44

Office of Naval Research
Contract N00014-67-A-0298-0006 NR-372-012

LINEAR STOCHASTIC DIFFERENTIAL GAMES

VE RI TAS

By

Robert Dietrich Behn

December 1968

Technical Report No. 578

Division of Engineering and Applied Physics
Harvard University • Cambridge, Massachusetts

305

Office of Naval Research

Contract N00014-67-A-0298-0006

NR-372-012

LINEAR STOCHASTIC DIFFERENTIAL GAMES

By

Robert Dietrich Behn

Technical Report No. 578

January 1969

The research reported in this document was made possible through support extended the Division of Engineering and Applied Physics, Harvard University by the U. S. Army Research Office, the U. S. Air Force Office of Scientific Research and the U. S. Office of Naval Research under the Joint Services Electronics Program by Contracts N00014-67-A-0298-0006, 0005, and 0008.

Division of Engineering and Applied Physics

Harvard University · Cambridge, Massachusetts

BLANK PAGE

LINEAR STOCHASTIC DIFFERENTIAL GAMES

By

Robert Dietrich Behn

Division of Engineering and Applied Physics

Harvard University · Cambridge, Massachusetts

ABSTRACT

The solution for a class of stochastic pursuit-evasion differential games between two dynamic systems is given; this class includes those games where one of the players has perfect knowledge of the state of the game while the other player is constrained to make noisy measurements on this state. The dynamic systems involved are linear and the performance index which is optimized is quadratic.

The strategy for the player with perfect information is not always a realizable one. It is shown that this player can implement his strategy, however, if the number of his control variables is as great as the number of the state variables involved in the pursuit and evasion. Thus the solution obtained is applicable for the classical interception game in euclidean space.

Several aspects of this game are studied in detail. The asymmetric roles of the pursuer and evader are discussed in general and relationships drawn between the deterministic and stochastic cases. It is pointed out that this game requires -- in reality -- the solution to a non zero-sum game since the two different information sets employed by the two players cause each player to evaluate the criterion differently. The "certainty-equivalence principle" which characterizes the standard stochastic control problem is shown to be applicable to this class of differential games.

Examples of the classical interception game are given and numerical results presented.

TABLE OF CONTENTS

		Page
ABSTRACT		i
LIST OF FIGURES		v
Chapter One	AN INTRODUCTION	
1.1	A Short Discussion of Game Theory	1-1
1.2	Some Essentials of Differential Games	1-5
1.3	A Short History	1-12
1.4	A Deterministic Pursuit-Evasion Game	1-15
1.5	A Preview	1-20
	References for Chapter One	1-24
Chapter Two	A STOCHASTIC PURSUIT-EVASION GAME	
2.1	Formulation of the Problem	2-1
2.2	The Solution	2-3
2.3	Proof of the Pursuer's Optimality	2-11
2.4	Proof of the Evader's Optimality	2-16
2.5	The Criterion	2-21
2.6	A Summary of the Optimized System	2-24
	References for Chapter Two	2-27
Chapter Three	REALIZATION OF THE PURSUER'S STRATEGY	
3.1	Two Realization Schemes	3-1
3.2	Reflections on the Pursuer's Strategy	3-9
3.3	Another Proof of the Evader's Optimality	3-17
3.4	The Meaning and Existence of L^{-1}	3-20
Appendix III-A	THE DISCRETE-TIME GAME	3-26
A.1	The Deterministic Problem	3-27
A.2	The Stochastic Problem	3-31
A.3	A Non-Optimal, Discrete-Time Game	3-35
Appendix III-B	MATHEMATICAL DETAILS OF THE EVADER'S OPTIMALITY PROOF	3-48
	References for Chapter Three	3-50

Page

Chapter Four A STUDY OF THE EQUATIONS FOR Γ_2 AND P

4.1 The Conjugate Point Condition 4-2
4.2 Strategies for Infinite Noise Variance 4-10
4.3 The Results for Zero Noise Variance 4-14
References for Chapter Four 4-17

Chapter Five THE ASYMMETRIC ROLES OF PURSUER AND EVADER

5.1 The Deterministic Case 5-2
5.2 The Pursuer with Noisy Measurements 5-15
5.3 The Asymmetry of the Two Stochastic Games 5-19
References for Chapter Five 5-25

Chapter Six INSIGHTS AND GENERALIZATIONS

6.1 Non Zero-Sum Properties of the Game 6-1
6.2 The Criterion Reconsidered 6-7
6.3 The Irrelevance of Uniqueness 6-12
6.4 The Certainty-Equivalence Principle 6-19
6.5 General Measurements 6-22
6.6 Process Noise 6-25
6.7 The Generalized Game 6-27
References for Chapter Six 6-31

Chapter Seven THE CLASSICAL INTERCEPTION PROBLEM

7.1 The Solution to the Classical Interception Problem 7-1
7.2 Proportional Navigation 7-10
7.3 Optimal Parameter Values and Feedback Gains 7-16
7.4 A Numerical Study of $\Gamma_2(t)$ and P(t) 7-28
7.5 Optimal Trajectories 7-49
References for Chapter Seven 7-68

Chapter Eight CONCLUDING REMARKS

8.1 A Summary of the Approach 8-1
8.2 The Closure Problem 8-5
8.3 A Summary of the Significant Results 8-9
References for Chapter Eight 8-11

ACKNOWLEDGEMENTS

LIST OF FIGURES

Figure Number		Page
3-1	Implementation of the Optimal Strategies	3-3
3-2	A General Feedback System	3-5
3-3	Realization II for the Pursuer	3-8
3-4	A General Dynamic System	3-14
3-5	The Optimized System as Viewed from the Nodes of the Pursuer's Second Feedback Loop	3-16
3-6	The Pursuer's Discrete, Non-Optimal Controller	3-46
3-7	A Reduction of the Pursuer's Non-Optimal Controller	3-47
5-1	Initial Configuration in Euclidean (x) Space	5-10
5-2	Initial Configuration in Reduced (y) Space	5-10
5-3	Possible Locations of $y'(t_o)$	5-12
5-4	The Pursuer's and Evader's Trajectories	5-14
5-5	Graphic Representation of Conjugate Points	5-24
7-1	The Geometry of the Proportional Navigation Problem	7-12
7-2	The Parameters $K^{-1}(t_f, t)$ and $\Gamma_2(t)$	7-19
7-3	The Feedback Gains	7-20
7-4	The Variance of the Error of the Evader's Estimate	7-22
7-5	The Parameter L(t) with $K^{-1}(t_f, t)$ and $-\Gamma_2(t)$	7-23
7-6	The Relative Criterion as a Function of the Evader's Measurement Noise Variance	7-26
7-7	The Pursuer's Gains G_1 and G_2 for Realization II	7-27

Figure Number		Page
7-8	Feedback Gain G_1 for the Pursuer's Realization II	7-29
7-9	Feedback Gain G_2 for the Pursuer's Realization II	7-30
7-10	The Error Variance P(t) for Various Values of Q. Evader with Noisy Measurements	7-32
7-11	The Parameter $\Gamma_2(t)$ for Various Values of Q. Evader with Noisy Measurements	7-33
7-12	The Parameter $\frac{P(t)}{Q}$ for Games of $(t_f - t_o)$ up to 20. Evader with Noisy Measurements	7-35, 7-36
7-13	The Parameter $\Gamma_2(t)$ for Games of $(t_f - t_o)$ up to 20. Evader with Noisy Measurements	7-38
7-14	The Error Variance P(t) for Various Values of Q. Pursuer with Noisy Measurements	7-40
7-15	The Parameter $\Gamma_2(t)$ for Various Values of Q. Pursuer with Noisy Measurements	7-41
7-16	The Parameter $\frac{P(t)}{Q}$ for Games of $(t_f - t_o)$ up to 2.4. Pursuer with Noisy Measurements	7-44
7-17	The Parameter $\Gamma_2(t)$ for Games of $(t_f - t_o)$ up to 2.4. Pursuer with Noisy Measurements	7-45
7-18	The Parameter $\frac{P(t)}{Q}$ for the Game with the Deterministic Conjugate Point at $(t_f - t) = 1.49$. Evader with Noisy Measurements	7-47
7-19	The Parameter $\Gamma_2(t)$ for the Game with the Deterministic Conjugate Point at $(t_f - t) = 1.49$. Evader with Noisy Measurements	7-48
7-20	The Parameter $\frac{P(t)}{Q}$ for the Game with the Deterministic Conjugate Point at $(t_f - t) = 1.49$. Pursuer with Noisy Measurements	7-50

Figure Number		Page
7-21	The Parameter $\Gamma_2(t)$ for the Game with the Deterministic Conjugate Point at $(t_f - t) = 1.49$. Pursuer with Noisy Measurements	7-51
7-22	Optimal Deterministic Trajectory I	7-53
7-23	The Reduced State for a Deterministic Game	7-55
7-24	The Controls for a Deterministic Game	7-56
7-25	Optimal Stochastic Trajectory I	7-58
7-26	The Reduced State and Estimate for a Stochastic Game	7-60
7-27	The Controls for a Stochastic Game	7-61
7-28	Table of Various Criteria	7-63
7-29	Optimal Deterministic Trajectory II	7-65
7-30	Optimal Stochastic Trajectory II	7-66

BLANK PAGE

CHAPTER ONE

AN INTRODUCTION

1.1 A Short Discussion of Game Theory

The Theory of Games -- the mathematical study of conflict -- was first conceived and investigated by von Neumann and Morgenstern in their now classical, though only two-decade-old, work, Theory of Games and Economic Behavior [1]. Their analysis begins with Zero-Sum, Two-Person Games, for this class permits the most detailed and most satisfying mathematical study. It is defined by two specifications:

i. exactly two players are involved,

ii. the competition between them is perfect in the sense that what one player loses, the other player gains.

This second specification needs, perhaps, further explanation.

Before the discussion of any particular game problem can begin, it is necessary to establish a payoff for each one of the players, denoted by J_1 and J_2. These two payoffs translate the outcome of the game into particular numerical values for the two players, with each player desiring to maximize his own payoff; the payoffs must be scalars to permit a conclusive comparison of any two possible outcomes. If the scales and zero orientation can be selected so that for every outcome of the game

$$-J_1 = J = +J_2 \quad , \tag{1:1}$$

the game is said to be zero-sum, i.e. the sum of the payoffs to the competing players is zero. This is the meaning of specification ii. above. Throughout this thesis, the non-subscripted symbol J shall refer to a zero-sum payoff (alternatively called the criterion) which one player attempts to minimize while the other to maximize.

The theory of games attempts to determine general rules by which the players, when confronted with a conflict situation, determine their optimal behavior. These rules, called strategies, are denoted here by U and V, for players one and two respectively. Thus the payoff is a function of the opposing players' strategies, $J(U, V)$.

As presented by von Neumann and Morgenstern, the general game theory problem is then to determine the optimal strategies, U^o and V^o, which minimize and maximize the payoff respectively, i.e.

$$J(U^o, V^o) = \underset{V}{\mathrm{Max}}\ \underset{U}{\mathrm{Min}}\ J(U, V) \tag{1:2}$$

or alternatively

$$J(U^o, V^o) = \underset{U}{\mathrm{Min}}\ \underset{V}{\mathrm{Max}}\ J(U, V) \quad . \tag{1:3}$$

Unfortunately, however, these two are not equivalent, min-max does not always equal max-min, for the Max and Min operators do not in general commute. If -- for a particular functional relationship between J, and U and V -- (1:2) and (1:3) do give the same result, the optimal strategies and the optimal values of the payoff satisfy the saddle-point relation

$$J(U^o, V) \leq J(U^o, V^o) \leq J(U, V^o) \quad . \tag{1:4}$$

Strategies U^o and V^o which satisfy (1:4) are said to provide the "solution" to the game, in particular a minimax solution. $J(U^o, V^o)$ is called the value of the game.

The rationale behind selecting (1:4) as defining the solution to the game is that the optimal strategy pair (U^o, V^o) is said to be in "equilibrium." One player cannot gain by changing his strategy if the other does not change. This solution also has the interesting sidelight that if the players are going to play U^o and V^o neither player will lose anything by announcing his strategy.

This concept of an equilibrium solution can be employed to solve non zero-sum games as well. Consider two different payoffs J_1 and J_2 where player one attempts to minimize J_1 and player two to maximize J_2. Any equilibrium solution (U^o, V^o) is one such that

$$J_1(U^o, V^o) \leq J_1(U, V^o) \quad , \tag{1:5a}$$

$$J_2(U^o, V) \leq J_2(U^o, V^o) \quad . \tag{1:5b}$$

Methods employed to obtain game theoretic solutions, as defined by either (1:4) or (1:5), are no end in themselves; either pure guess-work or logical arguments founded on rigorous mathematics may prove successful. The only test is whether (1:4), or (1:5), is satisfied or not. As used below, the combination of intuition plus mathematics may be disdained by some. Justification is found, however, in the fact that the strategies obtained are in equilibrium.

In control theory, the equation which determines the optimal control (strategy), U^o, is

$$J(U^o) \leq J(U) \quad . \tag{1:6}$$

That this equation is the defining one for control theory is not open to question. If the solution obtained is inappropriate, one need only modify the criterion or impose more realistic constraints.

The role of (1:4) (or (1:5)) in game theory is analogous to that of (1:6) in control theory. (From this view, games are two-sided control problems.) It is important to note, however, that the von Neumann-Morgenstern formula (1:4) is not the only possibility for defining the "solution" to a zero-sum game; reasonable men could obtain other rational standards. A strategy based on (1:4) does not always produce a result which cannot be improved upon, while (1:6) -- in a theoretical sense -- does. Luce and Raiffa [2] discuss this dilemma drawing on the insights of modern utility theory.

The opposition to using (1:4) to solve game theory problems usually arises in the form of the question: "Suppose one player doesn't play 'optimally'?" The implication is that if player two fails to employ strategy V^o, then player one may be able to do better by selecting a strategy other than the one obtained from a calculation based on (1:4). This is, of course, true; if player two employs a strategy other than V^o, player one can find another strategy which will give a lower payoff than $J(U^o, V^o)$.

However, two observations must be made. First, by playing U^o player one insures that the payoff will have a value no greater than $J(U^o, V^o)$. Second, if player two fails to employ V^o, the value of the payoff will decrease. Also, if player one -- on the assumption that player two will not select strategy V^o -- employs a strategy other than U^o, he opens the possibility that the payoff will be greater than $J(U^o, V^o)$, the value which is guaranteed by selecting strategy U^o.

This philosophical argument is equally applicable to non zero-sum games and could continue indefinitely. The author has accepted the von Neumann-Morgenstern approach, as defined by (1:4) and (1:5), as providing the correct strategies to be employed in any game situation. Those who would question the validity of the solutions described below should be careful to make sure that they are not really raising this philosophical question. Perhaps the best method for dealing with the possibility that one player will not play "optimally" is to assess the probability that he will not do so and solve a statistical as well as a game problem, or to place constraints on this player which appropriately model his incapacity to either calculate or employ the optimal strategy.

1.2 Some Essentials of Differential Games

(i) Strategies and Time Functions - For any two-person game, the solution consists of two strategies which satisfy the equilibrium condition, either (1:4) or (1:5). These two strategies determine the players' moves or behavior for every possible situation which may arise during the play of the game. The number of possible situations may be small, and consequently the strategies may delineate explicitly the possible moves. (If he hits me, I'll hit him back. Otherwise, I won't.)

However, if the number of situations which may occur during the game is quite large it would be unrealistic to require a strategy to explicitly state the optimal action for every situation. Rather a strategy will indicate a general relationship between the possible situations and the corresponding optimal actions. (If I hold 4 aces, I'll raise my opponent twice his bet.)

As used in this thesis, the term differential game shall mean a game in which the players must make a continuity of decisions, i.e. the decision process must produce a function. This does not mean that the strategy is such a function, but rather a set of rules by which the particular function is selected on the basis of the situations which confront each player. In this thesis, these functions have as their independent element time; the purpose of the strategies is to dictate how each player will select his own time function as a result of the data which he receives.

It is important to note the distinction between a player's strategy and the particular time function he will employ. In the game discussed and solved below, a player will probably use a different time function for every play of the game, simply because of the random nature of his information. His strategy, however, is the same provided that the parameters of the random processes encountered are the same. It is the strategy, not the time function, which must satisfy (1:4).

(ii) The Basic Equations - The strategies U and V produce the time functions u and v called controls. They effect the value of a state variable -- a vector, y, which describes the state of the game --

usually through a differential equation

$$\dot{y} = f_1(y, u, v, t) \quad , \qquad y(t_o) = y_o \tag{1:7}$$

which shall be called the game equation. This state vector acts as an intermediate variable between the control vectors and the payoff; though the criterion may depend on the controls, the nature of the state's entry is usually primary to the character of the solution for it describes the game. The most general form of the payoff includes both a function of the terminal value of the state variable and a functional of the state and two control variables.

$$J = f_2(y(t_f)) + \int_{t_o}^{t_f} f_3(y, u, v, t)dt \quad . \tag{1:8}$$

Both f_2 and f_3 are scalars and termed the terminal criterion and in-flight criterion respectively; f_3 is equivalent to the Lagrangian of control theory. All elements of this general form need not be present.

(iii) Open-Loop and Closed-Loop Strategies - The strategies may uniquely determine the controls before the game begins on the basis of the initial data, e.g.

$$u = u(t, y(t_o)) \quad . \tag{1:9}$$

The term open-loop strategies -- one borrowed from control theory -- is applied here, for there is no feedback loop involved. They determine control functions that are merely based on the initial state and the time they are to be applied.

The other possibility is for the controls to be determined as the game progresses, i.e. the value of the control function at time t is dependent on the value of the state at that time. The values of these controls at time t could be obtained from a specific relationship to the status of the game at time t. It is this relationship then that is the strategy,

$$u = u(t, y(t)) \quad , \tag{1:10}$$

and is called closed-loop. This term also has its roots in control theory. A closed feedback loop is constructed so that the control is dependent on the state, which is in turn dependent on the control, thus producing the closed-loop.

(iv) Necessary Conditions and Separability - The calculus of variations provides a method for solving control theory problems, by converting the problem of finding the minimum of a functional (i.e. the criterion) into one of finding the minimum of a special function which is called the Hamiltonian. Similarly the calculus of variations can be useful in obtaining solutions to differential games, by converting the problem of determining a functional minimax to one of selecting a function minimax.

The Hamiltonian for the differential game of (1:7)-(1:8) is given by

$$\mathcal{H}(y, u, v, t) \triangleq f_3(y, u, v, t) + \lambda^T(t) f_1(y, u, v, t) \quad . \tag{1:11}$$

What is now desired is a set of conditions which are necessary to ensure that the Hamiltonian satisfies the saddle-point condition with respect to particular time functions $u^o(t)$ and $v^o(t)$, during the entire play of the game.

$$\mathcal{H}(y, u^o, v, t) \leq \mathcal{H}(y, u^o, v^o, t) \leq \mathcal{H}(y, u, v^o, t) \quad . \tag{1:12}$$

These conditions are

$$\frac{\partial \mathcal{H}}{\partial u} = 0 \quad , \tag{1:13}$$

$$\frac{\partial \mathcal{H}}{\partial v} = 0 \quad , \tag{1:14}$$

$$\frac{\partial^2 \mathcal{H}}{\partial u^2} \geqslant 0 \quad , \tag{1:15}$$

$$\frac{\partial^2 \mathcal{H}}{\partial v^2} \leqslant 0 \quad . \tag{1:16}$$

These become necessary and sufficient when the partial inequalities of (1:15) and (1:16) become strict inequalities.

As was noted in Section 1.1, when attempting to satisfy a saddle-point condition, here (1:12), min-max will not always be found equal to max-min. It is well known (see for example, Bryson and Ho [3]) that here a simple condition for the two to be equal, i.e. for (1:12) to be satisfiable, is that the Hamiltonian be separable.

$$\mathcal{H}(y, u, v, t) = \mathcal{H}'(y, u, t) + \mathcal{H}''(y, v, t) \quad . \tag{1:17}$$

If this condition is satisfied, and if (1:13) and (1:14) determine values of v and u which satisfy (1:15) and (1:16), then these values do indeed provide a saddle-point for the Hamiltonian.

Separability means that there are no "cross terms" of u and v explicit in either the differential equation or the inflight criterion.

$$f_1(y, u, v, t) = f_1'(y, u, t) + f_1''(y, v, t) \quad , \tag{1:18}$$

$$f_3(y, u, v, t) = f_3'(y, u, t) + f_3''(y, v, t) \quad . \tag{1:19}$$

It guarantees that (1:13), which is used to determine u^o, is independent of v, and that (1:14), which is used to determine v^o, is independent of u. Thus (1:13) and (1:14) do not have to be solved simultaneously, but rather u^o and v^o can be obtained independently.

In order for the values of u^o and v^o which are obtained from (1:13) and (1:14) to provide a minimax of the functional J, it is furthermore necessary that λ be selected on the basis that

$$\dot{\lambda}^T = -\frac{\partial \mathcal{H}}{\partial y} \quad , \qquad \lambda^T(t_f) = \frac{\partial f_2(y(t_f))}{\partial y(t_f)} \quad . \tag{1:20}$$

Thus (1:13), (1:14) and (1:20) provide a method for obtaining a u^o and a v^o which must be used to obtain the optimal trajectory, i.e. these values of u and v must be used during the entire play of the game if the resulting play is to have been determined by optimal strategies. In particular (1:13) determines the time function $u^o(t)$, and (1:14) the time function $v^o(t)$. Consequently, this is a method for deriving open-loop strategies. However, these open-loop strategies can often be converted into closed-loop ones as is done in Section 1.4.

The fact that values of u(t) and v(t) are determined which satisfy (1:12) does not ensure that they satisfy (1:4). The separability conditions guarantee only a saddle-point for the Hamiltonian; what would really be desirable is the separability of J, a condition which is much more difficult to ascertain and generally not true. Consequently, solutions obtained by the above method provide only a candidate for a strategy pair which may satisfy (1:4). It is merely a technique -- like guessing -- for obtaining a possibility; any pair of strategies must always be checked directly by substituting into (1:4).

(v) Admissible Strategies - One difficulty encountered in differential games is that of defining the admissible strategies. In basketball, an unadmissible strategy for the leading team is to puncture the ball, deflate it and stuff it under a player's shirt in an effort to run out the clock. In this thesis, admissible strategies shall mean those strategies which produce control functions that result in a meaningful play of the game.

The question of admissible strategies cannot, in general, be settled by investigating each player individually; it is also important to investigate the problem of non-playable strategy pairs. It is possible that a particular strategy employed by player one is admissible against some of player two's strategies, but that against others will result in a play of the game which never terminates. For sudden-death overtime in some sports (e.g. soccer, lacrosse, hockey) a strictly defensive strategy by both teams (with each team huddled around its own goal) would be an unplayable pair; by these strategies, neither team would send a man to attempt to score, and thus the game would never end. In general then, it is necessary to define not the admissible strategies, but rather the "admissible strategy pairs."

It is important to note that this possibility does not occur in the game investigated below. All games are of fixed time duration and consequently any pair of strategies, selected from the two independent sets of admissible strategies, will result in the termination of the game when the clock runs out.

If the physics of a particular game demanded, the admissible strategies could be limited to open-loop ones. Then the solution to the game would consist of two specific time functions and these time functions themselves would have to satisfy the saddle-point condition. In general, however, closed-loop strategies are permitted. Specific examples of related open-loop and closed-loop strategies are found in Chapter 5.

(vi) Pursuit-Evasion Games - Pursuit-Evasion games are an important subclass of differential games. As the title suggests, this includes situations in which one player, the pursuer, attempts to minimize the "distance" between the two to implement capture; the other, the evader, attempts to escape. Thus, the state variable must somehow describe, either implicitly or explicitly, this distance, however it is measured. Also these elements of y(t) must enter into the terminal function f_2; other elements may be present in the total criterion, but this one is essential.

Thus for pursuit-evasion games, player one -- who desires to minimize the criterion which includes terminal separation -- is called the pursuer and player two is called the evader. Throughout this thesis, the subscripts p and e denote variables which relate to these two players, respectively.

1.3 A Short History

Differential games were first defined and studied by Isaacs in 1954 [4]. His initial work, accomplished without knowledge of both concurrent developments in control theory, independently obtained the Euler-Lagrange equation of the calculus of variations and the Bellman equation of dynamic programming. Eschewing rigorous mathematical analysis, he studied a number of specific differential games and, by cataloging their peculiarities, discovered a set of "singular behavior." This resulted in a number of different singular surfaces, which though all similar in nature to the singular surface in control theory, characterized a number of different types of phenomena. As the nature of the solution to a game problem is more complex than the solution of a control problem, so the nature of singular surface is more complex in game theory than in control theory. These surfaces are discussed in detail in Isaacs book [5] which is the most comprehensive treatment of differential games.

In 1957, Berkowitz and Fleming [6] applied classical variational techniques to simple differential games, with Berkowitz [7] expanding the applicable class of problems in 1963. The diverse approaches to control theory, as exemplified by the engineer and the mathematician, are mirrored in the field of differential games. Isaacs sought purely formal solutions to particular problems and in doing so discovered and catalogued numerous irregularities which he called singular surfaces. On the other hand Berkowitz and Fleming examined abstract mathematical games; in particular Berkowitz rigorously investigated the mathematical nature of solutions near these singular surfaces determining specific restrictions needed to guarantee existence.

In 1965, Ho, Bryson and Baron [8] used variational techniques to solve a general linear-quadratic pursuit-evasion game in phase space. A pursuer and an evader each controlled his own linear dynamic system. The terminal criterion was quadratic in the separation of the two systems at the terminal time; the inflight criterion consisted of two individual quadratic energy terms, one measuring the evader's control effort, the other the pursuer's. Feedback control strategies based on a predicted terminal miss were shown to be optimal. Baron [9] extended this work to include both energy and amplitude constraints on the two controls.

In 1966, Ho [10] solved a stochastic differential game, again employing variational techniques. Player one, only, controlled the state, and attempted to minimize a terminal error and simultaneously to confuse player two. The latter merely made noisy measurements on the state and attempted to minimize the error of his estimate. Since only one player actually controlled the state, the game was not of the pursuit-evasion type. It was solved sequentially, by first determining the form of player one's controller, and then using this result to determine the form of player two's estimator. The solution to this problem indicated that at a specified time, player one would change strategies from one with the main goal of confusing player two, to one which concentrated on minimizing the terminal criterion. Speyer [11] has investigated the conditions for singular control during the period of switching strategies.

Rhodes [12] has also studied stochastic games.

An excellent summary of the progress and various aspects of differential games is found in Simakova [13].

1.4 A Deterministic Pursuit-Evasion Game

The pursuit-evasion game solved by Ho, Bryson and Baron [8] is of particular interest to this thesis, for it is the deterministic model from which the stochastic game to be considered is taken.

This game is defined by two linear dynamic systems:

$$\dot{x}_p = F_p(t)x_p + \overline{G}_p(t)u(t) \quad , \qquad x_p(t_o) = x_{po} \quad , \tag{1:21}$$

$$\dot{x}_e = F_e(t)x_e + \overline{G}_e(t)v(t) \quad , \qquad x_e(t_o) = x_{eo} \quad , \tag{1:22}$$

and a quadratic payoff:

$$J = \frac{a^2}{2} \|x_p(t_f) - x_e(t_f)\|^2_{A^TA} + \frac{1}{2}\int_{t_o}^{t_f} [\|u(t)\|^2_{R_p(t)} - \|v(t)\|^2_{R_e(t)}]dt \ . \tag{1:23}$$

Here x_p is an n-vector describing the state of the pursuer's dynamic system, u(t) is an m-vector representing the control of the pursuer, $F_p(t)$ is an n x n matrix (the pursuer's system matrix) which is continuous in t, and $\overline{G}_p(t)$ is an n x m matrix (the pursuer's control matrix) which is also continuous in t. Similar statements apply to x_e, v(t), $F_e(t)$ and $\overline{G}_e(t)$, which describe the evader's dynamic system. $R_p(t)$ and $R_e(t)$ are respectively m x m and m' x m' dimensional, positive-definite, control weighting matrices.

The matrix A is of the form $[I_k:0]$, where I_k is the k-dimensional identity matrix, and where $1 \leq k \leq n$. The effect of A is to include in the criterion only those components of the two-state vectors which are relevant to the pursuit and evasion. Consequently, though there are n components of both state vectors only k of those components are "interesting" from the point of view of the criterion.

The a^2 term permits a weighting of the importance of terminal miss against control effort. The game is of fixed duration, $t_f - t_o$. u(t) and v(t) belong to m-dimensional and m'-dimensional euclidean spaces respectively.

The pursuer attempts to minimize the criterion (1:23), i.e. minimize the terminal miss distance and his own energy expended, while maximizing the energy expended by the evader. The evader attempts to maximize this same criterion. Thus the game is zero-sum.

Each player is assumed to have knowledge of the other's capabilities, i.e. the F, $\vec{G}$, and R matrices are known. For this deterministic problem, each player is also assumed to have knowledge of the other's state. Thus, this is a zero-sum game of complete information, and under certain conditions optimal strategies can be determined.

Since there are only k interesting state variables, it is worthwhile to reduce the dimension of the dynamics of (1:21) and (1:22). This is accomplished by defining a reduced state vector, denoted y(t), which represents the relevant terminal miss, $A[x_p(t_f) - x_e(t_f)]$, predicted at time t on the basis that no control will be applied during the interval (t, t_f).

$$y(t) \triangleq A[\Phi_p(t_f, t)x_p(t) - \Phi_e(t_f, t)x_e(t)] \quad . \tag{1:24}$$

Here Φ_p and Φ_e are the transition matrices obtained from F_p and F_e respectively. It should be noted that y(t) is, in itself, more meaningful than $x_p(t)$ and $x_e(t)$ -- even though y is k-dimensional, while x_p and x_e are n ($\geq$ k) dimensional -- for y(t) directly relates to the terminal miss term in the criterion.

Using (1:24), one can reduce (1:21), (1:22), and (1:23) to a simpler definition of the problem. The criterion is:

$$J = \frac{a^2}{2} \|y(t_f)\|^2 + \frac{1}{2} \int_{t_o}^{t_f} [\|u(t)\|^2_{R_p(t)} - \|v(t)\|^2_{R_e(t)}]dt \qquad (1:25)$$

subject to the differential constraint

$$\dot{y} = G_p(t_f, t)u(t) - G_e(t_f, t)v(t) \quad , \qquad y(t_o) = y_o \qquad (1:26)$$

where the new control matrix for the pursuer is given by:

$$G_p(t_f, t) = A\Phi_p(t_f, t)\overline{G}_p(t) \qquad (1:27)$$

and similarly for G_e. By use of his control, u(t), the pursuer still attempts to minimize this new form of the criterion, while the evader, with v(t), to maximize it. Note that the Hamiltonian is separable.

The only relevant restriction to be placed on the strategies which can be employed is that they guarantee that (1:26) is meaningful and integrable. Thus admissible strategies for both players shall be strategies which produce controls, either u(t) or v(t), which are bounded, and which are continuous in both t and y almost everywhere for $t_o \leq t \leq t_f$.

Ho, Bryson and Baron solve the problem of (1:25) and (1:26) by using the calculus of variation techniques outlined in Section 1.2 to obtain time functions, u*(t) and v*(t),

$$u^*(t) = -R_p^{-1}(t)G_p^T(t_f, t)K^{-1}(t_f, t_o)y(t_o) \qquad (1:28)$$

$$v^*(t) = -R_e^{-1}(t)G_e^T(t_f, t)K^{-1}(t_f, t_o)y(t_o) \qquad (1:29)$$

where

$$K(t_f, t_o) = \frac{I_k}{a^2} + M_p(t_f, t_o) - M_e(t_f, t_o) \tag{1:30}$$

and

$$M_p(t_f, t_o) = \int_{t_o}^{t_f} G_p(t_f, \tau) R_p^{-1}(\tau) G_p^T(t_f, \tau) d\tau \quad . \tag{1:31}$$

A similar expression defines $M_e(t_f, t)$. It should be noted that M_p and M_e are, from control theory, the reduced controllability matrices of the pursuer and evader respectively.

These time functions can be viewed as open-loop strategies, U* and V*, from which closed-loop strategies can be obtained. This is accomplished by noting that the current time can always be viewed as the initial time for defining a new problem. Thus, the relation between $u(t_o)$ and $y(t_o)$ as given by (1:28) can be taken as the relation between $u(t)$ and $y(t)$ for any $t_o \leq t \leq t_f$. Similarly for $v(t)$ and $y(t)$. Consequently, the optimal strategies, U^o and V^o, are feedback strategies of the form,

$$u^o(t) = C_p(t) y(t) \tag{1:32}$$

$$v^o(t) = C_e(t) y(t) \tag{1:33}$$

where the feedback matrix gains are given by

$$C_p(t) = -R_p^{-1}(t) G_p^T(t_f, t) K^{-1}(t_f, t) y(t) \tag{1:34}$$

$$C_e(t) = -R_e^{-1}(t) G_e^T(t_f, t) K^{-1}(t_f, t) y(t) \quad . \tag{1:35}$$

When these strategies are inserted into (1:21), the optimized game equation becomes

$$\dot{y} = -[G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T]K^{-1}y \quad , \qquad y(t_o) = y_o \quad . \tag{1:36}$$

The optimal value of the criterion is

$$J = \frac{a^2}{2} \|y(t_o)\|^2_{K^{-1}(t_f, t_o)} \quad . \tag{1:37}$$

It is shown in [9] that the existence of $K^{-1}(t_f, t)$ in the interval $t_o \leqslant t \leqslant t_f$ is equivalent to the nonexistence of a conjugate point for the game problem in this same interval. I. e. the existence of $K^{-1}(t_f, t)$ is a necessary and sufficient condition that U^o and V^o, as given by (1:32) and (1:33), satisfy the saddle-point condition (1:4). The existence of K^{-1} is guaranteed if

$$M_r(t_f, t) \triangleq M_p(t_f, t) - M_e(t_f, t) > 0 \quad . \tag{1:38}$$

M_r can be called the <u>relative controllability matrix</u>; the fact that it is positive-definite indicates that the pursuer is "more controllable" than the evader for every component of y.

The play of the game resulting from (1:28)-(1:29) is identical to the play from (1:32)-(1:33), simply because the use of either pair results in the constant vector

$$K^{-1}(t_f, t)y(t) = K^{-1}(t_f, t_o)y(t_o) \quad . \tag{1:39}$$

However, there are two reasons for employing the closed-loop rather than open-loop strategies. First, the conditions under which U^* and V^* are optimal in the sense that they satisfy (1:4) are merely a subset of the conditions under which U^o and V^o do. In particular there are conditions under which the pursuer's open-loop strategy permits the evader to employ a control other than (1:29) or (1:33) to get further away. By using the corresponding feedback strategy, however, the pursuer forces the evader to use the optimized control (either (1:29) or (1:33)), thus ensuring the value of the criterion given in (1:37). A detailed discussion of this point is found in Section 5.1.

Second, by employing a closed-loop control, either player takes the fullest possible advantage of any mistakes made by his opponent; he is better equiped to make inflight adjustments in his control. This is best seen by noting that even if U* and V* satisfy (1:4) the following are true.

$$J(U^o, V) \leq J(U^*, V) \tag{1:40}$$

$$J(U, V^*) \leq J(U, V^o) \quad . \tag{1:41}$$

1.5 A Preview

A logical extension of this deterministic, linear-quadratic pursuit-evasion game is one in which the players have imperfect knowledge of the states involved. In particular, this thesis reports the results of an investigation of the problem where one player is endowed with perfect information while the other is constrained to making measurements corrupted by additive, Gaussian noise. The difficulties encountered when both players have imperfect knowledge have not been resolved; a discussion of this problem is found in the concluding chapter.

The results obtained are in the form of optimal feedback strategies. In the case of the player making noisy measurements, the strategy is based on an optimal estimate of the state with the feedback matrix gain identical to the one this player employed in the deterministic problem. The other player's feedback control consists of two terms: one based on the value of the state, and the second on the error of his opponent's estimate. The first matrix gain which operates on the state vector is again identical to the one employed in the deterministic problem; the other gain is a new result which is obtained from the solution of a two-point boundary-value problem of two coupled Riccati equations.

As seen from the general nature of these two strategies, the solution to this game can be interpreted in light of the "certainty-equivalence principle" of (one-sided) stochastic control theory. (See Tou and Joseph, [14], and Franklin and Gunkel, [15].) This principle permits the optimal controller for the stochastic control problem to be divided into two separate units. The first is an optimal estimator of the state; the second is a feedback control scheme (the same one as for the corresponding deterministic problem) which operates on the optimal estimate as if it were the actual value of the state.

For the stochastic pursuit-evasion game studied here, the player with imperfect information can employ this same "certainty-equivalence principle." His optimal strategy consists of using the noisy measurements to obtain an optimal estimate of the state and then employing the feedback gains obtained for the deterministic problem (either (1:33) or (1:34)) in conjunction with this estimate.

The other player uses the optimal feedback strategy obtained in the deterministic problem in conjunction with the state which it knows perfectly. His strategy, however, includes a feedback gain applied to the error of his opponent's estimate. Under certain conditions, this error can be calculated exactly (see Chapter 3). This player's advantage results from the fact that his opponent must employ noisy measurements, and thus cannot realize a strategy which is as "good" as the one the opponent employed in the deterministic case.

One comment about differential games in general -- made so often that it has now become trite -- is that they reduce to optimal control problems if one player's control (strategy) is fixed. In this sense control theory is a subset of differential game theory. For this problem it is interesting to note the two control theory subsets to which this problem reduces.

If the control of the player with perfect information is fixed, the game is reduced to a stochastic control problem with measurements corrupted by white noise (see Section 2.4). When the control of the player who receives imperfect measurements on the state is fixed in a deterministic manner, the problem reduces to one of deterministic control theory. When, however, this player's control is fixed in some feedback manner based on an estimate obtained from the corrupted measurement, the problem reduces to a stochastic control problem with perfect information but also with colored process noise. It is thus interesting to note that what is viewed as measurement noise by one player is process noise to the other; see Section 2.6.

This thesis considers the various aspects of this particular game: schemes for one player to determine the error in the other's estimate; the difference between situations when the pursuer has the noisy measurements and when the evader does; the influence of the measurement noise variance on the strategies; and other generalizations. The classical interception problem and proportional navigation are discussed as examples of the application of the theory to concrete problems; numerical calculations are given to shed further light on the nature of the solutions.

This thesis does not concern itself with the rigorous mathematical problems of differential games. The objective is not an all-inclusive existence theorem but rather the solution to one particular problem, which may not only be useful in its own right, but also provide some insight for future progress.

References for Chapter One

[1] J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton, N. J., Princeton University Press, 1944.

[2] R. D. Luce and H. Raiffa, Games and Decisions, New York, John Wiley and Sons, Inc., 1957.

[3] A. E. Bryson, Jr. and Y. C. Ho, Optimization, Estimation and Control, Waltham, Massachusetts, Blaisdell Company, to be published, Chapter 9.

[4] R. Isaacs, "Differential games I, II, III, IV," RAND Corporation Research Memorandum RM-1391, 1399, 1411, 1468, 1954-1956.

[5] R. Isaacs, Differential Games, New York, John Wiley and Sons, Inc., 1965.

[6] L. D. Berkowitz and W. H. Fleming, "On differential games with integral payoff," Annals of Math. Study, No. 39, Princeton, N. J., Princeton University Press, 1957, pp. 413-435.

[7] L. D. Berkowitz, "A variational approach to differential games," Annals of Math. Study, No. 52, Princeton, N. J., Princeton University Press, 1964, pp. 127-173.

[8] Y. C. Ho, A. E. Bryson, Jr. and S. Baron, "Differential Games and Optimal Pursuit-Evasion Strategies," IEEE Transactions on Automatic Control, AC-10, No. 4, Oct. 1965, pp. 385-389.

[9] S. Baron, Differential Games and Optimal Pursuit-Evasion Strategies, Ph.D. Thesis, Harvard University, Cambridge, Massachusetts, 1966.

[10] Y. C. Ho, "Optimal Terminal Maneuver and Evasion Strategy," J. SIAM Control, Vol. 4, No. 3, 1966, pp. 421-428.

[11] J. L. Speyer, "A Stochastic Differential Game with Controllable Statistical Parameters," IEEE Transactions on System Science and Cybernetics, SSC-3, No. 1, June 1967, pp. 17-20.

[12] I. B. Rhodes, Optimal Control of Dynamic Systems by Two Controllers with Conflicting Objectives, Ph.D. Thesis, Stanford University, Palo Alto, California, 1967.

[13] E. N. Simakova, "Differential Games," Automation and Remote Control (Avtomatika i Telemekhanika), Vol. 27, No. 11, Nov. 1966, pp. 1980-1998.

[14] P. Joseph and J. Tou, "On Linear Control Theory," Transactions of the AIEE, Part III, September 1961.

[15] T. F. Gunckel and G. F. Franklin, "A General Solution for Linear Sampled-Data Control Systems," Journal of Basic Engineering, 1963, Vol. 85, Series D, No. 2, Transactions of the ASME, pp. 197-203.

BLANK PAGE

CHAPTER TWO

A STOCHASTIC PURSUIT-EVASION GAME

2.1 Formulation of the Problem

The dynamics of this stochastic problem are taken to be of the same form as for the deterministic case considered above, (1:16) and (1:17), which can be reduced to the single equation

$$\dot{y}(t) = G_p(t)u(t) - G_e(t)v(t) \quad , \qquad y(t_o) = y_o \quad , \tag{2:1}$$

where again y(t) is the predicted terminal miss as given by (1:19).†

For any stochastic problem, however, the form of the criterion (1:18) used in the deterministic game is inappropriate, for it implicitly assumes that $y(t_f)$ is known or can be determined exactly. Thus, this criterion must be modified by the use of an expected value operator.

$$J = E\{\frac{a^2}{2} \|y(t_f)\|^2 + \frac{1}{2}\int_{t_o}^{t_f} [\|u(t)\|^2_{R_p(t)} - \|v(t)\|^2_{R_e(t)}]dt\} \quad . \tag{2:2}$$

The problem to be solved here is one involving measurement noise. The pursuer is endowed with perfect knowledge of the state vector y(t), while the evader has only an estimate of the initial value of the state and noisy measurements of the state made during the play of the game.* The evader's initial estimate is denoted by $\hat{y}_o$, whose

---

† For the remainder of this thesis, the t_f argument will be omitted from G_p and G_e.

* This roughly reflects the practical situation where the pursuing dynamic system has been observing the evader for some time before the game starts and has considerable ground support for determining the state of the evader.

variance from the actual value of $y(t_o)$ is known to be P_o. The evader's measurements -- on the state of both the pursuer's system and his own -- are of the form:

$$z_1(t) = x_p(t) + w_1(t) \quad , \tag{2:3}$$

$$z_2(t) = x_e(t) + w_2(t) \quad , \tag{2:4}$$

where w_1 and w_2 are Gaussian white noise vectors, with zero mean and spectrum $Q_1(t)$ and $Q_2(t)$ respectively. The cross correlation between $w_1(t)$ and $w_2(\tau)$ is $C(t)\delta(t - \tau)$.

Since the dynamical system for the pursuit-evasion problem is given by (2:1) the evader's measurements must be reduced to measurements on y(t),

$$z = H(t)y(t) + w(t) \quad , \tag{2:5}$$

where w(t) is Gaussian white noise with zero mean and spectrum Q(t).

By premultiplying (2:3) by $A\Phi_p$, (2:4) by $A\Phi_e$, and subtracting we can define a new measurement vector:

$$z \triangleq A\Phi_p z_1 - A\Phi_e z_2 \tag{2:6}$$

which means z is of the form

$$z(t) = y(t) + w(t) \quad , \tag{2:7}$$

where

$$w(t) = A[\Phi_p w_1 - \Phi_e w_2] \tag{2:8}$$

has zero mean and spectrum

$$Q = A[\Phi_p Q_1 \Phi_p^T + \Phi_e Q_2 \Phi_e^T - \Phi_p C \Phi_e^T - \Phi_e C^T \Phi_p^T]A^T \quad . \tag{2:9}$$

In the solution obtained below it will be assumed that the measurements are of the general form (2:5) rather than the more limited case of (2:7). How this type of measurement can be obtained in general is discussed in Chapter 6.

The class of admissible strategies is defined as those U and V which give rise to controls

$$u = u(t, y(\tau): t_o \leq \tau \leq t) \tag{2:10}$$

and

$$v = v(t, z(\tau): t_o \leq \tau \leq t) \tag{2:11}$$

that are bounded and that are continuous almost everywhere for $t_o \leq t \leq t_f$. Note that these strategies utilize all available information. This class of strategy pairs is selected to insure that when the resulting controls u(t) and v(t) are inserted into (2:1), this differential equation is meaningful and integrable.

The problem then, is to find a pair of strategies (U^o, V^o), subject to (2:1), which satisfy the equilibrium condition (1:5) for all admissible strategies. Here J_1 is the value of (2:2) based on the evader's information set and J_2 the value of (2:2) for the pursuer's information set. The necessity for seeking a non zero-sum game solution is discussed in Sections 6.1 and 6.2.

2.2 The Solution

The approach used to obtain the solution to the problem outlined in the previous section consists of four basic steps:

[i] The structural form of the optimal strategies, U^o and V^o, are assumed to be given by

$$V: \; v = C_e(t)\hat{y}(t) \tag{2:12}$$

$$U: \; u = C_p(t)y(t) + D_p(t)\tilde{y}(t) \tag{2:13}$$

where $\hat{y}(t)$ is the evader's estimate of the state $y(t)$ which he constructs from his estimate of the initial state and from his measurements,

$$\hat{y}(t) \triangleq E[y(t)/\hat{y}_o, z(\tau): t_o \leq \tau \leq t] \tag{2:14}$$

and where $\tilde{y}(t)$ is the error in the evader's estimate

$$\tilde{y}(t) \triangleq y(t) - \hat{y}(t) \quad . \tag{2:15}$$

Here, $\hat{y}(t)$ is computed under the assumption that U^o and V^o are actually employed, and $\tilde{y}(t)$ is assumed to be somehow given to the pursuer by a mystical third party.

[ii] The values of the feedback gain matrices C_e, C_p, and D_p are determined by standard optimization techniques based on the procedure outlined in Section 1.2. This is the main concern of this section.

[iii] By solving two separate optimization problems, these strategies are shown to be optimal in the sense that they are in equilibrium. In Section 2.3, the strategy V^o determined in [ii] is assumed employed by the evader and the criterion is minimized over all possible U in the class of admissible strategies, thus verifying that this strategy pair satisfies the inequality of (1:5a). The other inequality is verified by a similar procedure in Section 2.4.

[iv] It is demonstrated that under certain conditions the pursuer can indeed calculate directly the error of the evader's estimate. This is the subject of Chapter 3.

The form of the evader's strategy assumed in (2:12) is sensible, since it reflects intuition gained in stochastic control theory. The first term in (2:13) comes from the form of the pursuer's strategy in the deterministic game; the second term comes from the suspicion that the pursuer should, somehow, be able to take advantage of the inaccuracies in his opponent's estimate, and the guess that a linear feedback relation might be appropriate.

Step [ii] is now carried out under the assumption that the evader's error is available to the pursuer, perhaps provided by some mystical third party. Methods by which the pursuer can calculate $\hat{y}$ are the subject of Chapter 3.

Using (2:12) and (2:13) to determine the values of u(t) and v(t) and using (2:15) to eliminate $\hat{y}(t)$,

$$J = E\left\{\frac{a^2}{2}\|y(t_f)\|^2 + \frac{1}{2}\int_{t_o}^{t_f}\left[\|y(t)\|^2_{C_p^T R_p C_p} + y^T(t)C_p^T R_p D_p \tilde{y}(t) + \tilde{y}(t)D_p^T R_p C_p y(t) + \|\tilde{y}(t)\|^2_{D_p^T R_p D_p} - \|y(t) - \tilde{y}(t)\|^2_{C_e^T R_e C_e}\right]dt\right\} \qquad (2:16)$$

which is to be minimaximized subject to the dynamical constraint

$$\dot{y}(t) = [G_pC_p - G_eC_e]y(t) + [C_pD_p + G_eC_e]\tilde{y}(t) \quad , \qquad y(t_o) = y_o \ . \qquad (2:17)$$

Since the forms of u and v are specified, the evader's estimate of the state is obtained from a Kalman-Bucy filter [1] which determines a Bayesian estimate or conditional mean of the state based on: a prior estimate of the initial state, y_o; a prior estimate of the variance of the error of this estimate, P_o; the measurements of the state, $z(\tau)$ from $\tau = t_c$ to $\tau = t$; and the dynamical equation (2:17).*

$$\dot{\hat{y}}(t) = [G_pC_p - G_eC_e]\hat{y}(t) + PH^TQ^{-1}[z - H\hat{y}(t)] \quad ,$$

$$\hat{y}(t_o) = \hat{y}_o \quad , \tag{2:18}$$

where P(t) is the variance of the error of the evader's estimate defined by

$$P(t) \triangleq E[\tilde{y}(t)\tilde{y}^T(t)] \quad . \tag{2:19}$$

P(t) is obtained from:

$$\dot{P} = [G_pC_p + G_pD_p]P(t) + P(t)[G_pC_p + G_pD_p]^T - P(t)H^TQ^{-1}HP(t) \quad ,$$

$$P(t_o) = P_o \quad , \tag{2:20}$$

and thus can be calculated and stored before the actual play of the game, once C_p, C_e, and D_p are determined.

The criterion can be further rewritten in the form

---

* From (2:15) it can be seen that any two of the three variables (y(t), $\hat{y}(t)$, $\tilde{y}(t)$) completely describe the state of the system. The selection of y(t) and $\tilde{y}(t)$ leads directly to results with useful interpretations. (2:18) may be more obvious, however, if (2:17) is rewritten in the form:

$$\dot{y}(t) = [G_pC_p + G_pD_p]y(t) - [G_eC_e + G_pD_p]\hat{y}(t) \quad , \qquad y(t_o) = y_o \quad .$$

$$J = \mathrm{Tr}\{ \frac{a^2}{2} Y(t_f) + \frac{1}{2} \int_{t_o}^{t_f} [Y(t)(C_p^T R_p C_p - C_e^T R_e C_e)$$

$$+ P(t)(C_p^T R_p D_p + D_p^T R_p C_p + D_p^T R_p D_p + C_e^T R_e C_e)]dt\} \qquad (2{:}21)$$

where Y(t) is defined by

$$Y(t) \triangleq E[y(t)y^T(t)] \qquad (2{:}22)$$

and is obtained from

$$\dot{Y} = G_p C_p Y + Y C_p^T G_p^T + G_p D_p P + P D_p^T G_p^T - G_e C_e [Y - P]$$

$$- [Y - P] C_e^T G_e^T \quad , \qquad Y(t_o) = Y_o \quad . \qquad (2{:}23)$$

where Y_o is given by $P_o + E[\hat{y}_o \hat{y}_o^T]$.

The stochastic game has now been transformed into a deterministic one by use of the expected value operator. For this new game, P and Y are the state variables, C_p and D_p the pursuer's control variables, and C_e the evader's control variable. (2:20) and (2:23) are the dynamical system or game equations. (2:21) is the deterministic criterion.

For this new problem, the objective is to determine the time functions $C_p(t)$, $D_p(t)$ and $C_e(t)$ which are effectively open-loop controls for this new game. Their role as feedback gains in the original problem means that their values are part of the final specification of the closed-loop strategies for the original game. When checking the equilibrium condition it is the original stochastic game which must be considered; this new deterministic game is only a convenient vehicle for obtaining a solution to the original.

The procedure for solving this new problem is simply the one given in Section 1.2. Adjoining (2:23) and (2:20) to (2:21) with multiplier matrices $\Lambda_1(t)$ and $\Lambda_2(t)$ respectively, the Hamiltonian, $\mathcal{H}$, for the problem can be defined.

$$\mathcal{H} \triangleq \mathrm{Tr}\{\tfrac{1}{2}Y[C_p^T R_p C_p - C_e^T R_e C_e] + \tfrac{1}{2}P[C_p^T R_p D_p + D_p^T R_p C_p$$

$$+ D_p^T R_p D_p + C_e^T R_e C_e] + \Lambda_1 \dot{Y} + \Lambda_2 \dot{P}\} \quad . \qquad (2{:}24)$$

Note that the Hamiltonian is separable; there are no cross terms between C_p and C_e, or between D_p and C_e.

As in Section 1.2, a set of necessary conditions for a saddle-point of the Hamiltonian is:

$$\frac{\partial \mathcal{H}}{\partial C_p} = 0 \quad , \qquad (2{:}25)$$

$$\frac{\partial \mathcal{H}}{\partial D_p} = 0 \quad , \qquad (2{:}26)$$

$$\frac{\partial \mathcal{H}}{\partial C_e} = 0 \quad . \qquad (2{:}27)$$

Carrying out the operations indicated, these three conditions become

$$R_p C_p Y + R_p D_p P + G_p^T(\Lambda_1 + \Lambda_1^T)Y + G_p^T(\Lambda_2 + \Lambda_2^T)P = 0 \quad , \qquad (2{:}28)$$

$$R_p(C_p + D_p)P + G_p^T(\Lambda_1 + \Lambda_1^T)P + G_p^T(\Lambda_2 + \Lambda_2^T)P = 0 \quad , \qquad (2{:}29)$$

$$-R_e C_e(Y - P) - G_e^T(\Lambda_1 + \Lambda_1^T)(Y - P) = 0 \quad . \qquad (2{:}30)$$

The solution of the two simultaneous equations (2:28) and (2:29) is

$$C_p = -R_p^{-1} G_p^T (\Lambda_1 + \Lambda_1^T) \quad , \tag{2:31}$$

$$D_p = -R_p^{-1} G_p^T (\Lambda_2 + \Lambda_2^T) \quad . \tag{2:32}$$

(2:30) directly gives

$$C_e = -R_e^{-1} G_e^T (\Lambda_1 + \Lambda_1^T) \quad . \tag{2:33}$$

It should be noted that the solution of (2:28), (2:29) and (2:30) for the control gains is possible because of the separability of the Hamiltonian. The equations for the pursuer's controls, (2:28) and (2:29), are independent of the parameters of the evader's system, i.e. G_e and R_e. Similarly, (2:30), which determines the evader's controls, is independent of G_p and R_p, the pursuer's parameters.

The Euler equations are:

$$\dot{\Lambda}_1^T = -\frac{\partial \mathcal{H}}{\partial Y} \quad , \qquad \Lambda_1(t_f) = \frac{a^2 I}{2} \quad , \tag{2:34}$$

$$\dot{\Lambda}_1^T = -\frac{\partial \mathcal{H}}{\partial P} \quad , \qquad \Lambda_2(t_f) = 0 \quad . \tag{2:35}$$

When the partial differentiations indicated in (2:34) and (2:35) are performed, these two conditions become

$$\dot{\Lambda}_1^T = -(C_p^T G_p^T - C_e^T G_e^T)\Lambda_1^T - \Lambda_1^T (G_p C_p - G_e C_e) - \frac{1}{2} C_p^T R_p C_p$$

$$+ \frac{1}{2} C_e^T R_e C_e \quad , \qquad \Lambda_1(t_f) = \frac{a^2 I}{2} \quad , \tag{2:36}$$

$$\dot{\Lambda}_2^T = [H^TQ^{-1}HP - (C_p^T + D_p^T)G_p^T]\Lambda_2 + \Lambda_2^T[PH^TQ^{-1}H - G_p(C_p + D_p)]$$

$$- [D_p^TG_p^T + C_e^TG_e^T]\Lambda_1^T - \Lambda_1^T[G_pD_p + G_eC_e] - \frac{1}{2}C_p^TR_pD_p$$

$$- \frac{1}{2}D_p^TR_pC_p - \frac{1}{2}D_p^TR_pD_p - \frac{1}{2}C_e^TR_eC_e \quad ,$$

$$\Lambda_2(t_f) = 0 \quad . \tag{2:37}$$

By defining a new set of multipliers,

$$\Gamma_1(t) \triangleq [\Lambda_1(t) + \Lambda_1^T(t)] \tag{2:38}$$

$$\Gamma_2(t) \triangleq [\Lambda_2(t) + \Lambda_2^T(t)] \tag{2:39}$$

and using (2:31), (2:32), and (2:33), the differential equations (2:36) and (2:37) can be replaced with:

$$\dot{\Gamma}_1 = \Gamma_1(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)\Gamma_1 \quad , \qquad \Gamma_1(t_f) = a^2I \, , \tag{2:40}$$

$$\dot{\Gamma}_2 = (\Gamma_2 + \Gamma_1)G_pR_p^{-1}G_p^T(\Gamma_2 + \Gamma_1) - \Gamma_1(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)\Gamma_1$$

$$+ \Gamma_2PH^TQ^{-1}H + H^TQ^{-1}HP\Gamma_2 \quad , \qquad \Gamma_2(t_f) = 0 \quad . \tag{2:41}$$

The solution to (2:40) is recognized as

$$\Gamma_1(t) = K^{-1}(t_f, t) \tag{2:42}$$

where $K(t_f, t)$ is given by (1:25). Thus, C_p and C_e are the same feedback gain matrices employed by the pursuer and the evader respectively in the deterministic problem, which was discussed in Section 1.4.

Employing (2:42) in (2:41), the differential equation for Γ_2 becomes

$$\dot{\Gamma}_2 = \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 + \Gamma_2 [G_p R_p^{-1} G_p^T K^{-1} + PH^T Q^{-1} H]$$

$$+ [K^{-1} G_p R_p G_p^T + H^T Q^{-1} HP] \Gamma_2 + K^{-1} G_e R_e^{-1} G_e^T K^{-1} \quad ,$$

$$\Gamma_2(t_f) = 0 \quad . \tag{2:43}$$

The optimal feedback strategies are:

$$U^o: \; u^o(t) = -R_p^{-1}(t) G_p^T(t) K^{-1}(t_f, t) y(t) - R_p^{-1}(t) G_p^T(t) \Gamma_2(t) \tilde{y}(t) \tag{2:44}$$

$$V^o: \; v^o(t) = -R_e^{-1}(t) G_e^T(t) K^{-1}(t_f, t) \hat{y}(t) \quad . \tag{2:45}$$

2.3 Proof of the Pursuer's Optimality

In this section, it is shown that the optimal pair, U^o and V^o, as given by (2:44) and (2:45) does indeed satisfy the inequality of (1:5a). This is accomplished by forming $J(U, V^o)$ and demonstrating that U^o minimizes this criterion.

Let v(t) in (2:1) be given by:

$$v(t) = -R_e^{-1}(t) G_e^T(t_f, t) K^{-1}(t_f, t) \xi(t) \tag{2:46}$$

where

$$\dot{\xi}(t) = [-G_p R_p^{-1} G_p^T + G_e R_e^{-1} G_e^T] K^{-1} \xi(t) + PH^T Q^{-1} [z - H\xi] \quad ,$$

$$\xi(t_o) = \hat{y}_o \tag{2:47}$$

and z(t) is given by (2:5).

