, \mu)$ and $x(t;\omega)$ is the unknown n-dimensional vector-valued random function defined on $\mathbb{R}_+$;
ii) under appropriate conditions the stochastic kernel $k(\tau,x(t;\omega);\omega)$ is an n-dimensional vector-valued random function defined on $\mathbb{R}_+$;
and
iii) for each $t \in \mathbb{R}_+$ and each random vector $x(t;\omega)$, the stochastic free term $h(t,x(t;\omega))$ is an n-dimensional vector-valued random variable.

We now state an existence theorem.

Theorem 2.2.2. *Assume that equation (2.2.1) satisfies the following conditions:*

(i) *$B \subseteq C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ and $D \subseteq C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ are Banach spaces stronger than $C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$;*
(ii) *the pair (B,D) is admissible with respect to the operator T given by $(Tx)(t;\omega) = \int_0^t x(t;\omega)d\tau$;*

(iii) $k(t,x(t;\omega);\omega)$ *is a mapping from the set* $D_\rho = \{x(t;\omega) \in D : \|x(t;\omega)\|_D \leqslant \rho,$ $\rho \geqslant 0\}$ *into the space B such that* $\|k(t,x(t;\omega);\omega) - k(t,y(t;\omega);\omega)\|_B \leqslant \lambda \|x(t;\omega) - y(t;\omega)\|_D$ *for* $x(t;\omega)$ *and* $y(t;\omega)$ *in* D_ρ *and* $\lambda \geqslant 0$ *a constant;*
and

(iv) $x(t;\omega) \to h(t,x(t;\omega))$ *is a mapping from* D_ρ *into D such that* $\|h(t,x(t;\omega) - h(t,y(t;\omega))\|_D \leqslant \gamma \|x(t;\omega) - y(t;\omega)\|_D$ *for some* $\gamma \geqslant 0$.

Then there exists a unique random solution of equation (2.2.1) in D_ρ *provided that* $\gamma + \lambda M < 1$ *where* $M = \|T\|_0$ *and* $\|h(t,x(t;\omega))\|_D + M\|k(t,x(t;\omega);\omega)\|_B \leqslant \rho$.

The conditions on the above theorem can be weakened somewhat. We prove the following

Corollary 2.2.1. *Assume that equation (2.1.1) satisfies the conditions of Theorem 2.2.3. Then there exists a unique random solution if* $\gamma + \lambda M \leqslant 1$ *where* $M = \|T\|_0$ *and*

$$\|h,(t,x(t;\omega))\|_D + M\|k,(t,x(t;\omega))\|_B \leqslant \rho.$$

Proof. Note that the operator $(Tx)(t;\omega) = \int_0^t x(\tau;\omega)d\tau$ is continuous from B to D, hence bounded. We shall define a contraction mapping on D_ρ and then apply Banach's fixed point theorem. Define the operator U from D_ρ into D by

$$(Ux)(t;\omega) = h(t,x(t;\omega)) + \int_0^t k(\tau,x(\tau;\omega);\omega)d\tau.$$

To show inclusions consider any $x(t;\omega) \in D_\rho$.

$$\begin{aligned}
\|(Ux)(t;\omega)\|_D &= \left\| h\left(t,x(t;\omega) + \int_0^t k(\tau,x(\tau;\omega);\omega)\,d\tau\right)\right\|_D \\
&\leqslant \|h(t,x(t;\omega))\|_D + \left\|\int_0^t k(\tau,x(\tau;\omega);\omega)\,d\tau\right\|_D \\
&\leqslant \|h(t,x(t;\omega))\|_D + M\|k(t,x(t;\omega);\omega)\|_B \leqslant \rho, \quad \text{by hypothesis.}
\end{aligned}$$

Hence $(Ux)(t;\omega) \in D_\rho$ or $UD_\rho \subseteq D_\rho$.

Now let $x(t;\omega)$ and $y(t;\omega)$be elements of D_ρ. Since $(Ux)(t;\omega)$ and $(Uy)(t;\omega)$ are elements of the Banach space D, $[(Ux)(t;\omega) - (Uy)(t;\omega)] \in D$.

Thus,

$$
\begin{aligned}
&\|(Ux)(t;\omega)-(Uy)(t;\omega)\|_D \\
&= \left\| h(t,x(t;\omega)) + \int_0^t k(\tau,x(\tau;\omega);\omega)d\tau - h\left(t,y(t;\omega) - \int_0^t k(\tau,y(\tau;\omega);\omega)d\tau\right) \right\|_D \\
&= \left\| h(t,x(t;\omega)) - h(t,x(t;\omega)) + \int_0^t [k(\tau,x(\tau;\omega);\omega) - k(\tau,y(\tau;\omega);\omega)d\tau] \right\|_D \\
&\leqslant \|h(t,x(t;\omega)) - h(t,y(t;\omega))\|_D + \left\| \int_0^t [k(\tau,x(\tau;\omega);\omega) - k(\tau,y(\tau;\omega);\omega)d\tau] \right\|_D \\
&\leqslant \gamma\|x(t;\omega)-y(t;\omega)\|_D + \|T\|_0\|k(t,x(t;\omega);\omega) - k(t,y(t;\omega);\omega)\|_B \\
&\leqslant \gamma\|x(t;\omega)-y(t;\omega)\|_D + M\|k(t,x(t;\omega);\omega) - k(t,y(t;\omega);\omega)\|_B \\
&\leqslant \gamma\|x(t;\omega)-y(t;\omega)\|_D + M\lambda\|x(t;\omega)-y(t;\omega)\|_D \\
&= (\gamma+M\lambda)\|x(t;\omega)-y(t;\omega)\|_D.
\end{aligned}
$$

Thus we see that we need only to require that $(\gamma+M\lambda)\leqslant 1$ for the condition of the contraction mapping principle to be satisfied. Then, by Banach's fixed point theorem, there exists a unique point $x(t;\omega)\in D_\rho$ such that

$$(Ux)(t;\omega) = h(t,x(t;\omega)) + \int_0^t k(\tau,x(\tau;\omega);\omega)\,d\tau = x(t;\omega).$$

□

2.3 Existence of a Solution for a Stochastic Linear Pursuit-Evasion Game

2.3.1 *A General Stochastic Linear Pursuit-Evasion Game*

Consider a stochastic linear pursuit-evasion game described by a stochastic transition equation in (2.1.1). The problem is to choose controls $v(t;\omega)$ and $u_v(t;\omega)$ such that $x(t_{u,v};\omega)\in M_\varepsilon$ for some finite time $t_{u,v}$ where the terminal set M_ε is defined by

$$M_\varepsilon = \left\{x(t;\omega); \|x(t;\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \varepsilon\right\}.$$

As mentioned previously, we give only one transition equation. In case we have two objects, called the pursuer and evader, then we can consider $x(t;\omega)$ as the distance between them. The evader tries to maximize this distance or maximize the time until $\|x(t;\omega)\|\leqslant\varepsilon$ while the pursuer tries to minimize these conditions. Thus, by a simple transformation, a pursuit-evasion game becomes a contest to bring a point in n-dimensional space into an ε-ball about the origin. The pursuer, using $u(t;\omega)$, tries to minimize the time required while the evader, using $v(t;\omega)$, tries to maximize the time. If possible, he would like it to be infinite.

The state space of a differential game can be thought of as divided into two regions. In one region one player is able to force a win on the other; while in the other region the reverse happens. Isaacs uses the term *barrier* to define the boundary between the two regions. The physical interpretation is that if the initial state is outside the barrier, then the state can never be brought to the origin. That is, escape always occurs outside the barrier. From the control theory point of view, this represents an uncontrollable region. Inside the barrier, in the controllable region, capture always occurs.

In the deterministic setting Pontryagin [139], Pshenichnity [142], Sakawa [160], and other researchers have all given conditions which are sufficient for a linear differential game to be completed. We now consider conditions for completing the most general stochastization of a linear pursuit-evasion game.

The stochastic transition equation (2.2.1) is the most general formalization of a stochastic linear pursuit-evasion game in the sense that all the functions involved are stochastic. It is more general because the random function $x(t;\omega)$ appears on the right hand side. Physically this means that the object (s) being controlled have energy of their own. We may think, for example, of an incoming guided missile Dix [46]. The missile has its own guidance system; and its mission is to descend to a certain altitude over a given city before exploding. The pursuer (enemy in this case) is also sending control signals to the missile while our own forces (the evader) are trying to jam the signals as well as the onboard controls.

2.3.2 *A Special Case of Equation (2.2.1)*

Equation (2.1.1) is equivalent to a vector stochastic integral equation of the form

$$x(t;\omega) = x_0(\omega)e^{A(\omega)t} + \int_0^t e^{A(\omega)(t-\tau)}\left[B(\omega)u(\tau,\omega) - C(\omega)v(\tau;\omega)\right]d\tau, \quad t \geqslant 0 \tag{2.3.1}$$

for which we now give conditions for the existence and uniqueness of a random solution. Referring to equation (2.2.1) we can make the following identifications:

$$\begin{aligned} h(t,x(t;\omega)) &= x_0(\omega)e^{A(\omega)t} \\ k(t,x(\tau;\omega);\omega) &= e^{A(\omega)(t-\tau)}\left[B(\omega)u(\tau;\omega) - C(\omega)v(\tau;\omega)\right]. \end{aligned}$$

We note that conditions (ii) and (iii) under equation (2.2.1) are satisfied. In particular

ii) the stochastic kernel is an n-dimensional vector valued random function from $\mathbb{R}_+$ into $L_2(\Omega, \mathrm{A}, \mu)$;

and

iii) the stochastic free term $x_0(\omega)e^{A(\omega)t}$ is an n-dimensional vector-valued random variable, i.e. for each $t \in \mathbb{R}_+$, $x_0(\omega)e^{A(\omega)t} \in L_2(\Omega, \mathrm{A}, \mu)$.

Note that the Banach space $C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ satisfies the definition of *stronger than itself*. Thus we can use the space $C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ in place of both B and D in Theorem 2.2.3. Clearly the pair (C_c, C_c) is admissible with respect to T given by $(Tx)(t;\omega) = \int_0^t x(\tau;\omega)d\tau$. Condition (iii) of Theorem 2.2.3 is satisfied vacuously since $x(t;\omega)$ does not appear explicitly in the stochastic kernel. That is,

$$\|k(t, x(t;\omega);\omega) - k(t, y(t;\omega);\omega)\|_{C_c} = 0 \;\; \mu - \text{a.e.}$$

We proceed by proving a theorem concerning the existence and uniqueness of a random solution for equation (2.3.1) and hence (2.1.1).

Theorem 2.3.1. *Given any $\rho \geqslant 0$, define the set D_ρ by*

$$D_\rho = \{x(t;\omega) \in C_c : \|x(t;\omega)\|_{C_c} \leqslant \rho\}.$$

There exists a unique random solution of equation (2.3.2) provided that

(i) *the initial condition $x(0;\omega) = x_0(\omega) \in D_\rho$*
and
(ii) $\left|e^{A(\omega)t}\right| \leqslant 1$.

Proof. The proof of this theorem will consist of showing that all the conditions of Corollary 2.2.4 are satisfied.

1) The Banach space $C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ satisfies the definition of stronger than itself.
2) The pair (C_c, C_c) is admissible with respect to the operator T given by
$$(Tx)(t;\omega) = \int_0^t x(\tau;\omega)d\tau.$$
3) The stochastic kernel is a mapping from the set D into the space C_c such that $\|k(t, x(t;\omega);\omega) - k(t, y(t;\omega);\omega)\|_{C_c} = 0$ μ-a.e. for $x(t;\omega)$ and $y(t;\omega)$ in D_ρ. We just take $\lambda = 0$.
4) The stochastic free term is such that
$$\left\|x_0(\omega)e^{A(\omega)t} - y_0(\omega)e^{A(\omega)t}\right\|_{C_c} \leqslant \gamma \|x_0(\omega) - y_0(\omega)\|_{C_c} \text{ for some } \gamma \geqslant 0.$$

We just take $\gamma = \left|e^{A(\omega)t}\right|$. The conclusion then reduces to: There exists a unique random solution of equation (2.3.1) in D_ρ provided that $\gamma \leqslant 1$. We have assumed that $\left|e^{A(\omega)t}\right| \leqslant 1$; hence, the theorem is proven. □

In the next section of this study we shall consider the existence and uniqueness of a random solution of a stochastic linear pursuit-evasion game with deterministic controls.

2.4 The Solution of a Stochastic Linear Pursuit-Evasion Game With Nonrandom Controls

In this section we shall be concerned with stochastic pursuit-evasion games described by stochastic linear differential equations of the form

$$\frac{d}{dt}x(t;\omega) = A(\omega)x(t;\omega) + Bu(t) - Cv(t), \quad t \geqslant 0 \tag{2.4.1}$$

where

i) $\omega \in \Omega$, where Ω is the supporting set of a complete measure space $(\Omega, \mathrm{A}, \mu)$;
ii) $x(t;\omega)$ is the unknown $(n \times 1)$ random state variable;
iii) $u(t) \in E^r$ is the control vector of the pursuer, $v(t;\omega) \in E^s$ is the control vector of the evader; and
iv) $A(\omega)$ is a $(n \times n)$ random matrix and B and C are respectively $(n \times r)$ and $(n \times s)$.

It is immediately obvious that equation (2.4.1) is a special case of equation (2.1.1). This equation is still general in the sense that $x(t;\omega)$ appears on the right hand side; but here we consider deterministic control vectors. Physically this means that the controllers are attempting to control a randomly varying object with non-random controls. Thinking of an incoming missile, the path which it is following cannot be fitted to a deterministic trajectory. On the other hand if we are thinking of $x(t;\omega)$ as some function of the distance between a pursuer and an evader, then $x(t;\omega)$ could be random because either or both of the players are following random paths or because the players cannot measure the distance accurately.
Mathematically this means that the state is being affected by some stochastic process $w(t;\omega)$, but since the players cannot observe Ω, they seek optimal deterministic controls.
The purpose of this chapter is threefold. We will find the smallest max-min completion time for the game (2.4.1) as well as optimal controls for the pursuer and evader. Finally we will give sufficient conditions for completion of the game in a finite time.

2.4.1 *Preliminaries*

The above random differential system (2.4.1) can easily be reduced to the stochastic vector integral equation

$$x(t;\omega) = \Phi(t;\omega)x_0(\omega) + \int_0^t \Phi(t-\tau;\omega)[Bu(\tau) - Cv(\tau)]d\tau.$$

with initial conditions

$$x(0;\omega) = x_0(\omega)$$

where the matrix $\Phi(t;\omega)$ is given by $\Phi(t;\omega) = e^{A(\omega)(t)}$.

The problem is to choose controls $v(t)$ and $u_v(t)$ such that $x(t_{u,v};\omega) \in M_\varepsilon$ for some finite time $t_{u,v}$, where M_ε was defined in Section 2.2 as an ε-ball about the origin.

We shall consider the random solution $x(t;\omega)$ and the stochastic free term $\Phi(t;\omega)$ as functions of the real argument t with values in the space $L_2(\Omega,\mathrm{A},\mu)$. The function $[Bu(t)-Cv(t)]$ is also a function of the real argument t whose values are in $L_2(\Omega,\mathrm{A},\mu)$. The stochastic kernel $\Phi(t-\tau;\omega)$ is an essentially bounded function with respect to μ for every t and τ, $0 \leqslant \tau \leqslant t < \infty$, with values in $L_\infty(\Omega,\mathrm{A},\mu)$. Thus the product $\Phi(t-\tau;\omega)[Bu(\tau)-Cv(\tau)]$ will always be in the space $L_2(\Omega,\mathrm{A},\mu)$. We shall assume that the mapping

$$(t,\tau) \to \Phi(t-\tau;\omega)$$

from the set

$$\Delta = \{(t,\tau) : 0 \leqslant \tau \leqslant t < \infty\}$$

into $L_\infty(\Omega,\mathrm{A},\mu)$ is continuous. That is,

$$\mu\text{-}\operatorname*{ess\,sup}_{\omega} |\Phi(t_n-\tau_n;\omega) - \Phi(t-\tau;\omega)| \to 0$$

as $n \to \infty$ whenever $(t_n,\tau_n) \to (t,\tau)$ as $n \to \infty$.

We shall define as *admissible controls* all measureable functions whose values belong (almost everywhere) to some given compact sets $U \subset E^r$ and $V \subset E^s$. $u(t) \in U$, $v(t) \in V$ for $t \geqslant 0$. Further, we shall assume that U is convex.

The terminal set, M_ε, is just an ε-ball about the zero element of $L_2(\Omega,\mathrm{A},\mu)$. As mentioned previously, the problem is to choose admissible controls $v(t)$ and $u_v(t)$ such that

$$\Phi(t_{u,v};\omega)x_0(\omega) + \int_0^{t_{u,v}} \Phi(t_{u,v}-\tau)[Bu_v(\tau)-Cv(\tau)]\,d\tau \in M_\varepsilon \tag{2.4.2}$$

for some $t_{u,v} \in \mathbb{R}_+$.

Definition 2.4.1. The game (2.4.1) is said to be *completed from an initial point* $x(0;\omega) = x_0(\omega)$, if, no matter what control $v(t)$ the evader chooses, the pursuer can choose a control $u_v(t)$ such that $x(t;\omega) \in M_\varepsilon$ for some finite time t.

We shall define the functions $H_U(\eta)$ and $H_V(\xi)$ by

$$\begin{aligned} H_U(\eta) &= \sup_{u\in U} \eta u; \\ H_V(\xi) &= \sup_{v\in V} \xi v \end{aligned} \tag{2.4.3}$$

where η and ξ are arbitrary $(r\times 1)$ and $(s\times 1)$ vectors. Then there exist vectors $u_\eta \in U$ and $v_\xi \in V$ such that

$$\begin{aligned} H_U(\eta) &= \sup_{u\in U} \eta u = \eta u_\eta; \quad \text{and} \\ H_V(\xi) &= \sup_{v\in V} \xi v = \xi v_\xi. \end{aligned} \tag{2.4.4}$$

It can be shown that the function $H_U(\eta)\,[H_V(\xi)]$ defined by (2.4.3) is continuous with respect to $\eta[\xi]$. Furthermore, if $u_\eta\,\left[V_\xi\right]$ is uniquely determined in some neighborhood of $\eta[\xi]$, then $u_\eta\,\left[V_\xi\right]$ is continuous in that neighborhood.

For convenience we shall define the $(n\times r)$ and $(n\times s)$ matrices $K(t;\omega)$ and $L(t;\omega)$ by

$$K(t;\omega) = \Phi(t;\omega)B;$$

$$L(t;\omega) = \Phi(t;\omega)C.$$

Equation (2.4.2) can now be rewritten as

$$\Phi(t_{u,v};\omega)x_0(\omega) + \int_0^{t_{u,v}} K(\tau;\omega)u_v(t_{u,v}-\tau)d\tau - \int_0^{t_{u,v}} L(\tau;\omega)v(t_{u,v}-\tau)d\tau \in M_\varepsilon \tag{2.4.5}$$

Theorem 2.4.1. *Given any admissible control $v(t)$, a necessary and sufficient condition for the existence of an admissible control $u_v(t)$ such that (2.3.5) holds for some finite time $t_{u,v} \geqslant 0$ is the existence of a $t \in \mathbb{R}_+$ such that*

$$-\varepsilon \leqslant \lambda\Phi(t;\omega)x_0(\omega) + \int_0^t H_U(\lambda K(\tau;\omega))d\tau - \int_0^t H_V(\lambda L(\tau;\omega))d\tau \tag{2.4.6}$$

for all $(1\times n)$ vectors $\lambda(\omega) = \lambda$ such that $\|\lambda\|_{L_2(\Omega,\mathrm{A},\mu)} = 1$.

Proof. Let λ be an arbitrary $(1\times n)$ vector such that $\|\lambda\|_{L_2(\Omega,\mathrm{A},\mu)} = 1$. Multiplying the left hand side of line (2.3.5) by $-\lambda$ on the left and applying Schwarz's inequality gives

$$-\lambda\Phi(t_{u,v};\omega)x_0(\omega) - \int_0^{t_{u,v}} \lambda K(\tau;\omega)u_v(t_{u,v}-\tau)d\tau + \int_0^{t_{u,v}} \lambda L(\tau;\omega)v(t_{u,v}-\tau)d\tau \leqslant \varepsilon.$$

Since the above inequality must hold for all $v(t)\in V$, it must hold for $\sup_{v\in V}\lambda L(t;\omega)v = H_V(\lambda L(t;\omega)) \geqslant \lambda L(t;\omega)v(t_{u,v}-t)$.

By definition, $H_U(\lambda K(t;\omega)) \geqslant \lambda K(t;\omega)u_v(t_{u,v}-t)$. Hence

$$\lambda\Phi(t_{u,v};\omega)x_0(\omega) + \int_0^{t_{u,v}} H_U(\lambda K(\tau;\omega))\,d\tau - \int_0^{t_{u,v}} H_V(\lambda L(\tau;\omega))d\tau \geqslant -\varepsilon.$$

Putting $t = t_{u,v}$ yields condition (2.4.6).

Now suppose that there is an admissible control $v(t)$ such that no admissible control $u_v(t)$ exists such that (2.4.5) holds for some finite time t. This means that the compact, convex set defined by

$$\left\{\int_0^t k(\tau;\omega)u(t-\tau)d\tau: \;\; u(-\tau)\in U\right\}$$

does not intersect the compact sphere

$$-\Phi(t;\omega)x_0(\omega)+\int_0^t L(\tau;\omega)v(t-\tau)d\tau+M_\varepsilon.$$

Therefore, there is a vector $\lambda \in L_2(\Omega, \mathrm{A}, \mu)$, $\|\lambda\|_{L_2(\Omega,\mathrm{A},\mu)}=1$, such that

$$-\lambda\Phi(t;\omega)x_0(\omega)+\int_0^t \lambda L(\tau;\omega)v(t-\tau)d\tau+\lambda a>\int_0^t K(\tau;\omega)u(t-\tau)d\tau \tag{2.4.7}$$

for all $u(t)\in U$, $0\leqslant \tau\leqslant t<\infty$, and for all $a\in M_\varepsilon$. Since inequality (2.4.7) must hold for a $u(t)\in U$ such that

$$\lambda K(\tau;\omega)u(t-\tau)=H_U(\lambda K(\tau;\omega))=\sup_{u\in U}\lambda K(\tau;\omega)u$$

and for a vector $\alpha=-\varepsilon\lambda'\in M_\varepsilon$, and since

$$\int_0^t H_V\left(\lambda L(\tau;\omega)\right)d\tau \geqslant \int_0^t \lambda L(\tau;\omega)v(t-\tau)d\tau,$$

$$\lambda\left(-\varepsilon\lambda'\right)>\lambda\Phi(t;\omega)x_0(\omega)+\int_0^t H_U(\lambda K(\tau;\omega))d\tau-\int_0^t H_V\left(\lambda L(\tau;\omega)\right)d\tau$$

contradicting inequality (2.4.6) □

Corollary 2.4.1. *Given any admissible control $v(t)$, a necessary and sufficient condition for the existence of an admissible control $u_v(t)$ such that (2.4.5) holds for some finite time $t_{u,v}\geqslant 0$ is that there exists a $t\in\mathbb{R}_+$ such that*

$$\inf_{\lambda\in Q}\left[\lambda\Phi(t;\omega)x_0(\omega)+\int_0^t H_U(\lambda K(\tau;\omega))d\tau-\int_0^t H_V\left(\lambda L(\tau;\omega)\right)d\tau\right]\geqslant-\varepsilon$$

where Q is a set of $(1\times n)$ vectors $\lambda\in L_2(\Omega,\mathrm{A},\mu)$ such that $\|\lambda\|_{L_2(\Omega,\mathrm{A},\mu)}=1$.

We shall denote by $u(t,\lambda)$ and $v(t,\lambda)$ the vectors $u\in U$ and $v\in V$ which maximize $\lambda K(t;\omega)u$ and $\lambda L(t;\omega)v$. That is,

$$H_U(\lambda K(t;\omega))=\sup_{u\in U}\lambda K(t;\omega)u=\lambda K(t;\omega)u(t,\lambda)$$

and

$$H_V(\lambda L(t;\omega))=\sup_{v\in V}\lambda L(t;\omega)v=\lambda L(t;\omega)v(t,\lambda).$$

Assume that for each $\lambda\in Q$, the controls $u(\tau,\lambda)$ and $v(\tau,\lambda)$ are uniquely determined for all $\tau\in[0,T]$ except on a set of measure zero. Then, see the remark following equation (2.4.4), the controls $u(\tau,\lambda)$ and $v(\tau,\lambda)$ are piecewise continuous on $[0,T]$.

The scalar function $F\left(t,\lambda;\omega,x_0(\omega)\right)$ will be defined by

$$\begin{aligned}
&F(t,\lambda;\omega,x_0(\omega))\\
&=\lambda\Phi(t;\omega)x_0(\omega)+\int_0^t H_U(\lambda K(\tau;\omega))d\tau-\int_0^t H_V(\lambda L(\tau;\omega))d\tau \qquad (2.4.8)\\
&=\lambda\Phi(t;\omega)x_0(\omega)+\lambda\int_0^t K(\tau;\omega)u(\tau,\lambda)(\lambda)d\tau-\lambda\int_0^t L(\tau;\omega)v(\tau,\lambda)d\tau.
\end{aligned}$$

Lemma 2.4.1. *The gradient vector with respect to* λ *of the function* $F(t,\lambda;\omega,x_0(\omega))$ *is given by*

$$\operatorname{grad}_\lambda F(t,\lambda;\omega,x_0(\omega)) = x(t,\lambda;\omega,x_0(\omega))$$

where

$$x(t,\lambda;\omega,x_0(\omega)) = \Phi(t;\omega)x_0(\omega) + \int_0^t K(\tau;\omega)u(\tau,\lambda)d\tau - \int_0^t L(\tau;\omega)v(\tau,\lambda)d\tau. \quad (2.4.9)$$

Moreover $\operatorname{grad}_\lambda F(t,\lambda;\omega,x_0(\omega))$ *is continuous in* t *and* λ.

Proof. Let γ be an arbitrary $(1\times n)$ vector. Then, from the definition of $u(t,\lambda)$,

$$\begin{aligned} H_U((\lambda+\gamma)K(t;\omega)) - H_U(\lambda K(t;\omega)) &\geqslant (\lambda+\gamma)K(t;\omega)u(t,\lambda) - \lambda K(t;\omega)u(t,\lambda) \\ &= \gamma K(t;\omega)u(t,\lambda), \end{aligned}$$

and

$$\begin{aligned} H_U((\lambda+\gamma)K(t;\omega)) - H_U(\lambda K(t;\omega)) &\leqslant (\lambda+\gamma)K(t;\omega)u(t,\lambda+\gamma) - \lambda K(t;\omega)u(t,\lambda+\gamma) \\ &= \gamma K(t;\omega)u(t,\lambda+\gamma). \end{aligned}$$

Integrating with respect to t we get

$$\begin{aligned} \gamma\int_0^t K(\tau;\omega)u(\tau,\lambda)\,d\tau &\leqslant \int_0^t H_U((\lambda+\gamma)K(\tau;\omega))d\tau - \int_0^t H_U(\lambda K(t;\omega))d\tau \\ &\leqslant \int_0^t K(\tau;\omega)u(\tau,\lambda+\gamma). \end{aligned} \quad (2.4.10)$$

Let $t_1,t_2,\ldots,t_N$ $(0<t_1<t_2<\cdots<t_N<t)$ be the points where $u(t,\lambda)$ is not continuous and define the following subintervals of $[0,t]$:

$$\begin{aligned} I_0(\varepsilon) &= [0,\varepsilon) \\ I_i(\varepsilon) &= (t_i-\varepsilon,t_i+\varepsilon), \quad i=1,2,\ldots,N \\ I_{N+1}(\varepsilon) &= (t-\varepsilon,t] \\ I(\varepsilon) &= [0,t] - \bigcup_{i=0}^{N+1} I_i(\varepsilon). \end{aligned}$$

By the continuity of u, for sufficiently small $\varepsilon>0$ there exists a $\delta=\delta(\varepsilon)>0$ such that if $\|\gamma\|<\delta(\varepsilon)$ and $t\in I(\varepsilon)$, then $\|u(t,\lambda+\gamma)-u(t,\lambda)\|<\varepsilon$.

Since U is compact (closed and bounded), there is a $k>0$ such that $\|u(t,\lambda+\gamma)-u(t,\lambda)\|<k$ if $t\in\bigcup_{i=0}^{N+1} I_i(\varepsilon)$.

Therefore,

$$\int_0^t \|u(t,\lambda+\gamma)-u(t,\lambda)\|d\tau < \varepsilon t + 2\varepsilon(N+1)k. \quad (2.4.11)$$

Inequalities (2.4.10) and (2.4.11) imply that

$$\text{grad}_\lambda \int_0^t \lambda K(\tau;\omega)u(\tau,\lambda)d\tau = \int_0^t K(\tau;\omega)u(\tau,\lambda)d\tau.$$

similarly,

$$\text{grad}_\lambda \int_0^t \lambda L(\tau;\omega)v(\tau,\lambda)d\tau = \int_0^t L(\tau;\omega)v(\tau,\lambda)d\tau.$$

Hence (2.4.9) is proven. The continuity of $\text{grad}_\lambda F(t,\lambda;\omega,x_0(\omega))$ is evident from the course of the proof. □

Since $F(t,\lambda;\omega,x_0(\omega))$ is continuous in λ and the set

$$Q = \left\{\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_n) \in L_2(\Omega,\mathrm{A},\mu) : \|\lambda\|_{L_2(\Omega,\mathrm{A},\mu)} = 1\right\}$$

is compact. Thus, there is a $\lambda \in Q$ which attains the infimum of $F(t,\lambda;\omega,x_0(\omega))$. Let us denote it by λ_t. That is

$$\inf_{\lambda\in Q} F(t,\lambda;\omega,x_0(\omega)) = F(t,\lambda_t;\omega,x_0(\omega)).$$

For convenience, when the initial condition is $x_0(\omega)$, we will write $F_\omega(t,\lambda)$ and $x_\omega(t,\lambda)$ instead of $F(t,\lambda;\omega,x_0(\omega))$and $x(t,\lambda;\omega,x_0(\omega))$ respectively.

Lemma 2.4.2. *We have*

$$\inf_{\lambda\in Q} F_\omega(t,\lambda) = F_\omega(t,\lambda_t) = -\|x_\omega(t,\lambda_t)\|_{L_2(\Omega,\mathrm{A},\mu)} \tag{2.4.12}$$

where $x_\omega(t,\lambda_t)$ is given by equation (2.4.9).

Proof. Since $\min F_\omega(t,\lambda)$ is sought for $\|\lambda\|^2_{L_2(\Omega,\mathrm{A},\mu)} - 1 = 0$ and t is fixed, define

$$\overline{F}_\omega(t,\lambda,\theta) = F_\omega(t,\lambda) + \theta\left(\|\lambda\|^2_{L_2(\Omega,\mathrm{A},\mu)} - 1\right)$$

where θ is a Lagrange multiplier. Put

$$\frac{\partial}{\partial\lambda_i}\overline{F}_\omega = x_i(t,\lambda) + 2\theta\lambda_i = 0, \quad i = 1,2,\ldots,n$$

where $x_i(t,\lambda)$ and λ_i denote the ith components of $x_\omega(t,\lambda)$ and λ respectively. Solving we get

$$\lambda_i = \frac{x_i}{2\theta}.$$

$$\|\lambda\|^2_{L_2(\Omega,\mathrm{A},\mu)} = \int_\Omega \sum_{i=1}^n \left(\frac{x_i}{2\theta}\right)^2 d\mu(\omega) = 1.$$

$$\int_\Omega \sum_{i=1}^n x_1^2\, d\mu(\omega) = 4\theta^2.$$

$$\|x_\omega(t,\lambda)\|_{L_2(\Omega,\mathrm{A},\mu)} = 2\theta.$$

Hence,

$$\lambda_t = \frac{-x'_\omega(t,\lambda)}{\|x_\omega(t,\lambda)\|_{L_2(\Omega,\mathrm{A},\mu)}}. \tag{2.4.13}$$

Substituting line (2.4.13) into line (2.4.8) gives the desired result (2.4.12). □

Lemma 2.4.3. *Let us assume that for any time $t > 0$ and for $\lambda_1, \lambda_2 \in Q$,*

$$\|x_\omega(t,\lambda_1)\|_{L_2(\Omega,\mathrm{A},\mu)} = \|x_\omega(t,\lambda_2)\|_{L_2(\Omega,\mathrm{A},\mu)}$$

implies that $\lambda_1 = \lambda_2$. Then,

$$\frac{d}{dt}F_\omega(t,\lambda_t) = \lambda_t A(\omega)\Phi(t;\omega)x_0(\omega) + H_U(\lambda_t K(t;\omega)) - H_V(\lambda_t L(t;\omega)).$$

Proof. Let δ be an arbitrary real number. Since the matrix $\Phi(t;\omega) = e^{A(\omega)t}$, we see that

$$\Phi(t+\delta;\omega) = \Phi(t;\omega) + \int_t^{t+\delta} A(\omega)\Phi(\tau;\omega)d\tau.$$

Then,

$$\begin{aligned} F_\omega(t+\delta,\lambda) &= F_\omega(t,\lambda) \\ &+ \int_t^{t+\delta}[\lambda A(\omega)\Phi(\tau;\omega)x_0(\omega) + H_U(\lambda K(\tau;\omega)) - H_V(\lambda L(\tau;\omega))]d\tau. \end{aligned} \tag{2.4.14}$$

Now, by the definition of λ_t,

$$F_\omega(t,\lambda_{t+\delta}) \geqslant F_\omega(t,\lambda_t) = \inf_{\lambda\in Q} F_\omega(t,\lambda).$$

Thus, from (2.4.14), we get

$$\begin{aligned} F_\omega(t+\delta,\lambda_{t+\delta}) - F_\omega(t,\lambda_t) &\geqslant \int_t^{t+\delta} \lambda_{t+\delta}A(\omega)\Phi(\tau;\omega)x_0(\omega)d\tau \\ &+ \int_t^{t+\delta}[H_U(\lambda_{t+\delta}K(\tau;\omega)) - H_V(\lambda_{t+\delta}L(\tau;\omega))]d\tau. \end{aligned} \tag{2.4.15}$$

On the other hand, $F_\omega(t+\delta,\lambda_{t+\delta}) \leqslant F_\omega(t+\delta,\lambda_t)$ implies that

$$F_\omega(t+\delta,\lambda_{t+\delta}) - F_\omega(t,\lambda_t) \leqslant F_\omega(t+\delta,\lambda_t) - F_\omega(t,\lambda_t). \tag{2.4.16}$$

Since F is continuous in t, inequalities (2.4.15) and (2.4.16) show the continuity of F in t and λ_t. That is,

$$F_\omega(t+\delta,\lambda_{t+\delta}) \to F_\omega(t,\lambda_t) \quad \text{as } \delta \to 0. \tag{2.4.17}$$

From equation (2.4.12) it is clear that the assumption of Lemma 2.4.5 implies the uniqueness of the $\lambda \in Q$ which attains the infimum of $F_\omega(t,\lambda)$. It then follows from the continuity of F_Ω, line (2.4.17), that

$$\lambda_{t+\delta} \to \lambda_t \quad \text{as } \delta \to 0. \tag{2.4.18}$$

If $\delta > 0$, we get from inequalities (2.4.15) and (2.4.16)

$$\begin{aligned} &\frac{1}{\delta}\int_t^{t+\delta}[\lambda_{\tau+\delta}A(\omega)\Phi(\tau;\omega)x_0(\omega) + H_U(\lambda_{\tau+\delta}K(\tau;\omega)) - H_V(\lambda_{\tau+\delta}L(\tau;\omega))]d\tau \\ &\qquad \leqslant \frac{1}{\delta}[F_\omega(t+\delta,\lambda_{t+\delta}) - F_\omega(t,\lambda_t)] \\ &\qquad \leqslant \frac{1}{\delta}[F_\omega(t+\delta,\lambda_t) - F_\omega(t,\lambda_t)]. \end{aligned} \tag{2.4.19}$$

In view of (2.4.18) and the continuity of $H_U(\lambda K(t;\omega))$ and $H_V(\lambda L(t;\omega))$ in λ and t, it follows from inequality (2.4.19) that

$$\frac{d}{dt}F_\omega(t,\lambda_t) = \lambda_t A(\omega)\Phi(t;\omega)x_0(\omega) + H_U(\lambda_t K(t;\omega)) - H_V(\lambda_t L(t;\omega)).$$

If $\delta < 0$, the same result holds. Thus the lemma is proven. □

We are now in a position to give conditions under which the game (2.3.1) will have a finite maximum completion time.

2.4.2 *Completion of the Game*

Suppose that $\|x_0(\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} > \varepsilon$ and there is a time $t \in \mathbb{R}_+$ such that

$$\inf_{\lambda\in Q} F_\omega(t,\lambda) = F_\omega(t,\lambda_t) = -\varepsilon. \tag{2.4.20}$$

Theorem 2.4.2. *No matter what admissible control $v(t)$, $t \in \mathbb{R}_+$, the evader chooses, the game can be completed in a time not greater than t_0, where t_0 is the smallest nonnegative time satisfying (2.4.20). Furthermore, no matter what admissible control $u(t)$, $t \in \mathbb{R}_+$, the pursuer chooses, the evader can choose a control $v(t)$ such that the game cannot be completed in a time less than t_0.*

Proof. Given an arbitrary control $v(t) \in V$, $t \in \mathbb{R}_+$, we shall define the function

$$\begin{aligned} F_V(t,\lambda;\omega,x_0(\omega)) &= \lambda\Phi(t;\omega)x_0(\omega) \\ &+ \lambda\int_0^t K(\tau;\omega)u(\tau;\lambda)\,d\tau - \lambda\int_0^t L(\tau;\omega)V(t-\tau)d\tau. \end{aligned} \tag{2.4.21}$$

From the definition of $v(t;\lambda)$ and equation (2.4.8) it is clear that

$$F_V(t,\lambda;\omega,x_0(\omega)) \geqslant F(t,\lambda;\omega,x_0(\omega))$$

for all $\lambda \in Q$. Hence,

$$\inf_{\lambda\in Q} F_v(t_0,\lambda;\omega,x_0(\omega)) \geqslant \inf_{\lambda\in Q} F_\omega(t_0,\lambda;\omega,x_0(\omega)) = -\varepsilon. \tag{2.4.22}$$

Let us also define the function

$$\begin{aligned} x_v(t,\lambda_t;\omega,x_0(\omega)) &= \Phi(t;\omega)x_0(\omega) \\ &+ \lambda\int_0^t K(\tau;\omega)u(\tau;\lambda_\tau)\,d\tau - \lambda\int_0^t L(\tau;\omega)v(t-\tau)d\tau. \end{aligned} \tag{2.4.23}$$

where $\lambda_t \in Q$ attains the infimum of $F_v(t,\lambda;\omega,x_0(\omega))$ when t and $x_0(\omega)$ are fixed. Then, by Lemma 2.4.4,

$$\inf_{\lambda\in Q} F_V(t,\lambda;\omega,x_0(\omega)) = F_V(t,\lambda_t;\omega,x_0(\omega)) = -\|x_v(t,\lambda_t;\omega,x_0(\omega))\|_{L_2(\Omega,\mathrm{A},\mu)}. \tag{2.4.24}$$

Since $x_v(t,\lambda_t;\omega,x_0(\omega))$ is continuous in time t, and equations (2.4.22) and (2.4.24) imply that

$$-\left\|x_v\left(t_0,\lambda_{t_0};\omega,x_0(\omega)\right)\right\|_{L_2(\Omega,\mathrm{A},\mu)} \geqslant -\varepsilon,$$

there exists a time t^*, $0 \leqslant t^* \leqslant t_0$, such that

$$-\|x_v(t^*,\lambda_{t^*};\omega,x_0(\omega))\|_{L_2(\Omega,\mathrm{A},\mu)} \geqslant -\varepsilon.$$

That is, the game can be completed in a time t^* which is not greater than t_0.

As in (2.4.21) we shall define another function $F_u(t,\lambda;\omega,x_0(\omega))$ by

$$F_u(t,\lambda;\omega,x_0(\omega)) = \lambda\Phi(t;\omega)x_0(\omega) + \lambda\int_0^t K(\tau;\omega)u(t-\tau)d\tau - \lambda\int_0^t L(\tau;\omega)u(\tau,\lambda)d\tau.$$

Now from the definition of $u(t,\lambda)$ and equation (2.4.8) we see that

$$F_u(t,\lambda;\omega,x_0(\omega)) \leqslant F(t,\lambda;\omega,x_0(\omega))$$

for all $\lambda \in Q$. Therefore,

$$\inf_{\lambda\in Q} F_u(t_0,\lambda;\omega,x_0(\omega)) \leqslant \inf_{\lambda\in Q} F(t_0,\lambda;\omega,x_0(\omega)) = -\varepsilon. \tag{2.4.25}$$

Following equation (2.4.23), let us define the function $x_u(t,\lambda_t;\omega,x_0(\omega))$ by

$$x_u(t,\lambda_t;\omega,x_0(\omega)) = \Phi(t;\omega)x_0(\omega) + \int_0^t K(\tau;\omega)u(t-\tau)d\tau - \int_0^t L(\tau;\omega)v(\tau,\lambda_\tau)\,d\tau$$

where $\lambda_t \in Q$ attains the infimum of $F_u(t,\lambda;\omega,x_0(\omega))$. Then again by Lemma 2.4.4,

$$\inf_{\lambda\in Q} F_u(t,\lambda;\omega,x_0(\omega)) - F_u(t,\lambda_t;\omega,x_0(\omega)) = -\|x_u(t,\lambda_t;\omega,x_0(\omega))\|_{L_2(\Omega,\mathrm{A},\mu)}. \tag{2.4.26}$$

Thus, by (2.4.25) and (2.4.26),

$$-\left\|x_u\left(t_0,\lambda_{t_0};\omega,x_0(\omega)\right)\right\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant -\varepsilon.$$

That is, the game cannot be completed in time less than t_0. Thus t_0 is the maximin completion time. □

The controls, $u(t) = u\left(t_0 - t,\lambda_{t_0}\right)$ and $v(t) = v\left(t_0 - t,\lambda_{t_0}\right)$ for $t \in [0,t_0]$, are optimal in the sense that the pursuer wants to complete the game as soon as possible and the evader wants to escape as long as possible. The time t_0 is the smallest maximin completion time of the game. When will a finite time t exist such that (3.3.1) holds?

Theorem 2.4.3. *If (i) the homogeneous stochastic differential equation*

$$\frac{d}{dt}x(t;\omega) = A(\omega)x(t;\omega) \tag{2.4.27}$$

is stochastically asymptotically stable; and (ii) $BU \supset CV$ where $BU = \{Bu : u \in U\}$ and $V = \{Cv; v \in V\}$ are subsets of E^n, then the game can be completed no matter what the initial condition $x_0(\omega) \in L_2(\Omega,\mathrm{A},\mu)$ may be.

Proof. Since $CV \subset BU$, whatever control $v(t) \in V$, $t \in \mathbb{R}_+$, the evader may choose, the pursuer can choose a control, such that

$$Bu(t) = Cv(t) \text{ for all } t \geqslant 0.$$

Since (2.4.27) is assumed to be stochastically asymptotically stable, there is a finite time t such that

$$\|x(t;\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \varepsilon.$$

Since $\Phi(t;\omega) = e^{A(\omega)t}$, $A(\omega)\Phi(t;\omega) = \Phi(t;\omega)A(\omega)$. That is, we can change the order of multiplication. Thus, the conclusion of Lemma 2.3.5 can be written as

$$\frac{d}{dt}F_\omega(t,\lambda_t) = \lambda_t\Phi(t;\omega)A(\omega)x_0(\omega) + \max_{\hat{u}\in BU} \lambda_t\Phi(t;\omega)\hat{u}(t) - \min_{\hat{v}\in CV} \lambda_t\Phi(t;\omega)\hat{v}(t) \quad (2.4.28)$$

$\square$

Theorem 2.4.4. *Assume that for any $t > 0$ and for any $\lambda_1,\ \lambda_2 \in Q$,*

$$\|x_\omega(t,\lambda_1)\|_{L_2(\Omega,\mathrm{A},\mu)} = \|x_\omega(t,\lambda_2)\|_{L_2(\Omega,\mathrm{A},\mu)}$$

implies that $\lambda_1 = \lambda_2$. If there exists a $\delta > 0$ such that

$$-A(\omega)x_0(\omega) + CV + M_\delta \subset BU; \quad (2.4.29)$$

and

$$\|\lambda_t\Phi(t;\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} \geqslant \delta \quad \textit{for all } t \in \mathbb{R}_+,$$

where $M_\delta = \left\{x(t;\omega) : \|x(t;\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \delta\right\}$, then the game starting from $x_0(\omega)$ can be completed.

Proof. Let $\gamma \in L_2(\Omega,\mathrm{A},\mu)$ be an arbitrary $(1\times n)$ vector such that $\|\gamma\|_{L_2(\Omega,\mathrm{A},\mu)} \geqslant \delta > 0$. Then

$$\max_{x(t;\omega)\in M_\delta} \gamma x(t;\omega) = \gamma x_\gamma(t;\omega) \geqslant \delta^2.$$

From relation (2.4.29), for arbitrary $x(t;\omega) \in M_\delta$ and $\hat{v}(t) \in CV$ there is a $\hat{u}(t) \in BU$ such that

$$-A(\omega)x_0(\omega) + \hat{v}(t) + x(t;\omega) = \hat{u}(t).$$

Hence, for all $\hat{v}(t) \in CV$ and for all γ such that $\|\gamma\|L_2(\Omega,\mathrm{A},\mu) \geqslant \delta$, there is a $\hat{u}(t) \in BU$ such that

$$\gamma\left(\hat{u}(y) - \hat{v}(t) + A(\omega)x_0(\omega)\right) \geqslant \delta^2 > 0.$$

The above inequality still holds for a $\hat{v}_\gamma$ such that

$$\gamma\hat{v}_\gamma(t) = \max_{\hat{v}(t)\varepsilon CV} \gamma\hat{v}(t).$$

Also

$$\gamma\hat{u}(t) \leqslant \gamma\hat{u}_\gamma = \max_{\hat{u}(t)\in BU} \gamma\hat{u}(t).$$

Hence, for all γ such that $\|\gamma\|_{L_2(\Omega,\mathrm{A},\mu)} \geqslant \delta > 0$,

$$\max_{\hat{u}(t)\varepsilon BU} \gamma\hat{u}(t) - \max_{\hat{v}(t)\in CV} \gamma\hat{v}(t) + \gamma A(\omega)x_0(w) \geqslant \delta^2.$$

Under the assumption of Theorem 2.4.8, Lemma 2.4.5 implies (2.4.28). Setting $\gamma = \lambda_t\Phi(t;\omega)$, we get

$$\frac{d}{dt}F_\omega(t,\lambda_t) \geqslant \delta^2 > 0 \text{ for all } t > 0.$$

Since $F_\omega(0,\lambda_0) = -\|x_0(\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} < -\varepsilon < 0$, it is clear that the game which starts from $x_0(\omega)$ can be completed if $x_0(\omega)$ satisfies relation (2.4.29). □

In Theorem 2.4.6 we gave a condition such that the stochastic linear pursuit-evasion game (2.4.1) will have a maximin completion time. Then, in Theorems 2.4.7 and 2.4.8 we gave sufficient conditions for completion of the game no matter what the starting state is. We now give an interactive procedure for determining the minimum completion time and the optimal controls.

2.4.3 *The Optimal Controls*

Assuming that the game (2.4.1) with initial condition $x(0,\omega) = x_0(\omega)$ can be completed, we can find the minimum completion time t_0 and the vector λ_{t_0} satisfying condition (2.4.20) as follows. Choose $\varepsilon > 0$.

1. Set $\lambda_1 = \dfrac{-x_0'(\omega)}{\|x_0(\omega)\|_{L_2(\Omega,\mathrm{A},\mu)(\Omega,\mathrm{A},\mu)}}$ and them compute $F_\omega(t,\lambda_1)$ for $t \geqslant 0$ up to the time t_1 such that $F(t_1,\lambda_1) = -\varepsilon$. Clearly $t_1 \leqslant t_0$.
2. Let $F_\omega(t_i,\lambda_i) = -\varepsilon$, $i = 1,2,\ldots$, and find $\min_{\lambda\in Q} F_\omega(t_i,\lambda)$ using the gradient method of Lemma 2.3.3. Call it $F_\omega(t_i,\lambda_{i+1})$. That is, $$\min_{\lambda\in Q} F_\omega(t_i,\lambda) = F_\omega(t_i,\lambda_{i+1}) \leqslant -\varepsilon.$$
3. Compute $F_\omega(t,\lambda_{i+1})$ for $t \geqslant t_i$ up to the time t_{i+1} such that $F_\omega(t_{i+1},\lambda_{i+1}) = -\varepsilon$. It is clear that $$F_\omega(t,\lambda_{i+1}) \geqslant F_\omega(t,\lambda_t) \text{ for all } t \in [0,t_i+1].$$
4. Repeat steps 2 and3 above for $i = 2,3,\ldots$.

Since $t_i \leqslant t_{i+1} \leqslant t_0$ for all i, $\lim_{i\to\infty} t_i$ exists. Let us denote it by $t_0^* \leqslant t_0$. We have $F_\omega(t_i,\lambda_i) = -\varepsilon$ for all $i = 1,2,\ldots$ and $\lim_{i\to\infty} t_i = t_0^* \leqslant t_0$. Since $F_\omega(t,\lambda_t)$ is continuous in t, we get

$$F_\omega\left(t_0^*,\lambda_{t_0^*}\right) = -\varepsilon.$$

But t_0 is the smallest nonnegative time satisfying line (2.4.20). Thus, $t_0^* = t_0$. Also, $\lambda_{i+1} = \lambda_{t_i} \to \lambda_{t_0}$ from the left. If λ_t is not continuous at t_0, let $\lambda_{t_0}^-$ denote the limit from below. That is, $\lim_{\delta+0} \lambda_{t_0-\delta} = \lambda_{t_0}^-$. Thus the optimal controls are $u(t) = u\left(t_0 - t, \lambda_{t_0}^-\right)$, $v(t) = v\left(t_0 - t, \lambda_{t_0}^-\right)$ for all $t \in [0,t_0]$.

With the iterative procedure described above one can program the game for an electronic computer. It is first necessary to check if the game can indeed be completed. For this it is an easy matter to program the Corollary 2.4.2. That is, we must first check to see if there exists a finite time $t \in \mathbb{R}_+$ such that

$$\inf_{\lambda\in Q}\left[\lambda\Phi(t;\omega)x_0(\omega) + \int_0^{+} H_U(\lambda K(\tau;\omega))d\tau - \int_0^t H_V(\lambda L(\tau;\omega))d\tau\right] \geqslant -\varepsilon$$

where Q is the set of all $(1\times n)$ vectors λ such that $\|\lambda\|_{L_2(\Omega,\mathrm{A},\mu)} = 1$.

In this section we have considered stochastic linear differential games of the form

$$\frac{d}{dt}x(t;\omega) = A(\omega)x(t;\omega) + BU(t) - Cv(t), \quad t \geqslant 0$$

which is a special case of equation (2.1.1)? Here we have taken constant matrices B and C and control sets $U(t)$ and $V(t)$ which are compact subsets of Euclidean spaces. The method of investigation was to first reduce the problem to the existence of a random solution to the stochastic vector integral equation

$$x(t;\omega) = \Phi(t;\omega)x_0(\omega) + \int_0^t \Phi(t-\tau;\omega)[Bu(\tau) - Cv(\tau)]d\tau$$

where $\Phi(t;\omega) = e^{A(\omega)t}$.

We then proved several theorems on completion of the game. Theorem 2.4.1 and the Corollary 2.4.2 give necessary and sufficient conditions for the existence of a control for the pursuer so that he can force completion of the game in a finite time. No matter what controls that two players choose, Theorems 2.4.6 gives a condition sufficient to guarantee the completion of the game and also gives the minimum completion time. Theorem 2.4.7 gives conditions on the control sets, which are independent of the initial condition, which guarantee completion of the game; while Theorem 2.4.8 gives conditions on the control sets and the initial condition which force completion of the game.

Finally we presented an iterative procedure which can be used to find the minimum completion time mentioned in Theorem 2.4.6 and to find the optimal controls to force completion in this time.

Chapter 3

Two Person Zero-Sum Differential Games-General Case

3.1 Introduction

The object of this Chapter is to present the concept of strategies and solutions as well as existence and uniqueness results for the two person zero-sum stochastic differential games. First, we will discuss some definitions and a brief survey of earlier works. Then, we will present the earlier work on stochastic differential games using martingale methods. Almost all of the material on this subsection comes from Elliott [47]. In the next subsection, we will briefly mention the recent results obtained on two person zero-sum stochastic differential games using the concept of viscosity solutions, Souganidis [181]. There are various other methods used in studying stochastic differential games. In Bensoussan and Lions [22], two player stochastic differential games with stopping is analyzed using the method of two sided variational inequalities. Also refer to Bensoussan and Friedman [23, 24] for more results in this direction. Also, a zero-sum Markov games with stopping and impulsive strategies is discussed in Stetner [187].

3.2 Two Person Zero-sum Games: Martingale methods

The evolution of the system is described by a stochastic differential equations

$$dx(t) = b(t,x,u_1,u_2)dt + \sigma(t,x)dB(t) \tag{3.2.1}$$

with

$$x(0) = x_0 \in \mathbb{R}^n, \quad t \in [0,1],$$

where B is an n-dimensional Brownian motion; $u_i \in \mathscr{U}_i$, $i = 1,2$ are control functions. There are two controllers, or players, I and II. The game is zero sum, if player I is choosing his control to maximize the payoff and player II is choosing his control to minimize the payoff. Let $\Im_t = \sigma\{x(s) : s \leqslant t\}$ be the σ-algebra generated on $\mathscr{C}$, the space of continuous functions from $[0,1] \to \mathbb{R}^n$, up to time t. Assume that $b : [0,1] \times \mathscr{C} \times \mathscr{U}_1 \times \mathscr{U}_2 \to \mathbb{R}^n$ and σ, a nonsingular $n \times n$ matrix, satisfy the usual measurability and growth conditions. Given an n-dimensional Brownian motion $B(t)$ on a probability space (Ω, P), these conditions on σ ensures the stochastic equation

$$x(t) = x_0 + \int_0^t \sigma(s,x) dB(t),$$

has unique solution with sample path in $\mathscr{C}$. Let $\Im_t = \sigma\{B(s) : s \leqslant t\}$.

Assume that the spaces $\mathscr{U}_1$ and $\mathscr{U}_2$ are compact metric spaces and suppose that b is continuous in variables $u_1 \in \mathscr{U}_1$ and $u_2 \in \mathscr{U}_2$. The admissible feedback controls $\mathscr{A}_{1s}^t$ for the player I, over $[s,t] \subset 0,1]$, are measurable functions $u_1 : [s,t] \times \mathscr{C} \to \mathscr{U}_1$ such that for each τ, $s \leqslant \tau \leqslant t$, $u_1(\tau, \cdot)$ is $\mathscr{F}_t$-measurable and for each $x \in \mathscr{C}$, and $u_1(\cdot, x)$ is Lebesgue measurable. The admissible feedback controls $\mathscr{A}_{2s}^t$ for the player II, over $[s,t] \subset [0,1]$, are measurable functions $u_2 : [s,t] \times \mathscr{C} \to \mathscr{U}_2$ with similar properties. Let $\mathscr{A}_i = \mathscr{A}_{i0}^1, i = 1,2$. For $u_i \in \mathscr{A}_{is}^t$, $i = 1,2$, write

$$b^{u_1,u_2}(\tau,x) = b\left(\tau, x, u_1(\tau,x), u_2(\tau,x)\right).$$

Then conditions on b ensure that

$$E\left[\exp \xi_s^t\left(b^{u_1,u_2}\right) \mid \mathbb{F}_s\right] = 1 \text{ a.s. } P,$$

where

$$\xi_s^t\left(f^{u_1,u_2}\right) = \int_s^t \{\sigma^{-1}(\tau,x) b^{u_1,u_2}(\tau,x)\}' dB(\tau) - 1/2 \int_s^t \left|\sigma^{-1}(\tau,x) b^{u_1,u_2}(\tau,x)\right|^2 d\tau.$$

For each $u_i \in \mathscr{A}_i$ a probability measure P_{u_1,u_2} is defined through

$$\frac{dP_{u_1,u_2}}{dP} = \exp \xi_0^1\left(b^{u_1,u_2}\right).$$

Then by the Girsanov's Theorem, we have the following result.

Theorem 3.2.1. *Under the measure P_{u_1,u_2} the process $w^{u_1,u_2}(t)$ is a Brownian motion on Ω, where*

$$dw^{u_1,u_2}(t) = \sigma^{-1}(t,x)\left(dx(t) - b^{u_1,u_2}(t,x)dt\right).$$

Corresponding to controls $u_i \in \mathscr{A}_i$, $i = 1,2$ the expected total cost is

$$J(u_1,u_2) = E_{u_1,u_2}\left[g(x(1)) + \int_0^1 h^{u_1,u_2}(t,x)dt\right] \tag{3.2.2}$$

where h and g are real valued and bounded, $g(x(1))$ is $\mathscr{F}_1$ measurable and h satisfies the same conditions as the components of b. Also E_{u_1,u_2} denotes the expectation with respect to P_{u_1,u_2}. For a zero sum differential game, player I wishes to choose u_1 so that $J(u_1,u_2)$ is maximized and player II wishes to choose u_2 so that $J(u_1,u_2)$ is minimized.

Now the principle of optimality will be derived. Suppose that player II uses the control $u_2(t,x) \in \mathscr{A}_2$ through out the game. Then if player I uses the control $u_1(t,x) \in \mathscr{A}_1$, the cost incurred from time t onwards, given $\mathscr{F}_t$ is independent of the controls used up to time t and is given by

$$\psi_t^{u_1,u_2} = E_{u_1 u_2}\left[g(x(1)) + \int_t^1 h^{u_1,u_2}(s,x)ds|\mathscr{F}_t\right].$$

Because $L^1(\omega)$ is a complete lattice, the spremium

$$W_t^{u_2} = \bigvee_{u_1 \in \mathbb{A}_1} \psi_t^{u_1,u_2}, \tag{3.2.3}$$

exists, and represents the best that player I can attain from t onwards, given that player II is using control u_2. Let $u_1(u_2)$ represent the response of player I to the control u_2 used by player II. Then we have

Theorem 3.2.2.

(a) *$u_1^*(u_2)$ is the optimal reply to u_2 iff*

$$W_t^{u_2} + \int_0^t h^{u_1^*,u_2}(s)ds,$$

is a martingale on $\left(\Omega, \Im_t, P_{u_1^*(u_2),u_2}\right)$.

(b) *In general, for $u_1 \in \mathscr{A}_1$,*

$$W_t^{u_2} + \int_0^t h^{u_1,u_2}(s)ds$$

is a super martingale on $(\Omega, \Im_t, P_{u_1,u_2})$.

From martingale representation results, one can see that u_1^* is optimal reply for player I iff there is a predictable process $g_t^{u_2}$, such that,

$$\int_0^1 |g_s^{u_2}|^2 ds < \infty \ \text{ a.s.},$$

and

$$W_t^{u_2} + \int_0^t h^{u_1^*,u_2}(s)ds = W_0^{u_2} + \int_0^t g_s^{u_2} dw_s^{u_1^*(u_2),u_2}.$$

For any other $u_1 \in \mathscr{A}_1$ the supermartingale $W_t^{u_2} + \int_0^t h^{u_1,u_2}(s)ds$ has a unique Doob-Meyer decomposition as

$$W_0^{u_2} + M_t^{u_1,u_2} + A_t^{u_1,u_2}, \tag{3.2.4}$$

where $M_t^{u_1,u_2}$ is a martingale on $(\Omega, \Im_t, P_{u_1,u_2})$ and $A_t^{u_1,u_2}$ is a predictable decreasing process. From the representation (3.2.4),

$$W_t^{u_2} + \int_0^t h^{u_1^*,u_2}(s)ds = W_0^{u_2} + \int_0^t g^{u_2}\sigma^{-1}\left(dx_s - b_s^{u_1,u_2}ds\right)$$
$$- \int_0^t \left[\left(g^{u_2}\sigma^{-1}b_s^{u_1^*(u_2),u_2} + h_s^{u_1^*(u_2),u_2}\right) - \left(g^{u_2}\sigma^{-1}b_s^{u_1,u_2} + h_s^{u_1^*(u_2),u_2}\right)\right] ds.$$

Again from Theorem 3.2.1, $dw_s^{u_1,u_2} = \sigma^{-1}\left(dx_s - b_s^{u_1,u_2}ds\right)$ is a Brownian motion on $(\Omega, \Im_t, P_{u_1,u_2})$ and hence the stochastic integral is a predictable process, so by uniqueness of the Doob-Meyer decomposition

$$M_t^{u_1,u_2} = \int_0^t g^{u_2} dw^{u_1,u_2},$$
$$A_t^{u_1,u_2} = \int_0^t \left[\left(g^{u_2}\sigma^{-1}b_s^{u_1^*(u_2),u_2} + h_s^{u_1^*(u_2),u_2}\right) - \left(g^{u_2}\sigma^{-1}b_s^{u_1,u_2} + h_s^{u_1^*(u_2),u_2}\right)\right] ds.$$

Since $A_t^{u_1,u_2}$ is decreasing one can obtain the following principle of optimality.

Theorem 3.2.3. *If $u_1^*(u_2)$ is the best reply for player I then, almost surely,*

$$g^{u_2}\sigma^{-1}b_s^{u_1^*(u_2),u_2} + h_s^{u_1^*(u_2),u_2} \geqslant g^{u_2}\sigma^{-1}b_s^{u_1,u_2} + h_s^{u_1^*(u_2),u_2}. \tag{3.2.5}$$

That is, if the optimal reply for player I exists, it is obtained by maximizing the Hamiltonian

$$g^{u_2}\sigma^{-1}b_s^{u_1,u_2} + h_s^{u_1,u_2}. \tag{3.2.6}$$

We will establish existence of optimal control $u_1^*(u_2) \in \mathscr{A}_1$ for player I in reply to any control $u_2 \in \mathscr{A}_2$ used by player II. Now we will make the payoff (3.2.1) into a completely terminal payoff by introducing a new state variable x_{n+1} and a new Brownian motion B_{n+1} on a probability space (Ω', P'). Suppose x_{n+1} satisfies the equation

$$dx_{n+1} = h(t, x, u_1, u_2)dt + dB_{n+1},$$
$$x_{n+1}(0) = 0.$$

The $(n+1)$-dimensional process (x, x_{n+1}) is defined on the product space $(\Omega^+, P^+) = (\Omega \times \Omega', P \times P')$. If we write

$$x^+ = (x, x_{n+1}), \quad b^+ = (b, h), \quad \sigma^+ = \begin{bmatrix} \sigma & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } w_{n+1} = B_{n+1},$$

then $w^+ = (w, w_{n+1})$ is an $(n+1)$-dimensional Brownian motion on Ω^+.
Define a new probability measure P_{u_1,u_2}^+ on Ω^+ by putting

$$\frac{dP_{u_1,u_2}^+}{dP} = \exp \xi_0^1\left(f_{u_1,u_2}^+\right).$$

Let $E^+_{u_1,u_2}$ denote the expectation with respect to $P^+_{u_1,u_2}$. Since w_{n+1} is a Brownian motion and h and g are independent of x_{n+1}, the expected payoff corresponding to the controls u_1 and u_2 is given by,

$$E^+_{u_1,u_2}\left[g\left(x(1)\right)+x_{n+1}(1)\right]=E_{u_1,u_2}\left[g\left(x(1)\right)+\int_0^1 h\left(s,x,u_1,u_2\right)ds\right].$$

Define

$$W^+_{u_2}(t)=\bigvee_{u_1\in\mathscr{U}_1} E^+_{u_1,u_2}[g\left(x(1)\right)+x_{n+1}(1)|\Im^+_t],$$

the suprimum being in $L^1\left(\Omega^+\right)$. Let C^+ denote the $\mathbb{R}^{n+1}$ valued continuous function on $[0,1]$ and $\Im^+_t$ the $\sigma-$field on C^+ generated up to time t. Let $\Phi^+=\{\phi:[0,1]\times C^+\to\mathbb{R}^{n+1}\}$ which satisfy

(i) for each $t\in[0,1]$, $\phi\left(t,\cdot\right)$ is $\Im^+_t$ measurable,

(ii) for each $x\in C^+$, $\phi\left(\cdot,x\right)$ is Lebesgue measurable,

and

(iii) $|(\sigma^+)^{-1}(t,x)\phi(t,x)|\leqslant M(1+\|x\|_t)$ where $\|x\|_t=\sup_{0\leqslant s\leqslant t}|x(s)|$.

Write $\mathscr{D}=\{\exp\xi^1_0\left(\phi\right):\phi\in\Phi^+\}$. Because ϕ has linear growth $E^+\exp\xi^1_0\left(\phi\right)=1$ for all $\phi\in\Phi^+$, where E^+ denotes the expectation with respect to P^+. Since $\mathscr{D}$ is weakly compact, we have the following result.

Theorem 3.2.4. *There is a function $H\in\Phi^+$, such that $(W^+_{u_2}(t),\Im^+_t,P^*)$ is a martingale. Here P^* is defined on Ω^+ by*

$$\frac{dP^*}{dP^+}=\exp\xi^1_0(H). \tag{3.2.7}$$

If there is an optimal reply $u_1^*(u_2)$ for player I, take $H=f^+_{u_1^*(u_2),u_2}$.

This result states that, even if there is not an optimal control, there is always a 'drift term' $H\in\Phi^+$ whose corresponding measure gives the maximum value function, that is,

$$\begin{aligned}W^+_{u_2}(t)&=\bigvee_{u_1\in\mathbb{U}_1} E^+_{u_1,u_2}\left[g(x(1))+x_{n+1}(1)\mid\Im^+_t\right]\\&=E^*\left[g(x(1))+x_{n+1}(1)\mid\Im^+_t\right]\end{aligned}$$

where E^* denotes expectation with respect to P^*.

Under P^*, using Girsanov's theorem, we are considering an $n+1$-dimensional Brownian motion w^* on (Ω^+,P^*) defined by

$$\begin{pmatrix}dw^*\\dw^*_{n+1}\end{pmatrix}=\begin{pmatrix}\sigma^{-1}&0\\0&1\end{pmatrix}\begin{pmatrix}dx-\widehat{H}dt\\dx_{n+1}-H_{n+1}dt\end{pmatrix}.$$

where $\widehat{H}$ denotes the first n coordinates of H.

Since $h(t,x,u_1(t,x),u_2(t,x))$ is independent of x_{n+1}, for any controls, the weak limit H_{n+1} is independent of x_{n+1}. So for any control $u_1 \in U_1$:

$$\begin{aligned}&E^{+}_{u_1,u_2}[g(x(1)) + \int_0^1 h(s,x,u_1,u_2)\,ds + w_{n+1}(1) - w_{n+1}(t)|\Im^{+}_t] + x_{n+1}(t)\\ &= E_{u_1,u_2}[g(x(1)) + \int_t^1 h(s,x,u_1,u_2)\,ds + w_{n+1}(1) - w_{n+1}(t)|\Im_t] + x_{n+1}(t).\end{aligned}$$

Taking suprimum to obtain $W^{+}_{u_2}$, we see

$$W^{+}_{u_2}(t) = W^{u_2}_t + \int_0^t H_{n+1}(s)ds + w^{*}_{n+1}(t).$$

Therefore,

$$W^{u_2}_t + \int_0^t H_{n+1}(s)ds + w^{*}_{n+1}(t) = E^{*}[g(x(1)) + x_{n+1}(1)|\Im^{+}_t].$$

Taking the expectation with respect to $\Im_t \subset \Im^{+}_t$ we have

$$W^{u_2}_t + \int_0^t H_{n+1}(s)ds = E^{*}[g(x(1)) + x_{n+1}(1)|\Im_t].$$

Hence, $W^{u_2}_t + \int_0^t H_{n+1}(s)ds$ is a martingale on $(\Omega, \Im_t, P^{*})$, and thus it can be represented as a stochastic integral, that is, $B^{u_2} + \int_0^t g^{*}dw^{*}$, with respect to $n-$dimensional Brownian motion w^{*} defined on $(\Omega, \Im_t, P^{*})$ by

$$dw^{*} = \sigma^{-1}dx - \sigma^{-1}Hdt.$$

Here, $B^{u_2} = W^{u_2}_0$ and g^{*} is a predictable process. Under any other control $u_1 \in U_1$, as in Theorem 3.2.2, $W^{u_2}_t + \int_0^t h^{u_1,u_2}_s ds$ is a supermartingale and hence

$$\begin{aligned}&W^{u_2}_t + \int_0^t h^{u_1,u_2}_s ds\\ &\quad = B^{u_2} + \int_0^t g^{*}dw^{u_1,u_2}_s + \int_0^t \left(g^{*}\sigma^{-1}b^{u_1,u_2}_s + h^{u_1,u_2}_s\right) - \left(g^{*}\sigma^{-1}\widehat{H}_s + H_{n+1}(s)\right)ds.\end{aligned} \tag{3.2.8}$$

Since w^{u_1,u_2}_s is a Brownian motion on (Ω, P_{u_1,u_2}) defined by

$$dw^{u_1,u_2}_s = \sigma^{-1}\left(dx_s - b^{u_1,u_2}_s ds\right),$$

the first integral on the right hand side of (3.2.8) is a stochastic integral and the second a decreasing process. Hence we have almost surely

$$g^{*}\sigma^{-1}\widehat{H} + H_{n+1} \geqslant g^{*}\sigma^{-1}b^{u_1,u_2} + h^{u_1,u_2}. \tag{3.2.9}$$

If there is a process $u_1^*(u_2)$ such that, almost surely,

$$g^*\sigma^{-1}\widehat{H}+H_{n+1}=g^*\sigma^{-1}f^{u_1^*,u_2}+h^{u_1^*,u_2}$$

then

$$W_t^{u_2}+\int_0^t h_s^{u_1^*,u_2}ds=B^{u_2}+\int_0^t g^*dw^+_{u_1^*(u_2),u_2}$$

and then, it is a martingale. Therefore, $u_1^*(u_2)$ would be an optimal reply to u_2.
For the above process g^*, since b and h are continuous in the control variables u_1 and u_2 and the control spaces are compact, there is a measurable feedback control $u_1^*(u_2)$ such that almost surely

$$g^*.\sigma^{-1}b^{u_1^*(u_2),u_2}+h^{u_1^*(u_2),u_2}\geqslant g^*.\sigma^{-1}b^{u_1,u_2}+h^{u_1,u_2}. \tag{3.2.10}$$

We will now show that such a control $u_1^*(u_2)$ is an optimal reply for Player I.
Let

$$\Gamma_s(u_1,u_2)=g^*.\sigma^{-1}b_s^{u_1,u_2}+h_s^{u_1,u_2}$$

and

$$\widehat{\Gamma}_s=g^*.\sigma^{-1}\widehat{H}_s+H_{m+1}(s),$$

and let $u_1^*(u_2)$ is selected as in (3.2.10) so that $\Gamma_s\left(u_1^*,u_2\right)\geqslant\Gamma_s(u_1,u_2)$. Then

$$W_t^{u_2}+\int_0^t h_s^{u_1,u_2}ds=B^{u_2}+\int_0^t g^*dw^+_{u_1,u_2}+\int_0^t\left(\Gamma_s(u_1,u_2)-\widehat{\Gamma}_s\right)ds.$$

Taking the expectations with respect to $\mu^+_{u_1,u_2}$ at $t=1$ we have

$$\begin{aligned}E^+_{u_1,u_2}\left[g(x(1))+\int_0^1 h_s^{u_1,u_2}ds\right]&=B^{u_2}+E^+_{u_1,u_2}\left[\int_0^1\left(\Gamma_s(u_1,u_2)-\widehat{\Gamma}_s\right)ds\right]\\&\leqslant B^{u_2}+E^+_{u_1,u_2}\left[\int_0^1\left(\Gamma_s\left(u_1^*(u_2),u_2\right)-\widehat{\Gamma}_s\right)ds\right].\end{aligned} \tag{3.2.11}$$

The left hand side of the inequality (3.2.11) is just $\psi_0^{u_1,u_2}$, so for any $n\in\mathbb{Z}^+$ there is a control $u_{1n}\in U_1$, such that,

$$-E^+_{u_{1n},u_2}\left[\int_0^1\left(\Gamma_s\left(u_1^*(u_2),u_2\right)-\widehat{\Gamma}_s\right)ds\right]<1/n.$$

Also, let

$$-X=\int_0^1\left(\Gamma_s\left(u_1^*(u_2),u_2\right)-\widehat{\Gamma}_s\right)ds.$$

Then the inequality (3.2.10) implies X is positive almost surely, and $E^+\phi_nX\to 0$, where $\phi_n=\exp\xi_0^1\left(f^+_{u_{1n},u_2}\right)$. Let $X^N=\min(N,X)$ for $N\in\mathbb{Z}^+$, so $0\leqslant X^N\leqslant X$ and $E^+\phi_NX^N\to 0$. By weak compactness of $\mathscr{D}$ there is a $\phi\in\mathscr{D}$ such that the ϕ_n converge to ϕ weakly, so

$$\lim_{n\to\infty}E^+\phi_nX^N=E^+\phi X^N=0.$$

Since $\phi > 0$ a.s., we have $X^N = 0$ a.s.. Therefore $X = 0$ a.s., and hence

$$\Gamma_s\left(u_1^*(u_2), u_2\right) = \widehat{\Gamma}_s \text{ a.s..}$$

Therefore, we conclude that an optimal reply $u_1^*(u_2)$ exists for player I in reply to any control $u_2 \in U_2$ used by player II.

We will now establish the existence, and obtain a characterization, of the optimal feedback control that player II should use if he chooses his control first. Assume that the player I will always play his best reply $u_1^*(u_2) \in U_1$ in response to any control $u_2 \in U_2$. Now the problem is how player II, who is trying to minimize the payoff (3.2.1), should choose a $u_2^* \in U_2$ such that

$$\inf_{u_2 \in U_2} \sup_{u_1 \in U_1} J(u_1, u_2) = \inf_{u_2 \in U_2} J\left(u_1^*(u_2), u_2\right).$$

For any $u_2 \in U_2$ and $t \in 0, 1]$, if player I plays $u_1^*(u_2)$, the expected terminal payoff is

$$\psi_{u_2}(t) = E_{u_1^*(u_2), u_2}\left[g(x(1) + \int_0^1 h^{u_1^*(u_2), u_2} ds | \mathfrak{I}_t\right].$$

Since $L^1(\omega)$ is a complete lattice the infimum (denoted by $\wedge$),

$$V_t^+ = \bigwedge_{u_2 \in U_2} \psi_{u_2}(t) \tag{3.2.12}$$

exists in $L^1(\omega)$. V_t^+ in (3.2.12) is called the *upper value function* of the differential game, and

$$V_0^+ = \inf_{u_2 \in U_2} \sup_{u_1 \in U_1} J(u_1, u_2)$$

is the *upper value* of the game. One can obtain the following result, Elliott [47].

Theorem 3.2.5.

(a) *$u_2^* \in U_2$ is optimal for player II if and only if*

$$V_t^+ + \int_0^t h^{u_1^*\left(u_2^*\right), u_2^*} ds$$

is a martingale on $\left(\Omega, \mathscr{A}_t, P_{u_1^*\left(u_2^*\right), u_2^*}\right)$.

(b) *In general, for $u_2 \in U_2$,*

$$V_t^+ + \int_0^t h^{u_1^*(u_2), u_2} ds$$

is a submartingale on $\left(\Omega, \mathscr{A}_t, P_{u_1^*(u_2), u_2}\right)$.

From the above martingale representation, $u_2^* \in U_2$ is optimal for player II playing first if and only if there is a predictable process g_t^* such that

$$\int_0^1 |g_s^*|^2 ds < \infty \text{ a.s.}$$

and

$$V_t^+ + \int_0^t h^{u_1^*(u_2^*),u_2^*} ds = B^* + \int_0^t g^* dw_s^*.$$

Here the w^* is the Brownian motion given by

$$dw^* = \sigma^{-1}\left(dx - b^{u_1^*(u_2^*),u_2^*} ds\right),$$

on $\left(\Omega, P_{u_1^*(u_2^*),u_2^*}\right)$. For a general $u_2 \in U_2$ the submartingale $V_t^+ + \int_0^t h^{u_1^*(u_2),u_2} ds$ has a unique Doob-Mayer decomposition $B^* + M_t^{u_2} + A_t^{u_2}$, where $M_t^{u_2}$ is a martingale on $\left(\Omega, P_{u_1^*(u_2),u_2}\right)$ and $A_t^{u_2}$ is a predictable increasing process. Also, if $u_2^* \in U_2$ is optimal for player II playing first, then almost surely

$$g^*.\sigma^{-1} b_s^{u_1^*(u_2^*),u_2^*} + h_s^{u_1^*(u_2^*),u_2^*} \leqslant g^*.\sigma^{-1} b_s^{u_1^*(u_2),u_2} + h_s^{u_1^*(u_2),u_2}.$$

Conversely, without a priori assuming there is an optimal control $u_2^* \in U_2$, one can obtain an integral representation for V_t^+, and show that the measurable strategy, obtained by minimizing a Hamiltonian $g^*.\sigma^{-1} b_s^{u_1^*(u_2),u_2} + h_s^{u_1^*(u_2),u_2}$, exists and is optimal. This leads to the following result.

Theorem 3.2.6. *There is a predictable process g^* and $u_2^* \in U_2$ is optimal if and only if u_2^* minimizes the Hamiltonian*

$$\Gamma_s\left(u_1^*(u_2), u_2\right) = g^*.\sigma^{-1} b_s^{u_1^*(u_2),u_2} + h_s^{u_1^*(u_2),u_2}, \ \textit{a.s. in } (s,\omega).$$

3.2.1 *The Isaacs condition*

We have seen that,

$$V_0^+ = \inf_{u_2 \in U_2} \sup_{u_1 \in U_1} J(u_1, u_2)$$

represents the best outcome that players I and II can ensure if player II chooses his feedback control first. Now, we will define the *lower value* of the game,

$$V_0^- = \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} J(u_1, u_2).$$

For $t \in 0,1]$, $x \in C$, $u_1 \in U_1$, $u_2 \in U_2$ and $p \in \mathbb{R}^n$ we can write

$$L(t,x,p;u_1,u_2) = p \cdot \sigma^{-1}(t,x) b(t,x,u_1,u_2) + h(t,x,u_1,u_2).$$

The game is said to satisfy the *Isaacs condition* if, for all such t, x, p,

$$\min_{u_2\in U_2}\max_{u_1\in U_1} L(t,x,p;u_1,u_2)=\max_{u_1\in U_1}\min_{u_2\in U_2} L(t,x,p;u_1,u_2). \tag{3.2.13}$$

We say the game satisfies a *saddle-point condition* if the upper and lower values of an 'infinitesimal' game are equal, then $V_0^+=V_0^-$. The result that follows states that the game has a value under Isaacs condition.

Theorem 3.2.7. *If the game satisfies the Isaacs condition then* $V_0^+=V_0^-$.

