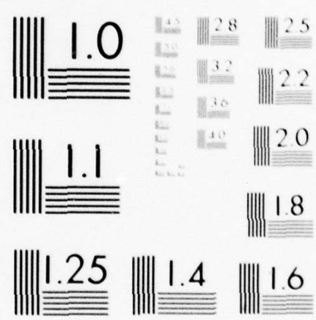
AD-A072 383 CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING A--ETC F/G 1/2
RANDOMIZED AND STOCHASTIC LINEAR, DISCRETE DIFFERENTIAL GAMES W--ETC(U)
JUL 78 K B BLEY, C T LEONDES F33615-77-C-3013

UNCLASSIFIED AFFDL-TR-78-87 NL

1 OF 2
AD
A072383



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963-A

AFFDL-TR-78-87

RANDOMIZED AND STOCHASTIC LINEAR, DISCRETE DIFFERENTIAL GAMES WITH QUADRATIC PAYOFF FUNCTIONS

UNIVERSITY OF CALIFORNIA, LOS ANGELES
SCHOOL OF ENGINEERING AND APPLIED SCIENCE
7620 BOELTER HALL, UCLA
LOS ANGELES, CALIFORNIA 90024

July 1978

TECHNICAL REPORT AFFDL-TR-78-87
Final Report for Period 1977-1978

Approved for public release; distribution unlimited.

AIR FORCE FLIGHT DYNAMICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

This technical report has been reviewed and is approved for publication.

Gerald W. Francis

JAMES W. MORRIS, Chief
Control Systems Development Branch
Flight Control Division

FOR THE COMMANDER

MORRIS A. OSTGAARD
Assistant for Research and Technology
Flight Control Division

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFFDL-TR-78-87

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): RANDOMIZED AND STOCHASTIC LINEAR, DISCRETE DIFFERENTIAL GAMES WITH QUADRATIC PAYOFF FUNCTIONS

5. TYPE OF REPORT & PERIOD COVERED: 5/77 5/78

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): K. B. Bley (C. T. Leondes, Principal Investigator)

8. CONTRACT OR GRANT NUMBER(s): F33 615-77-C-3013

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
School of Engineering and Applied Science
University of California, Los Angeles
Los Angeles, CA 90024

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2307-03-02

11. CONTROLLING OFFICE NAME AND ADDRESS:
U.S. Air Force Systems Command
AFFDL/FGL
WPAFB, OH 45433

12. REPORT DATE: July 1978

13. NUMBER OF PAGES: 124

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
Office of Naval Research
1030 E. Green St.
Pasadena, CA 91106
Attention: Mr. Perry Beilke

128p

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report):

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

Final rept. May 77 - May 78

18. SUPPLEMENTARY NOTES:

61102F

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Differential Games, Stochastic Systems, Flight Control, Optimal Control

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

This report deals with a limited, though still extensive, class of differential games. The system dynamics are given by linear difference equations which may have time varying, but deterministic, coefficients. In some of the work, an independent noise sequence is also allowed which then acts as a random forcing function. The system is allowed to change in response to each player's controls for a specified length of time -- a specified number of stages. The payoff is a quadratic function of the state and the two control variables at each stage.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

404 637

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Chapter 2 deals with multistage differential games (another name for differential games involving discrete time) having perfect information. The solution for the case of pure strategies and unbounded controls is derived and shown to have simple closed forms. Next, the solution to the case of pure strategies, but bounded controls, is developed. Finally, the case where the controls are bounded and pure strategies do not exist is investigated. It is shown that, under certain circumstances, randomized strategies are optimal and the requisite strategy is derived. An example illustrates a simple case. Chapter 3 deals with multistage stochastic differential games with pure strategies. In this case, each player has to determine his optimal control strategies based on noisy observations of the state, and so a game of imperfect information results. A combination of dynamic programming, variational arguments, and functional analysis leads to the game optimal control strategies based on each player's observations and his knowledge of the parameters of the system and the mean and variances of the noises in the system. The optimal control strategies are shown to be linear functions of past and present observations available to each player. Two examples are included. Chapter 4 is composed of several singlestage, scalar problems using the theory developed in Chapter 3. Varying amounts of information are made available to the players, and the corresponding optimal strategies are derived in closed form. It is shown that there is no separation of estimation and control, as in optimal control problems, when both players have noisy observations of the state. Chapter 5 developed the solution to the multistage stochastic differential game with pure stategies where the form of the solution is specified in advance. A two-state scalar example illustrates the fact that the specification of the form reduces the optimization problem to simple calculus. Chapter 6 is concerned with future work that appears interesting in the general area of the work covered by this report. It is felt that most of the work presented in the latter half of Chapter 2 and all the work in Chapters 3 through 5 represent new results. The major contribution of this research is in the application of relatively well-known mathematical results in the theory of games and in functional analysis to differential games in a mannter not heretofore performed.

PREFACE

In order to synthesize and effectively treat and control Air Force air-to-air and air-to-ground flight vehicle systems, it is absolutely necessary to have the groundwork contributed to in large part by the rather important results in this report whose development was motivated by these important Air Force flight dynamic vehicle issues.

TABLE OF CONTENTS

PRECEDING PAGE BLANK

TABLE OF CONTENTS (continued)

SECTION I

INTRODUCTION

1. HISTORY OF DIFFERENTIAL GAMES

Differential games came into existence in 1954 with the publication of several RAND Corporation reports written by Rufus Isaacs.[1-4] This work, and a great deal more, was incorporated into his book Differential Games[5] published in 1965.

The original impetus for his investigations was the problem of pursuit and evasion of hostile aircraft. Following his original work, the field of differential games became almost entirely the province of mathematicians who made no attempt to give any physical meaning to their results.[6-8] At the same time that American mathematicians were delving into the field, Soviet mathematicians also were exploring it (see the bibliography of Reference 9).

Differential games became better known to engineers in the middle 1960s with publication of Differential Games and its review,[10] and a small flurry of papers in engineering journals.[11-15] Following these came another small flurry of papers involving stochastic differential games.[16-18] Today the subject represents a rich field of investigation to those interested in control theory and the quantifiable aspects of conflict.

This exceedingly short summary is by no means exhaustive of the work that has been done.

2. THE THEORY OF GAMES

As the name differential games implies, it is a derivative of the mathematical theory of games first developed by von Neumann and Morgenstern.[19] Many good books on game theory (far easier to read and understand than is Theory of Games and Economic Behavior) are available.[20-23]

Games can have many forms depending on the number of players and the way in which winnings and losses are computed. The work herein involves only two players where, essentially, the loser pays the winner a specified amount after the play of a given game. Since the algebraic sum of the winner's game (positive) and loser's loss (negative) is zero, this type of game is known as a two-player, zero sum game.

Games may be presented either in extensive form – a set of rules and a succession of choices for each player, or in normal form – a matrix or function which relates the amount due to the winner to the choices made by the two players. The amount, as a function of the choices, is known as the payoff of the game.

The choices may involve only a limited number of individual elements – matrix games – or they may involve an infinite number of elements – infinite or continuous games. Each player wishes to find a strategy that allows him to make his choices (choose his control) so as to optimize the payoff. In general, these strategies are functions of the payoff. These strategies may be: 1) pure – given a fully specified payoff function, there is a single value of the control which should be chosen; or 2) randomized (mixed) — given a fully specified payoff function, there is a probability density function which should be chosen with the actual control found by a chance device having an output governed by the optimal probability density.

It can be shown that every matrix game, and some types of continuous games, admits of a pair of strategies, pure or mixed, such that, if the minimizing player uses his optimal strategy, the payoff to the maximizing player will be no greater than a certain amount. Conversely, if the maximizing player uses his optimal strategy, he is assured of receiving at least the same amount. This amount is known as the Value of the game. The condition that expresses the inequalities of the payoff is known as the saddle point condition. Control strategies that satisfy the saddle point condition are such that both players can compute them, or, equivalently, that both players may announce their strategies and still be assured that the payoff will be no worse, from each point of view, than the Value.

An important concept in game theory is information. In games of perfect information, each player knows the exact value of the payoff and all that has occurred in the past. Such games as chess and checkers are examples of games of perfect information. It can be shown that such games have saddle points for pure strategies for both sides. Games of imperfect information have some elements that are unknown or given by a probability distribution; bridge is such a game. Games of imperfect information may or may not have pure strategies that satisfy the saddle point condition.

3. DIFFERENTIAL GAMES

Differential games involve a payoff which is in some way related to a dynamical system. The two players attempt to optimize this payoff by choosing game optimal control strategies. The evolution of the state, as a function of the players' control variables, is described by differential equations (continuous time) or difference equations (discrete time); thus the name.

Because of the interaction among the state of the system, the dynamical relationship between the state and the controls, and the payoff, it is necessary that the game optimal control strategies be feedback strategies, ones that use information on the current state. Thus the central problem for the engineer is to find these feedback strategies. (Mathematicians must still deal with the many unsolved problems concerning the existence, uniqueness, and optimality of solutions.)

In general, there is no reason to assume that a given differential game has a Value or that its Value, when it exists, is obtained by using pure strategies. There are problems arising from the fact that the payoff is generally a continuous functional of the controls and, as noted in Section II, not all continuous payoffs admit of a Value, or, if they do, a Value resulting from pure strategies. Happily, there are cases where the structure of the payoff and the dynamical equations do result in a Value obtained by using pure strategies.[24]

Differential games can be divided into classes: those where observations of the state are perfect, and those where they are not. The latter are referred to as stochastic games. Up to the present, they have involved linear observations of the state corrupted by additive noise. The addition of such noise means that a deterministic payoff function is no longer meaningful; instead, it is replaced with an expected value which is then optimized.

It must be noted that the descriptions of the theory of games and of differential games are included mainly for orientation purposes; they are neither rigorous nor complete. Full expositions are to be found among the references already cited.

SECTION II

MULTISTAGE DIFFERENTIAL GAMES WITH PERFECT INFORMATION

1. INTRODUCTION

Many papers published in the last 15 years have dealt with differential games involving perfect information. Almost without exception, they have dealt with problems leading to pure strategies. Some of the exceptions are noted in References 25 through 27; in addition, Chapter 12 of Reference 5 discusses the subject.

In a sense, this is a very surprising turn of events since the whole theory of differential games is based on the theory of games — a discipline which is more concerned with randomized strategies than with pure ones. Further, until quite recently, the limited work done in the field of differential games had been performed by mathematicians (as opposed to engineers) who might have been thought to be more interested in questions of randomized strategies.

While engineers involved in doing research are quick to use some of the better known mathematical results in differential games involving pure strategies,[28] there seems to be no such inclination with regard to the theory of sequentially compounded two-person games.[21,23] Thus the available work in stochastic and recursive games has not received its due attention.

The work in this chapter is directed toward the complete solution of a multistage linear differential game with a quadratic payoff function. For convenience, only the scalar case is considered, but, at the cost of more work, the results go over identically when the dynamics are given in terms of matrix equations.

The solution is complete in that both deterministic and randomized control strategies, as required by the relative value of system parameters, are derived. To solve this problem, it is assumed that both players know the system parameters and are able to observe the true state of the system at all times.

2. DERIVATION OF PURE CONTROL STRATEGIES WHEN CONTROL MAGNITUDES ARE UNBOUNDED

The evolution of the state of the system is determined by a linear difference equation

$$z_{i-1} = k_i z_i + a_i u_i + b_i v_i \tag{1}$$

where

$$z_i = \text{state at the } i^{th} \text{ stage} \tag{2}$$

$$u_i = \text{minimizing player's control at the } i^{th} \text{ stage} \tag{3}$$

$$v_i = \text{maximizing player's control at the } i^{th} \text{ stage} \tag{4}$$

Both u_i and v_i can be functions of any or all of the past and present values of the state. (In game theory, the maximizing player is generally referred to as player I and the minimizing player as player II; these designations also are occasionally used here.)

The subscript indicates the stage number. Equation (1) is written in terms of time (stages)-to-go rather than the usual time measured from the initiation of the differential game. This is done because the stage number represents the number of times each player must choose a control — the actual value for u_i or v_i. Thus an

N stage game, which begins in state z_N and terminates in state z_o, requires N choices of each player's control.

The object of the differential game is to optimize, in a game sense, a quadratic payoff function

$$I_N(u, v) = \sum_{i=1}^{N} c_i z_{i-1}^2 + d_i u_i^2 - e_i v_i^2 \qquad (5)$$

That is, the minimizing player wishes to choose $u_N, \cdots, u_1$ so as to minimize I_N, while the maximizing player wishes to pick $v_N, \cdots, v_1$ so as to maximize it. Because there is only a single payoff functional, the problem falls within the general purview of zero sum game theory.

The Value of the game is given by

$$J_N = \underset{U,\, V}{\text{val}}\, J_N(u, v) = \underset{U,\, V}{\text{val}} \sum_{i=1}^{N} c_i z_{i-1}^2 + d_i u_i^2 - e_i v_i^2 \qquad (6)$$

where U and V represent the set of 2N controls $u_N, \cdots, u_1, v_N, \cdots, v_1$. The parameters of the system, a_i, b_i, c_i, d_i, e_i, and k_i, are all assumed to be real; a_i, b_i, d_i, e_i, and c_1 are assumed to be positive; $c_i (i \neq 1)$ is assumed to be nonnegative; and k_i can be positive, negative, or zero. These restrictions are required to produce meaningful results. If a_i or b_i is zero, the associated control can have no effect on the state at the next stage. Similarly, if d_i or e_i is zero, there is no penalty associated with the use of a control, and, if otherwise called for, infinite control magnitudes could be used. If c_1 were zero, the effect would be to terminate the game a stage early since neither player would wish to incur a penalty in the payoff by using his control when there was no attendant change due to a change in state.

On the other hand, if $c_i = 0 (i \neq 1)$, the game is still well defined as to the number of stages, and the control at that stage is not, in general, zero for either player. (This is the analog to terminal control problems in optimal control.) The controls u_i and v_i are also assumed to be real with no restrictions on their sign or magnitude.

The principle of optimality of the theory of dynamic programming,[29] along with simple variational principles, is used to find the desired control strategies.[30] (Control strategies are defined to be rules that tell each player the value he should choose for his control at each stage based on the information available to him. The control is defined to be the value actually chosen.)

As is usual in dynamic programming, the single-stage game is first solved. To do this, assume that pure strategies exist for both players, strategies that permit each player to use the actual value of the state. Denoting these game optimal strategies by overbars, it follows that the Value of the single-stage game, J_1, is achieved when

$$u_1 = \bar{u}_1(z_1) \tag{7}$$

$$v_1 = \bar{v}_1(z_1) \tag{8}$$

The saddle point condition states that, if one player uses his game optimal strategy and the other does not, then the payoff is as good or better than that which would have been achieved if both players had used their game optimal strategies. In terms of the payoff functional, (5),

$$I_1(\tilde{u}_1, \bar{v}_1) \geq I_1(\bar{u}_1, \bar{v}_1) = J_1 \geq I_1(\bar{u}_1, \tilde{v}_1)$$

where $\tilde{u}_1$ and $\tilde{v}_1$ are any real functions of z_1. Putting it another way,

$$I_1(\tilde{u}_1, \bar{v}_1) - I_1(\bar{u}_1, \bar{v}_1) \geq 0 \tag{10}$$

$$I_1(\bar{u}_1, \tilde{v}_1) - I_1(\bar{u}_1, \bar{v}_1) \leq 0 \tag{11}$$

If

$$\tilde{u}_1 = \bar{u}_1 + \epsilon\delta(z_1) \tag{12}$$

where ϵ is a small number and $\delta(z_1)$ is any real function of z_1, then (1), (5), (10), and (12) lead to

$$\begin{aligned} I_1(\bar{u}_1 + \epsilon\delta_1 \bar{v}_1) - I_1(\bar{u}_1, \bar{v}_1) &= c_1(k_1 z_1 + a_1\bar{u}_1 + a_1\epsilon\delta + b_1\bar{v}_1)^2 \\ &\quad + d_1(\bar{u}_1 + \epsilon\delta)^2 - e_1\bar{v}_1^2 \\ &\quad - c_1(k_1 z_1 + a_1\bar{u}_1 + b_1\bar{v}_1)^2 \\ &\quad - d_1\bar{u}_1^2 + e_1\bar{v}_1^2 \\ &= 2\left[a_1 c_1 k_1 z_1 + \left(a_1^2 c_1 + d_1\right)\bar{u}_1 \right. \\ &\quad \left. + a_1 b_1 c_1 \bar{v}_1\right]\epsilon\delta + \left(a_1^2 c_1 + d_1\right)\epsilon^2\delta^2 \geq 0 \end{aligned} \tag{13}$$

where the argument of δ has been dropped for brevity. Standard variational arguments lead to the following necessary conditions for (13) to hold

$$a_1 c_1 k_1 z_1 + (a_1^2 c_1 + d_1)\bar{u}_1 + a_1 b_1 c_1 \bar{v}_1 = 0 \tag{14}$$

$$a_1^2 c_1 + d_1 \geq 0 \tag{15}$$

A similar approach leads to

$$I_1(\bar{u}_1, \bar{v}_1 + \epsilon\Delta) - I_1(\bar{u}_1, \bar{v}_1) = 2\Big[b_1 c_1 k_1 z_1 + a_1 b_1 c_1 \bar{u}_1$$

$$- (e_1 - b_1^2 c_1)\bar{v}_1\Big]\epsilon\Delta - (e_1 - b_1^2 c_1)\epsilon^2 \Delta^2 \leq 0 \tag{16}$$

so that necessary conditions for (16) to hold are

$$b_1 c_1 k_1 z_1 + a_1 b_1 c_1 \bar{u}_1 - \left(e_1 - b_1^2 c_1\right)\bar{v}_1 = 0 \tag{17}$$

$$e_1 - b_1^2 c_1 \geq 0 \tag{18}$$

The simultaneous solution of (14) and (17) results in the game optimal control strategies for the single-stage game

$$\bar{u}_1 = - \frac{a_1 c_1 e_1 k_1}{\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1} \, z_1 \tag{19}$$

$$\bar{v}_1 = \frac{b_1 c_1 d_1 k_1}{\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 e_1^2} \, z_1 \tag{20}$$

Substitution of (1), (19), and (20) into (6) yields

$$J_1 = \frac{c_1 d_1 e_1 k_1^2}{\left(a^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1^2} \, z_1^2 \tag{21}$$

since $\bar{u}_1$ and $\bar{v}_1$ are precisely the pure control strategies which optimize, in a game theory sense, the payoff functional, (5).

Rewriting the denominator of (21) as

$$\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1^2 = c_1\left(a_1^2 e_1 - b_1^2 d_1\right) + d_1 e_1 \tag{22}$$

indicates that the a sufficient condition for pure strategies to exist is

$$a_1^2 e_1 - b_1^2 d_1 \geq 0 \tag{23}$$

(It should be clear that (21) will always be nonnegative since the maximizing player can always choose his control to be zero if it becomes apparent that other values would lead to a negative payoff.) When the equality in (23) holds, it is easy to show, using (19), (20), and (1), that

$$z_o = k_1 z_1 \tag{24}$$

so that the change of state is not affected when both players use their game optimal strategies. This might be called a case of equal efficiency.

Having found the game optimal pure strategies for a single-stage game, it is now possible to solve the two-stage game. Applying the principle of optimality yields

$$J_2 = \underset{u_2, v_2}{\text{val}} \left\{ c_2 z_1^2 + d_2 u_2^2 - e_2 v_2^2 + J_1 \right\} \tag{25}$$

Substituting (21) into (25) leads to

$$J_2 = \underset{u_2, v_2}{\text{val}} \left\{ \tilde{c}_2 z_1^2 + d_2 u_2^2 - e_2 v_2^2 \right\} \tag{26}$$

where

$$\tilde{c}_2 = c_2 + \frac{c_1 e_1 d_1 k_1^2}{\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1^2} \tag{27}$$

The use of (27) has changed the two-stage problem into a single-stage problem so that all that has gone before applies now with the subscript 2 replacing 1 and $\tilde{c}_2$ instead of c_1. Thus

$$\bar{u}_2 = - \frac{a_2 \tilde{c}_2 e_2 k_2}{\left(a_2^2 \tilde{c}_2 + d_2\right)\left(e_2 - b_2^2 \tilde{c}_2\right) + a_2^2 b_2^2 \tilde{c}_2^2} z_2 \tag{28}$$

$$\bar{v}_2 = \frac{b_2 \tilde{c}_2 d_2 k_2}{\left(a_2^2 \tilde{c}_2 + d_2\right)\left(e_2 - b_2^2 \tilde{c}_2\right) + a_2^2 b_2^2 \tilde{c}_2^2} z_2 \tag{29}$$

$$J_2 = \frac{\tilde{c}_2 d_2 e_2 k_2^2}{\left(a_2^2 \tilde{c}_2 + d_2\right)\left(e_2 - b_2^2 \tilde{c}_2\right) + a_2^2 b_2^2 \tilde{c}_2^2} z_2^2 \tag{30}$$

with sufficient conditions being

$$a_2^2 \tilde{c}_2 + d_2 > 0 \tag{31}$$

$$e_2 - b_2^2 \tilde{c}_2 > 0 \tag{32}$$

(The other necessary conditions involving the first variation are inherent in (28) and (29).)

The general solution for any number of stages is quite clear. For $i = 1, \ldots, N$

$$J_i = A_i z_i^2 \tag{33}$$

$$A_i = \frac{\tilde{c}_i d_i e_i k_i^2}{\left(a_i^2 \tilde{c}_i + d_i\right)\left(e_i - b_i^2 \tilde{c}_i\right) + a_i^2 b_i^2 \tilde{c}_i^2} \tag{34}$$

$$\tilde{c}_i = c_i + A_{i-1} \; ; \; A_o = 0 \tag{35}$$

$$\bar{u}_i = \frac{a_i}{d_i k_i} A_i z_i = - \frac{a_i}{2 d_i k_i} \frac{\partial J_i}{\partial z_i} \tag{36}$$

$$\bar{v}_i = \frac{b_i}{e_i k_i} A_i z_i = \frac{b_i}{2 e_i k_i} \frac{\partial J_i}{\partial z_i} \tag{37}$$

(The fact that the game optimal controls are specifically related to the Value of the game is no coincidence; the same behavior is exhibited in optimal control problems involving linear dynamics and quadratic cost functionals.)[31]

These results are, of course, not new. A slightly different approach was used in Reference 32. In the case of continuous, instead of discrete, dynamics, the analogous result was obtained in Reference 12 using straightforward variational arguments, and, in Reference 33, functional analysis techniques were used.

3. RANDOMIZED CONTROL STRATEGIES

Prior to this section and in the following chapters, only the pure strategy aspects of game theory are used to derive control strategies. In this section, a wider (but still very limited) appeal is made to other aspects. In particular, the following derivation is based on the theory of infinite convex games.[20, 34]

Roughly speaking, an infinite convex game is one where each player is free to choose his control from a region of an appropriately (finite) dimensioned, Euclidean space. The game is called infinite because of the infinite number of possible choices of a control within the region (as opposed to a finite number of possible choices in a matrix game). The game is called convex if the scalar payoff function is convex in the minimizing player's control variable.

It has been shown that, when the payoff is both convex in the minimizing control variable and continuous in both control variables, the game has a Value. Further, if the minimizing control must take its value from a compact and convex region of the control space, then the minimizing player has a pure strategy. Finally, if the minimizing player must choose his control from an arbitrary n dimensional region, then the optimal control strategy of the maximizing player will require randomization over, at most, n+1 points.