It is understood that $\xi(t)$ corresponds to $\hat{y}(t)$ in the game; however, to indicate that we are solving a one-sided problem, this revised notation is used. The term, $\xi(t)$, will be an optimal estimate of $y(t)$ (based on $z(t)$) only if the pursuer employs the strategy given by (2:44); otherwise it will be meaningless. The point of this analysis is to demonstrate that if the evader determines his control $v(t)$ from $z(t)$ by using (2:46) and (2:47) -- blindly ignoring the possibility that the pursuer may not use the control given by (2:44) -- that the optimal (i.e. minimizing) strategy for the pursuer is indeed (2:44).

The pursuer can employ a different strategy and thus insure that $\xi(t)$ is not an optimal estimate of $y(t)$, but this would accomplish nothing. As is shown in this section, this type of strategy could only mean an increase in the value of the payoff. Neither player cares if what the evader calls $\hat{y}(t)$ is actually an estimate of the state; $\hat{y}(t)$ is only an intermediate variable from which an optimal control vector, $v^o(t)$, is obtained.

Now define

$$\eta(t) \triangleq y(t) - \xi(t) \tag{2:48}$$

where $\eta(t)$ in this one-sided problem would correspond to $\tilde{y}(t)$ in the game, and is assumed known by the pursuer. The differential equation which governs $\eta(t)$ can be obtained by subtracting (2:47) from (2:1); then $y(t)$ and $\eta(t)$ can be combined to produce a new state vector and a new dynamical equation for this problem.

$$\begin{bmatrix} \dot{y} \\ \dot{\eta} \end{bmatrix} = \left[\begin{array}{c|c} G_eR_e^{-1}G_e^TK^{-1} & -G_eR_e^{-1}G_e^TK^{-1} \\ \hline G_pR_p^{-1}G_p^TK^{-1} & -G_pR_p^{-1}G_p^TK^{-1} - PH^TQ^{-1}H \end{array}\right] \begin{bmatrix} y \\ \eta \end{bmatrix}$$

$$+ \left[\begin{array}{c} G_p \\ \hline G_p \end{array}\right] u + \left[\begin{array}{c} 0 \\ \hline PH^TQ^{-1} \end{array}\right] w \; , \qquad \begin{bmatrix} y(t_o) \\ \eta(t_o) \end{bmatrix} = \begin{bmatrix} y_o \\ \tilde{y}_o \end{bmatrix} \tag{2:49}$$

Using (2:46) in (2:2) and employing the new state vector, a criterion for this one-sided problem is obtained

$$J = E\left\{ \frac{1}{2} \left\| \begin{bmatrix} y(t_f) \\ \eta(t_f) \end{bmatrix} \right\|^2_{S_f} + \frac{1}{2} \int_{t_o}^{t_f} \left[\|u\|^2_{R_p} + \left\| \begin{bmatrix} y(t) \\ \eta(t) \end{bmatrix} \right\|^2_{A} \right] \right\} dt \tag{2:50}$$

where

$$S_f = \left[\begin{array}{c|c} a^2 I & 0 \\ \hline 0 & 0 \end{array}\right] \tag{2:51}$$

and

$$A = \left[\begin{array}{c} -I \\ \hline I \end{array}\right] K^{-1}G_eR_e^{-1}G_e^TK^{-1}[I \mid -I] \tag{2:52}$$

The expected value operation indicated in (2:50) is undertaken on the basis that $y(t_o)$ and $\eta(t_o)$ are known but that their future values are functions of the random variable $w(t)$.

Now from stochastic control theory (see Bryson and Ho [2]), we know that if both y(t) and $\eta(t)$ are known at time t, the control which minimizes the criterion is a feedback one given by:

$$u = -R_p^{-1}[G_p^T \vdots G_p^T]S \begin{bmatrix} y \\ \eta \end{bmatrix} \tag{2:53}$$

where

$$\dot{S} = -SF - F^T S - A + S \begin{bmatrix} G_p \\ \hdashline G_p \end{bmatrix} R_p^{-1}[G_p^T \vdots G_p^T]S \quad ,$$

$$S(t_f) = S \tag{2:54}$$

and where F(t) is the system matrix in (2:49).

Partitioning S(t),

$$S(t) = \left[\begin{array}{c:c} S_{11}(t) & S_{12}(t) \\ \hdashline S_{12}^T(t) & S_{22}(t) \end{array}\right] \tag{2:55}$$

we can rewrite the minimizing control in the form:

$$u = -R_p^{-1} G_p^T[(S_{11} + S_{12}^T)y + (S_{12} + S_{22})\eta] \quad . \tag{2:56}$$

The differential equations for $S_{11}(t)$, $S_{12}(t)$, and $S_{22}(t)$ are:

$$\begin{aligned} \dot{S}_{11} = {} & [-S_{11}G_eR_e^{-1}G_e^TK^{-1} - K^{-1}G_eR_e^{-1}G_e^TS_{11} + K^{-1}G_eR_e^{-1}G_e^TK^{-1}] \\ & + [-S_{12}G_pR_p^{-1}G_p^TK^{-1} - K^{-1}G_pR_p^{-1}G_p^TS_{12} \\ & + (S_{11} + S_{12})G_pR_p^{-1}G_p^T(S_{11} + S_{12})] \quad , \\ & S_{11}(t_f) = a^2 I \quad , \end{aligned} \tag{2:57}$$

$$\dot{S}_{12} = (S_{11} - K^{-1})G_eR_e^{-1}G_e^TK^{-1} - K^{-1}G_eR_e^{-1}G_e^TS_{12}$$

$$+ (S_{11} + S_{12} - K^{-1})G_pR_p^{-1}G_p^TS_{22} + S_{12}(G_pR_p^{-1}G_p^TK^{-1}$$

$$+ PH^TQ^{-1}H) + (S_{11} + S_{12})G_pR_p^{-1}G_p^TS_{12} \quad ,$$

$$S_{12}(t_f) = 0 \quad , \tag{2:58}$$

$$\dot{S}_{22} = S_{22}[G_pR_p^{-1}G_p^TK^{-1} + PH^TQ^{-1}H] + [K^{-1}G_pR_p^{-1}G_p^T$$

$$+ H^TQ^{-1}HP]S_{22} + S_{12}^TG_eR_e^{-1}G_e^TK^{-1} + K^{-1}G_eR_e^{-1}G_e^TS_{12}$$

$$+ K^{-1}G_eR_e^{-1}G_e^TK^{-1} + (S_{22} + S_{12}^T)G_pR_p^{-1}G_p^T(S_{22} + S_{12}) \quad ,$$

$$S_{22}(t_f) = 0 \quad . \tag{2:59}$$

Now substitute

$$S_{11}(t) = K^{-1}(t_f, t) \quad , \tag{2:60}$$

$$S_{12}(t) = 0 \quad , \tag{2:61}$$

and

$$S_{22}(t) = \Gamma_2(t) \tag{2:62}$$

into (2:57), (2:58) and (2:59). Then (2:57) reduces to (2:40) the differential equation for $K^{-1}(t_f, t)$, (2:58) becomes identically zero, and (2:59) reduces to (2:43) the differential equation for $\Gamma_2(t)$. Thus (2:60), (2:61) and (2:62) are indeed the solutions to (2:57), (2:58) and (2:59) respectively. Q. E. D.

2.4 Proof of the Evader's Optimality

Assuming that the pursuer employs the strategy given by (2:44) the evader's strategy -- based on this knowledge -- will be determined. This will be shown to be the same feedback control as given in (2:45), thus proving that (2:44) and (2:45) satisfy the inequality of (1:5b).

Since a basic assumption has been that the pursuer is given the error of the evader's estimate, the implication is that the evader does calculate such an estimate. The evader is, of course, under no such obligation. To avoid this difficulty, let us assume that the mystical third party -- which was previously the vehicle for observing $\hat{y}(t)$ and transfering this information to the pursuer -- now observes the measurement, z(t). It is from this measurement that the evader would calculate an estimate if he desired. This third party can then construct the same filter that the evader employed in Section 2.2 above, and transfer this result to the pursuer.

Thus assume that the pursuer is given a $\hat{y}_p(t)$, calculated from the measurement z(t).

$$\dot{\hat{y}}_p(t) = -[G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T]K^{-1}\hat{y}_p + PH^TQ^{-1}[z - H\hat{y}_p(t)] \quad ,$$

$$\hat{y}_p(t_o) = \hat{y}_o \quad . \tag{2:63}$$

Now employing this "estimate," let the pursuer's control be

$$u(t) = -R_p^{-1}G_p^TK^{-1}y - R_p^{-1}G_p^T\Gamma_2(y - \hat{y}_p) \quad . \tag{2:64}$$

Then, since z(t) is given by (2:5), we can combine (2:63) and (2:1) using (2:64) to obtain an overall system equation with which the evader will have to contend.

$$\begin{bmatrix} \dot{y} \\ \dot{\hat{y}}_p \end{bmatrix} = \mathcal{F} \begin{bmatrix} y \\ \hat{y}_p \end{bmatrix} - \begin{bmatrix} G_e \\ \hdashline 0 \end{bmatrix} v + \begin{bmatrix} 0 \\ \hdashline PH^T Q^{-1} \end{bmatrix} w \tag{2:65}$$

where

$$\mathcal{F} = \left[\begin{array}{c:c} -G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) & G_p R_p^{-1} G_p^T \Gamma_2 \\ \hdashline PH^T Q^{-1} H & -[G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T] K^{-1} - PH^T Q^{-1} H \end{array}\right] . \tag{2:66}$$

The criterion to be <u>minimized</u> by the evader can be obtained by using (2:64) in (2:2) and can be written in terms of the new state vector, $\begin{bmatrix} y(t) \\ \hat{y}_p(t) \end{bmatrix}$.

$$-J = E\left\{ \frac{1}{2} \left\| \begin{bmatrix} y(t_f) \\ \hat{y}_p(t_f) \end{bmatrix} \right\|^2_{S_f} + \frac{1}{2} \int_{t_o}^{t_f} \left[\left\| \begin{bmatrix} y(t) \\ \hat{y}_p(t) \end{bmatrix} \right\|^2_{A(t)} + \| v(t) \|^2_{R_e} \right] dt \right\} \tag{2:67}$$

where

$$S_f = \left[\begin{array}{c:c} -a^2 I & 0 \\ \hdashline 0 & 0 \end{array}\right] , \tag{2:68}$$

and

$$A(t) = \left[\begin{array}{c|c} -(K^{-1} + \Gamma_2)G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) & +(K^{-1} + \Gamma_2)G_p R_p^{-1} G_p^T \Gamma_2 \\ \hline +\Gamma_2 G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) & -\Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 \end{array}\right] \tag{2:69}$$

The certainty-equivalence principle is a well known result (see Gunckel and Franklin [3]) of stochastic control theory. The principle is usually proven for discrete-time problems. However, the continuous analog can be obtained by letting the time interval approach zero. When this is done it is seen that a sufficient condition for certainty-equivalence to be applicable is that the control weighting matrix (in this case R_e) be positive-definite. Since R_e is positive-definite, the evader's optimal control to <u>minimize</u> (2:67) is a feedback control operating on the evader's optimal estimate of the state $\left[\begin{array}{c} y(t) \\ \hline \hat{y}_p(t) \end{array}\right]$. The feedback gain matrix is given by:

$$C_e = -R_e^{-1}[G_e^T \,\vdots\, 0]S(t) \tag{2:70}$$

where S(t) is obtained from

$$\dot{S}(t) = -S\mathcal{F} - \mathcal{F}^T S - A - S\left[\begin{array}{c} G_e \\ \hline 0 \end{array}\right] R_e^{-1}[G_e^T \,\vdots\, 0]S \quad ,$$

$$S(t_f) = S_f \quad . \tag{2:71}$$

Of course, any solution which employs the certainty-equivalence principle is dependent on the fact that S(t) remains finite.

Now it should be noted that the matrix S(t) can be partitioned

$$S \equiv \left[\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{12}^T & S_{22} \end{array}\right] \quad , \tag{2:72}$$

that since the evader knows z(t) he can reproduce $\hat{y}_p(t)$ exactly and that he can obtain an estimate of y(t) -- denoted $\hat{y}(t)$ -- from

$$\dot{\hat{y}}(t) = -G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)\hat{y} + G_pR_p^{-1}G_p^T\Gamma_2\hat{y}_p - G_ev$$
$$+ PH^TQ^{-1}(z - H\hat{y}) \quad , \qquad \hat{y}(t_o) = \hat{y}_o \quad . \tag{2:73}$$

Thus the evader's optimal control is given by

$$v(t) = -R_e^{-1}G_e^T(S_{11}\hat{y} + S_{12}\hat{y}_p) \quad . \tag{2:74}$$

The differential equations governing S are:

$$\dot{S}_{11} = S_{11}G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2) - S_{12}PH^TQ^{-1}H$$
$$- (K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2 - S_{11}) - H^TQ^{-1}HPS_{12}^T$$
$$- S_{11}G_eR_e^{-1}G_e^TS_{11} \quad , \qquad S_{11}(t_f) = a^2I \quad , \tag{2:75}$$

$$\dot{S}_{12} = -S_{11}G_pR_p^{-1}G_p^T\Gamma_2 + S_{12}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}$$
$$+ S_{12}PH^TQ^{-1}H + (K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T(S_{12} + \Gamma_2)$$
$$- H^TQ^{-1}HPS_{22} - S_{11}G_eR_e^{-1}G_e^TS_{12} \quad ,$$

$$S_{12}(t_f) = 0 \quad , \tag{2:76}$$

$$\dot{S}_{22} = -(S_{12}^T + \Gamma_2)G_pR_p^{-1}G_p^T\Gamma_2 + S_{22}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}$$

$$+ S_{22}PH^TQ^{-1}H - \Gamma_2G_pR_p^{-1}G_p^TS_{12} + K^{-1}(G_pR_p^{-1}G_p^T$$

$$- G_eR_e^{-1}G_e^T)S_{22} + H^TQ^{-1}HPS_{22} - S_{12}G_eR_e^{-1}G_e^TS_{12} \quad ,$$

$$S_{22}(t_f) = 0 \quad . \tag{2:77}$$

By adding (2:75) and (2:76) and also adding (2:76) and (2:77) it can be seen that

$$S_{11}(t) + S_{12}(t) = K^{-1}(t_f, t) \tag{2:78}$$

and that

$$S_{12}(t) + S_{22}(t) = 0 \quad . \tag{2:79}$$

When (2:78) is inserted into (2:74) the evader's control becomes

$$v(t) = -R_e^{-1}G_e^T[K^{-1}\hat{y} + S_{12}(\hat{y} - \hat{y}_p)] \quad . \tag{2:80}$$

But, using (2:80) in (2:73) and subtracting the result from (2:63), we find that

$$\hat{y}(t) \equiv \hat{y}_p(t) \quad , \qquad t_o \leqslant t \leqslant t_f \quad . \tag{2:81}$$

Thus the evader's strategy is indeed

$$v(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)\hat{y}(t) \quad . \tag{2;82}$$

Q. E. D.

In Chapter 3, a method is given whereby the pursuer can calculate the error of the evader's estimate, $\tilde{y}(t)$. Thus, the pursuer can directly obtain the information from which -- on the basis of his optimal strategy -- he calculates his optimal control vector, u(t). He is then no longer dependent upon the mystical third party. Since the above proof is dependent upon the method by which the pursuer obtains the information concerning $\tilde{y}(t)$, the inequality of (1:5b) may not be satisfied under these new conditions. Consequently, a second proof of the evader's optimality is given in Section 3.3, based on this method for directly calculating $\tilde{y}(t)$.

2.5 The Criterion

Having obtained the optimal strategies, a numerical value can be assigned to the criterion, based on the values of the system parameters and the initial conditions. When the optimal values of C_p, D_p and C_e are substituted into the expression for the criterion as given in (2:2), we obtain:

$$J = \mathrm{Tr}\{\frac{a^2}{2} Y(t_f) + \frac{1}{2}\int_{t_o}^{t_f} [(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}YK^{-1}$$

$$+ G_pR_p^{-1}G_p^T(K^{-1}P\Gamma_2 + \Gamma_2PK^{-1} + \Gamma_2P\Gamma_2)$$

$$+ G_eR_e^{-1}G_e^TK^{-1}PK^{-1}]dt\} \quad . \tag{2:83}$$

This expression can be written in terms of the initial values of Y(t) and P(t).

Consider the identically zero integral (see (2:23))

$$0 = \int_{t_o}^{t_f} K^{-1}[\dot{Y} + G_pR_p^{-1}G_p^T(K^{-1}Y + \Gamma_2P) + (YK^{-1} + P\Gamma_2)G_pR_p^{-1}G_p^T$$

$$- G_eR_e^{-1}G_e^TK^{-1}(Y - P) - (Y - P)K^{-1}G_eR_e^{-1}G_e^T]dt \quad . \tag{2:84}$$

Integrating the first term by parts, and employing (2:40) for the value of $\frac{d}{dt}K^{-1}$ (2:84) becomes

$$0 = a^2Y(t_f) - K^{-1}(t_f, t_o)Y(t_o) + \int_{t_o}^{t_f} [K^{-1}YK^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)$$

$$+ K^{-1}G_eR_e^{-1}G_e^TK^{-1}P + K^{-1}PK^{-1}G_eR_e^{-1}G_e^T$$

$$+ K^{-1}G_pR_p^{-1}G_p^T\Gamma_2P + K^{-1}P\Gamma_2G_pR_p^{-1}G_p^T]dt \quad . \qquad (2:85)$$

Consider also the identically zero integral (see (2:20))

$$0 = \int_{t_o}^{t_f} \Gamma_2[\dot{P} + G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)P + P(K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T$$

$$+ PH^TQ^{-1}HP]dt \quad . \qquad (2:86)$$

Again integrating by parts and employing (2:43), this becomes

$$0 = -\Gamma_2(t_o)P(t_o) + \int_{t_o}^{t_f} [-K^{-1}G_eR_e^{-1}G_e^TK^{-1}P - H^TQ^{-1}HP\Gamma_2P$$

$$+ \Gamma_2P\Gamma_2G_pR_p^{-1}G_p^T + \Gamma_2PK^{-1}G_pR_p^{-1}G_p^T$$

$$- K^{-1}G_pR_p^{-1}G_p^T\Gamma_2P]dt \quad . \qquad (2:87)$$

Taking the trace of (2:85) and (2:87), and using these results in (2:83), a more significant expression for the optimal criterion -- in terms of the system and initial parameters -- is obtained.

$$J = \mathrm{Tr}[\tfrac{1}{2}K^{-1}(t_f, t_o)Y(t_o) + \tfrac{1}{2}\Gamma_2(t_o)P(t_o) + \tfrac{1}{2}\int_{t_o}^{t_f} PH^TQ^{-1}HP\Gamma_2 dt] \quad . \qquad (2{:}88)$$

The first term in this expression for J corresponds directly to the total expression for J in the deterministic game, given by (1:37). The second term accounts for the initial uncertainty of the evader's estimate of the state. The integral term accounts for the uncertainty of the evader's estimate during the play of the game, which is known to exist even prior to the actual play.

Thus (2:88) is of fundamental significance for it directly relates the outcome of the stochastic game to the deterministic outcome and to the quality of the information available to the evader. The nature of the evader's information is expressed by $P(t_o)$, and $H(t)$ and $Q(t)$ for $t_o \leq t \leq t_f$; $P(t)$ for $t > t_o$, and $\Gamma_2(t_o)$ are derived quantities.

The differential equation for Γ_2, (2:43), can be viewed as a linear matrix differential equation, with $\Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 + K^{-1} G_e R_e^{-1} G_e^T K^{-1}$ as the driving term. Since R_p^{-1} and R_e^{-1} are both positive-definite, the driving expression is positive-semi-definite. Furthermore, the terminal value of Γ_2 is zero, and since Γ_2 can only become less negative-definite as time progresses (because of the positive nature of the driving term) Γ_2 must be negative-semi-definite for all $t < t_f$.

Now recall that the trace of the product of two positive-definite matrices is always positive (see [3]). Since $P(t_o)$ is positive-definite, $PH^TQ^{-1}HP$ positive-semi-definite and Γ_2 negative-semi-definite, and since all are symmetric, the second and third terms in (2:88) are non-positive. (They are negative if $H^TQ^{-1}H$, $G_pR_p^{-1}G_p^T$ and $G_eR_e^{-1}G_e^T$ are positive definite.) This reflects the decrease in capability of the evader -- the maximizing player.

2.6 A Summary of the Optimized System

For the linear-quadratic differential game formulated in Section 2.1, optimal strategies have been obtained in terms of feedback control laws, which are repeated here.

$$u^o(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f,t)y(t) - R_p^{-1}(t)G_p^T(t)\Gamma_2(t)\tilde{y}(t) \qquad (2{:}44)$$

$$v^o(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f,t)\hat{y}(t) \quad . \qquad (2{:}45)$$

When these are inserted in the system equation (2:1), this differential equation becomes

$$\dot{y} = -G_pR_p^{-1}G_p^TK^{-1}y - G_pR_p^{-1}G_p^T\Gamma_2\tilde{y} + G_eR_e^{-1}G_e^TK^{-1}\hat{y} \quad . \qquad (2{:}89)$$

This can be rewritten using (2:15) exclusively in terms of either y(t) and $\hat{y}(t)$, or y(t) and $\tilde{y}(t)$.

$$\dot{y} = -G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)y + [G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1}]\hat{y} \qquad (2{:}90)$$

$$\dot{y} = -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}y - [G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1}]\tilde{y} \, . \qquad (2{:}91)$$

The inputs to these equations results from either

$$\dot{\hat{y}} = -\{[G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T]K^{-1} + PH^TQ^{-1}H\}\hat{y} + PH^TQ^{-1}Hy + PH^TQ^{-1}w(t) \quad , \qquad (2{:}92)$$

or

$$\dot{\tilde{y}} = -[G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2) + PH^TQ^{-1}H]\tilde{y} - PH^TQ^{-1}w(t) \, . \qquad (2{:}93)$$

Note that (2:93) is self-contained, i.e. it is not coupled to other equations for it is only driven by the noise w(t).

Thus, the entire play of the game can be described by either one of two 2k-dimensional differential equations, which are obtained from either (2:90) and (2:92), or (2:91) and (2:93).

$$\begin{bmatrix} \dot{y} \\ \dot{\hat{y}} \end{bmatrix} = \left[\begin{array}{c|c} -G_pR_p^{-1}G_p^T(K^{-1}+\Gamma_2) & G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1} \\ \hline PH^TQ^{-1}H & -[G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T]K^{-1} - PH^TQ^{-1}H \end{array}\right] \begin{bmatrix} y \\ \hat{y} \end{bmatrix}$$

$$+ \left[\begin{array}{c} 0 \\ \hline PH^TQ^{-1} \end{array}\right] w(t) \quad , \qquad (2:94)$$

$$\begin{bmatrix} \dot{y} \\ \dot{\tilde{y}} \end{bmatrix} = \left[\begin{array}{c|c} -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1} & -G_pR_p^{-1}G_p^T\Gamma_2 - G_eR_e^{-1}G_e^TK^{-1} \\ \hline 0 & -G_pR_p^{-1}G_p^T(K^{-1}+\Gamma_2) - PH^TQ^{-1}H \end{array}\right] \begin{bmatrix} y \\ \tilde{y} \end{bmatrix}$$

$$+ \left[\begin{array}{c} 0 \\ \hline PH^TQ^{-1} \end{array}\right] w \quad . \qquad (2:95)$$

Note that as mentioned in Section 1.5, the white noise, w(t) -- though it is additive measurement noise to the evader -- is process noise to the optimized system as a whole.

At t_o, predictions of the future state of the game are determined by two (k by k)-dimensional matrix differential equations which are obtained from (2:23) and (2:20) by inserting the optimal values of C_p, D_p, and C_e.

$$\dot{Y}(t) = -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}Y(t) - Y(t)K^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)$$

$$- (G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1})P(t)$$

$$- P(t)(\Gamma_2G_pR_p^{-1}G_p^T + K^{-1}G_eR_e^{-1}G_e^T) \quad ,$$

$$Y(t_o) = Y_o \tag{2:96}$$

$$\dot{P}(t) = -G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)P(t) - P(t)(K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T$$

$$- P(t)H^TQ^{-1}HP(t) \quad , \qquad P(t_o) = P_o \quad . \tag{2:97}$$

Note that P(t) can be calculated independently of the value of Y(t).

Again, the optimized value of the criterion is given by

$$J = Tr[\tfrac{1}{2}K^{-1}(t_f, t_o)Y(t_o) + \tfrac{1}{2}\Gamma_2(t_o)P(t_o)$$

$$+ \frac{1}{2}\int_{t_o}^{t_f} PH^TQ^{-1}HP\Gamma_2\, dt] \quad . \tag{2:88}$$

References for Chapter Two

[1] R. E. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction," Journal of Basic Engineering, 1961, Vol. 83D, Transactions of the ASME, pp. 95-108.

[2] A. E. Bryson, Jr. and Y. C. Ho, Optimization, Estimation and Control, Waltham, Massachusetts, Blaisdell Company, to be published, Chapter 14.

[3] T. F. Gunckel and G. F. Franklin, "A General Solution for Linear Sampled-Data Control Systems," Journal of Basic Engineering, 1963, Vol. 85D, No. 2, Transactions of the ASME, pp. 197-203.

[4] R. Bellman, Introduction to Matrix Analysis, New York, McGraw-Hill Book Company, Inc., 1960, p. 99.

BLANK PAGE

CHAPTER THREE

REALIZATION OF THE PURSUER'S STRATEGY

3.1 Two Realization Schemes

The solution of Chapter 2 is dependent upon the ability of the pursuer to obtain exact knowledge of the error of the evader's estimate. Since the pursuer is assumed only to have perfect knowledge of the state, $y(t)$, and no knowledge of the evader's estimate, $\hat{y}(t)$, the pursuer cannot in general obtain this information. There are certain conditions, however, which insure that $\hat{y}(t)$ can be obtained directly from $y(t)$.

Substituting (2:44) into (2:1) gives

$$-G_e(t)R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)\hat{y}(t) = G_p(t)u(t) - \dot{y}(t) \quad . \tag{3:1}$$

Now, since the pursuer knows $G_p(t)$, $u(t)$, and $y(t)$ perfectly, he can -- by differentiating $y(t)$ with respect to time -- calculate the right-hand side of (3:1). The evader's estimate (and consequently the error of this estimate) can be calculated by the pursuer, if the inverse of $G_e(t)R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)$ exists.

Two necessary conditions for the existence of this inverse are:

i. The number of "interesting" state variables (i.e. the dimension of $y(t)$) must be less than or equal to the number of the evader's control variables (i.e. the dimension of $v(t)$), and

ii. The inverse of $G_e(t)R_e^{-1}(t)G_e^T(t)$ must exist for all values of $t < t_f$.

Note, condition ii. cannot be fulfilled unless condition i. is; also, the inverse of $R_e^{-1}(t)$ and $K^{-1}(t_f, t)$ must exist. Consequently, condition ii. itself is both a sufficient condition for the pursuer with a known u(t) to calculate $\tilde{y}(t)$.

Condition i. is useful, however, for it immediately restricts the class of problems to which the solution obtained above is applicable, without considering the nature of $G_e(t)$ in detail. For instance, the rendezvous game in n-dimensional euclidean space is not solved by the above, since the number of "interesting" state variables is 2n, while the number of controls is n. The class of problems to which a solution has been obtained does include, however, the important interception problem in n-dimensional euclidean space; here both the number of "interesting" state variables and the number of controls are n.

The difficulties encountered when the pursuer cannot calculate perfectly the error of the evader's estimate -- either because there are more interesting state variables than the evader has control variables, or because the pursuer does not himself have perfect knowledge of the state -- are discussed in Chapter 8.

Figure 3-1 displays a flow-chart for implementing the optimal strategies of the two players. The method shown for the pursuer to calculate u(t) is termed Realization I. The pursuer differentiates y(t) and feeds back u(t) to calculate $\hat{y}(t)$; then subtracting this from y(t), $\tilde{y}(t)$ is obtained.

Examination of the pursuer's controller for Realization I indicates an additional difficulty: the feedback loop, inside the controller itself,

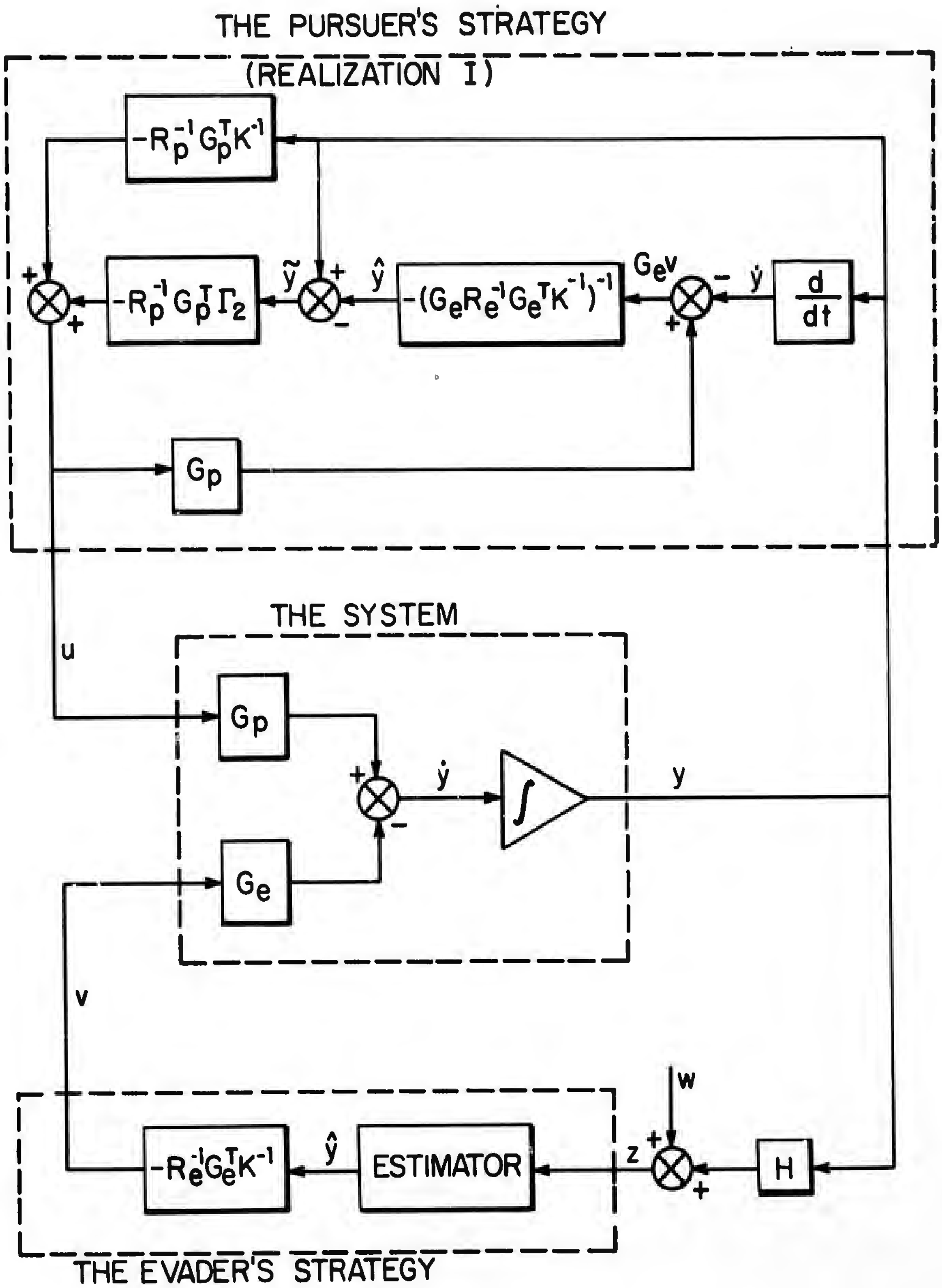


FIG. 3-1 IMPLEMENTATION OF THE OPTIMAL STRATEGIES

may be unstable. Adding the inputs which arrive at the summing junction which produces u(t) gives

$$u(t) = -R_p^{-1}G_p^T\Gamma_2(G_eR_e^{-1}G_e^TK^{-1})^{-1}G_pu(t) - R_p^{-1}G_p^T(\Gamma_2 + K^{-1})y$$

$$+ R_p^{-1}G_p^T\Gamma_2(G_eR_e^{-1}G_e^TK^{-1})^{-1}\dot{y}(t) \quad . \tag{3:2}$$

Solving this for u(t) produces

$$u(t) = M^{-1}(t)R_p^{-1}G_p^T[\Gamma_2(G_eR_e^{-1}G_e^TK^{-1})^{-1}\dot{y}(t) - (\Gamma_2 + K^{-1})y(t)] \quad , \tag{3:3}$$

where

$$M(t) = I + R_p^{-1}G_p^T\Gamma_2(G_eR_e^{-1}G_e^TK^{-1})^{-1}G_p \quad . \tag{3:4}$$

Actually, (3:3) can also be obtained by solving (2:44) and (3:1) for u(t) in terms of y(t) and $\dot{y}(t)$, using (2:15) to eliminate $\hat{y}(t)$ and $\tilde{y}(t)$.

The condition that ensures that the feedback loop internal to the pursuer's controller of Realization I be stable also ensures that M^{-1} exists. Consider the general feedback network of Figure 3-2, with vector input i(t) and vector output o(t). The input-output relation of this system is given by

$$o(t) = A(t)B(t)o(t) + A(t)i(t) \quad . \tag{3:5}$$

Observe that no dynamics are involved; the system is merely a feedback amplifier. For the scalar case, if A(t)B(t) is positive there is positive feedback.

For any realization of the feedback loop of Figure 3-2, the actual input-output relation will be

$$o(t + \delta) = A(t)B(t)o(t) + A(t)i(t) \tag{3:6}$$

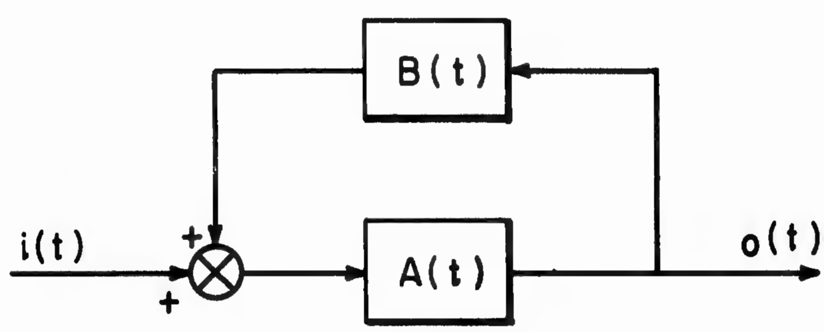


FIG. 3-2 A GENERAL FEEDBACK SYSTEM

for δ small but non-zero. Here the condition for stability is that all the eigenvalues, λ, of AB have magnitudes less than 1.0. (If any eigenvalue of AB does become +1.0, then the inverse of [I - AB] will certainly fail to be finite.) Thus it is necessary to check not only the existence of M^{-1}, but also the nature of the eigenvalues of $-R_p^{-1}G_p^T\Gamma_2(G_eR_e^{-1}G_e^TK^{-1})^{-1}G_p$ (the AB for the pursuer's internal feedback loop) to ensure stability.

Now note that $-R_p^{-1}G_p^T\Gamma_2(G_eR_e^{-1}G_e^TK^{-1})^{-1}G_p$ has a terminal value of zero. Thus when viewing time as progressing in a negative direction from the terminal time, M^{-1} will certainly be finite and the eigenvalues of AB will certainly have magnitudes less than one for at least some non-zero time interval.

Another expression for u(t) in terms of y(t) and $\dot{y}(t)$ can be obtained directly from the game system equation as given in (2:91). Solving this for $\tilde{y}(t)$ gives

$$\tilde{y}(t) = -L^{-1}(t)[G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T]K^{-1}y - L^{-1}(t)\dot{y}(t) \tag{3:7}$$

where

$$L(t) \triangleq G_p(t)R_p^{-1}(t)G_p^T(t)\Gamma_2(t) + G_e(t)R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t) \quad . \tag{3:8}$$

When (3:7) is inserted into (2:44), the pursuer's control is given by

$$u(t) = -R_p^{-1}G_p^T[I - \Gamma_2L^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)]K^{-1}y(t)$$

$$+ R_p^{-1}G_p^T\Gamma_2L^{-1}\dot{y}(t) \quad . \tag{3:9}$$

Thus $u^o(t)$ can be written as a feedback strategy based on $y(t)$ and $\dot{y}(t)$ multiplied by feedback gain matrices G_1 (Gain 1) and G_2 (Gain 2) respectively.

$$u^o(t) = G_1(t)y(t) + G_2(t)\dot{y}(t) \tag{3:10}$$

where G_1 can be given either by

$$G_1 = -R_p^{-1}G_p^T[I - \Gamma_2 L^{-1}(G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T)]K^{-1} \tag{3:11}$$

or

$$G_1 = -M^{-1}R_p^{-1}G_p^T(\Gamma_2 + K^{-1}) \quad , \tag{3:12}$$

and where G_2 can be given by either

$$G_2 = R_p^{-1}G_p^T\Gamma_2 L^{-1} \tag{3:13}$$

or

$$G_2 = M^{-1}R_p^{-1}G_p^T\Gamma_2(G_e R_e^{-1} G_e^T K^{-1})^{-1} \quad . \tag{3:14}$$

Thus (3:10) can be used to reduce the visual complexity of the pursuer's controller as shown in Figure 3-3. This is referred to as the pursuer's Realization II. If the inverse of G_p exists, it can be demonstrated by simple matrix manipulation that the two expressions for G_1, (3:11) and (3:12), are identical, as are the two expressions for G_2, (3:13) and (3:14). Consequently, the inverse of $L(t)$ will fail to exist at the same time that the inverse of $M(t)$ fails to exist; thus either (3:11) or (3:12) and either (3:13) or (3:14) can be employed. In further discussion, the use of (3:11) and (3:13) will be assumed.

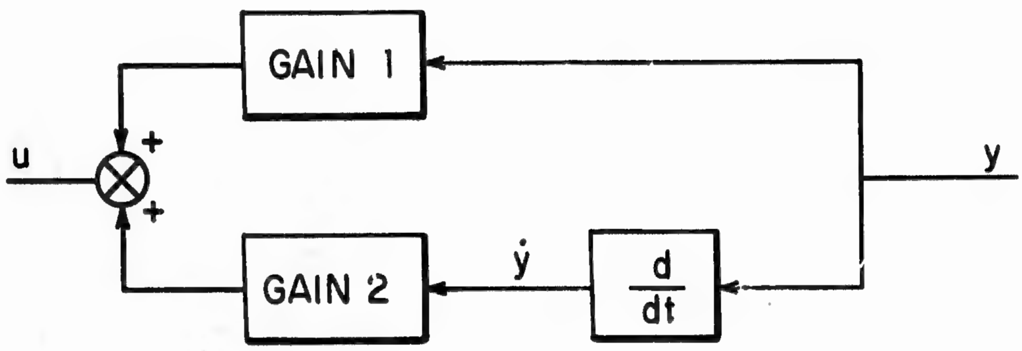


FIG. 3-3 REALIZATION II FOR THE PURSUER

Note that for Realization II to produce a finite output u(t) for finite inputs y(t) and $\dot{y}(t)$ it is only necessary that $L^{-1}(t)$ exist; there is no feedback loop and thus no stability requirement. The terminal value of L(t) is $a^2 G_e(t_f) R_e^{-1}(t_f) G_e^T(t_f)$ which is positive-definite and which certainly has an inverse. Thus when viewing time as progressing backwards from t_f, L(t) will certainly have an inverse for some finite time interval.

In conclusion then, a third condition,

iii. The inverse of L(t) must exist for all values of $t < t_f$,

is needed in addition to ii. to insure that the pursuer can implement the strategy obtained in Chapter 2. In other words, ii. is sufficient for the pursuer to be able to calculate $\hat{y}(t)$ (or $\tilde{y}(t)$) when he employs a u(t) independent of $\tilde{y}(t)$. But iii. is required to ensure that this information can actually be employed in realizing the pursuer's optimal strategy as defined by (2:44). This is true since $u^o(t)$ is a function of $\tilde{y}(t)$, and thus $u^o(t)$ cannot be used to solve for $\tilde{y}(t)$ without some extra restrictions.

3.2. Reflections on the Pursuer's Strategy

It should be pointed out that the above method of obtaining $\hat{y}(t)$ is similar to that of determining the input to a linear system by observing the output. In fact, the input v(t) can be determined exactly by differentiating the total system output, i.e. the vector y(t). This is not the question, however; what is needed is $\hat{y}(t)$ or $\tilde{y}(t)$. Unfortunately -- as can be seen from (2:45) -- $\hat{y}(t)$ cannot be obtained from v(t) unless condition ii. of Section 3.1 is satisfied.

This condition is necessary because it is possible that more than one value of $\hat{y}(t)$ -- in fact an infinite number -- can produce the same v(t). No further differentiating of the output can resolve this.

It is important to note why v(t) itself is not sufficient information upon which to base an optimal pursuit strategy. Certainly, v(t) is sufficient for the pursuer to determine the evader's deviation from his deterministically optimal control at time t. However, it is not the instantaneous control error which is important; if suddenly the evader is given perfect knowledge of the state so that this error can be corrected for all time greater than the current t, the result (from then on) will be identical to the deterministic game, despite the error at time t. What is significant is the fact that the pursuer knows that the evader's control will also deviate from the deterministic optimum <u>in the future</u> and it is this fact which permits the pursuer to make use of the error at time t.

This does not mean that the pursuer makes an explicit prediction of the evader's estimation error at future times. The pursuer does, however, assess the nature of this error at future times by employing P (the variance of the error) when calculating the control gain to be employed in conjunction with this error. This results from the fact that $\Gamma_2(t)$ is obtained by integrating (2:43) backwards from t_f, and that $P(\tau)$, $t \leq \tau \leq t_f$, is a factor in this integration.

The equation for the estimation error is given by (2:93). At time t, the pursuer could make a prediction of the evader's future error by dropping the white noise term, $PH^TQ^{-1}w$. Then using $\tilde{y}(t)$ as the initial condition, he could integrate this new equation forward in time. Note that to predict future deviations of the evader's control, it is necessary to integrate a k-dimensional differential equation; this requires k initial conditions. Therefore, v(t) must be k-dimensional (condition i.) and must not be confined to a manifold of less than k dimensions (condition ii.).

The thought of calculating the value of $\dot{y}(t)$ by differentiating y(t) and then employing it in a feedback control exactly at time t may, to some, seem farfetched. Since $\dot{y}(t)$ will probably be obtained from y(t) and $y(t - \delta t)$ by the approximation formula

$$\dot{y}(t) \approx \frac{y(t) - y(t - \delta t)}{\delta t} , \tag{3:15}$$

it would appear that the number used in the actual control would be closer to $\dot{y}(t - \delta t/2)$. Thus there would be a short delay in implementing this term in the feedback control. The obvious question is whether the actual outcome of the pursuit-evasion game is sensitive to this delay.

In Appendix A of this chapter, a discrete-time, pursuit-evasion game is considered. Here, the pursuer's control is based on the state at the <u>current</u> time interval, i, and the evader's estimation error at the <u>previous</u> time interval, i - 1. The control gain matrix used in conjunction with this delayed estimation error is, however, the one which is optimal for use with the estimation error at time t. It is shown that as the time interval approaches zero the expression for the criterion (found under such conditions) approaches the expression for the criterion obtained for the continuous problem in Section 2.5. The conclusion is that, by making δt in the approximation calculation (3:15) small enough, the pursuer can ensure a value of the criterion as close to that given by (2:88) as he desires.

By rewriting (3:7), an interesting observation about the pursuer's stochastic strategy can be made.

$$u(t) = -R_p^{-1}G_p^TK^{-1}y(t) + R_p^{-1}G_p^T\Gamma_2L^{-1}[(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}y(t) + \dot{y}(t)] \quad . \tag{3:16}$$

Recall that $-(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}y$ is the system's acceleration for the deterministic game when both players use their optimal strategies. See (1:36). Consequently it can be seen from (3:16) that the term added to the pursuer's deterministic feedback control law to form his stochastic one is proportional to deviations from the deterministically optimal acceleration, i.e.

$$u(t) = -R_p^{-1}G_p^TK^{-1}y(t) - R_p^{-1}G_p^T\Gamma_2L^{-1}[\dot{y}^{do}(t) - \dot{y}(t)] \tag{3:17}$$

where $\dot{y}^{do}(t)$ is the deterministically optimal acceleration. Consequently, the second term in the pursuer's stochastic strategy takes advantage of the evader's deviation from his deterministically optimal acceleration, which results from the evader's error in estimating the state y(t).

The procedure of feeding back $\dot{y}(t)$ along with y(t) is <u>not</u> common in control theory. For the general dynamic system of Figure 3-4, the input-output relation is given by

$$\dot{o}(t) = A(t)o(t) + B(t)i(t) \quad . \tag{3:18}$$

A usual control approach is to make the input some linear function of the output.

$$i(t) = C(t)o(t) \quad . \tag{3:19}$$

As the above realizations indicate, however, the solution obtained here also feeds back the derivative of the output.

$$i(t) = C(t)o(t) + D(t)\dot{o}(t) \quad . \tag{3:20}$$

Several observers have raised the point that this approach could, when inserted into the physical system supplied by nature, produce meaningless results. This would be the case if the gain in either one of the feedback loops (there are essentially two, the y loop and the $\dot{y}$ loop) was the inverse of the mathematical function seen when looking into the system from the two ends of the gain. From an electrical engineering viewpoint, the question is: "Have we decided to construct a feedback gain whose impedance is the negative of the impedance of the system when viewed from the two nodes between which this feedback gain is inserted?"

Inserting (3:20) into (3:18) gives

$$\dot{o}(t) = [A(t) + C(t)]o(t) + B(t)D(t)\dot{o}(t) \quad . \tag{3:21}$$

Then if D(t) was selected such that

$$D(t) = B^{-1}(t) \tag{3:22}$$

the result would be the absurdity of forcing the output to be identically zero. Note further that if

$$A(t) + C(t) \equiv 0 \quad , \tag{3:23}$$

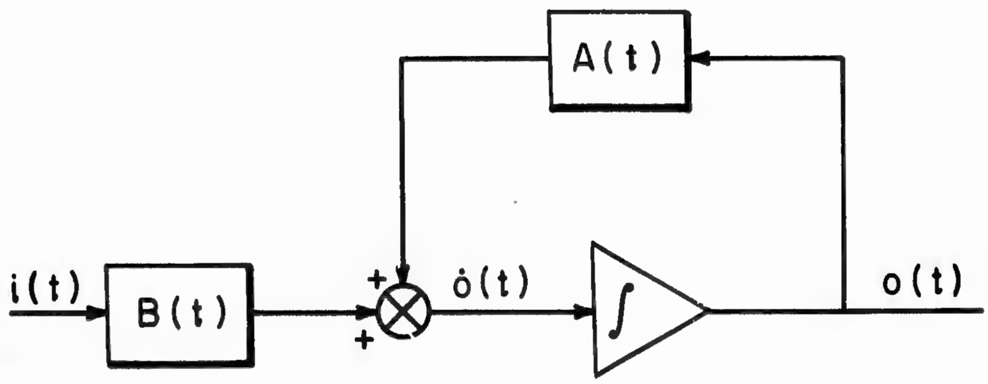


FIG. 3-4 A GENERAL DYNAMIC SYSTEM

the output could take on any value, with the actual one depending on the eccentricities of the physical system, whether it be nature or an analog computer.

It is obvious from (3:18) and (3:19) that this difficulty cannot be a consequence of ordinary feedback control based on the value of the state vector. The above solution, however, employs a feedback control based on the value of the time derivative of the state vector; the possibility of this difficulty must therefore be investigated.

Figure 3-5 shows the entire optimized system as viewed from the nodes of the pursuer's second feedback loop (the one containing G_2) where the input is i(t) and the output y(t). The dynamic equation for this system is

$$\begin{bmatrix} \dot{y} \\ \dot{\hat{y}} \end{bmatrix} = \left[\begin{array}{c:c} G_p G_1 & G_e R_e^{-1} G_e^T K^{-1} \\ \hdashline -PH^T Q^{-1} H & -(G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T) K^{-1} + PH^T Q^{-1} H \end{array}\right] \begin{bmatrix} y \\ \hat{y} \end{bmatrix}$$

$$+ \left[\begin{array}{c} G_p \\ \hdashline 0 \end{array}\right] i + \left[\begin{array}{c} 0 \\ \hdashline PH^T Q^{-1} \end{array}\right] w \quad . \qquad (3{:}24)$$

But we have not selected $G_2(t)$ to be $G_p^{-1}(t)$, see (3:13), and thus the absurdity -- which would have occurred if we had indeed attempted a feedback control of the form given by (3:22) -- does not exist.

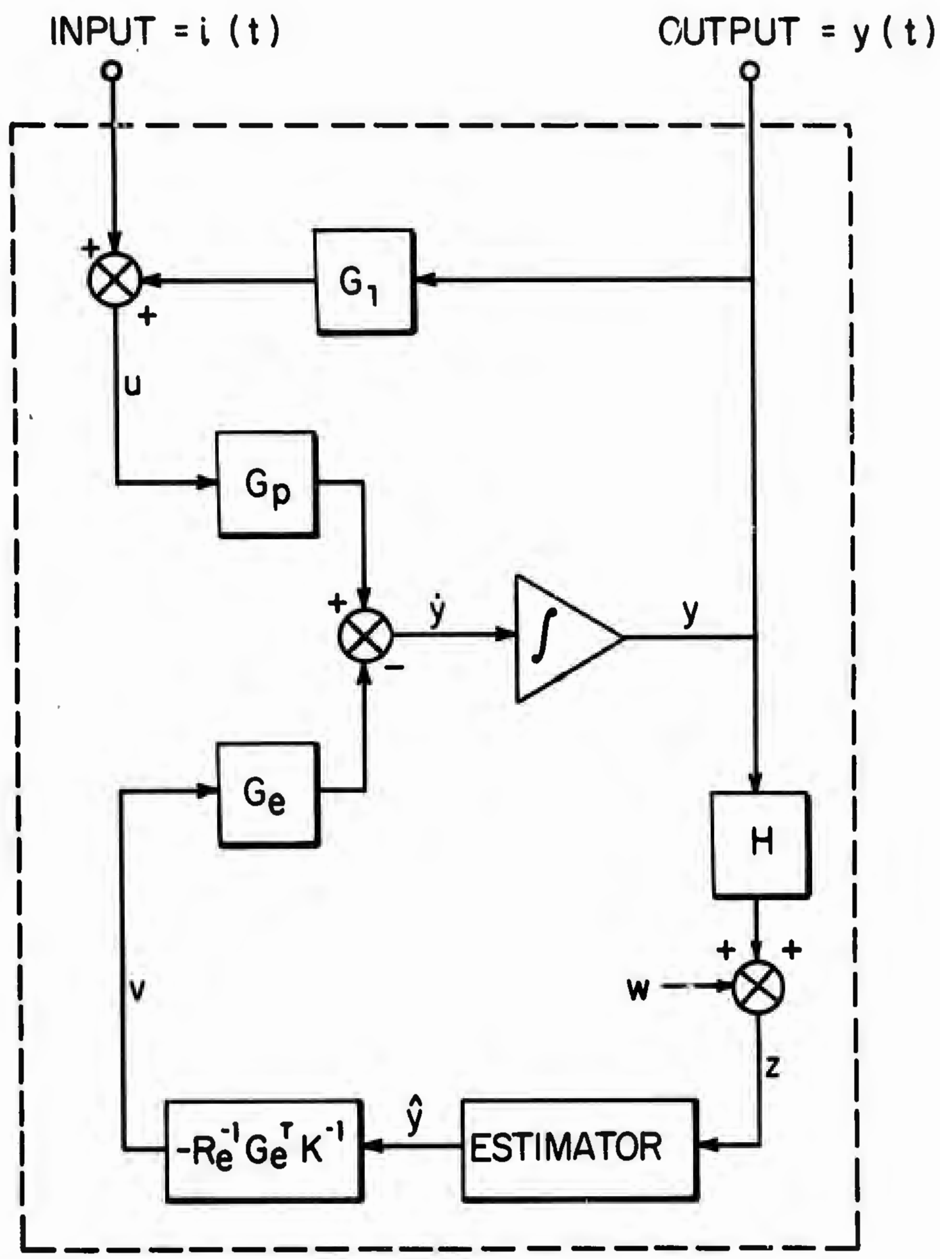


FIG. 3-5 THE OPTIMIZED SYSTEM AS VIEWED FROM THE NODES OF THE PURSUER'S SECOND FEEDBACK LOOP.

3.3 Another Proof of the Evader's Optimality

As pointed out in the previous section, the essence of the pursuer's strategy is to take advantage of his opponent's deviations from the deterministically optimal control acceleration. The realization schemes presented are two methods for calculating and utilizing this deviation. Since, the evader knows that the pursuer will be using one of these schemes -- rather than receiving $\hat{y}_p(t)$ from the mystical third party of Section 2.4 -- it may be possible for the evader to take advantage of the pursuer's computational technique to fool the pursuer.

Consequently, it is necessary to again prove the evader's optimality by ascertaining that indeed his strategy satisfies the left-hand inequality of (1:4) when U^o implicitly involves one of the realizations discussed above. The proof below is based on the assumption that Realization I is used by the pursuer. Since from a blackbox, input-output point of view, the two realizations are identical, this proof is sufficient for both realizations.

Let u(t) in (2:1) be given by:

$$u(t) = -R_p^{-1} G_p^T [K^{-1} y + \Gamma_2 \eta] \tag{3:25}$$

where $\eta(t)$ is obtained from v(t) in the following manner.

$$\eta(t) = y(t) + [G_e(t) R_e^{-1}(t) G_e^T(t) K^{-1}(t_f, t)]^{-1} G_e(t) v(t) \quad . \tag{3:26}$$

$G_e(t)v(t)$ appears explicitly in the pursuer's calculation of his optimal control for Realization I, see Figure 3-1. Observe that $\eta(t)$ of this one-sided problem corresponds to the error $\tilde{y}(t)$ in the game, and that the inverse in (3:26) does not exist unless condition ii. is satisfied.

Using (3:25) and (3:26) in (2:1), the system equation for this one-sided control problem is

$$\dot{y} = -G_pR_p^{-1}G_p^T[K^{-1} + \Gamma_2]y + [-G_pR_p^{-1}G_p^T\Gamma_2K(G_e^T)^{-1}R_e - G_e]v \ ,$$

$$y(t_o) = y_o \ . \tag{3:27}$$

The criterion for this problem is obtained by using (3:25) and (3:26) in (2:2).

$$-J = E\left\{\frac{-a^2}{2}\|y(t_f)\|^2 + \frac{1}{2}\int_{t_o}^{t_f}[y^T \mid v^T]\left[\begin{array}{c|c} A & N \\ \hline N^T & B \end{array}\right]\begin{bmatrix} y \\ v \end{bmatrix} dt\right\} \tag{3:28}$$

where

$$A = -(K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2) \ , \tag{3:29}$$

$$N = -(K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T\Gamma_2K(G_e^T)^{-1}R_e \ , \tag{3:30}$$

$$B = R_e - R_eG_e^{-1}K\Gamma_2G_pR_p^{-1}G_p^T\Gamma_2K(G_e^T)^{-1}R_e \ . \tag{3:31}$$

Using the measurement of the state obtained from (2:5), the evader can obtain an estimate of the state using a Kalman-Bucy Filter.

$$\dot{\hat{y}} = -G_pR_p^{-1}G_p^T[K^{-1} + \Gamma_2]\hat{y} + [-G_pR_p^{-1}G_p^T\Gamma_2K(G_e^T)^{-1}R_e - G_e]v$$

$$+ PH^TQ^{-1}(z - H\hat{y}) \ , \qquad \hat{y}(t_o) = \hat{y}_o \tag{3:32}$$

where P(t) is the variance of the error of this estimate but is given by the same equation as used in Chapter 2.

$$\dot{P} = -G_p R_p^{-1} G_p^T [K^{-1} + \Gamma_2] P - P[K^{-1} + \Gamma_2] G_p R_p^{-1} G_p^T$$

$$-PH^T Q^{-1} HP \quad , \qquad P(t_o) = P_o \quad . \tag{3:33}$$

The evader's control which minimizes (3:28) is obtained by using deterministic control theory (see for example Bryson and Ho [1]) in conjunction with the certainty-equivalence principle.

$$v \quad = -B^{-1}(N^T + \mathcal{G}^T S)\hat{y} \tag{3:34}$$

$$S(t) = -S(F - \mathcal{G}B^{-1}N^T) - (F - \mathcal{G}B^{-1}N^T)^T S - (A - NB^{-1}N^T)$$

$$+ S\mathcal{G}B^{-1}\mathcal{G}^T S \quad , \qquad S(t_f) = -a^2 I \quad . \tag{3:35}$$

In (3:34) and (3:35), F and $\mathcal{G}$ are, respectively, the system matrix and control matrix in the dynamic equation (3:27). Again in order for the certainty-equivalence principle to be applicable, B must be positive-definite, and S(t) must remain finite.

B consists of two terms. The first one, R_e, results from the evader's control weight. The second and negative one $[-R_e G_e^{-1} K \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 K (G_e^T)^{-1} R_e]$ results from the pursuer's control weight. If the pursuer attempts to overcompensate for the evader's estimation error, the evader can take advantage of this and do better by heading in a "non-optimal" direction to throw the pursuer off the track. This, in short, is the meaning of the extra condition that B be positive-definite. Efforts to prove that this is always true, or to relate B to L have been unsuccessful.

It can be shown that (see Appendix B of this chapter)

$$B^{-1}(N^T + \mathcal{B}^T S) = R_e^{-1} G_e^T K^{-1} \quad . \tag{3:36}$$

Q. E. D.

If the evader does not employ the feedback strategy given in (2:45), $\eta(t)$ as obtained from (3:26) will not equal the error in the evader's estimate. However, calculating the error in the evader's estimate is not -- for the pursuer -- an end in itself, but rather merely a step in obtaining the optimal control. Thus, if the evader deviates from the optimal strategy, the value of the criterion will be reduced, despite the fact that the pursuer will miscalculate the error of the evader's estimate.

3.4 The Meaning and Existence of L^{-1}

The question of the existence of L^{-1} is an intriguing problem, and still, unfortunately, a baffling one. As was pointed out in condition iii. of Section 3.1, the pursuer is capable of implementing his optimal strategy only if this inverse does exist. Consequently, it is desirable to characterize this matrix as best as possible. For convenience, the definition of L is repeated.

$$L = G_p R_p^{-1} G_p \Gamma_2 + G_e R_e^{-1} G_e K^{-1} \quad . \tag{3:8}$$

It should be first noted that L appears explicitly in the optimized system equations, see Section 2.6. In particular L(t) is the gain matrix for the driving term of the game equation, regardless of whether $\hat{y}(t)$ or $\tilde{y}(t)$ is the input vector, i.e. see both (2:90) and (2:91).

