May, 1976 Report ESL-R-662

EQUILIBRIA IN STOCHASTIC DYNAMIC GAMES OF STACKELBERG TYPE

by

David A. Castañon

This report is based on the unaltered thesis of David A. Castañon, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in May, 1976. The research was conducted at the Decision and Control Sciences group of the M.I.T. Electronic Systems Laboratory, and the Department of Applied Mathematics at M.I.T. with partial support provided by the Office of Naval Research Contract No. N00014-76-C-0346.

Electronic Systems Laboratory
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

EQUILIBRIA IN STOCHASTIC DYNAMIC GAMES
OF STACKELBERG TYPE

by

DAVID ALFRED CASTAÑON

B.S.E., Tulane University
1971

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May, 1976

Signature of Author ....David A. Castañón.......
Department of Mathematics, April 30, 1976

Certified by ....Nils R. Sandell Jr.......
Thesis Supervisor

Accepted by ....
Chairman, Departmental Committee

EQUILIBRIA IN STOCHASTIC DYNAMIC GAMES
OF STACKELBERG TYPE

by

David Alfred Castañón

Submitted to the Department of Mathematics on April 30, 1976 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

This dissertation presents a study of a class of stochastic dynamic games, characterized by alternate decision-making by the players. This class of games is termed Stackelberg games; the games considered involve only two players. Within this class of games, several types of solutions are discussed; two of these solutions are studied in detail: closed-loop equilibria and signaling-free equilibria.

The problem of finding closed-loop equilibria in general stochastic dynamic Stackelberg games is discussed in detail. Necessary conditions for the existence of equilibria are derived, based on similar developments in control theory. A stage-by-stage decomposition is used to provide sufficient conditions for the existence of an equilibrium solution.

A new class of equilibria, signaling-free equilibria, is defined by restricting the type of information available to the players. In a general formulation, sufficient conditions for the existence of equilibria are established, based on a stage-by-stage dynamic programming decomposition. Under the conditions of equivalent information for both players, an equivalence between signaling-free equilibria and closed-loop equilibria is established.

Signaling-free equilibria are studied for a special class of stochastic dynamic games, characterized by linear state and measurement equations, quadratic cost functionals, and gaussian uncertainties. For this class of games, signaling-free equilibrium strategies are established as affine functions of the available information. Under restrictions to the available information, signaling-free equilibria can be obtained in closed-form, described in a fashion similar to optimal control theory. The resulting equilibrium strategies resemble in part the "certainty-equivalence" principle in optimal control.

Various examples throughout the work illustrate the results

obtained. Additionally, some introductory remarks discuss conceptual algorithms to obtain these solutions. These remarks are part of a closing discussion on applications and future research involving the results of this dissertation.

Thesis Supervisor: Nils R. Sandell, Jr.

Title: Assistant Professor of Systems Science and Engineering

ACKNOWLEDGEMENT

I would like to express my appreciation to Professors Nils Sandell, Jr. and Michael Athans, for the support and valuable interaction during the thesis research. Without Professor Sandell's valuable supervision this research would not have reached its current state of development.

I would also like to thank Professors Yu Chi Ho of Harvard University and Professor Herman Chernoff for serving as my thesis readers. Their comments were extremely helpful for evaluating the thesis research.

The Publications and Drafting Departments of the M.I.T. Electronic Systems Laboratory deserve commendation for their efficient preparation of this manuscript. Special thanks go to Susanna Natti and Margaret Flaherty. I would also like to thank a number of close friends whose influence provided support and motivation during the thesis research.

I would like to thank the Applied Mathematics group at M.I.T. for their financial support as a teaching assistant and the valuable teaching and educational opportunities which they provided.

This research was conducted at the Decision and Control Sciences group of the M.I.T. Electronic Systems Laboratory, and the Department of Applied Mathematics at M.I.T., with partial support under the Office of Naval Research Contract No. N00014-76-C-0346.

TABLE OF CONTENTS

Page

CHAPTER I

I.1 Introduction

With the introduction and availability of sophisticated computing and modeling equipment in recent years, there has been a trend among academic disciplines to formulate mathematical models to study real-life situations. Among the physical sciences, this treatment is adequate since they deal with measurable quantities in Nature and obey certain known Mathematical relationships. Recently, the same treatment has been extended to such disciplines as Behavioral Sciences, Economics and Biology (S4, SH1). Interesting research has been conducted which attempts to characterize living systems in terms of language and ideas derived from the physical sciences. However, these systems are not understood to the degree where accurate laws of behavior can be formulated. There is conflict as to which are the relevant units of measure, variables and their proper scaling; these factors introduce a strong element of uncertainty associated with the models.

This trend in the social sciences has opened broad areas of application for Control Theory. Control Theory as a science concerns itself with making optimal decisions; this concept has a natural application to models of Economic and Social systems where policy decisions must be made. However, the decision problems which arise from these systems are different from the "classical" decision problems of the physical sciences. A "classical" decision problem is illustrated in Figure I.1. One decision maker observes the outputs of a system and uses them to choose inputs which cause the system to behave in a desired fashion.

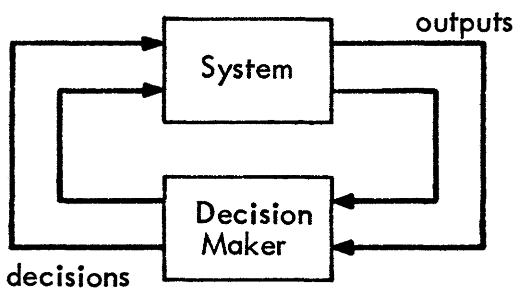

Figure I.1

An important aspect of models in the Behavioral Sciences is that decisions are often made by different persons using individual information. Thus, a theory of generalized decision making which subsumes "classical" decision problems must be developed to treat these problems. This dissertation deals with a subclass of mathematical problems originating from these models, which are classified under the name of games; these problems differ from classical decision problems in the fact that different information is available to the decision makers, and each decision maker may have different goals.

I.2 Game Theory

Game theory is a method for the study of decision making in situations of conflict (SH1). It deals with processes in which the individual decision-unit is not in complete control of the other decision units entering into the environment. The essence of a game is that it involves decision makers with different objectives and information whose fates are intertwined. A team problem (H2, H3, R1) is a game where all decision makers have a common objective. When decision makers act at more than one occasion, and there is an order introduced in these decisions, the optimization problem is known as a dynamic problem.

A game is described in terms of the players, the payoffs and the rules. A player is an autonomous decision-making unit in control of some resources. The rules of the game specify how these resources may be utilized, what information is available, and all other relevant aspects of the environment. A strategy is a set of instructions which states in advance how a player intends to select each move, taking into account the knowledge he will have available at the time he is called upon to select his move. It takes into account all contingencies and provides instructions on how to act under every circumstance. A payoff function for player i assigns a valuation to each complete play of the game, for player i; this payoff will in general depend on the rules of the game and the decisions of all the players.

There are two ways in which games are represented: in a compact manner called the normal form (VN1, L1), and in a descriptive manner called the extensive form (VN1, K1, K2, Au1). The normal form of the game identifies each complete play of the game by the decisions made, and hence it consists of a vector-valued function which assigns a payoff to each player for each possible combination of decisions (or strategies). It transcends the rules of the game, directly relating decisions to payoffs. The extensive form of the game is a detailed description of how the game is a detailed description of how the game is played which incorporates both temporal order and available information. Von Neumann and Morgernstern (VN1) introduced the notion of games in normal and extensive form. Kuhn and others (K1, K2, Au1, Wt2) developed the concept further.

The normal and extensive forms of a game describe the interconnection of the players' decision to their cost. However, there is no concept yet of what is a solution to a game. When only one decision maker is present, or there is a common goal, the solution to the game can be defined as the set of decision which optimizes this goal. In the general case where different goals exist, there is no unique concept of a solution; each solution concept carries certain assumptions as to the philosophy motivating the decision-makers. Thus, in order to pose properly the optimization problem in a game, the game must be specified in its normal or extensive form, and the desired solution concept must be specified also. Throughout the remainder of this work the term <u>solution concept</u> will denote the assumptions defining a solution to a game.

I.3 Summary of Thesis

Research in dynamic game theory to date is full of apparent contradictions and confusing results (Ba3, Sa2). A large part of these problems is due to the adaptation of concepts and techniques developed for nondynamic games and control theory to dynamic games. The present work is aimed at establishing various results using control-theory ideas in a rigorous framework for dynamic games, attempting to clarify many of these points of conflict. A mathematical formulation for the class of dynamic games considered is presented in Chapter II. Within this formulation, several solution concepts are defined, including some concepts which arise naturally out of a dynamic formulation. These concepts are looked at from a game-theoretic point of view, resulting in interpretations of these concepts in terms of extensive and normal forms of games. These interpretations provide considerable insight into the differences between the solution concepts. The chapter concludes with the development of some results characterizing these concepts.

In Chapter III, several examples are discussed which illustrate the results obtained in Chapter II. In particular, several of the games described illustrate the dominant assumptions of some solution concepts, and how this dominance is affected by uncertainties in the system.

Chapters IV and V deal with games of imperfect information. Two solution concepts are studied in this context; these concepts have the feature that decisions are not made simultaneously at any stage. Chapter V is a study of a subclass of problems of this type, whose solution can be expressed in closed form for one solution concept. Sufficient conditions are given for the existence and uniqueness of this solution.

Chapter IV addresses a general class of problems with imperfect information. Several results are established characterizing the solutions sought under two different solution concepts. Under some constraints in the information, these two solution concepts are similar. The validity of a dynamic programming (B1) approach towards obtaining these solutions is established.

Chapter VI contains several comments on open problems left in the area, and summarizes the results of the previous chapter.

I.4 Contributions of Thesis

The major contributions of this research are:

1. The formulation of the various solution concepts considered in a dynamic framework.
2. The development and clarification of the properties of these solutions.
3. The introduction of important examples.
4. The formulation of solution concepts for games of imperfect information.
5. The extension of dynamic programming techniques to games of imperfect information.
6. Formulation and conceptual solution of general games with imperfect information.
7. Exact solution for a subclass of games with imperfect information under one solution concept.

CHAPTER II

Dynamic Games

In this chapter a mathematical framework is provided for the study of dynamic games. Different solution concepts are studied within this framework; several relationships are derived between these solution concepts. Additionally, the chapter includes an overview of results available in dynamic games.

II.1 Mathematical Formulation

The systems studied are described by a state equation

$$x_{t+1} = f_t(x_t, u_t^1, \dots u_t^m, \theta_t); \quad t = 0, \dots N-1 \tag{2.1}$$

where $x_t \in X_t$, $u_t^i \in U_t^i$, $\theta_t \in (R^p, B^p)$; (X_t, G_t^0), (U_t^i, G_t^i) are measurable spaces, where the σ-fields G_t^i contains all sets of the form $\{u_t^i\}$, and $\{x_t\} \in G_t^0$ for $x_t \in X_t$, $u_t^i \in U_t^i$. R^p is the p-th dimensional real space, B^p the σ-field of Borel sets, and θ_t is a random variable taking its values in R^p. The function f_t is a measurable function with respect to the product σ-field $G_t^0 \times G_t^1 \times \dots G_t^m \times B^p$ to G_{t+1}^0.

There are m measurement equations, given by

$$y_t^i = g_t^i(x_t, \xi_t^i), \; i = 1, \dots, m, \; t = 0, \dots N-1 \tag{2.2}$$

where ξ_t^i are random variables with values in (R^q, B^q) and g_t^i is jointly measurable from the product σ-field $G_t^0 \times B^q$ into the measurable space (Y_t^i, D_t^i), where $\{y_t^i\} \in D_t^i$ for all $y_t^i \in Y_t^i$.

The equations can be interpreted as follows: equation (2.1) models some uncertain dynamical system and how it is affected by decision from the players. The variable x_t represents the possible states of the system, u_t^i are the inputs, and θ_t represents the stochastic effects present. Equation (2.2) represents the information acquired by each player at stage t. The total information available to player i at stage t is denoted z_t^i. Measurement uncertainties are modeled by the random vectors ξ_t^i. It is assumed that all of the uncertain variables are defined on an underlying probability space (Ω, F^0, P^0), and have a known joint probability distribution. Associated with the system is a set of equations describing the individual costs:

$$J^i = h_N^i(x_N) + \sum_{t=0}^{N-1} h_t^i(x_t, u_t^1, \ldots u_t^n); \quad i = 1, \ldots m \tag{2.3}$$

when h_t^i is an appropriately measurable function into the positive real numbers.

The information set z_t^i consists of the rules and parameters of the game (the passive information, which is the same to all players at all stages) and the active information gathered from observation of the other players' decisions and player i's measurements of the state of the system. It is assumed that every player can observe exactly the other players' past decisions; this assumption can be relaxed as discussed in Chapter VI. The information set z_t^i differs from z_T^j, $i \neq j$, $t \neq T$, only in the active information, so the symbol z_t^i will be used to represent the active information for player i at time t.

From its definition, $z_t^i \in Z_t^i = U_0^1 \times \ldots U_0^m \times \ldots U_{t-1}^m \times Y_0^i \times \ldots$ $\ldots Y_t^i$. A <u>control law</u> for each player i at time t is a map

$$\gamma_t^i: Z_t^i \rightarrow U_t^i$$

which is measurable from the product σ-field $D_0^i \times \ldots D_t^i \times G_0^1 \times \ldots G_{t-1}^m$ into G_t^i. Control laws are also referred to as <u>strategies</u>.

The set of admissible strategies Γ_t^i for player i at time t is a subset of the set of all strategies γ_t^i. The strategies $\left\{\gamma_k^i\right\}_{k=0,\ldots t-1}^{i=1,\ldots m}$ can be used together with equations (2.1) and (2.2) to define z_t^i as a random variable in terms of the probability space (Ω, F^0, P^0). The measurability of the maps produces a well-defined random variable z_t^i, measurable on the field $G_0^1 \times \ldots G_{t-1}^m \times D_0^i \times \ldots D_t^i$. This field induces a σ-field $F_t^i \subset F^0$ in Ω (D1) through the random variable z_t^i. The dependence of this induced field on $\left\{\gamma_k^i\right\}_{k=0,\ldots t-1}^{i=1,\ldots m}$ will be denoted by $F_t^i\left(\left\{\gamma_k^i\right\}_{k=0,\ldots t-1}^{i=1,\ldots m}\right)$. Thus, strategies can alternately be defined as functions $\gamma_t^i:\Omega \rightarrow U_t^i$ measurable on $F_t^i\left(\left\{\gamma_k^i\right\}_{k=0,\ldots t-1}^{i=1,\ldots m}\right)$. The σ-field $F_t^i\left(\left\{\gamma_k^i\right\}_{k=0,\ldots t-1}^{i=1,\ldots m}\right)$ represents the information available to player i at time t about the uncertainties in the system, following the strategies $\left\{\gamma_k^i\right\}_{k=0,\ldots t-1}^{i=1,\ldots m}$. The notation used to represent strategies in this alternate form will be

$$\gamma_t^i = \gamma_t^i\left(\omega, \left\{\gamma_k^j\right\}_{k=0,\ldots t-1}^{j=1,2\ldots m}\right)$$

to indicate the dependence of the strategies on the previous strategies. These two representations of admissible strategies will be used alternately throughout the remainder of the thesis.

A complete play of the game can be uniquely defined by specifying the values $\left\{u_t^i\right\}_{t=0,\ldots N-1}^{i=1,\ldots m}$ of all the decision and the values of all the random elements, as follows: Equation (2.1) can be used to define the values of x_t for all t. Then, equation (2.3) defines a cost vector $\underline{J} = (J^1, \ldots J^m)$. This cost vector is a function of the values $\left\{u_t^i\right\}_{t=0,\ldots N-1}^{i=1,\ldots m}$ and the uncertainties (which can be specified by choosing an element $\omega \in \Omega$); hence, by the measurability assumptions, this cost vector is a well-defined random variable in the probability space (Ω, F^0, P^0).

A complete play can be alternately defined by specifying the strategies $\left\{\gamma_t^i\right\}_{t=0,\ldots t-1}^{i=1,\ldots m}$ used by the players throughout the game; and a value for the random elements. Once the value of the random element ω is known, then y_0^i can be computed from equation (2.2). This value can be used to define z_0^i, and then the values $u_0^i = \gamma_0^i(z_0^i)$ can be computed for all i. Equation (2.1) then provides the value of x_1.

This procedure can be used recursively, defining x_t, then y_t^i and z_t^i, thereby specifying $u_t^i = \gamma_t^i(z_t^i)$. The values x_t and u_t^i are then used in equation (2.3) to define a cost vector in terms of ω, and the choice of strategies $\left\{\gamma_t^i\right\}_{t=0,\ldots N-1}^{i=1,\ldots m}$. The measurability assumptions stated earlier assure that this cost vector $\underline{J}\left(\left\{\gamma_t^i\right\}_{t=0,\ldots N-1}^{i=0,\ldots m}\right)$ is a random vector well-defined in the probability space (Ω, F^0, P^0).

There is a subtle difference between these two ways of specifying a complete play of the game. The earlier way treats the decisions as known constants, whereas the latter way treats u_t^i as a variable depending on the information; i.e. a random variable. The latter specification includes the earlier as a special case when only constant strategies $\left(\gamma_t^i(z_t^i) = u_t^i \text{ for all } z_t^i \in Z_t^i\right)$ are considered. This distinction is very important when dealing with the way fields are induced in Ω.

A game is called <u>deterministic</u> if the field induced on Ω by the original random elements x_0, $\left\{\xi_t^i, \theta_t\right\}_{t=0,\ldots N-1}^{i=1\ldots,m}$ is the trivial σ-field (ϕ, Ω). That is, there are no uncertainties in the state evolution and measurement equations. A <u>solution concept</u> is a set of criteria which a set of strategies $\left\{\gamma_t^i\right\}_{t=0,\ldots N-1}^{i=1,\ldots m}$ must meet in order to be considered a solution to the game. Solution concepts include assumptions about how each player chooses his strategy and what classes of strategies he chooses from. The remainder of this thesis deals with solution concepts for two types of games: games where decisions u_t^i are enacted simultaneously at each stage, termed Nash games in connection with the Nash equilibrium concept proposed by Nash in (N1, N2); and games where some u_t^i are enacted (and observed by the remaining players) prior to other u_t^i, termed Stackelberg games in reference to the work of Von Stackelberg (VS1). Throughout the remainder of this work the number of players is assumed to be two; in most cases, this represents no loss in generality. Additionally, it is assumed that each player has perfect recall of his own information; that is, for any t, the information vector z_{t+1}^i contains all of the elements of the information vector z_t^i for any play of the game.

As a final note, the symbol u_t^i will be used to denote both the value of $\gamma_t^i(z_t^i)$ and the constant strategy $\gamma_t^i(z_t^i) = u_t^i$ for all z_t^i; this distinction will be made when necessary to avoid confusion.

II.2 Solution Concepts in Dynamic Nash Games

Solution concepts are discussed in this chapter for two classes of games: Nash games and Stackelberg games. The basic difference between these two classes is that in Nash games decisions at stage t are chosen and enacted simultaneously by both players, while in Stackelberg games these conditions do not hold.

Three solution concepts are relevant for Nash games: the first discussed is the open-loop Nash equilibrium solution. Assume that each player seeks to minimize his cost.

Definition II.1. An open-loop equilbrium in a Nash game is a pair of decision sequences $\underline{u}^{1*} = \left\{u_t^{1*}\right\}_{t=0,\ldots N-1}$, $\underline{u}^{2*} = \left\{u_t^{2*}\right\}_{t=0,\ldots N-1}$ with the property that, for all $\underline{u}^i \in \prod_{t=0}^{N-1} U_t^i$,

$$E\{J^1(\underline{u}^{1*}, \underline{u}^{2*})\} \leq E\{J^1(\underline{u}^1, \underline{u}^{2*})\} \tag{2.4}$$

$$E\{J^2(\underline{u}^{1*}, \underline{u}^{2*})\} \leq E\{J^2(\underline{u}^{1*}, \underline{u}^2)\} \tag{2.5}$$

where the expectations are defined in terms of the original probability space (Ω, F^0, P^0) as described in Section II.1.

The open-loop Nash equilibria assume that no information is acquired through the play of the game; hence the admissible strategies are constant strategies. This solution concept is used primarily in deterministic games. In the presence of uncertainty, the closed-loop Nash equilibrium concept is more appropriate.

<u>Definition II.2.</u> A closed-loop equilibrium is a pair of admissible strategy sequences $\underline{\gamma}^{1*}$, $\underline{\gamma}^{2*}$, such that, for all admissible strategy sequences $\underline{\gamma}^{i} \in \prod_{t=0}^{N-1} \Gamma_t^i$,

$$E\{J^1(\underline{\gamma}^{1*}, \underline{\gamma}^{2*})\} \leq E\{J^1(\underline{\gamma}^{1}, \underline{\gamma}^{2*})\} \tag{2.6}$$

$$E\{J^2(\underline{\gamma}^{1*}, \underline{\gamma}^{2*})\} \leq E\{J^2(\underline{\gamma}^{1*}, \underline{\gamma}^{2})\} \tag{2.7}$$

where J^1, J^2 are defined as random variables in terms of (Ω, F^0, P^0) as described in Section II.1.

It should be noticed that the admissible strategies in Definition II.2 could be of either form discussed in Section II.1. To eliminate this ambiguity, let $\hat{\gamma}_t^i : Z_t^i \to U_t^i$ represent strategies which map available information to decisions, and $\gamma_t^i : \Omega \times \Gamma_0^1 \times \Gamma_0^2 \times \ldots \Gamma_{t-1}^1 \times \Gamma_{t-1}^2 \to U_t^i$ represent strategies expressed in terms of the original space (Ω, F^0, P^0).

The major difference between the open-loop and closed-loop equilibria is that in the open-loop concept admissible strategies do not depend on active information (they are constant); whereas in the closed-loop concept they may be nonconstant functions of the available information. The active information z_t^i has the values of the measurements $\left\{y_k^i\right\}_{k=0,\ldots t}$ and the past decisions by both players, $\left\{u_k^1, u_k^2\right\}_{k=0,\ldots t-1}$. Since this is a Nash game, it does not include knowledge of the current decision by the other player because both players enact their decisions simultaneously.

The last Nash solution concept of interest is the signaling-free Nash equilibrium concept. The philosophy behind this concept is that decisions can be made independently at each stage t and need not be made until the game is at stage t. Define the reduced cost J_t^i as:

$$J_t^i = h_N^i(x_N) + \sum_{j=t}^{N-1} h_j^i(x_j, u_j^1, u_j^2) \tag{2.8}$$

<u>Definition II.3.</u> A <u>signaling-free</u> strategy for player i at stage t in a Nash game is a map

$$\gamma_t^i : \Omega \times U_0^1 \times U_0^2 \times \ldots U_{t-1}^1 \times U_{t-1}^2 \to U_t^i$$

with the following properties:

(i) $\gamma_t^i \left(\cdot , \{u_k^j\}_{k=0,\ldots t-1}^{j=1,2} \right)$ is measurable from $F_t^i \left(\{u_k^j\}_{k=0,\ldots t-1}^{j=1,2} \right)$ to G_t^i.

F_t^i is the field induced by the measurements $\{y_k^i\}_{k=0,\ldots t}$ on Ω assuming constant strategy u_k^j.

(ii) For each $\omega \in \Omega$, $\gamma_t^i(\omega, \cdot)$ is measurable from $G_0^1 \times G_0^2 \times \ldots \times G_{t-1}^1 \times G_{t-1}^2$ to G_t^i.

(iii) $\gamma_t^i(\cdot, \cdot)$ is jointly measurable from $F^0 \times G_0^1 \times G_{t-1}^2$ to G_t^i.

The name signaling-free refers to the fact that information about the uncertainties in the space (Ω, F^0, P^0) is acquired only through the measurements y_t^i; the decisions u_t^i are assumed to be constant strategies in defining the information σ-field $F_t^i\left(\left\{u_k^j\right\}_{k=0,\ldots t-1}^{j=1,2}\right)$, thus they contain no information about Ω.

Notation: Let $\overline{\gamma}_t$ represent the sequence of strategies $\left\{\gamma_k^i\right\}_{k=0,\ldots t}^{i=1,2}$. Similarly, $\overline{u}_t$ is a sequence of constant strategies. Also, let $\underline{\gamma}_t^i$ represent the sequence of strategies $\gamma_t^i, \ldots \gamma_{N-1}^i$. Having introduced the concept of signaling-free strategies, the next step is to verify how the signaling-free assumptions get reflected in the costs. Equation (2.8) defines partial costs starting from any stage t. In particular, for any $\overline{u}_{N-2}$ define $J_{N-1}^i(\overline{u}_{N-2}, \gamma_{N-1}^1, \gamma_{N-1}^2)$ as a random variable on Ω using the state evolution equation (2.1); relate x_{t-1} to (Ω, F^0, P_0^0) through the constant strategies $\overline{u}_{N-2}$ in Equation 2.1. Then, for any $\omega \in \Omega$,

$$\begin{aligned} J_{N-1}^i(\overline{u}_{N-2}, \gamma_{N-1}^1, \gamma_{N-1}^2)(\omega) &= h_{N-1}^i\Big(x_{N-1}^{(\omega)}, \gamma_{N-1}^1(\omega, \overline{u}_{N-2}), \\ &\quad \gamma_{N-1}^2(\omega, \overline{u}_{N-2})\Big) + h_N^i\Big(f_{N-1}(x_{N-1}(\omega), \gamma_{N-1}^1(\omega, \overline{u}_{N-2}), \\ &\quad \gamma_{N-1}^2(\omega, \overline{u}_{N-2}), \omega)\Big). \end{aligned}$$

It is a well-defined random variable in terms of ω because of the measurability assumptions on h_t^i and f_t. With these preliminaries one can define appropriate costs.

Definition II.4. The expected cost-to-go for player i at stage t is a function I_t^i with the following properties:

(i) $I_t^i : \Omega \times U_0^1 \times U_0^2 \times \ldots \times U_{t-1}^1 \times U_{t-1}^2 \times \Gamma_t^1 \times \Gamma_t^2 \times \ldots$

$$\ldots \Gamma_{N-1}^1 \times \Gamma_{N-1}^2 \to R$$

(ii) $I_{N-1}^i(\omega, \bar{u}_{N-2}, \gamma_{N-1}^1, \gamma_{N-2}^2)$

$$= E\{J_{N-1}^i(\bar{u}_{N-2}, \gamma_{N-1}^1, \gamma_{N-1}^2) | z_{N-1}^i\} \qquad (2.9)$$

(iii) $I_t^i(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^2) = E\{h_t^i(x_t, u_t^1, u_t^2)$

$$+ I_{t+1}^i(\omega, \bar{u}_t, \underline{\gamma}_{t+1}^1, \underline{\gamma}_{t+1}^2) | z_t^i, u_t^j = \gamma_t^j(\omega, \bar{u}_{t-1}), j=1,2\} \qquad (2.10)$$

where the expectations in equations (2.9) and (2.10) are conditional expectations in terms of the fields $F_t^i(\bar{u}_{t-1})$, and the substitutions $u_t^i = \gamma_t^j(\omega, \bar{u}_{t-1})$ are used to define the random variable inside the conditional expectation in (2.10). It is assumed that $\bar{u}_{t-1}$ is consistent with z_t^i.

I_t^i is expressed in terms of conditional expectations defined on fields induced by the information $z_t^i(\omega)$. The value of these expectations can be expressed in terms of z_t^i, instead of ω. That is, there exists a function $\hat{I}_t^i : Z_t^i \times \Gamma_t^1 \times \Gamma_t^2 \times \ldots \times \Gamma_{N-1}^1 \times \Gamma_{N-1}^2 \to R$ such that (L1)

$$\hat{I}_t^i(z_t^i(\omega), \underline{\gamma}_t^1, \underline{\gamma}_t^2) = I_t^i(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^2) \quad \text{a.e.}$$

Both representations of the expected cost-to-go will be used further on.

The function $I_t^i(\cdot, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^2)$ is not defined uniquely in Ω; rather, it is defined within an equivalence class of P^0-null functions.

Hence it must be made clear when dealing that any equations dealing with the functions I_t^i or $\hat{I}_t^i$ will relate only to this equivalence class; this will be denoted by the words almost everywhere (a.e.). With these preliminaries, a signaling-free Nash equilibrium can be defined as:

Definition II.5.: A signaling-free equilibrium in a Nash game is a pair of admissible strategy sequences $\underline{\gamma}^{1*}$, $\underline{\gamma}^{2*}$, with the property that, for any admissible strategy sequences $\underline{\gamma}^1$, $\underline{\gamma}^2$,

$$I_t^i(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^{1*}, \underline{\gamma}_t^{2*}) \leq I_t^1(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^{2*}) \quad \text{a.e.} \tag{2.11}$$

$$I_t^2(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^{1*}, \underline{\gamma}_t^{2*}) \leq I_t^2(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^{1*}, \underline{\gamma}_t^2) \quad \text{a.e.} \tag{2.12}$$

for all $t \in 0, \ldots N-1$, all $\bar{u}_{t-1} \in \prod_{k=0}^{t-1} U_k^1 \times U_k^2$.

The conditional expectation in equations (2.9) and (2.10) is taken with the assumption that past decisions are known constants, as opposed to results of strategies which are unknown. This reflects the philosophy that decisions enacted at stage t are made independently of past strategies and that a strategy for stage t need not be chosen until stage t, based solely on the available information. In deterministic games, or in stochastic games where $z_t^1 = z_t^2$ for all t, signaling free Nash equilibria are also closed-loop equilibria. However, when $z_t^1 \neq z_t^2$ for all t, the possibility of transfer of information exists (S1, T1); this is similar to signaling strategies in team theory (S1, R1, Sa3). In this case, signaling-free Nash equilibria will not be closed-loop equilibria.

Various authors, notably Aumann and Maschler, and Ho (Au2, H4) have pointed out the inadequacy of the closed-loop equilibrium solution in dynamic games; the introduction and definition of signaling-free equilibria follows the spirit of these works. Under this solution concept, the players do not commit themselves to choosing a strategy until they have to enact it. Hence the equilibrium strategies must be equilibrium strategies for any part of the game starting at stage t. This is reflected in the definition by defining the equilibrium strategies as sequences which are in equilibrium for costs-to-go at all stages.

The other important aspect of the signaling-free assumption lies in the restrictions imposed on learning about the uncertainties through observation of the other player's decisions. This restriction arises from the interpretation of the solution: at stage t, player i seeks to minimize his expected cost-to-go, regardless of how previous decisions were made. The values $\bar{u}_{t-1}$ are known to both players at stage t, not the form of the strategies which produced those values. Under the signaling-free solution, the players make no further assumptions as to the strategies which produced these values, treating them as constant strategies.

II.3 Solution Concepts in Stackelberg Games

In the Stackelberg solution concepts decisions are not enacted simultaneously at each stage. This creates a difference in information between the two players, since the player who acts second knows the value of the first player's decisions. The player who acts first is called the <u>leader</u> and the other player is called the <u>follower</u>. Within the dynamic game context, three types of solution concepts are important in Stackelberg games: open-loop Stackelberg solutions, closed-loop Stackelberg solutions, and equilibrium solutions.

The open-loop Stackelberg solution assumes dominance of the game by the leader; the leader chooses all of his decisions and announces them to the follower, who then chooses his decisions. The dominance of the leader is reflected in the assumption that he can predict the follower's reactions to the leader's decisions.