Proof. Note that for $u_i\in U_i$, $i=1,2$

$$\Gamma_s(u_1,u_2)=L\left(s,x,g^*;u_1(t,x),u_2(t,x)\right),$$

where g^* is the predictable process introduced earlier. Also, for any $u_2\in U_2$, we proved that there exists a strategy $u_1^*(u_2)\in U_1$, such that,

$$\Gamma_s\left(u_1^*(u_2),u_2\right)=\max_{u_1\in U_1}\Gamma_s\left(u_1^*(u_2),u_2\right),$$

and then that there is a $u_2^*\in U_2$, such that,

$$\begin{aligned}\Gamma_s\left(u_1^*\left(u_2^*\right),u_2^*\right)&=\min_{u_2\in U_2}\Gamma_s\left(u_1^*(u_2),u_2\right)\ \text{ a.s.}\\&=\min_{u_2\in U_2}\max_{u_1\in U_1}\Gamma_s\left(u_1,u_2\right)\ \text{ a.s..}\end{aligned}$$

We also had a representation of the form

$$V_t^+ +\int_0^t h^{u_1^*(u_2^*),u_2^*}ds=B^*+\int_0^t g^*dw_s^*\ \text{ a.s.}$$

Because f and u_1 are continuous in u_1 and u_2 and U_1 and U_2 are compact, for any $u_1\in U_1$ there exists a strategy $u_2^*(u_1)\in U_2$ such that

$$\Gamma_s\left(u_1,u_2^*(u_1)\right)=\min_{u_2\in U_2}\Gamma_s(u_1,u_2)\ \text{ a.s..}$$

Similarly, there is a $u_1^*\in U_1$, such that,

$$\begin{aligned}\Gamma_s\left(u_1^*,u_2^*\left(u_1^*\right)\right)&=\max_{u_1\in U_1}\Gamma_s\left(u_1,u_2^*(u_1)\right)\ \text{ a.s.}\\&=\max_{u_1\in U_1}\min_{u_2\in U_2}\Gamma_s(u_1,u_2)\ \text{ a.s..}\end{aligned}$$

Since the Isaacs condition (3.2.13) holds, we have

$$\Gamma_s\left(u_1^*,u_2^*\left(u_1^*\right)\right)=\Gamma_s\left(u_1^*\left(u_2^*\right),u_2^*\right)\ \text{ a.s..}$$

Now, for any $u_2\in U_2$, we have

$$\Gamma_s\left(u_1^*,u_2^*\left(u_1^*\right)\right)\leqslant\Gamma_s\left(u_1^*,u_2\right)\ \text{ a.s..}$$

and for any $u_1 \in U_1$, we have

$$\Gamma_s\left(u_1, u_2^*\right) \leqslant \Gamma_s\left(u_1^*\left(u_2^*\right), u_2^*\right) \text{ a.s..}$$

Hence,

$$\Gamma_s\left(u_1, u_2^*\right) \leqslant \Gamma_s\left(u_1^*, u_2^*\right) \leqslant \Gamma_s\left(u_1^*, u_2\right) \text{ a.s..}$$

Therefore,

$$V_t^+ + \int_0^t h^{u_1^*, u_2^*} ds = B^* + \int_0^t g^* dw_s^{u_1^*, u_2^*} \text{ a.s.,}$$

where

$$dw_s^{u_1^*, u_2^*} = \sigma^{-1}\left(dx_s - b_s^{u_1^*, u_2^*} ds\right)$$

is a Brownian motion under $P_{u_1^*, u_2^*}$. For any other $u_1 \in U_1$, we can write

$$V_t^+ + \int_0^t h^{u_1, u_2^*} ds = B^* + \int_0^t g^* dw_s^{u_1, u_2^*} + \int_0^t \left(\Gamma_s\left(u_1, u_2^*\right) - \Gamma_s\left(u_1^*, u_2^*\right)\right) ds.$$

Taking the expectations at $t = 1$ with respect to P_{u_1, u_2^*}, results in,

$$E_{u_1, u_2^*}\left[g(x(1)) + \int_0^1 h_s^{u_1, u_2^*} ds\right] = J\left(u_1, u_2^*\right) \leqslant J^* = J\left(u_1^*, u_2^*\right).$$

Similarly, one can show that

$$J\left(u_1^*, u_2^*\right) \leqslant J\left(u_1^*, u_2\right).$$

Therefore, if Isaacs condition is satisfied

$$\sup_{u_1 \in U_1} \inf_{u_2 \in U_2} J(u_1, u_2) = \inf_{u_2 \in U_2} \sup_{u_1 \in U_1} J(u_1, u_2) = J^*,$$

hence the upper and lower value of the differential game are equal. Furthermore, we can also show that if the upper and lower values are equal then

$$\max_{u_1 \in U_1} \min_{u_2 \in U_2} L\left(t, x, g^*; u_1, u_2\right) = \min_{u_2 \in U_2} \max_{u_1 \in U_1} L\left(t, x, g^*; u_1, u_2\right) \text{ a.s..} \qquad \square$$

In this section, using the martingale methods we have proved the existence of a value for the game under the Isaacs condition as well as characterized the optimal strategies.

3.3 Two Person Zero-sum Games and Viscosity Solutions

The viscosity solution concept was introduced in the early 1980s by Michael Crandall and Pierre-Louis Lions, [43] as a generalization of the classical concept of a 'solution' to a partial differential equation (PDE). It has been found that the viscosity solution is the natural solution concept to use in differential games (the Isaacs equation) and in stochastic differential games.In this subsection, we present briefly some key elements of the viscosity solutions method for the theory of two person zero-sum stochastic differential games. For more details we refer to Fleming and Souganidis [64, 65]. For $s \in (t,T]$, consider the dynamics

$$dx_s = b(x_s, s, u_{1s}, u_{2s})\,ds + \sigma(x_s, s, u_{1s}, u_{2s})\,dw_s \tag{3.3.1}$$

with initial condition

$$x_t = x \quad (x \in \mathbb{R}^n),$$

where w is a standard m-dimensional Brownian motion. The payoff is given by

$$J_{x,t}(u_1,u_2) = E_{x,t}\left\{\int_t^T h(x_s, s, u_{1s}, u_{2s})ds + g(x_T)\right\}. \tag{3.3.2}$$

Here u_1 and u_2 are stochastic processes taking values in the given compact sets $U_1 \subset \mathbb{R}^k$ and $U_2 \subset \mathbb{R}^l$.

Assume that $b : \mathbb{R}^n \times (0,T] \times U_1 \times U_2 \to \mathbb{R}^n$ is uniformly continuous and satisfies, for some constant C_1 and all $t, \widehat{t} \in (0,T]$, $x, \widehat{x} \in \mathbb{R}^n$, $u_i \in U_i$, $i = 1,2$,

$$\begin{cases} |b(x,t,u_1,u_2)| \leqslant C_1, \\ \left|b(x,t,u_1,u_2) - b(\widehat{x},\widehat{t},u_1,u_2)\right| \leqslant C_1\left(|x-\widehat{x}| + |t-\widehat{t}|\right). \end{cases}$$

Also, let $h : \mathbb{R}^n \times (0,T] \times U_1 \times U_2 \to \mathbb{R}$ is uniformly continuous and satisfies, for some constant C_2,

$$\begin{cases} |h(x,t,u_1,u_2)| \leqslant C_2, \\ \left|h(x,t,u_1,u_2) - h(\widehat{x},\widehat{t},u_1,u_2)\right| \leqslant C_2\left(|x-\widehat{x}| + |t-\widehat{t}|\right), \end{cases}$$

and $g : \mathbb{R}^n \to \mathbb{R}^n$ satisfies

$$\begin{cases} |g(x)| \leqslant C_3, \\ |g(x) - g(\widehat{x})| \leqslant C_3\left(|x-\widehat{x}|\right). \end{cases}$$

Also the $n \times m$ matrix σ is bounded uniformly continuous and Lipschitz continuous with respect to x.

On a probability space $(\Omega, \mathfrak{I}, P)$, set

$$U_i(t) \equiv \{u_i : [t,T] \to U_i \ \text{ measurable}\}, \ \ i = 1,2.$$

These are the sets of all controls for players I and II. We consider the controls that agree a.e. are the same.
Define any mapping

$$\alpha : U_2(t) \to U_1(t)$$

to be a *strategy* for I (beginning at time t) provided for each $s \in [t,T]$ and u_2, $\widehat{u}_2 \in U_2(t)$

$$\text{if } \ u_2 = \widehat{u}_2 \ \text{ a.e. in } \ [t,s], \text{then } \ \alpha[u_2] = \alpha[\widehat{u}_2] \ \text{ a.s. in } \ [t,s].$$

Similarly, a mapping

$$\beta : U_1(t) \to U_2(t)$$

is a *strategy* for player II provided for each $s \in [t,T]$ and u_1, $\widehat{u}_1 \in U_1(t)$

$$\text{if } \ u_1 = \widehat{u}_1 \ \text{ a.e. in } \ [t,s], \text{then } \ \beta[u_1] = \beta[\widehat{u}_1] \ \text{ a.e. in } \ [t,s].$$

Denote by $\Gamma_i(t)$, $i = 1,2$, the set of all strategies for players I and II respectively, beginning at time t. At this point we note that there is some serious measurability problems that need to be addressed in the characterization of strategies for stochastic games. For a detailed account on the concept of measurability in the stochastic case and how to overcome this difficulty, we refer to Fleming [64]. Define the lower and upper values V and U by

$$V(x,t) = \inf_{\beta \in \Gamma_2(t)} \sup_{u_1 \in U_1(t)} J_{x,t}(u_1, \beta u_1])$$

and

$$U(t,x) = \sup_{\alpha \in \Gamma_1(t)} \inf_{u_2 \in U_2(t)} J_{x,t}(\alpha(u_2), u_2).$$

The U and V satisfy the dynamic programming principle which for simplicity is stated with $h \equiv 0$. The proof of this result rests on the results about uniqueness of viscosity solutions to fully nonlinear second-order pde as well as some appropriate discretization of the game in time but not in space and we refer the reader to Fleming [64].

Theorem 3.3.1. *Let $t, \tau \in [0,T]$ be such that $t \leqslant \tau$. for every $x \in \mathbb{R}^n$*

$$V(x,t) = \inf_{\beta \in \Gamma_2(t)} \sup_{u_1 \in U_1(t)} E_{x,t}\{V(x_\tau, \tau)\},$$

and

$$U(x,t) = \sup_{\alpha \in \Gamma_1(t)} \inf_{u_2 \in U_2(t)} E_{x,t}\{U(x_\tau, \tau)\}.$$

With this result, one can study the connections between U and V and the associated Bellman-Isaacs equations which are of the form

$$\begin{cases} y_t + H(D^2 y, Dy, x, t) = 0 \ \text{ in } \mathbb{R}^n \times 0, T], \\ \qquad\qquad\qquad y = g \ \text{ on } \mathbb{R}^n \times \{T\}, \end{cases} \tag{3.3.3}$$

with

$$\begin{aligned} H(A,p,x,t) &= H^-(A,p,x,t) \\ &= \max_{u_1 \in U_1} \min_{u_2 \in U_2} \left[\frac{1}{2} tr(a(x,t,u_1,u_2)A + b(x,t,u_1 u_2).p + h(x,t,u_1,u_2) \right] \end{aligned} \tag{3.3.4}$$

and

$$\begin{aligned} H(A,p,x,t) &= H^+(A,p,x,t) \\ &= \min_{u_2 \in U_2} \max_{u_1 \in U_1} \left[\frac{1}{2} tr(a(x,t,u_1,u_2)A + b(x,t,u_1 u_2).p + h(x,t,u_1,u_2) \right] \end{aligned} \tag{3.3.5}$$

where $a = \sigma\sigma^T$.

We will now give a result for the viscosity solution for (3.3.3) and a comparison principle.

Theorem 3.3.2. *A continuous function* $y : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ *is a viscosity solution* (*resp. supersolution*) *of* (3.2.3) *if*

$$y \leqslant g \ \ on \ \ \mathbb{R}^n \times \{T\},$$

(*respectively for,* $y \geqslant g$ *on* $\mathbb{R}^n \times \{T\}$), *and*

$$\phi_t(x,t) + H\big(D^2\phi(x,t), D\phi(x,t), x, t\big) \geqslant 0,$$

(*resp.* $\phi_t(x,t) + H(D^2\phi(x,t), D\phi(x,t), x, t) \leqslant 0$), *for every smooth function* ϕ *and any local maximum* (*respectively, minimum*) (x,t) *of* $y - \phi$.

Following result is obtained in Ishii [92].

Theorem 3.3.3. *Assume that the functions* b, g, h, *and* σ *are bounded and Lipschitz continuous. If* z *and* $\widetilde{z}$ (*resp.* y *and* $\widetilde{y}$) *are viscosity subsolution and supersolution of* (3.2.3) *with* H *given by* (3.2.4) (*resp. of* (3.2.3) *with* H *given by* (3.2.5)) *with terminal data* g *and* $\widetilde{g}$ *and if* $g \leqslant \widetilde{g}$ *on* $\mathbb{R}^n \times \{T\}$, *then* $z \leqslant \widetilde{z}$ (*resp.,* $y \leqslant \widetilde{y}$) *on* $\mathbb{R}^n \times [0,T]$.

Following is the main result for the zero-sum stochastic differential game problem with two players which is stated with out proof. The proof is given in Fleming and Souganidis [64] which is tedious and involve several approximation procedures.

Theorem 3.3.4. (i) *The lower value* V *is the unique viscosity solution of* (3.3.3) *with* H *as in* (3.3.4).

(ii) *The upper value* U *is the unique viscosity solution of* (3.3.3) *with* H *as in* (3.3.5).

For the dynamics of (3.3.3) with initial time $t = 0$, and for a discounted payoff

$$J(u_1,u_2) = E\left\{\int_0^\infty e^{-\lambda s} h(x(s),u_1(s),u_2(s))ds\right\},$$

the existence of value function is obtained by Swiech [190] using a different approach. The so called sub-and super optimality inequalities of dynamic programming are used in the proofs. In this approach to the existence of value functions, one start with solutions of the upper and lower Bellman-Isaacs equations which exist by the general theory and then prove that they must satisfy certain optimality inequalities which in turn yield that solutions are equal to the value functions. For further analysis of the subject problem see Swiech [190].

3.4 Stochastic differential games with multiple modes

In Ghosh and Marcus [76], two person stochastic differential games with multiple modes are studied. The state of the system at time t is given by a pair $(x(t),\theta(t)) \in \mathbb{R}^n \times S$, where $S = \{1,2,\ldots,N\}$. The discrete component $\theta(t)$ describes the various modes of the system. The continuous component $x(t)$ is governed by a "controlled diffusion process" with a drift vector which depends on the discrete component $\theta(t)$. Thus $x(t)$ switches from one diffusion path to another at random times as the mode, $\theta(t)$, changes. The discrete component $\theta(t)$ is a "controlled Markov chain" with transition rate matrix depending on the continuous component. The evolution of the process $(x(t),\theta(t))$ is given by the following equations

$$dx(t) = b(x(t),\theta(t),u_1(t),u_2(t))dt + \sigma(x(t),\theta(t))dw(t), \tag{3.4.1}$$

and

$$P(\theta(t+\delta t) = j \mid \theta(t) = i, x(s), \theta(s), s \leqslant t) = \lambda_{ij}(x(t))\delta t + \circ(\delta t), \quad i \neq j, \tag{3.4.2}$$

for $t \geqslant 0$, $x(0) = x \in \mathbb{R}^n$, $\theta(0) = i \in S$, where b, σ, λ are suitable functions. In a zero sum game player I is trying to maximize and player II is trying to minimize the expected payoff, that is,

$$J_{x,i}(u_1,u_2) = E_{x,i}\left[\int_0^\infty e^{-\alpha t} r(x(t),\theta(t),u_1(t),u_2(t))dt\right], \tag{3.4.3}$$

over their respective admissible strategies, where $\alpha > 0$ is the discount factor and $r : \mathbb{R}^n \times S \times U_1 \times U_2 \to \mathbb{R}$ is the payoff function and is defined by

$$r(x,i,u_1,u_2) = \int_{V_2}\int_{V_1} \bar{r}(x,i,v_1,v_2)u_1(dv_1)\,u_2(dv_2)\,.$$

Here V_l, $l = 1,2$ are compact metric spaces and $U_l = \mathscr{P}(V_l)$ the space of probability measures on V_l endowed with the topology of weak convergence and $\bar{r} : \mathbb{R}^n \times S \times V_1 \times V_2 \to \mathbb{R}$. Also let

$$\bar{b} : \mathbb{R}^n \times S \times V_1 \times V_2 \to \mathbb{R}^n$$
$$\sigma : \mathbb{R}^n \times S \to \mathbb{R}^{n \times n}$$

and

$$\lambda_{ij} : \mathbb{R}^n \to \mathbb{R}, \ \ 1 \leqslant i, j \leqslant N, \ \ \lambda_{ij} \geqslant 0, \ \ i \neq j, \quad \sum_{j=1}^{N} \lambda_{ij} = 0.$$

The following assumption is made.

(**A3.4.1**):

(i) For each $i \in S$, $\bar{b}(\cdot,i,\cdot,\cdot)$, $\bar{r}(\cdot,i,\cdot,\cdot)$ is bounded, continuous and Lipschitz in its first argument uniformly with respect to the rest.

(ii) For each $i \in S$, $\sigma(\cdot,i)$ is bounded and Lipschitz with the least eigen value of $\sigma\sigma'(\cdot,i)$ uniformly bounded away from zero.

(iii) For $i, j \in S$, $\lambda_{ij}(\cdot)$ is bounded and Lipschitz continuous.

Define

$$b_k(x,i,u_1,u_2) = \int_{V_1} \int_{V_2} \bar{b}_k(x,i,v_1,v_2) u_1(dv_1) u_2(dv_2), \ \ k = 1,\ldots,n$$

and

$$b(x,i,u_1,u_2) = [b_1(x,i,u_1,u_2),\ldots,b_n(x,i,u_1,u_2)]'.$$

If $u_l(\cdot) = v_l(x(\cdot),\theta(\cdot))$ for a measurable $v_l : \mathbb{R}^n \times S \to U_l$, then $u_l(\cdot)$ is called a Markov strategy for the lth player. Let M_l denote the set of Markov strategies for player l. A strategy $u_l(\cdot)$ is called pure if u_l is a Dirac measure, i.e., $u_l(\cdot) = \delta_{v_l}(\cdot)$, where $v_l(\cdot)$ is a V_l valued nonanticipative process. For $p \geqslant 1$ define

$$W^{2,p}_{\text{loc}}(\mathbb{R}^n \times S) = \left\{ f : \mathbb{R}^n \times S \to \mathbb{R} : \text{ for each } i \in S, \ \ f(\cdot,i) \in W^{2,p}_{\text{loc}}(\mathbb{R}^n) \right\}.$$

$W^{2,p}_{\text{loc}}(\mathbb{R}^n \times S)$ is endowed with the product topology of $\left(W^{2,p}_{\text{loc}}(\mathbb{R}^n)\right)^N$. For $f \in W^{2,p}_{\text{loc}}(\mathbb{R}^n \times S)$, we can write

$$L^{v_1,v_2} f(x,i) = L^{v_1,v_2}_i f(x,i) + \sum_{j=1}^{N} \lambda_{ij} f(x,j),$$

where

$$L^{v_1,v_2}_i f(x,i) = \sum_{j=1}^{n} \bar{b}_j(x,i,v_1,v_2) \frac{\partial f(x,i)}{\partial x_j} + \frac{1}{2} \sum_{j,k=1}^{n} a_{jk}(x,i) \frac{\partial^2 f(x,i)}{\partial x_j \partial x_k}.$$

Here, $a_{jk}(x,i) = \sum_{l=1}^{n} \sigma_{jl}(x,i)\sigma_{kl}(x,i)$.

Define

$$L^{u_1,u_2} f(x,i) = \int_{V_1}\int_{V_2} L^{v_1,v_2} f(x,i) u_1(dv_1) u_2(dv_2).$$

The Isaacs equation for this problem is given by

$$\begin{aligned}\inf_{u_2\in U_2}\sup_{u_1\in U_1}\left[L^{u_1,u_2}\phi(x,i) + r(x,i,u_1,u_2)\right] &= \sup_{u_1\in U_1}\inf_{u_2\in U_2}\left[L^{u_1,u_2}\phi(x,i) + r(x,i,u_1,u_2)\right]\\ &= \alpha\phi(x,i). \qquad (3.4.4)\end{aligned}$$

This is a quasilinear system of uniformly elliptic equations with weak coupling in the sense that the coupling occurs only in the zeroth order term. Now we will state the following results from Gosh and Marcus [76].

Theorem 3.4.1. *Under* (A3.4.1) *the equation* (3.4.4) *has a unique solution in* $C^2(\mathbb{R}^n \times S) \cap C_b(\mathbb{R}^n \times S)$.

The result that follows characterizes the optimal Markov strategies for both players.

Theorem 3.4.2. *Assume* (A3.4.1). *Let* $u_1^* \in M_1$ *be such that*

$$\begin{aligned}&\inf_{u_2\in U_2}\left[\sum_{j=1}^{n} b_j(x,i,u_1^*(x,i),u_2)\frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N}\lambda_{ij}(x)V(x,j) + r(x,i,u_1^*(x,i),u_2)\right]\\ &\quad = \sup_{u_1\in U_1}\inf_{u_2\in U_2}\left[\sum_{j=1}^{n} b_j(x,i,u_1,u_2)\frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N}\lambda_{ij}(x)V(x,j) + r(x,i,u_1,u_2)\right]\end{aligned}$$

for each i and a.e. in x. Then u_1^* *is optimal for player I. Similarly, let* $u_2^* \in M_2$ *be such that*

$$\begin{aligned}&\sup_{u_1\in U_1}\left[\sum_{j=1}^{n} b_j(x,i,u_1,u_2^*(x,i))\frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N}\lambda_{ij}(x)V(x,j) + r(x,i,u_1,u_2^*(x,i))\right]\\ &\quad = \inf_{u_2\in U_2}\sup_{u_1\in U_1}\left[\sum_{j=1}^{n} b_j(x,i,u_1,u_2)\frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N}\lambda_{ij}(x)V(x,j) + r(x,i,u_1,u_2)\right]\end{aligned}$$

for each i and a.e. x. Then u_2^* *is optimal for player II.*

This kind of game typically occurs in a pursuit-evasion problems where an interceptor tries to destroy a specific target. Due to swift movements of the evader and the corresponding response by the interceptor the trajectories keep switching rapidly.

In Kushner and Chamberlain [111], the problem of the numerical solution of the nonlinear partial differential equation associated with the subject game is considered. In general, due to the nonlinearities and to the nonellipticity or nonparabolicity of these equations, the available theory is not much helpful in choosing finite difference approximations, guaranteeing the convergence of the iterative procedures, or providing an interpretation of the approximation. For a specific problem, a finite difference scheme is given in Kushner and Chamberlain [111], so that the convergence of the iterative process is guaranteed. With the development of weak convergence theory for game problems, Ramachandran [158], and the numerical methods described in Kushner and Dupuis [112], it is possible to develop computational methods for stochastic differential games. This will be the topic of Chapter 8.

Chapter 4

Formal Solutions for Some Classes of Stochastic Linear Pursuit-Evasion Games

4.1 Introduction

As mentioned in Chapter 1, considerable attention has been given recently to pursuit-evasion games with linear dynamics and quadratic payoff. Consider the transition equation

$$\begin{aligned} &\dot{x} = G_P u - G_e v; \quad \text{with initial condition} \\ &x(0) = x_0. \end{aligned} \tag{4.1.1}$$

where

i) $x(t) \in \mathbb{R}^n$, is the state of the game;
ii) $u(t) \in \mathbb{R}^P$ is the pursuer's control chosen at time t;
iii) $v(t) \in \mathbb{R}^q$ is the evader's control chosen at time t;
and
iv) G_p and G_e are $(n \times p)$ and $(n \times q)$ time varying matrices.
The payoff of this game is given by

$$J = \frac{1}{2}\left\{x'(t)S(t)x(t) + \int_0^T \left[u'(t)B(t)u(t) - v'(t)C(t)v(t)\right] dt\right\} \tag{4.1.2}$$

where

v) $T = t_f$ is some prescribed terminal time. The initial time t_0 is taken to be zero without loss of generality.
vi) B and C are symmetric, positive definite, time-varying matrices;
and
vii) $S(t) = S_f$ is a symmetric and positive and semi-definite matrix which will be defined later.

Ho, Bryson, and Baron [89] have solved this game for the case where both players have perfect knowledge of the state of the game, $x(t)$. When a solution exists and it is given by

$$U^* : u = -B^{-1} G_P' s x;$$

and

$$V^* : v = -C^{-1} G_e' sx$$

where S is the solution to the matrix Riccati equation, given by

$$\dot{S} = S \left[G_P B^{-1} G_P' - G_e C^{-1} G_e' \right] S; S(t) = S_f.$$

If the solution is bounded on the interval $[0,T]$ then the strategies U^* and V^* are minimax.

4.2 Preliminaries

Consider a stochastic differential pursuit-evasion game of the form

$$\frac{d}{dt} x(t;\omega) = G_p(\omega)\, u(t;\omega) = G_e(\omega)\, v(t;\omega) \tag{4.2.1}$$

where

i) $\omega \in \Omega$, where Ω is the supporting set of a complete probability measure space $(\Omega, \mathrm{A}, \mu)$;

ii) $x(t;\omega) \in C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ is the is the n-dimensional state vector;

iii) $u(t;\omega) \in L_2(\Omega, \mathrm{A}, \mu)$ is the p-dimensional pursuer's control vector chosen at time $t \in \mathbb{R}_+$ for each $\omega \in \Omega$;

iv) $v(t;\omega) \in L_2(\Omega, \mathrm{A}, \mu)$ is the q-dimensional evader's control vector chosen at time $t \in \mathbb{R}_+$ for each $\omega \in \Omega$;

and

v) $G_P(\omega)$ and $G_3(\omega)$ are $(n \times p)$ and $(n \times q)$ dimensional random matrices. The payoff to be minimaximized is the expected value of equation (4.1.2). That is,

$$\begin{aligned} J = \frac{1}{2} E \Bigg\{ & x'(t,\omega) S_f(\omega) x(T;\omega) \\ & + \int_0^T \left[u'(t;\omega) B(t) u(t;\omega) - v'(t;\omega) C(t) v(t;\omega) \right] dt \Bigg\} \end{aligned} \tag{4.2.2}$$

where

vi) T is some prescribed termination time;

and

vii) B and C are symmetric and positive definite matrices.

It is seen that equation (4.2.1) is another special case of equation (2.1.1). In this Chapter we deal with stochastic differential games where the state variable does not appear on the right-hand side of the transition equation. Physically this means that the state exerts no control itself.

Integrating equation (4.2.1) with respect to t we obtain

$$x(t;\omega) = \int_0^t \left[G_P(\omega)u(\tau;\omega) - G_e(\omega)v(\tau;\omega)\right] dt, \quad t \geqslant 0. \tag{4.2.3}$$

Theorem 2.2.3 reveals that the game has a unique random solution if

$$\|G_P(\omega)u(t;\omega) - G_e(\omega)v(t;\omega)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \frac{P}{M},$$

where M is the norm of the operator T defined by

$$(Tx)(t;\omega) = \int_0^t x(\tau;\omega)\, d\tau.$$

Since the stochastic free term in (4.2.3) is identically zero, we take $\gamma \equiv 0$; and since the stochastic kernel does not explicitly involve $x(t;\omega)$, we can take $\lambda \equiv 0$.

We will now attempt a formal derivation of this random solution.

4.3 Formal solution for a Stochastic Linear Pursuit-Evasion game with perfect information

In this section we will assume that both players make perfect measurement of the state of the game. We will consider a multistage differential game formed by discretizing equation (4.2.1). By dividing the stochastic differential game into N short games, we can then apply the work of Ho, Bryson, and Baron [89] to approximate the optimal controls for each of these games. This iterative procedure generates a formal random solution to the stochastic game (4.2.1).

We shall divide the time interval $[0,T]$ into N small subintervals, each of length δ. By requiring the players to use constant controls during each subinterval, the transition and criterion equations can be expressed in multistage form, that is,

$$x[(i+1)\Delta;\omega] = x(i\Delta;\omega) + \Delta\left[G_P(\omega)u(i\Delta;\omega) - G_e(\omega)v(i\Delta;\omega)\right] \text{ for } i = 0,1,2,\ldots,N-1$$

and

$$\begin{aligned} J = \frac{1}{2}E\Bigg\{ & x'(N\Delta;\omega)S_f(\omega)x(N\Delta;\omega) \\ & + \sum_{i=0}^{N-1}\left[u'(i\Delta;\omega)B(i\Delta)u(i\Delta;\omega) - v'(i\Delta;\omega)C(i\Delta)v(i\Delta;\omega)\right]\Bigg\}. \end{aligned} \tag{4.3.1}$$

The time interval δ is assumed short enough so that $\omega \in \Omega$ does not change significantly during a subinterval, and the players are allowed to make perfect measurements of the state only at times i, $i = 0,1,2,\ldots,N-1$. Their controls must be based on these measurements.

Under the above assumptions, we have a series of deterministic games of the form

$$x(t;\omega_i) = G_P(\omega_i)u(t;\omega_i) - G_e(\omega_i)v(t;\omega_i),$$

for $t \in i\delta$ and $\omega_i \in \Omega$. At each instant, i, $i = 0,1,2,\ldots,N-1$, $\omega_i \in \Omega$ is chosen by nature and assumed fixed. The players observe the state of the game and choose their optimal controls for the next subinterval. As previously mention, Ho, Bryson, and Baron ([89] hbb) have determined the optimal controls for each of these short deterministic games. When a solution exists, it is given by

$$\begin{aligned} u(t;\omega_i) &= -B^{-1}(t)G_P'(\omega_i)S(t;\omega_i)x(t;\omega_i); \\ v(t;\omega_i) &= -C^{-1}(t)G_e'(\omega_i)S(t;\omega_i)x(t;\omega_i), \end{aligned} \tag{4.3.2}$$

for $t \in [i\Delta,(i+1)\Delta]$, $i = 0,1,2,\ldots,N-1$, $\omega_i \in \Omega$ and S a solution of

$$\dot{S} = S\left[G_P B^{-1} G_P' - G_e C^{-1} G_e'\right] S$$

and

$$S(i\Delta;\omega_i) = S_f(\omega_i) \quad i = 1,2,3,\ldots,N.$$

At the end of each subinterval the process is repeated until the terminal time $T = N\delta$ is reached.

Since the controls given by (4.3.2) are optimal, that is minimaximized the expected payoff, over the subintervals $i\Delta$, the stochastic controls

$$u(t;\omega) = u(t;\omega_i)$$

and

$$v(t;\omega) = v(t;\omega_i)$$

for $t \in i\delta$ and $\omega_i \in \Omega$, $i = 0,1,2,3,\ldots,N-1$, will be optimal for the game (4.2.1) in the sense that as $\Delta \to 0$ the expected payoff (4.3.1) will approach the minimax of equation (4.2.2).

Differential games and multistage games with perfect information have been the subject of many publications. Now, what if one or both players cannot make exact measurements? A logical extension is an investigation of a pursuit-evasion problem in which the players have imperfect knowledge of the states involved.

4.4 On Stochastic Pursuit-Evasion games with imperfect information

Differential games with noisy state observations have also been investigated by some authors, among them Behn and Ho [19] and Rhodes and Luenberger [167], under somewhat restricted situations. Yoshikawa [217] has solved a simple one-dimensional, two-stage game of the form

$$x_{i+1} = ax_i + u_i + v_i + \xi_i \text{ for } i = 0 \text{ and } 1$$

with payoff

$$\mathrm{J} = x_2^2 + \sum_{i=0}^{1} (b_i u_i^2 + c_i v_i^2)$$

and with the noisy state observations

$$y = x_1 + \eta$$

and

$$z = x_1 + \xi,$$

where $\xi_0,\ \xi_1,\ \eta$, and ξ are mutually independent zero mean noises; but has been unable to solve more general multistate games. The difficulty is that there appears to be infinite number of terms in the optimal strategies of each of the two players. That is, they are based on estimates of estimates of estimates ..., Behn and Ho [19] have termed this the *closure problem* in stochastic pursuit-evasion games and found conditions which are sufficient for closure.

Consider the optimization of the payoff

$$J = E\left\{ \frac{a^2}{2} \|y(t_f)\|^2 + \frac{1}{2} \int_{t_0}^{t_f} \left[\|u(t)\|^2 R_P - \|v(t)\|^2 R_e \right] dt \right\}, \tag{4.4.1}$$

subject to the differential constraint

$$\begin{aligned} y(t) &= G_P(t_f,t)u(t) - G_e(t_f,t)v(t) \\ y(t_0) &= y_0 \end{aligned} \tag{4.4.2}$$

where the pursuer can make perfect measurements; but the evader's measurements are given by

$$z(t) = H(t)y(t) + w(t),$$

where w is a Gaussian white $(0, Q(t))$ process. Assume that the controls are bounded and continuous so that the differential equation (4.4.2) is meaningful and Integrable. The optimal strategy pair is assumed given by

$$U^* :\ u(t) = C_P(t)y(t) + D_P(y)\widetilde{y}(t) \tag{4.4.3}$$

and

$$V^* : \ v(t) = C_e(t)\hat{y}(t) \tag{4.4.4}$$

where $\hat{y}(t)$ is the evader's optimal estimate of $y(t)$ based on the measurements $z(\tau)$, $t_o \leqslant \tau \leqslant t$. and $\widetilde{y}(t)$ is the error of the evader's estimate, $\widetilde{y}(t) = y(t) - \hat{y}(t)$. The values of the feedback gain matrices C_e, C_P, and D_p are then determined by standard optimization techniques.

Behn and Ho [19] showed that C_p and C_e are the same feed-back gain matrices employed by the players in the deterministic problem. The evader merely uses the feedback strategy employed in the deterministic game to operate on his optimal estimate $\hat{y}(t)$ of the state $y(t)$. From the pursuer's point of view, the optimal strategy is the deterministic feedback control plus a term to take advantage of the inaccuracy of the evader's measurements.

Using (4.4.3) and (4.4.4) to find the controls $u(t)$ and $v(t)$ and eliminating $\hat{y}(t)$, the criterion function (4.4.1) becomes

$$\begin{aligned} J = E\Bigg\{ \frac{a^2}{2}\|y(t)\|^2 + \frac{1}{2}\int_{t_0}^{t_f} & \|y(t)\|^2 C_P' R_P C_P \\ & + y'(t) C_P' R_P D_P \widetilde{y}(t) + \widetilde{y}(t) D_P' R_P C_P y(t) \\ & + \|\widetilde{y}(t)\|^2 D_P' R_P D_P - \|y(t) - \widetilde{y}(t)\| C_e' R_e C_e \big] \, dt \Bigg\}, \end{aligned}$$

subject to

$$\dot{y}(t) = [G_P C_P - G_e C_e] y(t) + [G_P D_P + G_e C_e] \widetilde{y}(t)$$

and

$$y(t_0) = y_0.$$

Behn and Ho [19] found that if the following two conditions are satisfied,

i) the dimension of $y(t)$ equals the dimension of $v(t)$;

 and

ii) $G_e^{-1}(t)$ exists for all $t < t_f$

then the investigation is still continuing on the existence of a random solution to equation (4.2.1) when one or both of the players have imperfect measurements.

4.5 Summary

The subject of this Chapter was the existence of a random solution of the stochastic linear-quadratic pursuit-evasion game of the form

$$\frac{d}{dt}x(t;\omega) = G_P(\omega)u(t;\omega) - G_e(\omega)v(t;\omega),$$

where the state has no effect on the right hand side of the equation.

Applying a theorem from the last Chapter, we found a sufficient condition for the game to have a unique random solution. By discretizing the game we were able to derive a formal random solution under the assumption that both players make perfect observations. We then presented the problem of the existence of a solution if one of the players cannot make perfect observations of the state of the game and pointed out the difficulties encountered.

Chapter 5

N-Person Noncooperative Differential Games

5.1 Introduction

In the previous four chapters we have presented the foundations for two-person zero sum differential games. In those cases, there were a single performance criterion which one player tries to minimize and the other tries to maximize. In applications, there are many situations in which more than two players and each player try to maximize (or minimize) his/her individual performance criterion, and the sum of all players' criteria is not necessarily zero nor is it a constant. Such cases are called N-person non-zero sum differential games. A non-zero-sum game is the game in which each player chooses a strategy as his/her best response to other players' strategies. An equilibrium, in this case, is a set of strategies such that when applied no player will profit from unilaterally changing his/her own strategy. In this Chapter, we will present some fundamental aspects of this case. First, we will present a pursuit-evasion case to get exposed to the idea of a non-zero sum game, and then extended to a general case.

5.2 A stochastic Pursuit-Evasion Game

5.2.1 *Two Person Non-Zero Sum Game*

In this section we shall consider a stochastic two person differential game of the general form given by

$$\frac{d}{dt}x(t;\omega) = f(t, x(t;\omega), u(t), v(t)), \tag{5.2.1}$$

where

(i) $\omega \in \Omega$ for Ω the supporting set of a complete probability measure space $(\Omega, \mathrm{A}, \mu)$;
(ii) $x(t;\omega) \in L_2(\Omega, \mathrm{A}, \mu)$ is the n-dimensional random state vector for each $t \geqslant 0$;

(iii) $u(t) \in E^P$ is the p-dimensional control vector of the first player (pursuer);

(iv) $v(t) \in E^q$ is the q-dimensional control vector of the second player (evader); and

(v) the initial conditions $x(0;\omega)$ are given by the known n-dimensional random vector with

$$x_0(\omega) = (x_{01}(\omega), \ldots, x_{0n}(\omega)) \in L_2(\Omega, \mathrm{A}, \mu).$$

We will assume as admissible control functions $u = u(t)$ and $v = v(t)$ which are measurable functions of t alone. That is, the controls are deterministic. Assuming an initial fixed time at $t = 0$, we will allow the terminal time $t_f(\omega)$ to vary randomly as a function of $\omega \in \Omega$ where Ω is some compact set. The assumption of a compact Ω is not restrictive in any way. We will consider an integral payoff for each player. That is,

$$J_1 = \int_\Omega G_i[t_f(\omega), x(t_f(\omega);\omega)]d\mu(\omega),$$

where G_i are real valued continuous functions for $i = 1, 2$.

The constraint set and boundary conditions will also be allowed to vary with $\omega \in \Omega$. That is, we shall define the constraint set A_Ω as a compact subset of the tx-space $\mathbb{R}_+ \times C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ and let the terminal set B_Ω be a closed subset of the tx-space $\mathbb{R}_+ \times C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$. The unions of these sets for all $\omega \in \Omega$ will be denoted by A and B. That is, the compact set $A = \bigcup_{\omega\in\Omega} A_\Omega$ and $B = \bigcup_{\omega\in\Omega} B_\Omega$.

Now, for each $t \in [0, T]$, where $T > 0$ is some fixed time, we shall define the set:

$$T_f = \left\{t_f(\omega) : 0 \leqslant t_f(\omega) \leqslant T,\ \omega \in \Omega\right\},$$

of all terminal times as a family τ of equicontinuous functions this is closed in the uniform topology.

The control sets$U(t)$ and $V(t)$ will be taken as subsets of the Euclidean spaces E^p and E^q, respectively. We shall define the sets

$$M_\omega = \{(t, x, u, v) : (t, x) \in A_\omega,\ u \in U,\ v \in V\}$$

and

$$M = \bigcup_{\omega\in\Omega} M_\omega = \{(t, x, u, v) : (t, x) \in A,\ u \in U,\ v \in V\},$$

as compact subsets of the space $\mathbb{R}_+ \times C_c \times E^p \times E^q$. The function $f = (f_1, f_2 \ldots, f_n)$ is continuous from M into $L_2(\Omega, \mathrm{A}, \mu)$. We shall assume further that f is *separable*. That is, there are functions g and h, such that,

$$f(t, x(t;\omega), u(t), v(t)) = g(t, x(t;\omega), u(t)) + h(t, x(t;\omega), v(t)).$$

Furthermore, we shall assume that g and h are Lipschitzian in x uniformly in t over u and v. That is, there exist finite constants λ_1 and λ_2, such that,

$$\|g(t,x,u) - g(t,y,u)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \lambda_1 \|x-y\|_{L_2(\Omega,\mathrm{A},\mu)}$$

and

$$\|h(t,x,v) - h(t,y,v)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \lambda_2 \|x-y\|_{L_2(\Omega,\mathrm{A},\mu)}.$$

Under the above assumptions, we have

$$\|f(t,x,u,v) - f(t,y,u,v)\|_{L_2(\Omega,\mathrm{A},\mu)} \leqslant \lambda \|x-y\|_{L_2(\Omega,\mathrm{A},\mu)},$$

for all $(t,x,u,v),\ (t,y,u,v) \in M$ where $\lambda = \lambda_1 + \lambda_2$. This guarantees that for each fixed $t \in [0,T]$, the state vector $x(t;\omega)$ is $C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu))$.

5.2.2 *Preliminaries*

Consider the class ψ of all triples $(x(t;\omega), u(t), v(t))$ for t, $t_f(\omega) \in [0,T]$ and $\omega \in \Omega$ which satisfy the following conditions.

(A5.2.1):

i) for each fixed $\omega \in \Omega$, $x(t;\omega)$ is absolutely continuous $[0,T]$;
ii) $u(t) \in U(t) \subset E^p$ is a measurable function for $t \in [0,T]$;
iii) $v(t) \in V(t) \subset E^q$ is a measurable function for $t \in [0,T]$;
iv) for each $\omega \in \Omega$ and $t\,\varepsilon[0,T]$, $(t, x(t;\omega)) \in A_\Omega$;
v) $x(0,\omega) = x_0(\omega)$;
vi) $t_f(\omega) \in \tau$;
vii) for each $\omega \in \Omega$, $(t_f(\omega), x(t_f(\omega);\omega)) \in B_\Omega$; and
viii) the ordinary differential equation

$$\frac{d}{dt}x(t;\omega) = f(t, x(t;\omega), u(t), v(t))$$

is satisfied μ-a.e. in $[0,T]$.

Definition 5.2.1. Any triple $(x(t;\omega), u(t), v(t))$ for t, $t_f(\omega) \in [0,T]$ which satisfies assumption (A5.2.1) is called an *admissible triple*. The random vector $x(t;\omega)$ is called a *random solution* and $u(t)$ and $v(t)$ are called *admissible controls*.

Let k_i, $i = 1,2$, be continuous functional defined on the set of continuous functionals

$$W = \left\{ (w_1(\omega), w_2(\omega)) : (t_f(\omega), w_2(\omega)) \in B_\omega \right\}$$

and assume that k_i is bounded from below on a subset W' of W, where

$$W' = \left\{(w_1(\omega), w_2(\omega)) : (t_f(\omega), w_2(\omega)) \in B_\omega \cap A_\omega\right\}.$$

Then the functionals

$$\begin{aligned} J_1[x,u] &= K_1[\eta(x)(\omega)] \\ &= K_1[t_f(\omega), x(t_f(\omega);\omega)] \end{aligned}$$

and

$$\begin{aligned} J_2[x,v] &= K_2[\eta(x)(\omega)] \\ &= K_2[t_f(\omega), x(t_f(\omega);\omega)] \end{aligned}$$

are called cost functionals. Player one exerts control on the state variable $x(t;\omega)$ through his control variable $u(t)$, so as to minimize $J_1[x,u]$ while player two uses his control, $v(t)$ to minimize $J_2[x,v]$. We are thus led to the following definition of optimal controls.

Definition 5.2.2. If there exists a triple $(x^*(t,w), u^*(t), v^*(t))$, such that,

$$J_1[x^*(t,w), u^*(t)] \leqslant J_1[x(t,w), u(t)]$$

and

$$J_2[x^*(t,w), v^*(t)] \leqslant J_2[x(t,w), v(t)],$$

for all triples $(x(t;\omega), u(t), v(t)) \in \psi$, then the triple $(x^*(t,w), u^*(t), v^*(t))$ is called an *optimal triple*. The controls $u^*(t)$ and $v^*(t)$ are called *optimal controls; and $x^*(t,w)$* is called an *optimal random solution*.

It should be noted that, although the optimal triple $(x^*(t;\omega), u^*(t), v^*(t))$ need not be unique in ψ, the value of the cost functionals are the same for all optimal triples. For ideas of the proof of the next result, we refer the reader to Nicholas [141].

Lemma 5.2.1. *Given a stochastic differential game as described above where $x(t;\omega)$ is uniformly continuous for $(t;\omega) \in [0,T] \times \Omega$ and given any sequence of admissible triples $\{x(t;\omega)_k u(t)_k, v(t)_k\}$, then $\{x(t;\omega)_k,\ k = 1,2,\ldots\}$ forms an equicontinuous and equibounded family of functions on $[0,T] \times \Omega$.*

5.2.3 *Main Results*

It will be necessary to impose some further requirements on the state equation (5.2.1) and on the control sets $U(t) \subset E^p$ and $V(t) \subset E^q$. Let us assume the following.

(A5.2.2):

(a) f is *completely separable.* That is, the random state vector and the controllers all act independently. $f(t,x,u,v) = f(t,x) + g(t,u) + h(t,v)$ for $(t,x,u,v) \in M$;

(b) $U(t)$ and $V(t)$ are compact sets for $t \in [0,T]$;

(c) $u(t)$ and $v(t)$ are upper semicontinuous functions of $t \in [0,T]$; and

(d) The sets $g(t,U(t))$ and $h(t,V(t))$ are convex subsets of the space $L_2(\Omega,\mathrm{A},\mu)$ where we define

$$g(t,U(t)) = \{y \in L_2(\Omega,\mathrm{A},\mu) : y = g(t,u),\ u \in U(t)\}$$

and

$$h(t,V(t)) = \{z \in L_2(\Omega,\mathrm{A},\mu) : z = h(t,v),\ v \in V(t)\}.$$

Then, $f(t,x,U(t),V(t))$ is a convex subset of $L_2(\Omega,\mathrm{A},\mu)$ for each $(t,x) \in A$.

(A5.2.3):

i) *The constraint sets A_Ω and $A = \bigcup_{\omega\in\Omega} A_\Omega$ is compact subsets of the tx-space $\mathbb{R}_+ \times C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu))$.*

ii) *The control sets $U(t)$ and $V(t)$ are compact subsets of E^p and E^q for every $t \in [0,T]$.*

iii) *The control functions $u(t)$ and $v(t)$ is upper semicontinuous functions of $t \in [0,T]$.*

iv) *M_Ω and $M = \bigcup_{\omega\in\Omega} M_\Omega$ are compact subsets of the space $\mathbb{R}_+ \times C_c(\mathbb{R}_+, L_2(\Omega,\mathrm{A},\mu)) \times E^p \times E^q$.*

v) *The function $f(t,x,u,v = f(t,x) + g(t,u) + h(t,v))$ is continuous from M into $L_2(\Omega,\mathrm{A},\mu)$.*

vi) *$g(t,U(t))$ and $h(t,V(t))$ are convex subsets of $L_2(\Omega,\mathrm{A},\mu)$ for every $(t,x) \in A$.*

vii) *$\{x(t;\omega)_k\}$ with $\{t_{fk}(\omega) : \omega \in \Omega\} \in \tau$ is a sequence of random solutions satisfying assumptions (A5.2.1) and converging uniformly to a function $x^*(t;\omega)$ which is absolutely continuous in $[0,T]$ for each $\omega \in \Omega$.*

viii) *$t_{fk}(\omega)$ converges uniformly to $t_f(\omega)$.*

Now, we have following result.

Theorem 5.2.1. *Under the above assumptions (A5.2.3), there exist measurable controls $u^*(t)$ and $v^*(t)$ such that the triple $(x^*(t;\omega), u^*(t), v^*(t))$ with stochastic terminal time*

$t_f(\omega)$ satisfies conditions (A5.2.1). That is, $x^*(t;\omega)$ is a random solution with stochastic terminal time $t_f(\omega) \in \tau$; and $J_1[x_k,u^*]$ and $J_2[x_k,u^*]$ converge uniformly to $J_1[x^*,u^*]$ and $J_2[x^*,u^*]$ respectively.

Proof. We will only sketch the proof here, see Nicholas [141]. We have,

$$x(t_{fk}(\omega);\omega)_k \xrightarrow{u} x^*(t_f(\omega);\omega)$$

and

$$\eta\left[x(t_f(\omega);\omega)_k\right](\omega) \xrightarrow{u} \eta\left[x^*(t_f(\omega);\omega)\right](\omega).$$

Thus,

$$J_1[x_k,u^*] \xrightarrow{u} J_1[x^*,u^*]$$

and

$$J_2[x_k,v^*] \xrightarrow{u} J_2[x^*,v^*].$$

□

Proof. We had to prove that there exist a measurable control functions $u^*(t) \in U(t)$ and $v^*(t) \in V(t)$, such that,

$$\frac{d}{dt}x^*(t;\omega) = f(t,x^*(t;\omega),u^*(t),v^*(t)), \tag{5.2.2}$$

μ-a.e. in $[0,T]$.
By assumption (vii), $x(t;\omega)_k \to x^*(t;\omega)$ where $x^*(t;\omega)$ is absolutely continuous in $[0,T]$. If, for each $\omega \in \Omega$ we consider the stochastic differential game with constraint set A_Ω, control sets $U(t)$ and $V(t)$, and transition equation (5.2.1), then each of the triples $(x(t;\omega)_k,u(t)_k,v(t)_k)$, $k = 1,2,\ldots$, belongs to the class of admissible triples ψ. Thus, by Cesari's closure Theorem, for each $\omega \in \Omega$, there exist measurable controls $u^*(t) \in U(t)$ and $v^*(t) \in V(t)$, such that,

$$\frac{d}{dt}x^*(t;\omega) = f(t,x^*(t;\omega),\, u^*(t),\, v^*(t)),$$

is μ-a.e. in $[0,T]$. In particular, $(x^*(t;\omega),u^*(t),v^*(t)) \in \psi$. The proof consists of showing that for any given $\omega_0 \in \Omega$, $\left(u^*_{\omega_0}(t),v^*_{\omega_0}(t)\right)$ generates all the random solutions. That is, for any $\omega \in \Omega$, and $\omega_0 \in \Omega$ fixed, we have

$$\frac{d}{dt}x^*(t;\omega) = f\left(t,x^*(t;\omega),u^*_{\omega_0}(t),v^*_{\omega_0}(t)\right),$$

μ-a.e. in $[0,T]$. Letting $u^*(t) = u^*_{\omega_0}(t)$ and $v^*(t) = v^*_{\omega_0}(t)$, this completes the proof. □

We can also state an existence theorem for pursuit-evasion games with state variable in $C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$, the space of continuous functions.

Theorem 5.2.2. *Consider the stochastic differential game satisfying conditions* (i)–(iv) *of* (A5.2.1). *If the class ψ of admissible triples is non-empty there exists an admissible triple $(x^*(t;\omega), u^*(t), v^*(t))$, such that*

$$J_1\left[x^*(t,w), u^*(t)\right] \leqslant J_1\left[x(t,w), u(t)\right]$$

and

$$J_2\left[x^*(t,w), v^*(t)\right] \leqslant J_2\left[x(t,w), v(t)\right],$$

for all admissible triples $(x(t;\omega), u(t), v(t)) \in \psi$.

Proof. We shall give a sketch of the proof. Given any admissible triple $(x(t;\omega), u(t), v(t)) \in \psi$, by the assumption, $\eta[x(t;\omega)](\omega) \in W'$. Since, $K_1[\eta[x(t;\omega)](\omega)]$ and $K_2[\eta[x(t;\omega)](\omega)]$ are assumed bounded from below on W', we have,

$$j_1 = \inf_{\psi} J_1[x(t;\omega), u(t)] > -\infty$$

and

$$j_2 = \inf_{\psi} J_2[x(t;\omega), v(t)] > -\infty.$$

Since there exists at least one admissible triple by the assumptions of the theorem, j_1 and j_2 are finite. Thus, there exists a minimizing sequence of admissible triples $\{x(t;\omega)_k, u(t)_k, v(t)_k\}$ with $t_{fk}(\omega) \in \tau$, such that,

$$J_1[x(t;\omega)_k, u(t)_k] \to j_1$$

and

$$J_2[x(t;\omega)_k, v(t)_k] \to j_2,$$

as $k \to \infty$.

Now, we apply Lemmas 5.2.1 and 5.2.2 to conclude that there are measurable controls $u^*(t) \in U(t)$ and $v^*(t) \in V(t)$, such that, the triple $(x^*(t;\omega), u^*(t), v^*(t))$ satisfies assumptions (i)–(vi) and (viii) of (A5.2.1) and, such that,

$$\eta[x^*(t;\omega)](\omega) = (t_f(\omega), x^*(t_f(\omega);\omega)) \in B_\omega.$$

That is, assumption (vii) of (A5.2.1) is also satisfied.

Thus,

$$(x^*(t;\omega), u^*(t), v^*(t)) \in \psi.$$

Finally, since K_1 and K_2 were assumed continuous on W, we have

$$J_1[x^*(t;\omega), u^*(t)] = \lim_{k\to\infty} J_1\left[x(t;\omega)_k, u(t)_k\right] = j_1$$

and

$$J_2[x^*(t;\omega), v^*(t)] = \lim_{k\to\infty} J_2\left[x(t;\omega)_k, v(t)_k\right] = j_2,$$

hence, the theorem is proven. □

We will now see that the above theorems can be extended further to N-person differential games where, $N > 2$.

5.2.4 *N-Person Stochastic Differential Games*

In this presentation we shall consider N-person stochastic differential games given by

$$\frac{d}{dt}x(t;\omega) = f(t, x(t;\omega), u_1(t), \ldots, u_N(t)), \tag{5.2.3}$$

where

i) $\omega \in \Omega$; and Ω is the supporting set of a complete probability measure space $(\Omega, \mathrm{A}, \mu)$;
ii) $x(t;\omega) \in L_2(\Omega, \mathrm{A}, \mu)$ is an n-dimensional random state vector for each $t \geqslant 0$;
iii) $u_i(t) \in E^{P_i}$ is the p_i-dimensional control vector for player i, $i = 1, 2, \ldots, N$;

and with initial conditions given by the known n-dimensional random vector $x(0;\omega) = x_0(\omega)$.

As before we will take as admissible controls u_i, $i = 1, 2, \ldots, N$, functions which are measurable functions of t alone; and the control sets $U_i(t)$ will be taken as subsets of the Euclidean spaces E^{P_i}.

The constraint set A_Ω will be assumed to be a compact subset of the space $\mathbb{R}_+ \times C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ and the terminal set B_Ω is assumed to be a closed subset of the same space.

We shall let the terminal time $t_f(\omega)$ vary with $\omega \in \Omega$ where Ω is compact; and we shall define the set

$$T_f = \left\{t_f(\omega) : 0 \leqslant t_f(\omega) \leqslant T,\ T > 0,\ \omega \in \Omega\right\},$$

of all termination times as a family τ of equicontinuous functions this is closed in the uniform topology. And, we shall consider integral cost functions given by

$$J_i = \int_{\Omega} G_i[t_f(\omega);\omega]d\mu(\omega),$$

where G_i are real valued continuous functions for $i = 1,2,\ldots,N$.
Finally, we shall define the sets

$$M_\omega = \{(t,x,u_1,\ldots,u_N) : (t,x) \in A_\omega,\ u_i \in U_i\}$$

and

$$M = \bigcup_{\omega\in\Omega} M_\omega = \{(t,x,u_1,\ldots,u_N) : (t,x) \in A_\omega,\ u_i \in U_i\},$$

as compact subsets of the space $\mathbb{R}_+ \times C_c \times E^{\sum P_i}$. We shall assume that the functions f which are continuous from M into $L_2(\Omega,\mathrm{A},\mu)$ are separable and Lipschitzian.
As before, we shall consider a class ψ of all $(N+1)$-tuples, $(x(t;\omega), u_1(t),\ldots,u_N(t))$, for $t, t_f(\omega) \in [0,T]$ and $\omega \in \Omega$ which satisfy the following conditions.

(A5.2.4):

i) for each fixed $\omega \in \Omega, x(t;\omega)$ is absolutely continuous in $[0,T]$;
ii) $u_i(t)$, $i = 1,2,\ldots,N$, are measurable functions for $t \in [0,T]$;
iii) for each $\omega \in \Omega$ and $t \in [0,T]$, $(t,x(t;\omega)) \in A_\Omega$;
iv) $u_i(t) \in U_i(t) \subset E^{P_i}$ for $t \in [0,T]$;
v) $x(0;\omega) = x_0(\omega)$;
vi) $t_f(\omega) \in \tau$;
vii) for each $\omega \in \Omega$, $(t_f(\omega), x(t_f(\omega);\omega)) \in B_\Omega$;
and
viii) the ordinary differential equation

$$\frac{d}{dt}x(t;\omega) = f(t,x(t;\omega),u_1(t),\ldots,u_N(t))$$

is satisfied μ-a.e. in $[0,T]$.

Definition 5.2.3. We shall define an *admissible* $(N+1)$-*tuple* as any $(N+1)$-tuple $(x(t;\omega),u_1(t),\ldots,u_N(t))$ which satisfies conditions (A5.2.4). Also, $x(t;\omega)$ will be called a *random solution*; and $u_1(t),\ldots,u_N(t)$ will be called admissible controls.
Let, K_i, $i = 1,2,\ldots,N$, be continuous functionals defined on the set W of continuous functions given by,

$$W = \left\{(w_1(\omega),w_2(\omega)) : (t_f(\omega),w_2(\omega)) \in B_\omega \text{ for each } \omega \in \Omega\right\};$$

and assume that K_i is bounded from below on W' where

$$W' = \{(w_1(\omega), w_2(\omega)) : (t_f(\omega), w_2(\omega)) \in B_\omega \cap A_\omega\} \subset W.$$

Then, the functionals

$$\begin{aligned} J_i[x, u_i] &= K_i[\eta(x)(\omega)] \\ &= K_i[t_f(\omega),\ x(t_f(\omega);\omega)] \end{aligned}$$

are called *cost functionals*, $i = 1, 2, \ldots, N$. We want to find the absolute minimum of $J_i[x(t;\omega), u(t)]$ in the class ψ for each i.

Definition 5.2.4. Given an admissible $(N+1)$-tuple $(x^*(t;\omega), u^*(t), \ldots, u_N^*(t))$. If

$$J_i[x^*(t;\omega), u_i^*(t)] \leqslant J_i[x(t;\omega), u_i(t)]$$

for all $i = 1, 2, \ldots, N$ and all $(x(t;\omega), u_1(t), \ldots, u_N(t)) \in \psi$, then $(x^*(t;\omega), u_1^*(t), \ldots, u_N^*(t))$ is called an *optimal* $[N+1]$*-tuple*. Similarly, $u_i^*(t)$ are called *optimal controls* and $x(t;\omega)$ is called an *optimal random solution*.

Although the optimal $[N+1]$-tuple need not be unique in ψ, the value of the cost functional, $J_i[x^*, u_i^*]$ is the same for all optimal pairs.

Instead of stating a formal lemma, we shall simply state that given a stochastic N-person differential game as described above where $x(t;\omega)$ is uniformly continuous for $(t, \omega) \in [0, T] \times \Omega$ that $\{x(t;\omega)_k, k = 1, 2, \ldots\}$ forms an equicontinuous and equibounded family of functions on $[0, T] \times \Omega$.

Let us make the following assumptions.

(A5.2.5):

a) Let f is completely separable function. That is,

$$f(t, x, u_1, \ldots, u_N) = g(t, x) + \sum_{i-1}^{N} h_i(t, u_i) \ \text{ for } \ (t, x, u_1, \ldots, u_N) \in M.$$

b) The control sets $U_i(t) \subset E^{P_i}$ are compact for $t \in [0, T]$.

c) The control functions $u_i(t)$ are upper semicontinuous functions of t in $[0, T]$.

d) The following subsets of $L_2(\Omega, \mathrm{A}, \mu)$, given by

$$h_i(t, u_i(t)) = \{y \in L_2(\Omega, \mathrm{A}, \mu) :\ y = h_i(t, u_i),\ u_i \in U_i(t)\}$$

are convex for each $t \in [0, T]$.

Then $f(t, x, U_1(t), \ldots, U_N(t))$ is a convex subset of $L_2(\Omega, \mathrm{A}, \mu)$ for each $(t, x) \in A$.

We shall now state without proof a closure theorem and existence theorem for stochastic N-person differential games with state variables $x(t;\omega) \in C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu))$ and stochastic termination times $t_f(\omega) \in \tau$.

Let us make the following assumptions.

(A5.2.6):

(i) *The constraint sets A_Ω and $A = \bigcup_{\omega \in \Omega} A_\omega$ are compact subsets of the tx-space, given by*
$$\mathbb{R}_+ \times C_c\left(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu)\right).$$

(ii) *The control sets $u_i(t)$ are compact subsets of E^{P_i} for every $i = 1, 2, \ldots, N$ and*
$$t \in [0, T].$$

(iii) *The control functions $u_i(t)$ are upper semi continuous functions of $t \in [0, T]$ for each $i = 1, 2, \ldots, N$.*

(iv) *M_Ω and $M = \bigcup_{\omega \in \Omega} M_\Omega$ are compact subsets of the space*
$$\mathbb{R}_+ \times C_c(\mathbb{R}_+, L_2(\Omega, \mathrm{A}, \mu)) \times E^{\Sigma p_i}.$$

(v) *The function $f(t, x, u_1, \ldots, u_N)$ is a completely separable function on M into $L_2(\Omega, \mathrm{A}, \mu)$.*

(vi) *The sets $h_i(t, U_i(t))$ are convex subsets of $L_2(\Omega, \mathrm{A}, \mu)$ for every $(t, x) \in A$ and $i = 1, 2, \ldots, N$.*

(vii) *$\{x(t;\omega)_k\}$ with $\{t_{fk}(\omega)\} \in \tau$ is a sequence of random solutions satisfying assumptions* (i)–(vi) *and* (vi) *of* (A5.2.4) *and converging uniformly to a function $x^*(t;\omega)$ which is absolutely continuous in $[0, T]$ for each $\omega \in \Omega$.*

and,

(viii) *$t_{fk}(\omega)$ converges uniformly to $t_f(\omega)$.*

Theorem 5.2.3. *Under the assumptions* (A5.2.6)*, there exist measurable controls $u_1^*(t), \ldots, u_N^*(t)$ such that the $[N+1]$-tuple $(x^*(t;\omega), u_1^*(t), \ldots, u_N^*(t))$ with stochastic terminal time $t_f(\omega)$ satisfies condition* (i) *to* (vi) *and* (viii) *of* (A5.2.4)*. That is, $x^*(t;\omega)$ is an admissible random solution with random terminal time $t_f(\omega) \in \tau$ and $J_i\left[x_k, u_i^*\right]$ converges uniformly to $J_i[x^*, u_i^*]$ for each $i = 1, 2, \ldots, N$.*

Theorem 5.2.4. *Consider the stochastic N-person differential game described above and satisfying conditions* (i) *to* (iv) *of* (A5.2.4)*. If the class of admissible $[N+1]$-tuples ψ is nonempty there exists an admissible $(x^*(t;\omega), u_1^*(t), \ldots, u_N^*(t))$ such that*
$$J_i\left[x^*(t;\omega), u_i^*(t)\right] \leqslant J_i\left[x(t;\omega), u_i(t)\right]$$
for all admissible $[N+1]$-tuples and all i, $i = 1, 2, \ldots, N$.

The reader has noted that in the description of the class ψ we required the existence of a finite time T such that $t_f(\omega) \in [0,T]$ for all $\omega \in \Omega$ and such that the random solution of the state equation (7.2.1) exists over the entire interval $[0,T]$. Physically this implies that if we ignore our boundary conditions we can extend the solutions beyond the stopping time $t_f(\omega)$ if $t_f(\omega) < T$. The assumption (F) we require $t_f(\omega) \in \tau$, a family of equicontinuous functions which is closed in the uniform topology. An example would be $t_f(\omega) = T_1$ (constant) for all $\omega \in \Omega$. Thus N-person stochastic differential games of prescribed duration are a special case of the games studied here.

5.3 General solution

Now we will deal with the stochastic differential game problem where N players are simultaneously controlling the evolution of a system. The approach that we are going to use in this section is based on occupation measures as described in Borkar and Ghosh [31]. In this framework the game problem is viewed as a multi decision optimization problem on the set of canonically induced probability measures on the trajectory space by the joint state and action processes. Each of the payoff criteria, such as discounted on the infinite horizon, limiting average, payoff up to an exit time, etc., are associated with the concept of an occupation measure so that the total payoff becomes the integral of some function with respect to this measure. Then the differential game problem reduces to a static game problem on the set of occupation measures, the dynamics of the game being captured in these measures. This set is shown to be compact and convex. A fixed point theorem for point-to-set mapping is used to show the existence of equilibrium in the sense of Nash.

Let V_i, $i = 1,2,\ldots,N$ be compact metric spaces and $U_i = \mathscr{P}(V_i)$ be the space of probability measures on V_i with Prohorov topology. Let $V = V_1 \times V_2 \times \cdots \times V_N$ and $U = U_1 \times U_2 \times \cdots \times U_N$. Let

$$\overline{m}(\cdot,\cdot) = [\overline{m}_1(\cdot,\cdot),\ldots,\overline{m}_d(\cdot,\cdot)]^T : \mathbb{R}^d \times V \to \mathbb{R}$$

and

$$\sigma = [[\sigma_{ij}(\cdot)]],\ \ 1 \leqslant i, j \leqslant d : \mathbb{R}^d \to \mathbb{R}^{d\times d},$$

be bounded continuous maps such that $\overline{m}$ is Lipschitz in its first argument uniformly with respect to the rest and σ is Lipschitz with the least eigenvalue of $\sigma\sigma^T(\cdot)$ be uniformly bounded away from zero. Define, for $x \in \mathbb{R}^d$, $u = (u_1,\ldots,u_N) \in U$, we have

$$m(\cdot,\cdot) = [m_1(\cdot,\cdot),\ldots,m_d(\cdot,\cdot)]^T : \mathbb{R}^d \times U \to \mathbb{R}^d,$$

by

$$m_i(x,u) = \int_{V_N} \cdots \int_{V_1} \overline{m}_i(x,y_1,\ldots,y_N)u_1(dy_1)\cdots u_N(dy_N)$$
$$\doteq \int_V \overline{m}_i(x,y)u(dy)$$

where $y \in V$. Let $x(\cdot)$ be an $\mathbb{R}^d$-valued process given by the following controlled stochastic differential equation of Ito type given by,

$$dx(t) = m(x(t),u(t))\,dt + \sigma(x(t))\,dw(t), t \geqslant 0, \tag{5.3.1}$$

with

$$x(0) = x_0,$$

where, (i) x_0 is a prescribed random variable, (ii) $w(\cdot) = [w_1(\cdot),\ldots,w_d(\cdot)]^T$ is a standard Wiener process independent of x_0, (iii) $u(\cdot) = (u_1(\cdot),\ldots,u_N(\cdot))$, where $u_i(\cdot)$ is a U_i-valued process satisfying : for $t_1 \geqslant t_2 \geqslant t_3$, $w(t_1) - w(t_2)$ is independent of $u(t)$, $t \leqslant t_3$. Such a process $u_i(\cdot)$ will be called an *admissible strategy* for the ith player. If $u_i(\cdot) = v_i(x(\cdot))$ for a measurable $v_i : \mathbb{R}^d \to U_i$, then, $u_i(\cdot)$ is called a *Markov strategy* for the ith player. A strategy $u_i(\cdot)$ is called pure if u_i is a Dirac measure, i.e., $u_i(\cdot) = \delta_{y_i(\cdot)}$, where $y_i(\cdot)$ is a V_i-valued process. If for each $i = 1,\ldots,N$, $u_i(\cdot) = v_i(x(\cdot))$ for some measurable $v_i : \mathbb{R}^d \to U_i$, then, (5.3.1) admits a unique strong solution which is a Feller process, Veretennikov [202]. Let A_i, M_i, $i = 1,2,\ldots,N$, denote the set of arbitrary admissible, respectively Markov strategies for the ith player. An N-tuple of Markov strategies $v = (v_1,\ldots,v_N) \in M$ is called *stable* if the corresponding process is positive recurrent and thus, has a unique invariant measure $\eta(v)$. For any $f \in W^{2,p}_{\text{loc}}(\mathbb{R}^d)$, $p \geqslant 2$, $x \in \mathbb{R}^d$, $u \in V$, let

$$(Lf)(x,u) = \frac{1}{2}\sum_{i,j,k=1}^{d} \sigma_{ik}(x)\sigma_{jk}(x)\frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \overline{m}_i(x,u)\frac{\partial f(x)}{\partial x_i}$$

and for any $v \in U$, we have,

$$(L_v f)(x) = \int_{V_N} \cdots \int_{V_1} (Lf)(x,y)v_1(x)(dy_1)\cdots v_N(x)(dy_N).$$

For an N-tuple $y = (y_1,\ldots,y_N)$, denote $y^{\widehat{k}} = (y_1,\ldots,y_{k-1},y_{k+1},\ldots,y_N)$ and $(y^{\widehat{k}},\widetilde{y}^k) = (y_1,\ldots,y_{k-1},\widetilde{y}_k,y_{k+1},\ldots,y_N)$.

For each $k = 1,\ldots,N$, let $\overline{r}_k : \mathbb{R}^d \times V \to \mathbb{R}$ be bounded continuous functions. When the state is x and actions $v \in V$ are chosen by the players then the player k receives a payoff $\overline{r}_k(x,v)$. For $x \in \mathbb{R}^d$, $u \in U$, let $r_k : \mathbb{R}^d \times U \to \mathbb{R}$ be defined by

$$r_k(x,u) = \int_{V_N} \cdots \int_{V_1} \overline{r}_k(x,y_1,\ldots,y_n)u_1(dy_1)\cdots u_N(dy_N).$$

Each player wants to maximize his accumulated income. We will now consider two evaluation criteria: discounted payoff on the infinite horizon, and ergodic payoff.

5.3.1 *Discounted Payoff on the Infinite Horizon*

Let $\lambda > 0$ be the discount factor and let $u \in A = A_1 \times \cdots \times A_N$. Let $x(\cdot)$ be the solution of (5.3.1) corresponding to u. The discounted payoff to player k for initial condition $x \in \mathbb{R}^d$ is defined by

$$R_\lambda^k[u](x) = E_u\left[\int_0^\infty e^{-\lambda t} r_k(x(t), u(t))dt \mid x(0) = x\right].$$

For an initial law $\pi \in \mathscr{P}(\mathbb{R}^d)$ the payoff is given by

$$R_\lambda^k[u](\pi) = \int_{\mathbb{R}^d} R_\lambda^k[u](x)\pi(dx). \tag{5.3.2}$$

An N-tuple of strategies $u^* = (u_1^*, \ldots, u_N^*) \in A_1 \times \cdots \times A_N$ is said to be a discounted equilibrium (in the sense of Nash) for initial law π if for any $k = 1, \ldots, N$, we have,

$$R_\lambda^k[u^*](\pi) \geqslant R_\lambda^k\left[u^{*\widehat{k}}, u_k\right](\pi), \tag{5.3.3}$$

for any $u_k \in A_k$. The existence of a discounted equilibrium will be shown later.

5.3.2 *Ergodic Payoff*

Let $u \in A$ and let $x(\cdot)$ be the corresponding process with initial law π. The ergodic payoff to player k is given by

$$C_k[u](\pi) = \liminf_{T \to \infty} \frac{1}{T} E_u\left[\int_0^T r_k(x(t), u(t))dt\right]. \tag{5.3.4}$$

The concept of equilibrium for the ergodic criterion is defined similarly. Under a Lyapunov stability condition (assumption (A5.3.1) introduced later) for all $v \in M$ will be stable. For such a v, (5.3.4) equals to

$$\rho_k[v] = \int_{\mathbb{R}^d} r_k(x, v(x))\eta[v](dx), \tag{5.3.5}$$

where $\eta[v] \in \mathscr{P}(\mathbb{R}^d)$ is the invariant measure of the process $x(\cdot)$ governed by v. It will be shown that there exists a $v^* \in M$, such that, for any $k = 1, \ldots, N$, we have,

$$\rho_k[v^*] \geqslant \rho_k\left[v^{*\widehat{k}}, v_k\right],$$

for any $v_k \in M_k$. Thus, v^* will be an ergodic equilibrium. Now we will explain the concept of occupation measures.

5.3.3 *Occupation Measures*

Let

$$M_k = \{v : \mathbb{R}^d \to U_k \mid v \text{ measurable}\}, \quad k = 1, 2, \ldots, N.$$

For $n \geqslant 1$, let Λ_n be the cube of side $2n$ in $\mathbb{R}^d$ with sides parallel to the axes and center at zero. Let B_n denote the closed unit ball of $L_\infty(\Lambda_n)$ with the topology obtained by relativizing to it the weak topology of $L_2(\Lambda_n)$. Then B_n is compact and metrizable, for example by the metric,

$$d_n(f,g) = \sum_{m=1}^{\infty} 2^{-m} \left| \int_{\Lambda_n} f e_m dx - \int_{\Lambda_n} g e_m dx \right|$$

where $\{e_m\}$ is an orthonormal basis of $L_2(\Lambda_n)$. Let $\{f_i\}$ be a countable dense subset of the unit ball of $C(V_k)$. Then, $\{f_i\}$ separates points of U_k. For each $v \in M_k$, define $g_{v_i} : \mathbb{R}^d \to \mathbb{R}$ by

$$g_{v_i}(x) = \int_{V_k} f_i dv(x), \quad i \geqslant 1,$$

and $g_{v_i n}(\cdot)$ denotes the restriction of $g_{v_i}(\cdot)$ to Λ_n, for each n. Define a pseudometric $d_k(\cdot,\cdot)$ on M_k by

$$d_k(v,u) = \sum_{i,n=1}^{\infty} 2^{-(n+1)} \frac{d_n(g_{v_i n}, g_{u_i n})}{[1 + d_n(g_{v_i n}, g_{u_i n})]}.$$

Replacing M_k by its quotient with respect to a.e. equivalence, $d_k(\cdot,\cdot)$ becomes a metric. The following result is given in detail in Borkar [30].

Theorem 5.3.1. *M_k is compact under the metric topology of $d_k(\cdot,\cdot)$. Let $f \in L_2(\mathbb{R}^d)$, $g \in C_b(\mathbb{R}^d \times V_k)$ and $v_n \to v$ in M_k. Then*

$$\int_{\mathbb{R}^d} f(x) \int_{V_k} g(x,\cdot) dv_n dx \to \int_{\mathbb{R}^d} f(x) \int_{V_k} g(x,\cdot) dv dx.$$

Conversely, if the above holds for all such f, g then $v_n \to v$ in M_k.

Endow M with the product topology of M_k. Let $v \in M$ and $x(\cdot)$ be the process governed by v with a fixed initial law. Let $L(v)$ denote the law of $x(\cdot)$.

Theorem 5.3.2. *The map $v \to L(v) : M \to \mathscr{P}(C[0,\infty); \mathbb{R}^d)$ is componentwise continuous, i.e., for each $k = 1, 2, \ldots, N$, if $v_k^n \to v_k^\infty$ in M_k, and $v_i \in M_i$, $i \neq k$, then $L(v^{\widehat{k}}, v_k^n) \to L(v^{\widehat{k}}, v_k^\infty)$ in $\mathscr{P}(C[0,\infty); \mathbb{R}^d)$.*

Now, we will introduce occupation measures for both discounted and ergodic payoff criterion. First consider the discounted case. Let $u \in A$ and $x(\cdot)$ be the corresponding process. The *discounted occupation measure* for initial condition $x \in \mathbb{R}^d$ denoted by $\nu_{\lambda x}[u] \in \mathscr{P}(\mathbb{R}^d \times V)$ is defined by,

$$\int_{\mathbb{R}^d \times V} f d\nu_{\lambda x}[u] =$$

$$\lambda^{-1} E_u \left[\int_0^\infty \int_{V_N} \cdots \int_{V_1} e^{-\lambda t} f(x(t), y_1, \ldots, y_N) u_1(t)(dy_1) \cdots u_N(t)(dy_N) dt \,\middle|\, x_0 = x \right]$$

for $f \in C_b(\mathbb{R}^d \times V)$ and for an initial law $\pi \in \mathscr{P}(\mathbb{R}^d)$, $\nu_{\lambda\pi}[u]$ is defined by

$$\int f d\nu_{\lambda\pi}[u] = \int_{\mathbb{R}^d} \pi(dx) \int_{\mathbb{R}^d \times V} f d\nu_{\lambda x}[u].$$

In terms of $\nu_{\lambda\pi}[u]$, (5.2.2) becomes,

$$R_\lambda^k[u](\pi) = \lambda \int \bar{r} d\nu_{\lambda x}[u].$$

Let

$$\nu_{\lambda\pi}[A] = \{\nu_{\lambda\pi}[u] | u \in A\},$$

$\nu_{\lambda\pi}[M_1, A_2, \ldots, A_N]$, $\nu_{\lambda\pi}[M_1, \ldots, M_N]$ are defined analogously. Then, from Borkar and Ghosh [31] we have the following result.

Theorem 5.3.3. *For any $k = 1, 2, \ldots, N$,*

$$\nu_{\lambda\pi}[M_1, \ldots, M_{k-1}, A_k, M_{k+1}, \ldots, M_N] = \nu_{\lambda\pi}[M_1, \ldots, M_N].$$

Let $v \in M$. By Krylov's inequality it can be shown that $\nu_{\lambda\pi}[v]$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ and hence has a density $\phi_{\lambda\pi}[v]$. For $f \in W^{2,p}_{\text{loc}}(\mathbb{R}^d)$ define

$$L_v^\lambda f(x) = (L_v f)(x) - \lambda f(x).$$

Then, $\phi_{\lambda\pi}[v]$ is the unique solution in $L_1(\mathbb{R}^d)$ and for every $f \in C_0^\infty(\mathbb{R}^d)$, we have

$$\int L_v^\lambda f(x) \phi(x) dx = -\int f(x) \pi(dx)$$

and

$$\int \phi(x) dx = 1, \quad \phi \geqslant 0.$$

Now from Borkar and Ghosh [31] we have following results.

Lemma 5.3.1. *Let $\nu_{\lambda\pi}[M_1, \ldots, M_N]$ be componentwise convex, i.e., for any fixed k and prescribed $v_i \in M_i$, $i \neq k$, we have*

$$\nu_{\lambda\pi}\left[v^{\widehat{k}}, M_k\right] = \left\{\nu_{\lambda\pi}\left[v^{\widehat{k}}, v_k\right] : v_k \in M_k\right\},$$

is convex.

Also, we have the following result.

Lemma 5.3.2. *Let $\nu_{\lambda\pi}[M_1,\dots,M_N]$ be componentwise compact, i.e., for any fixed k and prescribed $v_i \in M_i$, $i \neq k$, we have*

$$\nu_{\lambda\pi}\left[v^{\widehat{k}}, M_k\right] = \left\{\nu_{\lambda\pi}\left[v^{\widehat{k}}, v_k\right] : v_k \in M_k\right\},$$

is compact.

For the ergodic payoff criterion we will impose the following Lyapunov type stability condition.

(A5.3.1): There exists a twice continuously differentiable function $w : \mathbb{R}^d \to \mathbb{R}_+$, such that,

(i) $\lim_{\|x\|\to\infty} w(x) = \infty$ uniformly in $\|x\|$.

(ii) There exist $a > 0$, $\varepsilon_0 > 0$ such that for $\|x\| > a$,

$$Lw(x,u) < -\varepsilon_0 \text{ for all } u \in V$$

and

$$\|\nabla w\|^2 \geqslant (\overline{\lambda})^{-1},$$

where $\overline{\lambda}$ is the ellipticity constant of $\sigma\sigma^T$.

(iii) $w(x)$ and $\|\nabla w\|$ have polynomial growth.

For $v \in M$, let $x(\cdot)$ be the corresponding process. Also, for $\|x\| > a$, let

$$\tau_a = \inf\{t \geqslant 0 \mid \|x(t)\| = a\}.$$

The following result is a consequence of Assumption (A5.2.1).

Lemma 5.3.3.

(i) *All $v \in M$ are stable.*

(ii) $E_v[\tau_a \mid x(0) = x] \leqslant w(x)/\varepsilon_0$, *for any v.*

(iii) $\int w(x)\eta[v](dx) < \infty$ *for any v.*

(iv) *Under any v and $x \in \mathbb{R}^d$, with*

$$\lim_{t\to\infty} \frac{1}{t} E_v[w(x(t))] = 0.$$

and

(v) *The set $I = \{\eta[v] \mid v \in M\}$ is componentwise compact in $\mathscr{P}(\mathbb{R}^d)$.*

For $v \in M$, the ergodic occupation measure, denoted by $\nu_E[v] \in \mathscr{P}(\mathbb{R}^d \times V)$ is defined as

$$\nu_E[v](dx, dy_1, \ldots, dy_N) = \eta[v] \prod_{i=1}^{N} v_i(x)(dy_i).$$

Let

$$\nu_E[M] = \{\nu_E[v] | v \in M\}.$$

For $v \in M$, let $x(\cdot)$ be the process governed by v. Then,

$$\eta[v](dx) = \left(\int p(t, y, x) \eta[v](dy) \right) dx,$$

where $p(\cdot, \cdot, \cdot)$ is the transition density of $x(\cdot)$ under v. Thus, $\eta[v]$ itself has a density which we denote by $\varphi[v](\cdot)$. Then $\varphi[v]$ is the unique solution of the following. For every $f \in C_0^\infty(\mathbb{R}^d)$

$$\int L_v f(x) \phi(x) dx = 0$$

$$\int \phi(x) dx = 1, \quad \phi \geqslant 0.$$

As for the discounted case, we now have the following results.

Lemma 5.3.4. $\nu_E[M]$ *is componentwise convex and compact.*

For any fixed $k \in \{1, 2, \ldots, N\}$, let $v_i \in M_i$, $i \neq k$ and $u_k \in A_k$. Let $x(\cdot)$ be the process governed by $\left(\widehat{v^k}, u_k\right)$. Define $\mathscr{P}(\mathbb{R}^d \times V)$-valued empirical process ν_t as follows: For $B \subset \mathbb{R}^d$, $A_i \subset U_i$, $i = 1, \ldots, N$, Borel, and

$$\nu_t(B \times A_1 \times \cdots \times A_N) = \frac{1}{t} \int_0^t I\{x(s) \in B\} \prod_{\substack{i=1 \\ i \neq k}}^{N} v_i(x(s))(A_i) u_k(s)(A_k) ds.$$

Lemma 5.3.5. *The process* $\{\nu_t\}$ *is a.s., tight and outside a set of zero probability, each limit point* ν *of* $\{\nu_t\}$ *as* $t \to \infty$ *belongs to* $\nu_E[M]$.

5.3.4 *Existence of an Equilibrium*

We will proceed by making the following assumption.

(A5.3.2): $\overline{m}$ and $\overline{r}$ are of the form

$$\overline{m}(x, u_1, \ldots, u_N) = \sum_{i=1}^{N} \overline{m}^i(x, u_i)$$

and

$$\overline{r}(x, u_1, \ldots, u_N) = \sum_{i=1}^{N} \overline{r}_i(x, u_i)$$

where $\overline{m}^i : \mathbb{R}^d \times V_i \to \mathbb{R}^d$ and $\overline{r}_i : \mathbb{R}^d \times V_i \to \mathbb{R}$ and they satisfy the same conditions as $\overline{m}$ and $\overline{r}$.

