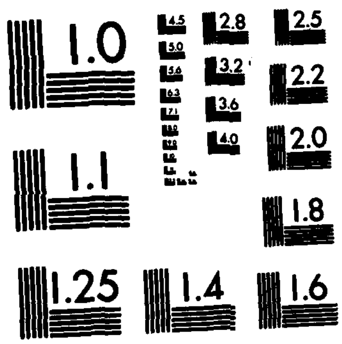
AD-A124 634 STOCHASTIC DIFFERENTIAL GAME TECHNIQUES(U) CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING ANDAPPLIED SCIENCE B MONS MAR 82 DASG60-80-C-0007 1/2

UNCLASSIFIED F/G 12/1 NL



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

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER None	2. GOVT ACCESSION NO. AD-A124634	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STOCHASTIC DIFFERENTIAL GAME TECHNIQUES		5. TYPE OF REPORT & PERIOD COVERED Final, 11/29/79-11/28/81
		6. PERFORMING ORG. REPORT NUMBER None
7. AUTHOR(s) Barend Mons School of Engineering & Applied Science, UCLA		8. CONTRACT OR GRANT NUMBER(s) DASG-60-80-C-0007
9. PERFORMING ORGANIZATION NAME AND ADDRESS UCLA, School of Engineering and Applied Science Los Angeles, California 90024		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS None
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Technology Center U.S. Army Ballistic Missile Defense Command Huntsville, Alabama		12. REPORT DATE March, 1982
		13. NUMBER OF PAGES 155
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as No. 11		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

~~Distribution of this document is not limited.~~

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

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

None

18. SUPPLEMENTARY NOTES

None

DTIC
ELECTE
FEB 22 1983
S D
B

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

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

See attached

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

83 02 018 084

AD A124634

DTIC FILE COPY

2

ABSTRACT

The development of readily computable strategies for differential games with noise corrupted measurements ~~has been~~ was hampered by the so called closure problem of stochastic differential games. The solutions required either an infinite dimensional dynamic system or the determination at each time t of the error in the opponent's state estimate.

In this dissertation, solutions to differential games with noise corrupted measurements ~~have been~~ were obtained that are readily computable.

As a consequence of the stochastic aspects of such games, the discussion ~~has been~~ was restricted to linear-quadratic differential games which are analyzed using function space techniques.

The solution to a linear-quadratic game with perfect information is obtained without the a priori assumption of a saddle-point solution and it is shown that the individual minimax and maximin solutions to such a game result in a set of strategies that satisfy the saddle-point condition, but with necessary and sufficient conditions that are more stringent than previously obtained.

Following recent developments, the concept of prior and delayed commitment strategies are introduced and the solutions obtained for a game where one player has perfect state information and the other player receives noise corrupted measurements. A pursuit-evasion example of such a game is developed and by solving it the numerical differences between the prior and delayed commitment solutions for this game are obtained.

The concept of delayed commitment games is then extended to differential games where both players have noise corrupted state measurements and solutions are obtained that are readily computable, thus playing to rest the closure problem of stochastic differential games.

STOCHASTIC DIFFERENTIAL GAME TECHNIQUES

by

Barend Mons

March, 1982

Submitted under contract
DASG-60-80-C-0007
Advances In Technology Development
for Exoatmospheric Intercept Systems

U.S. Army Ballistic Missile
Defense Command
Advanced Technology Center
Huntsville, Alabama

School of Engineering and Applied Science
University of California
Los Angeles, California

ABSTRACT

The development of readily computable strategies for differential games with noise corrupted measurements has been hampered by the so called closure problem of stochastic differential games. The solutions required either an infinite dimensional dynamic system or the determination at each time t of the error in the opponent's state estimate.

In this dissertation, solutions to differential games with noise corrupted measurements have been obtained that are readily computable.

As a consequence of the stochastic aspects of such games, the discussion has been restricted to linear-quadratic differential games which are analyzed using function space techniques.

The solution to a linear-quadratic game with perfect information is obtained without the a priori assumption of a saddle-point solution and it is shown that the individual minimax and maximin solutions to such a game result in a set of strategies that satisfy the saddle-point condition, but with necessary and sufficient conditions that are more stringent than previously obtained.

Following recent developments, the concept of prior and delayed commitment strategies are introduced and the solutions obtained for a game where one player has perfect state information and the other player receives noise corrupted measurements. A pursuit-evasion example of such a game is developed and by solving it the numerical differences between the prior and delayed commitment solutions for this game are obtained.

The concept of delayed commitment games is then extended to differential games where both players have noise corrupted state measurements and solutions are obtained that are readily computable, thus playing to rest the closure problem of stochastic differential games.

CONTENTS

CONTENTS (Continued)

FIGURES

TABLES

CHAPTER 1

INTRODUCTION

The theory of games may be described as the mathematical theory of decision-making by participants, or players, in a competitive environment. In a typical problem each player has some control over the outcome of a particular event, or game, and the theory is concerned with finding the optimal course of action, or strategy, taking into account the possible actions of the opponents. Although some game theoretic concepts can be traced over the past couple of centuries, modern game theory dates from 1944 with the publication of the now classical work, "Theory of Games and Economic Behavior," by von Neumann and Morgenstern [1].

In differential games the ideas of game theory are applied to dynamic conflict situations which can be described by differential equations (continuous time) or difference equations (discrete time). The dynamic system is under control of intelligent adversaries each seeking to optimize his own gain at the expense of that of his opponents, using all the available information to achieve his objective, and having no a priori knowledge of what the opponents are going to do. Differential game theory was first defined and studied by Isaacs [2 - 5] in 1954 at the Rand Corporation and it was only upon the publication of his book, "Differential Games" [6] in 1965, that the interest in the subject became widespread.

Fundamental to the analysis of a game is the formulation of a mathematical model, which includes the payoff, the allowable strategies and the available information sets upon which the players must base their decisions. If the interest is on detail, information and fine structure, the extensive form of a game is often used; while if the stress is on strategies and payoffs, the strategic or normal form of a game is usually employed.

A fundamental tenet of game theory is the Normalization Principle of von Neumann, which says that given a game in extensive form it can always be reduced to an equivalent game in normal form. Although the number of possible strategies in the normal form becomes rapidly enormous, the conceptual simplification makes it in practice a much simpler problem for computing optimal strategies. As a consequence most of the existing results in game theory are for games in normal form. However, there is still a major concern whether this approach is philosophically sound. Aumann and Maschler [7] recently re-examined the Normalization Principle and illustrate via a simple example some of the pitfalls in the passage from the extensive to the normal form of a game. Their results have immediate and serious consequences in differential games with imperfect state information. In effect, previously obtained results of games with imperfect information are useful and reasonable only if the players are irrevocably committed to a strategy determined at the beginning of the game (the prior commitment strategy). This severely limits their applicability, not to mention that, in general, these strategies can only be realized by infinite dimensional state estimators [8]. This paper is therefore concerned with determining the strategies (the delayed commitment strategies) for differential games with imperfect information where the players are not irrevocably committed to their prior commitment solution. The class of games are restricted to linear time varying differential games with noise corrupted measurements and a quadratic payoff function. The allowable strategies are closed-loop, based at each time t on all the available information up to that time and the final time T is fixed.

Chapter 2 presents the various concepts of game theory and a brief review of those aspects of modern optimal control theory that are pertinent to the later chapters. The theoretical development begins in Chapter 3, with a careful definition and analysis of a linear-quadratic differential game with perfect information.

Chapter 4 introduces the stochastic differential game and illustrates the prior commitment and delayed commitment strategy via a tutorial example.

The prior commitment solution obtained by Behn and Ho [9] and Rhodes and Luenberger [10] to a linear-quadratic differential game where the minimizing player has perfect measurements and the maximizing player has noise corrupted measurements of the state is presented in Chapter 5. The delayed commitment solution to this problem is then obtained and the results are compared with those of the prior commitment solution.

To illustrate the results obtained in Chapter 5 we analyze a pursuit-evasion example in Chapter 6 that also allows a finite dimensional solution using the prior commitment formulation. The solutions to both formulations have been obtained and their characteristics compared.

The delayed commitment formulation is then extended, in Chapter 7, to the case where both players have noise corrupted measurements and finite dimensional solutions, which are readily computable, are obtained for both players.

CHAPTER 2

GAME THEORETIC CONCEPTS AND MATHEMATICAL BACKGROUND

As pointed out in the Introduction, the study of differential games is the dynamical equivalent of the problems studied in classical game theory. Although many of the analytical methods for differential games are actually extensions of techniques developed in optimal control theory, the important concepts in differential games come mainly from general game theory.

The fundamental concepts of game theory are introduced in this chapter with a discussion of two basic game models. A brief review of those aspects of modern optimal control theory relevant to the sequel is then presented, and a general mathematical representation of a differential game formulated. The chapter is concluded with a discussion of the solution concepts of differential games.

2.1 GAME THEORETIC CONCEPTS

The success or failure of an analysis using game theory often hinges upon the ability to adequately model a physical situation. The way in which a game model is formulated depends upon our interests and the type of analysis to be performed. The two basic descriptions of a game of interest to us are:

1. the extensive form, and
2. the strategic or normal form.

The extensive form of a game can be illustrated by means of a diagram known as the game tree, shown in Figure 2.1 for a simple two-person game. In this representation of a game, the choice of the first

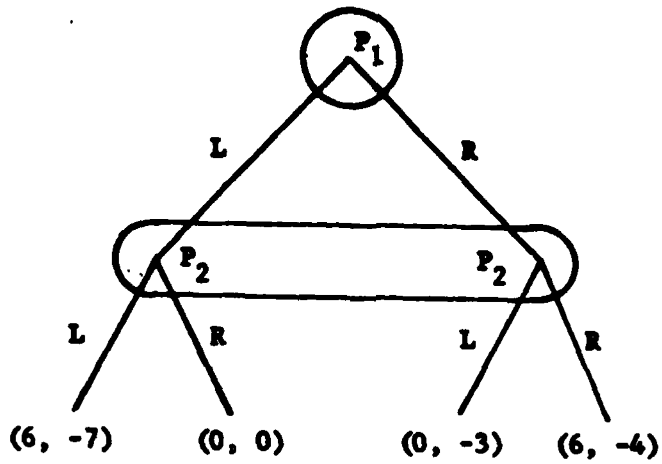


Figure 2.1. The Extensive Form of a Game

player amounts to selecting one of the two branches emanating from the point P_1. After player 1 has made his choice, the second player has to choose a branch at one of the two locations marked P_2. In our simple game, after both players have selected a branch, <u>the payoff</u> is given by the two numbers at the end of the branches. In order to indicate that both players move simultaneously we enclose both of the nodes at P_2 by a curve which indicates an <u>information set</u>. If the second player knows at the time he moves what the first player has chosen, we would then draw a separate information set around each of the nodes.

When engaged in a particular game, each player is faced with the problem of how best to play the game in order to maximize or minimize his expected payoff. A player's complete plan for playing a game is called a <u>strategy</u>, of which there are several different types. A <u>pure strategy</u> for player i is a rule for selecting a particular move at each of his information sets. A mixed strategy for player i is a probability distribution over the set of all pure strategies. A behavioral strategy for player i consists of a collection of probability distributions, one each over the set of possible choices at each of his information sets. A game for which the sum of the payoff's at each terminal node is equal to zero is called a zero-sum game, all other games are nonzero-sum. In 1912, Zermelo (see [1]) demonstrated the existence of an optimal pure strategy for two-person zero-sum games with perfect information, that is games in which all information sets contain a single node. Kuhn [11] extended this result to n-person general-sum games with perfect information. Kuhn also showed the existence of optimal behavioral strategies for games with perfect recall. A game has perfect recall if each player is aware, at each of his moves, of precisely what moves he picked prior to it, but may not know all the choices made by the other players. In 1928, von Neumann showed the existence of optimal mixed strategies for any two-person zero-sum game, which is the well-known Minimax or Fundamental Theorem of Game Theory.

Another of the fundamental tenets of game theory is the Normalization Principle of von Neumann, which says that given a game in extensive form it can always be reduced to an equivalent game in normal form involving only strategies and payoffs. The above example of a game in extensive form reduces in its normal form to a 2 x 2 matrix game shown in Figure 2.2. In this form, the dynamic and informational aspects of the original problem have been suppressed into the strategy which covers all contingencies of the players.

	Player 2's Choice 1	2
Player 1's Choice 1	6, -7	0, 0
2	0, -3	6, -4

Figure 2.2. The Normal Form of a Game

When a game is constrained by a system that evolves over time (or some other parameter) it is called a dynamic game. If the dynamic system representation takes the form of a difference equation, the game is known as a discrete differential or multistage game. The designation differential game is reserved for a dynamic game where the dynamic system representation is in the form of a differential equation. We will have more to say about the differential game representation in Section 2.3. At this stage it should be noted that implied in the formulation of a game is the assumption that the players "agree" on the structure of the model as well as what is important to both players as expressed by the payoff or payoff function.

In this paper we will be mainly concerned with two-person differential games with perfect, as well as with imperfect information. They represent an extension of optimal control theory, in that the optimal control problem can be considered as a one-sided game. That is a game with only one control input driving a dynamical system instead of two opposing controls as in two-person differential games. In terms of the matrix game of Figure 2.2, a one-player game would consist of simply a single row or column. The development in this paper will be from the optimal control system point of view and we will therefore first discuss the general optimal control problem in the following section.

2.2 REVIEW OF OPTIMAL CONTROL THEORY

In this section we will present a brief discussion of those aspects of modern optimal control theory that are pertinent to our discussion of differential games.

We will first formulate a general deterministic optimal control problem and discuss the basic methods of solution. We will then modify this problem to a stochastic optimal control problem, after which attention is focussed on the linear-quadratic-Gaussian problem. For this problem we discuss the Certainty Equivalence Principle or Separation Theorem, including the notions of controllability, observability and optimal estimation.

In the general optimal control problem one wishes to determine the p-component control vector u(t) that minimizes the given <u>cost functional</u>

$$J(t_o,x_o,u) = B(x(T),T) + \int_{t_o}^{T} F(x(t),u(t),t)dt \qquad (2.1)$$

subject to the constraints

$$\frac{dx}{dt} = \dot{x} = f(x(t),u(t),t) \ ; \ x(t_o) = x_o \qquad (2.2)$$

The n-component vector x is the <u>state vector</u> and Equation (2.2) is known as the <u>dynamic system equation</u>. The n-vector function f, as well as the scalar functions B and F are assumed to be sufficiently smooth in the sense that all the necessary partial derivatives exist. In addition, there may be magnitude or inequality constraints on the state and control variables, as well as restrictions on the terminal state. The terminal time T may be variable or fixed; here it is assumed fixed for simplicity.

The optimal control problem is then to find that control function u(t) (if it exists) defined on the interval $\left[t_o, T\right]$ that satisfies all the problem constraints and is optimal in the sense that it simultaneously minimizes the cost function. In other words, we wish to find the <u>allowable control function</u> $u^*(\cdot)$, such that for any control $u(\cdot)$ belonging to the allowable control function set U, there holds for all $t \in \left[t_o, T\right]$

$$J(t_o, x_o, u^*) \leq J(t_o, x_o, u) \tag{2.3}$$

Basically, four methods of approach are available to solve the optimal control problem; they are,

1. The classical <u>calculus of variations</u> approach, which leads to the Euler-Lagrange equations as the necessary conditions for the control to be optimal.
2. The <u>Maximum Principle of Pontryagin</u> approach, which provides the necessary conditions for optimality. It is usually the most direct method for problems involving magnitude constraints.
3. The <u>dynamic programming</u> approach, which leads to the Hamilton-Jacobi equations. Although the Hamilton-Jacobi equation cannot be easily solved in general, u(t) is determined as a function of x(t), or in other words, we find a feedback control law which is highly desirable.

4. The <u>functional analysis</u> approach. Its appeal stems primarily from its geometric character and is most useful for problems formulated on a fixed time interval.

In this paper we will almost exclusively use the functional analysis approach to obtain the solution to optimal control and differential game problems.

Frequently, it is required to obtain on-line <u>feedback or closed loop</u> control of the dynamic system; i.e., we seek a solution of the form $u(t) = u(x(t),t)$. However, restricting the allowable controls to belong to the set $U : u(t) = u(x(t),t)$ greatly complicates the determination of a solution. In fact, of the four basic approaches listed above, only the dynamic programming approach directly provides a closed loop solution. Otherwise, the dependence of the control $u(t)$ on $x(t)$ can be explicitly identified only for a linear dynamic system with a quadratic cost functional.

If the system dynamics (Equation (2.2)) are perturbed by random disturbances, and/or if the initial conditions are random, and/or if the only available information about the state $x(t)$ is available through noise corrupted measurements of the state variables, the deterministic optimal control problem becomes a stochastic optimal control problem. In this case, the criterion of optimality needs to be modified to that of minimizing the <u>expected value</u> of the cost functional.

Thus, by postulating that the only available information about the state of the system can be obtained by measurements of the form

$$z(t) = h(x(t),w(t),t) \tag{2.4}$$

where the <u>output vector</u> $z(t)$ is of dimension $m \leq n$, the function $h(\cdot,\cdot)$ is sufficiently smooth in each argument and $w(t)$ is a random noise process, it follows that we are dealing with a stochastic control problem. The conversion to a stochastic optimal control problem is completed by modifying the optimality criterion to that of minimizing the expected value of the cost functional; i.e.,

$$J(u) = E\left\{B(x(T),T) + \int_{t_o}^{T} F(x(t),u(t),t)dt\right\} \qquad (2.5)$$

Furthermore, it is necessary to seek a closed-loop solution, thus the allowable controls are of the form

$$u(t) = u(Z(t),t). \qquad (2.6)$$

where

$$Z(t) = \left\{(z(s),s)\ ;\ s \in t_o,t)\right\} \qquad (2.7)$$

i.e., the control u at time t depends on the past and present values of the measurement history $Z(t)$.

The class of problems for which a closed-form analytical solution to the stochastic optimal control problem has been found is the case of a linear system, a quadratic cost functional and white zero-mean Gaussian noise additively corrupting the measurements of the system output. For this special case, the optimal closed-loop solution is given by the important <u>Certainty Equivalence Principle</u> or <u>Separation Theorem</u>.

To review the Separation Theorem, we will consider the linear continuous time system described by the vector differential equation

$$\frac{dx}{dt} = \dot{x} = F(t)x(t) - G(t)u(t) \quad ; \quad x(t_o) = \bar{x}_o \tag{2.8}$$

to which are available measurements of the form

$$z(t) = H(t)x(t) + w(t) \tag{2.9}$$

where $x(t)$ is an n-dimensional state vector, $u(t)$ is a p-dimensional control vector, $z(t)$ is the output vector of dimension $m \leq n$ and the matrices $F(t)$, $G(t)$ and $H(t)$ have the appropriate dimension.

The initial state $x(t_o)$ is assumed a Gaussian random variable with mean $E\left|x(t_o)\right| = \bar{x}_o$ and $\mathrm{cov}\left|x(t_o),x(t_o)\right| = P_o$. The additive noise $w(t)$ is assumed white and Gaussian with zero mean, $\mathrm{cov}\left|w(t),w(\tau)\right| = W(t)\,\delta\,(t-\tau)$ and independent of the initial condition $x(t_o)$.

Consider also the quadratic cost functional

$$J(u) = 1/2\ E\left| x^T(T)x(T) + \int_{t_o}^{T} u^T(t)u(t)dt \right| \tag{2.10}$$

where the final time T is fixed and finite, and the superscript T denotes transposition.

Let the set U of allowable control functions be

$$U \ : \ u(t) = u(Z(t),t) \tag{2.11}$$

where

$$Z(t) = \left| (z(s),s) \ ; \ s \in (t_o,t) \right| \ , \tag{2.12}$$

then the objective is to find that $u^*(t) \in U$ such that

$$E\left|J(u^*(t))\right| \leq E\left|J(u(t))\right| \qquad (2.13)$$

for all $t \in \left|t_o, T\right|$.

The solution to this problem may be stated in three parts;

1. The optimal closed-loop solution to the corresponding deterministic optimal control problem; i.e., for $x(t_o)$ known exactly, $H(t) = I$ the identity matrix and $w(t) = 0$, may be written as

$$u^*(t) = G^T(t)S(t)x(t) \qquad (2.14)$$

where the n x n symmetric matrix $S(t)$ may be precomputed from the matrix Riccati equation.

$$\dot{S} = -S(t)F(t) - F^T(t)S(t) + S(t)G(t)G^T(t)S(t) \qquad (2.15)$$

with the terminal condition

$$S(T) = I \qquad (2.16)$$

If, in addition, (F,G) constitutes a controllable pair; i.e., if

$$\int_{t_o}^{T} \Phi(T,t)G(t)G^T(t)\Phi^T(T,t)dt > 0 \qquad (2.17)$$

where $\Phi(t,t_o)$ is the system state transition matrix which must satisfy the relation

$$\frac{\partial \Phi(t,t_o)}{\partial t} = F(t)\,\Phi(t,t_o) \quad ,$$

$$\Phi(t_o,t_o) = I \quad , \tag{2.18}$$

then $S(t)$ exists and is bounded for all $t \leq T$.

2. The optimal closed loop solution to the stochastic optimal control problem is

$$u^*(t) = G_1^T(t)S(t)\hat{x}(t) \tag{2.19}$$

where

$$\hat{x}(t) = E\left| x(t) \mid Z(t) \right| \tag{2.20}$$

with $Z(t)$ given by Equation (2.7), that is, $\hat{x}(t)$ is the expected value of $x(t)$ given the measurements $z(t)$ up to time t. The matrix $G_1^T(t)S(t)$ is the same as that of Equation (2.14) and is unchanged by the conversion of the deterministic optimal control problem to the stochastic optimal control problem.

3. The best estimate $\hat{x}(t)$ of the state $x(t)$ given the measurements $Z(t)$ is given by

$$\dot{\hat{x}} = F(t)\hat{x}(t) - G(t)u(t) + P(t)H^T(t)W^{-1}(t)$$
$$\left| z(t) - H(t)\hat{x}(t) \right| \;\; ; \;\; \hat{x}(t_o) = \bar{x}_o \tag{2.21}$$

where the n x n symmetric matrix satisfies the matrix Riccati equation

$$\dot{P} = F(t)P(t) + P(t)F^T(t) - P(t)H^T(t)W^{-1}(t)H(t)P(t) \quad ,$$
$$P(t_o) = P_o \tag{2.22}$$

If, in addition, (F,H) consitutes an observable pair; i.e., if

$$\int_{t_o}^{T} \Phi^T(t,t_o)H^T(t)H(t)\,\Phi(t,t_o)dt > 0 \qquad (2.23)$$

then P(t) exists and is bounded for all $t \in \left[t_o, T\right]$.

The two parts (1) and (2) illustrate the Certainty Equivalence Principle, which emphasizes the fact that, for linear systems with quadratic cost functions and subjected to additive white Gaussian noise inputs, the optimal feedback solution treats the conditional mean-state estimate, $\hat{x}(t)$, as the true state. The Separation Theorem expresses the fact that this problem can be solved via two separate problems; optimal estimation and control.

2.3 DIFFERENTIAL GAME FORMULATION

A two person differential game differs from the optimal control problem in that another set of control variables is available for manipulation. Each set of control variables, $u_1(t)$ and $u_2(t)$, can be thought of as being under control of an intelligent player or controller, and each player thus has control over only some of the relevant variables that decide the outcome of the game. The players are opponents, and if the objective of the one controlling $u_1(t)$ is to minimize the cost or payoff of the game, the objective of the one controlling $u_2(t)$ is to maximize it.

In general, the following situation arises for a two-person zero-sum game: For $i = 1, 2$ player i wishes to select his p_i component control vector $u_i(t)$ that optimizes

$$J(t_o,x_o;u_1,u_2) = B(x(T),T) + \int_{t_o}^{T} F(x(t),u_1(t),u_2(t),t)dt \qquad (2.24)$$

subject to the constraints

$$\frac{dx}{dt} = \dot{x} = f(x(t),u_1(t),u_2(t),t) \quad ; \quad x(t_o) = x_o \qquad (2.25)$$

and

$$u_1 \in U_1 \quad ; \quad u_2 \in U_2 \qquad (2.26)$$

As for the optimal control problem there may be inequality constraints on the state and control variables. To ensure termination of the game, the terminal time T is given explicitly in the above game.

The control variables u_1 and u_2 are called the strategies of player 1 and player 2 respectively, and are restricted to certain sets of admissible strategies U_1 and U_2, which depend, in general, on the specific problem to be solved. Equations (2.24) through (2.26) can be thought of as defining the rules of the game. The progress of the game is determined by the n-first order differential equations (2.25). Play starts at time t_o in the state x_o and terminates at time $t = T$. The game is zero-sum because there is a single payoff and the game is called <u>strictly competitive</u>. Furthermore, the game is one of perfect information since both players know the state $x(t)$ at any time $t \in [t_o,T]$. In the case of a two-person nonzero-sum game we may encounter a payoff function such as

$$J_i(t_o, x_o \,;\, u_1, u_2) = B_i(x(T), T) + \int_{t_o}^{T} F_i(x(t), u_1(t), u_2(t), t)\,dt \tag{2.27}$$

for $i = 1, 2$.

Since the players are assumed to have several strategies available for play, the central problem of game theory is the determination of which one to play.

2.4 SOLUTION CONCEPTS

In optimal control theory, the solutions are the allowable control functions that optimize the criterion function and there is no doubt about the meaning of a correct solution. In game theory, however, the presence of the opposing control introduces a dramatic new order of complication not usually found in the one-sided optimal control problem. When each player determines his optimal strategy, he must also take into account his opponent's actions toward the opposite end, the opponent's similar wariness of the other player's actions, and so forth. The basic difficulties are thus related to the available information sets and the rationales used by each player. In nonzero-sum games one can be faced with a great variety of relevant solution concepts involving coalitions, threats, enforceability of agreements, bargaining, etc. In this paper we will explore two solution concepts associated with nonzero-sum games, namely, Nash equilibrium and individual minimax solutions. In two-person zero-sum differential games, the problem of multiple solution concepts does not arise.

2.4.1 Equilibrium Solutions

If game theory is to recommend any specific pair of strategies for a two-person game, then each strategy must be the best possible against the other strategy in the pair; i.e., the pair must be an equilibrium point. Otherwise, a knowledgeable player will know what the theory recommends for the other player, and so will want to select a strategy that is better for him.

If we identify the players by

Player 1; minimizing player with control u_1

Player 2; maximizing player with control u_2

then a strategy pair (u_1^*, u_2^*) is in equilibrium if

$$J_1(u_1^*, u_2^*) \leq J_1(u_1, u_2^*) \quad \forall \; u_1 \in U_1 \tag{2.28}$$

and

$$J_2(u_1^*, u_2) \leq J_2(u_1^*, u_2^*) \quad \forall \; u_2 \in U_2 \tag{2.29}$$

In other words, the strategies are in equilibrium if no player has any positive reason for changing his strategy assuming that the other player is not going to change his strategy. In game theory such an equilibrium solution is known as a Nash equilibrium solution. Thus, if a player knows that the other player is committed to his equilibrium strategy, then he has reason to play the strategy which will give such an equilibrium pair and the game is stable in the sense that no player can unilaterally improve his payoff by changing his strategy.

For two-person zero-sum games, the Nash equilibrium solution leads to a saddle point on the cost surface in the control space and

$$J(u_1^*, u_2) \le J(u_1^*, u_2^*) \le J(u_1, u_2^*) \qquad (2.30)$$

In this case equilibrium pairs are both interchangeable and equivalent, in the sense that, if (u_1', u_2') and (u_1^*, u_2^*) are equilibrium pairs, then so are (u_1', u_2^*) and (u_1^*, u_2') and moreover

$$J(u_1', u_2') = J(u_1^*, u_2^*) = J(u_1', u_2^*) = J(u_1^*, u_2') \qquad (2.31)$$

This well-known result of equivalence and interchangeability [12] for zero-sum games with a saddle-point solution makes the question of uniqueness of the admissible strategies irrelevant. For, if two saddle-points exist, their values are equivalent, and the strategies which give those saddle-points could be played interchangeably without changing the value of the criterion.

Unfortunately, not every game has equilibrium strategy pairs. In general, if a game has no equilibrium strategy pairs, we usually see the players trying to outguess each other, keeping their strategies secret. This suggests, and is indeed true, that for finite games with complete information, equilibrium strategies do exist.

2.4.2 Minimax and Maximin Solutions

Most practical conflict situations are not games of perfect information since ignorance of an opponent's ultimate choice of control is generally an essential element of a conflict situation. In that case each player must approach the design of his own control prepared to limit the adverse cost resulting from his opponent's ultimate choice of control. This means that the minimizing player, player 1, must select u_1 so as to minimize the maximum possible cost, regardless of whether the maximizing player, player 2, ultimately selects u_2 such as to yield this cost.

Hence, from player 1's point of view, if he selects an arbitrary control u_1, then, regardless of the choice of player 2, he is assured of the cost being at most

$$J_1(u_1,u_2^*) = \max_{u_2} J_1(u_1,u_2) \tag{2.32}$$

Since player 1 is the minimizing player, he will select u_1 such that this choice minimizes the maximum cost, that is,

$$J_1(u_1^*,u_2^*) = \min_{u_1} \left| \max_{u_2} J_1(u_1,u_2) \right| \tag{2.33}$$

Thus, the minimax solution is the control u_1^*. Player 1 does not care what strategy his opponent ultimately selects, he is that much ahead if his opponent selects any strategy other than u_2^*, since the resulting cost would be less than, or at best, equal to $J_1(u_1^*,u_2^*)$:

$$J_1(u_1^*,u_2) \leq J_1(u_1^*,u_2^*) \tag{2.34}$$

Hence, $J_1(u_1^*,u_2^*)$ is the <u>loss ceiling</u> or the <u>security level</u> for player 1.

From the point of view of player 2, if he selects an arbitrary control u_2, then regardless of the control of player 1, he is assured of the cost being at least

$$J_2(u_1^*,u_2) = \min_{u_1} \quad J_2(u_1,u_2) \tag{2.35}$$

and since he is the maximizing player, he will select u_2 such that this choice maximizes the minimum cost; i.e.,

$$J_2(u_1^*,u_2^*) = \max_{u_2} \left| \min_{u_1} \quad J_2(u_1,u_2) \right| \tag{2.36}$$

Thus, the maximin solution is the control u_2^*. Player 2 also does not care what strategy his opponent ultimately selects, since if his opponent selects any strategy other than u_1^*, the resulting payoff to player 2, the maximizing player, will be greater than $J_2(u_1^*,u_2^*)$:

$$J_2(u_1^*,u_2^*) \leq J_2(u_1,u_2^*) \tag{2.37}$$

Hence, $J_2(u_1^*,u_2^*)$ is the <u>gain floor</u> or security level for player 2.

The controls u_1^* and u_2^*, derived on the basis of no a priori knowledge of each opponent's ultimate choice, are again stable solutions to the game. Assume, for example, that during a differential game, player 2 has calculated his security level by which he determined the control set $\left| u_1^*(t), u_2^*(t) \right|$ and subsequently found out that player 1 uses the strategy $u_1^*(t)$. Then, player 2 will be able to find another strategy $u_2'(t)$ which will give a payoff greater than $J_2(u_1^*,u_2^*)$. However, as soon as player 2 employs a strategy other than $u_2^*(t)$, there exists a strategy $u_1'(t)$ that together with

$u_2'(t)$ gives a payoff such that $J_2(u_1', u_2') < J_2(u_1^*,u_2^*)$. Hence if player 1 decides to secretly switch to $u_1'(t)$, player 2 has to accept a smaller payoff than if he had stayed with $u_2^*(t)$ in the first place. Thus, unless player 2 has reason to believe that player 1 is <u>irrevocably commited</u> to a strategy other than $u_1^*(t)$, there is no reason at all to play a strategy other than $u_2^*(t)$.

If a player reveals his strategy to his opponent the best he can hope for is the loss ceiling or the gain floor depending on whether the revealing player is player 1 or 2. For a two-person zero-sum game if it happens that $u_1^* = u_1^\circ$ and $u_2^* = u_2^\circ$, the minimax and maximin solutions have located the familiar saddle point solution and there is no point to secrecy.

2.4.3 <u>Open-Loop Versus Closed-Loop Control</u>

The fact that a player plays a maximin or a minimax strategy does not imply that he cannot take advantage of any non-optimal play of his opponent. In fact, the interim action of his opponent during the actual play of a differential game can not be ignored, and what is required are controls that depend explicitly on the state $x(t)$ of the game.

The indifference between open-loop and closed-loop control, as in the deterministic one-sided control problem, has its counterpart in differential games only in the determination of a priori strategies, in which case, $u_i(t) = u_i(x(t_o),t)$. During the actual play of the game it is mandatory that closed-loop control is used and $u_i(t) = u_i(x(t),t)$. Starr [13] has shown that for nonzero-sum differential games the open- and closed-loop equilibrium formulations give entirely different solutions.

CHAPTER 3

THE LINEAR-QUADRATIC PERFECT INFORMATION GAME

In this chapter we develop the solution to a differential game with perfect state information which is of fundamental importance to the delayed commitment strategy solutions of stochastic differential games discussed in later chapters.

The currently available control literature shows that a closed-loop solution for a stochastic optimal control problem seems to be available in closed-form only in the special case of a linear system, a quadratic cost functional and white Gaussian noise additively corrupting the system. It therefore seems unlikely that a closed-form solution for a stochastic differential game problem will be available unless we assume the same or more stringent restrictions for such a game problem. Since stochastic differential games will become our main interest, we will restrict our discussion in this chapter to a linear system with a quadratic payoff functional. Contrary to previously obtained results [14] , [15] , however, our solution will not be conditioned by the a priori assumption of a saddle point solution.