<u>Definition II.6.</u> Player 1's <u>rational reaction</u> to $\underline{\gamma}^2$, $V^{1c}(\underline{\gamma}^2)$, is the set of admissible strategies defined by

$$V^{1c}(\underline{\gamma}^2) = \{\underline{\gamma}^{10} \in \pi_t \Gamma_t^1 | E\{J^1(\underline{\gamma}^{10}, \underline{\gamma}^2)\} \leq E\{J^1(\underline{\gamma}^1, \underline{\gamma}^2)\} \text{ for all } \underline{\gamma}^1 \in \pi_t \Gamma_t^1\}$$

Player 2's rational reaction set is defined similarly. When only constant strategies are admissible, $V^1(\underline{u}^2)$ and $V^2(\underline{u}^1)$ are used to denote the players' rational reactions. The sets $\{(\underline{\gamma}^1, V^{2c}(\underline{\gamma}^1)), \underline{\gamma}^1 \in \pi_t \Gamma_t^1\}$ and $\{(\underline{\gamma}^2, V^{1c}(\underline{\gamma}^2)), \underline{\gamma}^2 \in \pi_t \Gamma_t^2\}$ are known as the <u>rational reaction sets</u> for players 2 and 1, respectively.

<u>Definition II.7.</u> An open-loop Stackelberg solution is a pair of decision sequences $\underline{u}^{1*}$, $\underline{u}^{2*}$ such that

(i) $\underline{u}^{2*} \in V^2(\underline{u}')$

(ii) $E\{J^1(u^{1*}, u^{2*})\} \leq E\{J^1(\underline{u}^1, \underline{u}^2)\}$

$$\text{for all } \underline{u}^1 \in \underset{t}{\pi}\, U_t^1;\ \underline{u}^2 \in V^2(\underline{u}^1). \qquad (2.13)$$

This definition differs slightly from the definitions given previously in the literature (Be1, Si1, Si2). The previous definitions were ambiguous if $V^2(\underline{u}^1)$ contained more than one element $\underline{u}^2$ for each $\underline{u}^1$. The open-loop Stackelberg concept is a dominant solution concept; here in the definition used here, it is implied that whenever two strategies yield an equal cost for the follower, he will choose the one that minimizes the leader's cost. This strengthens the dominance of the leader in this concept, while eliminating the ambiguity which existed in previous definitions.

As indicated in the Nash solutions, the open-loop solutions are used primarily when no information is acquired in the play of the game. When information is acquired, the closed-loop Stackelberg concept is more appropriate. The leader chooses and announces the strategy sequence $\underline{\gamma}^1$ to the follower who then chooses his strategy sequence $\underline{\gamma}^2$. Admissible strategies are defined as they were for the closed-loop Nash equilibrium concept in terms of Ω.

Definition II.8. A closed-loop Stackelberg solution is a pair of admissible strategy sequences $\underline{\gamma}^{1*}$, $\underline{\gamma}^{2*}$, such that $\underline{\gamma}^{2*} \in V^{2c}(\underline{\gamma}^{1*})$, and

$$E\{J^1(\underline{\gamma}^{1*}, \underline{\gamma}^{2*})\} \leq E\{J^1(\underline{\gamma}^1, \underline{\gamma}^2)\}$$

$$\text{for all } \underline{\gamma}^1 \in \underset{t}{\pi}\, \Gamma_t^1;\ \underline{\gamma}^2 \in V^{2c}(\underline{\gamma}^1). \qquad (2.14)$$

The closed-loop Stackelberg solution also assumes complete dominance by the leader; the major difference in the closed-loop and open-loop solutions is the choice of strategies, as the open-loop strategies are constant strategies. A difficulty associated with the closed-loop Stackelberg solution is that equation (2.13) defines a constrained functional minimization problem where the constraint set is not of simple form. Thus, solutions which satisfy (2.13) are difficult to find.

The third class of Stackelberg solutions is the equilibrium Stackelberg solutions. The basic difference between this and the two previous classes is that the leader is no longer assumed dominant. Thus, an equilibrium between the leader and follower is desired. The Stackelberg nature of the game arises in the fact that at each stage, the leader announces his decision to the follower before the follower makes his decision. Three types of equilibria are considered: open-loop, closed-loop and signaling-free. The definitions are analogous to the Nash equilibrium solutions, with the difference that the follower's information vector z_t^2 now includes knowledge of the leader's decision u_t^1.

Since the open-loop equilibria does not use any information other than that available before the start of the game, the open-loop equilibria are equilibria for both Stackelberg and Nash games. Since the information provided by the leader's decision is not used, it makes no difference whether they play simultaneously at each stage or whether the leader plays first.

Definition II.9. A closed-loop equilibrium strategy for a Stackelberg game is a function

$$\gamma_t^1 : \Omega \times \Gamma_0^1 \times \Gamma_0^2 \times \ldots \times \Gamma_{t-1}^1 \times \Gamma_{t-1}^2 \rightarrow U_t^1$$

or

$$\gamma_t^2 : \Omega \times \Gamma_0^1 \times \Gamma_0^2 \times \ldots \times \Gamma_{t-1}^1 \times \Gamma_{t-1}^2 \times \Gamma_t^1 \rightarrow U_t^2$$

such that $\gamma_t^1(\cdot, \bar{\gamma}_{t-1})$ is measurable on $F_t^1(\bar{\gamma}_{t-1})$, and $\gamma_t^2(\cdot, \bar{\gamma}_{t-1}, \gamma_t^1)$ is measurably on $F_t^2(\bar{\gamma}_{t-1}, \gamma_t^1)$ for all $\bar{\gamma}_t \in \Gamma_0^1 \times \Gamma_0^2 \times \ldots \Gamma_t^1 \times \Gamma_t^2$.

Equivalently, because of the definition of the fields F_t^i, a closed-loop strategy can be represented as a function of the information z_t^i. That is, a closed-loop strategy can be a function $\hat{\gamma}_t^1 : Z_t^1 \times \Gamma_0^1 \times \Gamma_0^2 \times \ldots \Gamma_{t-1}^1 \times \Gamma_{t-1}^2 \rightarrow U_t^1$ or $\hat{\gamma}_t^2 : Z_t^2 \times \Gamma_0^1 \times \ldots \times \Gamma_{t-1}^2 \times \Gamma_t^1 \rightarrow U_t^2$ such that

$$\hat{\gamma}_t^1(z_t^1(\omega), \hat{\bar{\gamma}}_{t-1}) = \gamma_t^1(\omega, \bar{\gamma}_{t-1}) \quad \text{a.e.}$$

$$\hat{\gamma}_t^2(z_t^2(\omega), \hat{\bar{\gamma}}_{t-1}, \hat{\gamma}_t^1) = \gamma_t^2(\omega, \bar{\gamma}_{t-1}, \gamma_t^1) \quad \text{a.e.}$$

Definition II.10. A closed-loop Stackelberg equilibrium is a pair of admissible strategy sequences $\underline{\gamma}^{1*}$, $\underline{\gamma}^{2*}$ such that, for all admissible $\underline{\gamma}^i$, i = 1, 2,

$$E\{J^1(\underline{\gamma}^{1*}, \underline{\gamma}^{2*})\} \leq E\{J^1(\underline{\gamma}^1, \underline{\gamma}^{2*})\} \tag{2.15}$$

$$E\{J^2(\underline{\gamma}^{1*}, \underline{\gamma}^{2*})\} \leq E\{J^2(\underline{\gamma}^{1*}, \underline{\gamma}^2)\} \tag{2.16}$$

Notice the difference in the field F_t^2 between the closed-loop equilibria in the Nash and the Stackelberg case. In the Stackelberg case, z_t^2 contains knowledge of u_t^1. Thus, the equilibrium is over a different set of admissible strategies than the Nash equilibria.

Definition II.11. A signaling-free strategy in a Stackelberg game for player 1 at stage t is a function $\gamma_t^1 : \Omega \times U_0^1 \times U_0^2 \times \ldots \times U_{t-1}^1 \times U_{t-1}^2 \to U_t^1$; for player 2 it is a function $\gamma_t^2 : \Omega \times U_0^1 \times U_0^2 \times \ldots \times U_{t-1}^1 \times U_{t-1}^2 \times U_t^1 \to U_t^2$ such that

(i) $\gamma_t^1(\cdot\,, \bar{u}_{t-1})$ and $\gamma_t^2(\cdot\,, \bar{u}_{t-1}, u_t^1)$ are measurable from $F_t^1(\bar{u}_{t-1})$ to G_t^1 and $F_t^2(\bar{u}_{t-1}, u_t^1)$ to G_t^2 respectively.

(ii) $\gamma_t^i(\omega, \cdot\,)$ is an appropriately measurable function.

(iii) $\gamma_t^i(\cdot\,, \cdot\,)$ is jointly measurable in F^0 and the appropriate fields to G_t^i.

In a fashion similar to Nash games, expected costs-to-go are defined as:

Definition II.12. The expected cost-to-go for player i at stage t is a function I_t^i such that:

(i) $I_t^1 : \Omega \times U_0^1 \times U_0^2 \times \ldots U_{t-1}^1 \times U_{t-1}^2 \times \Gamma_t^1 \times \Gamma_t^2 \times \ldots$
$\ldots \times \Gamma_{N-1}^1 \times \Gamma_{N-1}^2 \to R$

(ii) $I_t^2 : \Omega \times U_0^1 \times \ldots U_{t-1}^2 \times U_t^1 \times \Gamma_t^2 \times \Gamma_{t+1}^1 \times \ldots \times \Gamma_{N-1}^2 \to R$

(iii) $$I_{N-1}^2(\omega, \bar{u}_{N-2}, u_{N-1}^1, \gamma_{N-1}^2) = E\{J_{N-1}^2(u_{N-1}^1, \gamma_{N-1}^2) | z_{N-1}^2\} \tag{2.17}$$

(iv) $$I_{N-1}^1(\omega, \bar{u}_{N-2}, \gamma_{N-1}^1, \gamma_{N-1}^2) = E\{J_{N-1}^1(\gamma_{N-1}^1, \gamma_{N-1}^2) | z_{N-1}^1\} \tag{2.18}$$

$$\text{(v)} \quad I_t^2(\omega, \bar{u}_{t-1}, u_t^1, \underline{\gamma}_{t+1}^1, \underline{\gamma}_t^2) = E\{h_t^2(x_t, u_t^1, u_t^2) + I_{t+1}^2(\omega, \bar{u}_t, u_{t+1}^1, \underline{\gamma}_{t+2}^1, \underline{\gamma}_{t+1}^2) \mid z_t^2, u_{t+1}^1 = \gamma_{t+1}^1(\cdot), u_t^2 = \gamma_t^2(\cdot)\} \tag{2.19}$$

$$\text{(vi)} \quad I_t^1(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^2) = E\{h_t^1(x_t, u_t^1, u_t^2) + I_t^1(\omega, \bar{u}_t, \underline{\gamma}_{t+1}^1, \underline{\gamma}_{t+1}^2) \mid z_t^1, u_t^1 = \gamma_t^1(\cdot), u_t^2 = \gamma_t^2(\cdot)\} \tag{2.20}$$

where the expectations in (2.17)-(2.20) are conditional expectations defined in terms of the σ-fields $F_t^i(u_{t-1})$.

The expressions for the expected costs-to-go in Definition II.12 are similar to those obtained for Nash games in Definition II.4; the major difference lies in the information available to player 2, which includes the value u_t^1 at stage t. In particular, the conditional expectations are similarly defined, hence there exist functions $\hat{I}_t^1 : Z_t^1 \times {}_t^1 \times \Gamma_t^2 \times \ldots \times \Gamma_{N-1}^1 \times \Gamma_{N-1}^2 \to R$ and $\hat{I}_t^2 : Z_t^2 \times \Gamma_t^2 \times \Gamma_{t+1}^1 \times \ldots$ $\ldots \Gamma_{N-1}^2 \to R$ such that

$$\hat{I}_t^1(z_t^1(\omega), \underline{\gamma}_t^1, \underline{\gamma}_t^2) = I_t^1(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^2) \quad \text{a.e.}$$

$$\hat{I}_t^2(z_t^2(\omega), \underline{\gamma}_t^1, \underline{\gamma}_t^2) = I_t^2(\omega, \bar{u}_{t-1}, u_t^1, \underline{\gamma}_{t\_1}^1, \underline{\gamma}_t^2) \text{ a.e.}$$

Both of these representations of the expected costs-to-go will be used in other chapters. A signaling-free equilibrium can be defined as:

<u>Definition II.13.</u> A signaling-free equilibrium is a pair of admissible sequences $\underline{\gamma}^{1*}, \underline{\gamma}^{2*}$ such that for any admissible sequences

$\underline{\gamma}^i$,

$$I_t^1(\omega,\ \bar{u}_{t-1},\ \underline{\gamma}_t^{1*},\ \underline{\gamma}_t^{2*}) \leq I_t^1(\omega,\ \bar{u}_{t-1},\ \underline{\gamma}_t^1,\ \underline{\gamma}_t^{2*}) \quad \text{a.e.} \tag{2.21}$$

$$I_t^2(\omega,\ \bar{u}_{t-1},\ u_t^1,\ \underline{\gamma}_{t+1}^{1*},\ \underline{\gamma}_t^{2*}) \leq I_t^2(\omega,\ \bar{u}_{t-1},\ u_t^1,\ \underline{\gamma}_{t+1}^{1*},\ \underline{\gamma}_t^2) \text{a.e.} \tag{2.22}$$

for all $t \in 0, \ldots N-1$, all $\bar{u}_t$.

Note the inclusion of u_t^1 in (2.22), as opposed to equation (2.12). This is due to the fact that u_t^1 is known in z_t^2 in the Stackelberg game, while it is not known in the Nash game; hence the follower can treat this value as a known constant instead of a random variable produced by a strategy.

It is noteworthy that the concept of Stackelberg solutions in static games has two different dynamic interpretations: the so-called dominant interpretation, where the leader announces his entire strategy sequence at the beginning of the game; and the equilibrium interpretation, where at each stage the leader acts first, choosing a decision for that stage. This difference will be highlighted in Section II.5.

The comments made in the previous section about the Nash equilibrium strategies also apply to the Stackelberg equilibrium solutions. In cases of like information signaling-free equilibria are also closed-loop equilibria. This will be proven in a rigorous framework in Chapter IV. For deterministic games, additional relationships can be derived between the different solution concepts. These relationships are proven in Section II.6.

II.4 An Overview of Dynamic Game Literature

The mathematical theory of games was first presented in Von Neumann and Morgernstern's Theory of Games and Economic Behavior in 1944. Although some work in the area of games had been done previously, this book is acknowledged to be the introduction of game theory as a Mathematical science. It introduced the concepts of games in normal and extensive form, and discussed the ideas leading to various solution concepts of games. Their work was limited to finite games (games with finite decision spaces), but the concepts they introduced, such as equilibria, strategies and payoffs and their work in zero-sum games permeates all of the game literature.

Kuhn further developed the idea of games in extensive form (K1, K2) by describing the extensive form in a graph-theoretic fashion. In (K2), Kuhn establishes conditions for the use of a dynamic programming (B1) approach to obtaining equilibria in games presented in extensive form. Additionally, his formulation for extensive games included the concept of information available to a decision maker, in terms of information partitions. Although his work was restricted to finite games, Kuhn's results are the backbone of most of the theory of dynamic games with imperfect information.

The idea of looking for equilibrium points as solutions to n-person games is traced to the work of John F. Nash (N1, N2). In his work, Nash proposed the equilibrium points as solutions in static games where all players choose their strategies simultaneously. His name is commonly associated with equilibrium solution concepts in games. Contrary to standard terminology, in this dissertation Nash's name is used to indicate games where decisions are made simultaneously by all players.

The idea of games where decisions are not made simultaneously was used by H. Von Stackelberg (VS1). In his work, he discussed the solution of a static economic problem where one decision maker was constrained to wait for other decision maker's decisions before making his own. The term Stackelberg game is used to denote this class of games.

The theory of dynamic games was slow to evolve, with individual results appearing sporadically. Rufus Isaacs (I1, I2) published his works in Differential Games, where he used ideas associated with control theory and differential equations to solve games with dynamic evolution. Isaacs' work led to a concerted effort by control scientists and applied mathematicians to work on dynamic games. The games Isaacs studied were primarily deterministic zero-sum games.

Ian Rhodes, David Luenberger and others (Rh1, Rh2, Be1) considered the problem of zero-sum dynamic games with imperfect information, and developed solutions for special classes of problems. Starr and Ho (St1, St2) developed the theory one step further by considering non-zero sum games in a dynamic concept; the solution concept which was used in their work was the Nash equilibrium concept, since the games considered were games where decisions were made simultaneously at each stage. Chen and Cruz (Ch1) introduced non-zero sum dynamic Stackelberg games and defined the open-loop Stackelberg solution concept. Cruz and Simaan (S1, S2), Basar (Ba1) and others have also studied dynamic Stackelberg games.

The theory of dynamic games with imperfect information was slow in developing, because even in the simplest cases solutions were difficult to obtain. Willman (W1) studied the solution of a class of zero-sum dynamic games with general information structure; however, most of the work in this area dealt with dynamic games where the information structure was simplified. On the theoretical side, Witsenhausen (Wt2, 3, 4) and Ho (H5) developed a generalization of Kuhn's theory of extensive games to general stochastic dynamic games without the finite-decision restriction. This formulation is the basis for the modern theory of stochastic dynamic games.

The concepts of closed-loop and open-loop solutions were adapted from modern control theory to dynamic game theory, and relates to the class of admissible strategies. Simaan and Cruz (S1, S2) introduced the notion of "feedback" equilibria in Stackelberg games. Their work with these solutions was strictly deterministic. The feedback solutions discussed by Simaan and Cruz are essentially the signaling-free equilibria discussed in Section II.3, restricted to deterministic games. The work of Simaan and Cruz was extended to a class of stochastic games by Castanon and Athans (C1).

The idea of deriving information from other player's decision values can be traced back to Thompson (T1) and his work on signaling strategies. Sandell and Athans (Sa1, Sa3) have used this concept in some non-classical control problems. Aoki (Ao2) discusses a class of strategies which are essentially signaling-free strategies, while considering equilibria in Nash games. The restriction of the transfer of information imposed in the definition of signaling-free strategies in the previous section has been encountered before (Ao2), but never formalized in terms of probability spaces and fields as is done in this dissertation.

II.5 A Game-Theoretic View of Solution Concepts

In this section the various Stackelberg and Nash solution concepts are described in standard game theoretic terms. Throughout this section the games are assumed to be deterministic.

Definition II.14. A normal form of a two-player game is a functional $(J^1, J^2)(\underline{\gamma}^1, \underline{\gamma}^2)$ assigning a pair of real-valued costs to each pair of admissible strategies.

Definition II.15. An extensive geometric form for a game is a description with the following properties:

(i) There is a unique starting node x_0.

(ii) The game is divided into stages where only one player acts at each stage.

(iii) There is unique combination of decisions $\underline{u}^1$, $\underline{u}^2$ which will reach any node.

(iv) There is a partition of the nodes at each stage, where all nodes which cannot be distinguished given the available information belong to the same class.

Definition II.15 is similar to Kuhn's definition of game in extensive form, although it is considerably simpler because of the deterministic multistage structure of the formulation in Section II.1. The concepts of games in normal form and games in extensive geometric forms will be used for comparing the different solutions discussed in Sections II.2 and II.3.

The first concept considered is the open-loop Stackelberg solution whose definition is given by equation (2.13) in Section II.3. The follower's optimal decisions are functions of the leader's decisions $\underline{u}^1$; a map $\underline{u}^{20}(\underline{u}^1)$ which assigns a decision sequence $\underline{u}^2$ for each decision sequence $\underline{u}^1$ is known as a <u>reactive strategy</u>. A Stackelberg game expressed in normal form would be expressed in terms of decision sequences by the leader and reactive strategies for the follower, assigning a cost vector (determined by using the values as in Section II.2) to each pair $(\underline{u}^1, \underline{u}^{20}(\underline{u}^1))$. The open-loop Stackelberg solution is an equilibrium in this form of the game, since equation (2.13) implies

$$J^1(\underline{u}^{1*}, \underline{u}^{2*}(u^{1*})) \le J^1(\underline{u}^1, \underline{u}^{2*}(\underline{u}^1)) \tag{2.23}$$

$$J^2(\underline{u}^{1*}, \underline{u}^{2*}(\underline{u}^{1*})) \le J^2(\underline{u}^{1*}, \underline{u}^{20}(\underline{u}^{1*})) \tag{2.24}$$

for all admissible $\underline{u}^1$, $\underline{u}^{20}$.

However, not all $(\underline{u}^{1*}, \underline{u}^{2*}(\underline{u}^{1*}))$ satisfying (2.23) and (2.24) are open-loop Stackelberg solutions, because equation (2.24) must hold for all $\underline{u}^1$, not just $\underline{u}^{1*}$. This solution cannot be expressed in terms of a normal form. The basic reason is that the normal form of a game loses the concept of when decisions are made with respect to each other, concentrating instead on assigning payoffs once all decisions are made. The essential aspect of the open-loop Stackelberg solution is that the follower makes his choice of strategies after the leader. This property is lost in the normal form of the game.

The mathematical formulation given in Section II.2 describes a game in terms of the state of the game. To convert a game in state form to extensive form, three things are necessary: divide each stage into several stages so that only one decision maker (including the random element) acts at each stage; expand the state space so that two different sequences of decisions will never produce the same state, and convert the information equations into information sets. The geometric extensive form (K2) was developed for finite games, so in this discussion the games are also assumed to be finite.

The open-loop Stackelberg solution can be described in terms of a game whose extensive form has two stages: In the first stage the leader chooses a decision sequence, and in the second stage the follower chooses his decision sequence. The reader is referred to (K2) for a detailed discussion of finite games in extensive from. The resulting game is one of perfect information, so Theorem 3 of (K2) establishes the validity of a dynamic programming approach to obtaining an equilibrium to the game.

The equilibrium is obtained through dynamic programming as follows: Each node at stage 2 corresponds to a strategy sequence $\underline{u}^1$ by the leader at stage 1. For each node the follower computes his optimal decison $\underline{u}^2$. This corresponds to finding the reaction $V^2(\underline{u}^1)$. The leader chooses $\underline{u}^1$ at stage 1 knowing that the follower will choose $\underline{u}^2 \varepsilon V^2(\underline{u}^1)$. This is the process described by equation (2.13) in finding the open-loop Stackelberg solution. This solution is easily identified in a game in extensive geometric form, whereas it wasn't possible to do so in terms of a game expressed in normal form.

The closed-loop Stackelberg solution has similar properties to the open-loop Stackelberg solution, since the major difference lies in the definition of admissible decisions for each player; in the open-loop case, the players choose decision values, whereas in the closed-loop case the players choose strategies. As in the open-loop case, the closed-loop Stackelberg solution is an equilibrium pair of strategies obtained by dynamic programming in the extensive form of the game.

It is significant that the open and closed-loop Stackelberg solution concepts are described in terms of an extensive form consisting of two stages, yet the state representation of the games has N stages. This reduction arises because the leader chooses all of his strategies before the follower chooses any of his. Hence, the follower's choices are based on knowledge of past and future strategies by the leader. Thus, in choosing strategies, the leader's choices at stage t not only affect the future evolution of the state through the state equation but also the past evolution as it influences the choice of the follower's strategies at previous stages. This destroys the causal evolution of equation (2.1), and reduces the problem to a two-stage extensive form.

The open-loop equilibria were defined by equations (2.4) and (2.5). A game expressed in normal form would assign cost vectors to a pair $(\underline{u}^1, \underline{u}^2)$ of decision sequences. In terms of a game in normal form, the open-loop equilibrium is exactly any equilibrium. The game can be reformulated in extensive form in a manner such that 2N stages arise. The major feature of this extensive form is that, since the admissible strategies are constant strategies, the information partitions at each stage are trivial, so that all of the states at a stage are grouped together. Figure II.1 contains a typical representation of an open-loop Nash game in extensive form.

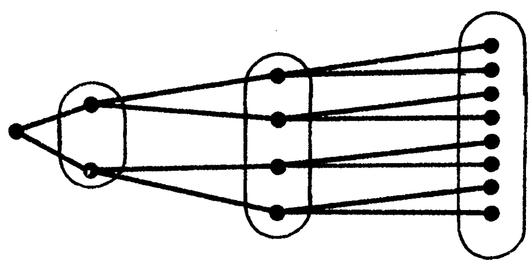

Figure II.1.

Unlike the open-loop equilibria, there is a difference between the Nash and Stackelberg closed-loop equilibria. The major difference lies in the fact that the follower sees the leader's decision at each stage before making his own; this is reflected clearly in the set of admissible strategies. In terms of games expressed in normal form, where cost vectors are assigned to each pair of admissible strategy sequences, both solution concepts are equilibria. Similarly, when the games are expressed in extensive form, the extensive form consists of 2N stages; the major difference lies in the information partition, where the Stackelberg game differs in that the follower now has more information. The Nash game is a simultaneous game (K2) so that at stages where the follower acts, the information partition must group together states where the only difference was the leader's last decision. In the Stackelberg game, the information partition separates all of the states.

Since the games are deterministic games, a dynamic programming approach can be taken to obtaining an equilibrium pair of strategies. The equilibrium pair of strategies so obtained is the signaling-free equilibrium. If the game were converted to normal form, the standard normal form consisting of two players is inadequate in representing a signaling-free equilibrium. Rather, a normal form corresponding to 2N players, with costs I_t^i, is the adequate normal form for a game under the signaling-free solution concept. Within this normal form, the signaling-free equilibrium is an equilibrium between 2N players, as will be shown in the next section. This is consistent with the assumption that decisions at each stage are made independently of other decisions, as opposed to choosing entire decision sequences as in the closed-loop concepts.

To summarize, the open-loop and closed-loop Stackelberg solution concepts are concepts described essentially in terms of a game with two stages, in extensive form. The open-loop and closed-loop equilibria are concepts which can be posed naturally in terms of the normal form of a game, or in a game in extensive form. The signaling-free equilibria correspond to closed-loop equilibria in deterministic games, and can be expressed in terms of a game in extensive form similar to the closed-loop equilibria. However, in terms of a game in normal form, the signaling-free solution represents an equilibrium between 2N individual players, as opposed to 2 in the closed-loop case.

II.6 Relationships between Solution Concepts

In this section several propositions are proven concerning the solution concepts discussed previously. There is strong emphasis on the costs achieved under different solution concepts. The results derived are primarily for deterministic games, although some results are established in the general case.

Proposition 1: In deterministic games, the leader achieves a cost under the closed-loop Stackelberg solution concept no higher than that under the open-loop Stackelberg solution concept if constant strategies are admissible and when both solutions exist.

Proof: Let J^{1*}_{CLS} be the leader's optimal expected cost under the closed-loop concept, J^{1*}_{OLS} the cost under the open-loop concept.

Let $\underline{u}^{10}$, $\underline{u}^{20}$ be optimal open-loop strategies. For any $\underline{u}^1 \in \pi_t U_t^1$, $V^2(\underline{u}^1)$ is a set determined by the solution of a deterministic optimal control problem:

$$\min_{\underline{u}^2} J^2(\underline{u}^1, \underline{u}^2) = h^2(x_N) + \sum_{t=0}^{N-1} h_t^2(x_t, u_t^1, u_t^2)$$

such that

$$x_{t+1} = f_t(x_t, u_t^1, u_t^2)$$

where the state x_t is known perfectly. From optimal control theory (Al, Sal),

$$\min_{\underline{u}^2} J^2(\underline{u}^1, \underline{u}^2) = \min_{\underline{\gamma}^2} J^2(\underline{u}^1, \underline{\gamma}^2(\cdot))$$

for all strategies $\gamma_t^2 : Z_t^2 \to U_t^2$, where z_t^2 consists of the value of all previous decisions u^i (no observations are needed because of the deterministic framework).

Hence $\underline{u}^{20} \in V^{2c}(\underline{u}^{10})$. Since by assumption $\underline{u}^{10}$ is an admissible strategy for player 1, equation (2.14) implies

$$J_{CLS}^{1*} = \min_{\substack{\underline{\gamma}^1 \\ \underline{\gamma}^2 \in V^{2c}(\underline{\gamma}^1)}} J^1(\underline{\gamma}^1, \underline{\gamma}^2) \le J^1(\underline{u}^{10}, \underline{u}^{20}) = J_{OLS}^{1*} .$$

<u>Proposition 2</u>: In general, the open-loop Stackelberg and closed-loop Stackelberg solutions yield lower expected costs for the

leaders than the open-loop Nash equilibria and closed-loop Nash equilibria respectively, when they exist.

Proof: Let J^{1*}_{CLE} be the leader's cost under the closed-loop Nash equilibrium concept, achieved by $\underline{\gamma}^{1e}$, $\underline{\gamma}^{2e}$. Equation (2.7) implies $\underline{\gamma}^{2e} \varepsilon V^{2c}(\underline{\gamma}^{1e})$. Equation (2.14) then implies

$$J^{1*}_{CLS} = \min_{\substack{\underline{\gamma}^1 \\ \underline{\gamma}^2 \varepsilon V^{2c}(\underline{\gamma}^1)}} E\{J^1(\underline{\gamma}^1, \underline{\gamma}^2)\} \leq J^{1*}_{CLE}$$

Similarly, let J^{1*}_{OLE} be the leader's cost under the open loop equilibrium concept, achieved by $\underline{u}^{1e}$, $\underline{u}^{2e}$. Equation (2.5) implies $\underline{u}^{2e} \varepsilon V^2(\underline{u}^{1e})$, so, by equation (2.13)

$$J^{1*}_{OLS} = \min_{\substack{\underline{u}^1 \\ \underline{u}^2 \varepsilon V^1(\underline{u}^1)}} E\{J^1(\underline{u}^1, \underline{u}^2)\} \leq J^{1*}_{OLE}$$

Proposition 3: In deterministic Nash and Stackelberg games, every open-loop equilibrium is also a closed-loop equilibrium if constant strategies are admissible.

Proof: Sandell (Sa2) proved this proposition for Nash games. The proof for Stackelberg games follows his proof. Let $(\underline{u}^{1e}, \underline{u}^{2e})$ be an open-loop equilibrium pair. From Proposition 1, $\underline{u}^{1e} \varepsilon V^{1c}(\underline{u}^{2e})$ and $\underline{u}^{2e} \varepsilon V^{2c}(\underline{u}^{1e})$ because of the equivalence of open- and closed-loop solutions in control theory for deterministic games. Thus, by definition of V^{1c} and V^{2c}, $(\underline{u}^{1e}, \underline{u}^{2e})$ satisfy equations (2.15) and

(2.16), since $\underline{u}^{1e}$ and $\underline{u}^{2e}$ are admissible strategies.

The implications of Proposition 3 are that, in deterministic games, whenever a closed-loop equilibrium is sought, it is not necessary to solve the functional minimization problems indicated in equations (2.15) and (2.16). Rather, the simpler problem of finding an open-loop equilibrium, when it exists, provides a feasible solution to the problem of finding a closed-loop equilibrium. The next proposition establishes another equivalence which provides an alternate method of finding closed-loop equilibria.

Proposition 4: The signaling-free equilibria in Nash and Stackelberg games are closed-loop equilibria in deterministic games.