A statement which is equivalent to, "L^{-1} fails to exist" is that "one of the eigenvalues of L is zero," with this eigenvalue corresponding to a particular (non-zero) eigenvector. Consequently, if L^{-1} fails to exist there will be a certain value of the error vector, $\tilde{y}$ (i.e. this particular eigenvector) which will not effect the state. In other words, at the same time when L^{-1} fails to exist, a certain value of $\tilde{y}(t)$ can make the driving term $L(t)\tilde{y}(t)$ identically zero.

The possibility of this event occurring is obviously insignificant; however, there is a more general implication. Any particular value of the error vector $\tilde{y}(t)$ can be decomposed into a unique linear combination of the various eigenvectors of L(t). This linear combination can be calculated by observing $L(t)\tilde{y}(t)$, which is available to the pursuer. However, when L(t) has a zero eigenvalue, the magnitude of the associated eigenvector cannot be determined, since this eigenvector does not effect the value of the vector $L(t)\tilde{y}(t)$ which is being observed.

Now recall that the pursuer can take advantage of deviations of the evader from his deterministically optimal control, not because the error at time t results in an irrevocable loss to the evader, but rather because it permits the pursuer to implicitly predict and take advantage of future deviations. Since the particular eigenvector in question does not effect the derivative of the state, it does not, in itself, produce an immediate deviation from the evader's optimal control. Its magnitude will, however, effect future estimation errors and consequently future deviations by the evader from his optimal control. Thus the requirement that L(t) have non-zero eigenvalues.

Since the terminal value of $\Gamma_2(t)$ is fixed at zero, the terminal value of L(t) is given by

$$L(t_f) = a^2 G_e(t_f) R_e^{-1}(t_f) G_e^T(t_f) \qquad (3:37)$$

which, unless $G_e(t_f)$ is zero, is positive-definite. Consequently, the existence of $L^{-1}(t)$ is guaranteed for $t_f - t_o$ sufficiently small. In this respect, condition iii. is similar to a conjugate point condition; if the time duration of the problem is short enough, the condition is satisfied.

Since $G_p R_p^{-1} G_p^T$, $G_e R_e^{-1} G_e^T$, and K^{-1} are positive-definite while Γ_2 is negative-definite, L(t) is the difference of two products of two positive-definite matrices each. Furthermore, the first three can be determined exclusively from the deterministic properties of the game. Only Γ_2 is dependent upon the stochastic character of the problem, and for Γ_2 sufficiently small, the condition is satisfied.

Not much more can be said, unfortunately, about the general nature of this condition. The differential equation for L(t) can be obtained from (2:40) and (2:43).

$$\begin{aligned}
\dot{L}(t) = {} & L(G_p R_p^{-1} G_p^T K^{-1}) + (G_p R_p^{-1} G_p^T K^{-1}) L + G_p R_p^{-1} G_p^T \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 \\
& - G_e R_e^{-1} G_e^T K^{-1} G_e R_e^{-1} G_e^T K^{-1} + G_p R_p^{-1} G_p^T [\Gamma_2 P H^T Q^{-1} H \\
& + H^T Q^{-1} H P \Gamma_2] + \frac{d[G_p R_p^{-1} G_p^T]}{dt} \Gamma_2 + \frac{d[G_e R_e^{-1} G_e^T]}{dt} K^{-1} ,
\end{aligned}$$

$$L(t_f) = a^2 G_e(t_f) R_e^{-1}(t_f) G_e^T(t_f) \ . \qquad (3:38)$$

Note immediately that the values of $\frac{d}{dt}[G_pR_p^{-1}G_p^T]$ and $\frac{d}{dt}[G_eR_e^{-1}G_e^T]$ can be such as to force one of the eigenvalues of L to zero, for there are no general restrictions on these terms. Even the relative controllability condition (1:38) represents only an integral constraint on $G_pR_p^{-1}G_p^T$ and $G_eR_e^{-1}G_e^T$; it does not effect their values or derivatives at any particular time and consequently adds no insight to the question of L^{-1}.

Consider now the special situation where

$$G_p(t)R_p^{-1}(t)G_p^T(t) = f(t)T_p \quad , \tag{3:39}$$

$$G_e(t)R_e^{-1}(t)G_e^T(t) = f(t)T_e \tag{3:40}$$

where T_p and T_e are independent of time. T_p and T_e can be assumed positive-definite, thus requiring that

$$f(t) \geqslant 0 \quad , \tag{3:41}$$

without adding further restrictions. Note that the relative controllability condition now requires that

$$T_p > T_e \quad . \tag{3:42}$$

The condition that L^{-1} exist reduces to the condition that $[T_p\Gamma_2 + T_eK^{-1}]^{-1}$ exists, and that (3:41) be a strict inequality.

Given (3:39) and (3:40), the differential equations for K^{-1} and Γ_2 can be rewritten.

$$\frac{d}{dt}[K^{-1}] = f(t)K^{-1}[T_p - T_e]K^{-1} \quad , \qquad K^{-1}(t_f) = a^2I \tag{3:43}$$

$$\dot{\Gamma}_2 = f(t)\Gamma_2 T_p \Gamma_2 + \Gamma_2[f(t)T_p K^{-1} + PH^T Q^{-1} H]$$

$$+ [f(t)K^{-1}T_p + H^T Q^{-1} HP]\Gamma_2 + f(t)K^{-1}T_e K^{-1} \quad ,$$

$$\Gamma_2(t_f) = 0 \quad . \tag{3:44}$$

From (3:43) and (3:44), the differential equation for the reduced L,

$$L_r \triangleq T_p \Gamma_2 + T_e K^{-1} \quad , \tag{3:45}$$

can be obtained.

$$\dot{L}_r = f(t)L_r\{\tfrac{1}{2}T_p(K^{-1} + \Gamma_2) + \tfrac{1}{2}(T_p - T_e)K^{-1}\}$$

$$+ f(t)\{\tfrac{1}{2}T_p(K^{-1} + \Gamma_2) + \tfrac{1}{2}(T_p - T_e)K^{-1}\}L_r$$

$$+ T_p(\Gamma_2 PH^T Q^{-1} H + H^T Q^{-1} HP\Gamma_2) \quad ,$$

$$L_r(t_f) = a^2 T_e \quad . \tag{3:46}$$

There still is no general statement which can be made about the existence of L_r^{-1}.

Consider the scalar case. (3:46) reduces to

$$\dot{L}_{r_s} = f'(t)L_{r_s} + d(t) \quad , \qquad L_{r_s}(t_f) = a^2 T_e \quad . \tag{3:47}$$

Because T_p, P, $H^T Q^{-1}$ are all positive-definite and Γ_2 is negative-definite, d(t) in (3:47) is negative. Consequently, when integrating (3:47) backwards from t_f, the driving term, d(t), is positive, forcing L_{r_s} to be larger and ensuring that

$$L_{r_s} > 0 \quad , \tag{3:48}$$

and thus that L^{-1} exists for the scalar problem where the time dependencies of all elements of $G_p R_p^{-1} G_p^T$ and $G_e R_e^{-1} G_e^T$ are identical.

The generalization of this result to the matrix case cannot be made and consequently the existence of L^{-1} must be immediately checked for each problem. Yet the proof for the scalar case is in itself significant. The two-dimensional example of Chapter 7 is reduced to a scalar one where these conditions, that ensure the existence of L^{-1}, are met. This proof then prevents the necessity of calculating $L(t)$ explicitly for such scalar cases.

APPENDIX III-A

THE DISCRETE-TIME GAME

The purpose of this appendix is to demonstrate that it is indeed computationally possible for the pursuer to determine $\dot{y}(t)$ and to then employ it in his feedback control so as to nearly obtain the value of the criterion given by (2:88). This is accomplished by solving a discrete-time problem consisting of the following three steps.

1. A discrete-time, deterministic game is solved.
2. A discrete-time, stochastic game is solved employing the gains already obtained in 1., adding only a new correction gain to the pursuer's control at time i based on the evader's estimation error at time i. The evader obtains his estimate from a discrete-time Kalman-Bucy filter.
3. The controls obtained in 2. are used, except that the pursuer's correction gain is applied to the evader's estimation error at time i - 1. This procedure is employed because it is always possible for the pursuer to calculate at time i, the evader's estimation error at time i - 1. The criterion is evaluated for this case and shown to approach the optimized, continuous-time criterion of Section 2.5 as the time interval for this discrete game approaches zero.

The aim of this appendix is not to completely solve the discrete-time stochastic game, complete with proof of the saddle-point condition. What follows below is sketchy in parts, though sufficient detail is provided so that the missing mathematical steps can be easily filled in. This mathematical austerity is partially justified on the grounds that the discrete-time and continuous-time games are very similar.

The real justification, however, is that the objective is merely to show that there exists for the pursuer some control which employs slightly delayed data (be it $\dot{y}$ or $\tilde{y}$) and that this slight error in the optimal control does not produce large deviations from the optimized value of the criterion. Thus the method by which this control is obtained need not be rigorous; the existence of such a control is what is important.

It is shown below that slightly delayed values of $\tilde{y}(t)$ can be employed without drastic losses. Since $\tilde{y}(t)$ is linearly proportional to the deviation of the acceleration vector $\dot{y}(t)$ from its deterministically optimal value (see (3:17)), this is equivalent to demonstrating that slightly delayed values of $\dot{y}(t)$ are not disasterous.

A. 1 The Deterministic Problem

Consider the difference equation

$$x(i+1) = x(i) + G_p(i)u(i) - G_e(i)v(i) \quad , \qquad i = 0, 1, \ldots, n-1 \quad , \tag{3:A1}$$

where x(i), is an n-dimensional state vector, and u(i) and v(i) are respectively the pursuer's and evader's m and m' dimensional control vectors. $G_p(i)$ and $G_e(i)$ are the m x n and m' x n control matrices. The criterion for this game is given by

$$J = \frac{a^2}{2} \|x(n)\|^2 + \frac{1}{2} \sum_{i=0}^{n-1} [\|u(i)\|^2_{R_p(i)} - \|v(i)\|^2_{R_e(i)}] \quad . \tag{3:A2}$$

Let

$$u(i) = C_p(i)x(i) \quad , \tag{3:A3}$$

$$v(i) = C_e(i)x(i) \tag{3:A4}$$

be the feedback form of the controls. Inserting (3:A3) and (3:A4) into (3:A1) gives

$$x(i+1) = [I + G_p(i)C_p(i) - G_e(i)C_e(i)]x(i) \quad ,$$

$$i = 0, 1, \ldots, n-1 \quad . \tag{3:A5}$$

This problem is to be solved by the method that must be applied to the stochastic game. Consequently, define

$$X(i) = x(i)x^T(i) \quad . \tag{3:A6}$$

Using (3:A5) and (3:A6) a difference relation which defines the propagation of X(i) can be obtained.

$$X(i+1) = [I + G_p(i)C_p(i) - G_e(i)C_e(i)]X(i)[I + C_p^T(i)G_p^T(i)$$

$$- C_e^T(i)G_e^T(i)] \quad , \qquad i = 0, 1, \ldots, n-1 \quad . \tag{3:A7}$$

(3:A6) in conjunction with (3:A3) and (3:A4) can be used in (3:A2) to give another expression for the criterion.

$$J = \mathrm{Tr}\{\frac{a^2}{2} X(n) + \frac{1}{2}\sum_{i=0}^{n-1} X(i)[C_p^T(i)R_p(i)C_p(i) - C_e^T(i)R_e(i)C_e(i)]\} \quad . \tag{3:A8}$$

Defining the Hamiltonian for this problem*

---

* In all cases the argument of all matrices and vectors is assumed to be i, unless otherwise indicated. All equations are assumed valid for $i = 0, 1, \ldots, n-1$ unless otherwise indicated.

$$\mathcal{H}(i) \triangleq \mathrm{Tr}\{\tfrac{1}{2} X(i)(C_p^T R_p C_p - C_e^T R_e C_e)$$

$$+ \Lambda^T(i+1)[I + G_p C_p - G_e C_e] X(i)[I + C_p^T G_p^T - C_e^T G_e^T]\} \qquad (3{:}A9)$$

two necessary conditions for stationarity of J can be easily found.

$$\frac{\partial \mathcal{H}(i)}{\partial C_p(i)} = 0 \quad , \qquad (3{:}A10)$$

$$\frac{\partial \mathcal{H}(i)}{\partial C_e(i)} = 0 \quad . \qquad (3{:}A11)$$

Carrying out the partial differentiation indicated in (3:A10) gives

$$0 = R_p C_p X + G_p^T \Lambda(i+1) G_p C_p X + G_p^T \Lambda^T(i+1) G_p C_p X$$

$$+ G_p^T \Lambda(i+1)[I - G_e C_e] X + G_p^T \Lambda^T(i+1)[I - G_e C_e] X \ . \qquad (3{:}A12)$$

Now certainly, if

$$0 = R_p C_p + G_p^T [\Lambda(i+1) + \Lambda^T(i+1)][I + G_p C_p - G_e C_e] \qquad (3{:}A13)$$

is true, (3:A12) is satisfied. From (3:A11), a similar procedure gives

$$0 = -R_e C_e - G_e^T [\Lambda(i+1) + \Lambda^T(i+1)][I + G_p C_p - G_e C_e] \ . \qquad (3{:}A14)$$

Another necessary condition for the stationarity of J is

$$\Lambda(i) = \frac{\partial \mathcal{H}(i)}{\partial X(i)} \quad , \qquad \Lambda^T(n) = \frac{a^2}{2} I \quad , \qquad (3{:}A15)$$

which, after the partial differentiation is taken, becomes

$$\Lambda(i) = \tfrac{1}{2}[C_p^T R_p C_p - C_e^T R_e C_e]$$

$$+ [I + C_p^T G_p^T - C_e^T G_e^T] \Lambda(i+1) [I + G_p C_p - G_e C_e] \ . \qquad (3{:}A16)$$

Since, it can be seen from (3:A16) that $\Lambda(i)$ and $\Lambda^T(i)$ are obtained from the same equation, $\Lambda(i)$ is symmetrical.

$$\Lambda^T(i) = \Lambda(i) \quad . \tag{3:A17}$$

Thus (3:A13) and (3:A14) become respectively

$$0 = \frac{1}{2}R_pC_p + G_p^T\Lambda(i+1)[I + G_pC_p - G_eC_e] \quad , \tag{3:A18}$$

$$0 = \frac{1}{2}R_eC_e + G_e^T\Lambda(i+1)[I + G_pC_p - G_eC_e] \quad . \tag{3:A19}$$

By premultiplying (3:A18) by C_p^T and subtracting this from (3:A16), and by premultiplying (3:A19) by C_e^T and subtracting this also, a simple difference equation for $\Lambda(i)$ is obtained.

$$\Lambda(i) = \Lambda(i+1)[I + G_pC_p - G_eC_e] \quad , \qquad \Lambda(n) = \frac{a^2}{2}I \quad . \tag{3:A20}$$

Using (3:A20) in (3:A18) and (3:A19) the optimized feedback control gains are given by

$$C_p(i) = -R_p^{-1}(i)G_p^T(i)\Gamma_1(i) \quad , \tag{3:A21}$$

$$C_e(i) = -R_e^{-1}(i)G_e^T(i)\Gamma_1(i) \quad , \tag{3:A22}$$

where

$$\Gamma_1(i) \triangleq 2\Lambda(i) \quad . \tag{3:A23}$$

Then, using (3:A21), (3:A22) and (3:A23) in (3:A20), a difference equation for $\Gamma_1(i)$ is obtained.

$$\Gamma_1(i) = \{\Gamma_1^{-1}(i+1) + [G_p(i)R_p^{-1}(i)G_p^T(i) - G_e(i)R_e^{-1}(i)G_e^T(i)]\}^{-1} \quad ,$$

$$\Gamma(n) = a^2I \quad . \tag{3:A24}$$

(3:A21), (3:A22) and (3:A24) define feedback control gains for the discrete-time game. These results are the exact analog of those obtained for the continuous-time case, see Section 1.4. It can, in fact be shown, that as the time interval involved in the difference equation (3:A24) approaches zero, that $\Gamma_1(i)$ approaches $K^{-1}(t_f, t)$.

A.2 The Stochastic Problem

The same difference equation applies here, (3:A1), but now the evader has an initial estimate of the state x(0), plus a series of measurements of the state which are corrupted by additive noise.

$$z(i) = H(i)x(i) + w(i) \quad , \qquad i = 1, 2, \ldots, n-1 \tag{3:A25}$$

where w(i) is Gaussian white noise with zero mean and variance Q(i).

Let the control vectors be of the form

$$u(i) = -R_p^{-1}(i)G_p^T(i)\Gamma_1(i)x(i) - R_p^{-1}(i)G_p^T(i)\Delta(i)\tilde{x}(i) \quad , \tag{3:A26}$$

$$v(i) = -R_e^{-1}(i)G_e^T(i)\Gamma_1(i)\hat{x}(i) \quad , \tag{3:A27}$$

where $\tilde{x}(i)$ is the error of the evader's estimate $\hat{x}(i)$, and $\Gamma_1(i)$ is given by (3:A24), or alternatively by

$$\Gamma_1(i) = \Gamma_1(i+1)\Psi(i) \tag{3:A28}$$

where

$$\Psi(i) = I - [G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T]\Gamma_1(i) \quad . \tag{3:A29}$$

Since $\Gamma_1(i)$ is predetermined, the only necessary optimization of the criterion

$$J = E\left[\frac{a^2}{2}\|x(n)\|^2 + \frac{1}{2}\sum_{i=0}^{n-1}\{\|u(i)\|^2_{R_p(i)} - \|v(i)\|^2_{R_e(i)}\}\right] \tag{3:A30}$$

takes place over the matrix $\Delta(i)$.

Using (3:A26) and (3:A27) in (3:A1) gives

$$x(i+1) = \Psi(i)x(i) - \{G_pR_p^{-1}G_p^T\Delta + G_eR_e^{-1}G_e^T\Gamma_1\}\tilde{x}(i) \quad . \tag{3:A31}$$

The Kalman-Bucy estimation equation for the evader's estimate is given by

$$\hat{x}(i+1) = \Psi(i)\hat{x}(i) + P(i+1)H^T(i+1)Q^{-1}(i+1)[z(i+1)$$

$$- H(i+1)\Psi(i)\hat{x}(i)] \quad , \qquad i = 1, 2, \ldots, n-1 \tag{3:A32}$$

where

$$P(i) \triangleq E[\tilde{x}(i)\tilde{x}^T(i)] \quad . \tag{3:A33}$$

Subtracting (3:A32) from (3:A33), and noting (from (3:A31)) that

$$x(i+1) - \Psi(i)\hat{x}(i) = [\Psi(i) - (G_pR_p^{-1}G_p^T\Delta - G_eR_e^{-1}G_e^T\Gamma_1)]\tilde{x}(i) \tag{3:A34}$$

the difference equation for $\tilde{x}(i)$ can be obtained.

$$\tilde{x}(i+1) = [I - P(i+1)H^T(i+1)Q^{-1}(i+1)H(i+1)]\Phi(i)\tilde{x}(i)$$

$$-P(i+1)H^T(i+1)Q^{-1}(i+1)H(i+1)w(i+1) \quad ,$$

$$i = 0, 1, \ldots, n-2 \tag{3:A35}$$

where

$$\Phi(i) \triangleq I - G_p(i)R_p^{-1}(i)G_p^T(i)[\Gamma_1(i) + \Delta(i)] = \Psi(i) - G_pR_p^{-1}G_p^T\Delta$$

$$- G_eR_e^{-1}G_e^T\Gamma_1] \quad . \tag{3:A36}$$

Postmultiplying (3:A36) by its transpose and taking expectations it can be shown that

$$P^{-1}(i+1) = \Phi^{T^{-1}}(i)P^{-1}(i)\Phi^{-1}(i) + H^T(i+1)Q^{-1}(i+1)H(i+1) \quad . \tag{3:A37}$$

Also, defining

$$X(i) = E\{x(i)x^T(i)\} \tag{3:A38}$$

and then postmultiplying (3:A37) by its transpose and again taking expectations gives

$$X(i+1) = \Psi(i)[X(i) - P(i)]\Psi^T(i) + \Phi(i)P(i)\Phi^T(i) \quad . \tag{3:A39}$$

Employing (3:A26) and (3:A27) in the criterion as given by (3:A30) and taking the expectation indicated gives

$$J = \mathrm{Tr}\{\frac{a^2}{2}X(n) + \frac{1}{2}\sum_{i=0}^{n-1}[x\Gamma_1 G_p R_p^{-1} G_p^T \Gamma_1 + P(\Gamma_1^T G_p R_p^{-1} G_p^T \Delta$$
$$+ \Delta^T G_p R_p^{-1} G_p^T \Gamma_1 + \Delta^T G_p R_p^{-1} G_p^T \Delta)$$
$$- (x - P)\Gamma_1 G_e R_e^{-1} G_e^T \Gamma_1]\} \quad . \tag{3:A40}$$

Then adjoining (3:A39) to (3:A40) with $\Lambda_1^T(i+1)$ and also adjoining (3:A37) with $\Lambda_2^T(i+1)$, a Hamiltonian can be defined.

$$\mathcal{H}(i) \triangleq \mathrm{Tr}\{\frac{1}{2}x(i)\Gamma_1(i)[G_p(i)R_p^{-1}(i)G_p^T(i) - G_e(i)R_e^{-1}(i)G_e^T(i)]\Gamma_1(i)$$
$$+ \frac{1}{2}P(i)[(\Gamma_1(i) + \Delta^T(i))G_p(i)R_p^{-1}(i)G_p^T(i)(\Gamma_1(i) + \Delta(i))$$
$$- \Gamma_1(i)(G_p(i)R_p^{-1}(i)G_p^T(i) - G_e(i)R_e^{-1}(i)G_e^T(i))\Gamma_1(i)]$$
$$+ \Lambda_1^T(i+1)[\Psi(i)(X(i) - P(i))\Psi^T(i) + \Phi(i)P(i)\Phi^T(i)]$$
$$+ \Lambda_2^T(i+1)[\Phi^{T^{-1}}(i)P(i)\Phi^{-1}(i) + H^T(i+1)Q^{-1}(i+1)H(i+1)]\} \quad . \tag{3:A41}$$

Note that the state variables are X(i) and $P^{-1}(i)$, and the control variable is $\Delta(i)$.

Now necessary conditions for the stationarity of the Hamiltonian are

$$\Lambda_1(i) = \frac{\partial \mathcal{H}(i)}{\partial X(i)} \quad , \qquad \Lambda_1(n) = \frac{a^2}{2} I \quad , \tag{3:A42}$$

$$\Lambda_2(i) = \frac{\partial \mathcal{H}}{\partial P^{-1}(i)} \quad , \qquad \Lambda_2(n) = 0 \quad , \tag{3:A43}$$

$$\frac{\partial \mathcal{H}}{\Delta(i)} = 0 \quad . \tag{3:A44}$$

Carrying out the partial differentiation indicated in (3:A42) gives

$$\Lambda(i) = \frac{1}{2}\,\Gamma_1(i)[G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T]\Gamma_1(i) + \Psi^T(i)\Lambda_1(i+1)\Psi(i) \quad . \tag{3:A45}$$

It can be shown by induction that

$$\Lambda_1(i) = \frac{1}{2}\Gamma_1(i) \quad . \tag{3:A46}$$

First assume that in (3:A45) $\Lambda_1(i+1)$ is given by $\frac{1}{2}\Gamma_1(i+1)$; then using (3:A29), (3:A45) reduces to (3:A46). Thus since both $\Lambda_1(n)$ and $\frac{1}{2}\Gamma_1(n)$ are $\frac{a^2}{2}I$, the induction proof is complete.

Carrying out the partial differentiation of (3:A43) gives

$$\Gamma_2(i) = \Phi^{-1}(i)\Gamma_2(i+1)\Phi^{T^{-1}}(i) + P(i)\{\Gamma_1 - [\Gamma_1 + \Delta]G_p R_p^{-1} G_p^T[\Gamma_1 + \Delta]$$

$$- \Phi^T\Gamma_1(i+1)\Phi\}P(i) \tag{3:A47}$$

where

$$\Gamma_2(i) = \Lambda_2(i) + \Lambda_2^T(i) \quad . \tag{3:A48}$$

Finally, taking the differentiation indicated in (3:A44) gives -- after extensive algebraic manipulation --

$$\Delta(i) = -P^{-1}(i)\Gamma_2(i)P^{-1}(i) \tag{3:A49}$$

or

$$\Gamma_2(i) = -P(i)\Delta(i)P(i) \quad . \tag{3:A50}$$

Using (3:A50) and (3:A37) in (3:A47) gives

$$\Delta(i + 1) = [Z^T(i + 1)]^{-1}\{\Delta(i)\Psi^{-1}(i) + \Gamma_1(i)[\Phi^{-1}(i) - \Psi^{-1}(i)]\}Z^{-1}(i + 1) ,$$

$$\Delta(n) = 0 \tag{3:A51}$$

where

$$Z(i + 1) \triangleq I - P(i + 1)H^T(i + 1)Q^{-1}(i + 1)H(i + 1) \quad . \tag{3:A52}$$

Again, observe the similarity between the results obtained here for the discrete-time game and the results obtained in Section 2.2 for the one in continuous-time. In particular, if the time interval between i and $i+1$ is allowed to approach zero, as n approaches infinity it can be shown that $\Delta(i)$ approaches $\Gamma_2(t)$ as given by (2:43).

A.3 A Non-Optimal, Discrete-Time Game

Again, the system difference equation shall be given by (3:A1) and the evader's control $v(i)$ by (3:A27). However, let the pursuer's control be given by

$$u(i) = -R_p^{-1}(i)G_p^T(i)[\Gamma_1(i)x(i) + \Delta(i)\tilde{x}(i - 1)] \quad ,$$

$$i = 1, 2, \ldots, n - 1 \tag{3:A53}$$

and

$$u(0) = -R_p^{-1}(0)G_p^T(0)\Gamma_1(0)x(0) \quad . \tag{3:A54}$$

Here it is assumed that $\Gamma_1(i)$ is given by (3:A24) and that $\Delta(i)$ is given by (3:A51) which must be solved simultaneously with (3:A37) which gives $P(i)$. The evader's estimate of the state is obtained from (3:A32).

By defining

$$\tilde{x}(-1) = 0 \tag{3:A55}$$

(3:A53) can be used to define $u(i)$ for $i = 0$ also. $P(i)$ is considered merely to be a multiplier in this section of the Appendix; the fact that the controls are non-optimal leaves it devoid of physical significance.

In A. 2 above, no thought was given to how the pursuer obtained $\tilde{x}(i)$ at time i. For the control employed here, however, this question is irrelevant. The pursuer employs $\tilde{x}(i - 1)$ at time i as if it actually was the error of the evader's estimate at time i. Since the pursuer knows $x(i)$, $x(i - 1)$, $u(i - 1)$ and the form of the evader's controller, he can always calculate $\hat{x}(i - 1)$, or equivalently $\tilde{x}(i - 1)$, provided the inverse of $G_e(i - 1)$ exists. Consequently, though the pursuer's control, as defined by (3:A53) and (3:A54), may be non-optimal there is no doubt that it is realizable.

Now define the following

$$M(i) \triangleq E[\tilde{x}(i)\tilde{x}^T(i)] \quad , \tag{3:A56}$$

$$N(i + 1) \triangleq E[\tilde{x}(i + 1)\tilde{x}^T(i)] \quad , \tag{3:A57}$$

$$X(i) \triangleq E[x(i)x^T(i)] \quad , \tag{3:A58}$$

$$S(i) \triangleq E[\tilde{x}(i)x^T(i)] \quad , \tag{3:A59}$$

$$U(i+1) \triangleq E[\tilde{x}(i)x^T(i+1)] \quad . \tag{3:A60}$$

These matrices have physical relevance and can be assigned the following initial conditions.

$$M(0) = P_o \quad , \tag{3:A61}$$

$$M(-1) = 0 \quad , \tag{3:A62}$$

$$N(0) = 0 \quad , \tag{3:A63}$$

$$X(0) = X_o \quad , \tag{3:A64}$$

$$S(0) = P_o \quad , \tag{3:A65}$$

$$U(0) = 0 \quad . \tag{3:A66}$$

The conditions (3:A62), (3:A63), and (3:A66) result from (3:A55). (3:A61) and (3:A64) are the definitions employed in A. 2 above. (3:A65) results from the fact that the evader's initial estimate is assumed uncoupled from the error of this estimate.

Using the assumed form of the controls, plus definitions (3:A56) through (3:A60) an expression for the criterion can be obtained from (3:A30).

$$\begin{aligned} J = \mathrm{Tr}\{ \frac{a^2}{2} X(n) + \frac{1}{2} \sum_{i=0}^{n-1} [X(i)\Gamma_1(i)G_p(i)R_p^{-1}(i)G_p^T(i)\Gamma_1(i) \\ + U(i)\Gamma_1(i)G_p(i)R_p^{-1}(i)G_p^T(i)\Delta(i) + U^T(i)\Delta(i)G_p(i)R_p^{-1}(i)G_p^T(i)\Gamma_1(i) \\ + M(i-1)\Delta(i)G_p(i)R_p^{-1}(i)G_p^T(i)\Delta(i) \\ + (X(i) - S(i) - S^T(i) - M(i))\Gamma_1(i)G_e(i)R_e^{-1}(i)G_e^T(i)\Gamma_1(i)\} \quad . \end{aligned} \tag{3:A67}$$

Inserting the controls (3:A27) and (3:A53) in the system difference equation gives

$$x(i+1) = \Psi(i)x(i) - G_eR_e^{-1}G_e^T\Gamma_1\tilde{x}(i) - G_pR_p^{-1}G_p^T\Delta(i)\tilde{x}(i-1) \quad ,$$

$$i = 0, 1, \ldots, n-1 \quad . \tag{3:A68}$$

Postmultiplying (3:A68) by the transpose of itself and taking expectations gives

$$\begin{aligned} X(i+1) = {} & \Psi X\Psi - \Psi S^T\Gamma_1G_eR_e^{-1}G_e^T - \Psi U^T\Delta G_pR_p^{-1}G_p^T \\ & - G_eR_e^{-1}G_e^TS\Psi^T + G_eR_e^{-1}G_e^T\Gamma_1M_1\Gamma_1G_eR_e^{-1}G_e^T \\ & + G_eR_e^{-1}G_e^T\Gamma_1N\Delta G_pR_p^{-1}G_p^T - G_pR_p^{-1}G_p^T\Delta U\Psi^T \\ & + G_pR_p^{-1}G_p^T\Delta N\Gamma_1G_eR_e^{-1}G_e^T + G_pR_p^{-1}G_p^T\Delta M(i-1)\Delta G_pR_p^{-1}G_p^T \; . \end{aligned} \tag{3:A69}$$

Subtracting the left side of (3:A69) from both sides, premultiplying the result by $\Gamma_1(i+1)$ and summing from $i = 0$ to $n-1$ gives

$$0 = \sum_{i=0}^{n-1} \Gamma_1(i+1)\{X(i+1) - [\Psi(i)X(i)\Psi(i) - \ldots\ldots]\} \tag{3:A70}$$

which can be rewritten as

$$0 = a^2X(n) - \Gamma(0)X(0) + \sum_{i=0}^{n-1}\Gamma(i)X(i) - \sum_{i=0}^{n-1}\Gamma_1(i+1)[\Psi(i)X(i)\Psi(i) - \ldots\ldots] \; . \tag{3:A71}$$

Then taking the trace and using (3:A28) to eliminate $\Gamma_1(i+1)$ gives

$$\mathrm{Tr}[\tfrac{1}{2}a^2X(n) + \tfrac{1}{2}\sum_{i=0}^{n-1}\{X\Gamma_1 G_p R_p^{-1} G_p^T \Gamma_1 + U\Gamma_1 G_p R_p^{-1} G_p^T \Delta$$

$$+ U^T \Delta G_p R_p^{-1} G_p^T \Gamma_1 + (X - S - S^T)\Gamma_1 G_e R_e^{-1} G_e^T \Gamma_1\}]$$

$$= \mathrm{Tr}[\tfrac{1}{2}\Gamma_1(0)X(0) + \tfrac{1}{2}\sum_{i=0}^{n-1}\Gamma_1 \Psi^{-1}\{G_e R_e^{-1} G_e^T \Gamma_1 M \Gamma_1 G_e R_e^{-1} G_e^T$$

$$+ G_e R_e^{-1} G_e^T \Gamma_1 N \Delta G_p R_p^{-1} G_p^T + G_p R_p^{-1} G_p^T \Delta N^T \Gamma_1 G_e R_e^{-1} G_e^T$$

$$+ G_p R_p^{-1} G_p^T \Delta M(i-1) \Delta G_p R_p^{-1} G_p^T\}] \quad . \qquad (3{:}A72)$$

Subtracting (3:A32) from (3:A68) gives the difference equation for the propagation of $\tilde{x}$.

$$\tilde{x}(i+1) = Z(i+1)[I - G_p(i)R_p^{-1}(i)G_p^T(i)\Gamma_1(i)]\tilde{x}(i)$$

$$- Z(i+1)G_p(i)R_p^{-1}(i)G_p^T(i)\Delta(i)\tilde{x}(i-1)$$

$$- P(i+1)H^T(i+1)Q^{-1}(i+1)w(i+1) \quad . \qquad (3{:}A73)$$

Postmultiplying (3:A68) by the transpose of (3:A73) and taking expectations gives the following difference equation for S(i).

$$S(i+1) = Z(i+1)\{[I - G_p R_p^{-1} G_p^T \Gamma_1]S\Psi^T - G_p R_p^{-1} G_p^T \Delta U \Psi^T$$

$$+ G_p R_p^{-1} G_p^T \Delta M(i-1)\Delta T_p - [I - G_p R_p^{-1} G_p^T \Gamma_1] M \Gamma_1 G_e R_e^{-1} G_e^T$$

$$+ G_p R_p^{-1} G_p^T \Delta N^T \Gamma_1 G_e R_e^{-1} G_e^T$$

$$- [I - G_p R_p^{-1} G_p^T \Gamma_1] N \Delta G_p R_p^{-1} G_p^T\} \quad . \qquad (3{:}A74)$$

Subtracting the left side of (3:A74) from both sides, premultiplying the result by $\Delta(i + 1)$ and summing from $i = 0$ to $n - 1$ gives

$$0 = \sum_{i=0}^{n-1} \Delta(i + 1)\{S(i + 1) - Z(i + 1)[(I - G_p R_p^{-1} G_p^T \Gamma_1)S\Psi^T - \ldots\ldots]\} \tag{3:A75}$$

which can be rewritten as

$$0 = -\Delta(0)S(0) + \sum_{i=0}^{n-1} \Delta(i)S(i) - \sum_{i=0}^{n-1} \Delta(i + 1)Z_1(i + 1)$$

$$\times \{[I - G_p R_p^{-1} G_p^T \Gamma_1]S\Psi^T - \ldots\ldots\} \quad . \tag{3:A76}$$

Taking the trace of (3:A76) and using (3:A51) to eliminate $\Delta(i + 1)Z(i + 1)$ gives

$$0 = \mathrm{Tr}[\tfrac{1}{2}\,\Delta(0)S(0) - \tfrac{1}{2} \sum_{i=0}^{n-1} \Delta S - Z^{T^{-1}}(i + 1)\{[\Gamma_1 + \Delta]\Phi^{-1}$$

$$- \Gamma_1 \Psi^{-1}\}\{[I - G_p R_p^{-1} G_p^T \Gamma_1][S\Psi^T - M\Gamma_1 G_e R_e^{-1} G_e^T$$

$$- N\Delta G_p R_p^{-1} G_p^T] + G_p R_p^{-1} G_p^T \Delta[-U\Psi^T + M(i - 1)\Delta G_p R_p^{-1} G_p^T$$

$$+ N^T \Gamma_1 G_e R_e^{-1} G_e^T]\}] \quad . \tag{3:A77}$$

Adding (3:A72) to (3:A77) and then adding $\frac{1}{2}\mathrm{Tr}\sum_{i=0}^{n-1} [M(i - 1)\Delta G_p R_p^{-1} G_p^T \Delta - M\Gamma_1 G_e R_e^{-1} G_e^T \Gamma_1]$ to both sides gives

$$J = \mathrm{Tr}[\tfrac{1}{2}\,\Delta(0)S(0) + \tfrac{1}{2}\Gamma_1(0)X(0)$$

$$+ \frac{1}{2}\sum_{i=0}^{n-1}\{B_1(i)S(i) + B_2(i)M(i) + B_3(i)N(i) + B_4(i)U(i)$$

$$+ B_5(i)M(i-1) + B_6(i)N^T(i)\}] \quad . \tag{3:A78}$$

where

$$B_1 = \Psi^T Z^{T^{-1}}(i+1)\{[\Gamma_1 + \Delta]\Phi^{-1} - \Gamma_1\Psi^{-1}\}\{I - G_p R_p^{-1} G_p^T \Gamma_1\} - \Delta \ , \tag{3:A79}$$

$$B_2 = -\Gamma_1 G_e R_e^{-1} G_e^T Z^{T^{-1}}(i+1)\{[\Gamma_1 + \Delta]\Phi^{-1} - \Gamma_1\Psi^{-1}\}\{I - G_p R_p^{-1} G_p^T \Gamma_1\}$$

$$- \Gamma_1 G_e R_e^{-1} G_e^T \Gamma_1 + \Gamma_1 G_e R_e^{-1} G_e^T \Gamma_1 \Psi^{-1} G_e R_e^{-1} G_e^T \Gamma_1 \ , \tag{3:A80}$$

$$B_3 = -\Delta G_p R_p^{-1} G_p^T Z^{T^{-1}}(i+1)\{[\Gamma_1 + \Delta]\Phi^{-1} - \Gamma_1\Psi^{-1}\}\{I - G_p R_p^{-1} G_p^T \Gamma_1\}$$

$$+ \Delta G_p R_p^{-1} G_p^T \Gamma_1 \Psi^{-1} G_e R_e^{-1} G_e^T \Gamma_1 \ , \tag{3:A81}$$

$$B_4 = -\Psi^T Z^{T^{-1}}(i+1)\{[\Gamma_1 + \Delta]\Phi^{-1} - \Gamma_1\Psi^{-1}\} G_p R_p^{-1} G_p^T \Delta \ , \tag{3:A82}$$

$$B_5 = -\Delta G_p R_p^{-1} G_p^T Z^{T^{-1}}(i+1)\{[\Gamma_1 + \Delta]\Phi^{-1} - \Gamma_1\Psi^{-1}\} G_p R_p^{-1} G_p^T \Delta$$

$$+ \Delta G_p R_p^{-1} G_p^T \Gamma_1 \Psi^{-1} G_p R_p^{-1} G_p^T \Delta + \Delta G_p R_p^{-1} G_p^T \Delta \ , \tag{3:A83}$$

$$B_6 = \Gamma_1 G_e R_e^{-1} G_e^T Z^{T^{-1}}(i+1)\{[\Gamma_1 + \Delta]\Phi^{-1} - \Gamma_1 \Psi^{-1}\} G_p R_p^{-1} G_p^T \Delta$$

$$+ \Gamma_1 G_e R_e^{-1} G_e^T \Gamma_1 \Psi^{-1} G_p R_p^{-1} G_p^T \Delta \quad . \tag{3:A84}$$

(3:A78) is one possible expression for the criterion resulting from the controls (3:A27) and (3:A53). The objective is to demonstrate that as the time interval between i and i + 1 approaches zero, it reduces to (2:88), the expression for the criterion for the optimized continuous game. Since $\Delta(0)$ approaches $\Gamma_2(t_o)$, $\Gamma_1(0)$ approaches $K^{-1}(t_f, t_o)$ as the interval approaches zero, the first two terms match. Consequently, it is only necessary to demonstrate that the summation of (3:A78) approaches the integral of (2:88).

The relationships between the discrete-time and continuous-time variables (see Bryson and Ho [1]) are given below.

$$\delta(i) = t \quad , \tag{3:A85}$$

$$n - 1 = \frac{N-1}{\delta} \quad , \tag{3:A86}$$

$$P(i) = P(t) \quad , \tag{3:A87}$$

$$\Gamma_1(i) = K^{-1}(t_f, t) \quad , \tag{3:A88}$$

$$\Delta(i) = \Gamma_2(t) \quad ,$$

$$H^T(i+1)Q^{-1}(i+1)H(i+1) = \delta H^T(t+\delta)Q^{-1}(t+\delta)H(t+\delta) \ , \tag{3:A90}$$

$$G_p(i)R_p^{-1}(i)G_p(i) = \delta G_p(t)R_p^{-1}(t)G_p(t) \quad , \tag{3:A91}$$

$$G_e(i)R_e^{-1}(i)G_e(i) = \delta G_e(t)R_e^{-1}(t)G_e(t) \quad . \tag{3:A92}$$

Using these, expansions can be obtained for the following inverses.

$$\Phi^{-1}(t) = I + \delta G_p(t)R_p^{-1}(t)G_p^T(t)[K^{-1}(t_f, t) + \Gamma_2(t)] + O(\delta^2) \quad , \qquad (3{:}A93)$$

$$\Psi^{-1}(t) = I + \delta[G_p(t)R_p^{-1}(t)G_p^T(t) - G_e(t)R_e^{-1}(t)G_e^T(t)]\Gamma_1(t) + O(\delta^2) \quad , \qquad (3{:}A94)$$

$$Z^{T^{-1}} = I + \delta H^T(t + \delta)Q^{-1}(t + \delta)H^T(t + \delta)P(t + \delta) + O(\delta^2) \ . \qquad (3{:}A95)$$

Using (3:A85) through (3:A95), the coefficients B_1 to B_6 can be expanded in terms of δ.

$$\begin{aligned} B_1(t) = \delta\{&H^T(t + \delta)Q^{-1}(t + \delta)H(t + \delta)P(t + \delta)\Gamma_2(t) \\ &+ K^{-1}(t_f, t)G_e(t)R_e^{-1}(t)G_e^T(t)[\Gamma_2(t) + K^{-1}(t_f, t)] \\ &+ \Gamma_2(t)G_p(t)R_p^{-1}(t)G_p^T(t)\Gamma_2(t)\} + O(\delta^2) \quad , \qquad (3{:}A96) \end{aligned}$$

$$B_2(t) = -\delta K^{-1}(t_f, t)G_e(t)R_e^{-1}(t)G_e^T(t)[K^{-1}(t_f, t) + \Delta(t)] + O(\delta^2) \qquad (3{:}A97)$$

$$B_3(t) = -\delta\Gamma_2(t)G_p(t)R_p^{-1}(t)G_p^T(t)\Gamma_2(t) + O(\delta^2) \quad , \qquad (3{:}A98)$$

$$B_4(t) = +B_3(t) \quad , \qquad (3{:}A99)$$

$$B_5(t) = -B_3(t) \quad , \qquad (3{:}A100)$$

$$B_6(t) = O(\delta^2) \quad . \qquad (3{:}A101)$$

Using (3:A96) to (3:A101) in the summation of (3:A78) which shall be denoted by SUM, gives

$$\text{SUM} = \text{Tr} \sum_{i=0}^{\frac{N-1}{\delta}} \delta \{ S(t)H^T(t+\delta)Q^{-1}(t+\delta)P(t+\delta)\Gamma_2(t)$$

$$+ K^{-1}(t_f, t)G_e(t)R_e^{-1}(t)G_e^T(t)[\Gamma_2(t) + K^{-1}(t_f, t)][S(t) - M(t)]$$

$$+ \Gamma_2(t)G_p(t)R_p^{-1}(t)G_p^T(t)\Gamma_2(t)[S(t) + M(t-\delta) - N(t) - U(t)]$$

$$+ O(\delta^2)\} \quad . \tag{3:A102}$$

As can be seen from the definitions of (3:A56), (3:A57), (3:A59) and (3:A60), $S(t)$, $M(t)$, $N(t)$, $U(t)$ and $M(t - \delta)$ all approach $P(t)$ as defined by (3:A33) as δ approaches zero. Consequently, as δ approaches zero and the summation of (3:A102) becomes an integral, the terms of order δ^2 approach zero gives

$$\lim_{\delta \to 0} \text{SUM} = \text{Tr} \int_{t_o}^{t_f} P(t)H^T(t)Q^{-1}(t)P(t)\Gamma_2(t)dt \quad . \tag{3:A103}$$

Q. E. D.

In summary then, the value of continuous-time optimized criterion can be approached by the discrete-time, non-optimized but realizable control given in (3:A53). Consequently, though the pursuer may not be able to determine $\dot{y}(t)$ or $y(t)$ exactly at time t, a slight delay in this calculation does not produce a large change in the value of the criterion.

This procedure of using slightly delayed values of $y(t)$ does not eliminate the problem of instability of the feedback loop internal to the pursuer's controller. Figure 3-6 indicates the form the pursuer's controller which produces the control u(i) as given in (3:A53). This can be reduced to the form shown in Figure 3-7 where the gain matrix A(i - 1) is given by

$$A(i - 1) = -R_p^{-1}(i)G_p^T(i)\Delta(i)[G_e(i - 1)R_e^{-1}(i - 1)G_e^T(i - 1)\Gamma_1(i - 1)]^{-1} \times G_p(i - 1) \tag{3:A104}$$

where

$$A(n - 1) = 0 \quad . \tag{3:A105}$$

The stability condition for this feedback loop is that all eigenvalues of A(i - 1) have magnitudes less than 1.0. Noting the correspondence between M(t) as defined in (3:4) and I + A(i - 1), it is seen that this stability condition for the discrete-time problem corresponds to the stability condition iii. of Section 3.1 for the game in continuous-time.

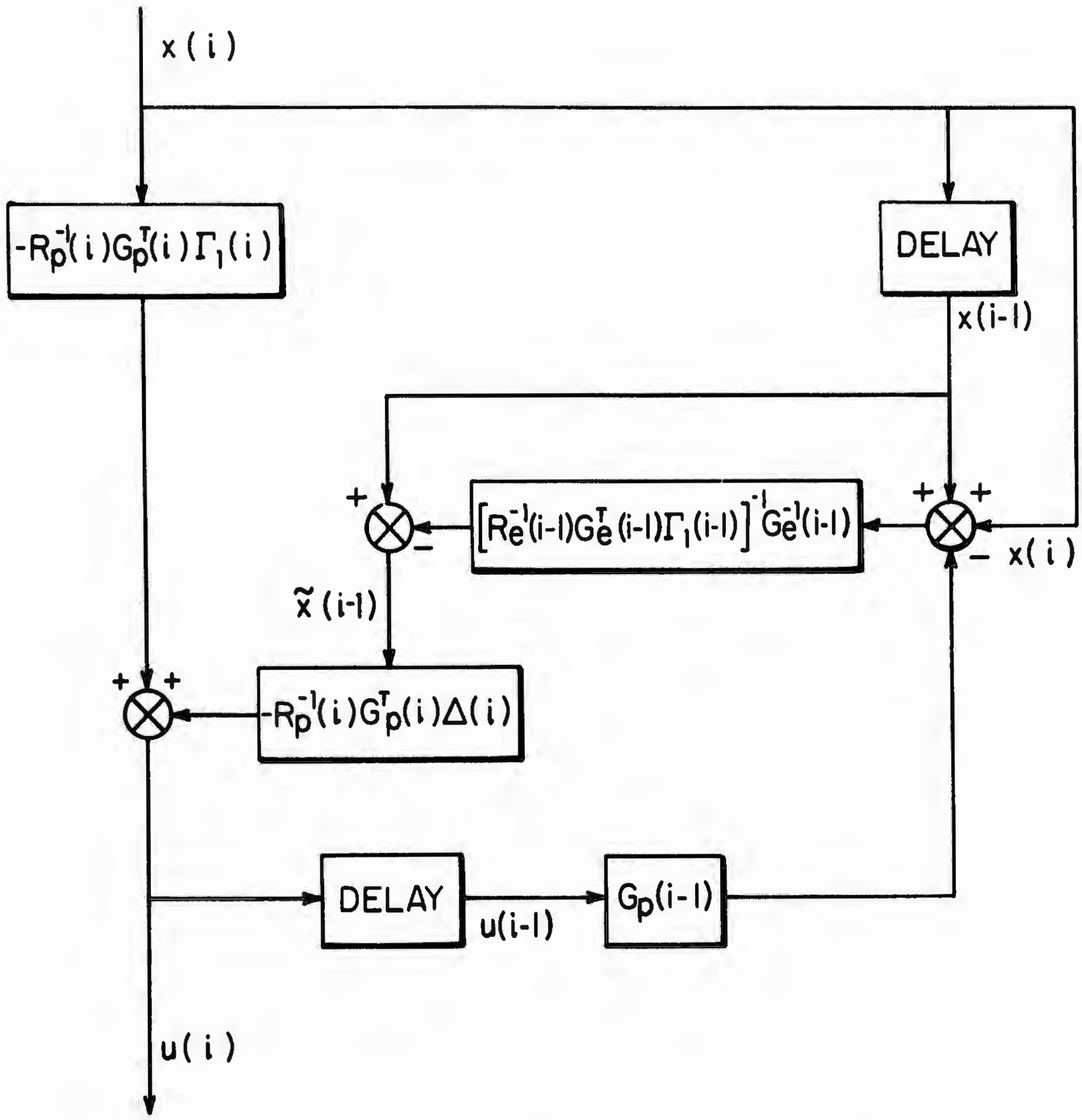


FIG. 3-6 THE PURSUER'S DISCRETE, NON-OPTIMAL CONTROLLER.

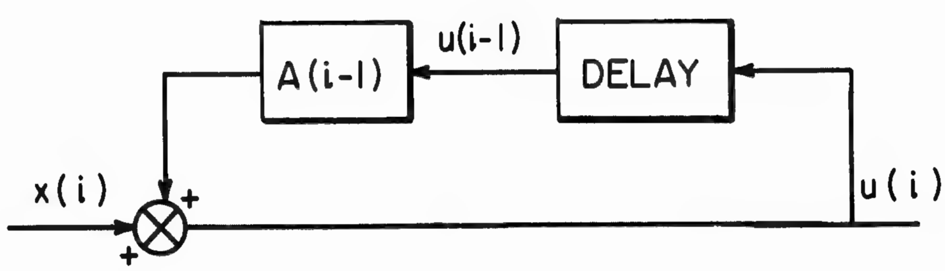


FIG. 3-7 A REDUCTION OF THE PURSUER'S NON-OPTIMAL CONTROLLER.

APPENDIX III-B

MATHEMATICAL DETAILS OF THE EVADER'S OPTIMALITY PROOF

The purpose of this appendix is to demonstrate that indeed equation (3:36)

$$B^{-1}(N^T + \mathcal{D}^T S) = R_e^{-1} G_e^T K^{-1} \tag{3:B1}$$

which gives the feedback gain matrix when determining the evader's optimality (Section 3. 3) is correct.

From Section 3. 3 the necessary definitions are

$$F \triangleq -G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) \quad , \tag{3:B2}$$

$$\mathcal{D} \triangleq -[G_p R_p^{-1} G_p^T \Gamma_2 + G_e R_e^{-1} G_e^T K^{-1}][G_e R_e^{-1} G_e^T K^{-1}]^{-1} G_e \quad , \tag{3:B3}$$

$$A \triangleq -(K^{-1} + \Gamma_2) G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) \quad , \tag{3:B4}$$

$$N = -(K^{-1} + \Gamma_2) G_p R_p^{-1} G_p^T \Gamma_2 [G_e R_e^{-1} G_e^T K^{-1}]^{-1} G_e \quad , \tag{3:B5}$$

$$\begin{aligned} B = {} & G_e^T [K^{-1} G_e R_e^{-1} G_e^T]^{-1} [K^{-1} G_e R_e^{-1} G_e^T K^{-1} - \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2] \\ & \times [G_e R_e^{-1} G_e^T K^{-1}]^{-1} G_e \quad . \end{aligned} \tag{3:B6}$$

Using these in the derived expression for the evader's feedback gain, $B^{-1}(N^T + \mathcal{D}^T S)$, gives

$$\begin{aligned} B^{-1}(N^T + \mathcal{D}^T S) = {} & R_e^{-1} G_e^T K^{-1} [K^{-1} G_e R_e^{-1} G_e^T K^{-1} - \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2]^{-1} \\ & \times [\Gamma_2 G_p R_p^{-1} G_p^T (-K^{-1} - S) - \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 \\ & - K^{-1} G_e R_e^{-1} G_e^T S] \quad . \end{aligned} \tag{3:B7}$$

Now observe that for

$$S(t) = -K^{-1}(t_f, t) \tag{3:B8}$$

(3:B7) reduces to (3:B1). Thus it is now necessary to prove that (3:B8) is true.

The differential equation which generates S(t), (3:35), can be rewritten in the convenient form

$$\dot{S} = -SF - F^T S + (S\mathfrak{B} + N)B^{-1}BB^{-1}(N^T + \mathfrak{B}S) - A \quad ,$$

$$S(t_f) = -a^2 I \quad . \tag{3:B9}$$

Inserting (3:B8) in the right-hand side of (3:B9) gives

$$\dot{S} = -K^{-1}(G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T)K^{-1} \quad ,$$

$$S(t_f) = -a^2 I \quad , \tag{3:B10}$$

which is exactly the negative of the expression for the time derivative of $K^{-1}(t_f, t)$, as given by (2:40).

Q. E. D.

References for Chapter Three

[1] A. E. Bryson, Jr. and Y. C. Ho, Optimization, Estimation and Control, Waltham, Massachusetts, Blaisdell Company, to be published, Chapter 14.

CHAPTER FOUR

A STUDY OF THE EQUATIONS FOR Γ_2 AND P

The parameters which define the optimized strategies of the two players are R_p, R_e, G_p, G_e, K^{-1}, Γ_2, P, H, and Q. Of these, all but K^{-1}, Γ_2, and P are system parameters given in the definition of the game. K^{-1} is a derived parameter of the associated deterministic game, and consequently is independent of the stochastic variables, H and Q.

Thus Γ_2 and P are the only parameters whose character is still unknown. Their values are determined by the two-point, boundary-value problem of (2:43) and (2:97). These coupled matrix Riccati equations are repeated here

$$\dot{\Gamma}_2 = \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 + \Gamma_2[G_p R_p^{-1} G_p^T K^{-1} + PH^T Q^{-1} H]$$

$$+ [K^{-1} G_p R_p^{-1} G_p^T + H^T Q^{-1} HP]\Gamma_2 + K^{-1} G_e R_e^{-1} G_e^T K^{-1} \quad ,$$

$$\Gamma_2(t_f) = 0 \qquad (4:1)$$

$$\dot{P} = -G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2)P - P(K^{-1} + \Gamma_2) G_p R_p^{-1} G_p^T - PH^T Q^{-1} HP \quad ,$$

$$P(t_o) = P_o \quad . \qquad (4:2)$$

Observe that although both equations are coupled to K^{-1}, the equation for K^{-1}, (2:40), is coupled to neither Γ_2 nor P. Consequently in studying (4:1) and (4:2), K^{-1} can be viewed as merely another time dependent matrix like $G_p(t)$, H(t), etc.

Since Γ_2 and P describe the play of the optimized stochastic game, it is important to investigate the nature of the solutions to (4:1)-(4:2). The characterization and relevance of Γ_2 and P are the subject of this chapter

Unfortunately, (4:1)-(4:2) cannot be written as a single Riccati equation. This would obviously simplify the investigation since it would reduce the question to an examination of a standard form. Even if a terminal condition

$$P(t_f) = P_f \tag{4:3}$$

is substituted for the initial condition of (4:2), Γ_2 and P cannot be written explicitly in terms of P_f (and $\Gamma_2(t_f) = 0$). Thus the results of this investigation must be derived from the general nature of the set of simultaneous equations (4:1)-(4:2), rather than directly from a well known equation with a specific solution.

Further comments on Γ_2 and P are found in Chapter 5, where a comparison is made between a new game where the pursuer makes noisy measurements, and the game which was solved in Chapter 2 where the evader makes the noisy measurements.

4.1 The Conjugate Point Condition

For problems in optimal control theory, the conjugate point condition indicates when a control -- which is derived by standard (first-order) optimization techniques -- fails to be optimal. It is always necessary to check this condition to ensure that the solution obtained is really an optimal one. The procedure for determining the existence and location of the conjugate point -- that point, when viewing time as progressing backwards from the terminal time, after which the solution is no longer valid -- is to consider the second-order accessory minimization problem. The conjugate point occurs when the identically zero control fails to be the solution or the unique solution for this accessory minimization problem.

For linear-quadratic control problems, the optimal control is a feedback one, related to the state by

$$u(t) = S(t)x(t) \tag{4:4}$$

where $S(t)$ is determined by a matrix Riccati equation which is obtained by the "sweep method" of solution. For this problem, it can be shown (see for example Bryson and Ho [1]) that the conjugate point occurs when and only when $S(t)$ fails to be finite.

As was mentioned in Section 1.4, the conjugate point for the deterministic game occurs when $K^{-1}(t_f, t)$ fails to exist. Baron [2] proved this by considering the second-order accessory minimization game. Consideration of this accessory minimization problem is required to determine sufficiency conditions for ensuring that the controls obtained are optimal; the fact that the controls are given by (1:32), (1:34) and (1:33), (1:35) are only necessary conditions and involve only first-order variations. The second-order variational terms are considered in the accessory minimization problem.

The conjugate point condition could also have been determined by considering either of the one-sided control problems which are obtained when one player declares his control and -- this given -- the other player optimizes to obtain his strategy. The conclusion from this approach must be identical, for the strategies obtained in this manner must be the same to satisfy the saddle-point condition. Consequently, if one player's strategy is not optimal from this control-theoretic view, it must be non-optimal when viewing the game as a whole. Once one player's strategy becomes non-optimal and fails to satisfy one side of the saddle-point condition, the other side of the condition becomes meaningless.

For either one-sided control problem associated with the deterministic game the S(t) which is determined by the sweep method solution is $K^{-1}(t_f, t)$.

In Chapters 2 and 3, the technique which was employed to ensure that the solutions obtained from the calculus of variations approach of Section 2.2 satisfied the saddle-point condition, was to again solve two, one-sided control problems. And too, the solution employed was obtained from the sweep method: (2:53) and (2:54); (2:70) and (2:71); (3:34) and (3:35). Consequently, the conjugate point for the game problem occurs when any of the S(t) matrices (as defined by either (2:54), (2:71), or (3:35)) fails to be finite.

Consider first the one-sided problem of Section 2.3: Proof of the Pursuer's Optimality. This is merely a linear-quadratic control problem where the state equation (2:49) is also driven by white noise. This latter fact is ignored, for the conjugate point occurs at the same place for both deterministic problems and stochastic ones involving process noise only. From (2:60), (2:61) and (2:62) it can be seen that the S(t) matrix is given by

$$S(t) = \left[\begin{array}{c:c} K^{-1}(t_f, t) & 0 \\ \hdashline 0 & \Gamma_2(t) \end{array}\right] . \tag{4:5}$$

Thus certainly there is a conjugate point when K^{-1} fails to be finite. This is most reasonable for there should be a conjugate point to the stochastic game when one occurs in the associated deterministic game. However, there is a new condition -- Γ_2 should be finite -- which reflects the new complexities of the stochastic game and gives some insight into the meaning of Γ_2. Certainly this latter restriction is also reasonable, for if Γ_2 fails to be finite, the pursuer would be adding an infinite correction for the evader's error.