The linear-quadratic differential game representation used in this paper is formulated and the solution is obtained using function space methods. Thus, the analysis is made in Hilbert space and follows the method of approach of Porter in [16] . It is then shown that the optimal strategies can be obtained from a matrix Riccati equation and can be computed prior to the actual game.

3.1 LINEAR-QUADRATIC GAME FORMULATION

Consider the linear continuous-time system governed by the vector differential equation

$$\dot{x}'(t) = \frac{dx'}{dt} = F'(t)x'(t) - G_1'(t)u_1'(t) + G_2'(t)u_2'(t); \quad x'(t_o) = x_o' \tag{3.1}$$

where the n vector $x'(t)$ is the system state; the control vectors $u_1'(t)$ and $u_2'(t)$ are of dimension p and q, respectively; and the matrices $F'(t)$, $G_1'(t)$ and $G_2'(t)$ have the appropriate dimensions. Consider also a quadratic cost (or payoff) functional

$$J(u_1,u_2) = 1/2\left\{ x'^T(T)Q_3x'(T) + \int_{t_o}^{T} u_1'^T(t)Q_1(t)u_1'(t)dt - \int_{t_o}^{T} u_2'^T(t)Q_2(t)u_2'(t)dt \right\} \tag{3.2}$$

where the matrices $Q_1(t)$, $Q_2(t)$ and Q_3 are symmetric positive definite; and the final time T is fixed and finite.

The payoff functional can be written more efficiently by use of the following transformations. Since $Q_1(t)$, $Q_2(t)$ and Q_3 are positive definite and symmetric, they may be factored as

$$Q_i = Q_i^{1/2T} Q_i^{1/2}, \qquad i = 1, 2, 3 \tag{3.3}$$

Then by the transformations

$$x'(t) \longrightarrow Q_3^{-1/2}x(t)$$

$$u_1'(t) \longrightarrow Q_1^{-1/2}(t)u_1(t) \tag{3.4}$$

$$u_2'(t) \longrightarrow Q_2^{-1/2}(t)u_2(t)$$

The system equation becomes

$$\dot{x} = Q_3^{1/2}F'(t)Q_3^{-1/2}x(t) - Q_3^{1/2}G_1'(t)Q_1^{-1/2}(t)u_1(t)$$

$$+ Q_3^{1/2}G_2'(t)Q_2^{-1/2}(t)u_2(t); \quad x(t_o) = Q_3^{1/2}x_o' = x_o \tag{3.5}$$

If we now define the new matrices

$$F(t) \triangleq Q_3^{1/2}F'(t)Q_3^{-1/2}$$

$$G_1(t) \triangleq Q_3^{1/2}G_1'(t)Q_1^{-1/2}(t) \tag{3.6}$$

$$G_2(t) \triangleq Q_3^{1/2}G_2'(t)Q_2^{-1/2}(t)$$

the system equation and payoff functional are respectively

$$\dot{x} = \frac{dx}{dt} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t); \quad x(t_o) = x_o \tag{3.7}$$

$$J(u_1,u_2) = 1/2\left\{x^T(T)x(T) + \int_{t_o}^{T}\left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right]dt\right\} \tag{3.8}$$

In view of the possibility of making the above transformations, we will consider Equation (3.7) as the defining system equation and Equation (3.8) as the payoff functional.

The above formulation involves a single dynamic system instead of the two separate systems of the pursuit-evasion problem as in [14]. However, this single system includes the pursuit-evasion problem as a special case, since the individual state vectors of the pursuit-evasion problem can be combined into a single state vector and the two opponents considered to constitute a single system.

Player 1, the minimizing player, attempts to minimize the payoff functional or criterion; i.e., he minimizes the term $x^T(T)x(T)$ in Equation (3.8) as well as his own expended energy, while maximizing the energy expended by player 2. Player 2, the maximizing player, attempts to maximize the same criterion. Thus, the game is zero-sum and since each player is assumed to have perfect knowledge of the system state it is more accurately a zero-sum game with perfect information.

The class of admissible strategies are defined as those U_1 and U_2 which give rise to the controls

$$\begin{aligned} U_1 &: u_1 = u_1(t,x(t)) \\ U_2 &: u_2 = u_2(t,x(t)) \end{aligned} \tag{3.9}$$

that are bounded and that are continuous almost everywhere for $t_o \leq t \leq T$.

It is well known that, for arbitrary $t = t_o$, x_o, $u_1(t)$ and $u_2(t)$, the solution to Equation (3.7) may be written as

$$x(t) = \Phi(t,t_o)x(t_o) - \int_{t_o}^{t} \Phi(t,\tau)G_1(\tau)u_1(\tau)d\tau$$

$$+ \int_{t_o}^{t} \Phi(t,\tau)G_2(\tau)u_2(\tau)d\tau \qquad (3.10)$$

where $\Phi(t,t_o)$ is the state transition matrix, i.e. it satisfies the relation

$$\frac{\partial \Phi(t,t_o)}{\partial t} = F(t)\Phi(t,t_o)$$

$$\Phi(t_o,t_o) = I \qquad (3.11)$$

As mentioned previously we will analyze the above problem using functional analysis techniques, although any other of the four basic methods of approach mentioned in Section 2.2 could have been used.

To reformulate the differential game in a suitable Hilbert space consider the controls $u_1(\cdot)$ and $u_2(\cdot)$ to be elements of the Hilbert spaces $H_1 = L_2^p\left[t_o,T\right]$ and $H_2 = L_2^q\left[t_o,T\right]$ respectively, where the space $L_2^r\left[t_o,T\right]$ is the space of r - vector functions which are defined and (Lebesgue -) square integrable over the interval $\left[t_o,T\right]$. The inner product on this space is defined as

$$< z, y > = \int_{t_o}^{T} z^T(t)y(t)dt \qquad (3.12)$$

and the norm is defined in terms of the inner product as

$$\|y\|^2 = < y, y > = \int_{t_o}^{T} y^T(t)y(t)dt \qquad (3.13)$$

Hence the two integrals in Equation (3.10) may be considered as linear operations on u_1 and u_2 respectively, and we can represent the dynamic system (3.7) in terms of linear transformations on suitable Hilbert spaces as

$$x(t) = \Phi(t)x_o - (T_1u_1)(t) + (T_2u_2)(t) \qquad (3.14)$$

where the linear operator $T_1 \;:\; L_2^p\left[t_o,T\right] \rightarrow E^n$ is defined by

$$(T_1u_1)(t) = \int_{t_o}^{t} \Phi(t,\tau)G_1(\tau)u_1(\tau)d\tau \qquad (3.15)$$

with a similar definition for T_2 and E^n is the n-dimensional Euclidean space. The terminal state can then be written as

$$x(T) = \Phi(T)x_o - (T_1u_1)(T) + (T_2u_2)(T) \qquad (3.16)$$

Dropping the argument T whenever t = T, the first term of the payoff functional (3.8) may then be written as

$$x^T(T)x(T) = < \Phi x_o - T_1u_1 + T_2u_2, \Phi x_o - T_1u_1 + T_2u_2 > \qquad (3.17)$$

The other terms of the payoff functional may similarly be expressed as inner products and we can write the payoff functional as

$$J(u_1,u_2) = 1/2 \left| < \Phi x_o - T_1u_1 + T_2u_2, \Phi x_o - T_1u_1 + T_2u_2 > + < u_1,u_1 > - < u_2,u_2 > \right| \tag{3.18}$$

which now includes the dynamic system since it has been used to develop this equation.

3.2 MINIMAX SOLUTION

For the minimax solution we have to find the $u_2^*(t)$ that maximizes (3.18) for arbitrary $u_1(t)$ and then that $u_1^*(t)$ that minimizes this maximum cost.

Forming the functional derivative of $J(u_1,u_2)$ with respect to u_2 and setting this derivative equal to zero, we obtain

$$\frac{\partial J(u_1,u_2)}{\partial u_2} = - u_2 + T_2^* \Phi x_o - T_2^* T_1 u_1 + T_2^* T_2 u_2 = 0 \tag{3.20}$$

(where the asterisk denotes the adjoint operator) or

$$u_2 = T_2^* \Phi x_o - T_2^* T_1 u_1 + T_2^* T_2 u_2 \tag{3.19}$$

The above equation requires u_2 to be in the range of T_2^*, thus we may write

$$u_2 = T_2^* \lambda_2. \tag{3.21}$$

Making this change of variable results in

$$T_2^* \lambda_2 = T_2^* \Phi x_o - T_2^* T_1 u_1 + T_2^* T_2 T_2^* \lambda_2 \tag{3.22}$$

which will hold whenever

$$\lambda_2 = \Phi x_o - T_1 u_1 + T_2 T_2^* \lambda_2 \tag{3.23}$$

or

$$\lambda_2 = (I - T_2 T_2^*)^{-1} (\Phi x_o - T_1 u_1) \tag{3.24}$$

Thus, whenever the indicated inverse exists, the candidate extremal control u_2^* is

$$u_2^* = T_2^* (I - T_2 T_2^*)^{-1} (\Phi x_o - T_1 u_1) \tag{3.25}$$

With $T : L_2 \left[t_o, T\right] \rightarrow E^n$ defined as in Equation (3.15) by

$$(Tu)(t) = \int_{t_o}^{t} \Phi(t,\tau) G(\tau) u(\tau) d\tau \tag{3.26}$$

the inner product Tu in E^n with an arbitrary vector $\zeta \in E^n$ is

$$\langle \zeta, Tu \rangle = \left[\zeta, \int_{t_o}^{t} \Phi(t,\tau) G(\tau) u(\tau) d\tau\right]$$

$$= \int_{t_o}^{t} \left[\zeta, \Phi(t,\tau) G(\tau) u(\tau)\right] d\tau$$

(Cont'd)

$$= \int_{t_o}^{t} \left[G^*(\tau)\, \Phi^*(t,\tau)\, \xi, u(\tau) \right] d\tau$$

$$= \langle T^* \xi, u \rangle \tag{3.27}$$

Hence the adjoint operators T_1^* and T_2^* are identified by the equations

$$(T_1^* \xi_1)(t) = G_1^T(t)\, \Phi^T(T,t)\, \xi_1 \tag{3.28}$$

and

$$(T_2^* \xi_2)(t) = G_2^T(t)\, \Phi^T(T,t)\, \xi_2 \tag{3.29}$$

Thus, Equation (3.25) can be written as

$$u_2^{\circ}(t) = G_2^T(t)\Phi^T(T,t)\left[I - \int_{t_o}^{T} \Phi(T,t)G_2(t)G_2^T(t)\Phi^T(T,t)dt \right]^{-1}$$

$$\cdot \left[\Phi(T,t_o)x_o - \int_{t_o}^{T} \Phi(T,t)G_1(t)u_1(t)dt \right] \tag{3.30}$$

For $u_2^{\circ}(t)$ to be indeed locally maximizing $\frac{\partial^2 J}{\partial u_2^2} < 0$ must be satisfied. Differentiating Equation (3.19) with respect to u_2 gives

$$-I + T_2^* T_2 < 0 \tag{3.31}$$

thus requiring that

$$I - \int_{t_o}^{T} \Phi^T(t,\tau)G_2^T(\tau)G_2(\tau)\Phi(t,\tau)d\tau > 0, \quad t_o \leq t \leq T \tag{3.32}$$

In addition, no conjugate points may exist on the extremal path, which is equivalent to requiring that $(I - T_2T_2^*) > 0$

or

$$I - \int_{t_o}^{T} \Phi(t,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(t,\tau)d\tau > 0, \quad t_o \leq t \leq T \tag{3.33}$$

which assures the existence of the inverse in Equation (3.30) so $u_2^*(t)$ exists over the entire interval $[t_o, T]$. If Equation (3.33) is <u>not</u> satisfied; i.e., if there exists a time $t_s < T$ for which the matrix

$$\left[I - \int_{t_o}^{T} \Phi(t,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(t,\tau)d\tau \right] \tag{3.34}$$

becomes singular, then the control $u_2^*(t)$ is no longer maximizing for $t > t_s$.

Assuming Equation (3.32) and thus Equation (3.33) to hold, the maximizing control $u_2^*(t)$ for arbitrary $u_1(t)$ is given by Equation (3.25) or

$$u_2^* = T_2^*(I - T_2T_2^*)^{-1}(\Phi x_o - T_1u_1)$$

$$= T_2^*D_2(\Phi x_o - T_1u_1) \tag{3.35}$$

where

$$D_2 = (I - T_2 T_2^*)^{-1} \tag{3.36}$$

Substituting u_2^* into (3.18) gives the following payoff functional

$$J(u_1) = \frac{1}{2}\Big\{ \langle \Phi x_o - T_1 u_1 + T_2 T_2^* D_2 (\Phi x_o - T_1 u_1), \Phi x_o - T_1 u_1 + T_2 T_2^* D_2 (\Phi x_o - T_1 u_1) \rangle + \langle u_1, u_1 \rangle - \langle T_2^* D_2 (\Phi x_o - T_1 u_1), T_2^* D_2 (\Phi x_o - T_1 u_1) \rangle \Big\} \tag{3.37}$$

which simplifies after some work to

$$J(u_1) = \frac{1}{2}\Big\{ \langle \Phi x_o - T_1 u_1, D_2 (\Phi x_o - T_1 u_1) \rangle + \langle u_1, u_1 \rangle \Big\} \tag{3.38}$$

and it is required to determine the $u_1^*(t)$ that minimizes this payoff.

Forming the functional derivative of $J(u_1)$ with respect to u_1 and setting this derivative equal to zero gives

$$\frac{\partial J(u_1)}{\partial u_1} = u_1 - T_1^* D_2 (\Phi x_o - T_1 u_1) = 0 \tag{3.39}$$

or

$$u_1 = - T_1^* D_2 T_1 u_1 + T_1^* D_2 \Phi x_o \tag{3.40}$$

But this equation requires u_1 to be in the range of T_1^*, thus we may write

$$u_1^* = T_1^* \lambda_1 \tag{3.41}$$

and after making this change of variables we obtain

$$T_1^* \lambda_1 = -T_1^* D_2 T_1 T_1^* \lambda_1 + T_1^* D_2 \Phi x_o \tag{3.42}$$

which will hold whenever

$$\lambda_1 = - D_2 T_1 T_1^* \lambda_1 + D_2 \Phi x_o \tag{3.43}$$

or

$$\lambda_1 = \left[I + (I - T_2 T_2^*)^{-1} T_1 T_1^*\right]^{-1} \left[I - T_2 T_2^*\right]^{-1} \Phi x_o$$

$$= \left[I - T_2 T_2^* + T_1 T_1^*\right]^{-1} \Phi x_o . \tag{3.44}$$

Thus, the minimax control for player 1 is

$$u_1^{\circ} = T_1^* \left[I - T_1 T_1^* - T_2 T_2^*\right]^{-1} \Phi x_o .$$

$$= T_1^* D \Phi x_o \tag{3.45}$$

where

$$D = \left[I + T_1 T_1^* - T_2 T_2^*\right]^{-1} \tag{3.46}$$

The indicated inverse exists if $(I - T_2 T_2^*) > 0$ as required for $u_2^{\circ}(t)$.

Substituting u_1° into Equation (3.35) gives as the corresponding optimal control for player 2.

$$\begin{aligned} u_2^{\circ} &= T_2^* D_2 (\Phi x_o - T_1 T_1^* D \Phi x_o) \\ &= T_2^* D \Phi x_o \end{aligned} \tag{3.47}$$

Evaluation of (3.18), using (3.45) and (3.47) yields the minimax cost from time t_o to completion at time T as

$$J(u_1^*,u_2^*) = \frac{1}{2} < \Phi x_o, \left[I + T_1 T_1^* - T_2 T_2^*\right]^{-1} \Phi x_o > \qquad (3.48)$$

With T_1, T_2 and T_1^*, T_2^* defined by equations (3.15) and (3.28), (3.29) respectively, the minimax solution is from equations (3.45) and (3.47).

$$u_1^*(t) = G_1^T(t)\Phi^T(T,t)\left[I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\Phi^T(T,\tau)d\tau - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(T,\tau)d\tau\right]^{-1}\Phi(T,t_o)x(t_o) \qquad (3.49)$$

and

$$u_2^*(t) = G_2^T(t)\Phi^T(T,t)\left[I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\Phi^T(T,\tau)d\tau - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(T,\tau)d\tau\right]^{-1}\Phi(T,t_o)x(t_o) \qquad (3.50)$$

The minimax cost to complete the process from the arbitrary time t_o is from equation (3.48)

$$J(u_1^*,u_2^*) = \frac{1}{2}\left\{x^T(t_o)\,\Phi^T(T,t_o)\left[I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\Phi^T(T,\tau)d\tau\right.\right.$$

(Cont'd)

$$- \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\, \Phi^T(T,\tau)d\tau \Bigg]^{-1} \Phi(T,t_o)x(t_o)\Bigg\} \quad (3.51)$$

The necessary and sufficient condition for the existence of the minimax solution is from Equation (3.34)

$$I - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\, \Phi^T(T,\tau)d\tau > 0, \; t_o \leq \tau \leq T \quad (3.52)$$

3.3 MAXIMIN SOLUTION

For the maximin solution it is required to find the $u_1^*(t)$ that minimizes (3.17) for arbitrary $u_2(t)$ and then that $u_2^*(t)$ that maximizes this minimum cost.

Forming the functional derivative of $J(u_1,u_2)$, i.e., Equation (3.18), with respect to u_1 and setting this derivative equal to zero we obtain

$$\frac{\partial J(u_1,u_2)}{\partial u_2} = u_1 - T_1^* \Phi x_o + T_1^* T_1 u_1 - T_1^* T_2 u_2 = 0 \quad (3.53)$$

or

$$u_1 = T_1^* \Phi x_o - T_1^* T_1 u_1 + T_1^* T_2 u_2 \quad (3.54)$$

This equation requires u_1 to be in the range of T_1^*, thus we may write

$$u_1 = T_1^* \lambda_1 \quad (3.55)$$

Making this change of variable results in

$$T_1^* \lambda_1 = T_1^* \Phi x_o - T_1^* T_1 T_1^* \lambda_1 + T_1^* T_2 u_2. \tag{3.56}$$

which will hold whenever

$$\lambda_1 = \Phi x_o - T_1 T_1^* \lambda_1 + T_2 u_2 \tag{3.57}$$

or

$$\lambda_1 = (I + T_1 T_1^*)^{-1} (\Phi x_o + T_2 u_2) \tag{3.58}$$

The indicated inverse always exists for $0 \leq t \leq T$, because the term added to I is at least positive semidefinite. Thus, there are no conjugate point difficulties associated with the maximin solution.

Furthermore, $\dfrac{\partial^2 J}{\partial u_1^2} = I + T_1^* T_1 > 0$, thus the control u_1^* is globally minimizing and is given by

$$u_1^* = T_1^* (I + T_1 T_1^*)^{-1} (\Phi x_o + T_2 u_2) \tag{3.59}$$

or

$$u_1^* = T_1^* D_1 (\Phi x_o + T_2 u_2) \tag{3.60}$$

where

$$D_1 = (I + T_1 T_1^*)^{-1} \tag{3.61}$$

Substituting u_2^* into (3.18) gives as payoff functional

$$J(u_2) = \frac{1}{2}\Big[<\Phi x_o - T_1T_1^*D_1(\Phi x_o + T_2u_2) + T_2u_2, \Phi x_o - T_1T_1^*D_1$$

$$(\Phi x_o + T_2u_2) + T_2u_2> + <T_1^*D_1(\Phi x_o + T_2u_2), T_1^*D_1$$

$$(\Phi x_o + T_2u_2)> - <u_2,u_2>\Big] \tag{3.62}$$

which reduces after some algebra to

$$J(u_2) = \frac{1}{2}\Big[<\Phi x_o + T_2u_2, D_1(\Phi x_o + T_2u_2)> - <u_2,u_2>\Big] \tag{3.63}$$

and it is required to determine that u_2^* which maximizes this payoff.

Proceeding as before,

$$\frac{\partial J}{\partial u_2} = - u_2 + T_2^*D_1(\Phi x_o + T_2u_2) = 0 \tag{3.64}$$

or

$$u_2 = T_2^*D_1T_2u_2 + T_2^*D_1\Phi x_o \tag{3.65}$$

This equation requires u_2 to be in the range of T_2^*, thus we may write

$$u_2 = T_2^*\lambda_2 \tag{3.66}$$

and on making this change of variable, we have

$$T_2^*\lambda_2 = T_2^*D_1T_2T_2^*\lambda_2 + T_2^*D_1\Phi x_o \tag{3.67}$$

which will hold whenever

$$\lambda_2 = D_1 T_2 T_2^* \lambda_2 + D_1 \Phi x_o \tag{3.68}$$

or

$$\lambda_1 = \left[I - (I + T_1 T_1^*)^{-1} T_2 T_2^*\right]^{-1} (I + T_1 T_1^*)^{-1} \Phi x_o.$$

$$= \left[I + T_1 T_1^* - T_2 T_2^*\right]^{-1} \Phi x_o. \tag{3.69}$$

Hence the maximin control for player 2 is

$$u_2^* = T_2^* \left[I + T_1 T_1^* - T_2 T_2^*\right]^{-1} \Phi x_o.$$

$$= T_2^* D \Phi x_o \tag{3.70}$$

Substituting u_2^* into Equation (3.59) gives

$$u_1^* = T_1^* D_1 \Phi x_o + T_1^* D_1 T_2 T_2^* D \Phi x_o$$

$$= T_1^* D \Phi x_o \tag{3.71}$$

Evaluation of (3.17), using (3.70) and (3.71) yields the maximin cost from time t_o to completion at time T as,

$$J(u_1^*, u_2^*) = \frac{1}{2} \left\langle \Phi x_o, \left[I + T_1 T_1^* - T_2 T_2^*\right]^{-1} \Phi x_o \right\rangle \tag{3.72}$$

With T_1, T_2 and T_1^*, T_2^* defined by Equations (3.15) and (3.28), (3.29) respectively, the <u>maximin solution</u> is from equations (3.70) and (3.71)

$$u_1^*(t) = G_1^T(t)\,\Phi^T(T,t)\left[I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\,\Phi^T(T,\tau)d\tau \right.$$

$$\left. - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\,\Phi^T(T,\tau)d\tau\right]^{-1} \Phi(T,t_o)x_o \qquad (3.73)$$

$$u_2^*(t) = G_2^T(t)\,\Phi^T(T,t)\left[I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\,\Phi^T(T,\tau)d\tau \right.$$

$$\left. - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\,\Phi^T(T,\tau)d\tau\right]^{-1} \Phi(T,t_o)x_o. \qquad (3.74)$$

The <u>maximin cost</u> to complete the process from the arbitrary time t_o is from Equation (3.72).

$$J(u_1^*,u_2^*) = \frac{1}{2}\left\{x^T(t_o)\,\Phi^T(T,t_o)\left[I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\,\Phi^T(T,\tau)d\tau \right.\right.$$

$$\left.\left. - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\,\Phi^T(T,\tau)d\tau\right]^{-1} \Phi(T,t_o)x(t_o)\right\} \qquad (3.75)$$

while the necessary and sufficient condition for the existence of the maximin solution is from Equation (3.69)

$$I + \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\,\Phi^T(T,\tau)d\tau - \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(T,\tau)d\tau > 0$$

$$t_o \le t \le T \qquad (3.76)$$

3.4 DISCUSSION

Comparison of the minimax solution (Equations (3.49) through (3.52) with the maximin solution (Equations (3.73) through (3.76)) shows that the solutions are identical. Hence, we have obtained the saddle point solution to the two-person zero-sum game, i.e.,

$$J(u_1^*, u_2) \leq J(u_1^*, u_2^*) \leq J(u_1, u_2^*) \tag{3.77}$$

If we define the symmetric matrices $M(T,t_o)$, $M_1(T,t_o)$ and $M_2(T,t_o)$ as

$$\begin{aligned} M(T,t_o) &= I + M_1(T,t_o) - M_2(T,t_o) \\ M_1(T,t_o) &= \int_{t_o}^{T} \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\,\Phi^T(T,\tau)d\tau \\ M_2(T,t_o) &= \int_{t_o}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\,\Phi^T(T,\tau)d\tau \end{aligned} \tag{3.78}$$

we can write the optimal solutions as

$$u_1^*(t) = G_1^T(t)\,\Phi^T(T,t)M^{-1}(T,t_o)\,\Phi(T,t_o)x_o \tag{3.79}$$

$$u_2^*(t) = G_2^T(t)\,\Phi^T(T,t)M^{-1}(T,t_o)\,\Phi(T,t_o)x_o \tag{3.80}$$

and it is obvious that the optimal controls are proportional to $\Phi(T,t_o)x_o$ which is the terminal miss if both controllers remain inactive and the system is allowed to run free. The time varying matrices reflect the control capabilities of both players.

From optimal control theory (Section 2.2), we know that the necessary and sufficient condition for the system to be controllable on $\left[t_o, T\right]$ by controller 1 with $u_2^*(\cdot) \equiv 0$ is

$$M_1(T,t_o) = \int_{t_o}^{T} \Phi(T,t)G_1(t)G_1^T(t)\,\Phi^T(T,t)dt > 0 \tag{3.81}$$

for all t in $\left[t_o, T\right]$, while the necessary and sufficient condition for the system to be controllable on $\left[t_o, T\right]$ by controller 2 with $u_1^*(\cdot) \equiv 0$ is

$$M_2(T,t_o) = \int_{t_o}^{T} \Phi(T,t)G_2(t)G_2^T(t)\,\Phi^T(T,t)dt > 0 \tag{3.82}$$

and we can define $M_1(T,t_o)$ and $M_2(T,t_o)$ as the reduced controllability matrices of player 1 and player 2 respectively.

The conditions for the existence of $M^{-1}(T,t_o)$ obtained in the maximin solution provides additional insight into the problem if we consider the limiting case of weighing the importance of terminal miss against control effort. In this case the payoff functional is written as

$$J(u_1,u_2) = \frac{1}{2}\left\{a^2x^T(T)x(T) + \int_{t_o}^{T}\left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right]dt\right\} \tag{3.83}$$

where the scalar a permits the required weighting, and the resulting $M(T,t_o)$ is then

$$M(T,t_o) = \frac{I}{a^2} + M_1(T,t_o) - M_2(T,t_o) \tag{3.84}$$

In the limiting case, i.e., $a^2 \to \infty$ in the sense that

$$\frac{a^2}{2}\,(x^T(T)x(T)) = \begin{cases} 0 & \text{if } x(T) = 0 \\ \\ \infty & \text{if } x(T) \neq 0 \end{cases} \tag{3.85}$$

the existence of $M(T,t_o)$ is guaranteed if

$$M_r(T,t_o) \stackrel{\Delta}{=} M_1(T,t_o) - M_2(T,t_o) > 0 \tag{3.86}$$

$M_r(T,t_o)$ is known as the relative controllability matrix, [14] the fact that it is positive definite indicates that the minimizing player, player 1, is "more controllable" than the maximizing player, player 2.

The initial time t_o is completely arbitrary, while the assumption of perfect information guarantees that $x(t)$ is available for any t. Hence the open-loop controls can be applied continuously and instantaneously to yield optimal feedback control laws by replacing t_o by t.

If we define

$$S(t) \stackrel{\Delta}{=} \Phi^T(T,t)\left[I + \int_t^T \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\Phi^T(T,\tau)d\tau \right.$$
$$\left. - \int_t^T \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(T,\tau)\right]^{-1}\Phi(T,t) \tag{3.87}$$

then we can write the optimal feedback controls for the linear-quadratic two-person differential game as

$$u_1^*(t) = G_1^T(t)S(t)x(t) \tag{3.88}$$

$$u_2^*(t) = G_2^T(t)S(t)x(t) \tag{3.89}$$

and the optimal cost to complete the game from the arbitrary time t is

$$J(u_1^*, u_2^*) = \frac{1}{2} x^T(t)S(t)x(t) \tag{3.90}$$

while the necessary and sufficient condition for existence of the solution is from the maximizing step of the minimax solution

$$I - \int_t^T \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\,\Phi^T(T,\tau)d\tau > 0; \quad t \leq \tau \leq T \tag{3.91}$$

This necessary and sufficient condition is more stringent than that of Ho, Bryson and Baron [14] and Rhodes and Luenberger [10] who claim

$$I + \int_t^T \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\,\Phi^T(T,\tau)d\tau$$
$$- \int_t^T \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\,\Phi^T(T,\tau)d\tau > 0; \quad t \leq \tau \leq T \tag{3.92}$$

as the necessary and sufficient solution. The difference occurs because their mathematics is conditioned by the a priori assumption of a saddle point solution.

With $S(t)$ defined as in Equation (3.87) we can determine the controls $u_1^*(t)$ and $u_2^*(t)$ from the matrix Riccati equation, developed below. Taking the derivative of $S(t)$ we obtain

$$\dot{S}(t) = \frac{\partial}{\partial t} \Phi^T(T,t)D(t)\Phi(T,t) + \Phi^T(T,t)\frac{\partial}{\partial t} D(t)\Phi(T,t)$$

$$+ \Phi^T(T,t)D(t)\frac{\partial}{\partial t}\Phi(T,t) \tag{3.93}$$

where

$$D(t) = \left[I + \int_t^T \Phi(T,\tau)G_1(\tau)G_1^T(\tau)\Phi^T(T,\tau)d\tau \right.$$

$$\left. - \int_t^T \Phi(T,\tau)G_2(\tau)\Phi^T(T,\tau)d\tau \right]^{-1} \tag{3.94}$$

But

$$\frac{\partial}{\partial t}\Phi^T(T,t) = \left[\frac{\partial}{\partial t}\Phi^{-1}(t,T)\right]^T = -\left[\Phi^{-1}(t,T)\frac{\partial}{\partial t}\Phi(t,T)\Phi^{-1}(t,T)\right]^T$$

$$= -F^T(t)\Phi^T(T,t); \tag{3.95}$$

$$\frac{\partial}{\partial t}\Phi(T,t) = \frac{\partial}{\partial t}\Phi^{-1}(t,T) = -\Phi(T,t)F(t) \tag{3.96}$$

and

$$\frac{\partial}{\partial t} D(t) = -D(t)\frac{\partial}{\partial t}D^{-1}(t)D(t) = D(t)\left[\Phi(T,t)G_1(t)G_1^T(t)\Phi^T(T,t)\right.$$

$$\left. - \Phi(T,t)G_2(t)G_2^T(t)\Phi^T(T,t)\right] D(t) \tag{3.97}$$

Substituting Equations (3.95) through (3.97) into Equation (3.93) we obtain the matrix Riccati equation

$$\dot{S}(t) = - S(t)F(t) - F^T(t)S(t) + S(t)\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right]S(t) \tag{3.98}$$

with boundary condition

$$S(T) = I \tag{3.99}$$

Note that the solution to the above equation can be obtained prior to the actual game. A summary of the optimal strategies for the linear-quadratic differential game with perfect information is presented in Table 3.1.

TABLE 3.1

SUMMARY OF OPTIMAL DETERMINISTIC STRATEGIES

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t); \quad x(t_o) = x_o$$

Player 1: Perfect measurements

Player 2: Perfect measurements

$$J = \frac{1}{2}\left\{x^T(T)x(T) + \int_{t_o}^{T} \left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right] dt\right\}$$

$$u_1^*(t) = G_1(t)S(t)x(t)$$

$$u_2^*(t) = G_2(t)S(t)x(t)$$

$$\dot{S} = -SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right]S; \quad S(T) = I$$

$$J(u_1^*, u_2^*) = \frac{1}{2}x^T(t)S(t)x(t)$$

Necessary and sufficient conditions

$$I - \int_{t}^{T} \Phi(T,\tau)G_2(\tau)G_2^T(\tau)\Phi^T(T,\tau)d\tau > 0$$

CHAPTER 4

INTRODUCTION TO STOCHASTIC DIFFERENTIAL GAMES AND DELAYED COMMITMENT STRATEGIES

For the differential game considered so far, we have assumed that we could make noiseless measurements of the system state vector and use those measurements in the system mechanization, i.e., we assumed a differential game with perfect information.

In many practical situations, however, the players have access only to noisy measurements, resulting in a game with imperfect information. Willman [8] has given a formal solution to this class of games, but, as an apparent consequence of this imperfect information, attempts to express these strategies in terms of finite-dimensional estimate vectors have been unsuccessful. A version of this game in which constraints are placed on the player's state estimators has been solved by Rhodes and Luenberger [17] .

A subclass of games with imperfect information where one of the player's measurements are corrupted by white noise and the other player has perfect measurements was solved in 1968 by Behn and Ho [9] for a pursuit-evasion game and in 1969 by Rhodes and Luenberger [10] for a more general game.

Harsanyi [18] , in 1967/1968, used a chance move as a mathematical device in the analysis of static games with imperfect information to reformulate the game into a game with perfect information, called the "Bayes-equivalent" of the original game. The players enter the game, so to speak, after chance has made its choice. In part II

[19], Harsanyi recognizes that this time gap is crucial when cooperative games with imperfect information are being played and shows that the normal form of a Bayesian game is, in many cases, a highly unsatisfactory representation of the game situation. He argued that the Bayesian games must be interpreted as games with delayed commitment.

In 1972, Aumann and Maschler [7] pointed out that the difficulties due to the time gap exist even if the players are playing a two-person zero-sum game with imperfect information and Ho [21] extended their results to stochastic two-person games.

In this chapter, we will define the differential game problem in which the two opposing players have access only to noise-corrupted output measurements and introduce the delayed commitment strategies via a simple example of a one-stage stochastic difference game.