The theory of infinite convex games also describes how to find solutions (control strategies) for each player. The application of this theory to the scalar case is carried out in the following work as an illustration. The pertinent theorems, suitably paraphrased in terms of notation, are taken from Chapter 12 of Reference 20 and given, without proof, in the Appendix.

Both payoff, (5), and the dynamics of the system, (1), remain unchanged. In this section, however, it is assumed that both u_i and v_i are limited, in absolute value, to be less than or equal to one. That is, both u_i and v_i belong to sets U and V such that

$$U = \{ u_i : |u_i| \le 1 \} \tag{38}$$

$$V = \{ v_i : |v_i| \le 1 \} \tag{39}$$

(Previously, it was assumed that u_i and v_i could take on any value.)

As before, the principle of optimality is used to derive the control strategies for all N stages. Before, at each stage, the Value was found from the (implicit) fact that

$$J_i = I_i(\bar{u}_i, \bar{v}_i) = \underset{u_i, v_i}{\text{val}} I_i(u_i, v_i) = \underset{u_i}{\min} \underset{v_i}{\max} I_i(u_i, v_i) = \underset{v_i}{\max} \underset{u_i}{\min} I_i(u_i, v_i) \tag{40}$$

In general, that is, when pure strategies do not exist for both players, this is not true. It is true, however, that for the game under consideration,

$$J_i = \underset{u_i}{\min} \underset{v_i}{\max} I_i(u_i, v_i) \tag{41}$$

as given in the theorem in the Appendix. Note that the order of the min-max operations is strictly a result of the convexity of the payoff in u_i which implies a pure strategy for player II. Having found J_i, $\bar{u}_i$ is found from the solution of

$$J_i = \underset{v_i}{\max} I_i(\bar{u}_i, v_i) \tag{42}$$

Equations (41) and (42) provide the tools needed to solve the problem at hand.

As usual, the single-stage game is solved first. Since

$$\frac{\partial^2 J_1}{\partial u_1^2} = 2\left(a_1^2 c_1 + d_1\right) \tag{43}$$

$$\frac{\partial^2 J_1}{\partial v_1^2} = -2\left(e_1 - b_1^2 c_1\right) \tag{44}$$

A pure strategy u_i for player II exists whenever

$$a_1^2 c_1 + d_1 > 0 \tag{45}$$

Player I's strategy then calls for randomization over, at most, two points (randomization over a single point is the same thing as a pure strategy). Completely analogous results are obtained for a payoff that is concave in the maximizing player's control variable, i.e., when

$$e_1 - b_1^2 c_1 > 0 \tag{46}$$

A pure strategy exists for player I, and player II's optimal strategy requires randomization over, at most, two points.

The rest of this chapter is concerned with the case where the payoff is strictly convex in u_i and where it may or may not be concave in v_i. (The last section dealt with a payoff which was convex in u_i – a pure strategy for player II – and which was concave in v_i – a pure strategy for player I.)

Making use of (1) and (5), it follows that

$$\begin{aligned} I_1(u_1, v_1) &= c_1(k_1 z_1 + a_1 u_1 + b_1 v_1)^2 + d_1 u_1^2 - e_1 v_1^2 \qquad (47) \\ &= A + Bv + Cv^2 \\ &= C\left(v_1 + \frac{B}{2C}\right)^2 + A - \frac{B^2}{4C} \end{aligned}$$

where

$$A = c_1(k_1 z_1 + a_1 u_1)^2 + d_1 u_1^2 \tag{48}$$

$$B = 2b_1 c_1 (k_1 z_1 + a_1 u_1) \tag{49}$$

$$C = c_1 b_1^2 - e_1 \tag{50}$$

The payoff, as a function of v_1, can take on either the shape of a straight line ($c = 0$) or a parabola ($c \neq 0$); if $c > 0$, then the parabola opens upward; if $c < 0$, it opens downward. If it opens downward, then the payoff is concave in v_i, and so a pure strategy exists for player I.

The parabola is symmetric about the line

$$v_1 = -\frac{B}{2c} = -\frac{b_1 c_1 (k_1 z_1 + a_1 u_1)}{b_1^2 c_1 - e_1} \tag{51}$$

and, of course, the maximum ($c < 0$) or minimum ($c > 0$) value of $I_1(u_1, v_1)$ occurs at the same value of v_1. It is important to know where the maximum or minimum value occurs as a function of u_1 and z_1.

4. DERIVATION OF PURE CONTROL STRATEGIES WHERE CONTROL MAGNITUDES ARE BOUNDED

It is instructive to see what happens to pure strategies when the magnitudes of u_i and v_i are constrained to belong to U and V, respectively, instead of taking on any value. The major result is, naturally, to complicate the form of the solution since the control strategies are no longer linear functions of the state. The theorems given in the Appendix are used to find the solution.

It has already been noted that pure strategies exist for both players whenever (45) and (46) hold (for this section the same assumption holds). However, because of the constraints on the magnitude of u_1 and v_1, it is not possible to simply apply the techniques of Section II to this problem. Instead, the optimal strategies for the single-stage game are found by:

1) Finding $v_1(u_1)$ such that $I_1(u_1, v_1(u_1)) = \max_{v_1} I_1(u_1, v_1)$

2) Finding $\bar{u}_1$ such that $\min_{u_1} I_1(u_1, v_1(u_1)) = J_1$

(The technique is the same whether the control strategies are pure or randomized.) Both $\bar{u}_1$ and $\bar{v}_1$, the optimal strategies, are functions of the state z_1.

From (51) it follows that v_1 is equal to minus one whenever

$$- \frac{b_1 c_1 (k_1 z_1 + a_1 u_1)}{b_1^2 c_1 - e_1} \leq -1 \tag{52}$$

or whenever

$$u_1 \leq \frac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \frac{k_1 z_1}{a_1} \tag{53}$$

Also v_1 is equal to plus one whenever

$$- \frac{b_1 c_1 (k_1 z_1 + a_1 u_1)}{b_1^2 c_1 - e_1} \geq 1 \tag{54}$$

which occurs when

$$u_1 \geq -\frac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \frac{k_1 z_1}{a_1} \tag{55}$$

When neither (53) nor (55) hold, v_1 is given by (51) directly. Thus $v_1(u_1)$ is given by

$$v_1(u_1) = \begin{cases} -1 & ; \ -1 \leq u_1 \leq \dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \\[2ex] -\dfrac{b_1 c_1 (k_1 z_1 + a_1 u_1)}{b_1^2 c_1 - e_1} & ; \ \dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} < u_1 < -\dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \\[2ex] +1 & ; \ -\dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \leq u_1 \leq 1 \end{cases} \tag{56}$$

It may happen that some of the sets of inequalities are incompatible in (56). For instance, it could be that

$$\frac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \frac{k_1 z_1}{a_1} \leq -1 \tag{57}$$

As an example, this could occur for $k_1 z_1$ sufficiently large and positive. The meaning of such an incompatibility is merely that such a value for $v_1(u_1)$ is never an optimal choice, since either u_1, from the middle inequality of (56), or +1 would be chosen. From a (somewhat) practical point of view, replace any expression in the inequalities of (56) whose absolute magnitude is greater than one by one times the algebraic sign of the expression. Disregard any values of $v_1(u_1)$ where the right side of an expression minus the left side (where no expression is larger than one in absolute magnitude) is zero.

Defining

$$D = \begin{cases} -1 & ; \quad \dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \le -1 \\ \dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} & ; \quad -1 < \dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} < 1 \\ +1 & ; \quad 1 \le \dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \end{cases} \tag{58}$$

$$E = \begin{cases} -1 & ; \quad -\dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \le -1 \\ -\dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} & ; \quad -1 < -\dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} < 1 \\ +1 & ; \quad 1 \le -\dfrac{b_1^2 c_1 - e_1}{a_1 b_1 c_1} - \dfrac{k_1 z_1}{a_1} \end{cases} \tag{59}$$

The Value of the single-stage game is given by

$$J_1 = \min \left\{ \min_{-1 \le u_1 \le D} c_1(k_1 z_1 + a_1 u_1 - b_1)^2 + d_1 u_1^2 - e_1 \; ; \right.$$

$$\left. \min_{D < u_1 < E} - \frac{c_1 e_1 (k_1 z_1 + a_1 u_1)^2}{b_1^2 c_1 - e_1} + d_1 u_1^2 \; ; \min_{E \le 1} c_1(k_1 z_1 + a_1 u_1 + b_1)^2 + d_1 u_1^2 - e_1 \right\} \tag{60}$$

Equation (60) is precisely the implementation of (41). For any value of the state, z_1, it is a straightforward task to find the value of u_1 which minimizes any of the pertinent expressions in (60). The least of the three is then the Value of the game. The optimal control for v_1 is found from (56). Thus the single-stage game is completely solved.

The solution to the two-stage game is still given by (25). Unfortunately, in the general case, it is no longer possible to express the control strategies and the Value by simple expressions similar to (28), (29), and (30). This comes as no particular surprise since it is a result of the nonlinearity of the control strategies and is not specifically related to differential games. The same behavior is exhibited when dynamic programming is used to solve optimal control problems involving constraints on the controls.

The same sort of a statement can be made concerning the solution to the N stage game. The problems encountered are those inherent in dynamic programming; no more theory is required. Accordingly, nothing more will be said about the case where the payoff is concave in the maximizing control variable.

5. DERIVATION OF RANDOMIZED CONTROL STRATEGIES

This section deals with a case where the payoff function is not strictly concave, that is, it is assumed that

$$b_1^2 c_1 - e_1 \geq 0 \tag{61}$$

(This is actually equivalent to saying that the payoff is convex in the maximizing control since every quadratic expression is either a concave or a convex function.)

It follows immediately from the second line of (47) that the maximum of $I_1(u_1, v_1)$ is achieved for

$$v_1 = \operatorname{sgn}[B] = \operatorname{sgn}[b_1 c_1(k_1 z_1 + a_1 u_1)] \tag{62}$$

where

$$\operatorname{sgn}[x] = \frac{x}{|x|} \quad ; x \neq 0 \tag{63}$$

and is to be defined for $x = 0$.

Consider first what happens when

$$\frac{k_1 z_1}{a_1} > 1 \tag{64}$$

When (64) holds, it is impossible for any choice of u_1 to change the sign of B; it must remain positive, which, by (62), means that the Value is given by

$$J_1 = \min_{u_1} c(k_1 z_1 + a_1 u_1 + b_1)^2 + d_1 u_1^2 - e_1 \tag{65}$$

The optimal value for u_1 is

$$\bar{u}_1 = \max\left[-1, \; -\frac{a_1 c_1 (k_1 z_1 + b_1)}{a_1^2 c_1 + d_1}\right] \tag{66}$$

When

$$\frac{k_1 z_1}{a_1} < -1 \tag{67}$$

then the same reasoning leads to the conclusion that

$$J_1 = \min_{u_1} c_1(k_1 z_1 + a_1 u_1 - b_1)^2 + d_1 u_1^2 - e \tag{68}$$

where

$$\bar{u}_1 = \min \left[1, - \frac{a_1 c_1 (k_1 z_1 - b_1)}{a_1^2 c_1 + d_1} \right] \tag{69}$$

In both (66) and (69), $\bar{u}_1$ does not take on a value on the boundary of U only if the weighting on the square of the control, d_1, is sufficiently large. Combining the two expressions shows that $\bar{u}_1$ is an interior point of U only if

$$d_1 > a_1 c_1 \left(|k_1 z_1| + b_1 \right) - a_1^2 c_1 > 0 \tag{70}$$

Even though the minimizing control may be forced to assume an interior value of U (not reducing the state component of the cost as much as possible), the maximizing player need never worry about a similar condition. This follows from the fact that the increase in the payoff due to a change in the state is greater than the cost incurred due to the use of the maximizing control — this is the real meaning of the payoff not being strictly concave in v_1. The concrete result is that player I always uses the largest magnitude of the control available to him for these two cases.

Finally, there is the case where

$$\left| \frac{k_1 z_1}{a_1} \right| \leq 1 \tag{71}$$

which is the most interesting of all, since it can lead to randomized control strategies. In this case, $\mathrm{sgn}[k_1 z_1 + a_1 u_1]$ may be equal to plus or minus one or sgn[0], depending on u_1. Because $\bar{v}_1$ is well defined when $k_1 z_1 + a_1 u_1$ is not equal to zero, it is possible to write

$$J_1 = \min \left\{ \min_{-1 \le u_1 < -\frac{k_1 z_1}{a_1}} c_1 (k_1 z_1 - a_1 u_1 - b_1)^2 + d_1 u_1^2 - e_1 ; \right.$$

$$u_1 = -\frac{k_1 z_1}{a_1} \quad \max_{v_1} c_1 (b_1 v_1)^2 + d_1 \left(\frac{k_1 z_1}{a_1} \right)^2 - e_1 ;$$

$$\left. \min_{-\frac{k_1 z_1}{a_1} < u_1 \le 1} c_1 (k_1 z_1 + a_1 u_1 + b_1)^2 + d_1 u_1^2 - e_1 \right\} \quad (72)$$

Consider the middle term first. Since $b_1^2 c_1 - e_1$ is non-negative,

$$\max_{v_1} (b_1 c_1 - e_1) v_1^2 + d_1 \left(\frac{k_1 z_1}{a_1} \right)^2 = b_1^2 c_1 - e_1 + d_1 \left(\frac{k_1 z_1}{a_1} \right)^2 \quad (73)$$

so that the maximum of v_1 (within V, of course) is achieved for v_1 equal to plus or minus one. Since these are precisely the values used for v_1 in the first and third terms of (72), the two half open intervals for u_1 can be replaced by closed intervals, and the second term can be discarded so that

$$J_1 = \min \left\{ \min_{-1 \le u_1 \le -\frac{k_1 z_1}{a_1}} c_1 (k_1 z_1 + a_1 u_1 - b_1)^2 + d_1 u_1^2 - e_1 ; \right.$$

$$\left. \min_{-\frac{k_1 z_1}{a_1} \le u_1 \le 1} c_1 (k_1 z_1 + a_1 u_1 + b_1)^2 + d_1 u_1^2 - e_1 \right\} \quad (74)$$

The result of (73) is effectively to define

$$|\mathrm{sgn}[0]| = 1 \tag{75}$$

although saying nothing about how an algebraic sign is to be attached.

The first term in (74) has a minimum, as a function of u_1, that need not lie between minus one and $-\frac{k_1 z_1}{a_1}$. Using ordinary calculus, the minimizing u_1 is found to be

$$u_1 = -\frac{a_1 c_1 (k_1 z_1 - b_1)}{a_1^2 c_1 + d_1} \tag{76}$$

which means that the absolute minimum of the term lies within the half-open interval $\left[-1, -\frac{k_1 z_1}{a_1}\right)$ when

$$-1 \leq -\frac{a_1 c_1 (k_1 z_1 - b_1)}{a_1^2 c_1 + d_1} < -\frac{k_1 z_1}{a_1} \tag{77}$$

The left-hand inequality always holds since it can be rewritten as

$$-1 - \frac{d_1}{a_1^2 c_1} \leq -\frac{k_1 z_1}{a_1} + \frac{b_1}{a_1} \tag{78}$$

where the first term on the right of (78) is always equal to or greater than minus one and the second term on the right is always positive, while the second term on the left is always negative.

Simple algebra shows that the right-hand inequality holds whenever $k_1 z_1$ is positive and

$$d_1 < \frac{a_1 b_1 c_1}{-\frac{k_1 z_1}{a_1}} \leq 0 \tag{79}$$

Since d_1 is assumed to be positive, (79) can never hold if $k_1 z_1 > 0$ and, for this case, the absolute minimum occurs for a value of u_1 greater than $-\frac{k_1 z_1}{a_1}$. If $k_1 z_1$ is negative and

$$d_1 > \frac{a_1 b_1 c_1}{-\frac{k_1 z_1}{a_1}} \tag{80}$$

then the absolute minimum occurs for a value of u_1 less than or equal to $-\frac{k_1 z_1}{a_1}$.

Investigation of the second term of (74) shows that the absolute minimum is achieved for values of u_1 less than one and falls within the interval $\left(-\frac{k_1 z_1}{a_1}, 1\right]$ whenever $k_1 z_1$ is positive and

$$d_1 > \frac{a_1 b_1 c_1}{\frac{k_1 z_1}{a_1}} \tag{81}$$

The absolute minimum occurs for

$$u_1 = -\frac{a_1 c_1 (k_1 z_1 + b_1)}{a_1^2 c_1 + d_1} \tag{82}$$

so that (76), (82), (80), and (81) can be summed up by saying that if

$$d_1 > \frac{a_1 b_1 c_1}{\left|\frac{k_1 z_1}{a_1}\right|} \tag{83}$$

then

$$\bar{u}_1 = -\frac{a_1 c_1 \left(k_1 z_1 + b_1 \operatorname{sgn}[k_1 z_1]\right)}{a_1^2 c_1 + d_1} \tag{84}$$

If (83) holds, then $\bar{v}_1$ is given by

$$\bar{v}_1 = \text{sgn}[k_1 z_1] \tag{85}$$

The game optimal minimizing control resulting from (84) means that $k_1 z_1 + a_1 \bar{u}_1$ has the same sign as $k_1 z_1$.

The Value of the game is then given by

$$J_1 = c_1 d_1 \frac{\left(|k_1 z_1| + b_1 \right)^2}{a_1^2 c_1 + d_1} - e_1 \tag{86}$$

If (83) does not hold, then the absolute minimum of each expression falls outside the half open intervals previously defined. The minimum is then achieved for

$$\bar{u}_1 = - \frac{k_1 z_1}{a_1} \tag{87}$$

with a corresponding Value of

$$J_1 = b_1^2 c_1 + d_1 \left(\frac{k_1 z_1}{a_1} \right)^2 - e_1 \tag{88}$$

Equation (75) already established that $\bar{v}_1$ will take on the values plus or minus one; having this knowledge, Theorem 2 of the Appendix can be used to find the optimal mixed strategy for the maximizing player. Using Theorem 2, it is seen that the optimal mixed strategy involves randomizing over the two points

$$\bar{v}_1^1 = -1 \tag{89}$$

$$\bar{v}_1^2 = +1 \tag{90}$$

with probabilities α and $(1 - \alpha)$, respectively, where α and $1 - \alpha$ are nonnegative. The optimal strategy is then given by a probability distribution:

$$\bar{v}_1 = \alpha H\left(\bar{v}_1^1\right) + (1 - \alpha) H\left(\bar{v}_1^2\right) \tag{91}$$

where $H(x)$ means the step function with the jump occurring at x. An expression for α is found from the requirement that

$$\alpha \left.\frac{\partial I_1}{\partial u_1}\right|_{\bar{u}_1, \bar{v}_1^1} + (1 - \alpha) \left.\frac{\partial I_1}{\partial u_1}\right|_{\bar{u}_1, \bar{v}_1^2} = 0 \tag{92}$$

and

$$\left.\frac{\partial I_1}{\partial u_1}\right|_{\bar{u}_1, \bar{v}_1^1} \leq 0 \tag{93}$$

$$\left.\frac{\partial I_1}{\partial u_1}\right|_{\bar{u}_1, \bar{v}_1^2} \geq 0 \tag{94}$$

Since

$$\frac{\partial I_1}{\partial u_1} = 2 a_1 c_1 (k_1 z_1 + a_1 u_1 + b_1 v_1) + 2 d_1 u_1 \tag{95}$$

$$\left.\frac{\partial I_1}{\partial u_1}\right|_{\bar{u}_1, \bar{v}_1^1} = - 2 a_1 b_1 c_1 - 2 d_1 \left(\frac{k_1 z_1}{a_1}\right) \tag{96}$$

$$\left.\frac{\partial I_1}{\partial u_1}\right|_{\bar{u}_1, \bar{v}_1^2} = 2 a_1 b_1 c_1 - 2 d_1 \left(\frac{k_1 z_1}{a_1}\right) \tag{97}$$

It follows that (96) must be nonpositive and (97) nonnegative. But those are precisely the requirement imposed by (83). Accordingly, α can be found from the solution of

$$- \alpha \left[a_1 b_1 c_1 + d_1 \left(\frac{k_1 z_1}{a_1} \right) \right] + (1 - \alpha) \left[a_1 b_1 c_1 - d_1 \left(\frac{k_1 z_1}{a_1} \right) \right] = 0 \tag{98}$$

so that

$$\alpha = \frac{a_1 b_1 c_1 - d_1 \left(\frac{k_1 z_1}{a_1} \right)}{2 a_1 b_1 c_1} \tag{99}$$

and the game optimal maximizing strategy, (91), is

$$\overline{v}_1 = \begin{cases} -1 & \text{with probability} \quad \dfrac{a_1 b_1 c_1 - d_1 \left(\frac{k_1 z_1}{a_1} \right)}{2 a_1 b_1 c_1} \\ \\ 1 & \text{with probability} \quad \dfrac{a_1 b_1 c_1 + d_1 \left(\frac{k_1 z_1}{a_1} \right)}{2 a_1 b_1 c_1} \end{cases} \tag{100}$$

The game optimal control strategies and the Value for the single-stage game, as a function of the state and the various system parameters, are summarized in Table 1.

Table 1 illustrates the nonlinear nature of the strategies as functions of both the state and the relative values of the system parameters. And, for the first time, a situation exists where a randomized strategy is optimal.

TABLE 1

STRATEGIES AND PAYOFF FOR

SINGLE-STAGE GAMES

$\left\lvert \frac{k_1 z_1}{a_1} \right\rvert > 1$	
$0 < d_1 \le a_1^2 c_1 \left(\left\lvert \frac{k_1 z_1}{a_1} \right\rvert + \frac{b_1}{a_1} - 1 \right)$	$d_1 > a_1^2 c_1 \left(\left\lvert \frac{k_1 z_1}{a_1} \right\rvert + \frac{b_1}{a_1} - 1 \right) > 0$
$\bar{u}_1 = -\operatorname{sgn}\left[\frac{k_1 z_1}{a_1}\right]$	$u_1 = \frac{a_1 c_1 (k_1 z_1 + b_1 \operatorname{sgn}[k_1 z_1])}{a_1^2 c_1 + d_1}$
$\bar{v}_1 = \operatorname{sgn}\left[\frac{k_1 z_1}{a_1}\right]$	$\bar{v}_1 = \operatorname{sgn}\left[\frac{k_1 z_1}{a_1}\right]$
$J_1 = a_1^2 c_1 \left[\left\lvert \frac{k_1 z_1}{a_1} \right\rvert - 1 + \frac{b_1}{a_1} \right]^2 + d_1 - e_1$	$J_1 = c_1 d_1 \frac{(\lvert k_1 z_1 \rvert + b_1)^2}{a_1^2 c_1 + d_1} - e_1$

$\left\lvert \frac{k_1 z_1}{a_1} \right\rvert \le 1$	
$d_1 < \frac{a_1 b_1 c_1}{\left\lvert \frac{k_1 z_1}{a_1} \right\rvert}$	$d_1 > \frac{a_1 b_1 c_1}{\left\lvert \frac{k_1 z_1}{a_1} \right\rvert}$
$\bar{u}_1 = \frac{k_1 z_1}{a_1}$	$\bar{u}_1 = -\frac{a_1 c_1 (k_1 z_1 + b_1 \operatorname{sgn}[k_1 z_1])}{a_1^2 c_1 + d_1}$
$\bar{v}_1 = \begin{cases} -1 \text{ with probability } \frac{a_1 b_1 c_1 - d_1 \left(\frac{k_1 z_1}{a_1}\right)}{2 a_1 b_1 c_1} \\ 1 \text{ with probability } \frac{a_1 b_1 c_1 + d_1 \left(\frac{k_1 z_1}{a_1}\right)}{2 a_1 b_1 c_1} \end{cases}$	$\bar{v}_1 = \operatorname{sgn}\left[\frac{k_1 z_1}{a_1}\right]$
$J_1 = b_1^2 c_1 + d_1 \left(\frac{k_1 z_1}{a_1}\right)^2 - e_1$	$J_1 = c_1 d_1 \frac{(\lvert k_1 z_1 \rvert + b_1)^2}{a_1^2 c_1 + d_1} - e_1$

Notice that e_1 does not appear except in the Value. This is to be expected since $|\bar{v}_1|$ always is equal to one. Accordingly, the penalty incurred through the use of v_1 is always the same.