Let $v \in M$. Fix a $k \in \{1,2,\ldots,N\}$ and $\pi \in \mathscr{P}(\mathbb{R}^d)$. Then by Lemma 5.2.3, we have

$$\sup_{u_k \in A_k} R_\lambda^k\left[v^{\widehat{k}}, u_k\right](\pi) = \sup_{\overline{v}_k \in M_k} R_\lambda^k\left[v^{\widehat{k}}, \overline{v}_k\right](\pi).$$

Since M_k is compact and $\overline{r}_k$ is continuous, the suprimum on the right hand side above can be replaced by the maximum. Then, there exists a $v_k^* \in M_k$, such that,

$$\sup_{u_k \in A_k} R_\lambda^k\left[v^{\widehat{k}}, u_k\right](\pi) = \max_{\overline{v}_k \in M_k} R_\lambda^k\left[v^{\widehat{k}}, \overline{v}_k\right](\pi) = R_\lambda^k\left[v^{\widehat{k}}, v_k^*\right](\pi). \tag{5.3.6}$$

This optimal discounted response strategy for player k, v_k^* can be chosen to be independent of π. Define $\widetilde{R}_\lambda^k[v] : \mathbb{R}^d \to \mathbb{R}$ by

$$\widetilde{R}_\lambda^k[v](x) = \max_{\overline{v}_k \in M_k} R_\lambda^k\left[v^{\widehat{k}}, \overline{v}_k\right](x).$$

Then, we can obtain the following result.

Lemma 5.3.6. *$\widetilde{R}_\lambda^k[v](\cdot)$ is the unique solution in $W^{2,p}_{\mathrm{loc}}(\mathbb{R}^d) \cap C_b(\mathbb{R}^d)$, $2 \leqslant p < \infty$, of*

$$\lambda\phi(x) = \sup_{\overline{v}_k}\left[L_{v^{\widehat{k}}, \overline{v}_k}\phi(x) + r\left(x, v^{\widehat{k}}(x), \overline{v}_k\right)\right]$$

in $\mathbb{R}^d$. A strategy $v_k^* \in M_k$ is discounted optimal response for player k given v if and only if

$$\begin{aligned}&\left[\sum_{i=1}^d m_i\left(x, v^{\widehat{k}}(x), v_k^*(x)\right)\frac{\partial \widetilde{R}_\lambda^k[v](x)}{\partial x_i} + r\left(x, v^{\widehat{k}}(x), v_k^*(x)\right)\right]\\ &= \sup_{v_k}\left[\sum_{i=1}^d m_i\left(x, v^{\widehat{k}}(x), v_k(x)\right)\frac{\partial \widetilde{R}_\lambda^k[v](x)}{\partial x_i} + r\left(x, v^{\widehat{k}}(x), v_k(x)\right)\right] \ \ a.e..\end{aligned}$$

Details of the next result can be found in Borkar and Ghosh [31] gives the existence of discounted equilibrium in the set of Markov strategies.

Theorem 5.3.4. *There exists a discounted equilibrium $v^* = (v_1^*, \ldots, v_N^*) \in M$.*

Proof. Let $v \in M$ and $\overline{v}_k \in U_k$. Set

$$F_k\left(x, v^{\widehat{k}}, \overline{v}_k\right) = \sum_{i=1}^d m_i\left(x, v^{\widehat{k}}(x), \overline{v}_k\right)\frac{\partial \widetilde{R}_\lambda^k[v](x)}{\partial x_i} + r\left(x, v^{\widehat{k}}(x), \overline{v}_k\right)$$

Let

$$G_k[v] = \left\{v_k^* \in M_k \mid F_k\left(x, v^{\widehat{k}}(x), v_k^*(x)\right) = \sup_{\overline{v}_k \in U_k} F_k\left(x, v^{\widehat{k}}, \overline{v}_k\right) \ \text{ a.e.}\right\}.$$

Then, $G_k[v]$ is non-empty, convex, closed and hence compact. Set

$$G[v] = \prod_{k=1}^N G_k[v].$$

Then, $G[v]$ is non-empty convex and compact subset of M. Thus, $v \to G[v]$ defines a point-to-set map from M to 2^M. This map is upper semicontinuous. Hence, by Fan's fixed point theorem [58], there exists a $v^* \in M$, such that, $v^* \in G[v^*]$. This v^* is a discounted equilibrium. □

Next we will discuss the existence results for the ergodic payoff. Let $v \in M$ and fix a $k \in \{1,2,\dots,N\}$. Let $v_k^* \in M_k$ be, such that,

$$\rho_k^*[v] \doteq \rho_k\left[v^{\widehat{k}}, v_k^*\right] = \max_{\overline{v}_k \in M_k} \rho v^{\widehat{k}}, v_k^*],$$

where $\rho_k[v]$ is defined in (5.3.5). If all but player k uses strategies $v^{\widehat{k}}$ then, by Lemma 5.3.8, player k can not obtain a higher payoff than $\rho_k^*[v]$ by going beyond M_k a.s. This v_k^* is said to be an *ergodic optimal* response for player k given v. Consider the following

$$\rho = \sup_{v^{\widehat{k}}, \overline{v}_k} \left[L\phi(x) + r\left(x, v^{\widehat{k}}(x), \overline{v}_k\right)\right] \tag{5.3.7}$$

where ρ is a scalar and $\phi : \mathbb{R}^d \to \mathbb{R}$. Then we have the following result.

Lemma 5.3.7. *The equation (5.3.7) has a uniqu e solution $(\phi_k[v], \rho_k^*[v])$ in the class of functions $W_{loc}^{2,p}(\mathbb{R}^d) \cap O(w(\cdot))$, $2 \leqslant p < \infty$, satisfying $\phi[v] = 0$. A $v_k^* \in M_k$ is ergodic optimal response for player k given v, if and only if,*

$$\begin{aligned}
&\left[\sum_{i=1}^{d} m_i\left(x, v^{\widehat{k}}(x), v_k^*(x)\right) \frac{\partial \phi_k[v](x)}{\partial x_i} + r\left(x, v^{\widehat{k}}(x), v_k^*(x)\right)\right] \\
&= \sup_{v_k} \left[\sum_{i=1}^{d} m_i\left(x, v^{\widehat{k}}(x), v_k(x)\right) \frac{\partial \phi_k[v](x)}{\partial x_i} + r\left(x, v^{\widehat{k}}(x), v_k(x)\right)\right] \quad a.e..
\end{aligned} \tag{5.3.8}$$

The following result is due to Borkar and Ghosh [31] that gives the existence of an ergodic equilibrium.

Theorem 5.3.5. *There exists an ergodic equilibrium $v^* \in M$.*

Proof. Let $v \in M$ and $\overline{v}_k \in U_k$. Set

$$J_k\left(x, v^{\widehat{k}}, \overline{v}_k\right) = \left[\sum_{i=1}^{d} m_i\left(x, v^{\widehat{k}}(x), \overline{v}_k(x)\right) \frac{\partial \phi_k[v](x)}{\partial x_i} + r\left(x, v^{\widehat{k}}(x), \overline{v}_k(x)\right)\right].$$

Let

$$H_k(v) = \left\{\widetilde{v}_k \in M_k \mid J_k\left(x, v^{\widehat{k}}, \widetilde{v}_k(x)\right) = \sup_{\overline{v}_k \in U_k} J_k\left(x, v^{\widehat{k}}, \overline{v}_k\right) \text{ a.e.}\right\}.$$

Set $H[v] = \prod_{k=1}^{N} H_k(v)$. Then $H(v)$ is a non-empty, convex, compact subset of M. As in the discounted case, an application of Fan's fixed point theorem yields a $v^* \in M$ such that $v^* \in H[v^*]$. This v^* is an ergodic equilibrium. □

In this section we have used a non-anticipative relaxed control framework to show the existence of an equilibrium for N-person stochastc differential game. Using this approach, one could also show the existence of value and optimal strategies for two person strictly competitive differential game that we have discussed in Section 2. Other payoff criteria could also be considered. Using the approach described here, one could obtain similar results for feedback randomized strategies.

Chapter 6

Weak Convergence in Two Player Stochastic Differential Games

6.1 Introduction

Much of stochastic game theory is concerned with diffusion models, as we have seen in Chapters 3 through 5. It is well known that such models are often only idealizations of the actual physical processes, which might be driven by a wide bandwidth process or be a discrete parameter system with correlated driving noises. Optimal strategies derived for the diffusion models would not be of much interest if they were not "nearly optimal" for the physical system which the diffusion approximates. Using the models of this Chapter, for many typical problem formulations, we can show that the optimal strategies derived for the "limit" system are also good strategies for the system which is driven by wide bandwidth noise processes. Such results not only gives robustness statement on the game problem, but also substantially simplifies the computational aspects, as we will see in Chapter 8. The results in this chapter will show that we need only to compute the value of the limiting system and proceed to obtain corresponding strategies and adapt these strategies in the n^{th} (actual) system instead of computing the saddle points or optimal strategies at each step and show the convergence.

In Section 6.2, we will briefly explain weak convergence preliminaries. For details on the weak convergence, we refer the reader to the excellent book by Kushner, [103]. Weak convergence methods for some popular payoff structures will be discussed in Section 6.3. Other extentions such as the multi mode case will be described in Section 6.4 and the partially observed case will be discribed in Section 6.5. Some deterministic approximations will be discussed in Section 6.6.

6.2 Weak Convergence Preliminaries

Let $D^d[0,\infty)$ denote the space of $\mathbb{R}^d$ valued functions which are right continuous and have left-hand limits endowed with the Skorohod topology. Following Kurtz [116], Kushner [103], we define the notion of "p-lim" and an operator $\widehat{A}^\varepsilon$ as follows. Let $\{\Im_t^\varepsilon\}$ denote the minimal σ-algebra over which $\{x^\varepsilon(s), \xi^\varepsilon(s),\ s \leqslant t\}$ is measurable, and let E_t^ε denote the expectation conditioned on $\Im_t^\varepsilon$. Let $\widetilde{M}$ denote the set of real valued functions of (ω, t) that are nonzero only on a bounded t-interval. Let

$$\overline{M}^\varepsilon = \left\{ f \in \widetilde{M};\, \sup_t E|f(t)| < \infty \text{ and } f(t) \text{ is } \Im_t^\varepsilon \text{ measurable} \right\}.$$

Let $f(\cdot), f^\Delta(\cdot) \in \overline{M}^\varepsilon$, for each $\Delta > 0$. Then $f = p\text{-}\lim_\Delta f^\Delta$, if and only if,

$$\sup_{t,\Delta} E\left| f^\Delta(t) \right| < \infty,$$

and $\lim_{\Delta\to 0} E|f(t) - f^\Delta(t)| = 0$, for each t. $f(\cdot)$ is said to be in the domain of $\widehat{A}^\varepsilon$, i.e., $f(\cdot) \in D\big(\widehat{A}^\varepsilon\big)$, and $\widehat{A}^\varepsilon f = g$ if

$$p\text{-}\lim_{\Delta\to 0} \left(\frac{E_t^\varepsilon f(t+\Delta) - f(t)}{\Delta} - g(t) \right) = 0.$$

If $f(\cdot) \in D\big(\widehat{A}^\varepsilon\big)$, then

$$f(t) - \int_0^t \widehat{A}^\varepsilon f(u)du \quad \text{is a martingale,}$$

and

$$E_t^\varepsilon f(t+s) - f(t) = \int_t^{t+s} E_t^\varepsilon \widehat{A}^\varepsilon f(u)du, \quad \text{w.p.1.}$$

The $\widehat{A}^\varepsilon$ operator plays the role of an infinitesimal operator for a non-Markov process. In our case, it becomes a differential operator by the martingale property and the definition of p-limit. We will use the terms like "tight", Skorohod imbedding, etc. with out explanation, the reader can obtain these from Kushner [103]. The following result will be used to conclude that various terms will go to zero in probability.

Note: If there is a strategy vector u involved, we can define $\widehat{\mathscr{A}^u}$ in the following manner. Let

$$b(x,u) = \sum_{i=1}^N b_i(x(t))u_i(t).$$

Define the operator $\widehat{\mathscr{A}^u}$ as follows,

$$\widehat{\mathscr{A}^u} f(x) = f_x(x)[\overline{a}(x) + b(x,u)].$$

Lemma 6.2.1. *Let $\xi(\cdot)$ be a ϕ-mixing process with mixing rate $\phi(\cdot)$, and let $h(\cdot)$ be a function of ξ which is bounded and measurable on $\Im_t^\infty$. Then, there exist K_i, $i = 1,2,3$, such that,*

$$|E(h(t+s)/\Im_0^t) - Eh(t+s)| \leqslant K_1\phi(s).$$

If $t < u < v$, and $Eh(s) = 0$ for all s, then,

$$|E(h(u)h(v)/\Im_\tau^t) - Eh(u)h(v)| \leqslant \begin{cases} K_2\phi(v-u), & u < \tau < v \\ K_3\phi(u-t), & t < \tau < u, \end{cases}$$

where $\Im_\tau^t = \sigma\{\xi(s); \tau \leqslant s \leqslant t\}$.

In order to obtain the weak convergence result, the following condition need to be verified:

$$\lim_{n\to\infty} \limsup_{\varepsilon\to 0} P\left(\sup_{t\leqslant T} |x^\varepsilon(t)| \geqslant n\right) = 0$$

for each $T < \infty$. Direct verification of this is very tenuous. Instead, one can utilize the method of K-truncation. This is as follows. For each $K > 0$, let

$$S_K = \{x : |x| \leqslant K\} \quad \text{be the } K\text{-ball.}$$

Let $x^{\varepsilon,K}(0) = x^\varepsilon(0)$, $x^{\varepsilon,K}(t) = x^\varepsilon(t)$, up until the first exit from S_k, and

$$\lim_{n\to\infty} \limsup_{\varepsilon\to 0} P\left(\sup_{t\leqslant T} |x^{\varepsilon,K}(t)| \geqslant n\right) = 0 \ \text{ for each } \ T < \infty.$$

Thus, $x^{\varepsilon,K}(t)$ is said to be the K-truncation of $x^\varepsilon(\cdot)$. Let

$$q^K(x) = \begin{cases} 1, & \text{for } x \in S_K \\ 0, & \text{for } x \in \mathbb{R}^d - S_{K+1} \\ \text{Smooth} & \text{otherwise.} \end{cases}$$

Define $a_K(x,\alpha) = a(x,\alpha)q^K(x)$ and $g_K(x,\xi) = g(x,\xi)q^K(x)$. Let $x^{\varepsilon,K}(\cdot)$ denote the process corresponding to the use of truncated coefficients. Then $x^{\varepsilon,K}(\cdot)$ is bounded uniformly in t and $\varepsilon > 0$.

To prove the main weak convergence results, we will use the following results from Kushner [103].

Lemma 6.2.2. *Let $\{y^\varepsilon(\cdot)\}$ be tight on $D^d[0,\infty)$. Suppose that for each $f(\cdot) \in C_0^3$, and each $T < \infty$, there exist $f^\varepsilon(\cdot) \in D(\widehat{A}^\varepsilon)$, such that,*

$$p\text{-}\lim\left(f^\varepsilon(\cdot) - f(y^\varepsilon(\cdot))\right) = 0 \tag{6.2.1}$$

and

$$p\text{-}\lim_{\varepsilon}\left(\widehat{A}^\varepsilon f^\varepsilon(\cdot) - \widehat{A}f(y^\varepsilon(\cdot))\right) = 0. \tag{6.2.2}$$

Then $y^\varepsilon(\cdot) \to y(\cdot)$, the solution of the martingale problem for the operator $\widehat{A}$.

Lemma 6.2.3. *Let the K-truncations $\{y^{\varepsilon,K}\}$ be tight for each K, and that the martingale problem for the diffusion operator A have a unique solution $y(\cdot)$ for each initial condition. Suppose that $y^K(\cdot)$ is a $K-$ truncation of $y(\cdot)$ and it solves the martingale problem for operator A^K. For each K and $f(\cdot) \in D$, let there be $f^\varepsilon(\cdot) \in D(A^\varepsilon)$ such that* (6.1.1) *and* (6.1.2) *hold with $y^{\varepsilon,K}(\cdot)$ and A^K replacing y^ε and A, respectively. Then $y^\varepsilon(\cdot) \to y(\cdot)$.*

Now we will outline a general method one can follow to show that a sequence of solutions to a wide band width noise driven ordinary differential equation (ODE) converge weakly to a diffusion, and identify the limit diffusion (Kushner [103], Ramachandran [158]). Let $z^\varepsilon(\cdot)$ be defined by

$$dz^\varepsilon = a(z^\varepsilon)dt + \frac{1}{\varepsilon}b(z^\varepsilon)\xi(t/\varepsilon^2)dt \tag{6.2.3}$$

where $\xi(\cdot)$ is a second order stationary right continuous process with left hand limits and integrable correlation function $\overline{R}(\cdot)$, and the functions $a(\cdot)$ and $b(\cdot)$ are continuous, $b(\cdot)$ is continuously differentiable and (6.2.3) has a unique solution. Define $\overline{R_0} = \int_{-\infty}^{\infty} E\xi(u)\xi'(0)du$ and assume that

$$E\left|\int_s^t du\left[E\left(\xi(u)\xi'(s)/\xi(\iota), \iota \geqslant 0\right) - \overline{R}(u-s)\right]\right| \to 0 \quad \text{as } t, s \to \infty.$$

Define the infinitesimal generator A and function $\overline{K} = \left(\overline{K_1}, \ldots\right)$ by

$$\begin{aligned} Af(z) &= f_z'(z)a(z) + \int_0^\infty E\left[f_z'(z)b(z)\xi(t)\right]_z' b(z)\xi(0)dt \\ &\equiv \sum_i f_{z_i}(z)\overline{K_i}(z) + \frac{1}{2}\operatorname{trace}\left\{f_{z_i z_j}(z)\right\}\left\{b(z)\overline{R_0}b(z)\right\}, \end{aligned} \tag{6.2.4}$$

where $\overline{K} = (\overline{K}_1, \ldots)$ are the coefficients of the first derivatives $(f_{z_1}, \ldots)$ in (6.1.4). The operator A is the generator of

$$dz = \overline{K}(z)dt + b(z)\overline{R}_0^{1/2}dw, \tag{6.2.5}$$

where $w(\cdot)$ is the standard Wiener process. In order to obtain that $z^\varepsilon(\cdot) \to z(\cdot)$ of (6.2.5), by martingale problem solution, it is enough to show that

$$p\text{-}\lim_{\varepsilon}\left(\widehat{A}^\varepsilon f^\varepsilon(\cdot) - Af(z^\varepsilon(\cdot))\right) = 0. \tag{6.2.6}$$

Then by Lemma 6.2.2, $z(\cdot)$ satisfies (6.2.5).

6.3 Some Popular Payoff Structures

In this section, we will discuss weak convergence methods for both average cost per unit time problem as well as the discounted payoff problem.

6.3.1 *Ergodic Payoff*

The average cost per unit time problem over an infinite time horizon for two person zero-sum stochastic differential games with diffusion model have been dealt with in the literature. For the diffusion models where payoff with expectations (not pathwise), existence of equilibrium has been proven in (Elliott and Davis [51]) and in the case of discounted and average cost cases the existence of equilibria in Markov strategies was established in Borkar and Ghosh [31]. We treat such a problem for wideband noise driven systems, which are 'close' to diffusion. The average is in the pathwise but not necessarily in the expected value sense (Ramachandran [158]). The 'pathwise' convergence result is of particular importance in applications, since we often have a single realization, then expectation is not appropriate in the cost function. In a typical application, we have a particular process with a wideband noise driving forces. Our interest is in knowing how well are the good policies for the 'limit' problem do for the actual 'physical', problem as well as various qualitative properties of the 'physical' process. Physical problem is better modeled by a wideband width noise driven process than the white noise process. However, owing to the wideband noise and appearance of the two parameters ε and T, convergence results of the 'almost sure' type are often rather meaningless from a practical point of view as well as nearly impossible to obtain. It is important that the convergence result obtained should not depend on the way in which $\varepsilon \to 0$ and $T \to \infty$. Where this is not the case, it would be possible that as $\varepsilon \to 0$, a larger and larger T is needed to closely approximate the limit value. In that case, the white noise limit (6.3.1) would not be useful for predictive or control purposes when the true model is given by (6.3.7). It will be shown that the optimal equilibrium policies of the limit diffusion when applied to the wide bandwidth processes, will be δ-equilibrium as the parameters $\varepsilon \to 0$ and $T \to \infty$, irrespective of the order in which the limit takes place. It is also shown that the δ-optimal pathwise discounted payoffs converge to the δ-equilibrium as both the discounted factor $\lambda \to 0$ and bandwidth goes to ∞. Apart from the fact that this gives a robustness statement for the diffusion model, one of the major advantage is by using the method of this work, it is enough to compute the optimal strategies for the limit diffusion and then use this strategies to the physical system in order to obtain near optimal strategies. The entire problem will be set in relaxed control framework. In the proofs, we will use the weak convergence theory.

6.3.2 *Problem Description*

Let the diffusion model be given in a non-anticipative relaxed control frame work. Let U_i, $i = 1,2$ be compact metric spaces (we can take U_i as compact subsets of $\mathbb{R}^d$), and $M_i = P(U_i)$, the space of probability measures on U_i with Prohorov topology.
For $m = (m_1, m_2) \in M = M_1 \times M_2$ and $U = U_1 \times U_2$, $x(\cdot) \in \mathbb{R}^d$ be an $\mathbb{R}^d$-valued process given by the following controlled stochastic differential equation

$$dx(t) = \int_{U_1} a_1(x(t),\alpha_1)m_{1t}(d\alpha) + \int_{U_2} a_2(x(t),\alpha_2)m_{2t}(d\alpha)dt + \overline{g}(x(t))dt + \sigma(x(t))dw(t)$$
$$x(0) = x_0 \tag{6.3.1}$$

where x_0 is a prescribed random variable. The pathwise average payoff per unit time for player 1 is given by

$$J[m](x) = \liminf_{T\to\infty} \frac{1}{T}\int_0^T \int r(x(s),\alpha)m_s(d\alpha)ds \tag{6.3.2}$$

and for the initial law π in $P(\mathbb{R}^d)$, it is given by

$$J[m](\pi) = \int_{\mathbb{R}^d} J[m](x)\pi(dX). \tag{6.3.3}$$

Let $w(\cdot)$ in (6.3.1) be a Wiener process with respect to a filtration $\{\Im_t\}$ and let Ω_i, $i = 1,2$ be a compact set in some Euclidean space. A measure valued random variable $m_i(\cdot)$ is an *admissible strategy* for the i^{th} player if $\int\int_0^t f_i(s,\alpha_i)m_i(dsd\alpha_i)$ is progressively measurable for each bounded continuous $f_i(\cdot)$ and $m_i([0,t]\times\Omega_i) = t$, for $t \geqslant 0$. If $m_i(\cdot)$ is admissible then there is a derivative $m_{it}(\cdot)$ (defined for almost all t) that is non-anticipative with respect to $w(\cdot)$ and

$$\int_0^t \int f_i(s,\alpha_i)m_i(dsd\alpha_i) = \int_0^t ds \int f_i(s,\alpha_i)m_{is}(d\alpha_i),$$

for all t with probability one (w.p.1). The results derived in this work are for the so called *Markov strategies*, which is a measure on the Borel sets of Ω_i for each x, and $m_i(c)$ is Borel measurable for each Borel measurable set C. We will denote by A_i the set of admissible strategies and M_{ai} the set of Markov strategies for the player i. One can introduce appropriate metric topology under which M_{ai} is compact, Borkar and Ghosh [31].
In relaxed control settings, one chooses at time t, a probability measure m_t on the control set M rather than an element $u(t)$ in U. We call the measure m_t the *relaxed control* at time t. Any ordinary control can be represented as a relaxed control via the definition of the derivative $m_t(d\alpha) = \delta_{u(t)}(\alpha)d\alpha$. Hence, if m_t is an atomic measure concentrated at a single point $m(t) \in M$ for each t, then the relaxed control will be called ordinary control. We will denote the ordinary control by $u_m(t) \in M$.

An admissible strategy $m_1^* \in A_1$ is said to be an *ergodic optimal* for initial law π if

$$J[m_1^*, \widetilde{m}_2](\pi) \geqslant \inf_{m_2 \in A_2} \sup_{m_1 \in A_1} J[m_1, m_2](\pi) = V^+(\pi), \tag{6.3.4}$$

for any $\widetilde{m}_2 \in A_2$. A strategy $m_1^* \in M_{a_1}$ is called discounted optimal for player I, if it is ergodic optimal for all initial laws. Similarly, $m_2^* \in A_2$ is discounted optimal for player II for an initial law π if

$$\begin{aligned} J(\widetilde{m}_1, m_2^*)(\pi) &\leqslant \sup_{m_1 \in A_1} \inf_{m_2 \in A_2} J[m_1, m_2](\pi) \\ &= V^-(\pi), \end{aligned} \tag{6.3.5}$$

for any $\widetilde{m}_1 \in A_1$. $m_2^* \in M_{a_2}$ is ergodic optimal for player II if (6.3.5) holds for all initial laws. If for any initial law π, $V^+(\pi) = V^-(\pi)$, then the game is said to have an ergodic equilibrium and we will denote it by $V(\pi)$. The policies $m_{1\delta}$ and $m_{2\delta}$ are said to be δ-*ergodic equilibrium* if

$$\sup_{m_1 \in A_1} J(m_1, m_{2\delta}) - \delta \leqslant V \leqslant \inf_{m_2 \in A_2} J(m_{1\delta}, m_2) + \delta. \tag{6.3.6}$$

The wide band noise system considered in this work is of the following type:

$$\begin{aligned} dx^\varepsilon = \Big[\int a_1(x^\varepsilon, \alpha_1) m_{1t}^\varepsilon(d\alpha_1) + \int a_2(x^\varepsilon, \alpha_2) m_{2t}^\varepsilon(d\alpha_2) dt + G(x^\varepsilon, \xi^\varepsilon(t)) \\ + \tfrac{1}{\varepsilon} g(x^\varepsilon, \xi^\varepsilon) dt\Big], \end{aligned} \tag{6.3.7}$$

and pathwise average payoff per unit time for player k is given by

$$J^\varepsilon[m^\varepsilon] = \liminf_{T \to \infty} \frac{1}{T} \int_0^T \int r(x^\varepsilon(s), \alpha) m_s^\varepsilon(d\alpha) ds. \tag{6.3.8}$$

Player I aims to maximize his accumulated income, while player II will minimize the same. An *admissible relaxed strategy* $m_k^\varepsilon(\cdot)$ for the k^{th} player with system (6.3.7) is a measure valued random variable satisfying $\int \int_0^t f(s, \alpha) m^\varepsilon(ds d\alpha)$ is progressively measurable with respect to $\{\Im_t^\varepsilon\}$, where $\Im_t^\varepsilon$ is the minimal σ-algebra generated by $\{\xi^\varepsilon(s), x^\varepsilon(s),\ s \leqslant t\}$. Also $m^\varepsilon([0,t] \times U) = t$ for all $t \geqslant 0$. Also, there is a derivative m_t^ε, where $m_t^\varepsilon(b)$ are $\Im_t^\varepsilon$ measurable for Borel B. The concept of δ-ergodic equilibrium for $x^\varepsilon(\cdot)$ is similarly defined as in (6.3.6).

Under the Lyapunov type stability condition (assumption A in Borkar and Ghosh [31]), the following result is proved.

Theorem 6.3.1. *For the stochastic differential game with ergodic payoff criterion has a value and both players have optimal strategies* $m^* = (m_1^*, m_2^*) \in M_{a1} \times M_{a2}$.

6.3.3 *Chattering Lemma*

In the relaxed control setting, each player chooses at time t a probability measure $m_i(t)$ on the control set M_i rather than an element $u_i(t) \in U_i$, $i = 1,2$. Since relaxed controls are devices with primarily a mathematical use, it is desirable to have a chattering type result for the game problem. In order for the relaxed control problem to be true extension of the original problem, the equilibrium among the relaxed control strategies must be the same as the equilibrium taken among the ordinary strategies when it exists. For this purpose, we extend the chattering results obtained for control problems as in Fleming [60], to two person zero-sum stochastic differential games. We note that $U_i \subseteq M_i$, because, if $m_i(t)$ is an atomic measure, concentrated at a single point $u(t)$ for each t, then we get an ordinary control policy as a special case of a relaxed control policy. Let (m_1^*, m_2^*) be the equilibrium policy pair in the relaxed controls and (u_1^*, u_2^*) be the equilibrium policy pair (if it exists) in the ordinary controls.

Theorem 6.3.2. *Under the conditions of Theorem 6.3.1,*

$$J(m_1^*, m_2^*) = J(u_1^*, u_2^*).$$

Proof.

(a) suppose $J(m_1^*, m_2^*) \geqslant J(u_1^*, u_2^*)$.

From Fleming [60], there exists a $u_\varepsilon \in U$, such that,

$$|J(m_1^*, u_2^*) - J(u_{1\varepsilon}, u_2^*)| < \varepsilon. \tag{6.3.9}$$

From the definition of $J(u_1^*, u_2^*)$ and $J(m_1^*, m_2^*)$, we have,

$$J(u_1^*, u_2^*) \geqslant J(u_{1\varepsilon}, u_2^*) \tag{6.3.10}$$

and

$$J(m_1^*, u_2^*) \geqslant J(m_1^*, m_2^*). \tag{6.3.11}$$

Adding (6.3.10) and (6.3.11), we have,

$$J(u_1^*, u_2^*) + J(m_1^*, u_2^*) \geqslant J(u_{1\varepsilon}, u_2^*) + J(m_1^*, m_2^*),$$

which implies,

$$J(m_1^*, u_2^*) - J(u_{1\varepsilon}, u_2^*) \geqslant J(m_1^*, m_2^*) - J(u_1^*, u_2^*) \geqslant 0 \quad \text{(by assumption)},$$

which implies,

$$\varepsilon > |J(m_1^*, u_2^*) - J(u_{1\varepsilon}, u_2^*)| \geqslant |J(m_1^*, m_2^*) - J(u_1^*, u_2^*)|,$$

and thus, we have $J(m_1^*, m_2^*) = J(u_1^*, u_2^*)$, as ε is arbitrary.

(b) Suppose $J(m_1^*, m_2^*) \leqslant J(u_1^*, u_2^*)$.
Let $u_{2\varepsilon} \in U_2$, such that

$$|J(u_1^*, m_2^*) - J(u_1^*, u_{2\varepsilon})| < \varepsilon$$

as before

$$J(u_1^*, u_2^*) \leqslant J(u_1^*, u_{2\varepsilon})$$

and

$$J(u_1^*, m_2^*) \leqslant J(m_1^*, m_2^*)$$

implies

$$0 \leqslant J(u_1^*, u_2^*) - J(m_1^*, m_2^*) \leqslant J(u_1^*, u_{2\varepsilon}) - J(u_1^*, m_2^*) < \varepsilon,$$

and thus,

$$J(m_1^*, m_2^*) = J(u_1^*, u_2^*).$$

Hence the proof. □

6.3.4 *Main Result*

Now, we will prove the weak convergence of the wideband system (6.3.7) to the diffusion system (6.3.1) and the δ-optimality of the equilibrium strategies of (6.3.1) applied to (6.3.7). We will use the following assumptions, which are very general. For a detailed description on these types of assumptions, we refer the reader to Kushner [103] and Kushner and Dupuis [112].

(A6.3.1): $a_i(\cdot,\cdot)$, $i = 1, 2$, $G(\cdot,\cdot)$, $g(\cdot,)$, $g_x(\cdot,\cdot)$ are continuous and are bounded by $O(1 + |x|)$. Also, $G_x(\cdot,\xi)$ is continuous in x for each ξ and is bounded. $\xi(\cdot)$ is bounded, right continuous, and $EG(x,\xi(t)) \to 0$, $Eg(x,\xi(t)) \to 0$ as $t \to \infty$, for each x. Also, $r(\cdot,\cdot)$ is bounded and continuous.

(A6.3.2): $g_{xx}(\cdot,\xi)$ is continuous for each ξ, and is bounded.

(A6.3.3): Let $W(x,\xi)$ denote either $\varepsilon G(x,\xi), G_x(x,\xi), g(x,\xi)$ or $g_x(x,\xi)$. Then for compact Q, we have

$$\varepsilon \sup_{x \in Q} \left| \int_{t/\varepsilon^2}^{\infty} E_t^{\varepsilon} W(x,\xi(s)) ds \right| \xrightarrow{\varepsilon} 0$$

in the mean square sense, uniformly in t.

(A6.3.4): Let g_i denote the *ith* component of g. There are continuous $\overline{g}_i(\cdot), b(\cdot) = \{b_{ij}(\cdot)\}$ such that

$$\int_t^\infty E g_{i,x}(x,\xi(s))g(x,\xi(t))\,ds \longrightarrow \overline{g}_i(x),$$

and

$$\int_t^\infty E g_i(x,\xi(s))g_j(x,\xi(t))\,ds \longrightarrow \frac{1}{2}b_{ij}(x),$$

as $t \to \infty$, and the convergence is uniform in any bounded x-set.

Note: Let $b(x) = \{b_{ij}(x)\}$. For $i \neq j$, it is not necessary that $b_{ij} = b_{ji}$. In that case define $\widetilde{b}(x) = \frac{1}{2}[b(x) + b'(x)]$, as the symmetric covariance matrix, then use b for the new $\widetilde{b}$. Hence, for notational simplicity, we will not distinguish between $b(x)$ and $\widetilde{b}(x)$.

(A6.3.5): For each compact set Q and all $i,\ j$, we assume

(a) $\sup_{x\in Q} \varepsilon^2 \left| \int_{t/\varepsilon^2}^\infty d\tau \int_\tau^\infty ds \left[E_{t/\varepsilon^2} g'_{i,x}(x,\xi(s))g(x,\xi(t)) - \mathrm{E}g'_{i,x}(x,x(s))g(x,x(t)) \right] \right| \to 0;$

and

(b) $\sup_{x\in Q} \varepsilon^2 \left| \int_{t/\varepsilon^2}^\infty d\tau \int_\tau^\infty ds \left[E_{t/\varepsilon^2} g_i(x,\xi(s))g_j(x,\xi(t)) - \mathrm{E}g_i(x,x(s))g_j(x,x(t)) \right] \right| \to 0;$

in the mean square sense as $\varepsilon \to 0$, uniformly in t.

Define $\overline{a}(x,\alpha) = a_1(x,\alpha_1) + a_2(x,\alpha_2) + \overline{g}(x)$ and the operator A^m as

$$A^m f(x) = \int A^\alpha f(x) m_x(d\alpha),$$

where

$$A^\alpha f(x) = f'_x(x)\overline{a}(x,\alpha) + \frac{1}{2}\sum_{i,j} b_{ij}(x) f_{x_i x_j}(x).$$

For a fixed control α, A^α will be the operator of the process that is the weak limit of $\{x^\varepsilon(\cdot)\}$.

(A6.3.6): The martingale problem for operator A^m has a unique solution for each relaxed admissible Markov strategy $m_x(\cdot)$, and each initial condition. The process is a Feller process. The solution of (6.3.7) is unique in the weak sense for each $\varepsilon > 0$. Also $b(x) = \sigma(x)\sigma'(x)$ for some continuous finite dimensional matrix $\sigma(\cdot)$.

For an admissible relaxed policy for (6.3.7) and (6.3.1), respectively, define the occupation measure valued random variables $P_T^{m,\varepsilon}(\cdot)$ and $P_T^m(\cdot)$ by, respectively,

$$P_T^{m,\varepsilon}(B \times C) = \frac{1}{T}\int_0^T I_{\{x^\varepsilon(t)\in B\}} m_t^\varepsilon(c)dt,$$

and

$$P_T^m(B \times C) = \frac{1}{T}\int_0^T I_{\{x(t)\in B\}} m_t(c)dt$$

where B and C are Borel subsets in $\mathbb{R}^d$ and $[0,t] \times U$, respectively.

Let $\{m^{\varepsilon}(\cdot)\}$ be a given sequence of admissible relaxed controls.

(A6.3.7): For a fixed $\delta > 0$,

$$\{x^{\varepsilon}(t),\ \text{small } \varepsilon > 0,\ t \in \text{ dense set in } [0,\infty),\ m^{\varepsilon} \text{ used}\}$$

are tight.

Note: The assumption (A6.3.7) implies that the set of measure valued random variables

$$\{P_T^{m^{\varepsilon},\varepsilon}(\cdot), \text{small } \varepsilon > 0,\ T < \infty\},$$

are tight.

(A6.3.8): For the *ergodic equilibrium* pair of Markov strategies $m^* = (m_1^*, m_2^*)$ with initial law π for (6.3.1) and (6.3.2), the martingale problem has a unique solution. The solution is a Feller process and there is a unique invariant measure $\mu(m^*)$.

Note: Existence of such an invariant measure is assured if the process is positive recurrent. Also, under the conditions of Theorem 6.3.1, the assumption (A6.3.8) will follow.

The following result gives the main convergence and δ-optimality result for the ergodic payoff criterion.

Theorem 6.3.3. *Assume* (A6.3.1) *to* (A6.3.8). *Let* $(m_1^{*\varepsilon}, m_2^{*\varepsilon})$ *be the policy pair* (m_1^*, m_2^*) *adaptively applied to* (6.3.7) *and* (6.3.8). *Then* $\{x^{\varepsilon}(\cdot), m_1^{*\varepsilon}, m_2^{*\varepsilon}\} \to (x(\cdot), m_1^*, m_2^*)$ (*in the Skorohod topology*) *and there is a Wiener process* $w(\cdot)$ *such that* $(x(\cdot), m_1^*, m_2^*)$ *is nonanticipative with respect to* $w(\cdot)$*, and* (6.3.1) *holds. Also,*

$$J^{\varepsilon}(m_1^{*\varepsilon}, m_2^{*\varepsilon}) \xrightarrow{P} J(m_1^*, m_2^*) = V(\pi). \tag{6.3.12}$$

In addition, let $(\widehat{m}_1^{\varepsilon}(\cdot), m_2^{\varepsilon}(\cdot))$ *be a* δ*-optimal strategy pair for player I and* $(m_1^{\varepsilon}(\cdot), \widehat{m}_2^{\varepsilon}(\cdot))$ *be* δ*-optimal pair for player II for* $x^{\varepsilon}(\cdot)$ *of* (6.3.7). *Then*

$$\underline{\lim}_{\varepsilon,T} P\{|J^{\varepsilon}(m_1^{*\varepsilon}, m_2^{*\varepsilon}) - J^{\varepsilon}(\widehat{m}_1^{\varepsilon}(\cdot), m_2^{\varepsilon}(\cdot))| < \delta\} = 1 \tag{6.3.13}$$

and

$$\underline{\lim}_{\varepsilon,T} P\{|J^{\varepsilon}(m_1^{*\varepsilon}, m_2^{*\varepsilon}) - J^{\varepsilon}(m_1^{\varepsilon}(\cdot), \widehat{m}_2^{\varepsilon}(\cdot))| < \delta\} = 1 \tag{6.3.14}$$

Proof. The correct procedure of the proof is to work with the truncated processes $x^{\varepsilon,K}(\cdot)$ and to use the piecing together the idea of Lemma 6.2.1 to get convergence of the original $x^{\varepsilon}(\cdot)$ sequence, unless $x^{\varepsilon}(\cdot)$ is bounded on each $[0,T]$, uniformly in ε. For notational simplicity, we ignore this technicality. Simply suppose that $x^{\varepsilon}(\cdot)$ is bounded in the following analysis. Otherwise, one can work with K-truncation. Let $\widehat{D}$ be a measure determining set of bounded real-valued continuous functions on $\mathbb{R}^d$ having continuous second partial derivatives and compact support. Let $m_t^\varepsilon(\cdot)$ be the relaxed Markov policies of (A6.3.7). Whenever convenient, we write $x^\varepsilon(t) = x$. For the test function $f(\cdot) \in \widehat{D}$, define the perturbed test functions (the change of variable $s/\varepsilon^2 \to s$ will be used through out the proofs) given by

$$f_0^\varepsilon(x,t) = \int_t^\infty E_t^\varepsilon f_x'(x) G(x, \xi^\varepsilon(s))\, ds = \varepsilon^2 \int_{t/\varepsilon^2}^\infty E_t^\varepsilon f_x'(x) G(x, \xi(s)) ds,$$

$$f_1^\varepsilon(x,t) = \frac{1}{\varepsilon} \int_t^\infty E_t^\varepsilon f_x'(x) g(x, \xi^\varepsilon(s))\, ds = \varepsilon \int_{t/\varepsilon^2}^\infty E_t^\varepsilon f_x'(x) g(x, \xi(s)) ds,$$

and

$$\begin{aligned} f_2^\varepsilon(x,t) = \frac{1}{\varepsilon^2} \int_t^\infty ds \int_s^\infty d\tau \Big\{ & E_t^\varepsilon [f_x'(x) g(x, \xi^\varepsilon(\tau))]_x' g(x, \xi^\varepsilon(s)) \\ & - E[f_x'(x) g(x, \xi^\varepsilon(\tau))]_x' g(x, \xi^\varepsilon(s)) \Big\} \\ = \varepsilon^2 \int_{t/\varepsilon^2}^\infty ds \int_s^\infty d\tau \Big\{ & E_t^\varepsilon [f_x'(x) g(x, \xi(\tau))]_x' g(x, \xi(s)) \\ & - E[f_x'(x) g(x, \xi(\tau))]_x' g(x, \xi(s)) \Big\}. \end{aligned}$$

From (A6.3.1), (A6.3.2), (A6.3.3), and (A6.3.5), $f_i^\varepsilon(\cdot) \in D(A^\varepsilon)$ for $i = 0, 1, 2$. Define the perturbed test function

$$f^\varepsilon(t) = f(x^\varepsilon(t)) + \sum_{i=0}^{2} f_i^\varepsilon(x^\varepsilon(t), t).$$

The reason for defining f_i^ε is to facilitate the averaging of the "noise" terms involving ξ^ε terms. By the definition of the operator A^ε and its domain $D(A^\varepsilon)$, we will obtain that $f(x^\varepsilon(\cdot))$ and the $f_i^\varepsilon(x^\varepsilon(\cdot), \cdot)$ are all in $D(A^\varepsilon)$, and

$$\begin{aligned} A^{m^\varepsilon, \varepsilon} f(x^\varepsilon(t)) = f_x'(x^\varepsilon(t)) \Bigg[& \sum_{i=1}^{2} \int a_i(x^\varepsilon(t), \alpha) m_{it}^\varepsilon(d\alpha) \\ & + G(x^\varepsilon(t), \xi^\varepsilon(t)) + \frac{1}{\varepsilon} g(x^\varepsilon(t), \xi^\varepsilon(t)) \Bigg]. \end{aligned} \tag{6.3.15}$$

From (6.3.15) we can obtain,

$$\begin{aligned} A^{m^\varepsilon, \varepsilon} f_0(x^\varepsilon(t)) &= -f_x'(x^\varepsilon(t)) G(x^\varepsilon(t), \xi^\varepsilon(t)) + \int_t^\infty ds [E_t^\varepsilon f_x'(x^\varepsilon(t)) G(x^\varepsilon(t), \xi^\varepsilon(s))]_x' \dot{x}^\varepsilon(t) \\ &= -f_x'(x^\varepsilon(t)) G(x^\varepsilon(t), \xi^\varepsilon(t)) + \varepsilon^2 \int_{t/\varepsilon^2}^\infty ds [E_t^\varepsilon f_x'(x^\varepsilon(t)) G(x^\varepsilon(t), \xi(s))]_x' \dot{x}^\varepsilon(t). \end{aligned} \tag{6.3.16}$$

Note that the first term in (6.3.16) will cancel with $f_{x'}G$ term of (6.3.15). The p-lim of the last term in (6.3.16) is zero. Thus, we have,

$$
\begin{aligned}
A^{m^\varepsilon,\varepsilon} f_1(x^\varepsilon(t)) &= -\frac{1}{\varepsilon} f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi^\varepsilon(t)) + \frac{1}{\varepsilon}\int_t^\infty ds[E_t^\varepsilon f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi^\varepsilon(s))]_x'\dot{x}^\varepsilon(t) \\
&= -\frac{1}{\varepsilon} f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi^\varepsilon(t)) + \varepsilon\int_{t/\varepsilon^2}^\infty ds[E_t^\varepsilon f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi(s))]_x'\dot{x}^\varepsilon(t). \qquad (6.3.17)
\end{aligned}
$$

The first term on the right of (6.3.17) will cancel with the $\frac{f_x'g}{\varepsilon}$ term in (6.3.15). The only component of the second term on the right of (6.3.17) whose $p\text{-}\lim_{\varepsilon}$ is not zero is

$$
\frac{1}{\varepsilon^2}\int_t^\infty ds\left\{E_t^\varepsilon[f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi^\varepsilon(s))]_x' g(x^\varepsilon(t),\xi^\varepsilon(t))\right\}.
$$

This term will cancel with the first term of (6.3.18). Thus,

$$
\begin{aligned}
A^{m^\varepsilon,\varepsilon} f_2(x^\varepsilon(t)) &= -\frac{1}{\varepsilon^2}\int_t^\infty ds\Big\{E_t^\varepsilon[f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi^\varepsilon(s))]_x' g(x^\varepsilon(t),\xi^\varepsilon(t)) \\
&\quad - E[f_x'(x^\varepsilon(t))g(x,\xi^\varepsilon(s))]_x' g(x,\xi^\varepsilon(t))\,|_{x=x^\varepsilon(t)}\Big\} + [f_2^\varepsilon(x,t)]_x' \overset{\varepsilon}{\dot{x}}|_{x=x^\varepsilon(t)} \\
&= -\int_{t/\varepsilon^2}^\infty ds\{E_t^\varepsilon[f_x'(x^\varepsilon(t))g(x^\varepsilon(t),\xi(s))]_x' g(x^\varepsilon(t),\xi^\varepsilon(t)) \\
&\quad - E[f_x'(x^\varepsilon(t))g(x,\xi(s))]_x' g(x,\xi^\varepsilon(t))\,|_{x=x^\varepsilon(t)}\} + [f_2^\varepsilon(x,t)]_x' \overset{\varepsilon}{\dot{x}}|_{x=x^\varepsilon(t)}. \qquad (6.3.18)
\end{aligned}
$$

The $p\text{-}\lim_{\varepsilon}$ of the last term of the right side of (6.3.18) is zero.

Evaluating $A^{m^\varepsilon,\varepsilon} f^\varepsilon(t) = A^{m^\varepsilon,\varepsilon}\left[f(x^\varepsilon(t)) + \sum_{i=0}^2 f_i^\varepsilon(x^\varepsilon(t),t)\right]$ and by deleting terms that cancel we have

$$
\begin{aligned}
A^{m^\varepsilon,\varepsilon} f^\varepsilon(t) &= f_x'(x^\varepsilon(t))\sum_{i=1}^2 \int a_i(x^\varepsilon(t),\alpha)m_{it}^\varepsilon(d\alpha) \\
&\quad + \int_{t/\varepsilon^2}^\infty E[f_x'(x^\varepsilon(t))g(x,\xi(s))]' g(x,\xi(t/\varepsilon^2))ds.
\end{aligned} \qquad (6.3.19)
$$

As a result, we get

$$
p\text{-}\lim(f^\varepsilon(t) - f(x^\varepsilon(\cdot))) = 0 \qquad (6.3.20)
$$

$$
p\text{-}\lim_{\varepsilon}\left|A^{m^\varepsilon,\varepsilon} f(x^\varepsilon(t)) - A^{m^{\varepsilon,\varepsilon}} f^\varepsilon(t)\right| = 0. \qquad (6.3.21)
$$

Hence, by Lemma 6.2.2,

$$
M_f^\varepsilon(t) = f^\varepsilon(t) - f^\varepsilon(0) - \int_0^t A^{m^\varepsilon} f^\varepsilon(s)ds
$$

is a zero mean martingale.

Let $[t]$ denote the greatest integer part of t. W can write

$$
\frac{M_f^\varepsilon(t)}{t} = \frac{1}{t}\left[\left(M_f^\varepsilon(t) - M_f^\varepsilon([t])\right) + M_f^\varepsilon(0)\right] + \frac{1}{t}\sum_{k=0}^{[t]-1}\left[M_f^\varepsilon(k+1) - M_f^\varepsilon(k)\right].
$$

Using the fact that $f(\cdot)$ is bounded and (6.3.21), and martingale property of $M_f^\varepsilon(\cdot)$, we get $E\left[\frac{M_f^\varepsilon(t)}{t}\right]^2 \to 0$ as $t \to \infty$ and $\varepsilon \to 0$, which in turn implies that $\frac{M_f^\varepsilon(t)}{t} \xrightarrow{P} 0$ as $t \to \infty$ and $\varepsilon \to 0$ in any way at all. From (6.3.21), and the fact that $\frac{M_f^\varepsilon(t)}{t}$, $\frac{f^\varepsilon(t)}{t}$, and $\frac{f^\varepsilon(0)}{t}$ all go to zero in probability implies that as $t \to \infty$ and $\varepsilon \to 0$, we have

$$\frac{1}{t}\int_0^t A^{m^\varepsilon} f(x^\varepsilon(s))ds \xrightarrow{P} 0. \tag{6.3.22}$$

By the definition of $P_T^{m^\varepsilon,\varepsilon}(\cdot)$, (6.3.22) can be written as

$$\int A^\alpha f(x) P_T^{m^\varepsilon,\varepsilon}(dxd\alpha) \xrightarrow{P} 0 \text{ as } T \to \infty \text{ and } \varepsilon \to 0. \tag{6.3.23}$$

For the policy $m^*(\cdot)$, choose a weakly convergent subsequence of set of random variables $\{P_T^{m^*,\varepsilon}(\cdot), \varepsilon, T\}$, indexed by ε_n, T_n, with limit $\widehat{\mu}(\cdot)$. Let this limit $\widehat{P}(\cdot)$ be defined on some probability space $(\widetilde{\Omega}, \widetilde{P}, \widetilde{\mathfrak{I}})$ with generic variable $\widetilde{\omega}$. Factor $\widehat{P}(\cdot)$ as $\widehat{P}(dxd\alpha) = m_x^*(d\alpha)\mu(dX)$. We can suppose that $m_x(c)$ are x-measurable for each Borel set C and $\widetilde{\omega}$. Now (6.3.23) implies that for all $f(\cdot) \in \widehat{D}$, we can write,

$$\int\int A^\alpha f(x) m_x^*(d\alpha)\widehat{\mu}(dX) = 0 \text{ for } \widetilde{P}\text{-almost all } \widetilde{\omega}. \tag{6.3.24}$$

Since $f(\cdot)$ is measure determining, (6.3.24) implies that almost all realizations of $\widehat{\mu}$ are invariant measures for (6.3.1) under the relaxed policies m^*. By uniqueness of the invariant measure, we can take $\mu(m^*,\cdot) = \widehat{\mu}(\cdot)$ does not depend on the chosen subsequence ε_n, T_n. By the definition of $P_T^{m^*,\varepsilon}(\cdot)$, we have

$$\begin{aligned}\frac{1}{t}\int_0^t \int r(x^\varepsilon(s),\alpha) m^{*\varepsilon}(d\alpha)ds &= \int_0^t \int r_k(x^\varepsilon(s),\alpha) P_T^{m^*,\varepsilon}(d\alpha dx)\\ &\xrightarrow{P} \int_0^t \int r(x,\alpha) m_x^*(d\alpha)\widehat{\mu}(dX) = J(m^*).\end{aligned}$$

Hence, we have (6.3.12). Let $\widetilde{m}^{\delta_1\varepsilon} = (\widehat{m}_1^\varepsilon(\cdot), m_2^\varepsilon(\cdot))$ and $\widetilde{m}^{\delta_1\varepsilon} = (m_1^\varepsilon(\cdot), \widehat{m}_2^\varepsilon(\cdot))$ are the δ-optimal strategies for players I and II, respectively. Now (6.3.13) and (6.3.14) follows using the fact that (6.3.12) holds for all the limits of the tight sets $\{P_T^{m^{\delta_i},\varepsilon}(\cdot); \varepsilon, T\}$, $i = 1,2$, the assumed uniqueness in (A6.3.8), and the definition of δ-optimality. □

It is important to note that, as a result of Theorem 6.3.3, if one needs a δ-optimal policy for the physical system, it is enough to compute for the diffusion model and use it to the physical system. There is no need to compute optimal policies for each ε.

6.3.5 *Discrete Games*

For the stochastic or the discrete parameter games, the system is given by

$$X_{n+1}^\varepsilon = X_n^\varepsilon + \varepsilon G(X_n^\varepsilon) + \varepsilon \sum_{i=1}^N \int a_i(X_n^\varepsilon, \alpha_i) m_{in}(d\alpha_i) + \sqrt{\varepsilon} g(X_n^\varepsilon, \xi_n^\varepsilon) \tag{6.3.25}$$

where $\{\xi_n^\varepsilon\}$ satisfies the discrete parameter version of (A6.3.2) and $m_{in}(\cdot)$, $i = 1,\dots,N$ be the relaxed control strategies depending only on $\{X_i, \xi_{i-1},\ i \leqslant n\}$. It should be noted that, in the discrete case, strategies would not be relaxed, one need to interpret this in asymptotic sense, i.e., the limiting strategies will be relaxed. Let E_n^ε denote the conditional expectation with respect to $\{X_i, \xi_{i-1},\ i \leqslant n\}$. Define $x^\varepsilon(\cdot)$ by $x^\varepsilon(t) = X_n^\varepsilon$ on $[n\varepsilon, n\varepsilon+\varepsilon)$ and $m_i(\cdot)$ by

$$m_i(B_i \times [0,t]) = \varepsilon \sum_{n=0}^{[t/\varepsilon]-1} m_{in}(B_i) + \varepsilon(t - \varepsilon t/\varepsilon) m_{[t/\varepsilon]}(B_i), \quad i = 1,\dots,N.$$

(A6.3.9):

(i) For V equal either $a(\cdot,\cdot)$, g or g_x, and for Q compact, $E\sup_x \left|\sum_{n+L_1}^{L} E_n^\varepsilon V(x,\xi_i^\varepsilon)\right| \to 0$, as L, n and $L_1 \to \infty$, with $L > n + L_1$ and $L - (n+L_1) \to \infty$.

(ii) There are continuous functions $c(i,x)$ and $c_0(i,x)$ such that for each x

$$\frac{1}{L}\sum_{n=\ell}^{\ell+L} E_\ell^\varepsilon g(x,\xi_{n+i}^\varepsilon) g'(x,\xi_n^\varepsilon) \xrightarrow{P} c(i,x)$$

and

$$\frac{1}{L}\sum_{n=\ell}^{\ell+L} E_\ell^\varepsilon g_x'(x,\xi_{n+i}^\varepsilon) g(x,\xi_n^\varepsilon) \xrightarrow{P} c_0(i,x)$$

as ℓ and $L \to \infty$.

(iii) For each $T < \infty$ and compact Q,

$$\varepsilon \sup_{x\in Q}\left|\sum_{j=n}^{T/\varepsilon}\sum_{k=j+1}^{T/\varepsilon}[E_n^\varepsilon g_{i,x}'(x,\xi_k)g(x,\xi_j) - Eg_{i,x}'(x,\xi_k)g(x,\xi_j)]\right| \to 0, \quad i \leqslant n,$$

and

$$\varepsilon \sup_{x\in Q}\left|\sum_{j=n}^{T/\varepsilon}\sum_{k=j+1}^{T/\varepsilon}[E_n^\varepsilon g'(x,\xi_k)g(x,\xi_j) - Eg'(x,\xi_k)g(x,\xi_j)]\right| \to 0,$$

in the mean as $\varepsilon \to 0$ uniformly in $n \leqslant T/\varepsilon$. Also, the limits hold when the bracketed terms are replaced by their (x-gradient$/\sqrt{\varepsilon}$).

Define,

$$\widetilde{a}(x) = \sum_{1}^{\infty} c_0(i,x)$$

and

$$\widetilde{c}(x) = c(0,x) + 2\sum_{1}^{\infty} c(i,x) = \sum_{-\infty}^{\infty} c(i,x).$$

With some minor modifications in the proof of Theorem 6.2.3, we can obtain the following result (Refer to Kushner [103] and Ramachandran [161], for convergence proofs in similar situation).

Theorem 6.3.4. *Assume* (A6.3.1) *to* (A6.3.3), (A6.3.6) *to* (A6.3.9). *Then the conclusions of Theorem* 6.23.3 *hold for model* (6.3.25).

6.3.6 *Discounted Payoff*

In this subsection, we will consider discounted payoff, rather than average payoff. As much as possible, we will use the same notation as in Section 6.3.1. The only changes will be highlighted.

Consider a system of the following type in relxed control setting.

$$dx^\varepsilon = \int a_1(x^\varepsilon,\alpha_1)m^\varepsilon_{1t}(d\alpha_1) + \int a_2(x^\varepsilon,\alpha_2)m^\varepsilon_{2t}(d\alpha_2)dt + \frac{1}{\varepsilon}g(x^\varepsilon,\xi^\varepsilon)dt$$
$$\text{with } x^\varepsilon(0) = x_0. \tag{6.3.26}$$

The total discounted payoff to player 1 is given by

$$J^\varepsilon[m^\varepsilon](x) = E_x\int_0^\infty\int e^{-\lambda t} r(x^\varepsilon(s),\alpha)m^\varepsilon_s(d\alpha)ds \tag{6.3.27}$$

and for the initial law π in $P(\mathbb{R}^n)$, it is given by

$$J^\varepsilon[m^\varepsilon](\pi) = \int_{\mathbb{R}^n} J^\varepsilon[m^\varepsilon](x)\pi(dX). \tag{6.3.28}$$

The diffusion model is given by

$$dx(t) = \int_{U_1} a_1(x(t),\alpha_1)m_{1t}(d\alpha) + \int_{U_2} a_2(x(t),\alpha_2)m_{2t}(d\alpha)dt + \overline{g}(x(t))dt + \sigma(x(t))dw(t)$$
$$x(0) = x_0, \tag{6.3.29}$$

with a total payoff to player 1 being

$$J[m](x) = E_x\int_0^\infty\int e^{-\lambda t} r(x(s),\alpha)m(d\alpha)ds, \tag{6.3.30}$$

and $J[m](\pi)$ defined as in (6.3.28). Discounted optimal strategy is defined same as in (6.3.4) and (6.3.5). Also δ-discounted equilibrium is defined as in (6.3.6), except that the J is as in (6.3.10). The *discounted occupation measure* for initial condition $x \in \mathbb{R}^n$ denoted by $\nu_{\lambda x}(m) \in P(\mathbb{R}^n \times U_1 \times U_2)$ is defined by

$$\int_{\mathbb{R}^n\times U} f d\nu_{\lambda x}[m] = \lambda^{-1}E_x\left[\int_0^\infty\int e^{-\lambda t} f(x(t),\alpha)m_t(d\alpha)dt\right]$$

and for initial law $\pi \in P(\mathbb{R}^n)$, $\nu_{\lambda\pi}[m]$ is defined as

$$\int f d\nu_{\lambda\pi}[m] = \int_{\mathbb{R}^n}\pi(dX)\int_{\mathbb{R}^n\times U} f d\nu_{\lambda x}[m].$$

Then $J[m](\pi)$ can be rewritten as

$$J[m](\pi) = \lambda\int r(x,\alpha)d\nu_{\lambda x}[m].$$

Let $\nu_{\lambda\pi}[A_1,A_2] = \{\nu_{\lambda\pi}(m) \mid m \in A_1\times A_2\}$. $\nu_{\lambda\pi}[M_{a_i},A_i]$ and $\nu_{\lambda\pi}[M_{a_1},M_{a_2}]$ are defined analogously. Now we will state following two results from Borkar and Ghosh [31], and Ramachandran [158].

Theorem 6.3.5. (i) $\nu_{\lambda\pi}[A_1,M_{a_2}] = \nu_{\lambda\pi}[M_{a_1},M_{a_2}] = \nu_{\lambda\pi}[M_{a_1},A_2]$.
(ii) $\nu_{\lambda\pi}[M_{a_1},M_{a_2}]$ *is component wise convex and compact.*

Theorem 6.3.6. *The stochastic differential game with system* (6.3.29) *admits a value and both players have optimal Markov strategies.*

Now we will state the main weak convergence result. The proof is similar to the proof of Theorem 6.2.3, Ramachandran [158].

Theorem 6.3.7. *Assume* (A6.3.1), (A6.3.4) *and that* $\xi^{\varepsilon}(t) = \xi(t/\varepsilon)$ *with* $\xi(\cdot)$ *being a stationary process which is strongly mixing, right continuous and bounded with mixing rate function* $\phi(\cdot)$ *satisfying* $\int_0^\infty \phi^{1/2}(s)ds < \infty$. *Let* $m^{\varepsilon}(\cdot) \to m(\cdot)$. *There is a* $w(\cdot)$ *such that* $m(\cdot)$ *is admissible strategy with respect to* $w(\cdot)$ *and* $(x^{\varepsilon}(\cdot), m^{\varepsilon}(\cdot)) \to (x(\cdot), m(\cdot))$, *where* $(x(\cdot), m(\cdot))$ *satisfies equation* (6.3.29).

Let $(\overline{m}_1, \overline{m}_2)$ be a value for the system (6.2.29), existence of which is guaranteed from Theorem 6.3.26. Also in Borkar and Ghosh [31], the value function is characterized as the unique solution of the Isaacs equation in $W^{2,p}_{\text{loc}}(\mathbb{R}^n) \cap C_b(\mathbb{R}^n)$ for $p \geqslant 2$.
(A6.3.10): Let (6.3.29) have a unique weak sense solution for the strategy $(\overline{m}_1, \overline{m}_2)$ and let the solution strategy be unique. Assume $(\overline{m}_1, \overline{m}_2)$is admissible for $x^{\varepsilon}(\cdot)$ of (6.3.26) for small ε.

Theorem 6.3.8. *Assume* (A6.3.10) *in addition to the assumptions of Theorem* 6.3.7. *Then, letting* $x^{\varepsilon}(\cdot)$ *denote the solution of* (6.3.26) *driven by the policy* $(\overline{m}_1, \overline{m}_2)$, *we have* $\{x^{\varepsilon}(\cdot), \overline{m}_1(\cdot), \overline{m}_2(\cdot)\} \to (x(\cdot), \overline{m}_1(\cdot), \overline{m}_2(\cdot))$ *and there is a Wiener process* $w(\cdot)$ *such that* $(x(\cdot), \overline{m}_1(\cdot), \overline{m}_2(\cdot))$ *is nonanticipative with respect to* $w(\cdot)$, *and* (6.3.29) *holds. Also,*

$$J^{\varepsilon}(\overline{m}_1, \overline{m}_2)(\pi) \longrightarrow J(\overline{m}_1, \overline{m}_2)(\pi) = V(\pi). \tag{6.3.31}$$

In addition, let $\widehat{m}_1^{\varepsilon}$ *and* $\widehat{m}_2^{\varepsilon}$ *be a* δ*-optimal strategy pair for player 1 and 2 respectively with* x^{ε} *of (6.3.26). Then*

$$\underline{\lim}_{\varepsilon} \left| \left[\sup_{m_1^{\varepsilon} \in A_1} J^{\varepsilon}(m_1^{\varepsilon}, \widehat{m}_2^{\varepsilon}) - J^{\varepsilon}(\overline{m}_1, \overline{m}_2) \right] \right| \leqslant \delta \tag{6.3.32}$$

and

$$\overline{\lim_{\varepsilon}} \left| \left[\inf_{m_2^{\varepsilon} \in A_2} J^{\varepsilon}(\widehat{m}_1^{\varepsilon}, m_2^{\varepsilon}) - J^{\varepsilon}(\overline{m}_1, \overline{m}_2) \right] \right| \leqslant \delta. \tag{6.3.33}$$

Proof. From Theorem 6.3.7 and by the uniqueness of $(x^{\varepsilon}(\cdot), \overline{m}_1(\cdot), \overline{m}_2(\cdot))$ converges weakly to $(x(\cdot), \overline{m}_1(\cdot), \overline{m}_2(\cdot))$. The limit satisfies equation (6.3.29) for some Wiener process $w(\cdot)$.
Fix N, and let S_N be the N-ball in $\mathbb{R}^n$. Let

$$\tau^{\varepsilon,N} = \inf\{t : x^{\varepsilon}(t) \notin S_N\} \text{ and } \tau^N = \inf\{t : x(t) \notin S_N\}.$$

By (A6.3.4), is nondegenerate. This with the properties of the Wiener process $w(\cdot)$, $\tau^N(\cdot)$ is continuous w.p.1. with respect to the measure induced by $x(\cdot)$. By weak convergence of the state processes, we have $\tau^{\varepsilon,N} \to \tau^N$ as $\varepsilon \to 0$. Let

$$J^N(m)(x) = E_x \int_0^{\tau^N} \int e^{-\lambda t} r(x,\alpha) m_t(d\alpha)dt$$

and let the occupation measure $\nu^N_{\lambda\pi}(m)$be the $\nu_{\lambda\pi}(m)$ corresponding to $J^N(m)(x)$. As $N \to \infty$, we have $\tau^N \to \infty$, w.p.1, since $x \in \mathbb{R}^n$. Hence all the steps below can be worked with the truncated version and then, take the limit as $N \to \infty$. For simplicity of notation, we will not carry the N-subscript.

By Theorem 6.3.5, we have

$$\begin{aligned}\lim_{\varepsilon} J^{\varepsilon}(\overline{m})(\pi) &= \lim_{\varepsilon} \lambda \int r(x^{\varepsilon},\alpha) d\nu_{\lambda\pi}[\overline{m}] \\ &= \lambda \int r(x,\alpha) d\nu_{\lambda\pi}[\overline{m}] = J(\overline{m}_1,\overline{m}_2)(\pi) = V(\pi).\end{aligned}$$

To show (6.3.32) and (6.3.33), we repeat the procedure with admissible strategies $(m_1^{\varepsilon}(\cdot), m_2^{\varepsilon}(\cdot))$ for which $\sup_{\varepsilon} J^{\varepsilon}(m_1^{\varepsilon}, m_2^{\varepsilon}) < \infty$. The limit $(x(\cdot), m_1(\cdot), m_2(\cdot))$ might depend on the chosen subsequence. For any convergent subsequence $\{\varepsilon_n\}$, we get

$$\lim_{\varepsilon=\varepsilon_n \to 0} J^{\varepsilon}(m_1^{\varepsilon}, m_2^{\varepsilon})(\pi) = J(m_1,m_2)(\pi).$$

By the weak convergence and the definition of δ-optimality, (6.3.32) and (6.3.33) follow. $\square$

Corollary 6.3.1. *Assume the conditions of Theorem* 6.3.8 *and the value* $V^{\varepsilon}(\widetilde{m}^{\varepsilon})$ *exists for* (6.3.26). *Let* $\overline{m}^{\delta}$ *be a* δ*-optimal policy for* (6.3.29). *Then*

$$\left| V^{\varepsilon}(\widetilde{m}^{\varepsilon}) - J^{\varepsilon}(\overline{m}^{\delta}) \right| \leqslant 2\delta.$$

6.3.7 *Payoff up to First Exit Time*

Another popular payoff structure is payoff up to the first exit time. Let $G \subset \mathbb{R}^n$ be bounded open set with C^2 boundary. Let $r \in C(\overline{G} \times U_1 \times U_2)$. Let $(m_1^{\varepsilon}, m_2^{\varepsilon})$ be admissible and $x^{\varepsilon}(\cdot)$ be the corresponding solution of (6.3.26) with initial law π supported in G. Let

$$\tau^{\varepsilon}(m) = \inf\{t \geqslant 0 : x^{\varepsilon}(t) \notin \overline{G}\}.$$

The payoff up to τ^{ε} to player 1 is

$$J^{\varepsilon}(m) = E \int_0^{\tau^{\varepsilon}(m)} \int r(x^{\varepsilon}(s),\alpha) m_s^{\varepsilon}(d\alpha) ds. \tag{6.3.34}$$

Let $x(\cdot)$ be the solution of (6.3.29) corresponding to $m = (m_1, m_2)$ and

$$\tau(m) = \inf\{t \geqslant 0 : x(t) \notin \overline{G}\},$$

and

$$J(m) = E\int_0^{\tau(m)}\int r(x(s),\alpha)m_s(d\alpha)ds. \tag{6.3.35}$$

The concept of optimal strategies and value are defined as in subsection 6.2.1. The "occupation measure up to $\tau(m)$" denoted by $\nu_{\tau\pi}[m] \in P\left(\overline{G}\times U_1\times U_2\right)$ is defined by

$$\int f d\nu_{\tau\pi}[m] = (E_m\tau)^{-1}E_m\left[\int_0^{\tau(m)}\int f(x(s),\alpha)m_s(d\alpha)ds\right].$$

We need to make the following additional assumptions.

(A6.3.11): $\xi^{\varepsilon}(\cdot)$ is bounded and Markov process.

(A6.3.12): There are $\delta > 0$ and $\beta > 0$ such that for initial condition π supported in G and admissible strategy m, $\inf_{\pi,m} P_\pi\{x(m,t) \notin N_\delta(G)$, some $t \leqslant T\} \geqslant \beta$, where $N_\delta(G)$ is a δ-neighborhood of G.

(A6.3.13): $a(x) = \{a_{ij}(x)\}$ for $x \in G$ is uniformly positive definite.

Theorem 6.3.9. *Assume* (A6.3.11)–(A6.3.13). *Then*

$$\sup_{\varepsilon} E_\pi \tau^{\varepsilon}(m^{\varepsilon}) < \infty. \tag{6.3.36}$$

In addition, if $(x^{\varepsilon}(\cdot),m^{\varepsilon}(\cdot)) \to (x(\cdot),m(\cdot))$, *then,* τ^{ε}, *the exit time also converge.*

Proof. To prove (6.3.36), it is enough to show that there is a $\beta_1 > 0$ such that for any admissible policy sequence $m^{\varepsilon}(\cdot)$, and for initial condition π,

$$\underline{\lim}_{\varepsilon} P_\pi\left\{x^{\varepsilon}(m^{\varepsilon},t) \notin \overline{G},\ \text{some } t \leqslant 2T\right\} \geqslant \beta_1. \tag{6.3.37}$$

It then follows that there is $\varepsilon_0 > 0$, such that,

$$\sup_{\varepsilon<\varepsilon_0} E_\pi \tau^{\varepsilon}(m) < \infty.$$

Suppose (6.3.37) is not true. Then there are $\varepsilon \to 0$ and initial condition π (supported in G), such that,

$$\lim_{\varepsilon} P_\pi\left\{x^{\varepsilon}(m^{\varepsilon},t) \notin \overline{G},\ t \leqslant 2T\right\} = 0. \tag{6.3.38}$$

There is a subsequence also indexed by ε, and an admissible strategy $m(\cdot)$ such that $\{x^{\varepsilon}(m^{\varepsilon},\cdot), m^{\varepsilon}(\cdot)\} \to (x(m,\cdot), m(\cdot))$. Then (6.3.38) contradicts (A6.3.13), Hence, (6.3.37). The last part of the result follows from (A6.3.13) and the weak convergence. □

Theorem 6.3.10. *Under the conditions of Theorem* 6.3.8 *and* (A6.3.11)–(A6.3.13), *the conclusions for Theorem* 6.3.7 *and Theorem* 6.3.8 *hold for model* (6.3.26) *with payoff* (6.3.34).

The results of this section can be directly applied to two person zero-sum differential games with pathwise discounted payoff structure, analogous to the results in Ramachandran [158]. Also, other payoff structures, such as finite horizon payoff, and payoff up to exit time can be handled by some minor modifications. If the coefficients in (6.3.7) are state dependent or even discontinuous, still we can obtain the results of this paper by adapting the methods of Ramachandran [161].

6.4 Two Person Zero-sum Stochastic Differential Game with Multiple Modes, Weak Convergence

In this Section, we are concerned with "near optimal" strategies for two person zero-sum stochastic differential game with multiple modes and driven by a wideband width noise process. Consider a system of following type in the relaxed control setting given by

$$
\begin{aligned}
dx^{\varepsilon} &= \sum_{l=1}^{2} \int b_l\left(x^{\varepsilon}, \theta^{\varepsilon}, \alpha_l\right) m_{lt}^{\varepsilon}(d\alpha_l)dt + \frac{1}{\varepsilon} g\left(x^{\varepsilon}, \theta^{\varepsilon}, \xi^{\varepsilon}\right) dt, \quad \text{and} \\
& P\left(\theta^{\varepsilon}\left(t+\delta t\right) = j | \theta^{\varepsilon}(t) = i, \theta^{\varepsilon}(s), x^{\varepsilon}(s), s \leqslant t\right) \\
& = \lambda_{ij}^{\varepsilon}(x^{\varepsilon}(t), \xi^{\varepsilon}(t))\delta t + o\left(\delta t\right), \quad i \neq j \\
& x^{\varepsilon}(0) = x_0
\end{aligned} \tag{6.4.1}
$$

where $\xi^{\varepsilon}(\cdot)$ is a wide bandwidth noise process, $\theta^{\varepsilon}(\cdot) \in S = \{1, 2, \ldots, N\}$ describes the various modes of the system, and x_0 a prescribed random variable with $x \in \mathbb{R}^d$, d-dimensional Euclidean-space.

We will use the following standard notation,

$$
\int r(x, i, \alpha) m_t(d\alpha) \equiv \int \int r(x, i, \alpha_1, \alpha_2) m_{1t}(d\alpha_1) d(\alpha_2).
$$

The β-discounted payoff to player I for initial condition (x, i) is given by

$$
J^{\varepsilon}\left(m^{\varepsilon}\right)(x, i) = E_{x,i}\left[\int_0^{\infty} \int e^{-\beta t} r(x^{\varepsilon}(t), i, \alpha) m_t^{\varepsilon}(d\alpha) dt\right], \tag{6.4.2}
$$

where $\alpha = (\alpha_1, \alpha_2)$ and $m = (m_1, m_2)$. When ever we need to emphasize on the strategy m^{ε}, we will use $x^{\varepsilon}\left(m^{\varepsilon}, \cdot\right)$ to denote the solution to (6.4.1).

In multi modal case, the state of the system at time is given by a pair $(x(t), \theta(t))$, where $\theta(t) \in S = \{1, 2, \ldots, N\}$. The discrete component $\theta(t)$ describes the various modes of the system. In pursuit-evasion games, when interceptor tries to destroy a specific target, this type of games occur naturally. Due to fast manueuvering by the evader and the corresponding response by the pursuier, the tragectories keep switching rapidly. In these cases models of the type (6.4.1) are more appropriate.

Suppose that the system (6.4.1) is "close" to a game problem modelled by the system (6.4.3), in the sense that if $m^\varepsilon(\cdot)$ is a sequence of "nice" strategies for (6.4.1), then there is a strategy $m(\cdot)$, and corresponding diffusion $x(m,\cdot)$ defined by (6.4.3), such that as $\varepsilon \to 0$, $x^\varepsilon(m^\varepsilon,\cdot)$ converges weakly to $x(m,\cdot)$ ($x^\varepsilon \to x$). We can write,

$$dx(t) = \left[\sum_{l=1}^{2}\int b_l(x(t),\theta(t),\alpha_l(t))m_{lt}(d\alpha_l) + \widetilde{b}(x(t),\theta(t))\right]dt + \sigma(x(t),\theta(t))dw(t), \text{ and}$$

$$P(\theta(t+\delta t)=j \mid \theta(t)=i,\, \theta(s), x(s), s \leqslant t) = \lambda_{ij}(x(t))\delta t + o(\delta t), \quad i \neq j \text{ with}$$

$$x(0)=x\in\mathbb{R}^d,\ \theta(0)=i\in S. \tag{6.4.3}$$

Here $w(\cdot) = [w_1(\cdot),\dots,w_d(\cdot)]'$ is a standard Wiener process. Also,

$$\lambda_{ij} \,:\, \mathbb{R}^d \to \mathbb{R}, \quad 1 \leqslant i,j \leqslant N, \quad \lambda_{ij}(\cdot) \geqslant 0, \quad i \neq j, \quad \sum_{j=1}^{N}\lambda_{ij}(\cdot) = 0.$$

The total discounted payoff to player I is given by

$$J(m)(x,i) = E_{x,i}\left[\int_0^\infty \int e^{-\beta t} r(x(t),i,\alpha)m_t(d\alpha)dt\right]. \tag{6.4.4}$$

Let $\overline{m}(\cdot) = (\overline{m}_1(\cdot),\overline{m}_2(\cdot))$ denote the optimal strategy for the limit diffusion (6.4.3), and let $\overline{m}^\delta(\cdot)$ be a δ-optimal strategy for (6.4.3). For the diffusion models the existence of equilibria in Markov strategies was established in Ghosh, Araposthatis, and Marcus [75] and we have discussed it in Section 3.5. Assume that $\overline{m}^\delta(\cdot)$ are admissible for $x^\varepsilon(\cdot)$ of (6.4.1) and let $V^\varepsilon(\overline{m}^\varepsilon)$ denote the value (when ever it exists, otherwise take upper and lower values) for (6.4.1). Under appropriate conditions, it will be shown that

$$\left|V^\varepsilon(\overline{m}^\varepsilon) - J^\varepsilon(\overline{m}^\delta)\right| \leqslant \delta, \tag{6.4.5}$$

for small $\varepsilon > 0$. The entire problem will be set in relaxed control framework.

6.4.1 *Problem Description*

For completeness, first we will summarize the results corresponding to the diffusion model from Chapter 3. In order to have smooth transition of notations, we will also change few of the notations from Chapter 3. Let U_l, $l = 1,2$ be compact metric spaces (we can take U_i as compact subsets of $\mathbb{R}^d$), and $M_l = P(U_l)$, the space of probability measures on U_i with the topology of weak convergence. Let $M = M_1 \times M_2$ and $U = U_1 \times U_2$. Let $S = \{1,2,\dots,N\}$. Let $w(\cdot)$ in (6.4.1) be a Wiener process with respect to a filtration $\{\Im_t\}$ and let Ω_l, $l = 1,2$ be a compact set in some Euclidean space. A measure valued random variable $m_l(\cdot)$ is an *admissible strategy* for the l^{th} player if $\int_0^t \int f_l(s,\alpha_l)m_l(dsd\alpha_l)$ is progressively measurable for each bounded continuous $f_l(\cdot)$ and $m_l([0,t]\times\Omega_l) = t$, for $t \geqslant 0$. If $m_l(\cdot)$ is admissible then there is a derivative $m_{lt}(\cdot)$ (defined for almost all t) that is non-anticipative with respect to $w(\cdot)$ and

$$\int_0^t \int f_l(s,\alpha_l) m_l(dsd\alpha_l) = \int_0^t ds \int f_l(s,\alpha_l) m_{ls}(d\alpha_l),$$

for all t with probability one (w.p.1.). If $m_l(\cdot) = u_l(x(\cdot),\theta(\cdot))$, for a measurable $u_l : \mathbb{R}^d \times S \to M_l$, then $m_l(\cdot)$ (or by an abuse of notation the map u_l itself) is called *Markov strategy*. The results derived in this subsection are for Markov strategies. A strategy $m_l(\cdot)$ is called *pure* if $m_l(\cdot)$ is a Dirac measure, i.e., $m_l(\cdot) = \delta_{u_l}(\cdot)$, where $u_l(\cdot)$ is a U_l-valued nonanticipative process. We will denote by A_l the set of admissible strategies and M_{al} the set of Markov strategies for the player l. One can introduce appropriate metric topology under which M_{al} is compact, see Borkar and Ghosh [31]. We will denote $A = A_1 \times A_2$, and $M_a = M_{a1} \times M_{a2}$. If for each $l = 1,2,,$ $m_l(\cdot)$ is a Markov strategy then (6.4.3) admits a unique strong solution which is a strong Feller process under the assumption (A3.4.1), see Ghosh, Araposthatis, and Marcus [75].

In relaxed control settings, one chooses at time t a probability measure m_t on the control set M rather than an element $u(t)$ in U. We call the measure m_t the relaxed control at time t. Any ordinary control can be represented as a relaxed control via the definition of the derivative $m_t(d\alpha) = \delta_{u(t)}(\alpha)d\alpha$. Hence, if m_t is an atomic measure concentrated at a single point $m(t) \in M$ for each t, then the relaxed control will be called ordinary control. We will denote the ordinary control by $u_m(t) \in M$.