4.1 GAMES WITH IMPERFECT STATE INFORMATION

As pointed out in Chapter 2, if the output measurements are corrupted by a random process we are faced with a stochastic problem. In order for a stochastic game to be mathematically tractable, the measurement noise must be describable by a finite set of sufficient statistics. In practice this means a linear system with quadratic cost and Gaussian noises corrupting the output measurements. The sufficient statistics are then the mean and covariance of the process.

Consider the linear system described by the vector differential equation

$$\dot{x}(t) = \frac{dx}{dt} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \qquad (4.1)$$

to which player 1, controlling $u_1(t)$, has available measurements of the form

$$z_1(t) = H_1(t)x(t) + w_1(t), \tag{4.2}$$

while player 2, controlling $u_2(t)$ has available the measurements

$$z_2(t) = H_2(t)x(t) + w_2(t) \tag{4.3}$$

The vector $x(t) \in E^n$ is the system state, $u_1(t) \in E^{P1}$ and $u_2(t) \in E^{P2}$ are the control vectors, $z_1(t) \in E^{m1}$ and $z_2(t) \in E^{m2}$ are the measurement vectors. The matrices $F(t)$, $G_1(t)$ and $G_2(t)$ have the appropriate dimensions, while the matrices $H_1(t)$ and $H_2(t)$ are respectively, $m_1 \times n$ and $m_2 \times n$ with $m_1, m_2 \leq n$. The noise processes $\left| w_1(t) \right|$ and $\left| w_2(t) \right|$ are white Gaussian, with properties

$$\begin{aligned} \operatorname{cov}\left[w_1(t), \quad w_1(\tau) \right] &= W_1(t)\,\delta(t-\tau) \\ \operatorname{cov}\left[w_2(t), \quad w_2(\tau) \right] &= W_2(t)\,\delta(t-\tau) \\ \operatorname{cov}\left[w_1(t), \quad w_2(\tau) \right] &= 0 \end{aligned} \tag{4.5}$$

The initial state $x(t_o)$ is a Gaussian random vector, uncorrelated for all t with $w_1(t)$ and $w_2(t)$, and having a mean of $\bar{x}_o$ and a covariance

$$\operatorname{cov}\left[x(t_o), \ x(t_o) \right] = P_o \tag{4.6}$$

The cost functional or payoff to the game is quadratic:

$$J(u_1, u_2) = \frac{1}{2} E \left\{ x^T(T)x(T) + \int_{t_o}^{T} u_1^T(t)u_1(t)dt - \int_{t_o}^{T} u_2^T(t)u_2(t)dt \right\} \tag{4.7}$$

where the final time T is fixed and the expectation is taken over all the underlying random quantities $(x(t_o), w_1(t), w_2(t))$. The simplified form of the payoff functional has been assumed, which can be obtained from a more general formulation using the transformation equations (3.4) of Chapter 3.

Let us now turn to the admissible strategies. Let $Z_i(t)$, $i = 1, 2$ be the output function measured by player i over the interval $[t_o, t)$, i.e.,

$$Z_i(t) = \left\{ (z_i(s), s) \; : \; s \in [t_o, t) \right\} \tag{4.8}$$

the class of admissible strategies are then restricted to those U_1 and U_2 which give rise to the feedback control laws

$$U_1 \;\; : \;\; u_1 = u_1(Z_1(t), t)$$

$$U_2 \;\; : \;\; u_2 = u_2(Z_2(t), t) \tag{4.9}$$

Thus, the admissible strategies can only depend on the past accumulative observation data. Equation (4.9) can be expressed equivalently for $i = 1, 2$ as

$$u_i(t) = \hat{\phi}_i(t, z_{t_i}); \; z_{t_i} \triangleq \left\{ z_i(s), s \in \left[t_o, t \right] \right\} \tag{4.10}$$

where $\hat{\phi}(\cdot \, , \, \cdot)$ is viewed as a mapping from $R \times C_m \left[t_o, T \right] \longrightarrow R^p$ and $C_m \left[t_o, T \right]$ is the class of continuous functions defined on $\left[t_o, T \right]$ with values in R^m.

The mapping $\phi_i(t, \cdot)$ for $i = 1, 2$ satisfies a Lipschitz condition:

$$\| \hat{\phi}(t,f) - \hat{\phi}(t,g) \| \leq \alpha \| f - g \|, \quad f,g \in C_m[t_o,T] \tag{4.11}$$

for all $t \in [t_o,T]$ where α is some constant. The Lipschitz condition is imposed for technical reasons; it gives a sufficient condition for the existence of $(x(t),z_1(t))$, $(x(t),z_2(t))$ in (4.1) through (4.3).

When each controller is allowed either perfect measurements or noise-corrupted measurements , a total of four problems may be formulated, of which, due to symmetry, three are basically different. Figure 4.1 shows the problem classification and indicates those discussed in this paper, together with some references to previous papers which examined solutions to those problems.

Player 2 / Player 1	Perfect Measurements	Noisy Measurements
Perfect Measurements	Closed-loop Game Chapter 3 [14, 15]	Chapters 5, 6 [9, 10]
Noisy Measurements		Chapter 7 [8, 17, 22]

Figure 4.1 Problem Classification

As shown in Chapter 3, the necessary and sufficient condition for a solution of the two person zero-sum game with perfect information is more stringent than previously determined. The solutions obtained so far to the games with noisy measurements are valid only under restricted circumstances. To discuss those restrictions and to develop the general solution concept to games with measurement noise we will now turn to our tutorial example.

4.2 A TUTORIAL EXAMPLE

The concept of delayed commitment strategies to stochastic differential games can best be explained, it is felt, by using a very simple one stage stochastic difference game example similar to that presented by Ho [21] . The notation $N(1, \sigma)$ indicates a normal process with a mean and a covariance equal to 1 and σ respectively.

Consider the scalar dynamic system

$$x_3 = x_2 - u_2 = x_1 + u_1 - u_2 \tag{4.12}$$

where $x_1 = x \sim N(0, \sigma)$, and u_1 and u_2 are the controls of player 1 and 2 respectively.

Consider also the performance criterion

$$J = \frac{1}{2} E \left| x_3^2 - x_1^2 + u_1^2 - 2u_2^2 \right| \tag{4.13}$$

which player 1 attempts to minimize and player 2 to maximize. Player 1 receives no measurements, while player 2 is given the measurements

$$z_2 = x + w_2, \qquad w_2 \sim N(0, 1) \tag{4.14}$$

where x and w_2 are independent. The class of admissible strategies for player 1 is

$$U_1 : u_1 = k_1 = \text{constant} \tag{4.15}$$

and for player 2 is

$$U_2 : u_2 = k_2 z_2 \; . \tag{4.16}$$

We can obtain the prior commitment strategy by substituting (4.12) into (4.13) which gives

$$J(u_1,u_2) = \frac{1}{2} E \left| 2u_1^2 - u_2^2 + 2xu_1 - 2xu_2 - 2u_1u_2 \right| \tag{4.17}$$

Then for

$$\begin{aligned} u_1 &= k_1 \\ u_2 &= k_2 z_2 \end{aligned} \tag{4.18}$$

we can write

$$\begin{aligned} J(k_1,k_2) &= \frac{1}{2} E \left| 2k_1^2 - k_2^2 z_2^2 + 2k_1 x - 2k_2 x z_2 - 2k_1 k_2 z_2 \right| \\ &= \frac{1}{2} \left| 2k_1^2 - k_2^2(\sigma + 1) - 2k_2 \sigma \right| \end{aligned} \tag{4.19}$$

Both the minimax and maximin solutions to this simple problem give

$$\begin{aligned} u_1^* &= 0 \\ u_2^* &= - \frac{\sigma}{\sigma + 1} z_2 \end{aligned} \tag{4.20}$$

Thus, u_1^*, u_2^* form a saddle point pair and

$$J(u_1^*, u_2^*) = \frac{\sigma 2}{2(\sigma + 1)} \tag{4.21}$$

and it has been assumed that u_1^*, u_2^* form <u>the</u> solution to the differential game.

However, consider the situation facing player 2 during the actual play of the game, <u>after</u> he has received the information z_2 and <u>before</u> anyone has acted. Player 2 now faces the payoff

$$J_2(u_1, u_2) = \frac{1}{2} E \left| 2u_1^2 - u_2^2 + 2xu_1 - 2xu_2 - 2u_1u_2 \mid z_2 \right| \tag{4.22}$$

and the secure strategy of this maximizing player is obtained by finding the maximin solution of Equation (4.22) subject to equations (4.12), (4.15) and (4.16).

For arbitrary u_2 the minimizing strategy u_1 obtained from the partial derivative of J_2 with respect to u_1 is

$$u_1 = \frac{1}{2} (u_2 - \hat{x}_2) \tag{4.23}$$

where

$$\hat{x}_2 = E \left| x \mid z_2 \right| \tag{4.24}$$

Substituting this result into Equation (4.22) gives

$$J_2(u_2) = \frac{1}{2} E \left| - \frac{3}{2} u_2^2 - xu_2 - x\hat{x}_2 + \frac{1}{2} \hat{x}_2^2 \mid z_2 \right| \tag{4.25}$$

and the maximizing u_2 is found to be

$$u_2^* = -\frac{1}{3}\hat{x}_2 = -\frac{1}{3}\frac{\sigma}{\sigma+1}z_2 \qquad (4.26)$$

and thus

$$u_2^* = -\frac{2}{3}\hat{x}_2 = -\frac{2}{3}\frac{\sigma}{\sigma+1}z_2 \qquad (4.27)$$

The resulting maximin solution is thus

$$u_1^* = -\frac{2}{3}E\left|x\mid z_2\right| \qquad (4.28)$$

$$u_2^* = -\frac{1}{3}E\left|x\mid z_2\right| \qquad (4.29)$$

Since z_2 can be regarded as part of the prior information and thus is a known number, u_1 and u_2 satisfy the restriction on the class of admissible strategies.

An analogous argument shows that the minimax strategies are the same as the maximin strategies. Hence u_1^* and u_2^* are not just the maximin solution, but they are a saddle point pair for J_2, i.e.,

$$J_2(u_1^*,u_2) \le J_2(u_1^*,u_2^*) \le J_2(u_1,u_2^*) \qquad (4.30)$$

and the resulting payoff is

$$J_2(u_1^*,u_2^*) = -\frac{1}{6}\frac{\sigma 2}{(\sigma+1)^2}z_2^2 \qquad (4.31)$$

On the other hand, if player 1 uses strategy u_1^* and player 2 uses u_2°, then

$$J_2(u_1^*,u_2^\circ) = -\frac{13}{72}\frac{\sigma^2}{(\sigma+1)^2} z_2^2, \tag{4.32}$$

Obviously, $J_2(u_1^*,u_2^*) \geq J_2(u_1^*,u_2^\circ)$ and we conclude that for all possible values of the observation z_2, the strategy u_2^* is actually a safer strategy for player 2 than u_2°.

The reason for this phenomena as first pointed out by Harsanyi [19] and then by Aumann and Machler [7] is inherent in the Normalization Principle of game theory. In the extensive form of the game a player makes his decision as to what control to use after receiving his measurements, while in the normal form of the game, this decision is effectively moved to before receiving those measurements.

In many games, the passage from the extensive to the normal form does not affect the course of action of the players and the two situations are formally equivalent. But, in our game, with imperfect information this passage changes the outlook of player 2. Indeed, if player 2 decides on a strategy before receiving the measurement z_2, he is justified in using the expected value of z_2 in his payoff function. However, when player 2 is informed, before making his decision, that a specific z_2 has been selected, there is no longer any justification for using the expected value of z_2. Thus, after the information is received, we really have a non zero-sum game facing the two players, with (4.17) the payoff for player 1 and (4.22) the payoff for player 2. It is this change in outlook that is ignored in the passage from the extensive to the normal form of the game.

In terms of Harsanyi's discussion the players "enter" the game after the "chance" (the measurement noise) has made its choice. During the play of a stochastic differential (difference) game at time t or t_k greater than t_o, the players effectively also "enter" the game having received the actual measurement (noise corrupted) up to that time.

Returning to our example, if player 2 has reason to believe that player 1 is committed to the strategy $u_1^* = 0$, then on solving the resulting one sided optimal control problem from player 2's point of view,

$$\max_{u_2} \quad J_2(u_1^*,u_2) = J_2(0,u_2) = \frac{1}{2} E \left| - u_2^2 - 2xu_2 \mid z_2 \right| \tag{4.33}$$

gives

$$u_2^* = - E \left| x \mid z_2 \right| = - \frac{\sigma}{\sigma + 1} z_2 \tag{4.34}$$

Similarly, if player 2 is committed to $u_2^* = - \frac{\sigma}{\sigma + 1} z_2$, the solution to the resulting one sided optimal control problem from player 1's point of view is

$$\min_{u_1} \quad J_1(u_1,u_2^*) = \frac{1}{2} E \left| u_1^2 - \frac{\sigma 2}{(\sigma + 1)^2} z_2^2 + 2 xu_1 \right.$$

$$\left. + 2 \frac{\sigma}{\sigma + 1} xz_2 + 2 \frac{\sigma}{\sigma + 1} u_1 z_2 \right|$$

(Cont'd)

$$= \frac{1}{2}\left[u_1^{\ 2} + \frac{\sigma 2}{\sigma + 1}\right] \tag{4.35}$$

which gives

$$u_1^{\bullet} = 0$$

Thus, the strategy pair $\left|u_1^{\bullet}, u_2^{\bullet}\right|$ is a Nash equilibrium solution to the non zero - sum game. Hence if player 2 knows a priori that player 1 will use $u_1^{\bullet}$, or if he can convince player 1 that he is using $u_2^{\bullet}$, then his optimal strategy will be $u_2^{\bullet}$. However, it is well known that Nash equilibrium strategies do not possess any minimax or guaranteed value properties in non zero - sum games, and without this a priori knowledge there is <u>no reason</u> at all to play $u_2^{\bullet}$ when u_2^{*} is safer and available.

CHAPTER 5

THE PERFECT/NOISY DIFFERENTIAL GAME

In this chapter we discuss the case where one of the players has perfect state information while the other player has only noisy measurements of the state. A physical example of such a problem would be the pursuit-evasion problem of a homing missile and an evading aircraft where the missile has considerable ground support via an up- and downlink to determine the state of the evader.

The problem is basically the same as that solved by Behn and Ho [9] as a pursuit-evasion differential game and extended by Rhodes and Luenberger [10] to more general differential games. Their solutions, however, are prior commitment solutions and assume that conditions are such that the player with perfect state information can deduce exactly, at each time t, the error in his opponent's state estimate.

In Section 1, of this chapter, we formulate the stochastic problem and discuss the prior commitment solution. The delayed commitment strategies for player 1 and player 2 are then obtained in Section 2 using function space techniques. It is then shown that the results can be interpreted in terms of matrix differential equations of the Riccati type.

The delayed commitment strategy optimality criteria are discussed in Section 3 and compared with those of the prior commitment strategy. The chapter is then concluded with a summary and discussion of the results obtained for the perfect/noise corrupted two-person differential game.

5.1 PROBLEM FORMULATION AND PRIOR COMMITMENT SOLUTION

The problem formulation differs only slightly from that presented in Chapter 4, in that only one player has noise corrupted measurements of the state vector, $x(t)$, during the game and an estimate of the initial condition, while the other player has perfect state information during the entire game. Thus, the linear continuous time dynamic system is described by the vector differential equation

$$\dot{x} = \frac{dx}{dt} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \tag{5.1}$$

and the quadratic cost functional is

$$J(u_1,u_2) = \frac{1}{2} E\left\{ x^T(T)x(T) + \int_{t_o}^{T} \left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right] dt \right\} \tag{5.2}$$

where the dimensions for the vectors and matrices are as discussed in Section 4.1 and the final time T is fixed.

Player 1 has perfect measurements of the state $x(t)$, while the measurements available to player 2 are of the form

$$z_2(t) = H_2(t)x(t) + v_2(t) \tag{5.3}$$

where the matrix $H_2(t)$ is $m_2 \times n$ with $m_2 \leq n$. The noise $v_2(t)$ is assumed white, zero-mean and Gaussian with covariance

$$\text{cov}\left[v_2(t), v_2(\tau) \right] = W_2(t)\, \delta(t - \tau) \tag{5.4}$$

The initial state $x(t_o)$ for player 2 is assumed to be a Gaussian random vector uncorrelated with $w_2(t)$ for all time $t \in \left[t_o, T\right]$ and having a mean $\bar{x}_o$ and covariance

$$\operatorname{cov}\left[x(t_o), x(t_o)\right] = P_o \tag{5.5}$$

The initial state for player 1 is $x(t_o) = x_o$.

Let $Z_2(t)$ be the output function measured by player 2 over the interval $[t_o, t)$, i.e.,

$$Z_2(t) = \left\{(z_2(s), s) : s \in [t_o, t)\right\} \tag{5.6}$$

The class of admissible strategies are restricted to those U_1 and U_2 which give rise to feedback control laws, i.e.,

$$\begin{aligned} U_1 &: u_1 = u_1(x(t), t) \\ U_2 &: u_2 = u_2(Z_2(t), t) \end{aligned} \tag{5.7}$$

Let the best linear estimate of the system state $x(t)$ given the measured output function $Z_2(t)$ be denoted $\hat{x}_2(t)$, i.e.,

$$\hat{x}_2(t) \triangleq E\left[x(t) \mid Z_2(t)\right] \tag{5.8}$$

The corresponding estimation error $\tilde{x}_2(t)$ is then

$$\tilde{x}_2(t) \triangleq x(t) - \hat{x}_2(t) \tag{5.9}$$

Since the random variables are normally distributed, the best linear estimate will also be the overall optimal estimate.

Previous prior commitment solutions require that conditions are such that the player with perfect state information (player 1) can deduce exactly at each time $t \in [t_o, T]$ the error in his opponent's state estimate, $\tilde{x}_2(t)$, or that this information is provided by some "mystical third party."

In the more general case, where player 1 cannot calculate nor is provided with $\tilde{x}_2(t)$, or equivalently $\hat{x}_2(t)$ from which $\tilde{x}_2(t) = x(t) - \hat{x}_2(t)$, he will have to build a filter from which he generates an estimate of his opponent's estimate, denoted $\hat{x}_{21}(t)$. Obviously, $\hat{x}_{21}(t)$, based on noisy data, will deviate from $\hat{x}_2(t)$ and player 2 should be able to take advantage of this error in player 1's estimate of the estimate of player 2, leading effectively to an additional term in player 2's control. However, such an additional correction term is based on noisy data and the opponent, player 1, should be able to take advantage of this error. However, the correction of player 1, in turn, is based on noisy data and player 2 should be able........... .

What we have just encountered, if the general problem is solved from the prior commitment point of view, is known as the <u>closure problem</u> in stochastic games. It expresses the fact that an infinite number of terms seem to be required in the optimal strategies of each of the two players.

For the differential game defined by Equations (5.1) through (5.7), and under the assumption that player 1 can determine exactly the error in player 2's estimate, $\tilde{x}_2(t)$, the prior commitment optimal strategies obtained by Behn and Ho and Rhodes and Luenberger are*

$$u_1'(t) = G_1^T \left[S(t)x(t) + N(t)\tilde{x}(t) \right]$$
$$u_2'(t) = G_2^T(t)S(t)\hat{x}(t) \tag{5.10}$$

where the symmetric gain matrix $S(t)$ satisfies the matrix Riccati equation

$$\dot{S} = -SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t) \right] S \tag{5.11}$$

with boundary condition

$$S(T) = I \tag{5.12}$$

and the symmetric gain matrix $N(t)$ satisfies the differential equation

$$\begin{aligned}\dot{N} = NF(t) - F^T(t)N - S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t) \right] S \\ + (S + N)G_1(t)G_1^T(t)(S + N) + NP(t)H_2^T(t)H_2(t) \\ + H_2^T(t)W_2^{-1}(t)H_2(t)P(t)N\end{aligned} \tag{5.13}$$

with boundary condition

$$N(T) = 0 \tag{5.14}$$

The symmetric error covariance matrix $P(t)$ satisfies

---

* In order to avoid confusion with the state estimate in the delayed commitment solution discussed below, the subscripts have been omitted from the state estimates and their errors in the prior commitment game.

$$\dot{P} = A(t)P + PA^T(t) - PH_2^T(t)W_2^{-1}(t)H_2(t)P \tag{5.15}$$

with boundary condition

$$P(t_o) = \text{cov}\left[x(t_o),x(t_o)\right] = P_o \tag{5.16}$$

where the matrix A(t) is defined by

$$A(t) = F(t) - G_1(t)G_1^T(t)\left[S(t) + N(t)\right] \tag{5.17}$$

Note that Equations (5.15) and (5.13) are coupled, so that the solution of this problem involves a nonlinear two-point boundary value problem given by these equations with boundary conditions (5.14) and (5.16). The solutions of the matrix Riccati type equations, i.e., Equations (5.11) through (5.17) can be obtained off-line, prior to the actual game.

The corresponding optimal expected cost from time t is

$$J(u_1',u_2') = \frac{1}{2}\ x^T(t)S(t)x(t) + \frac{1}{2}\tilde{x}^T(t)N(t)\tilde{x}(t)$$

$$+ \frac{1}{2}\ \text{tr}\left\{\int_t^T N(s)P(s)H_2^T(s)W_2^{-1}(s)H_2(s)P(s)ds\right\} \tag{5.18}$$

where tr $\left|\cdot\right|$ is the trace operator.

5.2 DELAYED COMMITMENT STRATEGIES

During the actual play of the game at time t, and from the point of view of player 1, the payoff functional becomes

$$J_1(u_1,u_2) = \frac{1}{2} E\left\{x^T(T)x(T) + \int_t^T \left[u_1^T(\tau)u_1(\tau) - u_2^T(\tau)u_2(\tau)\right] d\tau \mid X(t)\right\} \tag{5.19}$$

where

$$X(t) \triangleq \left\{(x(s), s) \ : \ s \ [t_o,t)\right\} \tag{5.20}$$

and while, as pointed out in Chapter 4, the strategy pair (u_1', u_2') presented in Section 5.1 still retains its equilibrium property, they are no longer secure strategies. In order for player 1 to determine his <u>secure</u> strategy, he has to find the saddle-point solution to Equation (5.19) subject to

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t); \quad x(t_o) = x_o \tag{5.21}$$

Player 2 is faced with the problem of extracting useful information from his past measurements on which to base his control. However, player 2's perfect estimate is $\hat{x}_2(t) = x(t)$ and for the purpose of calculating player 1's secure strategy we assume that the allowable strategy for player 2, in addition to being $Z_2(t)$ measurable, is also $X(t)$ measurable. In other words, we want to determine that $u_1^* \epsilon\ U_1$ and $u_2^* \epsilon\ U_1 \times U_2$ which are optimal in the sense that for all $t \epsilon \left[t_o,T\right]$

$$J_1(u_1^*,u_2) \leq J_1(u_1^*,u_2^*) \leq J_1(u_1,u_2^*) \tag{5.23}$$

where

$$U_1 \; : \; u_1 = u_1(x(t),t)$$
$$U_2 \; : \; u_2 = u_2(Z_2(t),t) \tag{5.24}$$

The delayed commitment game from player 1's point of view is then the same as that solved in Section (3.2) for which we obtained the saddle-point solution

$$u_1^*(t) = G_1^T(t)S(t)x(t) \tag{5.25}$$

with the corresponding optimal response for player 2

$$u_2^*(t) = G_2^T(t)S(t)x(t) \tag{5.26}$$

where $S(t)$ is the solution to Equation (5.11), i.e.,

$$\dot{S} = -SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right]S; \quad S(T) = I \tag{5.27}$$

The resulting security payoff for player 1, i.e., his <u>loss ceiling</u> at arbitrary time t is from Equation (3.90)

$$J_1(u_1^*,u_2^*) = \frac{1}{2}\,x(T)S(t)x(t) \tag{5.28}$$

Note that in real life when player 2 does not have a perfect estimate of the state $x(t)$, the payoff to player 1 can only be better, i.e., smaller than his loss ceiling.

If we now consider the game from the point of view of player 2, his payoff during the actual play of the game at time t is

$$J_2(u_1,u_2) = \frac{1}{2}\,E\left\{x^T(T)x(T) + \int_t^T \left[u_1^T(\tau)u_1(\tau) - u_2^T(\tau)u_2(\tau)\right] d\tau \,|\, Z_2(t)\right\} \tag{5.29}$$

and his secure strategy is obtained by finding the saddle-point solution to Equation (5.29) subject to

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \quad ; \quad x(t_o) = \bar{x}_o \tag{5.30}$$

For the purpose of determining the secure strategy of player 2, we assume that the allowable strategy for player 1 in addition to being $X(t)$ measurable is also $Z_2(t)$ measurable. Thus, player 2 wants to determine that $u_1^* \epsilon\ U_1 \times U_2$ and $u_2^* \epsilon\ U_2$ which are optimal in the sense that for all $t \epsilon \left| t_o, T \right|$.

$$J_2(u_1^*, u_2) \leq J_2(u_1^*, u_2^*) \leq J_2(u_1, u_2^*) \tag{5.31}$$

where U_1 and U_2 are defined by Equation (5.24).

In terms of the Hilbert space notation developed in Chapter 3 the payoff functional becomes

$$J_2(u_1,u_2) = \frac{1}{2} E \left\{ < \Phi x_o - T_1u_1 + T_2u_2, \Phi x_o - T_1u_1 + T_2u_2 > \right.$$

$$\left. + < u_1,u_1 > - < u_2,u_2 > \mid Z_2(t) \right\} \tag{5.32}$$

which includes the dynamic Equation (5.30), since it was used to develop the above payoff functional.

Thus, from player 2's point of view of a secure strategy, player 1 minimizes at arbitrary time $t = t_o$

$$\min_{u_1 \in U_1 \times U_2} \frac{1}{2} \left\{ <\Phi x_o - T_1 u_1 + T_2 u_2, \Phi x_o - T_1 u_1 + T_2 u_2> + <u_1, u_1> - <u_2, u_2> \right\} \tag{5.33}$$

From Section 3.3 we know that the globally minimizing control of player 1 is

$$u_1 = T_1^* D_1 (\Phi x_o + T_2 u_2) \tag{5.34}$$

where

$$D_1 = (I + T_1 T_1^*)^{-1} \tag{5.35}$$

Substituting Equation (5.34) into (5.32) gives

$$J_2(u_2) = \frac{1}{2} E \left\{ <\Phi x_o - T_1 T_1^* D_1 (\Phi x_o + T_2 u_2) + T_2 u_2, \Phi x_o - T_1 T_1^* D_1 (\Phi x_o + T_2 u_2) + T_2 u_2> + <T_1^* D_1 (\Phi x_o + T_2 u_2), T_1^* D_1 (\Phi x_o + T_2 u_2)> - <u_2, u_2> \mid Z_2(t) \right\} \tag{5.36}$$

which simplifies after some algebra to

$$J_2(u_2) = \frac{1}{2} E \left\{ <\Phi x_o + T_2 u_2, D_1 (\Phi x_o + T_2 u_2)> - <u_2, u_2> \mid Z_2(t) \right\} \tag{5.37}$$

Let us define.

$$P_2(t) \triangleq E \left\{ \left[x(t) - \hat{x}_2(t)\right] \left[x(t) - \hat{x}_2(t)\right]^T \mid Z_2(t) \right\} \tag{5.38}$$

and consider the term

$$E\left\{\langle\Phi x, D_1 \Phi x\rangle \mid Z_2(t)\right\} = E\left\{\Phi(x - \hat{x}_2 + \hat{x}_2), D_1 \Phi(x - \hat{x}_2 + \hat{x}_2)\rangle \mid Z_2(t)\right\}$$

$$= E\left\{\langle\Phi(x - \hat{x}_2), D_1 \Phi(x - \hat{x}_2)\rangle \mid Z_2(t)\right\} + 2E\left\{\langle\Phi(x - \hat{x}_2),\right.$$

$$\left. D_1 \Phi \hat{x}_2\rangle \mid Z_2(t)\right\} + E\left\{\langle\Phi \hat{x}_2, D_1 \Phi \hat{x}_2\rangle \mid Z_2(t)\right\} \tag{5.39}$$

But

$$E\left\{\langle\Phi(x - \hat{x}_2), D_1 \Phi \hat{x}_2\rangle \mid Z_2(t)\right\} = \langle\Phi \hat{x}_2, D_1 \Phi \hat{x}_2\rangle - \langle\Phi \hat{x}_2, D_1 \Phi \hat{x}_2\rangle = 0 \tag{5.40}$$

and

$$E\left\{\langle\Phi(x - \hat{x}_2), D_1 \Phi(x - \hat{x}_2)\rangle \mid Z_2(t)\right\} = E\left\{\langle\Phi^T D_1 \Phi(x - \hat{x}_2),\right.$$

$$\left.(x - \hat{x}_2)\rangle \mid Z_2(t)\right\} = \mathrm{tr}\left[\Phi^T D_1 \Phi P_2\right] \tag{5.41}$$

where $\mathrm{tr}[\cdot]$ is the trace operator.

Thus, in general, the payoff functional can be written as

$$J_2(u_2) = \frac{1}{2}\left\{\langle\Phi \hat{x}_2 + T_2 u_2, D_1(\Phi \hat{x}_2 + T_2 u_2)\rangle - \langle u_2, u_2\rangle\right.$$

$$\left. + \mathrm{tr}\left\{\Phi^T D_1 \Phi P_2\right\}\right. \tag{5.42}$$

and Equation (5.37) becomes

$$J_2(u_2) = \frac{1}{2}\left\{\langle\Phi \hat{x}_{2_o} + T_2 u_2, D_1(\Phi \hat{x}_{2_o} + T_2 u_2)\rangle - \langle u_2, u_2\rangle\right.$$

$$\left. + \mathrm{tr}\left[\Phi^T D_1 \Phi P_{2_o}\right]\right\} \tag{5.43}$$

where

$$\hat{x}_{2_o} = E\left|x(t_o)|Z_2(t)\right| = \bar{x}_o$$
$$P_{2_o} = \text{cov}\left|x(t_o),\ x(t_o)\right| = P_o \tag{5.44}$$

Since $\text{tr}\left[\Phi^T D_1 \Phi P_{2_o}\right]$ is independent of the control $u_2(t)$, maximizing $J_2(u_2)$ is equivalent to maximizing

$$\bar{J}_2(u_2) = \frac{1}{2}\left|<\Phi\hat{x}_{2_o} + T_2u_2, D_1(\Phi\hat{x}_{2_o} + T_2u_2)> - <u_2,u_2>\right| \tag{5.45}$$

From the results of Sections 3.3 and 3.4, we know that the resulting maximin control for player 2 is

$$u_2^* = T_2^* D\Phi x_{2_o} = T_2^*\left|I + T_1T_1^* - T_2T_2^*\right|^{-1}\Phi\hat{x}_{2_o} \tag{5.46}$$

or

$$x_2^*(t) = G_2^T(t)S(t)\hat{x}_{2_o}(t) \tag{5.47}$$

where S(t) is again the solution to Equation (5.11), i.e.,

$$\dot{S} = -SF(t) - F^T(t)S + S\left|G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right|S;\ S(T) = I \tag{5.48}$$

The corresponding optimal response for player 1 is from Equations (5.34) and (5.47).

$$u_1^* = T_1^* D_1(\Phi x_o + T_2T_2^* D\Phi\hat{x}_{2_o}) \tag{5.49}$$

However, the initial time t_o is completely arbitrary, thus if $\hat{x}_2(t)$ can be made available for any t, the open-loop controls (Equations (5.46) and (5.49) can be applied continuously and immediately to yield optimal feedback control laws by replacing t_o by t.

