1973

# Parameter identification applied to linear quadratic differential games

John D. Corrigan



Department: Electrical and Computer Engineering

PARAMETER IDENTIFICATION APPLIED TO

LINEAR QUADRATIC DIFFERENTIAL GAMES

by

John Donald Corrigan, 1942-

A DISSERTATION

Presented to the Faculty of the Graduate School of the

UNIVERSITY OF MISSOURI - ROLLA

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

IN

ELECTRICAL ENGINEERING

1973



T. Hauck
Advisor

John S. Pazdera, Jr.

Daniel K. Goodman

Frank J. Kern

D. Ronald Fannin

S. J. Pagano

153

# ABSTRACT

A two player zero-sum linear quadratic differential game is investigated for the case in which one of the players has incomplete a priori knowledge of the parameters of his opponent's dynamic system. This incomplete system parameter information game is shown to be playable since the ignorant player can make limiting estimates of the unknown parameters from the relative controllability condition for the game. Performance from the ignorant player's point of view is suboptimal.

It is also shown that parameter identification techniques can be applied by the ignorant player in order to directly identify the smart player's closed-loop parameters in the case in which the smart player's optimal control gains become time-invariant. The open-loop system parameters may then be estimated from the identified closed-loop parameters. Using these estimated open-loop parameters in the optimal control law results in an asymptotically optimal adaptive control strategy for the ignorant player.

Both continuous and discrete time parameter identification techniques were applied to the incomplete system parameter information game. In doing so, multivariable extensions were derived for previously developed single input/output continuous time and discrete time identification techniques. A multivariable combination response error and equation error continuous time learning model identification technique was also developed.

## ACKNOWLEDGEMENT

The author wishes to express his appreciation to Dr. T. L. Noack for having suggested the dissertation topic and for his continued interest in this research. A special thanks is extended to Dr. J. S. Pazdera for the many hours spent discussing the characteristics of continuous time learning model identification techniques and to Dr. D. K. Goodman for information on discrete time parameter identification techniques.

The author also would like to thank his colleagues and the management at McDonnell Aircraft Company for their cooperation during the period of this research.

Finally, the author wishes to thank his wife for her patience and encouragement during these almost eight years of both marriage and graduate study and also for having typed this manuscript.

# TABLE OF CONTENTS

# LIST OF ILLUSTRATIONS

# LIST OF TABLES

# I. INTRODUCTION

Games have been played since the beginning of time. A game exists, of course, whenever two or more persons who can affect the future strive for opposite or conflicting goals. Games can also be played by teams, which involves teamwork between players. Naturally, if the goals of the players or teams are not diametrically opposed, play of the game can involve negotiation and cooperation between players and teams.

Most games require that some form of mathematical score be kept. If it is strictly the magnitude of the score that is important, the game may be termed a quantitative game. However, if the score is utilized only in determining a winner and a loser, the game is actually a qualitative game. Most qualitative games are defined in terms of quantitative games.

Quite often the play of the game can be described mathematically. In this case, the best strategy for each player can usually be formulated mathematically or, at least, statistically.

Games whose play is governed by differential equations occur quite often. The intercept game is just one of many examples of this type of game. Differential games, as these games are usually termed, are of great interest in the field of control theory. This dissertation examines one particular type of differential game for the case in which one of the players has incomplete information concerning his opponent's dynamic system.

Before proceeding, it is informative to examine the research related to the field of game theory and, in particular, that of differential game theory.

John Von Neumann was probably the first to define games within the framework of modern mathematics [31]. Von Neumann and Morgenstern's book [32] on the application of game theory to the analysis of economic behavior is the classic work in the field.

Von Neumann's work stimulated a vast amount of research in the field of game theory, as documented by references [14], [15], [7], and [29]. Although these references do include limited research in the area of games whose outcome is governed by differential equations, attention was not fucused on differential game theory until after the publication of the research performed by Rufus Isaacs [13].

Isaacs' work concentrated primarily on two player differential games, commonly known as pursuit-evasion games. In most cases, these games were zero sum, i.e., the opponent's goals were directly opposite such that one player's gain was the other player's loss. Isaacs further defined these differential games to be either games of "kind" or games of "degree". In the game of kind, the player either achieves his goal, i.e., he wins, or he does not achieve his goal, i.e., he loses. In games of degree it is the margin by which the player wins or loses that counts. Since games of degree are much more amenable to mathematical analysis, most games of kind are redefined as games of degree. A guideline is then applied to the performance measure in order to determine the actual winner.

Blaquiere, Gerard, and Leitmann [5] examined differential games from a geometric viewpoint. In this work, games of degree and of kind are more properly termed quantitative and qualitative games, respectively. Also, as in Isaacs' earlier work, it was assumed that all players had complete or perfect information concerning the game.

Baron, Bryson, and Ho [2] were probably the first to investigate differential games from a control theory point of view. Their paper discusses the solution of a perfect information zero-sum pursuit-evasion game. Since their original paper, nonzero-sum differential games have been investigated by Ho and Starr [11] and Rhodes [23]. Likewise, stochastic differential games, i.e., games with system perturbations and/or noisy measurements have been investigated by Behn and Ho [3], [4], Willman [34], and Rhodes and Luenberger [24], [25].

Although Ho [10] has discussed a generalized control theory, differential games in which the players have incomplete information about their opponent's dynamic system have not received much attention. It is this aspect of differential game theory toward which this dissertation is oriented.

In order to make the analysis tractable, attention is limited to a quantitative, zero-sum, two player differential game in which one of the players has incomplete a priori knowledge of the parameters of his opponent's dynamic system. The systems of both players and their state measurements are assumed to be noise free. Generalization to include n player games, stochastic games, and nonzero-sum games is not attempted in this paper.

The perfect or complete information version of the differential game being investigated is defined in Chapter II. This game, which assumes linear time-invariant dynamic systems and a quadratic performance index is somewhat more general than that studied by Ho [2], [6]. Its solution, which is given without proof by Rhodes [23], is included and, for the sake of completeness, an original proof of this solution is given in Appendix A. The characteristics of the game are illustrated via digital simulation of an example game. The analytic solution to the Riccati equation utilized in this simulation is included in Appendix B.

The incomplete information differential game is investigated in Chapter III in order to determine the conditions for which the game is playable. Limiting estimates for the opponent's system parameters are obtained via an original derivation based on the relative controllabiltiy requirement for the game. The characteristics of this game are compared with those for the perfect information game via simulation using the original example game.

During the course of the game, the players must respond to each other's motion. This response inherently contains information on the players' system parameters. Identification of the opponent's system parameters by the ignorant player during the play of the game is the subject of Chapter IV.

While the sufficient conditions required for parameter identification are easily derived, the derivation does not lead to a practical identification technique. However, many other parameter identification techniques have been formulated. The works by Spence [27] and by Sage and Melsa [26] include reasonably complete bibliographies on this subject.

As discussed in Chapter IV, both continuous time learning model identification techniques and discrete time recursive identification techniques appear applicable to the incomplete system parameter information differential games. Applicable continuous time learning models include the multivariable response error model developed by Pazdera and Pottinger [21], as improved by Spence [27], and the multivariable extension of the single input/output equation error model developed by Lion [17]. Application of various combinations and extensions of these models in also possible. The response error learning model is included in Appendix C for the sake of completeness. Original multivariable versions of the equation error learning model, the generalized equation error learning model, and the combined response/equation error learning model are included in Appendices D, E, and F, respectively. An applicable discrete time identification technique is the multivariable extension of the single input/output recursive technique developed by Lee [16]. An original multivariable extension of this recursion technique is given in Appendix G.

In Chapter V an asymptotically optimal adaptive control strategy is formulated for the incomplete information game of Chapter III. This control strategy is based on the recursive discrete time identification technique discussed in Chapter IV and the optimal feedback control strategies derived in Chapter II.

Chapter VI presents a brief summary of the conclusions of the present research and indicates possible future areas of investigation.