An admissible strategy $m_1^* \in A_1$ is said to be an *discounted optimal for player I* if for $(x,i) \in \mathbb{R}^d \times S$,

$$J[m_1^*, \widetilde{m}_2](x,i) \geqslant \inf_{m_2 \in A_2} \sup_{m_1 \in A_1} J[m_1,m_2](x,i) \doteq V^+(x,i)$$

for any $\widetilde{m}_2 \in A_2$. The function $V^+ : \mathbb{R}^d \times S \to R$ is called *upper value function* of the game. A strategy $m_1^* \in M_{a_1}$ is called discounted optimal for player I, if it is discounted optimal for all initial laws. Similarly, $m_2^* \in A_2$ is *discounted optimal for player II* if

$$J(\widetilde{m}_1, m_2^*)(x,i) \leqslant \sup_{m_1 \in A_1} \inf_{m_2 \in A_2} J[m_1,m_2](x,i) \doteq V^-(x,i)$$

for any $\widetilde{m}_1 \in A_1$.The function $V^- : \mathbb{R}^d \times S \to R$ is called *lower value function* of the game. If $V^+(x,i) = V^-(x,i)$, then the game is said to admit a value for the discounted criterion and we will denote it by $V(x,i)$, which is called the *value function*. The policies $m_{1\delta}$ and $m_{2\delta}$ are said to be δ-*optimal strategies* for player I and II respectively if

$$\sup_{m_1 \in A_1} J(m_1, m_{2\delta})(x,i) - \delta \leqslant V(x,i) \leqslant \inf_{m_2 \in A_2} J(m_{1\delta}, m_2)(x,i) + \delta. \tag{6.4.6}$$

For $m \in A$ and $(x(\cdot), \theta(\cdot))$ the corresponding process, now we introduce the concept of $\beta-$ *discounted occupation measure* for initial condition $(x,i) \in \mathbb{R}^d \times S$ denoted by $\nu_{x,i}(m) \in$?$(\mathbb{R}^d \times S \times U_1 \times U_2)$ is defined by

$$\sum_{i=1}^{N} \int_{\mathbb{R}^d \times U} f \, d\nu_{x,i}[m] = \beta E_{x,i} \left[\int_0^\infty \int_U e^{-\beta t} f\left(x(t), \theta(t), \alpha\right) m_t(d\alpha) dt \right]$$

for $f \in C_b\left(\mathbb{R}^d \times S \times U\right)$. For notational convenience, we will suppress the dependence on the initial conditions and denote $\nu_{x,i}[m]$ by $\nu[m]$ when ever there is no confusion. In terms of$\nu[m]$, (6.4.4) becomes

$$J[m](x,i) = \beta^{-1} \sum_{j=1}^{N} \int_{\mathbb{R}^d \times U} r\left(x, j, \alpha\right) d\nu_{x,i}[m].$$

Let

$$\nu_{x,i}[A_1, A_2] = \{\nu_{x,i}(m)/m \in A_1 \times A_2\}.$$

$\nu_{x,i}[M_{a1}, A_2]$, $\nu_{x,i}[A_1, M_{a2}]$, $\nu_{x,i}[M_{a1}, M_{a2}]$, etc. are defined analogously. Following result is from Ghosh, Araposthatis, and Marcus [75] which basically states that for the two person zero-sum differential game no player can improve his/her payoff by going beyond Markov strategies

Lemma 6.4.1. *For any fixed initial condition*

$$\nu_{x,i}[A_1, M_{a2}] = \nu_{x,i}[M_{a1}, M_{a2}] = \nu_{x,i}[M_{a1}, A_2].$$

For $p \geqslant 1$, define

$$W^{2,p}_{\text{loc}}(\mathbb{R}^d \times S) = \left\{ f : \mathbb{R}^d \times S \to \mathbb{R} : \text{ for each } i \in S, \; f(\cdot, i) \in W^{2,p}_{\text{loc}}(\mathbb{R}^d) \right\}.$$

Let $b(x,i,\alpha) = b_1(x,i,\alpha_1) + b_2(x,i,\alpha_2) + \widetilde{b}(x)$. For $f \in W^{2,p}_{\text{loc}}\left(\mathbb{R}^d \times S\right)$ and $\alpha \in U$, define the operator

$$\mathscr{A}_i^\alpha f(x,i) = f_x'(x) b(x,i,\alpha) + \frac{1}{2} \sum_{l,j} a_{l,j}(x,i) f_{x_l x_j}(x,i)$$

where $a_{lj}(x,i) = \sum_{k=1}^{d} \sigma_{lk}(x,i) \sigma_{jk}(x,i)$, and

$$\mathscr{A}^\alpha f(x,i) = \mathscr{A}_i^\alpha f(x,i) + \sum_{j=1}^{N} \lambda_{ij} f(x,j).$$

For $m \in M$, define

$$\mathscr{A}^m f(x,i) = \int_U \mathscr{A}^\alpha f(x,i) m_t(d\alpha).$$

We will now state following results from Ghosh, Araposthatis, and Marcus [75].

Theorem 6.4.1. *Under* (A3.4.1) *the Isaacs equation*

$$\inf_{m_1 \in M_1} \sup_{m_2 \in M_2} [\mathscr{A}^m \phi(x,i) + r(x,i,m)] = \sup_{m_2 \in M_2} \inf_{m_1 \in M_1} [\mathscr{A}^m \phi(x,i) + r(x,i,m)]$$
$$= \beta \phi(x,i) \tag{6.4.7}$$

has a unique solution in $C^2(\mathbb{R}^d \times S) \cap C_b(\mathbb{R}^d \times S)$.

Consider the special case in which one player controls the game exclusively for each state $i \in S$. That is, we assume the following

(A6.4.1): Let $S_1 = \{i_1, \ldots, i_m\} \subset S$, $S_2 = \{j_1, \ldots, j_n\} \subset S$ be such that $S_1 \cap S_2 = \phi$ and $S_1 \cup S_2 = S$. Also assume that

$$b(x,i,m_1,m_2) = \overline{b}_1(x,i,m_1)$$

and

$$r(x,i,m_1,m_2) = \overline{r}_1(x,i,m_1)$$

for $i \in S_1$. Similarly for $i \in S_2$

$$b(x,i,m_1,m_2) = \overline{b}_2(x,i,m_2)$$

and

$$r(x,i,m_1,m_2) = \overline{r}_2(x,i,m_2)$$

where $\overline{b}_k$ and $\overline{r}_k$, $k = 1,2$, satisfy same conditions as b and r.

Now we have following result from Ghosh, Araposthatis, and Marcus [75].

Theorem 6.4.2. *Under* (A3.4.1), (A6.4.1)*, each player has Markov optimal pure strategies.*

Deriving similar results for (6.4.1) is very difficult and characterizing the optimal strategies is almost impossible. Our aim is to use the optimal strategies of the system (6.4.3) to system (6.4.1) and obtain weak convergence results. For completeness sake, we introduce some essential results from weak convergence theory and for more details we refer to Kushner [103].

Let $D^d[0,\infty)$ denote the space of $\mathbb{R}^d$ valued functions which are right continuous and have left-hand limits endowed with the Skorohod topology. Following Kushner [103], we define the notion of 'p-lim' and an operator $\widehat{A}^\varepsilon$ as follows. Let $\{\mathfrak{I}_t^\varepsilon\}$ denote the minimal σ-algebra over which $\{x^\varepsilon(s), \theta^\varepsilon(s), \xi^\varepsilon(s),\ s \leqslant t\}$ is measurable, and let E_t^ε denote the expectation conditioned on $\mathfrak{I}_t^\varepsilon$. Let $\widetilde{M}$ denote the set of real valued functions of (ω, t) that are nonzero only on a bounded t-interval. Let

$$\overline{M}^\varepsilon = \left\{ f \in \widetilde{M};\ \sup_t E|f(t)| < \infty, \text{ and } f(t) \text{ is } \mathfrak{I}_t^\varepsilon \text{ measurable} \right\}.$$

Define operator $\widehat{A}^{\varepsilon}$ and it's domain $D\big(\widehat{A}^{\varepsilon}\big)$ as in Section 6.2. For proof the main weak convergence result, Theorem 6.4.4, we will use Lemma (6.2.1) through Lemma(6.2.3), see Kuhner [103].

An *admissible relaxed strategy* $m_l^{\varepsilon}(\cdot)$ for the l^{th} player with system (6.4.1) is a measure valued random variable satisfying $\int\int_0^t f(s,\alpha)\, m^{\varepsilon}(dsd\alpha)$ is progressively measurable with respect to $\{\Im_t^{\varepsilon}\}$, where $\Im_t^{\varepsilon}$ is the minimal σ-algebra generated by $\{\xi^{\varepsilon}(s), x^{\varepsilon}(s), \theta^{\varepsilon}(s),\ s \leqslant t\}$. Also, $m^{\varepsilon}([0,t]\times U) = t$ for all $t \geqslant 0$. Also, there is a derivative m_t^{ε}, where $m_t^{\varepsilon}(b)$ are $\Im_t^{\varepsilon}$ measurable for Borel B. The concept of δ-optimality for the wideband noise driven system $(x^{\varepsilon}(\cdot), \theta^{\varepsilon}(\cdot))$ is similarly defined as in (6.4.6). We will use following assumptions in our analysis.

(A6.4.2): $\xi^{\varepsilon}(t) = \xi\left(t/\varepsilon^2\right)$, where $\xi(\cdot)$ is a stationary process which is strongly mixing, right continuous and bounded with mixing rate function $\phi(\cdot)$ satisfying $\int_0^{\infty}\lambda^{\varepsilon}(\cdot,\cdot)$ is bounded and Lipschitz continuous (uniformly in ε, x, ξ).

(A6.4.3): For each $i \in S$, $g(\cdot,i,\cdot)$, $g_x(\cdot,i,\cdot)$ are continuous (in x,ξ) and satisfy the uniform Lipshitz condition. For each x and i, $Eg(x,i,\xi) = 0$.

(A6.4.4): There are continuous functions $a(\cdot)$ and $\widetilde{b}(\cdot,\cdot)$ such that for each (x,i) and $T_1,\ T_2 \to \infty$ and $T_2 - T_1 \to \infty$, such that,

$$\int_{T_1}^{T_2} Eg_x(x,i,\xi(t))g(x,i,\xi(T_1))dt \to \widetilde{b}(x,i)$$

and

$$\int_{T_1}^{T_2} Eg(x,i,\xi(t))g'(x,i,\xi(T_1))dt \to \frac{1}{2}a(x,i).$$

The convergence is uniform in x. Also assume that there is a Lipshitz continuous square root for $a(x,\theta)$, that is, $a(x,\theta) = \sigma(x,\theta)\sigma'(x,\theta)$.

6.4.2 *Weak Convergence and near optimality*

Now, we will first prove the weak convergence of the wide bandwidth noise system to appropriate controlled diffusion. Then we will obtain convergence of payoffs and strategies as well as a result on near optimality.

Theorem 6.4.3. *Assume conditions* (A3.2.1), (A6.4.2) *to* (A6.4.4). *Let* $\widehat{m}^{\varepsilon}(\cdot) \to \widehat{m}(\cdot)$. *There is a* $w(\cdot)$ *such that* $\widehat{m}(\cdot)$ *is admissible strategy with respect to* $w(\cdot)$ *and* $(x^{\varepsilon}(\cdot), \theta^{\varepsilon}(\cdot), \widehat{m}^{\varepsilon}(\cdot)) \to (x(\cdot), \theta(\cdot), \widehat{m}(\cdot))$ *where*

$$\begin{aligned} dx(t) = &\left[\sum_{l=1}^{2}\int b_l(x(t),\theta(t),\alpha_l(t))m_{lt}(d\alpha_l) + \widetilde{b}(x(t),\theta(t))\right]dt \\ &+\sigma(x(t),\theta(t))dw(t). \end{aligned} \tag{6.4.8}$$

Proof. Since $U\times[0,t_1]$ is compact for each $t_1<\infty$, $\{\widehat{m}^{\varepsilon}(\cdot)\}$ is tight in $M_1(\infty)\times M_2(\infty)$. First we will prove the tightness of $\{x^{\varepsilon,K}(\cdot)\}$. Whenever there is no confusion, for notational convenience, we will use x^{ε} in place of $x^{\varepsilon,K}$ and $\widehat{A}^{\varepsilon}$ for $\widehat{A}_K^{\varepsilon,\widehat{m}}$. For $f(\cdot)\in C_0^3$, we have

$$\widehat{A}^{\varepsilon}f(x,i)=f_x'(x,i)\left[\sum_{l=1}^{2}\int b_l(x,i,\alpha_l)\,m_{lt}^{\varepsilon}(d\alpha_l)+\frac{1}{\varepsilon}g(x,i,\xi^{\varepsilon})\right]+\sum_{j=1}^{N}\lambda_{ij}^{\varepsilon}(x,\xi^{\varepsilon})\,f(x,j)\,.$$

Let there be a continuous function $\widetilde{\lambda}_{ij}(\cdot,\cdot)$ such that $\lambda_{ij}^{\varepsilon}(x,\xi)\to\widetilde{\lambda}_{ij}(x,\xi)$ uniformly on each compact (x,ξ)-set. Now, for each x define $\lambda_{ij}(x)$ by $\lambda_{ij}(x)=\int\widetilde{\lambda}_{ij}(x,\xi)P^x(d\xi)$, where we assume that there is a unique invariant probability measure $P^x(\cdot)$ corresponding to the transition function $P(\xi,l,\cdot\mid x)$, and for each compact set Q the set of invariant measures $\{P^x(\cdot),x\in Q\}$ is tight. We refer to Kushner [103] for a comment on such an assumption.
For arbitrary $T<\infty$ and for $t\leqslant T$, define

$$f_1^{\varepsilon}(t)=f_1^{\varepsilon}\left(x^{\varepsilon,K}(t),i,t\right)$$

where

$$\begin{aligned}f_1^{\varepsilon}(x,i,t)&=\frac{1}{\varepsilon}\int_t^T f_x'(x,i)E_t^{\varepsilon}g_K(x,i,\xi^{\varepsilon}(s))ds\\&=\varepsilon\int_{t/\varepsilon^2}^{T/\varepsilon^2} f_x'(x,i)E_t^{\varepsilon}g_K(x,i,\xi(s))ds.\end{aligned}$$

From, (A3.2.1) and (A6.4.2), $\lim_{\varepsilon}E|f_1^{\varepsilon}(t)|=0$. We have

$$\widehat{A}^{\varepsilon}f_1^{\varepsilon}(t)=-\frac{1}{\varepsilon}f_x'\left(x^{\varepsilon,K}(t),i,\xi^{\varepsilon}(s)\right)$$

$$+\frac{1}{\varepsilon}\int_t^T ds\left[f_x'\left(x^{\varepsilon,K}(t),i\right)E_t^{\varepsilon}g_K\left(x^{\varepsilon,K}(t),i,\xi^{\varepsilon}(s)\right)\right]_x'.\overset{\varepsilon}{x}(t)+o(1)$$

where p-$\lim_{\varepsilon}o(1)=0$ uniformly in t.
Define $f^{\varepsilon}(t)=f\left(x^{\varepsilon,K}(t),i\right)+f_1^{\varepsilon}(t)$. Writing x for $x^{\varepsilon,K}(t)$ and a scale change $s/\varepsilon^2\to s$, we have

$$\begin{aligned}\widehat{A}^{\varepsilon}f^{\varepsilon}(t)&=f_x'(x,i)\left[\int b_{1K}(x,i,\alpha_1)\widehat{m}_{1t}^{\varepsilon}(d\alpha_1)+\int b_{2K}(x,i,\alpha_2)\widehat{m}_{1t}^{\varepsilon}(d\alpha_2)\right]\\&+\sum_{j=1}^{N}\lambda_{ij}^{\varepsilon}(x,\xi^{\varepsilon})f(x,j)+\int_{t/\varepsilon^2}^{T/\varepsilon^2}dsE_t^{\varepsilon}[g_K(x,i,\xi(s))]_x'g_K(x,i,\xi^{\varepsilon}(t))\\&+\varepsilon\int_{t/\varepsilon^2}^{T/\varepsilon^2}dsE_t^{\varepsilon}[g_K(x,i,\xi(s))]_x'\\&\cdot\left[\int b_{1K}(x,i,\alpha_1)\widehat{m}_{1t}^{\varepsilon}(d\alpha_1)+\int b_{2K}(x,i,\alpha_2)\widehat{m}_{1t}^{\varepsilon}(d\alpha_2)\right]\\&+\varepsilon\sum_{j=1}^{N}\lambda_{ij}^{\varepsilon}(x,\xi^{\varepsilon})[f_x(x,j)g_K(x,j,\xi(s))]_x'.\end{aligned}\tag{6.4.9}$$

under (A3.2.1), (A3.2.2), and (A6.4.3), the third term in (6.4.12) is 0(1) and the next two terms go to zero in p-limit as $\varepsilon \to 0$. Then for each $T < \infty$, $\{\widehat{A}^{\varepsilon} f^{\varepsilon}(t), \varepsilon > 0, t \leqslant T\}$ is uniformly integrable and for $k > 0$, and

$$\lim_{\varepsilon} P\left\{\sup_{t\leqslant T} |f^{\varepsilon}(t) - f(x^{\varepsilon}(t))| \geqslant k\right\} = 0.$$

Now, by Lemma 6.2.3, $\{x^{\varepsilon,K}(\cdot)\}$ is tight in $D^d[0,\infty)$. Index by ε, a weakly convergent subsequence of $\{x^{\varepsilon,K}, \theta^{\varepsilon}, \widehat{m}^{\varepsilon}\}$, i.e.,

$$\{x^{\varepsilon,K}, \theta^{\varepsilon}, \widehat{m}^{\varepsilon}\} \Longrightarrow \{x^K(\cdot), \theta(\cdot), \widehat{m}(\cdot)\}.$$

There is progressively measurable $\widehat{m}_t(\cdot)$ such that $\widehat{m}_t(u) = 1$ and

$$\int_t \int 0 f(s,\alpha)\widehat{m}_s(d\alpha)ds = \int_t \int 0 f(s,\alpha)\widehat{m}(ds \times d\alpha)$$

for each continuous $f(\cdot)$. By Lemma 6.2.2, the proof will be complete if we verify (6.2.1) and (6.2.2). Now, treat (x,i) as parameters, we will average out the noise term only by using the perturbed test function methods as introduced in Kushner [103]. Define

$$f_2^{\varepsilon}(t) = \int_t^T \int_{v/\varepsilon^2}^{T/\varepsilon^2} \left[E_t^{\varepsilon} f_x'(x,i) g_{Kx}(x,i,\xi(s)) g_K(x,i,\xi^{\varepsilon}(v)) \right.$$
$$\left. - E f_x'(x,i) g_{Kx}(x,i,\xi(s)) g_K(x,i,\xi^{\varepsilon}(v))\right] dsdv$$

and

$$f_3^{\varepsilon}(t) = \int_t^T \int_{v/\varepsilon^2}^{T/\varepsilon^2} \left[E_t^{\varepsilon} f_{xx}'(x,i) g_K'(x,i,\xi(s)) g_K(x,i,\xi^{\varepsilon}(v)) \right.$$
$$\left. - E f_{xx}'(x,i) g_K'(x,i,\xi(s)) g_K(x,i,\xi^{\varepsilon}(v))\right] dsdv$$

with a scale change $v \to v/\varepsilon^2$, that is,

$$f_2^{\varepsilon}(t) = \varepsilon^2 \int_{t/\varepsilon^2}^{T/\varepsilon^2} \int_v^{T/\varepsilon^2} \left[E_t^{\varepsilon} f_x'(x,i) g_{Kx}(x,i,\xi(s)) g_K(x,i,\xi(v)) \right.$$
$$\left. - E f_x'(x,i) g_{Kx}(x,i,\xi(s)) g_K(x,i,\xi(v))\right] dsdv.$$

and

$$f_3^{\varepsilon}(t) = \varepsilon^2 \int_{t/\varepsilon^2}^{T/\varepsilon^2} \int_v^{T/\varepsilon^2} \left[E_t^{\varepsilon} f_{xx}'(x,i) g_K'(x,i,\xi(s)) g_K(x,i,\xi(v)) \right.$$
$$\left. - E f_{xx}'(x,i) g_K'(x,i,\xi(s)) g_K(x,i,\xi(v))\right] dsdv.$$

From

$$\lim_{\varepsilon} E \sup_{t\leqslant T} |f_i^{\varepsilon}(t)| = 0 \text{ for } i = 2,3.$$

Define

$$\widetilde{f}^{\varepsilon}(t) = f(x) + \sum_{i=1}^{3} f_i^{\varepsilon}(t). \tag{6.4.10}$$

Then, we have

$$p\text{-}\lim_{\varepsilon}\left(\widetilde{f}^{\varepsilon}(t)-f(x)\right)=0 \tag{6.4.11}$$

$$\begin{aligned}\widehat{A}^{\varepsilon} f_2^{\varepsilon}(t) &= o(1)+\int_{t/\varepsilon^2}^{T/\varepsilon^2}[Ef_x'(x,i)g_{Kx}(x,i,\xi(s))g_K(x,i,\xi^{\varepsilon}(t))\\ &-E_t^{\varepsilon} f_x'(x,i)g_{Kx}(x,i,\xi(s))g_K(x,i,\xi^{\varepsilon}(t))]ds+\varepsilon^2\int_{t/\varepsilon^2}^{T/\varepsilon^2}dv\int_{v}^{T/\varepsilon^2}ds\\ &\times E_t^{\varepsilon} f_x'(x,i)g_K'(x,i,\xi(s))g_K(x,i,\xi(v))\\ &-Ef_x'(x,i)g_K'(x,i,\xi(s))g_K(x,i,\xi(v))]_x.\overset{\varepsilon,K}{x}\end{aligned} \tag{6.4.12}$$

and

$$\begin{aligned}\widehat{A}^{\varepsilon} f_3^{\varepsilon}(t) &= o(1)+\int_{t/\varepsilon^2}^{T/\varepsilon^2}[Ef_{xx}'(x,i)g_K'(x,i,\xi(s))g_K(x,i,\xi^{\varepsilon}(t))\\ &-E_t^{\varepsilon} f_{xx}'(x,i)g_K'(x,i,\xi(s))g_K(x,i,\xi^{\varepsilon}(t))]ds+\varepsilon^2\int_{t/\varepsilon^2}^{T/\varepsilon^2}dv\int_{v}^{T/\varepsilon^2}ds\\ &\times E_t^{\varepsilon} f_{xx}'(x,i)g_K'(x,i,\xi(s))'g_K(x,i,\xi(v))\\ &-Ef_{xx}'(x,i)g_K'(x,i,\xi(s))'g_K(x,i,\xi(v))].\overset{\varepsilon,K}{x}.\end{aligned} \tag{6.4.13}$$

From (A6.4.2) and (A6.4.3), the p-limit of the last three terms on the right-hand side of (6.4.12) and (6.4.13) are all zero.

From (6.4.9), (6.4.10), (6.4.12) and (6.4.13), we can write

$$\begin{aligned}\widehat{A}^{\varepsilon}\widetilde{f}^{\varepsilon}(t) &= \widehat{A}^{\varepsilon} f^{\varepsilon}(t)+\widehat{A}^{\varepsilon} f_2^{\varepsilon}(t)+\widehat{A}^{\varepsilon} f_3^{\varepsilon}(t)=o(1)\\ &+f_x'(x,i)\left[\int b_{1K}(x,i,\alpha_1)\widehat{m}_{1t}^{\varepsilon}(d\alpha_1)+\int b_{2K}(x,i,\alpha_2)\widehat{m}_{1t}^{\varepsilon}(d\alpha_2)\right]\\ &+\sum_{j=1}^{N}\lambda_{ij}^{\varepsilon}(x,\xi^{\varepsilon})f(x,j)+\int_{t/\varepsilon^2}^{T/\varepsilon^2}E\left(f_x'(x,i)g_{Kx}(x,i,\xi(s))g_K(x,i,\xi^{\varepsilon}(t))\right)ds\\ &+\int_{t/\varepsilon^2}^{T/\varepsilon^2}E\left(f_{xx}(x,i)g_K(x,i,\xi(s)')g_K(x,i,\xi^{\varepsilon}(t))\right)ds,\end{aligned} \tag{6.4.14}$$

where $p\text{-}\lim_{\varepsilon} o(1)=0$ uniformly in t.

Equation (6.4.14) together with (A6.4.2) to (A6.4.4) yield (6.2.3). Applying Lemma 6.2.2, we have $x^{\varepsilon,K}(\cdot)\to x^K(\cdot)$.

Let $h(\cdot)$ be bounded continuous (except possibly on a countable set t_l and take t and $t+s$ outside this set) and let $f(\cdot)\in C_0^2$. Let q_1 and q_2 be arbitrary integers and $k_j(\cdot)$ arbitrary bounded and continuous functions. Taking limit as $\varepsilon\to 0$ and using the Skorokhod imbedding so that the weak convergence becomes w.p.1. in the topology of the space

$D^d[0,\infty)\times S\times M_1(\infty)\times M_2(\infty)$, we obtain

$$Eh\left(x^K(t_l),\theta(t_l),(k_j,\widehat{m})_{t_l},\ \ l\leqslant q_1,\ \ j\leqslant q_2\right).\left(f\left(x^K(t+s),i'\right)-f\left(x^K(t),i\right)\right.$$
$$\left.-\int_t^{t+s}\widehat{A}_K^{\widehat{m}}f\left(x^K(s),i\right)ds\right)=0. \tag{6.4.15}$$

Since $q_1,\ q_2,\ h(\cdot),\ k_j(\cdot),\ t_l,\ t,\ s$ are arbitrary, (6.4.18) implies that $x^K(\cdot)$ solves the martingale problem with operator $\widehat{A}_K^{\widehat{m}}$. It then follows that there is a standard Wiener process $w^K(\cdot)$ and $x^K(\cdot)$ is nonanticipative with respect to $w^K(\cdot)$ and satisfies (6.3.8) with $b(\cdot,\cdot,\cdot)$, $\sigma(\cdot,\cdot)$ and $w(\cdot)$ replaced by $b^K(\cdot,\cdot,\cdot)$, $\sigma^K(\cdot,\cdot)$ and $w^K(\cdot)$, respectively. Moreover, $\widehat{m}_i(A_i\times[0,t])$ and $\widehat{m}_{i,t}(A_i)$, $i=1,2$ are progressively measurable with respect to $w^K(\cdot)$. Hence $\widehat{m}(\cdot)$ is admissible (admissible Markov) strategies for the problem with coefficients b^K and σ^K. Define $\tau^K=\min\{t\geqslant 0:|x(t)|\geqslant K\}$. Let $w(\cdot)$ be any Wiener process such that $\widehat{m}_i(\cdot)$, $i=1,2$ are admissible with respect to $w(\cdot)$.For any given initial condition and with $(w(\cdot),\widehat{m}_1(\cdot),\widehat{m}_2(\cdot))$, (6.3.8) has unique solution whose distribution does not depend on the particular $w(\cdot)$. In addition, $P\{\tau^K\leqslant T\}\to 0$ as $K\to\infty$ for each $T<\infty$. Therefore, $\{x^\varepsilon(\cdot),\theta^\varepsilon(\cdot),\widehat{m}^\varepsilon(\cdot)\}$ is tight and converges weakly to a solution of (6.3.8). □

In the previous section, we have given the results concerning the existence of optimal strategy pair $(\overline{m}_1,\overline{m}_2)$ and a value for the system (6.4.8) under additional condition (A6.4.1). Also, the value function is characterized as the unique solution of the Isaacs equation in $C^2(\mathbb{R}^d\times S)\cap C_b(\mathbb{R}^d\times S)$. For the weak convergence methods, assumption (A6.4.1) is not crucial. Since it is possible that existence of optimal strategies for the system (6.4.8) could be proved, we will make following assumption.

(A6.4.5): Let (6.4.8) have a unique weak sense solution for strategy pair $(\overline{m}_1,\overline{m}_2)$ and let this strategy be unique. Assume $(\overline{m}_1,\overline{m}_2)$ is admissible for $x^\varepsilon(\cdot)$ of (6.4.1) for small ε.

Now, we will give a result on convergence of payoffs as well as the near optimality of the optimal controls of (6.4.8) to the system (6.4.1).

Theorem 6.4.4. *Assume conditions* (A3.4.1), (A6.4.2) *to* (A6.4.5). *Then, letting* $x^\varepsilon(\cdot)$ *denote the solution of* (6.4.1) *controlled by the policy pair* $(\overline{m}_1,\overline{m}_2)$, *we have* $\{x^\varepsilon(\cdot),\theta^e(\cdot),\overline{m}_1,\overline{m}_2\}\to(x(\cdot),\theta(\cdot),\overline{m}_1,\overline{m}_2)$ *and there is a Wiener process* $w(\cdot)$ *such that* $(x(\cdot),\theta(\cdot),\overline{m}_1,\overline{m}_2)$ *is nonanticipative with respect to* $w(\cdot)$, *and* (6.4.8) *holds. Also,*

$$J^\varepsilon(\overline{m}_1,\overline{m}_2)(x,i)\to J(\overline{m}_1,\overline{m}_2)(x,i)=V(x,i). \tag{6.4.16}$$

In addition, let $(\widehat{m}_1^\varepsilon,\widehat{m}_2^\varepsilon)$ *be a* δ*-optimal strategy pair for player I and II respectively with* $x^\varepsilon(\cdot)$ *of* (6.4.1). *Then*

$$\varlimsup_{\varepsilon}\left|\left[\sup_{m_1^\varepsilon\in A_1}J^\varepsilon(m_1^\varepsilon,\widehat{m}_2^\varepsilon)(x,i)-J^\varepsilon(\overline{m}_1,\overline{m}_2)(x,i)\right]\right|\leqslant\delta \tag{6.4.17}$$

and

$$\overline{\lim_{\varepsilon}}\left|\left[\sup_{m_2^\varepsilon\in A_2} J^\varepsilon(\widehat{m}_1^\varepsilon, m_2^\varepsilon)(x,i) - J^\varepsilon(\overline{m}_1,\overline{m}_2)(x,i)\right]\right| \leqslant \delta \tag{6.4.18}$$

Proof. From Theorem 6.4.4 and by the uniqueness, we have

$$(x^\varepsilon(\cdot),\theta^\varepsilon(\cdot),\overline{m}_1,\overline{m}_2) \to (x(\cdot),\theta(\cdot),\overline{m}_1,\overline{m}_2).$$

The limit satisfies equation (6.4.8) for some Wiener process $w(\cdot)$.
Fix K, and let S_k be the K-ball in $\mathbb{R}^d$. Let

$$\tau^{\varepsilon,K} = \inf\{t : x^\varepsilon(t) \notin S_K\}, \text{ and } \tau^K = \inf\{t : x(t) \notin S_K\}.$$

By (A6.4.4), $a(\cdot)$ is nondegenarate. This with the properties of Wiener process $w(\cdot)$, $\tau^K(\cdot)$ is continuous w.p.1. with respect to the measure induced by $x(\cdot)$. Now, since

$$\{x^\varepsilon(\cdot),\theta^\varepsilon(\cdot),\overline{m}_1,\overline{m}_2\} \to (x(\cdot),\theta(\cdot),\overline{m}_1,\overline{m}_2)$$

and by the continuity of $\tau^K(\cdot)$, $\tau^{\varepsilon,K} \Rightarrow \tau^K$ as $\varepsilon \to 0$. Let

$$J^K(m)(x,i) = E_{x,i}\int_0^{\tau^K}\int e^{-\lambda t} r(x,i,\alpha)m_t(d\alpha)dt$$

and let the occupation measure $\nu^K_{x,i}[m]$ be $\nu_{x,i}[m]$ corresponding to $J^K(m)(x,i)$. Now as $K \to \infty$, we have $\tau^K \to \infty$, w.p.1., since $x \in \mathbb{R}^d$. Hence, all steps below can first be worked with the truncated version and then, as in the proof of Theorem 6.4.3, take the limit $K \to \infty$. For simplicity of notation, we will not carry the $K-$ subscript.
By Lemma 6.4.1, we have

$$\begin{aligned}
\lim_{\varepsilon} J^\varepsilon(\overline{m})(x,i) &= \lim_{\varepsilon} \lambda \int r(x^\varepsilon,i,\alpha)d\nu_{x,i}[\overline{m}] \\
&= \lambda \int r(x,i,\alpha)d\nu_{x,i}[\overline{m}] \\
&= J(\overline{m}_1,\overline{m}_2)(x,i) = V(x,i).
\end{aligned}$$

To show (6.4.17) and (6.4.18), repeat the procedure with admissible strategies $(m_1^\varepsilon(\cdot), m_2^\varepsilon(\cdot))$ for which $\sup_\varepsilon J^\varepsilon(m_1^\varepsilon, m_2^\varepsilon)(x,i) < \infty$. The limit $(x(\cdot),\theta(\cdot),m_1(\cdot),m_2(\cdot))$ might depend on the chosen subsequence. For any convergent subsequence $\{\varepsilon_n\}$, we get

$$\lim_{\varepsilon=\varepsilon_n\to 0} J^\varepsilon(m_1^\varepsilon, m_2^\varepsilon)(x,i) = J(m_1,m_2)(x,i).$$

Now by the weak convergence and the definition of δ-optimality, (6.4.18) and (6.4.19) follows. □

From the previous result, we can now conclude that, whenever there is a value for the wideband noise driven system, the absolute difference between that value and the payoff using the δ-optimal strategies of the limit diffusion is negligible.

Corollary 6.4.1. *Assume conditions of Theorem* 6.4.5 *and that the value* $V^{\varepsilon}\left(\widetilde{m}^{\varepsilon}\right)$ *exists for* (6.4.1). *Let* $\overline{m}^{\delta}$ *be a* δ*-optimal policy for* (6.4.8). *Then*

$$\left|V^{\varepsilon}\left(\widetilde{m}^{\varepsilon}\right)(x,i)-J^{\varepsilon}\left(\overline{m}^{\delta}\right)(x,i)\right|\leqslant 2\delta. \tag{6.4.19}$$

Remark 6.4.1. If a value exists for the system (6.4.1) for each ε with the strategies $(\widetilde{m}_1^{\varepsilon},\widetilde{m}_2^{\varepsilon})$, then by the weak convergence and the uniqueness of the limit, we can write

$$J^{\varepsilon}\left(\widetilde{m}_1^{\varepsilon},\widetilde{m}_2^{\varepsilon}\right)(x,i)\to V(x,i).$$

In this subsection, we have showed that for a wideband noise driven system, using optimal policies of the limit diffusion will result in near optimal policies for the physical system if the parameter ε is small. This is a robustness statement on the diffusion model. Also, with the results of this paper, it is possible to develop numerical results as in Kushner and Dupuis [112]. It is also possible to derive this type of results for other payoff criterion such as ergodic payoffs, or payoffs to first exit time.

6.5 Partially Observed Stochastic Differential Games

In practical differential games difficulties are often encountered in obtaining information about the state of the system due to time lag, high cost of obtaining data, or simply asymmetry in availability of information due to the nature of the problems in a competitive environment. Stochastic differential games with imperfect state informations are inherently very difficult to analyze. In the literature, there are various information structures considered such as both players will have the same information as in the from broadcasting channel, Ho [88], Sun and Ho [184], or the two players will have available only noise-corrupted output measurements, Rhodes and Luenberger [166, 167]. There are various other possibilities, such as one player will have full information where as the other player will have only partial information or only a deterministic information. A fixed duration stochastic two-person nonzero-sum differential game in which one player has access to closed-loop nonanticipatory state information while the other player makes no observation is considered in Basar [10]. A comprehensive study on partially observed stochastic differential games is still far from being solved. In this subsection, we will present a linear system with quadratic cost functional and imperfect state information. Solution to the diffusion model is given and a weak convergence method is described. We will also deal with a form of nonlinearity.

The system under consideration is of the following type, where both players have the same information such as from a broadcasting channel.

$$dx = [A(t)x + B(t)u - C(t)v]dt + Ddw_1(t) \tag{6.5.1}$$

with observation data

$$y = Hxdt + Fdw_2(t) \tag{6.5.2}$$

and payoff

$$J(u,v) = E\left\{x'(t)Sx(t) + \int_0^T \left[u'Ru - v'Qv\right] dt\right\}. \tag{6.5.3}$$

Here, we are concerned with a partially observed two person zero-sum stochastic differential games driven by wide band noise. The actual physical system will be more naturally modeled by

$$\dot{x^{\varepsilon}} = Ax^{\varepsilon} + Bu - Cv + D\xi_1^{\varepsilon} \tag{6.5.4}$$

with observations

$$\dot{y^{\varepsilon}} = Hx^{\varepsilon} + \xi_2^{\varepsilon} \tag{6.5.5}$$

where ξ_i^{ε}, $i = 1,2$ are wide band noise processes. Let the payoff be given in linear quadratic form

$$J^{\varepsilon}(u^{\varepsilon}, v^{\varepsilon}) = E\left\{x^{\varepsilon'}(t)Sx^{\varepsilon}(t) + \int_0^T \left[u^{\varepsilon'}Ru^{\varepsilon} - v^{\varepsilon'}Qv^{\varepsilon}\right] dt\right\} \tag{6.5.6}$$

for some $T < \infty$.

Typically, one decides upon a suitable model (6.5.4), (6.5.5), (6.5.6), obtains a good or optimal policy pair, and uses this policy to the actual physical system. In this case, it is not clear the value of the determined policy for the physical system, as well as the value of the output of the filter for making estimates of functional of the physical process $x^{\varepsilon}(\cdot)$ which is approximated by $x(\cdot)$. The filter output will rarely be nearly optimal for use in making such estimates, and the policies based on the filter outputs will rarely be "nearly" optimal. In the case of game problem, very little attention has been devoted to such problems. Under quite broad conditions, we will obtain a very reasonable class of alternative filters and policies for the physical system with respect to which it is nearly optimal.

We begin with a discussion of game problem for the ideal white noise linear model (6.5.4), (6.5.5), (6.5.6) and use the Kalman-Bucy filter for this model to obtain an optimal strategy pair for the game problem. Then we will describe the wide bandwidth analogue and give results on filtering and near optimal policies. Also we will include the study on the asymptotic in time and bandwidth problem. Some extensions to partly nonlinear observations will also be given.

6.5.1 *The Diffusion Model*

Consider the linear quadratic Gaussian (LQG) games, that is,

$$dx = [A(t)x + B(t)u - C(t)v]dt + Ddw_1(t) \tag{6.5.7}$$

where $A,\ B,\ C,\ D$ are $n\times n$, $n\times m$, $n\times s$, and $n\times r$ matrices whose elements are continuous in $[0,T]$. Also, $x\in\mathbb{R}^n$ is the state vector with initial state x_0, which is normally distributed with mean $\bar{x}_0$ and variance M_0, $N(\bar{x}_0, M_0)$. Players I and II are endowed with measurements,

$$dy = dy_1 = dy_2 = Hxdt + Fdw_2(t), \tag{6.5.8}$$

where F is of full rank with $p\times q, q>p$ matrix. The objective functional is defined with

$$J(u,v) = E\{x'(t)Sx(t) + \int_0^T [u'Ru - v'Qv]dt\} \tag{6.5.9}$$

where $S\geqslant 0$, $R(t)>0$, $Q(t)>0$ are $n\times n$, $m\times m$, and $n\times s$ symmetric matrices whose elements are continuous on $[0,T]$. Let $R_0 = FF'$ be positive definite (denoted by $R_0>0$). Note that the $-v'Qv$ term is due to the fact that v is minimizing.

The policies u and v take values in compact sets U and V, and sets Ξ_1 and Ξ_2 denote the set of U and V-valued measurable (t,ω) functions on $[0,T]\times C[0,T]$ ($C[0,T]$ is the space of real valued continuous functions on $[0,T]$ with the topology of uniform convergence) which are continuous w.p.1., relative to the Wiener measure. Let Ξ_{1t} and Ξ_{2t} denote the sub class which depends only on the function values up to time t. Let $\Xi = \Xi_1\times\Xi_2$ and $\Xi_t = \Xi_{1t}\times\Xi_{2t}$. We view functions in Ξ as the data dependent policies with values $u(y(\cdot),t)$ and $v(y(\cdot),t)$ at time t and data $y(\cdot)$. Let $\overline{\Xi}$ denote the sub class of functions $(u,v)\in\Xi$ such that $(u(\cdot,t),v(\cdot,t))\in\Xi_t$ for all t and with the use of policies $(u(y,\cdot),v(y,\cdot))$, (6.5.7) has a unique solution in the sense of distributions. These pairs $(u(y,\cdot),v(y,\cdot))$ are the *admissible strategies. We* say that an admissible pair $(u^*(t),v^*(t))$ is a *saddle point* for the game iff

$$J(u(t),v^*(t)) \leqslant J(u^*(t),v^*(t)) \leqslant J(u^*(t),v(t)), \tag{6.5.10}$$

where $u(t)$ and $v(t)$ are any admissible control laws. We call $(u^*(t),v^*(t))$ the optimal strategic pair. Admissible strategies $\widehat{u}$ and $\widehat{v}$ are called δ-*optimal* for players I and II, respectively, if

$$\sup_u J(u,\widehat{v}) - \delta \leqslant J(u^*,v^*) \leqslant \inf_v J(\widehat{u},v) + \delta. \tag{6.5.11}$$

Let $G_t = \sigma\{y(s),\ s\leqslant t\}$ and $\widehat{x}(\tau) = E\{x(\tau)/G_\tau; u(\tau), v(\tau)\}$. For (6.5.7), (6.5.8), the classical Kalman-Bucy filter equations are given by

$$d\widehat{x} = (A\widehat{x} + Bu - Cv)\,dt + L(t)\,(dy - H\widehat{x}dt) \tag{6.5.12}$$

and

$$L(t) = P(t)H'(t)R_0^{-1}(t),$$

with $\widehat{x}_0 = \overline{x}_0$ and $P(t) = E\{(x(t) - \widehat{x}(t))\,(x(t) - \widehat{x}(t))'\}$ is the error covariance matrix and is the unique solution to the matrix Riccati equation:

$$\dot{P} = FP + PF' - PN(y)P + DD' \tag{6.5.13}$$

$P_0 = M_0$, $N(y) = H'R_0^{-1}H$, and the Riccati equation is of the form

$$\dot{\Sigma} = -\Sigma A - A'\Sigma + [\Sigma BR^{-1}B' - CQ^{-1}C']\Sigma \tag{6.5.14}$$

with the boundary condition

$$\Sigma(t) = S'(t)S(t).$$

The following result can be obtained from Ho [88] and Sun and Ho [184].

Theorem 6.5.1. *The optimal strategy pair for the problem* (6.5.7), (6.5.8), (6.5.9) *exists. The optimal pair at time t is*

$$u^*(t) = -R^{-1}(t)B'(t)\Sigma(t)\widehat{x}(t)$$

and

$$v^*(t) = -Q^{-1}(t)C'(t)\Sigma(t)\widehat{x}(t). \tag{6.5.15}$$

Furthermore,

$$\begin{aligned} J(u^*,v^*) = \int_0^T Tr\Sigma(s)[DD' + \left(B(s)R^{-1}B'(s) - C(s)Q^{-1}C'(s)\right)\Sigma(s)P(s)]ds \\ + Tr\Sigma_0 M_0, \end{aligned} \tag{6.5.16}$$

where P satisfies (6.5.13).

6.5.2 *Finite TimeFiltering and Game, Wide Band Noise Case*

Now consider the wide bandwidth analogue of the previous filtering and game problem. Let the system be defined by

$$\dot{x}^{\varepsilon} = Ax^{\varepsilon} + Bu - Cv + D\xi^{\varepsilon} \tag{6.5.17}$$

with observations $\dot{y}^{\varepsilon}(\cdot)$, where

$$\dot{y}^{\varepsilon} = Hx^{\varepsilon} + \xi_2^{\varepsilon}, \tag{6.5.18}$$

where $\int_0^t \xi_i^\varepsilon(s)ds = W_i^\varepsilon(t), i = 1,2$, $W_1^\varepsilon(\cdot)$ and $W_2^\varepsilon(\cdot)$ are mutually independent. Let $W_i^\varepsilon(\cdot) \to W_i(\cdot)$, standard Wiener processes. Let the corresponding objective functional be given by

$$J^\varepsilon(u,v) = E\{x^{\varepsilon'}(t)Sx^\varepsilon(t) + \int_0^T [u'Ru - v'Qv]dt. \tag{6.5.19}$$

In practice, with physical wide band observation noise and state process are not driven by the ideal white noise, one uses (6.5.12), (6.5.13) and the natural adjustment of (6.5.12), that is

$$\dot{x^\varepsilon} = (A\widehat{x}^\varepsilon + Bu - Cv) + L(t)[\dot{y^\varepsilon} - H\widehat{x}^\varepsilon]. \tag{6.5.20}$$

First of all we want to know in what way the triple (6.5.20), (6.5.13), (6.5.14) makes sense. In general, it is not an optimal filter for the physical observation. Instead of asking whether it is nearly optimal, we will ask, with respect to what class of alternative estimators is it nearly optimal when estimating specific functional of $x^\varepsilon(\cdot)$? Another problem is that if one obtains a policy (optimal or not) based on the white noise driven limit model, the policy will be a function of the outputs of the filters. The value of applying this to the actual wide band width noise system is not clear. If one use the model (6.5.7), (6.5.8), (6.5.12) to get a optimal (or nearly optimal) policy pair for the value (6.5.9), and apply this to the physical system, the question then is with respect to what class of comparison policies is such a policy nearly optimal? In both the cases, weak convergence theory can provide some answers. In subsequent results, in order to avoid lengthy calculations, we will not give the weak convergence proofs. Reader can obtain the necessary steps from Kushner [103] and Ramachandran [158].

Even when $W_2^\varepsilon(\cdot) \to W_2(\cdot)$, a non degenerate Wiener process, $y^\varepsilon(\cdot)$ might contain a great deal more information about $x^\varepsilon(\cdot)$ than $y(\cdot)$ does about $x(\cdot)$. We give the following example from Kushner and Runggaldier [115] for an extreme case when $B \equiv 0$ and $C \equiv 0$. We will represent the corresponding process $z^\varepsilon(\cdot)$.

Example 6.5.1. Let $t_i^\varepsilon, i > 0$ be a strictly increasing sequence of real numbers for each ε, such that $t_i^\varepsilon \xrightarrow{i} \infty$ and $\sup_i |t_{i+1}^\varepsilon - t_i^\varepsilon| \xrightarrow{\varepsilon} 0$. Define $\Delta_i^\varepsilon = t_{2i+1}^\varepsilon - t_{2i}^\varepsilon$, and for any $t > 0$, let $\sum_{t_i^\varepsilon \leqslant t} \Delta_i^\varepsilon \xrightarrow{\varepsilon} 0$. Define a new observation noise $\xi_y^\varepsilon(\cdot)$ by resetting $\xi_y^\varepsilon(t) = 0$ for $t \in (t_{2i}^\varepsilon, t_{2i+1}^\varepsilon)$, all i. The integral of the $\xi_y^\varepsilon(\cdot)$ still converges weakly to the Wiener process $W_2(\cdot)$. But $Hz^\varepsilon(\cdot)$ is exactly known for small ε.

The following result Kushner and Runggaldier [115], shows that we never gain information on going to the limit.

Lemma 6.5.1. *Let* $(Z_n, Y_n) \to (Z, Y)$. *Then*

$$\overline{\lim_n}\, E[Z_n - E(Z_n \mid Y_n)]^2 \leqslant E[Z - E(Z \mid Y)]^2$$

We shall now consider a class of estimators that play an integral part in the subject area. By earlier assumptions,we have $(Z^{\varepsilon}(\cdot), W_2^{\varepsilon}(\cdot)) \to (Z(\cdot), W_2(\cdot))$ as $\varepsilon \to 0$. By the weak convergence and independence of $z^{\varepsilon}(\cdot)$ and $\xi_2^{\varepsilon}(\cdot)$, $w_1(\cdot)$ is independent of $w_2(\cdot)$. The weak limit of $y^{\varepsilon}(\cdot)$ is $y(\cdot)$, and thus, the solution of (6.5.8).

Let $\aleph$ denote the class of measurable functions on $C[0,\infty)$ which are continuous w.p.1 relative to Wiener measure. Hence, they are continuous w.r.t. the measure of $y(\cdot)$. Let $\aleph_t$ denote the sub class which depends only on the function values up to time t. For arbitrary $f(\cdot) \in \aleph$ or in $\aleph_t$, $f(y^{\varepsilon}(\cdot))$ will denote an alternative estimator of a functional of $z^{\varepsilon}(\cdot)$. We consider $\aleph$ and $\aleph_t$ as a class of data processors.

We now obtain a robustness result. Let (m_t^{ε}, q) be the integral of a function $q(z)$ with respect to the Gaussian probability distribution with mean $\widehat{z}^{\varepsilon}(t)$ and the covariance $p(t)$. We will assume the following,

(A6.5.1): $\{(m_t^{\varepsilon}, q)^2, q^2(z^{\varepsilon}(t)), F^2(y^{\varepsilon}(\cdot))\}$ is uniformly integrable.

The following theorem states that, for a small ε, the Ersatz conditional distribution (see Kushner and Runggaldier [115]) is "nearly optimal" with respect to a specific class of alternative estimators.

Theorem 6.5.2. *Assume* (A6.5.1) *and that* $w_2^{\varepsilon}(\cdot) \to w_2(\cdot)$*, a standard Wiener process. Then*

$$(\widehat{z}^{\varepsilon}(\cdot), z^{\varepsilon}(\cdot), w_2^{\varepsilon}(\cdot)) \Rightarrow (\widehat{z}(\cdot), z(\cdot), w_2(\cdot)).$$

Also,

$$\lim_{\varepsilon} E[q(z^{\varepsilon}(t)) - F(y^{\varepsilon}(\cdot))]^2 \geqslant \lim_{\varepsilon} E[q(z^{\varepsilon}(t)) - (m_t^{\varepsilon}, q)]^2 \tag{6.5.21}$$

Proof. The weak convergence is clear from the assumptions. Since $F(\cdot)$ is w.p.1. continuous, we also have

$$(q(z^{\varepsilon}(t)), F(y^{\varepsilon}(\cdot)), (m_t^{\varepsilon}, q)) \Longrightarrow (q(z(t)), F(y(\cdot)), (m_t, q)).$$

Hence,

$$(m_t, q) = \int q(z) dN(\widehat{z}(t), P(t), dz)$$

and $N(\widehat{z}, P, \cdot)$ is the normal distribution with mean $\widehat{z}$ and covariance matrix P. Thus, we have

$$\lim_{\varepsilon} E[q(z^{\varepsilon}(t)) - F(y\varepsilon(\cdot))]^2 = E[q(z(t)) - F(y(\cdot))]^2$$

and

$$\lim_{\varepsilon} E\left[q\left(z^{\varepsilon}(t)\right)-\left(m_t^{\varepsilon}, q\right)\right]^2=E[q(z(t))-E[q(z(t)) \mid y(s),\ s \leqslant t]]^2 .$$

Since the conditional expectation is the optimal estimator, (6.5.21) follows. □

Now we will give the 'near optimality' result for the policies. Let M_{∞} (respectively M_{ε}) denote the class of U (respectively V) valued continuous functions $u(\cdot,\cdot)$ (respectively $v(\cdot,\cdot)$), such that, with the use of policy value $(u(\widehat{x}(t),t), v(\widehat{x}(t),t))$ at time t, (6.5.7), (6.5.12), has a unique (weak sense) solution. In Theorem 6.5.1, we have shown that there are optimal strategy pairs (u^*, v^*) and a value J^* for the system (6.5.7), (6.5.12) with payoff (6.5.9). Hence, we can assume the following.

(A6.5.2): Let the strategy pair $(u^*(\cdot,\cdot), v^*(\cdot,\cdot))$ be in M and let this strategy be unique. Assume (u^*, v^*) is admissible for $x^{\varepsilon}(\cdot), \widehat{x}^{\varepsilon}(\cdot)$ of (6.5.18), (6.5.20) for small ε.

Thus, we can proceed with the following important convergence result.

Theorem 6.5.3. *Assume* (A6.5.1), (A6.5.2). *Let $x^{\varepsilon}(\cdot)$ and $\widehat{x}^{\varepsilon}(\cdot)$ denote the process and its estimate with $(u^*(\cdot,\cdot), v^*(\cdot,\cdot))$ used. Then*

$$\{x^{\varepsilon}(\cdot), \widehat{x}^{\varepsilon}(\cdot), u^*, v^*\} \rightarrow (x(\cdot), \widehat{x}(\cdot), u^*, v^*)$$

and the limit satisfies (6.5.7), (6.5.12). *Also,*

$$J^{\varepsilon}(u^*, v^*) \longrightarrow J(u^*, v^*)=J^* . \tag{6.5.22}$$

In addition, let $\widehat{u}(\cdot,\cdot)$ and $\widehat{v}(\cdot,\cdot)$ be a δ-optimal strategy pair for players I and II, respectively, with $(x(\cdot), \widehat{x}(\cdot))$ of (6.5.7), (6.5.12). *Then*

$$\underline{\lim}_{\varepsilon}\left|\sup_{u \in \mathscr{M}_1} J\left(u(y^{\varepsilon},\cdot), \widehat{v}(\widehat{x}^{\varepsilon},\cdot)\right)-J^{\varepsilon}(u^*, v^*)\right| \leqslant \delta \tag{6.5.23}$$

and

$$\overline{\lim}_{\varepsilon}\left|\inf_{v \in \mathscr{M}_2} J\left(\widehat{u}(\widehat{x}^{\varepsilon},\cdot), v(y^{\varepsilon},\cdot)\right)-J^{\varepsilon}(u^*, v^*)\right| \leqslant \delta . \tag{6.5.24}$$

Proof. Weak convergence is strait forward. By the assumed uniqueness, the limit $(x(\cdot), \widehat{x}(\cdot), u^*, v^*)$ satisfies (6.5.7), (6.5.12). Also, by this weak convergence and the fact that $T<\infty$, by the bounded convergence,

$$\lim_{\varepsilon} J^{\varepsilon}\left(u^*, v^*\right)=J\left(u^*, v^*\right) .$$

To show (6.5.23) and (6.5.24), repeat the procedure with admissible strategies $\left(u\left(y^{\varepsilon},\cdot\right), v\left(y^{\varepsilon},\cdot\right)\right)$. The limit $(x(\cdot), u(y,\cdot), v(y,\cdot))$ might depend on the chosen subsequence. For any convergent subsequence $\{\varepsilon_n\}$, we obtain

$$\lim_{\varepsilon=\varepsilon_n \rightarrow o} J^{\varepsilon}\left(u\left(y^{\varepsilon},\cdot\right), v\left(y^{\varepsilon},\cdot\right)\right)=J(u, v) .$$

Now by the definition of δ-optimality (6.5.15), (6.5.23) and (6.5.24) follows. □

6.5.3 Large time Problem

When the filtering system with wide band noise operates over a very long time interval, there are two limits involved, since both $t \to \infty$ and $\varepsilon \to 0$. It is then important that the results do not depend on how $t \to \infty$ and $\varepsilon \to 0$. We will make the following assumptions.

(A6.5.3): A is stable, $[A,H]$ is observable and $[A,D]$ is controllable.

(A6.5.4): $\xi_i(\cdot), i = 1,2$ are right continuous second order stationary processes with integrable covariance function $\overline{S}(\cdot)$. Let$\xi_i^{\varepsilon}(t) = \frac{1}{\varepsilon}\xi_i\left(t/\varepsilon^2\right)$. Also, if $t_{\varepsilon} \to \infty$ as $\varepsilon \to 0$, then $W_2^{\varepsilon}(t_{\varepsilon}+.) - W_2^{\varepsilon}(t_{\varepsilon}) \to W_2(\cdot)$.

(A6.5.5): If $z^{\varepsilon}(t_{\varepsilon}) \Rightarrow z(0)$ (a random variable) as $\varepsilon \to 0$, then $z^{\varepsilon}(t_{\varepsilon}+.) \to z(\cdot)$ with initial condition $z(0)$. Also $\sup_{\varepsilon,t} E\left|z^{\varepsilon}(t)\right|^2 < \infty$.

(A6.5.6): For each $\varepsilon > 0$, there is a random process $\zeta^{\varepsilon}(\cdot)$ such that $\{\zeta^{\varepsilon}(t), t < \infty\}$ is tight and for each strategy pair $(u(\cdot), v(\cdot)) \in M$.We can write

$$\{x^{\varepsilon}(\cdot), \widehat{x}^{\varepsilon}(\cdot), z^{\varepsilon}(\cdot), \widehat{z}^{\varepsilon}(\cdot), \zeta^{\varepsilon}(\cdot), \xi_1^{\varepsilon}(\cdot), \xi_2^{\varepsilon}(\cdot)\}$$

is a right continuous homogeneous Markov-Feller process with left hand limits.

We have the following result for filtering from Kushner and Runggaldier [115].

Theorem 6.5.4. *Assume* (A6.5.3)–(A6.5.5) *and let $q(\cdot)$ be a bounded continuous function. Let $F(\cdot) \in \aleph_t$. Define $y^{\varepsilon}(s) = 0$, for $s \leqslant 0$, and $y^{\varepsilon}(-\infty, t, \cdot)$ to be the "reversed" function with values $y^{\varepsilon}(-\infty, t:\tau) = y^{\varepsilon}(t-\tau)$ for $0 \leqslant \tau < \infty$. Then, if $t_{\varepsilon} \to \infty$ as $\varepsilon \to 0$, we have*

$$\{z^{\varepsilon}(t_{\varepsilon}+\cdot), \widehat{z}^{\varepsilon}(t_{\varepsilon}+\cdot), W_2^{\varepsilon}(t_{\varepsilon}+\cdot) - W_2^{\varepsilon}(t_{\varepsilon})\} \longrightarrow (z(\cdot), \widehat{z}(\cdot), W_2(\cdot)), \tag{6.5.25}$$

where $z(\cdot)$ and $\widehat{z}(\cdot)$ are stationary. Also

$$\lim_{\varepsilon,t} E[q(z^{\varepsilon}(t)) - F\left(y^{\varepsilon}(-\infty, t; \cdot)\right)]^2 \geqslant \lim_{\varepsilon,t} E[q(z^{\varepsilon}(t)) - (m_t^{\varepsilon}, q)]^2. \tag{6.5.26}$$

The limit of (m_t^{ε}, q) is the expectation with respect to the stationary $(\widehat{z}(\cdot), P(0))$ system.

Now we will use an ergodic payoff functional of the form

$$\rho^{\varepsilon}(u,v) = \limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T k(x^{\varepsilon}(t), z^{\varepsilon}(t), u(t), v(t))dt\right] \tag{6.5.27}$$

and

$$\rho(u,v) = \limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T k(x(t), z(t), u(t), v(t))dt\right] \tag{6.5.28}$$

where $k(\cdot,\cdot,\cdot)$ is a bounded continuous function.

Ergodic optimal strategies for players I and II are defined similar to the finite horizon case. We will assume the following.

(A6.5.7): There is an optimal strategy pair $(u^*, v^*) \in M$ for (6.5.1), (6.5.2), and (6.5.2) with (6.5.1), (6.5.2) has a unique invariant measure $\mu^{(u,v)}(\cdot)$.

The assumptions are not very restrictive. For detailed discussion on these type of assumptions, we refer the reader to Kushner [103], and Kushner and Dupuis [112].

Theorem 6.5.5. *Assume* (A6.5.3)–(A6.5.7). *Then the conclusions of Theorem* 6.5.3 *hold for the model* (6.5.4), (6.5.5) *with payoff* (6.5.27).

Proof. For a fixed $(u,v) \in M$, we define

$$P_T^\varepsilon(\cdot) = \frac{1}{T} E_x \int_0^T P\{X^\varepsilon(t) \in ./X^\varepsilon(0)\} dt,$$

where $X^\varepsilon(\cdot)$ is the process corresponding to $(u(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot)), v(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot)))$. By (A6.5.6), $\{P_T^\varepsilon(\cdot), T \geqslant 0\}$ is tight. Also,

$$\rho^\varepsilon(u(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot)), v(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot))) = \limsup_T \int r(x,z,u(\widehat{x},\widehat{z}), v(\widehat{x},\widehat{z})) P_T^\varepsilon(dX),$$

where $X = (x,z,\widehat{x},\widehat{z})$. Let $T_n^\varepsilon \to \infty$ be a sequence such that it attains the limit $\limsup_T$, and for which $P_{T_n^\varepsilon}^\varepsilon(\cdot)$ converges weakly to a measure $P^\varepsilon(\cdot)$. Again by (A6.5.6), $P^\varepsilon(\cdot)$ is an invariant measure for $X^\varepsilon(\cdot)$. Also, by construction of $P^\varepsilon(\cdot)$,

$$\rho^\varepsilon(u(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot)), v(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot))) = \limsup_T \int r(x,z,u(\widehat{x},\widehat{z}), v(\widehat{x},\widehat{z})) P^\varepsilon(dX).$$

Now by a weak convergence argument and (A6.5.7), we have

$$\rho^\varepsilon(u(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot)), v(\widehat{x}^\varepsilon(\cdot), \widehat{z}^\varepsilon(\cdot))) \to \rho(\widehat{x},\widehat{z}) = \int r(x,z,u(\widehat{x},\widehat{z}), v(\widehat{x},\widehat{z})) \mu^{(u,v)}(dxdzd\widehat{x}d\widehat{z}).$$

The rest of the proof is similar (with minor modifications) to that of Theorem 6.5.3 and hence we omit. □

6.5.4 *Partly Nonlinear Observations*

The ideas of previous subsections are useful in the case of nonlinear observations. However, we need the limit system to be linear. Consider the observations with a normalizing term $(1/\varepsilon)$.

$$\dot{y}^\varepsilon = h(Hx^\varepsilon + \xi_2^\varepsilon(t))/\varepsilon \tag{6.5.29}$$

with

$$y^\varepsilon(0) = 0,$$

and

$$h(x) = sign(x).$$

We assume

(A6.5.8): $\xi_2^\varepsilon(t) = \frac{1}{\varepsilon}\xi_2(t/\varepsilon^2)$, where $\xi_2(\cdot)$ is a component of a stationary Gauss-Markov process whose correlation function goes to zero as $t \to \infty$.

Let $\upsilon_0^2 = E\left(\xi_2^\varepsilon(t)\right)^2$. Then the average of (6.5.30) over the noise ξ_2^ε is

$$\left(\frac{2}{\pi\upsilon_0^2}\right)^{\frac{1}{2}} Hx^\varepsilon(t) + \delta_\varepsilon$$

where $\delta_\varepsilon \to 0$ as $\varepsilon \to 0$, uniformly for $x^\varepsilon(t)$ in any bounded set. The limit observation system is given by

$$dy = \left(\frac{2}{\pi\upsilon_0^2}\right)^{\frac{1}{2}} Hxdt + 2\Gamma_0^{\frac{1}{2}} dw_2. \tag{6.5.30}$$

For (6.5.7), (6.5.30), the Kalman-Bucy filter equations are

$$d\widehat{x} = (A\widehat{x} + Bu - Cv)\,dt + L(t)\left(dy - \left(\frac{2}{\pi\upsilon_0^2}\right)^{\frac{1}{2}} H\widehat{x}dt\right) \tag{6.5.31}$$

and

$$L(t) = P(t)H'\left(\frac{2}{\pi\upsilon_0^2}\right)^{\frac{1}{2}} \frac{1}{4\Gamma_0},$$

where $P(t)$ satisfies the Riccati equations

$$\dot{P} = FP + PF' - PH'HP\left(\frac{1}{\Gamma_0}\right)\left(\frac{2}{\pi\upsilon_0^2}\right), \tag{6.5.32}$$

and (6.5.13), where

$$\Gamma_0 = \frac{1}{\pi}\int_0^\infty \sin^{-1}\kappa(t)dt,$$

with $\kappa(t)$ being the correlation function of $\xi_2(\cdot)$. Define

$$\dot{\widehat{x}}^\varepsilon = (A\widehat{x}^\varepsilon + Bu - Cv) + L(t)\left[\dot{y}^\varepsilon - \left(\frac{2}{\pi\upsilon_0^2}\right)^{\frac{1}{2}} H\widehat{x}^\varepsilon\right]. \tag{6.5.33}$$

Now we will give the main result of this subsection.

Theorem 6.5.6. *Assume* (A6.5.1), (A6.5.2)*, and* (A6.5.8)*. Then the conclusions of Theorem* 6.5.3 *and Theorem* 6.5.4 *continues to hold.*

Remark 6.5.1. All the analysis can be carried out for a "soft" limiter of the form $h(x) = sign(x)$ for $|x| > c > 0$, $h(x) = x/c$ for $|x| < c$.

In the present situation, we obtained filtering and near optimality results for linear stochastic differential games with wide band noise perturbations. It is clear from Example 6.5.1 that the limits of $\{u^\varepsilon(y^\varepsilon,\cdot), v(y^\varepsilon,\cdot)\}$ would not necessarily be dependent only on the limit data y-even when $y^\varepsilon(\cdot) \to y(\cdot)$. The case of partly nonlinear observations is also considered. Using the methods of this subsection, we can extend the results to the conditional Gaussian problem, in which, the coefficients of x^ε and ξ_2^ε in the observation equation (6.5.2) can depend on the estimate $\widehat{x}^\varepsilon$ and on $P^\varepsilon(\cdot)$.

6.6 Deterministic Approximations in Two-Person Differential Games

As we have seen in previous sections, considerable effort has been put into developing approximation techniques for such problems. One such approach use in the stochastic control literature is, in lieu of the original model, a model where the underlying processes are replaced by simpler ones (Fleming [60], Kushner [103], Kushner and Ramachandran [113], Kushner and Runggaldier [114], Lipster, Runggaldier, and Taksar [127]). In stochastic game problems such an effort was made in Ramachandran [161] using diffusion approximation techniques.In the present section, fluid approximation techniques (i.e., the simpler model is deterministic) to a two person zero sum differential game model will be developed. Consider a two person game problem described by a family of stochastic equations parametrized by a small parameter ε ($\varepsilon \downarrow 0$), with dynamics

$$dX^{\varepsilon}(t) = [a(X^{\varepsilon}(t), \xi^{\varepsilon}(t)) + b_1(x^{\varepsilon}(t))u_1^{\varepsilon}(t) + b_2(X^{\varepsilon}(t))u_2^{\varepsilon}(t)]\,dt + dM^{\varepsilon}(t) \tag{6.6.1}$$

and initial condition $X^{\varepsilon}(0)$. Here, $X^{\varepsilon} = (X^{\varepsilon}(t))$ is the controlled state process, $\xi = (\xi^{\varepsilon}(t))$ is the contamination process affecting the drift of X^{ε}, and $M = (M^{\varepsilon}(t))$ is the process representing the noise in the system. Also $u_1^{\varepsilon} = (u_1^{\varepsilon}(t))$ and $u_2^{\varepsilon} = (u_2^{\varepsilon}(t))$ are controls for players I and II, respectively. Given a finite horizon, $T > 0$, with each policy pair $u^{\varepsilon} = (u_1^{\varepsilon}, u_2^{\varepsilon})$, we associate the payoff to player I by

$$J^{\varepsilon}(u_1^{\varepsilon}, u_2^{\varepsilon}) = E\left\{\int_0^T [k(X^{\varepsilon}(t)) + p(u_1^{\varepsilon}(t)) + q(u_2^{\varepsilon}(t))]\,dt + r(X^{\varepsilon}(t))\right\}, \tag{6.6.2}$$

where $k(x),\ p(u_1),\ q(u_2)$, and $r(x)$ are nonnegative functions on the real line referred to as holding cost, control costs, and terminal cost functions, respectively. Our objective is to find value function V^{ε}, that is

$$V^{\varepsilon} = V^{\varepsilon}(u_1^{\varepsilon *}, u_2^{\varepsilon *}) = \inf_{u_2^{\varepsilon} \in A_2} \sup_{u_1^{\varepsilon} \in A_1} J^{\varepsilon}(u_1^{\varepsilon}, u_2^{\varepsilon}) = \sup_{u_1^{\varepsilon} \in A_1} \inf_{u_2^{\varepsilon} \in A_2} J^{\varepsilon}(u_1^{\varepsilon}, u_2^{\varepsilon}), \tag{6.6.3}$$

and the corresponding policy pair $(u_1^{\varepsilon *}, u_2^{\varepsilon *})$. The sets A_1, and A_2 will be defined in the next section. The process $\xi^{\varepsilon}(\cdot)$ is said to be *exogenous* or state independent if for each t and set B in $\sigma(\xi^{\varepsilon}(s), s > t)$, such that,

$$P\{B \mid \xi^{\varepsilon}(s), s \leqslant t\} = P\{B \mid \xi^{\varepsilon}(s), X^{\varepsilon}(s), s \leqslant t\}.$$

In order for a desired convergence to occur, the ‘rate of fluctuations’ of $\xi^{\varepsilon}(\cdot)$ must increase as $\varepsilon \to 0$. We consider the case in which the ‘intensity’ of the random noise disturbance M becomes very small with ε, while the ‘contaminating’ process ξ fluctuates with increasing speed. In this work, we assume that the controlled state process X is completely observed.

It is very hard to obtain optimal strategies and value satisfying (6.6.1) and (6.6.3). To this end, we will now introduce a deterministic model, which we will show to be the limiting model corresponding to (6.6.1) to (6.6.3) under appropriate ergodic conditions introduced in the next section.

Consider a two person zero-sum differential game problem where the dynamics of the limiting deterministic system is given by the following ordinary differential equation:

$$dx(t) = [\overline{a}(x(t)) + b_1(x(t))u_1(t) + b_2(x(t))u_2(t)]\,dt, \tag{6.6.4}$$

with

$$x(0) = x_0,$$

where $x(t)$ is deterministic controlled process, $u_1(t),\ u_2(t)$ are deterministic controls for players I and II, respectively. Define the payoff to player I by

$$j(u_1,u_2) = \int_0^T k(x(t)) + p(u_1(t)) + q(u_2(t))]dt + r(x(t)) \tag{6.6.5}$$

and

$$v = \inf_{u_2}\sup_{u_1} j(u_1,u_2) = \sup_{u_1}\inf_{u_2} j(u_1,u_2). \tag{6.6.6}$$

Here, player I maximizes $j(\cdot,\cdot)$ and player II minimizes $j(\cdot,\cdot)$. The linearity of controls is assumed, since nonlinear problems could rarely be solved analytically. With the use of the so called relaxed controls, we could allow nonlinear forms. However, for simplicity of presentation, in this work we will restrict to linear forms.

These type of results have two major benefits. From the theoretical point of view, one obtains a stability result for the optimal strategy pair of a deterministic system in the sense that this policy pair is asymptotically optimal for a large class of complicated problems of stochastic games. From a practical point of view, when a direct approach would be impossible, these results allow one to compute an asymptotically optimal strategy pair for a variety of stochastic game problems under quite general conditions.

6.6.1 *Preliminaries*

Let $F^{\varepsilon} = \{\Im_t^{\varepsilon}\}_{t\geqslant 0}$ denote the minimal σ-algebra over which $\{X^{\varepsilon}(s), \xi^{\varepsilon}(s), M^{\varepsilon}(s),\ s \leqslant t\}$, is measurable. For each ε let $(\Omega, \Im, F^{\varepsilon}, P)$ be a fixed stochastic basis, and where $(\Omega, \Im, P)$ is a complete probability space. Let E_t^{ε} denote the expectation conditioned on $\Im_t^{\varepsilon}$. Let U_1, U_2 be compact metric spaces with metric $d_i(\cdot)$. The control process $u_i^{\varepsilon}(t)$ with values in U_i is said to be *admissible strategy* for the i^{th} player if it is $\Im_t^{\varepsilon}$ adapted and $\int_0^T |u_i^{\varepsilon}(s)|ds < \infty$, a.s. Let A_i, $i=1,2$ denote the set of admissible strategies. Let $A=A_1\times A_2$. An admissible strategy $u_1^{\varepsilon *}\in A_1$ is said to be *optimal* for player I if

$$J^{\varepsilon}(u_1^{\varepsilon *},\widetilde{u}_2^{\varepsilon})\geqslant \inf_{u_2^{\varepsilon}\in A_2}\sup_{u_1^{\varepsilon}\in A_1} J^{\varepsilon}(u_1^{\varepsilon},u_2^{\varepsilon})=V^{\varepsilon +}, \tag{6.6.7}$$

for any $\widetilde{u}_2^{\varepsilon}\in A_2$. Similarly, an admissible strategy $u_2^{\varepsilon *}\in A_2$ is said to be *optimal* for player II if

$$J^{\varepsilon}(\widetilde{u}_1^{\varepsilon},u_2^{\varepsilon *})\leqslant \sup_{u_1^{\varepsilon}\in A_1}\inf_{u_2^{\varepsilon}\in A_2} J^{\varepsilon}(u_1^{\varepsilon},u_2^{\varepsilon})=V^{\varepsilon -}, \tag{6.6.8}$$

for any $\widetilde{u}_1^{\varepsilon}\in A_1$. If $V^{\varepsilon +}=V^{\varepsilon -}$ for each initial value, then the game is said to admit a value and we will denote the value by V^{ε}. Also, $u_{1\delta}$ and $u_{2\delta}$ are said to be *δ-optimal strategies* for player I and II, respectively, if

(6.6.9)

Similarly, we can define all of the above concepts for the deterministic model. Define the control space in the deterministic game by $\widetilde{A}_i=\left\{u_i : u_i \text{ is measurable and } \int_0^T |u_i(t)|dt<\infty\right\}$ and $\widetilde{A}=\widetilde{A}_1\times\widetilde{A}_2$. Note that $\widetilde{A}\subset A$.

We will use following assumptions.

(A6.6.1): $\xi^{\varepsilon}(t)=\xi(t/\varepsilon)$, where $\xi(\cdot)$ is a stationary process which is strong mixing, right continuous and bounded with mixing rate function $\phi(\cdot)$ satisfying $\int_0^{\infty}\phi(s)ds<\infty$.

(A6.6.2): $b_1(\cdot)$, $b_2(\cdot)$b are bounded and Lipshitz continuous $a(\cdot,\cdot)$ and its gradient $a_x(\cdot,\cdot)$ are continuous in (x,ξ) and satisfy uniform Lipschitz condition with same constant.

(A6.6.3): There is a continuously differentiable function $\overline{a}(\cdot)$ such that for each $t<T$ and x, we have

$$\int_t^T \left[E_t^{\varepsilon}a(x,\xi^{\varepsilon}(s))-\overline{a}(x)\right]ds\to 0$$

in probability as $\varepsilon\to 0$.

(A6.6.4): The cost functions $k(\cdot)$ and $r(\cdot)$ are continuous nonnegative satisfying

$$k(x),\quad r(x)\leqslant c_0\left(1+|x|^{\gamma}\right),\quad c_0,\quad \gamma>0.$$

Also,

$$p(u_1(t))+q(u_2(t))\geqslant c_2\left(|u_1|^{1+\gamma_2}+|u_2|^{1+\gamma_2}\right),\quad c_2,\gamma_2>0,$$

and

$$p(u_1)\ \text{and}\ q(u_2)\ \text{are nonnegative convex.}$$

(A6.6.5): The process $M^{\varepsilon}=(M^{\varepsilon}(t))_{t\geqslant 0}$ is a square integrable martingale with paths in the Skorokhod space, $D[0,\infty)$, whose predictable quadratic variations $\langle M^{\varepsilon}\rangle(t)$ satisfies

(i) $\langle M^{\varepsilon}\rangle(t) = \varepsilon\int_0^t m^{\varepsilon}(s)\,ds$, with bounded density $m^{\varepsilon}(s)$. That is, there exists a constat c_1 such that,

(ii) $m^{\varepsilon}(t) \leqslant c_1$; $t \leqslant T$, P-a.s..
The jumps $\Delta M^{\varepsilon}(s) \doteq M^{\varepsilon}(s) - \lim_{v\uparrow s} M^{\varepsilon}(s)$ are bounded, i.e., there exists a constant $K > 0$, such that,

(iii) $|\Delta M^{\varepsilon}(t)| \leqslant K; t \leqslant T, \quad \varepsilon \in (0,1]$.

(A6.6.6): $p\text{-}\lim_{\varepsilon\to 0} X^{\varepsilon}(0) = x_0, \quad x_0 \in R$.

Note that by the assumption (A6.6.2) equation (6.6.1) has a unique solution. Also, in (A6.6.1) if we replace stationarity with the ergodicity assumptions as in Liptser, Runggaldier, and Taksar [127], all the results of this paper continues to hold. In (A6.6.2), smoothness of $a(\cdot,\cdot)$ is assumed only to make the proof simpler. The case of non smooth dynamics can be carried out as in Kushner [103] by only assuming smoothness of $E_t^{\varepsilon} a(x,\xi^{\varepsilon}(s))$. As a result of (A6.6.5), we have $p\text{-}\lim_{\varepsilon\to 0}|M_t^{\varepsilon}| = 0$. We can use (A6.6.4) to avoid singular controls, as given in Lions and Souganidis [126].

6.6.2 *Fluid Approximation*

We will now give the main convergence result for the controlled state process and show that the limit satisfy (6.6.10). The proof will utilize the martingale methods and the so called perturbed test function method.

Theorem 6.6.1. *Suppose that* (A6.5.1)–(A6.5.6) *hold. Let* $X_0^{\varepsilon} \Rightarrow x_0$ *and* $u^{\varepsilon}(\cdot) = (u_1^{\varepsilon}(\cdot), u_2^{\varepsilon}(\cdot)) \to u(\cdot) \equiv (u_1(\cdot), u_2(\cdot))$*, where* $(u_1(\cdot), u_2(\cdot))$ *is an admissible strategy pair for* (6.6.4). *Then* $(X^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot)) \to (x(\cdot), u(\cdot))$*, where* $u(\cdot)$ *is measurable* (*admissible*) *process and*

$$dx(t) = [\overline{a}(x(t)) + b_1(x(t))u_1(t) + b_2(x(t))u_2(t))]\,dt \tag{6.6.10}$$

Also

$$J^{\varepsilon_n}\left(u_1^{\varepsilon_n}, u_2^{\varepsilon_n}\right) \to j(u_1,u_2). \tag{6.6.11}$$

Proof. Define a process $X^{\varepsilon,0}(\cdot)$ by

$$X^{\varepsilon,0}(t) = X^{\varepsilon,0}(0) + \int_0^t \left[a\left(X^{\varepsilon,0}(s),\xi^{\varepsilon}(s)\right) + b_1\left(X^{\varepsilon,0}(s)\right)u_1^{\varepsilon}(s) + b_2\left(X^{\varepsilon,0}(s)\right)u_2^{\varepsilon}(s)\right]ds \tag{6.6.12}$$

Let $Y^{\varepsilon}(s) = \sup_{s\leqslant t}\left|X^{\varepsilon}(s) - X^{\varepsilon,0}(s)\right|$. Then by (A6.6.2), we have,

$$Y^{\varepsilon}(t) \leqslant K\int_0^t Y^{\varepsilon}(s)d\left[s + \int_0^s |u_1^{\varepsilon}(w)|\,dw + \int_0^s |u_2^{\varepsilon}(w)|\,dw\right] + \sup_{s\leqslant t}|M^{\varepsilon}(s)|,\ \ t \leqslant T,$$

where K is the Lipschitz constant. Using the Gronwall-Bellman inequality we obtain

$$Y^{\varepsilon}(t) \leqslant K \sup_{s\leqslant t} |M^{\varepsilon}(s)| \exp\left\{ K\left[T + \int_0^T |u_1^{\varepsilon}(w)|\, dw + \int_0^T |u_2^{\varepsilon}(w)|\, dw \right] \right\}.$$

By (A6.6.5) (see Liptser, Runggaldier, and Taksar [127]), $\sup_{x\leqslant t} |M^{\varepsilon}(s)| \to 0$, $\varepsilon \to 0$ in probability and by (A6.6.2) and (A6.6.4)

$$\lim_{\delta\to 0} \limsup_{\varepsilon\to 0} P\left(\sup_{t,s\leqslant T:|t-s|\leqslant\delta} \int_s^t [|u_1(w)| + |u_2(w)|]dw > \eta \right) = 0.$$

Consequently $Y^{\varepsilon}(t) \to 0$, $\varepsilon \to 0$, in probability and the theorem remains true if its statements are proved only for $(X^{\varepsilon,0}(\cdot), u^{\varepsilon}(\cdot))$. We will prove the weak convergence for the process $(X^{\varepsilon,0}(\cdot), u^{\varepsilon}(\cdot))$. For notational convenience we will use $(X^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot))$ for $(X^{\varepsilon,0}(\cdot), u^{\varepsilon}(\cdot))$.