Proof: In deterministic games, the stochastic elements are trivial. Equivalently, F^0 in (Ω, F^0, P^0) is $F^0 = \{\phi, \Omega\}$. Hence, admissible γ_t^i for both the closed-loop and feedback equilibria depend only on the past decisions $\left\{u_k^j\right\}_{k=0,\ldots t-1}^{j=1,2}$ and the a priori parameters of the game. Thus, the sets of admissible strategies are the same.

Now, consider a Nash game. Equation (2.9) implies

$$\hat{I}_{N-1}^i(\gamma_{N-1}^1, \gamma_{N-1}^{2N}, z_{N-1}^i) = J_{N-1}^i(\gamma_{N-1}^1, \gamma_{N-1}^2)$$

and recursively, since the random elements are deterministic, $\hat{I}_t^i(\underline{\gamma}_t^1, \underline{\gamma}_t^2, z_t^i) = J_t^i(\underline{\gamma}_t^1, \underline{\gamma}_t^2)$ for all t. So, the costs I_0^i and J_0^i are the same in the Nash deterministic game. Then, from equations (2.11) and (2.12) for t = 0, any one equilibrium pair $\underline{\gamma}^{1e}$, $\underline{\gamma}^{2e}$ satisfies (2.6) and (2.7), hence it is a closed-loop equilibrium.

The proof for Stackelberg games is identical.

For finite-state, finite decision games expressed in Kuhn's (K2) extensive form, Kuhn (K2) has established a general result which gives conditions as to when the problem of finding closed-loop equilibria can be approached using a dynamic programming decomposition. Kuhn's results are not restricted to deterministic games; in Chapter IV the equivalence of closed-loop and signaling-free equilibria in stochastic games is discussed.

Proposition 4 does not hold in general games. The reason for this is that signaling-free equilibria are equilibria between 2N independent players; the causality condition that orders the decision-making process also suggests that each player needs to make assumptions of optimal play only from the players that act after him. This reduces the assumptions implicit in the solution concept, and allows for suboptimal strategies by the preceding players. However, in the closed-loop equilibria it is assumed that all strategies at all stages arise out of optimal play. Thus, it is significant that in deterministic games signaling-free equilibria are also closed-loop equilibria. This point will be discussed further for stochastic games in Chapters IV and V.

CHAPTER III

In this chapter several examples are introduced to clarify the concepts and results presented in the previous chapter. The examples presented consist of deterministic games, expressed in terms of the mathematical formulation of Chapter II. Some interesting results in non-zero sum dynamic games (Ba3) are discussed in the context of these examples. Section 1 deals with a finite-state, finite-decision game. The rest of the chapter deals with infinite state and decision spaces.

III.1. A Finite-State, Finite-Decision Game

Within the mathematical framework of Section II.1, define $U_1^1 = U_2^1 = U_1^2 = U_2^2 = \{0, 1\}$ as the sets of admissible decision by players 1 and 2 at stages 1 and 2 respectively. The sets of possible states x_t are $x_1 = \{0\}$, $x_2 = x_3 = \{0, 1\}$. This game is a deterministic game, so the stochastic elements in equations (2.1) and (2.2) are trivial. Define the state evolution equation as

$$x_{t+1} = (x_t + u_t^1 + u_t^2) \bmod 2 \tag{3.1}$$

The costs associated with player 1 can be expressed functionally as:

$$\begin{aligned} h_1^1(x, u^1, u^2) &= (2^{u^1} + 4^{u^2}) \bmod 3 + u^1 \\ h_2^1(x, u^1, u^2) &= x + u^1 + u^2 \\ h_3^1(x) &= x \end{aligned} \tag{3.2}$$

Player 2's costs are given by

$$
\begin{aligned}
h_1^2(x,\ u^1,\ u^2) &= (u^1 + 4^{u^2}) \bmod 3 \\
h_2^2(x,\ u^1,\ u^2) &= 2 + u^1 + u^1(1-u^2) + u^2(1-u^1)(-1)^x \\
h_3^2(x) &= 1
\end{aligned}
\tag{3.3}
$$

The expressions for the total cost are the same as equation (2.3); that is,

$$J^i = h_3^i(x_3) + h_2^i(x_2,\ u_2^1,\ u_2^2) + h_1^i(x_1,\ u_1^1,\ u_1^2) \tag{3.4}$$

The game described so far can be represented in terms of a state-transition map as in Figure III.1, where on each directed path, the numbers in parenthesis denote the values of the decision $(u^1,\ u^2)$ originating that transition, and the bracketed numbers the costs $[h^1,\ h^2]$ associated with that transition.

Figure III.1 contains a summary of the physical rules governing the play of the game. Additional assumptions concerning the order of play at each stage, and the approach towards optimality which each player takes need to be specified; these additional rules will be given by the various solution concepts discussed in Chapter II.

Eliminating the state transitions, a table can be constructed which assigns a pair $(J^1,\ J^2)$ of costs to each admissible decision sequence $(u_1^1,\ u_2^1,\ u_1^2,\ u_2^2)$. Identifying the rows as $u_1^1 u_2^1$, the columns as $u_1^2 u_2^2$, this table is shown in Table III.2.

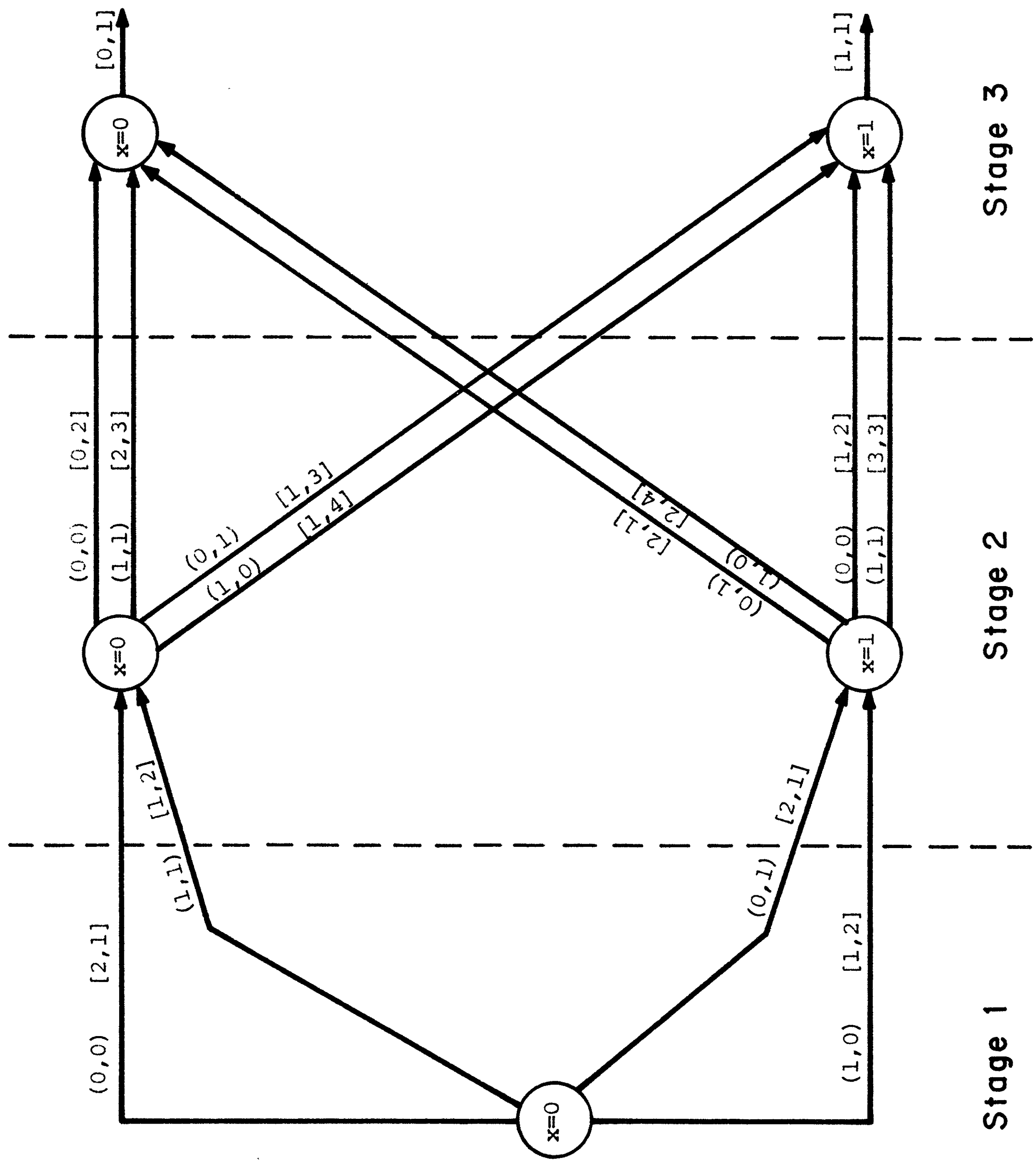
x=0
x=0
x=1
x=0
x=1
(0,0) [2,1]
(1,1) [1,2]
(0,1) [2,1]
(1,0) [1,2]
(0,0) [0,2]
(1,1) [2,3]
(0,1) [1,3]
(1,0) [1,4]
(1,0) [2,4]
(0,1) [2,1]
(0,0) [1,2]
(1,1) [3,3]
[0,1]
[1,1]
Stage 1
Stage 2
Stage 3

Figure III.1

$u_1^1u_2^1$ \ $u_1^2u_2^2$	00	01	10	11
00	(2,4)	(4,5)	(4,4)	(4,3)
01	(4,6)	(4,5)	(4,6)	(6,5)
10	(3,5)	(3,4)	(1,5)	(3,6)
11	(3,7)	(5,6)	(3,7)	(3,6)

Table III.2

An open-loop Nash equilibrium pair is described in Section II.2 by equations (2.4) and (2.5). Table III.2 is the normal form of the game for the open-loop Nash equilibrium concept (since the strategies are constant strategies). It is easy to verify by inspection that only two entries are open-loop Nash equilibria: The pair of decision sequences (10, 01) with associated costs (3,4), and the pair (11, 11) with associated costs (3,6).

Table III.2 can also be used in obtaining the solution under the open-loop Stackelberg concept described in Section II.3. Under this concept, with player 1 as leader, player 1 chooses a row first, then player 2 chooses a column. The rational reaction set of player 2 is then the entry in each row with the minimal second element. A quick inspection of Table III.2 yields one open-loop Stackelberg solution, satisfying equation (2.13): the pair (10, 01) with associated costs (3,4). As expected from Proposition 2 in Section II.6, the leader's cost 3 is no greater than the cost achieved under either of the open-loop Nash equilibria. It is important to note that Table III.2 is <u>not</u> the normal form of the Stackelberg game; additional rules, such as the order of choices, had to be specified in defining the open-loop Stackelberg solution. As indicated in Section II.5, the normal form for the game would be in terms of reactive strategies for player 2. The optimal reactive strategy for player 2 corresponds to his rational reaction set, and can be expressed as:

$$\begin{aligned} &\underline{u}^2(00) = 11; \quad \underline{u}^2(01) = 01 \\ &\underline{u}^2(10) = 01; \quad \underline{u}^2(11) = 11 \end{aligned} \tag{3.5}$$

Closed-loop and signaling-free equilibria are equilibria in terms of admissible strategies, which have yet to be defined. The sets of admissible strategies for the Nash game are

$$\Gamma_t^i = \{\text{all deterministic maps from } X_t \to U_t^i\} \tag{3.6}$$

It is readily seen that each player has four admissible strategies at stage 2. Identify these four strategies as a,b,c,d where

$$\begin{aligned} &a(0) = 0, \ a(1) = 0 \\ &b(0) = 0, \ b(1) = 1 \\ &c(0) = 1; \ c(1) = 0 \\ &d(0) = 1; \ d(1) = 1 \end{aligned} \tag{3.7}$$

At stage 1, since $x_1 = 0$, there are only two strategies; denote them by their value 0 or 1. Using this terminology, the rational reaction sets V^{ic} can be obtained to each strategy sequence used by a player.

The reaction sets are defined by (in the case of player 2),

$$V^{2c}(\underline{\gamma}^1) = \{\underline{\gamma}^{20} \in \Gamma_1^2 \times \Gamma_2^2 | J^2(\underline{\gamma}^1, \underline{\gamma}^{20}) \leq J^2(\underline{\gamma}^1, \underline{\gamma}^2) \quad \text{for } \underline{\gamma}^2 \in \Gamma_1^2 \times \Gamma_2^2\}. \tag{3.8}$$

Calculating V^{ic} is a simple minimization and can be done using Table III.2. Table III.3 describes the costs associated with each pair of strategy sequences.

	0a	0b	0c	0d	1a	1b	1c	1d
0a	(2,4)	(2,4)	(4,5)	(4,5)	(4,4)	(4,3)	(4,4)	(4,3)
0b	(2,4)	(2,4)	(4,5)	(4,5)	(4,6)	(6,5)	(4,6)	(6,5)
0c	(4,6)	(4,6)	(4,5)	(4,6)	(4,4)	(4,3)	(4,4)	(4,3)
0d	(4,6)	(4,6)	(4,5)	(4,6)	(4,6)	(6,5)	(4,6)	(6,5)
1a	(3,5)	(3,4)	(3,5)	(3,4)	(1,5)	(1,5)	(3,6)	(3,6)
1b	(3,7)	(5,6)	(3,7)	(5,6)	(1,5)	(1,5)	(3,6)	(3,6)
1c	(3,5)	(3,4)	(3,5)	(3,4)	(3,7)	(3,7)	(3,6)	(3,6)
1d	(3,7)	(5,6)	(3,7)	(5,6)	(3,7)	(3,7)	(3,6)	(3,6)

Table III.3

Using Table III.3, the rational reactions of each player can be computed. These are summarized in Table III.4.

Player 2's	Player 1's
V^{2c}(0a) = {1d,1b}	V^{1c}(0a) = {0a,0b}
(0b) = {0a,0b}	(0b) = {0a,0b}
(0c) = {1b,1d}	(0c) = {1a,1b,1c,1d}
(0d) = {0c,0d,1b,1d}	(0d) = {1a,1c}
(1a) = {0b,0d}	(1a) = {1a,1b}
(1b) = {1b,1a}	(1b) = {1a,1b}
(1c) = {0b,0d}	(1c) = {1a,1b,1c,1d}
(1d) = {0b,0d,1c,1d}	(1d) = {1a,1b,1c,1d}

Table III.4 Rational Reactions sets

The closed-loop Nash equilibria are pairs of admissible strategies $(\underline{\gamma}^{1N}, \underline{\gamma}^{2N})$ such that $\underline{\gamma}^{1N} \varepsilon V^{1c}(\underline{\gamma}^{2N} \varepsilon V^{2c}(\underline{\gamma}^{1N})$, as can be seen from equations (2.6) and (2.7) and the definition of the reaction sets. Searching through Table III.4 several equilibria are found. These are listed in Table III.5.

Equilibrium pairs	Costs
(0b,0a), (0b,0b)	(2,4)
(1a,0d), (1c,0d)	(3,4)
(1b,1a), (1b,1b)	(1,5)
(1d,1c), (1d,1d)	(3,6)

Table III.5 List of closed-loop equilibria

As expected from Proposition 3 of Section II.6, the open-loop Nash equilibria (1a,0d) and (1d,1d) are also closed-loop equilibria. Note the relative abundance of closed-loop equilibria; unless there is a previous agreement among the players as to which equilibrium is preferable, this solution concept is not uniquely defined even in this simple game. This is part of the motivation for the introduction of the third class of Nash equilibria in dynamic games, the signaling-free equilibria.

The signaling-free Nash equilibria can be obtained through backwards induction. Assume that the state at stage two is $x_2 = 0$, and consider the remaining game with costs J_2^i starting at $x_2 = 0$. This game is shown in Figure III.6,

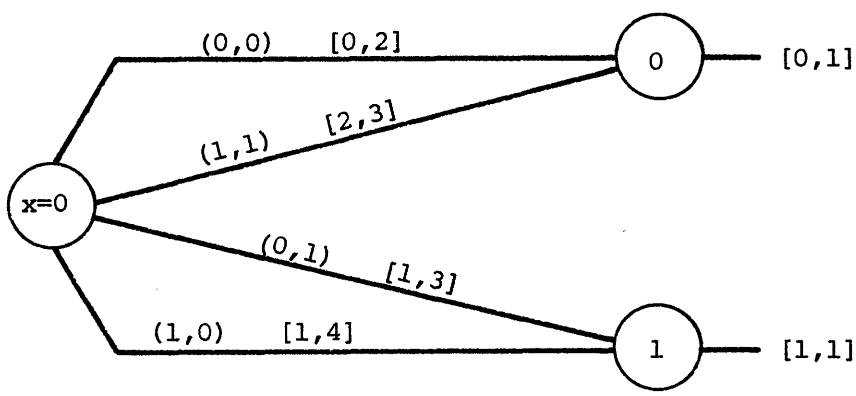

Figure III.6 Game starting at $x_2 = 0$

The Nash equilibrium solution for this truncated game is found by inspection to be $u_2^1 = 0$, $u_2^2 = 0$ with optimal costs (0,3). Similarly, for $x_2 = 1$, the optimal equilibrium decision are $u_2^1 = 0$, $u_1^2 = 1$ with costs [2,2]. Thus, at stage 2, the signaling-free Nash equilibrium strategies are $\gamma_2^1(x) = a$, $\gamma_2^2(x) = b$. Denote the costs-to-go at state x by the functions (I_2^1, I_2^2). Then, the associated costs-to-go are: $(I_2^1, I_2^2)(0) = (0,3)$ and $(I_2^1, I_2^2)(1) = (2,2)$.

The optimal strategies for stage 1 can now be determined, since since any combination of decisions at the first stage results in a transition to a state at the second stage; since the optimal strategies from stage 2 on are known, a complete cost can be associated with decisions at stage 1. The game at stage 1 can thus be reduced to the form expressed in Figure III.7.

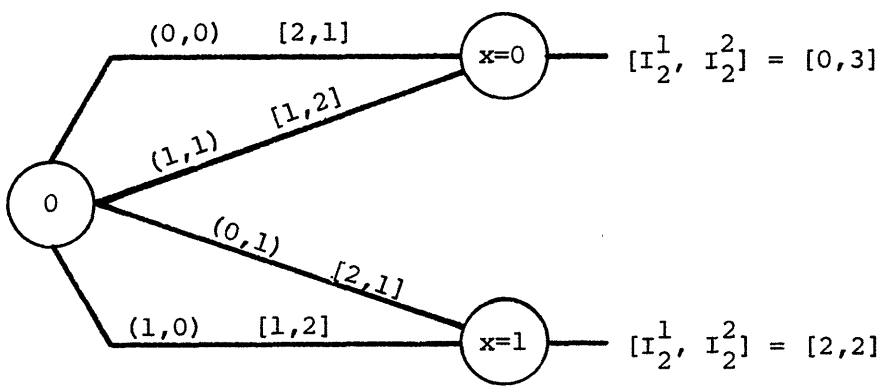

Figure III.7 Reduced game at stage 1

The reduced game of Figure III.7 has no equilibrium in terms of pure strategies. Although an equilibrium in terms of mixed strategies is guaranteed by Nash's extension of the minmax theorem (N1,N2), it is beyond the scope of this work as mixed strategies are not admissible strategies. Nevertheless, the fact that no signaling-free Nash equilibria exist does not imply that other closed-loop equilibria exist, as evidenced by Table III.5. Throughout a large section of the dynamic game literature, the search for closed-loop equilibria is performed by finding signaling-free equilibria recursively, overlooking the fact that these two solution concepts are not equivalent. Only in the special case of deterministic games and other cases discussed in Chapter IV can closed-loop equilibria be found in recursive fashion.

The closed-loop Stackelberg solution can be determined using Table III.4, since the follower is constrained to choose strategies in his rational reaction set $V^{2c}(\underline{\gamma}^1)$. Looking through the potential choices for $\underline{\gamma}^1$ and $\underline{\gamma}^2$, the optimal closed-loop Stackelberg solution pairs are (lb,lb) and (lb,la), both resulting in a cost vector of (1,5). This is the entry in Table III.3 with the lowest cost for the leader, underlying the leader's dominance under this solution concept.

The last two solution concepts that will be discussed in the context of this example are the closed-loop and the signaling-free Stackelberg equilibria. The difference between these two concepts and the Nash concepts previously discussed is the set of the follower's admissible strategies. For the Stackelberg concepts, Γ_1^2 and Γ_2^2 are defined as:

$$\Gamma_1^2 = \{\gamma_1^2 | \gamma_1^2 : X_1 \times U_1^1 \rightarrow U_1^2\}$$
$$\Gamma_2^2 = \{\gamma_2^2 | \gamma_2^2 : X_2 \times U_2^1 \rightarrow U_2^2\} \quad (3.8)$$

It is easy to see Γ_1^2 has four elements and Γ_2^2 has 16 elements. Denote the four elements of Γ_1^2 as a, b, c, and d, defined previously in equation (3.7), and the elements of Γ_2^2 by four binary numbers representing assignments of u_2^2 to elements of $X_2 \times U_2^1$. For example the strategy 0100 is the map $\gamma_2^2(x_2, u_2^1)$ described as:

$$\begin{array}{cc} x_2, u_2^1 & u_2^2 \end{array}$$
$$0100 \triangleq \left\{ \begin{array}{ccc} 0\ 0 & \rightarrow & 0 \\ 0\ 1 & \rightarrow & 1 \\ 1\ 0 & \rightarrow & 0 \\ 1\ 1 & \rightarrow & 0 \end{array} \right. \quad (3.9)$$

With this nomenclature, a table similar to Table III.4 can be obtained listing the rational reaction sets of each player. Table III.8 lists the possible closed-loop Stackelberg equilibria:

$\underline{\gamma}^1$	$\underline{\gamma}^2$	Costs
0b	a,0000 through a0111	(2,4)
1a,1c	a1010-a1111	(3,4)
	c0010,c0011	
	c0110,c0111	
	c1010,c1011	
	c1110,c1111	
1b	b0000-b0111	(1,5)
	d0000-d0111	
1d	b1100 $\rightarrow$ b1111	
	d1100 $\rightarrow$ d1111	

Table III.8 Closed-loop Stackelberg equilibria

Note the presence of the open-loop equilibria (1a,a1111) and (1d,d1111) in Table III.8, as implied by Proposition 3 of Section II.6. As in the closed-loop Nash solution concept, the abundance of closed-loop equilibria that are not cost equivalent and the redundancy of a lot of the strategies present make this solution concept undesirable for dynamic games.

The signaling-free equilibria can be found recursively in a fashion similar to the signaling-free Nash equilibria. The difference between these two concepts is the added assumption in the Stackelberg solution concept that the leader makes his decision known to the follower at each stage before the follower chooses his decision. Thus, using backwards induction, the optimal equilibrium strategies can be defined at stage 2 by

$$\gamma_2^{2*}(x_2,\ u_2^1) = \underset{u_2^2}{\text{argmin}} \left(h_2^2(x_2,\ u_2^1,\ u_2^2) + h_3^2(f_2(x_2,\ u_2^1,\ u_2^2))\right) \tag{3.10}$$

Similar equations can be written for γ_2^{1*}, γ_1^{2*}, and γ_1^{1*} in sequential order. Solution of these equations yields a unique signaling-free Stackelberg equilibrium pair:

$$\underline{\gamma}^{1*} = 1a,\ \underline{\gamma}^{2*} = c0011, \text{ with costs } (3,4) \tag{3.11}$$

As indicated by Proposition 4 of Section II.6, this equilibrium is also a closed loop equilibrium, shown in Table III.8.

Thus far, the solution concepts described in Chapter II have been applied to a simple finite example while verifying that the results in Section II.6 hold. Even in this simple multi-stage game, the inadequacy of some solution concepts is highlighted. In particular, the closed-loop equilibria are so numerous and have different associated costs that additional restrictions are needed in order to specify a unique solution. In deterministic games, Proposition 4 of Section II.6 establishes the signaling-free equilibria as closed-loop equilibria, thus providing an example of such additional restrictions. In the next sections, these solution concepts are used in some continuous games.

III.2 A Continuous State and Decision Multi-Stage Game

Consider the multi-stage game whose evolution equation is given as

$$x_1 = 1 + u_0 + v_0$$
$$x_2 = x_1 + u_1 \quad , \quad x_i \in R \tag{3.12}$$

The cost functions for player 1(u) and player 2(v) are:

$$J_1 = (x_2)^2 + (u_0)^2 + (u_1)^2$$
$$J_2 = (x_2)^2 + 1/2(u_1)^2 + (v_0)^2 \tag{3.13}$$

The closed-loop Nash equilibria of this game have been studied by Tamer Basar (Ba3, Ba4). He deals with a slightly more general version of this game, which he uses in showing the non-uniqueness of closed-loop Nash equilibria; in particular, he shows the existence of nonlinear equilibrium strategies in this game, a linear, quadratic cost problem. In order to exhibit this multiplicity of equilibria, he increases the set of admissible strategies to include "memory-dependent" strategies; mathematically, the sets of admissible strategies considered are

$$\Gamma_0^1 = \{\gamma_0^1 | \gamma_0^1 : X_0 \rightarrow R\} \qquad \Gamma_0^2 = \{\gamma_0^2 | \gamma_0^2 : X_0 \rightarrow R\}$$

$$\Gamma_1^1 = \{\gamma_1^1 | \gamma_1^1 : X_0^1 \times U_0^1 \times U_0^2 \times X_1 \rightarrow R,\ \gamma_1^1 \text{ is continuously differentiable}\}$$

The memory-dependent aspect of the strategies includes a certain redundancy which enables Basar to find strategies which are equal in value but different in representation. Although his conclusions that the closed-loop equilibrium concept in Nash games is not an adequate solution concept in this example are similar to those discussed in Section III.1, it is not necessary to deal with nonlinear strategies. It is shown below that even restricting to linear strategies does not achieve uniqueness of the closed-loop equilibria.

Let the set of admissible strategies Γ_1^1 be the set of all Borel measurable maps from the reals into the reals. The state space X_0 contains only one element, $X_0 = 1$, $X_1 = U_1^1 = U_0^1 = U_0^2 = R$. The claim is that Nash equilibrium strategies of the form $u_0 = b$, $u_1 = ax_1 + c$ and $v_0 = d$ exist for several choices of a, b, c, d. These strategies are neither nonlinear nor memory-dependent. In particular, Basar exhibits the strategy $u_1 = -\frac{1}{2}x_1$, $u_0 = -\frac{4}{15}$ and $v_0 = -\frac{1}{5}$ which is the signaling-free Nash strategy for this game with costs $J^{1*} = .2133$, $J^{2*} = .1467$.

In particular, let a = c. The rational reaction set of player 2 is $V^{2c}(b, a(x_1 + 1))$, determined as

$$V^{2c}(b,\ ax_1 + a) = \{v_0^* | J^2(b,\ ax_1 + a^2;\ v_0^*) \leq J^2(b,\ ax_1 + a;\ v_0) \quad \forall\, v_0 \in R\} \tag{3.14}$$

Now,
$$\begin{aligned} J^2(b, (ax_1 + a);\ v_0) &= (x_2)^2 + v_0{}^2 + \tfrac{1}{2}u_1^2 \\ &= (x_1 + ax_1 + a)^2 + v_0^2 + \tfrac{1}{2}(ax_1 + a)^2 \\ &= \{(1+v_0+b)(1+a) + a\}^2 + v_0^2 + \tfrac{1}{2}a^2(2+v_0+b)^2 \end{aligned}$$

This expression has a unique minimum in terms of v_0, given as the solution of

$$2(1 + a)((1 + v_0 + b)(1 + a) + a) + 2v_0 + a^2(2 + v_0 + b) = 0 \tag{3.15}$$

It is easy to inspect that, given constants a and b, (3.15) implies v_0 is uniquely defined. The rational reaction set $V^{1c}(v_0)$ can be similarly computed.

$$J^1(u_0, \gamma_1^1; v_0) = (1 + u_0 + v_0 + \gamma_1^1(1 + u_0 + v_0))^2 + u_0^2 + (\gamma_1^1(1 + u_0 + v_0))^2 \tag{3.16}$$

The minimization of (3.16) is a classical optimal control problem and can be obtained using dynamic programming. This procedure is:

$$\min_{u_0,\gamma_1^1} J^1(u_0, \gamma_1^1; v_0) = \min_{u_0} \{\min_{\gamma_1^1} J^1(u_0, \gamma_1^1; v_0)\}$$

$$= \min_{u_0} J^1(u_0, -\tfrac{1}{2} x_1; v_0)$$

Since $J_1^1(u_0, \gamma_1^1; v_0)$ is minimized by $\gamma_1^1 = -\frac{1}{2} x_1$

$$= \min_{u_0} \{\tfrac{1}{2}(1 + u_0 + v_0)^2 + u_0^2\}$$

The minimizing value of u_0 is $-\frac{1}{3}(1 + v_0)$. Furthermore, (3.16) is also minimized by $\gamma_1^1 = -\frac{1}{3}(1 + v_0) = -\frac{1}{2} x_1$ or any other strategy which has this value along the optimal trajectory. In particular, $\gamma_1^1 = -\frac{(v_0+1)}{5+2v_0}(x_1 + 1)$ is a strategy with the representation $\gamma_1^1 = a(x_1 + 1)$ which minimizes (3.16) and thus belongs to $V^{1c}(v_0)$. The closed-loop equilibrium is then defined by:

$$u_0 = -\frac{1}{3}(1 + v_0) = b \tag{3.17}$$

$$\gamma_1^1 = -\frac{(v_0+1)}{5+2v_0}(x_1 + 1) = a(x_1 + 1) \tag{3.18}$$

$$2(1 + a)\{(1 + v_0 + b)(1 + a) + a\} + 2v_0 + a^2(2 + v_0 + b) = 0 \tag{3.19}$$

Substitute (3.17) and (3.18) into (3.19) to get

$$\frac{2(v_0+4)}{(5+2v_0)^2}\left(\frac{2}{3}(v_0+1)(v_0+4)-(v_0+1)\right) + 2v_0 + \frac{(v_0 + 1)^2}{(5+2v_0)^2}\left(1 + \frac{2}{3}(v_0 + 1)\right) = 0 \tag{3.20}$$

Simplifying,

$$\frac{1}{3(2v_0 + 5)}\{2(v_0 + 4)(v_0 + 1) + 6(v_0)(2v_0 + 5) + (v_0 + 1)^2\} = 0$$

Expand the inner expression in powers of v_0

$$15\, v_0^2 + 42v_0 + 9 = 0 \tag{3.21}$$

Equation (3.21) has two real solutions, $v_0 = -.2338066$ and $v_0 = -2.5662$, each of which in combination with (3.17) and (3.18) would yield equilibrium strategies. Consider only the equilibrium where $v_0 = -.2338066$. The corresponding values of u_0 and a are $u_0 = -.2553968$ and $a = -.169048$. The trajectory produced by these decision is $x_0 = 1$, $x_1 = .510794$, and $x_2 = .255397$. Therefore, the costs associated with this equilibrium are

$$J^{1a} = u_0^2 + u_1^2 + x_2^2 = .19568$$
$$J^{2a} = v_0^2 + \frac{1}{2}u_1^2 + x_2^2 = .152507 \qquad (3.22)$$

In this new equilibrium, player 1 has reduced his cost, and player 2 has suffered an increase in cost over the signaling-free Nash equilibrium costs J^{1*}, J^{2*}. Another equilibrium for this game is obtained through the open-loop Nash equilibrium, as shown in Proposition 3 of Section II.6. The open-loop Nash equilibrium values are: $u_0 = u_1 = v_0 = -.25$, resulting in the trajectory $x_0 = 1$, $x_1 = .5$, $x_2 = .25$ and costs $J_1^0 = .1875$, $J_2^0 = .15625$ which again exhibit the behavior that the first player reduces his cost, player 2 increases his cost over the signaling-free equilibrium's costs.