# II. ZERO SUM LINEAR QUADRATIC DIFFERENTIAL GAMES

## A. Definition of the Game

Consider the zero sum differential game defined by the decoupled linear time-invariant systems

$$\dot{\underline{x}}_p(t) = F_p\underline{x}_p(t) + G_p\underline{u}_p(t) \quad ; \quad \underline{x}_p(0) = \underline{x}_{po} \tag{1}$$

$$\dot{\underline{x}}_e(t) = F_e\underline{x}_e(t) + G_e\underline{u}_e(t) \quad ; \quad \underline{x}_e(0) = \underline{x}_{eo} \tag{2}$$

and the quadratic time-invariant performance index

$$\begin{aligned} J(\underline{u}_p,\underline{u}_e) = \tfrac{1}{2}\int_0^{t_f} \left( [\underline{x}_e - \underline{x}_p]^T Q\, [\underline{x}_e - \underline{x}_p] + \underline{u}_p^T R_p \underline{u}_p - \underline{u}_e^T R_e \underline{u}_e \right) dt \\ + \tfrac{1}{2}\,[\underline{x}_e(t_f) - \underline{x}_p(t_f)]^T S\, [\underline{x}_e(t_f) - \underline{x}_p(t_f)] \end{aligned} \tag{3}$$

where the state vectors $\underline{x}_p$ and $\underline{x}_e$ are n-dimensional, the control vectors $\underline{u}_p$ and $\underline{u}_e$ are m-dimensional, $t_f$ is the final time, and $F_p$, $F_e$, $G_p$, $G_e$, Q, $R_p$, $R_e$, and S are constant matrices of appropriate dimensions. It is assumed that Q, $R_p$, $R_e$, and S are symmetric matrices and that $R_p$ and $R_e$ are positive definite, while Q and S are non-negative. The decoupled systems (1) and (2) are assumed to be completely state controllable. In addition, the states of both systems are assumed to be accessible, i.e., directly measurable by both players.

In the zero-sum differential game defined above the two players simultaneously try to minimize and maximize the performance index J. Consequently, the optimal control strategies, denoted as $\underline{u}_p^o$ and $\underline{u}_e^o$, must satisfy the saddle point condition

$$J(\underline{u}_p^o,\underline{u}_e) \leq J(\underline{u}_p^o,\underline{u}_e^o) \leq J(\underline{u}_p,\underline{u}_e^o) \tag{4}$$

where $J(\underline{u}_p^o, \underline{u}_e^o)$ is known as the value of the game.

The pursuit-evasion structure of this differential game is obvious. Hence, the subscripts p and e are used to denote those vectors and matrices associated with the Pursuer and the Evader, respectively.

B. Linear Quadratic Representation

The previously defined differential game may be rewritten by defining the augmented state and control vectors

$$\underline{x} = [\underline{x}_p^T \quad \underline{x}_e^T]^T \tag{5}$$

and

$$\underline{u} = [\underline{u}_p^T \quad \underline{u}_e^T]^T \quad . \tag{6}$$

The resulting linear quadratic representation is

$$\dot{\underline{x}} = \bar{F}\,\underline{x} + \bar{G}\,\underline{u} \quad ; \quad \underline{x}(0) = \underline{x}_o \quad ; \tag{7}$$

$$J(\underline{u}) = \tfrac{1}{2}\int_0^{t_f} (\underline{x}^T\bar{Q}\,\underline{x} + \underline{u}^T\bar{R}\,\underline{u})\,dt + \tfrac{1}{2}\,\underline{x}^T(t_f)\,\bar{S}\,\underline{x}\,(t_f) \tag{8}$$

where $\bar{F}$, $\bar{G}$, $\bar{Q}$, $\bar{R}$, and $\bar{S}$ are augmented matrices given by

$$\bar{F} = \begin{bmatrix} F_p & 0 \\ 0 & F_e \end{bmatrix} \quad ; \quad \bar{G} = \begin{bmatrix} G_p & 0 \\ 0 & G_e \end{bmatrix} \tag{9}$$

and

$$\bar{Q} = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} \quad ; \quad \bar{R} = \begin{bmatrix} R_p & 0 \\ 0 & -R_e \end{bmatrix} \quad ; \quad \bar{S} = \begin{bmatrix} S & -S \\ -S & S \end{bmatrix} \quad . \tag{10}$$

Equations (7) and (8) are the familiar representation for the linear quadratic optimal control problem with the major exception that $\bar{R}$ is indefinite rather than positive definite. Nevertheless, $\bar{R}$ is nonsingular since $R_p$ and $R_e$ are both positive definite. Therefore, the first order necessary conditions for the differential game, which is a two-sided optimal control problem [2], may be obtained via the calculus of variations [8], [33] in the same fashion as for the one-sided optimal control problem. However, the control strategies obtained may not satisfy the saddle point condition given by (4).

If the control strategies obtained for the two resulting one-sided optimal control problems are identical to those obtained for the two-sided problem, the saddle point condition (4) is satisfied and the optimal control strategies are deterministic or pure.

C. Optimal Feedback Control Strategies

The properties of the optimal solution to the previously defined differential game have been studied in detail for the special case in which $\overline{Q} = [0]$ by Ho and others [2], [6] and by Rhodes and Luenberger [24], [25] who mention that the theory is easily extended to the more general case in which $\overline{Q} \neq [0]$. Rhodes [23], while studying the properties of nonzero sum two player linear quadratic differential games, provided without proof the solution for a slightly more general differential game than that previously defined. The optimal solution to the previously defined game is presented below, while an original proof is included in Appendix A.

The first order necessary conditions for optimal solution of the differential game are given by the optimal control

$$\underline{u}^{o}(t) = -\overline{R}^{-1}\overline{G}^{T}\underline{y}(t) \tag{11}$$

where $\underline{y}$ is the costate vector defined by

$$\dot{\underline{y}}(t) = -\bar{Q}\,\underline{x}(t) - \bar{F}^T\underline{y}(t) \quad ; \quad \underline{y}(t_f) = \bar{S}\,\underline{x}(t_f) \quad . \tag{12}$$

Substituting (11) in (7) yields the state equation

$$\dot{\underline{x}}(t) = \bar{F}\,\underline{x}(t) - \bar{G}\,\bar{R}^{-1}\bar{G}^T\underline{y}(t) \quad ; \quad \underline{x}(0) = \underline{x}_o \tag{13}$$

which, together with the costate equation (12), forms a two point boundary value problem (TPBVP) which may be written in matrix form as

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{y}} \end{bmatrix} = \begin{bmatrix} \bar{F} & -\bar{G}\,\bar{R}^{-1}\bar{G}^T \\ -\bar{Q} & -\bar{F}^T \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} \quad \begin{matrix} ; & \underline{x}(0) = \underline{x}_o \\ ; & \underline{y}(t_f) = \bar{S}\,\underline{x}(t_f) \quad . \end{matrix} \tag{14}$$

The above TPBVP may be solved by defining

$$\underline{y}(t) = \bar{P}(t)\,\underline{x}(t) \tag{15}$$

which leads to the nonlinear matrix Riccati equation

$$\dot{\bar{P}} = -\bar{P}\,\bar{F} - \bar{F}^T\bar{P} + \bar{P}\,\bar{G}\,\bar{R}^{-1}\bar{G}^T\bar{P} - \bar{Q} \quad ; \quad \bar{P}(t_f) = \bar{S} \quad . \tag{16}$$

Assuming that a solution for $\bar{P}(t)$ exists, the optimal feedback control may be written as

$$\underline{u}^o(t) = -\bar{R}^{-1}\bar{G}^T\bar{P}(t)\,\underline{x}(t) \quad . \tag{17}$$

Partitioning $\bar{P}(t)$ such that

$$\bar{P}(t) = \begin{bmatrix} P_{pp}(t) & P_{pe}(t) \\ P_{ep}(t) & P_{ee}(t) \end{bmatrix} \tag{18}$$

and performing the operations indicated in (17), the pure optimal feedback control strategies become

$$\underline{u}_p^o(t) = - R_p^{-1} G_p^T P_{pp}(t)\, \underline{x}_p(t) - R_p^{-1} G_p^T P_{pe}(t)\, \underline{x}_e(t) \tag{19}$$

and

$$\underline{u}_e^o(t) = + R_e^{-1} G_e^T P_{ep}(t)\, \underline{x}_p(t) + R_e^{-1} G_e^T P_{ee}(t)\, \underline{x}_e(t) \tag{20}$$

where $P_{ep} = P_{pe}^T$ since $\overline{P}(t)$ is symmetric. Note also that if $\overline{P}(t)$ is non-negative, then $\underline{u}_p^o$ and $\underline{u}_e^o$ are respectively stabilizing and destabilizing feedback control strategies, as is characteristic of the pursuit-evasion type differential game.

The feedback control strategies given by (19) and (20) have been derived from the first order necessary conditions given by (11) and (12). As noted previously, in order for these strategies to be deterministic optimal control strategies, they must satisfy the saddle point condition given by (4). Sufficient conditions for $\underline{u}_p^o(t)$ and $\underline{u}_e^o(t)$ to be deterministic optimal feedback control strategies are:

(1) $R_p$ and $R_e$ are both positive definite $(R_p\, ,\, R_e > 0)$ ,

and

(2) $\overline{P}(t)$ exists for $t \in [0, t_f]$ .

From the calculus of variations, condition (1) is analogous to the "convexity" or "strengthened Legendre-Clebsch" condition, while condition (2) is the "no conjugate point" or "Jacobi" condition [2].