Define the perturbation $f_1^{\varepsilon}(t) = f_1^{\varepsilon}(X^{\varepsilon}(t), t)$, where

$$f_1^{\varepsilon}(x,t) = \int_t^T f_x(x)\, [E_t^{\varepsilon} a(x, \xi^{\varepsilon}(s)) - \overline{a}(x)]\, ds. \tag{6.6.13}$$

It is important to note that (6.6.13) averages only the noise, not the state $X^{\varepsilon}(\cdot)$. The state $x = X^{\varepsilon}(t)$ is considered as parameter in (6.6.13). Now,

$$\begin{aligned} f_1^{\varepsilon}(x,t) &= \int_t^T f_x(x)\, [E_t^{\varepsilon} a(x, \xi^{\varepsilon}(s)) - \overline{a}(x)]\, ds \\ &= \varepsilon \int_{t/\varepsilon}^{T/\varepsilon} f_x(x)\, [E_t^{\varepsilon} a(x, \xi^{\varepsilon}(s)) - \overline{a}(x)]\, ds. \end{aligned}$$

In view of Lemma 6.2.2, (A6.6.1) and (A6.6.2), for some $L > 0$,

$$\begin{aligned} \sup_{t\leqslant T} |f_1^{\varepsilon}(t)| &= \varepsilon \sup_{t\leqslant T} \left| \int_{t/\varepsilon}^{T/\varepsilon} f_x\, [E_t^{\varepsilon} a(x, \xi^{\varepsilon}(s)) - \overline{a}(x)] - [Ea(x, \xi^{\varepsilon}(s)) - \overline{a}(x)]\, ds \right| \\ &\leqslant L\varepsilon \sup_{t\leqslant T} \left(\int_{t/\varepsilon}^{T/\varepsilon} \phi\left(s - \frac{t}{\varepsilon}\right) ds \right) \\ &= O(\varepsilon). \end{aligned}$$

Hence,

$$\lim_{\varepsilon} \sup_{t<T} E\, |f_1^{\varepsilon}(t)| = 0. \tag{6.6.14}$$

Write $\widetilde{a}(x,\xi) = f_x(x)\, (a(x,\xi) - \overline{a}(x))$. We have

$$\widehat{A}^{\varepsilon} f_1^{\varepsilon}(t) = -\widetilde{a}(X^{\varepsilon}(t), \quad \xi^{\varepsilon}(t)) + \int_t^T (E_t^{\varepsilon} \widetilde{a}(X^{\varepsilon}(t), \quad \xi^{\varepsilon}(s)))_x \dot{X}^{\varepsilon}(t) ds + o(1)$$

where $p\text{-}\lim_\varepsilon o(1) = 0$ uniformly in t. Define the perturbed test function $f^\varepsilon(t) = f(X^\varepsilon(t)) + f_1^\varepsilon(t)$. For simplicity we write x for $X^\varepsilon(t)$. Then, we have

$$\begin{aligned}\widehat{A}^\varepsilon f_1^\varepsilon(t) &= f_x(x)\left[a(x,\xi^\varepsilon(t)) + b_1(x)u_1^\varepsilon(t) + b_2(x)u_2^\varepsilon(t)\right] - f_x(x)\left(a(x,\xi) - \overline{a}(x)\right)\\ &\quad + \int_t^T \left(E_t^\varepsilon \widetilde{a}(x,\xi^\varepsilon(s))\right)_x \left[a(x,\xi^\varepsilon(t)) + b_1(x)u_1^\varepsilon(t) + b_2(x)u_2^\varepsilon(t)\right] ds + o(1)\\ &= f_x(x)\left[\overline{a}(x) + b_1(x)u_1^\varepsilon(t) + b_2(x)u_2^\varepsilon(t)\right]\\ &\quad + \varepsilon \int_{t/\varepsilon}^{T/\varepsilon} \left(E_t^\varepsilon \widetilde{a}(x,\xi(s))\right)_x \left[a(x,\xi^\varepsilon(t)) + b_1(x)u_1^\varepsilon(t) + b_2(x)u_2^\varepsilon(t)\right] ds + o(1).\end{aligned} \tag{6.6.15}$$

Under (A6.6.2), the second term in (6.6.15) is $o(1)$ where $o(1)$ terms goes to zero in p-limit as $\varepsilon \to 0$. Then (6.6.14) and (6.6.15) imply that

$$p\text{-}\lim_{\varepsilon} \left[f^\varepsilon(\cdot) - f(X^\varepsilon(\cdot))\right] = 0$$

and

$$p\text{-}\lim_{\varepsilon} \left[\widehat{A}^\varepsilon f^\varepsilon(\cdot) - \widehat{A}^u f(X^\varepsilon(\cdot))\right] = 0 \ \text{ for } \ t \leqslant T.$$

Hence (6.6.10) is proved.

By the above methods, we can write

$$\int_0^T \left[k(X^\varepsilon(t)) + p(u_1^\varepsilon(t)) + q(u_2^\varepsilon(t))\right] dt \Longrightarrow \int_0^T \left[k(x(t)) + p(u_1(t)) + q(u_2(t))\right] dt \tag{6.6.16}$$

and

$$r(X^\varepsilon(t)) \to r(x(t)).$$

By (A6.6.2), each moment of $X^\varepsilon(t)$ is bounded uniformly in ε and $t \leqslant T$. By (A6.6.2) and (A6.6.4), the left hand terms in (6.6.16) are uniformly (in ε) integrable and the convergence in (6.6.11) follows. □

Remark 6.6.1. The condition in the theorem stating that $u^\varepsilon(\cdot) = (u_1^\varepsilon(\cdot), u_2^\varepsilon(\cdot)) \to u(\cdot) = (u_1(\cdot), u_2(\cdot))$ is a reasonable one. This follows, if $p(u_1(t)) + q(u_2(t)) \geqslant c_2(|u_1|^{1+\gamma} + |u_2|^{1+\gamma})$, $c_2,\ \gamma > 0$, and $p(u_1)$ and $q(u_2)$ are nonnegative convex, then mimicking the proof of Theorem 5.1 of Liptser, Runggaldier, and Taksar [127], we can obtain the weak convergence of Theorem 6.6.1. The analytic method used in [127], under their conditions, could also be adapted to prove Theorem 6.6.1.

6.6.3 δ-Optimality

In this section, we will prove near optimality and asymptotic optimality of the optimal strategy pair for the limit deterministic system.

Theorem 6.6.2. *Assume* (A6.6.1)–(A6.6.6). *Let* (u_1^*, u_2^*) *be the unique optimal strategy pair for* (6.6.4)–(6.6.5). *Then* $\{X^\varepsilon(\cdot), u_1^*, u_2^*\} \to (x(\cdot)u_1^*, u_2^*)$ *and the limit satisfies* (6.6.10). *Also*

$$J^\varepsilon(u_1^*, u_2^*) \to j(u_1^*, u_2^*) = v. \tag{6.6.17}$$

In addition, let $\widehat{u}_1^\varepsilon$ *and* $\widehat{u}_2^\varepsilon$ *be a* δ*-optimal strategy pair for player I and II respectively, with* $X^\varepsilon(\cdot)$ *of* (6.6.1). *Then,*

$$\liminf_{\varepsilon} \left| \sup_{u_1^\varepsilon \in U_1} J^\varepsilon(u_1^\varepsilon, \widehat{u}_2^\varepsilon) - J^\varepsilon(u_1^*, u_2^*) \right| \leqslant \delta \tag{6.6.18}$$

and

$$\liminf_{\varepsilon} \left| \sup_{u_2^\varepsilon \in U_2} J^\varepsilon(\widehat{u}_1^\varepsilon, u_2^\varepsilon) - J^\varepsilon(u_1^*, u_2^*) \right| \leqslant \delta. \tag{6.6.19}$$

Proof. By Theorem 6.6.1, the weak convergence is straight forward. By the assumed uniqueness, the limit satisfies (6.6.10). Also, by this weak convergence and the fact that $T < \infty$, by the bounded convergence,

$$\lim_{\varepsilon} J^\varepsilon(u_1^*, u_2^*) = j(u_1^*, u_2^*).$$

Now to show (6.6.18) and (6.6.19), repeat the procedure with admissible strategies u_1^ε and u_2^ε. The limit (u_1, u_2) might depend on the chosen subsequence. For any convergent subsequence, we obtain, $\lim_{\varepsilon=\varepsilon_n \to 0} J^\varepsilon(u_1^\varepsilon, u_2^\varepsilon) = j(u_1, u_2)$. Now by the definition of δ*-optimality* (6.6.18) and (6.6.19) follows. □

Note: If $(u_1^*(t), u_2^*(t))$ is the optimal strategy pair for (6.6.18) and (6.6.19), then $\{X^\varepsilon(t), u_1^*(t), u_2^*(t)\}_{0 \leqslant t \leqslant T}$ is the process associated with the policy pair $(u_1^\varepsilon(t), u_2^\varepsilon(t)) \equiv (u_1^*(t), u_2^*(t))$. Since $(u_1^*(t), u_2^*(t))$ is deterministic, corresponding $(u_1^\varepsilon(t), u_2^\varepsilon(t))$ is admissible control for the systems (6.6.1)–(6.6.3).

If for each ε, there is a value for the stochastic game, then the following result shows that they converge to the value of the deterministic game. To prove this we will now introduce a new game through which we will connect the values of stochastic and deterministic games. Define a continuous map ϕ, such that,

$$\phi : A \to \widetilde{A} \ \text{ such that if } \ u = (u_1, u_2) \in \widetilde{A}, \ \text{ then } \ \phi(u) = u$$

Note that example of one such map is $\phi(u) = Eu$ for $u \in A$. Clearly, if $u \in \widetilde{A}$, $\phi(u) = Eu = u$. Define the map $L_1 : \widetilde{A} \to \widetilde{A}$ by letting $L_1 u_2 (u_2 \in \widetilde{A}_2)$ to be:

$$j(L_1 u_2, u_2) = \sup_{u_1 \in \widetilde{A}_1} j(u_1, u_2), \quad u_2 \in \widetilde{A}_2.$$

Similarly, define $L_2 : \widetilde{A} \to \widetilde{A}$ by

$$\inf_{u_2 \in \widetilde{A}_2} j(u_1, u_2) \equiv j(u_1, L_2 u_1)$$

We will now make following simplifying assumption.

(A6.6.7): Let L_i, $i = 1, 2$ be continuous, that is, for any sequence $\{u_i^k\}$ of admissible controls in $\widetilde{A}_i$, such that, $u_i^k \to u_i \in \widetilde{A}_i$, $L_i(u_i^k) \to L_i(u_i)$, in $L_2(0, T)$ norm.

Remark 6.6.2. The continuity of L_i may be justified by the fact that the controls u_i are state dependent feedback controls.

Following result states that if the value exists for the stochastic game for each ε, then asymptotically (as $\varepsilon \to 0$) they coincide with the value of the limit deterministic model.

Theorem 6.6.3. *Assume* (A6.6.1)–(A6.6.7) *and that for each ε, the value exists for the stochastic game* (6.6.1) *to* (6.6.3). *Also the value for the deterministic game* (6.6.4) *to* (6.6.6) *exists. Then* $\lim_{\varepsilon \to 0} V^\varepsilon = v$.

Proof. For the proof, we introduce the following game which is played as follows. Player II chooses his control first, say u_2^ε, which is known to player I. Then player I chooses the control $L_1(\phi(u_2^\varepsilon))$. Define $\widetilde{u}_2^\varepsilon$ by

$$J^\varepsilon\left(L_1 \phi\left(\widetilde{u}_2^\varepsilon\right), \widetilde{u}_2^\varepsilon\right) = \inf_{\widetilde{u}_2^\varepsilon \in A_2} J^\varepsilon\left(L_1 \phi\left(\widetilde{u}_2^\varepsilon\right), \widetilde{u}_2^\varepsilon\right).$$

By relative compactness, $\widetilde{u}_2^\varepsilon \to u_2^+ \in \widetilde{A_2}$. By continuity and by the definition of ϕ and L_1, we have $L_1\left(\phi\left(\widetilde{u}_2^\varepsilon\right)\right) \to L_1\left(\phi\left(u_2^+\right)\right) = L_1\left(u_2^+\right)$.

Then

$$\begin{aligned} \lim_{\varepsilon \to 0} V^\varepsilon &\equiv \lim_{\varepsilon \to 0} \left[\inf_{u_2^\varepsilon \in A_2} \sup_{u_1^\varepsilon \in A_1} J^\varepsilon\left(u_1^\varepsilon\left(u_2^\varepsilon\right), u_2^\varepsilon\right) \right] \\ &\geqslant \lim_{\varepsilon \to 0} J^\varepsilon\left(L_1\left(\phi\left(\widetilde{u}_2^\varepsilon\right), \widetilde{u}_2^\varepsilon\right)\right) \\ &= j\left(L_1 u_2^+, u_2^+\right) \geqslant v^+, \end{aligned}$$

since $v^+ = \inf_{u_2 \in A_2} \sup_{u_1 \in A_1} j(L_1 u_2, u_2) = \inf_{u_2 \in A_2} j(L_1 u_2, u_2)$.

Similarly, when player I chooses first, proceeding as before with $\widetilde{u}_1^\varepsilon \to u_1^- \in \widetilde{A}_1$ we get,

$$\begin{aligned}
\lim_{\varepsilon\to 0} V^{\varepsilon -} &\equiv \lim_{\varepsilon\to 0}\left[\sup_{u_1^\varepsilon\in A_1}\inf_{u_2^\varepsilon\in A_2} J^\varepsilon\left(u_1^\varepsilon,\, u_2^\varepsilon\left(u_1^\varepsilon\right)\right)\right] \\
&\leqslant \lim_{\varepsilon\to 0}\left[\sup_{u_1^\varepsilon} J^\varepsilon\left(u_1^\varepsilon, L_2\left(\phi\left(u_1^\varepsilon\right)\right)\right)\right] \\
&\equiv \lim_{\varepsilon\to 0} J^\varepsilon\left(\widetilde{u}_1^\varepsilon, L_2\left(\phi\left(\widetilde{u}_1^\varepsilon\right)\right)\right) \leqslant v^-.
\end{aligned}$$

Since all the games have values

$$V_+^\varepsilon = V_-^\varepsilon = V^\varepsilon \text{ and } v^+ = v^- = v$$

Hence,

$$\lim_{\varepsilon\to 0} V^\varepsilon = v. \qquad \square$$

The following result is obtained direct from Theorem 6.6.2 and Theorem 6.6.3.

Theorem 6.6.4. *Assume* (A6.6.1)–(A6.6.7). *Let* $(u_1^*(t), u_2^*(t))$*,* $0 \leqslant t \leqslant T$ *be an optimal deterministic strategy for* (6.6.4), (6.6.5)*, then* $(u_1^*(t), u_2^*(t))$ *is asymptotically optimal for* (6.6.1), (6.6.3) *in the sense that*

$$\lim_{\varepsilon\to 0} |J^\varepsilon(u_1^*, u_2^*) - V^\varepsilon| = 0. \tag{6.6.20}$$

6.6.4 L^2-*Convergence*

In this section we consider a simpler physical system of the form given below by (6.6.21) and we will show that in place of the weak convergence of X^ε to x, under few additional conditions, we can in fact obtain an L^2-convergence. Rewrite the systems (6.6.12) and (6.6.4), respectively, in the integral form:

$$X^\varepsilon(t) = X^\varepsilon(0) + \int_0^t [a(X^\varepsilon(s), \xi^\varepsilon(s)) + b_1(X^\varepsilon(s))u_1^\varepsilon(s) + b_2(X^\varepsilon(s))u_2^\varepsilon(s)]\, ds \tag{6.6.21}$$

and

$$x(t) = x_0 + \int_0^t [\overline{a}(x(s)) + b_1(x(s))u_1(s) + b_2(x(s))u_2(s)]\, ds. \tag{6.6.22}$$

Note that using Theorem 6.6.3, it is enough to consider the system (6.6.12) instead of the system (6.6.1). The conditions on a, b_1, b_2 are the same.

Define $\|X^\varepsilon(t)\| = \sup_{0\leqslant t\leqslant T}\left\{E(X^\varepsilon(t))^2\right\}^{\frac{1}{2}}$.

(A6.6.8): Assume

(i) $\|X^\varepsilon(0) - x_0\| \to 0$ as $\varepsilon \to 0$.

(ii) $\|u_i^\varepsilon - u_i\| \to 0$ as $\varepsilon \to 0 \quad (i = 1,2)$.

(iii) $E\left(a\left(x\xi^\varepsilon\right) - \overline{a}(x)\right)^2 \leqslant \phi(\varepsilon, x) \to 0$ as $\varepsilon \to 0$.

Now, we can state the following result.

Theorem 6.6.5. *Assume* (A6.6.1), (A6.6.2), (A6.6.5), *and* (A6.6.8). *Then*

$$\|X^\varepsilon(t) - x(t)\| \to 0 \ \text{as} \ \varepsilon \to 0, \tag{6.6.23}$$

where $X^\varepsilon(\cdot)$ is the solution of (6.6.21) *and $x(\cdot)$ is the solution of* (6.6.22).

Proof.

$$\begin{aligned}
E\left|X^\varepsilon - x\right|^2 &\leqslant N\left\{E(X^\varepsilon(0) - x_0)^2 + \int_0^t E\left|a\left(X^\varepsilon, \xi^\varepsilon\right) - \overline{a}(x)\right|^2 ds\right. \\
&\quad \left. + \int_0^t E\left|b_1(X^\varepsilon)u_1^\varepsilon - b_1(x)u_1\right|^2 ds + \int_0^t E\left|b_2(X^\varepsilon)u_2^\varepsilon - b_2(x)u_2\right|^2 ds\right\} \\
&\leqslant N\left\{E(X^\varepsilon(0) - x_0)^2 + \int_0^t E|a(X^\varepsilon, \xi^\varepsilon) - \overline{a}(x, \xi^\varepsilon)|^2 ds + \int_0^t E|a(x, \xi^\varepsilon) - \overline{a}(x)|^2 ds\right. \\
&\quad + \int_0^t E|b_1(X^\varepsilon) - b_1(x)|^2 |u_1|^2 ds + \int_0^t E|b_1(X^\varepsilon)|^2 |u_1^\varepsilon - u_1|^2 ds \\
&\quad \left. + \int_0^t E|b_2(X^\varepsilon) - b_2(x)|^2 |u_2|^2 ds + \int_0^t E|b_2(X^\varepsilon)|^2 |u_2^\varepsilon - u_2|^2 ds\right\}.
\end{aligned}$$

Note that b_i's are bounded and Lipschtzian, and so is a. Hence,

$$\begin{aligned}
E|X^\varepsilon - x_0|^2 &\leqslant NK\left\{E\left|X^\varepsilon(0) - x_0\right|^2 + \int_0^t E\left|X^\varepsilon - x\right|^2 ds\right. \\
&\quad + \int_0^T E\left(a(x, \xi^\varepsilon)\right) - \overline{a}\,(x)\right)^2 ds + \int_0^t E\left|X^\varepsilon - x\right|^2 ds \\
&\quad + \int_0^t E\left(u_1^\varepsilon - u_1\right)^2 ds + \int_0^t E\left(u_2^\varepsilon - u_2\right)^2 ds.
\end{aligned} \tag{6.6.24}$$

Using the assumptions (A6.6.8) in equation (6.6.24) we get (for some K)

$$\begin{aligned}
E\left|X^\varepsilon - x\right|^2 &\leqslant KE\left[\left|X^\varepsilon(0) - x_0\right|^2 + \int_0^T \phi(\varepsilon, x)\, ds + \int_0^T E\left(u_1^\varepsilon - u_1\right)^2 ds\right. \\
&\quad \left. + \int_0^T E(u_2^\varepsilon - u_2)^2 ds + \int_0^t E(X(s)^\varepsilon - x(s))^2 ds\right]
\end{aligned}$$

Using the Grownwall-Bellman inequality, we have,

$$\begin{aligned}
E\left|X^\varepsilon(t) - x(t)\right|^2 &\leqslant KE\left[\left|X^\varepsilon(0) - x_0\right|^2 + \int_0^T \phi(\varepsilon, x)\, ds\right. \\
&\quad \left. + \int_0^T E\left(u_1^\varepsilon - u_1\right)^2 ds + \int_0^T E\left(u_2^\varepsilon - u_2\right)^2 ds\right] e^{KT}.
\end{aligned} \tag{6.6.25}$$

Equation (6.6.25) implies that,

$$\sup_{0 \leqslant t \leqslant T} E\left|X^\varepsilon(t) - x(t)\right|^2 \to 0 \ \text{as} \ \varepsilon \to 0.$$

Hence, $\|X^\varepsilon - x\| \to 0$ □

Once we have L^2-convergence, we can obtain pathwise convergence using the following arguments. Suppose there is no pathwise convergence of $X^{\varepsilon}(t,\omega)$ for $\omega \varepsilon A$, with $P(a) = \lambda > 0$. Then there is a sequence $\{\varepsilon_n\} \to 0$, such that for each ε_n, there is a $t_n \in (0,T]$, such that,

$$|X^{\varepsilon_n}(t_n,\omega) - x(t_n,\omega)| > \delta > 0, \quad \omega \in A.$$

Hence,

$$0 < \varepsilon^2 \delta < \int_A \left|X^{\varepsilon_n - x}\right|^2 dp \leqslant E\left|X^{\varepsilon_n} - x\right|^2 \leqslant \sup_{0 \leqslant t \leqslant T} E\left|X^{\varepsilon_n} - x\right|^2.$$

Since $\sup_{0 \leqslant t \leqslant T} \; E|X^{\varepsilon_n} - x|^2 \to 0$ as $\varepsilon_n \to 0$, this leads to a contradiction. The convergence of the payoffs and near optimality for this setup follows as in the earlier sections.

The type of asymptotic results derived in this chapter has two main benefits. From the theoretical point of view, one obtains a stability result for the optimal strategy pair of a diffusion or deterministic system in the sense that this policy pair is asymptotically optimal for a large class of complicated problems of stochastic games. From a practical point of view, when a direct approach would be impossible, these results allow one to compute an asymptotically optimal strategy pair for a variety of problems under quite general conditions. In Kushner and Depuis [112], such approximation techniques are utilized in developing numerical methods for stochastic control problems. For a class of pursuit-evation games, a nice computational approach is given in Raivio and Ehtamo [157].

Chapter 7

Weak Convergence in Many Player Games

7.1 Introduction

In this chapter, we will discuss the weak convergence methods for n-person games. The entire problem will be set in a relaxed control framework. The advantage is that the problem becomes linear in control variables. The main advantage of occupation measure setting is that the differential game problem reduces to a static game on the set of occupation measures, the dynamics of the game being captured in these measures. In the proofs, we will use the weak convergence theory. We will only explain the case of averege payoffs. Discounted and other payoffs structure can be dealt in a similar fasion.

7.2 Some Popular Payoffs

In this section, we will look at weak convergence with few of the popular payoof stuctures, such as, average payoff, pathwise discounted payoffs, and discrete games.

7.2.1 *Avergage Payoffs*

7.2.1.1 *Problem Description*

Let the diffusion model be given in a non-anticipative relaxed control frame work. Let U_i, $i = 1, \ldots, N$ be compact metric spaces (we can take U_i as compact subsets of $\mathbb{R}^d$), and $M_i = P(U_i)$, the space of probability measures on U_i with Prohorov topology. Use the notation $m^{\widehat{k}} = (m_1, \ldots, m_{k-1}, m_{k+1}, \ldots, m_N)$ and $\left(m^{\widehat{k}}, m^{\widetilde{k}}\right) = (m_1, \ldots, m_{k-1}, \widetilde{m}_k, m_{k+1}, \ldots, m_N)$.

For $m = (m_1, \ldots, m_N) \in M = M_1 \times \cdots \times M_N$ and $U = U_1 \times \cdots \times U_N$, $x(\cdot) \in \mathbb{R}^d$ be an $\mathbb{R}^d$-valued process given by the following controlled stochastic differential equation

$$dx(t) = \int_U a(x(t), \alpha) m_t(d\alpha)dt + \overline{g}(x(t))dt + \sigma(x(t))dw(t), \quad \text{with } x(0) = x_0, \tag{7.2.1}$$

where we use the notation $a(\cdot,\cdot) = (a_1(\cdot,\cdot), \ldots, a_N(\cdot,\cdot))' : \mathbb{R}^d \times U \to \mathbb{R}$, $\alpha = (\alpha_1, \ldots, \alpha_N)$, $\sigma = [[\sigma_{ij}]]$, $1 \leqslant i, j \leqslant d$: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$, and

$$\int_U a_i(x, \alpha) m_t(d\alpha) \doteq \int_{U_N} \cdots \int_{U_1} a_i(x, \alpha_1, \ldots, \alpha_N) m_{1t}(d\alpha_1) \ldots m_{Nt}(d\alpha_N).$$

The pathwise average payoff per unit time for player k is given by

$$J_k[m] = \liminf_{T\to\infty} \frac{1}{T} \int_0^T \int r_k(x(s), \alpha) m_s(d\alpha) ds. \tag{7.2.2}$$

Let $w(\cdot)$ in (7.2.1) be a Wiener process with respect to a filtration $\{\Im_t\}$ and let Ω_i, $i = 1, 2, \ldots, N$ be a compact set in some Euclidean space. A measure valued random variable $m_i(\cdot)$ is an *admissible strategy* for the i^{th} player if $\int \int_0^t f_i(s, \alpha_i) m_i(dsd\alpha_i)$ is progressively measurable for each bounded continuous $f_i(\cdot)$ and $m_i([0,t] \times \Omega_i) = t$, for $t \geqslant 0$. If $m_i(\cdot)$ is admissible then there is a derivative $m_{it}(\cdot)$ (defined for almost all t) that is non-anticipative with respect to $w(\cdot)$ and

$$\int_0^t \int f_i(s, \alpha_i) m_i(dsd\alpha_i) = \int_0^t ds \int f_i(s, \alpha_i) m_{is}(d\alpha_i)$$

for all t with probability one (w.p.1.). The results derived in this work are for so called *Markov strategies*, which is a measure on the Borel sets of Ω_i for each x, and $m_i(c)$ is Borel measurable for each Borel measurable set C. We will denote by A_i the set of admissible strategies and M_{ai} the set of Markov strategies for the player i. One can introduce appropriate metric topology under which M_{ai} is compact, reader can refer to Borkar and Ghosh [31].

An N-tuple of strategies $m^* = (m_1^*, \ldots, m_N^*) \in A_1 \times \cdots \times A_N$ is said to be *ergodic equilibrium* (in the sense of Nash) for initial law π if for $k = 1, \ldots, N$, we have

$$J_k[m^*](\pi) \geqslant J_k[m^{*\widehat{k}}, m_k](\pi),$$

for any $m_k \in A_k$. Fix a $k \in \{1, \ldots, N\}$. Let $m_k^* \in M_{ak}$ be, such that,

$$J_k^*[m] \doteq J_k[m^{\widehat{k}}, m_k^*] = \max_{m_k \in M_k} J[m^{\widehat{k}}, m_k].$$

If all but player k use strategies $m^{\widehat{k}}$ then player k can not get a higher payoff than $J_k^*[m]$ by going beyond M_{ak} a.s.. We say that m_k^* is *ergodic optimal response* for player k given m.

An N-tuple of strategies $m^{\delta} = \left(m_1^{\delta},\ldots,m_N^{\delta}\right)$ is a δ-*ergodic equilibrium* for initial law π if for any $k = 1,\ldots,N$, we have,

$$J_k[m^*](\pi) \geqslant \sup_{m_k \in A_k} J_k[m^{\widehat{k}}, m_k] - \delta.$$

The wide band noise system considered in this work is of the following type:

$$dx^{\varepsilon} = \int a\left(x^{\varepsilon},\alpha\right) m_t^{\varepsilon}(d\alpha)dt + G\left(x^{\varepsilon},\xi^{\varepsilon}(t)\right) + \frac{1}{\varepsilon} g\left(x^{\varepsilon},\xi^{\varepsilon}\right)dt, \tag{7.2.3}$$

and pathwise average payoff per unit time for player k is given by

$$J_k[m^{\varepsilon}] = \liminf_{T\to\infty} \frac{1}{T}\int_0^T \int r_k\left(x^{\varepsilon}(s),\alpha\right) m_s^{\varepsilon}(d\alpha)ds. \tag{7.2.4}$$

An *admissible relaxed strategy* $m_k^{\varepsilon}(\cdot)$ for the k^{th} player with system (7.2.3) is a measure valued random variable satisfying $\int\int_0^t f(s,\alpha)\, m^{\varepsilon}(dsd\alpha)$ is progressively measurable with respect to $\{\Im_t^{\varepsilon}\}$, where $\Im_t^{\varepsilon}$ is the minimal $\sigma-$ algebra generated by $\{\xi^{\varepsilon}(s), x^{\varepsilon}(s), s \leqslant t\}$. Also $m^{\varepsilon}([0,t]\times U) = t$ for all $t \geqslant 0$. Also, there is a derivative m_t^{ε}, where $m_t^{\varepsilon}(b)$ are $\Im_t^{\varepsilon}$ measurable for Borel B. We will use following assumptions, which are very general. For a detailed description on these types of assumptions, we refer the reader to Kushner [103] and Kushner and Dupuis [112].

We introduce the following assumptions.

(A7.2.1): Let $a_i(\cdot,\cdot)$, $G(\cdot,\cdot)$, $g(\cdot,)$, $g_x(\cdot,\cdot)$ are continuous and are bounded by $O(1+|x|)$. $G_x(\cdot,\xi)$ is continuous in x for each ξ and is bounded. Also, $\xi(\cdot)$ is bounded, right continuous, and $EG(x,\xi(t)) \to 0$, $Eg(x,\xi(t)) \to 0$ as $t \to \infty$, for each x.

(A7.2.2): Let $g_{xx}(\cdot,\xi)$ be a continuous function for each ξ, and is bounded.

(A7.2.3): Let $W(x,\xi)$ denote either $\varepsilon G(x,\xi)$, $G_x(x,\xi)$, $g(x,\xi)$ or $g_x(x,\xi)$. Then for compact Q, we have,

$$\varepsilon \sup_{x\in Q}\left|\int_{t/\varepsilon^2}^{\infty} E_t^{\varepsilon} W(x,\xi(s))ds\right| \xrightarrow{\varepsilon} 0,$$

in the mean square sense, uniformly in t.

(A7.2.4): Let g_i denote the i^{th} component of g. There are continuous $\overline{g}_i(\cdot), b(\cdot) = \{b_{ij}(\cdot)\}$ such that,

$$\int_t^{\infty} Eg_{i,x}(x,\xi(s))g\left(x,\xi(t)\right)ds \longrightarrow \overline{g}_i(x),$$

and

$$\int_t^{\infty} Eg_i(x,\xi(s))g_j\left(x,\xi(t)\right)ds \longrightarrow \frac{1}{2}b_{ij}(x),$$

as $t \to \infty$, and the convergence is uniform in any bounded x-set.

Note: Let $b(x) = \{b_{ij}(x)\}$. For $i \neq j$, it is not necessary that $b_{ij} = b_{ji}$. In that case define $\widetilde{b}(x) = \frac{1}{2}[b(x) + b'(x)]$ as the symmetric covariance matrix, then use b for the new $\widetilde{b}$. Hence, for notational simplicity, we will not distinguish between $b(x)$ and $\widetilde{b}(x)$.

(A7.2.5): For each compact set Q and all i, j,

$$\text{(a)} \quad \sup_{x \in Q} \varepsilon^2 \left| \int_{t/\varepsilon^2}^{\infty} d\tau \int_{\tau}^{\infty} ds \left[E_{t/\varepsilon^2} g'_{i,x}(x, \xi(s)) g(x, \xi(t)) - \mathrm{E} g'_{i,x}(x, x(s)) g(x, x(t)) \right] \right| \to 0;$$

and

$$\text{(b)} \quad \sup_{x \in Q} \varepsilon^2 \left| \int_{t/\varepsilon^2}^{\infty} d\tau \int_{\tau}^{\infty} ds \left[E_{t/\varepsilon^2} g_i(x, \xi(s)) g_j(x, \xi(t)) - \mathrm{E} g_i(x, x(s)) g_j(x, x(t)) \right] \right| \to 0;$$

in the mean square sense as $\varepsilon \to 0$, uniformly in t.

Now, define $\overline{a}(x, \alpha) = a(x, \alpha) + \overline{g}(x)$ and the operator A^m as

$$A^m f(x) = \int A^\alpha f(x) m_x(d\alpha),$$

where

$$A^\alpha f(x) = f'_x(x) \overline{a}(x, \alpha) + \frac{1}{2} \sum_{i,j} b_{ij}(x) f_{x_i x_j}(x).$$

For a fixed control α, A^α will be the operator of the process that is the weak limit of $\{x^\varepsilon(\cdot)\}$.

(A7.2.6): The martingale problem for operator A^m has a unique solution for each relaxed admissible Markov strategy $m_x(\cdot)$, and each initial condition. The process is a Feller process. The solution of (7.2.1) is unique in the weak sense for each $\varepsilon > 0$. Also $b(x) = \sigma(x)\sigma'(x)$ for some continuous finite dimensional matrix $\sigma(\cdot)$.

For an admissible relaxed policy for (7.2.3) and (7.2.1), respectively, define the occupation measure valued random variables $P_T^{m,\varepsilon}(\cdot)$ and $P_T^m(\cdot)$ by, respectively,

$$P_T^{m,\varepsilon}(B \times C) = \frac{1}{T} \int_0^T I_{\{x^\varepsilon(t) \in B\}} m_t^\varepsilon(c) dt,$$

and

$$P_T^m(B \times C) = \frac{1}{T} \int_0^T I_{\{x(t) \in B\}} m_t(c) dt.$$

Let $\{m^\varepsilon(\cdot)\}$ be a given sequence of admissible relaxed controls.

(A7.2.7): For a fixed $\delta > 0$,

$$\{x^\varepsilon(t), \text{ small } \varepsilon > 0, \ t \in \text{ dense set in } [0, \infty), \ m^\varepsilon \text{ used}\}$$

are tight.

Note: The assumption (A7.2.7) implies that the set of measure valued random variables

$$\{P_T^{m^\varepsilon,\varepsilon}(\cdot), \text{small } \varepsilon > 0, \ T < \infty\}$$

are tight.

(A7.2.8): For $\delta > 0$, there is an N-tuple of Markov strategies $m^\delta = \left(m_1^\delta, \dots, m_N^\delta\right)$ which is a δ-*ergodic equilibrium* for initial law π for (7.2.1) and (7.2.2), and for which the martingale problem has a unique solution for each initial condition. The solution is a Feller process and there is a unique invariant measure $\eta(m^\delta)$.

Note: Existence of such an invariant measure is assured if the process is positive recurrent. Also, a Lyapunov type stability condition as in Borkar and Ghosh [31] will assure the assumption (A7.2.8).

(A7.2.9): Let $r_k(\cdot,\cdot)$ be bounded and continuous function. Also,

$$r(x, m_1, \dots, m_N) = \sum_{k=1}^{N} r_k(x, m_k) \ \text{ and } \ a(x, m_1, \dots, m_N) = \sum_{k=1}^{N} a_k(x, m_k).$$

Borkar and Ghosh, [31], under the Lyapunov type stability condition and (A7.2.9), following result is proved.

Theorem 7.2.1. *There exists an ergodic equilibrium* $m^* = (m_1^*, \dots, m_N^*) \in M_{a1} \times \cdots \times M_{aN}$.

7.2.1.2 *Convergence Result*

The following result gives the main convergence and δ- optimality result for the ergodic payoff criterion.

Theorem 7.2.2. *Assume* (A7.2.1)–(A7.2.9). *Let* (7.2.3) *have a unique solution for each admissible relaxed policy and each* ε. *Then for* m^δ *of* (A7.2.8)*, following holds:*

$$\underline{\lim}_{\varepsilon, T} P\left\{J_k(m^\varepsilon) \geqslant J_k(m^\delta) - \delta\right\} = 1, \tag{7.2.5}$$

for any sequence of admissible relaxed policies $m^\varepsilon(\cdot)$.

Proof. The correct procedure of proof is to work with the truncated processes $x^{\varepsilon,K}(\cdot)$ and to use the piecing together idea of Lemma 6.2.3 to get convergence of the original $x^\varepsilon(\cdot)$ sequence, unless $x^\varepsilon(\cdot)$ is bounded on each $[0,T]$, uniformly in ε. For notational simplicity, we ignore this technicality. Simply suppose that $x^\varepsilon(\cdot)$ is bounded in the following analysis. Otherwise, one can work with K-truncation. Let $\widehat{D}$ be a measure determining set of bounded real-valued continuous functions on $\mathbb{R}^d$ having continuous second partial derivatives and compact support. Let $m_t^\varepsilon(\cdot)$ be the relaxed Markov policies of (A7.2.8). Whenever convenient, we write $x^\varepsilon(t) = x$. For the test function $f(\cdot) \in \widehat{D}$, define the perturbed test functions (the change of variable $s/\varepsilon^2 \to s$ will be used through out the proofs).

Thus,

$$
\begin{aligned}
f_0^\varepsilon(x,t) &= \int_t^\infty E_t^\varepsilon f_x'(x) G(x,\xi^\varepsilon(s))ds \\
&= \varepsilon^2 \int_{t/\varepsilon^2}^\infty E_t^\varepsilon f_x'(x) G(x,\xi(s))ds, \\
f_1^\varepsilon(x,t) &= \frac{1}{\varepsilon}\int_t^\infty E_t^\varepsilon f_x'(x) g(x,\xi^\varepsilon(s))ds \\
&= \varepsilon \int_{t/\varepsilon^2}^\infty E_t^\varepsilon f_x'(x) g(x,\xi(s))ds,
\end{aligned}
$$

and

$$
\begin{aligned}
f_2^\varepsilon(x,t) = \frac{1}{\varepsilon^2}\int_t^\infty ds \int_s^\infty d\tau \Big\{ & E_t^\varepsilon [f_x'(x) g(x,\xi^\varepsilon(\tau))]_x' g(x,\xi^\varepsilon(s)) \\
& - E\left[f_x'(x) g\left(x,\xi^\varepsilon(\tau)\right)\right]_x' g(x,\xi^\varepsilon(s))\Big\} \\
= \varepsilon^2 \int_{t/\varepsilon^2}^\infty ds \int_s^\infty d\tau \Big\{ & E_t^\varepsilon [f_x'(x) g(x,\xi(\tau))]_x' g(x,\xi(s)) \\
& - E\left[f_x'(x) g(x,\xi(\tau))\right]_x' g(x,\xi(s))\Big\}.
\end{aligned}
$$

From assumptions (A7.2.1), (A7.2.2), (A7.2.3), and (A7.2.5), $f_i^\varepsilon(\cdot) \in D(A^\varepsilon)$, for $i = 0,1,2$. Define the perturbed test function by

$$
f^\varepsilon(t) = f(x^\varepsilon(t)) + \sum_{i=0}^2 f_i^\varepsilon(x^\varepsilon(t),t).
$$

The reason for defining f_i^ε in such a form is to facilitate the averaging of the "noise" terms involving ξ^ε terms. By the definition of the operator A^ε and its domain $D(A^\varepsilon)$, we will obtain that $f(x^\varepsilon(\cdot))$ and the $f_i^\varepsilon(x^\varepsilon(\cdot),\cdot)$ are all in $D(A^\varepsilon)$, and

$$
\begin{aligned}
A^{m^\varepsilon,\varepsilon} f(x^\varepsilon(t)) = f_x'(x^\varepsilon(t)) \Bigg[& \sum_{i=1}^N \int a_i(x^\varepsilon(t),\alpha) m_t^\varepsilon(d\alpha) + G(x^\varepsilon(t),\xi^\varepsilon(t)) \\
& + \frac{1}{\varepsilon} g(x^\varepsilon(t),\xi^\varepsilon(t))\Bigg].
\end{aligned} \tag{7.2.6}
$$

From this expression we can obtain,

$$
\begin{aligned}
A^{m^\varepsilon,\varepsilon} f_0(x^\varepsilon(t)) &= -f_x'(x^\varepsilon(t)) G(x^\varepsilon(t),\xi^\varepsilon(t)) + \int_t^\infty ds [E_t^\varepsilon f_x'(x^\varepsilon(t)) G(x^\varepsilon(t),\xi^\varepsilon(s))]_x' x^\varepsilon(t) \\
&= -f_x'(x^\varepsilon(t)) G(x^\varepsilon(t),\xi^\varepsilon(t)) + \varepsilon^2 \int_{t/\varepsilon^2}^\infty ds [E_t^\varepsilon f_x'(x^\varepsilon(t)) G(x^\varepsilon(t),\xi(s))]_x' \dot{x}^\varepsilon(t).
\end{aligned} \tag{7.2.7}
$$

Note that the first term in (7.2.7) will cancel with $f_{x'}G$ term of (7.2.6). The p-lim of the last term in (7.2.7) is zero.

Also, we can write,

$$
\begin{aligned}
A^{m^\varepsilon,\varepsilon} f_1(x^\varepsilon(t)) &= -\frac{1}{\varepsilon} f_x'(x^\varepsilon(t)) g(x^\varepsilon(t),\xi^\varepsilon(t)) + \frac{1}{\varepsilon}\int_t^\infty ds [E_t^\varepsilon f_x'(x^\varepsilon(t)) g(x^\varepsilon(t),\xi^\varepsilon(s))]_x' \overset{\varepsilon}{x}(t) \\
&= -\frac{1}{\varepsilon} f_x'(x^\varepsilon(t)) g(x^\varepsilon(t),\xi^\varepsilon(t)) + \varepsilon \int_{t/\varepsilon^2}^\infty ds [E_t^\varepsilon f_x'(x^\varepsilon(t)) g(x^\varepsilon(t),\xi(s))]_x' \overset{\varepsilon}{x}(t)
\end{aligned} \tag{7.2.8}
$$

The first term on the right of (7.2.8) will cancel with the $\frac{f_x' g}{\varepsilon}$ term in (7.2.6). The only component of the second term on the right of (7.2.6) whose p-lim is not zero is

$$\frac{1}{\varepsilon^2}\int_t^{\infty} ds\left\{E_t^{\varepsilon}[f_x'(x^{\varepsilon}(t))g(x^{\varepsilon}(t),\xi^{\varepsilon}(s))]_x' g(x^{\varepsilon}(t),\xi^{\varepsilon}(t))\right\}.$$

This term will cancel with the first term of equation (7.2.8).

We can also, write,

$$\begin{aligned}A^{m^{\varepsilon},\varepsilon} f_2(x^{\varepsilon}(t)) &= -\frac{1}{\varepsilon^2}\int_t^{\infty} ds\Big\{E_t^{\varepsilon}[f_x'(x^{\varepsilon}(t))g(x^{\varepsilon}(t),\xi^{\varepsilon}(s))]_x' g(x^{\varepsilon}(t),\xi^{\varepsilon}(t))\\ &\quad -E[f_x'(x^{\varepsilon}(t))g(x,\xi^{\varepsilon}(s))]_x' g(x,\xi^{\varepsilon}(t))|_{x=x^{\varepsilon}(t)}\Big\}+[f_2^{\varepsilon}(x,t)]_x'\dot{x}^{\varepsilon}|_{x=x^{\varepsilon}(t)}\\ &= -\int_{t/\varepsilon^2}^{\infty} ds\Big\{E_t^{\varepsilon}[f_x'(x^{\varepsilon}(t))g(x^{\varepsilon}(t),\xi(s))]_x' g(x^{\varepsilon}(t),\xi^{\varepsilon}(t))\\ &\quad -E[f_x'(x^{\varepsilon}(t))g(x,\xi(s))]_x' g(x,\xi^{\varepsilon}(t))|_{x=x^{\varepsilon}(t)}\Big\}+[f_2^{\varepsilon}(x,t)]_x'\dot{x}^{\varepsilon}|_{x=x^{\varepsilon}(t)}.\end{aligned} \tag{7.2.9}$$

The p-lim of the last term of the right side of equation (7.2.9) is zero.

Evaluating

$$A^{m^{\varepsilon},\varepsilon} f^{\varepsilon}(t) = A^{m^{\varepsilon},\varepsilon}\left[f(x^{\varepsilon}(t))+\sum_{i=0}^{2} f_i^{\varepsilon}(x^{\varepsilon}(t),t)\right]$$

and by deleting terms that cancel yields

$$\begin{aligned}A^{m^{\varepsilon},\varepsilon} f^{\varepsilon}(t) &= f_x'(x^{\varepsilon}(t))\sum_{i=1}^{N}\int a_i(x^{\varepsilon}(t),\alpha)m_t^{\varepsilon}(d\alpha)\\ &\quad +\int_{t/\varepsilon^2}^{\infty} E[f_x'(x^{\varepsilon}(t))g(x,\xi(s))]' g(x,\xi(t/\varepsilon^2))ds.\end{aligned} \tag{7.2.10}$$

As a result of the above, we have,

$$p\text{-}\lim\left(f^{\varepsilon}(t)-f(x^{\varepsilon}(\cdot))\right)=0 \tag{7.2.11}$$

and

$$p\text{-}\lim_{\varepsilon}\left|A^{m^{\varepsilon},\varepsilon} f(x^{\varepsilon}(t)) - A^{m^{\varepsilon,\varepsilon}} f^{\varepsilon}(t)\right| = 0. \tag{7.2.12}$$

Hence, by Lemma 6.2.2, we have,

$$M_f^{\varepsilon}(t) = f^{\varepsilon}(t) - f^{\varepsilon}(0) - \int_0^t A^{m^{\varepsilon}} f^{\varepsilon}(s)ds,$$

is a zero mean martingale.

Let $[t]$ denote the greatest integer part of t. We proceed to write

$$\frac{M_f^{\varepsilon}(t)}{t} = \frac{1}{t}\left[\left(M_f^{\varepsilon}(t)-M_f^{\varepsilon}([t])\right)+M_f^{\varepsilon}(0)\right]+\frac{1}{t}\sum_{k=0}^{[t]-1}\left[M_f^{\varepsilon}(k+1)-M_f^{\varepsilon}(k)\right].$$

Using equation (7.2.12) and the fact that $f(\cdot)$ is bounded, and the martingale property of $M_f^{\varepsilon}(\cdot)$, we get $E\left[\frac{M_f^{\varepsilon}(t)}{t}\right]^2 \to 0$ as $t\to\infty$ and $\varepsilon\to 0$, which in turn implies that $\frac{M_f^{\varepsilon}(t)}{t}\xrightarrow{P}0$ as

$t \to \infty$ and $\varepsilon \to 0$ in any way at all. From equation (7.2.12), and the fact that $\frac{M_f^\varepsilon(t)}{t}$, $\frac{f^\varepsilon(t)}{t}$, and $\frac{f^\varepsilon(0)}{t}$ all go to zero in probability implies that as $t \to \infty$ and $\varepsilon \to 0$, we have,

$$\frac{1}{t}\int_0^t A^{m^\varepsilon} f\left(x^\varepsilon(s)\right) ds \xrightarrow{P} 0. \tag{7.2.13}$$

By the definition of $P_T^{m^\delta,\varepsilon}(\cdot)$, equation (7.2.13) can be written as

$$\int A^\alpha f(x) P_T^{m^\varepsilon,\varepsilon}(dxd\alpha) \xrightarrow{P} 0 \text{ as } T \to \infty, \text{ and } \varepsilon \to 0. \tag{7.2.14}$$

For the policy $m^\delta(\cdot)$, choose a weakly convergent subsequence of a set of random variables $\{P_T^{m^\delta,\varepsilon}(\cdot), \varepsilon, T\}$, indexed by ε_n, T_n, with limit $\widehat{\mu}(\cdot)$. Let this limit $\widehat{P}(\cdot)$ be defined on some probability space $\left(\widetilde{\Omega}, \widetilde{P}, \widetilde{\Im}\right)$ with generic variable $\widetilde{\omega}$. Factor $\widehat{P}(\cdot)$ as $\widehat{P}(dxd\alpha) = m_x^\delta(d\alpha)\mu(dx)$. We can suppose that $m_x(c)$ are $x-$measurable for each Borel set C and $\widetilde{\omega}$. Now equation (7.2.14) implies that for all $f(\cdot) \in \widehat{\mathrm{D}}$, we have,

$$\int\int A^\alpha f(x) m_x^\delta(d\alpha)\widehat{\mu}(dx) = 0 \text{ for } \widetilde{P}\text{-almost all } \widetilde{\omega}. \tag{7.2.15}$$

Since $f(\cdot)$ is measure determining, (7.2.15) implies that almost all realizations of $\widehat{\mu}$ are invariant measures for (7.2.1) under the relaxed policies m^δ. By uniqueness of the invariant measure, we can take $\mu\left(m^\delta, \cdot\right) = \widehat{\mu}(\cdot)$ does not depend on the chosen subsequence ε_n, T_n. By the definition of $P_T^{m^\delta,\varepsilon}(\cdot)$,

$$\begin{aligned}\frac{1}{t}\int_0^t \int r_k(x^\varepsilon(s), \alpha) m^\delta(d\alpha) ds &= \int_0^t \int r_k(x^\varepsilon(s), \alpha) P_T^{m^\delta,\varepsilon}(d\alpha dx)\\ &\xrightarrow{P} \int_0^t \int r_k(x, \alpha) m_x^\delta(d\alpha)\widehat{\mu}(dx) = J_k(m^\delta).\end{aligned}$$

Since $m^\delta(\cdot)$ is a δ-equilibrium policy, by the definition of $\delta-$equilibrium, for almost all $\widetilde{\omega}$ we have $J_k\left(m^\varepsilon\right) \geqslant J_k(m^\delta) - \delta$. Since this is true for all the limits of the tight set $\{P_T^{m^\delta,\varepsilon}(\cdot); \varepsilon, T\}$, (7.2.5) follows. □

It is important to note that, as a result of Theorem 7.2.2, if one needs a δ-optimal policy for the physical system, it is enough to compute for the diffusion model and use it to the physical system. There is no need to compute optimal policies for each ε.

7.2.2 *Pathwise Discounted Payoffs*

Define the pathwise discounted payoffs for the player k by

$$R_k^{\lambda,\varepsilon}\left(m^\varepsilon\right) = \lambda \int_0^\infty e^{-\lambda s} \int r_k\left(x^\varepsilon(s), \alpha\right) m_s(d\alpha) ds. \tag{7.2.16}$$

Now we will state the pathwise result for discounted payoff and suggest the necessary steps needed in the proof.

Theorem 7.2.3. *Let m^ε be a sequence of δ-optimal discounted payoffs and m^δ be δ-equilibrium policies for* (7.2.1). *Under the conditions of Theorem* 7.2.2*, following limits hold:*

$$R_k^{\lambda,\varepsilon}(m^\delta) \xrightarrow{P} J_k(m^\delta) \text{ as } \lambda \to 0, \varepsilon \to 0, \tag{7.2.17}$$

$$\underline{\lim}_{\varepsilon,T} P\{R_k^{\lambda,\varepsilon}(m^\varepsilon) \geqslant J_k(m^\delta) + \delta\} = 1. \tag{7.2.18}$$

Proof. The proof is essentially the same as that of Theorem 7.2.2. We will only explain the differences that are needed to follow. Define the discounted occupation measures by

$$P_\lambda^{m^\varepsilon,\varepsilon}(B \times C) = \lambda \int_0^\infty e^{-\lambda t} I_{\{x^\varepsilon(t) \in B\}} m_t(c) dt$$

and

$$P_\lambda^m(B \times C) = \lambda \int_0^\infty e^{-\lambda t} I_{\{x(t) \in B\}} m_t(c) dt.$$

Then equation (7.2.16) can be written as

$$R_k^{\lambda,\varepsilon}(m^\varepsilon) = \int r_k(x(s), \alpha) P_\lambda^{m^\varepsilon,\varepsilon}(dxd\alpha).$$

By the tightness condition (A7.2.7), the $\{P_\lambda^{m^\varepsilon,\varepsilon}(\cdot)\}$ and $\{P^{m^\delta,\varepsilon}(\cdot)\}$ are tight. Define,

$$f_\lambda^\varepsilon(t) = \lambda e^{\lambda t} f^\varepsilon(t).$$

This will be used in the place of $f^\varepsilon(\cdot)$ defined in Theorem 7.2.2. Then, we have

$$A^{m^\varepsilon,\varepsilon} f_\lambda^\varepsilon(t) = -\lambda^2 e^{\lambda t} f^\varepsilon(t) + \lambda e^{\lambda t} A^{m^\varepsilon,\varepsilon} f^\varepsilon(t).$$

Define the martingale by

$$f_\lambda^\varepsilon(t) - f_\lambda^\varepsilon(0) - \int_0^t A^{m^\varepsilon,\varepsilon} f_\lambda^\varepsilon(s) ds$$

$$= \lambda e^{\lambda t} f^\varepsilon(t) - \lambda f^\varepsilon(0) - \int_0^t [-\lambda^2 e^{\lambda s} f^\varepsilon(s) + \lambda e^{\lambda s} A^{m^\varepsilon,\varepsilon} f^\varepsilon(s)] ds.$$

As in Theorem 7.2.2, we conclude that

$$\lim_{(\lambda,\varepsilon)\to 0} \int\int A^\alpha f(x) P_\lambda^{m^\varepsilon,\varepsilon}(dxd\alpha) = 0.$$

Thus,

$$\lim_{(\lambda,\varepsilon)\to 0} \int\int A^\alpha f(x) P_\lambda^{m^\varepsilon,\varepsilon}(dxd\alpha) = 0.$$

Now choose weakly convergent subsequences of the $\{P_\lambda^{m^\varepsilon,\varepsilon}(\cdot)\}$ or $\{P_\lambda^{m^\delta,\varepsilon}(\cdot)\}$ and continue as in the proof of Theorem 7.2.2 to obtain (7.2.17) and (7.2.18). □

7.2.3 *Discrete Parameter Games*

The discrete parameter system is given by

$$X_{n+1}^{\varepsilon} = X_n^{\varepsilon} + \varepsilon G(X_n^{\varepsilon}) + \varepsilon \sum_{i=1}^{N} \int a_i(X_n^{\varepsilon}, \alpha_i) m_{in}(d\alpha_i) + \sqrt{\varepsilon} g(X_n^{\varepsilon}, \xi_n^{\varepsilon}) \tag{7.2.19}$$

where $\{\xi_n^{\varepsilon}\}$ satisfies the discrete parameter version of (A7.2.2) and $m_{in}(\cdot)$, $i = 1, \ldots, N$, the relaxed control strategies depending only on $\{X_i, \xi_{i-1},\ i \leqslant n\}$. It should be noted that, in the discrete case, strategies would not be relaxed, one need to interpret this in the asymptotic sense, i.e., the limiting strategies will be relaxed. Let E_n^{ε} denote the conditional expectation with respect to $\{X_i, \xi_{i-1},\ i \leqslant n\}$. Define, $x^{\varepsilon}(\cdot)$ by $x^{\varepsilon}(t) = X_n^{\varepsilon}$ on $[n\varepsilon, n\varepsilon + \varepsilon)$ and $m_i(\cdot)$ by

$$m_i(B_i \times [0,t]) = \varepsilon \sum_{n=0}^{[t/\varepsilon]-1} m_{in}(B_i) + \varepsilon(t - \varepsilon t/\varepsilon]) m_{[t/\varepsilon]}(B_i), \quad i = 1, \ldots, N.$$

(A7.2.10):

(i) For V equal to either $a(\cdot,\cdot)$, g or g_x, and for Q compact,

$$E \sup_x \left| \sum_{n+L_1}^{L} E_n^{\varepsilon} V(x, \xi_i^{\varepsilon}) \right| \to 0,$$

as L, n and $L_1 \to \infty$, with $L > n + L_1$ and $L - (n + L_1) \to \infty$.

(ii) There are continuous functions $c(i,x)$ and $c_0(i,x)$, such that, for each x

$$\frac{1}{L} \sum_{n=\ell}^{\ell+L} E_\ell^{\varepsilon} g(x, \xi_{n+i}^{\varepsilon}) g'(x, \xi_n^{\varepsilon}) \xrightarrow{P} c(i,x)$$

and

$$\frac{1}{L} \sum_{n=\ell}^{\ell+L} E_\ell^{\varepsilon} g_x'(x, \xi_{n+i}^{\varepsilon}) g(x, \xi_n^{\varepsilon}) \xrightarrow{P} c_0(i,x)$$

as ℓ and $L \to \infty$.

(iii) For each $T < \infty$ and compact Q,

$$\varepsilon \sup_{x \in Q} \left| \sum_{j=n}^{T/\varepsilon} \sum_{k=j+1}^{T/\varepsilon} [E_n^{\varepsilon} g_{i,x}'(x, \xi_k) g(x, \xi_j) - E g_{i,x}'(x, \xi_k) g(x, \xi_j)] \right| \to 0, \quad i \leqslant n,$$

and

$$\varepsilon \sup_{x \in Q} \left| \sum_{j=n}^{T/\varepsilon} \sum_{k=j+1}^{T/\varepsilon} [E_n^{\varepsilon} g'(x, \xi_k) g(x, \xi_j) - E g'(x, \xi_k) g(x, \xi_j)] \right| \to 0,$$

in the mean as $\varepsilon \to 0$ uniformly in $n \leqslant T/\varepsilon$. Also, the limits hold when the bracketed terms are replaced by their $x-$gradient/$\sqrt{\varepsilon}$.

Define,

$$\widetilde{a}(x) = \sum_{1}^{\infty} c_0(i,x)$$

and

$$\widetilde{c}(x) = c(0,x) + 2\sum_{1}^{\infty} c(i,x) = \sum_{-\infty}^{\infty} c(i,x).$$

With some minor modifications in the proof of Theorem 6.4.2, we can obtain the following result. The reader can find complete analysis in Kushner [103] and Ramachandran [161].

Theorem 7.2.4. *Assume* (A7.2.1) *to* (A7.2.3)*,* (A7.2.6) *to* (A7.2.9) *and* (7.2.10)*. Then the conclusions of Theorem* 7.2.2 *hold for model* (7.2.19)*.*

The results of this section can be directly applied to two person zero-sum differential games with pathwise payoff structure, analogous to the results in Ramachandran [158]. If the coefficients in (7.2.19) are state dependent or even discontinuous, still we can obtain the convergence results by adapting the methods of Ramachandran [161]. Also, other cost structures, such as finite horizon payoff, and payoff up to exit time can be handled by some minor modifications.

7.3 Deterministic Approximations in N-Person Differential Games

In this section, we will extend the results of Section 6.6 to many player case. Consider an N-person noncooperative dynamic game problem where the evolution of the system is given by the following deterministic ordinary differential equation:

$$dx(t) = \left[\overline{a}(x(t)) + \sum_{i=1}^{N} b_i(x(t))u_i(t)\right] dt \tag{7.3.1}$$

with

$$x(0) = x_0,$$

where $x(t)$ is deterministic controlled process, $u_i(t)$, $i = 1,2,\ldots,N$ are deterministic controls for each of the N-players. Let U_i, $i = 1,\ldots,N$, be compact metric spaces (we can take U_i as compact subsets of $\mathbb{R}^d$). Let $U = U_1 \times \cdots \times U_N$. Also, $u \in U$ is called an N-dimentional strategy vector. We denote $u_i(t) \in U_i$ as the i^{th} component of u and u_{-i} denotes the $N-1$ dimentional vector obtained by removing the i^{th} component of vector u, $i = 1,2,\ldots,N$. We define payoff to player k by

$$J_k(u_1,\ldots,u_N) = \int_0^T \left[k(x(t)) + \sum_{i=1}^{N} P_i(u_i(t))\right] dt + r(x(t)) \tag{7.3.2}$$

where $T < \infty$ is the fixed terminal time for the game. An N-tuple of strategies $u^* = (u_1^*, \ldots, u_N^*) \in U$ is said to be in equilibrium (in the sense of Nash) if for each $k = 1, \ldots, N$, we have

$$J_k[u^*] \geqslant J_k[u_{-k}^*, u_k],$$

for any $u_k \in U_k$. Fix a $k \in \{1, \ldots, N\}$. An N-tuple of strategies $u^\delta = (u_1^\delta, \ldots, u_N^\delta)$ is a δ-equilibrium if for any $k = 1, \ldots, N$, results in,

$$J_k\left[u^\delta\right] \geqslant \sup_{u_k \in U_k} J_k\left[u_{-k}^\delta, u_k\right] - \delta.$$

This concept of δ-equilibrium is important in the theory of approximation.

Since most of the physical systems are stochastic in nature, the deterministic models are only approximations to the real systems. Now consider a more realistic physical model for an N-person game problem described by a family of stochastic equations parametrized by a small parameter $\varepsilon(\varepsilon \downarrow 0)$, with dynamics

$$dX^\varepsilon(t) = \left[a\left(X^\varepsilon(t), \xi^\varepsilon(t)\right) + \sum_{i=1}^N b_i\left(X^\varepsilon(t)\right) u_i^\varepsilon(t)\right] dt + dM^\varepsilon(t), \tag{7.3.3}$$

and initial condition $X^\varepsilon(0)$. Here, $X^\varepsilon = (X^\varepsilon(t))$ is the controlled state process, $\xi = (\xi(t))$ is the contamination process affecting the drift of X^ε, and $M = (M^\varepsilon(t))$ is the process representing the noise in the system. Also $u_i^\varepsilon = (u_i^\varepsilon(t))$, $i = 1, \ldots, N$, are controls for each of the players. Given a finite horizon $T > 0$, with each strategy vector $u^\varepsilon = (u_1^\varepsilon, \quad u_2^\varepsilon, \ldots, u_N^\varepsilon)$, we associate the payoff to player k by

$$J_k^\varepsilon\left(u_1^\varepsilon, \quad u_2^\varepsilon, \ldots, u_N^\varepsilon\right) = E\left\{\int_0^T \left[k\left(X^\varepsilon(t)\right) + \sum_{i=1}^N p_i\left(u_i^\varepsilon(t)\right)\right] dt + r\left(X^\varepsilon(t)\right)\right\}, \tag{7.3.4}$$

where $k(x)$, $p_i(u_i)$, $i = 1, \ldots, N$ and $r(x)$ are nonnegative functions on the real line referred to as holding cost, control costs, and terminal cost functions, respectively. The Nash equilibrium and δ-equilibrium are defined analogously. Our objective is to find an N-tuple of strategies $u^\delta = u_1^\delta, \ldots, u_N^\delta$ that is, a δ-equilibrium if for each $k = 1, \ldots, N$, and the corresponding value function V_k^ε, that is,

$$V_k^\varepsilon = \max_{u_i^\varepsilon \in A_1} J_k^\varepsilon\left(u_k^\varepsilon, u_{-k}^{\varepsilon^*}\right) \tag{7.3.5}$$

The sets $A_1, A_2, \ldots, A_N$ will be defined in the next section. The process $\xi^\varepsilon(\cdot)$ is said to be *exogenous* or state independent if for each t and set B in $\sigma(\xi^\varepsilon(s),\ s > t)$, we have,

$$P\{B \mid \xi^\varepsilon(s),\ s \leqslant t\} = P\{B \mid \xi^\varepsilon(s),\ X^\varepsilon(s),\ s \leqslant t\}.$$

In order for desired convergence to occur, the "rate of fluctuations" of $\xi^\varepsilon(\cdot)$ must increase as $\varepsilon \to 0$. We consider the case in which the "intensity" of the random noise disturbance M

becomes very small with ε, while the "contaminating" process ξ fluctuates with increasing speed. In this study, we assume that the controlled state process X is completely observed. It is very hard to obtain optimal strategies and values satisfying (7.3.3) and (7.3.5). It is well known that only few stochastic game or stochastic control problems can be solved in closed form. For practical purposes one may just as well be interested in finding a near optimal or an asymptotically optimal strategy vector. Considerable effort has been put into developing approximation techniques for such problems. One such approach use in the stochastic control literature is, in lieu of the original model, a model where the underlying processes are replaced by simpler ones, for example, see (Fleming [60], Kushner [103], Kushner and ramachandran [113], Kushner and Runggaldier [114], and Liptser, Runggaldier and Taksar [127]). In stochastic game problems such an effort was made in Ramachandran [161], using diffusion approximation techniques.

In the present section, deterministic approximation techniques (i.e., the simpler model is deterministic) to a N-person non-zero sum differential game model will be developed. To this end, we will now introduce a deterministic model, which we will show to be the limiting model corresponding to (7.3.3) to (7.3.5) under introduced appropriate conditions.

With the use of the so called relaxed controls, we could allow nonlinear forms. However for simplicity of presentation, in this study we will restrict to the linearity in controls. These type of results have two major benefits. From the theoretical point of view, one obtains a stability result for the optimal strategy pair of a deterministic system in the sense that this strategy vector is asymptotically optimal for a large class of complicated problems of stochastic games. From a practical point of view, when a direct approach would be impossible, these results allow one to compute an asymptotically optimal strategy pair for a variety of stochastic game problems under quite general conditions.

7.3.1 *Main Convergence Results*

We will now present the main convergence result. The so called perturbed test function method will be utilized for the proof.

We will use following general assumptions similar to that in Section 6.6.

(A7.3.1): Let, $\xi^{\varepsilon}(t) = \xi\,(t/\varepsilon)$, where $\xi^{\varepsilon}(\cdot)$ is a stationary process which is strong mixing, right continuous and bounded with mixing rate function $\phi(\cdot)$ satisfying $\int_0^{\infty} \phi(s)ds < \infty$.

(A7.3.2): Let $b_i(\cdot)$, $i = 1,2,\ldots,N$, be bounded and Lipshitz continuous. $a(\cdot,\cdot)$ and its gradient $a_x(\cdot,\cdot)$ are continuous in (x,ξ) and satisfy uniform Lipschitz condition with the same constant.

(A7.3.3): There is a continuously differentiable function $\overline{a}(\cdot)$, such that, for each $t < T$ and x, we have,

$$\int_t^T \left[E_t^\varepsilon a\left(x, \xi^\varepsilon(s)\right) - \overline{a}(x)\right] ds \to 0$$

in probability as $\varepsilon \to 0$.

(A7.3.4): The cost functions $k(\cdot)$ and $r(\cdot)$ are continuous nonnegative satisfying

$$k(x), \quad r(x) \leqslant c_0\left(1 + |x|^\gamma\right), \quad c_0, \quad \gamma > 0.$$

Also, $\sum_{i=1}^N p_i\left(u_i(t)\right) \geqslant c_2\left(\sum_{i=1}^N |u_i|^{1+\gamma_2}\right)$, $c_2, \gamma_2 > 0$, and $p_i\left(u_i\right)$ are nonnegative convex.

(A7.3.5): The process $M^\varepsilon = \left(M^\varepsilon(t)\right)_{t \geqslant 0}$ is a square integrable martingale with paths in the Skorokhod space $D\left[0, \infty\right)$ whose predictable quadratic variations $\langle M^\varepsilon \rangle (t)$ satisfies

(i) $\langle M^\varepsilon \rangle (t) = \varepsilon \int_0^t m^\varepsilon(s)\, ds$ with bounded density $m^\varepsilon(s)$. That is, there exists a constat c_1 such that

(ii) $m^\varepsilon(t) \leqslant c_1$; $t \leqslant T$, P-a.s.
The jumps $\Delta M^\varepsilon(s) \doteq M^\varepsilon(s) - \lim_{v \uparrow s} M^\varepsilon(s)$ are bounded, i.e., there exists a constant $K > 0$ such that

(iii) $|\Delta M^\varepsilon(t)| \leqslant K$; $t \leqslant T$, $\varepsilon \in (0, 1]$.

(A7.3.6): The $p\text{-}\lim_{\varepsilon \to 0} X^\varepsilon(0) = x_0$, $x_\varepsilon \in \mathbb{R}$.

Note: These assumptions are general enough, but need not be most general. For instance, assumption (A7.3.2) could be relaxed to say that the equation (7.3.1) has a unique solution.

Theorem 7.3.1. *Suppose that* (A7.3.1)–(A7.3.6) *hold. Let* $X_0^\varepsilon \Rightarrow x_0$ *and* $u^\varepsilon(\cdot) \equiv (u_1^\varepsilon(\cdot), u_2^\varepsilon(\cdot), \ldots, u_N^\varepsilon) \to u(\cdot) \equiv (u_1(\cdot), u_2(\cdot), \ldots, u_N(\cdot))$, *where* $u(\cdot)$ *is an admissible strategy vector for* (7.3.1). *Then* $(X^\varepsilon(\cdot), u^\varepsilon(\cdot)) \to (x(\cdot), u(\cdot))$ *where* $u(\cdot)$ *is measurable* (*admissible*) *process and satisfies*

$$dx(t) = \left[\overline{a}(x(t)) + \sum_{i=1}^N b_i(x(t)) u_i(t)\right] dt. \tag{7.3.6}$$

Also,

$$J_k^{\varepsilon_n}\left(u_1^{\varepsilon_n}, u_2^{\varepsilon_n}, \ldots, u_N^{\varepsilon_n}\right) \longrightarrow J_k(u_1, u_2, \ldots, u_N). \tag{7.3.7}$$

Proof. Define a process $X^{\varepsilon,0}(\cdot)$ by

$$X^{\varepsilon,0}(t) = X^{\varepsilon,0}(0) + \int_0^t \left[a\left(X^{\varepsilon,0}(s), \xi^\varepsilon(s)\right) + \sum_{i=1}^N b_i\left(X^{\varepsilon,0}(s)\right) u_i^\varepsilon(s)\right] ds$$

Let $Y^{\varepsilon}(s) = \sup_{s\leqslant t}\left|X^{\varepsilon}(s) - X^{\varepsilon,0}(s)\right|$. Then by assumption (A7.3.5),

$$Y^{\varepsilon}(t) \leqslant K\int_0^t Y^{\varepsilon}(s)d\left[s+\sum_{i=1}^{N}\int_0^s |u_i^{\varepsilon}(w)|\,dw\right] + \sup_{s\leqslant T}|M^{\varepsilon}(s)|, \quad t\leqslant T,$$

where K is the Lipschitz constant. By the Gronwall-Bellman inequality we obtain

$$Y^{\varepsilon}(t) \leqslant K\sup_{s\leqslant T}|M^{\varepsilon}(s)|\exp\left\{K\left[T+\sum_{i=1}^{N}\int_0^T |u_i^{\varepsilon}(w)|\,dw\right]\right\}.$$

By assumption (A7.3.5) (see Liptser, Runggaldier, and Taksar [127]), $\sup_{s\leqslant T}|M^{\varepsilon}(s)| \to 0$, $\varepsilon \to 0$, in probability and by (A7.3.2) and (A7.3.4), we have,

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0} P\left(\sup_{t,s\leqslant T:|T-s|\leqslant\delta}\sum_{i=1}^{N}\int_s^t [|u_i(w)|]\,dw > \eta\right) = 0.$$

Consequently $Y^{\varepsilon}(t) \to 0$, $\varepsilon \to 0$, in probability and the theorem remains true if its statements are proved only for $\left(X^{\varepsilon,0}(\cdot), u^{\varepsilon}(\cdot)\right)$. The weak convergence for the process $\left(X^{\varepsilon,0}(\cdot), u^{\varepsilon}(\cdot)\right)$can be proved similar to the proof of Theorem 6.6.1 using perturbed test function method,for details refer to Ramachandran and Rao [164]. □

The following result states the near optimality and asymptotic optimality of the optimal strategy vector for the limit deterministic system.

Theorem 7.3.2. *Assume* (A7.3.1)–(A7.3.6). *Let* $(u_1^*, u_2^*, \ldots, u_N^*)$ *be the unique optimal strategy vector for* (7.3.1) *and* (7.3.2). *Then* $\{X^{\varepsilon}(\cdot), u_1^*, u_2^*, \ldots, u_N^*\} \to (x(\cdot), u_1^*, u_2^*, \ldots, u_N^*)$ *and the limit satisfies* (7.3.2). *Also*

$$J_k^{\varepsilon}(u_1^*, u_2^*, \ldots, u_N^*) \to J_k(u_1^*, u_2^*, \ldots, u_N^*), \quad k = 1,2,\ldots,N \tag{7.3.8}$$

In addition, let $\widehat{u}_k^{\varepsilon}$, $k = 1,2,\ldots,N$ *be a* δ*-optimal strategy vector for each player* k *with* $X^{\varepsilon}(\cdot)$ *of* (6.6.28). *Then*

$$\liminf_{\varepsilon}\left|\sup_{u_i^{\varepsilon}\in U_i} J_k^{\varepsilon}(u_1^{\varepsilon}, \widehat{u}_2^{\varepsilon}, \ldots, \widehat{u}_N^{\varepsilon}) - J_k^{\varepsilon}(u_1^*, u_2^*, \ldots, u_N^*)\right| \leqslant \delta, \quad k = 1,2,\ldots,N \tag{7.3.9}$$

Proof. By Theorem 7.3.1, the weak convergence is straight forward. By the assumed uniqueness, the limit satisfies (7.3.6). Also, by this weak convergence and the fact that $T < \infty$, by the bounded convergence, we have

$$\lim_{\varepsilon} J_k^{\varepsilon}(u_1^*, u_2^*, \ldots, u_N^*) = J_k(u_1^*, u_2^*, \ldots, u_N^*).$$

Now to show (7.3.9), we repeat the procedure with admissible strategies u_i^{ε}, $i = 1,2,\ldots,N$. The limit $(u_1, u_2, \ldots, u_N)$ might depend on the chosen subsequence. For any convergent subsequence, we obtain,

$$\lim_{\varepsilon=\varepsilon_n\to 0} J_k^{\varepsilon}\left(u_1^{\varepsilon}, u_2^{\varepsilon}, \ldots, u_N^{\delta}\right) = J_k(u_1, u_2, \ldots, u_N).$$

Now by the definition of δ-optimality (7.3.9) follows. □

Note: If $(u_1^*(t), u_2^*(t), \ldots, u_N^*)$ is the optimal strategies for equation (7.3.6), then

$$\{X^{\varepsilon}(t), u_1^*(t), u_2^*(t), \ldots, u_N^*\}_{0 \leqslant t \leqslant T}$$

is the process associated with policies

$$(u_1^{\varepsilon}(t), u_2^{\varepsilon}(t), \ldots, u_N^{\varepsilon}(t)) \equiv (u_1^*(t), u_2^*(t), \ldots, u_N^*(t)).$$

Since, $(u_1^*(t), u_2^*(t), \ldots, u_N^*(t))$ is deterministic, the corresponding $(u_1^{\varepsilon}(t), u_2^{\varepsilon}(t), \ldots, u_N^{\varepsilon}(t))$ is the admissible control for the systems (7.3.3)–(7.3.5).

If for each ε, there is a value for the stochastic game, then the following result shows that they converge to the value of the deterministic game. To prove this we will now introduce a new game through which we will connect the values of stochastic and deterministic games. Define a continuous map ϕ such that

$$\phi : A \to \widetilde{A} \text{ such that if } u = (u_1, u_2, \ldots, u_N) \in \widetilde{A}, \text{ then } \phi(u) = u.$$

Note that an example of one such map is $\phi(u) = Eu$, for $U \in A$. Clearly, if $u \in \widetilde{A}, \phi(u) = Eu = u$. Define the map $L_k : \widetilde{A} \to \widetilde{A}$ by letting $L_k u_{-k} \left(u_{-k} \in \widetilde{A}_{-k} \right)$ to be:

$$J_k(L_k u_{-k}, u_{-k}) = \sup_{u_k \in Ak} J_k(L_k u_{-k}, u_{-k}), u_{-k} \in \widetilde{A}_{-k}.$$

We will now make following simplifying assumption.

(A7.3.7): L_k, $k = 1, 2, \ldots, N$ are continuous, that is for any sequence $\{u_{-k,i}\}$ of admissible controls in $\widetilde{A}_{-k}$, such that, $u_{-k,i} \to u_{-k} \in \widetilde{A}_{-k}, L_k(u_{-k,i}) \to L_i(u_{-k})$, in the $L_2(0,T)$ norm.

Remark 7.3.1. The continuity of L_k may be justified by the fact that the controls u_k are state dependent feedback controls.

Following result states that if the value exists for the stochastic game for each ε, then asymptotically (as $\varepsilon \to 0$) they coincide with the value of the limit deterministic model.

Theorem 7.3.3. *Assume* (A7.3.1)–(A7.3.7) *and that for each ε, value exists for the stochastic game* (7.3.3) *to* (7.3.5). *Also the value for the deterministic game* (7.3.1) *to* (7.3.2) *exists. Then* $\lim_{\varepsilon \to 0} V_k^{\varepsilon} = v_k$.

Proof. For the proof, we introduce the following game which is played as follows. All players except k choose their controls first, say $u^{\varepsilon,\bar{k}}$, which is known to player k. Then player k chooses the control $L_k\left(\phi\left(u_{-k}^{\varepsilon}\right)\right)$. Define $\widetilde{u}_{-k}^{\varepsilon}$ to be optimal strategies for all players except player k with corresponding payoff for player k being $J^{\varepsilon}\left(L_1\phi\left(\widetilde{u}_{-k}^{\varepsilon}\right), \widetilde{u}_{-k}^{\varepsilon}\right)$ By relative compactness,

$$\widetilde{u}_{-k}^{\varepsilon} \longrightarrow \widetilde{u}_{-k}^* \in \widetilde{A}_{-k}.$$

By continuity and by the definition of ϕ and L_k L_k, we have

$$L_1\left(\phi\left(\widetilde{u}^{\varepsilon}_{-k}\right)\right) \to L_1\left(\phi\left(\widetilde{u}^{*}_{-k}\right)\right) = L_1\left(u^{*}_{-k}\right).$$

Now, using (6.6.33), we have,

$$\lim_{\varepsilon\to 0} V_k^{\varepsilon} = v_k. \qquad \square$$

The following result is directly from Theorem 7.3.2 and Theorem 7.3.3.

Theorem 7.3.4. *Assume* (A7.3.19)–(A7.3.7). *Let* $(u_1^*(t), u_2^*(t), \ldots, u_N^*)$, $0 \leqslant t \leqslant T$ *be an optimal deterministic strategy vector for* (7.3.1), (7.3.2), *then* $(u_1^*(t), u_2^*(t), \ldots, u_N^*)$ *is asymptotically optimal for* (7.3.3), (7.3.5) *in the sense that*

$$\lim_{\varepsilon\to 0} |J_k^{\varepsilon}(u_1^*, u_2^*, \ldots, u_N^*) - V_k^{\varepsilon}| = 0 \tag{7.3.10}$$

We can obtain L^2-convergence results and conclusions similar to that of Section 6.6, also refer Ramachandran and Rao [164].

Chapter 8

Some Numerical Methods

8.1 Introduction

In this chapter, we will explain some numerical methods to deal with two person stochastic differential games that are developed by Kushner ([107, 108] and [109]). Almost all of the materials of this chapter are extracted from these three excellent works of Kushner. The Markov chain approximation method is a powerful and more popularly used class of methods introduced in early 1990s by Kushner for the numerical solution of almost all of the standard forms of stochastic strategy problems [106, 109], Kushner and Ramachandran [113], and Kushner and Chamberlain [110]. The idea of the Markov chain approximation method is to first approximate the controlled diffusion dynamics by a suitable Markov chain on a finite state space with a discretization parameter $h > 0$, then approximate the payoff functions. One solves the game problem for the simpler chain model, and then proves that the value functions associated with equilibrium or δ-equilibrium strategies for the chain converge to the value functions associated with equilibrium or δ-equilibrium strategies for the diffusion model, as $\delta \to 0$. This is method is intuitive and it uses approximations which are "physically" close to the original problem. Extensions to approximations for two-person differential games with discounted, finite time, stopping time, and pursuit-evasion games were given in Kushner [108] for reflected diffusion models where the strategies for the two players are separated in the dynamics and payoff rate functions. An extension to two-player stochastic dynamic games with the same systems model, but where the payoff function is ergodic is given in Kushner [107].

For numerical purposes, we will confine the system to a bounded region. If the system state is not a priori confined to a bounded set, then we will bound the state space artificially by adding a reflecting boundary and then experimenting with the bounds. Thus, we assume that the systems model is confined to a state space G that is a convex polyhedron, and it is confined by a reflection on the boundary. In this chapter, we are not concerned with actual development of the algorithms for numerically solving the game for the chain model, only showing convergence of the solutions to the desired values as the discretization parameter goes to zero. The essential conditions for convergence of the numerical approximations are weak-sense existence and uniqueness of the solution to the strategized equations, almost everywhere continuity of the dynamical and payoff rate terms, and a local consistency condition.

8.2 Discounted Payoff Case

Let $w(\cdot)$ be a standard vector-valued Wiener process with respect to the filtration $\{\Im_t, t < \infty\}$, which might depend on the strategies. The *admissible strategies* for the two players are defined by $u_i(\cdot)$, $i = 1,2$, that are $\mathscr{U}_i$-valued, measurable and $\Im_t$-adapted processes. It should be noted that mere admissibility of $u_i(\cdot)$, $i = 1,2$, does not imply that they are acceptable strategies for the game, since the two players will have different information available depending on who "goes first". Nonetheless, for any strategies with the correct information dependencies, there will be a filtration with respect to which $w(\cdot)$ is a standard vector-valued Wiener process, and to which the strategies are adapted. The concept of admissibility will be used in getting approximations and bounds, independent of the strategies. For simplicity of numerical method development, the state process is restricted to the polyhedral region G in this chapter. It should be noted that the set G might not be part of the original problem statement. For solving the game problem numerically, it is usually a necessity. If the bounding set G is imposed for purely numerical purpose, then it must be large enough so that the basic features of the solution in the important region of the state space are not significantly affected. For this purpose, we now reformulate the system discussed in Chapter 3 to following reflected diffusion setup.

The dynamic model for the game process is the reflected stochastic differential equation given by

$$x(t) = x(0) + \sum_{i=1}^{2} \int_0^t b_i\left(x(s), u_i(s)\right) ds + \int_0^t \sigma\left(x(s)\right) dw(s) + z(t) \tag{8.2.1}$$

where $u_i(\cdot)$ is the strategy (payoff) for player i, $i = 1,2$. The process $z(\cdot)$ is due to the boundary reflections, and ensures that $x(t) \in G$. It has the representation

$$z(t) = \sum_i d_i y_i(t) \tag{8.2.2}$$

where $y(0) = 0$, the $y_i(\cdot)$ are continuous, nondecreasing and can increase only at t where $x(t)$ is on the i^{th} face of the boundary ∂G of the set $G \subset \mathbb{R}^n$.