As noted in the last section, the nonlinear nature of the game optimal control strategies makes it impossible to solve the multistage differential game explicitly in a few statements. The general solution is still given by functional equations of the form of (25) using the techniques outlined in this chapter.

6. EXAMPLE

This section illustrates the solution to a multistage differential game having randomized strategies. To do this, it is assumed that:

1) The system parameters a, b, c, d, e, and k are constant
2) The system parameters a, b, and e are positive
3) The initial state, z_N, is such that $\left|\frac{kz_N}{a}\right| < 1$
4) The control magnitudes are such that $a \geq b$
5) The weighting on the use of u_i is small enough so that $d < \frac{ab\tilde{c}}{\left|\frac{kz_N}{a}\right|}$, where $\tilde{c}$ is to be defined

Table 1 supplies the solutions to the single-stage game. Substituting J_1 into (25) and replacing "val" by "min max" yield

$$J_2 = \min_{u_2} \max_{v_2} \left[cz_1^2 + du_2^2 - ev_2^2 + b^2c^2 + d\left(\frac{kz_1}{a}\right)^2 - e \right]$$

$$= \min_{u_2} \max_{v_2} \left[\left(c + \frac{dk^2}{a}\right) z_1^2 + du_2^2 - ev_2^2 + b^2 c - e \right] \qquad (101)$$

Letting

$$\tilde{c} = c + \frac{dk^2}{a^2} \tag{102}$$

and substituting (1) and (102) into (101) leads to

$$J_2 = \min_{u_2} \max_{v_2} \left[\tilde{c} \, (kz_2 + au_2 + bv_2)^2 + du_2^2 - ev_2^2 + b^2 c - e \right] \tag{103}$$

which has the same form, except for a constant term, as (47). In other words, the same techniques used to solve the single-stage game will suffice to solve stage 2 of a two-stage game.

Since the constant term affects only the magnitude of the Value (and not the control strategies), it follows immediately that

$$\bar{u}_2 = - \frac{kz_2}{a} \tag{104}$$

$$\bar{v}_2 = \begin{cases} -1 & \text{with probability} \quad \dfrac{ab\tilde{c} - d\left(\dfrac{kz_2}{a}\right)}{2\,ab\tilde{c}} \\ \\ 1 & \text{with probability} \quad \dfrac{ab\tilde{c} - d\left(\dfrac{kz_2}{a}\right)}{2\,ab\tilde{c}} \end{cases} \tag{105}$$

and that the Value for the two-stage game is given by

$$J_2 = 2\,b^2 c - 2e + d\left(\frac{kb}{a}\right)^2 + d\left(\frac{kz_2}{a}\right)^2 \tag{106}$$

Repeated application of (25) leads to

$$J_N = d\left(\frac{kz_N}{a}\right)^2 + N(cb^2 - e) + (N-1)\, d\left(\frac{kb}{a}\right)^2 \tag{107}$$

with the game optimal control strategies being given by

$$\bar{u}_N = -\frac{kz_N}{a} \tag{108}$$

$$\bar{v}_N = \begin{cases} -1 & \text{with probability } \dfrac{ab\tilde{c} - d\left(\dfrac{kz_N}{a}\right)}{2\,ab\tilde{c}} \\ 1 & \text{with probability } \dfrac{ab\tilde{c} + d\left(\dfrac{kz_N}{a}\right)}{2\,ab\tilde{c}} \end{cases} \tag{109}$$

The Value of the game is a function only of the initial state, z_N, and the number of stages, as it should be. It is not a random variable. The same is not true for the trajectory. It describes a stochastic process taking on the values plus or minus b at each stage. Thus there are 2^N possible paths for the system to follow. It should be clear now why assumption 4 is required: without it the state at stage N - 1 would be such that the randomized control strategy would no longer be necessary. (This is the case, with d and e equal to zero, covered in Chapter III of Reference 35.)

VII. CONCLUDING COMMENTS

This chapter has dealt with the solution to multistage scalar games with linear dynamics and quadratic payoff functions. This was done for convenience rather than out of necessity, particularly in the case of randomized strategies. As Reference 34 indicates, the solution of convex games is, in theory, identical for finite dimensional systems; practically speaking, the added notational complexity would only obscure an already diffuse solution.

Results, completely analogous to those obtained in this chapter, are obtained under the assumption that the payoff is strictly concave in the maximizing control variable.

SECTION III

MULTISTAGE STOCHASTIC DIFFERENTIAL GAMES

1. INTRODUCTION

This chapter deals with the multistage, discrete time stochastic differential game. The dynamics are described by a linear difference equation having time varying deterministic coefficients, with the possible exception of a noisy forcing function. Both players wish to choose controls so as to either maximize (player I) or minimize (player II) the expected value of a quadratic cost functional. Neither player can observe the actual state of the system; instead, each player has an observation of the state which is corrupted by additive noise.

The purpose of this chapter is to derive game optimal strategies so as to allow the determination of the appropriate controls at each stage. It is assumed that pure strategies exist; necessary and sufficient conditions are derived for this to be true.

2. NOTATION

An N stage game is defined as one requiring that each player choose a value for his control at N instants of time. Time-to-go, rather than the usual forward flowing time, is treated as the independent variable. A subscript i indicates that the subscripted variable is at the i^{th} stage. Thus z_N represents the state at the N^{th}

(initial stage – N stages to go) stage, while z_o represents the state of the system at termination.

A superscript is used to indicate the set of present and all past values of the variable under discussion. Thus z^1 is the set of the present state, z_1 (the state when there is one stage-to-go), as well as all past values $z_2, z_3, \cdots, z_N$, i.e., $z^1 = (z_1, z_2, \cdots, z_N)$. Naturally, $z^N = z_N$.

A similar convention is used to indicate integration over a set of variables. Thus $dz_1\, dz_2 \cdots dz_N$ is written dz^1. Further, when several different variables of integration are used, only a single integral sign is used; the number of integrations is indicated by the differential. Thus

$$\int \cdots \int p(z_1, x_1, \cdots, x_N | y_1, \cdots, y_N)\, dz_1\, dx_1 \cdots dx_N$$

is written as

$$\int p(z_1, x^1 | y^1)\, d(z_1, x^1)$$

When no limits of integration are specified, the integration is considered to be from minus infinity to plus infinity.

Quadratic forms are given by

$$x^T A x = ||x||_A^2$$

so that a Gaussian probability density is given by

$$p(x) = \frac{1}{(2\pi)^{n/2}\ |\sigma|}\ \exp\left\{-\frac{1}{2}||x - \bar{x}||^2_{\sigma^{-2}}\right\}$$

where x is a vector with n components, $\bar{x}$ is the mean of x, and σ^2 is the n $\times$ n covariance vector of x. For simplicity, only twice the negative of the exponent is used when computations are required. Thus the probability density given above is denoted

$$p(x) = ||x - \bar{x}||^2_{\sigma^{-2}}$$

The Gaussian (normal) probability density is also denoted as

$$x\colon N(\bar{x}, \sigma^2)$$

where $\bar{x}$ and σ^2 are, respectively, the mean and covariance of the random variable x, i.e.,

$$\bar{x} = E\{x\}$$

$$\sigma^2 = E\left\{(x - \bar{x})(x - \bar{x})^T\right\}$$

where E is the expected value operator.

Various other subscripts are used to identify variables as required. Thus $\sigma^2_{\eta_2}$ indicates the covariance matrix of the random variable η at stage 2, when two stages-to-go remain.

3. DERIVATION OF PURE CONTROL STRATEGIES

The evolution of the state is described by a linear vector difference equation

$$z_i = k_{i+1} z_{i+1} + a_{i+1} u_{i+1} + b_{i+1} v_{i+1} + \lambda_{i+1}, \quad i = 0, 1, \cdots, N-1 \tag{1}$$

where

z_i = n component vector representing the state at stage i

u_i = m component vector representing the minimizing player's control at stage i

v_i = m' component vector representing the maximizing player's control at stage i

λ_i = n component vector representing the realization of an independent noise sequence at stage i

k_i, a_i, b_i = deterministic matrices of the appropriate dimensions

Both players make observations of linear combinations of the elements of the state vector, but each observation is corrupted by an additive, independent noise sequence so that

$$x_i = G_i z_i + \eta_i \tag{2}$$

$$y_i = H_i z_i + \xi_i \tag{3}$$

where

x_i = q component vector representing the minimizing player's observation at stage i

y_i = q' component vector representing the maximizing player's observation at stage i

η_i = q component vector representing the realization of an independent noise sequence

ξ_i = q' component vector representing the realization of an independent noise sequence

H_i, G_i = deterministic matrices of the appropriate dimensions

The Value of the multistage game, J_N, is given by

$$J_N = \underset{u^1,\ v^1}{\text{val}} E\left\{ \sum_{i=1}^{N} ||z_{i-1}||^2_{c_i} + ||u_i||^2_{d_i} - ||v_i||^2_{e_i} \right\}$$

$$= \min_{u^1}\ \max_{v^1}\ E\left\{ \sum_{i=1}^{N} ||z_{i-1}||^2_{c_i} + ||u_i||^2_{d_i} - ||v_i||^2_{e_i} \right\} \qquad (4)$$

("Min max" could be replaced by "max min" since only pure strategies are considered.) u_i and v_i can be any vectors of real numbers; however,

$$c_1, d_i, e_i > 0, \quad i = 1, 2, \cdots, N \qquad (5)$$

where the inequality in (5) means that each criterion parameter is a real positive definite matrix. Inequality signs are used as needed and should be read to mean positive definite (instead of greater than zero), positive semidefinite (instead of greater than or equal to zero), etc.

Only c_1 need be positive; $c_i (i \neq 1)$ need only be nonnegative. If c_i does equal zero, then the result is a terminal control problem. The remaining inequalities of (5) are required to make the game meaningful. If any parameter were to be zero, then either there would be no effect due to a control at that stage (a_i or b_i equal to zero) or there would be no cost for using control at that stage (d_i or e_i equal to zero).

For convenience, it is assumed that all the criteria parameters in (5) are symmetric.

Each of the various noises is assumed to have a Gaussian (normal) probability density as follows:

$$\lambda_i : N\left(0, \sigma^2_{\lambda_i}\right) \tag{6}$$

$$\eta_i : N\left(0, \sigma^2_{\eta_i}\right) \tag{7}$$

$$\xi_i : N\left(0, \sigma^2_{\xi_i}\right) \tag{8}$$

The initial state, z_N, also will be a Gaussian random variable:

$$z_N : N\left(m, \sigma^2_z\right) \tag{9}$$

It is assumed that both players know the information contained in (1) through (9). This does not mean that player I knows the actual value for x_i — the observation of the state at the i^{th} stage made by player II — but that he does know the structure of the observation, as given by (2), and the probability density of the additive noise, as given by (7).

To actually solve the multistage game, the principle of optimality of the theory of dynamic programming will be used. To do this, game optimization is first carried out for the controls chosen with one stage-to-go, u_1 and v_1. The technique used is similar, in terms of the basic structure, to that used to solve multistage stochastic optimal control problems.[36] The Value for this one-stage game is

$$J_1 = \min_{u_1} \max_{v_1} E\left\{ ||z_o||_{c_1} + ||u_1||^2_{d_1} - ||v_1||^2_{e_1} \right\}$$

$$= \min_{u_1} \max_{v_1} \int \left\{ ||z_o||^2_{c_1} + ||u_1||^2_{d_1} - ||v_1||^2_{e_1} \right\}$$

$$\times\ p(z_o, u_1, v_1)\, d(z_o, u_1, v_1) \tag{10}$$

where $p(z_o, u_1, v_1)$ is the joint (Gaussian) probability density function of z_o, u_1, and v_1. Since[37]

$$p(z_o, u_1, v_1) = \int p(z_0, z_1, u_1, v_1)\, dz_1$$

$$= \int p(z_o | z_1, u_1, v_1)\, p(z_1, u_1, v_1)\, dz_1 \tag{11}$$

Equation (10) can be rewritten as

$$J_1 = E\left\{\|\lambda_1\|^2_{c_1}\right\} + \min_{u_1}\max_{v_1} \int \left\{\|k_1 z_1 + a_1 u_1 + b_1 v_1\|^2_{c_1} + \|u_1\|^2_{d_1} - \|v_1\|^2_{e_1}\right\}$$

$$\times\ p(z_1, u_1, v_1)\ d(z_1, u_1, v_1) \tag{12}$$

Noting that

$$p(z_1, u_1, v_1) = \int p(z_1, x^1, y^1, u^1, v^1)\ d(x^1, y^1, u^2, v^2)$$

$$\int p(u_1, v_1 \,|\, z_1, x^1, y^1, u^2, v^2)\ p(z_1, x^1, y^1, u^2, v^2)$$

$$\times\ d(x^1, y^1, u^2, v^2) \tag{13}$$

where the entire past history of each player's observations, x^1 and y^1, and controls, u^2 and v^2, has been introduced, (12) can be rewritten as

$$J_1 = E\left\{\|\lambda_1\|^2_{c_1}\right\} + \min_{u_1}\max_{v_1} \int\!\!\int \left\{\|k_1 z_1 + a_1 u_1 + b_1 v_1\|^2_{c_1} + \|u_1\|^2_{d_1} - \|v_1\|^2_{e_1}\right\}$$

$$\times\ p(u_1, v_1 \,|\, z_1, x^1, y^1, u^2, v^2)\ d(u_1, v_1)$$

$$\times\ p(z_1, x^1, y^1, u^2, v^2)\ d(z_1, x^1, y^1, u^2, v^2) \tag{14}$$

It is at this point that admissible strategies are introduced. Equation (14) indicates that the probability density for u_1 and v_1 can be, if desired, conditioned on the actual value of the state, z_1; all or part of player I's past observations, y^1, and controls, v^2; and all or part of player II's past history of observations and controls, x^1 and u^2, respectively. Each choice of the characterization of information available leads to a different problem with different results. Further, if desired, the allowable structure of the controls may be specified. (This will be delineated in Chapter 5.)

A very reasonable set of controls may be found which uses only information reasonably available to each player. That is, playe I may choose a control strategy with one stage-to-go using only his own history of observations and controls (and, of course, his knowledge of the dynamics and payoff of the game as given by (1) through (9)), while player II chooses his control strategy based on his own history of observations and controls.

Assuming that nonrandomized (pure) control strategies exist for both players, and denoting them by an overbar, the game optimal control strategies are given by

$$\bar{u}_1 = \bar{u}_1(x^1, u^2) \tag{15}$$

$$\bar{v}_1 = \bar{v}_1(y^1, v^2) \tag{16}$$

Note that nothing has been said at this point about the strategies used to determine u^2 and v^2. In particular, note that no assumptions concerning their optimality have been made.

In view of (15) and (16), let

$$p(u_1, v_1 | z_1, x^1, y^1, u^2, v^2) = \delta(u_1 - \bar{u}_1)\,\delta(v_1 - \bar{v}_1) \tag{17}$$

where δ represents the Dirac delta (impulse) function. Substituting (17) into (14) and integrating over u_1 and v_1 yield

$$J_1 = E\left\{\|\lambda_1\|^2_{c_1}\right\} + \int\left\{\|k_1 z_1 + a_1\bar{u}_1 + b_1\bar{v}_1\|^2_{c_1} + \|\bar{u}_1\|^2_{d_1} - \|\bar{v}_1\|^2_{e_1}\right\}$$

$$\times\, p\left(z_1, x^1, y^1, u^2, v^2\right)\, d\left(z_1, x^1, y^1, u^2, v^2\right) \tag{18}$$

where the min max operation no longer need be performed since $\bar{u}_1$ and $\bar{v}_1$ are precisely those control strategies which satisfy the min max = max min requirement.

Variational arguments can be used to find the actual form of the game optimal controls based on the strategies allowed by (15) and (16). Define

$$I_1(\tilde{u}_1, \tilde{v}_1) = E\left\{\|\lambda_1\|^2_{c_1}\right\} + \int\left\{\|k_1 z_1 + a_1\tilde{u}_1 + b_1\tilde{v}_1\|^2_{c_1} + \|\tilde{u}_1\|^2_{d_1} - \|\tilde{v}_1\|^2_{e_1}\right\}$$

$$\times\, p\left(z_1, x^1, y^1, u^2, v^2\right)\, d\left(z_1, x^1, y^1, u^2, v^2\right) \tag{19}$$

where $\tilde{u}_1$ and $\tilde{v}_1$ are any admissible strategies. It immediately follows that

$$\min_{\tilde{u}_1} \max_{\tilde{v}_1} I_1(\tilde{u}_1, \tilde{v}_1) = I_1(\bar{u}_1, \bar{v}_1) = J_1 \tag{20}$$

Game optimal controls must satisfy the saddle point conditions which are:

$$I_1(\tilde{u}_1, \bar{v}_1) \geq J_1 \geq I_1(\bar{u}_1, \tilde{v}_1) \tag{21}$$

where J_1 is given by (18).

Let

$$\tilde{u}_1 = \bar{u}_1 + \epsilon\,\delta(x^1, u^2) \tag{22}$$

where ϵ is a small number and $\delta(x^1, u^2)$ is any real vector function, of the appropriate dimension, of x^1 and u^2. Using (22), the left-hand inequality of (21) can be written as

$$I_1(\bar{u}_1 + \epsilon\delta) - J_1 \geq 0 \tag{23}$$

where the arguments of δ have been omitted for brevity. Substituting (18) and (19) into (23) yields

$$I_1(\bar{u}_1+\epsilon\delta)-J_1 = \int\Big\{ \| k_1 z_1 + a_1\bar{u}_1 + a_1\epsilon\delta + b_1\bar{v}_1 \|^2_{c_1} + \|\bar{u}_1 + \epsilon\delta\|^2_{d_1} - \|\bar{v}_1\|^2_{e_1}$$

$$- \| k_1 z_1 + a_1\bar{u}_1 + b_1\bar{v}_1\|^2_{c_1} - \|\bar{u}_1\|^2_{d_1} + \|v_1\|^2_{e_1} \Big\}$$

$$\times\, p\left(z_1, x^1, y^1, u^2, v^2\right)\, d\left(z_1, x^1, y^1, u^2, v^2\right)$$

$$= \int\Big\{ 2\left[a_1^T c_1 k_1 z_1 + \left(a_1^T c_1 a_1 + d_1\right)\bar{u}_1 + a_1^T c_1 b_1 \bar{v}_1 \right]\epsilon\delta$$

$$+ \delta^T \left(a_1^T c_1 a_1 + d_1 \right)\delta\epsilon^2 \Big\}\, p\left(z_1, x^1, y^1, u^2, v^2\right)$$

$$\times\, d\left(z_1, x^1, y^1, u^2, v^2\right) \geq 0 \tag{24}$$

Using the chain rule of conditional probability densities and applying it to the probability density of (24) yields

$$p\left(z_1, x^1, y^1, u^2, v^2\right) = p\left(z_1 \mid x^1, y^1, u^2, v^2\right)\ p\left(x^1, y^1, u^2, v^2\right)$$

$$= p\left(z_1 \mid x^1, y^1, u^2, v^2\right)\ p\left(y^1 \mid x^1, u^2, v^2\right)\ p(x^1, u^2, v^2)$$

$$= p\left(z_1 \mid x^1, y^1, u^2, v^2\right)\ p\left(y^1 \mid x^1, u^2, v^2\right) p(v^2 \mid x^1, u^2)$$

$$\times\ p(x^1, u^2) \tag{25}$$

The mean associated with the first three conditional probability densities of (25) have the following meanings:

$p\left(z_1 \mid x^1, y^1, u^2, v^2\right)$ = minimum mean square estimate (MMSE) of the state, z_1, given all past and present observations and past controls for both players

$p(y^1 \mid x^1, u^2, v^2)$ = MMSE of all observations of player I, y^1, given player II's past and present observations and all of both players past controls

$p(v^2 \mid x^1, u^2)$ = MMSE of all past controls of player I, v^2, given only player I's past and present observations and past controls

It is important to note again that (25) involves estimation only. No assumptions concerning the game optimality of any of the past controls of either player have been made.

Substituting (25) into (24) yields, after some slight manipulation,

$$I_1(\bar{u}_1+\epsilon\delta,\bar{v}_1)-J_1=\int 2\left[\left\{a_1^T c_1 k_1 z_1+\left(a_1^T c_1 a_1+d_1\right)\bar{u}_1+a_1^T c_1 b_1 \bar{v}_1\right\}\right.$$

$$\times\, p\left(z_1 \,|\, x^1, y^1, u^2, v^2\right)\, dz_1$$

$$\left.\times\, p\left(y^1 \,|\, x^1, u^2, v^2\right)\, dy^1\, p\left(v^2 \,|\, x^1, u^2\right)\, dv^2\right]\epsilon\delta(x^1, u^2)$$

$$\times\, p(x^1, u^2)\, d(x^1, u^2)$$

$$+\left\{\delta^T(x^1, u^1)\left(a_1^T c_1 a_1+d_1\right)\delta(x^1, u^1)\,\epsilon^2\right\}$$

$$\times\, p\left(z_1, x^1, y^1, u^2, v^2\right)\, d\left(z_1, x^1, y^1, u^2, v^2\right) \geq 0 \qquad (26)$$

In view of (5) and, because ϵ and $\delta(x^1, u^2)$ are real, it is clear that the second integral of (26) is positive semidefinite. Invoking the standard variational arguments of the calculus of variations, it follows that the coefficient of $\epsilon\delta(x^1, u^2)$ in (26) must be equal to zero. If it were not zero, then $\delta(x^1, u^2)$ could be chosen to have the opposite sign as its coefficient. For ϵ small enough, the first integral would be larger in magnitude than the second and the inequality would not hold. Thus a necessary condition that a pure strategy exist for player II is that

$$\int \left\{ a_1^T c_1 k_1 z_1 + \left(a_1^T c_1 a_1 + d_1\right) \bar{u}_1 + a_1^T c_1 b_1 \bar{v}_1 \right\} p\left(z_1 \mid x^1, y^1, u^2, v^2\right) dz_1$$

$$\times\, p\left(y^1 \mid x^1, u^2, v^2\right) dy^1 \; p\left(v^2 \mid x^1, u^2\right) dv^2 = 0 \qquad (27)$$

(It is assumed that $p(x^1, u^2)$ is nonzero for all values of its arguments.)