The failure of $K^{-1}(t_f, t)$ to be finite shall be termed the <u>deterministic conjugate point condition</u>, i.e. the conjugate point condition for the deterministic game. When attempting to solve any stochastic game, it would naturally have to be assumed (or determined) that there was no conjugate point for the associated deterministic game. Thus the <u>stochastic conjugate point</u>, i.e. the conjugate point for the stochastic game, can be said to occur when $\Gamma_2(t)$ fails to be finite.

Consider next the problem of Section 3.3: <u>Proof of the Evader's Optimality</u> using Realization I. Here the S(t) matrix is merely $K^{-1}(t_f, t)$, see (3:B8), and thus no further restriction need be added.

For the third one-sided problem, Section 2.4: <u>Proof of the Evader's Optimality</u> with the mystical third party, the S(t) matrix is given by

$$S(t) = \left[\begin{array}{c|c} K^{-1}(t_f, t) - S_{12}(t) & S_{12}(t) \\ \hline S_{12}(t) & -S_{12}(t) \end{array}\right] \tag{4:6}$$

where $S_{12}(t)$ is obtained from

$$\begin{aligned} \dot{S}_{12} = S_{12}[PH^TQ^{-1}H - G_eR_e^{-1}G_e^TK^{-1} + G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)] \\ + [H^TQ^{-1}HP - K^{-1}G_eR_e^{-1}G_e^T + (K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T]S_{12} \\ + \Gamma_2G_pR_p^{-1}G_p^T\Gamma_2 + S_{12}G_eR_e^{-1}G_e^TS_{12} \quad , \\ S_{12}(t_f) = 0 \quad . \end{aligned} \tag{4:7}$$

Thus a conjugate point could also occur when $S_{12}(t)$ fails to exist. Certainly S_{12} will fail to be finite if Γ_2 does; however, the opposite is not necessarily true.

This is true despite the fact that the evader's optimal control as given by (2:82) is independent of $S_{12}(t)$. (2:82) was derived from (2:80) on the basis that when (2:80) was inserted into (2:73) it was found that $\hat{y}$ and $\hat{y}_p$ were identical. However, when $S_{12}(t)$ fails to exist, (2:73) does not make sense and thus these results cannot be obtained.

It seems contradictory at first that different conditions should result from the two procedures for optimizing the evader's strategy.

Yet, the implementation of the pursuer's general strategy is quite different for the two cases. In Section 2.4, the pursuer's strategy is based on data obtained from the mystical third party; in Section 3.3 he calculates the data directly from the evader's control. Consequently, when the pursuer implements his strategy as described in Section 3.3 he is more capable of taking advantage of the evader's actual deviations from the deterministic optimal control, than when the implementation of Section 2.4 is used. In Section 2.4, the pursuer can only determine what the evader's deviation "should" be, not what it actually is. By implementing his strategy in the manner of Section 3.3 the pursuer can ensure that the evader's control, as given by (2:44), will be optimal for less restrictive conditions.

By employing the implementation of Section 3.3, the pursuer is in effect using a closed-loop control feeding back the error of the evader's estimate. In Section 2.4, the pursuer's approach to using $\tilde{y}(t)$ is more in the open-loop category, for the value of $\tilde{y}(t)$ actually employed is a calculated one which may differ from the actual value as the game progresses.

Thus there exists an analogy between this situation and the deterministic pursuit-evasion game. In Section 5.1 it is shown that for the deterministic game there exist situations where the pursuer's closed-loop strategy is optimal while his open-loop one is not -- despite the fact that for the optimized trajectory they both lead to the same control time function. In other words, the conjugate point for the open-loop strategy lies closer to the terminal time than the conjugate point for the closed-loop strategy.

Here there is a similar result. S_{12} may fail to be finite, thus indicating that the pursuer's strategy as given in Section 2.4 is not optimal. However, unless $\Gamma_2(t)$ also fails to be finite, the strategy of Section 3.3 would still be valid. Again this is true despite the fact that the two strategies produce the same time functions u(t).

Since the pursuer is required in practice to employ one of the realization schemes described in Section 3.3, the problem posed by S_{12} failing to be finite will never be encountered. Future references to a stochastic conjugate point condition shall mean only that $\Gamma_2(t)$ fails to be finite.

Now consider that values of $\Gamma_2(t)$ and P(t) are needed to determine the values of the pursuer's feedback gain and the evader's estimator for a particular stochastic game with some $P(t_o)$. Assume that these values are to be obtained from (4:1)-(4:2) by integrating them backwards in time from t_f by selecting a terminal condition P_f (4:3). Then it can be seen from (4:2) that if $\Gamma_2(t)$ fails to be finite before time t_o, that P(t) will immediately fail to be finite too (unless P is identically zero).

However, such an integration process will <u>not</u> produce a value of $P(t_o)$ which is finite. Thus the values of $\Gamma_2(t)$ and P(t) obtained do not provide feedback and estimator gains for <u>any</u> game with time duration $t_f - t_o$; the definition of a game includes some finite value of $P(t_o)$. Consequently, a new terminal condition P_f must be selected and the integration performed again.

Now suppose that a particular P_f (and the accompanying integration) produces values of $\Gamma_2(t)$ and P(t) which are finite over the entire interval,

$t_o \leq t \leq t_f$, and also produces a finite $P'(t_o)$. This gives the feedback and estimator gains for optimal strategies for the game of duration $t_f - t_o$ with initial variance in the evader's estimation error of $P'(t_o)$. If the initial condition $P'(t_o)$ is finite, then $\Gamma_2(t)$ must be finite for $t_o \leq t \leq t_f$.

Consequently, the stochastic conjugate point condition adds no new restriction. If there exist values of $\Gamma_2(t)$ and $P(t)$, $t_o \leq t \leq t_f$, which satisfy (4:1)-(4:2) including the finite initial value $P(t_o)$, then $\Gamma_2(t)$, $t_o \leq t \leq t_f$, must be finite.

Recall that the conjugate point condition determines when a solution obtained by ordinary (first-order) optimization techniques is not the optimal one. However, the solution obtained for this problem automatically satisfies the conjugate point condition. This is true because the solution involves a two-point boundary-value problem, with a finite initial condition. This finite initial condition guarantees that the feedback gain will be finite.

Considering time moving in the negative direction, it is possible, however, that $\Gamma_2(t)$ will get very large as t approaches t_o. Thus it is certainly possible that the inverse of L(t) -- which is defined by (3:6) -- may fail to exist. Therefore, though Γ_2 is finite, it is necessary to determine if L^{-1} is also finite to be sure that a solution not only exists but can also be realized.

More discussion of the coupled equations for Γ_2 and P is found in Chapter 5. Some specific curves of $\Gamma_2(t)$ and $P(t)$ are found in Chapter 7.

4.2 Strategies for Infinite Noise Variance

The limiting case when the noise variance, Q, approaches infinity is of interest for it permits specific solutions to be obtained for the parameters P and Γ_2 and thus adds insight to the nature of the solution of the general problem.

As Q approaches infinity, it is certain that the term $H^TQ^{-1}H$ will approach zero, but because P(t) is dependent on Q, it is by no means certain that either

$$PH^TQ^{-1}H = 0 \tag{4:8}$$

or

$$PH^TQ^{-1}HP = 0 \quad . \tag{4:9}$$

For $PH^TQ^{-1}H$ or $PH^TQ^{-1}HP$ to be non-zero requires that P approach infinity. From (4:2) it can be seen that the driving term in the differential equation which describes P is always negative-definite; since P(t) is positive-definite the driving term is always forcing P(t) towards zero. Thus P(t) must remain finite, and consequently (4:8) and (4:9) are true.

Using (4:8) and (4:9), (4:1) and (4:2) become

$$\dot{\Gamma}_2 = \Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 + \Gamma_2 G_p R_p^{-1} G_p^T K^{-1} + K^{-1} G_p R_p^{-1} G_p^T \Gamma_2 + K^{-1} G_e R_e^{-1} G_e^T K^{-1} \quad , \qquad \Gamma_2(t_f) = 0 \tag{4:10}$$

$$\dot{P} = -G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) P - P(K^{-1} + \Gamma_2) G_p R_p^{-1} G_p^T \quad , \qquad P(t_o) = P_o \quad . \tag{4:11}$$

Here Γ_2 is decoupled from P, and thus (4:10) can be solved independently of (4:11).

By substitution, it can be shown that

$$\Gamma_2(t) = [\frac{I}{a^2} + M_p(t_f, t)]^{-1} - K^{-1}(t_f, t) \quad , \tag{4:12}$$

where $M_p(t_f, t)$ as given by (1:31), is the solution to (4:10). This method of solution may be unsatisfying to the reader, but it was the one actually employed.

Using (4:12), (4:11) becomes

$$\dot{P} = -G_p R_p^{-1} G_p^T [\frac{I}{a^2} + M_p]^{-1} P - P[\frac{I}{a^2} + M_p]^{-1} G_p R_p^{-1} G_p^T \quad ,$$

$$P(t_o) = P_o \quad . \tag{4:13}$$

The solution to the general linear differential equation

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) \quad , \qquad P(t_o) = P_o \tag{4:14}$$

is given by

$$P(t) = \Phi(t, t_o) P_o \Phi^T(t, t_o) \tag{4:15}$$

where

$$\frac{d}{dt} \Phi(t, t_o) = F(t)\Phi(t, t_o) \quad , \qquad \Phi(t_o, t_o) = I \quad . \tag{4:16}$$

Using $-G_p R_p^{-1} G_p^T [\frac{I}{a^2} + M_p(t_f, t)]^{-1}$ for F in (4:16) and solving gives

$$\Phi(t, t_o) = [\frac{I}{a^2} + M_p(t_f, t)][\frac{I}{a^2} + M_p(t_f, t_o)]^{-1} \quad . \tag{4:17}$$

Thus

$$P_\infty(t) = [\frac{I}{a^2} + M_p(t_f, t)][\frac{I}{a^2} + M_p(t_f, t_o)]^{-1} P_o [\frac{I}{a^2} + M_p(t_f, t_o)]^{-1}$$

$$\times [\frac{I}{a^2} + M_p(t_f, t)] \tag{4:18}$$

which is written here with the subscript ∞, to denote the value of Q.

Since, $M_p(t_f, t)$ is positive-definite, Γ_2 as given by (4:12) is always finite. Thus there certainly is no conjugate point for the stochastic game with infinite variance. Using (4:8) in the equations for the optimized system shows why.

Using (4:8), the evader's estimation equation (2:92) becomes

$$\dot{\hat{y}}(t) = -[G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T] K^{-1}(t_f, t)\hat{y}(t) \quad ,$$

$$\hat{y}(t_o) = \hat{y}_o \quad . \tag{4:19}$$

Taking the time derivative of the quantity, $K^{-1}(t_f, t)\hat{y}(t)$, gives zero.

Consequently the evader's control can be given by

$$v(t) = -R_e^{-1}(t) G_e^T(t) K^{-1}(t_f, t_o)\hat{y}(t_o) \tag{4:20}$$

without any loss of generality, for the evader can calculate $\hat{y}(t)$ for all t before the game begins, since he receives no data during the game. In reality then, the evader is employing an open-loop control -- in fact the same form of open-loop control as was employed in the deterministic game (1:29) except that $\hat{y}(t_o)$ replaces $y(t_o)$.

In this sense it is reasonable that the (non-existent) conjugate point occurs when $[\frac{I}{a^2} + M_p(t_f, t)]^{-1}$ fails to exist, for this is the same conjugate point condition which is associated with the optimality of the evader's open-loop strategy (1:29) in the deterministic game. Conjugate point conditions for the deterministic game are further discussed in Section 5.1.

Using (4:8) in (2:93), the equation for $\tilde{y}(t)$ becomes

$$\dot{\tilde{y}} = -G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2)\tilde{y} \quad , \qquad \tilde{y}(t_o) = \tilde{y}_o \quad . \tag{4:21}$$

This can also be integrated before the play of the game. Taking the time derivative of the quantity, $K^{-1}(t_f, t)y(t) + \Gamma_2(t)\tilde{y}(t)$ again gives zero. Thus the pursuer can also operate open-loop using the control

$$u(t) = -R_p^{-1}(t)G_p^T(t)[K^{-1}(t_f, t_o)y(t_o) + \Gamma_2(t_o)\tilde{y}(t_o)] \quad . \tag{4:22}$$

It can be shown that P(t) is a monotonically decreasing function of t, or equivalently that $P(t) < P_o$. From (4:18), $P_o - P(t)$ can be written as

$$\begin{aligned} P_o - P(t) = [\frac{I}{a^2} + M_p(t_f, t)]\{[\frac{I}{a^2} + M_p(t_f, t)]^{-1} P_o [\frac{I}{a^2} + M_p(t_f, t)]^{-1} \\ - [\frac{I}{a^2} + M_p(t_f, t_o)]^{-1} P_o [\frac{I}{a^2} + M_p(t_f, t_o)]^{-1}\} \\ \times [\frac{I}{a^2} + M_p(t_f, t)] \end{aligned} \tag{4:23}$$

which is positive-definite if

$$\begin{aligned} [\frac{I}{a^2} + M_p(t_f, t)]^{-1} P_o [\frac{I}{a^2} + M_p(t_f, t)]^{-1} > \\ [\frac{I}{a^2} + M_p(t_f, t_o)]^{-1} P_o [\frac{I}{a^2} + M_p(t_f, t_o)]^{-1} \quad . \end{aligned} \tag{4:24}$$

Observe that

$$[\frac{I}{a^2} + M_p(t_f, t)] < [\frac{I}{a^2} + M_p(t_f, t_o)] \quad , \qquad t > t_o \qquad (4:25)$$

and thus that

$$[\frac{I}{a^2} + M_p(t_f, t)]^{-1} > [\frac{I}{a^2} + M_p(t_f, t_o)]^{-1} \quad , \qquad t > t_o \quad . \qquad (4:26)$$

Consequently, (4:24) is true and the proof is complete.

That the variance of the error of the evader's estimate decreases with time -- despite the fact that the evader receives no information during the play of the game -- results from the basic stability of the differential equation governing $\tilde{y}$. Since the evader's control, $-R_e^{-1}G_e^TK^{-1}\hat{y}$, enters in an identical manner into both the equation for y and the one for $\hat{y}$, it does not effect the equation for $\tilde{y}$. Thus the system matrix of the P equation is $-G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)$ which, although it is not necessarily negative-definite, is certainly not positive-definite.

For finite values of Q(t), the pursuer will have more information than in this section. Thus the value of P(t) must be less than $P_\infty(t)$.

$$P(t) < P_\infty(t) \quad . \qquad (4:27)$$

4.3 The Results for Zero Noise Variance

As Q approaches zero, w(t) becomes identically zero and Q^{-1} approaches infinity. Then the only value of P(t) which will satisfy (4:2) is

$$P(t) \equiv 0 \quad , \qquad t < t_o \quad , \qquad (4:28)$$

for any non-zero value of P that would cause $\dot{P}$ to be $-\infty$. (4:28) indicates that

$$\tilde{y}(t) \equiv 0 \quad , \qquad t < t_o \quad . \tag{4:29}$$

This is certainly reasonable for as the evader is given perfect information of the state, the error of his estimate should become zero.

This means that both the pursuer's and the evader's control are given by their deterministic values

$$u(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) \quad , \tag{4:30}$$

$$v(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)y(t) \quad . \tag{4:31}$$

This is the required result: the optimal stochastic controls must approach the optimal deterministic controls as the noise level approaches zero.

By writing

$$\frac{d}{dt}[P^{-1}] = P^{-1}G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2) + (K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^TP^{-1}$$
$$+ H^TQ^{-1}H \quad , \qquad P^{-1}(t_o) = P_o^{-1} \tag{4:32}$$

it can be seen that

$$P^{-1}(t) = \Phi(t, t_o)P_o^{-1}\Phi^T(t, t_o) + \int_{t_o}^{t} \Phi(t, \tau)H^T(\tau)Q^{-1}(\tau)H(\tau)\Phi^T(t, \tau)d\tau \tag{4:33}$$

where

$$\frac{d}{dt}\Phi(t, t_o) = [K^{-1}(t_f, t) + \Gamma_2(t)]G_p(t)R_p^{-1}(t)G_p^T(t)\Phi(t, t_o) \quad ,$$
$$\Phi(t_o, t_o) = I \quad . \tag{4:34}$$

For Q very small the second term in (4:33) dominates and thus

$$P(t) = \left[\int_{t_o}^{t} \Phi(t,\tau)H^T(\tau)Q^{-1}(\tau)H(\tau)\Phi(t,\tau)d\tau\right]^{-1} \quad . \tag{4:35}$$

Now write

$$Q(t) = q\bar{Q}(t) \tag{4:36}$$

where q is a time-independent scaler. Then

$$P(t)H^T(t)Q^{-1}(t)H(t) = \frac{1}{q}\left[\frac{1}{q}\int_{t_o}^{t} \Phi(t,\tau)H^T(\tau)Q^{-1}(\tau)H(\tau)\Phi(t,\tau)d\tau\right]^{-1}$$

$$\times\ H^T(t)Q^{-1}(t)H(t) \quad . \tag{4:37}$$

Thus as q approaches zero and thus Q(t) approaches zero, $P(t)H^T(t)Q^{-1}(t)H(t)$ is non-zero and finite. Unfortunately, this cannot be written in terms of the basic parameters of the problem, and thus $\Gamma_2(t)$ cannot be given explicitly as it was when Q(t) approached infinity.

References for Chapter Four

[1] A. E. Bryson, Jr. and Y. C. Ho, Optimization, Estimation and Control, Waltham, Massachusetts, Blaisdell Company, to be published, Chapter 6.

[2] S. Baron, Differential Games and Optimal Pursuit-Evasion Strategies, Ph.D. Thesis, Harvard University, Cambridge, Massachusetts, 1966.

BLANK PAGE

CHAPTER FIVE

THE ASYMMETRIC ROLES OF PURSUER AND EVADER

The solutions to both the deterministic pursuit-evasion game (see Section 1.4) and the stochastic game indicated no basic difference in the roles of the two players. In the deterministic game, for example, the pursuer's and evader's feedback strategies were found to be

$$u(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) \quad , \tag{5:1}$$

$$v(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)y(t) \quad , \tag{5:2}$$

respectively. From these relations (and the fact that K^{-1} depends solely on R_p, R_e, G_p, and G_e) one might naïvely conclude that any difference between the capabilities of the pursuer and evader would depend strictly on the numerical values of R_p and R_e, and G_p and G_e.

Chapters 2, 3 and 4 above examine the solution to the pursuit-evasion game which has a basic asymmetry due to the evader's inability to make perfect measurements. Logically then, the optimal feedback strategies should have some asymmetric characteristics; they are given by

$$u(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) - R_p^{-1}(t)G_p^T(t)\Gamma_2(t)y(t) \quad , \tag{5:3}$$

$$v(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)\hat{y}(t) \quad . \tag{5:4}$$

However, an examination of the first-order variational process whereby this solution was obtained indicates no basic difference would arise if the pursuer was constrained to making imperfect measurements while the evader had perfect information. It would appear that the forms of (5:3) and (5:4) would simply be reversed.

There is, however, a fundamental asymmetry in the roles of the two players which produces some asymmetric properties of the solutions for both the deterministic and stochastic games. This asymmetry effects specifically the abilities of the two players to use open-loop control, and results from the different manner in which the two players view the criterion.

The pursuer is attempting to capture the evader and is minimizing both the terminal separation and his control effort. The evader is merely attempting to get away and is minimizing his control effort <u>but</u> maximizing the terminal separation. In short, the goals of the pursuer and the evader are <u>not</u> symmetrical.

This chapter is a study of asymmetry in both the deterministic and the stochastic games.

5.1 The Deterministic Case

In Ho, Bryson and Baron [1], the feedback strategies for the deterministic game were obtained by first determining open-loop control functions (see Section 1.4). These were converted to feedback strategies by noting the value of the feedback gain which produced the same trajectory in state space and then proving that these strategies indeed satisfied the saddle-point condition. The original open-loop time functions given by (1:28) and (1:29) are repeated here.

$$u(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t_o)y(t_o) \quad , \tag{5:5}$$

$$v(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t_o)y(t_o) \quad . \tag{5:6}$$

The conjugate point condition when closed-loop strategies are employed is that $K^{-1}(t_f, t)$ fails to be finite, as was pointed out in Chapter 4. However, suppose one player elects to employ an open-loop strategy; then where does the conjugate point occur? This problem is also discussed in Bryson and Ho [2].

Assume, for example, that the pursuer elects to employ the open-loop strategy of (5:5) and announces this fact to his opponent. It is then necessary that the evader determine if his open-loop strategy (5:6) is optimal, if his closed-loop strategy (5:2) is optimal, or if he should employ an entirely different control.

The evader maximizes a new criterion

$$J = \frac{a^2}{2} \|y(t_f)\|^2 - \frac{1}{2} \int_{t_o}^{t_f} \|v(t)\|^2_{R_e} dt \tag{5:7}$$

subject to the constraint

$$\dot{y} = -G_p(t)R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t_o)y(t_o) - G_e(t)v(t) \quad ,$$

$$y(t_o) = y_o \quad . \tag{5:8}$$

Note, that the term indicating the pursuer's control effort has been deleted from the criterion (5:8) for it is completely predetermined and thus not subject to the evader's influence.

The Hamiltonian for this problem is

$$\mathcal{H} \triangleq -\frac{1}{2}\|v\|^2_{R_e} + \lambda^T(t)\{-G_p(t)R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t_o)y_o$$

$$-G_e(t)v(t)\} \tag{5:9}$$

where $\lambda(t)$ is the adjoint vector. A necessary condition for the evader's control to be optimal is

$$\frac{\partial \mathcal{H}}{\partial v} = 0 \tag{5:10}$$

which gives the form of the control as

$$v(t) = -R_e^{-1}(t)G_e^T(t)\lambda(t) \quad . \tag{5:11}$$

The Euler equation is

$$\dot{\lambda}^T(t) = -\frac{\partial \mathcal{H}}{\partial y} \equiv 0 \quad , \qquad \lambda^T(t_f) = a^2 y^T(t_f) \tag{5:12}$$

which gives the value of the adjoint vector,

$$\lambda(t) = a^2 y(t_f) \quad . \tag{5:13}$$

Using (5:11) and (5:13) in (5:8) gives

$$\dot{y} = -G_p(t)R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t_o)y_o + a^2 G_e(t)R_e^{-1}(t)G_e^T(t)y(t_f) \quad ,$$

$$y(t_o) = y_o \quad . \tag{5:14}$$

Integrating with respect to t, gives the terminal value of the state as

$$y(t_f) = y_o - M_p(t_f, t_o)K^{-1}(t_f, t_o) + M_e(t_f, t_o)a^2 y(t_f) \quad . \tag{5:15}$$

(5:15) can be solved explicitly for $y(t_f)$.

$$y(t_f) = [I - a^2 M_e(t_f, t_o)]^{-1}[I - M_p(t_f, t_o)K^{-1}(t_f, t_o)]y_o \quad . \tag{5:16}$$

(5:16) can be reduced to

$$y(t_f) = \frac{1}{a^2} K^{-1}(t_f, t_o) y_o \quad . \tag{5:17}$$

Thus, (5:11), (5:13) and (5:17) indicate that the evader's optimal control is

$$v(t) = -R_e^{-1}(t) G_e^T(t) K^{-1}(t_f, t_o) y(t_o) \quad . \tag{5:6}$$

This is the original open-loop strategy; the closed-loop strategy (5:2) will produce the same trajectory.

This result is not valid however, if (see (5:16)) the inverse of $[I - a^2 M_e(t_f, t_o)]$ fails to exist. In fact, this represents the conjugate point condition for <u>this</u> one-sided control problem. This can be demonstrated by solving the same problem in a slightly different manner.

The evader is assumed to be permitted to play <u>after</u> the pursuer has played. Since the pursuer's strategy is assumed to be open-loop, the effect of the pursuer's control on the outcome is determined before the evader acts anyway, and thus this adds no new restriction. Here the initial condition from which the evader begins is given by

$$y_o' = y_o + \int_{t_o}^{t_f} G_p(t) u(t) dt \tag{5:18}$$

where u(t) is given by (5:5). Performing the integration indicated in (5:18) gives

$$y_o' = [I - M_p(t_f, t_o) K^{-1}(t_f, t_o)] y_o \quad . \tag{5:19}$$

Thus the differential equation which governs the new state variable y'(t) is

$$\dot{y} = -G_e(t) v(t) \quad , \qquad y'(t_o) = y_o' \quad . \tag{5:20}$$

The new criterion to be maximized is

$$J = \frac{a^2}{2} \| y'(t_f) \|^2 - \frac{1}{2} \int_{t_o}^{t_f} \| v(t) \|^2_{R_e(t)} dt \tag{5:21}$$

which is subject to the differential constraint (5:20). For any particular value of v(t), $t_o \leq t \leq t_f$, the $y'(t_f)$ obtained from (5:20) is identical to the $y(t_f)$ obtained from (5:8) and thus the value of the criterion given by (5:21) and (5:7) is the same.

From optimal control theory, it is known that the solution to (5:20)-(5:21) is given by

$$v(t) = -R_e^{-1}(t) G_e^T(t) S(t) y'(t) \tag{5:22}$$

where

$$\dot{S} = -S[G_e(t) R_e^{-1}(t) G_e^T(t)]S \quad , \qquad S(t_f) = a^2 I \quad . \tag{5:23}$$

The solution to (5:23) is

$$S(t) = [\frac{I}{a^2} - M_e(t_f, t)]^{-1} \tag{5:24}$$

and the failure of S(t) to be finite is the well-known conjugate point condition. Consequently, if the pursuer employs an open-loop control, and if S(t) as given by (5:24) fails to be finite, there exists a control for the evader <u>other</u> than (5:6) or (5:2) which maximizes the criterion.

It is also enlightening to determine the pursuer's optimal control if the evader employs an open-loop strategy. An analysis similar to the one above finds that the pursuer's strategy should be given by either (5:1) or (5:5); the conjugate point condition is that $[\frac{I}{a^2} + M_p(t_f, t)]^{-1}$ fails to be finite. However, since $M_p(t_f, t)$ is positive semi-definite, the conjugate point condition can never be satisfied. Consequently, the evader can always replace his feedback strategy (5:2) with the open-loop strategy (5:6) and still guarantee that under no conditions will the value of the criterion be less than that given by (2:88).

This difference in the ability of the two players to employ open-loop control is certainly intuitively reasonable. Since he is merely trying to get away, the evader -- after an initial look at the pursuer's location -- can simply run away blindly; the evader can always calculate where the pursuer should optimally be, and that the pursuer can get no closer. The pursuer, however, since he is attempting to catch a particular moving tar t cannot (always) close his eyes for fear the evader may try a tricky "non-optimal" maneuver, like running around behind him.

This point can best be demonstrated with an example. Consider a particular pursuit-evasion problem in two-dimensional euclidean space. This is the Classical Interception Problem which is the subject of Chapter 7.

$$\dot{x}_p = F_p x_p + \overline{G}_p u(t) \quad , \qquad x_p(t_o) = x_{po} \tag{5:25}$$

$$\dot{x}_e = F_e x_e + \overline{G}_e v(t) \quad , \qquad x_e(t_o) = x_{eo} \tag{5:26}$$

where

$$F_p = F_e = \left[\begin{array}{c|c} 0 & I_2 \\ \hline 0 & 0 \end{array}\right] \tag{5:27}$$

$$\overline{G}_p = \overline{G}_e = \begin{bmatrix} 0 \\ \hline I_2 \end{bmatrix} . \tag{5:28}$$

The control weighting terms in the criterion are assumed of the form

$$R_p = r_p I_2 \quad , \tag{5:29}$$

$$R_e = r_e I_2 \quad . \tag{5:30}$$

The pursuer's state vector is given by

$$x_p = \begin{bmatrix} x_{1p} \\ x_{2p} \\ x_{3p} \\ x_{4p} \end{bmatrix} \tag{5:31}$$

where the scalars x_{1p} and x_{2p} represent the pursuer's position in the two euclidean dimensions and x_{3p} and x_{4p} the pursuer's corresponding velocities. A similar vector defines the evader's state.

The transition matrices corresponding to F_p and F_e are given by

$$\Phi_p(t_f, t) = \Phi_e(t_f, t) = \left[\begin{array}{c|c} I_2 & (t_f - t)I_2 \\ \hline 0 & I_2 \end{array}\right] \tag{5:32}$$

and thus the control matrices for the reduced problem are

$$G_p(t) = G_e(t) = (t_f - t)I_2 \quad . \tag{5:33}$$

The state vector for the reduced problem is two-dimensional and given by

$$y(t) = [I_2 \mid (t_f - t)I_2][x_p(t) - x_e(t)] \tag{5:34}$$

and the state equation is

$$\dot{y}(t) = G_p(t)u(t) - G_e(t)v(t) \quad , \qquad y(t_o) = y_o \quad . \tag{5:35}$$

The initial conditions for this problem are taken to be

$$x_p(t_o) = 0 \quad , \tag{5:36}$$

$$x_e(t_o) = \begin{bmatrix} b \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad . \tag{5:37}$$

Thus in the two-dimensional euclidean space, the initial configuration of the two players is as shown in Figure 5-1 with no initial velocities. Using (5:34) these initial conditions can be transformed into initial conditions for the reduced problem.

$$y(t_o) = \begin{bmatrix} -b \\ 0 \end{bmatrix} \tag{5:38}$$

which are displayed in Figure 5-2. Note that y_1 represents the predicted terminal distance between the players in the first euclidean dimension, x_1, and y_2 the same in the second euclidean dimension, x_2.

In the reduced space of Figure 5-2, the pursuer's aim is to bring $y(t_f)$ close to the origin, while the evader attempts to move $y(t_f)$ away from the origin.

From Figure 5-2, it can be seen that the evader should employ no control in the y_2 (x_2) direction and positive control (note the negative sign which preceeds G_e in (5:35)) in the y_1 (x_1) direction in order to take y(t) in the negative direction along the y_1 axis.

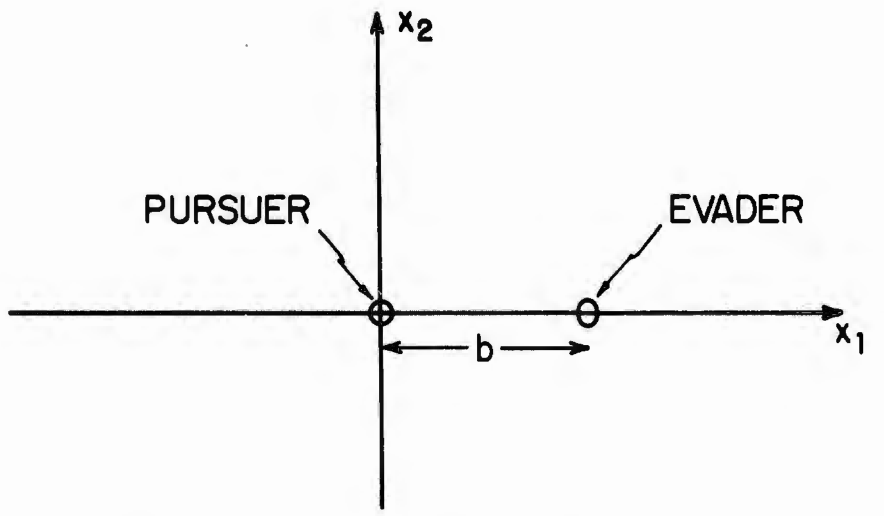


FIG. 5-1 INITIAL CONFIGURATION IN EUCLIDEAN (x) SPACE

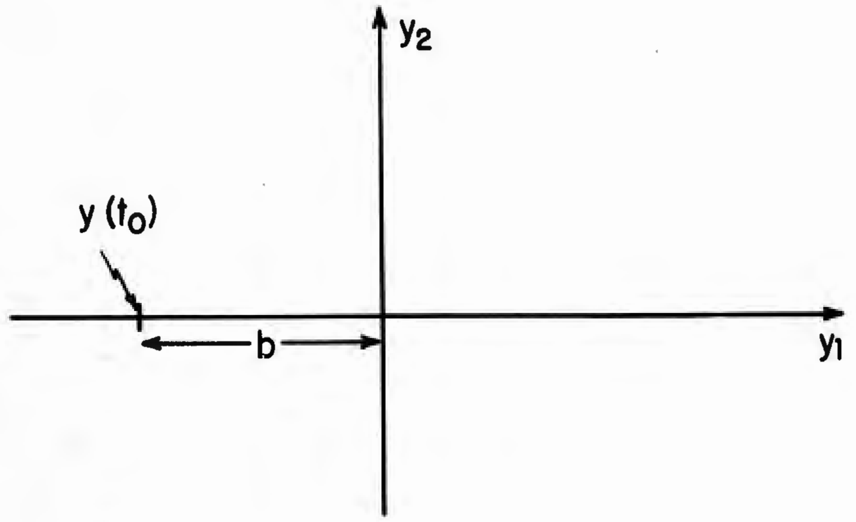


FIG. 5-2 INITIAL CONFIGURATION IN REDUCED (y) SPACE

However, suppose that the pursuer employs the open-loop control of (5:5). For the particular problem considered here, this time function is given by (see Section 7.1)

$$u(t) = \frac{3a^2 r_e (t_f - t)}{3r_p r_e + a^2 (r_e - r_p)(t_f - t_o)^3} \begin{bmatrix} b \\ 0 \end{bmatrix} . \tag{5:39}$$

Integrating $G_p(t)u(t)$ with respect to time from t_o to t_f gives

$$M_p(t_f, t_o)K^{-1}(t_f, t_o)y_o = \frac{a^2 r_e (t_f - t_o)^3}{3r_p r_e + a^2 (r_e - r_p)(t_f - t_o)^3} \begin{bmatrix} b \\ 0 \end{bmatrix} . \tag{5:40}$$

Using this in (5:19), $y'(t_o)$ is found to be

$$y'(t_o) = \left(-1 + \frac{a^2 r_e (t_f - t_o)^3}{3r_p r_e + a^2 (r_e - r_p)(t_f - t_o)^3} \right) \begin{bmatrix} b \\ 0 \end{bmatrix} . \tag{5:41}$$

The possible locations in y space of $y'(t_o)$ are shown in Figure 5-3. Note that if in (5:41)

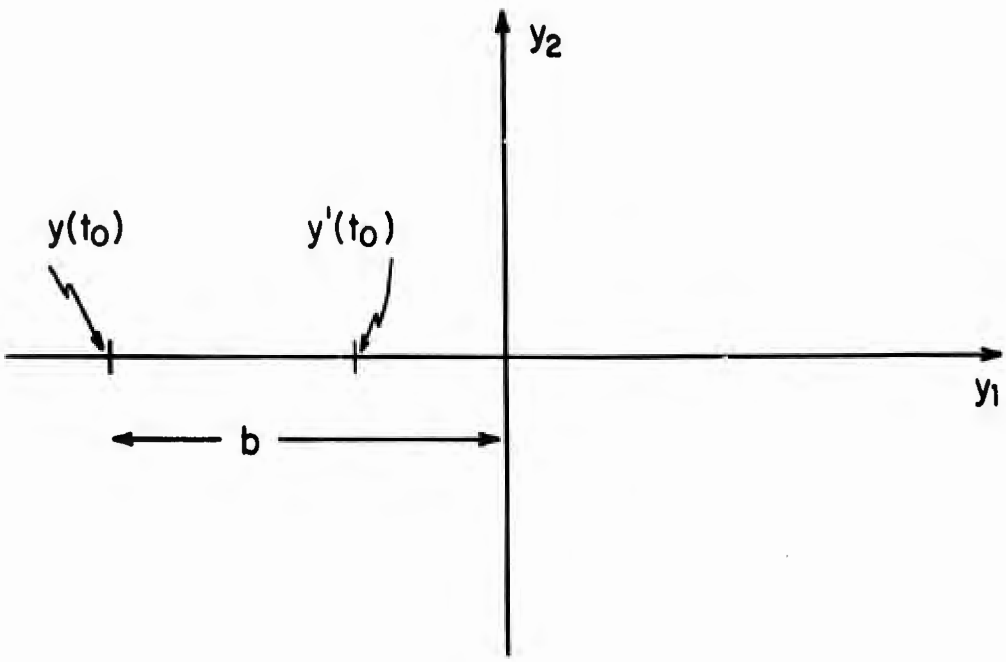
$$\frac{a^2 r_e (t_f - t_o)^3}{3r_p r_e + a^2 (r_e - r_p)(t_f - t_o)^3} > 1 \tag{5:42}$$

which can be simplified to

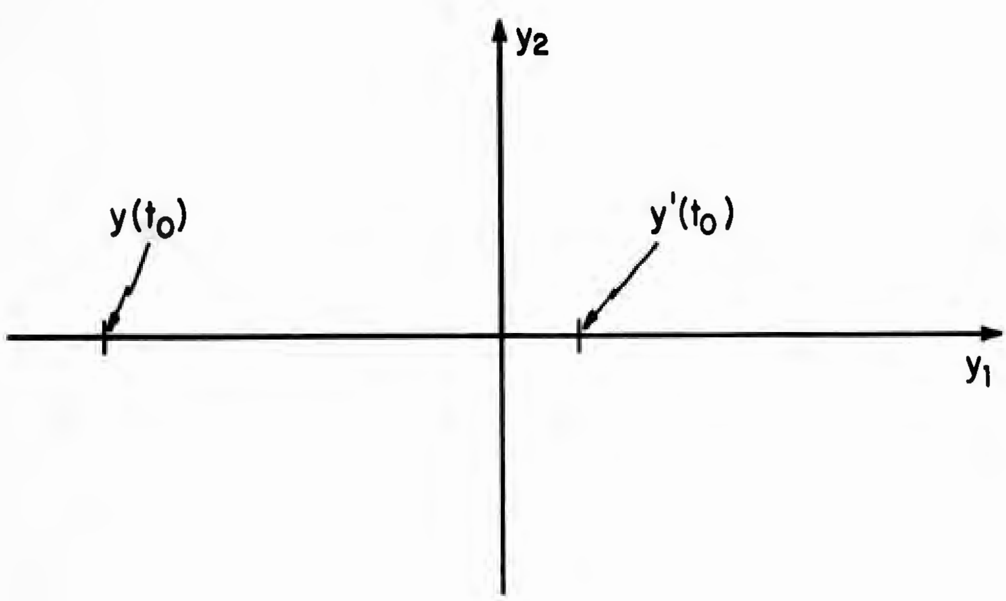
$$(t_f - t_o) > \frac{3r_e}{a^2} , \tag{5:43}$$

then $y'(t_o)$ will be located on the positive y_1 axis, as is indicated in Figure 5-3b.

However, if (5:43) is satisfied, it can be seen from Figure 5-3b that the evader should employ negative control (again recall the minus sign



(a)



(b)

FIG. 5-3 POSSIBLE LOCATIONS OF $y'(t_0)$

which preceeds G_e) so as to move in the positive y_1 direction. In the y-plane the evader desires to get the point y as far away from the origin as possible, and thus if y begins on the positive side, he certainly will want to move y in the positive direction.

In the x (euclidean) space this pair of controls will result in the trajectories indicated in Figure 5-4. The paths of the two players will cross, but the terminal separation will be larger than if the evader (using the same energy) moved in a positive x_1 direction.

Condition (5:43) does represent the conjugate point condition found above. Simple integration gives

$$M_e(t_f, t) = \frac{(t_f - t)}{3r_e} I_2 \quad . \tag{5:44}$$

Thus $[\frac{I}{a^2} - M_e(t_f, t^*)]^{-1}$ fails to exist for some t^* ($t_o \leq t^* \leq t_f$) when

$$\frac{(t_f - t^*)}{3r_e} = \frac{1}{a^2} \tag{5:45}$$

which will occur if (5:43) is satisfied.

In summary then, for the deterministic game initial data is sufficient for the evader to determine an optimal strategy, i.e. for the evader an open-loop strategy is always optimal. However, the pursuer must check to see if $[\frac{I}{a^2} - M_e(t_f, t)]^{-1}$ always exists before he uses an open-loop control. If $G_e(t)R_e^{-1}(t)G_e^T(t)$ is always non-singular, this one-sided conjugate point condition will occur if the time duration of the game is made long enough.

If the objective is escape, the optimal strategy is much simpler than it is if the goal is capture.

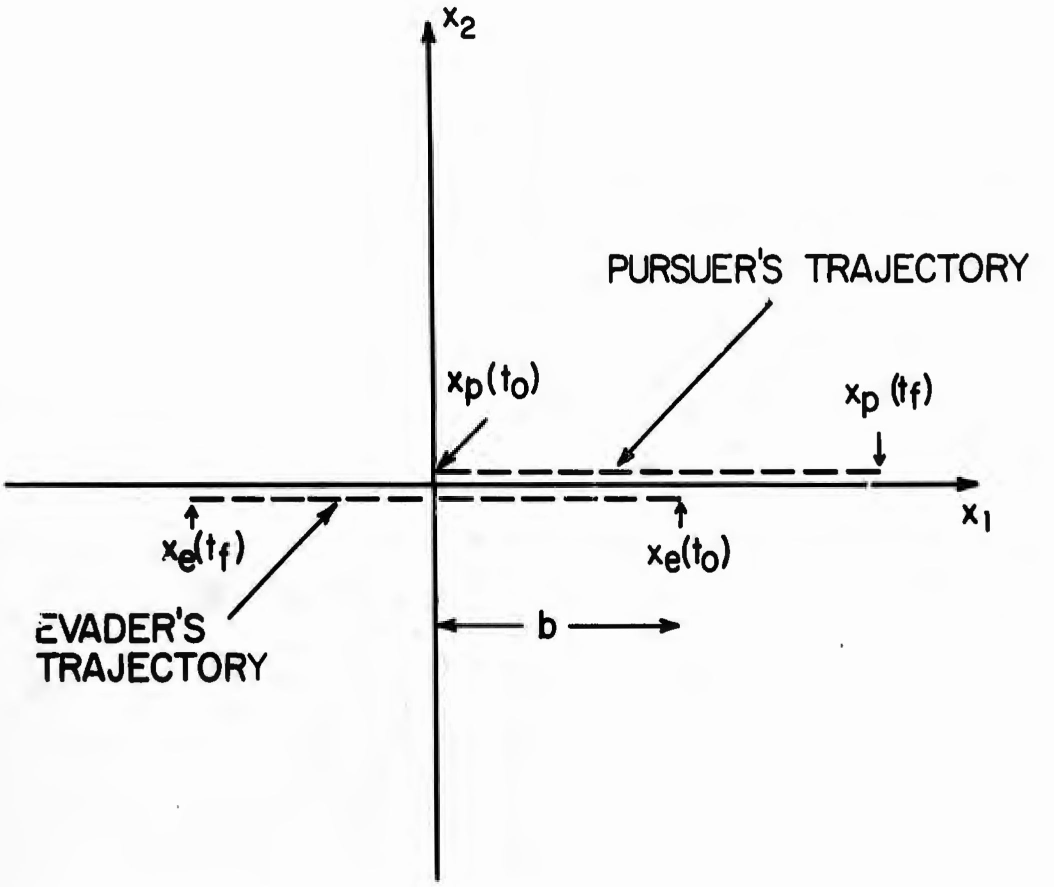


FIG. 5-4 THE PURSUER'S AND EVADER'S TRAJECTORIES

5.2 The Pursuer with Noisy Measurements

The problem considered in this section is one where the evader has perfect information while the pursuer makes noisy measurements of the state y(t) of the form

$$z(t) = H(t)y(t) + w(t) \tag{5:46}$$

where w(t) is Gaussian white noise with zero mean and variance Q(t). This is the analog of the problem solved in Chapter 2, with the quality of the information available to the two players reversed. Because of the similar nature of this problem to the one of Chapter 2, the details of the solution are not given. Rather this section summarizes the results of an analysis which is identical in form to that of Chapter 2.

One might expect the results here to be very similar to those obtained in Chapters 2, 3 and 4 for the game where the evader is making the noisy measurements. However, the precedent of Section 5.1 would also indicate that the results would not be identical. In particular, one would certainly expect some difficulty when the noise variance, Q, approaches infinity; then the pursuer would be effectively employing an open-loop strategy -- one which the previous section indicates has certain fallacies.

Again the game equation is given by

$$\dot{y} = G_p(t)u(t) - G_e(t)v(t) \quad , \qquad y(t_o) = y_o \tag{5:47}$$

and the criterion by

$$J = E\{\frac{a^2}{2}\|y(t_f)\|^2 + \frac{1}{2}\int_{t_o}^{t_f}[\|u(t)\|^2_{R_p} - \|v(t)\|^2_{R_e}]dt\} \quad . \tag{5:48}$$

The pursuer's and the evader's optimal strategies are

$$U^o:\ u^o(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)\hat{y}(t) \tag{5:49}$$

$$V^o:\ v^o(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)y(t) - R_e^{-1}(t)G_e^T(t)\Gamma_2(t)\tilde{y}(t) \tag{5:50}$$

where $\hat{y}(t)$ is the pursuer's estimate of the state, and $\tilde{y}(t)$ is the error of this estimate. The pursuer's estimate is obtained from the Kalman-Bucy filter which is given by

$$\dot{\hat{y}} = -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}\hat{y} + PH^TQ^{-1}(z - H\hat{y}) \quad ,$$

$$\hat{y}(t_o) = y_o \quad . \tag{5:51}$$

$K^{-1}(t_f, t)$ has the same definition as always, while the other two parameters, $\Gamma_2(t)$ and $P(t)$, are derived from the two-point boundary-value problem consisting of two coupled Riccati equations which are similar, but not identical, to the ones obtained in Chapter 2.

$$\dot{\Gamma}_2 = -\Gamma_2G_eR_e^{-1}G_e^T\Gamma_2 + \Gamma_2(PH^TQ^{-1}H - G_eR_e^{-1}G_e^TK^{-1})$$

$$+ (H^TQ^{-1}HP - K^{-1}G_eR_e^{-1}G_e^T)\Gamma_2 - K^{-1}G_pR_p^{-1}G_p^TK^{-1} \quad ,$$

$$\Gamma_2(t_f) = 0 \tag{5:52}$$

$$\dot{P} = G_eR_e^{-1}G_e^T(K^{-1} + \Gamma_2)P + P(K^{-1} + \Gamma_2)G_eR_e^{-1}G_e^T$$

$$- PH^TQ^{-1}HP \quad , \qquad P(t_o) = P_o \quad . \tag{5:53}$$

The equation which gives the optimized value of the criterion is the same as for the previous problem

$$J = \mathrm{Tr}\{\tfrac{1}{2}K^{-1}(t_f, t_o)Y(t_o) + \tfrac{1}{2}\Gamma_2(t_o)P(t_o) + \tfrac{1}{2}\int_{t_o}^{t_f} PH^TQ^{-1}HP\Gamma_2 dt\} \quad . \tag{5:54}$$

But it should be noted from (5:52) that Γ_2 is <u>positive</u>-definite for this problem. Thus the second and third terms in the criterion (5:54) are positive indicating a decrease in the pursuer's (he is the minimizing player) capabilities with respect to the deterministic problem.

The optimized game equation and the estimation equation can be written in the form of a 2k-dimensional vector differential equation.

$$\begin{bmatrix} \dot{y} \\ \dot{\hat{y}} \end{bmatrix} = \left[\begin{array}{c|c} +G_eR_e^{-1}G_e^T(K^{-1}+\Gamma_2) & -(G_pR_p^{-1}G_p^TK^{-1} + G_eR_e^{-1}G_e^T\Gamma_2) \\ \hline PH^TQ^{-1}H & -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1} - PH^TQ^{-1}H \end{array}\right] \begin{bmatrix} y \\ \hat{y} \end{bmatrix}$$

$$+ \left[\begin{array}{c} 0 \\ \hline PH^TQ^{-1} \end{array}\right] w(t) \quad , \quad \begin{bmatrix} y(t_o) \\ \hat{y}(t_o) \end{bmatrix} = \begin{bmatrix} y_o \\ \hat{y}_o \end{bmatrix} \quad . \tag{5:55}$$

The optimized game equation can also be written coupled to a differential equation which gives $\tilde{y}(t)$, the error of the pursuer's estimate.

$$\begin{bmatrix} \dot{y} \\ \dot{\tilde{y}} \end{bmatrix} = \left[\begin{array}{c|c} -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1} & (G_pR_p^{-1}G_p^TK^{-1} + G_eR_e^{-1}G_e^T\Gamma_2) \\ \hline 0 & G_eR_e^{-1}G_e^T(K^{-1}+\Gamma_2) - PH^TQ^{-1}H \end{array}\right] \begin{bmatrix} y \\ \tilde{y} \end{bmatrix}$$

$$- \left[\begin{array}{c} 0 \\ \hline PH^TQ^{-1} \end{array}\right] w(t) \quad , \quad \begin{bmatrix} y(t_o) \\ \tilde{y}(t_o) \end{bmatrix} = \begin{bmatrix} y_o \\ \tilde{y}_o \end{bmatrix} \quad . \tag{5:56}$$

Y(t) is obtained from

$$\dot{Y} = -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}Y - YK^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)$$

$$+ (G_pR_p^{-1}G_p^TK^{-1} + G_eR_e^{-1}G_e^T\Gamma_2)P$$

$$+ P(K^{-1}G_pR_p^{-1}G_p^T + \Gamma_2G_eR_e^{-1}G_e^T) \quad ,$$

$$Y(t_o) = Y_o \quad . \qquad (5:57)$$

The three conditions which together are necessary and sufficient for the evader to be able to calculate $\tilde{y}(t)$ and actually realize his optimal strategy as given by (5:50) are:

i. The number of the pursuer's control variables (i.e. the dimension of u(t)) must be equal to or greater than the number of "interesting" state variables (i.e. the dimension of y(t)).

ii. The inverse of $G_p(t)R_p^{-1}(t)G_p^T(t)$ must exist for all values of $t < t_f$.

iii. The inverse of L(t) must exist for all values of $t < t_f$.

Here, L(t) is defined as

$$L(t) \triangleq [G_p(t)R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t) + G_e(t)R_e^{-1}(t)G_e^T(t)\Gamma_2(t)] \quad . \qquad (5:58)$$

The condition required for the certainty-equivalence principle to hold when proving the pursuer's optimality (the proof corresponding to the one of Section 3.3) is that

$$R_p - R_pG_p^{-1}K\Gamma_2G_eR_e^{-1}G_e^T\Gamma_2K(G_p^T)^{-1}R_p > 0 \qquad (5:59)$$

which can be reduced to

$$K^{-1}G_pR_p^{-1}G_p^TK^{-1} - \Gamma_2G_eR_e^{-1}G_e^T\Gamma_2 > 0 \quad . \qquad (5:60)$$

Again the stochastic conjugate point occurs when $\Gamma_2(t)$ fails to be finite. And, as before, for a solution to exist $\Gamma_2(t)$ must be <u>finite</u> because of the finite initial condition in (5:53).

5.3 The Asymmetry of the Two Stochastic Games

As can be seen by comparing the results summarized in Section 3.2 for the stochastic game when the pursuer has noisy measurements with the results of Chapters 2, 3 and 4, the descriptions of the two solutions are similar but not identically symmetrical. This section analyzes these differences. In order to differentiate between the analogous parameters for the two different problems, the subscript e is used to denote the parameters resulting from the problem in which the evader makes noisy measurements; the subscript p indicates parameters for the pursuer-with-noisy-measurements problem.

Consequently, the Γ_2 and P parameters which describe the problem for the evader with noisy measurements are given by

$$\dot{\Gamma}_{2e} = \Gamma_{2e}G_pR_p^{-1}G_p^T\Gamma_{2e} + \Gamma_{2e}(P_eH^TQ^{-1}H + G_pR_p^{-1}G_p^TK^{-1})$$

$$+ (H_eQ^{-1}H^TP + K^{-1}G_pR_p^{-1}G_p^T)\Gamma_{2e} + K^{-1}G_eR_e^{-1}G_e^TK^{-1} \quad ,$$

$$\Gamma_{2e}(t_f) = 0 \quad , \tag{5:61}$$

$$\dot{P}_e = -G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_{2e})P_e - P_e(K^{-1} + \Gamma_{2e})G_pR_p^{-1}G_p^T$$

$$- P_eH^TQ^{-1}HP_e \quad , \qquad P_e(t_o) = P_{eo} \quad , \tag{5:62}$$

while the same parameters for the pursuer-with-noisy-measurements problem are calculated from

$$\dot{\Gamma}_{2p} = -\Gamma_{2p}G_eR_e^{-1}G_e^T\Gamma_{2p} + \Gamma_{2p}(P_pH^TQ^{-1}H - G_eR_e^{-1}G_e^TK^{-1})$$

$$+ (H^TQ^{-1}HP_p - K^{-1}G_eR_e^{-1}G_e^T)\Gamma_{2p} - K^{-1}G_pR_p^{-1}G_p^TK^{-1} \quad ,$$

$$\Gamma_{2p}(t_f) = 0 \quad , \tag{5:63}$$

$$\dot{P}_p = G_eR_e^{-1}G_e^T(K^{-1} + \Gamma_{2p})P + P(K^{-1} + \Gamma_{2p})G_eR_e^{-1}G_e^T$$

$$- P_pH^TQ^{-1}HP_p \quad , \qquad P_p(t_o) = P_{po} \quad . \tag{5:64}$$

The differential equation for Γ_{2p} is much more prone towards instability than the equation which defines Γ_{2e}. This can be seen by writing

$$\Gamma_{3e}(t) \triangleq K^{-1}(t_f, t) + \Gamma_{2e}(t) \quad , \tag{5:65}$$

$$\Gamma_{3p}(t) \triangleq K^{-1}(t_f, t) + \Gamma_{2p}(t) \tag{5:66}$$

and then writing the differential equations for these new parameters.

$$\dot{\Gamma}_{3e} = \Gamma_{3e}G_pR_p^{-1}G_p^T\Gamma_{3e} + \Gamma_{3e}P_eH^TQ^{-1}H + H^TQ^{-1}HP\Gamma_3$$

$$- [K^{-1}P_eH^TQ^{-1}H + H^TQ^{-1}HP_eK^{-1}] \quad ,$$

$$\Gamma_{3e}(t_f) = a^2I \tag{5:67}$$

$$\dot{\Gamma}_{3p} = -\Gamma_{3p}G_eR_e^{-1}G_e^T\Gamma_{3p} + \Gamma_{3p}P_pH^TQ^{-1}H + H^TQ^{-1}HP_p\Gamma_3$$

$$- [K^{-1}P_eH^TQ^{-1}H + H^TQ^{-1}HP_eK^{-1}] \quad , \qquad \Gamma_{3p}(t_f) = a^2I \quad . \tag{5:68}$$

In (5:67) the quadratic term is positive-definite, while in (5:68) it is negative-definite. Thus when integrating (5:67) backwards in time, the quadratic term will have a stabilizing effect as long as the eigenvalues of Γ_{3e} remain positive (i.e. as long as the driving term does not force one of the eigenvalues below zero). However, in (5:68) the quadratic term is destabilizing, for Γ_{3p} is positive-definite and thus when integrating backwards in time the quadratic term will drive the eigenvalues of Γ_{3p} even larger.

Consideration of the scalar case permits one to be even more specific. In this case, again when viewing time as progressing backwards, (5:67) is stable; the terminal value is positive, the quadratic term negative (thus stabilizing) while the driving term is positive ensuring that Γ_{3e} does not become negative. However, (5:68) is unstable. Here the terminal condition is again positive but the quadratic term is also positive and thus destabilizing; the driving term is also positive contributing to the instability. In fact, if $G_e R_e^{-1} G_e^T$ is non-zero for the entire duration of the game, and if the game interval is made long enough, then Γ_{3p} will fail to be finite.

If Q is allowed to approach infinity, Γ_{2e} and Γ_{2p} become

$$\Gamma_{2e}(t) = [\frac{I}{a^2} + M_p(t_f, t)]^{-1} - K^{-1}(t_f, t) \tag{5:69}$$

as was shown in Section 4.2, and

$$\Gamma_{2p}(t) = [\frac{I}{a^2} - M_e(t_f, t)]^{-1} - K^{-1}(t_f, t) \tag{5:70}$$

respectively. Using (5:65) and (5:66), these easily give the values of Γ_{3e} and Γ_{3p}.

$$\Gamma_{3e} = [\frac{I}{a^2} + M_p(t_f, t)]^{-1} \quad , \tag{5:71}$$

$$\Gamma_{3p} = [\frac{I}{a^2} - M_e(t_f, t)]^{-1} \quad . \tag{5:72}$$

These latter matrices are exactly the ones for determining the open-loop conjugate point condition for the deterministic game. And since K^{-1} is implicitly assumed to be finite, (5:69) and (5:70) can also play the same roles.

This is intuitively sensible. For the game with the evader making noisy measurements, as Q approaches infinity the evader's strategy approaches an open-loop one. Thus as Q approaches infinity the extra stochastic conjugate point condition (which is determined by optimizing the pursuer's control against a fixed evader's strategy) should approach the conjugate point condition obtained for the deterministic game when the evader operates strictly open-loop (and where the conjugate point is determined by optimizing for the pursuer against this fixed strategy). For the case where the pursuer makes noisy measurements, the opposite should hold true.

(5:71) and (5:72) add further insight to the question of stabilitv for equations (5:61) and (5:63). Note that (5:69) will always remain finite, indicating that (5:61) is certainly stable for Q infinite. However, (5:70) will fail to exist if $G_e R_e^{-1} G_e^T$ is non-zero (it is always positive-semi-definite) during the entire game, and if the duration of the game is made long enough. Thus as Q approaches infinity, (5:63) certainly will become unstable.

The graphic description of this phenomenon on the time axis is perhaps worth the proverbial thousand words. From Figure 5-5 it can be seen exactly where valid solutions can be found. Observe that the stability-instability property does not effect the existence of a solution, though it does effect the ease with which solutions can be determined; the examples of Chapter 7 illuminate this point better than further discussion.

One more point needs to be discussed in light of the asymmetry of the two problems. For each game there was found to be a matrix, L(t), which was required to be non-singular if the player with perfect information was to be able to implement his optimal strategy. For the game with the pursuer with noisy measurements, and for the one with the evader with noisy measurements these are given by

$$L_e(t) = G_p(t)R_p^{-1}(t)G_p^T(t)\Gamma_{2e}(t) + G_e(t)R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t) \qquad (5:73)$$

$$L_p(t) = G_p(t)R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t) + G_e(t)R_e^{-1}(t)G_e^T(t)\Gamma_{2p}(t) \qquad (5:74)$$

respectively.

Again it is difficult to state anything that is both exact and universal about the singularity of these two matrices. However, recall that Γ_{2e} is negative-definite while Γ_{2p} is positive-definite. Thus while L_e is the <u>difference</u> of two matrices which are each products of two positive-definite matrices, L_p is the <u>sum</u> of two matrices which are also each products of two positive-definite matrices. Certainly then for the scalar case, L_p will always be positive, while L_e could become zero.