The time-varying state equation for optimal play of the game, determined by substituting (17) into (7), becomes

$$\dot{\underline{x}}(t) = \overline{A}(t)\, \underline{x}(t) \qquad ; \qquad \underline{x}(0) = \underline{x}_o \tag{21}$$

where

$$\overline{A}(t) = \overline{F} - \overline{G}\ \overline{R}^{-1}\overline{G}^T\overline{P}(t) \quad . \tag{22}$$

The value of the game, i.e., the optimal performance, is obtained by substituting (17) into (8) and using the relationships defined above. Performing the indicated steps, the value of the game is

$$J(\underline{u}_p^o,\underline{u}_e^o) = \tfrac{1}{2}\,\underline{x}^T(0)\ \overline{P}(0)\ \underline{x}(0) \quad . \tag{23}$$

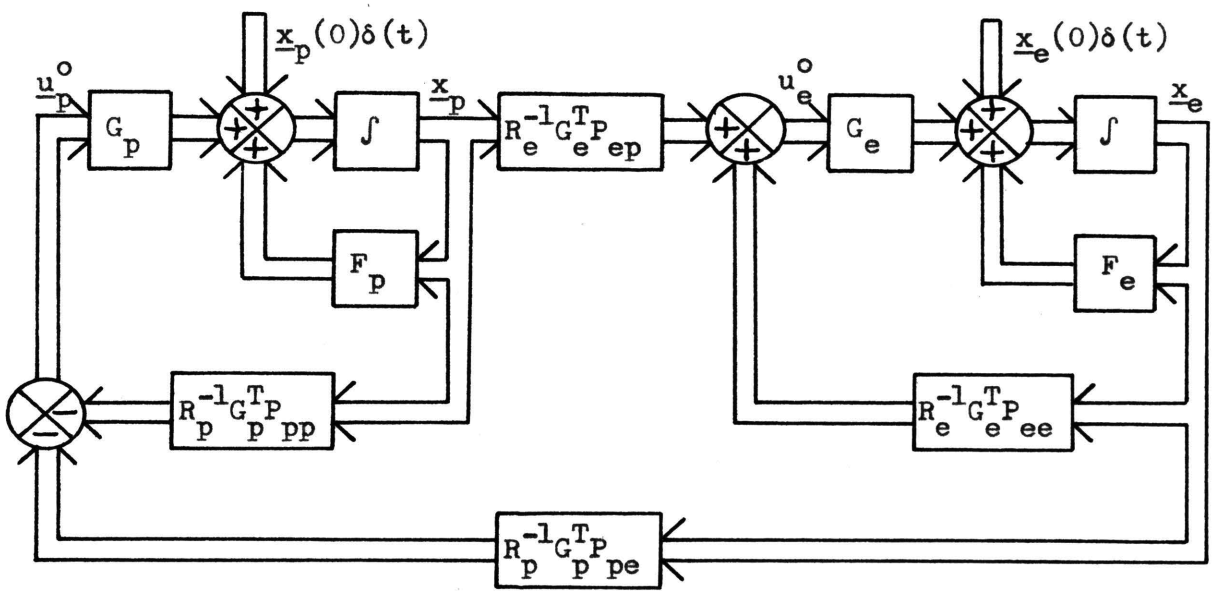
Solution of (21) for $\underline{x}(t)$ provides the possibility of utilizing (19) or (20) as open-loop rather than closed-loop feedback control strategies. As noted by Rhodes [23], the optimal open-loop feedback control strategy is identical to the optimal closed-loop feedback control strategy for the zero sum linear quadratic differential game, provided the opponent is also using an optimal feedback control strategy (either open-loop or closed-loop). However, should one player deviate from the optimal control strategy, the opponent cannot take advantage of this deviation unless he is utilizing a closed-loop feedback control strategy. Consequently, both players are assumed to utilize closed-loop feedback control strategies.

The block diagram for optimal play of the game is shown in Figure 1a using the optimal closed-loop feedback control strategies defined by (19) and (20). Figure 1b shows the closed-loop game, where the system matrices are obtained by partitioning $\overline{A}(t)$ and expanding (21) such that

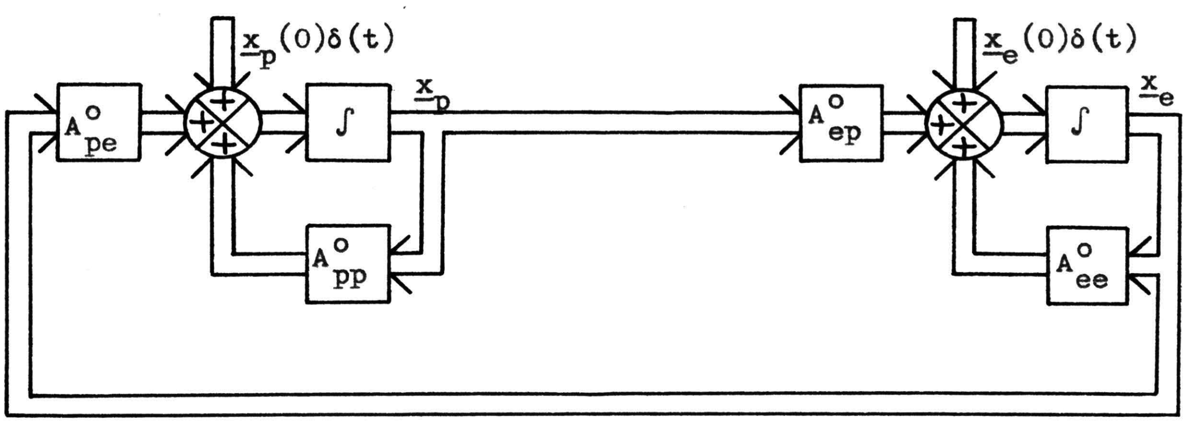
$$\dot{\underline{x}}_p = A^o_{pp}\underline{x}_p + A^o_{pe}\underline{x}_e \tag{24}$$

and

$$\dot{\underline{x}}_e = A^o_{ep}\underline{x}_p + A^o_{ee}\underline{x}_e \quad . \tag{25}$$




(a) Using Open-Loop Parameters and Optimal Closed-Loop Feedback Control Strategies




(b) Using Optimal Closed-Loop Parameters

**Figure 1 Perfect Information Differential Game Block Diagram**

## D. Matrix Riccati Equation Solution

Consider the solution for the matrix Riccati equation (16) suggested by Porter [22] in which $\overline{P}(t)$ is separated into time-invariant and time-varying symmetric matrices where the time-invariant matrix $\overline{P}_o$ satisfies the nonlinear algebraic matrix equation

$$\overline{P}_o\overline{F} + \overline{F}^T\overline{P}_o - \overline{P}_o\overline{G}\,\overline{R}^{-1}\overline{G}^T\overline{P}_o + \overline{Q} = [0] \quad . \tag{26}$$

The resulting solution of the matrix Riccati equation is given by

$$\begin{aligned}\overline{P}(t) = \overline{P}_o + e^{\overline{A}_o^T[t_f-t]}\overline{C}^T[I \\ + \int_t^{t_f} \overline{C}\, e^{\overline{A}_o[t_f-t]}\overline{G}\,\overline{R}^{-1}\overline{G}^T e^{\overline{A}_o^T[t_f-t]}\overline{C}^T\, dt]^{-1}\overline{C}\, e^{\overline{A}_o[t_f-t]}\end{aligned} \tag{27}$$

where

$$\overline{A}_o = \overline{F} - \overline{G}\,\overline{R}^{-1}\overline{G}^T\overline{P}_o \tag{28}$$

and

$$\overline{C}^T\overline{C} = \overline{S} - \overline{P}_o \quad . \tag{29}$$

From the above solution, the existence of $\overline{P}(t)$ for $t \in [0,t_f]$ is equivalent to the condition that

$$I + \int_t^{t_f} \overline{C}\, e^{\overline{A}_o[t_f-t]}\overline{G}\,\overline{R}^{-1}\overline{G}^T e^{\overline{A}_o^T[t_f-t]}\overline{C}^T\, dt > 0 \quad . \tag{30}$$

A sufficient condition for the existence of $\overline{P}(t)$, which results from the above requirement for positive definiteness, is that

$$\int_{t}^{t_f} \overline{C}\, e^{\overline{A}_o[t_f-t]}\, \overline{G}\, \overline{R}^{-1} \overline{G}^T e^{\overline{A}_o^T[t_f-t]}\, \overline{C}^T \, dt \geq 0 \quad . \tag{31}$$

As discussed by both Ho [2], [6] and Rhodes [23], [24], [25], this integral condition is associated with the relative controllability of the Pursuer and the Evader, the Pursuer being required to be more controllable than the Evader.

The problem of identification of an opponent's unknown system parameters during the early stages of the game is discussed later in Chapter IV. As shown there, convergence of the selected identification technique can be guaranteed if the opponent's closed-loop system is time-invariant during the identification period. Consequently, it is instructive at this point to examine the conditions for which the state equation for optimal play of the game becomes time-invariant.