Looking now at the right-hand inequality of (21), it is easy to see that

$$I_1(\bar{u}_1, \bar{v}_1 + \epsilon\Delta) - J_1 = \int 2 \left[\left\{ b_1^T c_1 k_1 z_1 + b_1^T c_1 a_1 \bar{u}_1 - \left(e_1 - b_1^T c_1 b_1\right) \bar{v}_1 \right\} \right.$$

$$\times\, p\left(z_1 \mid x^1, y^1, u^2, v^2\right) dz_1$$

$$\left. \times\, p\left(x^1 \mid y^1, u^2, v^2\right) dx^1 \; p\left(u^2 \mid y^1, v^2\right) du^2 \right] \epsilon \Delta\left(y^1, v^2\right)$$

$$\times\, p(y^1, v^2)\, d(y^1, v^2)$$

$$- \int \left\{ \Delta^T\left(y^1, v^2\right) \left(e_1 - b_1^T c_1 b_1\right) \Delta\left(y^1, v^2\right) \epsilon^2 \right\}$$

$$\times\, p\left(z_1, x^1, y^1, u^2, v^2\right) d\left(z_1, x^1, y^1, u^2, v^2\right) \leq 0 \qquad (28)$$

where $\Delta(y^1, v^2)$ is any real vector function of the appropriate dimension of y^1 and v^2. It immediately follows that the necessary condition for a pure strategy to exist for player I is that

$$\int \left\{ b_1^T c_1 k_1 z_1 + b_1^T c_1 a_1 \bar{u}_1 - \left(e_1 - b_1^T c_1 b_1\right) \bar{v}_1 \right\} p\left(z_1 \mid x^1, y^1, u^2, v^2\right) dz_1$$

$$\times\, p\left(x^1 \mid y^1, u^2, v^2\right) dx^1 \; p\left(u^2 \mid y^1, v^2\right) du^2 = 0 \qquad (29)$$

$$e_1 - b_1^T c_1 b_1 \geq 0 \qquad (30)$$

Equations (22) and (24) can be solved simultaneously for $\bar{u}_1$ and $\bar{v}_1$. To do this it is useful to introduce a set of linear transformations defined on suitable Hilbert spaces.[38] In each of the following transformations, $\alpha(\cdot)$ is an element in the domain of the transformation and $\beta(\cdot)$ is an element in its range.

$$\beta(x^1, y^1, u^2, v^2) = T_{11}\, \alpha(z_1) = \int \alpha(z_1)\, p\left(z_1 \mid x^1, y^1, u^2, v^2\right) dz_1 \qquad (31)$$

where

$$T_{11}: L_2\left[-\infty, \infty;\, p\left(z_1 \mid x^1, y^1, u^2, v^2\right)\right] \rightarrow L_2\left[-\infty, \infty;\, p\left(y^1 \mid x^1, u^2, v^2\right)\right] \qquad (32)$$

Both the domain and the range of the transformation are thus defined to be Hilbert spaces, with the appropriate conditional probability density taken as a measure on the space. By introducing such a measure, a number of functions, which would not ordinarily be L_2 when the limits of integration are plus and minus infinity, can be considered elements of a Hilbert space. In particular, the element z_1 is now L_2.

It should also be noticed that the range of T_{11} is multi-dimensional, since y^1 represents the N values of y_i. This does not add any conceptual difficulties although, as is seen later in an example, the practical problems of evaluation are increased.

The remaining transformations are

$$\beta(x^1, u^2, v^2) = T_{12}\, \alpha(y^1) = \int \alpha(y^1)\, p\left(y^1 \mid x^1, u^2, v^2\right) dy^1 \tag{33}$$

where

$$T_{12}\colon L_2\left[-\infty, \infty;\ p\left(y^1 \mid x^1, u^2, v^2\right)\right] \to L_2\left[-\infty, \infty;\ p(v^2 \mid x^1, u^2)\right] \tag{34}$$

$$\beta(x^1, u^2) = T_{13}\, \alpha(v^2) = \int \alpha(v^2)\, p(v^2 \mid x^1, u^2)\, dv^2 \tag{35}$$

where

$$T_{13}\colon L_2\left[-\infty, \infty;\ p(v^2 \mid x^1, u^2)\right] \to L_2\left[-\infty, \infty;\ p(x^1 \mid y^1, u^2, v^2)\right] \tag{36}$$

$$\beta(y^1, u^2, v^2) = T_{14}\, \alpha(x^1) = \int \alpha(x^1)\, p(x^1 \mid y^1, u^2, v^2)\, dx^1 \tag{37}$$

where

$$T_{14}\colon L_2\left[-\infty, \infty;\ p(x^1 \mid y^1, u^2, v^2)\right] \to L_2\left[-\infty, \infty;\ p(u^2 \mid y^1, v^2)\right] \tag{38}$$

$$\beta(y^1, v^2) = T_{15}\, \alpha(u^2) = \int \alpha(u^2)\, p(u^2 \mid y^1, v^2)\, du^2 \tag{39}$$

where

$$T_{15}\colon L_2\left[-\infty, \infty;\ p(u^2 \mid y^1, v^2)\right] \to L_2\left[-\infty, \infty;\ p(y^1 \mid x^1, u^2, v^2)\right] \tag{40}$$

and

$$\beta(x^1, y^1, u^2, v^2) = T_{16}\,\alpha(z_1) = \int \alpha(z_1)\, p\left(z_1 \mid x^1, y^1, u^2, v^2\right) dz_1 \quad (41)$$

where

$$T_{16}: L_2\left[-\infty, \infty;\ p\left(z_1 \mid x^1, y^1, u^2, v^2\right)\right] \rightarrow L_2\left[-\infty, \infty;\ p\left(x^1 \mid y^1, u^2, v^2\right)\right] \quad (42)$$

T_{16} and T_{11} differ only in the measure defined on the range space. For all practical purposes, they are identical since, for the elements of the domain of T_{11} and T_{16} encountered in this problem, an element in the range of one is also an element in the range of the other. This is, of course, not an intrinsic property of linear transformations but is a direct consequence of the simple structure of the problem under consideration.

Using (31) through (42), it is possible to write (27) and (29) as linear operator equations as follows:

$$a_1^T c_1 k_1 T_{13} T_{12} T_{11} z_1 + \left(a_1^T c_1 a_1 + d_1\right) \bar{u}_1 + a_1^T c_1 b_1 T_{13} T_{12} \bar{v}_1 = 0 \quad (43)$$

$$b_1^T c_1 k_1 T_{15} T_{14} T_{16} z_1 + b_1^T c_1 a_1 T_{15} T_{14} \bar{u}_1 - \left(e_1 - b_1^T c_1 b_1\right) \bar{v}_1 = 0 \quad (44)$$

Solving (43) and (44) simultaneously for $\bar{u}_1$ and $\bar{v}_1$ yields

$$\bar{u}_1 = -\left\{ I - \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 b_1 \left(b_1^T c_1 b_1 - e_1\right)^{-1} b_1^T c_1 a_1 T_{13} T_{12} T_{15} T_{14} \right\}^{-1}$$
$$\times \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 T_{13} T_{12} \left\{ T_{11} + b_1 \left(e_1 - b_1^T c_1 b_1\right)^{-1} b_1^T c_1 T_{15} T_{14} T_{16} \right\} k_1 z_1 \quad (45)$$

$$\overline{v}_1 = \left\{ I - \left(b_1^T c_1 b_1 - e_1\right)^{-1} b_1^T c_1 a_1 \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 b_1 T_{15} T_{14} T_{13} T_{12} \right\}^{-1}$$

$$\times \left(e_1 - b_1^T c_1 b_1\right)^{-1} b_1^T c_1 T_{15} T_{14} \left\{ T_{16} - a_1 \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 T_{13} T_{12} T_{11} \right\} k_1 z_1 \tag{46}$$

where I is an appropriately dimensioned unit matrix.

Since $T_{16} z_1$ is an element of the range space containing $T_{11} z_1$, (45) and (46) can be (slightly) rewritten as

$$\overline{u}_1 = - \left\{ I - \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 b_1 \left(b_1^T c_1 b_1 - e_1\right)^{-1} b_1^T c_1 a_1 T_{13} T_{12} T_{15} T_{14} \right\}^{-1}$$

$$\times \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 T_{13} T_{12} \left\{ I + b_1 \left(e_1 - b_1^T c_1 b_1\right)^{-1} b_1^T c_1 T_{15} T_{14} \right\} T_{11} k_1 z_1 \tag{47}$$

$$\overline{v}_1 = \left\{ I - \left(b_1^T c_1 b_1 - e_1\right)^{-1} b_1^T c_1 a_1 \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 b_1 T_{15} T_{14} T_{13} T_{12} \right\}^{-1}$$

$$\times \left(e_1 - b_1^T c_1 b_1\right)^{-1} b_1^T c_1 T_{15} T_{14} \left\{ I - a_1 \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 T_{13} T_{12} \right\} T_{11} k_1 z_1 \tag{48}$$

Equations (47) and (48) represent, in functional form, the game optimal control strategies to be used by each player. But, unless a way to evaluate the inverses is found, the solutions are formal and essentially meaningless. Happily, they can be evaluated, under certain circumstances, in an infinite series. This series, a

Neumann expansion,[38] converges (the inverse exists and has meaning) whenever the norm of the second term of the inverse is less than one. In other words,

$$(I-RT)^{-1} = I + RT + (RT)^2 + (RT)^3 + \cdots \tag{49}$$

whenever the norm of RT, denoted $||RT||$, is less than one. For the present case, the inverses exist whenever

$$|| \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 b_1 \left(b_1^T c_1 b_1 - e_1\right)^{-1} b_1^T c_1 a_1 T_{13} T_{12} T_{15} T_{14} || < 1 \tag{50}$$

$$|| \left(b_1^T c_1 b_1 - e_1\right)^{-1} b_1^T c_1 a_1 \left(a_1^T c_1 a_1 + d_1\right)^{-1} a_1^T c_1 b_1 T_{15} T_{14} T_{13} T_{12} || < 1 \tag{51}$$

where the norm of a transformation RT is given by[38]

$$||RT|| = \sup_{\alpha} \frac{||RT\alpha||}{||\alpha||} \tag{52}$$

and α is any nonzero element of the domain of the transformation RT. Naturally, the norms of α and $RT\alpha$ are computed according to the weighting function defined on the appropriate Hilbert space.

Thus sufficient conditions for the existence of game optimal controls are (43), (44), (50), and (51), and

$$a_1^T c_1 a_1 + d_1 > 0 \tag{53}$$

$$e_1 - b_1^T c_1 b_1 > 0 \tag{54}$$

where the strict inequality has replaced the positive semidefiniteness of (30).

When $\bar{u}_1$ and $\bar{v}_1$ exist, it is a straightforward (although extremely tedious) task to verify that the following assumption is valid: each involves only the past and present observations and past controls available to the appropriate player. The required admissible game optimal control strategies thus have been found for the last stage of an N stage game. There is still the matter of actually evaluating the various conditional probability densities; this is discussed later.

Having found the game optimal controls for the last stage, it is now possible to use the principle of optimality of the theory of dynamic programming to find the game optimal control strategies at stage 2, 3, ···, N.

In other words, the game optimal control strategies are chosen to optimize the payoff resulting from the application of controls at stages 1 and 2. Thus controls are applied at stage 2 which have an effect on the payoff and serve to change the state. Whatever the state resulting (and whatever the actual observations occurring at stage 1), game optimal control strategies $\bar{u}_1$ and $\bar{v}_1$ will be used. Symbolically, this is written

$$J_2 = \min_{u_2} \max_{v_2} \left[\iint \left\{ ||z_1||^2_{c_2} + ||u_2||^2_{d_2} - ||v_2||^2_{e_1} \right\} \times p(z_1, u_2, v_2)\, d(z_1, u_2, v_2) + J_1 \right] \tag{55}$$

Analogous to (11)

$$p(z_1, u_2, v_2) = \int p(z_1 | z_2, u_2, v_2)\, p(z_2, u_2, v_2)\, dz_2 \tag{56}$$

Substituting (56) into (55) and integrating over z_1 yield

$$J_2 = E\left\{||\lambda_2||^2_{c_2}\right\} + \min_{u_2}\max_{v_2}\left[\iint\left\{||k_2 z_2 + a_2 u_2 + b_2 v_2||^2_{c_2} + ||u_2||^2_{d_2} - ||v_2||^2_{e_2}\right\}\right.$$

$$\left.\times\ p(z_2, u_2, v_2)\ d(z_2, u_2, v_2) + J_1\right] \tag{57}$$

Consider J_1, as given by (18), where $\overline{u}_1$ and $\overline{v}_1$ are given by (47) and (48). Since

$$p\left(z_1, x^1, y^1, u^2, v^2\right) = \int p\left(z_1, z_2, x^1, y^1, u^2, v^2\right) dz_2 \tag{58}$$

and

$$p\left(z_1, z_2, x^1, y^1, u^2, v^2\right) = p\left(z_1, x_1, y_1 \mid z_2, x^2, y^2, u^2, v^2\right)$$

$$\times\ p\left(z_2, x^2, y^2, u^2, v^2\right) \tag{59}$$

where

$$p\left(z_1, x_1, y_1 \mid z_2, x^2, y^2, u^2, v^2\right) = p\left(x_1, y_1 \mid z_1, z_2, x^2, y^2, u^2, v^2\right)$$

$$\times\ p\left(z_1 \mid z_2, x^2, y^2, u^2, v^2\right) \tag{60}$$

From the definition of the manner in which observations are made, (2) and (3), and the independence of η_1 and ξ_1, it must be that

$$p\left(x_1, y_1 \mid z_1, z_2, x^2, y^2, u^2, v^2\right) = p(x_1 \mid z_1)\ p(y_1 \mid z_1) \tag{61}$$

Also, from the definition of how the state evolves, (1), it follows that

$$p\left(z_1 \mid z_2, x^2, y^2, u^2, v^2\right) = p(z_1 \mid z_2, u_2, v_2) \tag{62}$$

Substituting (58) through (62) back into (18) yields

$$J_1 = E\left\{\|\lambda_1\|^2_{c_1}\right\} + \int \gamma_1\left(z_2, x^2, y^2, u^2, v^2\right)$$

$$\times\, p\left(z_2, x^2, y^2, u^2, v^2\right) d\left(z_2, x^2, y^2, u^2, v^2\right) \tag{63}$$

where

$$\gamma_1\left(z_2, x^2, y^2, u^2, v^2\right) = \iint \left\{\|k_1 z_1 + a_1 \bar{u}_1 + b_1 \bar{v}_1\|^2_{c_1} + \|u_1\|^2_{d_1} - \|v_1\|^2_{e_1}\right\}$$

$$\times\, p(x_1 \mid z_1)\, dx_1 \;\; p(y_1 \mid z_1)\, dy_1$$

$$\times\, p\left(z_1 \mid z_2, u_2, v_2\right) dz_1 \tag{64}$$

Noting that

$$p\left(z_2, u_2, v_2\right) = \int p\left(z_2, x^2, y^2, u^2, v^2\right) d(x^2, y^2, u^3, v^3) \tag{65}$$

Equation (55) can be rewritten as

$$J_2 = E\left\{\|\lambda_1\|^2_{c_1} + \|\lambda_2\|^2_{c_2}\right\} + \min_{u_2} \max_{v_2} \iint \left\{\|k_2 z_2 + a_2 u_2 + b_2 v_2\|^2_{c_2} + \|u_2\|^2_{d_2}\right.$$

$$\left. - \|v_2\|^2_{e_2} + \gamma_1\left(z_2, x^2, y^2, u^2, v^2\right)\right\} p\left(z_2, x^2, y^2, u^2, v^2\right)$$

$$\times\, d\left(z_2, x^2, y^2, u^2, v^2\right) \tag{66}$$

To find the game optimal control strategies at stage 2, the same procedure used to find the game optimal controls at stage 1 is applied. That is, assume pure control strategies exist, as in (12). Defining

$$I_2(\tilde{u}_2, \tilde{v}_2) = E\left\{||\lambda_1||^2_{c_1} + ||\lambda_2||^2_{c_2}\right\} + \int\left\{||k_2 z_2 + a_2\tilde{u}_2 + b_2\tilde{v}_2||^2_{c_2} + ||\tilde{u}_2||^2_{d_2}\right.$$

$$\left. - ||\tilde{v}_2||^2_{e_2} + \gamma_1\left(z_2, x^2, y^2, \tilde{u}_2, u^3, \tilde{v}_2, v^3\right)\right\} p\left(z_2, x^2, y^2, u^3, v^3\right)$$

$$\times\, d\left(z_2, x^2, y^2, u^3, v^3\right) \tag{67}$$

analogous to (19), use the saddle point condition in both directions

$$I_2\left(\bar{u}_2 + \epsilon\delta(x^2, u^3)\, \bar{v}_2\right) - J_2 \geq 0 \tag{68}$$

$$I_2\left(\bar{u}_2, \bar{v}_2 + \epsilon\Delta(y^2, v^3)\right) - J_2 \leq 0 \tag{69}$$

Linear transformations similar to those given in (31) through (41) are used to generate a pair of simultaneous linear operator equations which are then solved simultaneously to find $\bar{u}_2$ and $\bar{v}_2$.

At this point, the procedure has become perfectly general. The procedure at any stage is then given by

$$J_i = E\left\{\sum_{j=1}^{i} ||\lambda_j||^2_{c_j}\right\} + \min_{u_i}\max_{v_i} \int\left\{||k_i z_i + a_i u_i + b_i v_i||^2_{c_i} + ||u_i||^2_{d_i}\right.$$

$$\left. - ||v_i||^2_{e_i} + \gamma_{i-1}\left(z_i, x^i, y^i, u^i, v^i\right)\right\} p\left(z_i, x^i, y^i, u^i, v^i\right)$$

$$\times\, d\left(z_i, x^i, y^i, u^i, v^i\right) \tag{70}$$

where

$$\gamma_{i-1}\left(z_i, x^i, y^i, u^i, v^i\right) = \int\left\{ \|k_{i-1} z_{i-1} + a_{i-1}\bar{u}_{i-1} + b_{i-1}\bar{v}_{i-1}\|^2_{c_{i-1}} + \|\bar{u}_{i-1}\|^2_{d_{i-1}} - \|\bar{v}_{i-1}\|^2_{e_{i-1}} + \gamma_{i-2}\left(z_{i-1}, x^{i-1}, y^{i-1}, \bar{u}_{i-1}, u^i, \bar{v}_{i-1}, v^i\right)\right\} \times p\left(z_{i-1}, x^{i-1}, y^{i-1}, u^i, v^i\right) d\left(z_{i-1}, x^{i-1}, y^{i-1}, u^i, v^i\right) \tag{71}$$

At each point the saddle point conditions

$$I_i\left(\bar{u}_i + \epsilon\delta(x^i, u^{i+1}),\ \bar{v}_i\right) - J_i \geq 0 \tag{72}$$

$$I_i\left(\bar{u}_i,\ \bar{v}_i + \epsilon\Delta(y^i, v^{i-1})\right) - J_i \leq 0 \tag{73}$$

are employed to generate the necessary conditions for game optimal strategies.

While the outline of the optimization problem is straightforward, the actual evaluation of the strategies, even for the easiest case of Gaussian random variables, is extremely tedious.

IV. GENERATION OF THE REQUIRED PROBABILITY DENSITY FUNCTIONS

Some of the required probability density functions are quite easy to express.[37] For example, from (2) and (3), using the notation discussed in Section II,

$$p(x_i \mid z_i) = \|x_i - G_i z_i\|^2_{\sigma_{\eta_i}^{-2}} \tag{74}$$

$$p(y_i | z_i) = \|y_i - H_i z_i\|^2_{\sigma_{\xi_i}^{-2}} \tag{75}$$

Also, from (1),

$$p(z_i | z_{i+1}, u_{i+1}, v_{i+1}) = \|z_i - k_{i+1} z_{i+1} - a_{i+1} u_{i+1} - b_{i+1} v_{i+1}\|^2_{\sigma_{\lambda_{i+1}}^{-2}} \tag{76}$$

The remaining conditional probabilities can be found recursively by using the chain rule for conditional densities in combination with Bayes's rule.[37]

An auxiliary conditional density function is first found

$$p\left(z_i, x^i, y^i | u^{i+1}, v^{i+1}\right) = \int p\left(z_i, z_{i+1}, x^i, y^i | u^{i+1}, v^{i+1}\right) dz_{i+1}$$

$$= \int p\left(z_i, x_i, y_i | z_{i+1}, x^{i+1}, y^{i+1}, u^{i+1}, v^{i+1}\right)$$

$$\times\, p\left(z_{i+1}, x^{i+1}, y^{i+1} | u^{i+1}, y^{i+1}\right) dz_{i+1} \tag{77}$$

But

$$p\left(z_i, x_i, y_i | z_{i+1}, x^{i+1}, y^{i+1}, u^{i+1}, v^{i+1}\right) = p\left(x_i, y_i | z_i, z_{i+1}, x^{i+1}, y^{i+1}, u^{i+1}, y^{i+1}\right)$$

$$\times\, p\left(z_i | z_{i+1}, x^{i+1}, y^{i+1}, u^{i+1}, v^{i+1}\right)$$

$$= p(x_i | z_i)\, p(y_i | z_i)$$

$$\times\, p(z_i | z_{i+1}, u_{i+1}, v_{i+1}) \tag{78}$$

where (74) through (76) are the justification for saying that the conditional densities of x_i and y_i, given z_i, depend on nothing else and that the conditional density for z_i, given z_{i+1}, u_{i+1}, and v_{i+1} is not changed if more information is available.

Also, it is easy to see that

$$p\left(z_{i+1}, x^{i+1}, y^{i+1} \mid u^{i+1}, v^{i+1}\right) = p\left(z_{i+1}, x^{i+1}, y^{i+1} \mid u^{i+2}, v^{i+2}\right) \quad (79)$$

since the value of a control chosen at the i+1 st stage can yield no information concerning either the state or an observation of the state. It must be reiterated that these conditional densities do not involve any assumptions concerning the optimality of controls chosen in the past. In particular, for purposes of estimation — which is what the conditional densities actually represent — no assumptions concerning strategies are required. When the probability density of one or more random variables is conditioned on one or more values of another variable, these conditioning variables enter the density function only as specific values. In this case, it means that one need only specify a set, any set, of values for u^{i+1} and v^{i+1} and then evaluate $p\left(z_{i+1}, x^{i+1}, y^{i+1} \mid u^{i+1}, v^{i+1}\right)$. When this is done, it can be seen that the values for u_{i+1} and v_{i+1} do not appear in the conditional density function.

Making use of (77) and (78) allows the rewriting of (77) as

$$p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right) = \int p(x_i \mid z_i)\, p(y_i \mid z_i)\, p\left(z_i \mid z_{i+1}, u_{i+1}, v_{i+1}\right) \times p\left(z_{i+1}, x^{i+1}, y^{i+1} \mid u^{i+2}, v^{i+2}\right) dz_{i+1} \quad (80)$$

which is the desired recursion relationship for $p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right)$. To start (80) off, note that

$$p\left(z_N, x^N, y^N \mid u^{N+1}, v^{N+1}\right) = p(z_N, x_N, y_N)$$

$$= p(x_N, y_N \mid z_N)\, p(z_N)$$

$$= p(x_N \mid z_N)\, p(y_N \mid z_N)\, p(z_N) \qquad (81)$$

where the first two conditional densities are given by (74) and (75) and the last one is given by (9). Thus it is possible to compute $p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right)$ for i + 1, 2, $\cdots$, N.