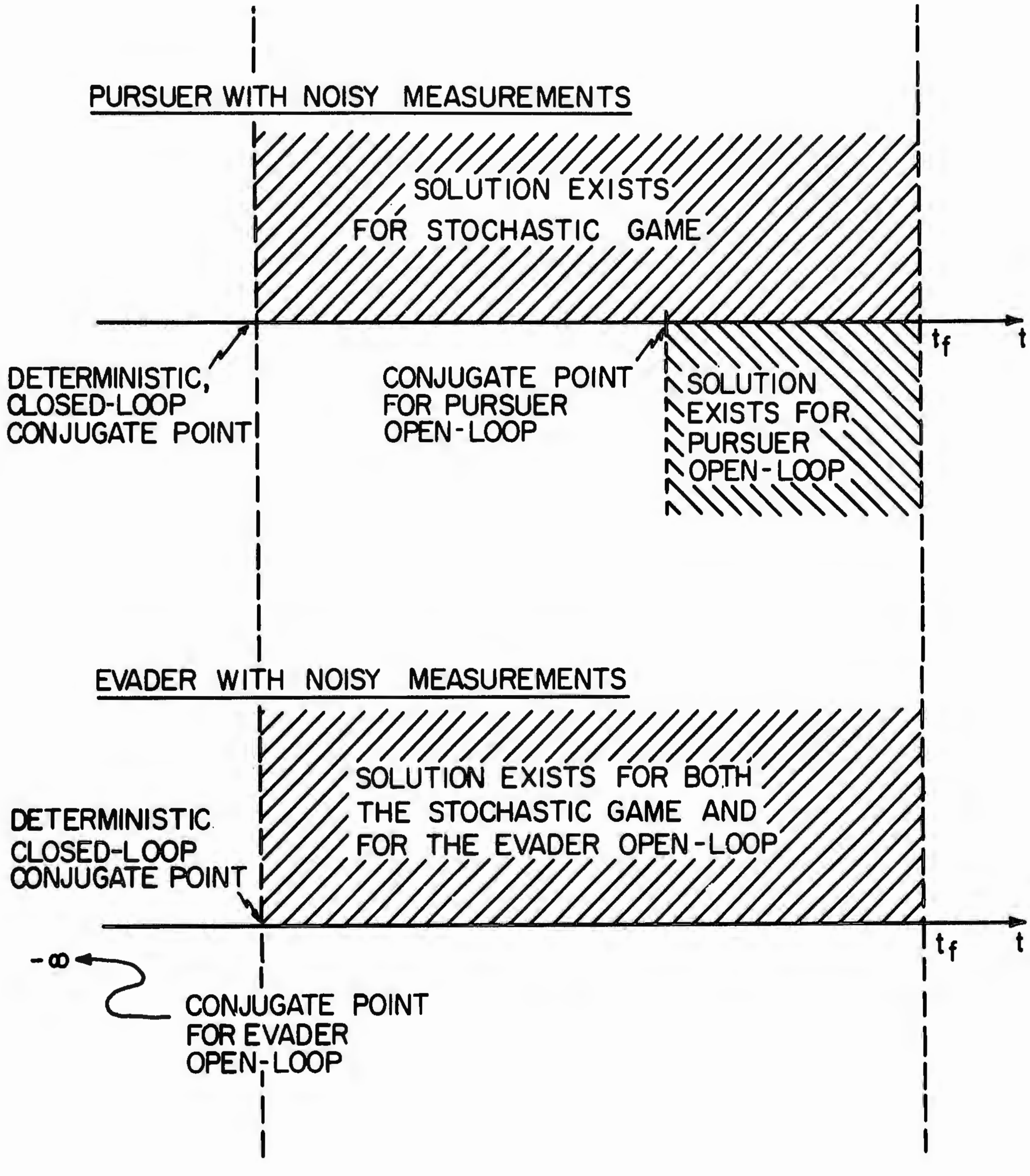


FIG. 5-5 GRAPHIC REPRESENTATION OF CONJUGATE POINTS

References for Chapter Five

[1] Y. C. Ho, A. E. Bryson, Jr. and S. Baron, "Differential Games and Optimal Pursuit-Evasion Strategies," IEEE Transactions on Automatic Control, AC-10, No. 4, Oct. 1965.

[2] A. E. Bryson, Jr. and Y. C. Ho, Optimization, Estimation and Control, Waltham, Massachusetts, Blaisdell Company, to be published, Chapter 9.

BLANK PAGE

CHAPTER SIX

INSIGHTS AND GENERALIZATIONS

6.1 Non Zero-Sum Properties of the Game

The outcome (actual payoff), $\mathcal{J}$, for the stochastic pursuit-evasion game studied in Chapter 2 is calculated at the conclusion of the play of the game by

$$\mathcal{J} = \frac{a^2}{2} \|y(t_f)\|^2 + \frac{1}{2} \int_{t_o}^{t_f} [\|u(t)\|^2_{R_p(t)} - \|v(t)\|^2_{R_e(t)}]dt \quad . \qquad (6:1)$$

This payoff can be viewed as the amount of dollars the pursuer gives the evader at the end of the game. In this sense the game is zero-sum; the actual payoff, once the game is completed, is a single number determined by (6:1).

The value of $\mathcal{J}$ is dependent on the actual control sequences u(t) and v(t), $t_o \leq t \leq t_f$, employed during the game. In general, however, the strategies U and V do not fix the values of these control sequences before that game begins, but rather assign values to them for every possible state of the game (i.e. for every possible spatial configuration of the two players) which is of course dependent on the realization of the white noise w(t). For the problem under consideration, then, the state cannot be predicted, for w(t) which corrupts the evader's measurements effects the state on which both player's control functions depend. Thus the numerical value of $\mathcal{J}$ which is realized at the end of the game is a function of: the initial state $y(t_o)$; the evader's initial estimate of the state, $\hat{y}_o$; the two strategies employed, U and V; and the actual value of the noise sequence, w(t), $t_o \leq t \leq t_f$.

$$\mathcal{J} = \mathcal{J}(y(t_o), \hat{y}_o, U, V, w(t):t_o \leq t \leq t_f) \quad . \tag{6:2}$$

Before the game starts and even during the play of the game, the outcome, $\mathcal{J}$, cannot be precisely determined.

For the particular pursuit-evasion game considered the two players have different information sets from which they can evaluate the state of the game. At time t, the pursuer knows the value of the state vector, y(t), while the evader has only the initial estimate $\hat{y}_o$ and the measurement of the state, $z(\tau)$, $t_o \leq \tau \leq t$. From this -- if the evader desires -- he can construct an estimate of the state,

$$\hat{y}(t) \triangleq E[y(t)/\hat{y}_o, z(\tau):t_o \leq \tau \leq t] \quad . \tag{6:3}$$

A problem immediately arises: since the two players have different information sets, they will each assess the value of the payoff, $\mathcal{J}$, differently. In particular at time t during the play of the game, the pursuer will expect the value of the outcome to be

$$J_p(t, U, V) = E[\mathcal{J}/y(\tau):t_o \leq \tau \leq t] \quad , \tag{6:4}$$

while the evader expects it to be

$$J_e(t, U, V) = E[\mathcal{J}/\hat{y}_o, z(\tau):t_o \leq \tau \leq t] \quad . \tag{6:5}$$

The pursuer's expectation is over the noise sequence $w(\tau)$, $t \leq \tau \leq t_f$, and is concerned with how it will effect the final value of the criterion through the evader's strategy. The evader's expectation is not only concerned with this noise sequence, but also with his assessment of the state $y(\tau)$, $t_o \leq \tau \leq t$.

Both players will probably be wrong. The actual outcome, $\mathcal{J}$, can only be determined after the game is completed; in fact, no one can determine $\mathcal{J}$'s value until he is given the noise sequence w(t), $t_o \leq t \leq t_f$. J_p

and J_e do represent, however, the pursuer's and evader's best estimate of $\mathcal{J}$. And what is important is that for the same strategy pair (U, V), J_p and J_e will in general not be equal.

Since at any time t for every strategy pair, (U, V), the two players will assign different values to the criterion, it appears that it will be necessary to solve a non zero-sum game. For if the pursuer knows that the evader's optimal strategy is V^o, then before the game begins he will look for a strategy U^o such that

$$J_p(t_o, U^o, V^o) \leq J_p(t_o, U, V^o) \quad . \tag{6:6a}$$

Conversely, the evader attempts to find a V^o such that

$$J_e(t_o, U^o, V) \leq J_e(t_o, U^o, V^o) \quad . \tag{6:6b}$$

Thus if the strategy pair (U^o, V^o) can be found which satisfies (6:6) then a solution has been obtained. The evader is satisfied since he has maximized his assessment of the criterion against U^o; the pursuer is also satisfied since he has minimized his assessment of the criterion against V^o. Thus U^o and V^o are in equilibrium on the basis of the two players' assessments of the payoff.

This is exactly the approach indicated in Section 1.1 as the appropriate one for solving non zero-sum games; equilibrium strategies which satisfy (1:5) provide the solution. However, this extra complexity -- besides being mentioned in Section 2.1 -- was not considered in detail in Chapter 2 when determining the "solution;" the distinction was ignored thus implicitly assuming that the solution was of the zero-sum form. It is the purpose of the first few sections of this chapter to point out that the answers given in Chapter 2 do in fact provide the solution to the non zero-sum game defined by (6:6).

It is important to note that this does not mean the game itself is non zero-sum. In the end the value of $\mathcal{J}$ as determined by (6:1) will be the same for both players, and the objectives of the players with respect to <u>this</u> payoff are completely in competition. However, because of the different information sets, the two players assess the end objective differently and thus the problem must really be solved from a non zero-sum viewpoint.

An analogous matrix game would be one in which the two players did not have perfect knowledge of the (monetary) value of the elements of a zero-sum payoff matrix. If they each thought the elements were different, they would attempt to do different things, though once their strategies had been declared and the actual values of the payoff elements revealed, any transfer of money would certainly be zero-sum. Unfortunately, this particular type of matrix game is not discussed in the literature.

The assessments of the criterion by the two opposing players are different only <u>during</u> the actual play of the game; before the game begins, the information sets -- identical for both players -- are the values of the system parameters and the nature of the different data that will be available <u>during</u> the game. The pursuer (and the evader) know that at the beginning of the game, the pursuer will be informed of the value of $y(t_o)$; but when determining his strategy the pursuer must design it in such a general way as to accommodate any value of $y(t_o)$. Consequently, it shall be assumed that before the game begins, J is calculated on the basis of (a) the system parameters, (b) two variables which specify in a probabilistic way the initial status of the game, $Y(t_o)$ and $P(t_o)$, where

$$Y(t) \triangleq E[y(t)y^T(t)] \quad , \tag{6:7}$$

$$P(t) \triangleq E[\tilde{y}(t)\tilde{y}^T(t)] \tag{6:8}$$

and (c) the knowledge of the two different information sets which will be employed <u>during</u> the game.

The game may still not be zero-sum in the sense that the two opponents assign different values to $Y(t_o)$ and $P(t_o)$. It has previously been assumed that the pursuer is aware of his opponent's system parameters, $F_e(t)$ and $\overline{G}_e(t)$. A second, equally plausible assumption is that the pursuer is also aware of the quality of the evader's information, i.e. the pursuer knows $P(t_o)$ and $Q(t)$. A critic may question the validity of this assumption for a particular problem, but not the answer derived from it.

Thus, before the game begins the expected value of the criterion for both players will be a function of the strategies, U and V, and the initial status of the game $Y(t_o)$ and $P(t_o)$, as given by

$$J_b \triangleq E[J/Y(t_o), P(t_o)] \quad . \tag{6:9}$$

The functional dependence of J is given by

$$J_b = J_b(U, V, Y(t_o), P(t_o)) \quad . \tag{6:10}$$

On the basis of this new evaluation of the criterion then, a solution is a pair of strategies U^o and V^o such that the following saddle-point condition is satisfied.

$$J_b(U^o, V) \leqslant J_b(U^o, V^o) \leqslant J_b(U, V^o) \quad . \tag{6:11}$$

Leaving the discussion of these non zero-sum properties of the stochastic pursuit-evasion game to this late chapter is justified because (a) it would have needlessly confused the presentation of Chapter 2, and (b) because the solution obtained there does indeed solve both the non zero-sum gam defined by (6:6) and the zero-sum game defined by (6:11). These solutions are repeated here.

$$U^o: \; u^o(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) - R_p^{-1}(t)G_p^T(t)\Gamma_2(t)\tilde{y}(t) \quad (6:12)$$

$$V^o: \; v^o(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)\hat{y}(t) \quad (6:13)$$

Consider first the non zero-sum game of (6:6). In Section 2.3, it was demonstrated that the feedback strategy (6:12) indeed minimized the criterion against the evader's optimal strategy (6:13). The assessment of the value of the criterion which was used here was based on the <u>pursuer's</u> information set, i.e. the minimization was performed with respect to J_p. Consequently, in Section 2.3 it was proven that (U^o, V^o) satisfies (6:6a).

In Section 2.4, the evader's feedback strategy which maximized the criterion against the pursuer's strategy (6:12), was shown to be (6:13). Here the criterion was evaluated on the basis of the <u>evader's</u> information set; the maximization was performed with respect J_e. Thus in Section 2.4, it was shown that (U^o, V^o) satisfies (6:6b).

Consequently, the results of Chapter 2 do provide a solution to the non zero-sum game defined by (6.6). This is true despite the fact that this aspect of the problem was not explicitly considered there.

Before the beginning of the game, the pursuer's evaluation of $\mathcal{J}$ is J_b. However, as soon as the pursuer is given the initial value of the state

$y(t_o)$, his evaluation changes to J_p. There is a discontinuity in the pursuer's assessment of $\mathcal{J}$ at time t_o.

No such discontinuity occurs for the evader's assessment. When (U^o, V^o) is employed, the value he assigns to the criterion is dependent only upon $Y(t_o)$ and $P(t_o)$. The value of $y(t_o)$ does influence his control $v(t_o)$, but not his assessment of the criterion. Consequently,

$$J_e(t_o, U^o, V^o) = J_b(U^o, V^o) \quad . \tag{6:14}$$

By proving in Section 2.4 that (6:6b) is satisfied, it was shown that the left-hand inequality of (6:11) is also satisfied.

To the pursuer, $J_b(U, V)$ is the expected value of $J_p(t_o, U, V)$ where the expectation is taken over the possible values of $y(t_o)$ and $\tilde{y}(t_o)$. Thus, since (6:6a) is true for all values of $y(t_o)$ and $\tilde{y}(t_o)$, then the right-hand inequality of (6:11) must also be satisfied.

Thus the results of Chapter 2 do really provide the solution to the stochastic pursuit-evasion game as defined by both (6:6) and (6:11).

6.2 The Criterion Reconsidered

Since, as it was pointed out in Section 6.1, the pursuer and the evader assign different values to the criterion, it is important to discover which assessment was considered in Section 2.5. Here the optimized value of the criterion is given by

$$J_b(U^o, V^o) = \mathrm{Tr}[\tfrac{1}{2}K^{-1}(t_f, t_o)Y(t_o) + \tfrac{1}{2}\Gamma_2(t_o)P(t_o)$$

$$+ \frac{1}{2}\int_{t_o}^{t_f} PH^TQ^{-1}HP\Gamma_2 dt] \quad . \tag{6:15}$$

This evaluation is based on the two variables, $Y(t_o)$ and $P(t_o)$, which describe probabilistically the initial status of the game. Consequently, (6:15) gives J_b and it is so denoted.

Certainly then $J_e(t_o)$ is also given by (6:15), see (6:14). In fact the analysis which produced (6:15) could be carried out for any time t. The point is that the evader's assessment of the criterion at t_o is independent of the value of y_o; it is based solely on the two variables $Y(t_o)$ and $P(t_o)$, which describe probabilistically the initial status of the game. Thus the generalization of (6:15) for any time gives the evader's evaluation of the criterion

$$J_e(t, U^o, V^o) = \mathrm{Tr}[\tfrac{1}{2}K^{-1}(t_f, t)Y(t) + \tfrac{1}{2}\Gamma_2(t)P(t) + \tfrac{1}{2}\int_t^{t_f} PH^TQ^{-1}HP\Gamma_2 dt] \quad . \tag{6:16}$$

Observe that for values of t other than t_o, the criterion evaluation (6:16) ignores the integral of the quadratic control terms from t_o to t; this part of the total criterion has already been determined and is thus of little significance to either player.

To determine the pursuer's assessment of the criterion during the game or once given the initial conditions $y(t_o)$ and $\tilde{y}(t_o)$, consider the form of the criterion given in Section 2.3, i.e. see (2:50). Using the pursuer's optimal control law as given by (2:56), (2:60), (2:61) and (2:62)

$$u(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) - R_p^{-1}(t)G_p^T(t)\Gamma_2(t)\eta(t) \quad , \tag{6:17}$$

this expression for the criterion becomes

$$J = \mathrm{Tr}\{\frac{a^2}{2}Y(t_f) + \frac{1}{2}\int_{t_o}^{t_f}[G_pR_p^{-1}G_p^T(K^{-1}YK^{-1} + K^{-1}M\Gamma_2 + \Gamma_2MK^{-1}$$

$$+ \Gamma_2N\Gamma_2) - K^{-1}T_eK^{-1}(Y - M - M^T + N)]dt\} \tag{6:18}$$

where

$$M(t) \triangleq E[y(t)\eta^T(t)/y(t_o), \eta(t_o)] \quad , \tag{6:19}$$

$$N(t) \triangleq E[\eta(t)\eta^T(t)/\eta(t_o)] \quad . \tag{6:20}$$

Using (6:17) in (2:49) gives a new system equation

$$\begin{bmatrix} \dot{y} \\ \eta \end{bmatrix} = \left[\begin{array}{c|c} -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1} & -(G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1} \\ \hline 0 & -G_pR_p^{-1}G_p^T(\Gamma_2 + K^{-1}) - PH^TQ^{-1}H \end{array}\right]\begin{bmatrix} y \\ \eta \end{bmatrix}$$

$$+ \left[\begin{array}{c} 0 \\ \hline PH^TQ^{-1} \end{array}\right] w \quad , \qquad \begin{bmatrix} y(t_o) \\ \eta(t_o) \end{bmatrix} = \begin{bmatrix} y_o \\ \tilde{y}_o \end{bmatrix} \tag{6:21}$$

from which the equation which determines Y(t), M(t), and N(t) is derived.

$$\left[\begin{array}{c|c} \dot{Y} & \dot{M} \\ \hline M^T & \dot{N} \end{array}\right] = \left[\begin{array}{c|c} -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1} & -G_pR_p^{-1}G_p^T\Gamma_2 - G_eR_e^{-1}G_e^TK^{-1} \\ \hline 0 & -G_pR_p^{-1}G_p^T(\Gamma_2 + K^{-1}) - PH^TQ^{-1}H \end{array}\right]$$

$$\times \left[\begin{array}{c|c} Y & M \\ \hline M^T & N \end{array}\right] + \left[\begin{array}{c|c} Y & M \\ \hline M^T & N \end{array}\right]$$

$$\times \left[\begin{array}{c|c} -K^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T) & 0 \\ \hline -\Gamma_2G_pR_p^{-1}G_p^T - K^{-1}G_eR_e^{-1}G_e^T & -(\Gamma_2 + K^{-1})G_pR_p^{-1}G_p^T - H^TQ^{-1}HP \end{array}\right]$$

$$+ \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & PH^TQ^{-1}HP \end{array}\right] ,$$

$$\left[\begin{array}{c|c} Y(t_o) & M(t_o) \\ \hline M^T(t_o) & N(t_o) \end{array}\right] = \left[\begin{array}{c} y_o \\ \hline \tilde{y}_o \end{array}\right][y_o^T \mid \tilde{y}_o^T] \tag{6:22}$$

From that part of (6:22) which determines Y(t), the following identically zero integral can be written.

$$0 = \int_{t_o}^{t_f} K^{-1}\{\dot{Y} + (G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}Y + YK^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)$$

$$+ (G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1})M^T + M(\Gamma_2G_pR_p^{-1}G_p^T$$

$$+ K^{-1}G_eR_e^{-1}G_e^T)\}dt \quad . \tag{6:23}$$

Integrating by parts results in

$$0 = a^2 Y(t_f) - K^{-1}(t_f, t_o)Y(t_o) + \int_{t_o}^{t_f} [K^{-1}YK^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)$$

$$+ K^{-1}(G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1})M^T$$

$$+ K^{-1}M(\Gamma_2G_pR_p^{-1}G_p^T + K^{-1}G_eR_e^{-1}G_e^T)]dt \quad . \tag{6:24}$$

Again from (6:22), this time from that part which determines N(t), another identically zero integral is constructed.

$$0 = \int_{t_o}^{t_f} \Gamma_2\{\dot{N} + [G_pR_p^{-1}G_p^T(\Gamma_2 + K^{-1}) + PH^TQ^{-1}H]N$$

$$+ N[(\Gamma_2 + K^{-1})G_pR_p^{-1}G_p^T + H^TQ^{-1}HP] + PH^TQ^{-1}HP\}dt$$

(6:25)

which, upon integrating by parts, becomes

$$0 = -\Gamma_2(t_o)N(t_o) + \int_{t_o}^{t_f} [-K^{-1}G_pR_p^{-1}G_p^T\Gamma_2N - H^TQ^{-1}HP\Gamma_2N$$

$$- K^{-1}G_eR_e^{-1}G_e^TN + \Gamma_2NH^TQ^{-1}HP + \Gamma_2N(\Gamma_2 + K^{-1})G_pR_p^{-1}G_p^T$$

$$- \Gamma_2PH^TQ^{-1}HP]dt \quad . \tag{6:26}$$

Taking the trace of (6:24) and of (6:26) and adding, the expression for the criterion, (6:18), reduces to

$$J = Tr[\tfrac{1}{2}K^{-1}(t_f, t_o)Y(t_o) + \tfrac{1}{2}\Gamma_2(t_o)N(t_o) + \tfrac{1}{2}\int_{t_o}^{t_f} PH^TQ^{-1}HP\Gamma_2dt] \quad .$$

(6:27)

Using the initial conditions from (6:22) and generalizing to any time t, the pursuer's evaluation of the optimized criterion can be written as

$$J_p(t, U^o, V^o) = \frac{1}{2}\|y(t)\|^2_{K^{-1}(t_f, t)} + \frac{1}{2}\|\tilde{y}(t)\|^2_{\Gamma_2(t)} + \frac{1}{2}\mathrm{Tr}\int_t^{t_f} PH^TQ^{-1}HP\Gamma_2 dt \quad . \tag{6:28}$$

In a sense then, a generalization of (2:88)

$$J = \mathrm{Tr}[\frac{1}{2}K^{-1}(t_f, t)Y(t) + \frac{1}{2}\Gamma_2(t)P(t) + \frac{1}{2}\int_t^{t_f} PH^TQ^{-1}HP\Gamma_2 dt] \tag{6:29}$$

can be used to evaluate the criterion from all points of view. The evader merely inserts his values of Y(t) and P(t) to make his calculation. The pursuer does the same if the game has not yet begun; once the game is in progress however, he employs

$$Y(t) = y(t)y^T(t) \quad , \tag{6:30}$$

$$P(t) = \tilde{y}(t)\tilde{y}^T(t) \tag{6:31}$$

which are -- on the basis of the pursuer's information set -- the correct values of these expectations.

6.3 The Irrelevance of Uniqueness

It was shown in Section 6.1, that the strategies U^o as given by (6:12) and V^o as given by (6:13) provide a solution to the stochastic pursuit-evasion game. They satisfy the saddle-point condition before the game begins,

$$J_b(U^o, V) \leq J_b(U^o, V^o) \leq J_b(U, V^o) \quad , \tag{6:11}$$

and they satisfy the equilibrium condition during the actual play of the game,

$$J_p(t, U^o, V^o) \leq J_p(t, U, V^o) \quad , \tag{6:6a}$$

$$J_e(t, U^o, V) \leq J_e(t, U^o, V^o) \quad . \tag{6:6b}$$

However, is this solution unique?

In particular, there might exist a pair of non-linear strategies, U* and V*, that provide a solution. In Section 2.3 all possible strategies for the pursuer -- both linear and non-linear -- were checked against V^o, and it was determined that there was no strategy, including a non-linear one, that was better than U^o. In Section 2.4, all the evader's possible strategies -- again both linear and non-linear -- were checked against U^o and it was also determined that there existed no strategy, not even a non-linear one, which was better than V^o. However, it was never determined that there was no pair of both non-linear strategies that satisfied the equilibrium conditions.

The question of uniqueness is not merely a mathematical nicety, but of fundamental importance. In control theory, any solution to a minimization problem is suitable as long as it provides the true minimum; the control theorist is completely happy with just one solution and cares little whether there are none or a hundred other solutions. In game theory, however, it is important to discover all the solutions or at least to characterize them all.

Consider what would happen if player one found the equilibrium solution (U^o, V^o), while his opponent by an entirely different thought process obtained the equilibrium solution (U^*, V^*). What would happen if the strategy pair (U^o, V^*) was played? Unless player one is convinced his solution is unique, he will have to worry about his opponent discovering a solution he did not even dream of. It is consequently important to either prove uniqueness, or determine some relationship between all the possible equilibrium points.

Suppose that there does exist a pair of non-linear strategies, U^* and V^*, that satisfy the saddle-point condition

$$J_b(U^*, V) \leq J_b(U^*, V^*) \leq J_b(U, V^*) \quad . \tag{6:32}$$

Then, from the left side of (6:32) and since all strategy pairs are playable

$$J_b(U^*, V^o) \leq J(U^*, V^*) \quad . \tag{6:33}$$

Also, inserting U^* into the right-hand inequality of (6:11) gives

$$J_b(U^o, V^o) \leq J_b(U^*, V^o) \quad . \tag{6:34}$$

From (6:33) and (6:34)

$$J_b(U^o, V^o) \leq J_b(U^*, V^*) \quad . \tag{6:35}$$

Again, inserting U^o into the right-hand side of (6:32) and V^* into the left-hand side of (6:11) gives

$$J_b(U^*, V^*) \leq J_b(U^o, V^*) \quad , \tag{6:36}$$

$$J_b(U^o, V^*) \leq J_b(U^o, V^o) \quad . \tag{6:37}$$

From these last two inequalities

$$J_b(U^*, V^*) \leq J_b(U^o, V^o) \quad . \tag{6:38}$$

On the basis of (6:35) and (6:38) it can be concluded that if there does exist a pair of non-linear strategies which produce a saddle-point (as in (6:32) above) the value of the criterion evaluated before the game begins at that saddle-point must be identical to the value obtained with the pair of linear strategies, U^o and V^o.

$$J_b(U^o, V^o) \equiv J_b(U^*, V^*) \quad . \tag{6:39}$$

In fact there will then be four saddle-points located at (U^o, V^o), (U^o, V^*), (U^*, V^o), and (U^*, V^*) with the value of the criterion the same for all four combinations of strategies.

$$J_b(U^o, V) \leq J_b(U^o, V^*) \leq J_b(U, V^*) \quad , \tag{6:40}$$

$$J_b(U^*, V) \leq J_b(U^*, V^o) \leq J_b(U, V^o) \quad . \tag{6:41}$$

The conclusion of all this is the well-known result of equivalence and interchangeability for zero-sum games. See Luce and Raiffa [1]. If two saddle-points exist their values are equivalent, and the strategies which give these saddle-points can be played interchangeably with still the same value of the criterion, as long as such combinations of strategies are admissible pairs.

$$J_b(U^o, V^o) = J_b(U^o, V^*) = J_b(U^*, V^o) = J_b(U^*, V^*) \quad . \tag{6:42}$$

Consequently, the question of uniqueness is irrelevant.

Unfortunately, there is no analogous procedure for non zero-sum games. Thus for this pursuit-evasion problem there is no general method to ascertain the nature of a pair of non-linear strategies. Assume, however, that the non-linear strategy pair (U^*, V^*) which satisfied (6:32) also satisfies the equilibrium condition at time t_o.

$$J_p(t_o, U^*, V^*) \leq J_p(t_o, U, V^*) \quad , \tag{6:43a}$$

$$J_e(t_o, U^*, V) \leq J_e(t_o, U^*, V^*) \quad . \tag{6:43b}$$

The question is how do the values $J_p(t_o, U^*, V^*)$ and $J_e(t_o, U^*, V^*)$ relate to $J_p(t_o, U^o, V^o)$ and $J_e(t_o, U^o, V^o)$.

First recall that at time t_o, the evader's evaluation of the criterion is not discontinuous

$$J_e(t_o, U, V) = J_b(U, V) \quad . \tag{6:44}$$

Thus from (6:42) and (6:44), it can be seen that

$$J_e(t_o, U^*, V^*) = J_e(t_o, U^o, V^*) = J_e(t_o, U^*, V^o) = J_e(t_o, U^o, V^o) \quad . \tag{6:45}$$

Consequently, the evader is not worried that the pursuer may find a different (non-linear) strategy pair than he does, and thus employ U^*. His expected value of the payoff does not change.

Also recall that the initial time t_o has no special characteristics to differentiate from the other values of t during the game. At any time t', the game can be temporarily halted, and the continuation of the game can be defined as an entirely new game. Thus the time t', at which the original game was stopped, becomes t_o for the newly defined game. Consequently, (6:45) can be generalized for all t.

Now consider the difficulty from the pursuer's point of view. Again assume that there exists a non-linear strategy pair (U^*, V^*) which satisfies both (6:32) and (6:43). The pursuer's first point of inquiry is whether his strategy U^o is minimizing against V^*, i.e. is the inequality

$$J_p(t_o, U^o, V^*) \leq J_p(t_o, U, V^*) \tag{6:46}$$

true? This is the primary question for the pursuer must be reassured that his strategy U^o is in equilibrium against a possible non-linear strategy that he has not discovered. After being so reassured, the pursuer would like to determine the relationship between the value of $J_p(t_o, U^o, V^*)$ and the value of $J_p(t_o, U^o, V^o)$.

Recall that $J_b(U, V)$ is merely the expected value of $J_p(t_o, U, V)$ where the expectation is taken over the possible values of the initial state $y(t_o)$. Thus the proper functional dependence of J_p is

$$J_p = J_p(t, U, V, y(t)) \quad . \tag{6:47}$$

Since from the original assumption of (6:43a) U^* minimizes J_p against V^*, the inequality (6:46) can be established if it can be shown that

$$J_p(t_o, U^o, V^*, y(t_o)) = J_p(t_o, U^*, V^*, y(t_o)) \quad . \tag{6:48}$$

Consider the set A consisting of values of $y(t_o)$ such that

$$J_p(t_o, U^o, V^*, y(t_o)) < J_p(t_o, U^*, V^*, y(t_o)) \quad , \tag{6:49}$$

and the complementary set B consisting of values of $y(t_o)$ such that

$$J_p(t_o, U^o, V^*, y(t_o)) \geq J_p(t_o, U^*, V^*, y(t_o)) \quad . \tag{6:50}$$

Then define the strategy U' which is equal to U^o when $y(t_o)$ is in A, and is equal to U^* when $y(t_o)$ is in B. Thus

$$J_p(t_o, U', V^*, y(t_o)) < J_p(t_o, U^*, V^*, y(t_o)) \quad . \tag{6:51}$$

Then taking the expectation of (6:51) over the possible initial state $y(t_o)$ gives

$$J_b(U', V^*) < J_b(U^*, V^*) \quad . \tag{6:52}$$

Since this contradicts the original assumption that U^* minimizes J_b against V^* (see the right-hand inequality of (6:32)) the conclusion is that the set A must be of measure zero.

Now consider two complementary subsets of B: the set C consisting of those values of $y(t_o)$ such that

$$J_p(t_o, U^o, V^*, y(t_o)) = J_p(t_o, U^*, V^*, y(t_o)) \tag{6:48}$$

and the set D which consists of those values of $y(t_o)$ such that

$$J_p(t_o, U^o, V^*, y(t_o)) > J_p(t_o, U^*, V^*, y(t_o)) \quad . \tag{6:53}$$

Now taking the expectation over all possible values of $y(t_o)$ in B gives

$$J_b(U^o, V^*) > J_b(U^*, V^*) \quad . \tag{6:54}$$

This violates the result given in (6:42), unless D is also of measure zero. Consequently, (6:48) is satisfied for all values of $y(t_o)$ except for a set of measure zero, and thus (6:46) is satisfied as well.

By the same general method it can be shown that

$$J_p(t_o, U^o, V^*, y(t_o)) = J_p(t_o, U^o, V^o, y(t_o)) \tag{6:55}$$

is true for all values of $y(t_o)$ except again for a set of measure zero. The generalization of these results for any time t during the game is the same as before.

Consequently, the pursuer is assured that if there does exist a pair of non-linear strategies (U^*, V^*) which does form a saddle-point pair for the zero-sum a priori problem, and an equilibrium pair for the in-flight problem, that his strategy U^o is both optimal against the resulting V^* and will give him the same (expected) criterion J_p.

Consequently, this game is also characterized by the properties of equivalence and interchangeability and thus the question of uniqueness is irrelevant. If another thought process produces a different solution, the results of the game played under those conditions will be the same and any crossing of strategies -- (U^o, V^*) or (U^*, V^o) -- does not effect the play for the unknowing player who sticks with his linear strategy.

The fact that these results could be obtained for the particular non zero-sum game considered -- when it is not true in general -- results from two facts. First of all recall that the <u>outcome</u> (actual payoff) of the game <u>is</u> zero-sum; if a transfer of payments is involved it is direct from one player to the other. Second, the initial assessments of the game by the two players were identical, thus permitting an expected value zero-sum game to be defined before the actual play began. The in-flight non zero-sum game is then related to this original game through the expected value operator. Using the a priori game as a base, and relating the in-flight game to it, permits the uniqueness-irrelevance conclusion.

6.4 The Certainty-Equivalence Principle

The optimal feedback strategies given in (6:12) and (6:13) indicate the applicability of the "certainty-equivalence principle" of optimal control theory to the stochastic differential game problem.

From (6:13) it can be seen that the evader merely uses the feedback strategy employed in the deterministic game to operate on his optimal estimate of the state y(t). Note that the estimation equation,

$$\dot{\hat{y}} = -[G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T] K^{-1} \hat{y} + P H^T Q^{-1} [z - H\hat{y}] \quad ,$$

$$\hat{y}(t_o) = \hat{y}_o \quad , \tag{6:56}$$

is the same estimation equation that would be employed by an imperfect outside observer of the deterministic game. However, because the pursuer employed a second, correction term in his control, the differential equation governing P is different.

$$\dot{P} = -G_p R_p^{-1} G_p^T (K^{-1} + \Gamma_2) P - P (K^{-1} + \Gamma_2) G_p R_p^{-1} G_p^T - P H^T Q^{-1} H P \quad ,$$

$$P(t_o) = P_o \quad . \tag{6:57}$$

$K^{-1}(t_f, t) + \Gamma_2(t)$ replaces $K^{-1}(t_f, t)$. Thus the evader must calculate $\Gamma_2(t)$. The coupled differential equations for Γ_2 and P may be solved simultaneously before the actual game begins, but the evader only needs to store the values of P(t), $t_o \leq t \leq t_f$.

For the pursuer, the optimal strategy consists of the same feedback control used in the deterministic game, plus an additional term to take advantage of the inaccuracy of the evader's knowledge of the state. From (6:12) it can be seen that whenever the evader's estimate of the state is exact, the pursuer's control will be exactly the same as in the deterministic case. Though the pursuer knows that the evader will deviate from the deterministic strategy in the future -- because of the noise in his future measurements -- the pursuer cannot utilize this information. The direction of this deviation is completely unpredictable; the measurement noise is zero mean.

The second term in the pursuer's strategy actually results from the immediate, non-deterministically optimal velocity vector of the evader. To the pursuer, the game has a 2k-dimensional state equation (2:95), governing the 2k-dimensional state vector $\begin{bmatrix} y(t) \\ \tilde{y}(t) \end{bmatrix}$. $\tilde{y}(t)$ is independent of y(t), as is shown in (2:93) which is repeated here.

$$\dot{\tilde{y}} = [-G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2) + PH^TQ^{-1}]\tilde{y} + PH^TQ^{-1}w \quad ,$$

$$\tilde{y}(t_o) = \tilde{y}_o \quad . \qquad (6:58)$$

When $\tilde{y}(t) \neq 0$, $\dot{y}(t)$ differs from the value it would have in the deterministic game; this is called a deviation from deterministic optimality. If the pursuer knows this deviation, he can predict values of the state vectors $y(\tau)$ and $\tilde{y}(\tau)$, $\tau > t$, and thus predict future deterministic non-optimalities.

To predict $\tilde{y}(\tau)$, $\tau > t$, w(t) is set equal to zero in (6:58) and the resulting differential equation is integrated forward in time from time t using $\tilde{y}(t)$ as the initial condition. Thus if $\tilde{y}(t)$ equals zero, any prediction of $\tilde{y}(\tau)$, $\tau > t$, will also be zero. Of course, this prediction is not perfect, for the actual equation for $\tilde{y}(t)$, (6:58), is driven by white noise w(t).

It should be noted that the evader cannot employ $\tilde{y}(t)$ in his controller. $\tilde{y}(t)$ is the error in the evader's estimate, and if the evader had any knowledge of this vector, he would use this knowledge to improve his estimate of y(t). By definition, the evader's estimate of $\tilde{y}(t)$ is zero.

6.5 General Measurements

The stochastic pursuit-evasion game solved in Chapter 2 constrained the evader to make noisy measurements on the state y(t) of the general form given by (2:5) which is repeated here.

$$z(t) = H(t)y(t) + w(t) \quad . \tag{6:59}$$

w(t) is zero-mean Gaussian white noise with variance Q(t).

The presentation of Section 2.1 considered the case where the evader's measurements on the states of the individual players were of the form:

$$z_1(t) = x_p(t) + w_1(t) \quad , \tag{6:60}$$

$$z_2(t) = x_e(t) + w_2(t) \quad . \tag{6:61}$$

This resulted in a measurement on the reduced state vector y(t) of the special form where the H of equation (6:59) was the identity matrix. This section demonstrates how the general measurement of (6:59) can be obtained.

Consider the general f-dimensional and g-dimensional measurements on the individual players' states

$$z_1 = \overline{H}_1 x_p + w_1 \quad , \tag{6:62}$$

$$z_2 = \overline{H}_2 x_e + w_2 \tag{6:63}$$

with w_1 and w_2 possessing the same characteristics as above. If either

$\overline{H}_1$ or $\overline{H}_2$ are not square, the necessary rows are added to produce square, non-singular* matrices H_1 and H_2. The corrupting noise of these additional measurements will be assumed to be white with variance of the form $q_1 I_{n-f}$ and $q_2 I_{n-g}$. After the analysis has been completed the scalars q_1 and q_2 can be permitted to approach infinity, thus eliminating any information gained from these additional measurements.

There now exist two new n-dimensional measurement vectors

$$z_1 = H_1 x_p + \overline{w}_1 \quad , \tag{6:64}$$

$$z_2 = H_2 x_e + \overline{w}_2 \tag{6:65}$$

where $\overline{w}_1$ and $\overline{w}_2$ are zero-mean, white noise vectors with spectra $\left[\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & q_1 I_{n-f} \end{array}\right]$ and $\left[\begin{array}{c|c} Q_2 & 0 \\ \hline 0 & q_2 I_{n-g} \end{array}\right]$ respectively. The cross correlation between $\overline{w}_1(t)$ and $\overline{w}_2(\tau)$ is $\left[\begin{array}{c|c} C\delta(t-\tau) & 0 \\ \hline 0 & 0 \end{array}\right]$.

A measurement of y(t) in the form given in (6:59) can be obtained by defining

$$z \triangleq A[\Phi_p H_1^{-1} z_1 - \Phi_e H_2^{-1} z_2] \quad , \tag{6:66}$$

where A is the matrix of the form $[I_k \mid 0]$ which weights the terminal separation of the two players in the original criterion (1:23) and Φ_p and Φ_e are the transition matrices associated with the individual dynamic systems of the pursuer (1:21) and the evader (1:22), respectively.

---

* If $\overline{H}_1$ (or $\overline{H}_2$) is square but singular, some of the resulting measurements are linearly dependent on others. Consequently, elements of the z_1 (or z_2) vector should either be combined to produce more accurate measurements, or deleted as superfluous.

Using (6:64) and (6:65) in conjunction with the definition of the reduced state (1:24), which is

$$y(t) \triangleq A[\Phi_p(t_f, t)x_p(t) - \Phi_e(t_f, t)x_e(t)] \quad , \tag{6:67}$$

(6:66) becomes

$$z = y + A[\Phi_p H_1^{-1}\overline{w}_1 - \Phi_e H_2^{-1}\overline{w}_2] \quad . \tag{6:68}$$

The zero-mean, white noise vector of (6:59) would be given by

$$w = A[\Phi_p H_1^{-1} w_1 - \Phi_e H_2^{-1} w_2] \tag{6:69}$$

with spectrum

$$Q = A\left\{\Phi_p H_1^{-1}\left[\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & q_1 I_{n-f}\end{array}\right] H_1^{T^{-1}} \Phi_p - \Phi_p H_1^{-1}\left[\begin{array}{c|c} C & 0 \\ \hline 0 & 0\end{array}\right] H_2^{T^{-1}} \Phi_e \right.$$

$$\left. - \Phi_e H_2^{-1}\left[\begin{array}{c|c} C^T & 0 \\ \hline 0 & 0\end{array}\right] H_1^{T^{-1}} \Phi_p + \Phi_e H_2^{-1}\left[\begin{array}{c|c} Q_2 & 0 \\ \hline 0 & q_2 I_{n-g}\end{array}\right] H_2^{T^{-1}} \Phi_e \right\} A \; . \tag{6:70}$$

(6:68) would indicate that the H matrix is I_k. However, when q_1 and q_2 are made to approach infinity, certain components of z, as given by (6:68), will become meaningless. Such components can be dropped, and upon rearranging the order of the elements in y, an H matrix of the form $[I \mid 0]$ is obtained. The variance matrix associated with w(t) in (6:59) is obtained from (6:70) by omitting all rows and columns which have diagonal elements that are functions of q_1 or q_2.

This procedure may produce a Q matrix in which all diagonal elements are dependent on q_1 or q_2, indicating that z of (6:68) contains no information pertinent to y. This is perfectly reasonable, since arbitrary measurements on x_p and x_e cannot necessarily be transformed into a useful measurement of y. It is always necessary to determine if specific measurements on x_p and x_e really permit the reduced dynamic system (2:1) to be observable.

6.6 Process Noise

Suppose the problem of Chapter 2 is further complicated by process noise in the dynamical system of (2:1)

$$\dot{y} = G_p(t)u(t) - G_e(t)v(t) + \theta(t) \quad . \tag{6:71}$$

Here $\theta(t)$ is a k-dimensional, zero-mean, Gaussian white noise vector with spectrum $\Theta(t)$.

The analysis of Chapter 2 is still valid with the only change in the results being in equations (2:96) and (2:97) which govern Y(t) and P(t).

$$\begin{aligned}\dot{Y} = &-(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}Y - YK^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T) \\ &- (G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}G_e^TK^{-1})P - P(\Gamma_2G_pR_p^{-1}G_p^T + K^{-1}G_eR_e^{-1}G_e^T) \\ &+ \Theta(t) \quad , \qquad Y(t_o) = Y_o \quad ,\end{aligned} \tag{6:72}$$

$$\begin{aligned}\dot{P} = &-G_pR_p^{-1}G_p^T(K^{-1} + \Gamma_2)P - P(K^{-1} + \Gamma_2)G_pR_p^{-1}G_p^T \\ &- PH^TQ^{-1}HP + \Theta(t) \quad , \qquad P(t_o) = P_o \quad .\end{aligned} \tag{6:73}$$

The optimal feedback strategies are still

$$U^o:\ u^o(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) - R_p^{-1}(t)G_p^T(t)\Gamma_2(t)\tilde{y}(t) \quad , \tag{6:12}$$

$$V^o:\ v^o(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)\hat{y}(t) \quad . \tag{6:13}$$

This is true because, although $\Theta(t)$ enters the Hamiltonian, when the Hamiltonian is differentiated with respect to either the controls C_p, C_e and D_p or the states Y and P the result is never a function of $\Theta(t)$.

However, the results of Chapter 3 are not valid. This is because the pursuer now cannot merely differentiate y(t), subtract $G_p(t)u(t)$ and multiply the results by $G_e^{-1}(t)$ to obtain the evader's control v(t). The noise vector, $\theta(t)$, which is unknown to the pursuer, prevents him from calculating the evader's control and thus the evader's estimation error. Consequently, though equilibrium strategies can be determined when the pursuer is provided with data on his opponent's estimation error by the mystical third party (Chapter 2), the pursuer cannot calculate this error (Chapter 3) in the presence of process noise.

Note that if the variance of the evader's measurement noise decreases to zero, optimal feedback strategies exist. Then (6:12) and (6:13) reduce to

$$U^o:\ u^o(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) \quad , \tag{6:74}$$

$$V^o:\ v^o(t) = -R_e^{-1}(t)G_e^T(t)K^{-1}(t_f, t)y(t) \tag{6:75}$$

which are the feedback strategies for the deterministic game. The pursuer no longer needs to calculate the evader's estimation error since it is zero.

For this case -- as in control theory -- the two player's merely ignore the process noise since it is zero-mean.

6.7 The Generalized Game

A more general linear stochastic differential game than the one of (2:1) and (2:2) in Chapter 2 can be defined and solved. The game equation has a non-zero system matrix F(t),

$$\dot{y} = F(t)y + G_p(t)u(t) - G_e(t)v(t) \quad , \qquad y(t_o) = y_o \tag{6:76}$$

and the in-flight criterion has a more generalized form.

$$J = E\left\{\frac{1}{2}\|y(t_f)\|^2_{\Gamma_f} + \frac{1}{2}\int_{t_o}^{t_f} [y^T \,\vdots\, u^T \,\vdots\, v^T] \left[\begin{array}{c:c:c} A & B & -C \\ \hdashline B^T & R_p & 0 \\ \hdashline -C^T & 0 & -R_e \end{array}\right] \begin{bmatrix} y \\ u \\ v \end{bmatrix} dt\right\} \tag{6:77}$$

Here again u(t) is the control of the player who attempts to minimize the criterion (6:77) while v(t) is the control of the player who attempts to maximize it. R_p and R_e are positive-definite and A is positive-semi-definite. This game has no direct relation to a pursuit-evasion problem; however, the minimizing player shall still be called the pursuer and the maximizing one the evader. The two zero submatrices of the partitioned weighting matrix of the in-flight criterion are so selected because it seems unreasonable that the penalty placed on one player's use of his own control should be a function of his opponent's control.

The evader is again assumed to have noisy measurements of the form

$$z(t) = H(t)y(t) + w(t) \tag{6:59}$$

where w(t) is a zero-mean white noise vector with spectrum $Q(t)$. The pursuer is again assumed to have perfect measurements of the state y(t).

Thus there are two different information sets, and thus two different definitions that can be placed on the expected value operator of (6:77). Yet, the purpose of this section is only to present the solution and -- since the analysis of Section 6.1 is valid for this more general problem too -- further elaboration of this point is not required.

Here the optimal feedback strategies are assumed to be of the form

$$u(t) = C_p(t)y(t) + D_p(t)\tilde{y}(t) \quad , \tag{6:78}$$

$$v(t) = C_e(t)\hat{y}(t) \quad . \tag{6:79}$$

The analysis to obtain the optimal values of C_p, D_p and C_e procedes in a manner identical to that of Section 2.2. The results are that:

$$C_p = -R_p^{-1}(G_p^T\Gamma_1 + B^T) \quad , \tag{6:80}$$

$$C_e = -R_e^{-1}(G_e^T\Gamma_1 + C^T) \quad , \tag{6:81}$$

$$D_p = -R_p^{-1}G_p^T\Gamma_2 \tag{6:82}$$

where

$$\dot{\Gamma}_1 = (BR_p^{-1}G_p^T - CR_e^{-1}G_e^T - F^T)\Gamma_1 + \Gamma_1(G_pR_p^{-1}B^T - G_eR_e^{-1}C^T - F)$$

$$- A + \Gamma_1(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)\Gamma_1 \quad ,$$

$$\Gamma_1(t_f) = \Gamma_f \quad , \tag{6:83}$$

and

$$\dot{\Gamma}_2 = [(\Gamma_1 G_p + B)R_p^{-1}G_p^T - F^T + H^TQ^{-1}HP]\Gamma_2$$

$$+ \Gamma_2[G_pR_p^{-1}(G_p^T\Gamma_1 + B) - F + PH^TQ^{-1}H]$$

$$+ \Gamma_2 G_pR_p^{-1}G_p^T\Gamma_2 + (\Gamma_1 G_e + C)R_e^{-1}(C^T + G_e\Gamma_1) \quad ,$$

$$\Gamma_2(t_f) = 0 \quad . \tag{6:84}$$

The evader's Kalman-Bucy filter for determining his estimate of the state is given by

$$\dot{\hat{y}} = [F - G_pR_p^{-1}(G_p\Gamma_1 + B^T) + G_eR_e^{-1}(G_e\Gamma_1 + C^T)]\hat{y}$$

$$+ PH^TQ^{-1}[z - H\hat{y}] \quad , \qquad \hat{y}(t_o) = \hat{y}_o \tag{6:85}$$

where P, the variance of the error of the evader's estimate, is obtained from

$$\dot{P} = [F - G_pR_p^{-1}(G_p^T\Gamma_1 + G_p^T\Gamma_2 + B^T)]P + P[F^T - (\Gamma_1 G_p + \Gamma_2 G_p + B)R_p^{-1}G_p^T]$$

$$- PH^TQ^{-1}HP \quad , \qquad P(t_o) = P_o \quad , \tag{6:86}$$

which is coupled to (6:84), the differential equation which determines Γ_2. Note, that when A, B, C, and F are all set equal to zero the above results all reduce appropriately to the corresponding results of Chapter 2.

Again the condition necessary for the pursuer to calculate the evader's estimate is that the inverse of $G_eR_e^{-1}G_e^T$ exist. The condition which is necessary for the pursuer to be able to <u>use</u> this data in his controller, as specified by (6:78), is that the inverse of $[G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}(G_e^T\Gamma_1 + C^T)]$ exist.

The additional conjugate point condition for the generalized stochastic problem is that the solution must have a finite $\Gamma_2(t)$, $t_o \leq t \leq t_f$. This is guaranteed by the finite initial condition on $P(t)$, (6:86).

When determining the evader's optimality, the condition that is necessary for the certainty-equivalence principle to hold is that

$$R_e - R_e(C + \Gamma_1 G_e)^{-1}\Gamma_2 G_p R_p^{-1} G_p^T \Gamma_2 (G_e^T\Gamma_1 + C^T)^{-1} R_e > 0 \ . \qquad (6:87)$$

The equation for determining $Y(t)$ is

$$\begin{aligned}\dot{Y} = {} & [F - G_pR_p^{-1}(G_p^T\Gamma_1 + B^T) + G_eR_e^{-1}(G_e^T\Gamma_1 + C^T)]Y \\ & + Y[F^T - (\Gamma_1 G_p + B)R_p^{-1}G_p^T + (\Gamma_1 G_e + C)R_e^{-1}G_e] \\ & + [G_pR_p^{-1}G_p^T\Gamma_2 + G_eR_e^{-1}(G_e^T\Gamma_1 + C^T)]P \\ & + P[\Gamma_2 G_pR_p^{-1}G_p^T + (C^T + \Gamma_1 G_e^T)R_e^{-1}G_e] \ , \\ & P(t_o) = P_o \ . \end{aligned} \qquad (6:88)$$

<u>References for Chapter Six</u>

[1] R. D. Luce and H. Raiffa, *Games and Decisions*, New York, John Wiley and Sons, Inc., 1957, p. 66.

BLANK PAGE

CHAPTER SEVEN

THE CLASSICAL INTERCEPTION PROBLEM

This chapter presents specific solutions and numerical results which illuminate the results of the previous chapters.

7.1 The Solution to the Classical Interception Problem

The classical interception problem in euclidean space, with the evader making noisy measurements, can be formulated in such a manner that the above analysis provides the solution. For this problem there are two state variables (position and velocity) and one control variable for each euclidean dimension. In general terms, the pursuer desires to capture the evader, i.e. to minimize the distance between them at the end of the game; the evader desires to escape, or to maximize this distance. Neither player cares about the difference in the velocities of the two players at the terminal time -- the game is one of interception, not rendezvous.

Since for each euclidean dimension there is one "interesting" state variable (position) and one control variable, condition i. of Section 3.1 is satisfied. It is shown below that conditions ii. and iii. are also satisfied.

The problem considered here is one in two euclidean dimensions; the extension to higher dimensions involves no extra complications. The pursuer's and evader's dynamical systems, respectively, are given by

$$\dot{x}_p = F_p x_p + \overline{G}_p u(t) \quad , \qquad x_p(t_o) = x_{po} \quad , \tag{7:1}$$

$$\dot{x}_e = F_e x_e + \overline{G}_e v(t) \quad , \qquad x_e(t_o) = x_{eo} \quad , \tag{7:2}$$

where

$$F_p = F_e = \left[\begin{array}{c:c} 0 & I_2 \\ \hdashline 0 & 0 \end{array}\right] , \tag{7:3}$$

$$\overline{G}_p = \overline{G}_e = \left[\begin{array}{c} 0 \\ \hdashline I_2 \end{array}\right] . \tag{7:4}$$

Thus both u(t) and v(t) are two-dimensional vectors.

The state vector of the pursuer is

$$x_p = \begin{bmatrix} x_{1p} \\ x_{2p} \\ x_{3p} \\ x_{4p} \end{bmatrix} \tag{7:5}$$

where x_{1p} and x_{2p} represent the pursuer's position in euclidean space and x_{3p} and x_{4p} the corresponding velocities. A similar statement is valid for the evader's state vector. Since $F_p = F_e$, the corresponding transition matrices are identical.

$$\Phi_p(t_f, t) = \Phi_e(t_f, t) = \left[\begin{array}{c:c} I_2 & (t_f - t)I_2 \\ \hdashline 0 & I_2 \end{array}\right] . \tag{7:6}$$

For this problem, the criterion is given by:

$$J = E\{\frac{a^2}{2}[(x_{1p} - x_{1e})^2 + (x_{2p} - x_{2e})^2]$$

$$+ \frac{1}{2}\int_{t_o}^{t_f} [\|u\|^2_{R_p} - \|v\|^2_{R_e}]dt\} . \tag{7:7}$$

Since the terminal separations $(x_{1p} - x_{1e})$ and $(x_{2p} - x_{2e})$ are the only ones of significance, the A matrix which reduces the dimension of the state vector is

$$A = [I_2 \,\vdots\, 0] \quad . \tag{7:8}$$

Thus from the definition (1:24) which gives the reduced state vector, y(t), in terms of x_p, x_e, Φ_p, Φ_e and A, it is found that

$$y(t) = [I_2 \,\vdots\, (t_f - t)I_2][x_p(t) - x_e(t)] \quad , \tag{7:9}$$

where the components of this two-dimensional state vector are:

$$y_1(t) = [x_{1p}(t) - x_{1e}(t)] + (t_f - t)[x_{3p}(t) - x_{3e}(t)] \quad , \tag{7:10}$$

$$y_2(t) = [x_{2p}(t) - x_{2e}(t)] + (t_f - t)[x_{4p}(t) - x_{4e}(t)] \quad . \tag{7:11}$$

(1:27) defines the control matrices for the reduced game, which for this problem are given by

$$G_p(t) = G_e(t) = (t_f - t)I_2 \quad . \tag{7:12}$$

It can be seen from (7:12) that the inverse of G_e exists for all $t < t_f$, thus satisfying condition ii. of Section 3. 1 which ensures that the pursuer can calculate the error in the evader's estimate.

It is now assumed that the energy weighting matrices have non-zero elements only along the diagonal, i. e.

$$R_p = \begin{bmatrix} r_{1p} & 0 \\ 0 & r_{2p} \end{bmatrix} \quad , \tag{7:13}$$

and similarly for R_e. Further R_p and R_e are assumed time independent.

Here observe that although the original problem is stationary (F_p, F_e, $\overline{G}_p$, $\overline{G}_e$, R_p, and R_e are all independent of time) the reduced problem is not. The time dependency of the two transition matrices causes the control matrices, for the reduced problem, G_p and G_e, to be functions of the time-to-go.

Then $K^{-1}(t_f, t)$ can be found to be

$$K^{-1} = \begin{bmatrix} \dfrac{3a^2 r_{1p} r_{1e}}{3r_{1p} r_{1e} + a^2(t_f - t)^3 (r_{1e} - r_{1p})} & 0 \\ 0 & \dfrac{3a^2 r_{2p} r_{2e}}{3r_{2p} r_{2e} + a^2(t_f - t)^3 (r_{2e} - r_{2p})} \end{bmatrix} \tag{7:14}$$

where in terms of the individual components, the following definition is found useful.

$$K^{-1}(t_f, t) \triangleq \begin{bmatrix} k_1^{-1}(t_f, t) & 0 \\ 0 & k_2^{-1}(t_f, t) \end{bmatrix} . \tag{7:15}$$

This particular form of K^{-1} is dependent on the fact that R_p and R_e are assumed to be independent of time; however, the fact that K^{-1} is diagonal depends only on the diagonal properties of R_p and R_e.

Thus the feedback gains for the deterministic game are given by

$$-R_p^{-1} G_p^T K^{-1} =$$

$$\begin{bmatrix} \dfrac{-3a^2 r_{1e}(t_f - t)}{3r_{1p}r_{1e} + a^2(t_f - t)^3(r_{1e} - r_{1p})} & 0 \\ 0 & \dfrac{-3a^2 r_{2e}(t_f - t)}{3r_{2p}r_{2e} + a^2(t_f - t)^3(r_{2e} - r_{2p})} \end{bmatrix} , \tag{7:16}$$

$$-R_e^{-1} G_e^T K^{-1} =$$

$$\begin{bmatrix} \dfrac{-3a^2 r_{1p}(t_f - t)}{3r_{1p}r_{1e} + a^2(t_f - t)^3(r_{1e} - r_{1p})} & 0 \\ 0 & \dfrac{-3a^2 r_{2p}(t_f - t)}{3r_{2p}r_{2e} + a^2(t_f - t)^3(r_{2e} - r_{2p})} \end{bmatrix} \tag{7:17}$$

The three matrices which define the quality of the evader's information are H(t), the evader's measurement matrix, Q(t), the covariance matrix of the measurement's noise, and P_o, the covariance of the error of the evader's initial estimate. Here, all three are assumed to be diagonal, and H and Q are further assumed to be time independent.

$$H \triangleq \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} , \tag{7:18}$$

$$\mathbf{Q} \triangleq \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} , \tag{7:19}$$

$$\mathbf{P} = \begin{bmatrix} P_{11o} & 0 \\ 0 & P_{22o} \end{bmatrix} . \tag{7:20}$$

Defining

$$\mathbf{P} \triangleq \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} , \tag{7:21}$$

$$\Gamma_2 \triangleq \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{bmatrix} , \tag{7:22}$$

the following set of simultaneous differential equations are found for the components of P(t) and $\Gamma_2(t)$ which are the matrices which define the nature of the evader's estimator and the second term in the pursuer's controller.

$$\dot{P}_{11} = -P_{11}\left[\frac{2(t_f - t)^2}{r_{1p}}(k_1^{-1} + \gamma_{11})\right] - \frac{P_{11}^2 h_1^2}{q_1} ,$$

$$P_{11}(t_o) = P_{11o} , \tag{7:23}$$

$$\dot{P}_{12} = 0 , \qquad P_{12}(t_o) = 0 , \tag{7:24}$$

$$\dot{P}_{22} = -P_{22}\left[\frac{2(t_f - t)^2}{r_{2p}}(k_2^{-1} + \gamma_{22})\right] - \frac{P_{22}^2 h_2^2}{q_2} \quad ,$$

$$P_{22}(t_o) = P_{22o} \quad , \tag{7:25}$$

$$\dot{\gamma}_{11} = 2\gamma_{11}\left[k_1^{-1}\frac{(t_f - t)^2}{r_{1p}} + \frac{P_{11}h_1^2}{q_1}\right] + \frac{\gamma_{11}^2(t_f - t)^2}{r_{1p}} + k_1^{-2}\frac{(t_f - t)^2}{r_{1e}} \quad ,$$

$$\gamma_{11}(t_f) = 0 \quad , \tag{7:26}$$

$$\dot{\gamma}_{12} = 0 \quad , \qquad \gamma_{12}(t_f) = 0 \quad , \tag{7:27}$$

$$\dot{\gamma}_{22} = 2\gamma_{22}\left[k_2^{-1}\frac{(t_f - t)^2}{r_{2p}} + \frac{P_{22}h_2^2}{q_2}\right] + \frac{\gamma_{22}^2(t_f - t)^2}{r_{2p}} + k_2^{-2}\frac{(t_f - t)^2}{r_{2e}} \quad ,$$

$$\gamma_{22}(t_f) = 0 \quad . \tag{7:28}$$

Note that (7:23) and (7:26) are coupled as are (7:25) and (7:28). Unfortunately, these cannot be solved in closed form.