Substituting Equations (5.46) and (5.49) into (5.30), the dynamic system for arbitrary t is

$$\dot{x}(t) = \left[F(t) - G_1(t)(T_1^* D_1 \Phi)(t)\right] x(t) + \left[G_2(t) - G_1(t)(T_1^* D_1 T_2)(t)\right]\left[(T_2^* D \Phi)(t)\right] \hat{x}_2(t); \quad x(t_o) = \bar{x}_o \tag{5.50}$$

and

$$z_2(t) = H_2(t)x(t) + w_2(t) \tag{5.51}$$

The linear-Gaussian assumptions imply that $\hat{x}_2(t)$ can be generated by a Kalman-Bucy filter based on a prior estimate of the initial state, a prior estimate of the variance of the error of this estimate; the measurements of the state up to time t; and the dynamic equation

$$\dot{\hat{x}}_2(t) = \left[F(t) - G_1(t)(T_1^* D \Phi)(t) + G_2(t)(T_2^* D \Phi)(t)\right] \hat{x}_2(t) + P_2(t)H_2^T(t)W_2^{-1}(t)\left[z_2(t) - H_2(t)\hat{x}_2(t)\right];$$
$$\hat{x}_2(t_o) = \bar{x}_o \tag{5.52}$$

where $P_2(t)$ is the variance of the error of player 2's estimate and is obtained from

$$\dot{P}_2(t) = \left[F(t) - G_1(t)(T_1^* D_1 \Phi)(t)\right] P_2(t) + P_2(t)\left[F(t) - G_1(t)(T_1^* D_1 \Phi)(t)\right]^T - P_2(t)H_2^T(t)W_2^{-1}(t)H_2(t)P_2(t);$$
$$P(t_o) = P_o \tag{5.53}$$

Hence, the closed-loop optimal controls for player 2 and the corresponding optimal closed-loop strategy for player 1 are:

$$u_2^* = T_2^* D \Phi \hat{x}_2$$
$$u_1^* = T_1^* D_1 \Phi x + T_1^* D_1 T_2 T_2^* D \Phi \hat{x}_2 = T_1^* D \Phi \hat{x}_2 + T_1^* D_1 \Phi \tilde{x}_2 \tag{5.54}$$

If we define the symmetrix matrix

$$N_1(t) \triangleq \Phi^T(T,t) D_1(t) \Phi(T,t) \tag{5.55}$$

where

$$D_1(t) = \left[I + \int_t^T \Phi(T,\tau) G_1(\tau) G_1^T(\tau) \Phi^T(T,\tau) d\tau \right]^{-1} \tag{5.56}$$

then taking the derivative of $N_1(t)$ with respect to t we obtain

$$\dot{N}_1(t) = \dot{\Phi}^T(T,t) D_1(t) \Phi(T,t) + \Phi^T(T,t) \dot{D}_1(t) \Phi(T,t) + \Phi^T(T,t) D_1(t) \dot{\Phi}(T,t) \tag{5.57}$$

But

$$\begin{aligned} \frac{d}{dt} D_1(t) &= - D_1(t) \frac{d}{dt} D_1^{-1}(t) D(t) \\ &= D_1(t) \Phi(T,t) G_1(t) G_1^T(t) \Phi^T(T,t) D_1(t) \end{aligned} \tag{5.58}$$

thus

$$\begin{aligned} \dot{N}_1(t) = &- F^T(t) \Phi^T(T,t) D_1(t) \Phi(T,t) + \Phi^T(T,t) D_1(t) \Phi(T,t) G_1(t) G_1^T(t) \\ &\Phi^T(T,t) D_1(t) \Phi(T,t) - \Phi^T(T,t) D_1(t) \Phi(T,t) F(t) \end{aligned} \tag{5.59}$$

or

$$\dot{N}_1(t) = - N_1(t) F(t) - F^T(t) N_1(t) + N_1(t) G_1(t) G_1^T N_1(t) \tag{5.60}$$

with boundary condition

$$N_1(T) = I \tag{5.61}$$

The optimal control for player 2 and corresponding optimal response for player 1 can then be written as

$$u_2^*(t) = G_2^T(t)S(t)\hat{x}_2(t) \tag{5.62}$$

$$u_1^*(t) = G_1^T(t)S(t)\hat{x}_2(t) + G_1^T(t)N_1(t)\tilde{x}_2(t) \tag{5.63}$$

where $S(t)$ and $N_1(t)$ are defined by Equations (5.11) and(5.60) respectively.

Furthermore, using Equations (5.11) and 5.60), the Kalman-Bucy filter (Equation (5.52)) and corresponding covariance equation (5.53) can be written as

$$\begin{aligned}\dot{\hat{x}}_2(t) = &\left| F(t) - G_1(t)G_1^T(t)S(t) + G_2(t)G_2^T(t)S(t)\right| \hat{x}_2(t) \\ &+ P_2(t)H_2^T(t)W_2^{-1}(t)\left| z_2(t) - H_2(t)\hat{x}_2(t)\right| ; \quad \hat{x}_2(t_o) = \bar{x}_o\end{aligned} \tag{5.64}$$

and

$$\begin{aligned}\dot{P}_2(t) = &\left| F(t) - G_1(t)G_1^T(t)N_1(t)\right| P_2(t) + P_2(t)\left| F(t) - G_1(t)G_1^T(t) \right. \\ &\left. N_1(t)\right|^T - P_2(t)H_2^T(t)W_2^{-1}(t)H_2(t)P_2(t); \; P_2(t_o) = P_o\end{aligned} \tag{5.65}$$

If we define

$$\begin{aligned}X_2(t) &\triangleq E\left| x(t)x(t)^T \mid Z_2(t)\right| \\ &= E\left|\left|\hat{x}_2(t) - \tilde{x}_2(t)\right| \left|\hat{x}_2(t) - \tilde{x}_2(t)\right|^T \mid Z_2(t)\right| \\ &= \hat{X}_2(t) + P_2(t)\end{aligned} \tag{5.66}$$

then on substituting the optimal strategies (Equations (5.62) and (5.63)) into the system equation (Equation (5.30)) we can write, after post-multiplying by $x^T(t)$, adding the transpose of the resulting equation and then taking the conditional mean of the resulting expression,

$$\dot{X}_2 = FX_2 + X_2F^T - G_1G_1{}^TS\hat{X}_2 + G_1G_1{}^TN_1P_2 + G_2G_2{}^TS\hat{X}_2$$
$$- \hat{X}_2SG_1G_1{}^T - P_2N_1G_1G_1{}^T + \hat{X}_2SG_2G_2{}^T \tag{5.67}$$

Substitution of the optimal strategies into the payoff functional (equation (5.29)) and using the trace operator allows us to write

$$J_2(u_1{}^*,u_2{}^*) = \frac{1}{2}\,\mathrm{tr}\left\{X_2(T) + \int_t^T \left[G_1G_1{}^TS\hat{X}_2S + G_1G_1{}^TN_1P_2N_1 - G_2G_2{}^TS\hat{X}_2S\right] dt\right. \tag{5.68}$$

If we now add the perfect differentials $\frac{d}{dt}(SX_2)$, $\frac{d}{dt}(N_1P_2)$ and $\frac{d}{dt}(-SP)$ into the integrand of Equations (5.68) and compensate by adding $S(t)X_2(t) - S(T)X_2(T) + \left|N_1(t) - S(t)\right| P_2(t) - \left|N_1(T) - S(T)\right|$ $P_2(T)$ outside the integral, most of the terms cancel and we obtain as the security payoff or gain floor for player 2

$$J_2(u_1{}^*,u_2{}^*) = \frac{1}{2}x^T(t)S(t)x(t) + \frac{1}{2}\tilde{x}_2(t)\left|N_1(t) - S(t)\right|\tilde{x}_2(t) \tag{5.69}$$
$$+ \frac{1}{2}\,\mathrm{tr}\left[\int_t^T \left|N_1(s) - S(s)\right| P_2(s)H_2{}^T(s)W_2{}^{-1}(s)H_2(s)P_2(s)ds\right]$$

The entire game from player 2's point of view can be described by a 2n-dimensional system consisting of the vectors $x(t)$ and $\hat{x}(t)$ or similarly of the vectors $x(t)$ and $\tilde{x}(t)$. The $\left|x(t),\hat{x}(t)\right|$ system is obtained by substituting Equations (5.62) and (5.63) into the system Equation (5.30) to give

$$
\begin{aligned}
\dot{x} &= F(t)x(t) - G_1(t)G_1^T(t)S(t)\hat{x}_2(t) - G_1(t)G_1^T(t)N_1(t)\tilde{x}_2(t) \\
&\quad + G_2(t)G_2^T(t)S(t)\hat{x}_2(t) \\
&= \left[F(t) - G_1(t)G_1^T(t)N_1(t)\right] x(t) + \left[G_1(t)G_1^T(t)N_1(t) - G_1(t)G_1^T(t)S(t)\right. \\
&\quad \left. + G_2(t)G_2^T(t)S(t)\right] \hat{x}_2(t)
\end{aligned}
\tag{5.70}
$$

The input to this equation is obtained from Equation (5.64) or, on substituting $z_2(t) = H_2(t)x(t) + w_2(t)$, from

$$
\begin{aligned}
\dot{\hat{x}}_2 &= \left[F(t) - G_1(t)G_1^T(t)S(t) + G_2(t)G_2^T(t)S(t) - P_2(t)H_2^T(t)W_2^{-1}(t)H_2(t)\right]\hat{x}_2 \\
&\quad + P_2(t)H_2^T(t)W_2^{-1}(t)H_2(t)x(t) + P_2(t)H_2^T(t)W_2^{-1}(t)H_2(t)w_2(t)
\end{aligned}
\tag{5.71}
$$

Then, from player 2's point of view the entire play of the game can be described by the 2n-dimensional differential equation

$$
\begin{aligned}
\begin{bmatrix} \dot{x} \\ \dot{\hat{x}}_2 \end{bmatrix} &= \left[\begin{array}{c|c} F - G_1G_1^TN_1 & G_1G_1^TN_1 - G_1G_1^TS + G_2G_2^TS \\ \hline P_2H_2^TW_2^{-1}H_2 & F - G_1G_1^TS + G_2G_2^TA - P_2H_2^TW_2^{-1}H_2 \end{array}\right] \begin{bmatrix} x \\ \hat{x}_2 \end{bmatrix} \\
&\quad + \left[\begin{array}{c} 0 \\ \hline P_2H_2^TW_2^{-1}H_2 \end{array}\right] w_2(t)
\end{aligned}
\tag{5.72}
$$

In the above system the white noise $w_2(t)$ which is additive measurement noise to player 2, appears as process noise to the 2n-dimensional system.

5.3 DISCUSSION

The prior commitment and delayed commitment solutions to the stochastic differential game discussed in this chapter are summarized in Tables 5.1 and 5.2 respectively.

In the prior commitment formulation, the optimal control for player 1 consists of the sum of a term that is the same as that of the corresponding deterministic differential game and a term that is a linear function of the error in his opponent's state estimate. The optimal control for player 2 satisfies the Separation Theorem. Determination of the feedback gain for the first term of player 1 and for player 2 requires the solution of a simple matrix Riccati equation with terminal boundary conditions. To determine the feedback gain of the second term of player 1's strategy, however, we are faced with the often difficult task of finding the solution of a nonlinear two point boundary value problem defined by the equations for $\dot{N}$ and $\dot{P}$ in Table 5.1.

In the case of the delayed commitment formulation, the secure strategy for player 1 is the same as for the deterministic game, while the secure strategy for player 2 satisfies the Separation Theorem. Determination of the feedback gains involves the simple solution of matrix Riccati equations with all the boundary conditions for each equation given at one point in time.

The secure delayed commitment payoff, J_1, for player 1 is identical to that obtained in Chapter 3 (Equation (3.90)) for the corresponding deterministic game. The difference between J_1 and the prior commitment payoff, J, can be written from Equations (5.28) and

Table 5.1 Summary of the Prior Commitment Strategies

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t), \quad x(t_o) \quad N(\bar{x}_o, P_o)$$

Player 1: Perfect measurements

Player 2: $z_2(t) = H_2(t)x(t) + w_2(t), \quad w_2 \sim N(0, W_2)$

$$J = \frac{1}{2} E \left\{ x^T(T)x(T) + \int_{t_o}^{T} \left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t) \right] dt \right\}$$

$$u_1'(t) = G_1^T(t)S(t)x(t) + G_1^T(t)N(t)\tilde{x}(t)$$

$$u_2'(t) = G_2^T(t)S(t)\hat{x}(t)$$

$$\dot{S} = -SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t) \right] S; \quad S(T) = I$$

$$\dot{N} = -NF(t) - F^T(t)N - S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t) \right] S$$

$$+ (S + N)G_1(t)G_1^T(t)(S + N) + NP(t)H_2^T(t)W_2^{-1}(t)H_2(t)$$

$$+ H_2^T(t)W_2^{-1}(t)H_2(t)P(t)N; \quad N(T) = 0$$

$$\dot{\hat{x}} = (F - G_1G_1^TS + G_2G_2^TS)\hat{x} + PH_2^TW_2^{-1}(z_2 - H_2\hat{x}); \quad \hat{x}(t_o) = \bar{x}_o$$

$$\dot{P} = AP + PA^T - PH_2^T(t)W_2^{-1}(t)H_2(t)P; \quad P(t_o) = P_o$$

$$A(t) = F(t) - G_1(t)G_1^T \left[S(t) + N(t) \right]$$

$$J(u_1', u_2') = \frac{1}{2} x^T(t)S(t)x(t) + \frac{1}{2} \tilde{x}^T(t)N(t)\tilde{x}(t)$$

$$+ \frac{1}{2} \operatorname{tr}\left[\int_{t}^{T} N(s)P(s)H_2^T(s)W_2^{-1}(s)H_2(s)P(s)ds \right]$$

Table 5.2 Summary of the Delayed Commitment Strategies

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t), \quad x(t_o) \sim N(\bar{x}_o, P_o)$$

Player 1: Perfect measurements

Player 2: $z_2(t) = H_2(t)x(t) + w_2(t), \quad w_2 \sim N(0, W_2)$

$$J = \frac{1}{2} E\left\{ x^T(T)x(T) + \int_{t_o}^{T} \left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t) \right] dt \right\}$$

Define: $Z_2(t) = \left\{ (z_2(s), s) \quad ; \quad s \in [t_o, t) \right\}$

$$U_1 : u_1 = u_1(x(t), t)$$

$$U_2 : u_2 = u_2(Z_2(t), t)$$

$$(Tu)(t) = \int_{t_o}^{t} \Phi(t,\tau)G(\tau)u(\tau)d\tau$$

$$(T^* \xi)(t) = G^T(t)\Phi^T(T,t)\xi$$

$$D = \left[I + T_1T_1^* - T_2T_2^* \right]^{-1}$$

$$D_1 = \left[I + T_1T_1^* \right]^{-1}, \quad D_2 = \left[I - T_2T_2^* \right]^{-1}$$

<u>Player 1</u>

$$J_1 = \frac{1}{2} \left\{ x^T(T)x(T) - \int_{t_o}^{T} \left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t) \right] dt \right\}$$

Table 5.2 (Continued)

$$u_1^* = T_1^* D \Phi x = G_1^T(t) S(t) x(t)$$

$$u_2^* = T_2^* D \Phi x = G_2^T(t) S(t) x(t)$$

$$\dot{S} = -SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right]S\ ;$$

$$S(T) = I$$

$$J_1(u_1^*, u_2^*) = \frac{1}{2} x^T(t) S(t) x(t).$$

<u>Player 2</u>

$$J_2 = \frac{1}{2} E\left[x^T(T)x(T) + \int_{t_o}^{T}\left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right] dt \,|\, Z_2(t)\right]$$

$$u_2^* = T_2^* D \Phi \hat{x}_2 = G_2^T(t) S(t) \hat{x}_2(t)$$

$$u_1^* = T_1^* D \Phi \hat{x}_2 + T_1^* D_1 \Phi \tilde{x}_2 = G_1^T(t) S(t) x(t)$$

$$+ G_1^T(t)\left[N_1(t) - S(t)\right]\tilde{x}_2(t)$$

$$\dot{S} = -SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right]S\ ;$$

$$S(T) = I$$

$$\dot{N}_1 = -N_1F(t) - F^T(t)N_1 + N_1G_1(t)G_1^T(t)N_1; \quad N_1(T) = I$$

$$\dot{\hat{x}}_2 = \left[F - G_1G_1^TS + G_2G_2^TS\right]\hat{x}_2 + P_2H_2^TW_2^{-1}\left[z_2 - H_2\hat{x}_2\right];$$

$$\hat{x}_2(t_o) = \bar{x}_o$$

Table 5.2 (Continued)

$$\dot{P}_2 = \left[F - G_1G_1^TN_1\right]P_2 + P_2\left[F - G_1G_1^TN_1\right]^T$$

$$- P_2H_2^TW_2^{-1}H_2P_2 \quad ; \quad P_2(t_o) = P_o$$

$$J_2(u_1^*,u_2^*) = \frac{1}{2}x^T(t)S(t)x(t) + \frac{1}{2}\tilde{x}_2^T(t)\left[N_1(t) - S(t)\right]\tilde{x}_2(t)$$

$$+ \frac{1}{2}\mathrm{tr}\left[\int_t^T \left[N_1(s) - S(s)\right]P_2(s)H_2^T(s)W_2^{-1}(s)H_2(s)P_2(s)ds\right]$$

(3.90) and using the trace operator as

$$J_1(u_1^*,u_2^*) - J(u_1',u_2') = -\frac{1}{2}\,\mathrm{tr}\left\{NP + \int_t^T N(s)P(s)H_2^T(s)W_2^{-1}(s)H_2(s)P(s)ds\right\} \qquad (5.73)$$

Taking the trace of Equation (5.13) and collecting terms, we can write

$$\begin{aligned} \mathrm{tr}\left|\dot{N}\right| = \mathrm{tr}\Big\{ & -N\left[F(t) - P(t)H_2^T(t)W_2^{-1}(t)H_2(t) - G_1(t)G_1^T(t)S\right] \\ & - \left[F(t) - P(t)H_2^T(t)W_2^{-1}(t)H_2(t) - G_1(t)G_1^T(t)S\right]^T N \\ & + NG_1(t)G_1^T(t)N + SG_2(t)G_2^T(t)S\Big\} \quad ; \quad \mathrm{tr}\left|N(T)\right| = 0 \end{aligned} \qquad (5.74)$$

The above equation can be viewed as a linear differential equation driven by the term $\mathrm{tr}\left|NG_1(t)G_1^T(t)N + SG_2(t)G_2^T(t)S\right|$, which is greater than or equal to zero. Since the terminal value, $\mathrm{tr}\left|N(T)\right|$, equals zero, it follows that $\mathrm{tr}\left|N(t)\right|$ can only become smaller than zero as time progresses, and we conclude from Equation (5.73) since $\mathrm{tr}\left|P\right|$ and $\mathrm{tr}\left|W_2^{-1}\right|$ are positive that

$$J(u_1',u_2') \le J_1(u_1^*,u_2^*) \qquad (5.75)$$

To study the relation between J and the delayed commitment secure payoff, J_2, for player 2, assume that $P(t) = P_2(t)$, then subtracting Equation (5.69) from Equation (3.90) we obtain

$$J(u_1',u_2') - J_2(u_1^*,u_2^*) = \frac{1}{2}\,\mathrm{tr}\left\{\left[N - (N_1 - S)\right]P\right.$$

(Cont'd)

$$+ \int_t^T \left[N(s) - (N_1(s) - S(s)) \right] P(s) H_2^T(s) W_2^{-1}(s) H_2(s) P(s) ds \right\} \tag{5.76}$$

But from Equations (5.13), (5.60) and (5.48), i.e., the $\dot{N}$, $\dot{N}_1$ and $\dot{S}$ equations, and using the trace operator we can write

$$\mathrm{tr}\left\{\dot{N} - (\dot{N}_1 - \dot{S})\right\} = \mathrm{tr}\left\{- \left[N - (N_1 - S)\right] F - F^T \left[N - (N_1 - S)\right] \right.$$
$$- N_1 G_1 G_1^T N_1 - (S + N) G_1 G_1^T (S + N) + NPH_2^T W_2^{-1} H_2$$
$$\left. + H_2^T W_2^{-1} H_2 PN \right\} \quad ; \quad \mathrm{tr}\left\{N(T) - (N_1(T) - S(T))\right\} = 0 \tag{5.78}$$

From our earlier observation $\mathrm{tr}\left\{N(t)\right\} \leq 0 \ \forall \ t \leq T$, and we can again view the above equation as a linear differential equation of $\mathrm{tr}\left\{N - (N_1 - S)\right\}$ driven by the term $\mathrm{tr}\left\{- N_1 G_1 G_1^T N_1 - (S + N) G_1 G_1^T (S + N) + NPH_2^T W_2^{-1} H_2 + H_2^T W_2^{-1} H_2 PN\right\}$ which is smaller than or equal to zero. Since the terminal value, $\mathrm{tr}\left\{N(T) - (N_1(T) - S(T))\right\}$, is equal to zero, it follows that $\mathrm{tr}\left\{N(t) - (N_1(t) - S(t))\right\}$ can only be greater than or equal to zero. Thus, all the terms in Equation (5.76) are ≥ 0 and hence

$$J_2(u_1^*, u_2^*) \leq J(u_1', u_2') \tag{5.79}$$

and as a result of Equation (5.75)

$$J_2(u_1^*, u_2^*) \leq J(u_1', u_2') \leq J_1(u_1^\bullet, u_2^\bullet) \tag{5.80}$$

Note, that if $W_2(t)$ is large, or if during the game

$$P_2(t) \longrightarrow 0 \quad ; \quad P(t) \longrightarrow 0$$

then

$$\dot{N}(t) \longrightarrow (\dot{N}_1(t) - \dot{S}(t))$$

and

$$J_2(u_1^*, u_2^*) \longrightarrow J(u_1', u_2')$$

The relationship between the various payoffs discussed above is shown in Figure 5.1

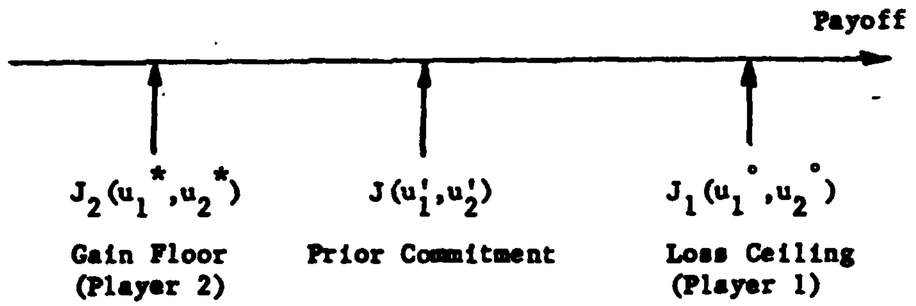


Figure 5.1 Relationship Between Prior Commitment and Delayed Commitment Payoffs

It is immediately clear from Figure 5.1 that if player 1 knows that player 2 is committed to strategy $u_2'(t)$, he should play $u_1'(t)$ and similarly for player 2. Thus if the players had to determine at $t = t_o$ the strategies they would have to play for the rest of the game, $u_1'(t)$ and $u_2'(t)$ would be the proper choice. However, as we have seen in our tutorial example (Section 4.2) as soon as the game has advanced to a time $t > t_o$, $u_1'(t)$ and $u_2'(t)$ become unsafe strategies, as compared to $u_1^*(t)$ and $u_2^*(t)$ respectively.

On the other hand, if either player 1 or player 2 commits himself to his secure strategy, he can only be assured of his secure payoff. Thus we find, as is usual with games with imperfect information, that the players should keep their strategies secret.

The <u>actual</u> payoff, J_o, can only be calculated at the conclusion of the game, i.e., when everything has become a fact, and it is calculated from

$$J_o = \frac{1}{2}\left\{x^T(T)x(T) + \int_{t_o}^{T}\left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right] dt\right\} \qquad (5.81)$$

which depends on the actual values of the control functions $u_1(t)$ and $u_2(t)$ employed during the game, which in turn depend on the strategies employed and the actual values of $w_2(t)$.

CHAPTER 6

A PURSUIT-EVASION EXAMPLE

One of the differential game problems most easily visualized is the problem of pursuit-evasion. In order to illuminate the results of the previous chapters we will, in this chapter, analyze a pursuit-evasion problem in two-dimensional Euclidian space where the pursuer, player 1, has perfect measurements of the state of his own system as well as that of the evader, while the evader, player 2, has only noise corrupted measurements. The problem satisfies Behn's [9] requirements for player 1 to determine exactly the error in player 2's state-estimate, and thus allows us to compare the prior and delayed commitment problem formulations.

As mentioned in Chapter 5, a physical example of this problem is a homing missile and an evading aircraft where the missile has an inertial reference unit which allows accurate determination of its state vector and, in addition, has considerable ground support via an up- and downlink to determine the state of the evader. The aircraft has only noise corrupted measurements of its own inertial reference system and of the missile from noise corrupted radar measurements.

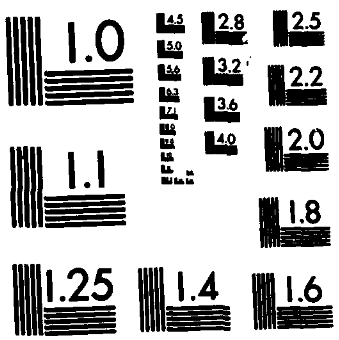
6.1 PROBLEM FORMULATION

The space diagram showing the geometric relationship between missile and airplane or target during the encounter are shown in Figure 6.1. The missile and target velocity, V_M and V_T respectively, are assumed to be constant. Gravity effects have been neglected and the encounter is assumed to be restricted to the x-y plane.

AD-A124 634 STOCHASTIC DIFFERENTIAL GAME TECHNIQUES(U) CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING AND APPLIED SCIENCE B MONS MAR 82 DASG60-80-C-0007 2/2

UNCLASSIFIED F/G 12/1 NL

END
FILMED
DTIC



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

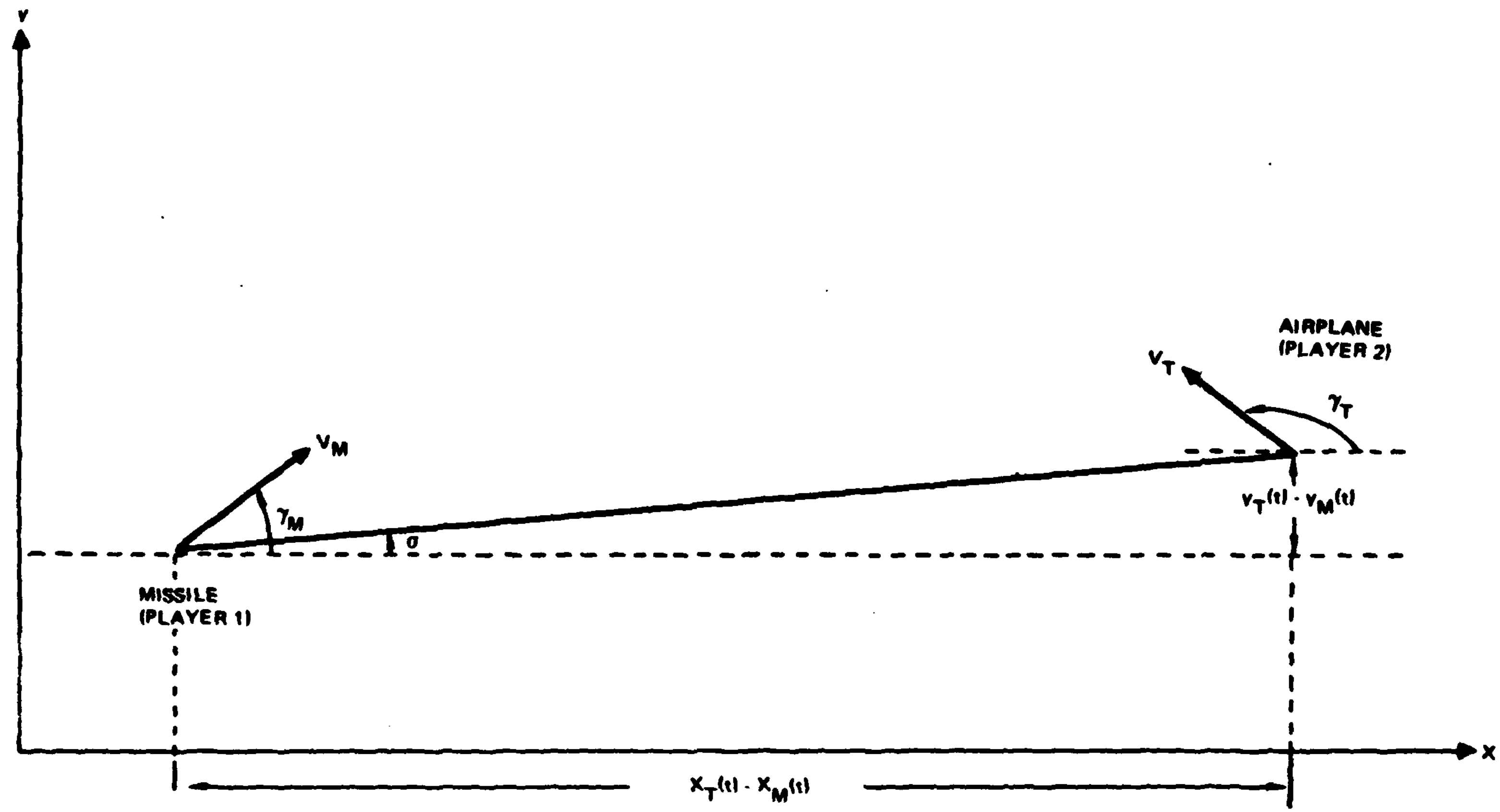


Figure 6.1 Geometry of the Pursuit - Evasion Problem

The fundamental relations governing missile and target paths are the velocity equations [22] .

$$
\begin{aligned}
\dot{x}_M &= V_M \cos \gamma_M \\
\dot{y}_M &= V_M \sin \gamma_M \\
\dot{x}_T &= V_T \cos \gamma_T \\
\dot{y}_T &= V_T \sin \gamma_T
\end{aligned}
\qquad (6.1)
$$

The angles are subject to change since both missile and target are, of course, free to maneuver in the x-y plane. At an arbitrarily selected time $t = t_o$, the angles γ_M and γ_T have some initial values γ_{MO} and γ_{TO} and at a later time t are perturbed by small amounts γ_m and γ_t to γ_M and γ_T respectively, while the line of sight has changed from zero at $t = t_o$ to σ . Under these conditions, the instantaneous angles of the velocity vectors are

$$\gamma_M(t) = \gamma_{MO} + \gamma_m(t), \qquad (6.2)$$

and

$$\gamma_T(t) = \gamma_{TO} + \gamma_t(t) \qquad (6.3)$$

so that the linear velocity components are

$$
\begin{aligned}
\dot{x}_M &= V_M \cos \gamma_{MO} - \gamma_m V_M \sin \gamma_{MO} \\
\dot{y}_M &= V_M \sin \gamma_{MO} + \gamma_m V_M \cos \gamma_{MO} \\
\dot{x}_T &= V_T \cos \gamma_{TO} - \gamma_t V_T \sin \gamma_{TO} \\
\dot{y}_T &= V_T \sin \gamma_{TO} + \gamma_t V_T \cos \gamma_{TO}
\end{aligned}
\qquad (6.4)
$$

where the small angle approximations $\sin\ \gamma = \gamma$ and $\cos\ \gamma = 1$ have been used.

If we assume that the missile and target are initially on a collision course, i.e.

$$V_T \sin \gamma_{TO} = V_M \sin \gamma_{MO} \tag{6.5}$$

then using this equation and Equation (6.4)

$$\dot{x}_T - \dot{x}_M = V_T \cos \gamma_{TO} - V_M \cos \gamma_{MO} - (\gamma_t - \gamma_m) V_M \sin \gamma_{MO} \tag{6.6}$$

If we neglect the difference term involving $\gamma_t - \gamma_m$, the closing velocity V_c is given approximately by

$$-V_c \sim \dot{x}_T - \dot{x}_M \sim V_T \cos \gamma_{TO} - V_M \cos \gamma_{MO} \tag{6.7}$$

and in view of the assumed constant velocities

$$x_T(t) - x_M(t) = V_c(T - t) \tag{6.8}$$

Since only the relative positions of the missile and target need to be known; i.e., $x_r(t) = x_T(t) - x_M(t)$ and $y_r(t) = y_T(t) - y_M(t)$, the relative missile-target position is uniquely specified by giving the time t and the projection of M and T on a line, L, perpendicular to the initial line of sight. Thus, the original problem has been changed from a two-dimensional intercept problem with unspecified final time T to a one-dimensional intercept problem with a final time

$$T = \frac{x_T(0) - x_M(0)}{V_c} \tag{6.9}$$

If we let y_1' and v_1' be the projection of the missile's position and velocity respectively on L, then we obtain as equations of motion for the missile

$$\begin{aligned} \dot{y}_1' &= v_1 \\ \dot{v}_1 &= \dot{\gamma}_m V_M \sin \gamma_{MO} \end{aligned} \tag{6.10}$$

with initial conditions

$$\begin{aligned} y'(t_o) &= y_M(0) \\ v(t_o) &= V_M \sin \gamma_{MO} \end{aligned} \tag{6.11}$$

By the same analysis we obtain a similar set for the target with the subscripts 1, m and M replaced by 2, t and T. But $\dot{\gamma} V = 32.2\ n$, where n is the lateral acceleration in G's, and we can write

$$\begin{aligned} \dot{y}_1' &= v_1 \\ \dot{v}_1 &= K_1 n_1 \end{aligned} \tag{6.12}$$

and

$$\begin{aligned} \dot{v}_2' &= v_2 \\ \dot{v}_2 &= K_2 n_2 \end{aligned} \tag{6.13}$$

where

$$\begin{aligned} K_1 &= 32.2 \cos \gamma_{MO}, \\ K_2 &= 32.2 \cos \gamma_{TO}, \end{aligned} \tag{6.14}$$

and

n_1, n_2 are the lateral missile and target accelerations respectively in G's.