For some filtration $\{\Im_t,\ t < \infty\}$ and standard vector valued $\Im_t$ Wiener process $w(\cdot)$, let each $r_i(\cdot)$, $i = 1,2$, be a probability measure on the Borel sets of $\mathscr{U}_i \times [0,\infty)$ such that $r_i(\mathscr{U}_i \times [0,t]) = t$ and $r_i(A \times [0,t])$ is $\Im_t$-measurable for each Borel set $A \subset \mathscr{U}_i$. Then, as before, $r_i(\cdot)$ is an *admissible relaxed strategy* for player i. For Borel sets $A \subset \mathscr{U}_i$, we will denote $r_i(A \times [0,t]) = r_i(A,t)$. For almost all (ω,t) and each Borel $A \subset \mathscr{U}_i$, one can define the derivative by

$$r_{i,t}(a) = \lim_{\delta \to 0} \frac{r_i(t,A) - r_i(t-\delta,A)}{\delta}.$$

Without loss of generality, we can suppose that the limit exists for all (ω,t). Then for all (ω,t), $r_{i,t}(\cdot)$ is a probability measure on the Borel sets of $\mathscr{U}_i$ and for any bounded Borel set B in $\mathscr{U}_i \times [0,\infty)$,

$$r_i(b) = \int_0^\infty \int_{\mathscr{U}_i} I_{\{(\alpha_i,t)\in B\}} r_{i,t}(d\alpha_i)dt.$$

An ordinary strategy $u_i(\cdot)$ can be represented in terms of the relaxed strategy $r_i(\cdot)$, defined by $r_{i,t}(a) = I_A(u_i(t))$, where $I_A(u_i)$ is unity if $u_i \in A$ and is zero otherwise. The weak topology will be used on the space of admissible relaxed strategies. Define the relaxed strategy $r(\cdot) = (r_1(\cdot) \times r_2(\cdot))$, with derivative $r_t(\cdot) = r_{1,t}(\cdot) \times r_{2,t}(\cdot)$. In this setup, the $r(\cdot)$ is a measure on the Borel sets of $(U_1 \times U_2) \times [0,\infty)$, with marginal's $r_i(\cdot)$, $i = 1,2$. Whenever there is no confusion, we will just write $r(\cdot) = (r_1(\cdot), r_2(\cdot))$. The pair $(w(\cdot), r(\cdot))$ is an *admissible pair* if each of the $r_i(\cdot)$ is admissible with respect to $w(\cdot)$.

In relaxed control terminology, we can rewrite (8.2.1) as

$$x(t) = x(0) + \sum_{i=1}^{2} \int_0^t \int_{\mathscr{U}_i} b_i(x(s),\alpha_i)\, r_{i,s}(d\alpha_i)\, ds + \int_0^t \sigma(x(s))\, dw(s) + z(t). \tag{8.2.3}$$

For $x(0) = x$ and $\beta > 0$, the payoff function is

$$J(x,r_1,r_2) = E\int_0^\infty e^{-\beta t}\left[\sum_{i=1}^{2}\int_{\mathscr{U}_i} k_i(x(s),\alpha_i) r_{i,t}(d\alpha_i)dt + c' dy(t)\right]. \tag{8.2.4}$$

Define $\alpha = (\alpha_1,\alpha_2)$, $u = (u_1,u_2)$,

$$b(x,\alpha) = b_1(x,\alpha_1) + b_2(x,\alpha_2)$$

and

$$k(x,\alpha) = k_1(x,\alpha_1) + k_2(x,\alpha_2).$$

Thus, for simplicity, we assume that both $b(\cdot)$ and $k(\cdot)$ are separable in control variables for every x.

Suppose that $(w(\cdot),r(\cdot))$ is admissible with respect to some filtration $\{\Im_t,\ t<\infty\}$ on a probability space. If there is a probability space on which with a filtration $\{\widetilde{\Im}_t,\ t<\infty\}$ and a $\widetilde{\Im}_t$-adapted triple $(\widetilde{x}(\cdot),\widetilde{w}(\cdot),\widetilde{r}(\cdot))$ where $(\widetilde{w}(\cdot),\widetilde{r}(\cdot))$ is admissible and has the same probability law as $(w(\cdot),r(\cdot))$, and the triple satisfies (8.2.3), then it is said that there is a *weak-sense solution* to (8.2.3) for $(w(\cdot),r(\cdot))$. Suppose that we are given two probability spaces (indexed by $i=1,2$) with filtration $\{\Im_t,\ t<\infty\}$ and on which are defined processes $\left(x^i(\cdot),w^i(\cdot),r^i(\cdot)\right)$, where $w^i(\cdot)$ is a standard vector valued $\Im_t^i$-Wiener process, $\left(w^i(\cdot),r^i(\cdot)\right)$ is an admissible pair, and $\left(x^i(\cdot),w^i(\cdot),r^i(\cdot)\right)$ solves (8.2.3). If equality of the probability laws of $\left(w^i(\cdot),r^i(\cdot)\right)$, $i=1,2$, implies equality of the probability laws of $\left(x^i(\cdot),w^i(\cdot),r^i(\cdot)\right)$, $i=1,2$, then we say that there is a *unique weak sense solution* to (8.2.3) for the admissible pair $\left(w^i(\cdot),r^i(\cdot)\right)$. For a relationship between values corresponding to ordinary and relaxed controls see Theorem 6.2.2, the chattering lemma.

Following are general assumptions, introduced by Kushner [108].

(A8.2.1): Let $G\subset\mathbb{R}^n$ be a bounded convex polyhedron with an interior and a finite number of faces. Let d_i be the direction of reflection to the interior on the i^{th} face, assumed constant for each i. On any edge or corner, the reflection direction can be any nonnegative linear combination of the directions on the adjacent faces. Let $d(x)$ denote the set of reflection directions at $x\in\partial G$. For an arbitrary corner or edge of ∂G, let $\overline{d}_i$ and $\overline{n}_i$ denote the direction of reflection and the interior normal, respectively, on the i^{th} adjacent face. Then there are constants $a_i>0$ (depending on the edge or corner) such that

$$a_i\left\langle\overline{n}_i,\overline{d}_i\right\rangle>\sum_{j:j\neq i}a_j\left|\left\langle\overline{n}_i,\overline{d}_j\right\rangle\right|\quad\text{for all } i. \tag{8.2.5}$$

Note: The condition (8.2.5) implies that the set of reflection directions on any set of intersecting boundary faces are linearly independent. This implies that the representation (8.2.2) is unique.

(A8.2.2): There is a neighborhood $N(\partial G)$ and an extension of $d(\cdot)$ to $\overline{N(\partial G)}$ such that: For each $\varepsilon>0$, there is $\mu>0$ which goes to zero as $\varepsilon\to0$ and such that if $x\in\overline{N(\partial G)}-\partial G$ and distance $(x,\partial G)\leqslant\mu$, then $d(x)$ is in the convex hull of $\{d(v);v\in\partial G,\ \text{distance}(x,v)\leqslant\varepsilon\}$.

(A8.2.3): Assume that $\mathscr{U}_i$, $i=1,2$, are compact subsets of some Euclidean space, and for (8.2.4), $c_i\geqslant0$.

(A8.2.4): The functions $k_i(\cdot)$ and $b_i(\cdot)$ are real-valued (resp. $\mathbb{R}^n$ valued) and continuous on $G\times\mathscr{U}_i$. Let $\sigma(\cdot)$ be a Lipschitz continuous matrix-valued function on G, with n rows and with the number of columns being the dimension of the Wiener process in (8.2.3). The $b_i(\cdot,\alpha_i)$ are Lipschitz continuous, uniformly in α_i.

If we are interested in only weak-sense solutions, condition (A8.2.5) and either (A8.2.6) or (A8.2.7) will replace (A8.2.4).

(A8.2.5): The functions $\sigma(\cdot)$, $b(\cdot)$, $k(\cdot)$ are bounded and measurable. Equation (8.2.4) has a unique weak-sense solution for each admissible pair $(w(\cdot), r(\cdot))$ and each initial condition.

(A8.2.6): The functions $\sigma(\cdot)$, $b(\cdot)$, and $k(\cdot)$ are continuous.

In assumption (A8.2.7), let $(w(\cdot), r(\cdot))$ be an arbitrary admissible pair and $x(\cdot)$ the corresponding solution.

(A8.2.7): There is a Borel set $D_d \subset G$, such that for $x \notin D_d$, $\sigma(\cdot)$, $b(\cdot)$, and $k(\cdot)$ are continuous, and for each $\varepsilon > 0$, there is $t_\varepsilon > 0$ which goes to zero as $\varepsilon \to 0$ and such that for any real T

$$\lim_{\varepsilon \to 0} \sup_{x(0)} \sup_{\text{admis.}\, r(\cdot)} \sup_{t_\varepsilon \leqslant t \leqslant T} P\{x(t) \in N_\varepsilon(D_d)\} = 0,$$

where $N_\varepsilon(D_d)$ is an ε-neighborhood of D_d.

Let $w(\cdot)$ be a standard vector-valued $\Im_t$-Wiener process. Let $\mathscr{U}_i$ denote the set of strategies (ordinary not relaxed) $u_i(\cdot)$ for player i that are admissible with respect to $w(\cdot)$. For computational purposes, we will Discretize and define a class of strategies as follows. For $\Delta > 0$, let $U_i(\Delta) \subset \mathscr{U}_i$ denote the subset of admissible strategies $u_i(\cdot)$ which are constant on the intervals $[k\Delta, k\Delta + \Delta)$, $k = 0, 1, \ldots$ and where $u_i(k\Delta)$ is $\Im_{k\Delta}$-measurable. Thus, δ is the length of time step. Let B be a Borel subset of U_1. Let $L_1(\Delta)$ denote the set of such piecewise constant strategies for player 1 that are represented by functions $Q_{1k}(B; \cdot)$, $k = 0, 1, \ldots$ of the conditional probability type given by

$$Q_{1k}(B; w(s), u(s), s < k\Delta) = P\big\{u_1(k\Delta) \in B \mid w(s), u_2(s),\ s < k\Delta;\ u_1(l\Delta),\ l < k\big\}, \quad (8.2.6)$$

where $Q_{1k}(B; \cdot)$ is a measurable function for each Borel set B.

If a rule for player 1 is given by the form (8.2.6), it will be written as $u_1\ (u_2)$ to emphasize its dependence is suppressed in the notation. Similarly define $L_2(\Delta)$ and the associated rules $u_2\ (u_1)$ for player 2. For relaxed strategies, $r_i(\cdot) \in \mathscr{U}_i$ means that $r_i(\cdot)$ is admissible, and $r_i(\cdot) \in U_i(\Delta)$ means that $r_i(\cdot)$ is admissible, the derivative $r_{i,t}(\cdot)$ is constant on the interval $[k\Delta, k\Delta + \Delta)$, and $r_{i,t}(\cdot)$ is $\Im_{k\Delta}-$measurable. Thus, the difference between $L_i(\Delta)$ and $U_i(\Delta)$ is that in the latter case, the strategy is determined by a conditional probability law such as (8.2.6). But, (A8.2.5) implies that it is the probability law of $(w(\cdot), u_1(\cdot), u_2(\cdot))$ (or, of $(w(\cdot), r_1(\cdot), r_2(\cdot))$ that determines the law of the solution and hence that of the payoff. Thus, we can always suppose that if the strategy for player 1 is determined by the a form such as (8.2.6), then (in relaxed strategy terminology) the law for $(w(\cdot), r_2(\cdot))$ is determined recursively by a conditional probability law, that is,

$$P\big\{\{w(s), r_2(s), k\Delta \leqslant s \leqslant k\Delta + \Delta\} \in \cdot \mid w(s), r_2(s), u_1(s), s < k\Delta\big\}.$$

Now, we are in a position to introduce the upper and lower values corresponding to these just introduced policies. For initial condition $x(0) = x$, define the *upper* and *lower values* for the game as

$$V^{+}(x) = \lim_{\Delta \to 0} \inf_{u_1 \in L_1(\Delta)} \sup_{u_2 \in U_2} J\left(x, u_1\left(u_2\right), u_2\right), \tag{8.2.7}$$

and

$$V^{+}(x) = \lim_{\Delta \to 0} \sup_{u_2 \in L_2(\Delta)} \inf_{u_1 \in U_1} J\left(x, u_1, u_2\left(u_1\right)\right). \tag{8.2.8}$$

The equation (8.2.7) can be interpreted as follows. For fixed $\Delta > 0$, consider the right side of (8.2.7). For each k, at time $k\delta$, player I uses a rule of the form (8.2.6) to decide on the constant action that it will take on $[k\Delta, k\Delta + \Delta)$. That is, player I "goes first". Player 2 can decide on his/her action at $t \in [k\Delta, k\Delta + \Delta)$ at the actual time that it is to be applied. Player 2 selects a strategy simply to be admissible. The operation yields admissible strategy $u(\cdot) = (u_1(\cdot), u_2(\cdot))$. With this strategy pair and under the assumption (A8.2.4) there is a unique solution to (8.2.3). The distribution of the set, (*solution, Wiener process, strategy*), does not depend on the probability space. Thus, the $\sup_{u_2 \in U_2}$ is well defined for each rule for player 1. Because player 1 can make decision more often, as $\Delta \to 0$, the inf sup is monotonically decreasing. The similar observation holds for (8.2.8). If the upper and lower values are equal, the game has a value, $V(x)$ and we say that there exists a *saddle point* for the game.

8.2.1 *The Markov Chain Approximation Method*

Now we will introduce a discrete time, discrete state controlled Markov chain to approximate the continuous time process given by (8.2.3). First, we will explain a computational procedure for control problem with single player. The Markov chain is designed for numerical purpose. Idea of the Markov chain approximation method is to find a controlled Markov chain ξ_n^h and an adaptation of the payoff function, such that the associated game problem is conveniently solvable, and the solution converges to the original game problem as the approximating parameter, $h \to 0$. Here, h will indicate the order of the spacing in the discretization of the state space for the Markov chain. The Markov chain approximation is natural for stochastic control problems, as discussed in Kushner and Dupuis [112]. It allows us to use physical intuition in the design of the algorithm. For a quick introduction to this topic, we refer the paper by Kushner, [106]. The approximating process is a Markov chain indexed by h, and the constraint on the chain is local consistency, where the "local" properties of the chain are close to those of the diffusion that it tries to approximate, for small h.

The Markov chain approximation method consists of two steps.

(i) Determine a finite-state controlled Markov chain (controlled by the strategies of two players) that has a continuous time interpolation that is an "approximation" of the process $x(\cdot)$.

(ii) Solve the optimization problem for the chain and a payoff function that approximates the one used for $x(\cdot)$.

Under a "local consistency" condition, the optimal payoff function $V^h(x)$ for the strategy led approximating chain converges to the optimal payoff function for the original problem. The optimal strategy for the original problem is also approximated. This method is a robust and effective way for solving optimal strategy problems governed by reflected jump-diffusions under very general conditions. An advantage of the approach is that the approximations "stay close" to the physical model and can be adjusted to exploit local features of the problem. Thus, this method involves, first defining an appropriate Markov chain, including obtaining suitable transition probabilities, so that the resulting chain satisfies the local consistency conditions. The optimization step will involve, iteratively solving corresponding Hamilton-Jacobi-Bellman equations (or the dynamic programming equations), such as using Gauss-Seidel numerical procedures. It is well known that, each of these steps presents its own challenges. These challenges will not be the topic of discussion in this Chapter, we refer to Puterman, [155].

To construct the Markov chain approximation, start by defining S_h, a discretization of $\mathbb{R}^n$. This can be done in many ways. For example, S_h might be a regular grid with the distance between points in any coordinate direction being h, or the distance between points in coordinate direction i might be $v_i h$, for some constants v_i. We are only interested the points in G and their immediate neighbors.

Now, define the approximating Markov chain ξ_n^h and its state space, which will be a subset of S_h. For convenience in coding for the reflecting boundary problem, the state space for the chain is usually divided into two parts:

(a) The first part is $G_h = G \cap S_h$, on which the chain approximates the diffusion part of (8.2.3), and

(b) If the chain tries to leave G_h, then it is returned immediately, consistently with the local reflection direction.

Thus, define ∂G_h^+ to be the set of points not in G_h to which the chain might move in one step from some point in G_h. The use of ∂G_h^+ simplifies the analysis and allows us to get a reflection process $z^h(\cdot)$ that is analogous to $z(\cdot)$ of (8.2.3). The set ∂G_h^+ is an approximation to the reflecting boundary. Thus, due to reflection terms in the dynamics of the controlled process, it is convenient to consider a slightly "enlarged" state space, namely $G_h \cup \partial G_h^+$, the points on this set is the only one of interest for the numerical work. This "approximating" reflection process is needed to get the correct form for the limits of the approximating process and for the components of the payoff function that are due to the boundary reflection.

Next, we will define *local consistency* for the controlled diffusion of (8.2.3) at $x \in G_h$. Let $u_{12}^h = (u_{1,n}^h, u_{2,n}^h)$ denote the actual strategies used at step n for approximating the chain ξ_n^h. Let $E_{x,n}^{h,\alpha}$ (respectively, $\mathrm{covar}_{x,n}^{h,\alpha}$) denote the expectation (respectively, the covariance) given all of the data up to step n, when $\xi_n^h = x$, $u_n^h = \alpha$. Then the chain satisfies the following local consistency conditions. There is a function $\Delta t^h(x,\alpha) > 0$ (called an *interpolation interval* that goes to zero as $h \to 0$), such that,

$$E_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] = b(x,\alpha)\Delta t^h(x,\alpha) + o\big(\Delta t^h(x,\alpha)\big),$$

$$\begin{aligned}\mathrm{covar}_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] &= E_{x,n}^{h,\alpha}\left[\left\{\left(\xi_{n+1}^h - x\right) - E_{x,n}^{h,\alpha}\left(\xi_{n+1}^h - x\right)\right\}\right.\\ &\qquad \left.\times\left\{\left(\xi_{n+1}^h - x\right) - E_{x,n}^{h,\alpha}\left(\xi_{n+1}^h - x\right)\right\}'\right]\\ &= a(x)\Delta t^h(x,\alpha) + \circ\left(\Delta t^h(x,\alpha)\right), \quad \text{where } a(x) = \sigma(x)\sigma'(x),\end{aligned}$$

$$\lim_{h\to 0}\sup_{x,\alpha}\Delta t^h(x,\alpha) = 0,$$

and

$$\left\|\xi_{n+1}^h - \xi_n^h\right\| \leqslant K_1 h, \tag{8.2.9}$$

for some real K_1.

With the straight forward methods as discussed by Kushner and Ramachandran [113], $\Delta t^h(\cdot)$ is obtained automatically as a byproduct of getting the transition probabilities and it will be used as an interpolation interval. Thus, in G the conditional mean first two moments of $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$ are very close to those of the "differences" of the solution $x(\cdot)$ of (8.2.3). The interpolation interval $\Delta t^h(x,\alpha)$ can always be selected so that it does not depend on the strategy α or on the state x. The expression (8.2.9) is the essential relationship that we will seek to satisfy in the construction of the approximating chains.

Remark 8.2.1. Note that the chain constructed in (8.2.9) has the "local properties" of the diffusion process (8.2.3) in the sense that

$$E_x(x(\delta)-x) = b(x,\alpha) + \circ(\delta),$$
$$E_x[x(\delta)-x][x(\delta)-x]' = \sigma(x)\sigma'(x) + \circ(\delta).$$

This is what "local consistency" (of the chain with the diffusion) means. The consistency condition (8.2.9) need not hold at all points. For instance, consider a case where the assumption (A8.2.7) holds: Let $k(\cdot)$, $\sigma(\cdot)$ be continuous, and let $b(\cdot)$ have the form $b(x,\alpha) = b_0(x) + \bar{b}(x,\alpha)$, where $b(\cdot)$ is continuous, but $b_0(\cdot)$ is discontinuous at $D_d \subset G$. If the assumption (A8.2.7) holds for D_d, then local consistency there is not needed. The assumption (A8.2.7) would hold if the "noise" $\sigma(x)dw$ "drives" the process away from the set D_d, no matter what the strategy, we refer the reader to Kushner and Dupuis [112].

From points in ∂G_h^+, the transitions of the chain are such that they move to G_h, with the conditional mean direction being a reflection direction at x. More precisely,

$$\lim_{h\to 0} \sup_{x\in\partial G_h^+} \text{distance}(x, G_h) = 0 \tag{8.2.10}$$

and there are $\theta_1 > 0$ and $\theta_2(h) \to 0$ as $h \to 0$, such that, for all $x \in \partial G_h^+$,

$$E_{x,n}^{h,\alpha}[\xi_{n+1}^h - x] \in \{a\gamma : \gamma \in d(x), \theta_2(h) \geqslant a \geqslant \theta_1 h\}, \tag{8.2.11}$$

and

$$\Delta t^h(x,\alpha) = 0 \text{ for } x \in \partial G_h^+.$$

The last line of (8.2.11) says that the reflection from states on ∂G_h^+ is instantaneous. Kushner and Dupuis [112] have given a general discussion of the straightforward methods of obtaining useful approximations in G as well as on the reflecting boundary. These methods continue to hold for the game problem.

The discretization of the payoff can be done in the following way. Define $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$ and $t_n^h = \sum_{t=0}^{n-1} \Delta t_t^h$. One choice of discounted payoff function for approximating chain and initial condition $x = x(0)$ is

$$J^h(x,u^h) = E\sum_{n=0}^{\infty} e^{-\beta t_n^h}\left[k(\xi_n^h, u_n^h)\Delta t_n^h I_{\{\xi_n^h \in G_h\}} + c'[\xi_{n+1}^h - \xi_n^h] I_{\{\xi_n^h \in G_h^+\}}\right]. \tag{8.2.12}$$

Let $p^h(x,y \mid u)$ denote the transition probability of the chain for $u = (u_1,u_2)$, $u_1 \in U_1$, $u_2 \in U_2$. The strategies for the game can be analogously defined as to what was done in (8.2.6). If player i goes first, his/her strategy is defined by a conditional probability law of the type given by

$$P\{u_{i,n}^h \in \cdot \mid \xi_t^h,\ l \leqslant n;\ u_t^h,\ l < n\}.$$

Let $U_i^h(1)$ be the class of such rules. If player i goes last, then its strategy is defined by a conditional probability law of the type

$$P\{u_{i,n}^h \in \cdot \mid \xi_t^h,\ l \leqslant n;\ u_t^h,\ l < n;\ u_{j,n}^h,\ j \neq i\}.$$

Denote the class of such strategies as $U_i^h(2)$. Let $\{\delta\widetilde{w}_n^h,\ n < \infty\}$ be mutually independent random variables and such that $\delta\widetilde{w}_n^h$ is independent of the "past" $\{\xi_l^h,\ l \leqslant n,\ u_l^h,\ l < n\}$. For further flexibility, the conditioning data can be augmented by $\{\delta\widetilde{w}_l^h,\ l \leqslant n\}$ as long as the Markov property

$$P\{\xi_{n+1}^h = \cdot \mid \xi_l^h, u_l^h, l \leqslant n\} = p^h\left(\xi_n^h, \xi_{n+1}^h \mid u_n^h\right)$$

holds. The same notation is used for admissible relaxed strategies. As in Chapter 6, define the upper values for the discretized system, respectively,

$$V^{+,h}(x) = \inf_{u_1 \in U_1^h(1)} \sup_{u_2 \in U_2^h(2)} J^h\left(x, u_1, u_2\right) \tag{8.2.13}$$

and

$$V^{-,h}(x) = \sup_{u_2 \in U_2^h(1)} \inf_{u_1 \in U_1^h(2)} J^h\left(x, u_1, u_2\right). \tag{8.2.14}$$

When interpreting the payoff function and interpolations, keep in mind that $\Delta t^h(x, \alpha) = 0$ for $x \in \partial G_h^+$. Owing to the local consistency, theoretically we can compute payoff $J^h(x, u^h)$ for any admissible strategy u^h using equation (8.2.12). Then, $V^{+,h}(x)$ and $V^{-,h}(x)$ can be found using equations (8.2.13) and (8.2.14). However, that is not very practical. Instead we can solve for $V^{+,h}(x)$ and $V^{-,h}(x)$ using the dynamic programming equations given in equations (8.2.15) and (8.2.16) using iteration methods For $x \in G_h$, the dynamic programming equation for the upper values is (for $\alpha = (\alpha_1, \alpha_2)$) given by

$$V^{+,h}(x) = \min_{\alpha_1 \in U_1} \left\{ \max_{u_2 \in U_2} E_x^\alpha \left[e^{-\beta \delta t^h(x,\alpha)} V^{+,h}(\xi_1^h) + k(x, \alpha) \Delta t^h(x, \alpha) \right] \right\} \tag{8.2.15}$$

and for $x \in \partial G_h^+$ it becomes

$$V^{+,h}(x) = E_x \left[V^{+,h}(\xi_1^h) + c'(\xi_1^h - x) \right]. \tag{8.2.16}$$

Here E_x^α denotes the expectation given initial state x (the reflection direction is not led by any strategy). The equations are analogous for lower value. Owing to the contraction implied by the discounting, there is unique solution to equations (8.2.15) and (8.2.16). It is possible that the transition probabilities could be constructed so that $\Delta t^h(\cdot)$ does not depend on α and we have the separated form given by,

$$p^h\left(x, y \mid \alpha\right) = p_1\left(x, y \mid \alpha\right) + p_2\left(x, y \mid \alpha\right).$$

Such a form is useful for establishing the existence of value for the chain, even though it is not necessary for the convergence of the numerical method, as explained by Kushner and Chamberlain [110, 111]. The equation (8.2.15) can be rewritten to reflect the transition probability as

$$V^{+,h}(x) = \min_{\alpha_1 \in U_1} \left\{ \max_{u_2 \in U_2} \sum_y \left[e^{-\beta \delta t^h(x,\alpha)} p^h(x,y \mid \alpha) V^{+,h}(y) + k(x,\alpha) \Delta t^h(x,\alpha) \right] \right\}.$$

Similar representation can be written for equation (8.2.16).

Define the positive and negative part of a real number by: $l^+ = \max[l,0]$, $l^- = \max[-l,0]$. In one dimension case, one of the possible ways to obtain the transition probabilities is the following,

$$p^h(x, x+h \mid \alpha) = \frac{\sigma^2(x)/2 + hb^+(x,\alpha)}{\sigma^2(x) + h|b(x,\alpha)|},$$
$$p^h(x, x-h \mid \alpha) = \frac{\sigma^2(x)/2 + hb^-(x,\alpha)}{\sigma^2(x) + h|b(x,\alpha)|},$$

and

$$\Delta t^h(x,\alpha) = \frac{h^2}{\sigma^2(x) + h|b(x,\alpha)|}.$$

For $y \neq x \pm h$, set $p^h(x,y \mid \alpha) = 0$. For the derivation of these transition probabilities and the higher dimensional versions, including such a system is locally consistent, we refer to Kushner and Dupuis [112].

We can rewrite (8.2.15) in the iterative form as follows. For any initial value $\{V_0^{+,h}(x),\ x \in G_h\}$, the sequence

$$V_{n+1}^{+,h}(x) = \min_{\alpha_1 \in U_1} \left\{ \max_{u_2 \in U_2} E_x^\alpha \left[e^{-\beta \Delta t^h(x,\alpha)} V_n^{+,h}(\xi_1^h) + k(x,\alpha) \Delta t^h(x,\alpha) \right] \right\}$$

and for $x \in \partial G_h^+$, we can write

$$V_{n+1}^{+,h}(x) = E_x \left[V_n^{+,h}(\xi_1^h) + c'(\xi_1^h - x) \right]$$

converges to $V^{+,h}(x)$, the unique solution of equation (8.2.15) as $n \to \infty$. Analogously, for any initial value $\{V_0^{-,h}(x), x \in G_h\}$, the sequence

$$V_{n+1}^{-,h}(x) = \max_{u_2 \in U_2} \left\{ \min_{\alpha_1 \in U_1} E_x^\alpha \left[e^{-\beta \delta t^h(x,\alpha)} V_n^{-,h}(\xi_1^h) + k(x,\alpha) \Delta t^h(x,\alpha) \right] \right\}$$

and for $x \in \partial G_h^+$, we have,

$$V_{n+1}^{-,h}(x) = E_x \left[V_n^{-,h}(\xi_1^h) + c'(\xi_1^h - x) \right],$$

converges to $V^{-,h}(x)$, the unique solution of equation (8.2.16) as $n \to \infty$. The computation of the discount factor $e^{-\beta \delta t^h(x,\alpha)}$ can be expensive. To simplify, we could use its first approximation $[1-\beta\delta t^h(x,\alpha)]$. There are many methods available for computing the $V^{\pm,h}$ such as Gauss-Seidel method, we refer to Puterman [155], among others.

If we are interested in obtaining optimal policies, it is possible to use the so called policy iterations such as, for $x \in G_h$ setting $u^h_{i,0}(x) = 0$, $i = 1,2$ and finding $u^h_{n+1}(x)$ through

$$u^h_{n+1}(x) = \underset{u_1 \in U_1}{\arg h} \underset{u_2 \in U_2}{\arg\max} \left\{ \sum_y \left[e^{-\beta\Delta t^h(x,\alpha)} p^h(x,y \mid \alpha) V^{+,h}_n(y) + k(x,\alpha)\Delta t^h(x,\alpha) \right] \right\}.$$

The convergence of the numerical scheme explained above will be given in Theorem 8.2.10. Due to this convergence, it is easy to give a stopping rule for the numerical scheme. For the rest of this section, we will deal with the convergence issues.

8.2.2 *Continuous Time Interpolations*

The chain ξ^h_n is defined in the discrete time, but $x(\cdot)$ is defined in the continuous time. It is important to observe that we only need the Markov chain for the numerical computations. However, for the proofs of convergence, the chain must be interpolated into a continuous time process which approximates $x(\cdot)$. This can be done similar to the discrete dynamics cases considered in Chapter 6. For completeness sake, we will now explain the necessary interpolations in the rest of this section.

The interpolation intervals are suggested by the $\Delta t^h(\cdot)$ in equation (8.2.9). There are two useful (and asymptotically equivalent) interpolations. The first interpolation $\xi^h(\cdot)$, is defined by (t^h_n is defined above the equation (8.2.12), that is,

$$\xi^h(t) = x(0) + \sum_{t^h_{i+1} \leqslant t} \left[\xi^h_{t+1} - \xi^h_t \right].$$

Given the current state x and strategy pair α, the next interpolation interval for $\xi^h(\cdot)$ is just $\Delta t^h(x,\alpha)$. Thus, $\xi^h(\cdot)$ is a semi-Markov process.

For simplification of proof, define an alternative and Markovian interpolation, $\psi^h(\cdot)$. Let $\{\Delta\tau^h_n,\ n < \infty\}$ be conditionally mutually independent and "exponential" random variables such that

$$P^{h,\alpha}_{x,n}\{\Delta\tau^h_n \geqslant t\} = e^{-t/\Delta t^h(x,\alpha)}.$$

Note that $\Delta\tau^h_n = 0$ if ξ^h_n is on the reflecting boundary ∂G^+_h. Define $\tau^h_0 = 0$, and for $n > 0$, set $\tau^h_n = \sum_{i=0}^{n-1} \Delta\tau^h_i$. The τ^h_n will be jump times of $\psi^h(\cdot)$. Now, define $\psi^h(\cdot)$ and the interpolated reflected processes by

$$\psi^h(t) = x(0) + \sum_{\tau_{i+1}^h \leqslant t} \left[\xi_{i+1}^h - \xi_i^h\right],$$

$$Z^h(t) = \sum_{\tau_{i+1}^h \leqslant t} \left[\xi_{i+1}^h - \xi_i^h\right] I_{\{\xi_i^h \in \partial G_h^+\}},$$

and

$$z^h(t) = \sum_{\tau_{i+1}^h \leqslant t} E_i^h \left[\xi_{i+1}^h - \xi_i^h\right] I_{\{\xi_i^h \in \partial G_h^+\}}.$$

Thus, $\psi^h(t) \in G_h$.

Define $\tilde{z}^h(\cdot)$ by $Z^h(t) = z^h(t) + \tilde{z}^h(t)$. The first part is composed of the "conditional mean" $E_i^h[\xi_{i+1}^h - \xi_i^h] I_{\{\xi_i^h \in \partial G_h^+\}}$, and the second part is composed of the perturbations about these conditional means. The process $z^h(\cdot)$ is a reflection term of the classical type. Both components can change only at t where $\psi^h(t)$ can leave G_h. Let $Z^h(t) - Z^h(t-) \neq 0$, with $\psi^h(t-) = x \in G_h$. Then by equation (8.2.11), $z^h(t) - z^h(t-)$ points in a direction of $d(N_h(x))$, where $N_h(x)$ is a neighborhood with radius that goes to zero as $h \to 0$. The process $\tilde{z}^h(\cdot)$ is the "error" due to the centering of the increments of the reflection term about their conditional means and has bounded (uniformly in x, h) by second moments and it converges to zero, as will be seen in Theorem 8.2.1. By assumptions (A8.2.1), (A8.2.2), and the local consistency condition (8.2.11), we can write (modulo an asymptotically negligible term)

$$z^h(t) = \sum_i d_i y_i^h(t),$$

where $y_i^h(0) = 0$, and $y_i^h(\cdot)$ is nondecreasing and can increase only when $\psi^h(t)$ is arbitrarily close (as $h \to 0$) to the i^{th} face of ∂G.

Define the continuous time interpolations $u_i^h(\cdot)$ of the strategies analogously. Let $r_i^h(\cdot)$ denote the relaxed strategy representation of $u_i^h(\cdot)$. The process $\psi^h(\cdot)$ is a continuous time Markov chain. When the state is x and strategy pair is α, the jump rate out of $x \in G_h$ is $[1/\Delta t^h(x,\alpha)]$. So the conditional mean interpolation interval is $\Delta t^h(x,\alpha)$; that is,

$$E_{x,n}^{h,\alpha}\left[\tau_{n+1}^h - \tau_n^h\right] = \Delta t^h(x,\alpha).$$

The payoff criterion (8.2.12) in a relaxed strategy terminology can be written as (modlulo an asymptotically negligible error), $x(0) = x$, and $r_i^h(\cdot)$ is the relaxed strategy representation of $u_i^h(\cdot)$, that is,

$$J^h(x,r^h) = E\int_0^\infty e^{-\beta t}\left[\sum_{i=1}^2 \int_{\mathscr{U}_i} k_i(\psi^h(s),\alpha_i) r_{i,t}^h(d\alpha_i)dt + c'dy^h(t)\right]. \tag{8.2.17}$$

In the numerical computations, the strategies are ordinary and not relaxed, but it will be convenient to use the relaxed strategy terminology when taking limits. From the proof of Theorem 8.2.10, there is $\rho^h \to 0$ as $h \to 0$, such that

$$V^{+,h}(x) \leqslant V^{-,h}(x) + \rho^h. \tag{8.2.18}$$

This implies that either the upper or lower numerical game gives an approximation to the original game.

The process $\psi^h(\cdot)$ has a representation which makes it appear close to (8.2.4), and which is useful in convergence proofs. Let $\xi_0^h = x$. If $a(\cdot)$ is not uniformly positive definite, then augment the probability space by adding a standard vector-valued Wiener process $\widetilde{w}(\cdot)$ where for each n, $\delta \widetilde{w}_{n+1}^h = \widetilde{w}\left(\tau_n^h + .\right) - \widetilde{w}(\tau_n^h)$ is independent of the "past" $\{\xi^h(s), u^h(s), \widetilde{w}(s),\ s \leqslant \tau_n^h\}$. Then, we can write

$$\psi^h(t) = x + \int_0^t b\left(\psi^h(s), u^h(s)\right) ds + \int_0^t \sigma\left(\psi^h(s)\right) dw^h(s) + Z^h(s) + \varepsilon^h(s), \tag{8.2.19}$$

where $\psi^h(t) \in G$. The process $\varepsilon^h(\cdot)$ is due to the $o(\cdot)$ term in (8.2.9) and is asymptotically unimportant in that, for any T, $\lim_h \sup_{x,x^h} \sup_{s\leqslant T} E|\varepsilon^h(s)|^2 = 0$. The process $w^h(\cdot)$ is a martingale with respect to the filtration induced by $\left(\psi^h(\cdot), u^h(\cdot), w^h(\cdot)\right)$, and converges weakly to a standard (vector-valued) Wiener process. The $w^h(t)$ is obtained from $\{\xi^h(s), \widetilde{w}(s),\ s \leqslant t\}$. All of the processes in equation (8.2.19) are constant on the intervals $[\tau_n^h, \tau_{n+1}^h)$.

Let $\left|z^h\right|(t)$ denote the variation process $z^h(\cdot)$ on the time interval $[0,T]$. We have the following result from Kushner and Dupuis [112].

Theorem 8.2.1. *Assume* (A8.2.1), (A8.2.2)*, the local consistency conditions, and let $b(\cdot)$ and $\sigma(\cdot)$ be bounded and measurable. Then for any $T < \infty$, there are $K_2 < \infty$ and δ_h, where $\delta_h \to 0$ as $h \to 0$, and which do not depend on the strategies or initial conditions, such that,*

$$E\left|z^h\right|(T) \leqslant K_2, \tag{8.2.20}$$

and

$$E \sup_{s\leqslant T} \left|\widetilde{z}^h(s)\right|^2 = \delta_h E\left|z^h\right|(T). \tag{8.2.21}$$

8.2.3 *Bounds and Approximations*

Assume (A8.2.1) and (A8.2.2) and let the components of the $\mathbb{R}^r$-valued function $\psi(\cdot)$ be right continuous and have left hand limits. Consider the equation $x(t) = \psi(t) + \bar{z}(t)$, $x(t) \in G$. Then we say that $\bar{x}(\cdot)$ solves the *Skorohod problem* if the following holds: The components of $\bar{z}(\cdot)$ are right continuous with $\bar{z}(0) = 0$, and $\bar{z}(\cdot)$ is constant on the time intervals where $\bar{x}(t)$ is in the interior of G. The variation $|\bar{z}|(t)$ of $\bar{z}(\cdot)$ on each $[0,t]$ is finite. There is measurable $\gamma(\cdot)$ with values $\gamma(t) \in d(\bar{x}(t))$, the set of reflection directions at$\bar{x}(t)$, such that $\bar{z}(t) = \int_0^t \gamma(s) d|\bar{z}|(s)$. Thus, $\bar{z}(\cdot)$ can only change when $\bar{x}(t)$ is on the boundary of G, and then its "increment" is in a reflection direction at $\bar{x}(t)$.

Theorem 8.2.2. *Assume* (A8.2.1) *and* (A8.2.2). *Let* $\psi(\cdot) \in D(\mathbb{R}^r;[0,\infty))$, *and consider the Skorohod problem* $\bar{x}(t) = \psi(t) + \bar{z}(t)$, $x(t) \in G$. *Then, there is a unique solution* $(\bar{x}(\cdot), \bar{z}(\cdot))$ *in* $D(\mathbb{R}^{2r};[0,\infty))$. *There is* $K < \infty$ *depending only on the* $\{d_i\}$ *such that*

$$|\bar{x}(t)| + |\bar{z}(t)| \leqslant K \sup_{s \leqslant t} |\psi(s)|, \tag{8.2.22}$$

and for any $\psi^i(\cdot) \in D(\mathbb{R}^{2r};[0,\infty))$, $i = 1,2$, *and corresponding solutions* $(\bar{x}^i(\cdot), \bar{z}^i(\cdot))$, *and*

$$|\bar{x}_1(t) - \bar{x}_2(t)| + |\bar{z}_1(t) - \bar{z}_2(t)| \leqslant K \sup_{s \leqslant t} |\psi_1(s) - \psi_2(s)|. \tag{8.2.23}$$

Consider (8.2.4) *where* $b(\cdot)$ *and* $\sigma(\cdot)$ *are bounded and measurable, and use the representation* (8.2.3) *for the reflection process* $z(\cdot)$. *Then for any* $T < \infty$ *there is a constant* K_1 *which does not depend on the initial condition or strategies and such that*

$$\sup_{x \in G} E|y(x)|^2 \leqslant K_1. \tag{8.2.24}$$

Suppose that the assumption (A8.2.4) holds. Then the bound (8.2.22) and Lipschitz condition (8.2.23) ensures unique strong sense solution to the stochastic differential equation (8.2.2) for any admissible strategies.

8.2.4 *Approximations under the condition* (A8.2.4)

For each admissible relaxed strategy $r(\cdot)$, let $r^\varepsilon(\cdot)$ be admissible relaxed strategies with respect to the same filtration and that satisfy

$$\lim_{\varepsilon \to 0} \sup_{r_i \in \mathscr{U}_i} E \sup_{t \leqslant T} \left| \int_0^t \int_{\mathscr{U}_i} \phi_i(\alpha_i)[r_{i,s}(d\alpha_i) - r^\varepsilon_{i,s}(d\alpha_i)] ds \right| = 0, \quad i = 1,2, \tag{8.2.25}$$

for each bounded and continuous real-valued nonrandom function $\phi_i(\cdot)$ and each $T < \infty$. For the future use, note that if equation (8.2.25) holds then it also holds for functions $\phi_i(\cdot)$ of (t, α_i) that are continuous except when t takes some value in the finite set $\{t_i\}$. Let $x(\cdot)$, and $x^\varepsilon(\cdot)$ denote the solutions to equation (8.2.4) corresponding to $r(\cdot)$ and $r^\varepsilon(\cdot)$, respectively, with the same Wiener process is used. In particular,

$$x^\varepsilon(t) = x(0) + \int_0^t \int_{U_1 \times U_2} b(x^\varepsilon(s), \alpha) r^\varepsilon_s(d\alpha) ds + \int_0^t \sigma(x^\varepsilon(s)) dw(s) + z^\varepsilon(t). \tag{8.2.26}$$

Define

$$\rho^{\varepsilon}(t) = \int_0^t \int_{U_1 \times U_2} b(x(s),\alpha)[r_s(d\alpha) - r_s^{\varepsilon}(d\alpha)]ds.$$

The process $x(\cdot)$, $x^{\varepsilon}(\cdot)$ and $\rho^{\varepsilon}(\cdot)$ depend on $r(\cdot)$, but this dependence is suppressed in the notation. The next result shows that the set $\{x(\cdot)\}$ over all admissible strategies is equi-continuous in probability in the sense that (8.2.27) holds, and that the payoffs corresponding to $r(\cdot)$ and $r^{\varepsilon}(\cdot)$ are arbitrarily close for small ε, uniformly in $r(\cdot)$.

Theorem 8.2.3. *Assume* (A8.2.1) *and* (A8.2.2) *and let* $b(\cdot), \sigma(\cdot)$ *be bounded and measurable. Then for each real* $\lambda > 0$,

$$\lim_{\Delta \to 0} \sup_{x(0)} \sup_{t} \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} P\left\{ \sup_{s \leqslant \Delta} |x(t+s) - x(t)| \geqslant \lambda \right\} = 0. \tag{8.2.27}$$

Additionally assume (A8.2.3), (A8.2.4), *also let* $(r(\cdot), r^{\varepsilon}(\cdot))$ *satisfy* (8.2.25) *for each bounded and continuous* $\phi_i(\cdot)$, $i = 1,2$, *and* $T < \infty$. *Define* $\Delta^{\varepsilon}(t) = \sup_{s \leqslant t} |x(s) - x^{\varepsilon}(t)|^2$. *Then for each t, we have*

$$\lim_{\varepsilon \to 0} \sup_{x(0)} \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} E \left| \sup_{s \leqslant t} \rho^{\varepsilon}(s) \right|^2 = 0, \tag{8.2.28}$$

$$\lim_{\varepsilon \to 0} \sup_{x(0)} \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} \left[E\Delta^{\varepsilon}(t) + E \sup_{s \leqslant t} |z(s) - z^{\varepsilon}(s)|^2 \right] = 0, \tag{8.2.29}$$

and

$$\lim_{\varepsilon \to 0} \sup_{x(0)} \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} |J(x,r) - J(x,r^{\varepsilon})| = 0. \tag{8.2.30}$$

Proof. Define $\psi(\cdot)$ by

$$\psi(t) = \int_0^t \int_{U_1 \times U_2} b(x(s),\alpha) r_s(d\alpha) ds + \int_0^t \sigma(x(s)) dw(s).$$

Then, we can write

$$x(t+\delta) - x(t) = [\psi(t+\delta) - \psi(t)] + [z(t+\delta) - z(t)].$$

By Theorem 8.2.2 there is a $K < \infty$, such that,

$$\sup_{s \leqslant \delta} |x(t+s) - x(t)| + \sup_{s \leqslant \delta} |z(t+s) - z(t)| \leqslant K \sup_{s \leqslant \delta} |\psi(t+s) - \psi(t)|.$$

Now using standard estimates for stochastic differential equations to evaluate the fourth moments of the right side of the last inequality yields, for some $K_1 < \infty$,

$$\lim_{\varepsilon \to 0} \sup_{x(0)} \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} E \sup_{s \leqslant \delta} |x(t+s) - x(t)|^4 \leqslant K_1 \delta^2, \tag{8.2.31}$$

which implies Kolmogorov's criterion for equi-continuity in probability, which is equation (8.2.27). Thus, we can write

$$x(t)-x^{\varepsilon}(t)=\int_0^t\int_{U_1\times U_2}[b(x(s),\alpha)-b(x^{\varepsilon}(s),\alpha)]r_s(d\alpha)ds+\rho^{\varepsilon}(t)$$
$$+\int_0^t[\sigma(x(s))-\sigma(x^{\varepsilon}(s))]dw(s)+z(t)-z^{\varepsilon}(t).$$

Then the Lipschitz condition, (8.2.23), together with standard estimates for stochastic differential equations, imply that there is a constant K not depending on $(r(\cdot),r^{\varepsilon}(\cdot))$ or the initial condition $x(0)$ and such that

$$E\Delta^{\varepsilon}(t)\leqslant K\left[E\sup_{s\leqslant t}|\rho^{\varepsilon}(s)|^2+(t+1)\int_0^t E\Delta^{\varepsilon}(s)ds+E\sup_{s\leqslant t}|z(s)-z^{\varepsilon}(s)|^2\right],$$

and

$$E\sup_{s\leqslant t}|z(s)-z^{\varepsilon}(s)|^2\leqslant K\left[E\sup_{s\leqslant t}|\rho^{\varepsilon}(s)|^2+(t+1)\int_0^t E\Delta^{\varepsilon}(s)ds\right]. \tag{8.2.32}$$

Suppose that in the definition of $\rho^{\varepsilon}(\cdot)$, the function $b(x(t),\alpha)$ was replaced by a bounded nonrandom function $\phi(t,\alpha)$ which is continuous except when t takes values in some finite set $\{t_i\}$. Then (8.2.28) and (8.2.29) would hold by equation (8.2.25) and use of Gronwall's inequality on the first line of (7.2.32), after the second line is substituted in to eliminate $z(\cdot)-z^{\varepsilon}(\cdot)$. The equi-continuity in probability (8.2.27) and the boundedness and continuity of $b(\cdot)$ imply that $b(x(t),\alpha)$ can be approximated arbitrarily well by replacing $x(t)$ by $xk\mu)$ for $t\in[k\mu,k\mu+\mu)$, $k=0,1,\ldots$, where μ can be chosen independently of $r(\cdot)$. Following this approximation and using equation (8.2.25) implies equations (8.2.28) and (8.2.29).

Now consider equation (8.2.30). By equations (8.2.28), (8.2.29), and the discounting, the parts of $J(x,r^{\varepsilon})$ that involve $k(\cdot)$ converges to corresponding parts of $J(x,r)$. As noted below equation (8.2.3), the linear independence of the reflection directions on any set of intersecting boundary faces which is implied (8.2.1) implies that $z(\cdot)$ uniquely determines $y(\cdot)$ with probability one. Thus, $y^{\varepsilon}(\cdot)$ converges to $y(\cdot)$ with probability one. This convergence, the uniform integrability of the set $\left\{|y^{\varepsilon}(t+1)-y^{\varepsilon}(t)|;t<\infty,\text{ for all } r(\cdot),\varepsilon>0\right\}$ (which is implied by equation (8.2.24) and the compactness of G), and the discounting, imply that the component of $J(x,r^{\varepsilon})$ involving $y^{\varepsilon}(\cdot)$ converges to the component of $J(x(0),r)$ involving $y(\cdot)$. □

The next result uses only weak-sense solutions and does not require the Lipschitz condition (A8.2.4). For the proof, we refer the reader to Kushner [108].

Theorem 8.2.4. *Assume* (A8.2.1)–(A8.2.3), (A8.2.5) *and* (A8.2.6). *Let* $r(\cdot)$ *and* $r^{\varepsilon}(\cdot)$, $\varepsilon>0$, *be admissible with respect to some Wiener process* $w^r(\cdot)$ *and satisfy* (8.2.25). *For each*

$\varepsilon > 0$, *there is a probability space with an admissible pair* $(\widetilde{w}^{r,\varepsilon}(\cdot), \widetilde{r}^{\varepsilon}(\cdot))$ *which has the same probability law as* $(w^r(\cdot), r^{\varepsilon}(\cdot))$ *and on which is defined a solution* $(\widetilde{x}^{r,\varepsilon}(\cdot), \widetilde{y}^{r,\varepsilon}(\cdot))$ *to* (8.2.4). *Let* $x^r(\cdot)$ *denote the solution to* (8.2.4), *corresponding to* $(w^r(\cdot), r(\cdot))$, *and let* $z^r(\cdot) = \sum_i d_i y_i^r(\cdot)$ *denote the associated reflection process. Let* $F(\cdot)$ *be a bounded and continuous real-valued function on the path space of canonical set* $(x(\cdot), y(\cdot), r(\cdot))$. *Then the approximation of the solutions by using* $r^{\varepsilon}(\cdot)$ *is uniform in that*

$$\lim_{\varepsilon \to 0} \sup_{x(0)} \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} \left| EF\left(\widetilde{x}^{r,\varepsilon}(\cdot), \widetilde{y}^{r,\varepsilon}(\cdot), \widetilde{r}^{\varepsilon}(\cdot)\right) - EF\left(\widetilde{x}^{r}(\cdot), \widetilde{y}^{r}(\cdot), \widetilde{r}(\cdot)\right) \right| = 0. \qquad (8.2.33)$$

Now, let $F(\cdot)$ *be only continuous with probability one with respect to the measure of any solution set* $(x(\cdot), y(\cdot), r(\cdot))$. *Then, if* $(x^n(\cdot), y^n(\cdot), r^n(\cdot))$ *converges weakly to* $F(x(\cdot), y(\cdot), r(\cdot))$. *Also* (8.2.33) *continues to hold.*

8.2.5 *Finite-Valued and Piecewise Constant Approximations* $r^{\varepsilon}(\cdot)$ *in* (8.2.25)

Now we will discuss some approximations of subsequent interest. They are piecewise constant and finite valued ordinary strategies. Consider the following discretization of the U_i. Given $\mu > 0$, partition $\mathscr{U}_i$ into a finite number of disjoint subsets C_i^l, $l \leqslant p_i$, each with diameter no greater than $\mu/2$. Choose a point $\alpha_i^l \in C_i^l$. Henceforth, let p_i be some given function of μ.

Now, given admissible strategies $(r_1(\cdot), r_2(\cdot))$, define the approximating admissible relaxed strategy $r_i^{\mu}(\cdot)$ on the strategy value space $\{\alpha_i^l, l \leqslant p_i\}$ by its derivative as $r_{i,t}^{\mu}(\alpha_i^l) = r_{i,t}(C_i^l)$. Denote the set of such strategies over all $\{C_i^l, \alpha_i^l,\ l \leqslant p_i\}$ by $U_i(\mu)$. Let $U_i(\mu, \delta)$ denote the subset of $U_i(\mu)$ that are ordinary strategies and constant on the intervals $[l\delta, l\delta + \delta)$, $l = 0, 1, \ldots$. Thus, we state without proof the following useful result.

Theorem 8.2.5. *Assume* (A8.2.1)–(A8.2.3), (A8.2.5), (A8.2.6), *and the above approximation of* $r_i(\cdot)$ *by* $r_i^{\mu}(\cdot) \in U_i(\mu)$, $i = 1, 2$. *Then* (8.2.25) *and* (8.2.30) *hold for* μ *replacing* ε, *no matter what the* $\{C_i^l,\ \alpha_i^l\}$. *The same result holds if we approximate only one of the* $r_i(\cdot)$.

8.2.6 *Finite-Valued, Piecewise-Constant and "Delayed" Approximations*

Let $r_i^{\mu}(\cdot) \in U_i(\mu)$, where the strategy-space values are $\{\alpha_i^l, l \leqslant p_i\}$. Let $\Delta > 0$. Define the "backward" differences

$$\Delta_{i,k}^l = r_i^{\mu}\left(\alpha_i^l, k\Delta\right) - r_i^{\mu}\left(\alpha_i^l, k\Delta - \Delta\right), \quad l \leqslant p_i,\ k = 1, 2, \ldots.$$

Define the piecewise constant ordinary strategies by $u_i^{\mu,\Delta}(\cdot) \in U_i(\mu,\Delta)$ on the interval $[k\Delta, k\Delta+\Delta)$ by

$$u_i^{\mu,\Delta}(t) = \alpha_i^t \text{ for } t \in \left[k\Delta + \sum_{\upsilon=1}^{l-1} \Delta_{i,k}^{\upsilon}, \quad k\Delta + \sum_{\upsilon=1}^{l} \Delta_{i,k}^{\upsilon}\right]. \tag{8.2.34}$$

Note that on the interval $[k\Delta, k\Delta+\Delta)$, $u^{\mu,\Delta}(\cdot)$ takes the value α_i^l on a time interval of length $\Delta_{i,k}^l$. Note also that the $u_i^{\mu,\Delta}(\cdot)$ are "delayed" in that the values of $r_i(\cdot)$ on $[k\Delta-\Delta, k\Delta)$ determine the values of $u^{\mu,\Delta}(\cdot)$ on $[k\Delta, k\Delta+\Delta)$. Thus, $u_i^{\mu,\Delta}(\cdot)$ is $\Im_{k\Delta}$-measurable. This delay will play an important role in the next two sections. Let $r_i^{\mu,\Delta}(\cdot)$ denote the relaxed strategy representation of $u_i^{\mu,\Delta}(\cdot)$.

The intervals $\Delta_{i,k}^l$ in (8.2.34) are just real numbers. For later use, it is important to have them to be some multiple of any $\delta > 0$, where Δ/δ is an integer. Consider the following method of performing this process. Divide $[k\Delta, k\Delta+\Delta)$ into Δ/δ subintervals of length δ each. To each value α_i^l first assign $[\Delta_{i,k}^l/\delta]$ (the integer part) subintervals of length δ. Then assign the remaining unassigned subintervals to the values α_i^l at random with probabilities proportional to the residual (unassigned) lengths $\Delta_{i,k}^l - [\Delta_{i,k}^l/\delta]\delta$, $i \leqslant p_i$. Call the resulting strategy $u_i^{\mu,\delta,\Delta}(\cdot)$, with relaxed strategy representation $r_i^{\mu,\delta,\Delta}(\cdot)$. Let $U_i(\mu,\delta,\Delta)$ denote the set of such strategies. If $u_i^{\mu,\delta,\Delta}(\cdot)$ is obtained from $r_i(\cdot)$ in this way, then we will henceforth write it as $u_i^{\mu,\delta,\Delta}(\cdot \mid r_i)$ to emphasize that fact. Similarly, if $u_i^{\mu,\Delta}(\cdot)$ is obtained from $r_i(\cdot)$, then it will be written as $u_i^{\mu,\Delta}(\cdot \mid r_i)$. Let $r_{i,t}^{\mu,\Delta}(\cdot \mid r_i)$ denote the time derivative of $r_i^{\mu,\Delta}(\cdot \mid r_i)$. As stated in the next theorem, for fixed μ and δ, $u_i^{\mu,\delta,\Delta}(\cdot \mid r_i)$ gives a good approximation $u_i^{\mu,\Delta}(\cdot \mid r_i)$ uniformly in $r_i(\cdot)$ and $\{\alpha_i^l\}$ in that (8.2.36) holds in the sense that for each $\mu > 0$, $\Delta > 0$, and bounded and continuous $\phi_i(\cdot)$, such that,

$$\lim_{\delta\to 0} \sup_{r_i \in \mathscr{U}_i} E \sup_{t \leqslant T} \left| \int_0^t \int_{\mathscr{U}_i} \phi_i(\alpha_i) \left[r_{i,s}^{\mu,\Delta}(d\alpha_i \mid r_i) - r_{i,s}^{\mu,\delta,\Delta}(d\alpha_i \mid r_i)\, ds \right| = 0, \quad i = 1,2. \tag{8.2.35}$$

This leads to the following interesting result.

Theorem 8.2.6. *Assume* (A8.2.1)–(A8.2.3), (A8.2.5) *and* (A8.2.6). *For $r_i(\cdot) \in \mathscr{U}_i$, let $r_i^{\mu,\Delta}(\cdot \mid r_i) \in U_i(\mu,\Delta)$ and $r_i^{\mu,\delta,\Delta}(\cdot \mid r_i) \in U_i(\mu,\delta,\Delta)$. Then* (8.2.25) *holds for $r_i^{\mu,\Delta}(\cdot \mid r_i)$ and (μ,Δ) replacing $r_i^{\varepsilon}(\cdot)$ and ε, respectively. Also,* (8.2.35) *holds and*

$$\lim_{\Delta\to 0} \lim_{\delta\to 0} \sup_x \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} \left| J(x, r_1, r_2) - J\left(x, r_1, u_2^{\mu,\delta,\Delta}(\cdot \mid r_2)\right) \right| = 0. \tag{8.2.36}$$

For each $\varepsilon > 0$, there are $\mu_\varepsilon > 0$ and $\delta_\varepsilon > 0$, such that, for $\mu \leqslant \mu_\varepsilon$ and $\delta \leqslant \delta_\varepsilon$ and $r_i(\cdot) \in U_i$, $i = 1,2$, there are $u_i^{\mu,\delta} \in U_i(\mu,\delta)$, such that, (8.4.4) holds for $u_i^{\mu,\delta}(\cdot)$ and (μ,δ) replacing $r_i^{\varepsilon}(\cdot)$ and ε, respectively, and

$$\sup_x \sup_{r_1 \in U_1} \sup_{r_2 \in U_2} \left| J(x, r_1, r_2) - J(x, r_1, u_2^{\mu,\delta}) \right| \leqslant \varepsilon. \tag{8.2.37}$$

The expressions (8.2.36) *and* (8.2.37) *hold with indices* 1 *and* 2 *interchanged.*

Under the assumption (A8.2.7) in lieu of assumption (A8.2.6), we have the following result.

Theorem 8.2.7. *If* (A8.2.7) *replaces* (A8.2.6) *in Theorem* 8.2.3–8.2.5*, their conclusions continue to hold.*

For the proof, we refer to Kushner [108]. Theorem 8.2.5–8.2.7 imply that the values defined by equations (8.2.7) and (8.2.8) would not change if relaxed strategies were used.
Next result states the existence of the value for the Game.

Theorem 8.2.8. *Assume* (A8.2.1)–(A8.2.3), (A8.2.5)*, and either* (A8.2.6) *or* (A8.2.7). *Then the game has a value in that* (8.2.9) *holds.*

8.2.7 *Near Optimal Policies*

We will construct particular ε-optimal minimizing and maximizing policies which will be needed in the proof of convergence of the numerical method. The constructed policies are for mathematical purposes only and presently do not have any computational value. Let $r^h(\cdot)$ denote the continuous time interpolation of the relaxed strategy representation of the optimal strategy approximating chain ξ_n^h. Then the optimal payoff $V^h(x)$ equals $J^h(x,r^h)$, and the corresponding set $\{\psi^h(\cdot),z^h(\cdot),w^h(\cdot),r^h(\cdot)\}$ is tight. The limit $(x(\cdot),z(\cdot),w(\cdot),r(\cdot))$ of any weakly convergent subsequence satisfies the (one player form of) (8.2.4). Hence, it cannot be better than an optimal solution for (8.2.4). This implies that $\liminf_h V^h(x) \geqslant V(x)$, is the minimal value of the payoff for (8.2.4).
To complete the convergence proof, we need to show that $\limsup_h V^h(x) \geqslant V(x)$. To do this, given an arbitrary $\varepsilon > 0$, a special ε-optimal strategy for (8.2.4) was constructed, that could be adapted for use of the chain. Let $r^\varepsilon(\cdot)$ denote the relaxed strategy form of this special ε-optimal strategy for (8.2.4), with Wiener process $w^\varepsilon(\cdot)$ and associated solution and reflection process $(x^\varepsilon(\cdot),z^\varepsilon(\cdot))$. Let $r^{\varepsilon,h}(\cdot)$ denote the relaxed strategy form of the adaption of this special strategy for use on the chain ξ_n^h, interpolated to continuous time, and let $\{\psi^{\varepsilon,h}(\cdot),z^{\varepsilon,h}(\cdot),w^{\varepsilon,h}(\cdot)\}$ denote the continuous time interpolation of the corresponding solution, reflection process and “pre-Wiener” process in the representation (8.2.19). Since $r^{\varepsilon,h}(\cdot)$ is no better than the optimal strategy for the chain, $V^h(x) \leqslant J^h\left(x,r^{\varepsilon,h}\right)$. By the method of construction of $r^{\varepsilon,h}(\cdot)$, the set $\{\psi^{\varepsilon,h}(\cdot),z^{\varepsilon,h}(\cdot),w^{\varepsilon,h}(\cdot),r^{\varepsilon,h}(\cdot)\}$ converged weakly to the set $\{\psi^\varepsilon(\cdot),z^\varepsilon(\cdot),w^\varepsilon(\cdot),r^\varepsilon(\cdot)\}$, with ε-optimal payoff $J(x,r^\varepsilon)$. Since ε is arbitrary, we have $\limsup_h V^h(x) \leqslant V(x)$, which completes the proof. $\square$
Such an ε-optimal strategy for (8.2.4) (whether minimizing or maximizing) for the player that goes first plays the same role for the problem discussed here.

Theorem 8.2.9. *Assume* (A8.2.1)–(A8.2.3), (A8.2.5) *and either* (A8.2.6) *or* (A8.2.7). *Then for each* $\varepsilon > 0$ *there is an optimal minimizing strategy law with the following properties. For positive* Δ, δ, *and* ρ, *let* δ/ρ *and* Δ/δ *be integers. The strategy is constant on the intervals* $[k\Delta, k\Delta+\Delta)$, $k = 0, 1, \ldots$, *finite-valued, the value at* $k\delta$ *is* $\Im_{k\Delta}$*-measurable, and for small* $\lambda > 0$ *it is defined by the conditional probability law* (*which defines the function* $q_{i,k}(\cdot)$), *thus,*

$$\begin{aligned} &P\big\{u_1(k\Delta) = \gamma \mid u_1(l\Delta), l < k; w(s), r_2(s),\ s < k\Delta\big\} \\ &\quad = P\big\{u_1(k\Delta) = \gamma \mid u_1(l\Delta), l < k; w(l\lambda), l\lambda < k\Delta, u_2^{\mu,\rho,\delta}(l\rho \mid r_2), l\rho < k\Delta\big\} \qquad (8.2.38) \\ &\quad = q_{1,k}\left(\gamma; w(l\lambda), l\lambda < k\Delta; u_1(l\Delta),\ l < k;\ u_2^{\mu,\rho,\delta}(l\rho \mid r_2),\ l\rho < k\Delta\right). \end{aligned}$$

The function $q_{1k}(\cdot)$ *is continuous in the w-arguments for each value of the others. Since the rule* (8.2.38) *depends on* $r_2(\cdot)$ *only via* $u_2^{\mu,\rho,\delta}(\cdot \mid r_2)$ *we write the rule as* $\overline{u}_1^{\varepsilon}\big(u_2^{\mu,\rho,\delta}(\cdot \mid r_2)\big)$. *In particular, for small* λ, μ, δ, *and large* δ/ρ *and* Δ/δ, *it satisfies the inequality*

$$\sup_{r_2 \in U_2} J\left(x, \overline{u}_1^{\varepsilon}\left(u_2^{\mu,\rho,\delta}(\cdot \mid r_2)\right), r_2\right) \leqslant V(x) + \varepsilon. \qquad (8.2.39)$$

Also, if $r_2^n(\cdot)$ *converges to* $r_2(\cdot)$, *then*

$$\limsup_n J\left(x, \overline{u}_1^{\varepsilon}\left(u_2^{\mu,\rho,\delta}(\cdot \mid r_2^n)\right), r_2^n\right) \leqslant V(x) + \varepsilon. \qquad (8.2.40)$$

For each $r_2(\cdot)$ *and* $l = 0, 1, \ldots$, *let* $\widetilde{u}_2^{\mu,\rho,\delta}(l\rho \mid r_2)$ *be a strategy that differs from* $u_2^{\mu,\rho,\delta}(l\rho \mid r_2)$ *by at most* μ *in absolute value. Then* (8.2.39) *and* (8.2.40) *hold for the perturbation* $\widetilde{u}_2^{\mu,\rho,\delta}(\cdot \mid r_2)$ *replacing* $u_2^{\mu,\rho,\delta}(\cdot \mid r_2)$.
Similarly, there is an ε*-optimal strategy rule of the same type for the maximizing player: In particular, and with the analogous terminology,*

$$\inf_{r_1 \in U_1} J\left(x, r_1, \overline{u}_2^{\varepsilon}\left(u_1^{\mu,\rho,\delta}(\cdot \mid r_1)\right)\right) \geqslant V(x) - \varepsilon, \qquad (8.2.41)$$

and (8.2.41) *continues to hold with the perturbation* $\widetilde{u}_1^{\mu,\rho,\delta}(\cdot \mid r_1)$ *replacing* $u_1^{\mu,\rho,\delta}(\cdot \mid r_1)$.

For the proof of this Theorem we refer the reader to Kushner, [108].

8.2.8 *Convergence of the Numerical Solutions*

The next result establishes the convergence of the numerical procedure. It supposes local consistency everywhere. In numerical examples, the sequence of optimal feedback strategies for the chain does converge as well. This would be the case if the optimal feedback strategies $\overline{u}_i^h(\cdot)$ for the chain converges to the feedback strategies $\overline{u}_i(\cdot)$, where the convergence is uniform and the limits are continuous outside of an arbitrary small neighborhood of a set D_d satisfying (A8.2.7), and the process (8.2.4) under $\overline{u}_i(\cdot)$ is unique in the weak sense. Then,

$$J(x, \overline{u}_1, \overline{u}_2) = V(x).$$

Now, we have the following main convergence result for the values resulting from Markov chain approximation, the proof of which can be found in Kushner, [108].

Theorem 8.2.10. *Assume the local consistency condition* (8.2.9)–(8.2.11), (A8.2.1)–(A8.2.3), (A8.2.5) *and either* (A8.2.6) *or* (A8.2.7). *Then* $V^{\pm,h}(x) \to V(x)$ *as* $h \to 0$.

8.2.9 *Stopping Time Problems and Pursuit-Evasion Games*

Suppose that player i, $i = 1,2$, now has a choice of an $\Im_t$-stopping time τ_i as well as of the strategies. Define $\tau = h\{\tau_1, \tau_2\}$. For a continuous function $g(\cdot)$, replace (8.2.5) by

$$J(x,r,\tau) = E\int_0^{\tau} e^{-\beta t}\left[\int_{\mathscr{U}_i}\sum_{i=1}^{2} k_i(x(t),\alpha_i) r_{i,t}(d\alpha_i)dt + c' dy(t)\right] + Ee^{-\beta\tau} g(x(\tau)). \quad (8.2.42)$$

Thus, in this model, the stopping payoff $g(x(\tau))$ does not depend on who selects the stopping time.

The strategy spaces such as $\mathscr{U}_i$, $U_i(\Delta)$, $L_i(\Delta)$, and $U_i(\mu,\delta,\Delta)$, etc., need to be extended so that they include the stopping times. Let $\overline{\mathscr{U}_i}$ be the set of pairs $(u_i(\cdot),\tau)$ where $u_i(\cdot) \in \mathscr{U}_i$ and τ is an $\Im_t-$ stopping time. Let $\overline{U}_i(\Delta)$ denote the subset where $u_i(\cdot) \in U_i(\Delta)$ and τ takes values $k\Delta$, $k = 0,1,\ldots$, where the set $\{w : \tau = k\Delta\}$ is $\Im_{k\Delta}$-measurable. Similarly, $\overline{U}_i(\mu,\delta,\Delta)$ denotes the subset of $\overline{U}_i(\Delta)$ where $u_i(\cdot) \in U_i(\mu,\delta,\Delta)$. Let $\overline{L}_1(\Delta)$ denote the set of strategies in $\overline{U}_1(\Delta)$ for player 1 which can be represented in the form of

$$\begin{aligned} &P\{\tau_1 > k\Delta | w(s), u_2(s)\ (s), s < t : u_1(l\Delta), l < k, \tau_1 \geqslant k\Delta\} \text{ and} \\ &P\{u_1(k\Delta) \in . | w(s), u_2(s)\ , s < t : u_1(l\Delta), l < k, \tau_1 \geqslant k\Delta\} \end{aligned} \quad (8.2.43)$$

Define $L_2(\Delta)$ analogously for player 2.

The definitions of the upper and lower values in (7.2.6) are replaced by, respectively,

$$\begin{aligned} V^+(x) &= \lim_{\Delta\to 0} \inf_{u_1,\tau_1 \in \overline{L}_1(\Delta)} \sup_{(u_2,\tau_2)\in \overline{U}_2} J(x,u_1,u_2,\tau) \text{ and} \\ V^-(x) &= \lim_{\Delta\to 0} \sup_{(u_2,\tau_2)\in \overline{U}_2} \inf_{u_1,\tau_1 \in \overline{U}_1} J(x,u_1,u_2,\tau) \end{aligned} \quad (8.2.44)$$

The first line of equation (8.2.44) is to be understood as follows. Suppose that the game has not stopped by time $k\delta$. Then at $k\delta$, player 1 goes first, and decides whether to stop or not, based on the data to time $k\Delta^-$. If it stops the game is over. If not, it selects the strategy value $u_1(k\Delta)$ (which it will use until $(k\Delta+\Delta)^-$ or until player 2 stops, whichever comes first based on data to time $k\Delta-$. If the game is not stopped at $k\delta$ by player 1, then player 2 has the opportunity to stop at any time on $[k\Delta, k\Delta+\Delta)$, with the decision to stop any time being based on all data to that time. Until it stops (if it does), it chooses admissible strategy values $u_2(\cdot)$. The procedure is then repeated at time $k\Delta+\delta$, and so forth. With these changes and minor modifications, the previous theorem continues to hold. In particular, Theorem 8.2.10 holds.

Consider the approximating Markov chain. Let player 1 go first, and let I_1 denote the indicator of the event that player 1 stops at the current step. Then the Bellman equation for the (for example) upper value is

$$\begin{aligned} V^{+,h}(x) = h\Big\{ g_1(x)I_1, (1-I_1)\max\Big[\max_{\alpha_2}\Big(E_x^{\alpha} e^{-\beta\Delta t^h(x,\alpha)}V^{+,h}(\xi_1^h) \\ +k(x,\alpha)\Delta t^h(x,\alpha)\Big), g_2(x)\Big]\Big\}. \end{aligned} \tag{8.2.45}$$

8.3 Ergodic Payoff case

The Markov chain approximation method that we discussed in Section 8.2 can be used for the ergodic payoff strategy problem. Both discounted and ergodic cases share the foundation in the theory of weak convergence. However, the approximations to the ergodic payoff strategy problem differ from the method developed in Section 8.2. We will only present the most needed modifications. For full detail, we refer to Kushner [107]. Most of the method of analysis used is parallel to the methods are given by Kushner [105]. Construction of the controlled Markov chain is similar to that discussed in Section 8.2.

Consider the system (8.2.1) (or (8.2.3)) and (8.2.2) now with ergodic payoff structure in relaxed control setting,

$$J_T[m](x) = \frac{1}{T}E_x^m \int_0^T k_m(x(s))ds + \frac{1}{T}E_x^m c'y(t)$$

where $k(x,\alpha) = k_1(x,\alpha_1)+k_2(x,\alpha_2)$, and $k_m(x) = \int_U k(x,\alpha)m(x,d\alpha)$. Payoff function of interest in this section is given by

$$J(m) = \lim_T J_T[m](x). \tag{8.3.1}$$

If player i uses relaxed strategy $r_i(\cdot)$, then we use the notation $k_{r_i}(x,t)$ to denote the $\int_{\mathcal{U}_i} k_i(x,\alpha_i)r_{i,t}(d\alpha_i)$. If player 1 selects his/her strategy first and uses a relaxed feedback strategy and player 2 selects strategy last and uses a relaxed control, then we define

$$J_T(x,m_1,r_2) = \frac{1}{T}E_x^{m_1,r_2}\int_0^T [k_{1,m_1}(x(s)) + k_{2,r_2}(x(s),s)]\,ds + \frac{1}{T}E_x^{m_1,r_2}c'y(t),$$

and

$$J(x,m_1,r_2) = \liminf_{T} J_T(x,m_1,r_2).$$

If player 2 selects his/her control first and uses a relaxed feedback strategy and player 1 uses a relaxed control strategy, define

$$J(x,r_1,m_2) = \limsup_{T} J_T(x,r_1,m_2).$$

It should be noted that in (8.3.1), we dropped the dependence on the initial position x because of Theorem 2.5 of Kushner [107], this will not depend on the initial condition.

Now, we will show the existence of optimal policies for the upper and lower values. Define the upper and lower values for the game as

$$\overline{J}^{+} = \inf_{\text{fb } m_1} \sup_{\text{rel } r_2} J(m_1,r_2) \tag{8.3.2}$$

and

$$\overline{J}^{-} = \sup_{\text{fb } m_2} \inf_{\text{rel } r_1} J(r_1,m_2) \tag{8.3.3}$$

where fb denotes relaxed feedback and *rel* denotes relaxed control strategies. We have following result that is due to Kushner, [107].

Theorem 8.3.1. *Assume* (A8.2.1), (A8.2.7). *For a sequence $\{m^n(\cdot)\}$ of relaxed feedback strategies, let $m^n(x,\cdot) \to m(x,\cdot)$ for almost all $x \in G$. Then, $J(m^n) \to J(m)$.*

For a fixed strategy $m_1(\cdot)$, maximize over $m_2(\cdot)$, and let $\{m_2^n(\cdot)\}$ be a maximizing sequence. Consider measures over the Borel sets of $G \times U$ which are defined by

$$m^n(x,d\alpha)dx = m_1(x,d\alpha_1)m_2^n(x,d\alpha_2)dx \tag{8.3.3}$$

and take a weakly convergent subsequence. The limit can be factored into the form

$$m_1(x,d\alpha_1)\overline{m}_2(x,d\alpha_2)dx, \tag{8.3.4}$$

where $\overline{m}_2(\cdot)$ is a relaxed feedback policy for player 2. *Since $\overline{m}_2(\cdot)$ depends on $m_1(\cdot)$, we write it as $\overline{m}_2(\cdot) = \overline{m}_2(\cdot;m_1)$. Then, given $m_1(\cdot)$, the relaxed feedback strategy $\overline{m}_2(\cdot;m_1)$ is maximizing for player* 2 *in that*

$$\sup_{m_2} J(m_1,m_2) = J(m_1,\overline{m}_2(m_1)).$$

The analogous result holds in the other direction, where player 2 *chooses first.*

Suppose that with $m_1(\cdot)$ fixed, player 2 is allowed to use relaxed controls and not simply relaxed feedback strategies. The following result says that the maximization over this larger class will not yield a better result for player 2. The analog of the result for player 2 choosing first also holds.

Theorem 8.3.2. *Assume the conditions of Theorem* 8.3.1. *Fix $m_1(\cdot)$ and let $m_2(\cdot, m_1)$ be an optimal relaxed feedback strategy and $r_2(\cdot)$ an arbitrary relaxed control for player* 2. *Then for each $x \in G$,*

$$J(x, m_1, r_2) \leqslant J(m_1, \overline{m}_2(m_1)).$$

Theorem 8.3.3. *Assume the conditions of Theorem* 8.3.1. *Let player* 1 *go first. Then it has an optimal strategy, denoted by $\overline{m}^+(\cdot)$. The analogous result holds if player* 2 *chooses first, and its optimal control is denoted by $\overline{m}(\cdot)$.*

Markov chain approximation is done exactly same as in Section 8.2, including the Theorem 8.2.1 continues to hold. The discretized sets and local consistency conditions are similar to that we discussed in Section 8.2. We can Discretize the ergodic payoff function and upper and lower values as follows. Relaxed feedback controls, when applied to the Markov chain, are equivalent to randomized controls. Let $u^h(\cdot) = \big(u_1^h(\cdot), u_2^h(\cdot)\big)$ be feedback strategies for the approximating chain. Then the payoff is given by

$$J_T^h(x, u^h) = J_T^h(x, u_1^h, u_2^h) = \frac{1}{T} E_x^{h,u^h} \int_0^T k_{u^h}(\psi^h(s))ds + E_x^{h,u^h} \frac{c' y^h(t)}{T}$$

and

$$J^h(u^h) = \lim_T J_T^h(x, u^h). \tag{8.3.5}$$

Let $m^h(\cdot)$ be a randomized strategy. Then the payoff function can be written as

$$J_T^h(x, m^h) = J_T^h(x, m_1^h, m_2^h) = \frac{1}{T} E_x^{h,m^h} \int_0^T k_{m^h}(\psi^h(s))ds + E_x^{h,m^h} \frac{c' y^h(t)}{T}$$

and

$$J^h(m^h) = \lim_T J_T^h(x, m^h). \tag{8.3.6}$$

With the relaxed feedback control representation of an ordinary feedback strategy, (8.3.5) is a special case of (8.3.6). Also, we can always take the strategies in (8.3.6) to be randomized feedback.

Suppose that player 1 chooses its control first and uses the relaxed feedback (or randomized feedback) strategy $m_1^h(\cdot)$. Then player 2 has a maximization problem for a finite state Markov chain. The approximating chain is ergodic for any feedback strategy, whether

randomized or not. Then, since the transition probabilities and cost rates are continuous in the control of the second player, the optimal control of the second player exists and is a pure feedback strategy (not randomized), Puterman [155]. The cost does not depend on the initial condition. The analogous situation holds if player 2 chooses its strategy first.
Let $m_i^h(\cdot)$ denote either a randomized feedback, relaxed feedback, or the relaxed feedback representation of an ordinary feedback control. Define the upper and lower values, resp.by

$$\overline{J}^{+,h} = \inf_{m_1^h} \sup_{m_2^h} J^h(m_1^h, m_2^h),$$

and

$$\overline{J}^{-,h} = \sup_{m_2^h} \inf_{m_1^h} J^h(m_1^h, m_2^h).$$

It should be noted that under our hypotheses, the upper and lower values might be different, although Theorem 8.3.4 says that they converge to the same value asymptotically. If the dynamics are separated in the sense that $P^h(x, y \mid \alpha)$ can be written as a function of (x, y, α_1) plus a function of (x, y, α_2), then $\overline{J}^{+,h} = \overline{J}^{-,h}$. Proof of the next result is available in [102], we will give it for the completeness sake.

Theorem 8.3.4. *Under the assumptions of Theorem* 8.3.1 *and suppose that*

$$\overline{J}^{+} = \overline{J}^{-} = \overline{J}. \tag{8.3.7}$$

Then

$$\overline{J}^{-} \leqslant \liminf_h \overline{J}^{-,h} \leqslant \limsup_h \overline{J}^{+,h} \leqslant \overline{J}^{+}. \tag{8.3.8}$$

Hence,

$$\lim_h \overline{J}^{+,h} = \lim_h \overline{J}^{-,h} = \overline{J} \tag{8.3.9}$$

and both the upper and lower values for the numerical approximation converge to the value for the original game.

Proof. Let player 1 choose its control first and let $\varepsilon > 0$. Let $\overline{m}_{\varepsilon,1}^{+}(\cdot)$ be an ε-smoothing of the optimal control $\overline{m}_1^{+}(\cdot)$ for player 1, when it chooses first. This implies that, given $\delta > 0$, there is $\varepsilon > 0$ such that $\overline{m}_{\varepsilon,1}^{+}(\cdot)$ is δ-optimal for player 1 for the original problem. Now, let player 1 use $\overline{m}_{\varepsilon,1}^{+}(\cdot)$ on the approximating chain, either as a randomized feedback or a relaxed feedback strategy. Given that player 1 chooses first and uses $\overline{m}_{\varepsilon,1}^{+}(\cdot)$, we have a simple strategy problem for player 2. As noted above, the optimal strategy for player 2 exists and is pure feedback, and we denote it by $\overline{u}_2^h(\cdot)$, with relaxed feedback control representation $\overline{m}_2^h(\cdot)$.

By the definition of the upper value, we have,

$$\overline{J}^{+,h} \leqslant \sup_{u_2^h} J^h(\overline{m}_{\varepsilon,1}^+, u_2^h) = \sup_{m_2^h} J^h(\overline{m}_{\varepsilon,1}^+, m_2^h) = J^h(\overline{m}_{\varepsilon,1}^+, \overline{u}_2^h), \tag{8.3.10}$$

where $u_2^h(\cdot)$ denotes an arbitrary ordinary feedback strategy, and $m_2^h(\cdot)$ an arbitrary randomized feedback strategy. The maximum value $J^h(\overline{m}_{1,\varepsilon}^+, \overline{u}_2^h)$ of the game problem for player 2 with player 1's strategy fixed at $\overline{m}_{1,\varepsilon}^+(\cdot)$ does not depend on the initial condition. Hence, without loss of generality, the corresponding continuous time interpolation, $\psi^h(\cdot)$ can be considered to be stationary. Then, using the continuity in (x, α_2) of $\int_{U_1} b(x,\alpha)\overline{m}_{1,\varepsilon}^+(x, d\alpha_1$ and of $\int_{U_1} k(x,\alpha)\overline{m}_{1,\varepsilon}^+(x, d\alpha_1)$, yields that there is a relaxed strategy $\overline{r}_2(\cdot)$ for the original problem such that

$$\limsup_h \overline{J}^{+,h} \leqslant \limsup_h J^h(\overline{m}_{1,\varepsilon}^+, \overline{u}_2^h) = J(\overline{m}_{1,\varepsilon}^+, \overline{r}_2) \leqslant \overline{J}^+ + \delta. \tag{8.3.11}$$

The last inequality of (8.3.11) follows from Theorem 8.3.2 and the δ-optimality of $\overline{m}_{1,\varepsilon}^+(\cdot)$ in the class of relaxed feedback strategies for player 1 if he/she chooses first.