The multiplicity of equilibria was observed by Basar (Ba3) while examining nonlinear equilibria. The conclusions were that additional rules had to be specified in order to define a solution. The signaling-free equilibria include such a set of additional rules. The behavior of the costs for the three equilibria found earlier can be explained in terms of the basic assumptions behind each equilibrium pair. In all three equilibria, it is optimal to have the value of the decision u_1 equal to $-\frac{1}{2}x_1$. This is easily verified looking at the various equilibria. What differs is the representation given to this value. For signaling-free equilibria, it is represented as $u_1 = -\frac{1}{2}x_1$, while in the others as $u_1 = a(x_1 + 1)$ and $u_1 = c$. The interpretation of these three representations is crucial in understanding the behavior of the costs: the representation $u_1 = -\frac{1}{2}x_1$ implies no assumptions at all about the previous decisions made in the game. It simply states that, no matter what is done previously, the decision u_1 will be made after seeing what x_1 is. The point is that this representation does not assume any type of optimal decision previously in order to be optimal.

In contrast, the other two representations assume certain optimal decisions u_0 and v_0 in order for them to have the value $u_1 = -\frac{1}{2}x_1$. The open-loop Nash equilibrium carries the strongest assumptions of any of the three, as it assumes exact values of u_0 and v_o for u_1 to be $-\frac{1}{2}x_1$. Thus, the last two representations of u_1 involve assumptions with respect to u_0 and v_0; limiting the classes of admissible strategies a priori for γ_1^1 to constant strategies, or as $a(x_1 + 1)$, affects the choice of v_0; this corresponds to allowing player 1 to make a decision (restricting his class of strategies) prior to player 2's choices. The reduction in cost is achieved because the restricted class of strategies is still able to produce the same decisions as were used in the unrestricted equilibrium solutions, and because player 2 accepts player 1's prior commitment to this class of strategies so his decision is influenced by the change in representation.

The behavior of the costs is similar to that exhibited in the closed-loop Stackelberg solution concept. The dominance of the leader enables him to specify the choices of his future strategies exactly before the follower can choose his present strategies. Thus, the leader influences the follower's choice through the future strategies, and thus achieves a reduced cost compared to equilibrium concepts. Since player 1 is the only decision maker in stage 2 of this game, the restriction to a given class of strategies exerts a similar influence on the follower's choice in the equilibria discussed earlier, and thus achieves a reduction in the leader's cost similar to that indicated by Proposition 2 of Section II.6.

This phenomenon is exhibited best considering memory-dependent strategies, as Basar did. The non-linear equilibria found in (Ba3) are of the form

$$u_1 = -\frac{1}{2}x_1 + K(x_1^2 - (1 + \bar{u}_0 + \bar{v}_0)^2) \tag{3.23}$$

where $\bar{u}_0$ and $\bar{v}_0$ are constants equal in value to the equilibrium values of u_0, v_0 and K is a constant. In the equilibrium trajectory, $u_1 = -\frac{1}{2}x_1$ as expected. However, restricting the admissible strategies to those in (3.23) affects the choices of u_0 and v_0 as for u_1 to be in equlibrium, u_0 and v_0 must take on the values $\bar{u}_0$, $\bar{v}_0$. The strategy in (3.23) poses a "threat" to player 2; his choice is influenced by the knowledge that in the future he is expected to play $\bar{u}_0$, $\bar{v}_0$. The clearest example of a threat would be restricting u_1 to strategies of the form

$$u_1 = -\frac{1}{2}x_1 + K(v_0 - c_0) \tag{3.24}$$

where K is very large and c_0 is a constant. As K increases beyond bounds, the equilibrium value of v_0 will approach c_0.

Player 1's reduction of cost is a byproduct of his choice of γ_1^1 prior to the choice of v_0, thus exerting influence on player 2's choice. The ambiguity in the closed-loop Nash equilibrium concept enables player 1 to choose γ_1^1 in this fashion. More appropriate is the closed-loop Stackelberg solution to the game, which specifies that this choice is made before v_0 is chosen.

For this game, the closed-loop Stackelberg solution is obtained as follows:

Restrict Γ_1^1 to be $\Gamma_1^1 = \{\gamma_1^1 | \gamma_1^1(x_1) = ax_1, \ a \in R\}$

Then, the follower's reaction set to strategies (u_0, ax_1) is given by

$$V^{2c} = \{v_0^* | J^2(u_0, ax_1; v_0^*) \leq J^2(u_0, ax_1; v_0) \quad \forall v_0 \in R\}$$

Now, $J^2(u_0, ax_1; v_0) = v_0^2 + ((1+a)^2 + \frac{1}{2} a^2)(1 + u_0 + v_0)^2$

Then, v_0^* is given by the solution of

$$2v_0^* + 2((1 + a)^2 + \frac{1}{2} a^2)(1 + u_0 + v_0^*) = 0.$$

Thus,

$$v_0^* = - \frac{(\frac{1}{2}a^2 + (1+a)^2)}{(1 + \frac{1}{2}a^2 + (1+a)^2)} (1 + u_0) \tag{3.25}$$

and

$$x_1 = 1 + u_0 + v_0^* = \frac{1}{(1 + \frac{1}{2}a^2 + (1+a)^2)} (1 + u_0)$$

It is easy to see from (3.25) that, even when γ_1^1 is restricted to linear strategies, the leader can achieve as low a cost as he desires by letting $u_0 = 0$ and a be very large. As $a \to \infty$, $x_1 \to 0$ and $ax_1 \to 0$ also.

Thus,

$$J^1_{CLS} = u_0^2 + x_2^2 + u_1^2 \to 0, \; J^2_{CLS} = v_0^2 + x_2^2 + \frac{1}{2} u_1^2 \to 1 \quad \text{as } a \to \infty$$

The interpretation of this behavior is as follows: the leader threatens to increase the state x_1 to $(1+a)x_0$; when a is very large, the follower's cost is increased greatly. Thus, as a increases, the follower will try to make x_1 as small as possible in lieu of the leader's threat. In making this threat, the leader achieves a reduction in cost similar to that seen previously in the Nash equilibria. Furthermore, this property is preserved even if player 2 was able to make decisions in stage 2, as long as player 1 is able to choose his strategy at stage 2 before player 2 chooses his strategy at stage 1.

The open-loop Stackelberg solution concept also assumes dominance by the leader, but the admissible strategies are restricted to constant strategies, so that the possible threats are limited. The open-loop Stackelberg solution in this example is found as follows:

The follower's reaction set $V^2(u_0, u_1)$ is found by minimizing

$$J^2(u_0, u_1, v_0) = u_1^2 + v_0^2 + (1 + u_0 + v_0 + u_1)^2$$

Thus,

$$V^2(u_0, u_1) = \{v_0 \in R \,|\, v_0 = -\frac{1}{2}(1 + u_0 + u_1)\}$$

Then, from equation (2.13),

$$(u_0^*, u_1^*) = \underset{\substack{u_0, u_1 \in R \\ v_0 \in V^2(u_0,u_1)}}{\text{argmin}} \; J^1(u_0, u_1; v_0) =$$

$$= \underset{u_0,u_1}{\text{argmin}} \{u_0^2 + u_1^2 + (1 + u_0 + u_1)^2 \tfrac{1}{4}\}$$

Differentiating,

$$2u_0 + \frac{1}{2}(u_0 + u_1 + 1) = 0$$

$$2u_1 + \frac{1}{2}(1 + u_0 + u_1) = 0$$

Therefore $u_0 = u_1 = -\frac{1}{6}$ and $v_0 = -\frac{1}{3}$.

The optimal costs are $J^{1*}_{OLS} = .16667$, $J^{2*}_{OLS} = .25$. Again, the leader's dominance is seen in the reduction of the leader's cost J^{1*}_{OLS} relative to the open-loop equilibrium solutions found earlier. In the next section an example with imperfect information is discussed to see how the dominance implied in some of these solution concepts is affected by uncertainties in the system.

III.3 A Game with Uncertainties

Consider the game whose evolution equation is given by

$$\begin{aligned} x_1 &= x_0 + u_0 + v_0 + \theta_0 \\ x_2 &= x_1 + u_1 \end{aligned} \tag{3.26}$$

where θ_0 is a zero-mean random variable with variance ε. The costs associated with this game are:

$$J^1 = E_\theta\{x_2^2 + u_0^2 + u_1^2\}$$

$$J^2 = E_\theta\{x_2^2 + v_0^2 + \tfrac{1}{2}u_1^2\}$$

The set of admissible strategies Γ_1^1 is the set of all Borel measurable functions $\gamma_1^1(x_1)$. That is, the value of x_1 is known to player 1 at stage 1, but not at the beginning of the game. Signaling-free Stackelberg equilibria for this game can be obtained as:

$$\gamma_1^{1*}(x_1) = \underset{\gamma_1^1 \in \Gamma_1^1}{\text{argmin}} \ \{E(x_2^2 + u_1^2)|x_1\}$$

$$= \underset{u_1}{\text{argmin}} \ \{x_1^2 + 2x_1u_1 + u_1^2 + u_1^2\} = -\frac{1}{2}x_1$$

$$\gamma_0^{2*}(x_0,u_0) = \underset{v_0}{\text{argmin}} \ E\{(x_2^2 + v_0^2 + \tfrac{1}{2}u_1^2|u_1,u_1 = -\frac{1}{2}x_1\}$$

$$= \underset{v_0}{\text{argmin}} \ E\{v_0^2 + \tfrac{3}{8}(x_0 + u_0 + v_0 + \theta_0)^2\}$$

$$= \underset{v_0}{\text{argmin}} \ \{v_0^2 + \tfrac{3}{8}(x_0 + u_0 + v_0)^2 + \varepsilon \cdot \tfrac{3}{8}\}$$

$$= -\frac{3}{11}(x_0 + u_0)$$

$$\gamma_0^{1*}(x_0) = \underset{u_0}{\text{argmin}} \ \{E(x_2^2 + u_1^2 + u_0^2)|u_1 = -\tfrac{1}{2}x_1, \ v_0 = -\tfrac{3}{11}(x_0+u_0)\}$$

$$= \underset{u_0}{\text{argmin}} \ E\{\tfrac{1}{2}(\tfrac{8}{11}(x_0 + u_0) + \theta_0)^2 + u_0^2)\}$$

$$= \underset{u_0}{\text{argmin}} \ \{\tfrac{1}{2}\varepsilon + \tfrac{32}{121}(x_0 + u_0)^2 + u_0^2\}$$

$$= -.20915 \ x_0$$

Let $x_0 = 1$ as in the previous section. Then

$$u_0^* = -.20915$$

$$v_0^* = -\frac{3}{11}(1_0 + u_0) = -.215686$$

$$x_1 = 1 + u_0^* + v_0^* + \theta_0 = .5752 + \theta_0$$

$$x_2 = \frac{1}{2}x_1 = -u_1 = .287582 + \theta_0/2$$

with costs $J_e^{1*} = .20915 + \frac{\varepsilon}{2}$, $J_e^{2*} = .17058 + \frac{3}{8}\varepsilon$.

In finding other closed-loop equilibria the approach of the previous section cannot be used, since the uncertainty in θ_0 makes the representation $u_1 = -\frac{1}{2}x_1$ unique. Basar (Ba4) points out that uncertainties in the system eliminate the multiplicity of equilibria exhibited in the previous section for the Nash game. For instance, the open-loop Nash equilibrium is no longer a closed-loop equilibrium, as the optimal reaction set $V^{1c}(.25)$ does not include the strategy $u_0 = .25$, $u_1 = .25$ because the strategy $u_0 = .25$, $u_1 = -\frac{1}{2}x_1$ achieves a lower value for the leader's cost.

The uniqueness of representation in γ_1^1 eliminates the possibility of influence introduced through the different representations discussed in the previous section. However, the dominance of the closed-loop Stackelberg solution concept is still present, although somewhat affected, as seen below. As before, the leader's strategy set is $\Gamma_1^1 = \{\gamma_1^1 | \gamma_1^1(x_1) = ax_1\}$. The follower's reaction set $V^{2c}(u_0, ax_1)$ is determined by $v_0^*(u_0, ax_1) \in V^{2c}(u_0, ax_1)$, where

$$v_0(u_0, ax_1) = \underset{v}{\operatorname{argmin}}\, E\{x_2^2 + v_0^2 + \frac{1}{2}u_1^2 | u_0, u_1 = ax_1\} =$$

$$= \underset{v}{\operatorname{argmin}}\ E\{(1+a)^2 x_1^2 + \frac{1}{2} a^2 x_1^2 + v_0^2\}$$

$$= \underset{v}{\operatorname{argmin}}\ \{(\frac{a^2}{2} + (1+a)^2)(x_0 + u_0 + v_0)$$

$$+ ((1+a)^2 + \frac{a^2}{2})\varepsilon + v_0^2\}$$

$$= - \frac{(\frac{a^2}{2} + (1+a)^2)}{1 + \frac{a^2}{2} + (1+a)^2} (x_0 + u_0).$$

Thus,

$$\min_{\substack{u_0,\gamma_1^1 \\ v_0 \,\varepsilon\, V^{2c}(u_0,ax_1)}} J^1(u_0,ax_1;\ v_0) = \min_{a,u_0} E\{(1+a)^2 x_1^2 + a^2 x_1^2$$

$$+ u_0^2 | v_0 \,\varepsilon\, V^{2c}(u_0,ax_1)\}$$

$$= \min_{a,u_0} \left\{ \frac{((1+a)^2 + a^2)(x_0+u_0)^2}{((1+a)^2 + \frac{a^2}{2} + 1)^2} + u_0^2 \right.$$

$$\left. + (1+a)^2\varepsilon + a^2\varepsilon \right\} \tag{3.27}$$

The difference between this minimization and that of Section III.2 is the presence of the term $(1+a)^2\varepsilon + a^2\varepsilon$ which goes to ∞ as a gets large, so that an unbounded solution does not exist. Differentiating (3.27) with respect to u_0, and equating to 0,

$$\frac{(1+a)^2 + a^2}{((1+a)^2 + \frac{a^2}{2} + 1)^2} (x_0 + u_0^*) + u_0^* = 0 \tag{3.28}$$

Thus,

$$u_0^* = - \frac{(a^2 + (1+a)^2)\ x_0}{((1+a)^2 + \frac{a^2}{2} + 1)^2 + a^2 + (1+a)^2}\ .$$

Substituting into (3.27) gives

$$J_{CLS}^{1*} = \min_a \left\{ \frac{((1+a)^2 + a^2)((1+a)^2 + \frac{a^2}{2} + 1)^2 + (a^2 + (1+a)^2)^2}{\left(((1+a)^2 + \frac{a^2}{2} + 1)^2 + a^2 + (a+1)^2\right)^2} \right.$$

$$\left. + (a^2 + (1+a)^2)\varepsilon \right\}$$

$$= \min_a \left\{ \frac{a^2 + (a+1)^2}{((1+a)^2 + \frac{a^2}{2} + 1)^2 + a^2 + (a+1)^2} + a^2\varepsilon + (a+1)^2\varepsilon \right\}$$

The minimizing value of a is a solution of

$$\frac{[(\frac{3}{2}a^2+2a+2)^2+2a^2+2a+1](4a+2)-(2a^2+2a+1)((3a^2+4a+4)(3a+2)+(4a+2))}{((\frac{3}{2}a^2+2a+2)^2+(2a^2+2a+1))^2}$$

$$+ (4a+2)\varepsilon = 0$$

This expression reduces to:

$$\frac{(-a)(\frac{3}{2}a^2 + 2a + 2)(6a^2 + 9a + 2)}{((\frac{3}{2}a^2 + 2a + 2)^2 + (2a^2 + 2a + 1))^2} + (4a + 2)\varepsilon = 0 \qquad (3.29)$$

If $\varepsilon = 0$, then (3.29) reduces to a quintic polynomial with three real roots, none of which is the true minimum as all of them yield costs of order 1. Assuming $0 < \varepsilon << 1$, (3.29) yields a 9th order polynomial. Treating this as a singular perturbation problem and trying to find the roots of this polynomial in terms of ε, the five original roots are changed little, since $\varepsilon << 1$. The other four roots are of large magnitude. Expanding (3.29),

$$-a(\tfrac{3}{2}a^2 + 2a + 2)(6a^2 + 9a + 2)$$
$$+ \varepsilon(4a + 2)((\tfrac{3}{2}a^2 + 2a + 1)^2 + 2a^2 + 2a + 1)^2 = 0 \qquad (3.30)$$

Since $|a| >> 1$, (3.30) can be approximated as

$$-9a^5 + 4\varepsilon \cdot \frac{81}{16} a^9 = 0 \qquad (3.31)$$

so, $\varepsilon a^4 - \frac{4}{9} = 0$. Thus, $a = (\frac{2}{3})^{\frac{1}{2}} \varepsilon^{-\frac{1}{4}} >> 1$.

This is only a zeroth order approximation, but sufficient for purposes of illustration. The cost achieved under these values of a is approximately

$$J_{CLS}^{1*} \approx \frac{\frac{4}{3}\ \varepsilon^{-\frac{1}{2}}}{\frac{1}{\varepsilon} + \frac{4}{3}\varepsilon^{-\frac{1}{2}}} + \frac{4}{3}\varepsilon^{\frac{1}{2}} \approx \frac{4}{3}\varepsilon^{\frac{1}{2}}\ (2 - \frac{4}{3}\varepsilon^{\frac{1}{2}}) \qquad (3.32)$$

A more detailed expansion of this cost yields a slight difference for the two possible signs of a; the negative sign yields a slightly lower cost. This expansion is not shown as it is not relevant to the question at hand. The important point is that, even in the presence of uncertainty, the leader was able to exercise a threat on the follower, but tempered by the parameters of the uncertainty. When no noise was present the leader was able to reduce his cost to 0 by threatening to increase the follower's cost without bounds. In the presence of even a minimal uncertainty, the leader can only threaten up to a point, after which it is not advantageous for him to do so. This is in contrast with the other types of threats discussed in the previous section, which were completely eliminated by the uncertainty in the system. Table III.9 summarizes the results of Sections III.2 and III.3.

III.4 Discussion

The example presented in this chapter were chosen to illustrate the various solution concepts discussed in the previous chapter, and to highlight several properties of some solution concepts. These examples include enough sophistication to cover most of the relevant cases. In general the major point stressed is the ambiguity of closed-loop equilibrium concepts when applied to dynamic deterministic games. As indicated in Chapter 2, these concepts are expressed in terms of games in normal form, which do not specify the order in which strategies are chosen, thus enabling one player by restricting his strategy choice, to exercise a threat over his opponent.

The cause of this ambiguity is the prior commitment of the players to their strategies. That is, when player 1 announces that he will restrict his choice at stage 1 a priori, will he still restrict his choice at stage 1 once he observes that, given what the players did

Game, Solution Concept	Player 1's Optimal Strategy and Cost		Player 2's Optimal Strategy and Cost	
Deterministic, signaling-free Nash equilibria	$u_0 = -\frac{4}{15}$ $u_1 = -\frac{1}{2} x_1$	$J^1 = .2133$	$v_0 = -\frac{1}{5}$	$J^2 = .1467$
Deterministic, open-loop Nash equilibria	$u_0 = -\frac{1}{4} = u_1$	$J^1 = .1875$	$v_0 = -\frac{1}{4}$	$J^2 = .15625$
Deterministic, closed-loop Nash equilibria	$u_0 = -.255397$ $u_1 = -.169(x_1 + 1)$	$J^1 = .19568$	$v_0 = -.2338$	$J^2 = .1525$
Deterministic, open-loop Stackelberg solution	$u_0 = -\frac{1}{6}$ $u_1 = -\frac{1}{6}$	$J^1 = .16667$	$v_0 = -\frac{1}{3}$	$J^2 = .25$
Deterministic, closed-loop Stackelberg solution	$u_0 = 0$ $u_1 = ax_1,\ a >> 1$	$J^1 \approx 0$	$v_0 \approx -1.$	$J^2 \approx 1.$
Stochastic, signaling-free Stackelberg equilibrium	$u_0 = -.20915$ $u_1 = -\frac{1}{2} x_1$	$J^1 = .20915 + \frac{\varepsilon}{2}$	$v_0 = -.2157$	$J^2 = .171 + \frac{3}{8}\varepsilon$
Stochastic, closed-loop Stackelberg equilibrium	$u_0 \approx -\frac{4}{3}\varepsilon^{+1/2}$ $u_1 = (\frac{2}{3})^{1/2}\varepsilon^{-1/4} x_1$	$J^1 \approx \frac{8}{3}\varepsilon^{1/2}$	$v_0 \approx -1+\frac{4}{3}\varepsilon^{1/2}$	$J^2 \approx 1 - \frac{4}{3}\varepsilon^{1/2}$

Table III.9

at stage 0, this strategy restriction is no longer optimal? The closed-loop equilibria assume that this prior commitment is there, as do the open- and closed-loop Stackelberg solutions. The signaling-free equilibria, on the other hand, do not assume prior commitment of any players to their strategies. The ultimate version of this effect is for a player to choose his strategy sequence completely before the other player has chosen his; this corresponds to the dominant solution in the closed-loop Stackelberg concept. Aumann and Maschler (Au2), Ho (H4) have discussed the inadequacies of the normal form of the game in representing dynamic games. These ambiguities indicate that solution concepts other than closed-loop equilibria, such as the signaling-free equilibria, are more appropriate. Furthermore, the examples show how even equilibrium solutions can have dominance properties in deterministic games. However, this dominance is eliminated in non-deterministic games as seen in the example of the previous section.

The dominance exhibited by the closed-loop Stackelberg solution concept is not eliminated by uncertainties in the system. The leader's dominance in deterministic games is complete, achieving the lowest possible cost as if both players were minimizing the leader's cost. However, in nondeterministic games, this dominance is weakened because of the uncertainties in the game, as seen in the previous section.

In nondeterministic games, the closed-loop equilibrium solution concept is less ambiguous than in deterministic games. However, the introduction of uncertainty brings up the notion of information and communication, problems which have not been considered yet. The next chapter will deal with closed-loop and signaling-free equilibria in non-deterministic games.

CHAPTER IV

IV.1 Introduction

Although the mathematical framework discussed in Chapter II incorporated the concept of uncertainty and information, most of the results presented in Chapters II and III have been restricted to deterministic games. In this chapter the closed-loop equilibrium and the signaling-free equilibrium concepts are applied to non-deterministic Stackelberg games. Several important differences between these concepts are highlighted by the stochastic nature of the games. Most of the work in the earlier sections of this chapter is based on Hans Witsenhausen's work (Wt2, 3, 4) on nonclassical control theory.

IV.2 A Standard Form for Closed-loop Equilibria (Wt4)

When dealing with closed-loop equilibria in nondeterministic games, it will be easier to discuss problems formulated in a manner which facilitates dealing with questions of measurability and information. The following standard form is adequate for the purposes of studying pure strategy closed-loop Stackelberg equilibria and is based on a similar form proposed by Witsenhausen for stochastic control (Wt4).

Let (X_t, Σ_t), (U_t, F_t), $t = 0, \ldots 2N$ be measurable spaces. For $t = 0, \ldots 2N-1$, let D_t be a σ-field contained in Σ_t. The state transition function $f_t:X_t \times U_t \rightarrow x_{t+1}$ is measurable from $\Sigma_t \times F_t \rightarrow \Sigma_{t+1}$.

Let π_0 be a probability measure on (X_0, Σ_0) and ϕ^1_{2N}, ϕ^2_{2N} be bounded, real-valued functions defined and measurable on (X_{2N}, Σ_{2N}). The set of admissible strategies Γ_t is the set of all functions $\gamma_t:X_t \rightarrow U_t$ such that γ_t is measurable from D_t to F_t. Let $\Gamma = \Gamma_0 \times \ldots \times \Gamma_{2N-1}$.

For any choice of $\underline{\gamma} \in \Gamma$, the equations $u_t = \gamma_t(x_t)$, $x_{t+1} = f_t(x_t, u_t)$ determine the variables u_t, x_t as functions of x_0 in the probability space (X_0, Σ_0, π_0). Thus, x_t and u_t are well-defined random variables with probabilities induced from this original probability space; similarly, the quantities $\phi^1_{2N}(x_{2N})$, $\phi^2_{2N}(x_{2N})$ are also well-defined random variables.

The problem of finding a closed-loop equilibrium between two players can be formulated as follows: Let $\Gamma^0 = \Gamma_1 \times \Gamma_3 \times \ldots \times \Gamma_{2N-1}$ and $\Gamma^e = \Gamma_0 \times \ldots \times \Gamma_{2N-2}$. It is assumed that the players enact their strategies alternately (in Stackelberg fashion) so that the first player acts at stages 0, 2, ...2N, and the second player acts at stages 1, 3, ...2N-1. Then, a closed-loop equilibrium is a pair of decision sequences $\underline{\gamma}^{0*} \in \Gamma^0$, $\underline{\gamma}^{e*} \in \Gamma^e$ such that

$$E^{\underline{\gamma}^{0*},\underline{\gamma}^{e*}}(\phi^1_{2N}(x_{2N})) \leq E^{\underline{\gamma}^{0*},\underline{\gamma}^{e}}(\phi^1_{2N}(x_{2N})) \quad \forall\ \underline{\gamma}^e \in \Gamma^e \tag{4.1}$$

$$E^{\underline{\gamma}^{0*},\underline{\gamma}^{e*}}(\phi^2_{2N}(x_{2N})) \leq E^{\underline{\gamma}^{0},\underline{\gamma}^{e*}}(\phi^2_{2N}(x_{2N})) \quad \forall\ \underline{\gamma}^0 \in \Gamma^0. \tag{4.2}$$

where the expectations are superscripted by $\underline{\gamma}$ to indicate the explicit dependence of the probability density of x_{2N} on the choice of strategy $\underline{\gamma}$.

An interpretation of the above standard form is the following. There is an initial random state x_0. This state undergoes deterministic state transitions f_t up to state x_{2N}. At each stage a decision maker bases his decision on partial information about the current state. Each decision maker seeks to minimize a function ϕ^i of the final state x_{2N}. In the next section it is shown that this standard form is general enough to include problems formulated in Chapter II.

IV.3 Conversion of Dynamic Games to Standard Form

The mathematical form presented in Section II.1 can be converted to standard form in the following fashion, where to avoid overlap in notation, terms referring to the standard form of the previous section are underlined.

Let $\underline{X}_0$ be the space containing all of the uncertainties in the formulation of Section II.1. That is, let $\underline{x}_0 \in \underline{X}_0$ be defined as $\underline{x}_0 = (x_0, 0, 0, \{\theta_t, \xi_t^1, \xi_t^2\}_{t=0,\ldots N-1})$. The probability density π_0 is defined as the joint probability density of the random variables included in $\underline{x}_0$. The σ-field $\underline{\Sigma}_0$ is then the product of the Borel fields B^q, B^p in the obvious way defined in Chapter II. The decision spaces $(\underline{U}_t, \underline{F}_t)$ are defined as $(\underline{U}_t, \underline{F}_{2t}) = (U_t^1, G_t^1)$; $(\underline{U}_{2t+1}, \underline{F}_{2t+1}) = (U_t^2, G_t^2)$.

Define the states $\underline{x}_{2t} \in \underline{X}_{2t}$ as $\underline{x}_{2t} = (x_0, \underline{u}_0, \ldots \underline{u}_{2t-1}, J_t^1,$

J_t^2, $\{\theta_i, \xi_i^1, \xi_i^2\}_{i=0,\ldots N-1})$ and $\underline{x}_{2t+1} = (\underline{x}_{2t}, \underline{u}_{2t})$, $t = 0,\ldots N-1$. The variables J_t^1, J_t^2 are real-valued variables related to the costs, as will been later. The σ-fields $\underline{\Sigma}_t$ are defined as the product σ-fields of the σ-fields of the individual components of $\underline{x}_t$, where the σ-fields of J_t^1, J_t^2 are the Borel σ-fields of the real line.

The state transition function can be defined as follows:

$$\underline{f}_{2t}(\underline{x}_{2t}, \underline{u}_{2t}) = (\underline{x}_{2t}, \underline{u}_{2t}) \tag{4.3}$$

$$\underline{f}_{2t+1}(\underline{x}_{2t+1}, \underline{u}_{2t+1}) = (x_0, \underline{u}_0, \ldots \underline{u}_{2t+1}, J_{t+1}^1, J_{t+1}^2, \{\theta_i, \xi_i^1, \xi_i^2\}_{i=0,\ldots N-1}) \tag{4.4}$$

where

$$x_{t+1} = f_t(x_t, \underline{u}_{2t}, \underline{u}_{2t+1}, \theta_t)$$

$$J_{t+1}^i = J_t^i + h_t^i(x_t, \underline{u}_{2t}, \underline{u}_{2t+1})$$

$$J_0^i = 0.$$

As required in the standard form, $\underline{f}_t$ is a deterministic function from $\underline{X}_t \times \underline{U}_t \to \underline{X}_{t+1}$. Furthermore, since f_t and h_t^i are appropriately measurable functions of their arguments, then $\underline{f}_t$ is measurable from $\underline{\Sigma}_t \times \underline{F}_t$ to $\underline{\Sigma}_{t+1}$.

The last concept which must be incorporated is the notion of available information to the players, embodied in the σ-fields $\underline{D}_t$. In the formulation in Section II.1, equation (2.2) defined measurements y_t^i which were available to the players. These measurements are deterministic functions of the new state variables $\underline{x}_{2t}$; these functions determine a subfield of $\underline{\Sigma}_t$ on which y_t^1 is measurable. Witsenhausen (Wt2) and Ho and Chu (H5) have discussed in detail how to define the subfields $\underline{D}_t$ in terms of the measurement equations (2.2). Briefly described, equation (2.2) can be rewritten as

$$y_t^i = g_t^i(x_t, \xi_t^i) = \hat{g}_t^i(\underline{x}_{2t})$$

The measurability assumptions on g_t^i enable the sequence $y_0^1, \ldots y_t^1$ to induce a field $\underline{D}$ on $\underline{X}_{2t}$. Similarly, $y_0^2, \ldots y_t^2$ induces a field on $\underline{X}_{2t+1}$.

The cost functions $\underline{\phi}_{2N}^1$ and $\underline{\phi}_{2N}^2$ are defined as

$$\underline{\phi}_{2N}^1(\underline{x}_{2N}) = J_N^1 + h_N^1(x_N)$$

$$\phi_{2N}^2(\underline{x}_{2N}) = J_N^2 + h_N^2(x_N)$$

The conversion to standard form is now complete. The concept of state has been expanded to incorporate the uncertainties of Section II.2. In terms of this new state, the concept of available information can be posed in terms of σ-fields of the state space, determined a priori because of the deterministic nature of the new evolution function $\underline{f}_t$ and the measurement equations. The notion of admissible strategies is then reduced to questions of measurability. In the next section this standard form is used in describing closed-loop Stackelberg equilibria.

IV.4 An Equivalent Deterministic Problem

A closed-loop equilibrium for a game expressed in standard form is still a stochastic problem, as indicated by the fact that the elements of X_0 are random elements. Following a similar development by Witsenhausen (Wt4), this problem can be reformulated as an equivalent deterministic problem by considering (in the notation of Section IV.2) the unconditional probability density π_t of x_t as the new state.