Now let $u_1'(t)$ and $u_2'(t)$ be the missile and target called for accelerations respectively, and let us assume the following missile and target system transfer functions

$$\frac{n_1(s)}{u_1'(s)} = \frac{1}{\tau_1 s + 1}$$
$$\frac{n_2(s)}{u_2'(s)} = \frac{1}{\tau_2 s + 1} \tag{6.15}$$

where both τ_1 and τ_2 are positive real numbers. The resulting equations of motion under the above used assumptions of

1. Constant target and missile velocity
2. σ, γ_m and γ_t are small angles

are then

$$\dot{y}_1'(t) = v_1(t)$$
$$\dot{v}_1(t) = K_1 n_1(t) \tag{6.16}$$
$$\dot{n}_1(t) = -\frac{n_1}{\tau_1} + \frac{u_1'}{\tau_1}$$

and

$$\dot{y}_2'(t) = v_2(t)$$
$$\dot{v}_2(t) = K_2 n_2(t) \tag{6.17}$$
$$\dot{n}_2(t) = -\frac{n_2}{\tau_2} + \frac{u_2'}{\tau_2}$$

If we define the vectors

$$y_1 = \begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \end{bmatrix} = \begin{bmatrix} y_1' \\ v_1 \\ n_1 \end{bmatrix} \text{ and } y_2 = \begin{bmatrix} y_{12} \\ y_{22} \\ y_{32} \end{bmatrix} = \begin{bmatrix} y_2' \\ v_2 \\ n_2 \end{bmatrix} \tag{6.18}$$

Then we can write the system for the missile or player 1 as

$$\dot{y}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & K_1 \\ 0 & 0 & -1/\tau_1 \end{bmatrix} y_1 + \begin{bmatrix} 0 \\ 0 \\ 1/\tau_1 \end{bmatrix} u_1' \tag{6.19}$$

with initial conditions

$$y_1(t_o) = \begin{bmatrix} y_M(0) \\ V_M \sin \gamma_{MO} \\ 0 \end{bmatrix} \tag{6.20}$$

or

$$\dot{y}_1 = F_1'y_1 + G_1'u_1' \quad ; \quad y_1(t_o) = y_{1_o} \tag{6.21}$$

and for the target or player 2 as

$$\dot{y}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & K_2 \\ 0 & 0 & -1/\tau_2 \end{bmatrix} y_2 + \begin{bmatrix} 0 \\ 0 \\ 1/\tau_2 \end{bmatrix} u_2' \tag{6.22}$$

with initial conditions

$$y_2(t_o) = \begin{bmatrix} y_T(0) \\ V_T \sin \gamma_{TO} \\ 0 \end{bmatrix} \tag{6.23}$$

or

$$\dot{y}_2 = F_2' y_2 + G_2' u_2' \quad ; \quad y_2(t_o) = y_{2_o} \tag{6.24}$$

and with a final time

$$T = \frac{x_T(t_o) - x_M(t_o)}{V_c} \tag{6.25}$$

Player 1 has perfect measurements of his own and his opponent's state vector, while the measurements of player 2 are of the form

$$\begin{aligned} z_1'(t) &= H_1(t) y_1(t) + w_1'(t) \\ z_2'(t) &= H_2(t) y_2(t) + w_2'(t) \end{aligned} \tag{6.26}$$

where $w_1'(t)$ and $w_2'(t)$ are Gaussian white noise vectors, with zero mean and with

$$\begin{aligned} \text{cov}\,[w_1'(t), w_1'(\tau)] &= W_1'(t)\,\delta(t - \tau) \\ \text{cov}\,[w_2'(t), w_2'(\tau)] &= W_2'(t)\,\delta(t - \tau) \\ \text{cov}\,[w_1'(t), w_2'(\tau)] &= C(t)\,\delta(t - \tau) \end{aligned} \tag{6.27}$$

In addition, let the payoff criterion be given by

$$J = \frac{1}{2} E \left\{ a^2 \left[y_1(T) - y_2(T) \right]^T \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & 0 \end{array} \right] \left[y_1(T) - y_2(T) \right] \right.$$

$$\left. + \int_{t_o}^{T} \left[u_1'^T(t) R_1(t) u_1'(t) - u_2'^T(t) R_2(t) u_2'(t) \right] dt \right\} \tag{6.28}$$

where both $R_1(t)$ and $R_2(t)$ are positive definite and a^2 is introduced to allow for weighting of terminal miss against energy.

The above formulation will now be recognized as the classical interception problem in Euclidean space; i.e., player 1, the pursuer, attempts to intercept with player 2, the evader, at some fixed time T, while the latter tries to do the opposite. Both players have limited energy sources and do not care about the difference in the velocities of the two players at the terminal time.

From the point of view of the criterion, the number of "interesting" variables are the same as the number of control variables. Hence, this formulation of the game basically satisfies Behn's criterion for the ability of player 1 to determine the error in the state estimate of player 2.

If we define

$$x'(t) = \left[I_1 \;\vdots\; 0 \quad 0 \right] \left[\Phi_1(T,t) y_1(t) - \Phi_2(T,t) y_2(t) \right] \tag{6.29}$$

where $\Phi_1(t,\tau)$ and $\Phi_2(t,\tau)$ are the transition matrices for player 1 and player 2 respectively, then $x'(t)$ represents the terminal miss predicted at time t on the basis that no control is applied during the interval $[t, T]$.

The above transformation allows us to reduce the dimension of the problem since on taking the derivative of Equation (6.29) and using (6.24) and (6.25) we obtain

$$\dot{x}'(t) = \left[I \;\vdots\; 0 \quad 0 \right] \left[\frac{\partial \Phi_1(T,t)}{\partial t} y_1(t) + \Phi_1(T,t)(F_1' y_1 + G_1' u_1') \right.$$

$$\left. - \frac{\partial \Phi_2(T,t)}{\partial t} y_2(t) - \Phi_2(T,t)(F_2' y_2 + G_2 u_1') \right]$$

$$= \left[I \;\vdots\; 0 \quad 0 \right] \left[\Phi_1(T,t) G_1' u' - \Phi_2(T,t) G_2' u_2' \right]$$

$$= - G_1'(t,T) u_1' + G_2'(t,T) u_2' \tag{6.30}$$

with

$$x'(t_o) = \left[I \;\vdots\; 0 \quad 0 \right] \left[\Phi_1(T,t_o) y_1(t_o) - \Phi_2(T,t_o) y_2(t_o) \right] \tag{6.31}$$

where

$$\begin{aligned} G_1'(t,T) &= - \left[I \;\vdots\; 0 \quad 0 \right] \left[\Phi_1(T,t) G_1' \right] \\ G_2'(t,T) &= - \left[I \;\vdots\; 0 \quad 0 \right] \left[\Phi_2(T,t) G_2' \right] \end{aligned} \tag{6.32}$$

while the performance criterion in terms of $x'(t)$ is

$$J = \frac{1}{2} E \left\{ a^2 x'^T(T) x'(T) + \int_{t_o}^{T} \left[u_1'^T(t) R_1(t) u_1'(t) - u_2'^T(t) R_2(t) u_2'(t) \right] dt \right\} \tag{6.33}$$

For our example, the transition matrix of the system of player 1 is

$$\Phi_1(T,t) = \begin{bmatrix} 1 & \bar{T} & -K_1\tau_1{}^2(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1) \\ 0 & 1 & K_1\tau_1(1 - e^{-\bar{T}/\tau_1}) \\ 0 & 0 & e^{-\bar{T}/\tau_1} \end{bmatrix} \tag{6.34}$$

where

$$\bar{T} = \text{time-to-go} = T - t$$

The transition matrix for player 2 is the same as that for player 1 with the subscripts 1 replaced by 2.

From Equation (6.32), $G_1'(t,T)$ and $G_2'(t,T)$ are scalars and are given by

$$G_1'(t,T) = -\begin{bmatrix}1 & 0 & 0\end{bmatrix}\Phi_1(T,t)\begin{bmatrix}0\\0\\1/\tau_1\end{bmatrix}$$

$$= + K_1\ \tau_1(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1) \tag{6.35}$$

and

$$G_2'(t,T) = -\begin{bmatrix}1 & 0 & 0\end{bmatrix}\Phi_2(T,t)\begin{bmatrix}0\\0\\1/\tau_2\end{bmatrix}$$

$$= + K_2\ \tau_2(1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2) \tag{6.36}$$

Thus,

$$\dot{x}'(t) = -K_1 \tau_1 (1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1) u_1' + K_2 \tau_2 (1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2) u_2' \tag{6.37}$$

with

$$x'(t_o) = \left[y_M(0) - y_T(0)\right] + \left[T - t_o\right]\left[V_M \sin \gamma_{MO} - V_T \sin \gamma_{TO}\right] \tag{6.38}$$

and

$$T = \frac{x_T(t_o) - x_M(t_o)}{V_c} \tag{6.39}$$

With the dynamical system reduced to Equation (6.37), the measurements of player 2 must be reduced to measurements on $x'(t)$. If we define

$$z_2''(t) \triangleq \left[1 \quad 0 \quad 0\right]\left[\Phi_1(T,t)H_1^{-1}(t)z_1'(t) - \Phi_2(T,t)H_2^{-1}(t)z_2'(t)\right] \tag{6.40}$$

then using Equations (6.26) and (6.29) we can write

$$\begin{aligned} z_2''(t) &= \left[1 \quad 0 \quad 0\right]\left[\Phi_1(T,t)(y_1 + H_1^{-1}w_1') - \Phi_2(T,t)(y_2 + H_2^{-1}w_2')\right. \\ &= x'(t) + \left[1 \quad 0 \quad 0\right]\left[\Phi_1(T,t)H_1^{-1}w_1' - \Phi(T,t)H_2^{-1}w_2')\right] \end{aligned} \tag{6.41}$$

or

$$z_2''(t) = x'(t) + w_2''(t) \tag{6.42}$$

where the zero-mean, white noise process $w_2''(t)$ is given by

$$w_2''(t) = \left[1 \quad 0 \quad 0\right]\left[\Phi_1(T,t)H_1^{-1}w_1' - \Phi_2(T,t)H_2^{-1}w_2'\right] \tag{6.43}$$

with

$$W_2'' = \begin{bmatrix}1 & 0 & 0\end{bmatrix} \Big\{ \Phi_1(T,t) H_1^{-1} W_1' H_1^{T^{-1}} \Phi_1(T,t) - \Phi_1 H_1^{-1} C H_2^{T^{-1}} \Phi_2(T,t)$$
$$- \Phi_2(T,t) H_2^{-1} C^T H_1^{T^{-1}} \Phi_1(T,t) + \Phi_2(T,t) H_2^{-1} W_2' H_2^{T^{-1}} \Phi_2(T,t) \Big\} \begin{bmatrix}1 & 0 & 0\end{bmatrix} \tag{6.44}$$

If we define the energy weighting matrices $R_1(t)$ and $R_2(t)$ which in this case are scalars by $r_{11}{}^2(t)$ and $r_{22}{}^2(t)$ respectively. then on using the transformations

$$\begin{aligned} x &= ax' \\ u_1 &= r_{11} u_1' \\ u_2 &= r_{22} u_2' \end{aligned} \tag{6.45}$$

we can write the system equations as

$$\dot{x} = - \frac{aG_1(T,t)}{r_{11}(t)} u_1(t) + \frac{aG_2(T,t)}{r_{22}(t)} u_2(t) \tag{6.46}$$

$$az_2''(t) = x(t) + aw_2''(t) \tag{6.47}$$

Defining

$$G_1(t) = \frac{aG_1'(T,t)}{r_{11}} = \frac{aK_1 \tau_1}{r_{11}} \left[1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1\right]$$

$$G_2(t) = \frac{aG_2'(T,t)}{r_{22}} = \frac{aK_2 \tau_2}{r_{22}} \left[1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2\right]$$

(Cont'd)

$$z_2(t) = az_2''(t)$$

$$w_2(t) = aw_2''(t)$$

$$W_2\,\delta\,(t-\tau) = a^2 W_2''\ (t-\tau). \qquad (6.48)$$

We have the original problem reduced to the notation used in this paper, i.e., the dynamic system is

$$\dot{x}(t) = -\frac{aK_1\,\tau_1}{\tau_{11}}\left[1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1\right]u_1(t) + \frac{aK_2\tau_2}{\tau_{22}}\left[1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2\right]u_2(t) \qquad (6.49)$$

$$z_2(t) = x(t) + w_2(t) \qquad (6.50)$$

with initial condition

$$x(t_o) = x_o = a\left\{\left[y_M(0) - y_T(0)\right] + \left[T - t_o\right]\left[V_M \sin\gamma_{MO} - V_T \sin\gamma_{TO}\right]\right\} \qquad (6.51)$$

for player 1 and an a priori estimate of $x(t_o)$ for player 2.

and with

$$T = \frac{x_T(0) - x_M(0)}{V_c} \qquad (6.52)$$

Or

$$\begin{aligned}\dot{x}(t) &= -G_1(t)u_1(t) + G_2(t)u_2(t) \quad ; \quad x(t_o) = x_o \\ z_2(t) &= H_2x(t) + w_2(t)\end{aligned} \qquad (6.53)$$

and the performance criterion is

$$J(u_1,u_2) = \frac{1}{2}E\left\{x^T(T)x(T) + \int_{t_o}^{T}\left[u_1^T(t)u_1(t) - u_2^T(t)u_2(t)\right]dt\right\} \qquad (6.54)$$

Note that in addition to our simplifying assumptions of small angles, γ and σ, and constant velocities, we have implicitly assumed that player 2's initial estimates are such that the final time T is the same for both players.

6.2 DELAYED COMMITMENT SOLUTION

With reference to Table 5.2, the delayed commitment strategy for player 1 is given by

$$u_1^*(t) = G_1^T(t)S(t)x(t) \tag{6.55}$$

where

$$\dot{S} = \Bigg[\frac{a^2K_1^2\,\tau_1^2}{r_{11}^2}\,(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1)^2 - \frac{a^2K_2^2\;{}_2^2}{r_{22}^2}\,(1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2)^2\Bigg]S^2 \tag{6.56}$$

with

$$S(T) = I \tag{6.57}$$

The above equation is separable and has a closed form solution; i.e.,

$$S^{-1}(t) = 1 + \int_t^T \Bigg[\frac{a^2K_1^2\;{}_1^2}{r_{11}^2}\,(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1)^2 - \frac{a^2K_2^2\;{}_2^2}{r_{22}^2}\,(1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2)^2\Bigg]dt \tag{6.58}$$

and S(t) is found to be

$$S(t) = 6r_{11}^2 r_{22}^2 / \Big\{ 6r_{11}^2 r_{22}^2 + a^2K_1^2 r_{22}^2 \Big[6\tau_1^2\bar{T} - 6\tau_1\bar{T}^2 + 2\bar{T}^3 + 3\tau_1^3(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_1^2\bar{T}e^{-\bar{T}/\tau_1}\Big] - a^2K_2^2 r_{11}^2 \Big[6\tau_2^2\bar{T} - 6\tau_2\bar{T}^2 + 2\bar{T}^3 + 3\tau_2^3(1 - e^{-2\bar{T}/\tau_2}) - 12\tau_2^2\bar{T}e^{-\bar{T}/\tau_2}\Big]\Big\} \tag{6.59}$$

Thus, the optimal delayed commitment control function for player 1 is given by

$$\begin{aligned} u_1^*(t) &= G_1^T(t)S(t)x(t) \\ &= 6aK_1\tau_1 r_{11} r_{22}^2 (1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1)x(t) / \\ &\quad \Big\{ 6r_{11}^2 r_{22}^2 + a^2K_1^2 r_{22}^2 \Big[6\tau_1^2\bar{T} - 6\tau_1\bar{T}^2 + 2\bar{T}^3 + 3\tau_1^3(1 - e^{-2\bar{T}/\tau_1}) \\ &\quad -12\tau_1^2\bar{T}e^{-\bar{T}/\tau_1}\Big] - a^2K_2^2 r_{11}^2 \Big[6\tau_2^2\bar{T} - 6\ {}_2\bar{T}^2 + 2\bar{T}^3 \\ &\quad + 3\tau_2^3(1 - e^{-2\bar{T}/\tau_2}) - 12\tau_2^2\bar{T}e^{-\bar{T}/\tau_2}\Big]\Big\} \end{aligned} \tag{6.60}$$

The corresponding optimal strategy for player 2 at time t is then

$$\begin{aligned} u_2^*(t) &= G_2^T(t)S(t)x(t) \\ &= 6aK_2\tau_2 r_{11}^2 r_{22}(1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2)x(t)/ \end{aligned}$$

(Cont'd)

$$\Big\{6r_{11}^{2}r_{22}^{2} + a^2K_1^{2}r_{22}^{2}\Big[6\tau_1^{2}\bar{T} - 6\tau_1\bar{T}^2 + 2\bar{T}^3 + 3\tau_1^{2}(1 - e^{-2\bar{T}/\tau_1})$$

$$- 12\tau_1^{2}\bar{T}e^{-\bar{T}/\tau_1}\Big] - a^2K_2^{2}r_{11}^{2}\Big[6\tau_2^{2}\bar{T} - 6\tau_2\bar{T}^2 + 2\bar{T}^3$$

$$+ 3\tau_2^{3}(1 - e^{-2\bar{T}/\tau_2}) - 12\tau_2^{2}\bar{T}e^{-T/\tau_2}\Big]\Big\} \tag{6.61}$$

and the secure payoff for player 1 at time t is

$$J_1(u_1^*,u_2^*) = \frac{1}{2}x^T(t)S(t)x(t)$$

$$= 3r_{11}^{2}r_{22}^{2}x^2(t)/\Big\{6r_{11}^{2}r_{22}^{2} + a^2K_1^{2}r_{22}^{2}\Big[6\tau_1^{2}\bar{T}$$

$$- 6\tau_1\bar{T}^2 + 2\bar{T}^3 + 3\tau_1^{3}(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_1^{2}\bar{T}e^{-\bar{T}/\tau_1}\Big]$$

$$- a^2K_2^{2}r_{11}^{2}\Big[6\tau_2^{2}\bar{T} - 6\tau_2\bar{T}^2 + 2\bar{T}^3 + 3\tau_2^{3}(1 - e^{-2\bar{T}/\tau_2})$$

$$- 12\tau_2^{2}\bar{T}e^{-T/\tau_2}\Big]\Big\} \tag{6.62}$$

The delayed commitment strategy for player 2 at time t is given by

$$u_2^*(t) = G_2^T(t)S(t)\hat{x}_2(t) \tag{6.63}$$

which is simply equation (6.61) with $x(t)$ replaced by $\hat{x}_2(t)$. The corresponding optimal strategy for player 1 at time t is then

$$u_1^*(t) = G_1^T(t)S(t)\hat{x}_2(t) + G_1^T(t)N_1(t)\tilde{x}_2(t) \tag{6.64}$$

where N (t) satisfies,

$$\dot{N}_1(t) = \frac{a^2 K_1^{\,2} \tau_1^{\,2}}{r_{11}^{\,2}} \left[1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1 \right]^2 N_1^{\,2} \tag{6.65}$$

with

$$N_1(T) = I \tag{6.66}$$

The above equation is separable and its solution is

$$N_1(t) = 6r_{11}^{\,2} \Big/ \left\{ 6r_{11}^{\,2} + a^2 K_1^{\,2} \left[6\tau_1^{\,2}\,\bar{T} - 6\tau_1 \bar{T}^2 + 2\bar{T}^3 + 3\tau_1^{\,3}(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_1^{\,2}\,\bar{T}^{-\bar{T}/\tau_1} \right] \right\} \tag{6.67}$$

Hence, the optimal response for player 1 at time t is

$$\begin{aligned}
u_1^*(t) = {} & 6aK_1\tau_1 r_{11} r_{22}^{\,2}(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1)\hat{x}_2(t) \Big/ \\
& \left\{ 6r_{11}^{\,2} r_{22}^{\,2} + a^2 K_1^{\,2} r_{22}^{\,2} \left[6\tau_1^{\,2}\,\bar{T} - 6\tau_1 \bar{T}^2 + 3\tau_1^{\,3}(1 - e^{-2\bar{T}/\tau_1}) \right.\right. \\
& \left. - 12\tau_1^{\,2}\bar{T}e^{-\bar{T}/\tau_1} \right] - a^2 K_2^{\,2} r_{11}^{\,2} \left[6\tau_2^{\,2}\,\bar{T} - 6\tau_2 \bar{T}^2 + 2\bar{T}^3 \right. \\
& \left.\left. +3\tau_2^{\,3}(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_2^{\,2}\,\bar{T}e^{-\bar{T}/\tau_2} \right] \right\} \\
& + 6aK_1\tau_1 r_{11}(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1)\hat{x}_2(t) \Big/ \\
& \left\{ 6r_{11}^{\,2} + a^2 K_1^{\,2} \left[6\tau_1^{\,2}\,\bar{T} - 6\tau_1 \bar{T}^2 + 2\bar{T}^3 + 3\tau_1^{\,3}(1 - e^{-2\bar{T}/\tau_1}) \right.\right. \\
& \left.\left. - 12\tau_1^{\,2}\,\bar{T}e^{-\bar{T}/\tau_1} \right] \right\}
\end{aligned} \tag{6.68}$$

If we assume that the noise variance for our example is given by

$$\mathrm{cov}\left[w_2(t), w_2(\tau)\right] = W_2, \tag{6.69}$$

and that the measurement matrix of player 2 is

$$H_2 = h_2 \quad \text{(a scalar)} \tag{6.70}$$

then the expected secure payoff for player 2 at time t is

$$\begin{aligned}
J_2(u_1^\circ, u_2^\circ) &= \frac{1}{2}\hat{x}_2^T(t)S(t)\hat{x}_2(t) + \frac{1}{2}\tilde{x}_2(t)^T N_1(t)\tilde{x}_2(t) \\
&\quad + \frac{1}{2}\,\mathrm{tr}\left[\int_t^T \left[N_1(s) - S(s)\right] P_2^2(s)h_2^2(s)W_2^{-1}ds\right] \\
&= 3r_{11}^2 r_{22}^2 \hat{x}_2^2(t) \Big/ \Big\{6r_{11}^2 r_{22}^2 + a^2K_1^2 r_{22}^2 \Big[6\tau_1^2\bar{T} - 6\tau_1\bar{T}^2 \\
&\quad + 2\bar{T}^3 + 3\tau_1^3(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_1^2\bar{T}e^{-\bar{T}/\tau_1} \\
&\quad - a^2K_2^2 r_{11}^2 \Big[6\tau_2^2\bar{T} - 6\tau_2\bar{T}^2 + 2\bar{T}^3 + 3\tau_2^3(1 - e^{-2\bar{T}/\tau_2}) \\
&\quad - 12\tau_2^2\bar{T}e^{-\bar{T}/\tau_2}\Big]\Big\} + 3r_{11}^2 x_2^2(t) \Big/ \Big\{6r_{11}^2 + a^2K_1^2 \\
&\quad \Big[6\tau_1^2\bar{T} - 6\tau_1^2\bar{T}^2 + 2\bar{T}^3 + 3\tau_1^3(1 - e^{-2\bar{T}/\tau_1}) \\
&\quad - 12\tau_1^2\bar{T}e^{-\bar{T}/\tau_1}\Big]\Big\} + \frac{1}{2}\,\mathrm{tr}\Big[\int_t^T \left[N_1(s) - S(s)\right] \\
&\quad P_2^2(s)h_2^2(s)W_2^{-1}ds\Big]
\end{aligned} \tag{6.71}$$

Note that we could obtain the gain coefficients of the controls for player 1 and player 2 in the delayed commitment solutions in closed form because all the differential equations involved in the computations are initial value problems. Furthermore, as can be seen from Table 5.2, the coefficients of the filtering equations can also be precomputed, i.e., they can be calculated off-line, and are again simple initial value problems.

6.3 PRIOR COMMITMENT SOLUTION

In the case of the prior commitment strategies, the gain matrix for the error term in player 1's control is coupled to the error covariance matrix of the Kalman filter. Let us assume that the covariance of the error of player 2's initial estimate is

$$P_o = p_o \tag{6.72}$$

then from Table 5.1, the following set of simultaneous differential equations are found for the second term in player 2's control and the nature of player 2's estimator,

$$\begin{aligned}\dot{N} = {} & \frac{a^2 K_2^{\,2} \tau_2^{\,2}}{r_{22}^{\,2}} \left| 1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2 \right|^2 S^2(t) \\ & + \frac{a^2 K_1^{\,2} \tau_1^{\,2}}{r_{11}^{\,2}} \left| 1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1 \right|^2 \left| N(t)^2 + 2S(t)N(t) \right| \\ & + \frac{2P^2 h^2}{W_2} N;\end{aligned} \tag{6.73}$$

$$N(T) = 0$$

$$\dot{P} = -2 \frac{a^2 K_1^{\,2} \tau_1^{\,2}}{r_{11}^{\,2}} \left| 1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1 \right|^2 \left| S(t) + N(t) \right| - \frac{P^2 h^2}{W_2}; \tag{6.74}$$

$$P(t_o) = p_o$$

Note that we are now faced with solving a <u>nonlinear</u> two-point boundary value problem. Experience has shown that such a problem if solved directly is very sensitive to the error of the unknowns or guessed initial conditions. Frequently, the guessed value of the missing initial condition has to be practically the correct value before the problem will converge. Hence, we have to resort to such computational techniques as quasilinearization or invariant imbedding to solve the above equations, thus greatly increasing the computational load as compared to simple initial value problems.

Leaving N(t) and P(t) undetermined, the optimal prior commitment strategy controls are then given by

$$
\begin{aligned}
u_1'(t) &= G_1^T(t)S(t)x(t) + G_1^T(t)N(t)\tilde{x}_2(t) \\
&= 6aK_1\tau_1 r_{11} r_{22}^2 (1 - e^{-T/\tau_1} - \bar{T}/\tau_1)x(t)/ \\
&\quad \Big\{ 6r_{11}^2 r_{22}^2 + a^2K_1^2 r_{22}^2 \Big[6\tau_1^2\,\bar{T} - 6\tau_1\bar{T}^2 + 2\bar{T}^3 \\
&\quad + 3\tau_1^3(1 - e^{-\bar{T}/\tau_1}) - 12\tau_1^2\,\bar{T}e^{-\bar{T}/\tau_1}\Big] - a^2K_2^2 r_{11}^2 \\
&\quad \Big[6\tau_2^2\,\bar{T} - 6\tau_2\bar{T}^2 + 2\bar{T}^3 + 3\tau_2^3(1 - e^{-2\bar{T}/\tau_2}) - 12\tau_2^2\bar{T}e^{-\bar{T}/\tau_2}\Big]\Big\} \\
&\quad + \frac{K_1 a\tau_1}{r_{11}}\,(1 - e^{-\bar{T}/\tau_1} - \bar{T}/\tau_1)N(t)\tilde{x}_2(t)
\end{aligned}
\qquad (6.75)
$$

and

$$
\begin{aligned}
u_2'(t) &= G_2^T(t)S(t)\hat{x}_2(t) \\
&= 6aK_2\tau_2 r_{11}^2 r_{22}(1 - e^{-\bar{T}/\tau_2} - \bar{T}/\tau_2)\hat{x}_2(t)/ \\
&\quad \Big\{ 6r_{11}^2 r_{22}^2 + a^2K_1^2 r_{22}^2 \Big[6\tau_1^2\bar{T} - 6\tau_1\bar{T}^2 + 2\bar{T}^3 \\
&\quad + 3\tau_1^3(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_1^2\bar{T}e^{-\bar{T}/\tau_1}\Big]
\end{aligned}
$$

(Cont'd)

$$-a^2K_2^{\,2}r_{11}^{\,2}\Big[6\tau_2^{\,2}\bar{T} - 6\tau_2\bar{T}^2 + 2\bar{T}^3 + 3\tau_2^{\,3}(1 - e^{-2\bar{T}/\tau_2})$$

$$-12\tau_2^{\,2}\,\bar{T}e^{-T/\tau_2}\Big]\Bigg\} \qquad (6.76)$$

The expected prior commitment payoff at time t is given by

$$J(u_1',u_2') = \frac{1}{2}x^T(t)S(t)x(t) + \frac{1}{2}\tilde{x}_2^{\,T}(t)N(t)\tilde{x}_2(t)$$

$$+\frac{1}{2}\,\mathrm{tr}\left[\int_t^T N(\tau)P(\tau)H_2^{\,T}(\tau)W_2^{\,-1}(\tau)H_2(\tau)P(\tau)d\tau\right]$$

$$= 3r_{11}^{\,2}r_{22}^{\,2}x^2(t)\Big/\Bigg\{6r_{11}^{\,2}r_{22}^{\,2} + a^2K_1^{\,2}r_{22}^{\,2}\Big[6\tau_1^{\,2}\bar{T} - 6\tau_1\bar{T}^2$$

$$+ 2\bar{T}^3 + 3\tau_1^{\,3}(1 - e^{-2\bar{T}/\tau_1}) - 12\tau_1^{\,2}\bar{T}e^{-\bar{T}/\tau_1}\Big] - a^2K_2^{\,2}r_{11}^{\,2}$$

$$\Big[6\tau_2^{\,2}\bar{T} - 6\tau_2\bar{T}^2 + 3\tau_2^{\,3}(1 - e^{-2\bar{T}/\tau_2}) - 12\ {}_2^{\,2}\bar{T}e^{-T/\tau_2}\Big]\Bigg\}$$

$$+\frac{1}{2}N(t)\tilde{x}_2^{\,2}(t) + \frac{1}{2}\int_t^T N(s)P^2(s)h_2^{\,2}W_2^{\,-1}ds \qquad (6.77)$$

6.4 NUMERICAL EXAMPLE

In this section we present a numerical example of the pursuit-evasion problem discussed in the previous sections.

In the selection of parameters, the specification of a^2, $R_1 = r_{11}^{\,2}$ and $R_2 = r_{22}^{\,2}$ in the performance criterion (Equation 6.33) has to be such that the terminal miss is acceptably small, and produces tolerable levels of control for the missile and the aircraft.

A choice that frequently results in acceptable levels are [23]:

$$\left[a^2\right]^{-1} = \text{maximum acceptable value of } \left[x_1' - x_2'\right]^2$$

$$\left[r_{11}^{\;2}\right]^{-1} = T \times \text{maximum acceptable value of } \left[u_1'\right]^2$$

$$\left[r_{22}^{\;2}\right]^{-1} = T \times \text{maximum acceptable value of } \left[u_2'\right]^2$$

If we assume a final time T of 10 sec. and a maximum missile acceleration of 10 G's, then $r_{11}^{\;2} = .001(G^2 - \text{sec.})^{-1}$. Similarly, for a maximum airplane acceleration of 5 G's, $r_{22}^{\;2} = .004\ (G^2 - \text{sec.})^{-1}$. Assuming a terminal separation of 5 ft., $a^2 = .04\ (\text{ft.}^2)^{-1}$. The constants and parameters used are summarized in Table 6.1.

By assuming that $r_{11}^{\;2} < r_{22}^{\;2}$ we assure that the relative controllability requirement discussed in Chapter 3 (Equation (3.86)) is satisfied. From the equation for $\dot{N}$ in Table 5.1 and player 2's estimation equations in both the prior commitment and delayed commitment solutions (Tables 5.1 and 5.2), we see that the range of possibilities of the nature of information available to player 2 depends on the ratio $\frac{PH_2}{W_2}$ in the prior commitment or $\frac{P_2H_2}{W_2}$ in the delayed commitment game. We have investigated the effect of the nature of the measurement information of player 2 to the game by varying W_2 over a range from 10 to 10^4 ft.^{2}.

To obtain the results for the prior commitment solution required the solution of a non-linear two-point boundary value problem. The quasilinearization technique was used to obtain the solution. It was found that four iterations were sufficient to converge to the solution.

All solutions were obtained on a Control Data Corp. 6400 digital computer using a fourth-order Runge-Kutta integration technique with

TABLE 6.1

CONSTANTS AND PARAMETERS USED IN A NUMERICAL EXAMPLE OF A PURSUIT-EVASION GAME

T	=	Final time	= 10 sec.
t_o	=	Initial time	= 0 sec.
K_1	=	$32.2 \cos \gamma_{MO}$	= 32.2 ft/sec^2 - G
K_2	=	$32.2 \cos \gamma_{TO}$	= 32.2 ft/sec^2 - G
τ_1	=	Missile time constant	= 1 sec.
τ_2	=	Airplane time constant	= 2 sec.
a^2	=	Terminal miss weighting factor	= .04 $(\text{ft}^2)^{-1}$
${r_{11}}^2$	=	Missile control weighting factor	= .001 $(G^2 - \text{sec})^{-1}$
${r_{22}}^2$	=	Airplane control weighting factor	= .004 $(G^2 - \text{sec})^{-1}$
P_o	=	Initial error covariance	= 100 ft^2
W_2	=	Measurement noise covariance	= $10 \longrightarrow 10^4$ ft.^{2}

an integration interval of .01 seconds. A listing of the computer program is presented in Appendix A. No attempt has been made to optimize the computer program.

The error variance of player 2, $P_2(t)$, in the delayed commitment game is shown for various values of W_2 in Figure 6.2. The error variance, $P(t)$, in the prior commitment game is shown in Figure 6.3 for the same range of values of W_2. The delayed - and prior commitment error variances differ at most by 3.2 percent.

The feedback gains $G_1(t)S(t)$, $G_2(t)S(t)$ and $G_1(t)N(t)$ for the example from zero to 7.5 seconds are shown in Figure 6.4 and on a less sensitive scale from 7.5 seconds to terminal time at 10 seconds in Figure 6.5. The curves for $G_1(t)S(t)$ and $G_2(t)S(t)$ are of course, independent of W_2, but it was found that $G_1(t)N(t)$ is also appropriate for all values of W_2 in the range from 10 to 10^4 ft.^{2}. Near the terminal time $G_1(t)N(t)$ is completely independent of W_2 and varies less than .1 percent at t = 5 seconds for the range of W_2 indicated above. This is clear from the equation for $\dot{N}$ in Table 5.1 which shows that W_2 effects $N(t)$ through the term P/W_2 and for the latter half of the game $P(t)$ is so small that W_2 cannot have an appreciable effect on $N(t)$. Only near the beginning of the game does $G_1(t)N(t)$ vary with W_2 but its value is so small that it cannot be displayed on Figure 6.4. At t = 0, the values for $G_1(t)N(t)$ are given in Table 6.2.

The curve for $G_1(t)\left|N_2(t) - S(t)\right|$ follows that of $G_1(t)N(t)$ so close as to be indistinguishable on Figures 6.4 and 6.5, the values are compared at various times for W_2 = 100 ft.^{2} in Table 6.3. After t = 6 seconds, the two values are identical to four decimal places.