Since the off-diagonal elements of Γ_2 are zero, the feedback matrix, with which the pursuer operates on the evader's estimation error, is given by

$$-R_p^{-1}G_p^T\Gamma_2 = \begin{bmatrix} \frac{-(t_f - t)}{r_{1p}}\gamma_{11} & 0 \\ 0 & \frac{-(t_f - t)}{r_{2p}}\gamma_{22} \end{bmatrix} \qquad (7:29)$$

with γ_{11} and γ_{22} determined by (7:26) and (7:27) respectively.

It is significant to observe that the feedback matrices are all diagonal and thus the resulting feedback control in one euclidean dimension is independent of the control and state in the other dimension. As defined above, the game is really two separate one-dimensional problems rather than a single two-dimensional game. The analysis and play can take place separately, and at different times.

In general then, a classical interception problem in n-dimensional euclidean space (with the stochastic properties defined above) can be reduced to n separate one-dimensional interception problems <u>if</u> there is no coupling in the dynamics and estimation between the various euclidean dimensions. The absence of coupling is guaranteed if:

a. The energy constraints placed on the controls do not involve cross products between the controls along different euclidean axes, i.e. R_p and R_e are diagonal.

b. The estimation process in one euclidean dimension is independent of those in the other dimensions, i.e. $P(t_o)$, H, and Q are diagonal.

It should now be recalled that it was shown in Section 3.4 that for a scalar game condition iii. of Section 3.1 was guaranteed to be satisfied for certain restrictions. These restrictions were that the time dependencies of $G_pR_p^{-1}G_p^T$ and $G_eR_e^{-1}G_e^T$ were the same, and that the relative controllability condition was satisfied. Indeed, the time dependencies of $G_pR_p^{-1}G_p^T$ and $G_eR_e^{-1}G_e^T$ are both $(t_f - t)^2$. Also, by assuming that

$$r_{1p} < r_{1e} \quad , \tag{7:30}$$

$$r_{2p} < r_{2e} \quad , \tag{7:31}$$

the relative controllability condition is certainly satisfied. Consequently, condition iii. -- which ensures that the inverse of L(t), defined by (3:8), exists and thus that the pursuer can actually utilize his calculation of the evader's estimation error in his controller -- is satisfied and the solution given here _is_ realizable.

It is worthwhile to note that for the scalar problem the individual values of P_{11}, h_1, and q_1 are not in themselves significant; only the factor $\frac{P_{11}h_1^2}{q_1}$ is important. This factor appears explicitly in the equation which determines γ_{11}, (7:26), and also in the evader's estimation equation; see (2:92). Multiplying (7:23) by $\frac{h_1^2}{q_1}$ and using the fact that this factor is time independent, a differential equation for $\frac{P_{11}h_1^2}{q_1}$ can be written.

$$\frac{d}{dt}\left[\frac{P_{11}h_1^2}{q_1}\right] = -\frac{P_{11}h_1^2}{q_1}\left[\frac{2(t_f - t)^2}{r_{1p}}(k_1^{-1} + \gamma_{11})\right] - \left(\frac{P_{11}h_1^2}{q_1}\right)^2 \quad ,$$

$$\frac{P_{11}(t_o)h_1^2}{q_1} = \frac{P_{11o}h_1^2}{q_1} \quad . \tag{7:32}$$

Thus this equation, in conjunction with (7:26), should be used to calculate the pertinent parameters, $\frac{P_{11}h_1^2}{q_1}$ and γ_{11}. The same is true for $\frac{P_{22}h_2^2}{q_2}$ and γ_{22}.

In the latter sections of this chapter, numerical examples for this game are given.

7.2 Proportional Navigation

One of the simplest strategies used in pursuit-evasion problems is the proportional navigation control law. This law gives a form for perturbation control for the pursuer when the pursuer and evader are on a nominal collision course of constant bearing, σ. That is, if the bearing of the pursuer's sighting on the evader, σ, deviates from its nominal value, the proportional navigation law states that the pursuer's optimal perturbation control in the direction perpendicular to his nominal flight path is of the form

$$u = n \cdot V_c \cdot \dot{\sigma} \quad . \tag{7:33}$$

Here V_c is the nominal closing (or relative) velocity between the two players, and n is the navigation constant. From experience the "best" values of n have been found to lie between 3 and 5. See Puckett and Ramo [1].

In Ho, Bryson and Baron [2], it was shown that the optimal (deterministic) pursuit strategy given in Chapter 1, reduced to the proportional navigation law when the players were on a collision course. Here, a^2 was permitted to approach infinity, indicating that the pursuer's objective is strictly capture while employing a minimum amount of energy. It was found that the navigation constant is given by

$$n = \frac{3}{(1 - r_p/r_e)} \quad . \tag{7:34}$$

Figure 7-1 displays the geometry of the proportional navigation problem in relative coordinates, i.e. in the coordinate system moving with the evader. (Consequently, the evader's nominal velocity is shown as zero.) The nominal value of the bearing σ is taken to be zero for convenience. Figure 7-1 also displays the positions and velocities of the two players perturbed from the nominal values.

For the deterministic case, the control to be applied by the pursuer in the direction perpendicular to the nominal line of sight is obtained from (7:16) and (7:11).

$$u_2^{do} = \frac{-3a^2 r_{2e}(t_f - t)[(x_{2p} - x_{2e}) + (t_f - t)(x_{4p} - x_{4e})]}{3r_{2p}r_{2e} + a^2(t_f - t)^3(r_{2e} - r_{2p})} \quad . \tag{7:35}$$

Letting a^2 approach infinity, (7:35) becomes

$$u_2^{do} = \frac{-3[(x_{2p} - x_{2e}) + (t_f - t)(x_{4p} - x_{4e})]}{(t_f - t)^2(1 - r_{2p}/r_{2e})} \quad . \tag{7:36}$$

Under the assumption that the deviations from the constant bearing are small, σ can be written as

$$\sigma = \frac{-(x_{2p} - x_{2e})}{V_c(t_f - t)} \quad . \tag{7:37}$$

Taking the time derivative of (7:37) gives

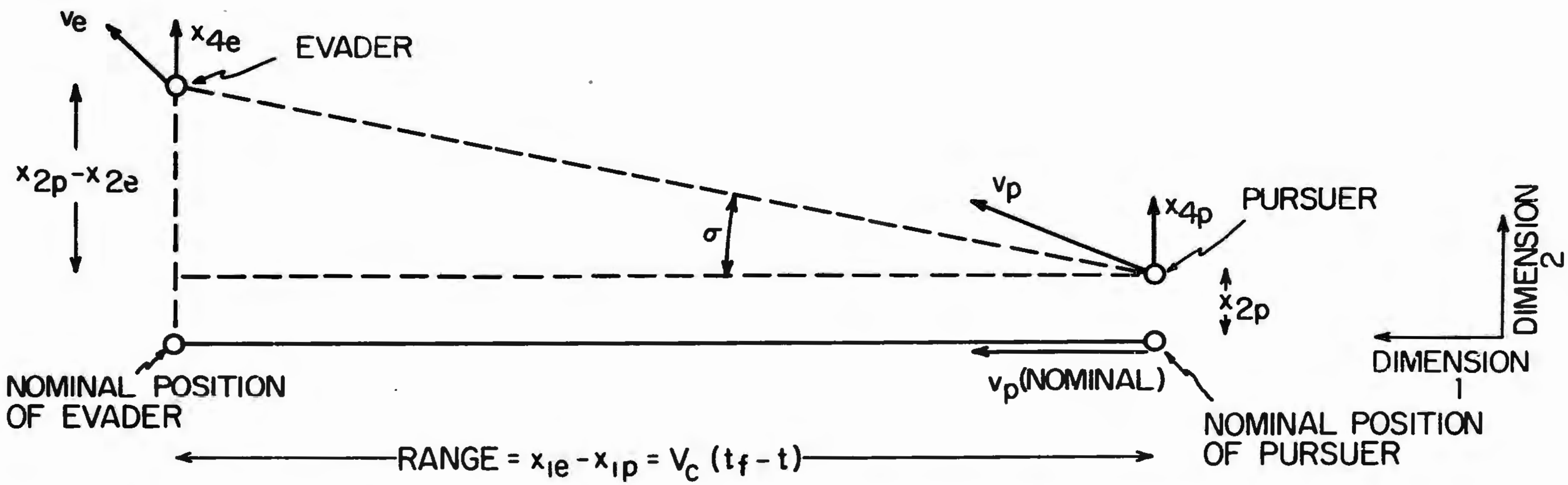


FIG. 7-1 THE GEOMETRY OF THE PROPORTIONAL NAVIGATION PROBLEM

$$\dot{\sigma} = \frac{-[(x_{2p} - x_{2e}) + (t_f - t)(x_{4p} - x_{4e})]}{(t_f - t)^2 V_c} \quad . \tag{7:38}$$

Now using (7:38) in (7:36) it can be seen that indeed u is of the form given in (7:33) with the navigation constant given by (7:34).

From (7:38) it can be seen that a non-zero $\dot{\sigma}$ can result from a number of causes. Either of the players could have been blown off course by an act of nature over which he had no control. The pursuer could have accidentally introduced some control perpendicular to the flight path (thus making x_{2p} and x_{4p}, the pursuer's position and velocity perpendicular to the nominal, non-zero), and his resulting perturbation control is an attempt to remedy this error. If these events do not occur, a non-zero $\dot{\sigma}$ can only result from a deviation of the evader from his optimal velocity in the direction perpendicular to the nominal flight path, x_{4e}; in this case, the pursuer's perturbation control takes advantage of the evader's deviation from his optimal <u>velocity</u>.

In Section 3.2 it was pointed out that the second term in the pursuer's control for the stochastic problem takes advantage of the evader's deviation from his optimal <u>acceleration</u>. This is not done when a standard proportional navigation strategy is used, since an acceleration deviation does not -- under ordinary circumstances -- permit predictions of future deviations.

Differentiating (7:38) with respect to time gives

$$\ddot{\sigma} = \frac{-(a_p - a_e)}{V_c(t_f - t)} + \frac{2\dot{\sigma}}{(t_f - t)} \tag{7:39}$$

where a_p and a_e are respectively the pursuer's and evader's accelerations in the direction perpendicular to the nominal flight path. From (7:4) it can be seen that a_p and a_e are exactly the pursuer's and evader's controls (u_2 and v_2) in this dimension. Thus if $\dot\sigma$ is zero, both a_p and a_e will be zero if the players behave optimally, and thus $\ddot\sigma$ will be zero -- a most reasonable result.

However, suppose that for some reason $\dot\sigma$ is not zero. Then a_p should be given by (7:33) in conjunction with (7:34), and a_e should be given by the "deterministically" optimal value of v_2,

$$v_2^{do} = \frac{-3[(x_{2p} - x_{2e}) + (t_f - t)(x_{4p} - x_{4e})]}{(t_f - t)^2(1 - r_{2e}/r_{2p})} = \frac{r_{2p}}{r_{2e}} u_2^{do} \quad . \tag{7:40}$$

However, if the evader's acceleration differs from this value by some error, $\tilde{a}_e$, where

$$\tilde{a}_e = v_2^{do} - a_e \quad , \tag{7:41}$$

then $\ddot\sigma$ will be given by

$$\ddot\sigma = \left[\frac{-(1 - r_{2p}/r_{2e})}{V_c} u_2^{do} + 2\dot\sigma + \frac{\tilde{a}_e}{V_c}\right]\frac{1}{(t_f - t)} \tag{7:42}$$

which reduces to

$$\ddot\sigma = [-\dot\sigma + \frac{\tilde{a}_e}{V_c}]\frac{1}{(t_f - t)} \quad . \tag{7:43}$$

(7:43) can be rewritten to give $\tilde{a}_e$ explicitly in terms of $\dot\sigma$ and $\ddot\sigma$.

$$\tilde{a}_e = V_c[(t_f - t)\ddot\sigma + \dot\sigma] \quad . \tag{7:44}$$

Now suppose that the evader's control (acceleration) differs from its deterministically optimal value because of noisy measurements which produce estimation errors. Then it can be seen that the second term in the pursuer's control will be proportional to the sum of (i) the first time derivative of the pursuer's bearing on his opponent and (ii) the second time derivative of this bearing times the time-to-go. It will be recalled from mechanics that acceleration is not proportionally related between rectangular and curvilinear coordinates; thus the $\ddot{\sigma}$ term in (7:44).

Because Γ_2 cannot be written explicitly, the second term of the pursuer's controller cannot be written in terms of the original parameters of the problem.

$$u = n \cdot V_c \cdot \dot{\sigma} - \frac{(t_f - t)\gamma_{22} V_c}{r_{2p}} [(t_f - t)\ddot{\sigma} + \dot{\sigma}] \quad . \tag{7:45}$$

(7:45) does <u>not</u> imply that if $\ddot{\sigma}$ is zero, the pursuer will add an extra correction for non-zero values of $\dot{\sigma}$. $\ddot{\sigma}$ and $\dot{\sigma}$ are intimately related by (7:39). If for some non-zero $\dot{\sigma}$, both players behave optimally, it can be seen from (7:44) that $\ddot{\sigma}$ will be given by $-\dot{\sigma}/(t_f - t)$. The second term in (7:45) will be non-zero only if $\ddot{\sigma}$ differs from this value, presumably because of the evader's estimation error.

There is a distinct similarity between the second term in (7:45) and the second term in the general pursuer's control as given by Realization II, of Section 3.1. In (3:14) the pursuer's second control term is a function of $[(G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T) K^{-1} y + \dot{y}]$. This function of y and $\dot{y}$ will be zero if the evader uses his deterministically optimal control. The fact that both y and $\dot{y}$ appear in this term certainly does not mean that they are independent; they are intimately related and under deterministically optimal conditions, this control term is zero. The same is true in (7:45).

7.3 Optimal Parameter Values and Feedback Gains

In this section a particular scalar pursuit-evasion problem is considered; it can be viewed as the game in either or both of the two euclidean dimensions analyzed in Section 7.1. Essentially this section is a parameter study, presenting numerical results* to illuminate the general nature of the results obtained in Chapters 2-6. This scalar problem is defined by the following numerical values which are assigned to the game parameters.

$$
\begin{aligned}
a^2 &= 10. \\
R_p &= 0.4 \\
R_e &= 1.0 \\
P_o &= 100. \\
t_o &= 0.0 \\
t_f &= 10.0 \\
H &= 1.0
\end{aligned}
\tag{7:46}
$$

The notation of the original problem (i.e. the capital letters) rather than

---

* All calculations were done on MINIC, a digital computer program for solving systems of differential equations on an IBM 7090/7094 computer with a Fortran IV IBSYS Monitor.

the scalar notation of Section 7.1 is used in this section, and units have been omitted since any consistent set can be used. The parameter Q is permitted to vary over several orders of magnitude, $1.0 \leq Q \leq 10.^5$.

Observe that there is no loss in generality in assuming that H has the value of 1.0. Recall that for any scalar estimation problem where information is obtained from the scalar measurement

$$z = Hy + w \quad , \tag{7:47}$$

the quality of the information obtained is completely determined by the ratio of H^2 to the variance of the measurement noise w, i.e. H^2/Q. If H is increased while this ratio remains constant, only the magnitude of the measurement is increased, not its information content. Thus for the scalar case, the quality of the measurement information can be varied over the entire range of possibilities by fixing H, and allowing Q to vary from zero to infinity.

Recall that at the end of Section 7.1 it was pointed out t[illegible]t the only pertinent combination of P and Q, to either the evader or pur[illegible]er when implementing their strategies, is P/Q (since H has already been set equal to one). Thus setting P_o equal to 100. does not represent any loss in generality either. The full spectrum of possibilities of the character of the information can be investigated by varying Q; this is done below for the range $1.0 \leq Q \leq 10.^5$. For cases when other values of P_o are needed, appropriate scaling from the calculations below will produce the desired result. This value of P means that the standard deviation of the evader's initial estimate is 10., a not unreasonable value for initial predicted terminal separations, $y(t_o)$, on the order of 50.

Note that R_p is less than R_e making the pursuer more controllable than the evader and thus ensuring that no deterministic conjugate point exists. Further observe that a^2 is an order of magnitude larger than both R_p and R_e.

In the previous section it was mentioned that as a^2 approaches infinity, the goal for the pursuer becomes strictly one of capture. For $a^2 = 10.0$, simple integration of the deterministic game equation shows that when both players employ their optimal strategies, $y(t_f)$ is $2.0 \times 10^{-4} y(t_o)$. Thus it can be seen that moderately large values of a^2 can ensure that capture "almost" occurs. Here the evader -- because he is less controllable than his opponent -- employs a strategy which 1) prevents capture from occuring till the terminal time, and 2) makes sure that the pursuer uses the maximum amount of energy necessary for capture.

Figure 7-2 displays $K^{-1}(t_f, t)$ and $\Gamma_2(t)$ for the scalar game with the parameters as given in (7:46). One should not jump to the hasty conclusion that -- because K^{-1} is so large near the terminal time -- most of the control in the deterministic case will be applied late in the game. Recall that the pursuer's feedback gain matrix, C_p, is composed of three factors in the form, $-R_p^{-1} G_p^T K^{-1}$. R_p is a constant and G_p is a linear function of the time-to-go, i.e. it is a linearly decreasing function of time. Thus the value of C_p, see Figure 7-3, is not <u>as</u> lopsidedly weighted at the end of the game. Also, since the pursuer is more controllable than the evader, y(t) will be larger at the beginning of the game,

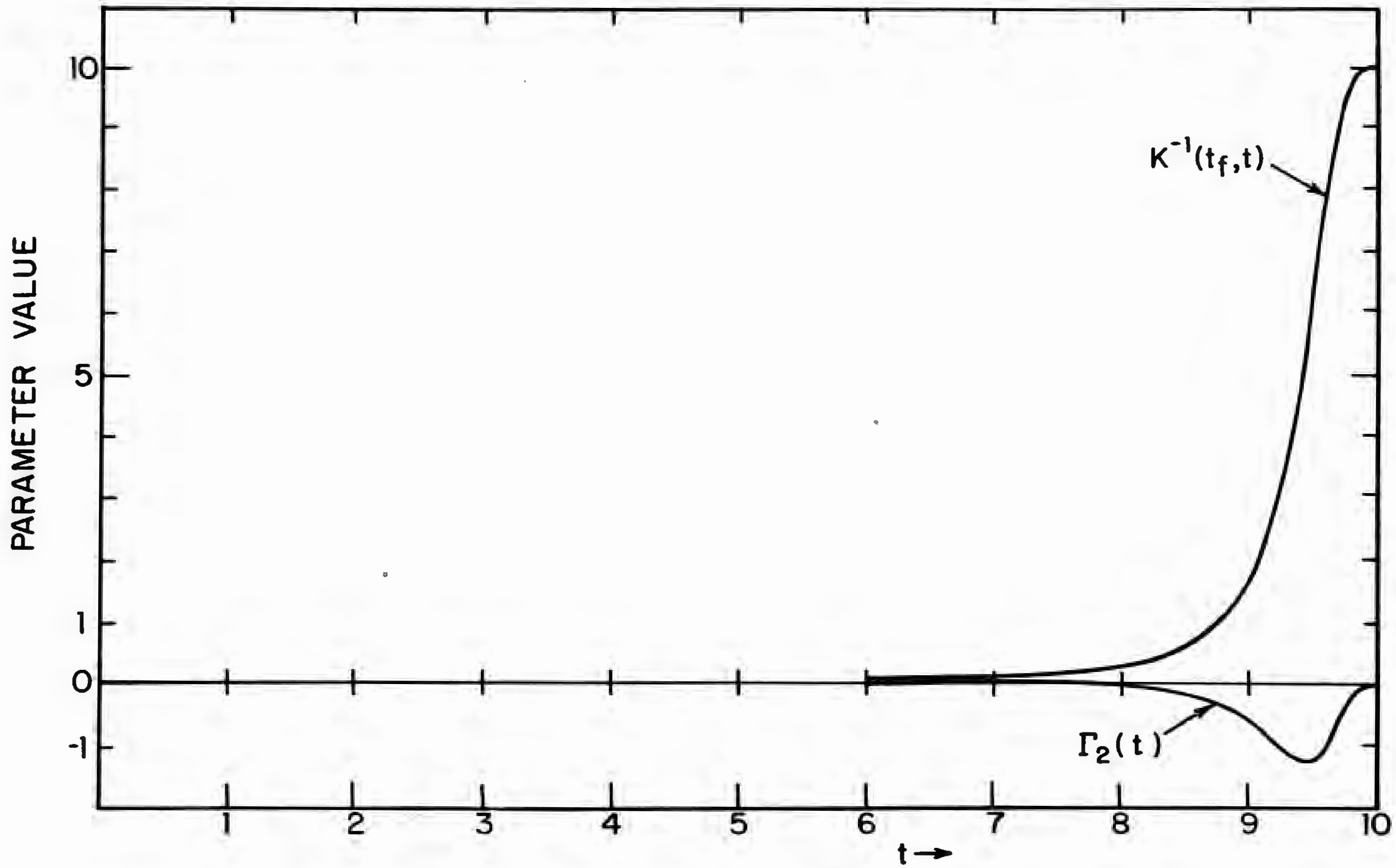


FIG. 7-2 THE PARAMETERS $K^{-1}(t_f, t)$ AND $\Gamma_2(t)$

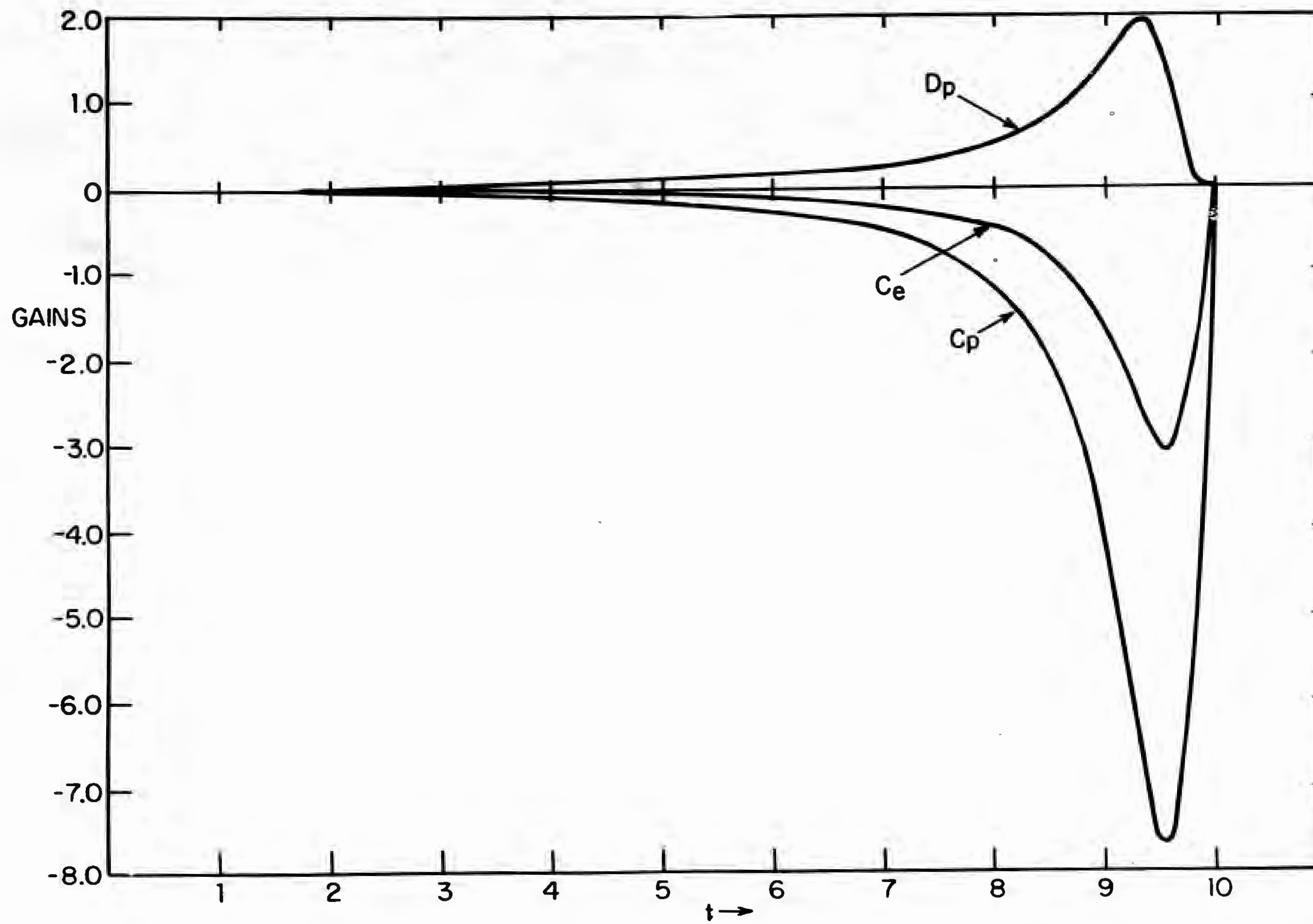


FIG. 7-3 THE FEEDBACK GAINS

further indicating that the control itself will not be applied exclusively near the end of the game. This discussion is continued in Section 7.5.

The curve for Γ_2 in Figure 7-2 is appropriate for all values of Q in the range 1.0 to 10^5. Γ_2 is completely independent of Q near the terminal time and varies by less than 1% at t = 5 for the values of Q indicated. Only near the beginning of the game does Q appreciably effect Γ_2; however, because Γ_2 is so small during this part of the game this difference cannot be displayed on a linear graph. The insensitivity of Γ_2 to Q means that, as given in Figure 7.3, the curve of D_p -- the pursuer's second feedback gain for the stochastic problem -- is appropriate for all values of Q.

The reason that Γ_2 is so independent of Q can be seen from Figure 7-4, which displays P(t) for various values of the parameter Q. The curve for $Q = 10^5$ is not displayed because it falls too close to the one for $Q = 10^4$ for a visual distinction to be made, while the one for Q = 1.0 the curve is very near the axes. Now recall from Section 4.2 that for Q large the evader is essentially operating open-loop; consequently, the curve for $Q = 10^4$ represents the upper bound of P. Thus during the second half of the game P(t) is very small no matter how large Q is; the P equation is very stable. Also -- as was pointed out at the end of Section 7.1 -- Q only effects Γ_2 through the term P/Q; see (7:32). Thus for P(t) very small, Q cannot have a large effect on Γ_2.

Figure 7-5 displays L(t). Defined as $G_p R_p^{-1} G_p^T \Gamma_2 + G_e R_e^{-1} G_e^T K^{-1}$ in Chapter 3, (3:8), this term must be positive (definite) for the pursuer's

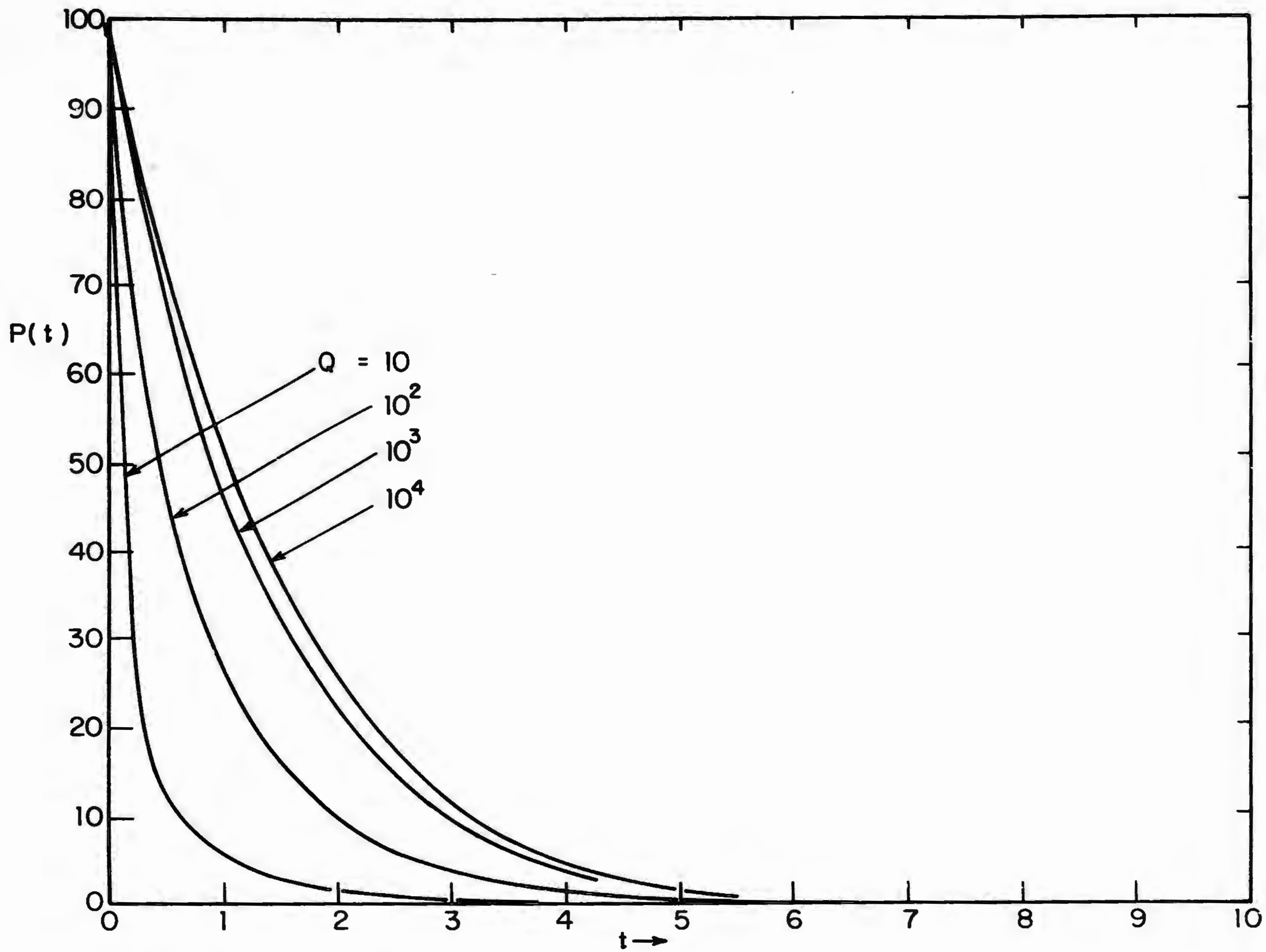


FIG. 7-4 THE VARIANCE OF THE ERROR OF THE EVADER'S ESTIMATE

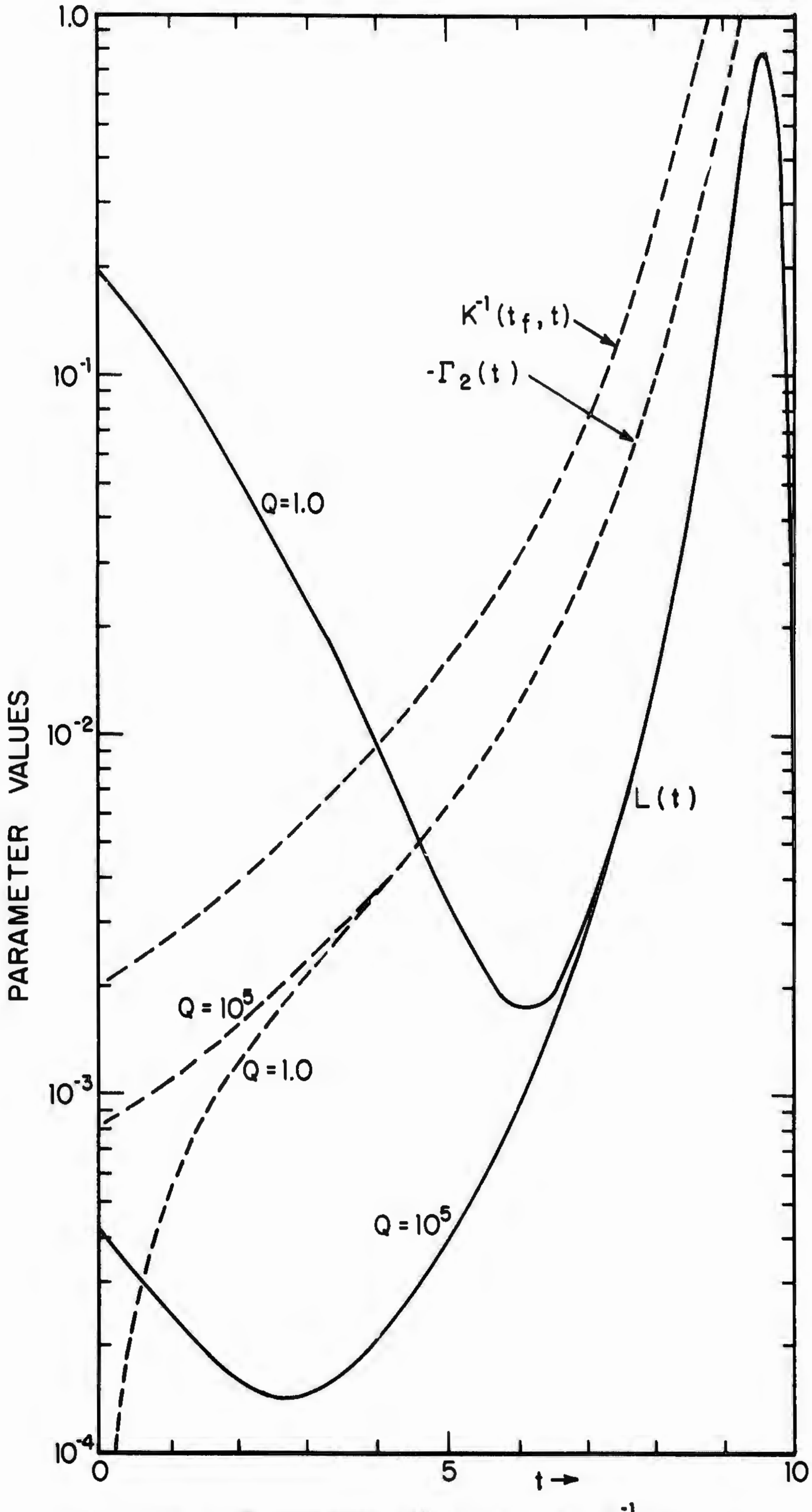


FIG. 7-5 THE PARAMETER L(t) WITH $K^{-1}(t_f, t)$ AND $-\Gamma_2(t)$

strategy to be realizable; note that it <u>is</u> always positive. This is a further graphical demonstration of the fact that L^{-1} always exists for the scalar game when $G_pR_p^{-1}G_p^T$ and $G_eR_e^{-1}G_e^T$ have the same time dependency. (See Section 3.4 for proof.)

$K^{-1}(t_f, t_o)$ and two curves for $-\Gamma_2(t)$ are also displayed in Figure 7-5. When viewing time as progressing backwards from t_f, observe that L(t) displays the effect of different values of Q before $\Gamma_2(t)$ does. This is because the $\Gamma_2(t)$ in L(t) is multiplied by $G_pR_p^{-1}G_p^T$ which for this problem is $2.5(t_f - t)^2$; this factor becomes quite large for $t_f - t$ large. Thus when adding $G_pR_p^{-1}G_p^T\Gamma_2$ and $G_eR_e^{-1}G_e^TK^{-1}$, which have nearly opposite values, small differences in $\Gamma_2(t)$, resulting from different values of Q, will show up sooner.

For this scalar game, the expression for the criterion is

$$J = \frac{1}{2}K^{-1}(t_f, t_o)Y(t_o) + \frac{1}{2}\Gamma_2(t_o)P(t_o) + \frac{1}{2}\int_{t_o}^{t_f} \frac{P^2(t)\Gamma_2(t)}{Q}\, dt \qquad (7:48)$$

where values of $Y(t_o)$ and $P(t_o)$ can be appropriately assigned by either player depending on his data. Since $Y(t_o)$ can be selected without effecting the gains, the second two terms alone represent the change in J because of the presence of measurement noise. Consequently, a reduced representation of J, called J_r or the reduced criterion, indicates more appropriately the dependence of J on Q.

$$J_r = \frac{1}{2}\Gamma_2(t_o)P(t_o) + \frac{1}{2}\int_{t_o}^{t_f} \frac{P^2(t)\Gamma_2(t)}{Q}\, dt \quad . \qquad (7:49)$$

From Figure 7-6, which displays J_r as a function of Q, several observations about the criterion can be made. J_r is always negative, indicating a reduction in the evader's or maximizing player's capability. For small measurement noise, J_r is approaching zero, for in this region the evader has almost perfect knowledge of the state, and thus deviates little from his deterministically optimal control. For large values of Q, J_r asymptotically approaches a constant; this reflects the fact that the evader is operating open-loop for large values of Q; consequently, the pursuer cannot take extra advantage of an increase in the variance of the measurement noise from $Q = 10^4$ to $Q = 10^5$ for the evader's control is now solely a function of his initial estimate.

Figure 7-7 displays the pursuer's gains G_1 and G_2 for $Q = 100$. These are employed by the pursuer in Realization II where his control is given by

$$u(t) = G_1(t)y(t) + G_2(t)\dot{y}(t) \quad . \tag{7:50}$$

Comparing this figure with Figure 7-3, it can be seen that G_1 and G_2 have values much larger than C_p and D_p. This does not mean that the control applied for this realization is greater. Since $G_p R_p^{-1} G_p^T$ is larger than $G_e R_e^{-1} G_e^T$, the deterministically optimal $\dot{y}$, as given by

$$\dot{y}^{do}(t) = -(t_f - t)^2\left(\frac{1}{R_p} - \frac{1}{R_e}\right)K^{-1}(t_f, t)y(t) \quad , \tag{7:51}$$

will be negative when y is positive, and unless the evader's estimate is very bad, the stochastically optimal $\dot{y}$ will also have the opposite sign from y. Consequently, the two terms in (7:50) are also of opposite sign, producing the appropriate control.

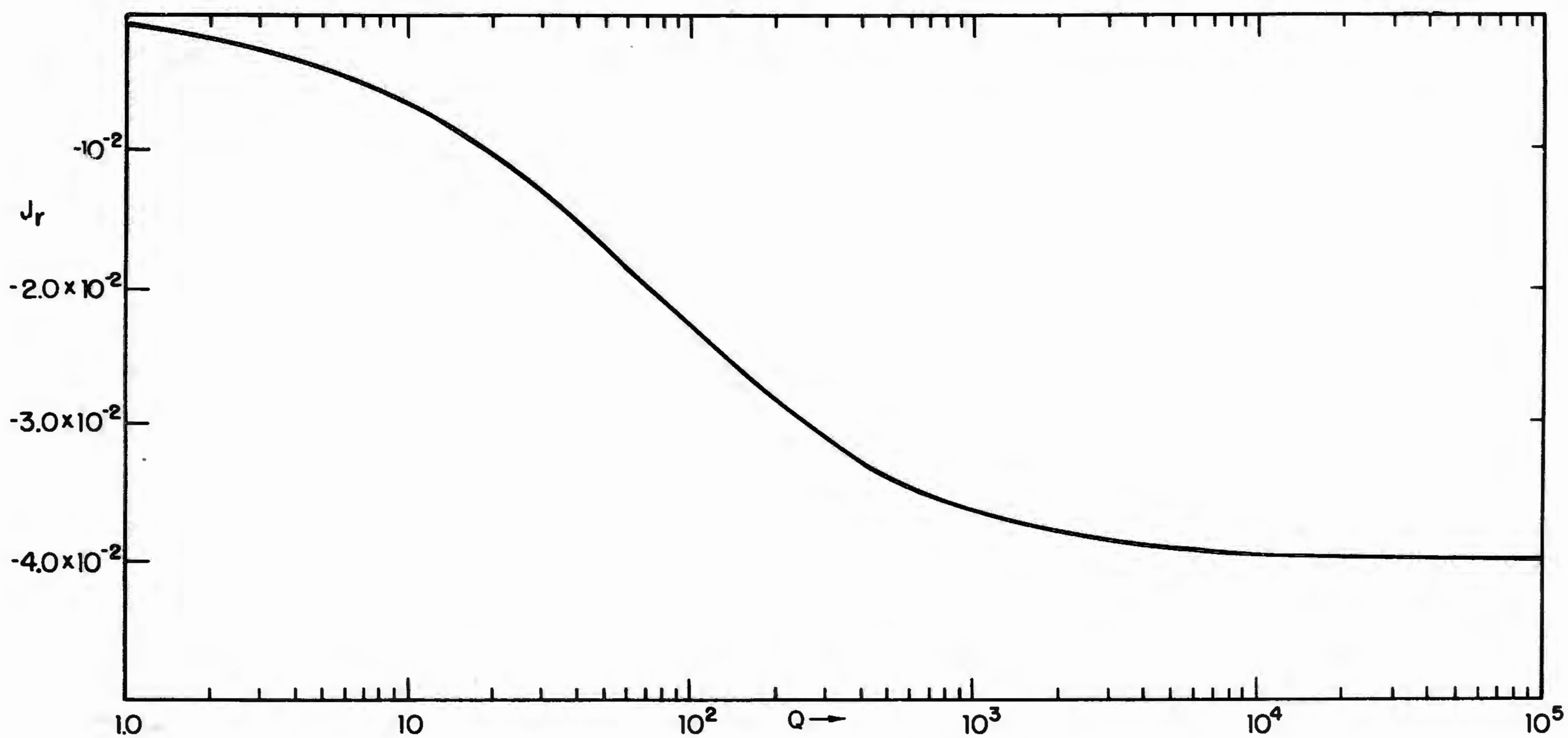


FIG 7-6 THE RELATIVE CRITERION AS A FUNCTION OF THE EVADER'S MEASUREMENT NOISE VARIANCE.

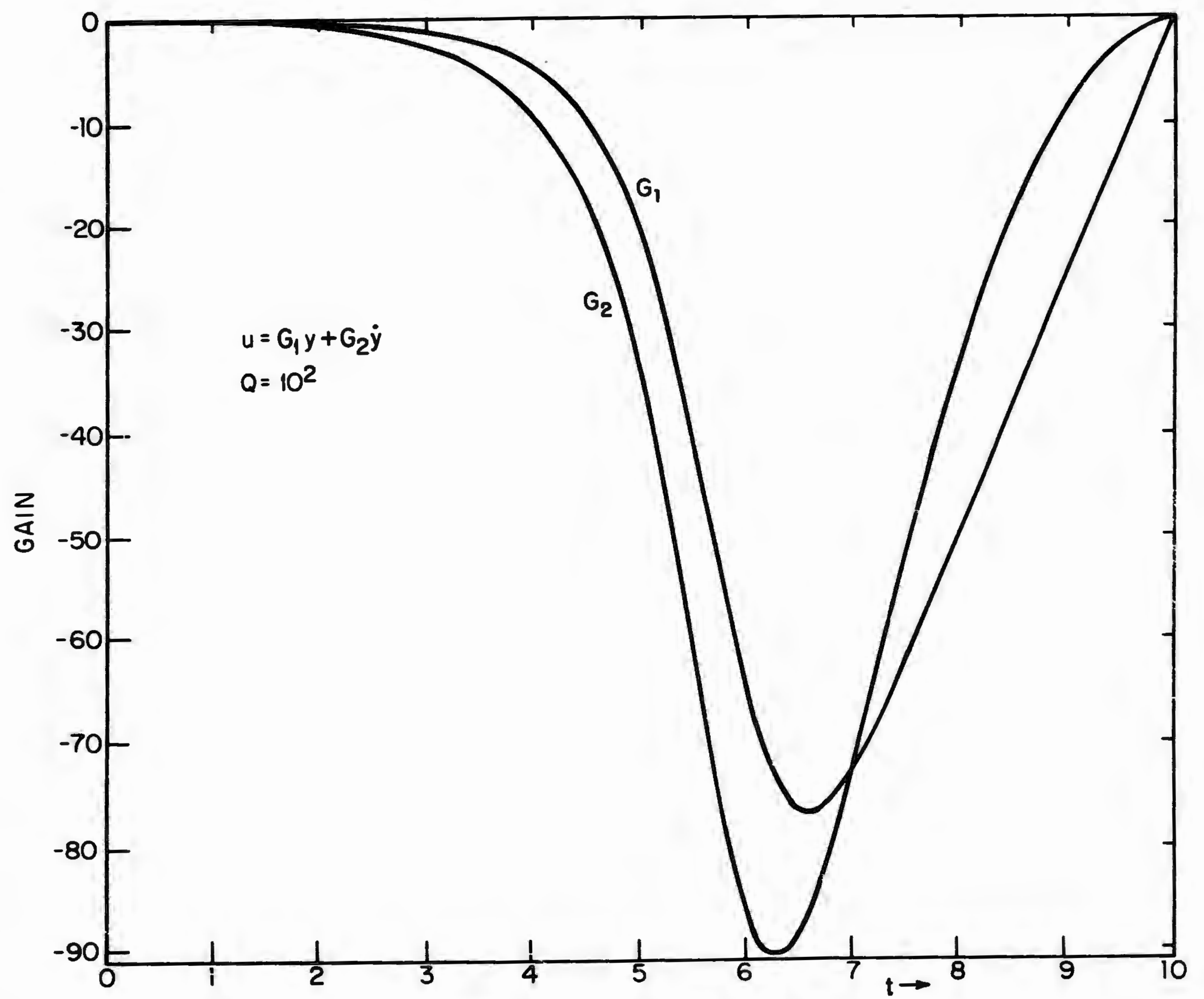


FIG. 7-7 THE PURSUER'S GAINS G_1 & G_2 FOR REALIZATION II

Figures 7-8 and 7-9 display families of curves for G_1 and G_2 which are dependent on Q. These curves display substantial differences as a function of Q, despite the fact that D_p in Figure 7-3 does not. However, observe the similarity between the curve families and recall the fact that the two terms in (7:50) are subtracted from one another. This means that very large differences in the two individual curve families do not result in very large control differences.

These curve families do not imply that Realization II is so sensitive to computation errors as to destroy the utility of this realization. G_1 can be thought of as consisting of two terms -- from (3:11)

$$G_1 = -R_p^{-1}G_p^TK^{-1} + R_p^{-1}G_p^T\Gamma_2L^{-1}(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}$$

$$= G_1' + G_1'' \qquad (7:52)$$

-- the first of which (G_1') is the standard deterministic gain for y. The second term (G_1'') multiplied by y equals the negative of the result of G_2 multiplied by the deterministically optimal $\dot{y}$. Thus the difference, $G_1''y - G_2(-\dot{y})$, is certainly sensitive to deviations of $\dot{y}$ from its deterministically optimal value, but this is its purpose.

7.4 A Numerical Study of $\Gamma_2(t)$ and P(t)

This section consists of a compilation of numerical data on the equations which govern Γ_2 and P for the scalar game, and consequently which determine the character of the solution for this problem. For the game with the evader making noisy measurements these coupled quadratic equations are:

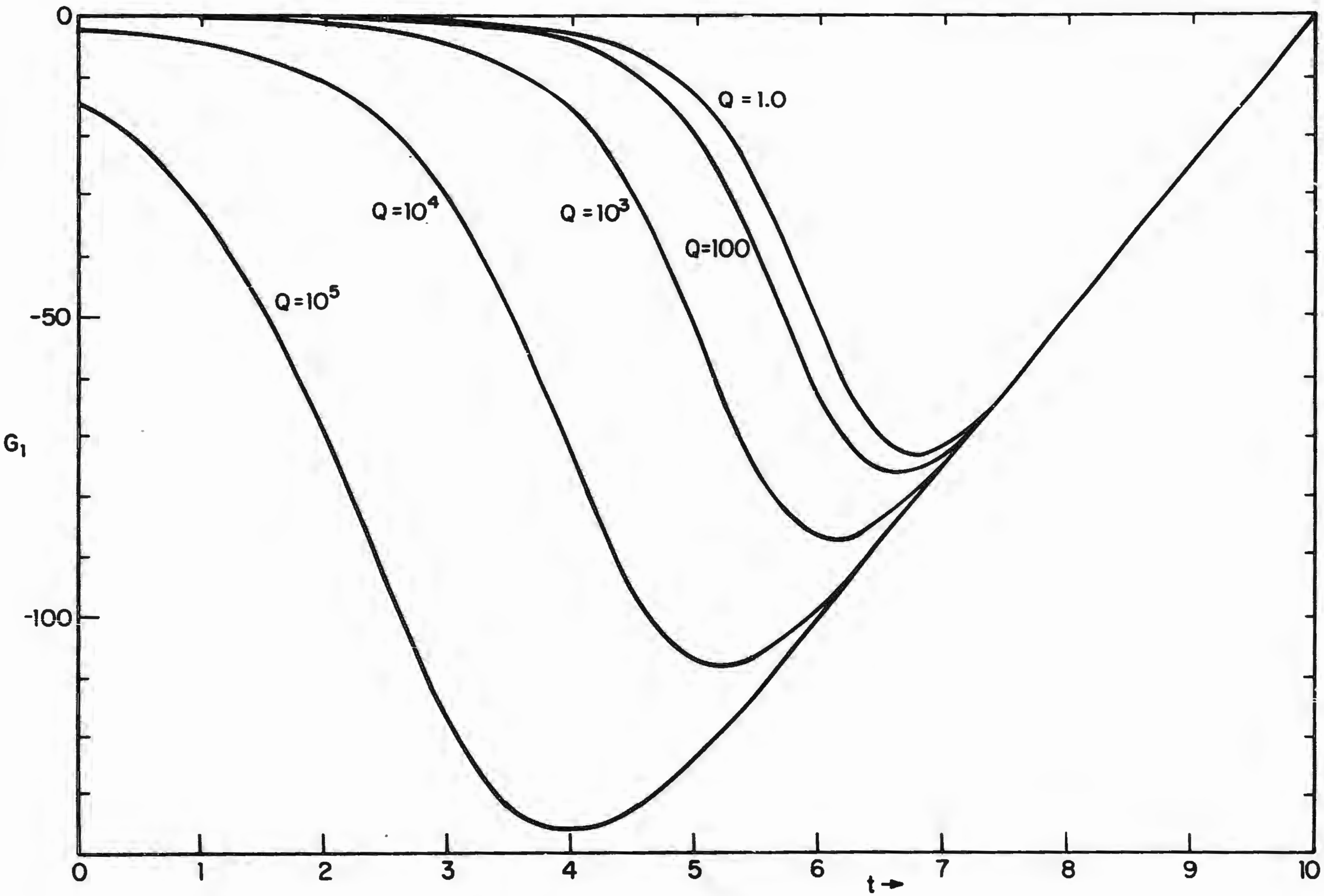


FIG. 7-8 FEEDBACK GAIN G_1 FOR THE PURSUER'S REALIZATION II

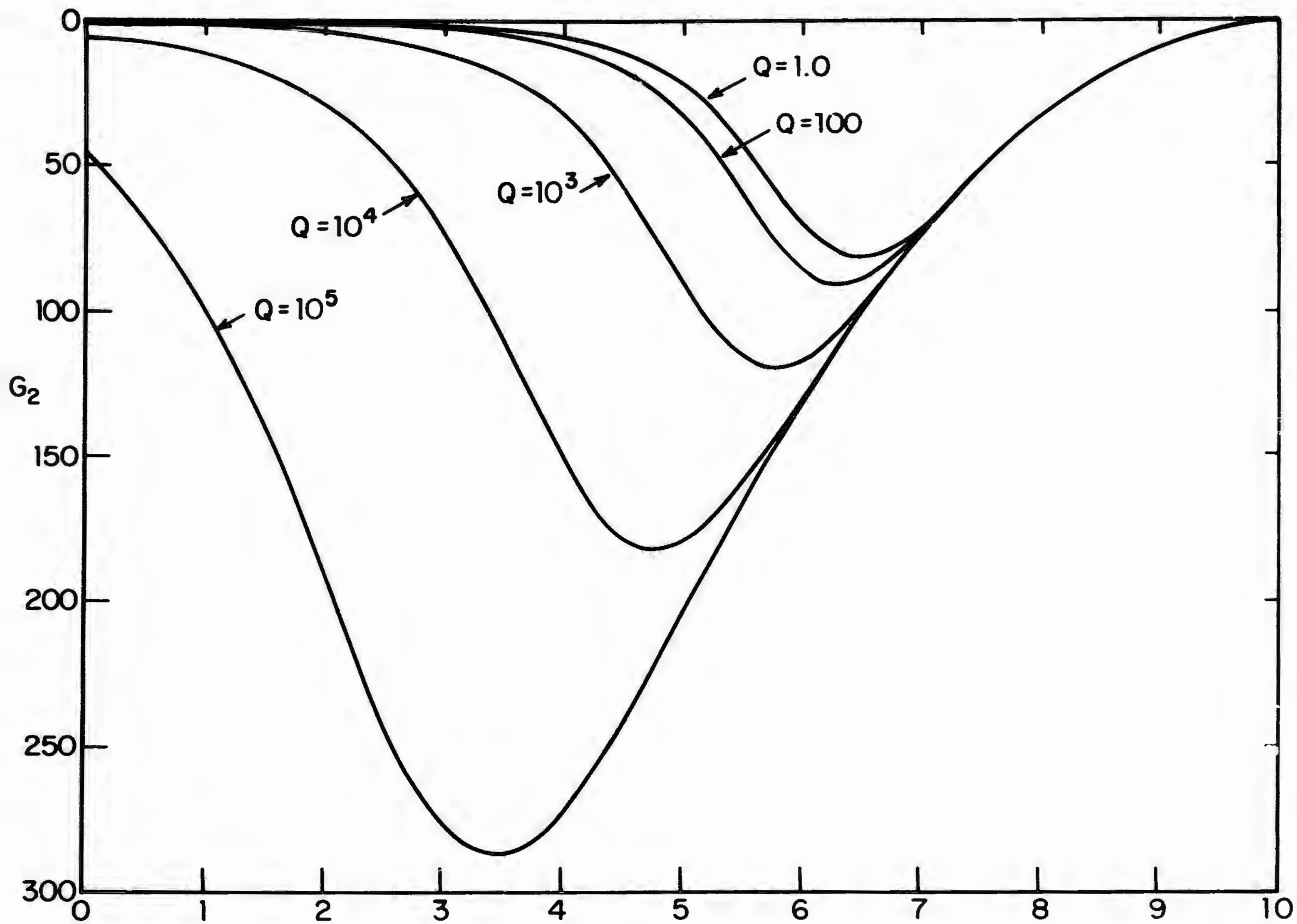


FIG. 7-9 FEEDBACK GAIN G_2 FOR THE PURSUER'S REALIZATION II

$$\dot{\Gamma}_2 = 2\left[\frac{G_p^2}{R_p} K^{-1} + \frac{P}{Q}\right]\Gamma_2 + \frac{G_p^2}{R_p}\Gamma_2^2 + \frac{G_e^2}{R_e}K^{-1} \quad ,$$

$$\Gamma_2(t_f) = 0 \quad , \tag{7:53}$$

$$\dot{P} = -2\frac{G_p^2}{R_p}(K^{-1} + \Gamma_2)P - \frac{P^2}{Q} \quad , \qquad P(t_o) = P_o \quad . \tag{7:54}$$

Consider the game where the values of the parameters are given by (7:46). Figures 7-10 and 7-11 display families of curves for P(t) and $\Gamma_2(t)$ respectively, where pertinent values of Q are the parameter. This is done on logarithmic scales to display the entire character of these two parameters.

Figure 7-10 points out that P(t) does indeed decay rapidly for all possible values of the noise variance; the curve labeled $Q = \begin{cases} 10^5 \\ 10^4 \end{cases}$ is the upper bound for P(t). When the game is a little more than half over, P(t) has already decreased by at least two orders of magnitude even if no measurements are taken. This is partially a function of the fact that the associated (optimized) deterministic game equation

$$\dot{y} = -(t_f - t)^2\left(\frac{1}{R_p} - \frac{1}{R_e}\right)K^{-1}(t_f, t)y \quad , \qquad y(t_o) = y_o \tag{7:55}$$

is quite stable for the parameter values of (7:46); y(t) is given by

$$y(t) = \frac{1}{2}(t_f - t)^3 K^{-1}(t_f, t_o)y_o \quad . \tag{7:56}$$

Since y(t) is naturally getting smaller, the variance of any estimate will decrease with it.