The required conditional probability densities are then:

$$p\left(z_i \mid x^i, y^i, u^{i+1}, v^{i+1}\right) = \frac{p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right)}{\int p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right) dz_i} \qquad (82)$$

$$p\left(y^i \mid x^i, u^{i+1}, v^{i+1}\right) = \frac{\int p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right) dz_i}{\int p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right) d\left(z_i, y^i\right)} \qquad (83)$$

$$p\left(x^i \mid y^i, u^{i+1}, v^{i+1}\right) = \frac{\int p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right) dz_i}{\int p\left(z_i, x^i, y^i \mid u^{i+1}, v^{i+1}\right) d\left(z_i, x^i\right)} \qquad (84)$$

The last two sets of conditional densities – $p(v^i \mid x^{i-1}, u^i)$ and $p(u^i \mid y^{i-1}, v^i)$ – must be handled somewhat differently.

$$p(v^i | x^{i-1}, u^i) = \frac{p(x_{i-1}, v_i, v^{i+1} | x^i, u^i)}{\int p(x_{i-1}, v_i, v^{i+1} | x^i, u^i)\, dv^i} \tag{85}$$

where

$$p(x_{i-1}, v_i, v^{i+1} | x^i, u^i) = p(x_{i-1}, v_i | x^i, u^i, v^{i+1})\, p(v^{i+1} | x^i, u^i) \tag{86}$$

Looking first at the second conditional probability density of the right side of (86), it is clear that

$$p(v^{i+1} | x^i, u^i) = p(v^{i+1} | x^i, u^{i+1}) \tag{87}$$

since knowledge of the value of a later control, for either player, can have no effect, can yield no information, on the estimate of the value for an earlier control unless there is some a priori known functional or statistical relationship.

The first conditional probability density on the right side of (86) can be written as

$$p(x_{i-1}, v_i | x^i, u^i, v^{i+1}) = p(x_{i-1} | x^i, u^i, v^i)\, p(v_i | x^i, u^i, v^{i+1}) \tag{88}$$

The first conditional density on the right side of (88) presents no real conceptual difficulties. It is generated by

$$p(x_{i-1} | x^i, u^i, v^i) = \frac{\int p(z_{i-1}, x^{i-1}, y^{i-1} | u^i, v^i)\, d(z_{i-1}, y^{i-1})}{\int p(z_{i-1}, x^{i-1}, y^{i-1} | u^i, v^i)\, d(z_{i-1}, x_{i-1}, y^{i-1})} \tag{89}$$

The second conditional density on the right side of (88) is not as straightforward. In fact,

$$p\left(v_i \mid x^i, u^i, v^{i+1}\right) = p(v_i) \tag{90}$$

To understand what is meant by (90), a clear understanding of the principle of optimality is required. Roughly, it states that, no matter what has occurred in the past, the best that can be done is to choose the controls in an optimal fashion in the future. With regard to the present problem, it means that a player need not (actually should not) assume that his opponent has used an optimal strategy or even that his opponent has used any nonrandom strategy whatsoever. Naturally, each player has a complete record of his own observations and past controls, but, even were he to be given a complete list of his opponent's control values, v^{i+1}, in the absence of any strategy which relates the opponent's control strategy to his (the opponent's) observations or to any other set of data, they can provide no hint as to what v_i will be. The only information that the opposing player can count on is data concerning physical bounds on the magnitude of the control available at the i^{th} stage. Accordingly, the a priori probability density for v_i is actually a uniform distribution over the physical limits known to exist. It should be stressed that this does not mean that one player believes that his opponent should have or would have chosen his i^{th} control from a uniform distribution; rather, it reflects the very limited knowledge available to a player about his opponent's real choice. It is merely the best, reliable information present.

Because this problem is addressed to Gaussian random variables, it makes sense to approximate the uniform distribution over a bounded set of values by a Gaussian distribution over an infinite set of values. A reasonable choice would be one with the same mean as the uniform distribution and with a variance such that the bounds of the uniform distribution are equal to plus and minus one standard deviation of the Gaussian distribution. Such a choice yields a relatively constant probability density function over the bounds of the uniform distribution. (The choice of a Gaussian random variable, instead of the uniform, is done only for the convenience associated with them. Theoretically, there is no reason why the uniform distribution should not be used.)

If no information as to capability is available, then the obvious choice for (90) is a density whose variance is, in the limit, infinite. Working with such variances leads to no difficulties.

Substituting (86), (87), (88) and (90) back into (85) leads to the required recursive formulation:

$$p\left(v^i \mid x^{i-1}, u^i\right) = \frac{p(v_i)\, p\left(x_{i-1} \mid x^i, u^i, v^i\right) p\left(v^{i+1} \mid x^i, u^{i+1}\right)}{\int p(v_i)\, p\left(x_{i-1} \mid x^i, u^i, v^i\right) p\left(v^{i+1} \mid x^i, u^{i+1}\right) dv^i} \tag{91}$$

where (89) is used to compute $p\left(x_{i-1} \mid x^i, u^i, v^i\right)$. To start (91), set

$$p\left(v^{N+1} \mid x^N, u^{N+1}\right) = p\left(v_{N+1} \mid x_N, u_{N+1}\right) = \delta\left(v_{N+1}\right) \tag{92}$$

where $\delta(v_{N+1})$ is the Dirac delta function.

5. EXAMPLE 1

The amount of work presently involved in actually solving an N stage stochastic game is tremendous. For this reason, example 1 involves only a two-stage scalar game since this serves to illustrate completely how the N stage vector game is solved.

All random variables are assumed to be Gaussian. Both players know the mean and variance of each random variable. Denoting each distribution by $N(\mu, \sigma^2)$, where μ is the mean and σ^2 the variance, the required a priori probability densities are

$$z_N : N\left(m, \sigma_z^2\right) \tag{93}$$

$$\lambda_i : N\left(0, \sigma_{\lambda_i}^2\right) \tag{94}$$

$$\eta_i : N\left(0, \sigma_{\eta_i}^2\right) \tag{95}$$

$$\xi_i : N\left(0, \sigma_{\xi_i}^2\right) \tag{96}$$

$$v_2 : N\left(\breve{v}_2, \sigma_{v_2}^2\right) \tag{97}$$

$$u_2 : N\left(\breve{u}_2, \sigma_{u_2}^2\right) \tag{98}$$

($\breve{v}_2$ and $\breve{u}_2$ here are merely the a priori means of the control capability and, for most problems, would be zero.) Both players know (93) through (98).

Using (93) through (98) and the formulas developed in Section IV, it is a straightforward task to generate the required conditional probability densities.

System dynamics and observations are given by (1), (2), and (3) with

$$G_i = H_i = 1 \tag{99}$$

To write out the conditional densities required to evaluate the transformations, a number of auxiliary variables are defined:

$$A = \frac{1}{\sigma^2_{\eta_2}} + \frac{1}{\sigma^2_{\xi_2}} + \frac{1}{\sigma^2_{z}} + \frac{k_2^2}{\sigma^2_{\lambda_2}} \tag{100}$$

$$D = \frac{1}{\sigma^2_{\lambda_2}} \tag{101}$$

$$E = \frac{1}{\sigma^2_{\eta_2}} \tag{102}$$

$$F = \frac{1}{\sigma^2_{\xi_2}} \tag{103}$$

$$G = \frac{k_2}{\sigma^2_{\lambda_2}} \tag{104}$$

$$L = \frac{1}{\sigma^2_{\eta_1}} \tag{105}$$

$$M = \frac{1}{\sigma^2_{\xi_1}} \tag{106}$$

$$N = \frac{m}{\sigma^2_z} \tag{107}$$

$$O = \frac{k_2 a_2}{\sigma^2_{\lambda_2}} \tag{108}$$

$$P = \frac{k_2 b_2}{\sigma^2_{\lambda_2}} \tag{109}$$

$$Q = \frac{a_2}{\sigma^2_{\lambda_2}} \tag{110}$$

$$R = \frac{b_2}{\sigma^2_{\dot{\lambda}_2}} \tag{111}$$

$$S = \frac{1}{\sigma^2_{v_2}} \tag{112}$$

$$T = \frac{1}{\sigma^2_{u_2}} \tag{113}$$

The first conditional density required is $p(z_1 | x^1, y^1, u_2, v_2)$, which is needed to evaluate T_{11} (or T_{16}) as given in (31).

$$T_{11} : p\left(z_1 | x^1, y^1, u_2, v_2\right) = -\frac{G^2 - A(L+D+M)}{A}$$

$$\times \left\{ z_1 + \frac{EGx_2 + FGy_2 + ALx_1 + AMy_1 + (AQ-OG)\,u_2 + (AR-GP)\,v_2 + GN}{G^2 - A(L+D+M)} \right\}^2 \tag{114}$$

or, making the obvious substitutions,

$$T_{11} : p\left(z_1 | x^1, y^1, u_2, v_2\right) = -\frac{G^2 - A(L+D+M)}{A}$$

$$\times \left\{ z_1 - \alpha_1 x_2 - \alpha_2 y_2 - \alpha_3 x_1 - \alpha_4 y_4 - \alpha_5 u_2 - \alpha_6 v_2 - \alpha_7 \right\}^2 \tag{115}$$

The evaluation of T_{12} is simplified by considering it to be the product of two other transformations, T_{12}^1 and T_{12}^2. Since

$$T_{12} : p\left(y^1 | x^1, u_2, v_2\right) = p\left(y_1 | x^1, y_2, u_2, v_2\right)\, p\left(y_2 | x^1, u_2, v_2\right) \tag{116}$$

it follows that

$$T_{12} = T_{12}^2\, T_{12}^1 \tag{117}$$

where

$$\beta\left(x^1, y_2, u_2, v_2\right) = T_{12}^{1}\alpha(y_1) = \int \alpha(y_1)\, p\left(y_1 \mid x^1, y_2, u_2, v_2\right) dy_1 \tag{121}$$

$$\beta\left(x^1, u_2, v_2\right) = T_{12}^{2}\alpha(y_2) = \int \alpha(y_2)\, p\left(y_2 \mid x^1, u_2, v_2\right) dy_2 \tag{122}$$

$$T_{12}^{1} : L_2\left[-\infty, \infty; p\left(y_1 \mid x^1, y_2, u_2, v_2\right)\right] \rightarrow L_2\left[-\infty, \infty; p\left(y_2 \mid x^1, u_2, v_2\right)\right] \tag{123}$$

$$T_{12}^{2} : L_2\left[-\infty, \infty; p\left(y_2 \mid x^1, u_2, v_2\right)\right] \rightarrow L_2\left[-\infty, \infty; p\left(v_2 \mid x^1, u_2\right)\right] \tag{124}$$

$$T_{12}^{1} : p\left(y_1 \mid x^1, y_2, u_2, v_2\right) = \frac{M\left[G^2 - A(L+D)\right]}{G^2 - A(L+D+M)}$$

$$\times \left\{ y_1 + \frac{EGx_2 + FGy_2 + ALx_1 + (AQ-DG)\,u_2 + (AR-GP)\,v_2 + GN}{G^2 - A(L+D)} \right\}^2 \tag{125}$$

or

$$T_{12}^{1} : p\left(y_1 \mid x^1, y_2, u_2, v_2\right) = \frac{M\left[G^2 - A(L+D)\right]}{G^2 - A(L+D+M)}$$

$$\times \left\{ y_1 - \alpha_8 x_2 - \alpha_9 y_2 - \alpha_{10} x_1 - \alpha_{11} u_2 - \alpha_{12} v_2 - \alpha_{13} \right\}^2 \tag{126}$$

$$T_{12}^{2} : p\left(y_2 \mid x^1, u_2\right) = \frac{F\left[G^2 + (F-A)(L+D)\right]}{G^2 - A(L+D)}$$

$$\times \left\{ y_2 + \frac{E(L+D)\,x_2 + LGx_1 + \left[GQ - (L+D)\,O\right]u_2 + \left[GR - (L+D)\,P\right]v_2 + (L+D)N}{G^2 + (F-A)(L+D)} \right\}^2 \tag{127}$$

or

$$T_{12}^{2}: p\left(y_2 \mid x^1, u_2\right) = \frac{F\left[G^2+(F-A)(L+D)\right]}{G^2 - A(L+D)}$$

$$\times \left\{ y_2 - \alpha_{14} x_2 - \alpha_{15} x_1 - \alpha_{16} u_2 - \alpha_{17} v_2 - \alpha_{18} \right\}^2 \qquad (118)$$

$$T_{13}^{2}: p\left(v_2 \mid x^1, u_2\right) = \frac{S\left[G^2+(F-A)(L+D)\right]+b_2^2 L\left[G^2+(F-A)\right]}{G^2+(F-A)(L+D)} \left\{ v_2 - \left(- LEPx_2 \right.\right.$$

$$-b_2 L\left[G^2+(F-A)\right] x_1 - a_2 b_2 L\left[G^2+(F-A)\right] u_2$$

$$+S\left[G^2+(F-A)(L+D)\right] \tilde{v}_2$$

$$\left.\left. -b_2 LGN\right) \Big/ \left(S\left[G^2+(F-A)(L+D)\right]+b_2^2 L\left[G^2+(F-A)\right]\right) \right\}^2$$

(119)

or

$$T_{13}: p\left(v_2 \mid x^1, u_2\right) = \frac{S\left[G^2+(F-A)(L+D)\right]+b_2^2 L\left[G^2+(F-A) D\right]}{G^2+(F-A)(L+D)}$$

$$\times \left\{ v_2 - \alpha_{19} x_2 - \alpha_{20} x_1 - \alpha_{21} u_2 - \alpha_{22} \right\}^2 \qquad (120)$$

Defining T_{14}^{1} and T_{14}^{2} analogously to (117)

$$T^{1}_{14}: p\left(x_1 \mid x_2, y^1, u_2, v_2\right) = \frac{L\left[G^2 - A(M+D)\right]}{G^2 - A(L+D+M)}$$

$$\times \left\{ x_1 + \frac{EGx_2 + FGy_2 + AMy_1 + (AQ-OG)u_2 + (AR-GP)u_2 + GN}{G^2 - A(M+D)} \right\}^2 \tag{128}$$

or

$$T^{1}_{14}: p\left(x_1 \mid x_2, y^1, u_2, v_2\right) = \frac{L\left[G^2 - A(M+D)\right]}{G^2 - A(L+D+M)}$$

$$\times \left\{ x_1 - \beta_1 x_2 - \beta_2 y_2 - \beta_3 y_1 - \beta_4 u_2 - \beta_5 v_1 - \beta_6 \right\}^2 \tag{129}$$

$$T^{2}_{14}: p\left(x_2 \mid y^1, u_2, v_2\right) = \frac{E\left[G^2 + (E-A)(M+D)\right]}{G^2 - A(M+D)}$$

$$\times \left\{ x_2 + \frac{F(M+D)y_2 + MGy_1 + \left[GQ - (M+D)O\right]u_2 + \left[GR - (M+D)P\right]v_2 + (M+D)N}{G^2 + (E-A)(M+D)} \right\}^2 \tag{130}$$

or

$$T^{2}_{14}: p\left(x_2 \mid y^1, u_2, v_2\right) = \frac{E\left[G^2 + (E-A)(M+D)\right]}{G^2 - A(M+D)}$$

$$\times \left\{ x_2 - \beta_7 y_2 - \beta_8 y_1 - \beta_9 u_2 - \beta_{10} v_2 - \beta_{11} \right\}^2 \tag{131}$$

and, finally,

$$T_{15}: p\left(u_2 \mid y^1, v_2\right) = \frac{T\left[G^2+(E-A)(M+D)\right]+a_2^2 M\left[G^2+(E-A)\right]}{G^2+(E-A)(M+D)} \left\{u_2 - \left(-MFOy_2 \right.\right.$$

$$-a_2 M\left[G^2-(E-A)\right] y_1 - a_2 b_2 M\left[G^2+(E-A)\right] v_2 + T\left[G^2+(E-A)(M+D)\right] \tilde{u}_2$$

$$\left.\left. -a_2 MGN\right) \Big/ \left(T\left[G^2+(E-A)(M+D)\right] + a_2^2 M\left[G^2+(E-A)\right]\right)\right\}^2$$

(132)

or

$$T_{15}: p\left(u_2 \mid y^1, v_2\right) = \frac{T\left[G^2+(E-A)(M+D)\right] + a_2^2 M\left[G^2+(E-A)\right]}{G^2+(E-A)(M+D)}$$

$$\times \left\{x_2 - \beta_{12} y_2 - \beta_{13} y_1 - \beta_{14} v_2 - \beta_{15}\right\}^2 \tag{133}$$

It is now possible to evaluate $\bar{u}_1$ and $\bar{v}_1$. To illustrate how this is done, consider $\bar{u}_1$ as given by (47).

$$T_{13} T_{12} \left[1 + \frac{b_1^2 c_1}{e_1 - b_1^2 c_1} T_{15} T_{14}\right] T_{11} z_1 = \theta_5 x_2 + \theta_6 x_1 + \theta_7 u_2 + \theta_8 \tag{134}$$

where

$$\theta_5 = \alpha_1 + \gamma_4 \alpha_8 + (\gamma_2 + \gamma_4 \alpha_9)\alpha_{14} + \left[\gamma_6 + \gamma_4 \alpha_{12} + (\gamma_2 + \gamma_4 \alpha_9)\alpha_{17}\right]\alpha_{19} \tag{135}$$

$$\theta_6 = \alpha_3 + \gamma_4 \alpha_{10} + (\gamma_2 + \gamma_4 \alpha_9)\alpha_{15} + \left[\gamma_6 + \gamma_4 \alpha_{12} + (\gamma_2 + \gamma_4 \alpha_9)\alpha_{17}\right]\alpha_{20} \tag{136}$$

$$\theta_7 = \alpha_5+\gamma_4\alpha_{11}+(\gamma_2+\gamma_4\alpha_9)\alpha_{16}+\left[\gamma_6+\gamma_4\alpha_{12}+(\gamma_2+\gamma_4\alpha_9)\alpha_{17}\right]\alpha_{21} \qquad (146)$$

$$\theta_8 = \gamma_7+\gamma_4\alpha_{13}+(\gamma_2+\gamma_4\alpha_9)\alpha_{18}+\left[\gamma_6+\gamma_4\alpha_{12}(\gamma_2+\gamma_4\alpha_9)\alpha_{17}\right]\alpha_{22} \qquad (147)$$

$$\gamma_2 = \alpha_2 + \frac{b_1^2\, c_1}{e_1 - b_1^2\, c_1}\,\theta_1 \qquad (148)$$

$$\gamma_4 = \alpha_4 + \frac{b_1^2\, c_1}{e_1 - b_1^2\, c_1}\,\theta_2 \qquad (149)$$

$$\gamma_6 = \alpha_6 + \frac{b_1^2\, c_1}{e_1 - b_1^2\, c_1}\,\theta_3 \qquad (150)$$

$$\gamma_7 = \alpha_7 + \theta_4 \qquad (151)$$

$$\theta_1 = \alpha_2+\alpha_3\beta_2+(\alpha_1+\alpha_3\beta_1)\,\beta_7+\left[\alpha_5+\alpha_3\beta_4+(\alpha_1+\alpha_3\beta_1)\,\beta_9\right]\beta_{12} \qquad (152)$$

$$\theta_2 = \alpha_4+\alpha_3\beta_3+(\alpha_1+\alpha_3\beta_1)\,\beta_8+\left[\alpha_5+\alpha_3\beta_4+(\alpha_1+\alpha_3\beta_1)\alpha_9\right]\beta_{13} \qquad (153)$$

$$\theta_3 = \alpha_6+\alpha_3\beta_5+(\alpha_1+\alpha_3\beta_1)\,\beta_{10}+\left[\alpha_5+\alpha_3\beta_4+(\alpha_1+\alpha_3\beta_1)\alpha_9\right]\beta_{14} \qquad (154)$$

$$\theta_4 = \alpha_7+\alpha_3\beta_6+(\alpha_1+\alpha_3\beta_1)\,\beta_{11}+\left[\alpha_5+\alpha_3\beta_4+(\alpha_1+\alpha_3\beta_1)\,\beta_9\right]\beta_{15} \qquad (155)$$

The next concern is with the elements of the expansion appearing in the Neumann series. Equation (134) indicates that the transformations operate on $\theta_5 x_2 + \theta_6 x_1 + \theta_7 u_2 + \theta_8$. Actually performing the indicated transformations leads to

$$T_{13}T_{12}T_{15}T_{14}(\theta_5 x_2+\theta_6 x_1+\theta_7 u_2+\theta_8) = (\theta_5\varphi_5+\theta_6\varphi_1+\theta_7\varphi_9)x_2$$

$$+(\theta_5\varphi_6+\theta_6\varphi_2+\theta_7\varphi_{10})x_1+(\theta_5\varphi_7+\theta_6\varphi_3+\theta_7\varphi_{11})u_2+\theta_5\varphi_8+\theta_6\varphi_4+\theta_7\varphi_{12}+\theta_8 \quad (137)$$

where

$$\varphi_1 = \theta_{10}\alpha_8+(\theta_9+\theta_{10}\alpha_9)\alpha_{14}+\left[\theta_{11}+\theta_{10}\alpha_{12}+(\theta_9+\theta_{10}\alpha_9)\alpha_{17}\right]\alpha_{19} \quad (138)$$

$$\varphi_2 = \theta_{10}\alpha_9+(\theta_9+\theta_{10}\alpha_9)\alpha_{15}+\left[\theta_{11}+\theta_{10}\alpha_{12}+(\theta_9+\theta_{10}\alpha_9)\alpha_{17}\right]\alpha_{20} \quad (139)$$

$$\varphi_3 = \theta_{10}\alpha_{11}+(\theta_9+\theta_{10}\alpha_9)\alpha_{16}+\left[\theta_{11}+\theta_{10}\alpha_{12}+(\theta_9+\theta_{10}\alpha_9)\alpha_{17}\right]\alpha_{21} \quad (140)$$

$$\varphi_4 = \theta_{12}+\theta_{10}\alpha_{13}+(\theta_9+\theta_{10}\alpha_9)\alpha_{18}+\left[\theta_{11}+\theta_{10}\alpha_{12}+(\theta_9+\theta_{10}\alpha_9)\alpha_{17}\right]\alpha_{22} \quad (141)$$

$$\theta_9 = \beta_2+\beta_1\beta_7+(\beta_4+\beta_1\beta_9)\beta_{12} \quad (142)$$

$$\theta_{10} = \beta_3+\beta_1\beta_8+(\beta_4+\beta_1\beta_9)\beta_{13} \quad (143)$$

$$\theta_{11} = \beta_5+\beta_1\beta_{10}+(\beta_4+\beta_1\beta_9)\beta_{14} \quad (144)$$

$$\theta_{12} = \beta_6+\beta_1\beta_{11}+(\beta_4+\beta_1\beta_9)\beta_{15} \quad (145)$$

$$\varphi_5 = \theta_{13} + \theta_{16}\alpha_{19} \tag{165}$$

$$\varphi_6 = \theta_{14} + \theta_{16}\alpha_{20} \tag{166}$$

$$\varphi_7 = \theta_{15} + \theta_{16}\alpha_{21} \tag{167}$$

$$\varphi_8 = \theta_{17} + \theta_{16}\alpha_{22} \tag{168}$$

$$\theta_{13} = (\beta_8 + \beta_9\beta_{13})\alpha_8 + \left[\beta_2 + \beta_9\beta_{12} + (\beta_8 + \beta_9\beta_{13})\alpha_9\right]\alpha_{14} \tag{169}$$

$$\theta_{14} = (\beta_8 + \beta_9\beta_{13})\alpha_{10} + \left[\beta_2 + \beta_9\beta_{12} + (\beta_8 + \beta_9\beta_{13})\alpha_9\right]\alpha_{15} \tag{170}$$

$$\theta_{15} = (\beta_8 + \beta_9\beta_{13})\alpha_{11} + \left[\beta_2 + \beta_9\beta_{12} + (\beta_8 + \beta_9\beta_{13})\alpha_9\right]\alpha_{16} \tag{171}$$

$$\theta_{16} = \beta_{10} + \beta_9\beta_{14} + (\beta_8 + \beta_9\beta_{13})\alpha_{12} + \left[\beta_2 + \beta_9\beta_{12} + (\beta_8 + \beta_9\beta_{13})\alpha_9\right]\alpha_{17} \tag{172}$$

$$\theta_{17} = \beta_{11} + \beta_9\beta_{15} + (\beta_8 + \beta_9\beta_{13})\alpha_{13} + \left[\beta_2 + \beta_9\beta_{12} + (\beta_8 + \beta_9\beta_{13})\alpha_9\right]\alpha_{18} \tag{173}$$

$$\varphi_9 = \beta_{13}\alpha_8 + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{14} + \left[\beta_{14} + \beta_{13}\alpha_{12} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{17}\right]\alpha_{19} \tag{174}$$

$$\varphi_{10} = \beta_{13}\alpha_{10} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{15} + \left[\beta_{14} + \beta_{13}\alpha_{12} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{17}\right]\alpha_{20} \tag{175}$$

$$\varphi_{11} = \beta_{13}\alpha_{11} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{16} + \left[\beta_{14} + \beta_{13}\alpha_{12} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{17}\right]\alpha_{21} \tag{176}$$

$$\varphi_{12} = \beta_{15} + \beta_{13}\alpha_{13} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{18} + \left[\beta_{14} + \beta_{13}\alpha_{12} + (\beta_{12} + \beta_{13}\alpha_9)\alpha_{17}\right]\alpha_{22} \tag{177}$$