Let Π_t be the Banach space of signed measures on (X_t, Σ_t), with the norm being the total variation norm. Similarly let Φ_t be defined as $L_\infty^\varepsilon(X_t, R)$, the space of bounded real-valued Σ_t-measurable functions with the supremum norm. For $\phi \in \Phi_t$, $\pi \in \Pi_t$, define the relation

$$\langle\phi, \pi\rangle = \int_{X_t} \phi(x)\, d\pi(x) \tag{4.5}$$

Equation (4.5) establishes a form of duality between Φ_t and Π_t as each of these spaces defines a norm determining set of bounded linear functionals on the other. The probability measures π_t represent a closed convex subset of Π_t. The given π_0 belongs to this subset of Π_0. For any t, $\gamma_t \in \Gamma_t$, define $g_t(\cdot, \gamma_t): X_t \to X_{t+1}$ by $g_t(x_t; \gamma_t) = f_t(x_t, \gamma_t(x_t))$. Similarly, define $T_t(\gamma_t): \Pi_t \to \Pi_{t+1}$ as

$$(T_t(\gamma_t)\pi)(E) = \pi\{x \mid g_t(x; \gamma_t) \in E\} \quad \forall\ E \in \Sigma_{t+1}. \tag{4.6}$$

Note g_t is measurable from Σ_t to Σ_{t+1}, and $T_t(\gamma)$ is the probability measure induced in X_{t+1} by the use of strategy γ_t in the function

$g(\cdot\,;\gamma_t)$.

For any t,γ, $T_t(\gamma)$ is a map of unit norm. Define dually $T_t^*(\gamma):\Phi_{t+1} \to \Phi_t$ by

$$[T_t^*(\gamma)\Phi_{t+1}](x) = \phi_{t+1}(g_t(x;\gamma)) \qquad \forall\, x \in X_t. \tag{4.7}$$

The maps $T_t(\gamma)$ and $T_t^*(\gamma)$ are adjoint in terms of the relationship (4.5) as the rules for change of variable yield

$$\langle \phi,\, T_t(\gamma)\pi \rangle = \langle T_t^*(\gamma)\phi,\, \pi \rangle \qquad \forall\, \phi \in \Phi_{t+1},\ \pi \in \Pi_t.$$

In the equivalent deterministic problem, π_t is the state, with state space Π_t. The admissible decisions are elements of Γ_t. The state evolution equation is given by equation (4.6). The costs are linear functions of the final state, given by

$$J^i = \langle \phi_{2N}^i,\, \pi_{2N} \rangle\ ;\ i = 1,2. \tag{4.8}$$

In this problem, player 1 chooses his strategies for even stages and player 2 chooses his for odd stages. Two sequences of costates ϕ_t^1, ϕ_t^2 can be defined in terms of equation (4.7) for each decision sequence $\underline{\gamma} = (\underline{\gamma}^e, \underline{\gamma}^0)$ as:

$$\phi_t^i = T_t^*(\gamma_t)\phi_{t+1}^i;\ i = 1,2;\ t = 0,\ldots 2N-1$$

where ϕ_{2N}^i is given in equation (4.8).

<u>Theorem IV.1</u>: A necessary condition for $\underline{\gamma}^* = (\underline{\gamma}^{e*}, \underline{\gamma}^{0*})$ to represent a closed-loop equilibrium to the game is: Let ϕ_t^{10}, ϕ_t^{20} and π_t^0 be the associated costates and states with $\underline{\gamma}^*$ arising from equations (4.7)

and (4.9). Then,

(i) For t even, $J^{1*} = J^1(\underline{\gamma}^*) = \langle \phi_t^{10}, \pi_t^0 \rangle$

$$= \min_{\gamma_t} \langle \phi_{t+1}^{10}, T_t(\gamma_t)\pi_t^0 \rangle$$

(ii) For t odd, $J^{2*} = J^2(\underline{\gamma}^*) = \langle \phi_t^{20}, \pi_t^0 \rangle$

$$= \min_{\gamma_t} \langle \phi_{t+1}^{20}, T_t(\gamma_t)\pi_t^0 \rangle.$$

<u>Proof:</u> From equations (4.7), (4.8), (4.9), it follows that $J^i(\underline{\gamma}^*) = \langle \phi_{2N}^i, \pi_{2N}^0 \rangle = \langle \phi_t^i, \pi_t^0 \rangle$ for all t. Assume t is even. Consider

$$\underline{\gamma}^1 = (\gamma_0^*, \ldots \gamma_{t-1}^*, \gamma_t, \gamma_{t+1}^*, \ldots \gamma_{2N-1}^*) \quad \text{where } \gamma_t \in \Gamma_t.$$

From (4.6), $\pi_t = T_{t-1}(\gamma_{t-1})T_{t-2}(\gamma_{t-2})\ldots T_0(\gamma_0)\pi_0$. Then π_t is the same for $\underline{\gamma}^1$ and $\underline{\gamma}^*$. Similarly, (4.7) implies $\phi_{t+1}^i = T_{t+1}^*(\gamma_{t+1}) \ldots T_{2N-1}^*(\gamma_{2N-1})\phi_{2N}^i$ so ϕ_{t+1}^i is the same for $\underline{\gamma}^1$ and $\underline{\gamma}^*$.

Now, $J^1(\underline{\gamma}^1) = \langle \phi_{t+1}^{10}, \pi_{t+1} \rangle = \langle \phi_{t+1}^{10}, T_t(\gamma_t)\pi_t^0 \rangle$

$$\geq J^1(\underline{\gamma}^*) = \langle \phi_{t+1}^{10}, T_t(\gamma_t^*)\pi_t^0 \rangle$$

because $\underline{\gamma}^*$ is an equilibrium as in equation (4.1) so, (i) is necessary at the equilibrium point. A similar argument for t odd establishes (ii).

A notion which is very useful is that of determining the coreachable sets for each player; that is, the sets of costates which he can achieve using his strategies. In contrast with control theory, these sets will depend in general on the decisions of the other player. Define the coreachable sets $\rho_t(\phi^1_{2N}; \underline{\Upsilon}^0_t)$ and $\rho_t(\phi^2_{2N}; \underline{\Upsilon}^e_t)$ as

$$\rho_t(\phi^1_{2N}; \underline{\Upsilon}^0_t) = \bigcup_{\gamma_t \in \Gamma_t} T^*_t(\gamma_t)\rho_{t+1}(\phi^1_{2N}; \underline{\Upsilon}^0_t) \text{ if } t \text{ even}$$

$$= T^*_t(\gamma_t)\rho_{t+1}(\phi^1_{2N}; \underline{\Upsilon}^0_t) \quad \text{if } t \text{ odd, where } \gamma_t \in \underline{\Upsilon}^0$$

$$\rho_t(\phi^2_{2N}; \underline{\Upsilon}^e_t) = \bigcup_{\gamma_t \in \Gamma_t} T^*{}_t(\gamma_t)\rho_{t+1}(\phi^2_{2N}; \underline{\Upsilon}^e_t) \text{ if } t \text{ odd}$$

$$\rho_t(\phi^2_{2N}; \underline{\Upsilon}^e_t) = T^*_t(\gamma_t)\rho_{t+1}(\phi^2_{2N}; \underline{\Upsilon}^e_t) \text{ if } t \text{ even, where } \gamma_t \in \underline{\Upsilon}^e \tag{4.10}$$

with the initial conditions $\rho_{2N}(\phi^1_{2N}; \underline{\Upsilon}^0_{2N}) = \{\phi^1_{2N}\}$ and $\rho_{2N}(\phi^2_{2N}; \underline{\Upsilon}^e_{2N}) = \{\phi^2_{2N}\}$.

The elements of $\rho_t(\phi^1; \underline{\Upsilon}^0)$ and $\rho_t(\phi^2; \underline{\Upsilon}^e)$ have the same norm as ϕ^1 and ϕ^2 for all t, hence the coreachable sets are bounded. Since the criteria are linear in the costates as seen in Theorem IV.1, the achievable performance by each player in the face of the other player's decision sequence can be characterized in terms of the support function of the coreachable sets.

Define the support function of $\rho_t(\phi; \underline{\Upsilon})$ as

$$V^1_t(\pi; \phi^1_{2N}; \underline{\Upsilon}^0_t) = \sup_{\phi \in \rho_t(\phi^1_{2N};\underline{\Upsilon}^0_t)} \langle\phi, \pi\rangle$$

$$V^2_t(\pi, \phi^2_{2N}; \underline{\Upsilon}^e_t) = \sup_{\phi \in \rho_t(\phi^2_{2N}; \underline{\Upsilon}^e_t)} \langle\phi, \pi\rangle \tag{4.11}$$

The boundedness of the coreachable sets implies V_t^i are Lipschitz continuous with constants $||\phi_{2N}^i||$. With this terminology it is now possible to specify a procedure which yields a pair of strategy sequences which are a closed-loop Stackelberg equilibrium in the game.

<u>Theorem IV.2</u>: The sequence of strategies $\underline{\gamma}^* = (\underline{\gamma}^{e*}, \underline{\gamma}^{0*})$ obtained recursively by:

$$\text{(i)} \quad \text{for } t \text{ even, } \gamma_t^*(\pi) = \arg\sup_{\gamma_t \in \Gamma_t} V_{t+1}^1(T_t(\gamma_t)\pi; -\phi_{2N}^1; \underline{\gamma}_{t+1}^{0*})$$

$$\text{(ii)} \quad \text{for } t \text{ odd, } \gamma_t^*(\pi) = \arg\sup_{\gamma_t \in \Gamma_t} V_{t+1}^2(T_t(\gamma_t)\pi; -\phi_{2N}^2; \underline{\gamma}_{t+1}^{e*})$$

where the suprema exist and are admissible strategies for all $\pi \in \Pi_t$, and since π_0 is known, the values of π_t and γ_t^* are obtained from (i) and (ii) to yield $\underline{\gamma}^*$, is a closed-loop Stackelberg equilibrium to the game (i.e., satisfies equations (4.1) and (4.2)).

<u>Proof</u>: Equation (4.10) defines a recursive relation between coreachable sets. Assume t is even; then

$$V_t^1(\pi; -\phi_{2N}^1; \underline{\gamma}_{t+1}^0) = \sup_{\phi \in \rho_t(-\phi_{2N}^1, \underline{\gamma}_{t+1}^0)} \langle \phi, \pi \rangle$$

$$= \sup_{\phi \in \rho_{t+1}(-\phi_{2N}^1, \underline{\gamma}_{t+1}^0)} \sup_{\gamma_t \in \Gamma_t} \langle T_t^*(\gamma_t)\phi, \pi \rangle$$

from equation (4.10)

$$= \sup_{\gamma_t \in \Gamma_t} \sup_{\phi \in \rho_{t+1}(-\phi^1_{2N},\, \gamma^0_{t+1})} \langle \phi,\, T_t(\gamma)\pi \rangle$$

$$= \sup_{\gamma_t \in \Gamma_t} V^1_{t+1}(T_t(\gamma_t)\pi;\ -\phi^1_{2N};\ \underline{\gamma}^0_{t+1}) \tag{4.12}$$

Also, $V^2_t(\pi;\ -\phi^2_{2N};\ \underline{\gamma}^e_t) = \sup_{\phi \in \rho_t(-\phi^2_{2N};\ \underline{\gamma}^e_t)} \langle \phi,\, \pi \rangle$

$$= \sup_{\phi \in \rho_{t+1}(-\phi^2_{2N};\ \underline{\gamma}^e_{t+1})} \langle T^*_t(\gamma^e_t)\phi,\, \pi \rangle$$

$$= V^2_{t+1}\ (T^*_t(\gamma^e_t)\pi; -\ \phi^2_{2N};\ \underline{\gamma}^e_{t+2}) \tag{4.13}$$

Similarly, for t odd, equation (4.10) implies

$$V^2_t(\pi;\ -\phi^2_{2N};\ \underline{\gamma}^e_{t+1}) = \sup_{\gamma_t \in 0_t} V^2_{t+1}(T_t(\gamma_t)\pi;\ -\phi^2_{2N},\ \underline{\gamma}^e_{t+1}) \tag{4.14}$$

$$V^1_t(\pi;\ -\phi^1_{2N};\ \underline{\gamma}^0_t) = V^1_{t+1}(T_t(\gamma^0_t)\pi;\ -\phi^1_{2N},\ \underline{\gamma}^0_{t+2}) \tag{4.15}$$

Equations (4.12)-(4.15) define recursive relations for the support functionals V^1 and V^2.

Now, assume $\underline{\gamma}^* = (\underline{\gamma}^{e*},\ \underline{\gamma}^{0*})$ satisfy the conditions of Theorem IV.2. Consider the problem of maximizing the expected value $-J^1(\underline{\gamma}^e,\ \underline{\gamma}^{0*})$. From the definition of support functions, the optimal value is $V^1_0(\pi_0;\ -\phi^1_T;\ \underline{\gamma}^{0*})$. If $\gamma^*_t(\pi)$ satisfies the conditions of Theorem IV.2, then equations (4.12) to (4.15) establish that

$$V^1_t(\pi;\ -\phi^1_{2N};\ \underline{\gamma}^{0*}) = V^1_{t+1}(T_t(\gamma^*_t(\pi))\pi;\ -\phi^1_{2N};\ \underline{\gamma}^{0*}_{t+1}) \tag{4.16}$$

for all t. In particular,

$$V_0^1(\pi_0; -\phi_T^1; \underline{\gamma}^{0*}) = V_{2N}^{1}(T_{2N-1}(\gamma_{2N-1}^*)T_{2N-2}(\gamma_{2N-2}^*)\cdots$$
$$\cdots T_0(\gamma_0^*)\pi_0; -\phi_{2N}^1)$$

The expression on the right is the value resulting from using $\underline{\gamma}^{e*}$ obtained in the conditions of Theorem IV.2, while the expression on the left is the optimal value attainable. Hence

$$E\{J^1(\underline{\gamma}^{e*}, \underline{\gamma}^{0*})\} = -V_0^1(\pi_0; -\phi_{2N}^1; \underline{\gamma}^{0*}) \leq E\{J^1(\underline{\gamma}^{e}, \underline{\gamma}^{0*})\}$$
$$\forall\, \underline{\gamma}^e \in \Gamma^e$$

An identical argument establishes the inequality for $\underline{\gamma}^0 \in \Gamma^0$. That is,

$$E\{J^2(\underline{\gamma}^{e*}, \underline{\gamma}^{0*})\} = -V_0^2(\pi_0; -\phi_{2N}^2; \underline{\gamma}^{e*}) \leq E\{J^2(\underline{\gamma}^{e*}, \underline{\gamma}^{0})\}.$$

Theorem IV.2 establishes the validity of a dynamic-programming approach at finding closed-loop equilibria in non-deterministic games. The expressions V_t^1 and V_t^2 are the optimal costs-to-go from a given state π_t assuming knowledge of the other player's future strategies. The alternating framework used by the players enables a recursive algorithm to be carried out. At stage 2N-1, player 2 can choose his optimal strategy $\gamma_{2N-1}^{2*}(\pi)$. Using this knowledge, player 1 can then choose his strategy $\gamma_{2N-1}^{1*}(\pi)$. Equations (4.12)-(4.15) provide the recursive relations for defining the evolution of the support functions used in parts (i) and (ii) of Theorem IV.2, in terms of the optimal decisions γ_t^*. Thus, Theorem IV.2 provides sufficient conditions for the existence of a closed-loop Stackelberg equilibrium; namely, the existence of the various suprema described for the set of admissible strategies. If all these suprema are admissible strategies, then Theorem IV.2 offers a constructive method for obtaining a closed-loop equilibrium.

IV.5 Example

Consider the following finite-state game, with one stage. Initially the state is 0. The next state is one of a possible 8 states. There are no transitional costs, only a cost associated with each final state, which is the same for both players. The state transition function is stochastic, described by

$$f_0(x_0,\ u_0^1,\ u_0^2,\ \theta_0) = (\theta_0,\ u_0^1,\ u_0^2) \tag{4.17}$$

where $U_0^1 = U_0^2 = \{0,\ 1\}$ and θ_0 is random, taking values 0 and 1 with equal probabilities. The information structure is such that player 2 does not know the value of θ_0, but player 1 does. Finally, player 1 is the leader, and makes his choice first.

The extensive form of this game is shown in Figure IV.1, where the costs have been indicated at the final states. This game problem can be reduced to the standard form of Section IV.2. The initial state is $(x_0,\ \theta_0)$ with the σ-field product of $\{\phi,\ \{0\}\}$ for x_0 and $\{\phi,\ \{0\},\ \{1\},\ \{0,1\}\}$ for θ_0. The probability density is the product of two marginal densities, assigning therefore values $\frac{1}{2}$ to outcomes $(x_0,0)$ and

$(x_0, 1)$.

The state at the next stage is (θ_0, u_0^1), the transition described in the obvious way. The information field D_1 is now the produce of the power set of U_0^1 and $\{\phi, \{0,1\}\}$ for θ_0. Notice D_1 is a proper subset of Σ_1, the power set of $X_0 \times U_0$.

The transition to the next stage is again trivial, mapping $X_1 \times U_1 \to X_2$ by $f_1((\theta_0, u_0^1), u_0^2) = (\theta_0, u_0^1, u_0^2)$. Σ_2 is again the power set of $\{0, 1\}^3$. The admissible strategies Γ_0 and Γ_1 are the sets of all maps from X_0 and X_1 respectively into $\{0, 1\}$, measurable on D_0 and D_1. Only pure strategies will be considered admissible.

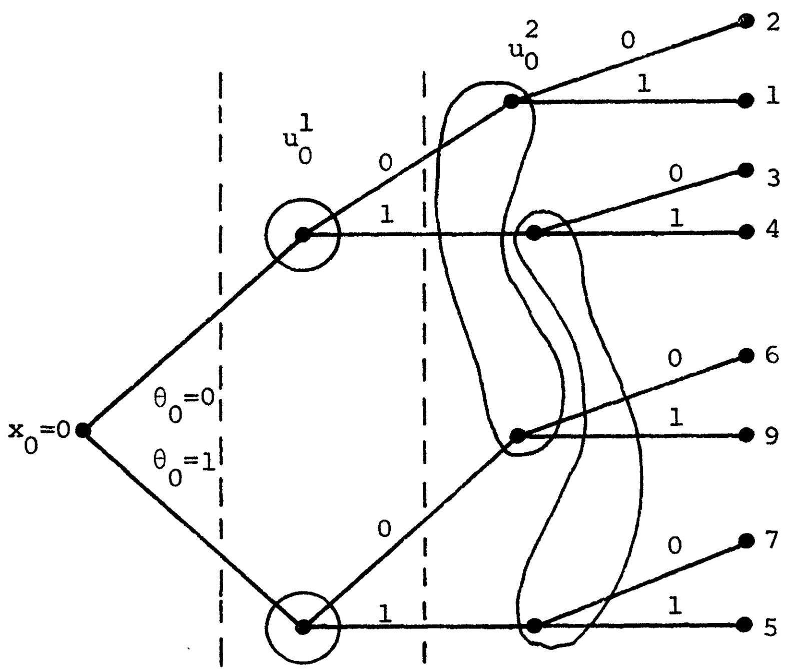

Figure IV.2 Extensive form of game

The numbers at the end of the branches of the tree in the extensive form represent the costs associated with the particular outcome of the game. Hence, the play $\theta_0 = 0$, $u_0^1 = 1$, $u_0^2 = 0$ results in a cost of 3. The cost is identical for both players.

Using the terminology of Section IV.2, the costs, in terms of the probability π_2 induced in X_2, are

$$J^i = \sum_{i=1}^{8} \pi_2(x_2)\phi(x_2)$$

where ϕ is the cost associated with terminal state x_2. Using Theorem IV.2,

$$\begin{aligned} \gamma_1^*(\pi_1) &= \sup_{\gamma \in \Gamma_1} \langle T_1(\gamma_1)\pi_1; -\phi\rangle \\ &= \sup_{\gamma \in \Gamma_1} \left(\sum_{i=1}^{8} - \pi_2(x_2)\phi(x_2)\right), \quad \pi_2 = T_1(\gamma_1)\pi_1. \end{aligned}$$

Since only pure strategies are admissible (this restriction is not essential; by a detailed argument the same results can be obtained for mixed strategies or behavior strategies [K2]), there are four admissible γ_1, listed as

a $\gamma_1(\theta_0, u_0^1) = 0$

b $\gamma_1(\theta_0, u_0^1) = u_0^1$

c $\gamma_1(\theta_0, u_0^1) = 1 - u_0^1$

d $\gamma_1(\theta_0, u_0^1) = 1$

The resulting costs for each of these strategies in terms of π_1 are tabulated in Table IV.3.

Strategy	Cost
a	$-(2\pi_1(0,0) + 6\pi_1(1,0) + 7\pi_1(1,1) + 3\pi_1(0,1))$
b	$-(2\pi_1(0,0) + 6\pi_1(1,0) + 5\pi_1(1,1) + 4\pi_1(0,1))$
c	$-(\pi_1(0,0) + 9\pi_1(1,0) + 7\pi_1(1,1) + 3\pi_1(0,1))$
d	$-(\pi_1(0,0) + 9\pi_1(1,0) + 5\pi_1(1,1) + 4\pi_1(0,1))$

Table IV.2 Costs associated with γ_1

The space of all probability distributions can be divided into four regions, over which one of these strategies is optimal. However, the problem can be simplified when only pure strategies are admissible, since now there are only four possible distributions π_1. Table IV.3 lists these distributions, together with the corresponding optimal strategy γ_1 and cost-to-go V_1^{1*}.

	π_1	Optimal γ_1	Cost V_1^{1*}
1	$\pi_1^1(0,0) = \frac{1}{2}$, $\pi_1^1(1,0) = \frac{1}{2}$	a or b	-4
2	$\pi_1^2(0,0) = \frac{1}{2}$, $\pi_1^2(1,1) = \frac{1}{2}$	d	-3
3	$\pi_1^3(0,1) = \frac{1}{2}$, $\pi_1^3(1,0) = \frac{1}{2}$	a	-4.5
4	$\pi_1^4(0,1) = \frac{1}{2}$, $\pi_1^4(1,1) = \frac{1}{2}$	b or d	-5

Table IV.3

Similarly, $\gamma_0^* = \underset{\gamma_0 \in \Gamma_0}{\operatorname{argsup}} \quad (V_1^{1*}(T_0(\gamma_0)\pi_0;\ -\phi;\ \gamma_1^1)$.

Again, there are four possible strategies, identifiable as

a	$\gamma_0(\theta_0) = 0$
b	$\gamma_0(\theta_0) = \theta_0$
c	$\gamma_0(\theta_0) = 1 - \theta_0$
d	$\gamma_0(\theta_0) = 1$

Each of these strategies maps π_0 into a distribution in Table IV.3, where V_1^* is also computed. Hence, γ_0^* is such that $T_0(\gamma_0^*)\pi_0 = \pi_1^2$. This strategy is $\gamma_0^* = b$.

Thus the equilibrium pair of strategies are $\gamma_0^* = b$, $\gamma_1^* = d$ with an associated cost of 3. Even though this example is very simple, it highlights the difficulties in applying the techniques of Section IV.4 in constructing closed-loop equilibria. The spaces π_t of all probability distributions are very large, and hence the maximizations of Theorem IV.2, which must be carried out for each element of the space, are unrealistic. In this example, through independent arguments the cardinality of π_1 was reduced to four elements. These arguments will not hold in general; the determination of closed-loop equilibria is an untractable problem in most cases.

IV.6 Signaling-free Equilibria in Stochastic Stackelberg Games

When the nature of the game is stochastic, so that the information available to each decision maker at stage t can be different, signaling-free equilibria in Stackelberg games are no longer closed-loop equilibria. Signaling-free equilibria differ from closed-loop equilibria in that the players wait until stage t to choose their decision at stage t (rather than committing themselves to a strategy prior to the game), and do not assume that previous decisions were the output of strategies. These differences were presented in mathematical form in Chapter II, as well as some basic results for deterministic games. This section discusses signaling-free equilibria in stochastic games. Throughout the remainder of this chapter, it is assumed that the state, decision and observation spaces X_t, U_t^i, Y_t^i are finite-dimensional Euclidean spaces, with associated fields the Borel σ-fields.

In Chapter II, equations (2.17) through (2.22) define signaling-free equilibria in Stackelberg games. These equations are:

$$I_t^1(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^{1*}, \underline{\gamma}_t^{2*}) \leq I_t^1(\omega, \bar{u}_{t-1}, \underline{\gamma}_t^1, \underline{\gamma}_t^{2*}) \quad \text{a.e.} \tag{4.14}$$

$$I_t^2(\omega, \bar{u}_{t-1}, u_t^1, \underline{\gamma}_{t+1}^{1*}, \underline{\gamma}_t^{2*}) \leq I_t^2(\omega, \bar{u}_{t-1}, u_t^1, \underline{\gamma}_t^{1*}, \underline{\gamma}_t^2) \quad \text{a.e.} \tag{4.15}$$

for all admissible $\underline{\gamma}^i \in \prod_{t=0}^{N-1} \Gamma_t^i$, where I_t^i is defined by

$$I_{N-1}^2(\omega, \bar{u}_{N-2}, u_{N-1}^1, \gamma_{N-1}^2) = E\{J_{N-1}^2(\omega, \bar{u}_{N-2}, u_{N-1}^1, \gamma_{N-1}^2) \mid z_{N-1}^2\} \quad \text{a.e.} \tag{4.16}$$

$$I^1_{N-1}(\omega,\bar{u}_{N-2},\underline{\gamma}_{N-1}) = E\{J^1_{N-1}(\omega,\bar{u}_{N-2},\underline{\gamma}_{N-1})\,|\,z^1_{N-1}\} \quad \text{a.e.} \tag{4.17}$$

$$\begin{aligned} I^2_t(\omega,\bar{u}_{t-1},u^1_t,\underline{\gamma}^1_{t+1},\underline{\gamma}^2_t) &= E\{h^2_t(x_t,u^1_t,u^2_t) \\ &\quad + I^2_{t+1}(\omega,\bar{u}_t,u^1_{t+1},\underline{\gamma}^1_{t+2},\underline{\gamma}^2_{t+1})\,|\,z^2_t, \\ &\quad u^1_{t+1} = \gamma^1_t(\omega,\bar{u}_t),\ u^2_t = \gamma^2_t(\omega,\bar{u}_{t-1},u^1_t)\} \quad \text{a.e.} \end{aligned} \tag{4.18}$$

$$\begin{aligned} I^1_t(\omega,\bar{u}_{t-1},\underline{\gamma}^1_t,\underline{\gamma}^2_t) &= E\{h^1_t(x_t,u^1_t,u^2_t) + I^1_{t+1}(\omega,\bar{u}_t,\underline{\gamma}^1_{t+1},\underline{\gamma}^2_{t+1})\,|\,z^1_t, \\ &\quad u^1_t = \gamma^1_t(\omega,\bar{u}_{t-1}),\ u^2_t = \gamma^2_t(\omega,\bar{u}_{t-1},u^1_t)\} \quad \text{a.e.} \end{aligned} \tag{4.19}$$

where $z^1_{t+1} = (z^1_t,\ u^1_t,\ u^2_t)$

$z^2_{t+1} = (z^2_t,\ u^2_t,\ u^1_{t+1})$.

Admissible strategies γ^1_t are jointly measurable functions of ω and past decisions $\bar{u}_{t-1}$, measurable for each sequence of decisions $\bar{u}_{t-1}$ on the fields $F^1_t\{\bar{u}_{t-1}\}$, to reflect the fact that previous strategies are assumed to be constant. Admissible γ^2_t satisfy equivalent measurability constraints. The expectations in equations (4.16)-(4.19) are conditional expectations with respect to the probability measure P^0 in $(\Omega,\ F^0,\ P^0)$; the elements inside the expectations are well-defined random variables defined on $(\Omega,\ F^0,\ P^0)$ as discussed in Chapter II. The following lemma establishes a different representation of admissible strategies:

<u>Lemma IV.3</u> (D1). Let $x_1(\omega),\ldots x_n(\omega)$ be n real-valued random variables defined on a probability space $(\Omega,\ F,\ P)$. Let B be the smallest

σ-field contained in F where $\{x_i(\omega)\}$ are measurable; i.e., the field induced on Ω by the Borel fields through $\{x_i(\omega)\}$. Let $f(\omega)$ be any B-measurable function. Then, $\exists$ a Borel-measurable function $g(x_1,\ldots x_n)$ such that

$$f(\omega) = g(x_1(\omega), \ldots x_n(\omega)) \qquad \text{a.e.}$$

The proof of this lemma is given in (D1) and (Bℓ1). This lemma establishes the existence of Borel-measurable strategies $\hat{\gamma}^i_t : Z^i_t \to U^i_t$ such that $\hat{\gamma}^i_t(z^i_t(\omega)) = \gamma^i_t(\omega, \bar{u}_{t-1})$ a.e.

As pointed out in Chapter II, equations (4.16)-(4.19) are conditional expectations, hence defined only up to an equivalence class of a.e. equal functions. The elements I^i_t refer to one arbitrary member of this equivalence class; in terms of this formulation, conditions can be stated for finding the optimal signaling-free strategies.

<u>Theorem IV.4:</u> Suppose there exists a pair of admissible strategy sequences $\underline{\gamma}^{10}$, $\underline{\gamma}^{20}$, satisfying

$$I^1_t(\omega,\bar{u}_{t-1},\underline{\gamma}^{10}_t,\underline{\gamma}^{20}_t) \leq I^1_t(\omega,\bar{u}_{t-1},\gamma^1_t,\underline{\gamma}^{10}_{t+1},\underline{\gamma}^{20}_t) \quad \text{a.e.} \tag{4.20}$$

$$I^2_t(\omega,\bar{u}_{t-1},u^1_t,\underline{\gamma}^{10}_{t+1},\underline{\gamma}^{20}_t) \leq I^2_t(\omega,\bar{u}_{t-1},u^1_t,\underline{\gamma}^{10}_{t+1},\gamma^{20}_t,\underline{\gamma}^{20}_{t+1}) \quad \text{a.e.} \tag{4.21}$$

except for ω in a fixed P^0-null set E^i_t independent of $\bar{u}_t$, for all admissible $\underline{\gamma}^i_t$, t, $\bar{u}_t$. Then $\underline{\gamma}^{10}$,$\underline{\gamma}^{20}$ constitute a signaling-free equilibrium.