Reference to (22) shows that the state equation becomes time-invariant if, and only if, $\overline{P}(t)$ becomes time-invariant. Since the time-invariant portion of $\overline{P}(t)$ is $\overline{P}_o$, the state equation for optimal play during this period is given by

$$\dot{\underline{x}}(t) = \overline{A}_o \underline{x}(t) \quad ; \quad \underline{x}(0) = \underline{x}_o \quad . \tag{32}$$

Examination of (27) shows that there are two cases for which $\overline{P}(t)$ becomes time-invariant. Consider first the case in which $\overline{C} = [0]$ so that $\overline{C}^T\overline{C} = \overline{S} - \overline{P}_o = [0]$. Although $\overline{P}(t)$ is time invariant for the whole interval $[0,t_f]$, the requirement that $\overline{S} = \overline{P}_o$. is extremely restrictive. Consider, therefore, the second case in which it is assumed that the remaining playing time of the game is long, i.e., $[t_f-t] \rightarrow \infty$ . Examination of (27) shows that

$$\lim_{[t_f - t] \to \infty} \bar{P}(t) = \bar{P}_o \tag{33}$$

if $\bar{A}_o$ is stable. Consequently, rather than place restrictions on $\bar{S}$, only differential games for which $\bar{A}_o$ is stable will be considered throughout the remainder of this dissertation.

The assumption that $\bar{A}_o$ is stable allows an analytic solution to the Riccati equation [1], [19], [30], avoiding the integration shown in (27). Since this analytic solution was used for simulation of the differential game, it is described in Appendix B.

E. Simulation

The solution to the state equation for optimal play of the game (21) is easily written symbolically [20] as

$$\underline{x}(t) = \Phi(t)\, \underline{x}(0) \tag{34}$$

where the fundamental or state transition matrix $\Phi(t)$ satisfies

$$\dot{\Phi}(t) = \bar{A}(t)\Phi(t) \quad ; \quad \Phi(0) = I \quad . \tag{35}$$

Although (35) is easily solved for $\Phi(t)$ when $\bar{A}(t)$ is time-invariant, the solution for $\Phi(t)$ when $\bar{A}(t)$ is time-varying can be found analytically in only a few special cases. For the general time-varying case, a computer solution for the state transition matrix used in (34), or alternately, directly for the game (21) must be obtained.

Examination of (27) shows that even though only differential games having stable $\bar{A}_o$ matrices are being considered, $\bar{P}(t)$, and consequently $\bar{A}(t)$, are time-varying as $t \to t_f$ except for the special case in which $\bar{S} = \bar{P}_o$. Since $\bar{S}$ has not been restricted, $\bar{A}(t)$ is generally time-varying, thereby requiring a computed solution for the play of the game.

Digital simulation of the game was chosen because of the complexity of the required calculations. In order to maintain accuracy, a fourth order Runge-Kutta integration routine [9] with fixed step size was utilized. Note that the relative accuracy of the digital simulation may be determined by comparing the value of the game obtained analytically from (23) with that obtained by performing the integration indicated by (3) or (8) during the simulation. Consequently, the adequacy of the fixed step size is easily verified.

The general characteristics of the differential game are illustrated by the following example game involving players with first order dynamic systems. This game also provides the reference for comparison of the results obtained in the following chapters.

Consider the example differential game defined by the decoupled linear time-invariant system:

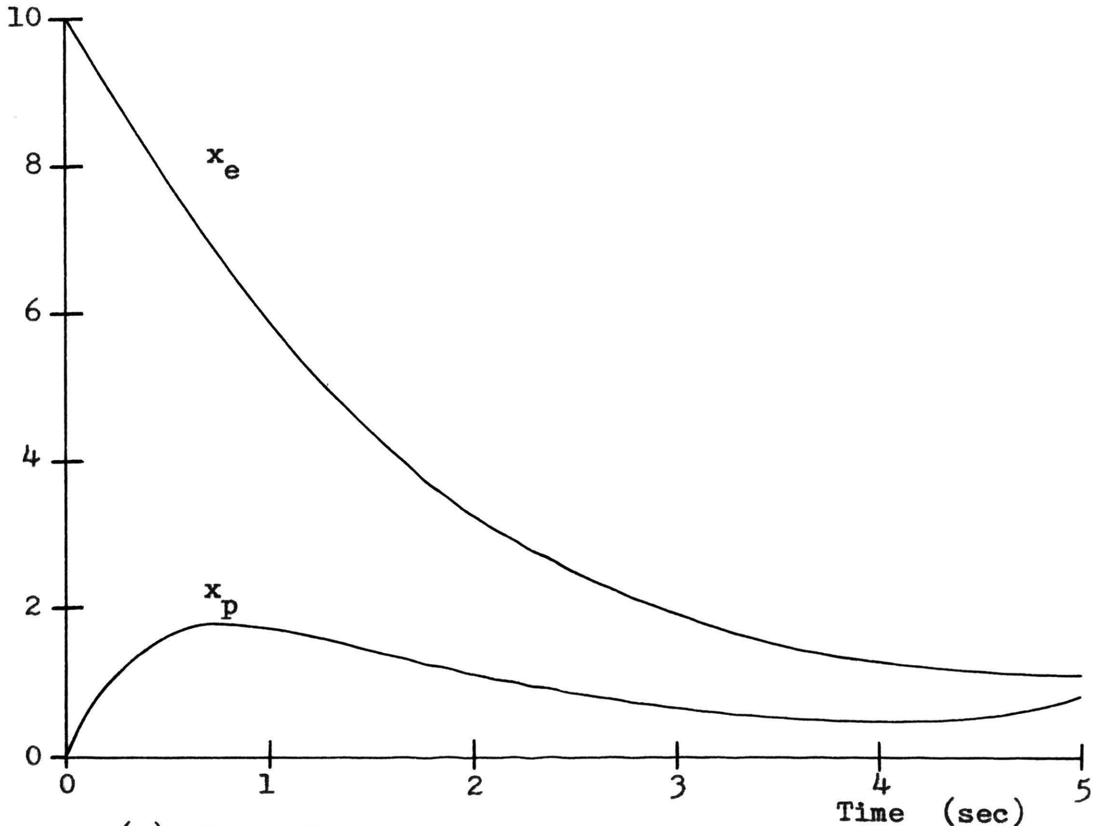
$$\dot{x}_p = -2x_p + u_p \quad ; \quad x_p(0) = 0 \tag{36}$$

$$\dot{x}_e = -x_e + u_e \quad ; \quad x_e(0) = 10 \tag{37}$$

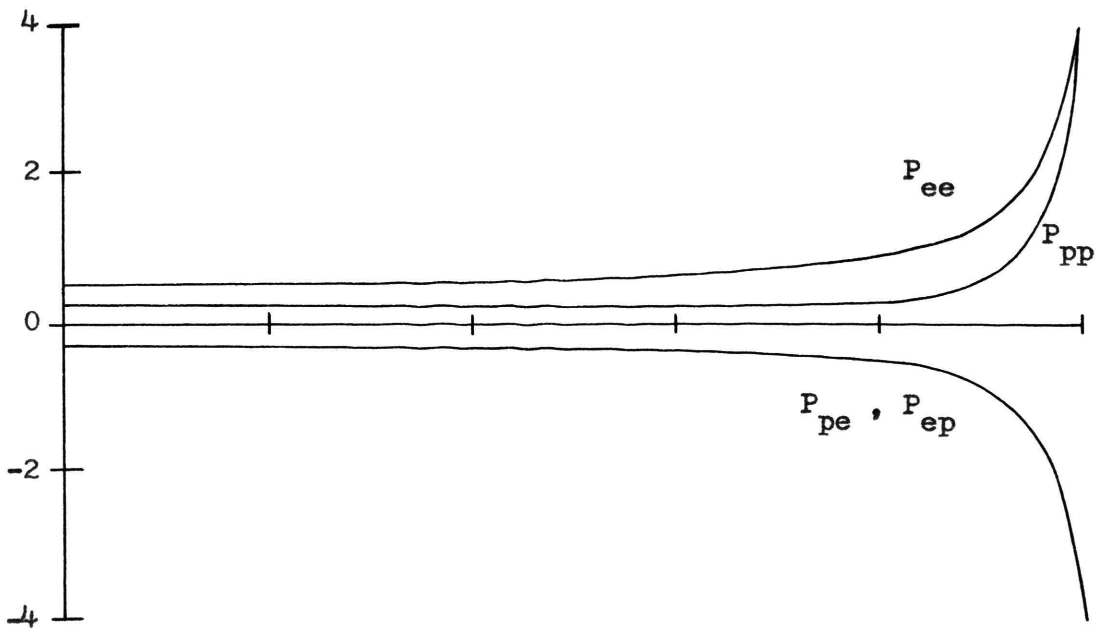
and the quadratic time-invariant performance index

$$J(u_p, u_e) = \tfrac{1}{2}\int_0^5 \left[ (x_e - x_p)^2 + \tfrac{1}{2}u_p^2 - u_e^2 \right] dt + 2\left[x_e(5) - x_p(5)\right]^2 \tag{38}$$

The time histories of the states $x_p$ and $x_e$ obtained by digital simulation using a fixed 0.1 second integration step size are shown in Figure 2a. Note that the Pursuer has made significant progress toward




(a) State Trajectories




(b) Riccati Equation Solution

Figure 2. Example Perfect Information Game

capture of the Evader during the 5 seconds play of the game, the resulting terminal miss being only 0.3382 compared with the initial displacement of 10.