Now define

$$\theta_5(i+1) = \theta_5(i)\varphi_5+\theta_6(i)\varphi_1+\theta_7(i)\varphi_9 \quad ; \quad \theta_5(o) = \theta_5 \tag{156}$$

$$\theta_6(i+1) = \theta_5(i)\varphi_6+\theta_6(i)\varphi_2+\theta_7(i)\varphi_{10} \quad ; \quad \theta_6(o) = \theta_6 \tag{157}$$

$$\theta_7(i+1) = \theta_5(i)\varphi_7+\theta_6(i)\varphi_3+\theta_7(i)\varphi_{11} \quad ; \quad \theta_7(o) = \theta_7 \tag{158}$$

$$\theta_8(i+1) = \theta_5(i)\varphi_8+\theta_6(i)\varphi_4+\theta_7(i)\varphi_{12} \quad ; \quad \theta_8(o) = \theta_8 \tag{159}$$

$$WT = \frac{a_1^2 b_1^2 c_1^2}{\left(a_1^2 c_1+d_1\right)\left(b_1^2 c_1-e_1\right)} T_{13}T_{12}T_{15}T_{14} \tag{160}$$

with

$$\| WT \| < 1 \tag{161}$$

and

$$\pi_1 = -\frac{a_1^2 c_1}{a_1^2 c_1+d_1} \sum_{i=o}^{\infty} W^i \theta_5(i) \tag{162}$$

$$\pi_2 = -\frac{a_1^2 c_1}{a_1^2 c_1+d_1} \sum_{i=o}^{\infty} W^i \theta_6(i) \tag{163}$$

$$\pi_3 = -\frac{a_1^2 c_1}{a_1^2 c_1+d_1} \sum_{i=o}^{\infty} W^i \theta_7(i) \tag{164}$$

$$\pi_4 = -\frac{a_1^2 c_1}{a_1^2 c_1 + d_1} \sum_{i=o}^{\infty} W^i \theta_8(i) \tag{178}$$

The game optimal strategy for the minimizing player is then given by

$$\bar{u}_1 = \pi_1 x_2 + \pi_2 x_1 + \pi_3 u_2 + \pi_4 \tag{179}$$

A similar evaluation of (48) leads to

$$\bar{v}_1 = \pi_5 y_2 + \pi_6 y_1 + \pi_7 v_2 + \pi_8 \tag{180}$$

At this point, it is seen that the game optimal strategies, under the assumptions given, are linear in the observations and past controls. Because of the complexity of the computations needed to achieve (179) and (180), there is little that can be seen in the way of structure beyond linearity.

Substituting (179) and (180) into (64) leads to

$$\begin{aligned} \gamma_1(z_2, x_2, y_2, u_2, v_2) = {} & \mu_o + c_1 \left\{ \mu_1 z_2 + \mu_2 x_2 + \mu_3 u_2 + \mu_4 y_2 + \mu_5 v_2 + \mu_6 \right\}^2 \\ & + d_1 \left\{ \mu_7 x_2 + \mu_8 z_2 + \mu_9 u_2 + \mu_{10} v_2 + \mu_{11} \right\}^2 - e_1 \left\{ \mu_{12} y_2 + \mu_{13} z_2 + \mu_{14} u_2 + \mu_{15} v_2 + \mu_{16} \right\}^2 \end{aligned} \tag{181}$$

where

$$\mu_o = (a_1^2 c_1 + d_1) \pi_2^2 \sigma_{\eta_1}^2 - (e_1 - b_1^2 c_1) \pi_6^2 \sigma_{\xi_1}^2 + \left[c_1 (k_1 + a_1 \pi_2 + b_1 \pi_6)^2 + d_1 \pi_2^2 - e_1 \pi_6^2 \right] \sigma_{\wedge_2}^2 \tag{183}$$

$$\mu_1 = (k_1 + a_1 \pi_2 + b_1 \pi_6)\, k_2 \tag{183}$$

$$\mu_2 = a_1 \pi_1 \tag{184}$$

$$\mu_3 = a_1 \pi_3 + a_2 (k_1 + a_1 \pi_2 + b_1 \pi_6) \tag{185}$$

$$\mu_4 = b_1 \pi_5 \tag{186}$$

$$\mu_5 = b_1 \pi_7 + b_2 (k_1 + a_1 \pi_2 + b_1 \pi_6) \tag{187}$$

$$\mu_6 = a_1 \pi_4 + b_1 \pi_8 \tag{188}$$

$$\mu_7 = \pi_1 \tag{189}$$

$$\mu_8 = k_2 \pi_2 \tag{190}$$

$$\mu_9 = \pi_3 + a_2 \pi_2 \tag{191}$$

$$\mu_{10} = b_2 \pi_2 \tag{192}$$

$$\mu_{11} = \pi_4 \tag{193}$$

$$\mu_{12} = \pi_5 \tag{194}$$

$$\mu_{13} = k_2 \pi_6 \tag{195}$$

$$\mu_{14} = a_2 \pi_6 \tag{203}$$

$$\mu_{15} = \pi_7 + b_2 \pi_6 \tag{204}$$

$$\mu_{16} = \pi_8 \tag{205}$$

The Value of the game, with two stages to go, is given by (66).

$$J_2 = c_1 \sigma_{\lambda_1}^2 + c_2 \sigma_{\lambda_2}^2 + \mu_o + \min_{u_2} \max_{v_2} \int \Big\{ c_2 (k_2 z_2 + a_2 u_2 + b_2 v_2)^2 + d_2 u_2^2 - e_2 v_2^2$$

$$+ c_1 (\mu_1 z_2 + \mu_2 x_2 + \mu_3 u_2 + \mu_4 y_2 + \mu_5 v_2 + \mu_6)^2$$

$$+ d_1 (\mu_7 x_2 + \mu_8 z_2 + \mu_9 u_2 + \mu_{10} v_2 + \mu_{11})^2$$

$$- e_1 (\mu_{12} y_2 + \mu_{13} z_2 + \mu_{14} u_2 + \mu_{15} v_2 + \mu_{16})^2 \Big\}$$

$$\times p(z_2, x_2, y_2, u_2, v_2)\, d(z_2, x_2, y_2, u_2, v_2) \tag{206}$$

The assumption of the existence of pure strategies

$$\bar{u}_2 = \bar{u}_2 (x_2) \tag{207}$$

$$\bar{v}_2 = \bar{v}_2 (y_2) \tag{208}$$

and the satisfaction of the saddlepoint condition, (68) and (69), lead to the following sufficient conditions for game optimal strategies:

$$(a_2c_2k_2+c_1\mu_1\mu_3+d_1\mu_8\mu_9-e_1\mu_{13}\mu_{14})T_{22}T_{21}z_2+(a_2^2c_2+d_2+c_1\mu_3^2+d_1\mu_9^2-e_1\mu_{14}^2)\bar{u}_2$$

$$+(a_2b_2c_2+c_1\mu_3\mu_5+d_1\mu_9\mu_{10}-e_1\mu_{14}\mu_{15})T_{22}\bar{v}_2+(c_1\mu_3\mu_2+d_1\mu_7\mu_9)x_2$$

$$+(c_1\mu_3\mu_4-e_1\mu_{12}\mu_{14})T_{22}y_2+c_1\mu_3\mu_6+d_1\mu_9\mu_{11}-e_1\mu_{14}\mu_{16} = 0 \qquad (196)$$

$$(b_2c_2k_2+c_1\mu_1\mu_4+d_1\mu_8\mu_{10}-e_1\mu_{13}\mu_{15})T_{24}T_{21}z_2+(a_2b_2c_2+c_1\mu_3\mu_5+d_1\mu_9\mu_{10}$$

$$-e_1\mu_{14}\mu_{15})T_{24}\mu_2-(e_2-b_2^2c_2-c_1\mu_5^2-d_1\mu_{10}^2+e_1\mu_{15}^2)\bar{v}_2$$

$$+(c_1\mu_2\mu_5+d_1\mu_7\mu_{10})T_{24}x_2+(c_1\mu_4\mu_5-e_1\mu_{12}\mu_{15})y_2$$

$$+c_1\mu_5\mu_6+d_1\mu_{10}\mu_{11}-e_1\mu_{15}\mu_{16} = 0 \qquad (197)$$

$$a_2^2c_2+d_2+c_1\mu_3^2+d_1\mu_9^2-e_1\mu_{14}^2 > 0 \qquad (198)$$

$$e_2-b_2^2c_2-c_1\mu_5^2-d_1\mu_{10}^2+e_1\mu_{15}^2 > 0 \qquad (199)$$

where

$$\beta(x_2, y_2) = T_{21}\alpha(z_2) = \int \alpha(z_2)\, p(z_2|x_2, y_2)\, dz_2 \qquad (200)$$

$$T_{21}: L_2\left[-\infty, \infty; p\left(z_2|x_2, y_2\right)\right] \to L_2\left[-\infty, \infty; p\left(y_2|x_2\right)\right] \qquad (201)$$

$$\beta(x_2) = T_{22}\alpha(y_2) = \int \alpha(y_2)\, p\left(y_2|x_2\right) dy_2 \qquad (202)$$

$$T_{22}: L_2\left[-\infty, \infty; p(y_2|x_2)\right] \rightarrow L_2\left[-\infty, \infty; p(x_2|y_2)\right] \qquad (210)$$

$$\beta_2(y_2) = T_{24}\alpha(x_2) = \int \alpha(x_2)\, p(x_2|y_2)\, dx_2 \qquad (211)$$

$$T_{24}: L_2\left[-\infty, \infty; p(x_2|y_2)\right] \rightarrow L_2\left[-\infty, \infty; p(y_2|x_2)\right] \qquad (212)$$

Equations (202) and (203) are solved simultaneously to yield the game optimal strategies. This is not done here since the operator algebra does not provide any fresh insights.

An interesting point can be made by comparing (203), (204), (53), and (54). Only system parameters at stage 1 entered into the latter, whereas system parameters at stages 1 and 2 appear in the former. Also, it appears that the variances of the observation noises show up in (203) and (204). In effect, this means that pure strategies can exist only where the dynamics allow them and when the observation noise variances make the observations meaningful. The exact dependence can be made clear only by actually evaluating (203) and (204).

6. EXAMPLE 2

Example 1 indicated that the game optimal controls used all information available to each player to determine the control strategy. In effect, the information was used to better define the state at each stage. While not obvious, the key piece of information is the assumed a priori distribution representing physical limitations on the opposing player's available control magnitude. These assumptions allow a player to generate an a priori estimate of the state at each stage, which is then combined with the current observation to produce the a posteriori estimate of the state.

What happens if the two players have no knowledge of their opponent's capabilities? In this case, the a priori control distribution may be taken as one which, in the limit, has infinite variance. The result is a particularly simple separation solution: the game optimal control strategy at each stage, except the N^{th} where an a priori distribution of the state is assumed available, is the deterministic control strategy with current observation taking the place of the true state. At the N^{th} stage, a more usual strategy (involving noise variances) is used.

To show that such is actually the case, a different set of linear transformations is used to define the necessary conditions. Instead of (25), consider the following:

$$\begin{aligned} p(z_1, x^1, y^1, u^2, v^2) &= p(y_1 | z_1, x^1, y_2, u_2, v_2) p(z_1, x^1, y_2, u_2, v_2) \\ &= p(y_1 | z_1) p(y_2, v_2 | z_1, x^1, u_2) p(z_1 | x^1, u_2) p(x^1, u_2) \\ &= p(y_1 | z_1) p(y_2 | z_1, x^1, u_2, v_2) p(v_2 | z_1, x^1, u_2) \\ &\quad \times p(z_1 | x^1, u_2) p(x_1, u_2) \end{aligned} \tag{209}$$

where the independence of the observation noise justifies the statement that $p(y_1 | z_1, x^1, y_2, u_2, v_2) = p(y_1 | z_1)$. Using this decomposition of the joint density does not lead to any nice characterization of the resulting conditional densities although it is equally valid. Its virtue, as is seen later, rests in the conditioning of z_1 upon x^1 and u_2, the information available to player II. In this case, the past is discarded and, because of the decomposition, the actual evaluation of the functional form of the control strategies becomes almost trivial.

Using (212), the control strategies are

$$\overline{u}_1 = -\frac{a_1 c_1 k_1}{a_1^2 c_1 + d_1}\left[1 - \frac{a_1^2 b_1^2 c_1^2}{\left(a_1^2 c_1 + d_1\right)\left(b_1^2 c_1 - e_1\right)} T_{14} T_{13} T_{12} T_{11} T_{18} T_{17} T_{16} T_{15}\right]^{-1} T_{14}$$

$$\times \left[1 + \frac{b_1^2 c_1}{e_1 - b_1^2 c_1} T_{13} T_{12} T_{11} T_{18}\right] z_1 \qquad (221)$$

$$\overline{v}_1 = \frac{b_1 c_1 k_1}{e_1 - b_1^2 c_1}\left[1 - \frac{a_1^2 b_1^2 c_1^2}{\left(a_1^2 c_1 + d_1\right)\left(b_1^2 c_1 - e_1\right)} T_{18} T_{17} T_{16} T_{15} T_{14} T_{13} T_{12} T_{11}\right]^{-1} T_{18}$$

$$\times \left[1 - \frac{a_1^2 c_1}{a_1^2 c_1 + d_1} T_{17} T_{16} T_{15} T_{14}\right] z_1 \qquad (222)$$

where

$$\dot{\beta}\left(z_1, x^1, u_2, v_2\right) = T_{11}\alpha(y_2) = \int \alpha(y_2)\, p\left(y_2 \mid z_1, x^1, u_2, v_2\right) dy_2 \qquad (223)$$

$$\beta\left(z_1, x^1, u_2\right) = T_{12}\alpha(v_2) = \int \alpha(v_2)\, p\left(v_2 \mid z_1, x^1, u_2\right) dv_2 \qquad (224)$$

$$\beta(z_1) = T_{13}\alpha(y_1) = \int \alpha(y_1)\, p\left(y_1 \mid z_1\right) dy_1 \tag{213}$$

$$\beta\left(x^1, u_2\right) = T_{14}\alpha(z_1) = \int \alpha(z_1)\, p\left(z_1 \mid x^1, u_2\right) dz_1 \tag{214}$$

$$\beta\left(z_1, y^1, u_2, v_2\right) = T_{15}\alpha(x_2) = \int \alpha(x_2)\, p\left(x_2 \mid z_1, y^1, u_2, v_2\right) dx_2 \tag{215}$$

$$\beta\left(z_1, y^1, v_2\right) = T_{16}\alpha(u_2) = \int \alpha(u_2)\, p\left(u_2 \mid z_1, y^1, v_2\right) du_2 \tag{216}$$

$$\beta(z_1) = T_{17}\alpha(x_1) = \int \alpha(x_1)\, p\left(x_1 \mid z_1\right) dx_1 \tag{217}$$

$$\beta\left(y^1, v_2\right) = T_{18}\alpha(z_1) = \int \alpha(z_1)\, p\left(z_1 \mid y^1, v_2\right) dz_1 \tag{218}$$

and all transformations have their domain and range in the appropriate Hilbert space. It is clear that these are not the same transformations defined in (31) through (42), although they are an equivalent set.

The required conditional densities are:

$$T_{11}: p\left(y_2 \mid z_1, x^1, u_2, v_2\right) = -\frac{F^2 - AF}{A}\left\{y_2 + \frac{Ex_2 + Gz_1 + N - Ou_2 - Pv_2}{F - A}\right\}^2 \tag{219}$$

$$T_{12}: p\left(v_2 \mid z_1, x^1, u_2\right) = \frac{b_2^2\left[G^2 + (F-A)D\right]}{G^2 + (F-A)(L+D)}$$

$$\times \left\{v_2 - \left[\frac{1}{b_2} z_1 - \frac{a_2}{b_2} u_2 + \frac{EP}{b_2^2\left[G^2 + (F-A)D\right]} x_2 + \frac{GN}{b_2\left[G^2 + (F-A)D\right]}\right]\right\}^2 \tag{220}$$

$$T_{13}: p(y_1 | z_1) = M\{y_1 - z_1\}^2 \tag{225}$$

$$T_{14}: p(z_1 | x^1, u_2) = L\{z_1 - x_1\}^2 \tag{226}$$

$$T_{15}: p(x_2 | z_1, y^1, u_2, v_2) = -\frac{E^2 - AE}{A}\left\{x_2 + \frac{Fy_2 + Gz_1 + N - Ou_2 - Pv_2}{E-A}\right\}^2 \tag{227}$$

$$T_{16}: p(u_2 | z_1, y^1, v_2) = \frac{a_2^2\left[G^2 + (E-A)D\right]}{G^2 + (E-A)(M+D)}\left\{u_2 - \left[\frac{1}{a_2} z_1 - \frac{b_2}{a_2} v_2 + \frac{FO}{a_2^2\left[G^2 + (E-A)D\right]} x_2 + \frac{GN}{a_2\left[G^2 + (E-A)D\right]}\right]\right\}^2 \tag{228}$$

$$T_{17}: p(x_1 | z_1) = L\{x_1 - z_1\}^2 \tag{229}$$

$$T_{18}: p(z_1 | y^1, v_2) = M\{z_1 - y_1\}^2 \tag{230}$$

The conditional densities for the state, (226) and (230), involve only the observation at stage 1 because the entire past has been lost, as it were, by the assumption of infinite variance on the opposing player's control at stage 2. In concept, this is similar to the way in which a Kalman filter is initialized.[39] The practical result is to place all weighting on the current observation. It is the best estimate of the state at that point.

Now consider (213). Using (223) through (230), it follows that

$$T_{14}\left[1+\frac{b_1^2 c_1}{e_1-b_1^2 e_1} T_{13}T_{12}T_{11}T_{18}\right] z_1 = \left[1+\frac{b_1^2 c_1}{e_1-b_1^2 c}\right] x_1 \tag{231}$$

$$T_{18}T_{17}T_{16}T_{15}T_{14}T_{13}T_{12}T_{11}\, z_1 = x_1 \tag{232}$$

so that

$$\bar{u}_1 = -\frac{a_1 c_1 e_1 k_1}{\left(a_1^2 c_1+d_1\right)\left(e_1-b_1^2 c_1\right)+a_1^2 b_1^2 c_1}\, x_1 \tag{233}$$

In the same manner, it is easy to show that

$$\bar{v}_1 = \frac{b_1 c_1 d_1 k_1}{\left(a_1^2 c_1+d_1\right)\left(e_1-b_1^2 c_1\right)+a_1^2 b_1^2 c_1}\, y_1 \tag{234}$$

Note that (235) and (234) are precisely the control strategies (detailed in Chapter 2) for the deterministic case. They could have been obtained using (47) and (48), with S and T set to zero, but it would have been difficult to obtain the results in closed form.

It is a straightforward task to use (64) to show that

$$J_2 = c_1\sigma_{\lambda_1}^2 + c_2\sigma_{\lambda_2}^2 + \frac{a_1^2 c_1^2 e_1^2 k_1^2 \left(a_1^2 c_1 + d_1\right)\sigma_{\eta_1}^2 - b_1^2 c_1^2 d_1^2 k_1^2 \left(e_1 - b_1^2 c_1\right)\sigma_{\xi_1}^2}{\left[\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1^2\right]^2}$$

$$+ \frac{c_1 d_1 e_1 k_1^2}{\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1}\,\sigma_{\lambda_2}^2 + \min_{u_2}\max_{v_2}\int \Bigg\{\Bigg[c_2$$

$$+ \frac{c_1 d_1 e_1 k_1^2}{\left(a_1^2 c_1 + d_1\right)\left(e_1 - b_1^2 c_1\right) + a_1^2 b_1^2 c_1^2}\Bigg]\left(k_2 z_2 + a_2 u_2 + b_2 v_2\right)^2 + d_2 u_2^2 - e_2 v_2^2\Bigg\}$$

$$\times\, p(z_2, u_2, v_2)\, d(z_2, u_2, v_2) \qquad (235)$$

The game optimal control strategies for (235) are easy to find and so are not derived here.

VII. CONCLUDING COMMENTS

A major factor in the solution of the multistage stochastic differential game is the shared knowledge of the two players. Both players know the value for all parameters of the dynamical equations and the payoff. All density functions are fully known to both sides. Given this type of structural knowledge, it should be clear that other types of strategies involving other information sets could as easily be used.

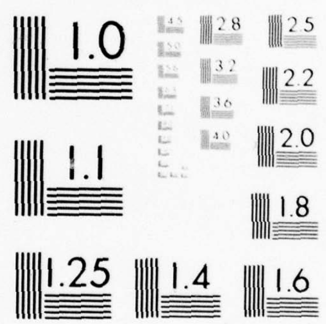
Admissible control strategies, other than those specified by (17), can be handled in a similar manner. The main difference, practically speaking, is in the form of the linear transformations that arise from a consideration of the necessary conditions. Some examples are considered in Chapter 4.

AD-A072 383 CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING A--ETC F/G 1/2
RANDOMIZED AND STOCHASTIC LINEAR, DISCRETE DIFFERENTIAL GAMES W--ETC(U)
JUL 78 K B BLEY, C T LEONDES F33615-77-C-3013

UNCLASSIFIED AFFDL-TR-78-87 NL

2 OF 2
AD
A072383

END
DATE
FILMED
9-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SECTION IV

SINGLE-STAGE STOCHASTIC DIFFERENTIAL GAMES

I. INTRODUCTION

Chapter 3 derived all the theory required to solve multistage stochastic differential games having pure strategies. This chapter is limited to examples of single-stage scalar games involving Gaussian random variables. Consequently, the theory already in hand is used.

This chapter has two purposes: 1) to show that the solutions to the single-stage game have a closed form solution (which may or may not be true for the multistage case), and 2) to exhibit the game optimal control strategies that result when different assumptions are made concerning the information available to each player (different admissible strategies).

All subscripts referring to the stage number are absent since there can be no ambiguity. Further, shorthand notation, which is obvious in context, is introduced as required for convenience.

Finally, in each of the following examples the maximizing player (player I) is assumed to have only noisy observations of the state.