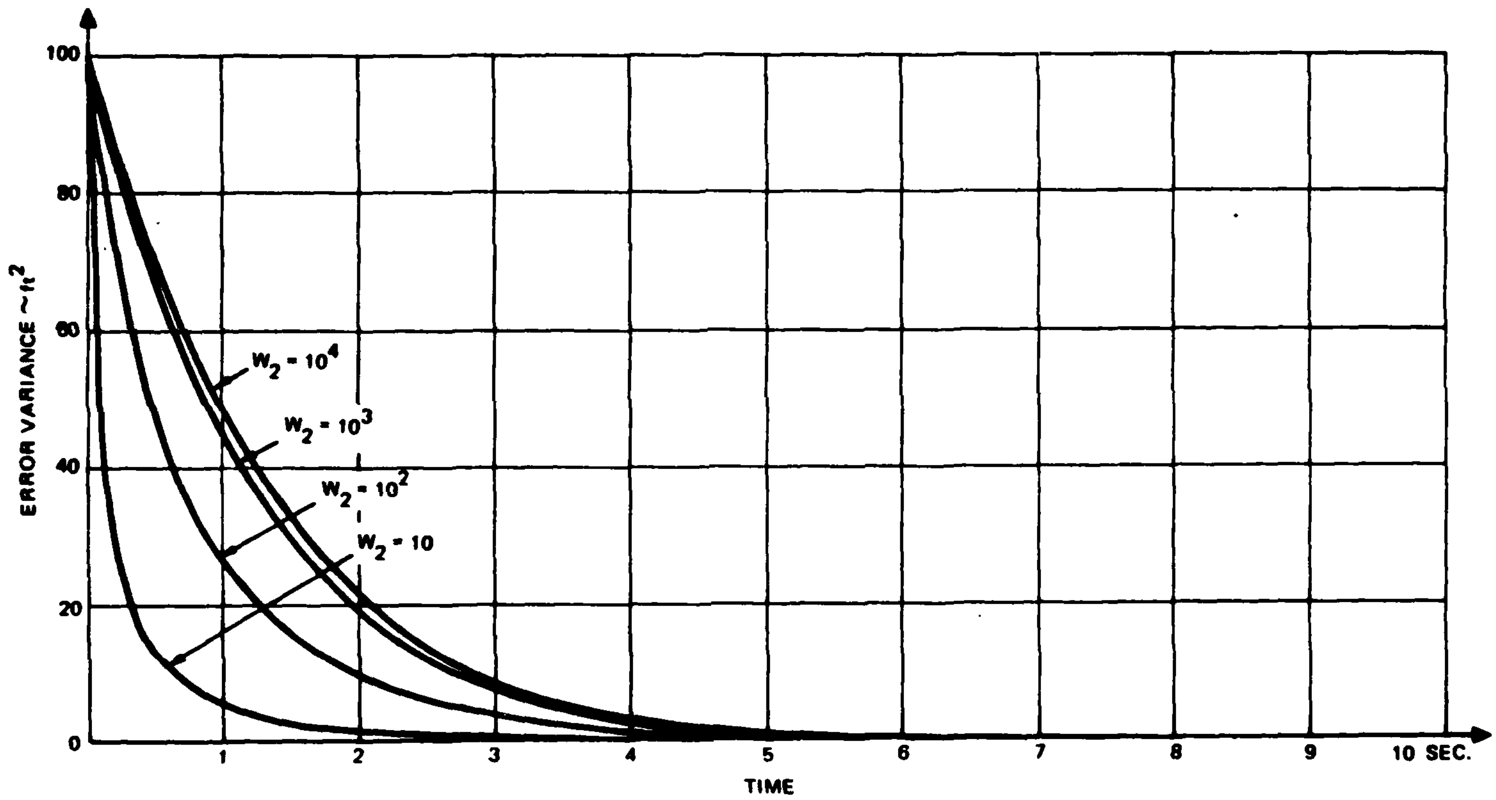


Figure 6.2 Error Variance of Player 2 in Delayed Commitment Game

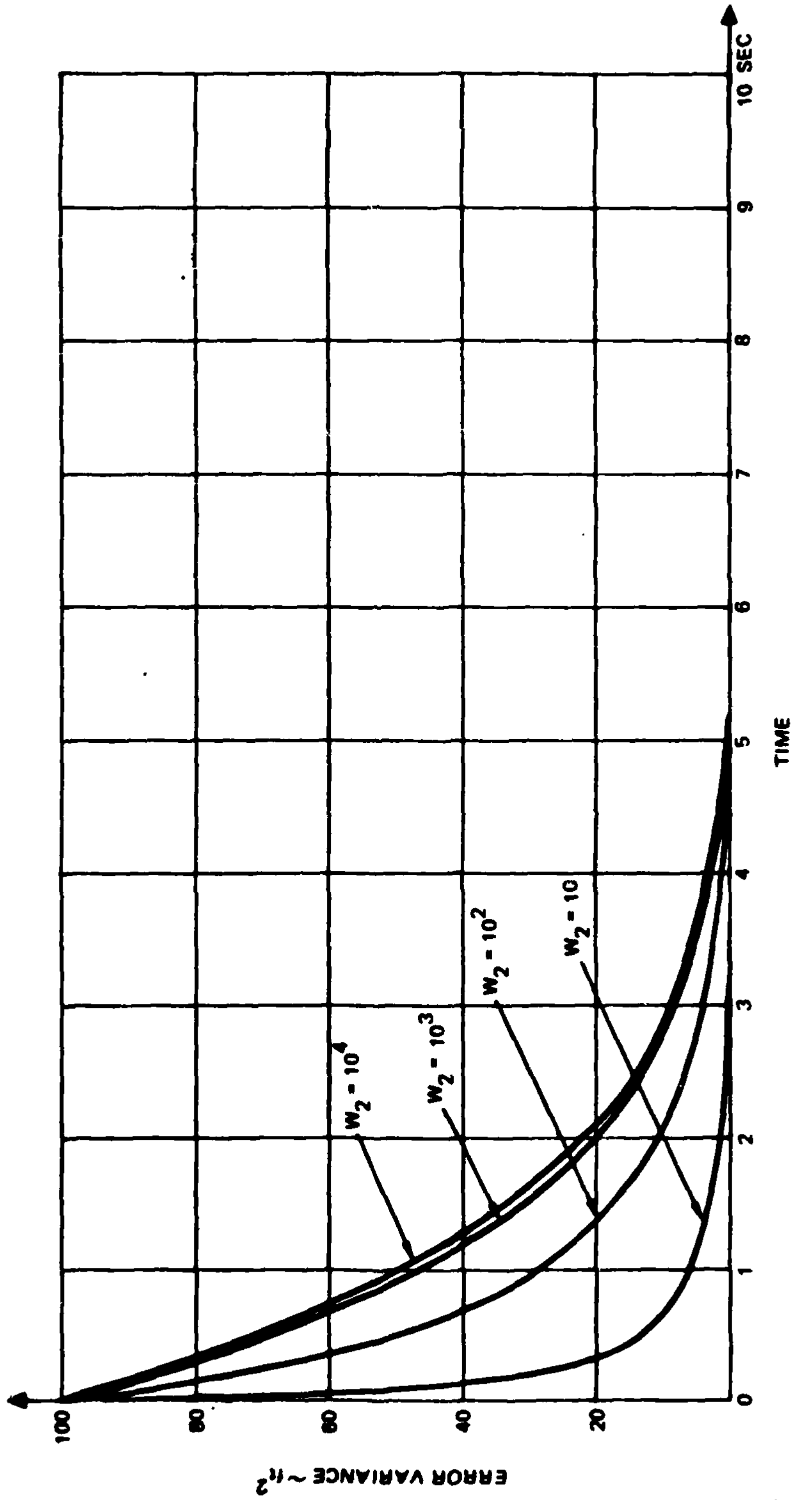


Figure 6.3 Error Variance of Player 2 in Prior Commitment Game

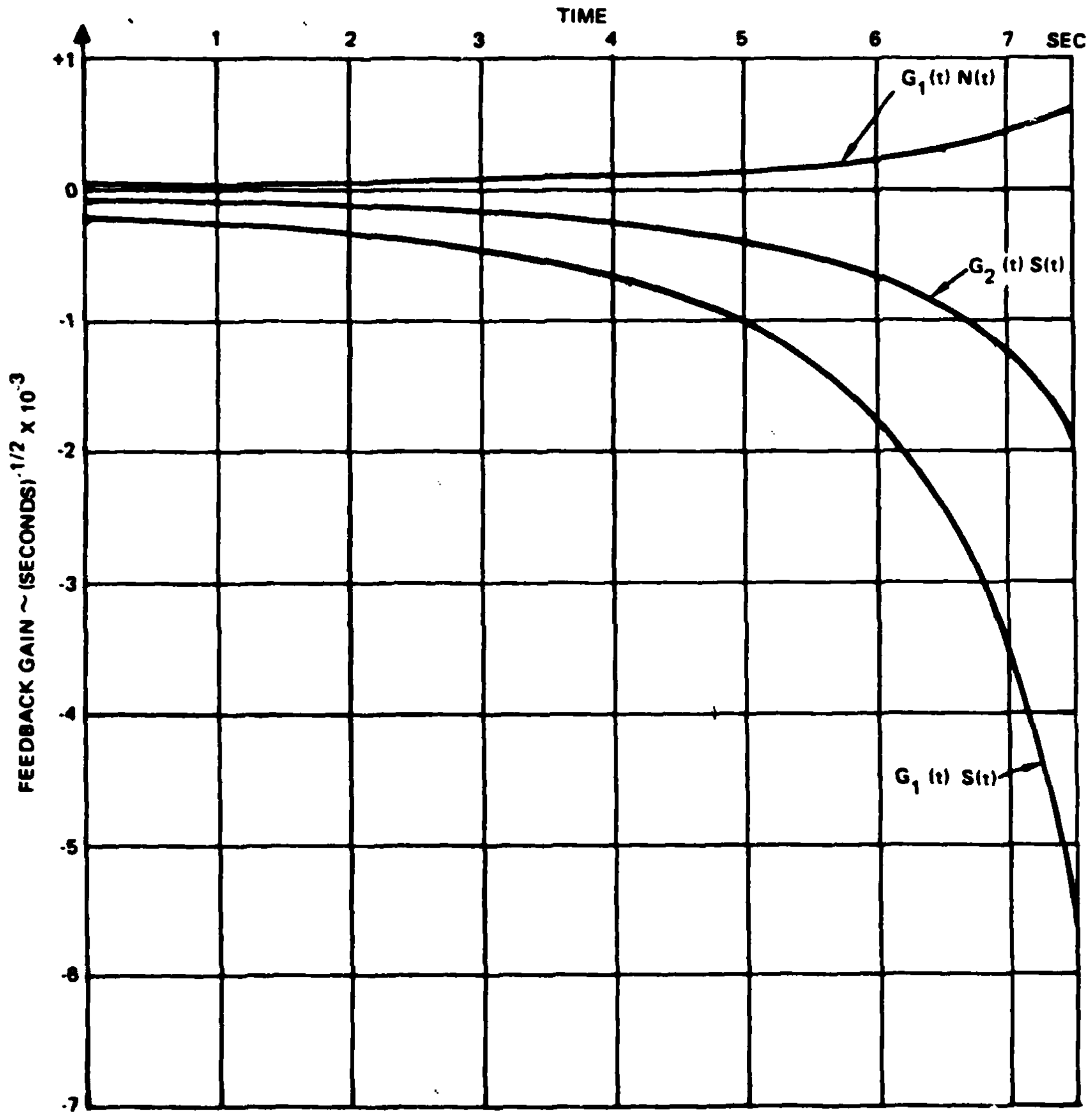


Figure 6.4 Feedback Gains Versus Time

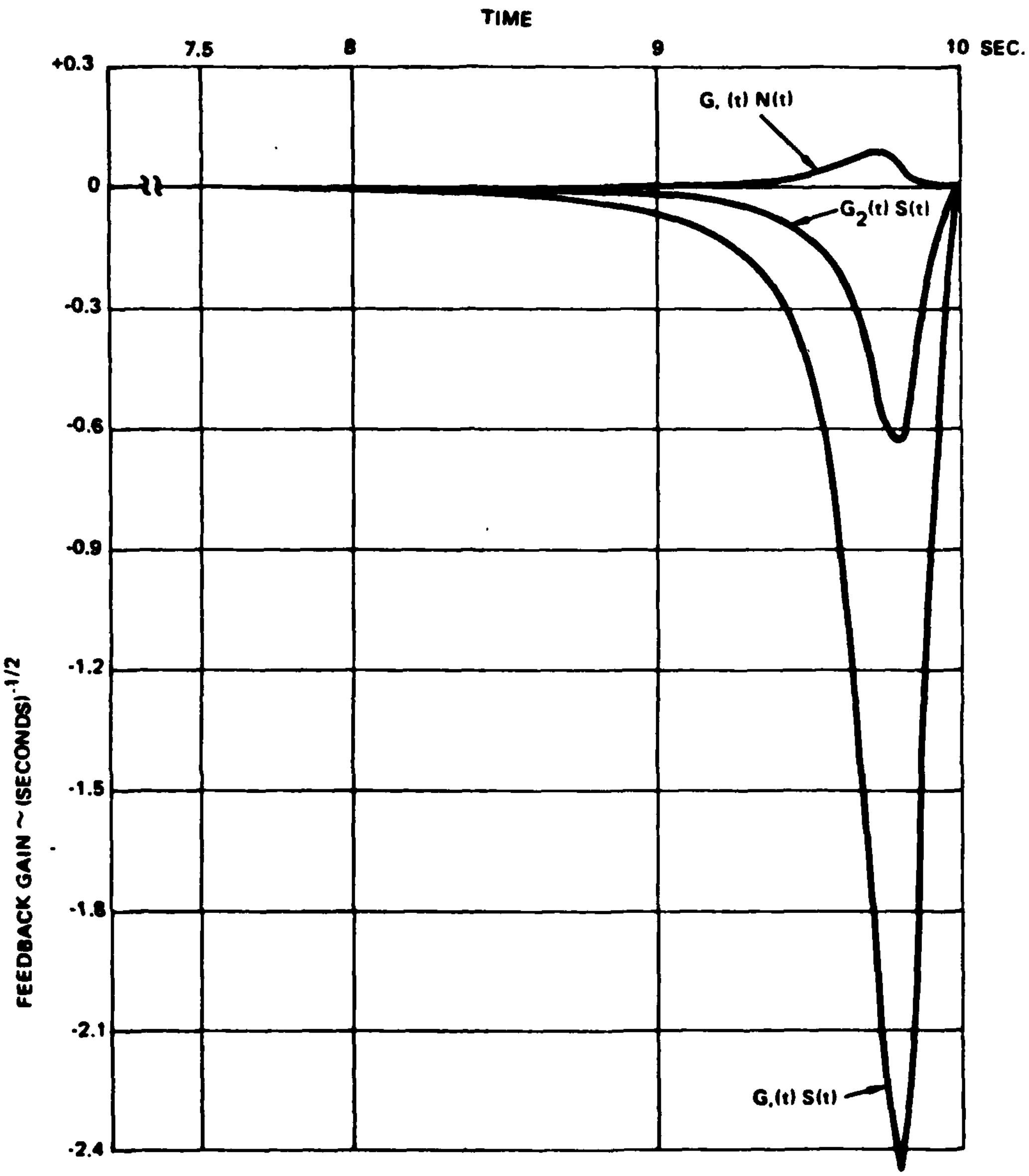


Figure 6.5 Feedback Gains Versus Time

TABLE 6.2

VALUES OF $G_1(t)N(t)$ AT $t = 0$

W_2 FT^2	$G_1(0)\ N(0)$ $(SEC)^{-1/2}$
10	.1864 E-05
10^2	.1432 E-04
10^3	.3416 E-04
10^4	.3944 E-04

TABLE 6.3

COMPARISON OF $G_1(t)N(t)$ WITH

$G_1(t)\left[N_1(t) - S(t)\right]$ FOR $W_2 = 100\ FT^2$

TIME SEC.	$G_1(t)N\ (t)$ $(SEC)^{-1/2}$	$G_1(t)\left[N_1(t) - S(t)\right]$ $(SEC)^{-1/2}$
0	.1432 E-04	.4013 E-04
.5	.2444 E-04	.4408 E-04
1.0	.3418 E-04	.4866 E-04
1.5	.4365 E-04	.5403 E-04
2.0	.5313 E-04	.6037 E-04
3.0	.7394 E-04	.7714 E-04
4.0	.1014 E-03	.1026 E-03
5.0	.1443 E-03	.1447 E-03
6.0	.2252 E-03	.2252 E-03

The difference between the prior commitment payoff, J, and the delayed commitment payoff for player 1, J_1, shows the dependence of J on W_2 and is defined in this paper as the relative criterion of the prior commitment game. The relative criterion for the delayed commitment game is obtained by taking the difference between J_2 and J_1.

The relative payoffs for a W_2 of 10^3 ft.^{2} are shown in Figure 6.6. The relative payoffs are always negative, indicating a reduction in player 2's payoff compared to the perfect information game. Furthermore, the relative payoff for the delayed commitment game (J_2 - J_1) is more negative than that of the prior commitment game (J - J_1) indicating the relationship between the payoffs as discussed in Section 5.3 (see Figure 5.1).

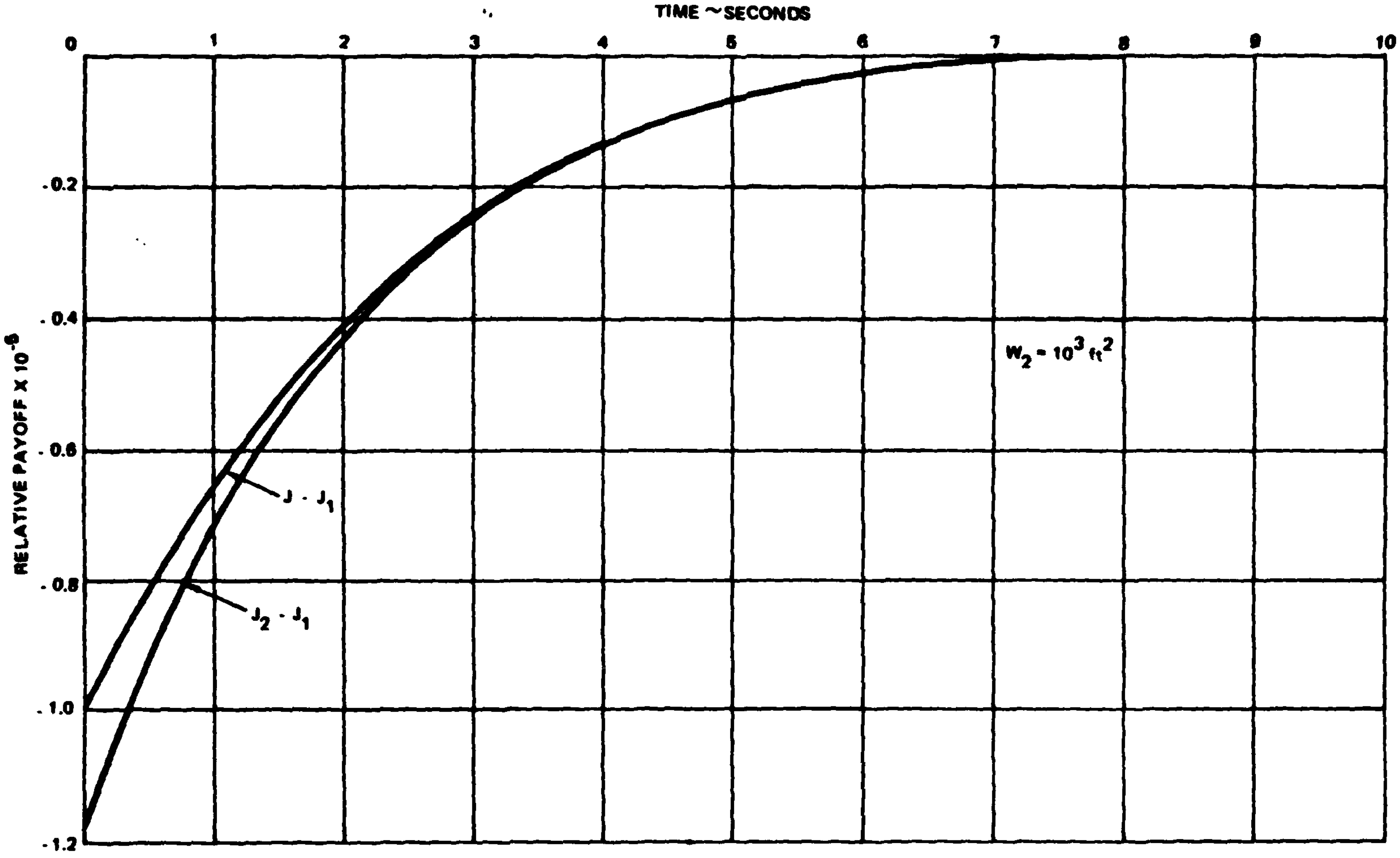


Figure 6.6 Relative Payoff Versus Time

CHAPTER 7

THE NOISY/NOISY DIFFERENTIAL GAME

In this chapter we extend the presentation of the previous chapters where either both players or only one player had perfect state information to the case where both players have noise corrupted measurements.

Since both players are faced with the problem of extracting useful information from their noise corrupted measurements, and neither player can determine exactly his opponent's estimation error, we are led in the prior commitment formulation to the addition of correction terms in each player's controller and thus initiate the vicious cycle of estimates of estimates.

The problem formulation for this chapter is as defined in Section 4.1. The basic equations are repeated below, but for a more careful definition the reader is referred to the above mentioned section. The dynamic system is described by

$$\begin{aligned} \dot{x}(t) &= \frac{dx}{dt} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \\ z_1(t) &= H_1(t)x(t) + w_1(t) \qquad\qquad (7.1) \\ z_2(t) &= H_2(t)x(t) + w_2(t) \end{aligned}$$

The noise processes $\left|w_1(t)\right|$ and $\left|w_2(t)\right|$ are white Gaussian, with properties

$$\mathrm{cov}\left|w_1(t),\quad w_1(\tau)\right| = W_1(t)\,\delta(t-\tau)$$
$$\mathrm{cov}\left|w_2(t),\quad w_2(\tau)\right| = W_2(t)\,\delta(t-\tau)$$
$$\mathrm{cov}\left|w_1(t),\quad w_2(\tau)\right| = 0 \tag{7.2}$$

For simplicity it is assumed that both players consider the initial condition $x(t_o)$ to be a Gaussian random variable, uncorrelated for all t with $w_1(t)$ and $w_2(t)$, and having a mean of $\bar{x}_o$ and a covariance

$$\mathrm{cov}\left|x(t_o),x(t_o)\right| = P_o \tag{7.3}$$

The cost functional or payoff to the game is quadratic:

$$J(u_1,u_2) = \frac{1}{2}\,E\left|x^T(T)x(T) + \int_{t_o}^{T} u_1^T(t)u_1(t)dt - \int_{t_o}^{T} u_2^T(t)u_2(t)dt\right| \tag{7.4}$$

The class of admissible strategies are restricted to those U_1 and U_2 which give rise to the feedback control laws

$$\begin{aligned} U_1 &: u_1 = u_1(Z_1(t),t) \\ U_2 &: u_2 = u_2(Z_2(t),t) \end{aligned} \tag{7.5}$$

The delayed commitment strategy to the above defined stochastic differential game is obtained in Section 1 for player 1 and in Section 2 we obtain the delayed commitment solution for player 2.

7.1 DELAYED COMMITMENT SOLUTION FOR PLAYER 1

From the point of view of the minimizing player, player 1, the performance criterion during the actual play of the game at time t is

$$J_1(u_1,u_2) = \frac{1}{2} E \left\{ x^T(T)x(T) + \int_t^T \left[u_1^T(\tau)u_1(\tau) - u_2^T(\tau)u_2(\tau) \right] d\tau \mid Z_1(t) \right\} \tag{7.6}$$

and he obtains his secure strategy by finding the saddle-point solution to above equation subject to

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \quad ; \quad x(t_o) = \bar{x}_o \tag{7.7}$$

Similarly to our assumption in Chapter 5 we assume, for the purpose of determining player 1's secure strategy solution, that the allowable strategy for player 2 in addition to being $Z_2(t)$ measurable is also $Z_1(t)$ measurable. Thus we want to determine that $u_1^* \in U_1$ and $u_2^* \in U_1 \times U_2$ which are optimal in the sense that for all $t \in [t_o, T]$

$$J_1(u_1^*,u_2) \leq J_1(u_1^*,u_2^*) \leq J_1(u_1,u_2^*) \tag{7.8}$$

Hence from player 1's point of view of a secure strategy, player 2 maximizes at $t > t_o$

$$\max_{u_2 \in U_1 \times U_2} \frac{1}{2} E \left\{ x^T(T)x(T) + \int_t^T \left[u_1^T(\tau)u_1(\tau) - u_2^T(\tau)u_2(\tau) \right] d\tau \mid Z_1(t), Z_2(t) \right\} \tag{7.9}$$

subject to

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \quad ; \quad x(t_o) = \bar{x}_o \tag{7.10}$$

But for arbitrary $t = t_o, u_1(t)$ and $u_2(t)$ we can write the solution to the system Equation (7.7) as

$$x(t) = \Phi(t,t_o)x_o + \int_{t_o}^{T} \left| - \Phi(T,\tau)G_1(\tau)u_1(\tau) + \Phi(T,\tau)G_2(\tau)u_2(\tau) \right| d\tau \tag{7.11}$$

where $\Phi(t,t_o)$ is the state transition matrix which must satisfy the relation

$$\frac{\partial \Phi(t,t_o)}{\partial t} = F(t)\,\Phi(t,t_o) \tag{7.12}$$

$$\Phi(t_o,t_o) = I$$

Hence in terms of the Hilbert space notation developed in Chapter 3 we can write

$$\max_{u_2 \in U_1 \times U_2} \frac{1}{2} E \left| \langle \Phi x_o - T_1u_1 + T_2u_2, \Phi x_o - T_1u_1 + T_2u_2 \rangle + \langle u_1,u_1 \rangle - \langle u_2,u_2 \rangle \mid Z_1(t),Z_2(t) \right| \tag{7.13}$$

If we now define

$$P(t) = E \left| (x - \hat{x})(x - \hat{x})^T \mid Z_1(t),Z_2(t) \right| \tag{7.14}$$

where

$$\hat{x} = E \left| x(t) \mid Z_1(t),Z_2(t) \right| \tag{7.15}$$

and consider the term $E \left| \langle \Phi x_o, \Phi x_o \rangle \mid Z_1(t)Z_2(t) \right|$ of Equation (7.13), then we can write for arbitrary $t = t_o$

$$E\left|<\Phi x, \Phi x> \mid Z_1(t), Z_2(t)\right| = E\left|<\Phi(x - \hat{x} + \hat{x}), \Phi(x - \hat{x} + \hat{x})>\right.$$

$$\left.\mid Z_1(t), Z_2(t)\right|$$

$$= E\left|<\Phi(x - \hat{x}), \Phi(x - \hat{x})> + <\Phi(x - \hat{x}), \Phi\hat{x}>\right.$$

$$\left. + <\Phi\hat{x}, \Phi(x - \hat{x})> + <\Phi\hat{x}, \Phi\hat{x}> \mid Z_1(t), Z_2(t)\right| \qquad (7.16)$$

But, the two middle terms in the above expression are equal to zero, while the first term can be written as tr $\left|\Phi^T \Phi P\right|$, thus

$$E\left|<\Phi x, \Phi x> \mid Z_1(t), Z_2(t)\right| = \mathrm{tr}\left|\Phi^T \Phi P\right| + <\Phi\hat{x}, \Phi\hat{x}> \qquad (7.17)$$

and we can rewrite Equation (7.13) as

$$\max_{u_2 \in U_1 \times U_2} \frac{1}{2}\left|<\Phi\hat{x}_o - T_1u_1 + T_2u_2, \Phi\hat{x}_o - T_1u_1 + T_2u_2>\right.$$

$$\left. + <u_1, u_1> - <u_2, u_2>\right| + \frac{1}{2}\mathrm{tr}\left|\Phi^T \Phi P_o\right| \qquad (7.18)$$

However, tr $\left|\Phi^T \Phi P_o\right|$ is independent of the control u_2, thus maximizing Equation (7.18) with respect to $u_2(t)$ for arbitrary $u_1(t)$ is equivalent to maximizing $\bar{J}_1(u_1, u_2)$, where

$$\bar{J}_1(u_1, u_2) = \frac{1}{2}\left|\Phi\hat{x}_o - T_1u_1 + T_2u_2, \Phi\hat{x}_o - T_1u_1 + T_2u_2>\right.$$

$$\left. + <u_1, u_1> - <u_2, u_2>\right| \qquad (7.19)$$

From the results of Chapter 3 we know that, whenever the inverse of $(I - T_2T_2^*)$ exists, the candidate extremal control u_2° is

$$
\begin{aligned}
u_2^* &= T_2^*(I - T_2T_2^*)^{-1}(\Phi\hat{x}_o - T_1u_1) \\
&= T_2^*D_2(\Phi\hat{x}_o - T_1u_1),
\end{aligned}
\tag{7.20}
$$

Furthermore, the linear-Gaussian assumptions imply that $\hat{x}_o$ can be generated for any time t by a Kalman-Bucy filter using a prior estimate of the initial state, $\hat{x}_o$, a prior estimate of the variance of the error of this estimate, P_o, the noise corrupted measurements $z_1(t)$ and $z_2(t)$ of the state up to time t and the dynamic equations

$$
\begin{aligned}
\dot{\hat{x}}(t) &= F(t)\hat{x}(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \\
&\quad + P(t)\left[H_1^T(t) \;\vdots\; H_2^T(t)\right] \left[\begin{array}{c|c} W_1^{-1}(t) & 0 \\ \hline 0 & W_2^{-1}(t) \end{array}\right] \left[\begin{array}{c} z_1(t) - H_1(t)\hat{x}(t) \\ \hline z_2(t) - H_2(t)\hat{x}(t) \end{array}\right]
\end{aligned}
\tag{7.21}
$$

with

$$\hat{x}(t_o) = \bar{x}_o$$

and

$$
\begin{aligned}
\dot{P}(t) &= F(t)P(t) + P(t)F^T(t) \\
&\quad - P(t)\left[H_1^T(t) \quad H_2^T(t)\right] \left[\begin{array}{c|c} W_1^{-1}(t) & 0 \\ \hline 0 & W_2^{-1}(t) \end{array}\right] \left[\begin{array}{c} H_1(t) \\ \hline H_2(t) \end{array}\right] P(t)
\end{aligned}
\tag{7.22}
$$

with

$$P(t_o) = P_o$$

Thus we can write

$$u_2 = T_2^* D_2(\Phi \hat{x} - T_1 u_1) \tag{7.23}$$

Substituting Equation (7.23) into Equation (7.6) we obtain as payoff functional for player 1 at arbitrary time $t = t_o$

$$J_1(u_1) = \frac{1}{2} E \Big| \langle \Phi\hat{x}_o - T_1u_1 + T_2T_2^*D_2(\Phi\hat{x}_o - T_1u_1), \Phi\hat{x}_o - T_1u_1 + T_2T_2^*D_2(\Phi\hat{x}_o - T_1u_1)\rangle + \langle u_1,u_1\rangle - \langle T_2T_2^*D_2(\Phi\hat{x}_o - T_1u_1), D_2(\Phi\hat{x}_o - T_1u_1)\rangle \mid Z_1 \Big| \tag{7.24}$$

which player 1 seeks to minimize.