<u>Proof:</u> At stage N-1, clearly the strategies γ^{10}_{N-1}, γ^{20}_{N-1} satisfy equations (4.14) and (4.15). Hence assume inductively that (4.14) and (4.15) hold for $\underline{\gamma}^{10}_{T+1}$, $\underline{\gamma}^{20}_{T+1}$; that is, $\underline{\gamma}^{10}$ and $\underline{\gamma}^{20}$ satisfy the signaling-

free inequalities for $t \geq T + 1$, almost everywhere except for fixed P^0-null sets E_t^i. Then, by assumption, for $t = T$,

$$I_t^1(\omega,\bar{u}_{t-1},\underline{\gamma}_t^{10},\underline{\gamma}_t^{20}) \leq I_t^1(\omega,\bar{u}_{t-1},\gamma_t^1,\underline{\gamma}_{t+1}^{10},\underline{\gamma}_t^{20}) \text{ a.e.}$$

$$= E\{h_t^1(x_t,u_t^1,u_t^2) + I_{t+1}^1(\omega,\bar{u}_t,\underline{\gamma}_{t+1}^{10},\underline{\gamma}_{t+1}^{20})|z_t^1,\ u_t^1 = \gamma_t^1(\omega,\bar{u}_{t-1}),$$

$$u_t^2 = \gamma_t^{20}(\omega,\bar{u}_{t-1},u_t^1)\} \text{ a.e. by equation (4.19)}$$

$$\leq E\{h_t^1(x_t,u_t^1,u_t^2) + I_{t+1}^1(\omega,\bar{u}_t,\underline{\gamma}_{t+1}^{1},\underline{\gamma}_{t+1}^{20})|z_t^1,u_t^1 = \gamma_t^1(\omega,\bar{u}_{t-1}),$$

$$u_t^2 = \gamma_t^{20}(\omega,\bar{u}_{t-1},u_t^1)\} \text{ a.e. by equation (4.14) and the inductive assumption}$$

$$= I_t^1(\omega,\bar{u}_{t-1},\underline{\gamma}_t^1,\ \underline{\gamma}_t^{20}) \text{ for admissible } \underline{\gamma}_t^1.$$

Hence, $\underline{\gamma}^{10}$, $\underline{\gamma}^{20}$ satisfy equation (4.14) for $t = T$. A similar argument establishes that $\underline{\gamma}^{10}$, $\underline{\gamma}^{20}$ satisfy equation (4.15) for $t = T$. Hence, by induction, $\underline{\gamma}^{10}$ and $\underline{\gamma}^{20}$ satisfy equations (4.14) and (4.15) for all t, and hence constitute a signaling-free equilibrium.

Theorem IV.4 establishes the validity of a dynamic programming approach in finding signaling-free equilibria in stochastic Stackelberg games; furthermore, since all signaling-free equilibria must satisfy equations (4.20) and (4.21), Theorem IV.4 yields necessary and sufficient conditions for the existence of signaling-free equilibria. The implications are that these equilibria can be characterized alternately in terms of an equilibrium between 2N independent players, each acting only once during the game. This property was mentioned earlier in discussions of signaling-free equilibria in deterministic games, in

Chapter II.

Theorem IV.4 yields an obvious constructive way of obtaining a signaling-free equilibrium: Define the strategies γ_t^{10}, γ_t^{20} recursively by

$$\gamma_t^{10}(\omega, \bar{u}_{t-1}) = \underset{u^1}{\operatorname{argmin}}\ I_t^1(\omega, \bar{u}_{t-1}, u^1, \underline{\gamma}_{t+1}^{10}, \underline{\gamma}_t^{20}) \quad \text{a.e.} \tag{4.22}$$

$$\gamma_t^{20}(\omega, \bar{u}_{t-1}, u_t^1) = \underset{u^2}{\operatorname{argmin}}\ I_t^2(\omega, \bar{u}_{t-1}, u_t^1, \underline{\gamma}_{t+1}^{10}, u^2, \underline{\gamma}_{t+1}^{20}) \text{ a.e.} \tag{4.23}$$

The random variables minimized on the right side of the equation treat u^i as constants, whereas I_t^i is defined as a conditional expectation where u_t^i are strategies dependent on ω. This substitution is possible because these strategies γ_t^i are measurable on the conditioning σ-fields defined by z_t^i (Brl). The only problem with this constructive technique is that the strategies γ_t^{10}, γ_t^{20} so obtained need not be admissible strategies; that is, there is no guarantee that they will be appropriately measurable functions. Striebel (Strl) discusses properties of the class of admissible strategies which guarantee the existence of an appropriate minimizing γ_t^{i0}. For the special case where the field induced by z_t^i is separable (generated by a countable collection), (4.22) and (4.23) yield admissible strategies.

For a special subclass of games called LQG games, the assumptions of Theorem IV.4 are satisfied. This results from the Gaussian nature of the uncertainty in LQG games. Furthermore, the constructive technique described in Theorem IV.4 yields admissible strategies which are appropriately measurable. Chapter V studies in detail signaling-free equilibria in LQG games.

The costs represented by equations (4.16)-(4.19) are in general different from the truncated costs which occur in the closed-loop equilibrium concept. Hence, proposition II.4 of Chapter II does not hold in general stochastic games: A signaling-free equilibrium pair of strategy sequences does not have a corresponding pair of closed-loop strategy sequences which constitute a closed-loop equilibrium. However, in some special cases a correspondence exists between these two concepts. Define a game of equivalent information as a game where $F_t^1(\overline{\gamma}_{t-1}) = F_t^2(\overline{\gamma}_{t-1}, \gamma_t^1)$ for all t, all admissible $\overline{\gamma}_t$. The following theorem relates closed-loop equilibria to signaling-free equilibria:

Theorem IV.5. In games of equivalent information, every signaling-free equilibrium $\underline{\gamma}^{10}$, $\underline{\gamma}^{20}$ has a corresponding closed-loop equilibrium $\underline{\gamma}^{1e}$, $\underline{\gamma}^{2e}$ such that

$$E\{J^i(\underline{\gamma}^{1e}, \underline{\gamma}^{2e})\} = E\{I_0^i(\omega, \underline{\gamma}^{10}, \underline{\gamma}^{20})\}$$

Before proceeding with the proof of this theorem, some preliminary results need to be established.

Definition IV.1. The actual cost-to-go J_t^i are functions such that:

(i) $$J_t^1 : \Omega \times U_0^1 \times U_0^2 \times \dots U_{t-1}^1 \times U_{t-1}^2 \times \Gamma_t^1 \times \Gamma_t^2 \times \dots \times \Gamma_{N-1}^1 \times \times \Gamma_{N-1}^2 \to R^+$$

(ii) $$J_t^2 : \Omega \times U_0^1 \times \dots U_{t-1}^1 \times U_t^1 \times \Gamma_t^2 \times \dots \times \Gamma_{N-1}^1 \times \Gamma_{N-1}^2 \to R^+$$

(iii) $J_t^i(\omega, \overline{u}_{t-1}, \underline{\gamma}_t)$ is the cost associated with a play of the game with strategies $(\overline{u}_t, \underline{\gamma}_t)$ and the original random variables determined by ω.

The next theorem relates the notion of expected cost-to-go to actual cost-to-go.

Theorem IV.6. In games of equivalent information,

$$I_t^1(\omega, \overline{u}_{t-1}, \underline{\gamma}_t) = E\{J_t^1(\omega, \overline{u}_{t-1}, \underline{\gamma}_t) | z_t^1\} \quad \text{a.e.} \tag{4.24}$$

$$I_t^2(\omega, \overline{u}_{t-1}, u_t^1, \gamma_t^2, \underline{\gamma}_{t+1}) = E\{J_t^2(\omega, \overline{u}_{t-1}, u_t^1, \gamma_t^2, \underline{\gamma}_{t+1} | z_t^2\} \quad \text{a.e.} \tag{4.25}$$

for all admissible $\underline{\gamma}$, all t. The conditional expectations are with respect to the fields $F_t^1\{\overline{u}_{t-1}\}$.

Proof: For t = N-1, equations (4.24) and (4.25) are identical to equations (4.16) and (4.17), which define I_{N-1}^i. Hence the theorem is true for t = N-1. Assume inductively that the theorem is true for $t \geq T+1$. Then,

$$E\{J_T^1(\omega, \overline{u}_{T-1}, \underline{\gamma}_T) | z_T^1\} =$$

$$E\{h_T^1\Big(x_T(\omega), \gamma_T^1(\omega, \overline{u}_{T-1}), \gamma_t^2(\omega, \overline{u}_{T-1}, \gamma_T^1(\omega, \overline{u}_{T-1}))\Big) +$$

$$J_{T+1}^1(\omega, \overline{u}_{T-1}, \underline{\gamma}_T) | z_T^1\} \quad \text{a.e. by definition}$$

Since the game has equivalent information, both γ_t^1 and γ_t^2 are measurable on the conditioning field $F_T^1\{\overline{u}_{T-1}\}$. Hence, their value can be substituted in the expectations. That is, let $u_T^1 = \gamma_T^1(\omega, \overline{u}_{T-1})$, $u_T^2 = \gamma_T^2(\omega, \overline{u}_{T-1}, u_T^1)$ for all ω corresponding to the event $z_T^1(\omega, \overline{u}_{T-1}) = z_T^1$. Then,

$$E\{J_T^1(\omega, \bar{u}_{T-1}, \underline{\gamma}_T) \mid z_T^1(\omega, \bar{u}_{T-1}) = z_T^1\} = E\{h_T^1(x_T, u_T^1, u_T^2) +$$
$$J_{T+1}^1(\omega, \bar{u}_{T-1}, u_T^1, u_T^2, \underline{\gamma}_{T+1}) \mid z_T^1(\omega, \bar{u}_{T-1}) = z_T^1,\ u_T^1 = \gamma_T^1(\omega, \bar{u}_{T-1}),$$
$$u_T^2 = \gamma_T^2(\omega, \bar{u}_{T-1}, u_T^1)\} \quad \text{a.e.}$$
$$= E\Big\{h_T^1(x_T, u_T^1, u_T^2) + E\{J_{T+1}^1(\omega, \bar{u}_{T-1}, u_T^1, u_T^2, \underline{\gamma}_{T+1}) \mid z_{T+1}^{1a}\} \mid z_T^1,$$
$$u_T^1 = \gamma_T^1(\omega, \bar{u}_{T-1}),\ u_T^2 = \gamma_T^2(\omega, \bar{u}_{T-1})\Big\} \quad \text{a.e.}$$

because the conditioning field specified by $z_{T+1}^{1a} = (z_T^1, u_T^1, u_T^2, y_{T+1}^1)$, $F_{T+1}^1(\bar{u}_t) \supset F_T^1(\bar{u}_{t-1})$

$$= E\{h_T^1(x_T, u_T^1, u_T^2) + I_{T+1}^1(\omega, \bar{u}_T, \underline{\gamma}_{T+1}^1) \mid z_T^1,\ u_T^1 = \gamma_T^1(\omega, \bar{u}_{T-1}),$$
$$u_T^2 = \gamma_T^2(\omega, \bar{u}_{T-1}, u_T^1)\} \quad \text{a.e.}$$

because of the inductive assumption on I_{T+1}^1 and J_{T+1}^1

$$= I_T^1(\omega, \bar{u}_{T-1}, \underline{\gamma}_T) \quad \text{a.e. from equation (4.19).}$$

A similar argument establishes (4.25) for $t = T$. Hence, by induction, the theorem holds for all t.

Notice where the condition of equivalent information is used. In general games, γ_t^2 is not measurable on $F_t^1\{\bar{u}_{t-1}\}$, so that the value $u_t^2 = \gamma_t^2(\omega, \bar{u}_{t-1}, u_t^1)$ cannot be substituted in the conditioning expectation. Theorem IV.5 can now be proved.

Proof of Theorem IV.5

Consider $\gamma_t^{10}(\omega, \bar{u}_{t-1})$. By Lemma IV.4 and the measurability assumption in Chapter II, there exists a function $\hat{\gamma}_t^{10}: Z_t^1 \to U_t^1$, Borel measurable, such that $\hat{\gamma}_t^{10}(\overline{y}_t^1(\omega, \overline{u}_{t-1}), \overline{u}_{t-1}) = \gamma_t^{10}(\omega, \overline{u}_{t-1})$ a.e. Similarly define $\hat{\gamma}_t^{20}: Z_t^2 \to U_t^2$.

For each ω, define the random variables $y_t^{ie}(\omega)$, $u_t^{ie}(\omega)$ as the outcomes of the play of the game with strategies $\hat{\underline{\gamma}}_t^{10}$, $\hat{\underline{\gamma}}_t^{20}$. Clearly,

$$J^1(\underline{\gamma}^{10}, \underline{\gamma}^{20}) = J^1(\hat{\underline{\gamma}}^{10}, \hat{\underline{\gamma}}^{20}) \quad \text{a.e.}$$

Recursively define strategies $\gamma_t^{2c}(\omega, \overline{\gamma}_{t-1}^c, \gamma_t^{1c}) = \hat{\gamma}_t^{20}(\overline{y}_t^{2e}(\omega), \overline{u}_{t-1}^e(\omega), u_t^{1e}(\omega))$ a.e. and $\gamma_t^{1c}(\omega, \overline{\gamma}_{t-1}^c) = \hat{\gamma}_t^{10}(y_t^{1e}(\omega), \overline{u}_{t-1}^e(\omega))$ a.e. Then, the construction of γ_t^{1c}, γ_t^{2c} imply that γ_t^{1c} is measurable on the field induced by the random variables $\overline{y}_t^1$, $\overline{u}_{t-1}$ defined through the strategies $\overline{\gamma}_{t-1}^c$. That is, γ_t^{1c} is measurable on $F_t^1(\overline{\gamma}_{t-1}^c)$. Similarly, γ_t^{2c} is measurable on $F_t^2(\overline{\gamma}_{t-1}^c, \gamma_t^{1c})$. Thus, γ_t^{1c} and γ_t^{2c} are admissible closed-loop strategies. Furthermore,

$$J^i(\underline{\gamma}^{1c}, \underline{\gamma}^{2c}) = J^i(\underline{\gamma}^{10}, \underline{\gamma}^{20}) \quad \text{a.e.}$$

by construction. Thus, from Theorem IV.6,

$$E\{E(J^i(\underline{\gamma}^{1c}, \underline{\gamma}^{2c}) \mid z_0^i\} = E\{I_0^i(\omega, \underline{\gamma}^{10}, \underline{\gamma}^{20})\}$$

$$= E\{J^i(\underline{\gamma}^{1c}, \underline{\gamma}^{2c})\} \quad \text{a.e.}$$

Assume there is an admissible closed-loop strategy $\hat{\underline{\gamma}}^{1b}$ such that

$$E\{J^1(\hat{\underline{\gamma}}^{1b}, \hat{\underline{\gamma}}^{2c})\} < E\{J^1(\hat{\underline{\gamma}}^{1c}, \hat{\underline{\gamma}}^{2c})\}.$$

The representation $\hat{\gamma}_t^{1b}: Z_t^1 \to U_t^1$ is used here for convenience. Then, define $y_t^{1d}(\omega, \overline{u}_{t-1})$ as the random variable produced by the use of constant strategies $\overline{u}_{t-1}$ in the play of the game. Define $\gamma_t^{1d}(\omega, \overline{u}_{t-1})$ by

$$\gamma_t^{1d}(\omega, \bar{u}_{t-1}) = \hat{\gamma}_t^{1b}(\bar{y}_t^{1d}(\omega, \bar{u}_{t-1}), \bar{u}_{t-1}) \quad \text{a.e.}$$

Then, by construction, γ_t^{1d} is a jointly measurable function of ω and $\bar{u}_{t-1}$; as a function of ω it is measurable on $F_t^1\{\bar{u}_{t-1}\}$. Thus γ_t^1 is an admissible signaling-free strategy. But, the way γ_t^{1d} was constructed,

$$J^1(\underline{\gamma}^{1d}, \underline{\gamma}^{20}) = J^1(\hat{\underline{\gamma}}^{1b}, \hat{\underline{\gamma}}^{2c}) \quad \text{a.e.}$$

Hence,

$$\begin{aligned} E\{J^1(\underline{\gamma}^{1d}, \underline{\gamma}^{20})\} &= E(J^1(\hat{\underline{\gamma}}^{1b}, \hat{\underline{\gamma}}^{20})) \\ &< E(J^1(\underline{\gamma}^{10}, \underline{\gamma}^{20})) \\ &= E(I_0^1(\omega, \underline{\gamma}^{10}, \underline{\gamma}^{20})) \end{aligned}$$

In particular, since $I_0^1(\omega, \underline{\gamma}^{10}, \underline{\gamma}^{20}) \leq I_0^1(\omega, \underline{\gamma}^{1d}, \underline{\gamma}^{20})$ a.e. by the equilibrium property of $\underline{\gamma}^{10}$, $\underline{\gamma}^{20}$, and

$$\begin{aligned} E\{J^1(\underline{\gamma}^{1d}, \underline{\gamma}^{20})\} &= E\{E\{J^1(\underline{\gamma}^{1d}, \underline{\gamma}^{20}) \mid z_0^1\}\} \\ &= E\{I_0(\omega, \underline{\gamma}^{1d}, \underline{\gamma}^{20})\} \quad \text{by Theorem IV.6,} \end{aligned}$$

then this leads to a contradiction. Thus, for all admissible closed-loop strategies $\hat{\underline{\gamma}}^1$,

$$E\{J^1(\hat{\underline{\gamma}}^1, \hat{\underline{\gamma}}^{2c})\} \geq E\{J^1(\hat{\underline{\gamma}}^{1c}, \hat{\underline{\gamma}}^{2c})\}$$

A similar argument shows

$$E\{J^2(\hat{\underline{\gamma}}^{1c}, \hat{\underline{\gamma}}^{2c})\} \geq E\{J^2(\hat{\underline{\gamma}}^{1c}, \hat{\underline{\gamma}}^2)\}$$

for all admissible $\hat{\underline{\gamma}}^2 : Z_t^2 \to U_t^2$.

Hence, $(\hat{\underline{\gamma}}^{1c}, \hat{\underline{\gamma}}^{2c})$ is a closed-loop equilibrium.

IV.7 Example

Consider the game in extensive form presented in Section IV.5. This is not a game of equivalent information. The signaling-free Stackelberg equilibrium can be obtained from Theorem IV.3

$$\begin{aligned}
\gamma_0^{2*}(u_0^1) &= \underset{u^2}{\operatorname{argmin}}\ E\{Jh_N(x_N)\,|\,u_0^1\} \\
&= 0 \quad \text{if } u_0^1 = 0, \text{ cost-to-go} \quad 4 \\
&= 1 \quad \text{if } u_0^1 = 1, \text{ cost-to-go} \quad 4.5
\end{aligned}$$

$$\begin{aligned}
\gamma_0^{1*}(\theta_0) &= \underset{u}{\operatorname{argmin}}\ E(J^1(u,\ \gamma_0^2(u))\,|\,\theta_0) \\
&= 0 \text{ if } \theta_0 = 0,\ I_0^1 = 2 \\
&= 1 \text{ if } \theta_0 = 1,\ I_0^1 = 5
\end{aligned}$$

In terms of strategies in Section IV.5,

$$\gamma_0^{2*}(u_0^1) = u_0^1, \quad \gamma_0^{1*}(\theta_0) = \theta_0$$

with expected cost of 3.5. The signaling-free equilibrium is not a closed-loop equilibrium, as the strategy $\gamma_0^{2*}(u_0^1) = 1$ yields a lower expected cost for player 2. Similarly, suppose now that the extensive game is of the form shown in Figure IV.4.

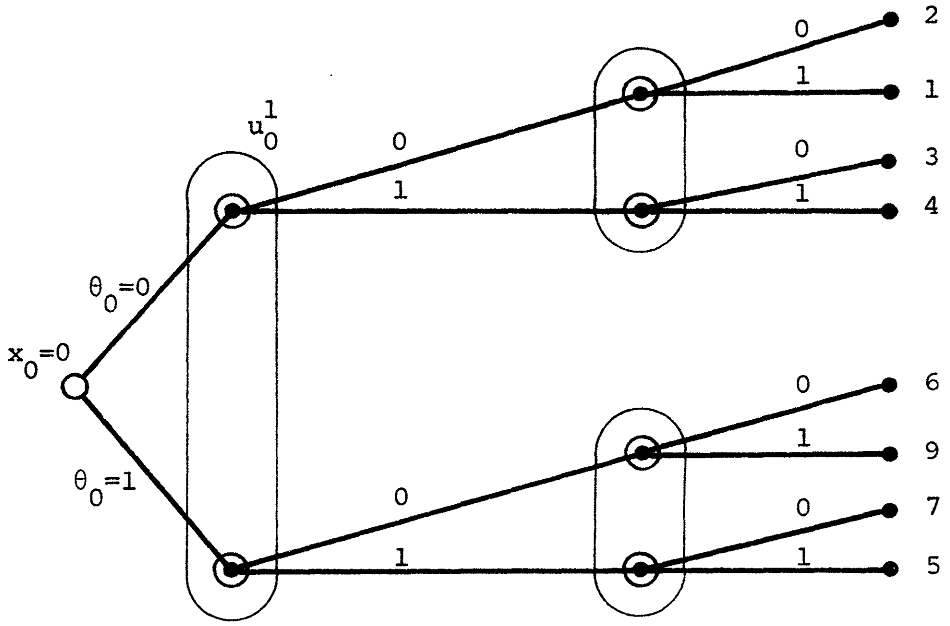

Figure IV.4

This game can be decomposed in a manner indicated by Kuhn (K2); the decomposition is similar to searching for a signaling-free equilibrium. The admissible strategies for player 1 are now

$$\gamma_0^1 = 0,\ \gamma_0^1 = 1.$$

The signaling-free equilibrium strategies for player 2 are the same. That is,

$$\gamma_0^{2*} = u_0^1.$$

It is easy to inspect that $\gamma_0^{1*} = 0$, with expected cost of 4. Furthermore, consider

$$E(J^2(\gamma_0^{1*},\ \gamma_0^2)) = 4 \quad \text{if } \gamma_0^2(0) = 0$$

$$= 5 \quad \text{if } \gamma_0^2(0) = 1.$$

Since $\gamma_0^{2*}(0) = 0$, then, for all γ_0^{2*} admissible

$$E(J^2(\gamma_0^{1*}, \gamma_0^2)) \geq E(J^2(\gamma_0^{1*}, \gamma_0^{2*}))$$

Hence the signaling-free equilibrium is a closed-loop equilibrium in this case.

In general, the assumption that the players will not make any assumptions as to the strategy used by other players in the past prevents the players from deriving additional information about ω from observation of each other's decisions. Hence, the signaling-free equilibria will be different from closed-loop equilibria, where learning is permitted. In the special case where the players have equivalent information about the uncertainties, then nothing can be learned from observation of the other player's decisions. In this special case, Theorem IV.5 establishes the equivalence of signaling-free equilibria and closed-loop equilibria.

CHAPTER V

Signaling-Free Equilibria in LQG Stackelberg Games

V.1 Introduction

This chapter studies signaling-free equilibria in a class of games where the state evolution equation (2.1) is linear in the states, controls, and noise, the observation equations (2.2) are linear, and the costs (2.3) are quadratic. Furthermore, the uncertainties are assumed to have a joint Gaussian probability density. For this important class of games, signaling-free equilibria can be found exactly. The first part of the chapter deals with the solution of games with nested information; that is, when one player (the leader) has knowledge of all the information available to the follower. The second part deals with the solution of games with general information.

Throughout this section, the following notation will be used: Matrices are denoted by upper case letters, vectors by lower case letters, stages by arguments of a function, (t). The notation for a function z on its value at stage t is z(t). Euclidean n-space is R^n. Lower case greek letters denote Gaussian random variables of zero mean, while upper case greek letters denote their covariance. M' is the transpose of M, tr(M) its trace. Gaussian random variables will be indicated as $x = N(\bar{x},\Sigma)$ where $\bar{x}$ is its mean, Σ its covariance. With this notation, the class of problems considered is described in the following section.

V.2 Problem Statement

Problem 1: Consider the stochastic multistage game where $x(t)$, the system state at stage t, evolves as:

$$x(t+1) = A(t)x(t) + B(t)u(t) + C(t)v(t) + \theta(t), \quad t=0,\dots N-1 \tag{5.1}$$

$$y_1(t) = H_1(t)x(t) + \xi_1(t) \tag{5.2}$$

$$y_2(t) = H_2(t)x(t) + \xi_2(t), \quad t=0, \dots N-1$$

where $x(t) \in R^n$, $u(t) \in R^m$, $v(t) \in R^\ell$, $y_1(t) \in R^p$, $y_2(t) \in R^q$ for all t. The vectors $\theta(t)$, $\xi_1(t)$, $\xi_2(t)$ and $x(0)$ are independent gaussian random vectors for all t, where $x(0) = N(x_0, \Sigma_0)$; $\theta(t) = N(0, \Theta(t))$; $\xi_i(t) = N(0, \Xi_i(t))$. The matrices A, B, C, H_1 and H_2 are bounded for all t and of appropriate dimensions. In equation (5.1), $u(t)$ represents the decision of the leader, $v(t)$ the decision of the follower. Under these assumptions the state vector x is a discrete Gaussian Markov process whose evolution is described by (1).

At the start of the game, each player seeks to mimimize the expected value of costs J_i, where

$$J_i(x,u,v,0) = x'(N)Q_i(N)x(N) + \sum_{t=0}^{N-1} (x'(t)Q_i(t)x(t) + u'(t)R_i(t)u(t) + v'(t)S_i(t)v(t)); \quad i = 1,2 \tag{5.3}$$

At a given stage t, the information $z_i(t)$ includes exact knowledge of the system dynamics (1), the measurements rules of both players (2) and

and the cost functionals of both players (3). Additionally, they include exact knowledge of all decisions made by each player up to stage t-1 and the statistics of random elements $\theta(t)$, $\xi_1(t)$ and $\xi_2(t)$ for all t. The leader's and follower's information $z_i(t)$ contain the values of all measurements up to stage t available to the leader and follower, respectively. Additionally, the Stackelberg nature of the game implies that the follower's information $z_2(t)$ contains the exact value of the leader's decision at time t, $u(t)$.

Signaling-free equilibria for Stackelberg games are defined in Section II.3 by equations (2.17)-(2.22). Furthermore, theorem IV.4 gives necessary and sufficient conditions for a sequence of admissible strategies to be a signaling-free equilibrium in terms of a constructive procedure described in equations (4.20) and (4.21). In terms of the LQG problem statement, these equations can be written as

$$\begin{aligned}\gamma_t^{10}(z_1(t)) = \underset{u}{\operatorname{argmin}}\ E\{&x'(t)Q_1(t)x(t) + u'R_1(t)u \\ &+ v's_1(t)v + J_1^*(t+1)\,|\,z_1(t),\ v=\gamma_1^{20}(z_2(t))\}\end{aligned} \tag{5.4}$$

$$\begin{aligned}\gamma_t^{20}(z_2(t)) = \underset{v}{\operatorname{argmin}}\ E\{&x'(t)Q_2(t)x(t) + u_1'(t)R_2(t)u_1(t) \\ &+ v's_2(t)v + J_2^*\ (t+1)\,|\,z_2(t),\ u(t+1) = \gamma_{t+1}^{10}(z_1(t))\}\end{aligned} \tag{5.5}$$

where $J_1^*(t+1) = \hat{I}_1(z_1(t), \underline{\gamma}_{t+1}^{10}, \underline{\gamma}_{t+1}^{20})$ and $J_2^*(t+1) = \hat{I}_2(z_2(t), \underline{\gamma}_{(t+1)}^{10}, \underline{\gamma}_t^{20})$. The solution of equations (5.4) and (5.5) can be obtained recursively. The expected costs-to-go at the Nth stage for the players are:

$$J_1^*(N) = E\{x'(N)Q_1(N)x(N) \mid z_1(N)\} \tag{5.6}$$

$$J_2^*(N) = E\{x'(N)Q_2(N)x(N) \mid z_2(N)\} \tag{5.7}$$

The solution of these equations is now straightforward. The next sections discuss the solution of these equations for some special cases of the problem stated previously.

V.3 The Deterministic Case

Suppose both players have perfect information of the state, as a special case of problem 1. Then, since $z_1(N)$ and $z_2(N)$ both contain knowledge of $x(N)$, (5.6) and (5.7) become

$$J_1{}^*(N) = x'(N)Q_1(N)x(N) \tag{5.8}$$

$$J_2{}^*(N) = x'(N)Q_2(N)x(N) \tag{5.9}$$

Make the inductive assumption that $J_\ell{}^*(t) = x'K_\ell(t)x + \Pi_\ell(t)$, $\ell = 1,2$ for some deterministic matrix $K_\ell(t)$ and function $\Pi_\ell(t)$. This assumption is established in appendix 2 using the principles of dynamic programming. The equilibrium strategies are:

$$u^*(t) = -W(t)^{-1}Y(t)x(t) = \gamma_t^{10}(x(t)) \tag{5.10}$$

$$\begin{aligned} v_0(u,t) &= -(S_2(t) + C'(t)K_2(t+1)C(t))^{-1}C'(t)K_2(t+1)(A(t)x(t)+B(t)u(t)) \\ &= -\Lambda(t)(A(t)x + B(t)u) = \gamma_t^{20}(x(t),\ u(t)) \end{aligned} \tag{5.11}$$

where, dropping the argument t for brevity,

$$W = R_1 + B'\Lambda'S_1\Lambda B + B'(I-C\Lambda)'K_1(t+1)(I-C\Lambda)B \tag{5.12}$$

$$Y = B'\Lambda'S_1\Lambda A + B'(I-C\Lambda)'K_1(t+1)(I-C\Lambda)A \tag{5.13}$$

$$L = Q_1 + A'\Lambda'S_1\Lambda A + A'(I-C\Lambda)'K_1(t+1)(I-C\Lambda)A \tag{5.14}$$

This assumes the required inverses exist. Further on sufficient conditions for the existence of these inverses are stated. The optimal costs-to-go are

$$J_1^*(t) = x'K_1(t)x + \Pi_1(t) \tag{5.15}$$

$$J_2^*(t) = x'K_2(t)x + \Pi_2(t) \tag{5.16}$$

where

$$K_1(t) = L(t) - Y'(t)W^{-1}(t)Y(t); \quad K_1(N) = Q_1(N) \tag{5.17}$$

$$\Pi_1(t) = \Pi_1(t+1) + \mathrm{Tr}\left\{\Theta(t)K_1(t+1)\right\} \;; \Pi_1(N) = 0 \tag{5.18}$$

$$\begin{aligned} K_2(t) = Q_2 + (A-BW^{-1}Y)'(K_2(t+1))(I-C\Lambda)(A - BW^{-1}Y) \\ + Y'W^{-1}R_2W^{-1}Y; \; K_2(N) = Q_2(N) \end{aligned} \tag{5.19}$$

$$\Pi_2(t) = \Pi_2(t+1) + T\,\{\Theta(t)K_2(t+1)\}; \; \Pi_2(N) = 0 \tag{5.20}$$

The right hand side of the four equations depend only on known parameters of the game (i.e., $A(t)$, $B(t)$, etc.) and future values of the cost-to-go matrices $K_1(t)$ and $K_2(t)$, $\Pi_1(t)$ and $\Pi_2(t)$. Thus, they are solved backwards in time using the initial conditions $K_1(N) = Q_1(N)$, $K_2(N) = Q_2(N)$, $\Pi_1(N) = 0 = \Pi_2(N)$.