The value of the game obtained analytically from (23) is 26.3672, while that obtained at the end of the digital simulation is 26.3673, which is in error by less than 0.0004%, indicating that the digital simulation for this example game is sufficiently accurate.

The solution to the nonlinear matrix Riccati equation (16) for the example game is shown in Figure 2b. Note that during the early stages of the game this solution is essentially time-invariant and closely approximates the solution to the algebraic Riccati equation (26). As a result, the optimal feedback control gains given in (19) and (20) as well as the closed-loop system matrix given by (22) are also time-invariant during the early stages of the example game.

# III. INCOMPLETE SYSTEM PARAMETER INFORMATION GAMES

## A. Description of the Game

The solution of the differential game discussed in Chapter II assumed that complete information about all aspects of the game was available to both players, which may not always be true. Consider, for instance, the case in which one of the players does not have perfect a priori knowledge of his opponent's system parameters. In this case, optimal play of the game by the player with incomplete information is not possible.

If, of course, this player has a statistical knowledge of his opponent's parameters, he may play a stochastic optimal control strategy designed to optimize the expected value of the performance index given by (3) and (8). He could also simply choose to use the mean value of each parameter and compute in the manner illustrated in Chapter II a deterministic, but perhaps suboptimal, control strategy.

In this paper, it will be assumed that no statistical information about an opponent's parameters is available. In this case, the ignorant player must use some other criterion for selection of control strategies.

## B. Selection of Strategies

The sufficient condition for a solution of the differential game as given by (31) can also be used to determine the upper or lower limits appropriate to an opponent's system parameters. Interpretation of (31) is difficult, however, when the parameter limits of both $F_p$ and $G_p$ or $F_e$ and $G_e$, whichever is the case, are unknown. Consider, therefore, $F_p$ or $F_e$ limits such that $F_p = F_e$ . Since the Pursuer's system would normally be designed to respond more quickly than the Evader, the above limits are realistic.

Using these limits the dimension of the differential game can be reduced, allowing easier interpretation of the limits implied by (31). In order to proceed, define

$$\underline{z} = \underline{x}_p - \underline{x}_e \tag{39}$$

and assume that

$$F = F_p = F_e \tag{40}$$

so that the differential game is defined by the linear time-invariant system

$$\dot{\underline{z}} = F\,\underline{z} + G_p\underline{u}_p - G_e\underline{u}_e \quad ; \quad \underline{z}(0) = \underline{x}_p(0) - \underline{x}_e(0) \tag{41}$$

and the quadratic time-invariant performance index

$$J(\underline{u}_p,\underline{u}_e) = \tfrac{1}{2}\int_0^{t_f} (\underline{z}^T Q\,\underline{z} + \underline{u}_p^T R_p \underline{u}_p - \underline{u}_e^T R_e \underline{u}_e)\,dt + \tfrac{1}{2}\,\underline{z}^T(t_f)\,S\,\underline{z}(t_f) \quad . \tag{42}$$

Proceeding in a manner completely analogous to that of Chapter II., it is easily shown the optimal control strategies for this reduced dimension game are given by

$$\underline{u}_p^*(t) = -\,R_p^{-1}G_p^T P(t)\,\underline{z}(t) \tag{43}$$

and

$$\underline{u}_e^*(t) = -\,R_e^{-1}G_e^T P(t)\,\underline{z}(t) \tag{44}$$

where P is the solution to the matrix Riccati equation

$$\dot{P} = -\,P\,F - F^T P + P\,H\,P - Q \quad ; \quad P(t_f) = S \tag{45}$$

and H is defined as

$$H = G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T \quad . \tag{46}$$

The solution for the matrix Riccati equation (45) is easily shown to be

$$P(t) = P_o + e^{A_o^T[t_f - t]} C^T \Big[ I + \int_t^{t_f} C\, e^{A_o[t_f-t]} H\, e^{A_o^T[t_f-t]} C^T \, dt \Big]^{-1} C\, e^{A_o[t_f-t]} \tag{47}$$

where $P_o$ satisfies

$$P_o F + F^T P_o - P_o H P_o + Q = [0] \tag{48}$$

and

$$A_o = F - H P_o \tag{49}$$

and

$$C^T C = S - P_o \quad . \tag{50}$$

The resulting sufficient condition for the existence of P(t) is given by

$$\int_t^{t_f} C\, e^{A_o[t_f-t]} [G_p R_p^{-1} G_p^T - G_e R_e^{-1} G_e^T]\, e^{A_o^T[t_f-t]} C^T dt \geq 0 \quad . \tag{51}$$

Thus, the limit on $G_p$ or $G_e$, whichever is the case, is given by

$$G_p R_p^{-1} G_p^T = G_e R_e^{-1} G_e^T \tag{52}$$

which causes the equality of (51) to be satisfied. Defining

$$R_p = R_p^{\frac{1}{2}} R_p^{\frac{1}{2}} \tag{53}$$

and

$$R_e = R_e^{\frac{1}{2}} R_e^{\frac{1}{2}} \quad , \tag{54}$$

the limits implied by (52) can be restated as

$$G_pR_p^{-\frac{1}{2}} = G_eR_e^{-\frac{1}{2}} \quad . \tag{55}$$

Consequently, if the Pursuer is the ignorant player, his pessimistic limiting estimate of the Evader's parameters would be

$$\hat{F}_e = F_p \tag{56}$$

and

$$\hat{G}_e = G_pR_p^{-\frac{1}{2}}R_e^{\frac{1}{2}} \quad . \tag{57}$$

Likewise, if the Evader is the ignorant player, his optimistic limiting estimate of the Pursuer's parameters would be

$$\hat{F}_p = F_e \tag{58}$$

and

$$\hat{G}_p = G_eR_e^{-\frac{1}{2}}R_p^{\frac{1}{2}} \quad . \tag{59}$$

Lacking better information on which to base a control strategy, the ignorant player is assumed to utilize the appropriate above limiting estimate in computing the feedback control strategy given by (43) or (44). Notice that use of the limiting estimate simplifies (45) since H = 0 so that its solution is now given by

$$P(t) = P_o + e^{F^T[t_f-t]}[S - P_o]\, e^{F[t_f-t]} \tag{60}$$

where $P_o$ satisfies

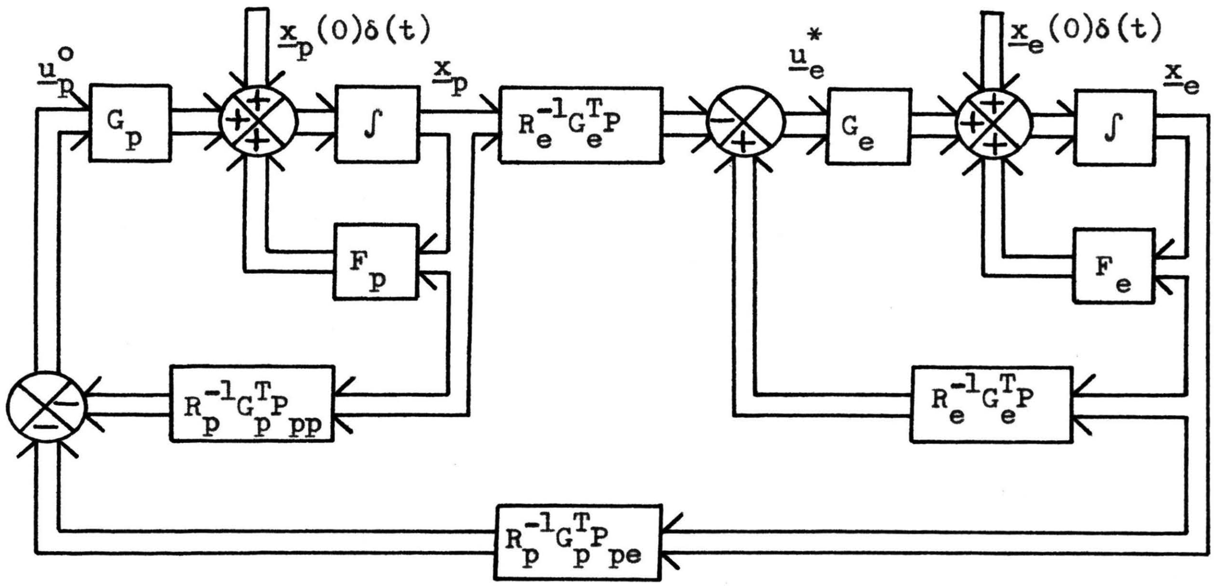
$$P_oF + F^TP_o + Q = [0] \quad . \tag{61}$$