Now, let player 2 choose first, Then there is an analogous result with analogous notation: In particular, given $\delta > 0$, there is an $\varepsilon > 0$ and an ε-smoothing $\overline{m}_{2,\varepsilon}^-(\cdot)$ of the optimal strategy, and a relaxed strategy $\overline{r}_1(\cdot)$ for the original game problem such that

$$\liminf_h \overline{J}^{-,h} \geqslant \liminf_h J^h\left(\overline{u}_1^h, \overline{m}_{2,\varepsilon}^-\right) = J\left(\overline{r}_2, \overline{m}_{2,\varepsilon}^-\right) \geqslant \overline{J}^- - \delta. \tag{8.3.11}$$

Hence, since δ is arbitrary, (8.3.8) holds. This, with (8.3.7), yields the theorem. □

Now we will show the existence of the value, (8.3.7). Without loss of generality, we can assume that if the $\overline{m}_i^h(\cdot)$, $i = 1,2$, are relaxed feedback strategies for each h and the $\overline{m}_i^h(x,\cdot)$ are defined for almost all x, then there is always a subsequence and relaxed feedback strategies $\overline{m}_i(\cdot)$, $i = 1,2$, for which $J^h(m_1^h, m_2^h) \to J(m_1, m_2)$.

To get the approximating process, time will be discretized but not space. Let $\Delta > 0$ denote the time discretization interval. We need to construct process whose n-step transition functions $P^\Delta(x, n\Delta, . \mid \alpha)$ have densities that are mutually absolutely continuous with respect to Lebesgue measure, uniformly in $(\Delta, \text{strategy}, t_0 \leqslant n\Delta \leqslant t_1)$ for any $0 < t_0 < t_1 < \infty$.

Consider the following procedure. Start with the process (8.2.1), but with the strategies held constant on the intervals $[l\Delta, l\Delta + \Delta)$, $l = 0, 1, \ldots$. The discrete approximation will be the samples at times $l\Delta$, $l = 0, 1, \ldots$. The policies are chosen at $t = 0$, with one of the players selected to choose first, just as for the original game. Let u_i^Δ, $i = 1,2$, denote the strategies, if in pure feedback (not relaxed or randomized) form. In relaxed control notation write the strategies as $m_i^\Delta(\cdot)$, $i = 1,2$. These strategies are used henceforth, whenever control is applied. The chosen strategies are applied at random as follows. At each time, only one of the players will use his/her strategy. At each time $l\Delta$, $l = 0,1,\ldots$, flip a fair coin. With probability 1/2, player 1 will use his/her strategy during the interval $[l\Delta, l\Delta+\Delta)$ and player 2 not. Otherwise, player 2 will use his/her strategy, and player 1 not. The values of the strategies during the interval will depend on the state at its start. The optimal strategies will be feedback. Define $x^\Delta(t) = x(l\Delta)$ on $[l\Delta, l\Delta+\Delta)$. For pure (not randomized or relaxed) feedback strategies $u_i^\Delta(\cdot)$, $i = 1,2$, the system is given by

$$dx = b^\Delta\left(x, u^\Delta\left(x^\Delta\right)\right)dt + \sigma(x)dw + dz, \tag{8.3.12}$$

where the value of $b^\Delta(\cdot)$ is determined by the coin tossing randomization procedure at the times $l\Delta$, $l = 0,1,\ldots$, In particular, at $t \in [l\Delta, l\Delta+\Delta)$, $b^\Delta\left(x, m^\Delta\left(x^\Delta\right)\right)$ is $2b_i\left(x(t), u_i^\Delta\left(x^\Delta(t)\right)\right)$, for either $i = 1$ or $i = 2$ according to the random choice made at $l\delta$. If the strategy is relaxed feedback, then write the model as

$$dx = b^\Delta\left(x, m^\Delta\left(x^\Delta\right)\right)dt + \sigma(x)dw + dz, \tag{8.3.13}$$

where at $t \in [l\Delta, l\Delta+\Delta)$, $2b_i\left(x(t), m_i^\Delta\left(x^\Delta(t)\right)\right)$ is $2\int_{\mathcal{U}_i} b_i\left(x(t), \alpha_i\right) m_i^\Delta\left(x\left(l\Delta\right), d\alpha_i\right)2$, for either $i = 1$ or $i = 2$ according to the random choice made at $l\delta$. Following the Girsanov transformation, the Wiener process $w(\cdot)$ should be indexed by the strategies $u^\Delta(\cdot)$ or $m^\Delta(\cdot)$, but we omit it for notational simplicity.

Let $E_{x(l\Delta)}^{\Delta,i,\alpha_i}$ denote the expectation of functionals on $[l\Delta, l\Delta+\Delta)$ when player i acts on that interval and uses actionα_i. Let $P_i^\Delta\left(x, .\,|\alpha_i\right)$ denote the measure of $x\left(\Delta\right)$, given that the initial condition is x, player i acts and uses strategy action α_i. The conditional mean increment in the total cost function on the time interval $[l\Delta, l\Delta+\Delta)$ is, for $u_i^\Delta\left(x\left(l\Delta\right)\right) = \alpha_i$, $i = 1,2$, given by

$$C^\Delta(x(l\Delta), \alpha) = \frac{1}{2}\sum_{i=1,2} E_{x(l\Delta)}^{\Delta,i,\alpha_i}\left[\int_{l\Delta}^{l\Delta} 2k_i(x(s), \alpha_i))ds + c'(y(l\Delta+\Delta) - y(l\Delta))\right]. \tag{8.3.14}$$

Note that $C^\Delta(x,\alpha)$ is the sum of two terms, one depending on (x,α_1) and the other on (x,α_2). The weak sense uniqueness of the solution to (8.3.1) for any strategy and initial condition implies the following result.

Theorem 8.3.5. *Assume conditions of Theorem* 8.3.1. *Then for each* $\Delta > 0$, $C^\Delta(\cdot)$ *is continuous and the measures* $P_i^\Delta(\cdot)$ *are weakly continuous in that for any bounded and continuous real-valued function* $f(\cdot)$, $\int f(y)P_i^\Delta(x, dy \mid \alpha)$ *and* $C^\Delta(x,\alpha)$ *are continuous in* (x,α).

The reason for choosing the acting strategies at random at each time $l\Delta$, $l=0,1,\ldots$, is that the randomization "separates" the cost rates and dynamics in the strategies for the two players. By separation, we mean that both the payoff function and transition function are the sum of two terms, one depending on (x,α_1) and the other on (x,α_2). This separation is important since it gives the "Isaacs condition" which is needed to assure the existence of a value for the game for the discrete time process, as seen in Theorem 8.3.6. Proceeding formally at this point, let $\mu^{\Delta}_{m^{\Delta}}(\cdot)$ denote the invariant measure under the strategy $m^{\Delta}(\cdot)$. Define the stationary strategy increment

$$\lambda^{\Delta}(m^{\Delta}) = \int_G \mu^{\Delta}_{m^{\Delta}}(dx)\left[\int_U C(x,\alpha)m^{\Delta}(x,d\alpha)\right].$$

Note that, due to the scaling, $\lambda^{\Delta}\left(m^{\Delta}\right)$ is an average over an interval of length δ: hence $\lambda^{\Delta}\left(m^{\Delta}\right) = \Delta J^{\Delta}\left(m^{\Delta}\right)$. Suppose for the moment that there is an optimal strategy $\overline{m}_i^{\Delta}(\cdot)$, $i=1,2$, for each $\Delta>0$ and define $\overline{\lambda}^{\Delta} = \lambda^{\Delta}\left(\overline{m}^{\Delta}\right)$. The "separation" is easily seen from the formal Isaacs equation for the value of the discrete time problem, namely,

$$\begin{aligned}&\overline{\lambda}^{\Delta} + \overline{g}^{\Delta}(x) = \\ &\inf_{\alpha_1}\sup_{\alpha_2}\left[\frac{1}{2}\int \overline{g}^{\Delta}(x+y)\,P_1^{\Delta}(x,dy\,|\,\alpha_1) + \frac{1}{2}\int \overline{g}^{\Delta}(x+y)\,P_2^{\Delta}(x,dy\,|\,\alpha_2) + C^{\Delta}(x,\alpha)\right].\end{aligned} \tag{8.3.15}$$

where $\overline{g}^{\Delta}(\cdot)$ is the relative value or potential function.

Theorem 8.3.6. *Under the conditions of Theorem* 8.3.1*, equation* (8.3.7) *holds.*

Proof. We will work with the approximating process $x(l\Delta)$, $l=0,1,\ldots$ just described, where $x(\cdot)$ is defined by (8.3.12) with the piecewise constant control, and verify the conditions imposed in the formal discussion at the beginning of the section. Results from Kushner [105] will be exploited whenever possible. The result (8.3.13) holds (with δ replacing h) for the same reasons that it holds for the numerical approximating process described earlier. For any sequence of relaxed strategies $m_i^{\Delta}(\cdot)$, $i=1,2$, there is a subsequence (indexed by Δ) and $\overline{m}_i^{\Delta}(\cdot)$, $i=1,2$, such that,

$$m_1^{\Delta}(x,d\alpha_1)\,m_2^{\Delta}(x,d\alpha_2)\,dx \to \overline{m}_1(x,d\alpha_1)\,\overline{m}_2(x,d\alpha_2)\,dx.$$

One needs to show the analog of (8.3.13), namely (along the same subsequence, indexed by Δ)

$$J^{\Delta}(m^{\Delta}) \to J(\overline{m}). \tag{8.3.16}$$

The process $\{x(l\Delta)\}$ based on (8.3.12) inherits the crucial properties of (8.3.1), as developed by Kushner [105]. In particular, for each positive δ and n the n-step transition probability $P^{\Delta}\left(x,n\Delta,\cdot \mid m^{\Delta}\right)$ is mutually absolutely continuous with respect to Lebesgue measure, uniformly in the strategy and in $x \in G$, $n\Delta \in [t_0,t_1]$, for any $0 < t_0 < t_1 < \infty$, and it is a strong Feller process. The invariant measures are mutually absolutely continuous with respect to Lebesgue measure, again uniformly in the strategy. Then the proof of (8.3.16) is very similar to the corresponding proof for (8.3.1) given in Kushner [105] and the details are omitted. There are strategies $\overline{m}_1^{\Delta,+}(\cdot)$ which are optimal if player 1 chooses its control first (i.e., for the upper value), and $\overline{m}_2^{\Delta,-}(\cdot)$ which are optimal if player 2 chooses its strategy first (i.e., for the lower value).

We will concentrate on showing that

$$\overline{J}^{+,\Delta} = \overline{J}^{-,\Delta}. \tag{8.3.17}$$

By the (uniform in the strategies) mutual absolute continuity of the one step transition probabilities for each $\Delta > 0$, the process satisfies a Doeblin condition, uniformly in the strategy. Hence, it is uniformly ergodic, uniformly in the strategy, Meyn and Tweedie [134]. In particular it follows that there are constants K_δ and ρ_δ, with $\rho_\Delta < 1$, such that,

$$\sup_{x,m^{\Delta}} \left| E_x^{\Delta,m^{\Delta}} \int_U C\big(x(n\Delta),\alpha\big) m^{\Delta}\big(x(n\Delta),d\alpha\big) - \lambda^{\Delta}\big(m^{\Delta}\big) \right| \leqslant K_\Delta\big[\rho_\Delta\big],$$

where $\lambda^{\Delta}\big(m^{\Delta}\big)$ is defined above by (8.3.15).

Define the relative value function by

$$g^{\Delta}\big(x,m^{\Delta}\big) = \sum_{l=0}^{\infty} \left[E_x^{\Delta,m^{\Delta}} C\big(x(l\Delta), m^{\Delta}(x(n\Delta))\big) - \lambda^{\Delta}\big(m^{\Delta}\big) \right].$$

The summands converge to zero exponentially, uniformly in $\big(x,m^{\Delta}(\cdot)\big)$. Also, by the strong Feller property the summands (for $l > 0$) are continuous. Define $g^{\Delta,+}(x) = g^{\Delta}\big(x,\overline{m}^{\Delta,+}\big)$ and $g^{\Delta,-}(x) = g^{\Delta}\big(x,\overline{m}^{\Delta,-}\big)$. Then, a direct evaluation yields

$$\overline{\lambda}^{\Delta,+} + g^{\Delta,+}(x) = E_x^{\Delta,\overline{m}^{\Delta,+}} \left[g^{\Delta,+}(x(\Delta)) + C^{\Delta}\big(x,\overline{m}^{\Delta,+}(x)\big) \right]. \tag{8.3.18}$$

Next we will show that under $\overline{m}_1^{\Delta,+}(\cdot)$ (and for almost all x), that

$$\overline{\lambda}^{\Delta,+} + g^{\Delta,+}(x) = \sup_{\alpha_2} \left[E_x^{\Delta,\overline{m}^{\Delta,+},\alpha_2} g^{\Delta,+}(x(\Delta)) + C^{\Delta}\big(x,\overline{m}_1^{\Delta,+}(x),\alpha_2\big) \right]. \tag{8.3.19}$$

By (8.3.18), (8.3.19) holds for almost all x with the equality replaced by the inequality $\leqslant$. The function in brackets in (8.3.19) is continuous in α_2, uniformly in $x \in G$. Suppose that (8.3.19) does not hold on a set $A \subset G$ of Lebesgue measure $l(a) > 0$. Let $\overline{m}_2^{\Delta}(\cdot)$ denote the (relaxed feedback strategy representation of the) maximizing strategy in (8.3.19). Then

$$\overline{\lambda}^{\Delta,+} + g^{\Delta,+}(x) \leqslant \left[E_x^{\Delta,\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}} g^{\Delta,+}(x(\Delta)) + C^{\Delta}\big(x,\overline{m}_1^{\Delta,+}(x),\overline{m}_2^{\Delta}(x)\big) \right], \tag{8.3.20}$$

with strict inequality for $x \in A$. Now, integrate both sides of (8.3.20) with respect to the invariant measure $\mu^{\Delta}_{\{\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}\}}(\cdot)$ corresponding to the strategy $\left(\overline{m}_1^{\Delta}(\cdot),\overline{m}_2^{\Delta}(\cdot)\right)$ and note that

$$\int g^{\Delta,+}(x)\mu^{\Delta}_{\{\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}\}}(dx) = \int \left[E_x^{\Delta,\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}} g^{\Delta,+}(x(\Delta))\right]\mu^{\Delta}_{\{\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}\}}(dx). \tag{8.3.21}$$

Also, by definition, we have,

$$\lambda^{\Delta}\left(\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}\right) = \int C^{\Delta}\left(x,\overline{m}_1^{\Delta,+}(x),\overline{m}_2^{\Delta}(x)\right)\mu^{\Delta}_{\{\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}\}}(dx).$$

Then, canceling the terms in (8.3.21) from the integrated inequality and using the fact that the invariant measure is mutually absolutely continuous with respect to Lebesgue measure we have, $\overline{\lambda}^{\Delta,+} < \lambda^{\Delta}\left(\overline{m}_1^{\Delta,+},\overline{m}_2^{\Delta}\right)$, which contradicts the optimality of $\overline{m}_2^{\Delta,+}(\cdot)$ for player 2, if player 1 selects his/her strategy first. Thus, (8.3.19) holds.

Next, given that (8.3.19) holds, let us show that for almost all x, we have,

$$\overline{\lambda}^{\Delta,+} + g^{\Delta,+}(x) = \inf_{\alpha_1}\sup_{\alpha_2} E_x^{\Delta,\overline{m}^{\Delta,+},\alpha_1,\alpha_2}\left[g^{\Delta,+}(x(\Delta)) + C^{\Delta}(x,\alpha_1,\alpha_2)\right]. \tag{8.3.22}$$

By (8.3.19), this last equation holds if $\overline{m}_1^{\Delta,+}(\cdot)$ replaces α_1 and the *inf* is dropped. Suppose that (8.3.22) is false. Then there are $A \subset G$ with $l(a) > 0$ and $\varepsilon > 0$ such that for $x \in A$ the equality is replaced by the inequality $\geqslant$ plus ε, with the inequality $\geqslant$ holding for almost all other $x \in G$. More particularly, let $\widehat{m}_1^{\Delta,+}(\cdot)$ denote the minimizing strategy for player 1 in (8.3.22). Then we have, for almost all x and any $m_2^{\Delta}(\cdot)$,

$$\overline{\lambda}^{\Delta,+} + g^{\Delta,+}(x) \geqslant E_x^{\Delta,\overline{m}^{\Delta,+},\widehat{m}_1^{\Delta},m_2^{\Delta}}\left[g^{\Delta,+}(x(\Delta)) + C^{\Delta}\left(x,\widehat{m}_1^{\Delta}(x),m_2^{\Delta}(x)\right)\right] + \varepsilon I_{\{x\in A\}}. \tag{8.3.23}$$

Now, repeating the procedure used to prove (8.3.19), integrate both sides of (8.3.23) with respect to the invariant measure associated with $\left(\widehat{m}_1^{\Delta}(\cdot),m_2^{\Delta}(\cdot)\right)$, use the fact that the invariant measure is mutually absolutely continuous with respect to Lebesgue measure, uniformly in the strategies, and cancel the terms which are analogous to those in (8.3.21), to show that

$$\overline{\lambda}^{\Delta,+} > \sup_{m_2^{\Delta}} \lambda^{\Delta}\left(\widehat{m}_1^{\Delta},m_2^{\Delta}\right).$$

This implies that $\overline{m}_1^{\Delta,+}(\cdot)$ is not optimal for player 1 if it selects his/her strategy first, a contradiction. Thus, (8.3.22) holds. The analogous procedure can be carried out for the lower value where player 2 selects his/her strategy first.

Now, the fact that the dynamics and payoff rate are separated in the strategy implies that $\inf_{\alpha_1}\sup_{\alpha_2} = \sup_{\alpha_2}\inf_{\alpha_1}$ in (8.3.22). Thus, (8.3.22) holds with the order of the *sup* and *inf* inverted. By working with the equation (8.3.22) with the *sup* and *inf* inverted and following an argument similar to that used to prove (8.3.22), we can show that $\overline{\lambda}^{\Delta,+} = \overline{\lambda}^{\Delta,-}$ and that $\overline{m}_i^{\Delta}(\cdot)$ is optimal for player i. $\square$

8.4 Non-zero-Sum Case

In the previous two sections, we have introduced a numerical method for zero-sum stochastic differential games under different payoff structures. In this section, we will extend the Markov chain approximation method to numerically solve a class of non-zero-sum stochastic differential games the strategies for the two players are separated in the dynamics and cost function. As before, we will show that equilibrium values for the approximating chain converge to equilibrium values for the original process and that any equilibrium value for the original process can be approximated by a δ-equilibrium for the chain for arbitrarily small $\delta > 0$. The numerical method solves a stochastic game for a finite-state Markov chain. This section is based on the publication of Kushner [109]. Here, the state space G and the boundary absorption are selected to simplify the development of the non-zero-sum case. We can replace the boundary absorption by boundary reflection, if the reflection directions satisfy the conditions in Section 8.2. For simplicity of notations, we will describe two person games. The method can be easily adapted to $n \geqslant 2$-person games. For the non-zero-sum game, as opposed to the zero-sum case, the players are not strictly competitive and have their own value functions, accordingly some modifications are necessary from the previous two sections.

In the two-person zero-sum game of last two sections, the advantage is that the policies are determined by a minimax operation and that there is a single payoff function, so that one player's gain is another's loss. The non-zero-sum game does not have this property, where each player has his/her own value function, and one seeks Nash equilibria and not *min max*=*max min* (that is, saddle point) solutions. Unlike the single player problem, we must work with strategies and not simply controls, at least for one of the players at a time. Furthermore, it is not too common that there is a unique equilibria, and we are forced to look at the structure of the chain much more closely and (for the purposes of the proof, not for the numerics) try to approximate it so that it has a "diffusion" form with a driving process that does not depend heavily on the strategy, with minimal change in the values. This requires that we work with strong-sense, rather than with the weak-sense solutions that were described in Kushner and Dupuis, [112].

8.4.1 *The Model*

Consider systems of the form, where $x(t) \in \mathbb{R}^{\nu}$, Euclidean ν-space, given by

$$x(t) = x(0) + \int_0^t \sum_{i=1}^{2} b^i\left(x(s), u_i(s)\right) ds + \int_0^t \sigma\left(x(s)\right) dw(s), \tag{8.4.1}$$

where Player i, $i = 1, 2$, has controls $u_i(\cdot)$, and $w(\cdot)$ is a standard vector-valued Wiener process. The control stops at the first time τ that the boundary of a set G is hit ($\tau = \infty$, if the boundary is never reached). Let $\beta > 0$ and let E_x^u denote the expectation given the use of strategy $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ and initial condition $x(0) = x$. Then the payoff function for Player i is given by

$$J_i\left(u\right) = E_x^u \int_0^{\tau} e^{-\beta t} k_i\left(x(s), u_i(s)\right) ds + E_x^u e^{-\beta t} g_i\left(x(\tau)\right) \tag{8.4.2}$$

Let $b(\cdot) = b^1(\cdot) + b^2(\cdot)$, and $k(\cdot) = k_1(\cdot) + k_2(\cdot)$. The following condition is assumed to hold. Similar to (A8.2.4), we assume the following.

(A8.4.1): The functions $b^i(\cdot)$ and $\sigma(\cdot)$ are bounded and continuous and Lipschitz continuous in x, uniformly in u. The controls $u_i(\cdot)$ for Player i take values in U_i, a compact set in some Euclidean space, and the functions $k_i(\cdot)$ and $g_i(\cdot)$ are bounded and continuous.

A strategy $u_i(\cdot)$ is said to be in $\mathscr{U}_i$, the set of admissible strategies for Player i, if it is measurable, non-anticipative with respect to $w(\cdot)$, and it is U_i-valued. For a topological space S, let $D[S; 0, \infty)$ denote the S-valued functions on $[0, \infty)$ that are right-continuous and have left-hand limits, endowed with the Skorokhod topology, see Ethier and Kurtz [55] for more discussion. If $S = \mathbb{R}^n$, then we can write $D[S; 0, \infty) = D^n[0, \infty)$, to reflect the dimensionality.

For $\phi(\cdot)$ in $D^n[0, \infty)$, define the function $\hat{\tau}(\phi)$ with values in the compactified infinite interval $\overline{\mathbb{R}}^+ = [0, \infty]$ by $\hat{\tau}(\phi) = \infty$, if $\phi(t) \in G^0$, the interior of G, for all $t < \infty$, and otherwise use

$$\hat{\tau}(\phi) = \inf\left\{t : \phi(t) \notin G^0\right\}$$

We refer to Kushner [109] for a discussion on the need of the following assumption and when it will be satisfied.

(A8.4.2): For a continuous real-valued function $\Phi(\cdot)$ on $\mathbb{R}^n$, define $G = \{x : \Phi(x) \leqslant 0\}$, and suppose that it is the closure of its interior. $\{x : \Phi(x) < 0\}$. For each initial condition and control, the function $\hat{\tau}(\cdot)$ is continuous (as a map from $D^n[0, \infty)$ to the compactified interval $[0, \infty]$) with probability one relative to the measure induced by the solution to the system (8.4.1).

8.4.2 *Randomized Stopping*

Many times, the original game problem might be defined in an unbounded space. The space is truncated only for numerical reasons. Thus, the boundaries in game problems need not be fixed. The "randomized stopping" alternative discussed next exploits these ideas and assures (A8.4.2). Under randomized stopping, the probability of stopping at time t (if the process has not yet been stopped) goes to unity as $x(t)$ at the same time approaches the boundary, ∂G. This can be formalized as follows.

Let $N_\varepsilon(\partial G)$ be the ε-neighborhood of the boundary and G^0 is the interior of G. For $\varepsilon > 0$, let $\overline{\lambda}(\cdot) > 0$ be a continuous function on the set $N_\varepsilon(\partial G) \cap G^0$. Let $\overline{\lambda}(x) \to \infty$ as x approaches to ∂G. Then, stop $x(\cdot)$ at time t with stopping rate $\overline{\lambda}(x(t))$ and stopping cost (or payoff) $g_i(x(t))$ for Player i. Such a randomized stopping is equivalent to adding an additional (and state dependent) discount factor which is active near the boundary.

Recall the relaxed control concept from Section 8.2. Define the "*product*" relaxed control $r(\cdot)$, by product of its derivatives, $r'(\cdot,t) = r_1'(\cdot,t) \times r_2'(\cdot,t)$. Thus, $r(\cdot)$ is a product measure, with marginal's $r_i(\cdot)$, $i = 1,2$. We will usually write $r(\cdot) = (r_1(\cdot), r_2(\cdot))$ without ambiguity. The pair $(w(\cdot), r(\cdot))$ is called an *admissible pair* if each of the $r_i(\cdot)$ is admissible with respect to $w(\cdot)$. In relaxed control terminology, (8.4.1) and (8.4.2) can be written as

$$\begin{aligned} x(t) &= x(0) + \sum_{i=1}^{2} \int_0^t \int_{\mathscr{U}_i} b^i(x(s),\alpha_i)\, r_i'(d\alpha_i,s)ds + \int_0^t \sigma(x(s))dw(s) \\ &= x(0) + \int_0^t \int_{\mathscr{U}_i} b(x(s),\alpha_i)\, r'(d\alpha_i,s)ds + \int_0^t \sigma(x(s))dw(s). \end{aligned} \tag{8.4.3}$$

and

$$J_i(x,r) = E_x^r \int_0^t e^{-\beta t} \int_{\mathscr{U}_i} k_i(x(s),\alpha_i)\, r_i'(d\alpha_i,s)ds + E_x^r e^{-\beta t} g_i(x(\tau)). \tag{8.4.4}$$

Now consider the discrete time form given by

$$\begin{aligned} x^\Delta(n\Delta+\Delta) = x^\Delta() &+ \int_{n\Delta}^{n\Delta+\Delta} \int_U b(x\Delta(n\Delta),\alpha) r'(d\alpha,s)\, ds \\ &+ \sigma\left(x^\Delta(n\Delta)\right)[w(n\Delta+\Delta) - w(\Delta)]. \end{aligned} \tag{8.4.5}$$

We can define the continuous time interpolation $x^\Delta(\cdot)$ either by constants $x^\Delta(t) = x^\Delta(n\Delta)$ for $t \in [n\Delta, n\Delta+\Delta)$, or by

$$x^\Delta(t) = x^\Delta(n\Delta) + \int_{n\Delta}^t \int_U b\big(x^\Delta(n\Delta),\alpha\big) r'(d\alpha,s)ds + \int_{n\Delta}^t \sigma\big(x^\Delta(n\Delta)\big)dw(t), \tag{8.4.6}$$

where it is assumed that $r(t,\cdot)$ is adapted to $\mathscr{F}_{n\Delta-}$, for $t \in [n\Delta, n\Delta+\Delta)$.

The associated payoff function $J_i^\Delta(x,r)$ is (8.4.4) with $x^\Delta(\cdot)$ replacing $x(\cdot)$. Let $r^\Delta(\cdot), r(\cdot)$ be admissible relaxed controls with respect to $w(\cdot)$ with $r^\Delta(\cdot) \to r(\cdot)$ w.p.1. (in the weak topology) and $r^\Delta(\cdot)$ adapted as above. Then, as $\Delta \to 0$, the sequence of solutions $\{x^\Delta(\cdot)\}$ of (8.4.6) also converges w.p.1, uniformly on any bounded time interval and the limit $(x(\cdot), r(\cdot), w(\cdot))$ solves (8.4.3). By the assumption (A8.4.2), the first hitting times of the boundary also converge w.p.1 to that of the limit. The payoffs converge as well. The analogous result holds if the randomized stopping alternative is used.

For the discrete time system (8.4.5) or (8.4.6), the relaxed control can be approximated by a randomized ordinary control, as follows. Let $r(\cdot)$ be a relaxed control that is admissible with respect to $w(\cdot)$. Let $\widetilde{u}^\delta_{i,n}$ be a random variable with the distribution

$$r^\Delta_{i,n}(\cdot) = E_{n\Delta}\big[r_i(\cdot,\,|\, n\Delta, n\Delta + n\Delta)\big]/\Delta,$$

where $E_{n\Delta}$ denotes the conditional expectation given $\mathscr{F}_{n\Delta-}$. Set $\widetilde{u}^\Delta_n = \big(\widetilde{u}^\Delta_{1,n}, \widetilde{u}^\Delta_{2,n}\big)$, and define its continuous-time interpolation (with intervals Δ)$\widetilde{u}^\Delta(\cdot)$, and define $\widetilde{x}^\Delta(0) = x^\Delta(0) = x(0) = x$ and

$$\widetilde{x}^\Delta(n\Delta + \Delta) = \widetilde{x}^\Delta(n\Delta) + \Delta b\big(\widetilde{x}^\Delta(n\Delta), \widetilde{u}^\Delta_n\big) + \sigma\big(\widetilde{x}^\Delta(n\Delta)\big)\big[w(n\Delta + \Delta) - w(n\Delta)\big]. \tag{8.4.7}$$

Let $\widetilde{x}^\Delta(t)$ denote the continuous time interpolation. Define $r^\Delta_n(\cdot) = r^\Delta_{1,n}(\cdot) r^\Delta_{2,n}(\cdot)$, and let $r^\Delta(\cdot)$ be the relaxed control with derivative $r^\Delta_n(\cdot)$ on $[n\Delta, n\Delta + \Delta)$. In Theorem 8.4.1, $r^\Delta(\cdot)$ is used for $x^\Delta(\cdot)$ in (8.4.6). This leads to following result that implies that in the continuous limit, randomized controls turn into relaxed controls.

Theorem 8.4.1. *Assume condition* (A8.4.1) *and use $r^\Delta_n(\cdot)$ in (8.4.5) and (8.4.6). Then for any $T < \infty$,*

$$\lim_{\Delta\to 0} \sup_{x(0)\in G} \sup_{r\in U} E \sup_{t\leqslant T} \big|x^\Delta(t) - x(t)\big|^2 = 0, \tag{8.4.8}$$

and

$$\lim_{\Delta\to 0} \sup_{x(0)\in G} \sup_{r\in U} E \sup_{t\leqslant T} \big|x^\Delta(t) - \widetilde{x}^\Delta(t)\big|^2 = 0. \tag{8.4.9}$$

Under the additional condition (A8.4.2) *the payoff for* (8.4.5) *and* (8.4.7) *converge* (*uniformly in $x(0)$, $r(\cdot)$*) *to those for* (8.4.3) *as well.*

8.4.3 *Comment on proof*

Let $\delta x^\Delta_n = x^\Delta(n\Delta) - \widetilde{x}^\Delta(n\Delta)$. Then, we can write

$$\begin{aligned}\delta x^\Delta_{n+1} = \delta x^\Delta_n + \Delta \int_U \Big[b\big(x^\Delta(n\Delta), \alpha\big) - b\big(\widetilde{x}^\Delta(n\Delta), \alpha\big)\Big]\, r^\Delta_n(d\alpha)\\ + \Big[\sigma\big(x^\Delta(n\Delta)\big) - \sigma\big(\widetilde{x}^\Delta(n\Delta)\big)\Big]\big[w(n\Delta + \Delta) - w(n\Delta)\big] + N^\Delta_n,\end{aligned}$$

where

$$N_n^\Delta = \Delta\left[\int_U b\big(\widetilde{x}^\Delta(n\Delta),\alpha\big)r_n^\Delta(d\alpha) - b\big(\widetilde{x}^\Delta(n\Delta),\widetilde{u}_n^\Delta\big)\right]$$

is an $\mathscr{F}_{n\Delta}$-martingale difference by the definition of $\widetilde{u}_n^\delta$ via the conditional distribution given by $\mathscr{F}_{n\Delta}$. Also, $E_{n\Delta}\left|N_n^\Delta\right|^2 = O(\Delta^2)$. The proof of the uniform (in the control and initial condition) convergence to zero of the absolute difference $\left|x^\Delta(\cdot) - \widetilde{x}^\Delta(\cdot)\right|$ and of the differences between the integrals

$$E\int_0^t e^{-\beta t}k\big(\widetilde{x}^\Delta(s),\widetilde{u}^\Delta(s)\big)ds,$$

and

$$E\int_0^t\int_U e^{-\beta t}k\big(x^\Delta(s),\alpha\big)r^{\Delta'}(d\alpha,s)ds,$$

can then be completed by using the Lipschitz condition and this martingale and conditional variance property. This implies (8.4.9). An analogous argument can be used to get (8.4.8) for each $r(\cdot)$ and $x(0)$. The facts that condition (A8.4.2) holds for (8.4.3) and that (8.4.8) hold simply that the stopping times for $x^\Delta(\cdot)$, $\widetilde{x}^\Delta(\cdot)$ converge to those for (8.4.3) as well as for each $x(0)$ and $r(\cdot)$.

The uniformity in (8.4.9) and in the convergence of the costs can be proven by an argument by contradiction that goes roughly as follows. Suppose, for example, that the uniformity in (8.4.9) does not hold. Then, for intervals and relaxed controls $r^m(\cdot)$, $\quad m = 1,2,\ldots$, define $r_n^{m,\Delta_m}(\cdot)$ as $r_n^\Delta(\cdot)$ was, but based on $r^m(\cdot)$, and let $r_n^{m,\Delta_m}(\cdot)$ denote the interpolation of the associated relaxed control. Let $\Delta_m \to 0$. Let $x^m(\cdot)$ solve (8.4.3) and $x^{m,\Delta_m}(\cdot)$ solve (8.4.6), both under $r^m(\cdot)$. Let $\widetilde{x}^{m,\Delta_m}(\cdot)$ solve (8.4.7) under $r^{m,\Delta_m}(\cdot)$. Suppose that, for some $T < \infty$,

$$\limsup_{m\to\infty} E\sup_{t\leqslant T}\left|x^{m,\Delta_m}(t) - \widetilde{x}^{m,\Delta_m}(t)\right|^2 > 0.$$

Take an arbitrary weakly-convergent subsequence of $x^m(\cdot)$, $x^{m,\Delta_m}(\cdot)$, $\widetilde{x}^{m,\Delta_m}(\cdot)$, $r^m(\cdot)$, $r^{m,\Delta_m}(\cdot)$, $\hat{w}(\cdot)$, also indexed by m and with (weak-sense) limit denoted by $x(\cdot)$, $\hat{x}(\cdot)$, $\widetilde{x}(\cdot)$, $r(\cdot)$, $\hat{r}(\cdot)$, $\hat{w}(\cdot)$. Then it is easy to show that $x(\cdot) = \hat{x}(\cdot) = \widetilde{x}(\cdot)$ and $r(\cdot) = \hat{r}(\cdot)$, and that $\hat{w}(\cdot)$ is a standard Wiener process, and that $x(\cdot)$, $\hat{x}(\cdot)$, $\widetilde{x}(\cdot)$, $r(\cdot)$, $\hat{r}(\cdot)$ are non-anticipative with respect to $\hat{w}(\cdot)$, and that the limit set satisfies (8.4.3). Assume, without loss of generality, that Skorokhod representation is used as described in Ethier and Kurtz, [55], so that we can suppose that the original and the limiting processes are all defined on the same probability space and that convergence is with probability l in the Skorokhod topology. Then, it results in

$$\lim_{m\to\infty} E\sup_{t\leqslant T}\left|\widetilde{x}^{m,\Delta_m}(t) - \hat{x}(t)\right|^2 = 0,$$

and

$$\lim_{m\to\infty} E \sup_{t\leqslant T} \left| x^{m,\Delta_m}(t) - \hat{x}(t) \right|^2 = 0,$$

a contradiction to the assertion that the uniformity in $x(0)$ and $r(\cdot)$ in (8.4.9) does not hold.

□

8.4.4 *Approximating the Controls*

For each admissible relaxed control $r(\cdot)$ and $\varepsilon > 0$, let $r_i^\varepsilon(\cdot)$ be admissible relaxed controls with respect to the same filtration and Wiener process $w(\cdot)$, with derivatives $r_i^{\varepsilon'}(\cdot)$, and in that it satisfies

$$\lim_{\varepsilon\to\infty} \sup_{r_i\in\mathscr{U}_i} E \sup_{t\leqslant T} \left| \int_0^t \int_{\mathscr{U}_i} \phi_i(\alpha_i) \left[r_i'(d\alpha_i,s) - r_i^{\varepsilon'}(d\alpha_i,s)ds \right] \right| = 0, \quad i = 1,2, \tag{8.4.10}$$

for each bounded and continuous real-valued nonrandom function $\phi_i(\cdot)$ and each $T < \infty$. Let $x(\cdot)$ and $x^\varepsilon(\cdot)$ denote the solutions to (8.4.3) corresponding to $r(\cdot)$ and $r^\varepsilon(\cdot)$, respectively, with the same $w(\cdot)$ used, but perhaps different initial conditions. In particular, define $x^\varepsilon(\cdot)$ by

$$x^\varepsilon(t) = x^\varepsilon(0) + \int_0^t \int_U b\left(x^\varepsilon(s), \alpha\right) r^{\varepsilon'}(d\alpha, s)ds + \int_0^t \sigma\left(x^\varepsilon(s)\right) dw(s). \tag{8.4.11}$$

Similar to Section 8.2 (Theorem 8.2.6), it is shown by Kushner [109] that the solution $x(\cdot)$ is continuous in the controls in the sense that

$$\lim_{\varepsilon\to 0} \sup_{\substack{x(0),x^\varepsilon(0): \\ |x^\varepsilon(0)-x(0)|\to 0}} \sup_{r\in U} E \sup_{s\leqslant t} \left| \delta x^\varepsilon(s) \right|^2 = 0$$

holds, and that the payoffs corresponding to $r(\cdot)$ and $r^\varepsilon(\cdot)$ are arbitrarily close for small ε, uniformly in $r(\cdot)$.

Now, similar to Section 8.2, some approximations of subsequent interest will be described. It will be seen that we can confine attention to control processes that are just piecewise constant and finite-valued ordinary admissible controls. Consider the following discretization of the set U_i. Let $U_i \in \mathbb{R}^{c_i}$, Euclidean c_i-space. Given $\mu > 0$, partition $\mathbb{R}^{c_i}$ into disjoint (hyper)cubes $\{R_i^{\mu,l}\}$ with diameters μ. The boundaries can be assigned to the subsets in any way. Define $U_i^{\mu,l} = U_i \cap R_i^{\mu,l}$, for the finite number (p_i^μ) of non-empty intersections. Choose a point $\alpha_i^{\mu,l} \in U_i^{\mu,l}$. Now, given admissible $\left(r_1(\cdot), r_2(\cdot)\right)$, define the approximating admissible relaxed control $r_i^\mu(\cdot)$ on the control value space $U_i^\mu = \left\{\alpha_i^{\mu,l}, l \leqslant p_i^\mu\right\}$ by its derivative as $r_i^{\mu'}\left(\alpha_i^{\mu,l}, t\right) = r_i'\left(U_i^{\mu,l}, t\right)$. Denote the set of such controls by $\mathscr{U}_i(\mu)$. It can be shown that the payoffs corresponding to $r(\cdot)$ and $r^\mu(\cdot)$ are arbitrarily close for small μ, uniformly in $r(\cdot)$.

Let $r_i^{\mu}(\cdot) \in \mathscr{U}_i(\mu)$, where the control value space for Player i is U_i^{μ}. Let $\Delta > 0$. Define

$$\Delta_{i,k}^{\mu,l} = r_i^{\mu}\left(\alpha_i^{\mu,l}, k\Delta\right) - r_i^{\mu}\left(\alpha_i^{\mu,l}, k\Delta - \Delta\right), \quad l \leqslant p_i^{\mu}, \quad k = 1, \ldots.$$

Now, consider the piecewise constant ordinary controls $u_i^{\mu,\Delta}(\cdot) \in \mathscr{U}_i(\mu)$ on the interval $\left[k\Delta, k\Delta + \Delta\right)$ by

$$u_i^{\mu,\Delta}(t) = \alpha_i^{\mu,l} \text{ for } t \in \left[k\Delta + \sum_{v=1}^{l-1} \Delta_{i,k}^{\mu,v},\ k\Delta + \sum_{v=1}^{l} \Delta_{i,k}^{\mu,v}\right). \tag{8.4.12}$$

Note that on the interval $\left[k\Delta, k\Delta + \Delta\right), u_i^{\mu,\Delta}(\cdot)$ takes the value $\alpha_i^{\mu,l}$ on a time interval of length $\Delta_{i,k}^{\mu,l}$. Also, observe that the $u_i^{\mu,\Delta}(\cdot)$ are "delayed," in that the values of $r_i(\cdot)$ on $\left[k\Delta - \Delta, k\Delta\right)$ determine the values of $u_i^{\mu,\Delta}(\cdot)$ on $\left[k\Delta, k\Delta + \Delta\right)$. Thus, $u_i^{\mu,\Delta}(t)$, $t \in [k\Delta, k\Delta + \Delta)$, is $\mathscr{F}_{k\Delta}$-measurable. Let $r_i^{\mu,\Delta}$ denote the relaxed control representation of $u_i^{\mu,\Delta}(\cdot)$, with time derivative $r_i^{\mu,\Delta'}(\cdot)$. Let $\mathscr{U}_i(\mu, \delta)$ denote the subset of $\mathscr{U}_i(\mu)$ that are ordinary controls and constant on the intervals $[l\delta, l\delta + \delta)$, $l = 0, 1, \ldots$.

The intervals $\Delta_{i,k}^{\mu,l}$ in (8.4.12) are real numbers. For later use, it is important to have them be some multiple of some small $\delta > 0$, where Δ/δ is an integer. We shall discuss one method of with this approach. Divide $[k\Delta, k\Delta + \Delta)$ into Δ/δ subintervals of length δ each. Working in order $l = 1, 2, \ldots$, for each value $\alpha_i^{\mu,l}$, we first assign (the integer part) $\left[\Delta_{i,k}^{\mu,l}/\delta\right]$ successive subintervals of length δ. The total fraction of time that is unassigned on any bounded time interval will go to zero as $\delta \to 0$, and how the control values are assigned to them will have little effect. However, for specificity consider the following method. The unassigned length for value $\alpha_i^{\mu,l}$ is

$$L_{i,k}^{\mu,\delta,l} = \Delta_{i,k}^{\mu,l} - \left[\Delta_{i,k}^{\mu,l}/\delta\right]\delta, \quad i \leqslant p_i^{\mu}.$$

Define the sum $S_{i,k}^{\mu,\delta} = \sum_l L_{i,k}^{\mu,\delta,l}$, which must be an integral multiple of δ. Then assign each unassigned δ-interval at random with value $\alpha_{i,k}^{\mu,l}$ chosen with probability $L_{i,k}^{\mu,\delta,l}/S_{i,k}^{\mu,\delta}$. By Theorem 8.4.1, this assignment and randomization approximates the original relaxed control.

Let $\mathscr{U}_i(\mu, \delta, \Delta)$ denote the set of such controls. If $u_i^{\mu,\delta,\Delta}(\cdot)$ is obtained from $r_i(\cdot)$ in this way, then it is a function of $r_i(\cdot)$, but this functional dependence will be omitted in that notation. Let $r_i^{\mu,\Delta,\delta'}(\cdot)$ denote the time derivative of $r_i^{\mu,\Delta,\delta}(\cdot)$.

The next theorem states that, for fixed μ and small δ, $u_i^{\mu,\delta,\Delta}(\cdot)$ well approximates the effects of $u_i^{\mu,\delta}(\cdot)$ and $r_i(\cdot)$, uniformly in $r_i(\cdot)$ and $\left\{\alpha_i^{\mu,l}\right\}$. In particular, (8.4.10) holds in the sense that, for each $\mu > 0$, $\delta > 0$, and bounded and continuous $\phi_i(\cdot)$, for $i = 1, 2$, we have,

$$\lim_{\delta \to 0} \sup_{r_i \in U_i} E \sup_{t \leqslant T} \left| \int_0^t \int_{U_i} \phi_i(\alpha_i) \left[r_i^{\mu,\delta,\Delta'}(d\alpha_i, s) - r_i^{\mu,\delta,\Delta'}(d\alpha_i, s)\right] ds \right| = 0. \tag{8.4.13}$$

Theorem 8.4.2. *Assume* (A8.4.1)–(A8.4.2). *Let* $r_i(\cdot) \in \mathscr{U}_i$*,* $i = 1,2$*. Given* $(\mu, \delta, \Delta) > 0$*, approximate as above the theorem to get* $r_i^{\mu,\delta,\Delta}(\cdot) \in \mathscr{U}_i(\mu, \delta, \Delta)$*. Then* (8.4.10) *holds for* $r_i^{\mu,\delta,\Delta}(\cdot)$ *and* (μ, δ, Δ) *replacing* $r_i^{\varepsilon}(\cdot)$ *and* ε*, respectively. Also,* (8.4.13) *holds. In Particular, given* $\varepsilon > 0$*, there are* $\mu_\varepsilon > 0$*,* $\delta_\varepsilon > 0$*,* $\delta_\varepsilon > 0$ *and* $\kappa_\varepsilon > 0$*, can be defined so that*

$$\sup_x \sup_{r_1} \sup_{r_2} \left| J_i(x, r_1, r_2) - J_i\big(x, r_1, u_2^{\mu,\delta,\Delta}\big) \right| \leqslant \varepsilon. \tag{8.4.14}$$

The expression (8.4.14) *holds with the indices* 1 *and* 2 *interchanged or if both controls are approximated.*

Consider the discrete-time system (8.4.5) *with either the interpolation that is piecewise constant or* (7.4.6)*. Then for* $\mu_\varepsilon > 0$*,* $\delta_\varepsilon > 0$*,* $\delta_\varepsilon > 0$ *and* $\kappa_\varepsilon > 0$ *can be defined so that*

$$\sup_x \sup_{r_1} \sup_{r_2} \left| J_i(x, r_1, r_2) - J_i^{\Delta}\big(x, r_1, u_2^{\mu,\delta,\Delta}\big) \right| \leqslant \varepsilon. \tag{8.4.15}$$

The expression (8.4.15) *holds with the indices* 1 *and* 2 *interchanged or if both controls are approximated and/or further delayed by* Δ*.*

8.4.5 *Equilibria and Approximations*

A strategy $c_1(\cdot)$ for Player 1 is a mapping from $\mathscr{U}_2$ to $\mathscr{U}_1$ with the following property. If admissible controls $r_2(\cdot)$ and $\widetilde{r}_2(\cdot)$ satisfy, $r_2(\cdot) = \widetilde{r}_2(\cdot)$ for $s \leqslant t$, then $c_1(r_2)s$, $s \leqslant t$, and with an analogous definition for Player 2 strategies. Let $\mathscr{C}_i$ denote the set of such strategies or mappings for Player i, $i = 1,2$. An Elliott-Kalton strategy is a generalization of a feedback control, [52]. The current control action that it yields for any player is a function only of the past control actions, and does not otherwise depend on the form of the strategy of the other player.

A pair $\overline{c}_i(\cdot) \in \mathscr{C}_i$, $i = 1,2$, is said to be an ε-*equilibrium strategy pair* if for all admissible controls $r_i(\cdot)$, $i = 1,2$,

$$\begin{aligned} J_1(x, \overline{c}_1, \overline{c}_2) &\geqslant J_1(x, r_1, \overline{c}_2) - \varepsilon, \quad \text{and} \\ J_2(x, \overline{c}_1, \overline{c}_2) &\geqslant J_2(x, \overline{c}_1, r_2) - \varepsilon. \end{aligned} \tag{8.4.16}$$

The notation $J_1(x, c_1, c_2)$ implies that each Player i, $i = 1,2$ uses its strategy $c_i(\cdot)$. When writing $J_1(x, c_1, c_2)$, it is assumed that the associated process is well defined. This will be the case here, since Theorem 8.4.2 implies that it is sufficient to restrict attention to strategies whose control functions are piecewise constant, finite-valued and can depend only on slightly delayed values of the other players control realizations. If (8.4.16) holds with $\varepsilon = 0$, then we have an equilibrium strategy pair. The strategies can be either ordinary or relaxed. The notation $J_2(x, c_1, r_2)$ implies that Player 1 uses its strategy $c_1(\cdot)$ and Player 2 uses the relaxed control strategy $r_2(\cdot)$.

The above definition of strategy does not properly allow for randomized controls, where the realized responses given by the strategy of a player to a fixed control process of the other player might differ, depending on the random choices that it makes. So we also allow randomized strategies.

(A8.4.3): For each small $\varepsilon > 0$ there is an ε-equilibrium Elliott-Kalton strategy $(\bar{c}_1^{\varepsilon}(\cdot), \bar{c}_2^{\varepsilon}(\cdot))$ under which the solution to (8.4.1) or (8.4.3) is well defined.

The following approximation result will be a key item in the development. For a proof, we refer to Kushner [109].

Theorem 8.4.3. *Assume condition* (A8.4.1) *and* (A8.4.2). *Given $\varepsilon_1 > 0$, there are positive numbers μ, δ, Δ, where Δ/δ is an integer, such that the values for any strategy pair $(c_1(\cdot), c_2(\cdot))$ with $c_i(\cdot) \in \mathscr{C}_i$ and under which the solution to* (8.4.3) *is well defined, can be approximated within ε_1 by strategy pairs $c_i^{\mu,\delta,\Delta}(\cdot)$, $i = 1, 2$, of the following form. The realizations of $c_i^{\mu,\delta,\Delta}(\cdot)$* (*which depend on the other player's strategy or control*) *are ordinary controls in $\mathscr{U}_i(\mu, \delta, \Delta)$, and we denote them by $u_i^{\mu,\delta,\Delta}(\cdot)$. For integers n, k, and $k\delta \in [n\Delta, n\Delta + \Delta)$ and α_i taking values in U_i^{μ}, we have,*

$$\begin{aligned} & P\left\{u_i^{\mu,\delta,\Delta}(k\delta) = \alpha_i | w(s),\ s \leqslant k\delta;\ u_j^{\mu,\delta,\Delta}(l\delta),\ j = 1, 2,\ l < k\right\} \\ & = P\left\{u_i^{\mu,\delta,\Delta}(k\delta) = \alpha_i | w(l\delta),\ l \leqslant n;\ u_j^{\mu,\delta,\Delta}(l\delta),\ j = 1, 2,\ l\delta < n\Delta\right\} \\ & = p_{i,k}\left(\alpha_i;\ w(l\delta),\ l \leqslant n;\ u_j^{\mu,\delta,\Delta}(l\delta),\ j = 1, 2,\ l\delta < n\Delta\right), \end{aligned} \tag{8.4.17}$$

which defines the functions $p_{i,k}(\cdot)$. For each positive value of μ, δ, Δ, the functions $p_{i,k}(\cdot)$ can be taken to be continuous in the w-arguments, for each value of the other arguments. Suppose that the control process realizations for Player i are $\mathscr{U}_i(\mu, \delta, \Delta)$, but those of the other player are general relaxed controls. Then we interpret (8.4.17)*, as applied to that control, as being based on its discretized approximation as derived above Theorem* 8.4.2.

8.4.6 *A Convenient Representation of the Values in* (8.4.17)

It will be useful for the convergence proofs if the random selections implied by the conditional probabilities in (8.4.17) were systematized as follows. Let $\{\theta_k\}$ be random variables that are mutually independent and uniformly distributed on $[0, 1]$. The $\{\theta_k,\ k \geqslant l\}$ will be independent of all system data before time $l\delta$. For each i, n, k, divide $[0, 1]$ into (random) subintervals whose lengths are proportional to the conditional probability of the $\alpha_i^{\mu,l}$ as given by equation (8.4.17), and select $u_i^{\mu,\delta,\Delta}[k\delta] = \alpha_i^{\mu,l}$ if the random selection of θ_k on $[0,1]$ falls into that subinterval. The same random variables $\{\theta_k\}$ are used for both players, and for all conditional probability rules of the form given by (8.4.17). This representation is used for theoretical purposes only.

8.4.7 *The Markov Chain Approximation Method*

As we had mentioned earlier in Sections 8.2 and 8.3, the method of Markov chain approximation is the main tool in numerical methods for the game problem and it consists of two steps. Let $h > 0$ be an approximation parameter.

(i) The first step is the determination of a finite-state controlled Markov chain ξ_n^h that has a continuous-time interpolation that is an "approximation" of the process $x(\cdot)$, the solution of (8.4.1).

(ii) The second step solves the optimization problem for the chain and a payoff function that approximates the one used for $x(\cdot)$.

Such approximations should "stay close" to the physical model and should be able to adjust to exploit local features. Under a natural "local consistency" condition, it will be shown that the optimal payoff function for the chain converges to the minimal cost function for the original problem. The book by Kushner and Dupuis [112] contains a comprehensive discussion of Markov chain approximation methods including many automatic and simple methods for getting the transition probabilities and other properties of the chain.

The simplest state space for the chain for model (8.4.1) and (8.4.2) is based on the regular h-grid S_h in $\mathbb{R}^v$. Define $G_h = S_h \cap G$ and $G_h^0 = S_h \cap G^0$, where G^0 is the interior of G. On G_h^0 the chain "approximates" the diffusion part of (8.4.1) or (8.4.3). Let ∂G_h denote the points in $\left(S_h - G_h^0\right)$ that can be reached by the Markov chain in one step from G_h^0 under an admissible strategy. These are the boundary points, and the process stops upon first reaching ∂G_h. It is only the points in $G_h^0 \cup \partial G_h$ that are of interest for the numerical development.

Next, we define the basic condition of *local consistency*. Let $u_n^h = \left(u_{1,n}^h, u_{2,n}^h\right)$ denote the strategies that are used at step n. Let ξ_n^h be the corresponding controlled Markov chain. Define, $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$ and let $E_{x,n}^{h,\alpha}$ denote the expectation given the data up to step n (when ξ_n^h has just been computed) with $\xi_n^h = x$ and control value $\alpha = u_n^h$ to be used on the next step. The following steps are for us to relate the chain to the system (8.4.1). For the game problem, $\alpha = (a_1, \alpha_2)$ with $\alpha_i \in \mathscr{U}_i$. Define $a(x) = \sigma(x)\sigma'(x)$. Suppose that there is a function $\Delta t^h(\cdot)$, such that, this defines the functions $b^h(\cdot)$ and $a^h(\cdot)$. Such a function $\Delta t^h(\cdot)$ can be obtained automatically when the transition probabilities are calculated, as given by Kushner and Dupuis [112].

$$E^{h,\alpha}_{x,n}\delta\xi^h_n \equiv b^h(x,\alpha)\Delta t^h(x,\alpha) = b(x,\alpha)\Delta t^h(x,\alpha) + \sigma\big(\Delta t^h(x,\alpha)\big),$$
$$\mathrm{cov}^{h,\alpha}_{x,n}\left[\Delta\xi^h_n - E^{h,\alpha}_{x,n}\Delta\xi^h_n\right] \equiv a^h(x,\alpha)\Delta t^h(x,\alpha) = a(x)\Delta t^h(x,\alpha) + o\big(\Delta t^h(x,\alpha)\big), \quad (8.4.18)$$

and
$$\lim_{h\to 0}\sup_{\substack{x\in G\\ \alpha\in U}} \delta t^h(x,\alpha) = 0.$$

It can be seen that the chain has the "local properties" (conditional mean change and conditional covariance) of the diffusion process of the expression (8.4.1). We can always select the transition probabilities such that the intervals $\Delta t^h(x,\alpha)$ do not depend on the control variable, although the general theory discussed in Kushner and Dupuis [112] does not require it. Such a simplification is often done in applications only to simplify the coding. Let $p^h(x,y\mid\alpha)$ denote the probability that the next state is y given that the current state is x and strategy pair $\alpha=(a_1,\alpha_2)$ is used.

Thus, under the given condition that the controls are separated in $b(\cdot)$, in that,

$$b(x,\alpha) = b^1(x,\alpha_1) + b^2(x,\alpha_2),$$

and if desired one can construct the chain so that the controls are "separated" in that the one-step transition probability is of the form

$$p^h(x,y\mid\alpha) = p^h_1(x,y\mid\alpha_1) + p^h_2(x,y\mid\alpha_2). \quad (8.4.19)$$

Similar to the expression (8.2.12), we could discretize the payoff function, $J_i(u)$ of (8.4.2). The payoff functions are the analogs of (8.4.2) or (8.4.4). The cost rate for Player i, $i=1,2$ is $k_i(x,\alpha_i)\Delta t^h(x,\alpha)$. The stopping costs are $g_i(\cdot)$, and τ^h denotes the first time that the set G^0_h is exited. Let $J^h_i(x,u^h_1,u^h_2)$ denote the expected cost for Player i, $i=1,2$ under the control sequences $u^h_i = \left\{u^h_{i,n},\ n\geqslant 0\right\}$, $i=1,2$. The numerical problem is to solve the game problem for the approximating chain. For this, we can obtain dynamic programming equation as discussed in Section 5.3, and then iteratively solve, such as using Gauss-Seidel procedure, as explained in Section 8.2.

The rest of this section will deal with the convergence aspect. For the convergence proof, it is useful to have the chains for each h defined on the same probability space, no matter what the strategies. This is done as follows. Let $\{X_n\}$ be a sequence of mutually independent random variables, uniformly distributed on the interval $[0,1]$ and such that $\{X_l,\ l\geqslant n\}$ is independent of $\left\{\xi^h_l,\ u^h_l,\ l\leqslant n\right\}$. For each value of $x=\xi^h_n$, $\alpha=u^h_n$, arrange the finite number of possible next states y in some order, and divide the interval $[0,1]$ into successive subintervals whose lengths are $p^h(x,y \mid \alpha)$. Then, for $x = \xi_n^h$, $\alpha = u_n^h$, select the next state according to where the (uniformly distributed) random choice for $\{X_n\}$ falls. The same random variables $\{X_n\}$ will be used in all cases, for all controls and all values of h.

The simplest case for illustrative purposes is the one-dimensional case and where h is small enough so that $h|b(\alpha,x)| \leqslant \sigma^2(x)$. Then we can use the transition probabilities and interval, for $x \in G_h^0$, to obtain

$$\begin{aligned} p^h(x,x\pm h \mid \alpha) &= \frac{\sigma^2(x) \pm hb(x,\alpha)}{2\sigma^2(x)}, \\ \delta t^h(x,\alpha) &= \frac{h^2}{\sigma^2(x)}, \quad \text{and} \\ \Delta t_n^h &= \frac{h^2}{\sigma^2\left(\xi_n^h\right)}. \end{aligned} \tag{8.4.20}$$

For $y \neq x \pm h$, set $p^h(x,y \mid \alpha) = 0$.

Let $\mathscr{F}_n^h$ denote the minimal σ-algebra that measures the control and state data to step n, and let E_n^h denote the expectation conditioned on. An *admissible strategy* for Player i at step n is a $\mathscr{U}_i$-valued random variable that is $\mathscr{F}_n^h$-measurable. Let $\mathscr{U}_i^h$ denote the set of the admissible control processes for Player i, $i = 1,2$.

A relaxed strategy for the chain can be defined as follows. Let $r_{i,n}^h(\cdot)$ be a probability distribution on the Borel sets of $\mathscr{U}_i$ such that $r_{i,n}^h(A)$ is $\mathscr{F}_n^h$-measurable for each Borel set $A \in \mathscr{U}_i$. Then the $r_{i,n}^h(\cdot)$ are said to be relaxed strategies for Player i, $i = 1,2$ at step n. As for the model (8.4.3), an ordinary control at step n can be represented by the relaxed control at step n defined by $r_{i,n}^h(a) = I_{\{u_{i,n}^h \in A\}}$ for each Borel set $A \subset U_i$. Define $r_n^h(\cdot)$ by,

$$r_n^h(A_1 \times A_2) = r_{1,n}^h(A_1) r_{2,n}^h(A_2),$$

where the A_i are Borel sets in U_i. The associated transition probability is

$$\int_U p^h(x,y \mid \alpha) r_n^h(d\alpha).$$

If $r_{i,n}^h(a)$ can be written as a measurable function of ξ_n^h for each Borel set A, then the control is said to be relaxed feedback. Under any feedback (or relaxed feedback or randomized feedback) control, the process ξ_n^h is a Markov chain. More general controls, under which there is more "past" dependence and the chain is not Markovian, will be used as well. Let C_i^h denote the set of control strategies for ξ_n^h.

For the proofs of convergence, we use a continuous-time interpolation $\xi^h(\cdot)$ of $\left\{\xi_n^h\right\}$ that will approximate $x(\cdot)$. This will be a continuous-time process that is constructed as follows.

Define,

$$\Delta t_n^h = \Delta t^h\left(\xi_n^h, u_n^h\right),$$

and

$$t_n^h = \sum_{i=0}^{n-1} \delta t_i^h.$$

Also, let $\xi^h(t) = \xi_n^h$ on $\left[t_n^h, t_{n+1}^h\right)$ and the continuous-time interpolations $u_i^h(\cdot)$ of the control actions for Player i, $i = 1,2$ by

$$u_i^h(t) = u_{i,n}^h, \quad t_n^h \leqslant t < t_{n+1}^h,$$

and let its relaxed control representation be denoted by $r_i^h(\cdot)$. Let $r^h(\cdot) = (r_1^h(\cdot), r_2^h(\cdot))$, with time derivative $r^{h'}(\cdot)$. We shall use U_i^h for the set of continuous time interpolations of the control for Player i, $i = 1,2$ as well.

For simplicity of convergence results, we will use an alternative interpolation. For each h, let v_n^h, $n = 0,1,\ldots$, be mutually independent and exponentially distributed random variables with unit mean, and that are independent of $\{\xi_n^h, u_n^h,\ n \geqslant 0\}$. Define,

$$\Delta \tau_n^h = v_n^h \delta t_n^h,$$

and

$$\tau_n^h = \sum_{i=0}^{n-1} \delta \tau_i^h.$$

Also, let $\psi^h(t) = \xi_n^h$ and $u_\psi^h(t) = u_n^h$ on $\left[\tau_n^h, \tau_{n+1}^h\right)$. Now, we proceed to decompose $\psi^h(\cdot)$ in terms of the continuous-time compensator and a martingale. Since the intervals between jumps are $\Delta_n^h v_n^h$, where v_n^h is exponentially distributed and independent of $\mathscr{F}_n^h$, the jump rate of $\psi^h(\cdot)$ when in state x and under control value α is $1/\Delta t^h(x,\alpha)$. Given a jump, the distribution of the next state is given by the $p^h(x,y \mid \alpha)$, and the conditional mean change is $b^h(x,\alpha)\Delta t^h(x,\alpha)$. Thus, we can write

$$\psi^h(t) = x(0) + \int_0^t b^h\left(\psi^h(s), u_\psi^h(s)\right)ds + M^h(t), \tag{8.4.21}$$

where the martingale $M^h(t)$ has quadratic variation process

$$\int_0^t \alpha^h\left(\psi^h(s), u_\psi^h(s)\right)ds.$$

Under any feedback (or randomized feedback) control, the process $\psi^h(\cdot)$ is a continuous-time Markov chain.

It can be shown that there is a martingale $\hat{w}^h(\cdot)$, with respect to the filtration generated by the state and control processes, possibly augmented by an "independent" Wiener process, such that,

$$M^h(t) = \int_0^t \sigma^h\big(\psi^h(s), u^h_\psi(s)\big)d\hat{w}^h(s) = \int_0^t \sigma^h\big(\psi^h(s), u^h_\psi(s)\big)d\hat{w}^h(s) + \varepsilon^h(t), \quad (8.4.22)$$

where $\sigma^h(\cdot)\big[\sigma^h(\cdot)\big]' = \alpha^h(\cdot)$, recall the definition of $\alpha^h(\cdot)$ in (8.4.18), that $\hat{w}^h(\cdot)$ has quadratic variation It, where I represents the identity matrix, and converges weakly to a standard Wiener process. The martingale $\varepsilon^h(\cdot)$ is due to the difference between $\sigma(x)$ and $\sigma^h(x)$, also, recall the $o(\Delta t^h)$ terms in (8.4.18) and the fact that

$$\lim_{h\to 0}\sup_{u^h} E \sup_{s\leqslant t}\left|\varepsilon^h(s)\right|^2 = 0 \quad (8.4.23)$$

for each t. Thus, the $r^h_\psi(\cdot)$ is the relaxed control representation of $u^h_\psi(\cdot)$, given by

$$\psi^h(t) = x(0) + \int_0^t\int_U b^h\big(\psi^h(s), \alpha\big) r^{h'}_\psi(d\alpha, s)ds + \int_0^t \sigma^h\big(\psi^h(s)\big)d\hat{w}^h(s) + \varepsilon^h(t). \quad (8.4.24)$$

The processes $\xi^h(\cdot)$ and $\psi^h(\cdot)$ are asymptotically equivalent, as it will be seen in the following theorem, so that any asymptotic results for one are also asymptotic
results for the other. We will use $\xi^h(\cdot)$. Under local consistency, both the time scales with intervals Δt^h_n and $\Delta\tau^h_n$ are asymptotically equivalent.
By equation (8.4.18), we can write,

$$\xi^h_{n+1} = \xi^h_n + b^h\big(\xi^h_n, u^h_n\big)\Delta t^h_n + \beta^h_n,$$

where β^h_n is a martingale difference with

$$E^h_n\big[\beta^h_n\big]\big[\beta^h_n\big]' = \alpha^h\big(\xi^h_n, u^h_n\big)\Delta t^h_n.$$

There are martingale differences δw^h_n with conditional (given F^h_n) covariance $\Delta t^h_n I$, such that, $\beta^h_n = \sigma^h\big(\xi^h_n, u^h_n\big)\delta w^h_n$. Let $w^h(\cdot)$ denote the continuous time interpolation of $\sum_{i=0}^{n-1}\delta w^h_n$ with intervals Δt^h_n. Then, with similar notation, we can write,

$$\begin{gathered}\xi^h(t) = x(0) + \int_0^t b^h(\xi^h(s), u^h(s))ds + \int_0^t \sigma^h(\xi^h(s))dw^h(s) + \varepsilon^h(t), \text{ and}\\ \int_0^t \sigma^h(\xi^h(s), u^h(s))dw^h(s) = \int_0^t \sigma^h(\xi^h(s))dw^h(s) + \varepsilon^h(t),\end{gathered} \quad (8.4.25)$$

where $\varepsilon^h(\cdot)$ satisfies equation (8.4.23) and is due to the $O(\Delta t^h)$ approximation of $\alpha^h(x,\alpha)$ by $\sigma(x)\sigma(x)'$.

8.4.8 *On the Construction of* $\delta w^h(\cdot)$

Full details for the general method of constructing $w^h(\cdot)$, are given in Kushner and Dupuis [112], we shall briefly discuss the process. Suppose that $\sigma(\cdot) = \sigma$ is a constant. Also, we shall assume that the components of x can be partitioned as $x = (x_1, x_2)$, and σ can be partitioned as

$$\sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix},$$

where the dimension of x_1 is d_1, and σ_1 is a square and invertible matrix of dimension d_1. Partition the $a^h(\cdot)$ in the second line of equation (8.4.18) as

$$a^h(x,\alpha) = \begin{bmatrix} a_1^h(x,\alpha) & a_{1,2}^h(x,\alpha) \\ a_{2,1}^h(x,\alpha) & a_2^h(x,\alpha) \end{bmatrix}.$$

Thus, as $h \to 0$, $a_1^h(\cdot) \to \sigma_1[\sigma]'$, and all other components go to zero, all uniformly in (x,α). Write the analogous partition by $w^h(\cdot) = \big(w_1^h(\cdot), w_2^h(\cdot)\big)$. For any Wiener process $w_2(\cdot)$ that is independent of the other random variables, we can let $w_2^h(\cdot) = w_2(\cdot)$. The only important component of $w^h(\cdot)$ is $w_1^h(\cdot)$ and we can write

$$\begin{aligned} \delta w_{1,n}^h &\equiv w_1^h\big(t_{n+1}^h\big) - w_1^h\big(t_n^h\big) \\ &= \Big[a_1^h\big(\xi_n^h, u_n^h\big)\Big]^{-1/2} \left[\xi_{1,n+1}^h - \xi_{1,n}^h - \int_{t_n^h}^{t_{n+1}^h} \int_U b_1^h\big(\xi_n^h,\alpha\big) r^{h'}(s,d\alpha)ds\right] \\ &= [\sigma_1]^{-1} \left[\xi_{1,n+1}^h - \xi_{1,n}^h - \int_{t_n^h}^{t_{n+1}^h} \int_U b_1^h\big(\xi_n^h,\alpha\big) r^{h'}(s,d\alpha)ds\right] + \delta\varepsilon_n^{1,h}, \end{aligned} \tag{8.4.26}$$

where $\delta\varepsilon_n^{1,h}$ is due to the approximation of $a_1^h(\cdot)$ by $\sigma_1[\sigma_1]'$ and its continuous time interpolation satisfies expression (8.4.23). If an ordinary control is used, then the double integral is simply $b_1\big(\xi_n^h, u_n^h\big)\Delta t_n^h$.

8.4.9 *First Approximations to the Chain*

Consider the representation (8.4.25), and for μ, δ, Δ, as used in Theorem 8.4.2 and the $r^h(\cdot) = (r_1^h(\cdot), r_2^h(\cdot))$ in (8.4.25), define the approximation $u_i^{\mu,\delta,\Delta,h}(\cdot)$, $i = 1,2$, analogously to what was done above Theorem 8.4.2. For the process $w^h(\cdot)$ that appears in expression (8.4.25) under the original control $r^h(\cdot)$, define the process

$$\xi^{\mu,\delta,\Delta,h}(t) = x(0) + \int_0^t b\big(\xi^{\mu,\delta,\Delta,h}(s), u^{\mu,\delta,\Delta,h}(s)\big)ds + \int_0^t \sigma\big(\xi^{\mu,\delta,\Delta,h}(s)\big)dw^h(s). \tag{8.4.27}$$

Let $r_i^{\mu,\delta,\Delta,h}(\cdot)$ denote the relaxed control representation of $u_i^{\mu,\delta,\Delta,h}(\cdot)$. The process defined by (8.4.27) is not a Markov chain even if the controls are feedback, since the $w^h(\cdot)$ is obtained from the process (8.4.25) under $r^h(\cdot)$ and not under the $r_i^{\mu,\delta,\Delta,h}(\cdot)$, $i=1,2$. Let $W_i^{\mu,\delta,\Delta,h}\big(x, r_i^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h}\big)$ denotes the cost for the process (8.4.27). Also, define the discrete time system by

$$\begin{aligned}\widetilde{\xi}^{\mu,\delta,\Delta,h}(n\Delta+\Delta) = \widetilde{\xi}^{\mu,\delta,\Delta,h}(n\Delta) + \int_{n\Delta}^{n\Delta+\Delta} b\big(\widetilde{\xi}^{\mu,\delta,\Delta,h}(n\Delta), u^{\mu,\delta,\Delta,h}(s)\big)ds \\ + \sigma\big(\widetilde{\xi}^{\mu,\delta,\Delta,h}(n\Delta)\big)\big[w^h(n\Delta+\Delta) - w^h(n\Delta)\big],\end{aligned} \tag{8.4.28}$$

with initial condition $x(0)$ and piecewise-constant continuous-time interpolation denoted by $\widetilde{\xi}^{\mu,\delta,\Delta,h}(\cdot)$. Let $\widetilde{W}_i^{\mu,\delta,\Delta,h}\big(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h}\big)$ denote the associated cost. Thus, we have the following analog of Theorem 8.4.2.

Theorem 8.4.4. *Assume condition* (A8.4.1). *Given* $(\mu,\delta,\Delta) > 0$, *approximate* $r_i^h(\cdot)$ *as noted above to get* $r_i^{\mu,\delta,\Delta,h}(\cdot)$. *Given* $\varepsilon > 0$ *and* $t < \infty$, *there are* $\mu_\varepsilon > 0$, $\delta_\varepsilon > 0$, $\Delta_\varepsilon > 0$ *and* $\kappa_\varepsilon > 0$, *such that, for* $\mu \leqslant \mu_\varepsilon$, $\delta \leqslant \delta_\varepsilon$, $\delta \leqslant \Delta_\varepsilon$ *and* $\delta/\Delta \leqslant \kappa_\varepsilon$,

$$\limsup_{h\to 0} \sup_{x, r_1^h, r_2^h} E \sup_{s \leqslant t} \big|\xi^{\mu,\delta,\Delta,h}(s) - \xi^h(s)\big| \leqslant \varepsilon, \tag{8.4.29}$$

and if (A8.4.2) *holds in addition, then*

$$\lim \sup_{h\to 0} \sup_{x, r_1^h, r_2^h} \Big| J_i^{\mu,\delta,\Delta,h}\big(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h}\big) - J_i^h\big(x, r_1^h, r_2^h\big)\Big| \leqslant \varepsilon. \tag{8.4.30}$$

The expressions (8.4.29) *and* (8.4.30) *hold if only one of the controls is approximated, and also if* $\xi^{\mu,\delta,\Delta,h}(\cdot)$ *and* $W_i^{\mu,\delta,\Delta,h}(\cdot)$ *are replaced by* $\widetilde{\xi}^{\mu,\delta,\Delta,h}(\cdot)\widetilde{J}_i^{\mu,\delta,\Delta,h}(\cdot)$, *respectively.*

Again, we refer to the complete works of Kushner [109] for the proof.

8.4.10 *Representations of the Chain With Control-Independent Driving Noise*

Consider the case where the driving noise $w^h(\cdot)$ depends on the path and control. We will need to factor $w^h(\cdot)$ as $w^h(\cdot) = \overline{w}^h(\cdot) + \zeta^h(\cdot)$ where $\overline{w}^h(\cdot)$ does not depend on the control and $\zeta^h(\cdot)$ is "asymptotically negligible." We will work with the model, where

$$\sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix},$$

the dimension of x_1 is d_1, and σ_1 is a square and invertible matrix of dimension d_1. Let $b_i(\cdot)$ denote the i^{th} component of $b(\cdot)$.

Case 1. This case arises when one uses the so-called central-difference approximation to get the transition probabilities. Suppose that $d_1 = v$, so that σ is invertible. For $a = \sigma\sigma'$, suppose that $a_{i,i} - \sum_{j:j\neq i} |a_{i,j}| \geqslant 0$. Let e_i denote the unit vector in the i^{th} coordinate direction. A "central-difference" version of the canonical form of the transition probabilities and interpolation interval is given by

$$\begin{gathered}
p^h(x, x \pm e_i h \mid \alpha) = \frac{q_{i,i} \pm h b_i(x,\alpha)/2}{Q}, \quad \delta t^h(x,\alpha) = \Delta t^h = \frac{h^2}{Q}, \\
p^h(x, x + e_i h + e_i h \mid \alpha) = p^h(x, x - e_i h - e_j h \mid \alpha) = \frac{a^+_{i,j}}{2Q}, \\
p^h(x, x + e_i h - e_j h \mid \alpha) = p^h(x, x - e_i h + e_j h \mid \alpha) = \frac{a^-_{i,j}}{2Q}, \text{ and} \\
Q = \sum_i a_{i,i} - \sum_{i,j:i\neq j} \frac{|a_{i,i}|}{2}, \quad q_{ii} = \frac{a_{i,i}}{2} - \sum_{j:j\neq i} \frac{|a_{i,j}|}{2}.
\end{gathered} \tag{8.4.31}$$

Now, let us consider that $q_{i,i} - h|b_i(x,\alpha)| \geqslant 0$. A strait forward computation using (8.4.31) shows that $b^h(x,\alpha) = b(x,\alpha)$ and $a^h(x,\alpha) = \sigma\sigma' + O(\Delta t^h)$. Also, by equation (8.4.31) we can write $\Delta t^h_n = \Delta t^h$. In one dimension, (8.4.31) reduces to (8.3.20), where $q_{1,1} = \sigma^2/2$.

Case 2. This case arises when one uses a central-difference approximation for the non-degenerate part and a one-sided or "upwind" approximation for the degenerate part Suppose that s can be partitioned as in expression (8.4.19).

For example,

$$\sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where the dimension of x_1 is d_1, and σ_1 is a square and invertible matrix of dimension d_1. The problem concerns the effect of the degenerate part.

The following canonical model for such cases is motivated by the general model of Kushner and Dupuis [112]. Define $\overline{b} = \sup_{x,\alpha} \sum_{i=d_1+1}^{v} |b_i(x,\alpha)|$. For this case, redefine

$$\Delta t^h = \Delta t^h(x,\alpha) = h^2/\left[Q + h\overline{b}\right].$$

Proceed to use the form of (8.4.31) for $i \leqslant d_1$, with Q replaced by $Q^h = Q + h\overline{b}$. For $i = d_1 + 1, \ldots, v$, use

$$p^h(x, x \pm e_i h \mid \alpha) = \frac{h b_i(x,\alpha)}{Q^h},$$

and

$$p^h(x, x \mid \alpha) = \frac{h\overline{b} - h\sum_{i=d_1+1}^{v} |b_i(x,\alpha)|}{Q^h}.$$

We still have $a^h(x,\alpha) = \sigma\sigma' + O(\Delta t^h)$ and $b^h(x,\alpha) = b(x,\alpha)$. Let E^h_n denote the expectation given all the data up to step n. The proof of the next result can be found in Kushner, [109].

Theorem 8.4.5. *Use either of the models Case* 1 *or Case* 2 *described above. Then we can write,* $\delta w_n^h = \delta \overline{w}_n^h + \delta \zeta_n^h$ *where the components are martingale differences. The* $\delta \overline{w}_n^h$ *are* $\left\{\delta \overline{w}_n^h,\ l \geqslant n\right\}$ *independent of* $\left\{\xi_1^h, u_1^h,\ l \leqslant n\right\}$*, and the components have values* $O(h)$*. Also, for either case* $E_n^h \delta \overline{w}_n^h \left[\delta \overline{w}_n^h\right]' = \Delta t^h$*, and* $E_n^h \delta \zeta_n^h \left[\delta \zeta_n^h\right]' = O\left(h\Delta t^h\right)$*,* $E_n^h \delta \zeta_n^h \left[\delta \overline{w}_n^h\right]' = O\left(h\Delta t^h\right)$*.*

In the next result, $\sigma(\cdot)$ is just the constant σ. Theorem 8.4.5 implies that $\xi^h(\cdot)$ can be written in the form of

$$\xi^h(t) = x(0) + \int_0^t \int_U b\left(\xi^h(s), \alpha\right) r^{h'}(d\alpha, s)\, ds + \int_0^t \sigma d\overline{w}^h(s) + \varepsilon_2^h(t), \tag{8.4.32}$$

where $\varepsilon_2^h(\cdot)$ equals $\varepsilon_1^h(\cdot)$ plus a stochastic integral with respect to $\zeta^h(\cdot)$,and satisfies (8.4.23). The quadratic variation process of $\overline{w}^h(\cdot)$ is It, where I is the identity matrix. Now, we have the following result.

Theorem 8.4.6. *Assume* (A8.4.1) *and the models of Theorem* 8.4.5*. Define*

$$\overline{\xi}^h(t) = x(0) + \int_0^t \int_U b(\overline{\xi}^h(s), \alpha)\, r^{h'}(d\alpha, s) ds + \int_0^t \sigma d\overline{w}^h(s). \tag{8.4.33}$$

then, for each $t > 0$*,*

$$\lim_{h\to 0} \sup_{x(0), r^h} E \sup_{s \leqslant t} \left|\xi^h(s) - \overline{\xi}^h(s)\right|^2 = 0. \tag{8.4.34}$$

If (A8.4.2) *is assumed as well, then the costs for the two processes are arbitrarily close, uniformly in the control and initial condition.*

Also, given $(\mu, \delta, \Delta) > 0$*, let* $u_i^{\mu,\delta,\Delta,h}(\cdot)$ *be the delayed and discretized approximation of* $r_i^h(\cdot)$ *that would be defined by the procedure above Theorem* 8.4.2*, with relaxed control representation of the pair* $(i = 1, 2)$ *of approximations being* $r^{\mu,\delta,\Delta,h}(\cdot)$*. Define the system*

$$\overline{\xi}^{\mu,\delta,\Delta,h}(t) = x(0) + \int_0^t \int_U b\left(\overline{\xi}^{\mu,\delta,\Delta,h}(s), \alpha\right) r^{\mu,\delta,\Delta,h'}(d\alpha, s)\, ds. \tag{8.4.35}$$

Then, for $t > 0$ *and* $\gamma > 0$ *there are positive numbers* μ_γ*,* δ_γ*,* Δ_γ*,* h_γ*,* κ_γ *such that for* $\mu \leqslant \mu_\gamma$*,* $\delta \leqslant \delta_\gamma$*,* $\Delta \leqslant \Delta_\gamma$*,* $h \leqslant h_\gamma$*,* $\delta/\Delta \leqslant \kappa_\gamma$*, we have,*

$$\sup_{r^h, x(0)} E \sup_{s \leqslant t} \left|\overline{\xi}^{\mu,\delta,\Delta,h}(s) - \overline{\xi}^h(s)\right|^2 \leqslant \gamma. \tag{8.4.36}$$

If (A8.4.2) *is assumed as well, then for small* (μ, δ, Δ, h) *the costs are arbitrarily close, uniformly in the control and initial condition.*

The next result states that an approximate equilibrium for the diffusion model (8.4.1) or (8.4.3) is an approximate equilibrium for the chain and vice versa. This can be proved using the techniques we discussed in Chapter 6. For complete details, see Kushner [109].