If we define

$$P_1(t) \triangleq E \Big| (x - \hat{x}_1)(x - \hat{x}_1)^T \mid Z_1(t) \Big|$$

where

$\hat{x}_1(t) = E \Big| x(t) \mid Z_1(t) \Big|$ and recalling that the double expectation-first given more information, then less information (information is taken away) - is the same as the expectation given the less information only see [24], then we can write $J_1(u_1)$ as

$$J_1(u_1) = \frac{1}{2} \Big| \langle \Phi\hat{x}_{1_o} - T_1u_1 + T_2T_2^*D_2(\Phi\hat{x}_{1_o} - T_1u_1), \Phi\hat{x}_{1_o} - T_1u_1 + T_2T_2^*D_2(\Phi\hat{x}_{1_o} - T_1u_1)\rangle + \langle u_1,u_1\rangle - \langle T_2^*D_2(\Phi\hat{x}_{1_o} - T_1u_1), T_2^*D_2(\Phi\hat{x}_{1_o} - T_1u_1)\rangle \Big| + \frac{1}{2} \operatorname{tr} (\Phi^T \Phi P_1) \tag{7.25}$$

Minimizing the above expression with respect to $u_1(t)$ is equivalent to minimizing $\bar{J}_1(u_1)$, where

$$\bar{J}_1(u_1) = \frac{1}{2} \Big| < \Phi \hat{x}_{1_o} - T_1 u_1 + T_2 T_2^* D_2 (\Phi \hat{x}_{1_o} - T_1 u_1), \Phi \hat{x}_{1_o} - T_1 u_1 + T_2 T_2^* D_2 (\Phi \hat{x}_{1_o} - T_1 u_1) > + (u_1, u_1) - < T_2 T_2^* D_2 (\Phi \hat{x}_{1_o} - T_1 u_1), D_2 (\Phi \hat{x}_{1_o} - T_1 u_1) > \Big| \tag{7.26}$$

Again drawing upon the results obtained in Chapter 3, we know that the minimizing control for player 1 is

$$\begin{aligned} u_1^\circ &= T_1^* \left[I + T_1 T_1^* - T_2 T_2^* \right]^{-1} \Phi \hat{x}_{1_o} \\ &= T_1^* D \Phi \hat{x}_{1_o} \end{aligned} \tag{7.27}$$

The dynamic system from player 1's point of view can then be written from Equations (7.7) and (7.21) as

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \left[\begin{array}{c|c} F & G_2 T_2^* D_2 \Phi \\ \hline P H_2^T W_2^{-1} H_2 & F + G_2 T_2^* D_2 \Phi - P H_1^T W_1^{-1} H_1 - P_{12} H_2^T W_1^{-1} H_2 \end{array} \right] \begin{bmatrix} x \\ \hat{x} \end{bmatrix} - \left[\begin{array}{c} G_1 + G_2 T_2^* D_2 T_1 \\ \hline G_1 + G_2 T_2^* D_2 T_1 \end{array} \right] u_1 + \left[\begin{array}{c} 0 \\ \hline P H_1^T W_1^{-1} \end{array} \right] z_1 + \left[\begin{array}{c} 0 \\ \hline P H_2^T W_2^{-1} \end{array} \right] w_2 \tag{7.28}$$

with initial conditions

$$\begin{aligned} x(t_o) &= \bar{x}_o \\ \hat{x}(t_o) &= \bar{x}_o \end{aligned} \tag{7.29}$$

and with measurements

$$z_1(t) = H_1(t)\, x(t) + v_1(t) \tag{7.30}$$

It then follows from the linear-Gaussian assumptions that the optimal estimates of x(t) and $\hat{x}(t)$ given $Z_1(t)$, i.e.,

$$\hat{x}_1(t) = E\left\{x(t) \mid Z_1(t)\right\}$$

$$\hat{x}_{12}(t) = E\left\{\hat{x}(t) \mid Z_1(t)\right\} = E\left\{E\left\{x(t) \mid Z_1(t), Z_2(t)\right\} \mid z_1(t)\right\} \qquad (7.31)$$

$$= E\left\{x(t) \mid Z_1(t)\right\} = \hat{x}_1(t)$$

are obtained from (7.32)

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_{12} \end{bmatrix} = \left[\begin{array}{c|c} F & G_2T_2^*D_2\Phi \\ \hline PH_2^TW_2^{-1}H_2 & F + G_2T_2^*D_2\Phi - PH_1^TW_1^{-1}H_1 - PH_2^TW_1^{-1}H_2 \end{array}\right] \begin{bmatrix} \hat{x}_1 \\ \hat{x}_{12} \end{bmatrix}$$

$$- \begin{bmatrix} G_1 + G_2T_2^*D_2T_1 \\ G_1 + G_2T_2^*D_2T_1 \end{bmatrix} u_1 + \begin{bmatrix} 0 \\ PH_1^TW_1^{-1} \end{bmatrix} z_1$$

$$+ \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \begin{bmatrix} H_1^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 - H_1\hat{x}_1 \\ 0 \end{bmatrix}$$

with initial conditions

$$\hat{x}_1(t_o) = \bar{x}_o$$
$$\hat{x}_{12}(t_o) = \bar{x}_o \qquad (7.33)$$

while the error covariance matrix satisfies

$$\dot{P}_1 = \begin{bmatrix} \dot{P}_{11} & \dot{P}_{12} \\ \dot{P}_{12} & \dot{P}_{22} \end{bmatrix} = \left[\begin{array}{c|c} F & G_2T_2^{*}D_2\Phi \\ \hline PH_2^{T}W_2^{-1}H_2 & F + G_2T_2^{*}D_2\Phi - PH_1^{T}W_1^{-1}H_1 - PH_2^{T}W_2^{-1}H_2 \end{array}\right]$$

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} + \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \left[\begin{array}{c|c} F & G_2T_2^{*}D_2\Phi \\ \hline PH_2^{T}W_2^{-1}H_2 & F + G_2T_2^{*}D_2\Phi - PH_1^{T}W_1^{-1}H_1 - PH_2^{T}W_2^{-1}H_2 \end{array}\right]^T + \begin{bmatrix} 0 & 0 \\ 0 & PH_2^{T}W_2^{-1}H_2P \end{bmatrix}$$

$$- \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \begin{bmatrix} H_1^{T}W_1^{-1}H_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \tag{7.34}$$

with

$$P_1(t_o) = \begin{bmatrix} P_o & 0 \\ 0 & 0 \end{bmatrix} \tag{7.35}$$

The optimal minimizing control for player 1 can thus be written

$$u_1^{*} = T_1^{*}D\Phi\hat{x}_1 \tag{7.36}$$

and the corresponding optimal response of player 2 is then

$$u_2^{*} = T_2^{*}D_2\Phi\hat{x} - T_2^{*}D_2T_1T_1^{*}\Phi\hat{x}_1 \tag{7.37}$$

The Kalman-Bucy filter (Equation (7.21)) and its corresponding error covariance matrix (Equation (7.22)) can be simplified by the following observations. Rewriting the conditional estimates (Equation (7.32)) we obtain

$$\dot{\hat{x}}_1 = F\hat{x}_1 + G_2T_2{}^*D_2\Phi\hat{x}_{12} - \left[G_1 + G_2T_2{}^*D_2T_1\right] u_1 + P_{11}H_1{}^TW_1{}^{-1}\left[z_1 - H_1\hat{x}_1\right] \quad ; \quad \hat{x}_1(t_o) = \bar{x}_o \tag{7.38}$$

and

$$\dot{\hat{x}}_{12} = F\hat{x}_{12} + G_2T_2{}^*D_2\Phi\hat{x}_{12} - \left[G_1 + G_2T_2{}^*D_2T_1\right] u_1 + P_{12}H_1{}^TW_1{}^{-1}\left[z_1 - H_1\hat{x}_1\right] + PH_1{}^TW_1{}^{-1}\left[z_1 - H_1\hat{x}_1\right] + PH_2{}^TW_2{}^{-1}H_2\hat{x}_1 - PH_2{}^TW_2{}^{-1}H_2\hat{x}_{12} \quad ; \quad \hat{x}_{12}(t_o) = \bar{x}_o \tag{7.39}$$

and we observe that, since

$$\hat{x}_1(t) \triangleq E\left[x \mid Z_1(t)\right] = \hat{x}_{12}(t) \triangleq E\left[E\left[x \mid Z_1(t)Z_2(t)\right] \mid Z_1(t)\right] \tag{7.40}$$

it implies that

$$P_1(t) = P_{11}(t) = P(t) + P_{12}(t) \tag{7.41}$$

and thus

$$\dot{P}_{11}(t) = \dot{P}(t) + \dot{P}_{12}(t) \tag{7.42}$$

Using Equation (7.42) we then obtain from the error covariance matrix (Equation (7.34))

$$\begin{aligned}\dot{P}_{12} = {} & FP_{12} + G_2T_2{}^*D_2\Phi P_{22} + P_{11}H_2{}^TW_2{}^{-1}H_2P + P_{12}F^T \\ & + P_{12}\left[G_2T_2{}^*D_2\Phi\right]^T - P_{12}H_1{}^TW_1{}^{-1}H_1P - P_{12}H_2{}^TW_2{}^{-1}H_2P \\ & - PH_1{}^TW_1{}^{-1}H_1P_{12} - P_{12}H_1{}^TW_1{}^{-1}H_1P_{12} \quad ; \quad P_{12}(t_o) = 0 \end{aligned} \tag{7.43}$$

and

$$\begin{aligned}\dot{P}_{22} = {} & PH_2{}^TW_2{}^{-1}H_2P_{12} + FP_{22} + G_2T_2{}^*D_2\Phi P_{22} - PH_1{}^TW_1{}^{-1}H_1P_{22} \\ & - PH_2{}^TW_2{}^{-1}H_2P_{22} + P_{12}H_2{}^TW_2{}^{-1}H_2P + P_{22}F^T \\ & + P_{22}\left[G_2T_2{}^*D_2\;\Phi\right]^T - P_{22}H_1{}^TW_1{}^{-1}H_1P - P_{22}H_2{}^TW_2{}^{-1}H_2P \\ & + PH_2{}^TW_2{}^{-1}H_2P - P_{12}H_1{}^TW_1{}^{-1}H_1P_{12} \quad ; \quad P_{22}(t_o) = 0 \end{aligned} \tag{7.44}$$

where the last equation simplifies to

$$\begin{aligned}\dot{P}_{22} = {} & FP_{22} + G_2T_2{}^*D_2\Phi P_{22} + P_{11}H_2{}^TW_2{}^{-1}H_2P + P_{22}F^T \\ & + P_{22}\left[G_2T_2{}^*D_2\,\Phi\right]^T - P_{22}H_1{}^TW_1{}^{-1}H_1P - P_{22}H_2{}^TW_2{}^{-1}H_2P \\ & - PH_1{}^TW_1{}^{-1}H_1P_{22} - P_{12}H_1{}^TW_1{}^{-1}H_1P_{12} \quad ; \quad P_{22}(t_o) = 0 \end{aligned} \tag{7.45}$$

Comparing Equation (7.43) with Equation (7.45) we see that

$$P_{12}(t) = P_{22}(t) \tag{7.46}$$

Thus the Kalman-Bucy filter and its error covariance matrix Equations (7.32) and (7.34) respectively, reduce to

$$\dot{\hat{x}}_1 = (F + G_2T_2{}^*D_2\Phi)\hat{x}_1 - (G_1 + G_2T_2{}^*D_2T_1)u_1$$

$$+ P_{11}H_1{}^TW_1{}^{-1}(z_1 - H_1\hat{x}_1) \quad ; \quad \hat{x}_1(t_o) = \bar{x}_o \tag{7.47}$$

and

$$\dot{P}_{11} = FP_{11} + P_{11}F^T + G_2T_2{}^*D_2\Phi P_{12} + P_{12}(G_2T_2{}^*D_2\Phi)^T$$

$$- P_{11}H_1{}^TW_1{}^{-1}H_1P_{11} \quad ; \quad P_{11}(t_o) = P_o \tag{7.48}$$

$$\dot{P}_{12} = FP_{12} + P_{12}F^T + G_2T_2{}^*D_2\Phi P_{12} + P_{12}(G_2T_2{}^*D_2\Phi)^T$$

$$- P_{11}H_1{}^TW_1{}^{-1}H_1P_{12} - P_{12}H_1{}^TW_1{}^{-1}H_1P_{11} + P_{12}H_1{}^TW_1{}^{-1}H_1P_{12}$$

$$+ P_{11}H_2{}^TW_2{}^{-1}H_2P_{11} - P_{11}H_2{}^TW_2{}^{-1}H_2P_{12} - P_{12}H_2{}^TW_2{}^{-1}H_2P_{11}$$

$$+ P_{12}H_2{}^TW_2{}^{-1}H_2P_{12} \quad ; \quad P_{12}(t_o) = 0 \tag{7.49}$$

The above results can be written in terms of solutions to matrix Riccati equations. It was shown in Chapter 3 (Equation (3.87)) that $T_1{}^*D\Phi\hat{x}_1$ can be written as $G_1{}^T(t)S(t)\hat{x}_1(t)$, where $S(t)$ satisfies Equation (3.98). In Chapter 5 we found (Equations (5.55) through (5.61)) that $T_1{}^*D_1\Phi\tilde{x}$ could be written as $G_1{}^T(t)N_1(t)\tilde{x}(t)$. By a completely parallel argument we can show that we can write $T_2{}^*D_2\Phi\hat{x}$ as $G_2{}^T(t)N_2(t)\hat{x}(t)$, where $N_2(t)$ satisfies

$$\dot{N}_2(t) = -N_2F(t) - F^T(t)N_2 + N_2G_2(t)G_2(t)N_2 \quad ; \quad N_2(T) = I \qquad (7.50)$$

If we further define

$$R_2(t) \triangleq \Phi^T(T,t)D_2(t)R_{21}(t)D(t)\Phi(T,t) \qquad (7.51)$$

where

$$R_{21}(t) \triangleq T_1T_1^* \qquad (7.52)$$

then on taking the derivative of $R_2(t)$ with respect to t we obtain by using

$$\begin{aligned} \dot{\Phi}^T(T,t) &= -F^T(t)\,\Phi^T(T,t) \\ \dot{\Phi}(T,t) &= -\,\Phi(T,t)F(t) \end{aligned} \qquad (7.53)$$

$$\begin{aligned} \dot{R}_2 = &- F^T\Phi^TD_2R_{21}D\Phi - \Phi^TD_2\Phi G_2G_2^T\Phi^TD_2R_{21}D\Phi \\ &- \Phi^TD_2\Phi G_1G_1^T\Phi^TD\Phi + \Phi D_2R_{21}D\Phi G_1G_1^T\Phi^TD\Phi \\ &- \Phi^TD_2R_{21}D\Phi G_2G_2^T\Phi^TD\Phi - \Phi^TD_2R_{21}D\Phi F \end{aligned} \qquad (7.54)$$

Substituting Equation (7.51) and the defining equations for S(t) and $N_2(t)$, i.e.,

$$\begin{aligned} S(t) &\triangleq \Phi^T(T,t)D(t)\Phi(T,t) \\ N_2(t) &\triangleq \Phi^T(T,t)D_2(t)\Phi(T,t) \end{aligned} \qquad (7.55)$$

the resulting equation is

$$\dot{R}_2 = - R_2F(t) - F^T(t)R_2 - N_2G_2(t)G_2^T(t)R_2 - N_2G_1(t)G_1^T(t)S$$

$$+ R_2 \left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right] S \quad ; \quad R_2(T) = 0 \qquad (7.56)$$

From Equations (7.36), (7.47), (7.48) and (7.49) the optimal delayed commitment strategy for player 1 is then given by the following set of equations

$$u_1^*(t) = G_1^T(t)S(t)\hat{x}_1(t) \qquad (7.57)$$

$$(7.58)$$

$$\dot{S} = - SF(t) - F^T(t)S + S\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right] S \quad ; \quad S(T) = I$$

$$(7.59)$$

$$\dot{\hat{x}}_1(t) = \left[F(t) - G_1(t)G_1^T(t)S + G_2(t)G_2^T(t)N_2(t) - G_2(t)G_2^T(t)R_2(t)\right]\hat{x}_1(t)$$

$$+ P_{11}H_1^T(t)W_1^{-1}(t)\left[z_1(t) - H_1(t)\hat{x}_1(t)\right] \quad ; \quad \hat{x}_1(t_o) = \bar{x}_o$$

$$\dot{P}_{11} = F(t)P_{11} + P_{11}F^T(t) + G_2(t)G_2^TN_2(t)P_{12} + P_{12}N_2(t)G_2(t)G_2^T(t)$$

$$- P_{11}H_1^T(t)W_1^{-1}(t)H_1(t)P_{11} \quad ; \quad P_{11}(t_o) = P_o \qquad (7.60)$$

$$\dot{P}_{12} = F(t)P_{12} + P_{12}F^T(t) + G_2(t)G_2^T(t)N_2(t)P_{12} + P_{12}N_2(t)G_2(t)G_2^T(t)$$

$$- P_{11}H_1^TW_1^{-1}(t)H_1(t)P_{12} - P_{12}H_1^T(t)W_1^{-1}(t)H_1(t)P_{11}$$

$$+ P_{12}H_1^T(t)W_1^{-1}(t)H_1(t)P_{12} + P_{11}H_2^T(t)W_2^{-1}(t)H_2(t)P_{11}$$

$$- P_{11}H_2^T(t)W_2^{-1}(t)H_2(t)P_{12} - P_{12}H_2^T(t)W_2^{-1}(t)H_2(t)P_{11} \qquad \text{(Cont'd)}$$

$$+ P_{12}H_2^T(t)W_2^{-1}(t)H_2(t)P_{12} \quad ; \quad P_{12}(t_o) = 0 \tag{7.61}$$

$$\dot{N}_2 = - N_2F(t) - F^T(t)N_2 + N_2G_2(t)G_2^T(t)N_2 \quad ; \quad N_2(T) = I \tag{7.62}$$

$$\dot{R}_2 = - R_2F(t) - F^T(t)R_2 - N_2G_2(t)G_2^T(t)R_2 - N_2G_1(t)G_1^T(t)S$$
$$+ R_2 \left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right] S \quad ; \quad R_2(T) = 0 \tag{7.63}$$

Note that the above matrix Riccati type equations do not present a two point boundary value problem but can all be solved using either forward - or backward integration. This solution can take place "on-line" with a digital computer during the actual game.

7.2 DELAYED COMMITMENT SOLUTION FOR PLAYER 2

If we now consider the game from the point of view of the maximizing player, player 2, his performance criterion during the game at time t is

$$J_2(u_1,u_2) = \frac{1}{2} E\left\{ x^T(T)x(T) + \int_t^T \left[u_1^T(\tau)u_1(\tau) - u_2^T(t)u_2(\tau)\right] d\tau \mid Z_2(t)\right\} \tag{7.64}$$

and his secure strategy can be determined by finding the saddle-point solution to this equation subject to

$$\dot{x} = F(t)x(t) - G_1(t)u_1(t) + G_2(t)u_2(t) \quad ; \quad x(t_o) = \bar{x}_o \tag{7.65}$$

To determine player 2's secure strategy solution we assume that the allowable strategy for player 1, in addition to being $Z_1(t)$ measurable, is also $Z_2(t)$ measurable and we seek that $u_2^* \epsilon U_2$ and $u_1^* \epsilon U_1 \times U_2$ which are optimal in the sense that for all t t_o,T

$$J_2(u_1^*, u_2) \leq J_2(u_1^*, u_2^*) \leq J_2(u_1, u_2^*) \tag{7.66}$$

By a completely parallel argument as used for the solution of the game from player 1's point of view, Equation (7.23), and replacing max. by min. and player 1 by player 2, we obtain as the candidate extremal control for player 1

$$u_1 = T_1^* D_1(\Phi \hat{x} + T_2 u_2), \tag{7.67}$$

with the Kalman-Bucy filter given by Equations (7.21) and (7.22).

Substituting Equation (7.67) into Equation (7.64) we obtain as payoff functional for player 2 at arbitrary time $t = t_o$.

$$\begin{aligned} J_2(u_2) = \frac{1}{2} E \Big\{ &< \Phi \hat{x}_o - T_1 T_1^* D_1(\Phi \hat{x}_o + T_2 u_2) + T_2 u_2, \Phi \hat{x}_o \\ &- T_1 T_1^* D_1(\Phi \hat{x}_o + T_2 u_2) + T_2 u_2 > \\ &+ < T_1 T_1^* D_1(\Phi \hat{x}_o + T_2 u_2), D_1 \Phi \hat{x}_o + T_2 u_2 > \\ &+ < u_2, u_2 > | \; Z_2(t) \Big\} \end{aligned} \tag{7.68}$$

where

$\hat{x}_2(t) = E\left| x(t) \mid Z_2(t) \right|$ and again recalling that the double expectation first over a "finer" and then over a "coarser" set (i.e., containing fewer sets or less information) is the same as the expectation over the coarser set, we can write

$$\begin{aligned} J_2(u_2) = \frac{1}{2}\Big| &< \Phi\hat{x}_{2_o} - T_1T_1^*D_1(\Phi\hat{x}_{2_o} + T_2u_2) + T_2u_2, \\ &\Phi\hat{x}_{2_o} - T_1T_1^*D_1(\Phi\hat{x}_{2_o} + T_2u_2) + T_2u_2 > \\ &+ < T_1^*D_1(\Phi\hat{x}_{2_o} + T_2u_2), T_1^*D_1(\Phi\hat{x}_{2_o} + T_2u_2)> \\ &- < u_2, u_2 > \Big| + \frac{1}{2}\,\mathrm{tr}\,(\Phi^T\Phi P_2) \end{aligned} \tag{7.69}$$

From the results of Chapter 3 we then obtain

$$\begin{aligned} u_2^* &= T_2^*\left[I + T_1T_1^* - T_2T_2^* \right]^{-1} \Phi\,\hat{x}_{2_o} \\ &= T_2^*D\Phi\hat{x}_{2_o} \end{aligned} \tag{7.70}$$

The dynamic system from player 2's point of view can now be written from Equations (7.7) and (7.21) as (7.71)

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} \quad \left[\begin{array}{c|c} F & -G_1T_1^*D_1 \\ \hline PH_1^TW_1^{-1}H_1 & F - G_1T_1^*D_1 \quad -PH_1^TW_1^{-1}H_1 - PH_2^TW_2^{-1}H_2 \end{array}\right] \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

$$+ \begin{bmatrix} G_2 - G_1T_1^*D_1T_2 \\ \hline G_2 - G_1T_1^*D_1T_2 \end{bmatrix} u_2 + \begin{bmatrix} 0 \\ \hline PH_2^TW_2^{-1} \end{bmatrix} z_2 + \begin{bmatrix} 0 \\ \hline PH_1^TW_1^{-1} \end{bmatrix} w_1$$

with initial conditions

$$x(t_o) = \bar{x}_o$$
$$\hat{x}(t_o) = \bar{x}_o \tag{7.72}$$

and with measurements

$$z_2(t) = H_2(t)x(t) + v_2(t) \tag{7.73}$$

The linear-Gaussian assumptions assure us that the optimal estimates of $x(t)$ and $\hat{x}(t)$ given $Z_2(t)$, which are denoted by $\hat{x}_2(t)$ and $\hat{x}_{21}(t)$ respectively, are available from

$$\begin{bmatrix} \dot{\hat{x}}_2 \\ \dot{\hat{x}}_{21} \end{bmatrix} = \left[\begin{array}{c|c} F & -G_1T_1{}^*D_1\Phi \\ \hline PH_1{}^TW_1{}^{-1}H_1 & F - G_1T_1{}^*D_1\Phi - PH_1{}^TW_1{}^{-1}H_1 - PH_2{}^TW_2{}^{-1}H_2 \end{array}\right] \begin{bmatrix} \hat{x}_2 \\ \hat{x}_{21} \end{bmatrix}$$
$$+ \begin{bmatrix} G_2 - G_1T_1{}^*D_1T_2 \\ G_2 - G_1T_1{}^*D_1T_2 \end{bmatrix} u_2 + \begin{bmatrix} 0 \\ PH_2{}^TW_2{}^{-1} \end{bmatrix} z_2$$
$$+ \begin{bmatrix} P_{11} & P_{21} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} H_2{}^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W_2{}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_2 - H_2\hat{x}_2 \\ 0 \end{bmatrix} \tag{7.74}$$

with initial conditions

$$\hat{x}_2(t_o) = \bar{x}_o$$
$$\hat{x}_{21}(t_o) = \bar{x}_o \tag{7.75}$$

with the corresponding error covariance matrix

$$\dot{P}_2 = \begin{bmatrix} \dot{P}_{11} & \dot{P}_{21} \\ \dot{P}_{21} & \dot{P}_{22} \end{bmatrix} = \left[\begin{array}{c|c} F & - G_1 T_1{}^* D_1 \Phi \\ \hline PH_1{}^T W^{-1} H_1 & F - G_1 T_1{}^* D_1 \Phi - PH_1{}^T W_1{}^{-1} H_1 - PH_2{}^T W_2{}^{-1} H_2 \end{array}\right]$$

$$\begin{bmatrix} P_{11} & P_{21} \\ P_{21} & P_{22} \end{bmatrix} + \begin{bmatrix} P_{11} & P_{21} \\ P_{21} & P_{22} \end{bmatrix} \left[\begin{array}{c|c} F & - G_1 T_1{}^* D_1 \Phi \\ \hline PH_1{}^T W_1{}^{-1} H_1 & F - G_1 T_1{}^* D_1 \Phi - PH_1{}^T W_1{}^{-1} H_1 - PH_2{}^T W_2{}^{-1} H_2 \end{array}\right]^T + \begin{bmatrix} 0 & 0 \\ 0 & PH_1{}^T W_1{}^{-1} H_1 P \end{bmatrix}$$

$$- \begin{bmatrix} P_{11} & P_{21} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} H_2{}^T W_2{}^{-1} H_2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_{11} & P_{21} \\ P_{21} & P_{22} \end{bmatrix} \tag{7.76}$$

with

$$P_2(t_o) = \begin{bmatrix} P_o & 0 \\ 0 & 0 \end{bmatrix} \tag{7.77}$$

The optimal control for player 2 is thus

$$u_2{}^* = T_2{}^* D \Phi \hat{x}_2 \tag{7.78}$$

while the corresponding optimal response of player 1 is

$$u_1^* = T_1^* D_1 \hat{x} + T_1^* D_1 T_2 T_2^* D \Phi \hat{x}_2 \tag{7.79}$$

The Kalman-Bucy filtering equations (7.74) and (7.76) can again be simplified by rewriting the conditional estimates, Equation (7.74) as

$$\dot{\hat{x}}_2 = F\hat{x}_2 - G_1 T_1^* D_1 \Phi \hat{x}_{21} + (G_2 - G_1 T_1 D_1 T_2) u_2$$

$$+ P_{11} H_2^T W_2^{-1} (z_2 - H_2 \hat{x}_2) \quad ; \quad x_2(t_o) = \bar{x}_o \tag{7.80}$$

and

$$\dot{\hat{x}}_{21} = F\hat{x}_{21} - G_1 T_1^* D_1 \Phi \hat{x}_{21} + (G_2 - G_1 T_1 D_1 T_2) u_2$$

$$+ P_{21} H_2^T W_2^{-1} (z_2 - H_2 \hat{x}_2) + P H_2^T W_2^{-1} (z_2 - H_2 \hat{x}_2)$$

$$+ P H_1^T W_1^{-1} H_1 \hat{x}_2 - P H_1^T W_1^{-1} H_1 \hat{x}_{21} \quad ; \quad \hat{x}_{21}(t_o) = \bar{x}_o \tag{7.81}$$

then since

$$\hat{x}_2(t) \triangleq E\left| x \mid Z_2(t) \right| = \hat{x}_{21} \triangleq E\left| E\left| x \mid Z_1(t) Z_2(t) \right| \mid Z_2(t) \right| \tag{7.82}$$

we find that

$$P_2(t) = P_{11}(t) = P(t) + P_{21}(t) \tag{7.83}$$

thus

$$\dot{P}_{11}(t) = \dot{P}(t) + \dot{P}_{21}(t) \tag{7.84}$$

Using Equation (7.84) we obtain from Equation (7.76)

$$
\begin{aligned}
\dot{P}_{21} = {} & FP_{21} - G_1T_1^{*}D_1\Phi P_{22} + P_{11}H_1^{T}W_1^{-1}H_1P + P_{21}F^{T} \\
& - P_{21}\left[G_1T_1^{*}D_1\Phi\right]^{T} - P_{21}H_1^{T}W_1^{-1}H_1P - P_{21}H_2^{T}W_2^{-1}H_2P \\
& - PH_2^{T}W_2^{-1}H_2P_{21} - P_{21}H_2^{T}W_2^{-1}H_2P_{21} \quad ; \quad P_{21}(t_o) = 0
\end{aligned}
\tag{7.85}
$$

$$
\begin{aligned}
\dot{P}_{22} = {} & PH_1^{T}W_1^{-1}H_1P_{21} + FP_{22} - G_1T_1^{*}D_1\Phi P_{22} - PH_1^{T}W_1^{-1}H_1P_{22} \\
& - PH_2^{T}W_2^{-1}H_2P_{22} - P_{21}H_1^{T}W_1^{-1}H_1P + P_{22}F^{T} \\
& - P_{22}\left[G_1T_1^{*}D_1\Phi\right]^{T} - P_{22}H_1^{T}W_1^{-1}H_1P - P_{22}H_2^{T}W_2^{-1}H_2P \\
& + PH_1^{T}W_1^{-1}H_1P - P_{21}H_2^{T}W_2^{-1}H_2P_{21} \quad ; \quad P_{22}(t_o) = 0
\end{aligned}
\tag{7.86}
$$

which simplifies to

$$
\begin{aligned}
\dot{P}_{22} = {} & FP_{22} - G_1T_1^{*}D_1\Phi P_{22} + P_{11}H_1^{T}W_1^{-1}H_1P + P_{22}F^{T} \\
& - P_{22}\left[G_1T_1^{*}D_1\Phi\right]^{T} - P_{22}H_1^{T}W_1^{-1}H_1P - P_{22}H_2^{T}W_2^{-1}H_2P \\
& + PH_1^{T}W_2^{-1}H_1P_{21} - P_{21}H_2^{T}W_2^{-1}H_2P_{21} \quad ; \quad P_{22}(t_o) = 0
\end{aligned}
\tag{7.87}
$$

Then comparing Equation (7.85) with Equation (7.87) we observe that

$$P_{21}(t) = P_{22}(t) \tag{7.88}$$

and we can write for Equation (7.80) and (7.76)

$$\dot{\hat{x}}_2 = (F - G_1T_1{}^*D_1\Phi)\hat{x}_2 + (G_2 - G_1T_1{}^*D_1T_2)u_2$$

$$+ P_{11}H_2{}^TW_2{}^{-1}(z_2 - H_2\hat{x}_2) \quad ; \quad \hat{x}_2(t_o) = \bar{x}_o \tag{7.89}$$

and

$$\dot{P}_{11} = FP_{11} + P_{11}F^T - G_1T_1{}^*D_1\Phi P_{21} - P_{21}(G_1T_1{}^*D_1\Phi)^T$$

$$- P_{11}H_2{}^TW_2{}^{-1}H_2P_{11} \quad ; \quad P_{11}(t_o) = P_o \tag{7.90}$$

$$\dot{P}_{21} = FP_{21} + P_{21}F^T - G_1T_1{}^*D_1\Phi P_{21} - P_{21}(G_1T_1{}^*D_1\Phi)^T$$

$$+ P_{11}H_1{}^TW_1{}^{-1}H_1P_{11} - P_{11}H_1{}^TW_1{}^{-1}H_1P_{21} - P_{21}H_1{}^TW_1{}^{-1}H_1P_{11}$$

$$+ P_{21}H_1{}^TW_1{}^{-1}H_1P_{21} - P_{11}H_2{}^TW_2{}^{-1}H_2P_{21} - P_{21}H_2{}^TW_2{}^{-1}H_2P_{11}$$

$$+ P_{21}H_2{}^TW_2{}^{-1}H_2P_{21} \quad ; \quad P_{21}(t_o) = 0 \tag{7.91}$$

If we define

$$R_1(t) = \Phi^T(T,t)D_1(t)R_{12}(t)D(t)\Phi(T,t) \tag{7.92}$$

where

$$R_{12}(t) = T_2T_2{}^*$$

then using

$$\dot{\Phi}^T(T,t) = -F^T(t)\Phi^T(T,t)$$

$$\dot{\Phi}^T(T,t) \quad \cdot \Phi^T(T,t)F(t) \tag{7.93}$$

we obtain after taking the derivative of $R_1(t)$ with respect to t

$$\dot{R}_1 = - F^T \Phi^T D_1 R_{12} D\Phi + \Phi^T D_1 \Phi\, G_1 G_1^T \Phi^T D_1 R_{12} D\Phi$$

$$- \Phi^T D_1 \Phi\, G_2 G_2^T \Phi^T D\Phi + \Phi^T D_1 R_{12} D\Phi\, G_1 G_1^T \Phi^T D\Phi$$

$$- \Phi^T D_1 R_{12} D\Phi\, G_2 G_2^T \Phi^T D\Phi - \Phi^T D_1 R_{12} D\Phi F \qquad (7.94)$$

Substituting Equation (7.92) and the defining Equations for $S(t)$ and $N_1(t)$, i.e.,

$$\begin{aligned} S(t) &\triangleq \Phi^T(T,t) D(t) \Phi(T,t) \\ N_1(t) &\triangleq \Phi^T(T,t) D_1(t) \Phi(T,t) \end{aligned} \qquad (7.95)$$

we obtain

$$\dot{R}_1 = - R_1 F(t) - F^T(t) R_1 + N_1 G_1(t) G_1^T(t) R_1 - N_1 G_2(t) G_2^T(t) S$$

$$+ R_1 \left[G_1(t) G_1^T(t) - G_2(t) G_2^T(t) \right] S \quad ; \quad R_1(T) = 0 \qquad (7.96)$$

Using Equations (7.95) and (7.96) in Equations (7.78), (7.86), (7.87) and (7.89) the optimal delayed commitment strategy for player 2 is then given by the following set of equations.

$$u_2^*(t) = G_2^T(t) S(t) \hat{x}_2(t) \qquad (7.97)$$

$$\dot{S} = - SF(t) - F^T(t) S + S \left[G_1(t) G_1^T(t) - G_2(t) G_2^T(t) \right] S \; ; \; S(T) = I \qquad (7.98)$$

$$\dot{\hat{x}}_2(t) = \left[F(t) - G_1(t) G_1^T(t) N_1(t) + G_2(t) G_2^T(t) S(t) - G_1(t) G_1^T(t) R_1(t) \right] \hat{x}_2(t)$$

$$+ P_{11} H_2^T(t) W_2^{-1}(t) \left[z_2(t) - H_2(t) \hat{x}_2(t) \right] \; ; \; \hat{x}_2(t_o) = \bar{x}_o \qquad (7.99)$$

$$\dot{P}_{11} = F(t)P_{11} + P_{11}F^T(t) - G_1(t)G_1^T(t)N_1(t)P_{21} - P_{21}N_1(t)G_1(t)G_1^T(t)$$
$$- P_{11}H_2^T(t)W_2^{-1}(t)H_2(t)P_{11} \quad ; \quad P_{11}(t_o) = P_o \tag{7.100}$$

$$\dot{P}_{21} = F(t)P_{21} + P_{21}F^T(t) - G_1(t)G_1^T(t)N_1(t)P_{21} - P_{21}N_1(t)G_1(t)G_1^T(t)$$
$$+ P_{11}H_1^T(t)W_1^{-1}(t)H_1(t)P_{11} - P_{11}H_1^T(t)W_1^{-1}(t)H_1(t)P_{21}$$
$$- P_{21}H_1^T(t)W_1^{-1}(t)H_1(t)P_{11} + P_{21}H_1^T(t)W_1^{-1}(t)H_1(t)P_{21}$$
$$- P_{11}H_2^T(t)W_2^{-1}(t)H_2(t)P_{21} - P_{21}H_2^T(t)W_2^{-1}(t)H_2(t)P_{11}$$
$$+ P_{21}H_2^T(t)W_2^{-1}(t)H_2(t)P_{21} \quad ; \quad P_{21}(t_o) = 0 \tag{7.101}$$

$$\dot{N}_1 = - N_1F(t) - F^T(t)N_1 + N_1G_1(t)G_1^T(t)N_1 \quad ; \quad N_1(T) = I \tag{7.102}$$

$$\dot{R}_1 = - R_1F(t) - F^T(t)R_1 + N_1G_1(t)G_1^T(t)R_1 - N_1G_2(t)G_2^T(t)S$$
$$+ R_1\left[G_1(t)G_1^T(t) - G_2(t)G_2^T(t)\right]S \quad ; \quad R_1(T) = 0 \tag{7.103}$$

The above solutions for player 2 are very similar to those obtained for player 1 and are "simple" in that they can be directly solved using forward and backward integration with a digital computer.