Thus far, only the effect of incomplete system parameter information on the selection of control strategies by the ignorant player has been discussed. Consider, therefore, the selection of control strategies by the opposing player who is assumed to be smart, i.e., assumed to have complete a priori knowledge of his opponent's system parameters. If the smart player is also aware of both the ignorant player's predicament and selected control strategy, then the optimal control strategy is given by the solution to the resulting one-sided optimal control problem. This, of course, presupposes a great deal of additional knowledge. Thus, throughout the remainder of this dissertation the smart player assumes that his opponent is also smart, which results in the use of the optimal control strategy previously derived in Chapter II. The smart player does, however, use this optimal strategy in a closed-loop feedback control manner to take advantage of an opponent's ignorance.

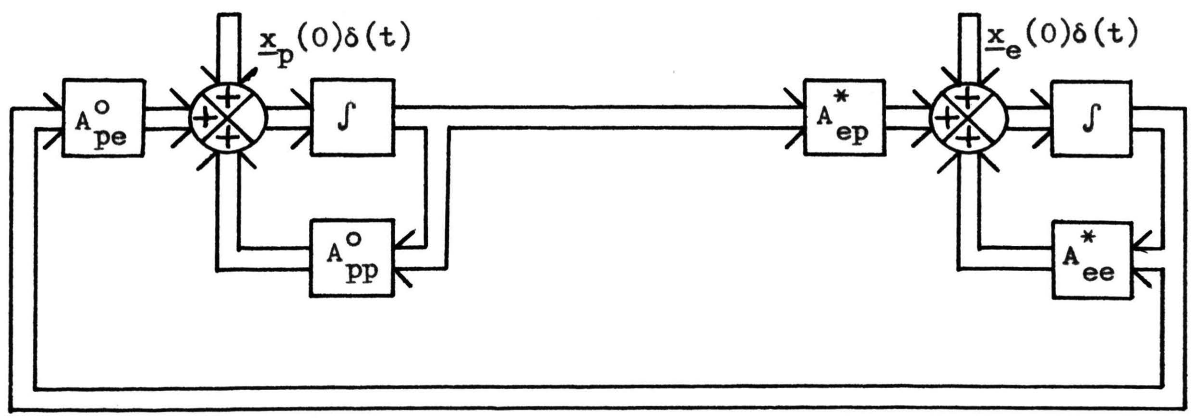
The block diagram for the incomplete information game, assuming the Evader to be the ignorant player, is shown in Figure 3.

## C. Simulation

The digital simulation described in Chapter II was modified in order to allow simulation of incomplete system parameter information games. Using the example game defined by (36), (37), and (38), the incomplete information game was simulated assuming first the Evader and then the Pursuer to be the ignorant player. The resulting performance indices and terminal miss distances for these incomplete information games as well as for the previous perfect information game are given in Table I. As shown, lack of complete system parameter information is costly. The trajectories for these incomplete information games are given for comparison to those of the perfect information game in Figures 4a and 5a, respectively. The comparisons between matrix Riccati equation solutions are given in Figures 4b and 5b.




(a) Using Open-Loop Parameters



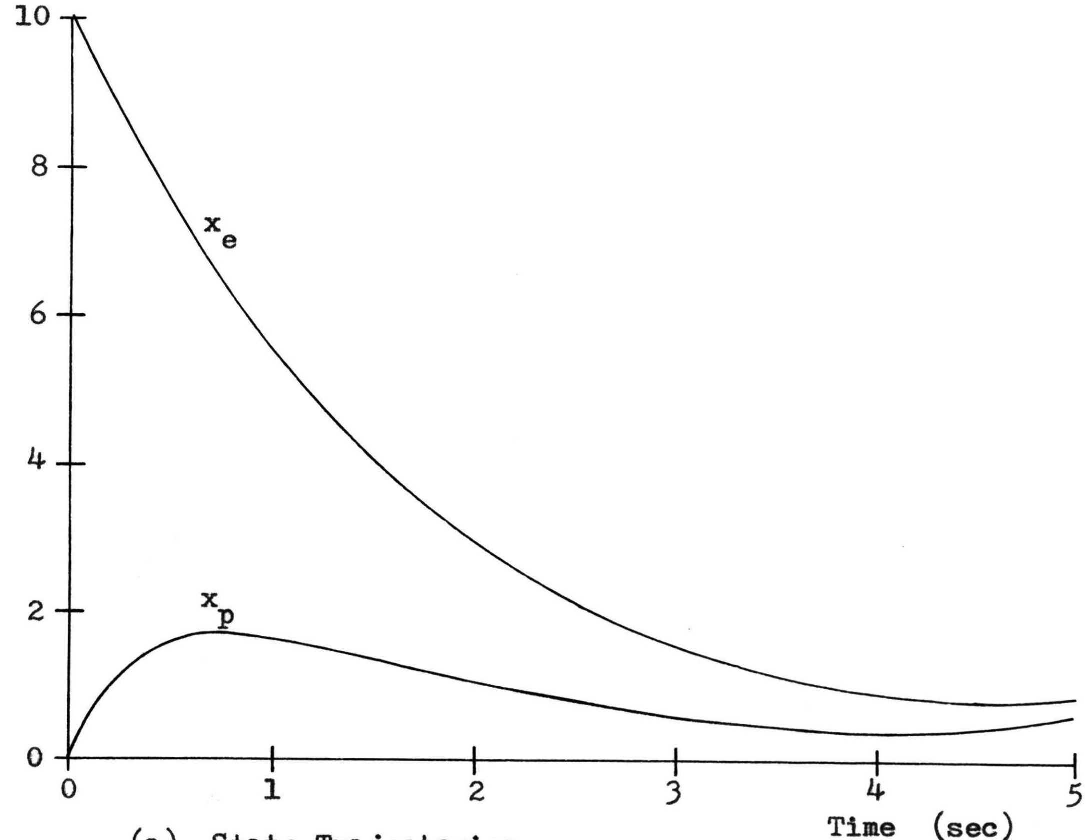

(b) Using Closed-Loop Parameters

**Figure 3** Incomplete System Parameter Information Game Block Diagram - Evader Ignorant

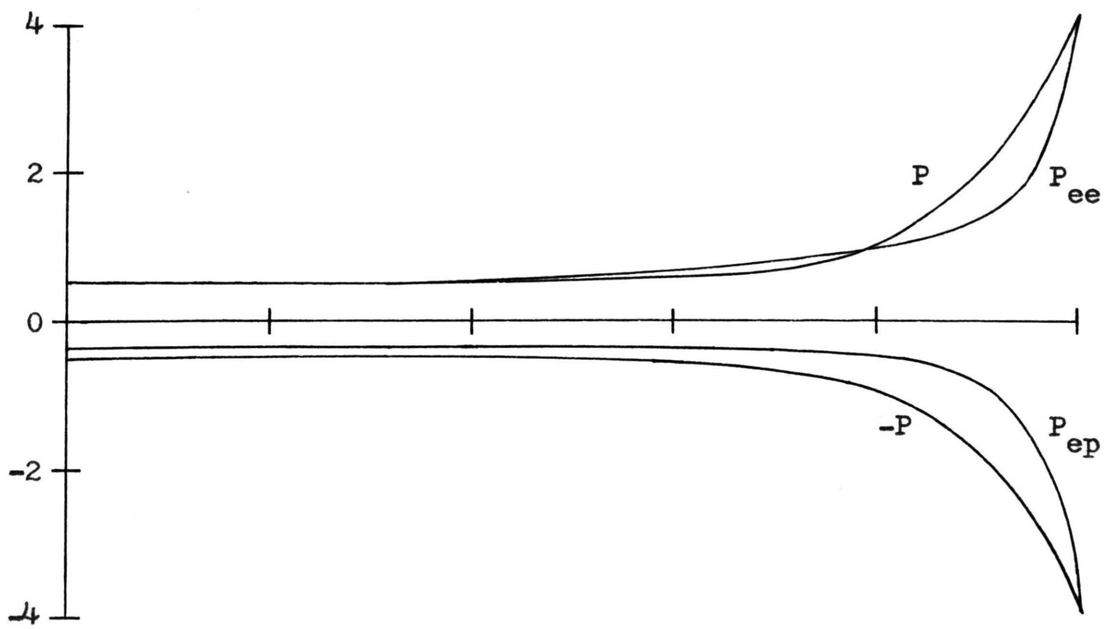
Table I.