The solutions of these equations provide equilibrium feedback matrices for the implementation of the adaptive Stackelberg strategies. Figure 1 contains a block representation of the equilibrium strategies:

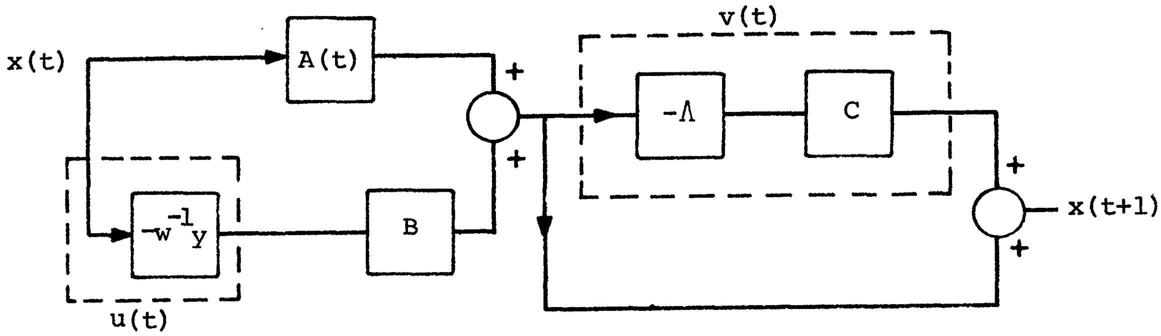

Figure V.1

Notice that the follower implements his optimal strategy after the leader does so. Still, the existence of a solution is not guaranteed unless W^{-1} exists for all t, and $(S_2 + C'K_2(t+1)C)^{-1}$ exists for all t. Furthermore, the solutions given must minimize the expressions for $J_1(x,u,v,t)$ and $J_2(x,u,v,t)$.

Sufficient conditions for the existence of the inverses and the minimum are given as follows: Let Q_1 and S_1 be positive semidefinite for all t, and $R_1(t)$ positive definite. The expected cost-to-go is thus non-negative so $K_1(t) \geq 0$. Thus, from (20), W is positive definite, hence invertible. Similarly, if we let Q_2 and R_2 be positive semidefinite, and S_2 be positive definite, then $(S_2 + C'K_2(t+1)C)$ is invertible. In addition, these conditions ensure that the functionals minimized are strictly convex over the whole space, so the minimums are unique. Thus, the optimal strategies (5.4), (5.5) are the unique equilibrium signaling-free Stackelberg strategies for problem 2, under the conditions previously described. The next section examines the

the more general case where measurement noise is present in the problem.

V.4 The Stochastic Case

Consider the special case of problem 1 where the leader knows both $y_1(t)$ and $y_2(t)$ in (2), whereas the follower knows only $y_2(t)$. So, for any t, $z_1(t) \supset z_2(t)$, implying that the information sets are nested. In addition, note that neither player can assume the other has played optimally in the past. This assumption is implicit in the definition of signaling-free equilibria; it is stated here explicitly so that the possibility of information transfer between players through their controls is avoided completely.

The optimal strategies for this problem are derived in Appendix 2 as

$$v_0(u,t) = -\Lambda(t)\ (A(t)\hat{x}_2 + B(t)u) = \gamma_t^{20}(z_2(t)) \tag{5.21}$$

$$u^*(t) = -W(t)^{-1}Y(t)\hat{x}_1 - W(t)^{-1}M(t)\ (\hat{x}_2 - \hat{x}_1) = \gamma_t^1\ (z_1(t)) \tag{5.22}$$

$$J_1^*(t) = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2-\hat{x}_1 \end{pmatrix}' \begin{pmatrix} K_{1A}(t) & K_{1B}(t) \\ K_{1B}'(t) & K_{1C}(t) \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2-\hat{x}_1 \end{pmatrix} + \Pi_1(t) \tag{5.23}$$

$$J_2^*(t) = \hat{x}_2'K_2(t)\hat{x}_2 + \Pi_2(t) \tag{5.24}$$

where $\hat{x}_1(t) = E(x|z_1(t))$, $\hat{x}_2(t) = E(x|z_2(t))$, $W(t)$, $\Lambda(t)$ and $Y(t)$ are defined in (5.11), (5.12) and (5.13) as in the deterministic case with $K_{1A}(t)$ replacing $K_1(t)$. The following equations are derived in Appendix 2:

$$K_{1A}(t) = Q_1 + A'\Lambda'S_1\Lambda A + A'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A-Y'W^{-1}Y \tag{5.25}$$

$$\begin{aligned} K_{1B}(t) = {} & A'\Lambda'S_1\Lambda A + A'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A \\ & + A'(I-C\Lambda)'(K_{1B}(t+1)-K_{1A}(t+1))A \\ & - A'(I-C\Lambda)'K_{1B}(t+1)G_2(t+1)H_2(t+1)A - Y'W^{-1}M \end{aligned} \tag{5.26}$$

$$\begin{aligned} M(t) = {} & B'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A + B'(I-C\Lambda)'(K_{1B}(t+1)-K_{1A}(t+1))A \\ & + B'\Lambda'S_1 A - B'(I-C\Lambda)'K_{1B}(t+1)G_2(t+1)H_2(t+1)A \end{aligned} \tag{5.27}$$

$$\begin{aligned} K_{1C}(t) = {} & -M'W^{-1}M + A'\Lambda'(S_1 + C'K_{1A}(t+1)C)\Lambda A \\ & - 2A'\Lambda'C'K_{1B}(t+1)(I-G_2(t+1)H_2(t+1))A \\ & + A'(I-G_2(t+1)H_2(t+1))'K_{1C}(t+1)(I-G_2(t+1)H_2(t+1))A \end{aligned} \tag{5.28}$$

$$\begin{aligned} \Pi_1(t) = {} & \Pi_1(t+1) + Tr\{(\textstyle\sum_1(t+1\ t) - \sum_1(t+1)\ (K_{1A}(t+1)-2K_{1B}(t+1)+K_{1C}(t+1))\} \\ & + Tr\{G_2(t+1)H_2(t+1)\textstyle\sum_1(t+1|t)(2K_{1B}(t+1)-2K_{1C}(t+1))\} \\ & + Tr\{\textstyle\sum_1(t)Q_1(t)\} + Tr\ G_2(t+1)(\Xi_2(t+1)+H_2\sum_1(t+1|t)H_2') \\ & G_2'K_{1C}(t+1)\} \end{aligned} \tag{5.29}$$

$$\begin{aligned} K_2(t) = {} & Q_2 + (A - BW^{-1}Y)'(K_2(t+1) - \Lambda'C'K_2(t+1))(A - BW^{-1}Y) \\ & + Y'W^{-1}R_2W^{-1}Y \end{aligned} \tag{5.30}$$

$$\begin{aligned} \Pi_2(t) = {} & \Pi_2(t+1) + Tr\{\textstyle\sum_2(t)Q_2(t)\} + Tr\{(\sum_2(t+1|t) - \sum_2(t+1))K_2(t+1)\} \\ & + Tr\{(\textstyle\sum_2(t) - \sum_1(t))\ (M-Y)'W^{-1}R_2+B'(K_2(t+1)-\Lambda'C'K_2(t+1)) \\ & BW^{-1}(M-Y)\} \end{aligned} \tag{5.31}$$

where the initial conditions are $K_{1B}(N) = K_{1C}(N) = 0$, $K_{1A}(N) = Q_1(N)$, $K_2(N) = Q_2(N)$, $\Pi_1(N) = Tr\{\sum_1(N)Q_1(N)\}$, $\Pi_2(N) = Tr\}\sum_2(N)Q_2(N)\}$. As in the deterministic case, existence and uniqueness of the optimal Stackelberg strategies follows from assuming $S_2 > 0$, $R_1 > 0$, S_1, $R_2, Q_1, Q_2 \geq 0$ for all stages. The right hand sides of equations (33)-(39) can be determined at stage t if the cost-to-go matrices $K_{1A}(t+1)$, $K_{1B}(t+1)$, $K_{1C}(t+1)$ and $K_2(t+1)$ are known. Initial conditions at stage N are known, so these equations can be solved backwards a priori, because the only terms involved are the parameters of the game $A(t)$, $B(t)$, etc. which are known to both players, and the covariance of the estimates, which can be computed a priori.

There are several important aspects in the solution. First, the recursive relations (5.25) and (5.30) are identical to the relations (5.17) and (5.19) in the deterministic case, with the same initial conditions, so that the solutions $K_{1A}(t)$ and $K_2(t)$ in (5.25), (5.30) are equal to $K_1(t)$ and $K_2(t)$ in (5.17) and (5.19). Thus, as far as the follower is concerned, he is playing a "separation principle" strategy which consists of the optimal deterministic feedback law of his best estimate of the state. The leader, on the other hand, knows both his own estimate and the follower's estimate, so that his optimal strategy includes a term taking advantage of the difference in estimates. When both estimates are the same, the leader also plays as in the separation principle of optimal control.

A special case of problem 1 would be the case where the follower has no measurements, in which case $G_2 = 0$. Another case of interest is where the leader has perfect information on the state, in addition to knowing the measurements of the follower. In both of these cases the information sets are nested, so $z_1 \supset z_2$. If the information sets were not nested, then both players would be unable to estimate the other player's estimate at each stage. Thus, at each stage, the optimal strategies would include, in addition to their own estimates, terms involving the estimate of the other player's estimates of the state in the future. Carried out to N stages, the augmented state vector would be roughly of dimension 2nN for each player.

This leads to estimators of much larger dimension than the system itself, hence it is impractical. An interesting variation of problem 1 is where, instead of knowing the other player's past decisions exactly, there is some channel noise involved, so a noise-corrupted version of the decisions is known. If this noise is zero-mean, additive, Gaussian and statistically independent of the other noises in the system, then the equilibrium signalling-free strategies remain unaffected, but the expected optimal cost increases. The reason for this is that neither player extracts information from the other player's past decisions. Thus, past decisions merely affect the mean value of the random vector x(t), and to include additive channel noise with the above properties would just leave x(t) as a Gauss-Markov process with increased covariance. It has no effect on the equilibrium because of its independent with the other random elements in the system implies that the conditional expectations of x(t) given $z_i(t)$ would not be affected.

V.5 LQG Games with Nonnested Information

This section discusses signalling-free equilibria in the LQG games of section V.2 when the players' information is nonnested. The main result of this section is theorem V.1, stated below:

Theorem V.1: For LQG games with nonnested information, signaling-free equilibrium strategies are affine functions of the active information $z_i(t) = (u(0), v(0), y_i(0), \ldots, y_i(t))$ where the coefficients are predetermined by the parameters of the game.

Proof of Theorem V.1:

The proof is by induction on N, using theorem IV.3 to find the equilibrium strategies. From equations (5.4) and (5.5),

$$\gamma^{20}_{N-1} = \underset{K}{\operatorname{argmin}}\, E\{x'Q_2x + u'R_2u + v'S_2v + x'(N)Q_2(N)x(N)|z_2(N-1)\}$$

where the argument N-1 has been ommitted for brevity.

Since $x(N) = Ax + Bu + Cv + \Theta$, the above minimization can be carried out as in appendix 2, yielding

$$\gamma^{20}_{N-1}(z^2_2(N-1)) = -(S_2+C'Q_2(N)C)^{-1}C'Q_2(N)\{AE\{x|z_2(N-1)\} + Bu\} \quad (5.32)$$
$$= \Lambda^2_{N-1}(z_2(N-1))$$

where Λ^2_{N-1} is an affine operator mapping $z_2(N-1)$ into U^2_{N-1} whose coefficients depend on the parameters of the game, as shown in Appendix 1. Let λ^i_t be the class of affine operators Λ^i_t from $Z^i_t \rightarrow U^i_t$ such that the coefficients are computed apriori from the parameters of the game.

From Appendix 1, it also follows that

$$\hat{I}^{2}_{N-1}(z_2(N-1),\ \gamma^{20}_{N-1}) = z_2'(N-1)K^2_{N-1}z_2(N-1)+L^2_{N-1}z_2(N-1)+\Pi^2_{N-1}$$

where K^2_{N-1}, L^2_{N-1} and Π^2_{N-1} are determined by the parameters of the game. Similarly, equation (5.4) implies

$$\gamma^{1*}_{N-1} = \underset{u}{\operatorname{argmin}}\ E\{u'R_1u+x'Q_1x + x(N)'Q_1(N) + v'S_1v \mid z_1(N-1),\ u,\ v = \Lambda^2_{N-1}(z_2(N-1))\} \qquad (5.33)$$

$$= \underset{u}{\operatorname{argmin}}\ E\{2x'(Q_1+A'Q_1(N)A)x + u'(R_1+B'Q_1(N)B)u$$

$$+ v'(S_1+C'Q_1(N)C)v + 2u'B'Q_1(N)(Ax+Cv+\theta)+2v'C'Q_1(N)(Ax+\theta)$$

$$+ 2x'A'Q_1(N)\theta + \theta'Q_1(N)\theta \mid u,\ \Lambda^2_{N-1}(z_2(N-1))=v\}$$

The expectations in (5.33) can be considered individually. For any matrix M, $E\{x'Mx \mid z_1(N-1)\}$ is shown in appendix 1 to be a general quadratic in $z_1(N-1)$.

$$E\{u'Mx \mid z_1(N-1),\ u\} = u'ME\{x \mid z_1(N-1)\}$$

$$E\{x'M\theta\} = 0,\ E\{\theta'M\theta\} = \mathrm{Tr}\{\Theta M\}$$

$$\hat{v} = E\{v \mid z_1^*(N-1),\ u,v = \Lambda^2_{N-1}(z_2(N-1))\} = \Lambda^2_{N-1}(E\{z_2(N-1) \mid z_1(N-1),u)$$

Entries in $z_2(N-1)$ are of the form $y_2(t)$, $u(t)$, $v(t)$. $u(t)$ and $v(t)$ are known constants in $z_1(N-1)$, so from equation (5.2),

$$E\{y_2(t) \mid z_1(N-1)\} = H_2(t)E(x(t) \mid z_1(N-1))\varepsilon\ \lambda^1_{N-1}$$

Thus, since every entry of $E\{z_2(N-1)|z_1(N-1)\}$ belongs to λ^1_{N-1}, then $E\{v|z_1(N-1),u,v = \Lambda^2_{N-1}(z_2(N-1))\}$ is a general quadratic in terms of $z_1(N-1)$ and u. Similarly,

$$E\{v'Mv|z_1(N-1),\ u,\ v = \Lambda^2_{N-1}(z_2(N-1))\} =$$

$$\hat{v}'M\hat{v} \ + \mathrm{Tr}\{ME\}$$

where $E = E\{(v-\hat{v})(v-\hat{v})'|z_1(N-1),u,v = \Lambda^2_{N-1}(z_2(N-1))\}$. Consider a typical entry of $(v-\hat{v})(v'-\hat{v}')$. It is of the forms 0, or

$$\alpha(y_2(t)-\overline{y}_2(t))\ \ (y_2(\tau)-\overline{y}_2(\tau))'$$

since $v \in \lambda^2_{N-1}$, so that $v-\hat{v}$ has entries 0 (corresponding to u(t), v(t)) or $\alpha(y_2(t) - \hat{y}_2(t))$.

It is shown in appendix 1 that $E\{(y_2(t)-\hat{y}_2(t))(y_2(T)-\hat{y}_2(T))'|z_1(N-1)\}$ is a quadratic in $z_1(N-1)$ which can be computed apriori from the parameters of the game.

Also, $E\{v'Mx|z_1(N-1),\ u,\ v = \Lambda^2_{N-1}(z_2(N-1))\} =$

$$E\{E(v'Mx|z_1(N-1),\ z_2(N-1))|z_1(N-1),\ u,\ v = \Lambda^2_{N-1}(z_2(N-1))\}$$

$$= E\{v'ME\{x|z_1(N-1),\ z_2(N-1)\}|z_1(N-1),u,v = \Lambda^2_{N-1}(z_2(N-1))\}$$

From Appendix 1, $E\{x|z_1(N-1),\ z_2(N-1)\}$ is an affine function of $z_1(N-1)$ and $z_2(N-1)$. Using the results from the previous expectations yields

$E\{v'Mx|z_1(N-1),\ u,\ v = \Lambda^2_{N-1}(z_2(N-1))\}$ is a general quadratic in $z_1(N-1)$, u. Thus, (5.33) can be expanded; when expectations are taken every term is at most a general quadratic in $z_1(N-1)$, and u then (5.33) can be rewritten as

$$\gamma^{10}_{N-1}\ (z_1(N-1)) = \underset{u}{\text{argmin}} \begin{pmatrix} z_1(N-1) \\ u \end{pmatrix}' \begin{pmatrix} M & L \\ L' & N \end{pmatrix}_{N-1} \begin{pmatrix} z_1(N-1) \\ u \end{pmatrix} + LL_{N-1} \begin{pmatrix} z_1(N-1) \\ u \end{pmatrix} + \Pi_{N-1} \qquad (5.34)$$

where M, L, N, LL and Π are determined a priori from the parameters of the game, and the matrix

$$\begin{pmatrix} M & L \\ L' & N \end{pmatrix}_{N-1}$$

is positive semidefinite, and the matrix N is positive definite. This is guaranteed by the definiteness assumptions discussed in section V.3.

Hence, since for all $z_1(N-1)$, (5.34) is a minimization in Euclidean space,

$$\gamma^{10}_{N-1}\ (z_1(N-1)) = \Lambda^1_{N-1}(z_1(N-1))$$

where $\Lambda^1_{N-1} \ \varepsilon\ \lambda^1_{N-1}$ and

$$J^*_1(N-1) = z_1'(N-1)K^1_{N-1}z_1(N-1) + L^1_{N-1}z_1(N-1) + \Pi^1_{N-1} \qquad (5.35)$$

Now, assume inductively that, for $t \geq T + 1$,

$$\hat{I}^1_t(z_1(t), \underline{\gamma}^{10}_t) = z_1'(t)K^1_t z_1(t) + L^1_t z_1(t) + \Pi'_t \tag{5.36}$$

$$\hat{I}^2_t(z_2(t), \gamma^{10}_{t+1}, \underline{\gamma}^{20}_t) = z_2'(t)K^2_t z_2(t) + L^2_t z_2(t) + \Pi^2_t$$

and $\gamma^{i0}_t(z_i(t)) \in \lambda^i_t$. Then equations (5.4) and (5.5) imply

$$\begin{aligned}\gamma^{2*}_T(z_2(T)) = \underset{u}{\operatorname{argmin}}\ & E\{x'Q_2x + u'R_2u + v'S_2v + \\ & + z_2'(T+1)K^2_{T+1}z_2(T+1) + L^2_{T+1}z_2(T+1) \\ & + \Pi^2_{T+1} | z_2(t),\ v,\ u(T+1) = \gamma^{10}_{T+1}(z_1(T+1))\}\end{aligned} \tag{5.37}$$

where the subscript T was ommitted for brevity. Taking expectations as before, the first three terms are general qudratics in $z_2(T)$, $U(T)$

The third term becomes

$$E\{z_2^1(T+1)Mz_2(T+1)|z_2(T), v, u^1_T = \Lambda^1_{T+1}(z_1(T+1))\}$$

$$= E \begin{pmatrix} z_2(T) \\ v \end{pmatrix}' M^a \begin{pmatrix} z_2(T) \\ v \end{pmatrix} + 2\begin{pmatrix} z_2(T) \\ v \end{pmatrix}' M^b \begin{pmatrix} y_2(T+1) \\ u(T+1) \end{pmatrix}$$

$$+ \begin{pmatrix} y_2(T+1) \\ u(T+1) \end{pmatrix}' M^c \begin{pmatrix} y_2(T+1) \\ u(T+1) \end{pmatrix} \Bigg| z_2(T),\ u_1(T+1) = \Lambda^1_{T+1}(z_1(T+1))\}$$

Since $z_2(T+1) = \begin{pmatrix} z_2(T) \\ v(T) \\ y_2(T+1) \\ u(T+1) \end{pmatrix}$ and $M = \begin{pmatrix} M^a & M^b \\ M^b & M^c \end{pmatrix}$.

The first two terms are general quadratics in $\begin{pmatrix} z_2(T) \\ v \end{pmatrix}$. The last term can be shown in a fashion identical to the case for t=N-1, to be a general quadratic of $\begin{pmatrix} z_2(T) \\ v \end{pmatrix}$, since $y_2(T+1) = H_2(T+1)\ x(T+1) + \xi_2(T+1)$ so $E\left\{y_2(T+1) \mid \begin{pmatrix} z_2(T) \\ v \end{pmatrix}\right\}$ is an affine function of $\begin{pmatrix} z_2(T) \\ v \end{pmatrix}$. Thus (5.37) can be rewritten as

$$\gamma_T^{20}(z_2(T)) = \underset{v}{\operatorname{argmin}} \left\{\begin{pmatrix} z_2(T) \\ v \end{pmatrix}' K_T^{20} \begin{pmatrix} z_2(T) \\ v \end{pmatrix} + L_T^{20} \begin{pmatrix} z_2(T) \\ v_2 \end{pmatrix} + \Pi_T^{20}\right.$$

$$= \Lambda_T^2\ (z_T^2) \text{ where } \Lambda_T^2 \ \varepsilon\ \lambda_T^2$$

because the function to be minimized is a convex quadratic functional of v over Euclidean space.

Similarly,

$$\gamma_T^{10}(z_1(T)) = \underset{u}{\operatorname{argmin}}\ E\left\{x'Q_1x + u'R_1u + v'S_1v \right. \tag{5.38}$$

$$+ z_1(T+1)'\ K_{T+1}^1(z_1(T+1) + L_{T+1}^1 z_1(T+1)$$

$$\left. + \Pi_{T+1}^1 \middle| \begin{pmatrix} z_1(T) \\ u \end{pmatrix},\ v = \Lambda_T^2(z_T^2)\right\}$$

Taking expectations, the first two terms yield general quadratics in $\begin{pmatrix} z_1(T) \\ u \end{pmatrix}$

$$E\{v'S_1v \mid \begin{pmatrix} z_1(T) \\ u \end{pmatrix},\ v = \Lambda_T^2(z_2(T))\} = \bar{v}'S_1\bar{v} + Tr(\textstyle\sum_T^{21} S_1)$$

where $\bar{v} = E\{\Lambda_T^2(z_2^2(T)) \mid z_1(T),\ u\}$

$$\textstyle\sum_T^{21} = E\{(v-\bar{v})(v-\bar{v})' \mid z_1(T),\ u,\ v=\Lambda_T^2(z_T^2)\}$$

From appendix 1,

$\overline{v}'s_1\overline{v}$ is a general quadratic in $(z_1(T), u)$

$\sum_T^{21}$ is a constant determined a priori.

Also, $E\{z_1(T+1)'Mz_1(T+1) \mid z_1(T), u, v=\Lambda_T^2(z_2(T))\}$

$$= \begin{pmatrix} z_1(T) \\ u \end{pmatrix}' M^a \begin{pmatrix} z_1(T) \\ u \end{pmatrix} + 2\begin{pmatrix} z_1(T) \\ u \end{pmatrix}' M^b E\left\{\begin{matrix} v \\ y_1(T+1) \end{matrix} \middle| z_1(T), u, v=\Lambda_T^2(z_2(T))\right\}$$

$$+ E\left\{\begin{pmatrix} v \\ y_1(T+1) \end{pmatrix}' M^c \begin{pmatrix} v \\ y_1(T+1) \end{pmatrix} \middle| z_1(T), u, v=\Lambda_T^2(z_2(T))\right\}$$

which is also a general quadratic in $(z_1(T), u)$ by a similar argument based on appendix 1. Therefore, (5.38) can be expressed as

$$\gamma_T^{10}(z_1(T)) = \underset{u}{\text{argmin}} \begin{pmatrix} z_1(T) \\ u \end{pmatrix}' K_T^{10} \begin{pmatrix} z_1(T) \\ u \end{pmatrix} + L_T^{10}\begin{pmatrix} z_1(T) \\ u \end{pmatrix} + \Pi_T^{10}$$

$$= \Lambda_T^1(z_1(T)) \text{ where } \Lambda_T^1 \in \lambda_T^1$$

and $J_1^*(T)$ is a general quadratic in $z_1(T)$, so

$$J_1^*(T) = z_1'(T)K_T^1(z_1(T)) + L_T^1 z_1(T) + \Pi_T^1$$

Hence the theorem is proved by induction.

Theorem V.3 describes signaling-free equilibrium solutions of LQG Stackelberg games without nested information. Similar to the results of section V.4 for the nested information case, the equilibrium strategies are linear functions of the available information. In section V.4, these functions were linear in terms of two n-dimensional vectors $\hat{x}_1(t)$ and $\hat{x}_2(t)$ which summarized the information at stage t. In the nonnested case, the equilibrium strategies cannot be expressed in terms of vectors of lesser dimension than $z_1(t)$ or $z_2(t)$. The strategies are still linear maps of the information but; the information cannot be summarized in terms of two n-dimensional vectors. The added structure in the nested case permits the development of the solution in a fashion similar to the separation principle of optimal control; in the case of nonnested information, the equilibrium strategies are linear, but not related to the "separation principle".

CHAPTER VI

VI.1 Summary

Chapters I and II describe a general formulation of stochastic dynamic games. Within this formulation, several concepts of solution are studied, describing relative properties of these solutions. Significantly, these are properties associated with the general definition of the solutions as opposed to properties of the solutions in particular games. Two solutions are of particular interest: closed-loop equilibrium solutions and signaling-free equilibrium solutions. The concept of signaling-free equilibria is introduced to represent equilibrium solutions in games where a type of information restriction is present.

Chapter III discusses two examples which illustrate the results of Chapter II. The first example has a finite-state, finite-decision structure; it illustrates the methodology used in obtaining solutions, and points out the abundance of closed-loop equilibria with different costs. The second example has a two-state continuous-decision structure; it highlights the fact that in dynamic games solution concepts with dominant assumptions, such as the closed-loop Stackelberg solution, threats are a natural part of the optimal solution. Even in equilibrium solutions, a milder version of threats exists because of the ambiguity present in when strategies are chosen.

Chapter IV examines the problem of obtaining closed-loop and signaling-free equilibria in general stochastic dynamic games. For closed-loop equilibria, games are expressed in a standard form. Using this standard form, necessary conditions for equilibria are established, in terms of an equivalent deterministic problem which considers the unconditional probability density of the state as the new state of the system. A constructive technique for finding equilibria is established, thereby yielding sufficient conditions for a sequence of strategies to be an equilibrium.

The second part of Chapter IV deals with signaling-free equilibria. A constructive technique is described which provides a sufficient condition for equilibrium of a pair of strategy sequences. In special games with equivalent information for both players, it is shown that every signaling-free equilibrium is equivalent to a closed-loop equilibrium. This equivalence does not hold in general games, as shown by a counterexample.

Chapter V deals with signaling-free equilibria in a special class of games, called LQG games. Under the assumptions of LQG games, signaling-free equilibria can be obtained constructively. For general information it is shown that equilibrium strategies are affine functions of the available information, and the expected costs-to-go are quadratic functions of the available information. Under the restriction of nested information between players, the equilibrium strategies can be expressed as linear functions of a constant-dimension sufficient statistic. For the player with least information, equilibrium strategies follow the "certainty-equivalence" principle of optimal control theory.

VI.2 Discussion

There are numerous decision-making problems in various disciplines which can be modeled as dynamic games. Furthermore, each problem contains an intrinsic set of assumptions which indicate what solutions the different decision-makers are trying to achieve. The major thrust of this dissertation was to explore a class of solutions (equilibria) in dynamic problems characterized by alternate decision-making (Stackelberg games). There is considerable argument relating to the value of equilibria as solutions in non-zero sum games (H4 , Au2, Sh1). Equilibrium solutions are important in a large class of problems.

Prior to this research, most of the work dealing with equilibria concerned itself with closed-loop equilibria. The classical theory of games includes existence theorems for closed-loop equilibria in finite-state, finite-decision games (K2); additional work has been done in establishing existence and uniqueness of equilibria for classes of dynamic games (Ba2). Using some of the properties of Stackelberg games, this dissertation provides general theorems relating to necessary or sufficient conditions for existence of equilibria in an abstract formulation. Additionally, it provides the basis for constructive approaches to obtain equilibrium solutions. Hence, the work in the early part of Chapter IV provides a direct generalization of classical finite-state game theory to abstract games.

The second class of equilibria discussed in this dissertation is the class of signaling-free equilibria. This solution concept has not been studied before except under very special conditions (Ao2). However, it is a solution concept which rises naturally in dynamic games. Closed-loop equilibria are formulated in terms of both players' complete strategy sequences. At stage t, players have observed past decisions by all other players. Closed-loop equilibria assume that those decisions were the outcome of equilibrium strategies, whereas signaling-free equilibria make no assumptions as to which strategy was used in the past. Closed-loop equilibria thus represent a solution which is completely determined before the game is ever played, in terms of strategies; the players are assumed to have an unbreakable commitment to their equilibrium strategies. Signaling-free equilibria, on the other hand, make no assumption as to past strategies; hence, in the presence of non-equilibrium play by a player, the remaining strategies are still in equilibrium. Signaling-free equilibria therefore represent a solution which is determined as the game is played; the a priori commitment to strategies is not an essential part of the equilibrium solution.

The second half of Chapter IV studies signaling-free equilibria in general dynamic games. A constructive approach towards obtaining these equilibria is described using dynamic programming ideas. Several conditions must be met before this constructive approach can be completed, concerning questions of measurability and existence of minima. Specific games, such as LQG games discussed in Chapter V, meet these conditions naturally.

The two classes of equilibria are closely related in many cases. Particularly, in games of equivalent information, signaling-free equilibria correspond to closed-loop equilibria. For LQG games, the conditions guaranteeing the existence of signaling-free equilibria are met. The equilibrium strategies are affine functions of the available information at each stage for each player. Under a particular information constraint, these optimal strategies can be expressed in terms of a pair of sufficient statistics in a manner reminiscent of the "separation principle" in optimal control theory. It is important that, for this class of games, signaling-free equilibria can be readily obtained.

VI.3 Areas of future research

There are three main directions in which to continue the research topics discussed in this dissertation. The first direction lies in the area of applications. There are numerous examples of decentralized control theory, hierarchical systems, and large-scale systems which can be modeled as dynamic games. Using the constructive techniques of Chapter IV (in a fashion similar to Chapter V), one obtains equilibrium strategies in this important class of games. Some of the other solution concepts discussed in Chapters II and III are useful in examining large-scale systems. The special structure of problems in hierarchical systems and large-scale systems can be exploited to develop stronger results than those developed for general games in this dissertation. Also, in a fashion similar to Sandell's (Sa3) development of Finite State, Finite Memory control problems, the results can be specialized to FSFM games.

The second direction is closely related to applications; it consists of developing numerical methodology for obtaining approximations to solutions of dynamic games. This is essential to solving practical game problems. Chapter V provides algorithms which obtain signaling-free equilibria for LQG games. However, for general games, no practical methodology is given. This problem is currently under study.

The final direction of research lies in the theoretical area of exploring existence and uniqueness questions of signaling-free equilibria. The theorems discussed in Chapter IV have some restrictive assumptions which are not easy to verify in general problems. The theoretical issue lies in weakening those assumptions, or replacing them by assumptions which are easier to verify in general dynamic games. There are fundamental questions in functional analysis which relate to these problems, and must be examined closely.

APPENDIX 1

Discrete Time Estimation of LQG Systems

Consider the dynamical system described in equation (a):

$$x_{t+1} = A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2 + \theta_t; \quad t = 0, \ldots N-1 \tag{a}$$

with measurement equations

$$y_t^i = H_t^i x_t + \xi_t^i \quad ; \; i = 1,2; \; t = 0, \ldots N-1 \tag{b}$$

where x_t is the state vector at stage t, $x_t \in R^n$; A_t, B_t^i, H_t^i are real-valued matrices of appropriate dimension; u_t^1 is a <u>known constant</u> vector in R^p, u_t^2 is a <u>known constant</u> vector in R^q; y_t^i are observation vectors in R^{m_i}. The variables θ_t, ξ_t^1 and ξ_t^2 are random variables of appropriate dimension, such that for any t_1, t_2, t_3, θ_{t_1}, $\xi_{t_2}^1$, $\xi_{t_3}^2$ are mutually independent.