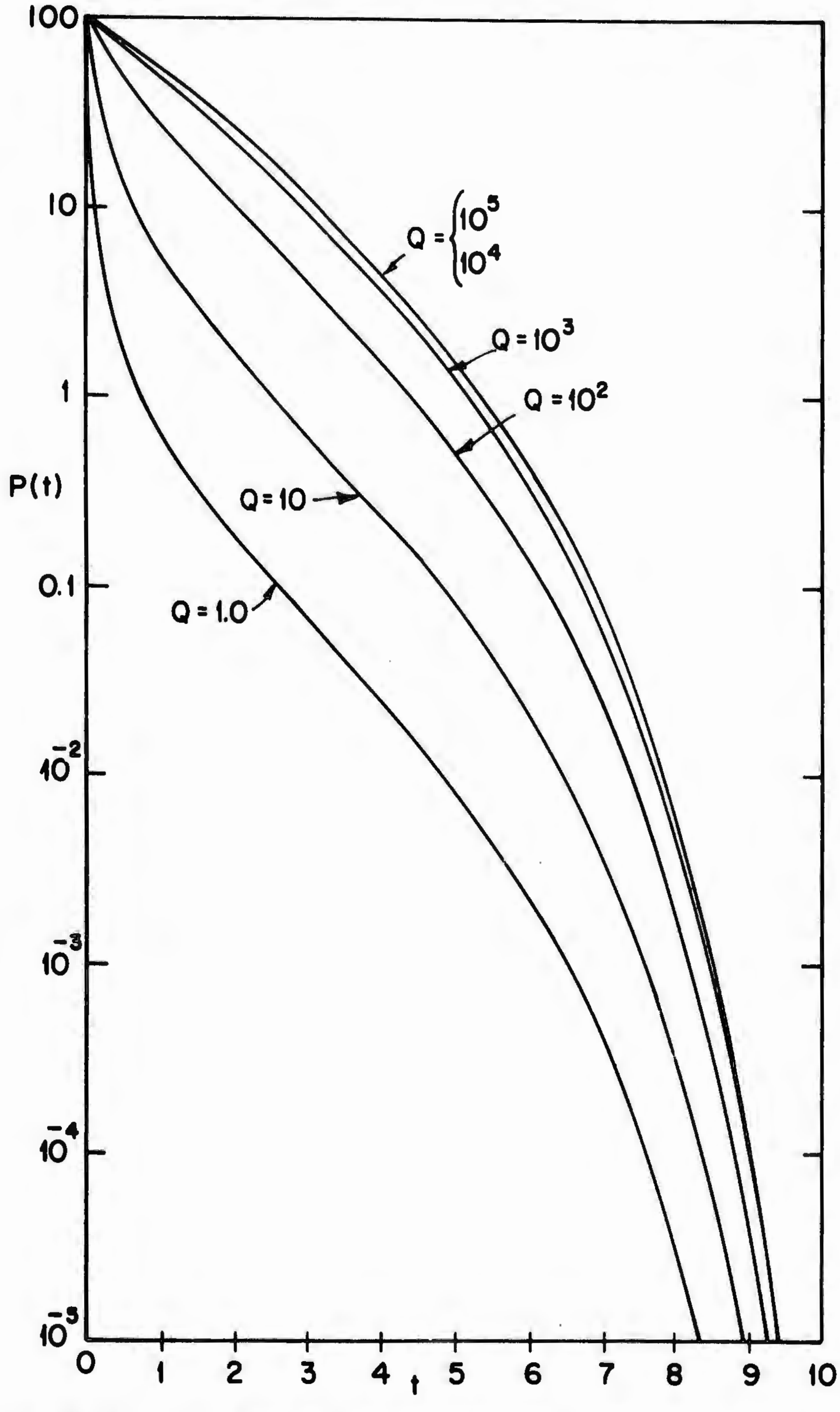


FIG. 7-10 THE ERROR VARIANCE P(t) FOR VARIOUS VALUES OF Q. EVADER WITH NOISY MEASUREMENTS. $a^2 = 10$, $Rp = 0.4$, $Re = 1.0$, $t_0 = 0$, $t_f = 10$.

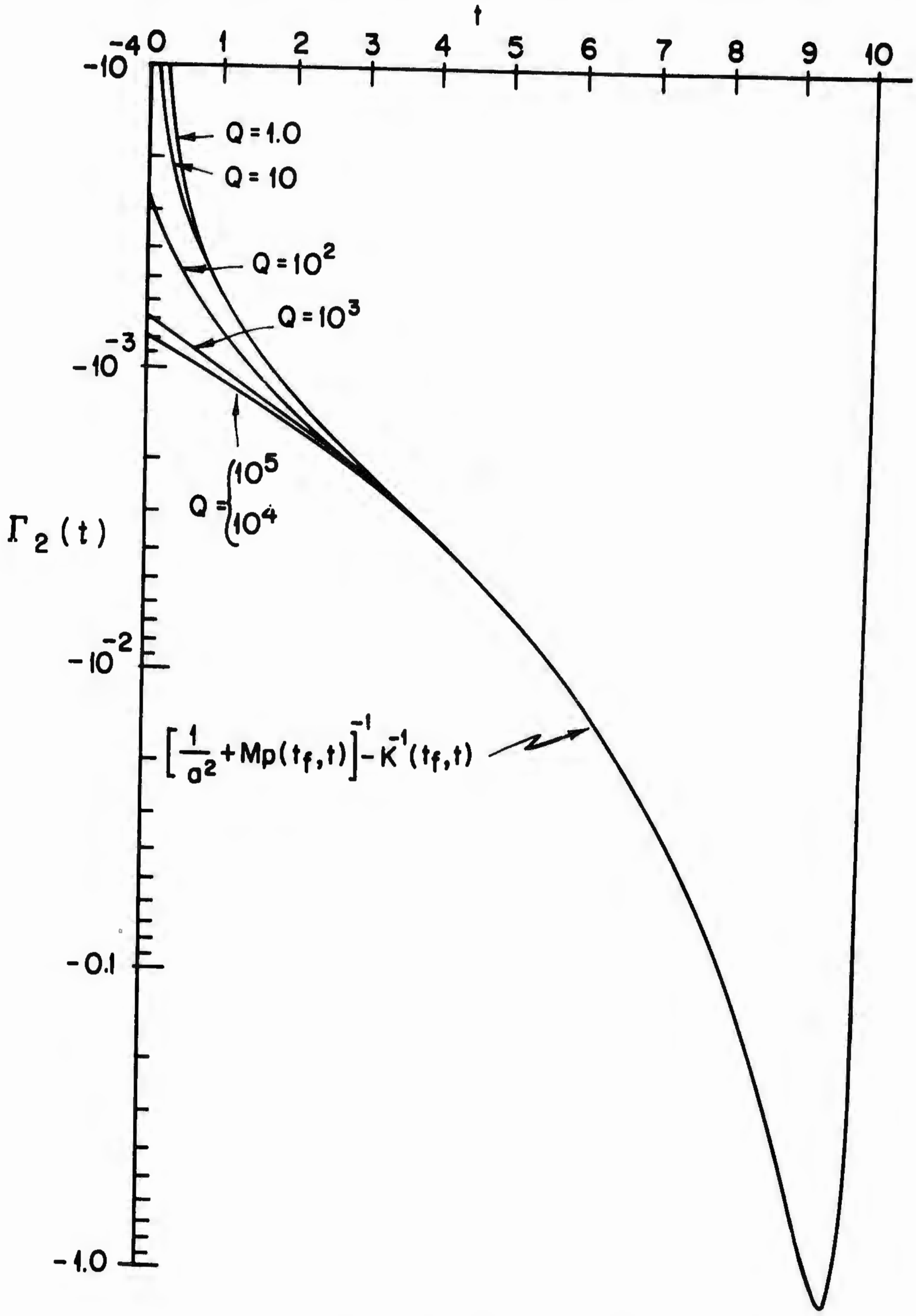


FIG. 7-11 THE PARAMETER $\Gamma_2(t)$ FOR VARIOUS VALUES OF Q. EVADER WITH NOISY MEASUREMENTS. $a^2=10$, $R_p=0.4$ $R_e=1.0$ $t_0=0.$ $t_f=10.$

Figure 7-11 displays the effect of Q on Γ_2, which a linear graph could not do. The curve labeled $Q = \begin{cases} 10^5 \\ 10^4 \end{cases}$ represents a lower bound on Γ_2; this curve is in fact the one obtained in Chapter 4, (4:12), for the value of $\Gamma_2(t)$ when Q equals infinity.

$$\Gamma_2(t) = [\frac{1}{a^2} + M_p(t_f, t)]^{-1} - K^{-1}(t_f - t) \quad . \tag{7:57}$$

As Q approaches zero -- so that the stochastic game approaches a deterministic one -- $\Gamma_2(t_o)$ approaches zero. This certainly must be true if the optimized value of the reduced criterion, J_r, is to approach zero, or what is equivalent, that the optimized stochastic criterion is to approach the deterministic value. J_r, (7:49), is composed of two terms. The second, $\frac{1}{2}\int_{t_o}^{t_f} \frac{P^2\Gamma_2}{Q} dt$, certainly approaches zero as Q does since P approaches zero as Q; see (4:8) and Figure 7-10. The first term, $\frac{1}{2}P_o\Gamma_2(t_o)$, must also approach zero as Q. Since P_o is fixed by the definition of the problem, $\Gamma_2(t_o)$ must approach zero with Q if this latter condition is to be satisfied.

Figures 7-10 and 7-11 are particularly applicable for a game of time duration 10. and initial estimation error variance of 100., though other problems can be discussed through scaling. Figures 7-12 and 7-13 provide a more general picture by displaying $\frac{P(t)}{Q}$ and $\Gamma_2(t)$ respectively, in a format that is applicable for games of time duration less than 20. The horizontal scales are plotted in time-to-go, i.e. $(t_f - t)$.

It was necessary to divide Figure 7-12 into two parts, 7-12a and 7-12b, so that all appropriate values of $\frac{P(t)}{Q}$ could be displayed. The

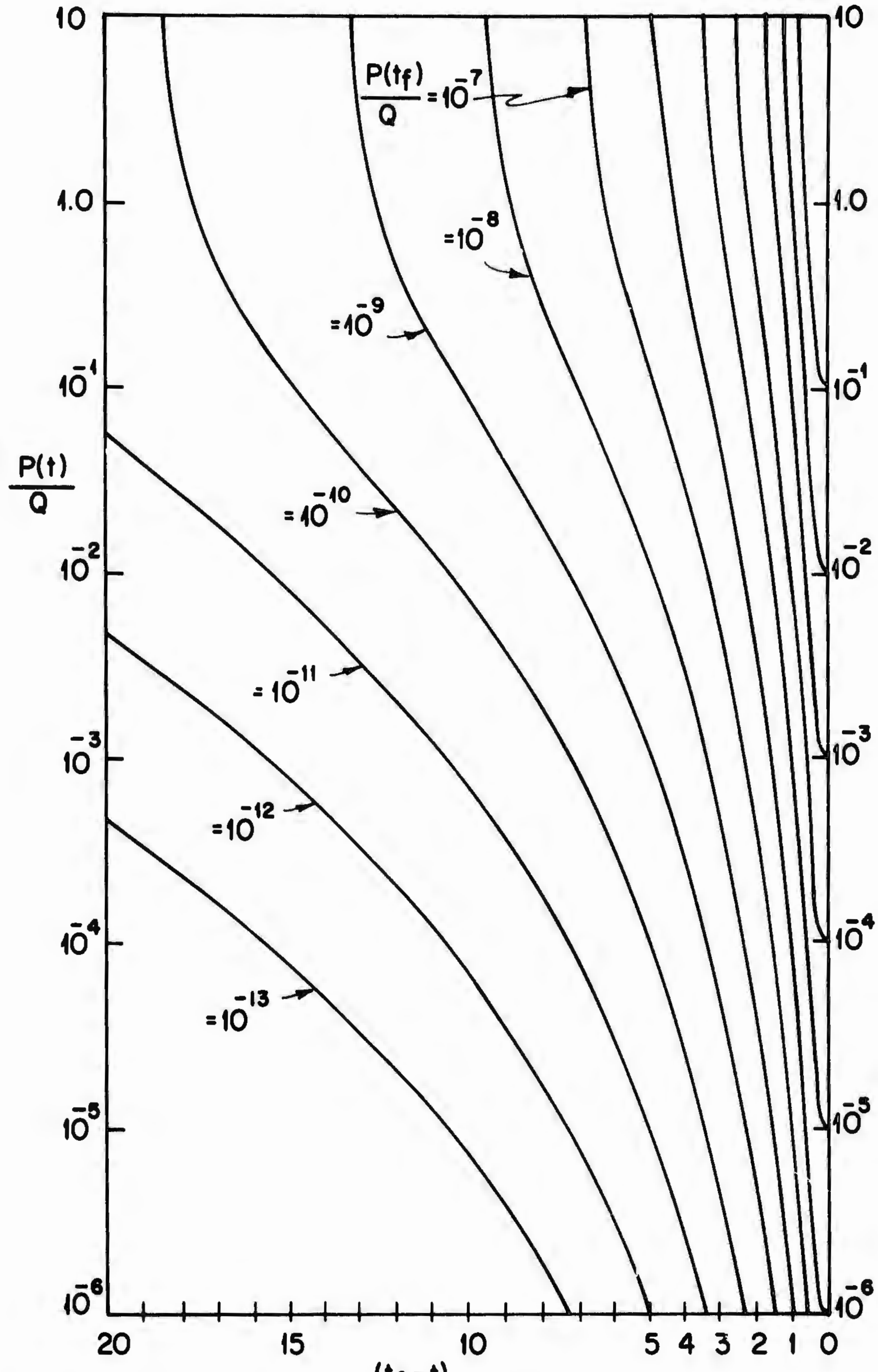


FIG. 7-12A THE PARAMETER $\frac{P(t)}{Q}$ FOR GAMES OF $(t_f - t_o)$ UP TO 20. EVADER WITH NOISY MEASUREMENTS. $a^2 = 10$, $Rp = 0.4$, $Re = 1.0$

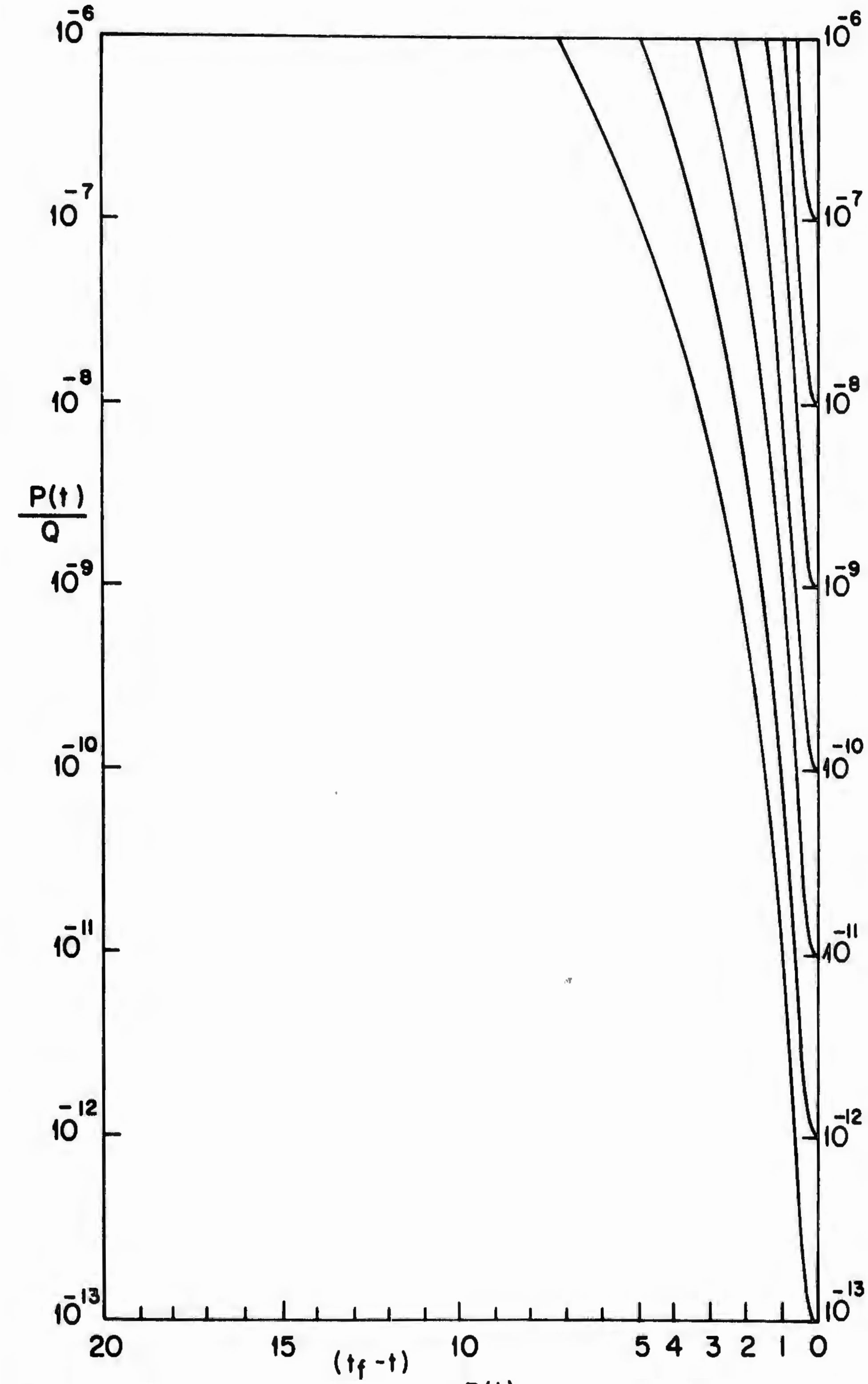


FIG. 7-12B THE PARAMETER $\frac{P(t)}{Q}$ FOR GAMES OF $(t_f - t_o)$ UP TO 20. EVADER WITH NOISY MEASUREMENTS. $a^2 = 10$, $Rp = 0.4$, $Re = 1.0$.

interesting range of initial conditions is $10^2 \geq \frac{P_o}{Q} \geq 10^{-3}$. But since the game can be of any time duration, data on $\frac{P(t)}{Q}$ must be available over this range for all t. Data for $\frac{P(t)}{Q} \geq 10.$ is omitted, however, since all curves are almost vertical in this region.

Figure 7-12 demonstrates again how P(t) decays with time, here for a substantially larger range of initial conditions and time durations of the game. Consequently, this characteristic -- while it is a function of the values of a^2, R_p and R_e -- is not dependent on Q, P_o or $t_f - t$. Observe for $\frac{P(t_o)}{Q} > 1.0$, that $\frac{P(t)}{Q}$ drops off quite quickly, certainly more so than when $\frac{P(t_o)}{Q} < 1.0$. This is because for large values of the ratio $\frac{P(t_o)}{Q}$ the measurements will help to improve the estimate. For small values of this ratio the measurements will have little effect; the estimate will be based largely on $\hat{y}(t_o)$ and the decay will result mostly from the natural stability of (7:55).

Figure 7-13 demonstrates that the possible curves for $\Gamma_2(t)$ can be quite varied, though this was not obvious from the single problem discussed above. Again observe that $[\frac{1}{a^2} + M_p(t_f, t)]^{-1} - K^{-1}(t_f, t)$ is the limiting lower bound on $\Gamma_2(t)$; since both $[\frac{1}{a^2} + M_p(t_f, t)]^{-1}$ and $K^{-1}(t_f, t)$ approach zero as $(t_f - t)$ gets large, so does this lower bound.

The labels on the various curves in 7-13 can be used to identify them with the corresponding curve in 7-12. Viewing time as progressing backwards, one can see that $\Gamma_2(t)$ keeps near $[\frac{1}{a^2} + M_p(t_f, t)]^{-1} - K^{-1}(t_f, t)$ until $\frac{P(t)}{Q}$ gets large enough to force $\Gamma_2(t)$ to zero. $\Gamma_2(t)$ breaks away from the bounding curve when $\frac{P(t)}{Q}$ reaches 10^{-2}.

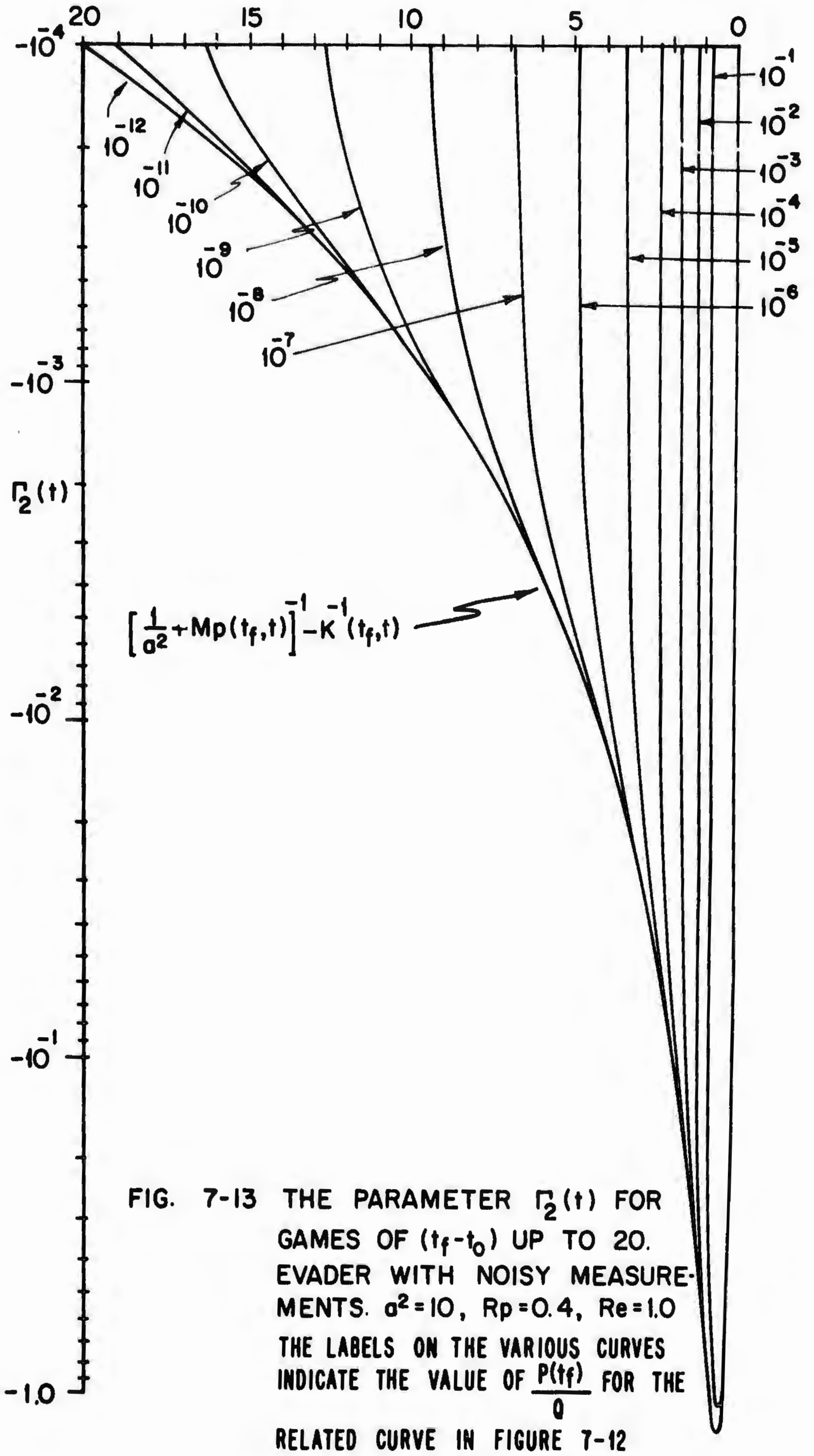


FIG. 7-13 THE PARAMETER $\Gamma_2(t)$ FOR GAMES OF $(t_f - t_0)$ UP TO 20. EVADER WITH NOISY MEASUREMENTS. $a^2 = 10$, $Rp = 0.4$, $Re = 1.0$

THE LABELS ON THE VARIOUS CURVES INDICATE THE VALUE OF $\frac{P(t_f)}{Q}$ FOR THE RELATED CURVE IN FIGURE 7-12

Observe that all possible initial conditions $\frac{P_o}{Q} = \frac{P(t_o)}{Q}$ are crossed by the family of curves in Figure 7-12. Further note that for every such possible initial condition there exists an initial value of $\Gamma_2(t_o)$ which is finite and that the curve of $\Gamma_2(t)$, $t_o \leq t \leq t_f$, originating at this point always remains finite. Consequently, it can be seen graphically that the conjugate point condition (that $\Gamma_2(t)$ fails to be finite) is never satisfied because of the finite initial condition on the pair of equations (7:53)-(7:54).

For the problem where the pursuer is the player making the noisy measurements, Section 5.2, the equations for $\Gamma_2(t)$ and P(t) are (5:52)-(5:53). For the scalar game, these reduce to

$$\dot{\Gamma}_2 = 2\left[\frac{P}{Q} - \frac{G_e^2}{R_e}K^{-1}\right]\Gamma_2 - \frac{G_e^2}{R_e}\Gamma_2 - \frac{G_p^2}{R_p}K^{-1} \quad ,$$

$$\Gamma_2(t_f) = 0 \quad , \tag{7:58}$$

$$\dot{P} = 2\frac{G_e^2}{R_e}(K^{-1} + \Gamma_2)P - \frac{P^2}{Q} \quad , \qquad P(t_o) = P_o \quad . \tag{7:59}$$

For this game, also consider the parameter values given by (7:46). Figures 7-14 and 7-15 display families of curves for P(t) and $\Gamma_2(t)$ respectively, where pertinent values of Q are the parameter. Logarithmic scales are again used to display the full range of values.

Figure 7-14 indicates that for large noise variances, $Q = 10^2$, 10^3, or 10^4, that P(t) initially increases with time, despite the fact that the pursuer is making measurements. From Figure 7-15 it can be seen that

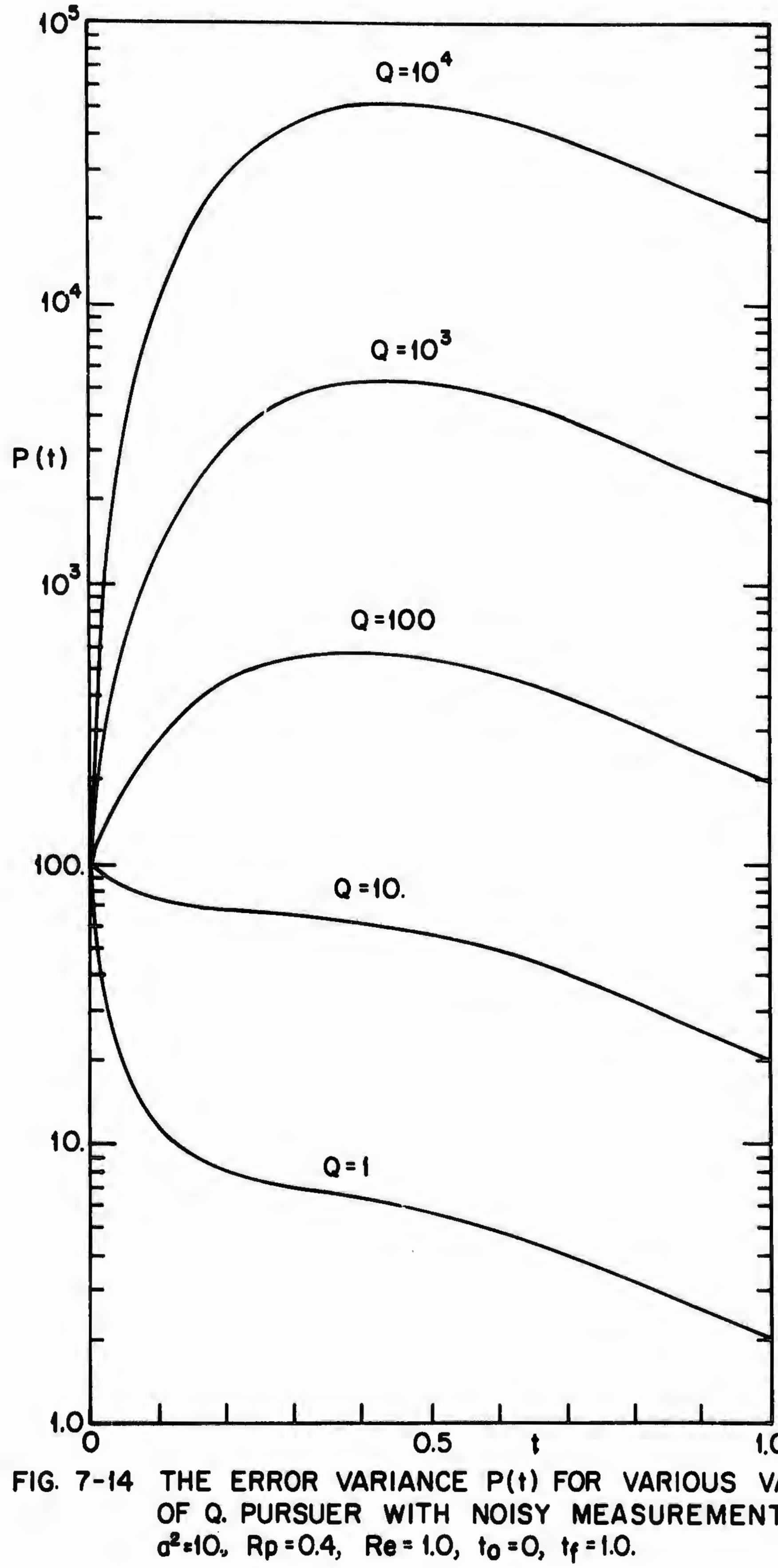


FIG. 7-14 THE ERROR VARIANCE P(t) FOR VARIOUS VALUES OF Q. PURSUER WITH NOISY MEASUREMENTS. $a^2 = 10.$, $Rp = 0.4$, $Re = 1.0$, $t_0 = 0$, $t_f = 1.0$.

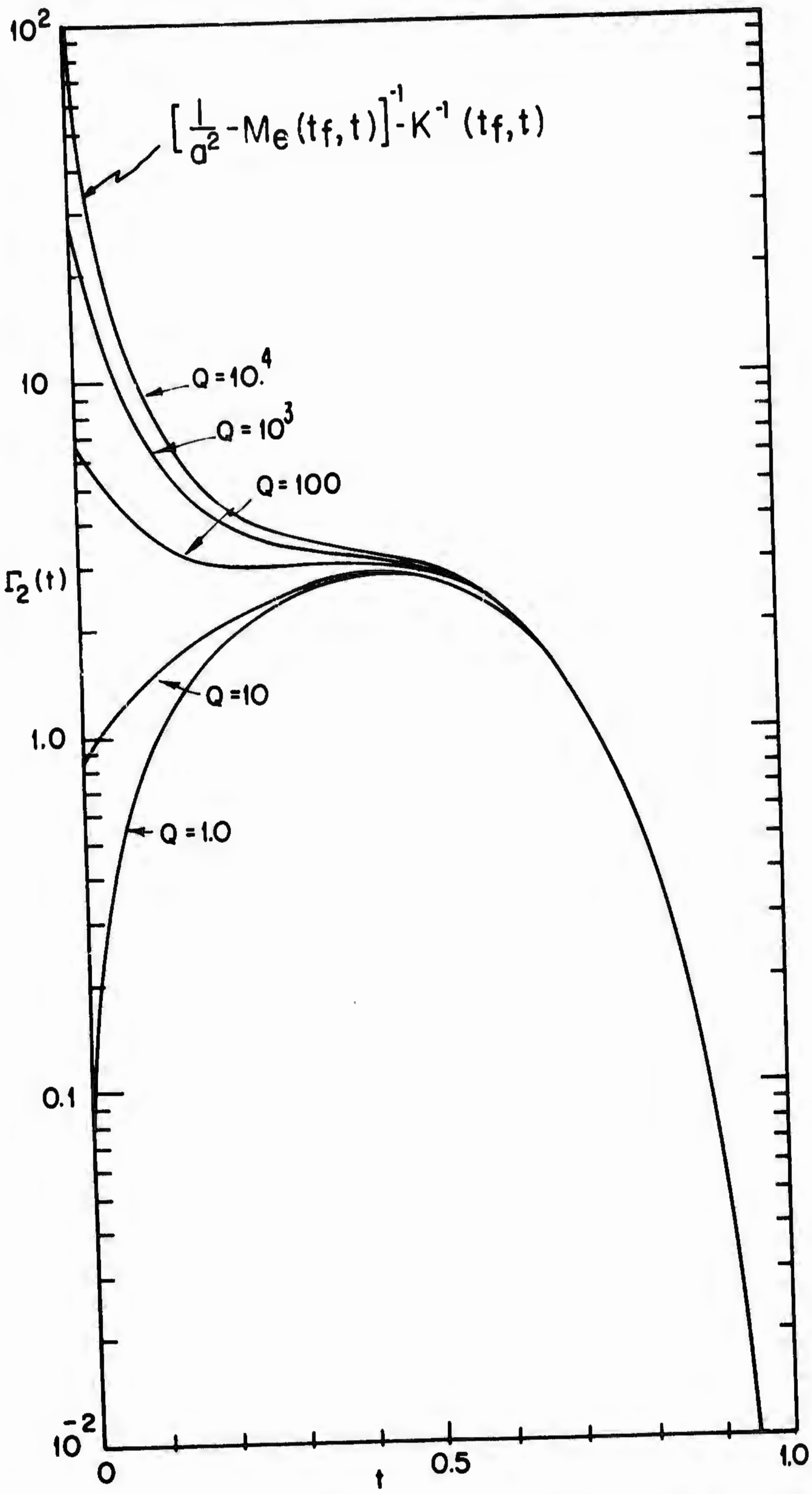


FIG. 7-15 THE PARAMETER $\Gamma_2(t)$ FOR VARIOUS VALUES OF Q. PURSUER WITH NOISY MEASUREMENTS. $a^2 = 10.$, $R_p = 0.4$, $R_e = 1.0$, $t_o = 0.$, $t_f = 1.0$

the evader uses large initial values of Γ_2 for large Q, thus taking substantial advantage of errors in the pursuer's estimate. These two characteristics are a consequence of the fact that if the measurements are not sufficiently accurate, to provide continual correction of the pursuer's estimate, the pursuer will initially direct himself towards a spot which is apt to be quite far from the evader's location; furthermore the pursuer will not correct this error very quickly. Unless the pursuer's initial estimate is fortuitously perfect, y(t) can actually be increasing with time, despite the fact that the pursuer is more controllable than the evader. In short, the pursuer is all muscle and no brains.

It can be seen from Figure 7-14 that $P(t_f)$ is linearly dependent upon Q.

$$P(t_f) \doteq 2Q \quad . \tag{7:60}$$

In fact, after t reaches 0.4 the variance of the pursuer's estimation error is essentially independent of $P(t_o)$; the curves are almost parallel and are displayed on an order of magnitude for every order of magnitude change in Q.

It appears after the game has progressed for a sufficiently long time for the effect of initial conditions to disappear, that a kind of equilibrium is established. The evader does not attempt to get too far away; if he did so the pursuer would be able to use nearly the proper control even though his estimation was quite imperfect, for the standard deviation of his estimation error would be small compared to y. The level of this equilibrium is dependent upon the quality of the measurements -- for these are used by the pursuer to sustain his knowledge of what the evader is doing -- not the initial estimate.

Figure 7-15 demonstrates that as Q approaches zero $\Gamma_2(t_o)$ also approaches zero; again insuring that the criterion approaches its deterministic value.

Figures 7-16 and 7-17 provide the general picture of the game for the pursuer with noisy measurements with durations up to 2.4. From Figure 7-16 it can be seen that for games of durations greater than 1.0

$$\frac{P(t_f)}{Q} \doteq 2.0 \quad . \tag{7:61}$$

When integrating backwards, the differential equation governing P(t), (7:59), is very unstable; only terminal conditions of $\frac{P(t)}{Q}$ near two produce curves that fail to go to infinity or zero immediately.

Consider games of time durations greater than 1.0. For any initial value $\frac{P(t_o)}{Q}$, $\frac{P(t)}{Q}$ quickly goes to an equilibrium value somewhere between 1.0 and 10. If the initial error variance is substantially larger than the measurement noise variance, the measurements quickly improve the estimate. If the initial error variance is very small, compared with the measurement noise variance, the estimate will become worse with time, because the pursuer will be going in the wrong direction while collecting data that is not accurate enough to provide sufficient information as to how the game is progressing.

Figure 7-17 displays $\Gamma_2(t)$ for the game with the pursuer making noisy measurements. Observe as $\frac{P(t_f)}{Q}$ approaches zero that $\Gamma_2(t)$

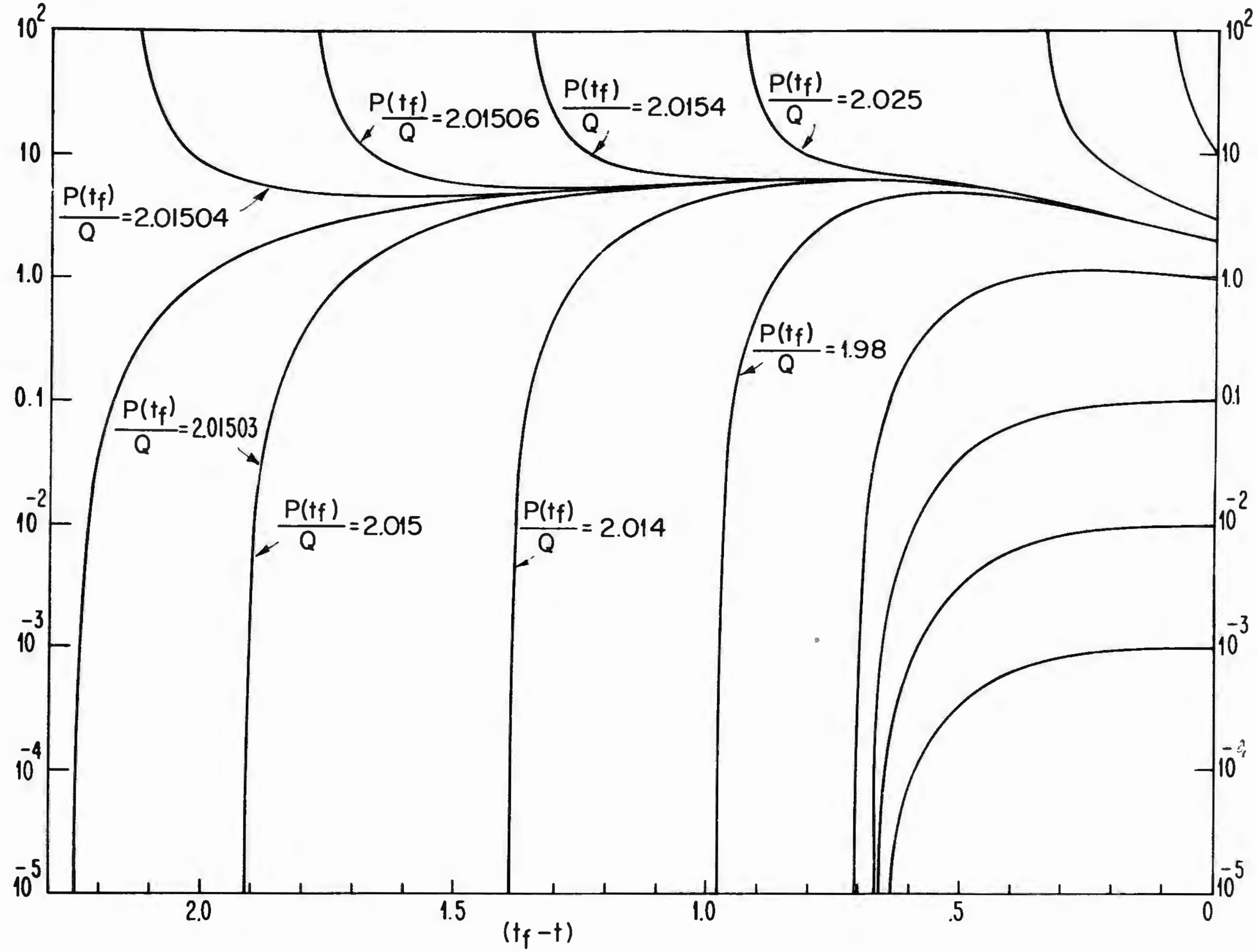


FIG. 7-16 THE PARAMETER $P(t)/Q$ FOR GAMES OF $(t_f - t_0)$ UP TO 2.4. PURSUER WITH NOISY MEASUREMENTS. $a^2 = 10.$, $R_p = 0.4$, $R_e = 1.0$.

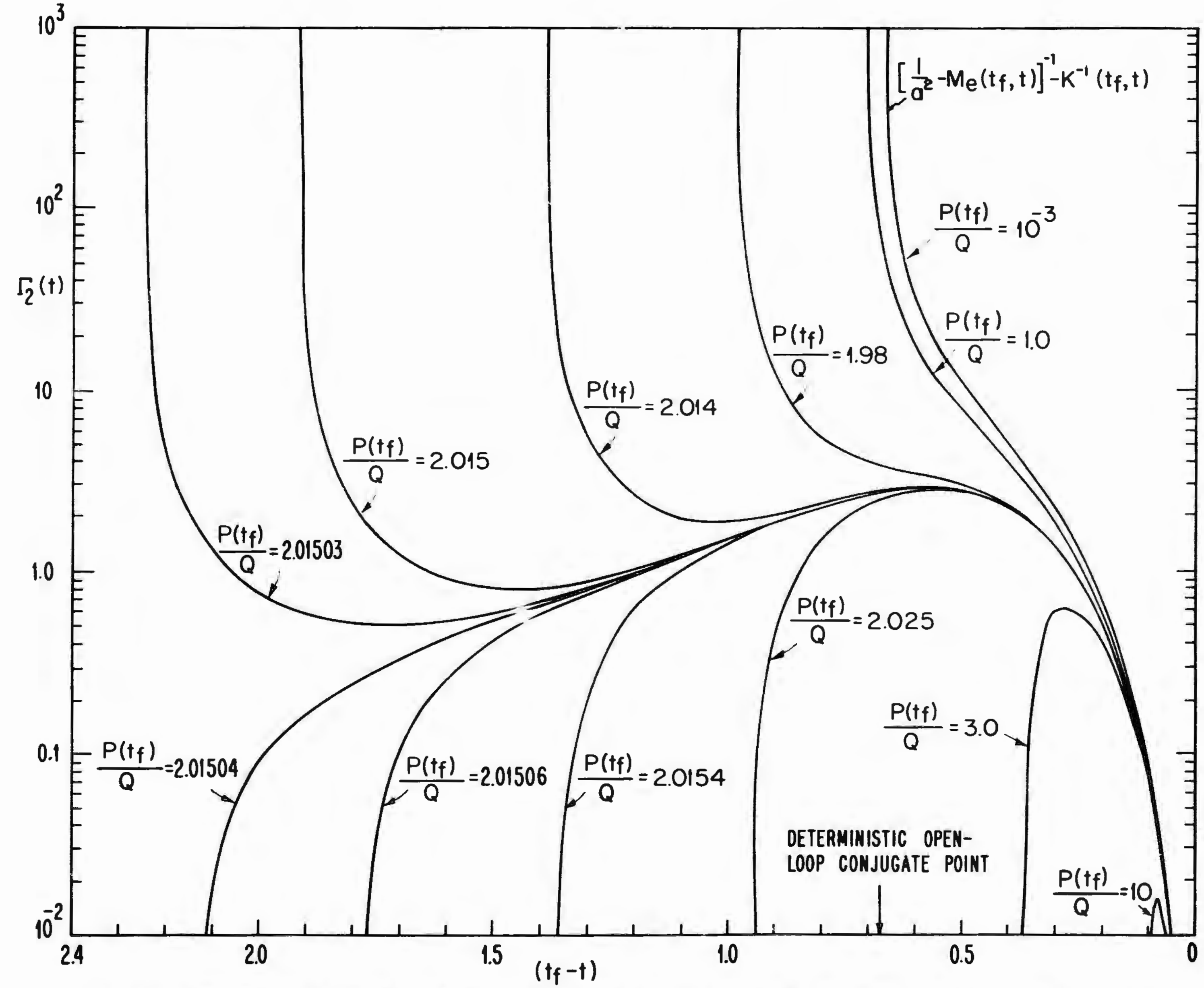


FIG. 7-17 THE PARAMETER $\Gamma_2(t)$ FOR GAMES OF $(t_f - t_c)$ UP TO 2.4. PURSUER WITH NOISY MEASUREMENTS. $a^2 = 10$, $R_p = 0.4$, $R_e = 1.0$

approaches $[\frac{1}{a^2} - M_e(t_f, t)]^{-1} - K^{-1}(t_f - t)$. The curve for this latter term goes to infinity at the conjugate point for the game with the pursuer operating open-loop.

Note in Figure 7-17 that some curves for $\Gamma_2(t)$ do go to infinity. This does not mean that the stochastic conjugate point has occurred. Rather this corresponds to $\frac{P(t)}{Q}$ dropping off to zero (when viewing time as progressing backwards). For every finite value of $\frac{P(t^*)}{Q}$, there exists a finite value of $\Gamma_2(t^*)$ and the curve of $\Gamma_2(t)$ passing through $\Gamma_2(t^*)$ is finite for all values of $t \geq t^*$.

A comparison of Figures 7-12 and 7-13 with Figures 7-16 and 7-17 provides a visual demonstration of the asymmetry of the two games: the game with the evader making noisy measurements and the one with the pursuer making the noisy measurements.

Figures 7-18 and 7-19 display $\frac{P(t)}{Q}$ and $\Gamma_2(t)$ for the game with the evader making noisy measurements, but also with a deterministic conjugate point at $(t_f - t) \doteq 1.49$. With the exception of the fact that

$$R_p = 1.1 \quad , \tag{7:62}$$

all the parameters are the same as given in (7:46). This change, however, means that the evader is more controllable than the pursuer; $K^{-1}(t_f, t)$ goes to infinity when $t = 1.49$. Note that the equations $P(t)$ and $\Gamma_2(t)$ are still (7:53) and (7:54).

The character of the family of curves for $\frac{P(t)}{Q}$ in Figure 7-18 is not substantially different from the curves in Figure 7-12. The only difference is that all curves go to infinity at the deterministic conjugate point, because $K^{-1}(t_f, t)$ does.

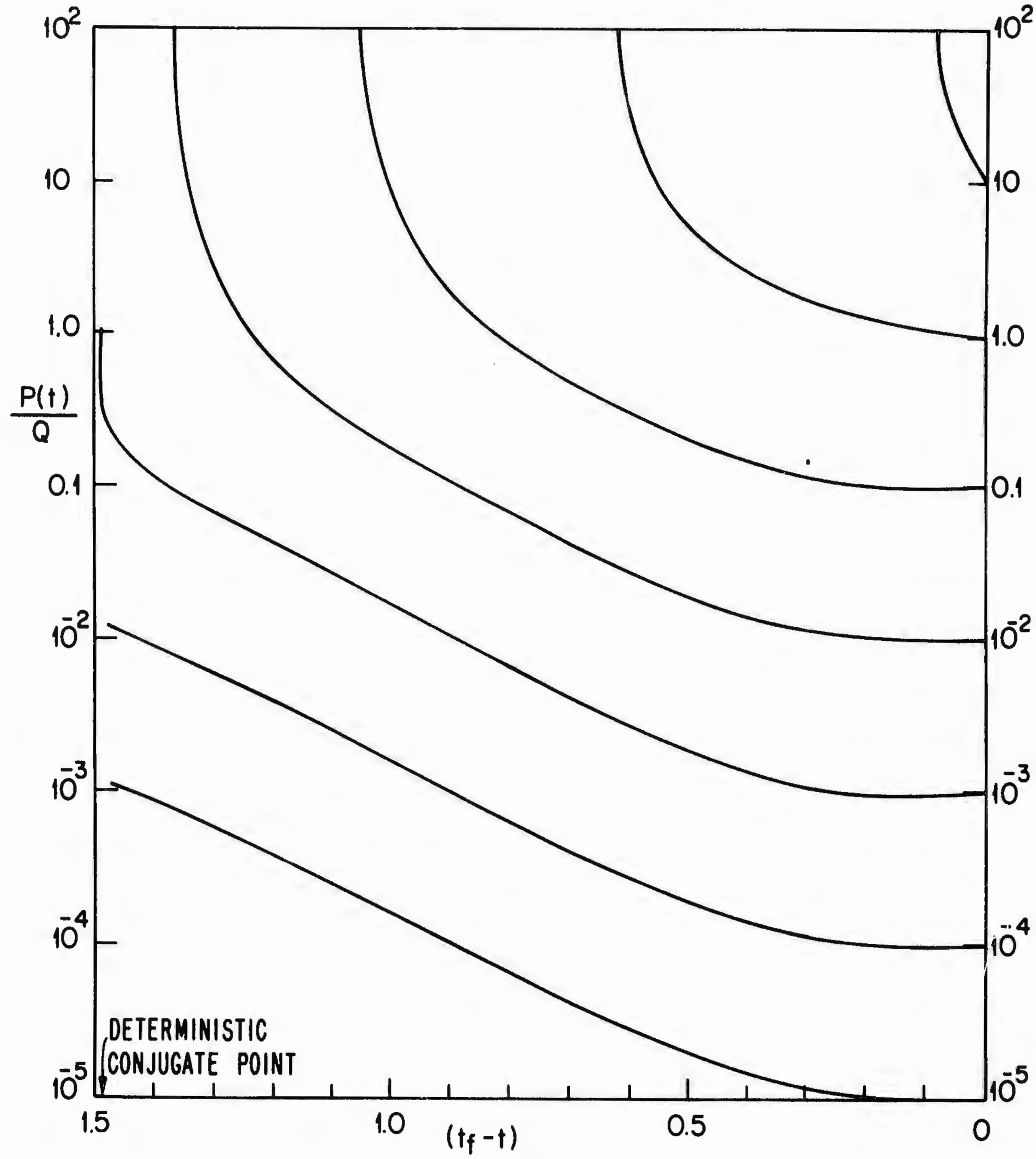


FIG. 7-18 THE PARAMETER P(t)/Q FOR THE GAME WITH A DETERMINISTIC CONJUGATE POINT AT $(t_f - t) = 1.49$. $a^2 = 10.$, $R_p = 1.1$, $R_e = 1.0$.

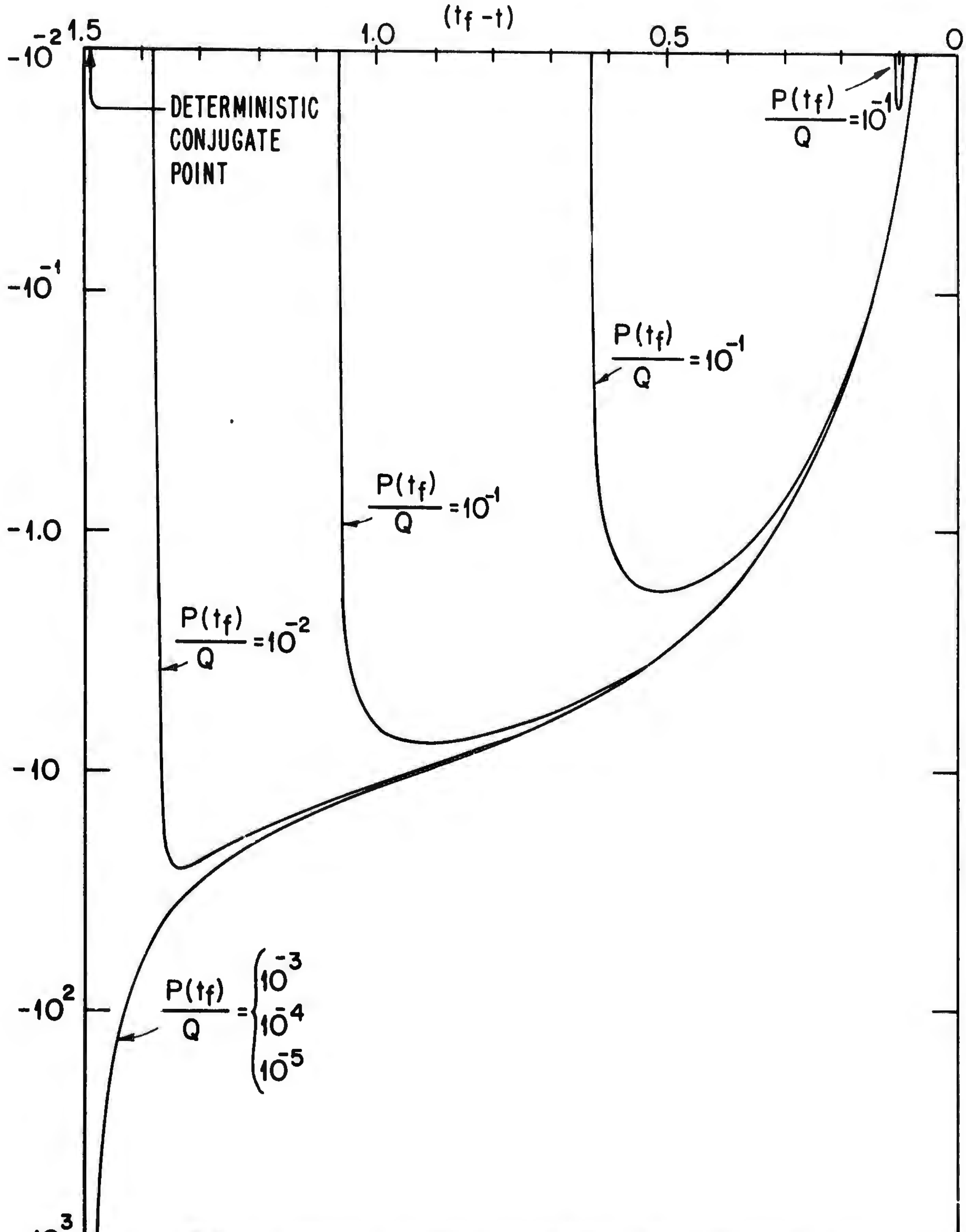


FIG. 7-19 THE PARAMETER $\Gamma_2(t)$ FOR THE GAME WITH A DETERMINISTIC CONJUGATE POINT AT $(t_f - t) = 1.49$. EVADER WITH NOISY MEASUREMENTS. $a^2 = 10$, $R_p = 1.1$, $R_e = 1.0$.

The curves for $\Gamma_2(t)$ of Figure 7-19 do differ from those of Figure 7-13 in that some do go to infinity. This results from the fact that as $\frac{P(t_f)}{Q}$ gets small, $\Gamma_2(t)$ approaches $[\frac{1}{a^2} + M_p(t_f, t)]^{-1} - K^{-1}(t_f, t)$. However, for the game with a deterministic conjugate point, this latter term does go to infinity since $K^{-1}(t_f, t)$ does.

For the game with the pursuer making noisy measurements plus a deterministic conjugate point at $(t_f - t) = 1.49$, Figures 7-20 and 7-21 display $\frac{P(t)}{Q}$ and $\Gamma_2(t)$. The only major differences between these curve families and those given in Figures 7-16 and 7-17 is that the conjugate point appears to act as a barrier forcing the curve families towards the terminal time.

7.5 Optimal Trajectories

This section presents some optimal trajectories -- both stochastic samples and the associated deterministic path -- for the interception problem in two-dimensional euclidean space which was discussed in Section 7.1. Here the values of the system and estimation parameters are the same in the two dimensions and given by:

$$
\begin{aligned}
&a^2 = 10. \\
&r_{1p} = r_{2p} = r_p = 0.4 \quad , \\
&r_{1e} = r_{2e} = r_e = 1.0 \quad , \\
&P_{11o} = P_{22o} = P_o = 100. \quad , \qquad\qquad (7:63) \\
&q_1 = q_2 = 100. \quad , \\
&t_o = 0. \quad , \\
&t_f = 10. \quad , \\
&h_1 = h_2 = 1.0 \quad .
\end{aligned}
$$

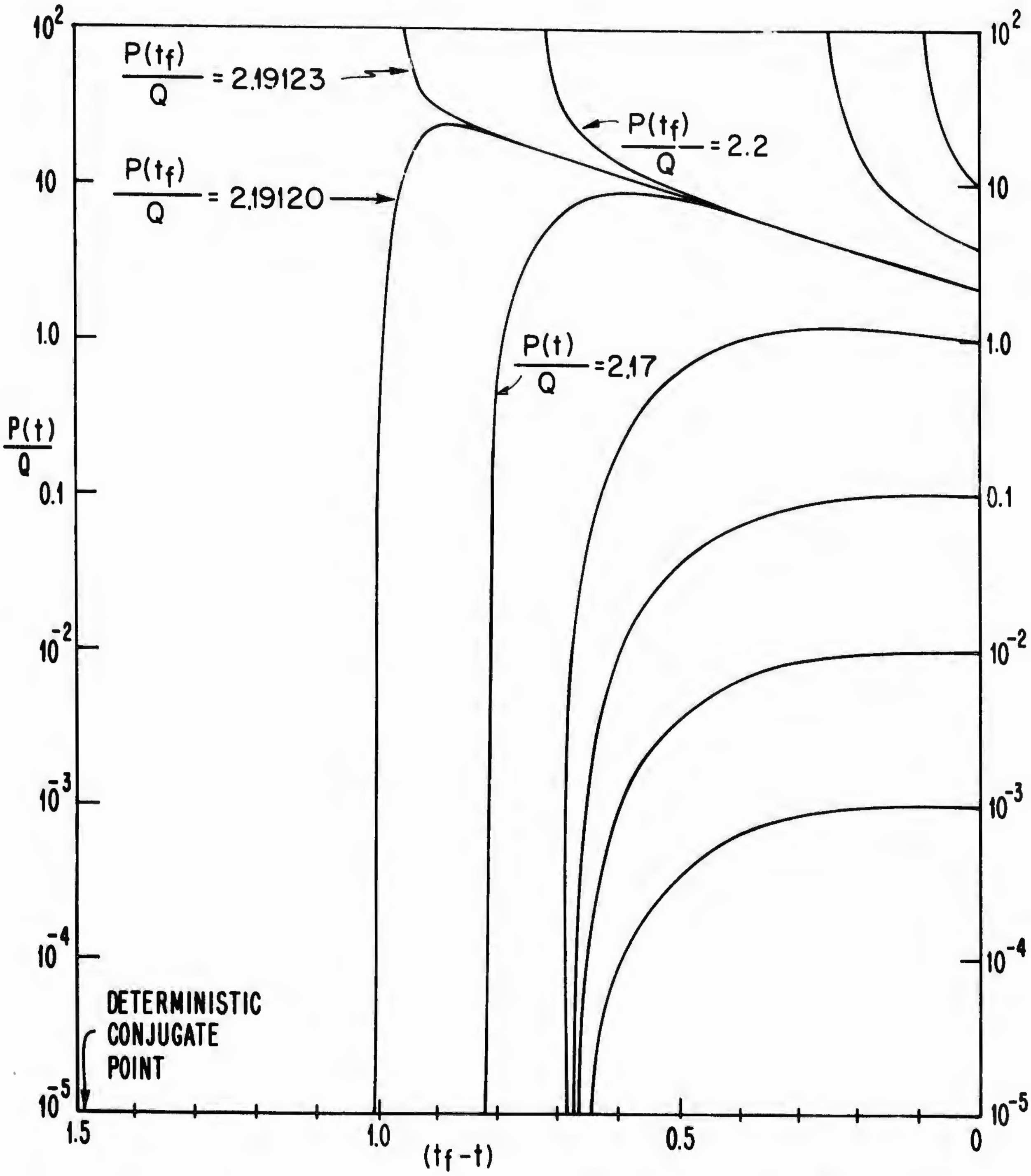


FIG. 7-20 THE PARAMETER P(t)/Q FOR THE PURSUER WITH NOISY MEASUREMENTS. DETERMINISTIC CONJUGATE POINT AT $(t_f - t) = 1.49$, $a^2 = 1.0$, $R_p = 1.1$, $R_e = 1.0$

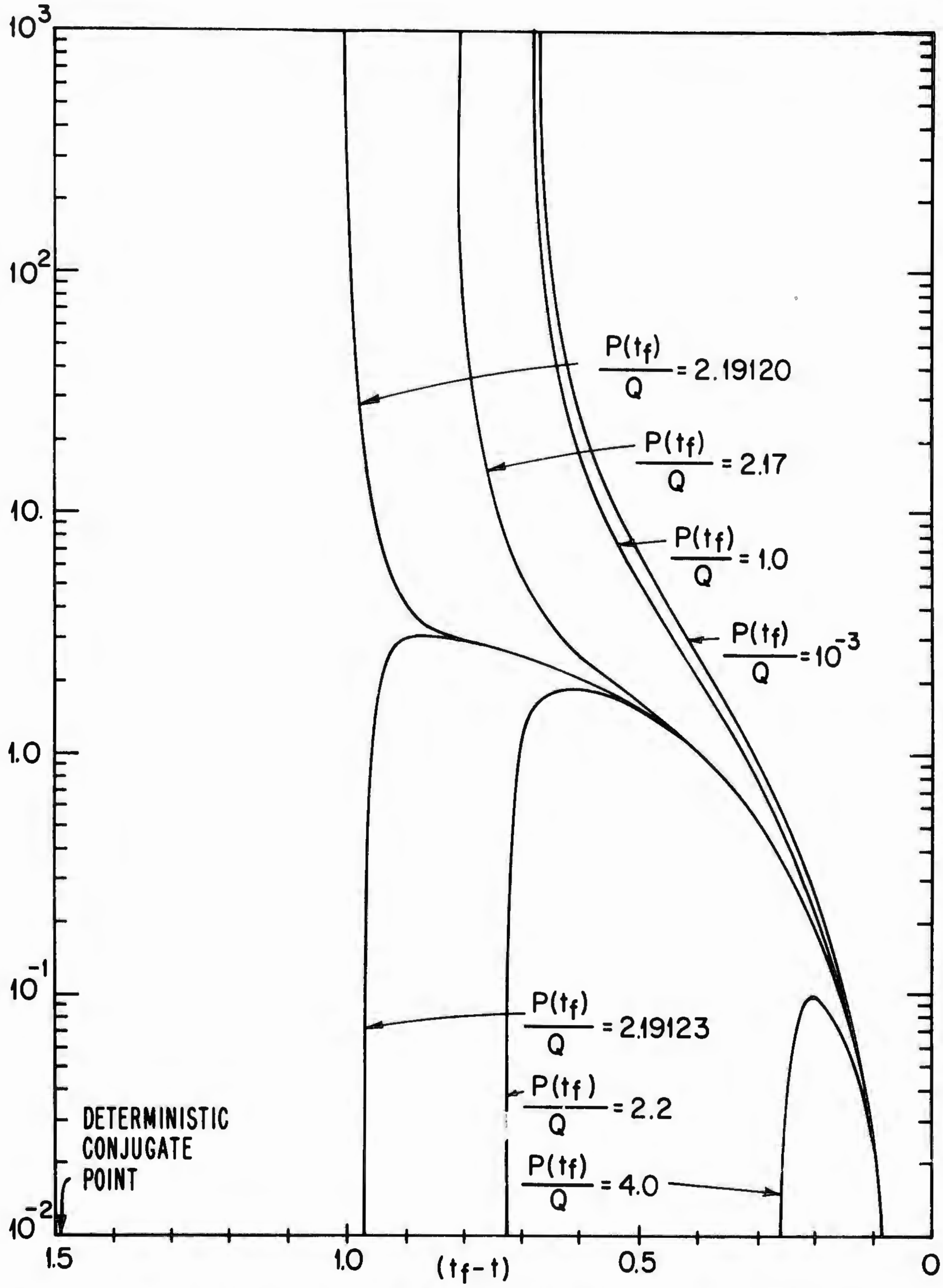


FIG. 7-21 THE PARAMETER $\Gamma_2(t)$ FOR THE PURSUER WITH NOISY MEASUREMENTS. DETERMINISTIC CONJUGATE POINT AT $(t_f - t) = 1.49$, $a^2 = 1.0$, $R_p = 1.1$, $Re = 1.0$

Note that these are the same parameters which were used in the previous section for the numerical discussion of $\Gamma_2(t)$ and $P(t)$.