If the $c\sigma_\lambda^2$ term is neglected, then

$$J = \min_u \max_v \int \left\{ c(kz+au+bv)^2 + du^2 - ev^2 \right\} p(z,x,y,u,v)\, d(z,x,y,u,v) \tag{1}$$

or, since

$$p(z, x, y, u, v) = p(u, v \mid z, x, y)\, p(z, x, y) \tag{2}$$

$$J = \int \left\{ c(kz+a\bar{u}+b\bar{v})^2 + d\bar{u}^2 - e\bar{v}^2 \right\} p(z, x, y)\, d(z, x, y) \tag{3}$$

where $\bar{u}$ and $\bar{v}$ are the game optimal admissible control strategies.

Defining

$$I = (\tilde{u}, \tilde{v}) = \int \left\{ c(kz+a\tilde{u}+b\tilde{v})^2 + d\tilde{u}^2 - e\tilde{v}^2 \right\} p(z, x, y)\, d(z, x, y) \tag{4}$$

the saddle point conditions are

$$I(\bar{u}+\epsilon\delta, \bar{v}) - I(\bar{u}, \bar{v}) = 2\int \left\{ ackz+(a^2c+d)\bar{u}+abc\bar{v} \right\} \epsilon\delta p(z, x, y)\, d(z, x, y)$$

$$+\int \left\{ a^2c+d \right\} \epsilon^2\delta^2 p(z, x, y)\, d(z, x, y) \geq 0 \tag{5}$$

$$I(\bar{u}, \bar{v}+\epsilon\Delta) - I(\bar{u}, \bar{v}) = 2\int \left\{ bckz+abc\bar{u}-(e-b^2c)\bar{v} \right\} \epsilon\Delta p(z, x, y)\, d(z, x, y)$$

$$-\int \left\{ e-b^2c \right\} \epsilon^2\Delta^2 p(z, x, y)\, d(z, x, y) \leq 0 \tag{6}$$

In the following, it is assumed that

$$a^2c + d > 0 \tag{7}$$

$$e - b^2c > 0 \tag{8}$$

So that the game optimal control strategies are found from the simultaneous solution of

$$\int \left\{ ackz + (a^2c+d)\bar{u} + abc\bar{v} \right\} \epsilon \delta p(z, x, y)\, d(z, x, y) = 0 \tag{9}$$

$$\int \left\{ bckz + abc\bar{u} - (e - b^2c)\bar{v} \right\} \epsilon \Delta p(z, x, y)\, d(z, x, y) = 0 \tag{10}$$

About all that can be said at this point is that ϵ is a small number (not zero) and that Δ is any real function of y. Until admissible strategies are defined for player II, it is impossible to say what δ is a function of, although, in all cases, it is assumed to be a real quantity.

As in Chapter 3, it is assumed that both players know the structure of the game, the class of admissible strategies, the values of all system parameters, and the mean and variances of all distributions.

2. EXAMPLE 1. THE MINIMIZING PLAYER HAS NOISY OBSERVATIONS

Example 1 is the single-stage case which corresponds to the derivations and examples of Chapter 3. In this case,

$$p(u, v \mid z, x, y) = \delta(u - \bar{u}(x))\, \delta(v - \bar{v}(y)) \tag{11}$$

so that it is convenient to decompose p(z, x, y) into

$$p(z, x, y) = p(z \mid x, y)\, p(y \mid x)\, p(x) \tag{12}$$

and

$$p(z, x, y) = p(z \mid x, y)\, p(x \mid y)\, p(y) \tag{13}$$

Defining the following linear transformations

$$\beta(x, y) = T_1\alpha(z) = \int \alpha(z)\, p(z\,|\,x, y)\, dz \tag{14}$$

$$\beta(x) = T_2\alpha(y) = \int \alpha(y)\, p(y\,|\,x)\, dy \tag{15}$$

$$\beta(y) = T_3\alpha(x) = \int \alpha(x) p(x\,|\,y)\, dx \tag{16}$$

(with suitable domain and range, of course) the necessary conditions for game optimal control strategies are, (9) and (10),

$$ackT_2T_1z + (a^2c + d)\overline{u} + abcT_2\overline{v} = 0 \tag{17}$$

$$bckT_3T_1z + abcT_3\overline{u} - (e - b^2c)\overline{v} = 0 \tag{18}$$

so that

$$\overline{u} = -\frac{ack}{a^2c+d}\left[1 - \frac{a^2b^2c^2}{(a^2c+d)(b^2c-e)}\, T_2T_3\right]^{-1} T_2\left[1 + \frac{b^2c}{e-b^2c}\, T_3\right]T_1z \tag{19}$$

$$\overline{v} = \frac{bck}{e-b^2c}\left[1 - \frac{a^2b^2c^2}{(a^2c+d)(b^2c-e)}\, T_3T_2\right]^{-1} T_3\left[1 - \frac{a^2c}{a^2c+d}\, T_2\right]T_1z \tag{20}$$

The required conditional densities are

$$T_1{:}p(z\,|\,x, y) = \frac{\sigma_\xi^2\sigma_z^2+\sigma_\eta^2\sigma_z^2+\sigma_\eta^2\sigma_\xi^2}{\sigma_\eta^2\sigma_\xi^2\sigma_z^2}\left\{z - \frac{\sigma_\xi^2\sigma_z^2x+\sigma_\eta^2\sigma_z^2y+\sigma_\eta^2\sigma_\xi^2m}{\sigma_\xi^2\sigma_z^2+\sigma_\eta^2\sigma_z^2+\sigma_\eta^2\sigma_\xi^2}\right\}^2 \tag{21}$$

$$T_2 : p(y|x) = \frac{\sigma_z^2 + \sigma_\eta^2}{\sigma_\xi^2\sigma_z^2 + \sigma_\eta^2\sigma_z^2 + \sigma_\eta^2\sigma_\xi^2}\left\{ y - \frac{\sigma_z^2 x + \sigma_\eta^2 m}{\sigma_z^2 + \sigma_\eta^2} \right\}^2 \tag{22}$$

$$T_3 : p(x|y) = \frac{\sigma_z^2 + \sigma_\xi^2}{\sigma_\xi^2\sigma_z^2 + \sigma_\eta^2\sigma_z^2 + \sigma_\eta^2\sigma_\xi^2}\left\{ x - \frac{\sigma_z^2 y + \sigma_\xi^2 m}{\sigma_z^2 + \sigma_\xi^2} \right\}^2 \tag{23}$$

Rewriting (21) through (23) as

$$T_1 : p(z|x, y) = \alpha_1 \left\{ z - \theta_1 x - \theta_2 y - \theta_3 \right\}^2 \tag{24}$$

$$T_2 : p(y|x) = \alpha_2 \left\{ y - \theta_4 x - \theta_5 \right\}^2 \tag{25}$$

$$T_3 : p(x|y) = \alpha_3 \left\{ x - \theta_6 y - \theta_7 \right\}^2 \tag{26}$$

it follows that

$$T_1 z = \theta_1 x + \theta_2 y + \theta_3 \tag{27}$$

where the θ_i are not the θ_i of Chapter 3.

$$\left[1 + \frac{b^2c}{e - b^2c} T_3\right] T_1 z = \theta_1 x + \theta_2 y + \theta_3 + \frac{b^2c}{e - b^2c}\left[\theta_1(\theta_6 y + \theta_7) + \theta_2 y + \theta_3\right]$$

$$= \theta_1 x + \left[\left(1 + \frac{b^2c}{e - b^2c}\right)\theta_2 + \frac{b^2c}{e - b^2c}\,\theta_1\theta_6\right] y + \theta_3 + \frac{b^2c}{e - b^2c}(\theta_1\theta_7 + \theta_3)$$

$$= \theta_1 x + \theta_8 y + \theta_9 \tag{28}$$

$$T_2\left[1+\frac{b^2c}{e-b^2c}T_3\right]T_1 z = \theta_1 x + \theta_8(\theta_4 x + \theta_5) + \theta_9$$

$$= (\theta_1 + \theta_8\theta_4)x + \theta_8\theta_5 + \theta_9$$

$$= \theta_{10}x + \theta_{11} \qquad (29)$$

Where the inverse is expanded, terms involving powers of T_2T_3 appear which operate on (29).

$$T_3 x = \theta_6 y + \theta_7 \qquad (30)$$

$$T_2T_3 = \theta_6(\theta_4 x + \theta_5) + \theta_7$$

$$= \theta_6\theta_4 x + \theta_6\theta_5 + \theta_7$$

$$= \theta_{12}x + \theta_{13} \qquad (31)$$

Thus, if

$$WT = \frac{a^2b^2c^2}{(a^2c+d)(b^2c-e)}\ T_2T_3 \qquad (32)$$

$$||WT|| < 1 \qquad (33)$$

then

$$[1-WT_2T_3]^{-1}T_2\left[1+\frac{b^2c}{e-b^2c}T_3\right]T_1 z = \theta_{10}x+\theta_{11}+W[\theta_{10}(\theta_{12}x+\theta_{13})+\theta_{11}]$$

$$+ W^2[\theta_{10}\theta_{12}(\theta_{12}x+\theta_{13}) + \theta_{10}\theta_{13} + \theta_{11}]$$

$$+ W^3[\theta_{10}\theta_{12}^2(\theta_{12}x+\theta_{13}) + \theta_{10}\theta_{12}\theta_{13} + \theta_{10}\theta_{13} + \theta_{11}]$$

$$+ W^4[\theta_{10}\theta_{12}^3(\theta_{12}x+\theta_{13}) + \theta_{10}\theta_{12}^2\theta_{13} + \theta_{10}\theta_{12}\theta_{13} + \theta_{10}\theta_{13} + \theta_{11}]$$

$$+ W^5[\theta_{10}\theta_{12}^4(\theta_{12}x+\theta_{13}) + \theta_{10}\theta_{12}^3\theta_{13} + \theta_{10}\theta_{12}^2\theta_{13} + \theta_{10}\theta_{12}\theta_{13} + \theta_{10}\theta_{13} + \theta_{11}] + \ldots$$

$$= \theta_{10}x\left[1+W\theta_{12}+W^2\theta_{12}^2+W^3\theta_{12}^3+W^4\theta_{12}^4+W^5\theta_{12}^5+\ldots\right]$$

$$+ \theta_{11}[1+W+W^2+W^3+W^4+W^5+\ldots]$$

$$+ \theta_{10}\theta_{13}\Big[1+W+W^2+W^3+W^4+\cdots+\theta_{12}(1+W+W^2+W^3+\cdots)$$

$$+ W^2\theta_{12}^2(1+W+W^2+\cdots)+\cdots\Big]$$

$$= \frac{\theta_{10}}{1-W\theta_{12}}x + \frac{\theta_{11}}{1-W} + \frac{W\theta_{10}\theta_{13}}{(1-W)(1-W\theta_{12})} \tag{34}$$

so that

$$\bar{u} = \frac{ack}{a^2c+d}\left[\frac{\theta_{10}}{1-W\theta_{12}}x + \frac{\theta_{11}}{1-W} + \frac{W\theta_{10}\theta_{13}}{(1-W)(1-W\theta_{12})}\right] \tag{35}$$

where

$$\theta_{10} = \left[1 + \frac{b^2c}{e-b^2c}\,\frac{\sigma_z^2}{\sigma_z^2+\sigma_\xi^2}\right]\frac{\sigma_z^2}{\sigma_z^2+\sigma_\eta^2} \tag{36}$$

$$\theta_{11} = \frac{m}{\sigma_z^2+\sigma_\eta^2}\left[\sigma_\eta^2 + \frac{b^2c}{e-b^2c}\,\frac{\sigma_\xi^2\sigma_z^2+\sigma_\eta^2\sigma_z^2+\sigma_\eta^2\sigma_\xi^2}{\sigma_z^2+\sigma_\xi^2}\right] \tag{37}$$

$$\theta_{12} = \left(\frac{\sigma_z^2}{\sigma_z^2+\sigma_\eta^2}\right)\left(\frac{\sigma_z^2}{\sigma_z^2+\sigma_\xi^2}\right) \tag{38}$$

$$\theta_{13} = \left(\frac{\sigma_z^2}{\sigma_z^2+\sigma_\xi^2}\right)\left(\frac{m\sigma_\eta^2}{\sigma_z^2+\sigma_\eta^2}\right) + \frac{m\sigma_\xi^2}{\sigma_z^2+\sigma_\xi^2} \tag{39}$$

(Note that

$$|W\theta_{12}| < 1 \tag{40}$$

so that the infinite series involving $W\theta_{12}$ in (34) converge with the sum being given by a closed form solution.)

A similar exercise leads to

$$\bar{v} = \frac{bck}{e-b^2c}\left[\frac{\theta_{14}}{1-W\theta_{16}}\,y + \frac{\theta_{15}}{1-W} + \frac{W\theta_{14}\theta_{17}}{(1-W)(1-W\theta_{16})}\right] \tag{41}$$

where

$$\theta_{14} = \left[1 - \frac{a^2 c}{a^2 c+d} \frac{\sigma_z^2}{\sigma_z^2+\sigma_\eta^2}\right] \frac{\sigma_z^2}{\sigma_z^2+\sigma_\xi^2} \tag{42}$$

$$\theta_{15} = \frac{m}{\sigma_z^2+\sigma_\xi^2}\left[\sigma_\xi^2 - \frac{a^2 c}{a^2 c+d} \frac{\sigma_\xi^2\sigma_z^2+\sigma_\eta^2\sigma_z^2+\sigma_\eta^2\sigma_\xi^2}{\sigma_z^2+\sigma_\eta^2}\right] \tag{43}$$

$$\theta_{16} = \left(\frac{\sigma_z^2}{\sigma_z^2+\sigma_\eta^2}\right)\left(\frac{\sigma_z^2}{\sigma_z^2+\sigma_\xi^2}\right) \tag{44}$$

$$\theta_{17} = \left(\frac{\sigma_z^2}{\sigma_z^2+\sigma_\eta^2}\right)\left(\frac{m\sigma_\xi^2}{\sigma_z^2+\sigma_\xi^2}\right) + \frac{m\sigma_\eta^2}{\sigma_z^2+\sigma_\eta^2} \tag{45}$$

A comparison of (35) and (41) with the example of a single-stage game, with control strategies specified to be linear (Reference 40), shows that (except for a typographical error) the two solutions are identical if the mean of the a priori estimate of the state is zero (m = 0). However, this example proves, by construction of the solution, that (under the assumptions of linear dynamics, quadratic payoff function, Gaussian random variables, and noisy observations for both players) there is no nonlinear pure control strategy that can do better.

Consider what happens when player I is unable to affect the outcome of the game (b = 0). In this case, the single-stage game degenerates to a single-stage minimization problem with solution given by

$$\bar{u} = - \frac{ack}{a^2c+d} \left(\frac{\sigma_z^2 x + \sigma_\eta^2 m}{\sigma_z^2 + \sigma_\eta^2} \right) \tag{46}$$

where the terms within the parentheses comprise the minimum variance estimate of the true state — an example of the well-known separation property of this class of stochastic control problems.[36] Note that this happy situation does not exist in (35) since system parameters and both players' observation noise variances are inextricably mixed together. In other words, even in this simplest of stochastic differential games, a separation theorem does not exist.

Setting the observation noise variances to zero (perfect observations) leads directly to the deterministic solution found in Chapter 2.

If only one of the players, say player II, has perfect observations, then the control strategies are still given by (35) and (41), except that

$$\sigma_\eta^2 = 0 \tag{47}$$

In this case, player II's control strategy is not identical with the deterministic one of Chapter 2. Player II's game optimal strategy still involves terms reflecting the noisy nature of player I's observations. Again, there is no separation theorem.

3. EXAMPLE II. THE MINIMIZING PLAYER HAS PERFECT OBSERVATIONS AND KNOWS THE MAXIMIZING PLAYER'S OBSERVATION

This single-stage game corresponds, in terms of information content, to the work presented in Reference 41. In this case,

$$p(u, v \mid z, x, y) = \delta(u - \bar{u}(z, y))\, \delta(v - \bar{v}(y)) \tag{48}$$

where x has been dropped from consideration since x is identical to z at all times (perfect information). Instead of (13), the remaining joint density can be written

$$p(z, y) = p(z \mid y)\, p(y) \tag{49}$$

so that only one linear transformation is required

$$\beta(y) = T\alpha(z) = \int \alpha(z)\, p(z \mid y)\, dz \tag{50}$$

which results in

$$ackz + (a^2 c + d)\,\bar{u} + abc\bar{v} = 0 \tag{51}$$

$$bckTz + abcT\bar{u} - (e - b^2 c) = 0 \tag{52}$$

Solving (51) and (52) simultaneously leads to

$$\bar{u} = -\frac{ack}{a^2 c + d}\, z - \frac{abc}{a^2 c + d}\left[\frac{bcdk}{(a^2 c + d)(e - b^2 c) + a^2 b^2 c^2}\right] Tz \tag{53}$$

$$\bar{v} = \frac{bcdk}{(a^2 c + d)(e - b^2 c) + a^2 b^2 c^2}\, Tz \tag{54}$$

Since the conditional mean of z given y is precisely the minimum variance estimate of the state, $\hat{z}$, (54) indicates that, at last, there is a separation theorem for the maximizing player. If the error, $\tilde{z}$, between the true state, z, and the best estimate of the state, $\hat{z}$, is introduced, then (53) can be rewritten as

$$\bar{u} = - \frac{acdk}{(a^2c+d)(e-b^2c)+a^2b^2c^2} z + \frac{abc}{a^2c+d} \left[\frac{bcdk}{(a^2c+d)(e-b^2c)+a^2b^2c^2} \right] \tilde{z} \tag{55}$$

where

$$\tilde{z} = z - \hat{z} \tag{56}$$

Equation (55) shows that player II's optimal control strategy can be broken into parts: one part which is identical to that used in the fully deterministic case, and a second part which is proportional to the error in player I's estimate of the true state. As usual, a linear control strategy results.

This example could be extended to the multistage case, if desired, but the results would not match those obtained in Reference 41. Even though both players have linear control strategies under either formulation, the imposition of the requirement that the strategies be linear changes the essential character of the solution; a great deal more information is available to both players if they know the form (the structure) of the strategies. In effect, the variance of the estimate of the opposing player's past controls is reduced since mere capability must no longer be considered alone. Instead, the estimate depends on the ability of each player to estimate his opponent's observation --- a situation which is much easier to handle.

SECTION V

MULTISTAGE STOCHASTIC DIFFERENTIAL GAMES WITH SPECIFIED CONTROL STRATEGIES

I. INTRODUCTION

This chapter investigates the case of a multistage stochastic differential game wherein both players have only noisy observations of the true state. Unlike the work presented in Chapter 3, the form of the control strategies for both sides is specified. In particular, the strategies are specified to be linear functions of the present and past observations available to each player.

The assumption of the form of the control strategy has a major impact on the method of solution required. Previously, the problem was one of functional optimization over the class of all strategies, and the methods of functional analysis were used; now, the problem is reduced to optimization over a set of parameters, and the ordinary calculus suffices.

The method is best presented by performing a two-stage example. The extension to an N stage game is then obvious.

II. DERIVATION OF THE GAME OPTIMAL LINEAR STRATEGIES

The problem is to choose a set of parameters in the pure control strategies which optimize (in a game sense) a quadratic payoff functional

$$J_2 = \min_{\alpha^1} \max_{\beta^1} E\left\{\sum_{i=1}^{2} c_i z_{i-1}^2 + d_i \tilde{u}_i^2 - e_i \tilde{v}_i^2\right\} \tag{1}$$

(The payoff function is written in terms of scalar states and control variables for the convenience of doing an example problem, but everything to be said goes over immediately to the vector case.) The tilde over the two control strategies, $\tilde{u}_i$ and $\tilde{v}_i$, is meant to denote that the control strategies are restricted to those having a certain form, which is linear in this example. The min and max operators are to be evaluated over the set of parameters α_1, α_{11}, α_{12}, α_2, and α_{22} and β_1, β_{11}, β_{12}, β_2 (denoted by α and β, respectively) since

$$\tilde{u}_i = \alpha_1 + \sum_{j=i}^{2} \alpha_{ij} x_j \tag{2}$$

$$\tilde{v}_i = \beta_i + \sum_{j=i}^{2} \beta_{ij} y_j \tag{3}$$

Thus, solving the last stage game first,

$$J_1 = \min_{\alpha} \max_{\beta} \int \left\{ c_1 z_o^2 + d_1 \tilde{u}_1^2 - e_1 \tilde{v}_1^2 \right\} p(z_o, \tilde{u}_1, \tilde{v}_1)\, d(z_o, \tilde{u}_1, \tilde{v}_1)$$

$$= c_1 \sigma_{\lambda_1}^2 + \min_{\alpha} \max_{\beta} \int \left\{ c_1 (k_1 z_1 + a_1 \tilde{u}_1 + b_1 \tilde{v}_1)^2 + d_1 \tilde{u}_1^2 - e_1 \tilde{v}_1^2 \right\}$$

$$\times\, p(z_1, \tilde{u}_1, \tilde{v}_1) d(z_1, \tilde{u}_1, \tilde{v}_1) = c_1 \sigma_{\lambda_1}^2 + \min_{\alpha} \max_{\beta} \int \Big\{ c_1 \Big[k_1 z_1$$

$$+ a_1(\alpha_1 + \alpha_{11} x_1 + \alpha_{12} x_2) + b_1(\beta_1 + \beta_{11} y_1 + \beta_{12} y_2)^2 \Big]^2 + d_1 [\alpha_1 + \alpha_{11} x_1$$

$$+ \alpha_{12} x_2]^2 - e_1 [\beta_1 + \beta_{11} y_1 + \beta_{12} y_2]^2 \Big\}\, p(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2)$$

$$\times\, d(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2) \tag{4}$$

where α and β in (4) are the set α_1, α_{11}, α_{12}, β_1, β_{11}, and β_{12}, respectively.

Since (4) is a problem in ordinary calculus, the usual sufficient conditions for the minimization and maximation of a function of several variables are applicable, namely

$$\frac{\partial J_1}{\partial \alpha} = 0 \tag{5}$$

$$\frac{\partial J_1}{\partial \beta} = 0 \tag{6}$$

$$\frac{\partial^2 J_1}{\partial \alpha^2} > 0 \tag{7}$$

$$\frac{\partial^2 J_1}{\partial \beta^2} < 0 \tag{8}$$

where (5) and (6) are vector equations and (7) and (8) are matrix equations.