The random variables θ_t, ξ_t^1, ξ_t^2 are assumed to be Gaussian in density, with mean 0 and covariance Θ_t, Ξ_t^1, Ξ_t^2 respectively, denoted (for θ_t) by $\theta_t = N(0, \Theta_t)$. Furthermore, the initial condition x_0 is assumed to be unknown and independent of other variables, where $x_0 = N(\underline{x}_0, \Sigma_0)$. Assume also that Θ_t, $\Xi_t^i > 0$ for all t. The problem at hand can now be expressed as follows: Consider the LQG game of Chapter V. Given a sequence $(y_0^1, \ldots y_T^1)$ of measurements, and decisions $(u_T^1, \ldots u_{T-1}^2)$, denoted by z_T^1, find the values of

$$\hat{x}_t^T = E\{x_t | z_T^1\}$$

$$\Sigma_t^T = E\{(x_t-\hat{x}_t^T)(x_t-\hat{x}_t^T)' | z_T^1\}$$

for all $t \leq T$.

This problem is standard in estimation theory (As1, Ao1) The solution can be described as follows.

$$\Sigma_{t+1}^t = A_t\Sigma_t^t(A_t)' + \Theta_t \tag{c}$$

$$\hat{x}_{t+1}^t = A_t\hat{x}_t^t + B_t^1u_t^1 + B_t^2u_t^2 \tag{d}$$

$$\hat{x}_{t+1}^{t+1} = \hat{x}_{t+1}^t + \Sigma_{t+1}^{t+1}(H_{t+1}^1)'(\Xi_{t+1}^1)^{-1}(y_{t+1}^1 - H_{t+1}^1\hat{x}_{t+1}^t) \tag{e}$$

$$\Sigma_{t+1}^{t+1} = \Sigma_{t+1}^t - \Sigma_{t+1}^t(H_{t+1}^1)'(H_{t+1}^1\Sigma_{t+1}^t(H_{t+1}^1)' + \Xi_{t+1}^1)^{-1}H_{t+1}^1\Sigma_{t+1}^t \tag{f}$$

$$\hat{x}_t^T = \hat{x}_t^t - \Sigma_t^tA_t'(\Sigma_{t+1}^t)^{-1}(\hat{x}_{t+1}^t - \hat{x}_{t+1}^T) \tag{g}$$

where $t < T$.

$$\Sigma_t^T = \Sigma_t^t - \Sigma_t^tA_t'(\Sigma_{t+1}^t)^{-1}(\Sigma_{t+1}^t - \Sigma_{t+1}^T)(\Sigma_t^tA_t'(\Sigma_{t+1}^t)^{-1})^T \tag{h}$$

The initial conditions are given by

$$\Sigma_0^t = \Sigma_0, \quad \hat{x}_0^t = \underline{x}_0 \quad \text{for } t = -1. \tag{i}$$

Equations (c), (f) and (h) depend strictly on the a priori parameters of the game: the system matrices and the noise covariances. With the initial conditions given in (i), these three equations can be solved recursively to obtain Σ_{t+1}^T for any value of $t \leq T$. Equations (d), (e), and (g) are equations containing terms from the active information u_t^i, y_t^1, and parameters of the game. As indicated before, the covariances Σ_t^T can be computed from the a priori parameters of the game. Thus, equations (e) and (g) can be rewritten as

$$\hat{x}_{t+1}^{t+1} = \hat{x}_{t+1}^{t} + K_{t+1}^{1}(y_{t+1}^{1}(y_{t+1}^{1} - H^{1}\hat{x}_{t+1}^{t}) \qquad (j)$$

$$\hat{x}_{t}^{T} = \hat{x}_{t}^{t} - K_{t}^{2}(\hat{x}_{t+1}^{t} - \hat{x}_{t+1}^{T}) \qquad (k)$$

where K_t^1, K_t^2 are matrices determined by the parameters of the game.

Recursive solution of equations (d) and (j) establishes $\hat{x}_{t+1}^{t}$ as an affine combination of $\{u_k^1, u_k^2, y_k^1\}_{k=0,\ldots t}$ and $\hat{x}_t^t$ as an affine combination of $\{u_k^1, u_k^2, y_k^1\}_{k=0,\ldots t-1}$ and y_t^1, where the coefficients are determined by the parameters of the game. Similarly $\hat{x}_t^T$ can be expressed as an affine combinations of $\{y_k^1, u_k^1, u_k^2\}_{k=0,\ldots T-1}$, y_T^1 with coefficients determined by the a priori parameters of the game.

In particular,

$$E(y_t^2 | z_T^1) = H_t^2 \ \hat{x}_t^T$$

which is an affine function of z_T^1.

Also, for $t \leq T$

$$E\{y_t^{2\prime} M y_t^2 | z_T^1\} = (H_t^2 \ \hat{x}_t^T)' M (H_t^2 \ \hat{x}_t^T) + \mathrm{Tr}\{M(H_t^2 \Sigma_t^T H_t^{2\prime} + \Xi_t^2)\}$$

which is a general quadratic expression in terms of z_t^1 with coefficients determined a priori.

Similar relationships can be established for player 2's information $\{y_0^2, \ldots y_t^2\}$ and $\{u_0^1, u_0^2, \ldots u_{t-1}^1, u_{t-1}^2, u_t^1\}$.

APPENDIX 2

Consider the problem of Section V.4.

Let $\hat{x}_i(t) = E\{x(t)|z_i(t)\}$, $i = 1,2$.

$$\Sigma_i(t) = E\{(x(t) - \hat{x}_i(t))(x(t)-\hat{x}_i(t))'|z_i(t)\}$$

$$\Sigma_i(t+1|t) = E\{\Sigma_i(t+1)|z_i(t)\}$$

$$\overline{x}_i(t+1) = \hat{x}_i(t+1|t) = E\{x(t+1)|z_i(t)\}$$

$$J_i^*(N) = \hat{x}_i'Q_i(N)\hat{x}_i + Tr\{\Sigma_i(N)Q_i(N)\} \qquad \text{(a)}$$

where $\Sigma_i(N) = E\{(x - \hat{x}_i)(x - \hat{x}_i)'|z_i(N)\}$.

Assume that, at stage t, the optimal expected costs-to-go are given by

$$J_1^*(t) = \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t)-\hat{x}_1(t) \end{pmatrix}' \left(\begin{array}{c|c} K_{1A}(t) & K_{1B}(t) \\ \hline K_{1B}'(t) & K_{1C}(t) \end{array}\right) \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t)-\hat{x}_1(t) \end{pmatrix} + \Pi_1(t) \qquad \text{(b)}$$

$$J_2^*(t) = \hat{x}_2'(t)K_2(t)\hat{x}_2(t) + \Pi_2(t). \qquad \text{(c)}$$

These assumptions are consistent with (a). Proceed by induction to show that the above assumptions are indeed true. From (5.4) and (5.5), the optimal strategies are given by:

$$\gamma_t^{20}(z_2(t)) = \underset{v}{\text{argmin}}\ E\{x'Q_2x + u'R_2u + v'S_2v + \Pi_2(t+1) + \hat{x}_2'(t+1)K_2(t+1)\hat{x}_2(t+1)|z_2(t)\} \qquad \text{(d)}$$

Taking expectations, equation (d) becomes

$$\gamma_t^{20}(z_2(t)) = \underset{v}{\text{argmin}}\ [\hat{x}_2{}'Q_2\hat{x}_2 + u'R_2u + v'S_2v$$
$$+ x_2'(t+1|t)K_2(t+1)x_2(t+1|t) + \Pi_2(t+1)$$
$$+ \text{Tr}\{\Sigma_2Q_2\} + \text{Tr}\{(\Sigma_2(t+1|t) - \Sigma_2(t+1))K_2(t+1)\}]$$

From equation (d) in Appendix 1, (d) becomes

$$\gamma_t^{20}(z_2(t)) = \underset{v}{\text{argmin}}\ \{\hat{x}_2{}'(Q_2 + A'K_2(t+1)A)\hat{x}_2$$
$$+ u'(R_2 + B'K_2(t+1)B)u + v'(S_2 + C'K_2(t+1)C)v$$
$$+ 2v'C'K_2(t+1)(A\hat{x}_2 + Bu) + 2u'B'K_2(t+1)A\hat{x}_2$$
$$+ b_2(t)\} \tag{e}$$

where $b_2(t) = \Pi_2(t+1) + \text{Tr}(\Sigma_2Q_2) + \text{Tr}\{(\Sigma_2(t+1|t - \Sigma_2(t+1))K_2(t+1)\}$.
Differentiate now with respect to v and obtain the minimum:

$$\gamma_t^{20}(z_2(t)) = -(S_2 + C'K_2(t+1)C)^{-1}C'K_2(t+1)(A\hat{x}_2 + Bu)$$
$$= -\ \Lambda(A\hat{x}_2 + Bu) \tag{f}$$

Notice that $\Lambda(t)$ is defined as it was in equation (5.11). Using equation (5.4), one gets

$$\gamma_t^{10}(z_1(t)) = \underset{u}{\text{argmin}}\ \ E\Big\{x'Q_1x + u'R_1u + v_0{}'(u,t)S_1v_0(u,t) + \Pi_1(t+1)$$
$$+ \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2-\hat{x}_1 \end{pmatrix}' \begin{pmatrix} K_{1A}(t+1) & K_{1B}(t+1) \\ K_{1B}'(t+1) & K_{1C}(t+1) \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2-\hat{x}_1 \end{pmatrix} \Bigg| z_1(t)\Big\} \tag{g}$$

For any matrix L, one can determine the following expectations:

$$E\{x_1'(t+1)L\hat{x}_1(t+1)\,|\,z_1(t)\} = \overline{x}'(t+1)L\overline{x}(t+1) + \mathrm{Tr}\{\left(\Sigma_1(t+1)\,|\,t) - \Sigma_1(t+1)\right)L\} \tag{h}$$

$$E\{\hat{x}_1'(t+1)L\hat{x}_2(t+1)\,|\,z_1(t)\} = E\{E[\hat{x}_1'(t+1)L\hat{x}_2(t+1)\,|\,z_1(t) \cup z_2(t+1)]\,|\,z_1(t)\} \tag{i}$$

because $z_1(t) \cup z_2(t+1)$ induces a refinement of the partition induced by $z_1(t)$ alone. The inner expectation can be rewritten using the results of Appendix 1 as

$$E\{\hat{x}_1'(t+1)L\hat{x}_2(t+1)\,|\,z_1(t) \cup z_2(t+1)\}=E\{\hat{x}_1(t+1)\,|\,z_1(t) \cup z_2(t+1)\}'L\hat{x}_2(t+1) \tag{j}$$

Since $z_2(t+1) \cup z_1(t)$ involves adding the information of the measurement $y_2(t+1)$ to $z_1(t)$, then

$$E\{\hat{x}_1\,|\,z_2(t+1) \cup z_1(t)\} = E\{x\,|\,z_2(t+1) \cup z_1(t)\}$$

$$= \overline{x}_1(t+1) + \tilde{\Sigma}(t+1)H_2'(t+1)\Xi_2^{-1}(t+1)\left(y_2(t+1) - H_2(t+1)\overline{x}_1(t+1)\right) \tag{k}$$

where $\tilde{x}_1 = E\{x\,|\,z_2(t+1) \cup z_1(t)\}$, $\tilde{\Sigma} = E\{(x-\tilde{x}_1)(x-\tilde{x}_1)'\,|\,z_1(t) \cup z_2(t+1)\}$. The argument t+1 is dropped for brevity. From Appendix 1, one also knows

$$\hat{x}_2 = \overline{x}_2 + \Sigma_2 H_2 \Xi_2^{-1}(y_2 - H_2\overline{x}_2)$$

So, (i) becomes

$$E\{\hat{x}_1'L\hat{x}_2\,|\,z_1(t)\} =$$

$$E\{\left(\overline{x}_1+\tilde{\Sigma}H_2'\Xi_2^{-1}(y_2-H_2\overline{x}_1\right)'L\left(\overline{x}_2+\Sigma_2 H_2'\Xi_2^{-1}(y_2-H_2\overline{x}_2)\right)\,|\,z_1(t)\} \tag{l}$$

Now, y_2 conditioned on $z_1(t)$ is a gaussian random variable with mean $H_2\overline{x}_1$ and covariance $H_2\Sigma_1(t+1|t)H_2' + \Xi_2$. So

$$E\{\hat{x}_1'L\hat{x}_2|z_1(t)\} = \overline{x}_1'L\overline{x}_2 + \overline{x}_1'LG_2H_2(\overline{x}_1-\overline{x}_2) + \mathrm{Tr}\{\Sigma_1(t+1|t)LG_2H_2\}$$

$$\text{where } G_2 = \Sigma_2H_2'\Xi_2^{-1} \tag{m}$$

$$E\{\hat{x}_2'L\hat{x}_2|z_1(t)\} = E\{\left(\overline{x}_2 + G_2(Y_2-H_2\overline{x}_2)\right)'L\left(\overline{x}_2+G_2(Y_2-H_2\overline{x}_2)\right)|z_1(t)\}$$

$$= \overline{x}_2'L\overline{x}_2 + 2\overline{x}_2'LG_2H_2(\overline{x}_1-\overline{x}_2) + \mathrm{Tr}\{G_2\Xi_2G_2'L+GH_2\Sigma_1(t+1|t)H_2'G_2'L\} \tag{n}$$

Expand (g) and take expected values using (h), (m) and (n):

$$\begin{aligned}\gamma_t^{10}(z_1(t)) = \underset{u}{\operatorname{argmin}} \Big\{ & \hat{x}_1'(t)Q_1(t)\hat{x}_1(t) + u'(t)R_1(t)u(t) \\ & + \left(A(t)\hat{x}_2(t) + B(t)u(t)\right)'\Lambda'(t)S_1(t)\Lambda(t)\left(A(t)\hat{x}_2(t)+B(t)u(t)\right) \\ & + \overline{x}_1'(K_{1A}-2K_{1B}+K_{1C})\overline{x}_1 + 2\overline{x}_1'(K_{1B}-K_{1C})\overline{x}_2 \\ & + 2\overline{x}_1'(K_{1B}-K_{1C})G_2H_2(\overline{x}_1-\overline{x}_2) + \overline{x}_2'K_{1C}\overline{x}_2 + 2\overline{x}_2'K_{1C}G_2H_2(\overline{x}_1-\overline{x}_2) \\ & + (\overline{x}_2-\overline{x}_1)'H_2'G_2'K_{1C}G_2H_2(\overline{x}_2-\overline{x}_1) + \Pi_1(t+1) + \mathrm{Tr}\{\Sigma_1(t)Q_1(t)\} \\ & + \mathrm{Tr}\{\left(\Sigma_1(t+1|t) - \Sigma_1(t+1)\right)\ (K_{1A}-2K_{1B} + K_{1C}) \\ & + 2G_2H_2\Sigma_1(t+1|t)(K_{1B}-K_{1C}) + G_2\left(\Xi_2 + H_2\Sigma_1(t+1|t)H_2'\right)G_2'K_{1C}\}\Big\}\end{aligned} \tag{o}$$

$$\begin{aligned}\text{Since } \overline{x}_1(t+1) &= A(t)\hat{x}(t) + B(t)u(t) + C(t)\gamma_t^{20}(z_2(t)) \\ &= (I-C\Lambda)A\hat{x}_1 + (I-C\Lambda)Bu - C\Lambda A(\hat{x}_2-\hat{x}_1) \\ \overline{x}_2(t+1) &= (I-C\Lambda)A\hat{x}_1 + (I-C\Lambda)Bu + (I-C\Lambda)A(\hat{x}_2-\hat{x}_1)\end{aligned}$$

where the augment t was dropped for brevity. Substituting into (o), and differentiating, one finds $\gamma_t^{10}(z_1(t))$ to be

$$\gamma_t^{10}(z_1(t)) = -W^{-1}(t)Y(t)\hat{x}_1(t) - W^{-1}(t)M(t)\left(\hat{x}_2(t)-\hat{x}_1(t)\right) \qquad \text{(p)}$$

where

$$W(t) = R_1 + B'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)B + B'\Lambda'S_1\Lambda B$$

$$Y(t) = B'\Lambda'S_1\Lambda A + B'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A$$

$$M(t) = B'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A + B'(I-C\Lambda)'\Big(K_{1B}(t+1) - K_{1A}(t+1)\Big)A + B'\Lambda'S_1\Lambda A - B'(I-C\Lambda)'K_{1B}(t+1)G_2(t+1)H_2(t+1)A$$

So, from (b),

$$J_1^*(t) = \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t)-\hat{x}_1(t) \end{pmatrix}' \begin{pmatrix} K_{1A}(t) & K_{1B}(t) \\ K'_{1B}(t) & K_{1C}(t) \end{pmatrix} \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t)-\hat{x}_1(t) \end{pmatrix} + \Pi_1(t) \qquad \text{(q)}$$

verifying the induction assumption made in (b), where

$$K_{1A}(t) = Q_1 + A'\Lambda'S_1\Lambda A + A'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A - Y'W^{-1}Y$$

$$\begin{aligned} K_{1B}(t) = {} & A'\Lambda'S_1\Lambda A + A'(I-C\Lambda)'K_{1A}(t+1)(I-C\Lambda)A \\ & + A'(I-C\Lambda)'\Big(K_{1B}(t+1)-K_{1A}(t+1)\Big)A \\ & - A'(I-C\Lambda)'K_{1B}(t+1)G_2(t+1)H_2(t+1)A - Y'W^{-1}M \end{aligned}$$

$$\begin{aligned} K_{1C}(t) = {} & -M'W^{-1}M + A'\Lambda'\Big(S_1+C'K_{1A}(t+1)C\Big)\Lambda A - A'\Lambda'C'\Big(K_{1B}(t+1) \\ & - K_{1B}(t+1)G_2(t+1)H_2(t+1)\Big)A + A'\Big(K_{1B}(t+1)\ G_2(t+1)H_2(t+1) \\ & - K_{1B}(t+1)\Big)'\ C\ \Lambda\ A + A'\Big(I-G_2(t+1)H_2(t+1)\Big)'K_{1C}(t+1)\ \cdot \\ & \cdot\ \Big(I-G_2(t+1)H_2(t+1)\Big)A \end{aligned}$$

$$\begin{aligned} \Pi_1(t) = {} & \Pi_1(t+1) + \mathrm{Tr}\Big\{\Big(\Sigma_1(t+1|t) - \Sigma_1(t+1)\Big)\Big(K_{1A}(t+1) - 2K_{1B}(t+1) \\ & + K_{1C}(t+1)\Big)\Big\} + 2\ \mathrm{TR}\{G_2(t+1)H_2(t+1)\Sigma_1(t+1|t)\,\cdot \end{aligned}$$

$$\cdot\left(K_{1B}(t+1) - K_{1C}(t+1)\right)\} + Tr\{\Sigma_1(t)Q_1(t)\}$$

$$+ Tr\{G_2(t+1)\left(H_2(t+1)\Sigma_2(t+1|t)H_2'(t+1)\right.$$

$$\left.+ \Xi_2(t+1)\right)G_2'(t+1)K_{1C}(t+1)\} \qquad (r)$$

Similarly, the optimal cost-to-go for player 2, the follower, is given by:

$$J_2{}^* = E\{\hat{x}_2'\left(Q_2 + A'K_2(t+1)A\right)\hat{x}_2 + u^{*\prime}\left(R_2 + B'K_2(t+1)B\right)u^*$$

$$+ 2u^{*\prime}B'K_2(t+1)A\hat{x}_2 - (A\hat{x}_2 + Bu^*)'\Lambda'C'K_2(t+1)(A\hat{x}_2 + Bu^*)$$

$$+ b_2(t)|Z_2(t)\}$$

$$= \hat{x}_2'K_2(t)\hat{x}_2 + \Pi_2(t) \qquad (s)$$

This follows because $E\left(\hat{x}_1|z_2(t)\right) = \hat{x}_2$, because $Z_2(t) \subset Z_1(t)$. The optimal cost-to-go matrix is given by

$$K_2(t) = Q_2 + (A - BW^{-1}Y)'\left(K_2(t+1) - \Lambda'C'K_2(t+1)\right)(A - BW^{-1}Y)$$

$$+ Y'W^{-1}R_2W^{-1}Y \qquad (t)$$

and

$$\Pi_2(t) = \Pi_2(t+1) + Tr(\Sigma_2Q_2) + Tr\{\left(\Sigma_2(t+1|t) - \Sigma_2(t+1)\right)K_2(t+1)\}$$

$$+ Tr\{\left(\Sigma_2(t) - \Sigma_1(t)\right)(M-Y)'W^{-1}\left(R_2 + B'\left(K_2(t+1)\right.\right.$$

$$\left.\left.- K_2(t+1)C\Lambda\right)B\right)W^{-1}(M-Y)\}. \qquad (u)$$

BIBLIOGRAPHY

(Ao1) M. Aoki, Optimization of Stochastic Systems, Academic Press, New York, 1967.

(Ao2) M. Aoki, "On Decentralized Linear Stochastic Problems with Quadratic Cost," IEEE Transactions on Automatic Control, Vol. AC-18, No. 3, 1973.

(As1) K. J. Astrom, "Optimal Control of Markov Processes with Incomplete State Information," Journal of Mathematical Analysis and Applications, Vol. 10, 1965.

(Al) M. Athans and P. L. Falb, Optimal Control, McGraw-Hill, New York, 1966.

(Au1) R. J. Aumann, "Mixed and Behavior Strategies in Infinite Extensive Games," Advances in Game Theory, Princeton University Press, Princeton, New Jersey, 1964.

(Au2) R. J. Aumann and P. Maschler, "Some Thoughts on the Minimax Principle", Management Science, Vol. 18, 1972.

(Ba1) T. Basar, "On the Relative Leadership Property of Stackelberg Strategies", Journal of Optimization Theory and Applications, Vol. 11, No. 6, 1973.

(Ba2) T. Basar, "Equilibrium Solutions in Two-person Quadratic Decision Problems with Static Information Structures", IEEE Trans. on Auto. Control, Vol. AC-20, June 1975.

(Ba3) T. Basar, "A Counterexample in Linear Quadratic Games: Existence of Nonlinear Nash Solutions", JOTA, Vol. 14, October 1974.

(Ba4) T. Basar, "A New Class of Nash Strategies for M-Person Differential Games with Mixed Information Structures", presented at 6th Triennial World Congress of IFAC, August 1975, Cambridge, Massachusetts.

(Be1) R. D. Behn and Y. C. Ho, "On a Class of Linear Stochastic Differential Games," IEEE Trans. on Auto. Control, Vol. AC-13, No. 3, 1968.

(B1) R. Bellman, Dynamic Programming, Princeton University Press, Princeton, New Jersey,1957.

(Bℓ1) A. V. Balakrishnan, Stochastic Differential Systems I, Lecture Notes in Economics and Mathematician Systems, No. 84, Springer Verlag, New York, 1973.

(Br1) R. R. Bahadur and P. J. Bickel, "Substitution in Conditional Expectation",Annals of Mathematical Statistics, Vol. 39, No. 2, 1968.

(Br2) R. R. Bahadur, "Statistics and Subfields", Annals of Mathematical Statistics, Vol. 26, 1955.

(Ca1) J. H. Case, "Toward a Theory of Many Player Differential Games", SIAM Journal of Control, Vol. 7, No. 2, 1969.

(C1) D. Castanon and M. Athans, "On Stochastic Dynamic Stackelberg Strategies", MIT ESL Report ESL-P-561, June 1974.

(Ch1) C. I. Chen and J. B. Cruz, Jr., "Stackelberg Solution for Two-person Games with Biased Information Patterns", IEEE Trans. on Auto. Control, Vol. AC-17, No. 6, 1972.

(Co1) C. Y. Chong, On the Decentralized Control of Large-Scale Systems, Ph.D. Thesis, MIT, 1973.

(Cu1) K. C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems--Part II," IEEE Trans. on Auto. Control, Vol. AC-17, No. 1, 1972.

(D1) J. L. Doob, Stochastic Processes, John Wiley and Sons, New York, 1953.

(Fr1) A. Friedman, "Differential Games of Pursuit in Banach Space", Journal of Mathematical Analysis and Applications, Vol. 25, No. 1, 1969.

(Ha1) P. R. Halmos, Measure Theory, Van Nostrand, New York, 1950.

(H1) Y. C. Ho, "Toward Generalized Control Theory", IEEE Trans. on Auto. Control, Vol. AC-10, 1965.

(H2) Y. C. Ho and K. C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems--Part I," IEEE Trans. on Auto. Control, Vol. AC-17, No. 1, 1972.

(H3) Y. C. Ho and K. C. Chu, "Equivalence of Information Structure in Static and Dynamic Team Problems," IEEE Trans. on Auto. Control, Vol. AC-18, 1973.

(H4) Y. C. Ho, "On the Minimax Principle and Zero-Sum Stochastic Differential Games," JOTA, Vol. 13, No. 3, March 1974.

(H5) Y. C. Ho and K. C. Chu, "Information Structure in Dynamic Multi-person Games," Automatica, Vol. 10, 1974.

(I1) R. Isaacs, "Differential Games," I, II, III, IV, Rand Corporation RM-1391, Rm1399, Rm1411, Rm1486, 1954.

(I2) R. Isaacs, Differential Games, John Wiley, 1965.

(J1) D. Jacobson and D. Q. Mayne, Differential Dynamic Programming, American Elsevier Publications, New York, 1970.

(K1) H. W. Kuhn, "Extensive Games," Proceedings of the National Academy of Sciences, U.S.A., Vol. 36, 1950.

(K2) H. W. Kuhn, "Extensive Games and the Problem of Information", Contributions to the Theory of Games II, Princeton University Press, Princeton, New Jersey, 1953.

(Kul) H. J. Kushner, Introduction to Stochastic Control, Holt, Rinehart and Winston, New York, 1971.

(Lol) M. Loeve, Probability Theory, Van Nostrand, New York, 1963.

(L1) R. D. Luce and H. Raiffa, Games and Decisions, John Wiley, New York, 1951.

(M1) J. Marshak and R. Radner, The Economic Theory of Teams, Yale University Press, New Haven, Conn., 1971.

(Mel) M. D. Mesorovic, D. Macko and Y. Takahara, Theory of Hierarchical Multilevel Systems, Academic Press, New York, 1970.

(N1) J. F. Nash, "Equilibrium Points in N-Person Games", Proceedings of the National Academy of Sciences, USA, Vol. 36, 1950.

(N2) J. F. Nash, "Noncooperative Games", Annals of Mathematics, No. 54, No. 2, 1951.

(R1) R. Radner, "Team Decision Problems", Annals of Mathematical Statistics, Vol. 33, 1962.

(Ra1) H. Raiffa and R. Schlaifer, Applied Statistical Decision Theory, Harvard University Press, Cambridge, Mass., 1961.

(Rh1) I. B. Rhodes and D. G. Luenberger, "Differential Games with Imperfect State Information", IEEE Trans. on Auto. Control, Vol. AC-14, No. 1, 1969.

(Rh2) I. B. Rhodes and D. G. Luenberger, "Nondeterministic Differential Games with Constrained State Estimators", IEEE Trans. on Auto. Control, Vol. AC-14, No. 1, 1969.

(Sa1) N. Sandell, Jr., and M. Athans, "Solutions of Some Non-Classical LQG Stochastic Decision Problems," IEEE Transactions of Auto. Control, Vol. AC-19, No. 2, 1974.

(Sa2) N. Sandell, Jr., "On Open Loop and Closed Loop Nash Strategies", IEEE Trans. on Auto. Control, Vol. AC-19, June, 1974.

(Sa3) N. Sandell, Jr., "Control of Finite-State, Finite-Memory Stochastic Systems", Ph.D. dissertation, Electrical Engineering, M.I.T., 1974.

(S1) M. Simaan and J. B. Cruz, Jr., "On the Stackelberg Strategy in Nonzero Sum Games", JOTA, Vol. 11, No. 6, June 1973.

(S2) M. Simaan and J. B. Cruz, Jr., "Additional Aspects of the Stackelberg Strategy in Nonzero Sum Games," JOTA, Vol. 11. No. 6, June 1973.

(S3) M. Simaan and J. B. Cruz, Jr., "Sampled Data Nash Controls in Nonzero Sum Differential Games", International Journal of Control, Vol. 17, No. 6, 1973.

(S4) M. Simaan and T. Takayama, "Dynamic Duopoly Game: Differential Game Theoretic Approach," presented at 6th triennial IFAC Congress, Cambridge, Mass., August 1975.

(S5) M. Simaan, "Stackelberg vs. Nash Strategies in Nonzero Sum Games," Proceedings of 8th Princeton Conference on Systems and Information Sciences, March 1974.

(Sh1) M. Shubik, Game Theory - Related Approaches to Social Behavior, John Wiley, New York, 1964.

(St1) A. W. Starr and Y. C. Ho, "Nonzero Sum Differential Games," JOTA, Vol. 3, No. 3, 1969.

(St2) A. W. Starr and Y. C. Ho, "Further Properties of Nonzero Sum Differential Games," JOTA, Vol. 3, No. 4, 1971.

(Str1) C. Striebel, Optimal Control of Discrete Time Stochastic Systems, Springer Verlag, Berlin, 1975.

(T1) G. L. Thompson, "Signaling Strategies in N-Person Games", Contributions to the Theory of Games II, Princeton University Press, Princeton, New Jersey, 1953.

(V1) P. P. Varaiya, "On the Existence of Solutions to a Differential Game", Siam Journal of Control, Vol. 5, 1967.

(Vn1) J. Von Neumann and O. Morgernstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, New Jersey, 1947.

(Vs1) H. Von Stackelberg, The Theory of the Market Economy, Oxford University Press, Oxford, England, 1952.

(W1) W. W. Willman, "Formal Solutions for a Class of Stochastic Pursuit Evasion Games," IEEE Trans. on Auto. Control, Vol. AC-14, 1969.

(Wt1) H. S. Witsenhausen, "A Counterexample in Stochastic Optimal Control", Siam Journal of Control, Vol. 6, 1968.

(Wt2) H. S. Witsenhausen, "On Information Structures, Feedback and Causality", Siam Journal on Control, Vol. 9, 1971.

(Wt3) H. S. Witsenhausen, "Separation of Estimation and Control for Discrete Time Systems," Proc. IEEE, Vol. 59, No. 11, 1971.

(Wt4) H. S. Witsenhausen, "A Standard Form for Sequential Stochastic Control," Mathematical Systems Theory, Vol. 7, 1973.

(Wt5) H. S. Witsenhausen, "On the Relations Between the Values of a Game and its Information Structure", Information and Control, Vol. 19, No. 3, October 1971.

BIOGRAPHICAL NOTE

David A. Castañón was born in La Habana, Cuba, on November 24, 1950. He attended high school in Humacao, Puerto Rico, graduating in 1967. He received the BSE degree with honors from Tulane University in New Orleans, Louisiana in 1971.

Mr. Castañón has been a teaching assistant at M.I.T. in the Applied Mathematics group (1971-1975). He was a summer instructor in Mathematics (1974, 1975). The past year he has been associated with the Electronic Systems Laboratory at M.I.T., working in adaptive control of aircraft under NASA grant NSG-1018.

Mr. Castañón has published several journal articles. He is a member of Phi Eta Sigma and Tau Beta Pi.