Consider the deterministic game defined by (7:63) and the initial conditions on the pursuer's and evader's state vectors

$$x_{po} = \begin{bmatrix} 0. \\ 0. \\ 0. \\ 0. \end{bmatrix} , \qquad (7:64)$$

$$x_{eo} = \begin{bmatrix} 20. \\ 10. \\ 0. \\ 0. \end{bmatrix} , \qquad (7:65)$$

The first two components of these vectors indicate the players' position on the x_1 and x_2 axes, while the other two components indicate the corresponding velocities. These translate, using (7:10) and (7:11), to the initial condition for y which is

$$y_o = \begin{bmatrix} -20. \\ \\ -10. \end{bmatrix} . \qquad (7:66)$$

Figure 7-22 gives the optimal deterministic trajectory for these initial conditions. Note that since neither player has an initial velocity, the pursuer heads directly towards the original location of the evader, while the evader merely heads directly away from the pursuer's initial location. There is a very small terminal separation, but it cannot be displayed on the graph. (Recall from Section 7.3 that y is reduced by a factor of 2.0×10^{-4} in each dimension during the play of this game.)

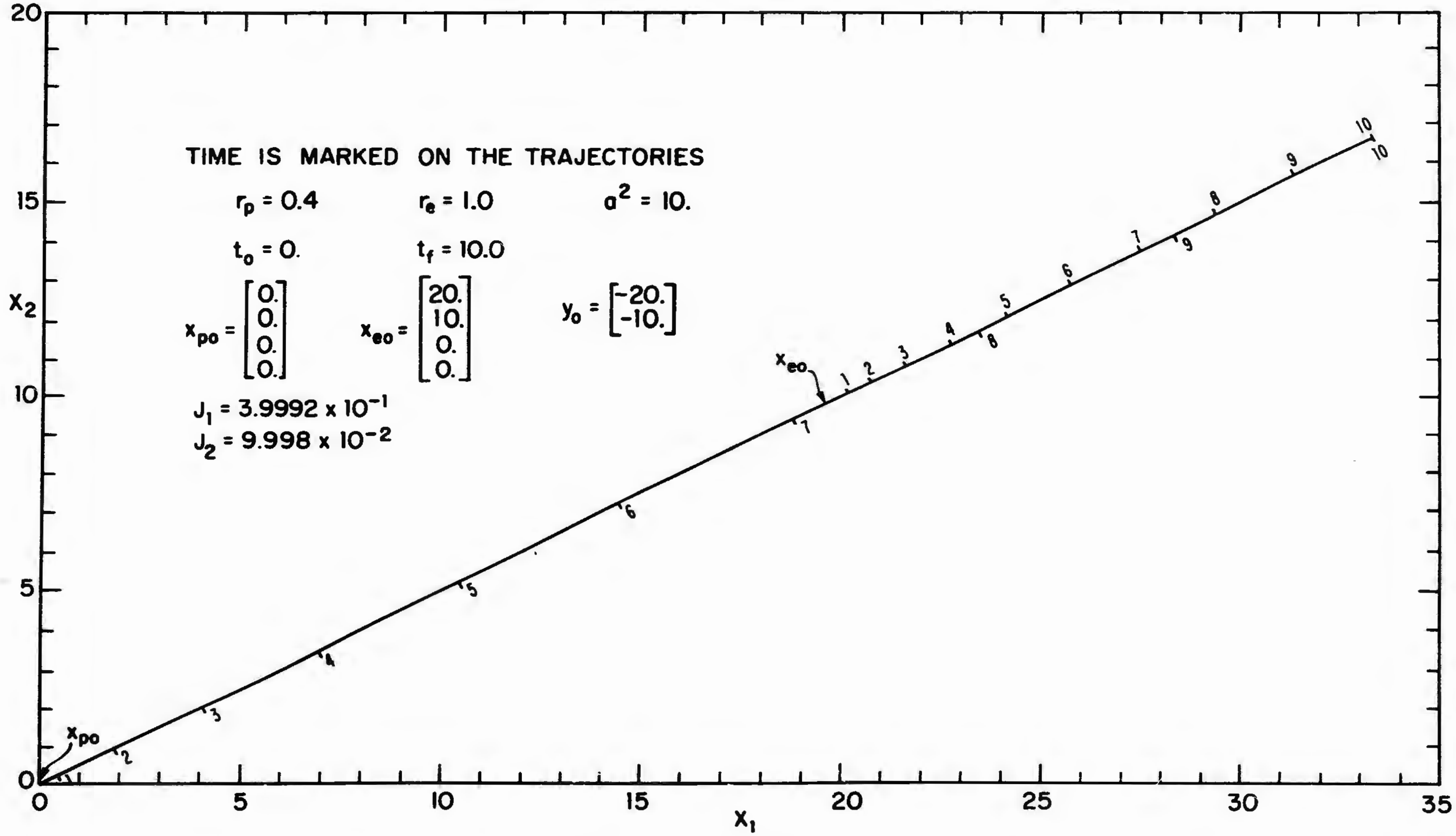


FIG. 7-22 OPTIMAL DETERMINISTIC TRAJECTORY I

The values of the criterion in the two separate dimensions are: $J_1 = 3.9992 \times 10^{-1}$ and $J_2 = 9.998 \times 10^{-2}$. These can either be computed by integrating $u^2(t)$ and $v^2(t)$ during an actual play of the game, or from (1:37) and the value of $K^{-1}(t_f, t_o)$ which is 1.9996×10^{-3}.

Figure 7-23 gives the y(t) corresponding to the game along the x_1 axis in Figure 7-22. The shape of this curve is appropriate for all deterministic games with the parameters as given in (7:63) regardless of initial conditions; this is due to the fact that $K^{-1}(t_f, t)y(t)$ is independent of time and $K^{-1}(t_f, t)$ is fixed in the form displayed in Figure 7-2.

Figure 7-24 displays the pursuer's and evader's controls for this deterministic game. Note that since $K^{-1}(t_f, t)y(t)$, r_p, and r_e are all constant, the time dependencies of u(t) and v(t) are the same as $G_p(t)$ and $G_e(t)$ respectively. However, since $G_p(t)$ and $G_e(t)$ are both proportional to time-to-go, the controls decrease linearly with time. This is true despite the fact that $C_p(t)$ and $C_e(t)$, the deterministic feedback gains, peak sharply near the terminal time; see Figure 7-3.

Now consider the stochastic game, with which the deterministic game of Figure 7-22 is associated, where the evader is making noisy measurements. It might be argued that the standard deviation of the initial estimate (standard deviation = $\sqrt{P_o} = 10.$) is too large compared to the initial values of y(t) to make the problem realistic. However, if the standard deviation is small compared with y_o, the problem is essentially a deterministic one, since the evader's estimation error will cause

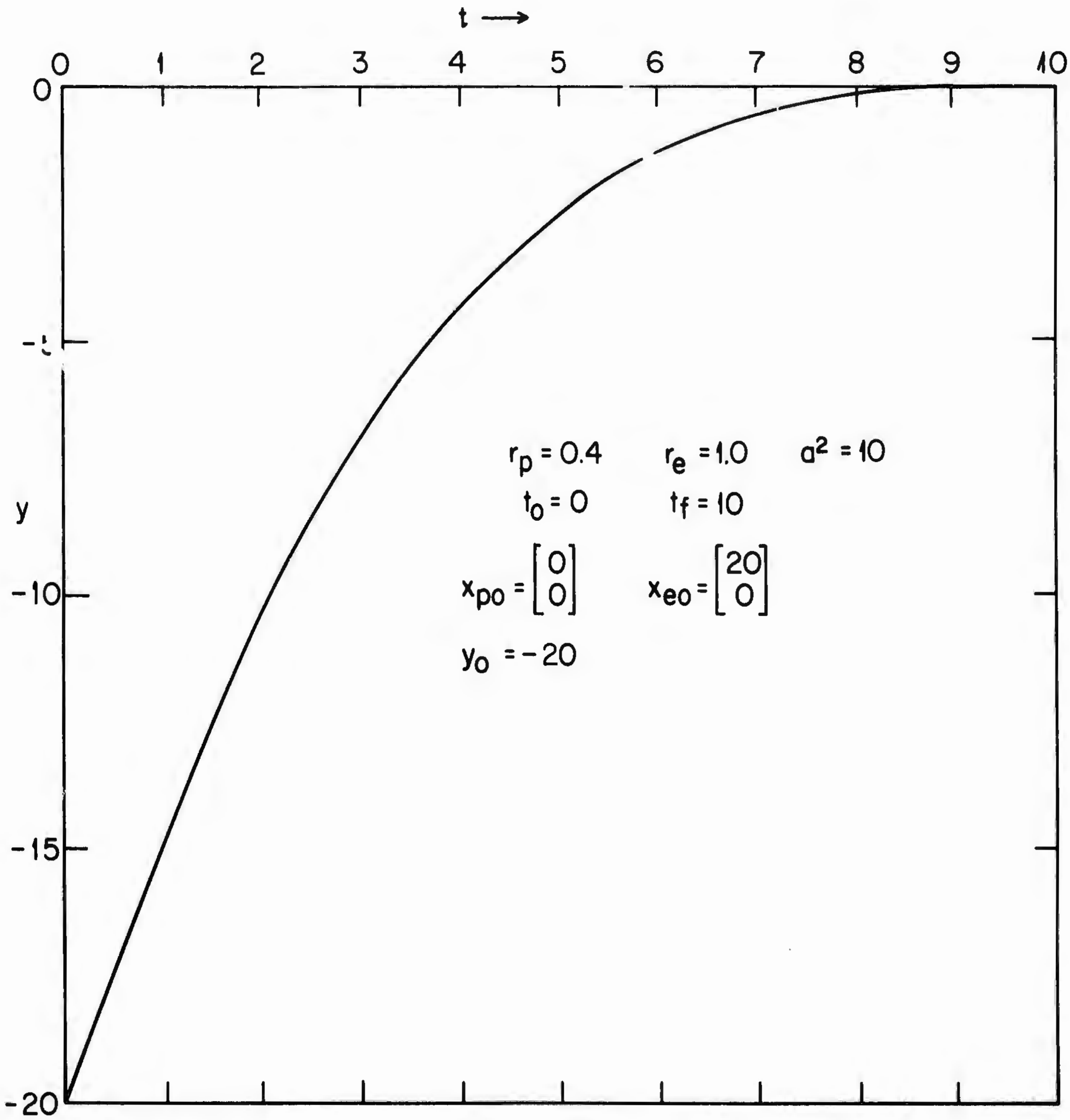


FIG. 7-23 THE REDUCED STATE FOR A DETERMINISTIC GAME

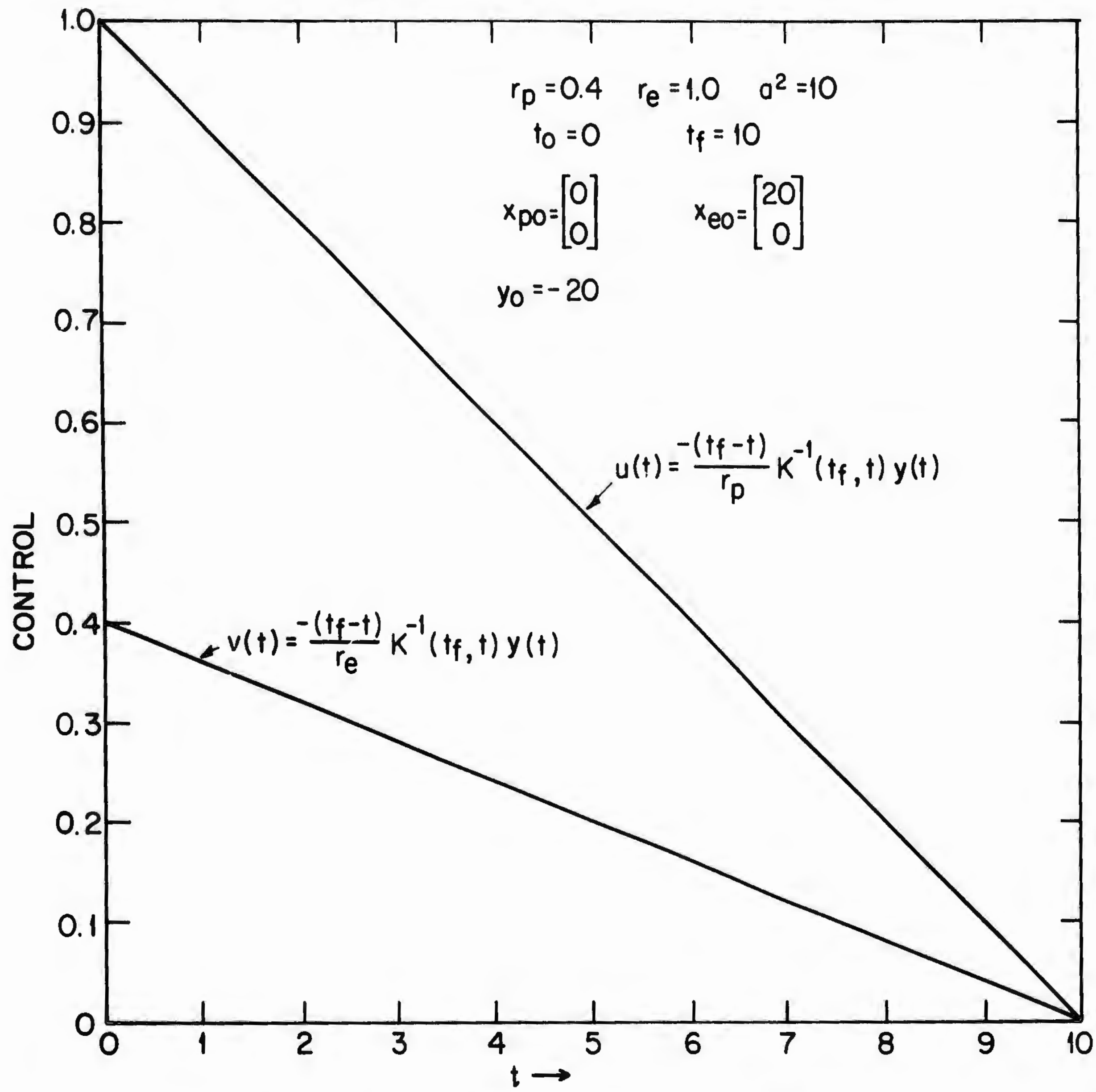


FIG. 7-24 THE CONTROLS FOR A DETERMINISTIC GAME

little deviation from the deterministically optimal control. The value of P_o used was selected to ensure that the deviations from the deterministic game would be quite visible graphically.

Figure 7-25 displays an optimal trajectory for a particular sample of the noise process w(t), with the evader making a perfect initial estimate of the reduced state, i.e. $y_o = \hat{y}_o$. Observe that there is substantial deviation from the deterministic trajectory, indicated by a dotted line, which is also the expected stochastic trajectory.

Since the evader's initial estimate is known, the most accurate determination of the criterion for either dimension is the one given by (6:28); for the scalar case this is

$$J_p(t_o, U^o, V^o) = \frac{1}{2}K^{-1}(t_f, t_o)y^2(t_o) + \frac{1}{2}\Gamma_2(t_o)\tilde{y}^2(t_o) + \frac{1}{2}\int_{t_o}^{t_f} \frac{P^2(t)\Gamma_2(t)}{q}\, dt \quad . \tag{7:67}$$

This is precisely the pursuer's evaluation of the criterion after being given the initial data. The first term is merely the deterministic criterion and the second is zero since the initial estimate is perfect. The third term has been calculated numerically and is given by

$$\frac{1}{2}\int_{t_o}^{t_f} \frac{P^2(t)\Gamma_2(t)}{q}\, dt = -8.3953 \times 10^{-3} \quad . \tag{7:68}$$

This number is valid for all games with parameters given by (7:63).

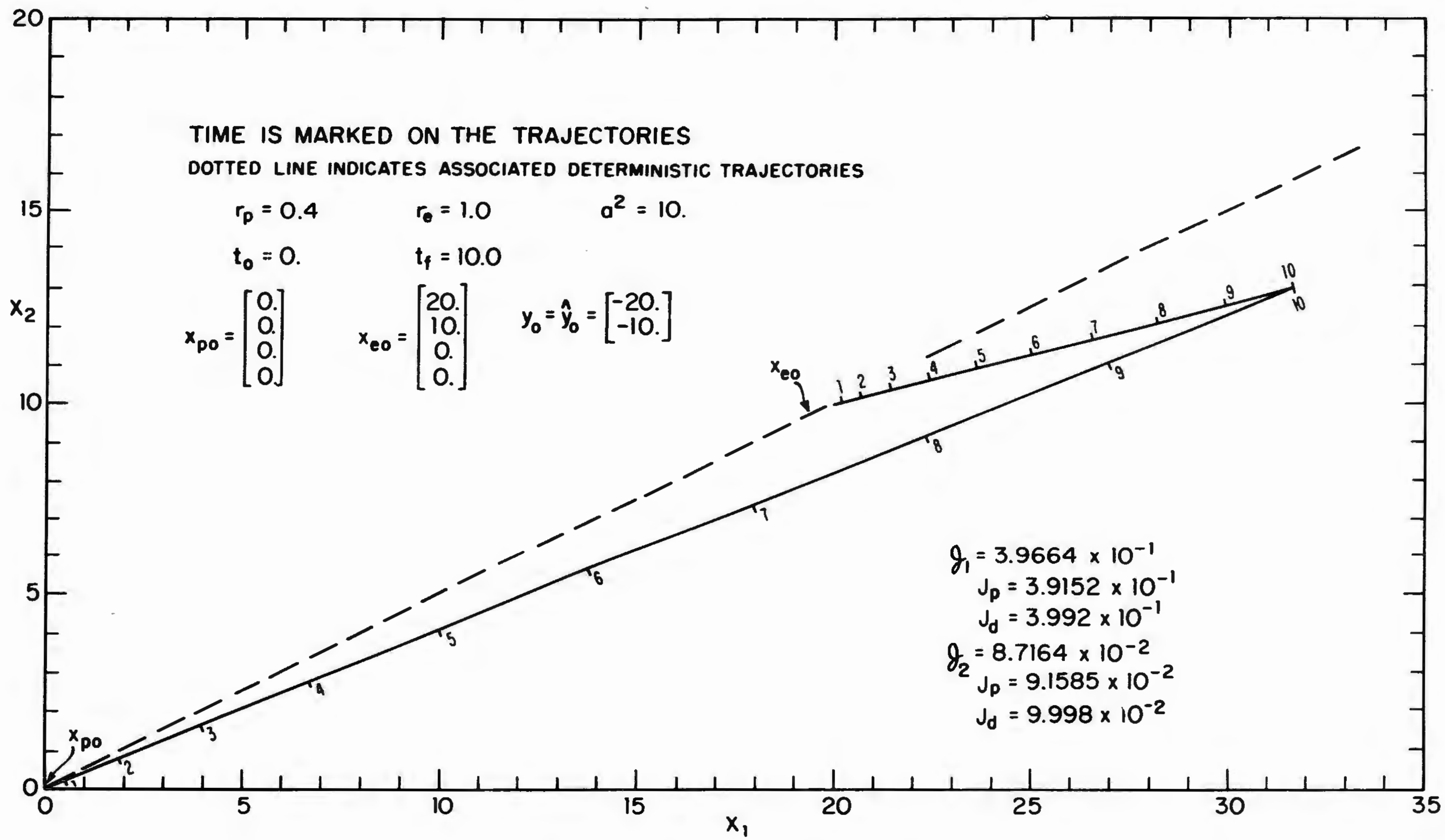


FIG. 7-25 OPTIMAL STOCHASTIC TRAJECTORY I

The values of J_p in the x_1 and x_2 dimensions are 3.9152×10^{-1} and 9.1585×10^{-2} respectively; these are merely 8.3953×10^{-3} subtracted from the deterministic criterion. Observe that the two outcomes, $\mathcal{J}_1 = 3.9664 \times 10^{-1}$ and $\mathcal{J}_2 = 8.7164 \times 10^{-2}$, are slightly higher and lower respectively than the corresponding J_p.

Figure 7-26 displays y(t) and $\hat{y}(t)$ corresponding to the sample trajectory along the x_1 axis in Figure 7-25. The deterministic value of y(t) does not differ enough from the stochastic value to be displayed at the scale used. The curve of $\hat{y}(t)$ which is displayed was obtained by connecting samples of $\hat{y}(t)$ taken at intervals of $\Delta t = 0.05$; the actual process would have been impossible to display. From the optimized stochastic system equations, (2:94), it can be seen that $\hat{y}(t)$ is once integrated white noise, while y(t) is twice integrated. Thus the curve for y(t) is relatively smooth compared with $\hat{y}(t)$.

Figure 7-27 displays the controls for this play of the stochastic game. Again because of the random nature of the controls, the curves were obtained by connecting samples taken at intervals of $\Delta t = 0.05$. Note that u(t) is less than its deterministically optimal value when v(t) is also. The controls are (usually) less than the values for the associated deterministic game because the estimate y(t) is (usually) less than the actual value.

The above game has also been played using certain non-optimal strategies for the same sample of the noise process w(t). The resulting trajectories were too near the optimal ones to be displayed graphically,

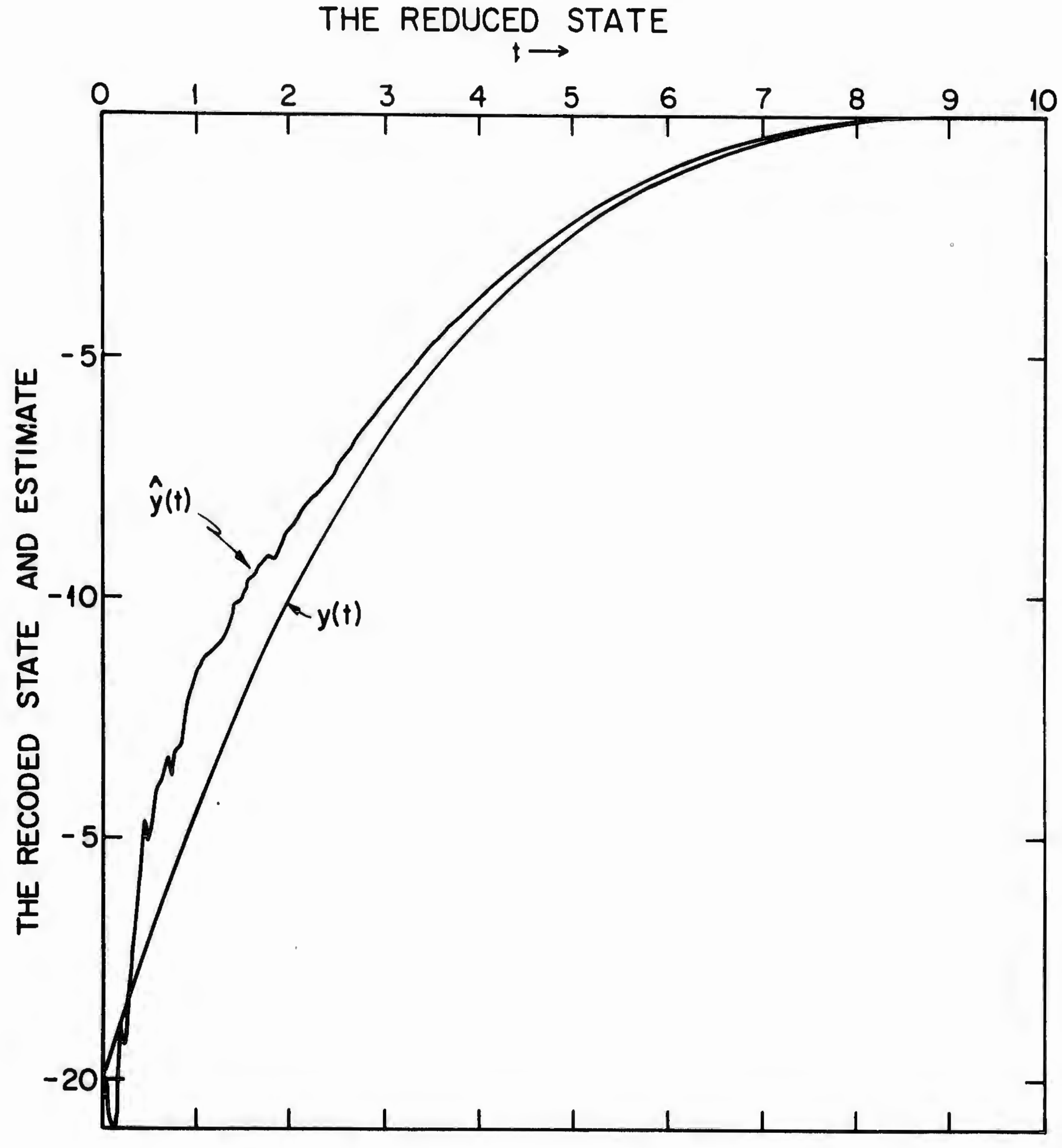


FIG. 7-26 THE REDUCED STATE AND ESTIMATE FOR A STOCHASTIC GAME

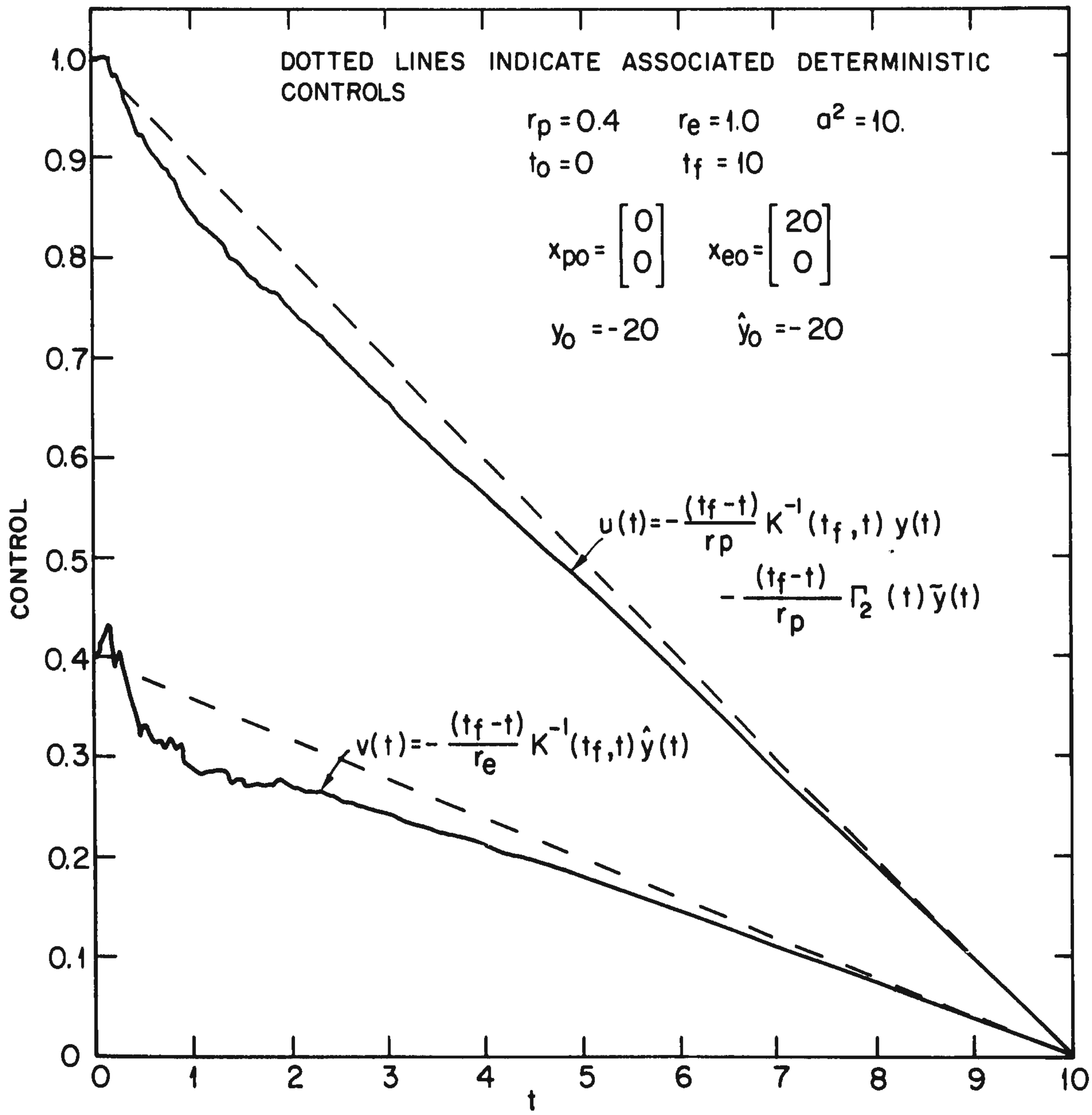


FIG. 7-27 THE CONTROLS FOR A STOCHASTIC GAME

though the outcomes ($\mathcal{C}$) did differ. Figure 7-28 displays these and other evaluations of the criterion in tabular form.

Consider the case where the pursuer employs a non-optimal strategy. Here the pursuer fails to take advantage of his opponent's estimation errors and merely employs his deterministic strategy,

$$u(t) = -R_p^{-1}(t)G_p^T(t)K^{-1}(t_f, t)y(t) \quad . \tag{7:69}$$

The evader still employs his optimal strategy as previously determined. The result is that the outcome increases in both the x_1 and x_2 dimensions; the pursuer, who is playing non-optimally, is the minimizing player.

For the evader playing non-optimally, his strategy was altered to employ a different calculation of P(t). Instead of using (7:23), the dependence on $\Gamma_2(t)$ was omitted.

$$\dot{P} = -2\frac{(t_f - t)^2}{r_p} K^{-1}(t_f, t)P - \frac{P^2(t)}{q} \quad . \tag{7:70}$$

The resulting estimation equation is the one appropriate for an outside, noisy observer of the deterministic game.

From the last column in Figure 7-28, it can be seen that the criterion decreases in one dimension but increases in the other. Thus in one instance the evader actually gains (increases the value of the outcome) by playing non-optimally. This does not mean that the results of this thesis are invalid. The strategies obtained herein are only optimal from the point of view of the expected value of the criterion. They are optimal on the average, though for particular noise sequences other strategies may produce better outcomes.

	Deterministic Game (I)	Stochastic Game (I)			
Dimension in Figures 7-22 & 7-25	J	J_p Predicted given y_o & $\hat{y}_o$	Actual Outcome	For Pursuer Non-Optimal	For Evader Non-Optimal
x_1	3.9992×10^{-1}	3.9152×10^{-1}	3.9664×10^{-1}	3.9706×10^{-1}	3.9654×10^{-1}
x_2	9.998×10^{-2}	9.1585×10^{-2}	8.7164×10^{-2}	8.9905×10^{-2}	8.7613×10^{-2}

Figure 7-28

Table of Various Criteria

The trajectories of Figures 7-22 and 7-25 may appear to be slightly dull since the initial velocities for the two players are zero. Consequently consider the deterministic game where the initial conditions are given by

$$x_{po} = \begin{bmatrix} 0. \\ 0. \\ -2. \\ -1. \end{bmatrix} , \tag{7:71}$$

$$x_{eo} = \begin{bmatrix} 2. \\ 20. \\ 1. \\ -0.5 \end{bmatrix} . \tag{7:72}$$

The optimal trajectories are found in Figure 7-29.

The vectors originating at the initial location of the two players indicate their respective velocities. Note the pursuer must reverse his direction of travel in the x_1 dimension, while the evader must slow his progress towards the pursuer in the x_2 direction. Again, the terminal separation cannot be displayed graphically.

Figure 7-30 displays a trajectory for the stochastic game associated with the deterministic one of Figure 7-29. The initial value of the state is

$$y_o = \begin{bmatrix} -32. \\ -5. \end{bmatrix} \tag{7:73}$$

while the evader's estimate of this state is

$$\hat{y}_o = \begin{bmatrix} -20. \\ -10. \end{bmatrix} . \tag{7:74}$$

Again the stochastic trajectory differs from the deterministic one, even more so because the evader's initial estimate is in error.

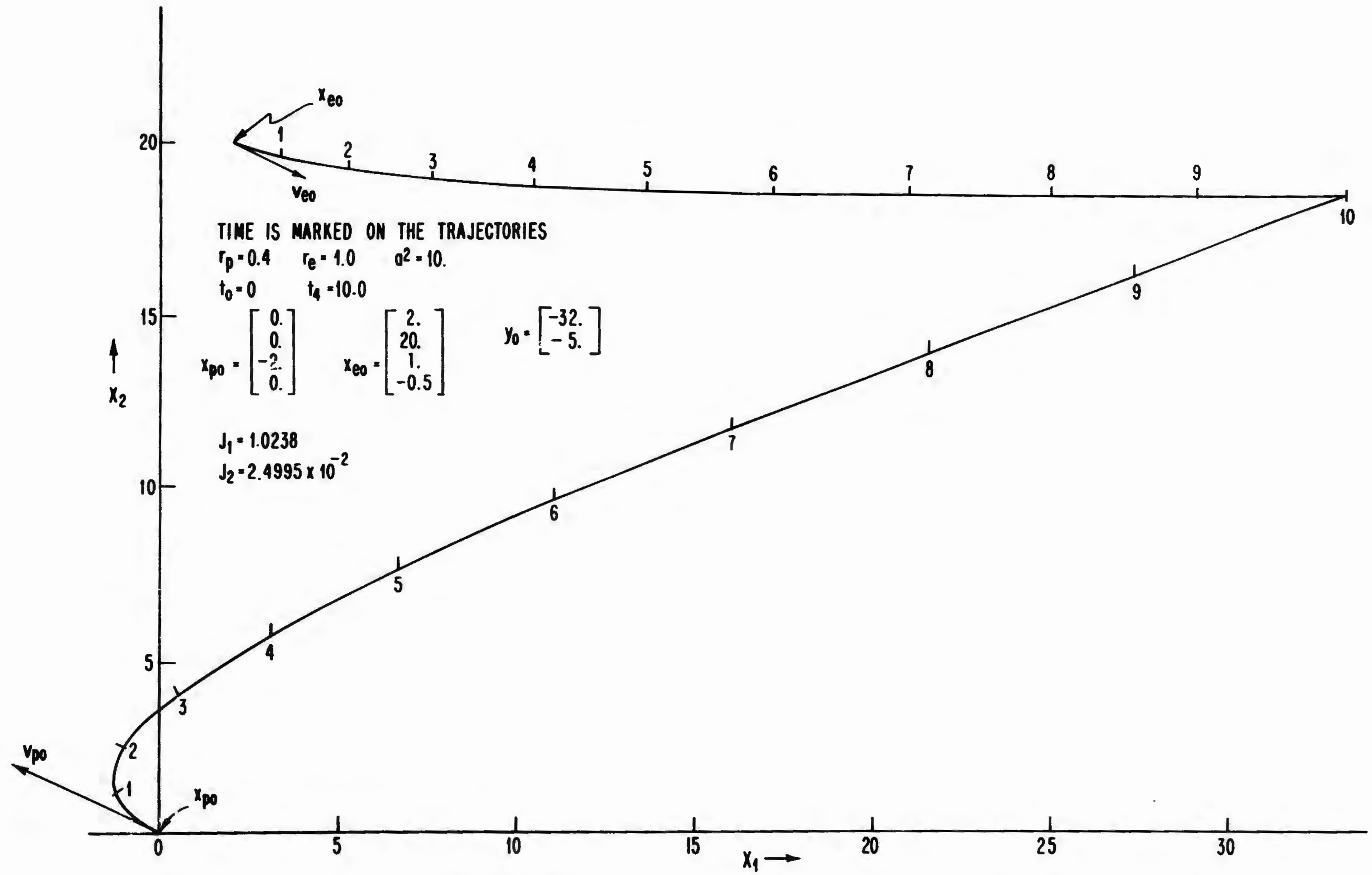


FIG. 7-29 OPTIMAL DETERMINISTIC TRAJECTORY II

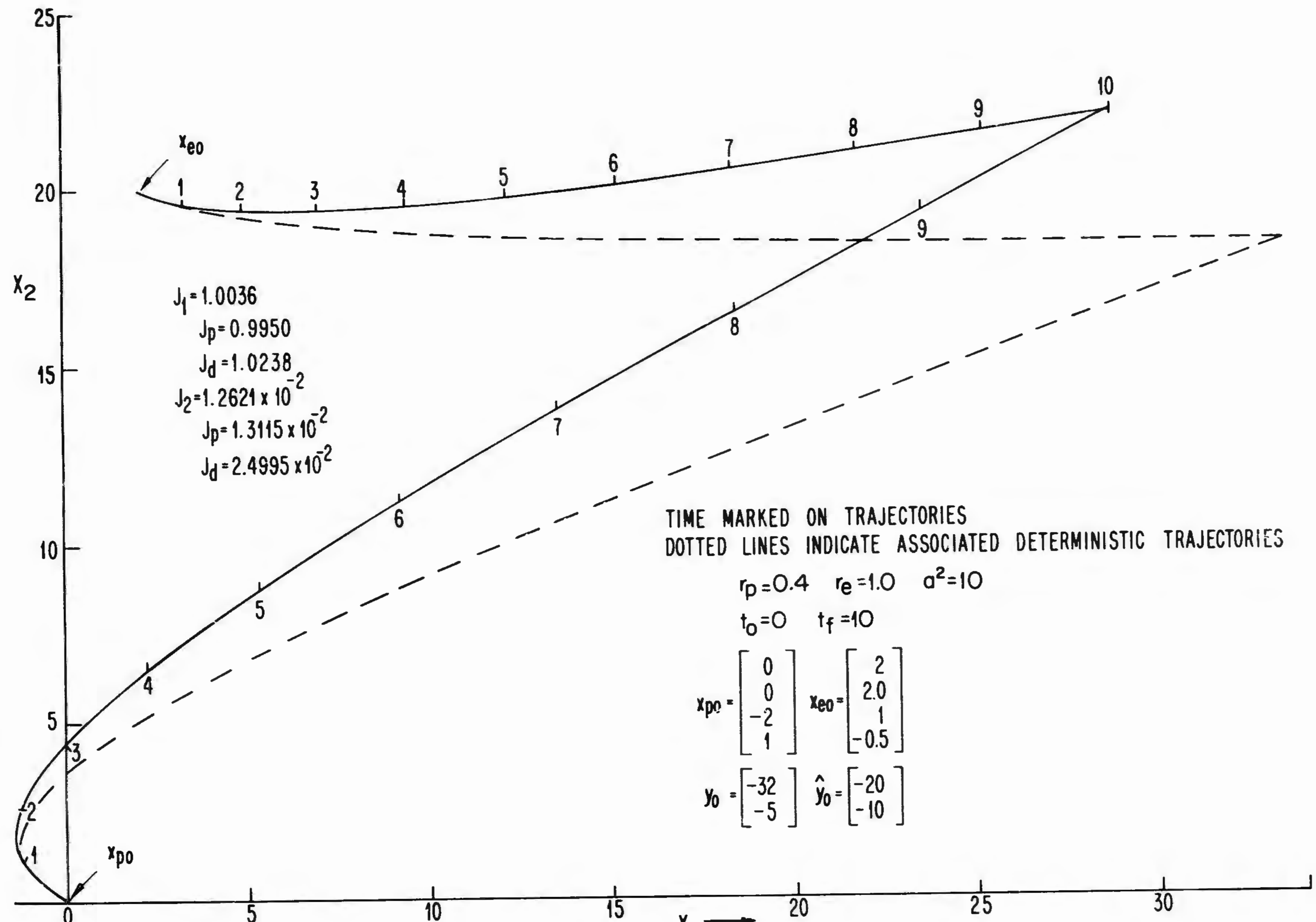


FIG. 7-30 OPTIMAL STOCHASTIC TRAJECTORY II

To predict the value of the criterion, for this case, the term $\frac{1}{2}\Gamma_2(t_o)\tilde{y}^2(t_o)$ must be included. These expected values, J_p, are 0.9950 and 1.3115×10^{-2} in the x_1 and x_2 dimensions respectively. Note that the outcomes, 1.0036 and 1.2621×10^{-2}, are near these expected values.

References for Chapter Seven

[1] A. Puckett and S. Ramo, Guided Missile Engineering, New York, McGraw-Hill, 1959, pp. 176-180.

[2] Y. C. Ho, A. E. Bryson, Jr. and S. Baron, "Differential Games and Optimal Pursuit-Evasion Strategies," IEEE Transactions on Automatic Control, AC-10, No. 4, Oct. 1965, pp. 385-389.

CHAPTER EIGHT

CONCLUDING REMARKS

8.1 A Summary of the Approach

There are three significant steps which are fundamental to the solution of the stochastic differential game considered in this thesis.

1) It must be realized that the pursuer can take advantage of the error of his opponent's estimate of the state; in particular, the pure guess that the pursuer's control is of the form

$$u(t) = C_p(t)y(t) + D_p(t)\tilde{y}(t) \tag{8:1}$$

is essential for the solution of this problem. The discovery that the evader's control is of the form

$$v(t) = C_e(t)\hat{y}(t) \tag{8:2}$$

takes less luck for this follows more or less directly from the certainty-equivalence principle of optimal control theory.

2) The stochastic problem is converted to a deterministic one using the assumptions of the control forms in (8:1) and (8:2).

3) Now standard calculus-of-variations techniques can be employed to <u>simultaneously</u> optimize both the pursuer's and the evader's controls. This gives the values for the feedback gains as determined by the equations for K^{-1}, Γ_2, and P.

Then, as is standard for any problem in game theory, the strategies must be shown to satisfy the saddle-point or equilibrium condition. It is essential to note, however, that the equilibrium inequalities

$$J_p(U^o, V^o) \leq J_p(U, V^o) \quad , \tag{8:3a}$$

$$J_e(U^o, V) \leq J_e(U^o, V^o) \tag{8:3b}$$

cannot be used to solve the problem in a sequential manner, first optimizing one player's control (8:3a) and then the other's (8:3b). In fact, even knowing the answer, the author has not found a different, logical approach which will produce the correct solution.

To see this, attempt to solve this problem by first letting the evader's control be of the form $-R_e^{-1}G_e^TK^{-1}\hat{y}$. $\hat{y}$ would be obtained from the estimation equation

$$\dot{\hat{y}} = -(G_pR_p^{-1}G_p^T - G_eR_e^{-1}G_e^T)K^{-1}\hat{y} + PH^TQ^{-1}(z - H\hat{y}) \quad ,$$

$$\hat{y}(t_o) = y_o \quad . \tag{8:4}$$

Certainly this approach is both reasonable and on the right track. All the terms in the estimation equation (8:4) -- with the exception of P -- are either given in the definition of the problem or derived in the deterministic game. Since this approach is assumed to be based on no previous analysis of the problem, the only form of the equation governing P that could "logically" be used would be the one employed by an inaccurate, outside observer of the deterministic game.

Thus P would be derived from

$$\dot{P} = -G_pR_p^{-1}G_p^TK^{-1}P - PK^{-1}G_pR_p^{-1}G_p^T - PH^TQ^{-1}HP \quad ,$$

$$P(t_o) = P_o \tag{8:5}$$

which is similar to the one obtained in Chapter 2, (2:97). The dependence of P on Γ_2 is omitted however, for the very existence of Γ_2 is yet unknown.

The next step in this method of solution would be for the pursuer to attempt to satisfy (8:3a) by minimizing J_p against the evader's strategy as determined by (8:4) and (8:5). Here he would essentially follow the analysis of Section 2. 3, Proof of the Pursuer's Optimality. Since (8:5) differs from (2:97), the evader would obviously obtain a value for his control which would be different from the one obtained in Chapter 2. As can be seen from Section 2. 3, the pursuer would determine a control that had the form of (8:1), but that had different values from the feedback matrices that were determined in Chapter 2.

Now the evader would attempt to satisfy (8:3b) by maximizing J_e against the pursuer's strategy which resulted from the previous paragraph. His approach would be similar to that of Section 2. 4, Proof of the Evader's Optimality. But because the values of the pursuer's feedback gains would differ from those in Section 2. 4, the evader would not obtain the exact results of this section. Recall that in Section 2. 4, the pursuer was given a $\hat{y}_p$, by the mystical third party, which was based on the assumption that the evader played "optimally." When the evader determined his strategy, he found that $\hat{y}_p$ was indeed identical to his optimal estimate. However, if the pursuer's feedback gains are not $-R_p^{-1}G_p^TK^{-1}$ and $-R_p^{-1}G_p^T\Gamma_2$, this will not in general be true, and consequently, the evader will employ a feedback strategy consisting not only of the term dependent on $\hat{y}$, but also on a second term which is dependent on $\hat{y}_p$ which is different.

Suppose the attempt to satisfy (8:3b) and obtain the evader's optimal strategy follows the approach of Section 3.3, Another Proof of the Evader's Optimality, instead of the one of Section 2.4. Then the result will be of the form given in (8:2), but C_e will not be $-R_e^{-1}G_e^TK^{-1}$.

Thus at the end of one iteration of this sequential approach to the solution, neither method for optimizing the evader's control has produced the correct result. A series of iterations might converge to the correct solution, but the labor involved and the possibly awkward form make this approach impractical. Unless the form and values of one of the player's strategies are guessed correctly, the sequential approach seems doomed on a bier of algebra.

Possibly a discrete-time, dynamic programming approach similar to the one discussed in Appendix 3A could be employed to obtain the correct solution by considering one time interval at a time. However, the algebraic complexities of this Appendix warn that the result may be difficult to obtain from first principles, and obscure once determined.

The particular sequential approach -- consisting of alternately employing the analyses of Section 2.3 and Section 3.3 -- can be viewed as an alternative to guessing the forms of the optimal strategies as given in (8:1) and (8:2). Note that the forms of the feedback controls do not change over one cycle of the iteration; only the values of the feedback gains do. Consequently, this approach demonstrates that if only the appropriate values of the feedback gains can be found, the equilibrium condition (8:3) can be satisfied. Once this has been established, simultaneous optimization can determine these values. Thus, in a sense, the sequential approach provides an alternative to the guess employed in Step 1).

Consequently, it appears that the only approach which produces a closed-form solution is the one where 1) the correct form of the feedback strategies is somehow obtained, 2) the problem is converted to a deterministic one, and 3) the criterion is optimized simultaneously with respect to both players' controls.

8.2 The Closure Problem

Chapters 2 through 7 of this thesis present the solution and implications of the pursuit-evasion game in which one of the two players has noisy measurements of the state of the game, while the other player has perfect information as to this state. This particular problem was tackled after attempts to solve the game where both players have noisy measurements proved unsuccessful. It is worthwhile to note why this second, more general problem (here denoted Problem II) is much more difficult to solve than the first.

The problem solved in this thesis does not include all games where one player makes noisy measurements. Unless one is willing to accept the implausible existence of the mystical third party, it is necessary to check the three conditions of Section 3.1 to ensure that the solution is applicable.

i. The number of interesting state variables must equal the number of the evader's control variables.

ii. The inverse of $G_e R_e^{-1} G_e^T$ must exist.

iii. The inverse of L must exist.

A problem which satisfies these three conditions is denoted as Problem IA.

If conditions i. and ii. are not satisfied*, the problem is entirely different, for the pursuer cannot obtain the data necessary to implement the optimal strategy of Chapter 2; this problem is called Problem IB. The difficulty encountered in solving Problem IB (in particular for the cases where the dimension of the reduced state vector is smaller than the dimension of the evader's control vector) is essentially the same one which obscures the solution to Problem II.

It is interesting to see how the analysis of Problem IB proceeds. The first step is to assume that the evader's controller is as given in Chapter 2. Then from the analysis of Section 2.3 and the certainty-equivalence principle of control theory, it can be seen that the pursuer would employ a strategy similar to the one obtained in Chapter 2. However, the inverse of $G_e R_e^{-1} G_e^T$ does not exist and thus the pursuer cannot calculate $\hat{y}(t)$; he must build a Kalman-Bucy filter -- based on the evader's estimation equation -- from which he obtains an estimate of the evader's estimate, denoted $\hat{\hat{y}}$. Thus his control would look like

$$u(t) = -R_p^{-1} G_p^T K^{-1} y - R_p^{-1} G_p^T \Gamma_2 (y - \hat{\hat{y}}) \quad . \tag{8:6}$$

---

* Condition iii. determines the time duration of the game for which the pursuer can realize his strategy; it does not effect the existence of a general solution, however. Conditions i. and ii. ensure that a solution will exist if the game's time duration is short enough.

Obviously, this control would deviate from the one employed if the pursuer could calculate $\hat{y}$, or if he was given this data by the mystical third party.

Consequently, the evader should be able to take advantage of this error in the pursuer's estimate of the evader's estimate of the state. The above analysis suggests that the proper approach would be to add an additional term to the evader's control. However, such a correction term would be based on noisy data, and thus would be "imperfect." The pursuer could then take advantage of such an imperfection.

The difficulty that is encountered is now obvious; there appears to be an infinite number of terms in the optimal strategies of each of the two players. These terms are based on estimates of estimates of estimates of ..... This problem shall be called the "closure problem" of stochastic differential games. When conditions i. and ii. of Chapter 3 are satisfied, closure is possible for the game in which only one player makes noisy measurements; in this case the strategies can be written in closed-form consisting of a finite number of terms.

The closure problem is also a part of the second problem when both players make noisy measurements on the state. Here the first step would be to assume that the forms of the players' strategies are

$$u(t) = C_p(t)\hat{y}_p(t) + D_p(t)\tilde{y}_e \quad , \qquad (8:7)$$

$$v(t) = C_e(t)\hat{y}_e(t) + D_e(t)\tilde{y}_p \qquad (8:8)$$

where $\hat{y}_p$ and $\tilde{y}_p$ are the pursuer's estimate and estimation error, respectively; similar statements apply to $\hat{y}_e$ and $\tilde{y}_e$. However, since neither player can determine his opponent's estimation error, these strategies are nonrealizable. Consequently, correction terms must be added, initiating the vicious cycle of estimates of estimates.

It can be argued that this sequential approach is improper -- that one only need guess the correct form of both players' strategies and then optimize simultaneously to obtain the proper values. After all, didn't the author just criticize the sequential method of solution as invalid even for Problem IA? The argument would continue to point out that the above analysis does not prove that optimal strategies do not exist.

Certainly it is possible that closed-form, linear strategies exist for these two problems. Perhaps it is only that the author has not been clever enough to guess them correctly. However, recall that the sequential process of Section 8.1 (which consisted of using the approach of Section 2.3 to optimize the pursuer's strategy followed by the approach of Section 3.3 to optimize the evader's strategy) produced results that always had the form of (8:1) and (8:2); only the values of the feedback gains differed from the optimal ones. The problem of convergence was one of numerical value, not one of mathematical form. Consequently, sequential analysis for Problem IA leads directly to strategy forms which are closed upon sequential iteration, and thus permits the values to be determined by simultaneous optimization.

The analysis above for Problems II and IB immediately reveals a need for an infinite number of correction terms, which prevents the use of the calculus-of-variations for simultaneous optimization. This discussion is presented to point out the problem of extrapolating the logic used to obtain the results in this thesis to guess strategies for these two other problems. If the optimal control laws for either Problem II or IB are to be linear, it would appear that the number of simple and "logical" choices for strategy forms would be finite. Yet no solutions have been found. Perhaps an entirely new approach is needed to solve Problems II and IB. The author is aware that both Rhodes [1] and Willman [2] are working on this.

8.3 A Summary of the Significant Results

Chapter 2 presented the optimal feedback strategies for a stochastic pursuit-evasion game in which one player had perfect information of the game's state, while his opponent had only noisy measurements. The values of these strategies were determined, and shown to satisfy the equilibrium condition.

Chapter 3 demonstrated how the player with perfect information could realize his strategy which called for him to feedback his opponent's estimation error. Conditions under which the player with perfect information could calculate this estimation error were determined.

Chapter 4 discussed the coupled equations for Γ_2 and P which are fundamental to the solution. It was pointed out that the additional conjugate point condition for the stochastic game was that Γ_2 remain finite. However, it was also observed that because of the nature of the equations determining Γ_2 and P and because of the finite value of $P(t_o)$ this would always be true.

Chapter 5 considered the asymmetric character of the capabilities of the pursuer and evader for both the deterministic and stochastic game. The limiting relationship between the stochastic game and the open-loop deterministic game was demonstrated as the measurement noise variance approached infinity.

Chapter 6 pointed out the need to carefully consider the non zero-sum properties of this game which result from the two players' different information sets. It was demonstrated, however, that the strategies obtained in Chapter 2 did satisfy the equilibrium condition. Further, it was shown that if other, non-linear strategies satisfied the equilibrium condition, they gave the same value for all evaluations of the criterion and satisfied the equilibrium condition when played against the strategies of the original solution. Consequently, the certainty-equivalence principle was found to apply for this game.

Chapter 7 demonstrated that the solution obtained in this thesis was applicable to the interception game in euclidean space. Numerous graphs illuminated this solution, and illustrated the points made in previous chapters.

References for Chapter Eight

[1] I. B. Rhodes, private communication, 1968.

[2] W. W. Willman, private communication, 1968.

ACKNOWLEDGEMENTS

To Professor Yu Chi Ho I wish to express my sincere thanks for his guidance and assistance in the preparation of this work. His generous contributions of time and insight were a constant source of encouragement and the principle motivation for this work.

I am grateful to my collegues, Dr. Philip M. Newbold, Dr. Jason L. Speyer and Mr. Warren W. Willman, for the time they devoted to listening to my problems and helping me unravel an unending ball of difficulties. I also wish to thank Professor Arthur E. Bryson, Jr. and Ian B. Rhodes who read and commented on the manuscript.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

Division of Engineering and Applied Physics
Harvard University
Cambridge, Massachusetts 02138

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

LINEAR STOCHASTIC DIFFERENTIAL GAMES

4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Interim technical report

5. AUTHOR(S) (First name, middle initial, last name)

Robert Dietrich Behn

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1969	266	29

8a. CONTRACT OR GRANT NO.

N00014-67-A-0298-0006

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

Technical Report No. 578

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted by the U. S. Government.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Office of Naval Research

13. ABSTRACT

The solution for a class of stochastic pursuit-evasion differential games between two dynamic systems is given; this class includes those games where one of the players has perfect knowledge of the state of the game while the other player is constrained to make noisy measurements on this state. The dynamic systems involved are linear and the performance index which is optimized is quadratic.

The strategy for the player with perfect information is not always a realizable one. It is shown that this player can implement his strategy, however, if the number of his control variables is as great as the number of the state variables involved in the pursuit and evasion. Thus the solution obtained is applicable for the classical interception game in euclidean space.

Several aspects of this game are studied in detail. The asymmetric roles of the pursuer and evader are discussed in general and relationships drawn between the deterministic and stochastic cases. It is pointed out that this game requires -- in reality -- the solution to a non zero-sum game since the two different information sets employed by the two players cause each player to evaluate the criterion differently. The "certainty-equivalence principle" which characterizes the standard stochastic control problem is shown to be applicable to this class of differential games.

Examples of the classical interception game are given and numerical results presented.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified

Security Classification

A-31408

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Stochastic						
Differential Game						
Linear-Quadratic						
Pursuit-Evasion						
Certainty-Equivalence Principle						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0101-807-6821

Unclassified
Security Classification

A-31409