Theorem 8.4.7. *Assume* (A8.4.1)*,* (A8.4.2)*, and* (A8.4.3)*. An* ε*-equilibrium value for* (8.4.1) *or* (8.4.3) *is an* ε_1*-equilibrium value for the approximating Markov chain, where* $\varepsilon_1 \to 0$ *as* $\varepsilon \to 0$*.*

8.4.11 *The Converse Result*

If the ε-equilibrium value for the chain is unique for arbitrarily small ε, then the converse result is true; namely, that ε-equilibrium values for the chain are ε_1-equilibrium values for (8.4.3), where $\varepsilon_1 \to 0$ as $\varepsilon \to 0$, and we are done, since Theorem 8.4.7, implies that the ε-equilibrium values for the diffusion are also unique for small ε, and that the numerical solutions will converge to the desired value. If the ε-equilibrium value for the chain is not unique for arbitrarily small ε, then we will show that this "converse" assertion is true for the models used in Theorem 8.4.5. When $\sigma(\cdot)$ is constant, we have next result, due to Kushner [109]. To show the converse result when $\sigma(\cdot)$ depends on x, needs further research.

Theorem 8.4.8. *Assume* (A8.4.1) *and* (A8.4.2) *and the models used in Theorem* 8.4.5*, where $\sigma(\cdot)$ is constant. Then for any $\varepsilon > 0$ there is $\varepsilon_1 \to 0$ which goes to zero as $\varepsilon \to 0$ such that an ε-equilibrium value for the chain ξ_n^h for small h is an ε_1-equilibrium value for* (8.4.3).

In this Chapter, we have summarized some of the numerical methods for stochastic differential games that are based on Markov chain approximation method. These results are originally derived by Kushner, [107, 108, 109], and we refer the reader to these works for more details. It is important to observe that the basic Markov chain methods are similar for most types of the game problems and the basic philosophy of these approximations for both control and game problems are same. Majority of discussion in this Chapter as well as different works in numerical methods for control and game problems in literature deal with convergence aspect. Efficient coding of the methods developed here needs further attention. Some basics of coding in the case of control of heavy traffic queues (that also can be adapted to other forms of control problem) is discussed in Kushner and Ramachandran, [113]. For a general development of numerical methods based on Markov chain approximation, we refer to an excellent book by Kushner and Dupuis, [112].

Chapter 9

Applications to Finance

9.1 Introduction

Stochastic differential game models are increasingly used in various fields ranging from environmental planning, market development, natural resources extraction, competition policy, negotiation techniques, capital accumulation, investment and inventory management, to name a few. Military applications of differential games such as aircraft combat and missile control are well known. There are tremendous amount of work in the field of mathematical finance and economics, Basak *et al.* [8], Basu *et al.* [18], Ramachandran *et al.* [162], [168], Samuelson [173], Shell [176], Sorger [179], Wan [206], Yeung [214], among others. In this chapter, we will discuss a couple of such applications.

In Yavin [208], stochastic differential game techniques are applied to compare the performance of a medium-range air-to-air missile for different values of the second ignition time in a two-pulse rocket motor. The measure of performance is the probability that it will reach a lock-on-point with a favorable range of guidance and flight parameters, during a fixed time interval. A similar problem is considered in Yavin and de Villiers [212].

In mathematical finance, it is common to model investment opportunities through game theory. For example, if two investors (players) who have available two different, but possibly correlated, investment opportunities, could be modeled as stochastic dynamic investment games in continuous time, Browne [33]. There is a single payoff function which depends on both investors' wealth processes. One player chooses a dynamic portfolio strategy in order to maximize his/her expected payoff while his/her opponent is simultaneously choosing a dynamic portfolio strategy so as to minimize the same quality. This leads to a stochastic differential game with controlled drift and variance. Consider games with payoffs that depend on the achievement of relative performance goals and/or shortfalls. Browne [33] provide conditions under which a game with a general payoff function has an achievable value, and gave an explicit representation for the value and resulting equilibrium portfolio strategies in that case. It is shown that non perfect correlation is required to rule out trivial solutions. This result allows a new interpretation of the market price of risk in a Black-Scholes world. Games with discounting strategies are also discussed as are games of fixed duration related to utility maximization. In Basar [12], a stochastic model of monetary policy and inflation in continuous-time has been studied. We refer the reader to Smith [178] for a review of: (i) the development of the general equilibrium option pricing model by Black and Scholes, and the subsequent modifications of this model by Merton [133] and others; (ii) the empirical verification of these models; and (iii) applications of these models to value other contingent claim assets such as the debt and equity of a levered firm and dual purpose mutual funds.

Economists are interested in bargaining not only because many transactions are negotiated but also because, conceptually, bargaining is precisely the opposite of the "perfect competition" among infinitely many traders, in terms of which economists often think about the markets. With the advances in game theory, attempts were made to develop theories of bargaining which would predict particular outcomes in the contract curve. John Nash initiated work on this direction. Nash's approach of analyzing bargaining with complementary models –abstract models which focus on outcomes, in the spirit of "cooperative" game theory, and more detailed strategic models, in the spirit of "non-cooperative" game theory –has influenced much of the game theoretic applications in economics. We refer to Gaidov [73], and Roth [168, 169] for more details as well as details on some new approaches based on experimental economics. For a study on stochastic differential games in economic modeling, refer to Haurie [86]. We will now describe the idea of Nash equilibrium applied to the study of institutional investor speculation. The material described in the next subsection mainly comes from Yeung [214]. Later, we will also discuss a competitive advertising under stochastic perturbations.

9.2 Stochastic Equity Investment Model with Institutional Investror Speculation

In recent times, we have witnessed mounting concern and interest in the growing power of institutional investors (fund houses of various kinds) in financial markets. The shares of corporations have been increasingly concentrated in the hands of institutional investors and these investors have become the major holders of corporate stock. Since the asset prices are mainly influenced by trading, a large volume of speculative buying and selling by institutional investors often produce a profound effect on market volatility. The asset prices might fluctuate for reasons having to do more with speculative activities than with information about true fundamental values which leads to study investment behavior in a strategically interactive framework. Since the financial assets are traded continuously, it is reasonable to assume that the price dynamics are continuous time stochastic process.

Let $R(s)$ be the gross revenue/earning of a firm at time $s \in [0,\infty)$ and let m be the corresponding outlay generating this return. The net return/earnings of the stock of the firm at time s are then $R(s) - m$. The value of the firm at any time t with the discount rate r can be expressed as

$$V(t) = \int_t^{\infty} [R(s) - m] \exp[-r(s-t)] ds. \tag{9.2.1}$$

The value $V(t)$, normalized with respect to the total number of shares issued, reflects actually the price of the firm's stock and is denoted by $P(t)$. The future gross revenues are not known with certainty and vary over time according to the following dynamics:

$$dR(s) = k\big[\overline{R} - R(s)\big] ds + R(s)\mho dw(s), \tag{9.2.2}$$

where $w(s)$ is a Wiener process. The term $\mho$ is a scalar factor governing the magnitude of the stochastic element. Gross revenue tends to perturb around a central tendency $\overline{R}$, and k is the positive parameter gauging the rate adjustment of gross revenues toward their central tendency. Hence, the net return of the firm is centered around $\overline{R} - m$. Also, $R(s)$ remains positive if its initial value is positive. To simplify the derivation of a closed form solution, the proportion of m to $\overline{R}$ is assumed to be equal to $k/(r+k)$.

An issue concerning institutional investors is that they are capable of initiating large block transactions. Since asset prices are influenced largely by trading, a large volume of speculative buying and selling by institutional investors often produces a significant effect on market volatility. The following model reflects the sensitivity of market price to institutional investor's actions. Let there be n institutional investors in the market. In Yeung [214], it is assumed that n is less than three and the price dynamics is given by the following expression,

$$dP(s) = \left\{ -a \left[\sum_{j=1}^{n} u_j \right]^{1/3} - (k/r)[rP(s) - (\overline{R} - m)] \right\} ds + P(s) dw(s), \tag{9.2.3}$$

where u_j is the quantity of stock sold by institutional investor j. Negative u_j represents quantity of stock purchased. The parameter a is used to gauge the sensitivity of the market price to the large trader's action. The dynamics given by (9.2.3) show that institutional buying would create an upward pressure on equity, price and that institutional selling would exert a downward pressure. Denoting the quantity of the stock held by institutional investor i at time s by $x_i(s)$ and the discount rate by r, the ith investor seeks to maximize the payoff given by

$$J_i(u_i,P,R,x,t) = E_0\left[\int_0^{\infty}\left\{P(s)u_i(s)+[R(s)-m]x_i(s)\right\}\exp(-rs)ds\right], \tag{9.2.4}$$

subject to the stock dynamics

$$dx_i(s) = -u_i(s)ds, \tag{9.2.5}$$

earning variation (9.2.2) and price dynamics as expressed in (9.2.3). The term $P(s)u_i(s)$ represents the revenue/outlay from selling/buying of stocks at time s, and the dividend yield is $[R(s)-m]x_i(s)$. Equation (9.2.5) shows that the quantity of stock held by institutional investor i varies according to their buying and selling the stock.
Now we consider the equilibrium outcome in the equity market defined by (9.2.2), (9.2.3), (9.2.4) and (9.2.5). For the solution concept, we will adopt a feedback Nash equilibrium (FNE) notion. The institutional investors use feedback buying and selling strategies, which at each point of time s depend on the observed values of stock price, the firm dividend, and the quantity of stock held by each institutional investor. Let $x=(x_1,\ldots,x_n)$ be the vector of stock holdings of institutional investors.

Definition 9.2.1. A *feedback buying and selling strategy* of institutional investor i is decisions rule $u_i(s)=\Phi(P,R,x,s)$, such that, Φ_i is uniformly Lipschitz continuous in P, R, and x at every instant s in the game horizon. The *set of feasible feedback strategies* for institutional investor i is denoted by A_i.

These feedback strategies satisfy the property that investors actions are based on observed market information at each time instant. The maximized payoff of the ith institutional investor is denoted by

$$V^i(P,R,x,t) = \max_{u_i\in A_i} J_i(u_i,P,R,x,t). \tag{9.2.6}$$

By the principle of optimality, $V^i(P,R,x,t)$ must satisfy the following Hamilton-Jacobi-Bellman (HJB) equations, that is,

$$\begin{aligned} V_t^i = \max_{u_i\in A_i}\Bigg\{ & V_{x_i}^i u_i - [Pu_i + (R-m)x_i\exp(-rt) \\ & + V_P^i\left(-a\left(\sum_{j=1}^n u_j\right)^3 - (k/r)\left[rP-\left(\overline{R}-m\right)\right]\right) \\ & + V_R^i\left[k\left(\overline{R}-R\right)\right] + (1/2)V_{PP}^i\mho^2\mathbb{R}^2 + V_{PR}^i\mho^2 PR\right]\Bigg\}, \end{aligned} \tag{9.2.7}$$

$i = 1,\ldots,n$. Equations (9.2.7) characterize the maximized payoffs and give conditions from which the optimal feedback strategies of the institutional investors are derived. From this, the following set of first order equations are obtained, given by

$$P\exp(-rt) = V^i_{x_i} + V^i_P 3a\left(\sum_{j=1}^{n} u_j\right)^2, \quad i = 1,\ldots,n. \tag{9.2.8}$$

The left hand side term of (9.2.8) is the price (in present value) of a unit of the firm's stock. The term $V^i_{x_i}$ measures the change in maximized payoff due to marginal change in the quantity of stock held by the institutional investor i. The term V^i_P is the change in the investor i maximizing the payoff brought about by a marginal change in price and can be interpreted as the marginal value of maintaining price at P. The marginal effect on the stock price brought about by buying and selling is represented by the term $3a\left(\sum_1^n u_j\right)^2$. The right hand side of equation (9.2.8) reflects the marginal cost (gain) of selling (buying) and the left hand side shows the marginal cost (gain) of selling (buying). In an optimal situation, institutional investors would buy or sell up to the point where the marginal gain equals the marginal cost of trading the stock. Since the marginal effect of one institutional investor buying and selling on the stock price is related to the actions of other institutional investors, the optimal strategies are interrelated. The best (optimal) response/reaction functions of the institutional investor i to the actions of the competitors at time t can be expressed as

$$u_i = \left\{\left[P\exp(-rt) - V^i_{x_i}\right]/3aV^i_P\right\}^{1/2} - \sum_{\substack{j\neq i\\ j=1}}^{n} u_j. \tag{9.2.9}$$

The derivation of institutional investor i's optimal strategy at any time is a decision making process which takes into consideration three types of factors: (i) current observed market information $(P(t), R(t), x(t), r)$, (ii) optimal strategies chosen by competing institutional investors, and (iii) marginal value of holding the stock and marginal value of maintaining price at P. The first type of factor is available at each instant of time. The second factor is derived from the premise that investors are rational and they choose their actions with full consideration of their competitor rational behavior. The third type of factor is the result of inter temporal optimization.

Substituting u_i, $i = 1,\ldots,n$, that are obtained in equation (9.2.9) into the Hamilton-Jacobi-Bellman (HJB) equations (9.2.7), one gets a set of parabolic partial differential equations. Now, the task is to find a set of twice differentiable functions $V^i : \mathbb{R}^3 \times [0,\infty) \to \mathbb{R}$ that are governed by this set of partial differential equations. The smooth functions yield the optimal payoffs of the institutional investors and proceed to solve the game. The optimal payoffs are obtained in Yeung [203] as

$$V^i(P,R,x,t) = \left\{A[P - R/(r+k)]^{4/3} + [R/(r+k)]x_i\right\}\exp(-rt), \quad i = 1,\ldots,n, \tag{9.2.10}$$

where A is a constant given by,

$$A = \left\{ \left[a^{-1/2}((1/2n)-(1/6)) \right] \div \left[r+(3/4)k-(2/9)\mho^2 \right] \right\}^{2/3}.$$

The value function $V^i(P,R,x,t)$ yields the equilibrium payoff of institutional investor i. Following Samuelson [161], it is assumed that $\mho^2 \leqslant k$. This assumption guarantees that A is positive. From (9.2.10), one can derive two marginal valuation measures. The institutional investor i's marginal value of maintaining the price at P can be derived as

$$V_P^i = (4A/3)[P-R/(r+k)]^{1/3}\exp(-rt). \tag{9.2.11}$$

The investor's marginal value of holding the stock can be obtained by

$$V_{x_i}^i = [R/(r+k)]\exp(-rt). \tag{9.2.12}$$

The marginal value of stock holding is always positive. It is increasing in the current earnings and reflects the fact that higher yields raise the value of holding the stock. At the same time, it is negatively related to the discount rate and exhibits the property that the gain from investing in the stock decline as the discount rate raises. Also from equation (9.2.11), the investor marginal value of maintaining the price at P is positive (negative) when P is greater (less) than $R/(r+k)$.

Now we can derive a feedback Nash equilibrium of the equity market with speculating investors. Substituting V_P^i in (9.2.11) and $V_{x_i}^i$ in equation (9.2.12) into the optimal strategies given in equation (9.2.9), the feedback Nash equilibrium buying and selling strategies of institutional investor i is obtained by

$$\Phi_i(P,R,x,t) = (1/n)(1/4Aa)^{1/2}[P-R/(r+k)]^{1/3}, \quad i=1,\ldots,n. \tag{9.2.13}$$

The set of feedback buying and selling strategies in equation (9.2.13) constitutes a feedback Nash equilibrium of the equity market as characterized by (9.2.2), (9.2.3), (9.2.4), and (9.2.5). These buying and selling strategies are decision rules contingent upon the current values of the price and earnings.

To examine the impact of the institutional investor speculation on stock price volatility, substitute the feedback strategies in (9.2.13) into (9.2.3) to obtain the equilibrium price dynamics, that is,

$$\begin{aligned} dP(s) = \big\{ &-a(1/4aA)^{3/2}[P(s)-R(s)/(r+k)] \\ &-(k/r)[rP(s)-r\overline{R}/(r+k)]\big\} + P(s)\mho dw(s). \end{aligned} \tag{9.2.14}$$

These along with equation (9.2.2) characterize the joint behavior of the stock price and earnings of the firm. In Samuelson [173], for the equity market with numerous ordinary investors, the change in stock price of the firm is modeled by

$$dP(s) = -(k/r)[rP(s) - \big(\overline{R}-m\big)]ds + P(s)\mho dw(s). \tag{9.2.15}$$

A comparison between (9.2.14) and (9.2.3) shows additional movements, symbolized as the first term in the right-hand side of (9.2.14), in the price dynamics caused by institutional investors. In Yeung [214], an analysis is given to show that the prices tend to rise in spite of the fact that they have been valued above their intrinsic value and prices tend to drop although $P(s)$ is below its intrinsic value in the presence of institutional speculation. *Hence, one could conclude that the market is more volatile in the presence of institutional speculation.* The following results are proved by Yeung [214]: (i) The greater the discrepancy between P and $R/(r+k)$, the higher the profit of an institutional investor, and (ii) The greater the degree of uncertainty in the market, the higher the speculative profits. This implies that institutional investors are more attracted to markets with high uncertainty, like emerging markets.

9.3 Competitive Advertising under Uncertainty

Analysis of advertising policies has always been occupying a front-and-center place in market research, Chintagunta et al. [39], Erickson [54], Prasad and Sethi [153], and Sorger [179], among others. In this section, we will present an application of stochastic differential games to an optimal advertising spending in a duopolistic market where each firm's market share depends on its own and its competitor's advertising decisions, under random disturbances. All of the material of this section is covered in Prasad and Sethi [153]. A differential game model of advertising is used in which the dynamic behavior is based on the classic Vidale-Wolfe advertising model, [204], and the Lanchester model of combat, as well as being perturbed by a Brownian motion. The combination of the large amounts of money spent on advertising and potential inefficiencies in the advertising budgeting process motivates the interest in better understanding of optimal advertising budgeting.

We examine a duopoly market in a mature product category where the two firms compete for market share using advertising as the dominant marketing tool. The firms are strategic in their behavior. That is, they take actions that maximize their objective while also considering the actions of the competitor. Additionally, they interact dynamically for the foreseeable future. This is in part due to the carry-over effect of advertising. This means that advertising today's spending will continue to influence sales several days or months down the line. Each firm's advertising acts to increase its market share while the competitor's advertising acts to reduce other firm's market share. It should be observed that due to the inherent randomness in the marketplace and in the choice behavior of customers, marketing and competitive activities alone do not govern market shares in a deterministic manner. The market for cola drinks, dominated by Coke, Pepsi and their Cola Wars, provides us with an example of a market with such features, Erickson [54].

For a competitive market with stochastic disturbances and other features as described above, the objective then is to find optimal advertising expenditures over time for the two firms. Due to the carry-over effect of advertising, the optimal advertising spending over time need to be determined using dynamic optimization methods. For this purpose, we formulate a stochastic differential game model. It will be shown that there is a unique equilibrium where the optimal advertising for both firms follows a simple rule.

9.3.1 *The Model*

Consider a duopoly market in a mature product category where total sales are distributed between the two firms, denoted as firm 1 and firm 2, which compete for the market share through advertising spending. Denote the market shares of firms 1 and 2 at time t as $x(t)$ and $y(t)$, respectively.

We shall use the following notation where the subscript $i \in \{1,2\}$ is to reference the two firms. Let, $x(t) \in [0,1]$ is the market share for firm 1 with $x(0) = x_0$, and $y(t) = 1 - x(t)$ is the market share for firm 2 with $y(0) = 1 - x_0$. Also, $u_i(x(t), y(t))$, $t \geqslant 0$ is the advertising rate by firm i at time t, and $\rho_t > 0$ is the advertising effectiveness parameter for firm i. Let $\delta > 0$ be the market share decay or chum parameter, $r_t > 0$ is the discount rate for firm i, $C(u_i(t))$ is the cost of advertising, parameterized as $c_i u_i(t)^2$, $c_i > 0$, $\sigma(x(t), y(t))dw(t)$ is the disturbance function with standard white noise, V_i is the value function for firm i, α_i, β_i are the components of the value function, and $R_i = \rho_i^2/4c$, $W_i = r_i + 2\delta$, $A_i = \frac{\beta_i \rho_i^2}{2c_i} + \delta$ are some useful intermediate terms.

The model dynamics are given by the system of equations,

$$\begin{cases} dx = \left[\rho_1 u_1(x,y)\sqrt{1-x} - \rho_2 u_2(x,y)\sqrt{x} - \delta(x-y)\right]dt + \sigma(x,y)dw, \;\; x(0) = x_0, \\ dy = \left[\rho_2 u_2(x,y)\sqrt{1-y} - \rho_1 u_1(x,y)\sqrt{y} - \delta(y-x)\right]dt - \sigma(x,y)dw, \;\; y(0) = y_0. \end{cases} \tag{9.3.1}$$

The market share is nondecreasing with its own advertising, and non-increasing with the competitor's advertising expenditure. Consistent with the literature on the subject matter, non-competitive decay is proportional to market share. As we previously discussed, this churn (or decay) is caused by influences other than competitive advertising, such as a lack of perceived differentiation between brands, so that market shares tend to converge in the absence of advertising. Note that in a duopoly situation, the decay of market share for one firm is a gain in market share for the other. The market shares are subject to a white noise, $\sigma(x,y)dw$.

Since $dx+dy=0$ and since $x(0)+y(0)=1$, this implies that $x(t)+y(t)=1$ for all $t \geqslant 0$. Thus, $y(t)=1-x(t)$. Consequently, we need only use the market share of firm 1 to completely describe the market dynamics. Thus, $u_i(x,y)$, $i=1,2$ and $\sigma(x,y)$ can be written as $u_i(x,1-x)$ and $\sigma(x,1-x)$. With abuse of notation, we will use $u_i(x)$ and $\sigma(x)$ in place of $u_i(x,1-x)$ and $\sigma(x,1-x)$, respectively. Thus,

$$dx=[\rho_1 u_1(x)\sqrt{1-x}-\rho_2 u_2(x)\sqrt{x}-\delta(2x-1)]dt+\sigma(x)dw, \quad x(0)=x_0, \tag{9.3.2}$$

with $0 \leqslant x_0 \leqslant 1$.

An important consideration when choosing a formulation is that the market share should remain bounded within $[0,1]$, which can be problematic given the stochastic disturbances. In our model it is easy to see that $x \in [0,1]$ almost surely (i.e., with probability 1) for $t>0$, as long as $u_i(x)$ and $\sigma(x)$ are continuous functions which satisfy Lipschitz conditions on every closed subinterval of $(0,1)$ and further that $u_i(x) \geqslant 0$, $x \in [0,1]$ and $\sigma(x)>0$, $x \in (0,1)$ and $\sigma(0)=\sigma(1)=0$. With these assumptions, we have a strictly positive drift at $x=0$ and a strictly negative drift at $x=1$, that is,

$$\rho_1 u_1(0)\sqrt{1-0}+\delta>0, \quad \text{and} \quad -\rho_2 u_2(1)-\delta<0. \tag{9.3.3}$$

Then from Gihman and Skorohod (1973) (Theorem 2, pp. 149, 157–158), $x=0$ and $x=1$ are natural boundaries for the solutions of equation (9.3.2) with $x_0 \in [0,1]$, i.e., $x \in (0,1)$ almost surely for $t>0$.

Let m_i denote the industry sales volume multiplied by the per unit profit margin for firm i. The objective functions for the two firms are given by

$$\begin{cases} V_1(x_0)=\max\limits_{u_1\geqslant 0} E\displaystyle\int_0^\infty e^{-r_1 t}\left[m_1 x(t)-c_1 u_1(t)^2\right]dt, \text{ and} \\ V_2(x_0)=\max\limits_{u_2\geqslant 0} E\displaystyle\int_0^\infty e^{-r_2 t}\left[m_2(1-x(t))-c_2 u_2(t)^2\right]dt, \text{ such that,} \\ dx=\left[\rho_1 u_1(x)\sqrt{1-x}-\rho_2 u_2(x)\sqrt{x}-\delta(2x-1)\right]dt+\sigma(x)dw, \\ x(0)=x_0 \in [0,1]. \end{cases} \tag{9.3.4}$$

Thus, each firm seeks to maximize its expected, discounted profit stream subject to the market share dynamics.

Now, we want to find the closed-loop Nash equilibrium strategies, for which, we form the Hamilton-Jacobi-Bellman (HJB) equation for each firm. That is,

$$\begin{aligned} r_1 V_1 = \max_{u_1} \Bigg\{ & m_1 x - c_1 u_1^2 \\ & + V_1'\left(\rho_1 u_1\sqrt{1-x}-\rho_2 u_2^*\sqrt{x}-\delta(2x-1)\right)+\frac{\sigma(x)^2 V_1''}{2}\Bigg\} \end{aligned} \tag{9.3.5}$$

and

$$r_2V_2 = \max_{u_2} \Big\{ m_2(1-x) - c_2u_2^2$$
$$+ V_2'\big(\rho_1 u_1^* \sqrt{1-x} - \rho_2 u_2 \sqrt{x} - \delta(2x-1)\big) + \frac{\sigma(x)^2 V_2''}{2} \Big\}, \tag{9.3.6}$$

where $V_i' = \frac{dV_i}{dx}$, $V_i'' = \frac{d^2V_i}{dx^2}$ and u_1^* and u_2^* denote the competitor's advertising policies in equations (9.3.5) and (9.3.6), respectively. We obtain the optimal feedback advertising decisions

$$u_1^*(x) = \max\left(0, \frac{V_1'(x)\rho_1\sqrt{1-x}}{2c_1}\right) \quad \text{and} \quad u_2^*(x) = \max\left(0, -\frac{V_2'(x)\rho_2\sqrt{x}}{2c_2}\right). \tag{9.3.7}$$

Since $0 \leqslant x \leqslant 1$ and since it is reasonable to expect $V_1' \geqslant 0$ and $V_2' \leqslant 0$, we can reduce the advertising decisions (9.3.7) to

$$u_1^*(x) = \frac{V_1'(x)\rho_1\sqrt{1-x}}{2c_1} \quad \text{and} \quad u_2^*(x) = -\frac{V_2'(x)\rho_2\sqrt{x}}{2c_2}, \tag{9.3.8}$$

which hold as we shall see later. Substituting (9.3.8) in equations (9.3.5) and (9.3.6), we obtain the Hamilton-Jacobi equations as

$$r_1V_1 = m_1x + \frac{V_1'^2\rho_1^2(1-x)}{4c_1} + \frac{V_1'V_2'\rho_2^2x}{2c_2} - V_1'\delta(2x-1) + \frac{\sigma(x)^2V_1''}{2} \tag{9.3.9}$$

and

$$r_2V_2 = m_2(1-x) + \frac{V_2'^2\rho_2^2x}{4c_2} + \frac{V_1'V_2'\rho_2^2(1-x)}{2c_1} - V_2'\delta(2x-1) + \frac{\sigma(x)^2V_2''}{2}. \tag{9.3.10}$$

Now consider the following particular forms for the value functions $V_1 = \alpha_1 + \beta_1 x$ and $V_2 = \alpha_2 + \beta_2(1-x)$. These value functions are used in equations (9.3.9) and (9.3.10) to determine the unknown coefficients α_1, β_1, α_2, β_2. Equating powers of x in equation (9.3.9) and powers of $(1-x)$ in equation (9.3.10), we obtain the following system of equations that can be solved for the unknown coefficients.

$$r_1\alpha_1 = \frac{\beta_1^2\rho_1^2}{4c_1} + \beta_1\delta, \tag{9.3.11}$$

$$r_1\beta_1 = m_1 - \frac{\beta_1^2\rho_1^2}{4c_1} - \frac{\beta_1\beta_2\rho_2^2}{2c_2} - 2\beta_1\delta, \tag{9.3.12}$$

$$r_2\alpha_2 = \frac{\beta_2^2\rho_2^2}{4c_2} + \beta_2\delta, \tag{9.3.13}$$

and

$$r_2\beta_2 = m_2 - \frac{\beta_2^2\rho_2^2}{4c_2} - \frac{\beta_1\beta_2\rho_1^2}{2c_1} - 2\beta_2\delta. \tag{9.3.14}$$

A unique solution to these equations, together with the requirements that $\beta_1 > 0$ and $\beta_2 > 0$, will be shown to exist. Since for firms having different parameter values, the solutions are more complicated. First we will consider the case of two symmetric firms. The case of asymmetric firms will be dealt after that.

9.3.2 *Symmetric Firms*

For the symmetric case take, $\alpha = \alpha_1 = \alpha_2$, $\beta = \beta_1 = \beta_2$, $m = m_1 = m_2$, $c = c_1 = c_2$, $\rho = \rho_1 = \rho_2$ and $r = r_1 = r_2$. Then, the four equations in (9.3.11)–(9.3.14) reduce to the following two,

$$\begin{aligned} r\alpha &= \frac{\beta^2\rho^2}{4c} + \beta\delta, \quad \text{and} \\ r\beta &= m - \frac{3\beta^2\rho^2}{4c} - 2\beta\delta. \end{aligned} \tag{9.3.15}$$

There are two solutions for β. One is negative, which makes no sense. Thus, the remaining positive solution is the only correct one. This in turn gives the corresponding α. The solution is

$$\begin{aligned} \alpha &= \frac{(r-\delta)\left(W - \sqrt{W^2+12Rm}\right) + 6Rm}{18Rr}, \quad \text{and} \\ \beta &= \frac{\sqrt{W^2+12Rm} - W}{6R}, \end{aligned} \tag{9.3.16}$$

where $R = (\rho^2/4c)$, $W = r + 2\delta$. We can now see that with the solution for the value function, the strategies specified in equation (9.3.7) reduce to that in (9.3.8). This validates the choice of (9.3.8) in deriving the value function. Note that when the margin $m = 0$, the firm makes zero profit, i.e., the value functions $V_1 = \alpha + \beta x$ and $V_2 = \alpha + \beta(1-x)$ are identically zero. In turn, this implies that the coefficients α, β, a, and b are each zero when $m = 0$.

We will summarize the analytical results of comparative statistics with symmetric firms in the Table 9.3.1, Prasad and Sethi [153].

When there is a marginal increase in the value of advertising (r increases) or a reduction in its cost (c decreases), then, the amount of advertising increases. However, contrary to what one would expect to see in a monopoly model of advertising, the value function decreases. This occurs because in this type of market all advertising occurs from competitive

Table 9.3.1 Comparative Statistics with Symmetric Firms

Variables	Parameters				
$R=\left(\rho^2/4c\right),\, W=r+2\delta$	c	ρ	m	δ	r
$\alpha=\dfrac{(r-\delta)\left(W-\sqrt{W^2+12Rm}\right)+6Rm}{18Rr}$	+	−	+	+	−
$\beta=\dfrac{\sqrt{W^2+12Rm}-W}{6R}$	+	−	+	−	−
$u_1^*=\dfrac{\left(\sqrt{W^2+12Rm}-W\right)\rho\sqrt{1-x}}{12Rc}$	−	+	+	−	−
Value function, $V_1=\alpha+\beta x$	+	−	+	?	−

$+=$increase, $-=$decrease, $?=$ambiguous.

motivations, since the optimal advertising expenditure would be zero if a single firm were to own both identical products. Advertising does not increase the size of the marketing pie but only affects its allocation. Thus, the increase in advertising causes a decrease in the value function.

However, the same logic does not apply when m increases, or r decreases. In these cases, it is true that the wasteful advertising is increased, but it is also true that the size of the pie has increased. Although intuitively it is difficult to predict that the latter effect should dominate the former, it turns out to be the case that an increase in m or decrease in r improves the value function.

The churn parameter δ reduces competitive intensity. Hence, it might be expected that an increase in δ should increase the profitability by reducing advertising. In fact, only the constant α part of the value functions increases and it is unclear what happens to the value functions overall. We can derive the exact conditions under which there is an increase or a decrease in the value function of a firm due to an increase in δ. We find that if the market share of a firm is less than half, the effect on the firm's value function is always positive. However, if the market shares of a firm is greater than half, its value function can decrease because of an increase in δ if

$$x>\frac{\sqrt{(r+2\delta)^2+12Rm}-(r+2\delta)}{6r}+\frac{1}{2}$$

is satisfied. The reason is that when a firm has a market share advantage over its rival, δ helps the rival unequally by tending to equalize market shares.

9.3.3 *Asymmetric Firms*

We now return to the general case of asymmetric firms. For asymmetric firms, we re-express equations (9.3.11)–(9.3.14) in terms of a single variable β_1 which is determined by the solution to the quadratic equation (9.3.17), that is,

$$\begin{aligned} 3R_1^2\beta_1^4 + 2R_1\left(W_1 + W_2\right)\beta_1^3 + \left(4R_2m_2 - 2R_1m_1 - W_1^2 + 2W_1W_2\right)\beta_1^2 \\ + 2m_1\left(W_1 - W_2\right)\beta_1 - m_1^2 = 0, \end{aligned} \tag{9.3.17}$$

$$\alpha_1 = \frac{\beta_1}{r_1}\left(\beta_1 R_1 + \delta\right), \tag{9.3.18}$$

$$\beta_2 = \frac{m_1 - \beta_1^2 R_1 - \beta_1 W_1}{2\beta_1 R_2}, \tag{9.3.19}$$

and

$$\alpha_2 = \frac{\beta_2}{r_2}\left(\beta_2 R_2 + \delta\right), \tag{9.3.20}$$

where $R_1 = \frac{\rho_1^2}{4c_1}$, $R_2 = \frac{\rho_2^2}{4c_2}$, $W_1 = r_1 + 2\delta$, $W_2 = r_2 + 2\delta$.
Once we obtain the correct value of β_1 out of the possible four solutions, the other coefficients can be obtained by solving for α_1 and β_2 and then, in turn, obtain α_2.
We now collect the main results of the analysis into Proposition 9.3.1 which is due to Prasad and Sethi [153], where a complete proof is given.

Proposition 9.3.1. *For the advertising game described in* (9.3.14)*:*

(a) *There exists a unique closed-loop Nash equilibrium solution to the differential game.*
(b) *Optimal advertising is*

$$u_1^*(x) = \frac{\beta_1\rho_1\sqrt{1-x}}{2c_1}, \quad u_2^*(x) = \frac{\beta_2\rho_2\sqrt{1-y}}{2c_2},$$

where in the symmetric firm case, from equation (9.3.16),

$$\beta_1 = \beta_2 = \frac{\sqrt{W^2 + 12Rm} - W}{6R},$$

and in the asymmetric firm case, β_1 is given by the real positive solution out of the four possible roots, and β_2 is given by and (9.3.19).

We see that the optimal advertising policy is to spend in proportion to the competitor's market share. *The firm that is in a disadvantageous position fights harder than its opponent and it should succeed in wresting market share from the opponent.* Spending is decreasing in own market share, thus, the advertising-to-sales ratio is higher for the lower share firm.

Table 9.3.2 Comparative Statistics with Asymmetric Firms

Variables	Parameters				
	$c_i,\ c_j$	$\rho_i,\ \rho_j$	$m_i,\ m_j$	δ	$r_i,\ r_j$
α_i	$?, +$	$?, -$	$+, -$	$?$	$-, +$
β_i	$?, +$	$?, -$	$+, -$	$-$	$-, +$
u_i^*	$-, +$	$+, -$	$+, -$	$-$	$-, +$
$V_i(x)$	$?, +$	$?, -$	$+, -$	$?$	$-, +$

Legend: $+=$increase, $-=$decrease, $?=$ambiguous

Many firms do advertising budgeting based on the affordable method, the percentage-of-sales method, and the competitive-parity method. These methods would suggest that the firm with lower market share should spend less on advertising. This is in contradiction to the optimal advertising policy derived here, Prasad and Sethi [153]. In fact, the methods derived in this section have been shown true using empirical methods in Chintagunta and Vilcassim [39]. Table 9.3.2 below is also due to Prasad and Sethi [153], which gives comparative statistics with asymmetric firms.

A comparison of the comparative statics in Table 9.3.1 and Table 9.3.2 shows the following main features. First, due to the additional complexity of the asymmetric case, there are a few more ambiguous effects. However, secondly, it appears that the change in its own parameters have the same effect in the asymmetric case as a change in these parameters had for the symmetric case. This is to be expected since the first order effects likely dominate the second order effects, thus, yielding the same results as in the symmetric case. It becomes clear that a beneficial increase in its own parameters (ρ_i, c_i, m_i, r_i) have a negative effect on the competitor's profits. Finally, the results for the amount of advertising u_i^* are completely unambiguous and follow the same intuition as in the symmetric case. Note that the optimal advertising policy does not depend on the noisiness of the selling environment. This is a consequence of the linear form of the value function.

Next, we will examine the market share paths analytically. Inserting the values of the strategies into the equations of motion (9.3.1), one obtains the following set of equations,

$$dx = \left(\frac{\beta_1\rho_1^2}{2c_1} + \delta - x\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right)dt + \sigma(x)dw, \quad x(0) = x_0, \quad \text{and}$$
$$dy = \left(\frac{\beta_2\rho_2^2}{2c_2} + \delta - y\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right)dt - \sigma(1-y)dw, \quad y(0) = 1 - x_0. \tag{9.3.21}$$

These equations can be rewritten as stochastic integral equations, that is,

$$x(t) = x_0 + \int_0^t \left(\frac{\beta_1\rho_1^2}{2c_1} + \delta - x(s)\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right)ds + \int_0^t \sigma(x)dw, \quad \text{and}$$
$$y(t) = (1-x_0) + \int_0^t \left(\frac{\beta_2\rho_2^2}{2c_2} + \delta - y(s)\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right)ds - \int_0^t \sigma(1-y)dw. \tag{9.3.22}$$

The mean evolution path turns out to be independent of the nature of the stochastic disturbance. That is,

$$E[x(t)] = x_0 + \int_0^t \left(\frac{\beta_1\rho_1^2}{2c_1} + \delta - E[x(s)]\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right)ds, \quad \text{and}$$
$$E[y(t)] = (1-x_0) + \int_0^t \left(\frac{\beta_2\rho_2^2}{2c_2} + \delta - E[y(s)]\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right)ds. \tag{9.3.23}$$

These equations can be expressed as ordinary differential equations in $E[x(t)]$ and $[E[y(t)]$ with the solutions given by

$$E[x(t)] = e^{-\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)t} x_0 + \left(1 - e^{-\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)t}\right) \frac{\frac{\beta_1\rho_1^2}{2c_1} + \delta}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta}, \quad \text{and}$$
$$E[y(t)] = e^{-\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)t} (1-x_0) + \left(1 - e^{-\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)t}\right) \frac{\frac{\beta_2\rho_2^2}{2c_2} + \delta}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta}. \tag{9.3.24}$$

The long run equilibrium market shares $(\bar{x}, \bar{y})$ are obtained by taking the limit as $t \to \infty$ and are given by

$$\bar{x} = \frac{\frac{\beta_1\rho_1^2}{2c_1} + \delta}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta} \quad \text{and} \quad \bar{y} = \frac{\frac{\beta_2\rho_2^2}{2c_2} + \delta}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta}. \tag{9.3.25}$$

Thus, the expected market shares converge to the form resembling the attraction models commonly used in marketing. However, while an attraction model would rate the attractiveness of each firm based on its lower cost, higher productivity of advertising, and higher advertising, it would exclude exogenous market phenomena such as churn.

To further characterize the evolution path, calculate the variance of the market shares at each point in time. A specification of the disturbance function is required for this characterization. We will use $\sigma(x)dw = \sigma\sqrt{x(1-x)}dw$, where σ is a positive constant, and it can be seen that market shares will remain in $(0,1)$.

An application of Itô's formula to equation (9.3.21) provides the following result.

$$d\big(x(t)^2\big) = \left[2x\left(\frac{\beta_1\rho_1^2}{2c_1} + \delta - x\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)\right) + \sigma^2 x(1-x)\right]dt + 2x\sigma\sqrt{x(1-x)}dw. \tag{9.3.26}$$

Rewriting this as a stochastic integral, taking the expected value, and rewriting it as a differential equation, we obtain

$$\frac{dE[x(t)^2]}{dt} = \left(\frac{\beta_1\rho_1^2}{2c_1} + 2\delta + \sigma^2\right)E[x(t)] - \left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 4\delta + \sigma^2\right)E[x(t)^2]. \tag{9.3.26}$$

Inserting the solution for $E[x(t)]$ from (9.3.24), we obtain a first order linear differential equation in the second moment $E[x(t)^2]$, that is,

$$\begin{aligned}&\frac{dE[x^2]}{dt} + \left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 4\delta + \sigma^2\right)E[x^2]\\ &= \frac{\left(\frac{\beta_1\rho_1^2}{2c_1} + 2\delta + \sigma^2\right)\left(\frac{\beta_1\rho_1^2}{2c_1} + \delta\right)}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta} + e^{-\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)t}\\ &\quad\times\left(\left(\frac{\beta_1\rho_1^2}{2c_1} + 2\delta + \sigma^2\right)x_0 - \frac{\left(\frac{\beta_1\rho_1^2}{2c_1} + \delta\right)\left(\frac{\beta_1\rho_1^2}{2c_1} + 2\delta + \sigma^2\right)}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta}\right).\end{aligned} \tag{9.3.27}$$

The solution of (9.3.27) is given by

$$\begin{aligned}E[x(t)^2] &= x_0 e^{-2\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta + \frac{\sigma^2}{2}\right)t}\\ &\quad+ \frac{\left(\frac{\beta_1\rho_1^2}{2c_1} + \delta\right)\left(\frac{\beta_1\rho_1^2}{2c_1} + \delta + \frac{\sigma^2}{2}\right)}{\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta + \frac{\sigma^2}{2}\right)\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)}\left(1 - e^{-2\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta + \frac{\sigma^2}{2}\right)t}\right)\\ &\quad+ \frac{e^{-\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta\right)t} - e^{-2\left(\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta + \frac{\sigma^2}{2}\right)t}}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta + \sigma^2}\\ &\quad\times\left(\left(\frac{\beta_1\rho_1^2}{2c_1} + 2\delta + \sigma^2\right)x_0 - \frac{\left(\frac{\beta_1\rho_1^2}{2c_1} + \delta\right)\left(\frac{\beta_1\rho_1^2}{2c_1} + 2\delta + \sigma^2\right)}{\frac{\beta_1\rho_1^2}{2c_1} + \frac{\beta_2\rho_2^2}{2c_2} + 2\delta}\right).\end{aligned} \tag{9.3.28}$$

We can calculate the convergence of the second moment, as the influence of the initial condition disappears. That is,

$$\lim_{t\to\infty} E[x(t)^2] = \frac{\left(\frac{\beta_1\rho_1^2}{2c_1}+\delta\right)\left(\frac{\beta_1\rho_1^2}{2c_1}+\delta+\frac{\sigma^2}{2}\right)}{\left(\frac{\beta_1\rho_1^2}{2c_1}+\frac{\beta_2\rho_2^2}{2c_2}+2\delta+\frac{\sigma^2}{2}\right)\left(\frac{\beta_1\rho_1^2}{2c_1}+\frac{\beta_2\rho_2^2}{2c_2}+2\delta\right)}. \tag{9.3.29}$$

Written in this form, it becomes clear that when $\sigma = 0$ the expression is just $\bar{x}^2$ so that the variance is appropriately zero in the absence of the stochastic effect. More generally, when $\sigma = 0$, $E[x(t)^2] = (E[x(t)])^2$ holds for all t. For $\sigma > 0$ the standard deviation of the solution $x(t)$ is

$$\sqrt{E[x(t)^2] - (E[x(t)])^2}.$$

Similar results can be obtained for the second firm, as discussed in Prasad and Sethi [153]. We present the results for the mean and variance of the long-run market share in the following proposition.

Proposition 9.3.2. *For the advertising game described by equation* (9.3.4)*, we have*

(a) *The mean market shares in the long run are given by* (9.3.25)*,*

$$\bar{x} = \frac{\frac{\beta_1\rho_1^2}{2c_1}+\delta}{\frac{\beta_1\rho_1^2}{2c_1}+\frac{\beta_2\rho_2^2}{2c_2}+2\delta} \quad \textit{and} \quad \bar{y} = \frac{\frac{\beta_2\rho_2^2}{2c_2}+\delta}{\frac{\beta_1\rho_1^2}{2c_1}+\frac{\beta_2\rho_2^2}{2c_2}+2\delta}.$$

(b) *The variance of the market shares in the long run are obtained from* (9.3.25) *and* (9.3.29) *as* $E[x(t)^2] - (E[x(t)])^2$ *and for both firms are given by*

$$\frac{\left(\frac{\beta_1\rho_1^2}{2c_1}+\delta\right)\left(\frac{\beta_2\rho_2^2}{2c_2}+\delta\right)\frac{\sigma^2}{2}}{\left(\frac{\beta_1\rho_1^2}{2c_1}+\frac{\beta_2\rho_2^2}{2c_2}+2\delta+\frac{\sigma^2}{2}\right)\left(\frac{\beta_1\rho_1^2}{2c_1}+\frac{\beta_2\rho_2^2}{2c_2}+2\delta\right)^2}.$$

In Prasad and Sethi [153], a particular case is analyzed and also it is shown that the densities of the stationary distributions of the market shares are given by a Beta probability density. We will not discuss these results here.

Due to the results of this section, it is particularly important to note that the morphing of the Vidale-Wolfe sales decay term into decay caused by competitive advertising and non-competitive 'churn' that acts to equalize market shares in the absence of advertising. We have presented closed-loop Nash equilibria for symmetric as well as asymmetric competitors. For all cases, explicit solutions and comparative statics were presented. The analysis suggests another counter-intuitive result that brands with smaller market share should spend more aggressively on advertising than larger brands, Prasad and Sethi [153]. This finding is contrary to the conventional practice of some firms to maintain share-of-voice proportional to market share (which implies smaller brands should spend less aggressively). Thus, the result of this section suggests that managers should re-consider the validity of their decision rules in ever changing dynamic markets. However, one must be careful to limit the conclusions of optimality to only those markets for which the model applies. For instance, advertisement expenditure or advertising policies that are optimal in a monopoly setting would not be optimal in a competitive setting.

References

[1] S.I. Aihara and A. Bagchi, Linear-quadratic stochastic differential games for distributed parameter systems. Pursuit-evasion differential games, Comput. math. Appl., 13, 1987, 247–259.

[2] R. Ardanuy, Stochastic diferential games:the linear quadratic zero sum case, Sankhy: The Indian Journal of Statistics, Vol. 57, Series A, Pt. 1, 1995, 161–165.

[3] E. Altman, and O. Pourtallier, Approximating Nash equilibria in Nonzero-sum games, International Game Theory Review, Vol. 2, Nos. 2&3, 2000, 155–172.

[4] R. Atar, and A. Budhiraja, Stochastic differential game for the inhomogeneous $\infty -$ Laplace equation, The Annals of Probability, Vol. 38, No. 2, 2010, 498–531.

[5] R. Bafico, On the definition of stochastic differential games and the existence of saddle points, Ann. Mat. Pura Appl., 96, 1972, 41–67.

[6] M. Bardi, M. Falcone, and P. Soravia. Numerical methods for pursuit evasion games via viscosity solutions. In M. Bardi, T.E.S. Raghavan, and T. Parthasarathy, editors, *Stochastic and Differential Games: Theory and Numerical Methods*. Birkhäuser, Boston, 1998.

[7] M. Bardi, T.E.S. Raghavan, and T. Parthasarathy (Eds.), Stochastic and differential games: Theory and numerical methods, Birkhäuser, 1999.

[8] G.K. Basak, M.K. Ghosh, and D. Mukherjee, Equilibrium and stability of a stock market game with big traders, Diffferential Equations and Dynamical Systems, 17, No. 3, 2009, 283–299.

[9] T. Basar, Lecture Notes on Non-Cooperative Game Theory, July 26, 2010; `http://www.hamilton.ie/ollie/Downloads/Game.pdf` [last accessed: 4-11-2011].

[10] T. Basar, Existence of unique Nash equilibrium solutions in nonzero-sum stochastic differential games, Differential games and control theory II, Proceedings of second Kingston conference, 1976, 201–228.w33

[11] T. Basar, Nash equilibrium of risk-sensitive nonlinear differential games, J. Optim. Theory Appl., 100, 1999, 479–498.

[12] T. Basar, A continuous-time model of monetary policy and inflation: a stochastic differential game, Decision processes in economics (Modena, 1989), 3-17, Lecture Notes in Econom. and Math. Systems, 353, Springer, 1991.

[13] T. Basar, On the existence and uniqueness of closed-loop sampled data Nash controls in linear-quadratic stochastic differential games, Optimal Techniques, K. Iracki, K. Malonowski, and S. Walukiewicz (Eds.), Lecture Notes in Control and Information Sciences, Springer Verlag, Vol. 22, 1980, 193–203.

[14] T. Basar and P. Bernhard, $H_{\{\infty\}}$-Optimal control and related minimax design problems: A dynamic game approach, 2nd Edition, Birkhäuser, 1995.

[15] T. Basar and A. Haurie, Feedback equilibria in differential games with structural and modal uncertainties, in Advances in Large Scale Systems, vol. 1, Editor: J.B. Cruz, Jr, 1984, 163–301.

[16] T. Basar and A. Haurie (Eds.), *Advances in dynamic games and applications*, Birkhäuser, 1994.

[17] T. Basar and G.J. Olsder, *Dynamic noncooperative game theory*, 2nd Edition, Academic Press, 1995.

[18] A. Basu and M.K. Ghosh, Stochastic differential games with multiple modes and applications to portfolio optimization, Stochastic Analysis and Applications, 25, 2007, 845–867.

[19] R.D. Behn and Y.C. Ho, On a class of linear stochastic differential games, IEEE Trans. Automatic Control, AC-13, 1968, 227–240.

[20] M. Benaïm, and J.W. Weibull, Deterministic approximation of stochastic evolution in games, `http://www.bu.edu/econ/workingpapers/papers/Jorgen`, 2002. [Last assessed 8-23-2011].

[21] V.E. Benes, Existence of optimal strategies based on a specific information for a class of stochastic decision problems, SIAM J. Control, 8, 1970, 179–188.

[22] A. Bensoussan and J.L. Lions, Stochastic differential games with stopping times, Differential games and control theory II, Proceedings of second Kingston conference, 1976, 377–399.

[23] A. Bensoussan and A. Friedman, Nonlinear variational inequalities and differential games with stopping times, J. Functional Analysis, 16, 1974, 305–352.

[24] A. Bensoussan and A. Friedman, Nonzero-sum stochastic differential games with stopping times and free boundary problems, Trans. Amer. Math. Soc., 231, 1977, 275–327.

[25] L.D. Berkovitz, Two person zero sum differential games: an overview, The theory and application of differential games, Editor: J.D. Grote, Proceedings of the NATO advanced study institute held at the University of Warwick, Coventry, England, 27 August-6 September, 1974. 359–385.

[26] L.D. Berkowitz, L.D. A variational approach to differential games, RAND Report RM-2772, 1961.

[27] L.D. Berkowitz, and W.H. Flemimg, On differential games with integral payoff, Contributions to the Theory of Games III. Princeton, N.J.: Princeton University Press, 1957, 413–435.

[28] A.T. Bharaucha_Reid, Random algebraic equations, Probabilistic Methods in Applied Mathematics, Vol. 2, 1970, 1–52.

[29] A.T. Bharaucha_Reid, On the theory of random equations, Proceedings of Sixth Symposium in Applied Mathematics, 1964, 40–69.

[30] V.S. Borkar, Optimal control of diffusion processes, Pitman research Notes in Math. Series 203, Logman, Harlow, 1989.

[31] V.S. Borkar and M.K. Ghosh, Stochastic differential games: An occupation measure based approach, Journal of Optimization Theory and Applications, 73, 1992, 359–385.

[32] V.S. Borkar and M.K. Ghosh, Stochastic games: An occupation measure based approach, Preprint.

[33] Browne, S., Stochastic differential portfolio games, J. Applied Probability, 37, 2000, 126–147.

[34] A.E. Bryson, Jr., and Y.C. Ho, *Applied Optimal Control*, Waltham, Mass: Blaisdell Pub. Co., 1969.

[35] Buckdahn Rainer, Pierre Cardaliaguet, and Catherine Rainer, Nash equilibrium payoffs for nonzero-sum stochastic differential games, SIAM J. on Control and Optimization, Vol. 43, No.2, 2004, 624–642.

[36] R. Buckdahn, P. Cardaliaguet, and C.Rainier. Nash equilibrium payoffs for nonzero-sum stochastic differential games. SIAM J. Control and Optimzaton, 43, 2002, 624–642.

[37] J.H. Case, Toward a theory of many player differential games, SIAM J. Control, Vol. 7, 1969, 179–197.

[38] J.H. Case, A differential games in economics, Management Science, vol. 17, 1971, 394–410.

[39] P.K. Chintagunta, and N.J. Vilcassim, An Empirical Investigation of Advertising Strategies in a Dynamic Duopoly, *Management Science* **38** (9), 1992, 1230—1244

[40] R.J. Chitashvili and N.V. Elbakidze, Optimal stopping by two players, in Statistics and Control of Stochastic Processes, Steklov seminar, N.V. Krylov, R.Sh. Lipster and A.A. Novikov (Eds.), 1984, 10–53.

[41] E. Cockayne, Plane pursuit with curvature constraints, SIAM J. Applied Mathematics, vol. 15, 1967, 1511–1516.

[42] B. Conolly, and A. Springall, The overall outcome of a certain Lanchester combat model, Interim Technical report 67/1, Department of Statistics and Statistical Laboratory, Virginia Polytechnic Institute, Blacksburg, 1967.

[43] M.G. Crandall and P.L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. AMS, 277, 1983, 1–42.

[44] W.B. Davenport, Signal to noise ratios in band pas limiters, J. Appl. Phys., 24, 1953.

[45] G.J. Disher (Ed.), New trends in Dynamic games and applications, Annals of the ISDG, 3, Birkhauser, 1995.

[46] J.P. Dix, Game-theoretic applications, IEEE Spectrum, 1968, 108–117.

[47] R.J. Elliott, The existence of optimal strategies and saddle points in stochastic differential games, Differential games and applications; Proceedings of a workshop Enschede, 1977; Lec. notes in control and Information Sciences, 3, 1977; pp. 123–135.

[48] R.J. Elliott. Stochastic differential games and alternate play, Proceedings of the International Symposium on Control Theory at I.N.R.I.A. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, 107 (1974): 97–106.

[49] R.J. Elliott. Introduction to differential games II. Stochastic games and parabolic equations, The Theory and Application of Differential Games (J. Grote and D. Reidel, eds.) Dordrecht, Holland (1975): 34–43.

[50] R.J. Elliott. The existence of value in stochastic differential games, S.I.A.M. Journal of Control 14 (1976): 85–94.

[51] R.J. Elliott and M.H.A. Davis, Optimal play in a stochastic differential game, SIAM J. Control Opt. 19, 1981, 543–554.

[52] R.J. Elliott and N.J. Kalton. *Existence of value in differential games*, Mem. AMS, **126**. Amer. Math. Soc, Providence, RI, 1974.

[53] J.C. Engwerda, W.A. van den Broek, and J.M. Schumacher, Feedback Nash eqilibria in uncertain infinite time horizon differential games, `http://arno.uvt.nl/show.cgi?fid=4752`, 2011.

[54] G.M. Erickson, G.M., Advertising Competition in a Dynamic Oligopoly with Multiple Brands, Operations Research 57 (5), 2009, 1106–1113.

[55] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.

[56] L.C. Evans and P.E. Souganidis, Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations, Ind. Univ. math. J., 33, 1984, 773–797.

[57] C.O. Ewald, The Malliavin calculus and stochastic differential games with information asymmetry, PROCEEDINGS OF THE SECOND CONFERENCE ON GAME THEORY AND APPLICATIONS, Hongwei Gao, Leon A. Petrosyan, eds., pp. 26–30, World Academic Union Ltd, September 2007.

[58] K. Fan, Fixed points and minimax theorems in locally convex topological linear spaces, Proc. Nat. Acad. Sci., U.S.A. 38, 1952, pp. 121–126.

[59] J.A. Filar and K. Vrieze, Competitive Markov Decision Processes, Springer-Verlag, 1997.

[60] W.H. Fleming, Generalized solutions in optimal stochastic control, Proc. URI Conf. on Control, 1982, 147–165.

[61] W.H. Fleming, The convergence problem for differential games, J. Math. Analysis and Applications, 3, 1961, 102–116.

[62] W.H. Fleming and D.H. Hernandez, On the value of stochastic differential games, Communications on Stochastic Analysis, Vol. 5, No. 2, 2011, 341–351.

[63] W.H. Fleming and H.M. Soner, Controlled Markov processes and viscosity solutions, Springer-Verlag, 1993.

[64] W.H. Fleming and P.E. Souganidis, On the existence of value functions of two-player, zero-sum stochastic differential games, Indiana Univ. Math. J., 38, 1989, 293–314.

[65] W.H. Fleming and P.E. Souganidis, Two player, zero sum stochastic differential games, Analyse Mathématque et applications, gauthier-Villars, 1988, 151–164.

[66] A. Friedman, Computation of saddle points for differential games of pursuit and evasion, Archive for Rational Mechanichs and Analysis, vol. 40, 1971, 79–119.

[67] A. Friedman, Differential games, Wiley, 1971.

[68] A. Friedman, Stochastic differential equations and applications, Vol.2, Academic Press, 1976.

[69] A. Friedman, Stochastic differential games, J. Differential Equations, 11, 1972, 79–108.

[70] S.D. Gaidov, Nash Equilibrium in stochastic differential games, Computers and MAthematics with applications, 12A, 1986, 761–768.

[71] S.D. Gaidov, Z-equilibrium in Many-player stochastic differential games, ARCHIVUM MATHEMATICUM (BRNO), Tomus 29, 1993, 123–133.

[72] S.D. Gaidov, On the Nash-bargaining solution in stochastic differential games, Serdica, 16, 1990, 120–125.

[73] S. D. Gaidov, Mean-square strategies in stochastic differential games, Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform., 18, 1989, 161–168.

[74] L. Gawarecki and V. Mandrekar, On the existance of weak variational solutions to stochastic differentialequations, Communications on Stochastic Analysis, Vol. 4, No. 1, 2010, 1–20.

[75] M.K. Ghosh, A. Araposthatis, and S.I. Marcus, Optimal control of switching diffusions with application to flexible manufacturing systems, SIAM J. Control Optim., 31, 1993, 1183–1204.

[76] M.K. Ghosh and S.I. Marcus, Stochastic differential games with multiple modes, Stochastic Analysis and Applications, 16, 1998, 91–105.

[77] M.K. Ghosh and K.S. Kumar, Zero-sum stochastic differential games with reflecting diffusions, Mat. Apl. Comput., 16, 1997, 237–246.

[78] X. Guo and O. Hernández-Lerma, Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates, Bernoulli, Vol. 11, No. 6, 2005, 1009–1029.

[79] P. Hagedorn, H.W. Knobloch, and G.J. Olsder (Eds.), *Differential games and applications*, Proceedings of a workshop Enschede, 1977.

[80] S. Hamadène, Backward-forward SDE's and stochastic differential games, Stochastic Process. Appl., 77, 1998, 1–15.

[81] S. Hamadène and J.P. Lipeltier, Backward equations, stochastic control and zero-sum stochastic differential games, Stochastics Stochastics Rep., 54, 1995, 221–231.

[82] R. P. Hämalainen and H. Ehtamo (Eds.), *Advances in Dynamic Games and Applications*, Annals of the ISDG, Vol. 1, Birkhauser, 1994.

[83] R. P. Hämalainen and H. Ehtamo (Eds.), Advances in Dynamic Games and Applications, Annals of the ISDG, Vol. 1, Birkhauser, 1994.

[84] R. P. Hämalainen and H. Ehtamo (Eds.), Dynamic games in economic analysis, Lecture notes in control and Information sciences, 157, Springer-Verlag, 1991.

[85] R. P. Hämalainen and H. Ehtamo (Eds.), Differential games-Developments in modelling and computation, Lecture Notes in Control and Information Sciences, Vol. 156, Springer-Verlag, 1991.

[86] A. Haurie, Stochastic differential games in economic modeling, Lecture Notes in Control and Inform. Sci., 197, Springer, 1994, 90–108.

[87] Y.C. Ho, Optimal terminal maneuver and evasion strategy, SIAM J. Control, 4, 1966, 421–428.

[88] Y.C. Ho, On maximum principle and zero-sum stochastic differential games, JOTA, 13, 1974.

[89] Y.C. Ho, A.E. Bryson, Jr., and S. Baron, "Differential games and optimal pursuit-evasion strategies, " IEEE Trans. Automatic Control, vol. AC-10, 1965, pp. 385–389.

[90] R. Isaacs, Differential Games I, II, III, IV, The RAND Corporation, Research Memoranda RM-1391, RM-1399, RM-1411, RM-1486 (1954).

[91] R. Isaacs, *Differential Games*, John Wiley and Sons, New York, 1965.

[92] H. Ishii, On uniqueness and existence of viscosity solutions for fully nonlinear second order elliptic pde, Comm. Pure Appl. Math., 42, 1989, 14–45.

[93] S. Jørgensen and D.W.K. Yeung, Stochastic differential game model of a common property fishery, J. Optim. Theory Appl., 90, 1996, 381–403.

[94] R.E. Kalman and R.S. Bucy, New results in linear filtering and prediction theory, Trans. ASME, J. Basic Engrg., ser. D, 83, 1961, 95–108.

[95] N.J. Kalton, N.N. Krasovskii, and A.I. Subbotin, *Positional differential games*, Nauka, 1974, and Springer, 1988.

[96] V. Kaitala, Equilibria in a stochastic resource management game under imperfect information, European Journal of Operations Research, 71, 1993, 439–453.

[97] D.L. Kelendzeridze, Theory of optimal pursuit strategy, Soviet Mathematics, vol. 2, 1961, 654–656.

[98] N.N. Krasovskii, Game problems in Dynamics, I, Engineering Cybernetics, No. 5. 1969, 1–10.

[99] N.N. Krasovskii, Game problems in Dynamics, II, Engineering Cybernetics, No. 1. 1970, 1–11.

[100] N.N. Krasovskii and A.I. Subbotin, *Game theoretical control problems*, Springer-Verlag, 1988.

[101] N.V. Krylov, *Controlled Diffusion Processes*, Springer, New York, 1980.

[102] H.J. Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York, 1977.

[103] H.J. Kushner, *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*, MIT Press, 1984.

[104] H.J. Kushner, *Weak convergence methods and singularly perturbed stochastic control and filtering problems*, Birkhauser, 1990.

[105] H.J. Kushner, *Heavy Traffic Analysis of Controlled Queueing and Communication Networks*. Springer-Verlag, Berlin and New York, 2001.

[106] H.J. Kushner, A numerical methods for stochastic control problems in continuous time, SIAM J. Control and Optimization, Vol. 28, No. 5, 1990, 999–1048.

[107] H.J. Kushner, Numerical approximations for stochastic differential games: the ergodic case, SIAM Journal on Control Optim., vol. 42, 2004, 1911–1933.

[108] H.J. Kushner, Numerical methods for stochastic differential games. *SIAM J. Control Optim.*, vol. 41, 2002, 457–486.

[109] H.J. Kushner, Numerical Approximations for Nonzero-Sum Stochastic Differential Games, SIAM J. Control Optim. Vol. 46, 2007, 1942–1971.

[110] H.J. Kushner and S.G. Chamberlain, On stochastic differential games: sufficient conditions that a given strategy be a saddle point, and numerical procedures for the solution of the game, J. Math. Anal. Appl. 26, 1969, 560–575.

[111] H.J. Kushner and S.G. Chamberlain, Finite state stochastic games: existence theorems and computational procedures", IEEE Trans. Automat. Control, 14, 1969, 248–255.

[112] H.J. Kushner and P.G. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Springer-Verlag, 1992, Second edition, 2001.

[113] H.J. Kushner and K.M. Ramachandran, Optimal and approximately optimal control policies for queues in heavy traffic, SIAM J. Control Optim., 27, 1989.

[114] H.J. Kushner and W.J. Runggaldier, Nearly optimal state feedback controls for stochastic systems with wideband noise disturbances, SIAM J. Control Opt., 25, 1987.

[115] H.J. Kushner and W. Runggaldier, Filtering and control for wide bandwidth noise driven systems, LCDS Report #86-8, 1986.

[116] T.G. Kurtz, Semigroups of conditional shifts and approximations of Markov processes, Annals of Probability, 4, 1975.

[117] R. Lachner , M. H. Breitner , H. J. Pesch, Three-Dimensional Air Combat: Numerical Solution of Complex Differential Games, Annals of the International Society of Dynamic Games: New Trends in Dynamic Games and Applications, 1996.

[118] R. Lachner , M. H. Breitner , H. J. Pesch, Efficient Numerical Solution of Differential Games with Application to Air Combat , Report No. 466, Deutsche Forschungsgemeinschaft, Schwerpunkt "Anwendungsbezogene Optimierung und Steuerung, 1993.

[119] J. Lehoczky, S. Shreve, Absolutely continuous and singular control, Stochastics, 17, 1986.

[120] G. Leitmann, Multicriteria decision making and differential games, Plenum Press, 1976.

[121] C. K. Leong, W. Huang, A Stochastic Differential Game of Capitalism, Journal of Mathematical Economics, 2010.

[122] J. Lewin, *Differential games*, Springer, 1994.

[123] D. Li, and J. B. Cruz, A Two-Player Stochastic Pursuit-Evasion Differential Game, `http://www2.ece.ohio-state.edu/~lido/CDC07_game.pdf`, [last accessed April 5, 2011].

[124] P.L. Lions and P.E. Souganidis, Differential games, optimal control and directional derivatives of viscosity solutions of Bellman, s and Isaacs equations, SIAM J. of Control and Optimization, 23, 1985, 566–583.

[125] P.L. Lions and P.E. Souganidis, Differential games, optimal control and directional derivatives of viscosity solutions of Bellman, s and Isaacs equations II, SIAM J. of Control and Optimization, 24, 1986, 1086–1089.

[126] P.L. Lions and P.E. Souganidis, Viscosity solutions of second-order equations, stochastic control and stochastic differential games, Stochastic differential systems, stochastic control theory and applications, W. Fleming and P.L. Lions (Eds.), Springer-Verlag, 1988, 293–309.

[127] R.SH. Liptser, W.J. Runggaldier, and M. Taksar, Deterministic approximation for stochastic control problems, SIAM J. Control Opt. 34, 1996.

[128] R.S. Lipster and A.N. Shiryaev, *Statistics of Random Processes*, Springer-Verlag, 1977.

[129] P.T. Liu, On a problem of stochastic differential games, Ph.D. dissertation, State University of New York, Stony Brook, 1968.

[130] D. Lund and B. Øksendal (Eds.) Stochastic models and option values: Applications to resources, environment and investment problems, North-Holland, 1991.

[131] P. Mannucci, Nonzero-sum stochastic differential games with discontinous feedback, SIAM J. Control Optim., Vol. 43, No. 4, 2004, 1222–1233.

[132] L. Meier, III, A new technique for solving pursuit evasion differential games, IEEE Trans. Automatic Control, vol. AC-14, 1969, 352–359.

[133] R.C. Merton, Theory of finance from the perspective of continuous time, Journal of Financial and Quantitative Analysis, 1975, 659–674.

[134] S.P. Meyn and R.I. Tweedie. Markov Chains and Stochastic Stability. Springer-Verlag, Berlin and New York, 1994.

[135] H. Morimoto and M. Ohashi, On linear stochastic differential games with average cost criterions, J. Optim. Theory Appl., 64, 1990, 127–140.

[136] T. Morozan, Stability of some linear stochastic systems, J. Differential equations, vol. 3, 1967, 153–169.

[137] L. Mou, Two-person zero-sum linear quadratic stochastic differential games by a Hilbert space method, J. of Industrial and Management Optimization, vol. 2, No. 1, 2006.

[138] R. B. Myerson, Nash equilibrium and the history of economic theory, 1999, `http://docs.google.com/viewer?a=v&q=cache:nq-Kbgqw1XAJ:citeseerx.ist.psu.edu/viewdoc/download%3Fdoi%3D10.1.1.130.7955%26rep%3Drep1%26type%3Dpdf+applications+of+nash+equilibrium&hl=en&gl=us&pid=bl&srcid=ADGEESgDQBCn-2E8bo4TNaVYHNM7xBij7GBa-D6H1ba0drlNZ67FG_LYlJFwG7CsP3TbnQMEo1ItD-hzLywlIVS6IPZHhtnXpChS1FVGUVEW0G4z6Fz49WGIedn4ZqUHFZ4PeLCCoraK&sig=AHIEtbRA1vNQH8vu1RlQTybWsIl1rjXCVQ`, *last accesed, July 9, 2011.*

[139] John Nash, Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences* 36(1), 1950, 48–49.

[140] John Nash, Non-Cooperative Games, *The Annals of Mathematics* 54(2), 1951, 286–295.

[141] W.G. Nicholas, Stochastic differential games and control theory, Thesis, Virginia Polytechnic Institute and State University, Blackburg, Virginia, 1971.

[142] M.S. Nikol'skii, Nonstationary linear differential games, Vestnik Moskov. Univ. ser I Mat. Meh., vol 24, 1969, 65–73.

[143] M. Nisio, Stochastic differential games and viscosity solutions of Isaacs equations, Nagoya Math. J., 110, 1988, 163–184.

[144] M. Nisio, On infinite-dimensional stochastic differential games, Osaka J. math., 35, 1998, 15–33.

[145] G.J. Olsder, On observation costs and information structures in stochastic differential games, Differential games and applications; Proceedings of a workshop Enschede, 1977; Lec. notes in control and Information Sciences, 3, 1977; pp. 172–185.

[146] G.J. Olsder (Ed.), *New tends in dynamic games and applications*, Birkäuser, 1995.

[147] L.A. Petrosyan, *Differential Games of Pursuit*, World Scientific, Singapore, 1993.

[148] L. Petrosyan, and N.A. Zaccour, Time-consistent Shapley value allocation of pollution cost reduction, Journal of Economic Dynamics and Control, 27, No. 3, 2003, 381–398.

[149] C. Plourde, and D. Yeung, Harvesting of a transboundary replenishable fish stock: a non-cooperative game solution, Marine Resource Economics, 6, 1989, 54–71.

[150] L.S. Pontyagin, On the theory of differential games, Uspekhi Mat. Nauk, 21, 1966, 219–274.

[151] L.S. Pontryagin, Linear Differential games, Mathematical Theory of Control, A.V. Balakrishnan and L.W. Neustadt, Eds., NEW YORK: Academic Press, 1967, 330–334.

[152] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkerlidze, and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, New York, Inter-Science, 1962.

[153] A. Prasad, and S.P. Sethi, Competitive advertising under uncertainity: A stochastic differential game approach, Journal of Optimization Theory and Applications **123** (1), 2004, 163–185.

[154] B.N. Pshenichniy, Linear differential games, Mathematical Theory of Control, A.V. Balakrishnan and L.W. Neustadt, eds., New York: Academic Press, 1967, 335–341.

[155] M.L. Puterman, *Markov Decision Processes*, Wiley, 1994.

[156] R.K. Ragade, and I.G. Sarma, A game theoretic approach to optimal control in the presence of uncertainty, IEEE Trans. Automatic Control, vol. AC-12, 1967, 395–402.

[157] T. Raivio and H. Ehtamo, On the numerical solution of a class of pursuit-evation games, in Advances in Dynamic Games and Applications, Birkhäuser, 2000.

[158] K.M. Ramachandran, Stochastic differential games with a small parameter, Stochastics and Stochastics Reports, 43, 1993, 73–91.

[159] K.M. Ramachandran, N-Person stochastic differential games with wideband noise perturbation , Journal of Combinatorics, & Information System Sciences, Vol. 21, Nos. 3-4, 1996, pp. 245–260

[160] K.M. Ramachandran, Weak convergence of partially observed zero-sum stochastic differential games, Dynamical systems and Applications, Vol 4, No. 3, 1995, 329–340.

[161] K.M. Ramachandran, Discrete parameter singular control problem with state dependent noise and non-smooth dynamics, Stochastic Analysis and Applications, 12, 1994, 261–276.

[162] K.M. Ramachandran, Stochastic differential games and applications, *Hand Book of Stochastic Analysis and Applications*, Editors: D. Kannan and V. Lakshmikantham, Marcel Dekker, Inc., pp. 475–534, Chapter 8, October, 2001.

[163] K.M. Ramachandran and A.N.V. Rao, Deterministic approximation to two person stochastic game problems, to appear in Dynamics of Continuous, Discrete and Impulsive Systems, 1998.

[164] K.M. Ramachandran and A.N.V. Rao, N-person stochastic differential games with wideband noise perturbations: Pathwise average cost per unit time problem, Preprint, 1999.

[165] K.M. Ramachandran and G.Yin, Nearly optimal state feedback controls for delay differential equations with a small parameter, Journal of Mathematical Analysis and Applications, Vol. 172, No. 2, 1993, 480–499.

[166] I.B. Rhodes and D.G. Luenberger, Differential games with imperfect state information, IEEE Trans. Automatic control, AC-14, 1969, 29–38.

[167] I.B. Rhodes and D.G. Luenberger, Stochastic differential games with constrained state estimators, IEEE Trans. on Automatic Control, AC-14, 1969, 476–481.

[168] A.E. Roth (Ed.), *Game-Theoretic models of bargaining*, Cambridge, 1985.

[169] A.E. Roth, Bargaining experiments, Handbook of experimental Economics, J. Kagel and A.E. Roth (Eds.), Princeton University Press, 1995, 253–348.

[170] E. Roxin, C.P. Tsokos: On the Definition of a Stochastic Differential Game, Mathematical Systems Theory 4(1), 1970, 60–64.

[171] Y. Sakawa, On linear differential games, Mathematical Theory of Control, A.V. Balakrishnan and L.W. Neustadt, eds., new York: Academic Press, 1967, 373–385.

[172] Y. Sakawa, Solution of linear pursuit-evasion games, SIAM J. Control, vol.8, 1970, 100–112.

[173] P.A. Samuelson, Rational theory of warrant pricing, Industrial management Review, 6, 1965, 13–31.sam]

[174] L.S. Shapley, A value for n-person games, In Controbutions to the Theory of Games, Prinston University Press II, 1953, 307–317.

[175] L.S. Shapley, Stochastic games, Proceedings of the National Academy of Science U.S.A., 39, 1953, 1095–1100.

[176] K. Shell, The theory of Hamiltonian dynamical systems, and an application to economics, The Theory and Application of Differential Games (J.D. Grote, eds.) D. Reidel Publishing Company, 1975, 189–199.

[177] R. Sircar, Stochastic Differential Games and Applications to Energy and Consumer Goods Markets, `http://www.impa.br/opencms/pt/eventos/extra/2010_rio/attach/ronnie_sircar.pdf`, last accessed: July 9, 2011.

[178] C.W. Smith, Jr, Option Pricing: A review, Journal of Financial Economics, 3, 1976, 3–51.

[179] G. Sorger, Competitive Dynamic Advertising: A Modification of the Case Game, *Journal of Economic Dynamics and Control* **13** (1): 1989, 55–80

[180] P.E. Souganidis, Approximation schemes for viscosity solutions of Hamilton-Jacobi equations with applications to differential games, J. of Nonlinear Analysis, T.M.A., 9, 1985, 217–257.

[181] P.E. Souganidis, Two player, zero-sum differential games and viscosity solutions, Stochastic and differential games: Theory and numerical methods, M. Bardi, T.E.S. Raghavan, and T. Parthasarathy (Eds.), Birkhäuser, 1999, 69–104.

[182] J.L. Speyer, A stochastic differential game with controllable statistical parameters, IEEE Trans. Systems Sci. Cybernetics SSC-3, 1967, 17–20.

[183] J.L. Speyer, S. Samn, and R. Albanese, A stochastic differential game theory approach to human operators in adversary tracking encounters, IEEE Trans. Systems

Man Cybernet., 10, 1980, 755–762.

[184] F.K. Sun and Y.C. Ho, Role of information in the stochastic zero-sum differential game, Multicriteria decision making and Differential games, G. Leitmann (Ed.), Plenum Press, 1976.

[185] A.W. Starr, and Y.C. Ho, Nonzero-sum differential games, J. Optimization Theory and Applications, vol. 3, 1969, 184–206.

[186] A.W. Starr, and Y.C. Ho, Further properties of nonzero-sum differential games, J. Optimization Theory and Applications, vol. 3, 1969, 207–219.

[187] L. Stetner, Zero-sum Markov games with stopping and impulsive strategies, Appl. Math. Optim., 9, 1982, 1–24.

[188] D. W. Stroock and S. R. S. Varadhan. On degenerate elliptic and parabolic operators of second order and their associated diffusions. Comm. Pure Appl.Math., Vol. 25, 1972, 651–713.

[189] A. Swiech, Risk-sensitive control and differential games in infinite dimensions, preprint, 1999.

[190] A. Swiech, Another approach to existence of value functions of stochastic differential games, preprint, 1996.

[191] K. Szajowski, Markov stopping games with random priority, Zeitschrift für Operations research, , 1993, 69–84.

[192] M. Tidball. Undiscounted zero-sum differential games with stopping times. In G.J. Olsder, editor, *New Trends n Dynamic Games and Applications*. Birkhäuser, Boston, 1995.

[193] M. Tidball, and E. Altman, Approximations in dynamic zero-sum games I, SIAM Journal on Control and Optimization, Vol. 34, 311–328, 1996.

[194] M. Tidball and R.L.V. González, Zero-sum differential games with stopping times: Some results and about its numerical resolution. In T. Basar and A. Haurie, editors, *Advances n Dynamic Games and Applications*. Birkhäuser, Boston, 1994.

[195] M. Tidball, O. Pourtallier, and E. Altman, Approximations in dynamic zer-sum games II, SIAM Journal on Control and Optimization, Vol. 35, No.6, 1997, 2101–2117.

[196] C.P. Tsokos, On a stochastic integral equation of Volterra type, Math. Systems Theory, vol. 3, 1969, 222–231.

[197] K. Uchida, On existence of a Nash equilibrium point in N-person nonzero sum stochastic differential games, SIAM J. Control Optim., 16, 1978, 142–149.

[198] P.P. Varaiya, N-person stochastic differential games, The Theory and Application of Differential Games (J. Grote and D. Reidel, eds.) Dordrecht, Holland (1975): 97–107.

[199] P.P. Varaiya, N-player stochastic differential games, SIAM J. Control Optim., 4, 1976, 538–545.

[200] P.P. Varaiya and J. Lin, Existence of saddle points in differential games, S.I.A.M. Jour. Control, 7 (1969), 141–157.

[201] J. VonNeumann and O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press, 1944.

[202] A.Ju. Veretennikov, On strong solution and explicit formulas for solutions of stochastic integral equations, Math. USSR-Sb. 39, (1981), pp. 387–403.

[203] T.L. Vincent, An evolutionary game theory for differential equation models with reference to ecosystem management, Advances in dynamic games and applications, T. Basar and A. Haurie (Eds.), Birkhäuser, 1994, 356–374.

[204] M.L. Vidale, and H.B. Wolfe, An Operations Research study of sales response to advertising, *Operations Research* 5, 1957, 370–381.

[205] B. Wernerfelt, Uniqueness of Nash equilibrium for linear-convex stochastic differential games, J. Optim. Theory Appl., 53, 1987, 133–138.

[206] S. Wan, Stochastic differential portfolio games with Duffie-Kan interest rate, Kybernetes, Vol. 39 Iss: 8, 2010, pp. 1282–1290.

[207] W. Willman, Formal solution of a class of stochastic differential games, IEEE Trans. on Automatic Control, AC-14, 1969, 504–509.

[208] Y. Yavin, The numerical solution of three stochastic differential games, Comput. Math. Appl., 10, 1984, 207–234.

[209] Y. Yavin, Computation of Nash equilibrium pairs of a stochastic differential game, Optimal Control Appl. methods, 2, 1981, 443–464.

[210] Y. Yavin, Computation of suboptimal Nash strategies for a stochastic differential game under partial observation, Internat. J. systems Sci., 13, 1982, 1093–1107.

[211] Y. Yavin, Applications of stochastic differential games to the suboptimal design of pulse motors. Pursuit-evasion differential games, III, Comput. Math. Appl., 26, 1993, 87–95.

[212] Y. Yavin and R de Villiers, Application of stochastic differential games to medium-range air-to-air missiles, J. Optim. Theory Appli., 67, 1990, 355–367.

[213] D. Yeung, A feedback Nash equilibrium solution for noncooperative innovations in a stochastic differential framework, Stochastic Anal. Appl., 9, 1991, 195–213.

[214] D.W.K. Yeung, A stochastic differential game of Institutional Investor speculation, J. Optim. Theory Appli., 102, 1999, 463–477.

[215] D.W.K. Yeung, A differential game of industrial pollution management, Annals of Operations Research, 37, 1992, 297–311.

[216] D.W.K. Yeung and M.T. Cheung, Capital accumulation subject to pollution control: a differential game with a feedback Nash equilibrium, Advances in dynamic games and applications, T. Basar and A. Haurie (Eds.), Birkhäuser, 1994, 289–300.

[217] T. Yoshikawa, An example of stochastic multistage games with noisy state observations, IEEE Trans. Automatic Control, vol. AC-15, 1970, 455–458.

[218] K. Yosida, *Functional Analysis*, Springer, 1980.