Incomplete Information Game Performance Comparison

| Ignorant Player | Evader | Neither | Pursuer |
|---|---|---|---|
| Performance Index (J) | 26.1389 | 26.3672 | 27.1900 |
| Terminal Miss $[x_e(5)-x_p(5)]$ | 0.2321 | 0.3382 | 0.4569 |

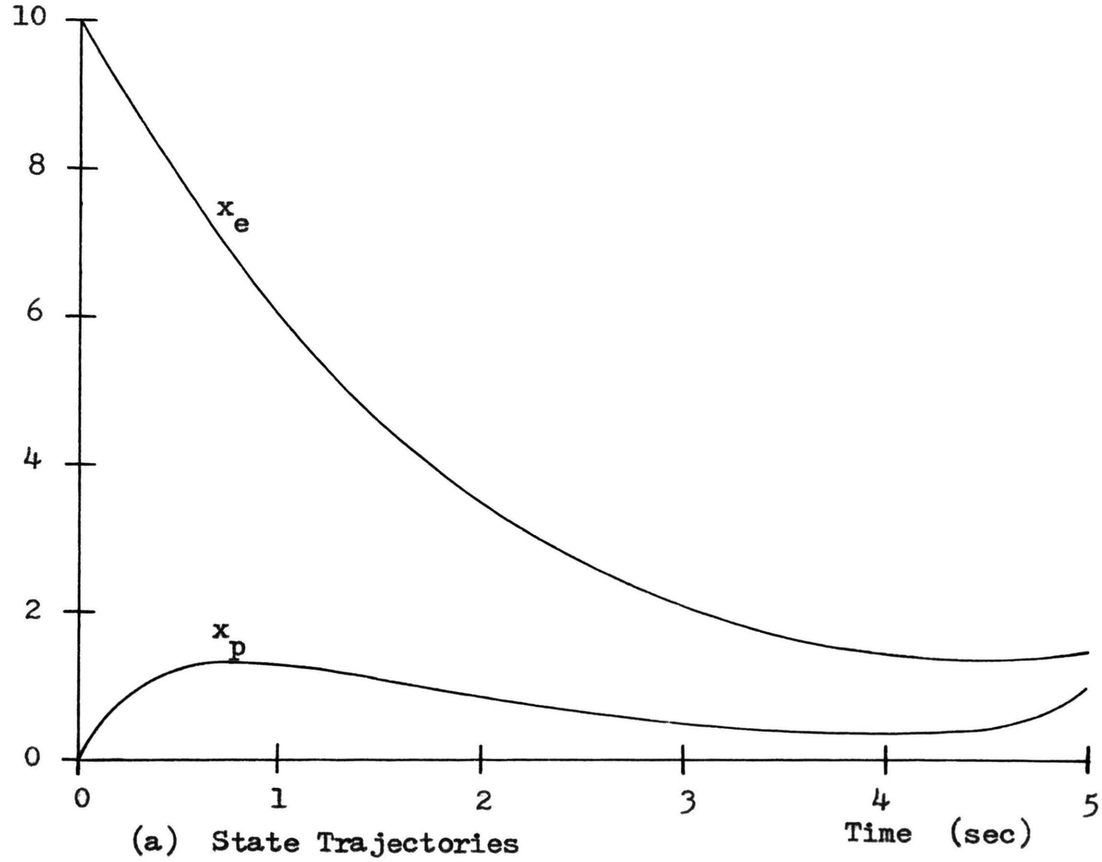



(a) State Trajectories

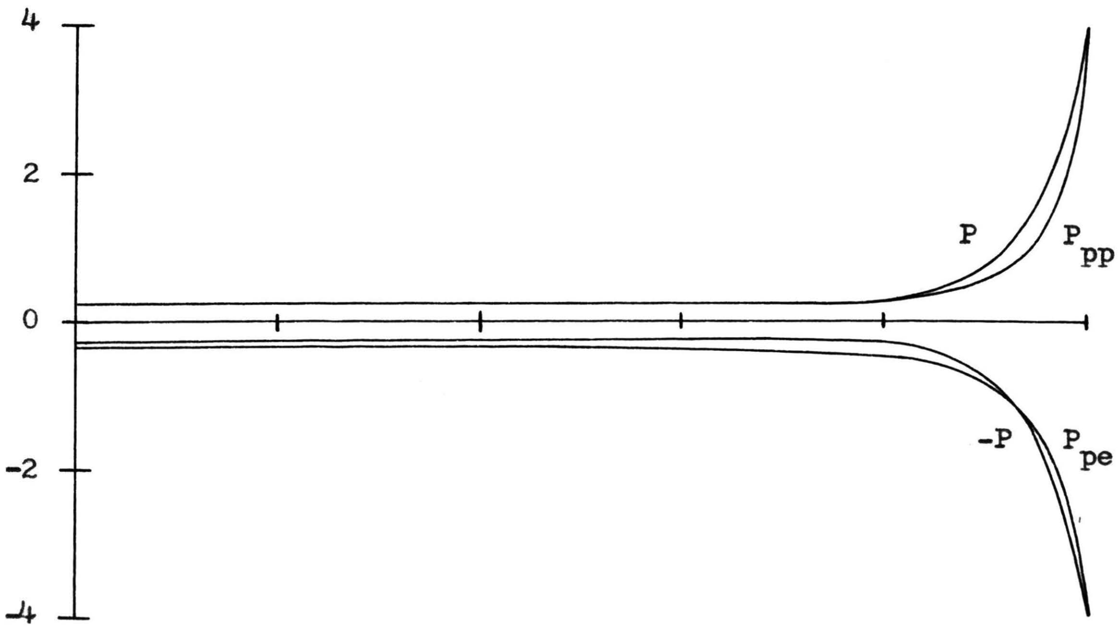



(b) Riccati Equation Solution Comparison

Figure 4. Example Incomplete Information Game - Evader Ignorant




(a) State Trajectories




(b) Riccati Equation Solution Comparison

Figure 5. Example Incomplete Information Game - Pursuer Ignorant

# IV. PARAMETER IDENTIFICATION

## A. Differential Game Observables

In the discussion of incomplete information games in Chapter III, it was implicitly assumed that the ignorant player's knowledge of the opponent's system parameters remained constant throughout the play of the game, i.e., the a posteriori knowledge of these parameters was no better than the a priori knowledge. Although the ignorant player's a priori knowledge may be incomplete, he is not, as may have been previously implied, incapable of learning. Consider, therefore, the possibility that the ignorant player may be able to identify the smart player's system parameters during the course of the game.

While many system identification techniques can be formulated, all are based upon the system's outputs and the associated control inputs. As illustrated in Figure 3a, each player's outputs, which in this case are the states, must be accessible in order for the players to apply any form of closed-loop feedback control strategy. However, even though the players respond to each other's motion by virtue of their closed-loop feedback control strategies, a player's control inputs would not normally be accessible to the opposing player. Consequently, even though it may be validly assumed that the smart player utilizes the optimal closed-loop feedback control strategy, the ignorant player cannot directly identify the smart player's open-loop parameters.

Since the opponent's optimal control inputs are not available, the identification technique must be based upon the player's accessible states. As illustrated in Figure 3b, the player's states provide the correct relationship for direct identification of the smart player's closed-loop system parameters.

B. Closed-Loop System Parameter Identification

In order to simplify discussion throughout the remainder of this dissertation, the Evader is assumed to be the ignorant player desiring to identify the Pursuer's closed-loop system parameters. The opposite case is easily obtained by interchanging the player's names and subscripts wherever they appear and changing the sign of all terms previously containing $R_p$ or $R_e$.

The Pursuer's optimal closed-loop system as given by (24) is approximately time-invariant if (1) $\overline{A}_o$ is a stable matrix and (2) the remaining playing time, i.e., $[t_f - t]$ is long. In this case, (24) may be written as

$$\dot{\underline{x}}_p = A^o_{ppo}\underline{x}_p + A^o_{peo}\underline{x}_e \tag{62}$$

since A(t) given by (22) becomes time-invariant, i.e., $\overline{A}(t) \rightarrow \overline{A}_o$.

In order to estimate $A^o_{ppo}$ and $A^o_{peo}$, the closed-loop system must be observable, controllable, as well as identifiable. The observability condition is satisfied by the prior assumption that the player's states are accessible. The controllability condition is dependent on the Pursuer's utilization of the optimal closed-loop feedback control strategy for the complete information game. The identifiability condition, however, must be examined.