Recalling that Willman [8] showed that for the class of games discussed in this chapter,the strategies could only be realized with infinite dimensional dynamic systems, we observe that the point of view of delayed commitment strategies leads to solutions which are readily computable.

CHAPTER 8

SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE WORK

In this dissertation the problem of prior and delayed commitment strategies to differential games with noise corrupted state measurements is discussed. It is pointed out that the prior commitment solution, which has led previous researchers to define the closure problem, is valid only under restricted circumstances.

The delayed commitment solutions are then obtained for a differential game where one player has perfect state information and the other player has only noise corrupted measurements of the state and is extended to a differential game where both players have noise corrupted measurements in Chapter 7. In both cases, the resulting secure strategies do again satisfy the familiar Separation Theorem of stochastic optimal control.

Of particular significance is the fact that the governing equations do not result in an often difficult to solve non-linear two-point boundary value problem, but are readily computable with a digital computer.

A detailed example of a pursuit-evasion game is presented in Chapter 6. It discusses a missile and an airplane system where the missile (or player 1) has perfect state measurements and the airplane (or player 2) has noise corrupted measurements. Both the prior commitment and delayed commitment solutions have been obtained and the results compared.

An immediate and direct extension of the research presented in this dissertation is to extend the results to differential games, where in addition to noise corrupting the measurements, additive white Gaussian noise, independent of the measurement noise and of the initial estimate of the state, is present in the system dynamics. Of course, if the noises are not white but Markov with rational spectra, they can be modelled as outputs of a dynamic system which is driven by white noise and by adjoining this dynamic model to the system equations an augmented system is obtained with white noise disturbances.

From the game theoretic point of view the realization that the zero-sum assumption has to be abandoned during the actual stochastic game offers several interesting analytic and conceptual concepts not found in zero-sum differential games. We have used the minimax solution concept, however, non-inferior (or Pareto optimal) strategies or solution concepts involving coalitions, bargaining, etc., can be envisioned.

REFERENCES

1. von Neumann, John, and O. Morgenstern. Theory of Games and Economic Behavior, Princeton University Press, Princeton, New Jersey, 1943.

2. Isaacs, R. P. "Differential Games - I: Introduction," RAND Corporation, Research Memorandum, RM-1391, November 1954.

3. Isaacs, R. P. "Differential Games - II: The Definition and Formulation," RAND Corporation, Research Memorandum, RM-1399, November 1954.

4. Isaacs, R. P. "Differential Games - III: The Basic Principles of the Solution Process," RAND Corporation, Research Memorandum, RM-1411, December 1954.

5. Isaacs, R. P. "Differential Games - IV: Mainly Examples," RAND Corporation, Research Memorandum, RM 1468, March 1955.

6. Isaacs, Rufus P. Differential Games, J. Wiley and Sons, Inc., New York, 1965.

7. Aumann, R. J. and M. Maschler. "Some Thoughts on the Minimax Principle," Management Science, Vol. 18, No. 5, Part 2, pp. 54-63, January 1972.

8. Willman, W. W. "Formal Solutions for a Class of Stochastic Pursuit-Evasion Games," IEEE Trans. on Automatic Control, Vol. AC-14, No. 5, pp. 504-509, October 1969.

9. Behn, R. D., and Y. C. Ho. "On a Class of Linear Stochastic Differential Games," IEEE Trans. on Automatic Control, Vol. AC-13, No. 3, pp. 227-240, June 1968.

10. Rhodes, I. B., and D. G. Luenberger. "Differential Games with Imperfect State Information," IEEE Trans. on Automatic Control, Vol. AC-14, No. 1, pp. 29-38, February 1969.

11. Kuhn, H. W. "Extensive Games," Proc. Nat. Acad. Sci., Vol. 36, pp. 570-576, October 1950.

12. Luce, Robert D. and H. Raiffa. Games and Decisions, J. Wiley and Sons, Inc., New York, 1957.

13. Starr, A. W. "Nonzero-Sum Differential Games: Concepts and Models," Harvard University, Division of Engineering and Applied Physics, Technical Report No. 590, June 1969.

14. Ho, Y. C., A. E. Bryson, and S. Baron. "Differential Games and Optimal Pursuit-Evasion Strategies," IEEE Trans. on Automatic Control, Vol. AC-10, pp. 385-389, October 1965.

15. Berkovitz, L. D. "Variational Approach to Differential Games," Advances in Game Theory. (Annals of Mathematics Studies, 52), Princeton University Press, Princeton, New Jersey, 1964, pp. 127-174.

16. Porter, W. A. "On Function Space Pursuit-Evasion Games," SIAM J. Control, Vol. 5, No. 4, pp. 555-574, April 1967.

17. Rhodes, I. B., and D. G. Luenberger. "Stochastic Differential Games with Constrained State Estimators," IEEE Trans. on Automatic Control, Vol. AC-14, No. 5, pp. 476-481, October 1969.

18. Harsanyi, J. C. "Games with Incomplete Information Played by Bayesian Players - Part I, The Basic Model," Management Science, Vol. 14, No. 3, pp. 159-182, November 1967.

19. Harsanyi, J. C. "Games with Incomplete Information Played by Bayesian Players - Part II, Bayesian Equilibrium Points," Management Science, Vol. 14, No. 5, pp. 320-334, January 1968.

20. Harsanyi, J. C. "Games with Incomplete Information Played by Bayesian Players - Part III, The Basic Probability Distribution of the Game," Management Science, Vol. 14, No. 7, pp. 486-502,

21. Ho, Y. C. "On the Minimax Principle and Zero Sum Stochastic Differential Games," Proceedings of the 1972 IEEE Conf. on Decision and Control and 11th Symposium on Adaptive Processes, New Orleans, Louisiana, December 1972, IEEE, New York, 1972, pp. 333-339.

22. Peterson, Edwin L. Statistical Analysis and Optimization of Systems, J. Wiley and Sons, Inc., New York, 1961.

23. Bryson, Jr., Arthur E. and Y. C. Ho. Applied Optimal Control, Ginn and Co., Waltham, Massachusetts, 1969.

24. Kushner, Harold J. Introduction to Stochastic Control, Holt, Rinehart and Winston, Inc., New York, 1971.

APPENDIX A

-COMPUTER PROGRAM LISTING FOR THE NUMERICAL EXAMPLE OF SECTION 6.4

- QUASILINIARIZATION ITERATION 1

```
      SUBROUTINE FUNEV
      COMMON TIME,DELT,NSTART,NFIRST,NEXIT,IPASS,ROMCON(2094)
      REAL K1,K2,KT1,KT2,N,ND
      REAL N0,NP1D,NP1,NH1D,NH1
      DATA TF,AS,KT1,KT2,TAU1,TAU2,R11S,R22S/10.,.04,32.2,32.2,1.,2.,.00
     11,.004/
      IF(NSTART)30,50,10
   10 READ(5,20)NP1,PP1,W2,NH1,PH1
   20 FORMAT(4E20.0)
      CALL INTG(NP1D,NP1)
      CALL INTG(PP1D,PP1)
      CALL INTG(NH1D,NH1)
      CALL INTG(PH1D,PH1)
      CALL PRINT(10H      S(T),10H,G12.4     ,S,1,0.)
      CALL PRINT(10H      NP1D,10H,G12.4     ,NP1D,3,0.)
      CALL PRINT(10H      NP1 ,10H,G12.4     ,NP1,1,0.)
      CALL PRINT(10H      PP1D,10H,G12.4     ,PP1D,3,0.)
      CALL PRINT(10H      PP1 ,10H,G12.4     ,PP1,1,0.)
      CALL PRINT(10H      NH1D,10H,G12.4     ,NH1D,3,0.)
      CALL PRINT(10H      NH1 ,10H,G12.4     ,NH1,1,0.)
      CALL PRINT(10H      PH1D,10H,G12.4     ,PH1D,3,0.)
      CALL PRINT(10H      PH1 ,10H,G12.4     ,PH1,1,0.)
   30 CALL PRINT(10H     K1(T),10H,G12.4     ,K1,5,0.)
      CALL PRINT(10H     K2(T),10H,G12.4     ,K2,5,0.)
      RETURN
   50 TGO=TF-TIME
      T1=1.-EXP(-TGO/TAU1)-TGO/TAU1
      T2=1.-EXP(-TGO/TAU2)-TGO/TAU2
      S=6.*R11S*R22S/(6.*R11S*R22S+AS*KT1**2*R22S*(6.*TAU1**2*TGO-6.*TAU
     11*TGO**2+2.*TGO**3+3.*TAU1**3*(1.-EXP(-2.*TGO/TAU1))-12.*TAU1**2*T
     2GO*EXP(-TGO/TAU1))-AS*KT2**2*R11S*(6.*TAU2**2*TGO-6.*TAU2*TGO**2+2
     3.*TGO**3+3.*TAU2**3*(1.-EXP(-2.*TGO/TAU2))-12.*TAU2**2*TGO*EXP(-TG
     40/TAU2)))
      K1=AS*KT1**2*TAU1**2*T1**2/R11S
      K2=AS*KT2**2*TAU2**2*T2**2/R22S
      P0=1000.*EXP(-.5*TIME)
      N0=-.000000015
      NP1D=2.*(K1*N0+K1*S+P0/W2)*NP1+2.*N0*PP1/W2-K1*N0*N0-2.*N0*P0/W2+K
     12*S*S
      PP1D=-2.*K1*P0*NP1-2.*(K1*N0+K1*S+P0/W2)*PP1+P0*P0/W2+2.*K1*N0*P0
      NH1D=2.*(K1*N0+K1*S+P0/W2)*NH1+2.*N0*PH1/W2
      PH1D=-2.*K1*P0*NH1-2.*(K1*N0+K1*S+P0/W2)*PH1
      IF(NFIRST)60,80,60
```

```
60 A1=-NP1/NH1
   WRITE(6,70)NP1,PP1,W2,A1
70 FORMAT(1H0,4X,4HN   = G14.6,5X,4HP   =  G14.6,5X,5HW2 = G14.6,5HA1
  1=G22.14)
80 RETURN
   END
```

- ITERATION 2

```
   SUBROUTINE FUNEV
   COMMON TIME,DELT,NSTART,NFIRST,NEXIT,IPASS,ROMCON(2094)
   REAL K1,K2,KT1,KT2,N,ND
   REAL NO,NP1,NP1D,NH1,NH1D,NP2,NP2D,NH2,NH2D
   DATA TF,AS,KT1,KT2,TAU1,TAU2,R11S,R22S/10.,.04,32.2,32.2,1.,2.,.00
  11,.004/
   IF(NSTART)30,50,10
10 READ(5,20)NP1,PP1,W2,NP2,PP2,NH2,PH2
20 FORMAT(4E20.0)
   CALL INTG(NP1D,NP1)
   CALL INTG(PP1D,PP1)
   CALL INTG(NP2D,NP2)
   CALL INTG(PP2D,PP2)
   CALL INTG(NH2D,NH2)
   CALL INTG(PH2D,PH2)
   CALL PRINT(10H      S(T),10H,G12.4    ,S,1,0.)
   CALL PRINT(10H      NP1D,10H,G12.4    ,NP1D,3,0.)
   CALL PRINT(10H      NP1 ,10H,G12.4    ,NP1,1,0.)
   CALL PRINT(10H      PP1D,10H,G12.4    ,PP1D,3,0.)
   CALL PRINT(10H      PP1 ,10H,G12.4    ,PP1,1,0.)
   CALL PRINT(10H      NP2D,10H,G12.4    ,NP2D,3,0.)
   CALL PRINT(10H      NP2 ,10H,G12.4    ,NP2,1,0.)
   CALL PRINT(10H      PP2D,10H,G12.4    ,PP2D,3,0.)
   CALL PRINT(10H      PP2 ,10H,G12.4    ,PP2,1,0.)
   CALL PRINT(10H      NH2D,10H,G12.4    ,NH2D,3,0.)
   CALL PRINT(10H      NH2 ,10H,G12.4    ,NH2,1,0.)
   CALL PRINT(10H      PH2D,10H,G12.4    ,PH2D,3,0.)
   CALL PRINT(10H      PH2 ,10H,G12.4    ,PH2,1,0.)
30 CALL PRINT(10H     K1(T),10H,G12.4    ,K1,5,0.)
   CALL PRINT(10H     K2(T),10H,G12.4    ,K2,5,0.)
   RETURN
50 TGO=TF-TIME
   T1=1.-EXP(-TGO/TAU1)-TGO/TAU1
   T2=1.-EXP(-TGO/TAU2)-TGO/TAU2
   S=6.*R11S*R22S/(6.*R11S*R22S+AS*KT1**2*R22S*(6.*TAU1**2*TGO-6.*TAU
  11*TGO**2+2.*TGO**3+3.*TAU1**3*(1.-EXP(-2.*TGO/TAU1))-12.*TAU1**2*T
  2GO*EXP(-TGO/TAU1))-AS*KT2**2*R11S*(6.*TAU2**2*TGO-6.*TAU2*TGO**2+2
  3.*TGO**3+3.*TAU2**3*(1.-EXP(-2.*TGO/TAU2))-12.*TAU2**2*TGO*EXP(-TG
  4O/TAU2)))
   K1=AS*KT1**2*TAU1**2*T1**2/R11S
   K2=AS*KT2**2*TAU2**2*T2**2/R22S
   NO=-.000000015
```

```
      P0=1000.*EXP(-.5*TIME)
      NP1D=2.*(K1*N0+K1*S+P0/W2)*NP1+2.*N0*PP1/W2-K1*N0*N0-2.*N0*P0/W2+K
     12*S*S
      PP1D=-2.*K1*P0*NP1-2.*(K1*N0+K1*S+P0/W2)*PP1+P0*P0/W2+2.*K1*N0*P0
      NP2D=2.*(K1*NP1+K1*S+PP1/W2)*NP2+2.*NP1*PP2/W2-K1*NP1*NP1-2.*NP1*P
     1P1/W2+K2*S*S
      PP2D=-2.*K1*PP1*NP2-2.*(K1*NP1+K1*S+PP1/W2)*PP2+PP1*PP1/W2+2.*K1*N
     1P1*PP1
      NH2D=2.*(K1*NP1+K1*S+PP1/W2)*NH2+2.*NP1*PH2/W2
      PH2D=-2.*K1*PP1*NH2-2.*(K1*NP1+K1*S+PP1/W2)*PH2
      IF(NFIRST)60,80,60
   60 A2=-NP2/NH2
      WRITE(6,70)NP1,PP1,W2,A2
   70 FORMAT(1H0,4X,4HN  = G14.6,5X,4HP  =  G14.6,5X,5HW2 = G14.6,5HA2
     1=G22.14)
   80 RETURN
      END


      - ITERATION 3

      SUBROUTINE FUNEV
      COMMON TIME,DELT,NSTART,NFIRST,NEXIT,IPASS,ROMCON(2094)
      REAL K1,K2,KT1,KT2,N,ND
      REAL N0,NP1,NP1D,NH1,NH1D,NP2,NP2D,NH2,NH2D
      REAL NP3,NH3,NP3D,NH3D
      DATA TF,AS,KT1,KT2,TAU1,TAU2,R11S,R22S/10.,.04,32.2,32.2,1.,2.,.00
     11,.004/
      IF(NSTART)30,50,10
   10 READ(5,20)NP1,PP1,W2,NP2,PP2,NP3,PP3,NH3,PH3
   20 FORMAT(4E20.0)
      CALL INTG(NP1D,NP1)
      CALL INTG(PP1D,PP1)
      CALL INTG(NP2D,NP2)
      CALL INTG(PP2D,PP2)
      CALL INTG(NP3D,NP3)
      CALL INTG(PP3D,PP3)
      CALL INTG(NH3D,NH3)
      CALL INTG(PH3D,PH3)
      CALL PRINT(10H      S(T),10H,G12.4    ,S,1,0.)
      CALL PRINT(10H      NP1D,10H,G12.4    ,NP1D,3,0.)
      CALL PRINT(10H      NP1 ,10H,G12.4    ,NP1,1,0.)
      CALL PRINT(10H      PP1D,10H,G12.4    ,PP1D,3,0.)
      CALL PRINT(10H      PP1 ,10H,G12.4    ,PP1,1,0.)
      CALL PRINT(10H      NP2D,10H,G12.4    ,NP2D,3,0.)
      CALL PRINT(10H      NP2 ,10H,G12.4    ,NP2,1,0.)
      CALL PRINT(10H      PP2D,10H,G12.4    ,PP2D,3,0.)
      CALL PRINT(10H      PP2 ,10H,G12.4    ,PP2,1,0.)
      CALL PRINT(10H      NP3D,10H,G12.4    ,NP3D,3,0.)
      CALL PRINT(10H      NP3 ,10H,G12.4    ,NP3,1,0.)
      CALL PRINT(10H      PP3D,10H,G12.4    ,PP3D,3,0.)
      CALL PRINT(10H      PP3 ,10H,G12.4    ,PP3,1,0.)
```

```
      CALL PRIN~(10H      NH3D,10H,G12.4     ,NH3D,3,0.)
      CALL PRINT(10H      NH3 ,10H,G12.4     ,NH3,1,0.)
      CALL PRINT(10H      PH3D,10H,G12.4     ,PH3D,3,0.)
      CALL PRINT(10H      PH3 ,10H,G12.4     ,PH3,1,0.)
   30 CALL PRINT(10H     K1(T),10H,G12.4     ,KT1,5,0.)
      CALL PRINT(10H     K2(T),10H,G12.4     ,KT2,5,0.)
      RETURN
   50 TGO=TF-TIME
      T1=1.-EXP(-TGO/TAU1)-TGO/TAU1
      T2=1.-EXP(-TGO/TAU2)-TGO/TAU2
      S=6.*R11S*R22S/(6.*R11S*R22S+AS*KT1**2*R22S*(6.*TAU1**2*TGO-6.*TAU
     11*TGO**2+2.*TGO**3+3.*TAU1**3*(1.-EXP(-2.*TGO/TAU1))-12.*TAU1**2*T
     2GO*EXP(-TGO/TAU1))-AS*KT2**2*R11S*(6.*TAU2**2*TGO-6.*TAU2*TGO**2+2
     3.*TGO**3+3.*TAU2**3*(1.-EXP(-2.*TGO/TAU2))-12.*TAU2**2*TGO*EXP(-TG
     4O/TAU2)))
      K1=AS*KT1**2*TAU1**2*T1**2/R11S
      K2=AS*KT2**2*TAU2**2*T2**2/R22S
      N0=-.000000015
      P0=1000.*EXP(-.5*TIME)
      NP1D=2.*(K1*N0+K1*S+P0/W2)*NP1+2.*N0*PP1/W2-K1*N0*N0-2.*N0*P0/W2+K
     12*S*S
      PP1D=-2.*K1*P0*NP1-2.*(K1*N0+K1*S+P0/W2)*PP1+P0*P0/W2+2.*K1*N0*P0
      NP2D=2.*(K1*NP1+K1*S+PP1/W2)*NP2+2.*NP1*PP2/W2-K1*NP1*NP1-2.*NP1*P
     1P1/W2+K2*S*S
      PP2D=-2.*K1*PP1*NP2-2.*(K1*NP1+K1*S+PP1/W2)*PP2+PP1*PP1/W2+2.*K1*N
     1P1*PP1
      NP3D=2.*(K1*NP2+K1*S+PP2/W2)*NP3+2.*NP2*PP3/W2-K1*NP2*NP2-2.*NP2*P
     1P2/W2+K2*S*S
      PP3D=-2.*K1*PP2*NP3-2.*(K1*NP2+K1*S+PP2/W2)*PP3+PP2*PP2/W2+2.*K1*N
     1P2*PP2
      NH3D=2.*(K1*NP2+K1*S+PP2/W2)*NH3+2.*NP2*PH3/W2
      PH3D=-2.*K1*PP2*NH3-2.*(K1*NP2+K1*S+PP2/W2)*PH3
      IF(NFIRST)60,80,60
   60 A3=-NP3/NH3
      WRITE(6,70)NP1,NP2,W2,A3
   70 FORMAT(1H0,4X,4HN  = G14.6,5X,4HP  =  G14.6,5X,5HW2 = G14.6,5HA3
     1=G22.14)
   80 RETURN
      END


      - ITERATION 4

      SUBROUTINE FUNEV
      COMMON TIME,DELT,NSTART,NFIRST,NEXIT,IPASS,ROMCON(2094)
      REAL K1,K2,KT1,KT2,N,ND
      REAL N0,NP1,NP1D,NH1,NH1D,NP2,NP2D,NH2,NH2D
      REAL NP3,NH3,NP3D,NH3D
      REAL NP4,NH4,NP4D,NH4D
      DATA TF,AS,KT1,KT2,TAU1,TAU2,R11S,R22S/10.,.04,32.2,32.2,1.,2.,.00
     11,.004/
      IF(NSTART)30,50,10
```

```
10 READ(5,20)NP1,PP1,W2,NP2,PP2,NP3,PP3,NP4,PP4,NH4,PH4
20 FORMAT(4E20.0)
   CALL INTG(NP1D,NP1)
   CALL INTG(PP1D,PP1)
   CALL INTG(NP2D,NP2)
   CALL INTG(PP2D,PP2)
   CALL INTG(NP3D,NP3)
   CALL INTG(PP3D,PP3)
   CALL INTG(NP4D,NP4)
   CALL INTG(PP4D,PP4)
   CALL INTG(PH4D,PH4)
   CALL INTG(NH4D,NH4)
   CALL PRINT(10H       S(T),10H,G12.4     ,S,1,0.)
   CALL PRINT(10H       NP1D,10H,G12.4     ,NP1D,3,0.)
   CALL PRINT(10H       NP1 ,10H,G12.4     ,NP1,1,0.)
   CALL PRINT(10H       PP1D,10H,G12.4     ,PP1D,3,0.)
   CALL PRINT(10H       PP1 ,10H,G12.4     ,PP1,1,0.)
   CALL PRINT(10H       NP2D,10H,G12.4     ,NP2D,3,0.)
   CALL PRINT(10H       NP2 ,10H,G12.4     ,NP2,1,0.)
   CALL PRINT(10H       PP2D,10H,G12.4     ,PP2D,3,0.)
   CALL PRINT(10H       PP2 ,10H,G12.4     ,PP2,1,0.)
   CALL PRINT(10H       NP3D,10H,G12.4     ,NP3D,3,0.)
   CALL PRINT(10H       NP3 ,10H,G12.4     ,NP3,1,0.)
   CALL PRINT(10H       PP3D,10H,G12.4     ,PP3D,3,0.)
   CALL PRINT(10H       PP3 ,10H,G12.4     ,PP3,1,0.)
   CALL PRINT(10H       NP4D,10H,G12.4     ,NP4D,3,0.)
   CALL PRINT(10H       NP4 ,10H,G12.4     ,NP4,1,0.)
   CALL PRINT(10H       PP4D,10H,G12.4     ,PP4D,3,0.)
   CALL PRINT(10H       PP4 ,10H,G12.4     ,PP4,1,0.)
   CALL PRINT(10H       NH4D,10H,G12.4     ,NH4D,3,0.)
   CALL PRINT(10H       NH4 ,10H,G12.4     ,NH4,1,0.)
   CALL PRINT(10H       PH4D,10H,G12.4     ,PH4D,3,0.)
   CALL PRINT(10H       PH4 ,10H,G12.4     ,PH4,1,0.)
30 CALL PRINT(10H      K1(T),10H,G12.4     ,KT1,5,0.)
   CALL PRINT(10H      K2(T),10H,G12.4     ,KT2,5,0.)
   RETURN
50 TGO=TF-TIME
   T1=1.-EXP(-TGO/TAU1)-TGO/TAU1
   T2=1.-EXP(-TGO/TAU2)-TGO/TAU2
   S=6.*R11S*R22S/(6.*R11S*R22S+AS*KT1**2*R22S*(6.*TAU1**2*TGO-6.*TAU
  11*TGO**2+2.*TGO**3+3.*TAU1**3*(1.-EXP(-2.*TGO/TAU1))-12.*TAU1**2*T
  2GO*EXP(-TGO/TAU1))-AS*KT2**2*R11S*(6.*TAU2**2*TGO-6.*TAU2*TGO**2+2
  3.*TGO**3+3.*TAU2**3*(1.-EXP(-2.*TGO/TAU2))-12.*TAU2**2*TGO*EXP(-TG
  4O/TAU2)))
   K1=AS*KT1**2*TAU1**2*T1**2/R11S
   K2=AS*KT2**2*TAU2**2*T2**2/R22S
   N0=-.000000015
   P0=1000.*EXP(-.5*TIME)
   NP1D=2.*(K1*N0+K1*S+P0/W2)*NP1+2.*N0*PP1/W2-K1*N0*N0-2.*N0*P0/W2+K
  12*S*S
   PP1D=-2.*K1*P0*NP1-2.*(K1*N0+K1*S+P0/W2)*PP1+P0*P0/W2+2.*K1*N0*P0
   NP2D=2.*(K1*NP1+K1*S+PP1/W2)*NP2+2.*NP1*PP2/W2-K1*NP1*NP1-2.*NP1*P
```

```
     1P1/W2+K2*S*S
      PP2D=-2.*K1*PP1*NP2-2.*(K1*NP1+K1*S+PP1/W2)*PP2+PP1*PP1/W2+2.*K1*N
     1P1*PP1
      NP3D=2.*(K1*NP2+K1*S+PP2/W2)*NP3+2.*NP2*PP3/W2-K1*NP2*NP2-2.*NP2*P
     1P2/W2+K2*S*S
      PP3D=-2.*K1*PP2*NP3-2.*(K1*NP2+K1*S+PP2/W2)*PP3+PP2*PP2/W2+2.*K1*N
     1P2*PP2
      NP4D=2.*(K1*NP3+K1*S+PP3/W2)*NP4+2.*NP3*PP4/W2-K1*NP3*NP3-2.*NP3*P
     1P3/W2+K2*S*S
      PP4D=-2.*K1*PP3*NP4-2.*(K1*NP3+K1*S+PP3/W2)*PP4+PP3*PP3/W2+2.*K1*N
     1P3*PP3
      NH4D=2.*(K1*NP3+K1*S+PP3/W2)*NH4+2.*NP3*PH4/W2
      PH4D=-2.*K1*PP3*NH4-2.*(K1*NP3+K1*S+PP3/W2)*PH4
      IF(NFIRST)60,80,60
   60 A4=-NP4/NH4
      WRITE(6,70)NP1,NP2,W2,A4
   70 FORMAT(1H0,4X,4HN    = G14.6,5X,4HP    =   G14.6,5X,5HW2 = G14.6,5HA4
     1=G22.14)
   80 RETURN
      END


      - FINAL SOLUTION

      SUBROUTINE FUNEV
      COMMON TIME,DELT,NSTART,NFIRST,NEXIT,IPASS,ROMCON(2094)
      REAL K1,K2,KT1,KT2,N,ND
      REAL N0,NP1,NP1D,NH1,NH1D,NP2,NP2D,NH2,NH2D
      REAL NP3,NH3,NP3D,NH3D
      REAL NP4,NH4,NP4D,NH4D
      REAL ND2,ND21,J1,J1D,J2,J2D,JR1,JR2,JR11,JR12,JR21,JR22,NND2
      DATA TF,AS,KT1,KT2,TAU1,TAU2,R11S,R22S/10.,.04,32.2,32.2,1.,2.,.00
     11,.004/
      IF(NSTART)30,50,10
   10 READ(5,20)NP1,PP1,W2,NP2,PP2,NP3,PP3,NP4,PP4,P2
   20 FORMAT(4E20.0)
      CALL INTG(NP1D,NP1)
      CALL INTG(PP1D,PP1)
      CALL INTG(NP2D,NP2)
      CALL INTG(PP2D,PP2)
      CALL INTG(NP3D,NP3)
      CALL INTG(PP3D,PP3)
      CALL INTG(NP4D,NP4)
      CALL INTG(PP4D,PP4)
      CALL INTG(P2D,P2)
      CALL INTG(J1D,J1)
      CALL INTG(J2D,J2)
      CALL PRINT(10H      S(T),10H,G12.4    ,S,1,0.)
      CALL PRINT(10H      NP1D,10H,G12.4    ,NP1D,3,0.)
      CALL PRINT(10H      NP1 ,10H,G12.4    ,NP1,1,0.)
      CALL PRINT(10H      PP1D,10H,G12.4    ,PP1D,3,0.)
      CALL PRINT(10H      PP1 ,10H,G12.4    ,PP1,1,0.)
```

```
      CALL PRINT(10H       NP2D,10H,G12.4      ,NP2D,3,0.)
      CALL PRINT(10H       NP2 ,10H,G12.4      ,NP2,1,0.)
      CALL PRINT(10H       PP2D,10H,G12.4      ,PP2D,3,0.)
      CALL PRINT(10H       PP2 ,10H,G12.4      ,PP2,1,0.)
      CALL PRINT(10H       NP3D,10H,G12.4      ,NP3D,3,0.)
      CALL PRINT(10H       NP3 ,10H,G12.4      ,NP3,1,0.)
      CALL PRINT(10H       PP3D,10H,G12.4      ,PP3D,3,0.)
      CALL PRINT(10H       PP3 ,10H,G12.4      ,PP3,1,0.)
      CALL PRINT(10H       NP4D,10H,G12.4      ,NP4D,3,0.)
      CALL PRINT(10H       NP4 ,10H,G12.4      ,NP4,1,0.)
      CALL PRINT(10H       PP4D,10H,G12.4      ,PP4D,3,0.)
      CALL PRINT(10H       PP4 ,10H,G12.4      ,PP4,1,0.)
      CALL PRINT(10H       ND2 ,10H,G12.4      ,ND2,1,0.)
      CALL PRINT(10H       ND21,10H,G12.4      ,ND21,1,0.)
      CALL PRINT(10H       NND2,10H,G12.4      ,NND2,1,0.)
      CALL PRINT(10H       G1S ,10H,G12.4      ,G1S,1,0.)
      CALL PRINT(10H       G2S ,10H,G12.4      ,G2S,1,0.)
      CALL PRINT(10H       G1N2,10H,G12.4      ,G1N2,1,0.)
      CALL PRINT(10H       G1N1,10H,G12.4      ,G1N1,1,0.)
      CALL PRINT(10H       G1N ,10H,G12.4      ,G1N,1,0.)
      CALL PRINT(10H       P2D ,10H,G12.4      ,P2D,3,0.)
      CALL PRINT(10H       P2  ,10H,G12.4      ,P2,1,0.)
      CALL PRINT(10H       J1D ,10H,G12.4      ,J1D,3,0.)
      CALL PRINT(10H       J1  ,10H,G12.4      ,J1,1,0.)
      CALL PRINT(10H       J2D ,10H,G12.4      ,J2D,3,0.)
      CALL PRINT(10H       J2  ,10H,G12.4      ,J1,1,0.)
      CALL PRINT(10H       JR11,10H,G12.4      ,JR11,1,0.)
      CALL PRINT(10H       JR12,10H,G12.4      ,JR12,1,0.)
      CALL PRINT(10H       JR21,10H,G12.4      ,JR21,1,0.)
      CALL PRINT(10H       JR22,10H,G12.4      ,JR22,1,0.)
      CALL PRINT(10H       JR1 ,10H,G12.6      ,JR1,1,0.)
      CALL PRINT(10H       JR2 ,10H,G12.6      ,JR2,1,0.)
   30 CALL PRINT(10H      K1(T),10H,G12.4      ,K1,5,0.)
      CALL PRINT(10H      K2(T),10H,G12.4      ,K2,5,0.)
      A=SQRT(AS)
      R11=SQRT(R11S)
      R22=SQRT(R22S)
      I=1
      RETURN
   50 TGO=TF-TIME
      T1=1.-EXP(-TGO/TAU1)-TGO/TAU1
      T2=1.-EXP(-TGO/TAU2)-TGO/TAU2
      S=6.*R11S*R22S/(6.*R11S*R22S+AS*KT1**2*R22S*(6.*TAU1**2*TGO-6.*TAU
     11*TGO**2+2.*TGO**3+3.*TAU1**3*(1.-EXP(-2.*TGO/TAU1))-12.*TAU1**2*T
     2GO*EXP(-TGO/TAU1))-AS*KT2**2*R11S*(6.*TAU2**2*TGO-6.*TAU2*TGO**2+2
     3.*TGO**3+3.*TAU2**3*(1.-EXP(-2.*TGO/TAU2))-12.*TAU2**2*TGO*EXP(-TG
     4O/TAU2)))
      K1=AS*KT1**2*TAU1**2*T1**2/R11S
      K2=AS*KT2**2*TAU2**2*T2**2/R22S
      ND2=6.*R11S/(6.*R11S+AS*KT1**2*(6.*TAU1**2*TGO-6.*TAU1*TGO**2+2.*T
     1GO**3+3.*TAU1**3*(1.-EXP(-2.*TGO/TAU1))-12.*TAU1**2*TGO*EXP(-TGO/T
     2AU1)))
```

3-8

DT