Applying (5) and (6) to (4) yields a set of linear algebraic equations in $\bar{\alpha}$ and $\bar{\beta}$, the optimal values for α and β

$$\frac{\partial J_1}{\partial \alpha_1} = 2\int \Big\{ a_1 c_1 \Big[k_1 z_1 + a_1(\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2) + b_1(\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2) \Big]$$

$$+ d_1 \Big[\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2 \Big] \Big\} \, p\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) \, d\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) = 0 \tag{9}$$

$$\frac{\partial J_1}{\partial \alpha_{11}} = 2\int \Big\{ a_1 c_1 x_1 \Big[k_1 z_1 + a_1(\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \alpha_{12} x_2) + b_1(\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2) \Big]$$

$$+ d_1 x_1 \Big[\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2 \Big] \Big\} \, p(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2) \, d(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2) = 0 \tag{10}$$

$$\frac{\partial J_1}{\partial \alpha_{12}} = 2\int \Big\{ a_1 c_1 x_2 \Big[k_1 z_1 + a_1 (\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2) + b_1 (\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2) \Big]$$

$$+ d_1 x_2 \Big[\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2 \Big] \Big\} p\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) d\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) = 0 \tag{11}$$

$$\frac{\partial J_1}{\partial \beta_1} = 2\int \Big\{ b_1 c_1 \Big[k_1 z_1 + a_1 (\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2) + b_1 (\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2) \Big]$$

$$- e_1 \Big[\bar{\beta}_1 + \bar{\beta}_{11} x_1 + \bar{\beta}_{12} x_2 \Big] \Big\} p\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) d\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) = 0 \tag{12}$$

$$\frac{\partial J_1}{\partial \beta_{11}} = 2\int \Big\{ b_1 c_1 y_1 \Big[k_1 z_1 + a_1 (\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{x}_{12} x_2) + b_1 (\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2) \Big]$$

$$- e_1 y_1 \Big[\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2 \Big] \Big\} p(z_1, x^1 y^1, \tilde{u}_2, \tilde{v}_2)\, d(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2) = 0 \tag{13}$$

$$\frac{\partial J_1}{\partial \beta_{12}} = 2\int \Big\{ b_1 c_1 y_2 \Big[k_1 z_1 + a_1 (\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2) + b_1 (\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2) \Big]$$

$$- e_1 y_2 \Big[\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2 \Big] \Big\} p\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) d\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) = 0 \tag{14}$$

Since

$$p\big(z_1, x^1, y^1, \tilde{u}_2, \tilde{v}_2\big) = \int p(x_1 \mid z_1)\, p(y_1 \mid z_1)\, p(z_1 \mid z_2, \tilde{u}_2, \tilde{v}_2)$$

$$\times\, p(z_2, x_2, y_2, \tilde{u}_2, \tilde{v}_2)\, dz_2 \tag{15}$$

under the assumptions of independent observation and process noise, (9) through (15) can be rewritten as

$$a_1 c_1 k_1 E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) + (a_1^2 c_1 + d_1) \Big\{ \bar{\alpha}_1 + E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, \bar{\alpha}_{11}$$

$$+ E(x_2)\bar{\alpha}_{12} \Big\} + a_1 b_1 c_1 \Big\{ \bar{\beta}_1 + E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\bar{\beta}_{11} + E(y_2)\bar{\beta}_{12} \Big\} = 0 \qquad (16)$$

$$a_1 c_1 k \Big\{ \sigma_{\lambda_2}^2 + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)^2\right] \Big\} + (a_1^2 c_1 + d_1) \Big\{ E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, \bar{\alpha}_1$$

$$+ \left(\sigma_{\eta_1}^2 + \sigma_{\lambda_2}^2 + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)^2\right]\right)\bar{\alpha}_{11} + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, x_2\right]\bar{\alpha}_{12} \Big\}$$

$$+ a_1 b_1 c_1 \Big\{ E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\bar{\beta}_1 + \left(\sigma_{\lambda_2}^2 + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)^2\right]\right)\bar{\beta}_{11}$$

$$+ E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, y_2\right]\bar{\beta}_{12} \Big\} = 0 \qquad (17)$$

$$a_1 c_1 k_1 E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, x_2\right] + (a_1^2 c_1 + d_1) \Big\{ E(x_2)\bar{\alpha}_1$$

$$+ E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, x_2\right]\bar{\alpha}_{11} + E(x_2^2)\bar{\alpha}_{12} \Big\} + a_1 b_1 c_1 \Big\{ E(x_2)\bar{\beta}_1$$

$$+ E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)\, x_2\right]\bar{\beta}_{11} + E(x_2 y_2)\,\bar{\beta}_{12} \Big\} = 0 \qquad (18)$$

$$b_1 c_1 k_1 E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) + a_1 b_1 c_1 \left\{ \bar{\alpha}_1 + E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) \bar{\alpha}_{11} + E(x_2) \bar{\alpha}_{12} \right\}$$

$$- \left(e_1 - b_1^2 c_1 \right) \left\{ \bar{\beta}_1 + E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) \bar{\beta}_{11} + E(y_2) \bar{\beta}_{12} \right\} = 0 \qquad (19)$$

$$b_1 c_1 k_1 \left\{ \sigma^2_{\lambda_2} + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)^2 \right] \right\} + a_1 b_1 c_1 \left\{ E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) \bar{\alpha}_1 \right.$$

$$\left. + \left(\sigma^2_{\lambda_2} + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)^2 \right] \right) \bar{\alpha}_{11} + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) x_2 \right] \bar{\alpha}_{12} \right\}$$

$$- \left(e_1 - b_1^2 c_1 \right) \left\{ E(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) \bar{\beta}_1 + \left(\sigma^2_{\xi_1} + \sigma_{\lambda_2}^{\ 2} + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2)^2 \right] \right) \bar{\beta}_{11} \right.$$

$$\left. + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) y_2 \right] \bar{\beta}_{12} \right\} = 0 \qquad (20)$$

$$b_1 c_1 k_1 E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) y_2 \right] + a_1 b_1 c_1 \left\{ E(y_2) \bar{\alpha}_1 + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) y_2 \right] \bar{\alpha}_{11} \right.$$

$$\left. + E(x_2 y_2) \bar{\alpha}_{12} \right\} - \left(e_1 - b_1^2 c_1 \right) \left\{ E(y_2) \bar{\beta}_1 + E\left[(k_2 z_2 + a_2 \tilde{u}_2 + b_2 \tilde{v}_2) y_2 \right] \bar{\beta}_{11} \right.$$

$$\left. + E\left(y_2^2 \right) \bar{\beta}_{12} \right\} = 0 \qquad (21)$$

where

$$E(\ ) = \int (\ )\, p(z_2, x_2, y_2, \tilde{u}_2, \tilde{v}_2)\, d(z_2, x_2, y_2, \tilde{u}_2, \tilde{v}_2) \qquad (22)$$

At this point, it can be seen that $\bar{\alpha}_1$, $\bar{\alpha}_{11}$, $\bar{\alpha}_{12}$, $\bar{\beta}_1$, $\bar{\beta}_{11}$, and $\bar{\beta}_{12}$ can be found, in terms of system parameters and the expected values of quantities appearing at stage 2, from the simultaneous solution of six linear equations, (16) through (21), in six unknowns. Because of the amount of algebra involved, no attempt has been made to solve the equations explicitly.

Having actually solved the equations, the principle of optimality is used to find the game optimal values for α_2, α_{22}, β_2, and β_{22}.

$$
\begin{aligned}
J_2 &= \underset{\alpha}{\min}\ \underset{\beta}{\max}\ E\left\{c_2 z_1^2 + d_2 \tilde{u}_2^2 - e_2 \tilde{v}_2^2 + J_1\right\} \\
&= \underset{\alpha}{\min}\ \underset{\beta}{\max}\left[\iint\left\{c_2 z_1^2 + d_2 \tilde{u}_2^2 - e_2 \tilde{v}_2^2\right\} p(z_1, \tilde{u}_2, \tilde{v}_2)\, d(z_1, \tilde{u}_2, \tilde{v}_2) + J_1\right] \\
&= c_1 \sigma_{\lambda_2}^2 + \underset{\alpha}{\min}\ \underset{\beta}{\max}\left[\iint\left\{c_2\left[k_2 z_2 + a_2(\alpha_2 + \alpha_{22} x_2) + b_2(\beta_2 + \beta_{22} y_2)\right]^2\right.\right. \\
&\quad \left. + d_2(\alpha_2 + \alpha_{22} x_2)^2 - e_2(\beta_2 + \beta_{22} y_2)^2\right\} p(z_2, k_2, y_2)\, d(z_2, k_2, y_2) \\
&\quad \left. + \gamma_1(\alpha_2, \alpha_{22}, \beta_2, \beta_{22})\right]
\end{aligned}
\tag{23}
$$

where

$$
\begin{aligned}
\gamma_1(\alpha_2, \alpha_{22}, \beta_2, \beta_{22}) &= c_1 \sigma_{\lambda_1}^2 + \iint\left\{c_1\left[k_1 z_1 + a_1(\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2)\right.\right. \\
&\quad \left. + (\bar{\beta}_1 + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2)\right]^2 + d_1(\bar{\alpha}_1 + \bar{\alpha}_{11} x_1 + \bar{\alpha}_{12} x_2)^2 - e_1(\bar{\beta}_1 \\
&\quad \left. + \bar{\beta}_{11} y_1 + \bar{\beta}_{12} y_2)^2\right\} p(z^1, x^1, y^1, \tilde{u}_2, \tilde{v}_2)\, d(z^1, x^1, y^1, \tilde{u}_2, \tilde{v}_2)
\end{aligned}
\tag{24}
$$

and

$$p\left(z_1, x^1, k^1, \tilde{u}_2, \tilde{v}_2\right) = p(x_1 \mid z_1)\, p(y_1 \mid z_1)$$

$$\times\, p(z_1 \mid z_2, \tilde{u}_2, \tilde{v}_2)\, \delta(\tilde{u}_2 - \alpha_2 - \alpha_{22} x_2)$$

$$\times\, \delta(\tilde{v}_2 - \beta_2 - \beta_{22} y_2)\, p(x_2 \mid z_2)\, p(y_2 \mid z_2)\, p(z_2) \tag{25}$$

Substituting (24) and (25) into (23) and then carrying out all the indicated integrations lead to a characterization of J_2 in terms of α_2, α_{22}, β_2, and β_{22}. Again setting the first partial derivatives with respect to each of the four variables equal to zero, along with the required positive definiteness of the matrix of second partials with respect to α_2 and α_{22} and the negative definiteness of the matrix of second partials with respect to β_2 and β_{22}, leads to the optimal values for $\bar{\alpha}_2$, $\bar{\alpha}_{22}$, $\bar{\beta}_2$, and $\bar{\beta}_{22}$ in terms of system parameters and the a priori variances and means of the various random variables.

3. CONCLUDING COMMENTS

The specification of a certain form for the control strategies reduces the conceptual difficulties associated with solving multistage stochastic differential games, but it does little to reduce the difficulty of actually finding the correct values for the control strategy parameters. For instance, at stage i of an N stage game, there are 2 (N-i+2) control strategy parameters to be found by solving a like number of simultaneous equations (and, at this point, there is no way of telling whether the γ_i resulting from optimization at the i^{th} stage leads to an equation which is quadratic in α and β at the $i + 1^{st}$ stage).

Having found $\bar{\alpha}$ and $\bar{\delta}$ at the i^{th} stage, it is still necessary to investigate the eigenvalues of two $N - i + 2$ by $N - i + 2$ matrices which may or may not be functions of $\bar{\alpha}$ and $\bar{\delta}$.

Nevertheless, a straightforward method of actually solving a multistage stochastic differential game having pure strategies has been developed. It is at least possible, though tedious, to solve such a game analytically and thus determine the effects of various values in system parameters and in the a priori distributions.

This work may be compared to that presented in Reference 40 which, in part, solves the same problem. A major difference between the two is in the handling of the various conditional densities that arise. In Reference 40, they are summarized in terms of Kalman filters, while here they are introduced directly. It appears that recursive filtering, while probably leading to the same answer, adds a fair measure of both conceptual and practical difficulties.

Finally, there is the question of the relationship between the multistage stochastic differential games when the control strategies are and are not specified as to form. Since neither closed form analytical nor numerical results are available, one can only speculate as to the differences in the payoff.

The resulting linear control strategies (with and without a linear strategy being prescribed a priori) are not identical. The conclusion is that the control strategy, which is optimal over the set of all linear strategies, is not equivalent to the control strategy (also linear) which is optimal over any control strategies. This seeming contradiction can, however, be resolved.

The reason for the difference between the two strategies is precisely the difference of the information available. In both cases, all available information is used. In this sense, the knowledge that an opponent is limited to using only one form of a solution is merely an additional piece of information. Thus, just as changes in information led to different linear strategies in the examples in Chapter 4, so too do changes in information in multistage games lead to changes in strategies.

This is an example of the difference between stochastic differential games and stochastic optimal control. Unlike stochastic optimal control, there is no separation between estimation of the state and control. And, as noted above, even for the case of linear, deterministic dynamics, quadratic payoff functions, and Gaussian random variables, the optimal strategy, over all strategies, is linear but not equal to the optimal linear strategy.

By setting the appropriate quantities to zero, optimal control problems may be considered to be special cases of differential games; the same statement is not true in reverse. Differential games are not, in general, mere extensions of optimal control.

SECTION VI

FUTURE WORK

The work presented in this dissertation leads one inevitably to consider future areas of research.

For games of perfect information, the question of what combinations of payoff function and dynamics lead to pure strategies is a natural one to ask. It also would be useful to know under what conditions randomized strategies exist and how they are to be found.

The corresponding questions for continuous time games are also worth asking. Does it, in fact, make any sense to talk about randomized strategies when a new control must be chosen at every instant of time?

Much work remains to be done for stochastic games. Simple extensions of the work done herein would include the closed form solution, if one exists, for the multistage vector game of Chapter 3. Numerical solutions should be of interest in any event.

Also, the solution to continuous time differential games, of the type studied in Chapter 3, would be interesting. It is not immediately clear that the same use of conditional probability densities and simple linear operators would produce answers.

Still, in the realm of pure strategies, it would be useful to extend the results to nonlinear problems and to payoff functions that are not quadratic. With regard to randomized strategies, is it possible to apply these techniques to stochastic games or must new ones be developed?

A great deal of information concerning the structure of the problem is assumed available to both players. Further work might consider the effect of less information or information in the form of probability densities. In the same vein, it would be interesting to know if there is a suitable corollary to adaptive control in the game situation.

Obviously, there is a great deal of work yet to be done.

REFERENCES

1. Isaacs, R., "Differential Games I: Introduction," RAND Corporation, Research Memorandum, RM-1391, November 30, 1954.

2. Isaacs, R., "Differential Games II: The Definition and Formulation," RAND Corporation, Research Memorandum, RM-1399, November 30, 1954.

3. Isaacs, R., "Differential Games III: The Basic Principles of the Solution Process," RAND Corporation, Research Memorandum, RM-1411, December 21, 1954.

4. Isaacs, R., "Differential Games IV: Mainly Examples," RAND Corporation, Research Memorandum, RM-1486, March 25, 1955.

5. Isaacs, Rufus, Differential Games, John Wiley and Sons, Inc., New York, 1965.

6. Dresher, M., A.W. Tucker, and P. Wolfe (eds.), Contributions to the Theory of Games, Vol. III, Princeton University Press, Princeton, New Jersey, 1964.

7. Dresher, M., L.S. Shapley, and A.W. Tucker (eds.), Advances in Game Theory, Princeton University Press, Princeton, New Jersey, 1964.

8. Balakrishnan, A.V. and L.W. Neustadt (eds.), Mathematical Theory of Control, Academic Press, New York, 1967.

9. Simakova, E.N., "Differential Games," Automation and Remote Control," Vol. 27, No. 11, p. 1980-1998, November, 1966.

10. Ho, Y.C., "Book Review of Differential Games," IEEE Transaction on Automatic Control, Vol. AC-10, No. 4, p.501-503, October, 1965.

11. Billik, B.H., "Some Optimal Low-Acceleration Rendezvous Maneuvers," AIAA Journal. Vol. 2, No. 3, p. 510-516, March, 1964.

12. Ho, Y.C., A.E. Bryson, Jr., and S. Baron, "Differential Games and Optimal Pursuit - Evasion Strategies," IEEE Transactions on Automatic Control, Vol. AC-10, No. 4, p. 385-389, October, 1965.

13. Wong, R.E., "Some Aerospace Differential Games," Journal of Spacecraft and Rockets, Vol. 4, No. 11, p. 1460-1465, November, 1967.

14. Meschler, P.A., "On a Goal-Keeping Differential Game," IEEE Transactions on Automatic Control, Vol. AC-12, No. 1, p. 15-21, February, 1967.

15. Chattopadhyay, R., "On Differential Games," International Journal of Control, Vol. 6, No. 3, p. 287-295, September, 1967.

16. Ho, Y.C., "Optimal Terminal Maneuver and Evasion Strategy," J. SIAM Control, Vol. 4, No. 3, p. 421-428, August, 1966.

17. Speyer, J.L., "A Stochastic Differential Game with Controllable Statistical Parameters," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-3, No. 1, p. 17-20, June, 1967.

18. Behn, R.D. and Ho, Y.C., "On a Class of Linear Stochastic Differential Games," IEEE Transactions on Automatic Control, Vol. AC-13, No. 3, p. 227-240, June, 1967.

19. von Neumann, John and Oskar Morgenstern, Theory of Games and Economic Behavior, 2nd ed., Princeton University Press, Princeton, New Jersey, 1947.

20. McKinsey, J.C.C., Introduction to the Theory of Games, McGraw-Hill Book Company, Inc., New York, 1952.

21. Luce, R. Duncan and Howard Raiffa, Games and Decisions, Introduction and Critical Survey, John Wiley and Sons, Inc., New York, 1957.

22. Dresher, Melvin, Games of Strategy: Theory and Applications, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1961.

23. Owen, Guillermo, Game Theory, W.B. Saunders Company, Philadelphia, 1968.

24. Fleming, W.H., "The Convergence Problem for Differential Games," Journal of Mathematical Analysis and Applications, Vol. 3, No. 1, p. 102-116, August, 1961.

25. Berkovitz, L.D. and M. Dresher, "A Game-Theory Analysis of Tactical Air War," Operations Research, Vol. 7, No. 5, p. 599-620, September-October, 1959.

26. Berkovitz, L.D. and M. Dresher, "A Multimove Infinite Game With Linear Payoff," Pacific Journal of Mathematics, Vol. 10, No. 3, p. 743-765, Fall, 1960.

27. Berkovitz, L.D. and M. Dresher, "Allocation of Two Types of Aircraft in Tactical Air War: A Game-Theoretic Analysis," Operations Research, Vol. 8, No. 5, p. 694-706, September-October, 1960.

28. Berkovitz, L.D., "A Variational Approach to Differential Games," in Advances in Game Theory (M. Dresher, A.W. Tucker, and P. Wolfe, eds.), Princeton University Press, Princeton, New Jersey, 1964.

29. Bellman, Richard E., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

30. Bliss, Gilbert A., Lectures on the Calculus of Variations, University of Chicago Press, Chicago, 1946.

31. Smith, F.T., "An Introduction to the Application of Dynamic Programming to Linear Control Systems," RAND Corporation, Research Memorandum, RM-3526-PR, February, 1963.

32. Ciletti, M.D., "On a Class of Deterministic Differential Games With Imperfect Information," Department of Electrical Engineering, Notre Dame University, Technical Report, No. EE-679, December 1, 1967.

33. Porter, W.A., "On Function Space Pursuit - Evasion Games," J. SIAM Control, Vol. 5, No. 4, p. 555-574, November, 1967.

34. Bohnenblust, H.F., S. Karlin and L.S. Shapley, "Games With Continuous, Convex Pay-Off," in Contributions to the Theory of Games, Vol. I, (H.W. Kuhn and A.W. Tucker, eds.), Princeton University Press, Princeton, New Jersey, p. 181-192, 1950.

35. Meschler, P.A., "Differential Game-Theoretic Analysis of Maneuvering Threats," Defense Research Corporation, Technical Memorandum, No. 346/1, June, 1966.

36. Aoki, Masanao, Optimization of Stochastic Systems, Academic Press, New York, 1967.

37. Papoulis, Athanasios, Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Company, New York, 1965.

38. Porter, William A., Modern Foundations of Systems Engineering, Macmillan Company, New York, 1966.

39. Liebelt, Paul B., An Introduction to Optimal Estimation, Addison-Wesley Publishing Co., Reading, Mass., 1967.

40. Willman, W.W., "Formal Solutions for a Class of Stochastic Pursuit-Evasion Games," Division of Engineering and Applied Physics, Harvard University, Technical Report, No. 575, November, 1968.

41. Behn, R.D. and Y.C. Ho, "On a Class of Linear Stochastic Differential Games," Division of Engineering and Applied Physics, Harvard University, Technical Report, No. 542, October, 1967.

APPENDIX

The following two theorems are taken almost directly from Chapter 12 of Reference 20, with appropriate changes in rotation.

<u>Theorem I (12.2)</u>: Let I_i be the payoff function of a continuous game, and suppose that I_i is continuous in both variables and that $I_i(u_i, v_i)$ is strictly convex in u_i for every v_i. Then there is a unique optimal strategy for the second player, which is a step function of first order; i.e., there is a number $\bar{u}_i$ in the closed interval $[-1, 1]$ such that the (unique) optimal strategy for the second player is the step function $H(\bar{u}_i)$. ($H(\bar{u}_i)$ is a probability density function). The Value J_i of the game is given by the formula

$$J_i = \min_{-1 \le v_i \le 1} \; \max_{-1 \le u_i \le 1} I_i(u_i, v_i)$$

and the constant $\bar{u}_i$ is the unique solution of the equation

$$\max_{-1 \le v_i \le 1} I_i(v_i, \bar{u}_i) = J_i$$

Theorem I provides a means for finding the minimizing player's game optimal control strategy and the Value of the game. The following theorem does the same for the maximizing player.

<u>Theorem II (12.5)</u>: Let I_i be the payoff function of a continuous game, and suppose that I_i is continuous in both variables, that $\partial I_i(u_i, v_i)/\partial u_i$ exists for each u_i and v_i in $[-1, 1] \times [-1, 1]$, and that $I_i(u_i, v_i)$ is a

strictly convex function of u_i for each v_i. Let $H(\overline{u}_i)$ be the unique optimal strategy for the second player, and let J_i be the Value of the game. If $\overline{u}_i = -1$ or 1, then there is an optimal strategy $H(\overline{v}_i)$ for the first player; the constant $\overline{v}_i$ can be taken to be any number satisfying the conditions

$$0 \le \overline{v}_i \le 1 \quad ,$$

$$I_i(\overline{u}_i, \overline{v}_i) = J_i \quad ,$$

$$\left.\frac{\partial I_i}{\partial u_i}\right|_{\overline{u}_i, \overline{v}_i} \begin{cases} \ge 0 & \text{if } \overline{u}_i = -1 \\ \le 0 & \text{if } \overline{u}_i = 1 \end{cases}$$

If $-1 < \overline{u}_i < 1$, then there is an optimal strategy for the first player, which has the form

$$\alpha H(\overline{v}_i^{\,1}) + (1 - \alpha) H(\overline{v}_i^{\,2})$$

and the constants α, $\overline{v}_i^{\,1}$, and $\overline{v}_i^{\,2}$ can be taken to be any numbers satisfying the conditions

$$-1 \le \overline{v}_i^{\,1} \le 1 \quad , \quad -1 \le \overline{v}_i^{\,1} \le 1 \quad , \quad 0 \le \alpha \le 1$$

$$I_i(\overline{u}_i, \overline{v}_i^{\,1}) = J_i \quad , \quad I_i(\overline{u}_i, \overline{v}_i^{\,2}) = J_i$$

$$\left.\frac{\partial I_i}{\partial u_i}\right|_{\overline{u}_i, \overline{v}_i^{\,1}} \le 0 \quad , \quad \left.\frac{\partial I_i}{\partial u_i}\right|_{\overline{u}_i, \overline{v}_i^{\,2}} \ge 0$$

$$\alpha \left. \frac{\partial I_i}{\partial u_i} \right|_{\bar{u}_i, \bar{v}_i^1} + (1-\alpha) \left. \frac{\partial I_i}{\partial u_i} \right|_{\bar{u}_i, \bar{v}_i^2} = 0$$

Analogous theorems exist for payoffs which are strictly concave in v_i.

*U.S.Government Printing Office: 1979 — 657-002/669