## 1. Identifiability Condition

The i'th derivative of (62) is given by

$$\overset{(i)}{\underline{x}}_p = A^o_{ppo} \overset{(i-1)}{\underline{x}}_p + A^o_{peo} \overset{(i-1)}{\underline{x}}_e \quad . \tag{63}$$

Adjoining 2n successive derivatives like (63), one obtains

$$\left[ \overset{(2n)}{\underline{x}}_p \ \cdots \ \ddot{\underline{x}}_p \ \dot{\underline{x}}_p \right] = \left[ A^o_{ppo} \quad A^o_{peo} \right] \left[ \overset{(2n-1)}{\underline{x}} \ \cdots \ \dot{\underline{x}} \ \ \underline{x} \right] \tag{64}$$

where $\underline{x}$ is the augmented state vector defined by (5). Therefore, an equation for estimation of $A^o_{ppo}$ and $A^o_{peo}$ is given by

$$\left[ \hat{A}^o_{ppo} \quad \hat{A}^o_{peo} \right] = \left[ \overset{(2n)}{\underline{x}}_p \ \cdots \ \ddot{\underline{x}}_p \ \dot{\underline{x}}_p \right] \left[ \overset{(2n-1)}{\underline{x}} \ \cdots \ \dot{\underline{x}} \ \ \underline{x} \right]^{-1} \tag{65}$$

where the circumflex has been used to distinguish between estimated and actual parameters.

From (65) one sees that the Pursuer's closed-loop parameters are identifiable if $\left[ \overset{(2n-1)}{\underline{x}} \ \cdots \ \dot{\underline{x}} \ \ \underline{x} \right]$ is nonsingular, i.e.,

$$\det \left[ \overset{(2n-1)}{\underline{x}} \ \cdots \ \dot{\underline{x}} \ \ \underline{x} \right] \neq 0 \quad . \tag{66}$$

Condition (66) may be recognized [18] as the Wronskian of $\underline{x}$, i.e., $W(\underline{x})$. Therefore, the Pursuer's closed-loop system parameters are identifiable during any time interval for which all of the states of the augmented state vector $\underline{x}$ are linearly independent. The identification time interval required is simply that needed in order to accurately construct the 2n derivatives. Theoretically, of course, this identification time interval can be made arbitrarily short for noise free measurements.

The condition for linear independence of the states of $\underline{x}$ may be examined by considering the early stages of the incomplete information differential game shown in Figure 3, whose state equation may be written as

$$\dot{\underline{x}} = \begin{bmatrix} A^o_{ppo} & A^o_{peo} \\ A^*_{epo} & A^*_{eeo} \end{bmatrix} \underline{x} \quad ; \quad \underline{x}(0) = \underline{x}_o \quad . \tag{67}$$

It can be shown [20] that the states of $\underline{x}$ are, at least initially, linearly independent if (1) the eigenvectors of (67) are distinct and (2) the initial condition, i.e., $\underline{x}_o$, has a non-zero projection on all of the eigenvectors. Thus, an alternate way of stating the condition for identifiability [16] is that the initial conditions must excite all modes of the system. The identification must, of course, be completed before any of these modal responses dies out. In fact, as any of the modal responses approaches zero, the matrix whose inverse is required in (65) becomes almost singular, thereby causing computational difficulties.

Notice also that if the condition for identifiability holds, then all of the parameters of (67) may be identified. However, $A^*_{peo}$ and $A^*_{eeo}$ are already known by the Evader and, therefore, need not be identified.

If the condition for identifiability is not satisfied, the Evader can alter the play of the game in order to make the states linearly independent by modifying his control strategy such that

$$\underline{u}_e = \underline{u}^*_e + \underline{u}^t_e \tag{68}$$

where $\underline{u}^t_e$ is a special test input. As a result of this modification, the vector equation for the imperfect information game becomes

$$\begin{bmatrix} \dot{\underline{x}}_p \\ \dot{\underline{x}}_e \end{bmatrix} = \begin{bmatrix} A^o_{pp} & A^o_{pe} \\ A^*_{ep} & A^*_{ee} \end{bmatrix} \begin{bmatrix} \underline{x}_p \\ \underline{x}_e \end{bmatrix} + \begin{bmatrix} 0 \\ G_e \end{bmatrix} \underline{u}^t_e \quad . \tag{69}$$

Assuming the differential game described by (69) is completely state controllable, introduction of the test signal $\underline{u}_e^t$ by the Evader alters the states of both players as well as the Pursuer's control strategy, as shown by (19). The performance index (3) for the game is also affected since it is a function of the states and control strategies of both players. Since the cost to the Evader is, in general, adversely affected by the introduction of such a test signal, its use should be avoided if possible.

The parameter identification technique given by (65) requires 2n derivatives of $\underline{x}_p$ and 2n-1 derivatives of $\underline{x}_e$. Although the differentiation involved can be carried out, measurement noise inherent in all physical systems, even though relatively minor, can cause large errors in computing high order derivatives and, consequently, in the parameter estimates obtained from (65). Therefore, even though (65) allows the sufficient conditions for identification to be stated, its actual use in any physical system is usually impractical.

2. Continuous Time Learning Models

Since the learning model technique avoids both the differentiation and the matrix inversion required by (65), it normally provides a more practical method for parameter identification. Therefore, the various learning models described in the literature were surveyed.

Two learning models, which are distinguishable by the error measure employed, appear applicable to the present identification problem. The first, based on response error, is the multivariable model originally developed by Pazdera and Pottinger [21]. In the context of the incomplete information game under consideration, its response error is given by the solution of

$$\dot{\underline{e}} = D\,\underline{e} + \Delta A^{o}_{ppo}\underline{x}_p + \Delta A^{o}_{peo}\underline{x}_e \tag{70}$$

where D is a stable matrix and the parameter misalignments are defined by

$$\Delta A^{o}_{ppo} = A^{o}_{ppo} - \hat{A}^{o}_{ppo} \tag{71}$$

and

$$\Delta A^{o}_{peo} = A^{o}_{peo} - \hat{A}^{o}_{peo} \quad . \tag{72}$$

The development of the response error learning model, including the Spence [27], [28] modification to provide improved response characteristics, is outlined in Appendix C.

The second learning model, based on a generalized equation error, is the multivariable extension of the single input/output model originally developed by Lion [17]. In the present context, its equation error is given by

$$\underline{e} = \Delta A^{o}_{ppo}\,\tilde{\underline{x}}_p + \Delta A^{o}_{peo}\tilde{\underline{x}}_e \tag{73}$$

where the tilde is used to indicate filtered state variables. An original development of the multivariable equation error learning model is included in Appendix D.

The two continuous learning models described above are actually quite similar in that both utilize the same form of parameter adjustment and also error measures which are dependent on the parameter misalignments. The major differences in these models can be seen by considering their error measures. As shown the equation error (73) is a linear function of the parameter misalignment while the response error (70) has only a functional dependence on the parameter misalignment. On the other hand, the response error is based directly on the system input and output vectors while filtered state vectors are required for the equation error learning model. Both the functional dependence shown by (70) and the filtering requirements shown by (73) result from having the learning models avoid the use of the generally inaccessible state vector derivative $\dot{\underline{x}}_p$.

Parameter identification for either learning model technique is complete at time $t_{id}$ if, and only if, the parameter misalignments are null for all $t \geq t_{id}$. From consideration of the parameter adjustment laws used by these techniques, it is seen that the error vector must also be null for $t \geq t_{id}$. It is easily shown (See Appendices C and D) that both techniques are asymptotically stable in the sense that parameters will continue to be adjusted until the error vector is null. From the error equations for the two methods it is obvious that even though zero parameter misalignments implies zero error, zero error implies only that

$$[\Delta A^o_{ppo} \quad \Delta A^o_{peo}] \begin{bmatrix} \underline{x}_p \\ \underline{x}_e \end{bmatrix} = \underline{0} \tag{74}$$

for the response error learning model and

$$[\Delta A^o_{ppo} \quad \Delta A^o_{peo}] \begin{bmatrix} \tilde{\underline{x}}_p \\ \tilde{\underline{x}}_e \end{bmatrix} = \underline{0} \tag{75}$$

for the equation error learning model. It follows that a sufficient condition for identifiability for either learning model is that the states of the augmented state matrix must be linearly independent during the identification interval $[0, t_{id}]$.

By comparison one sees that the condition for identifiability for the learning technique is essentially the same as that for the computational technique given by (65). The only difference is the length of the time interval required. While the interval for the computational technique (65) is, at least theoretically, infinitesimally short, the interval for the learning model is governed by the gain setting of the parameter adjustment laws and the dynamic response of the system itself.

The length of time required for identification is not usually an overiding consideration since in many cases a test input may be utilized in order to generate the desired linear independence. However, in the present application, the identification time is extremely important since the game is of finite duration. The identification time is made critical by the fact that for the incomplete information games being considered the states are linearly independent during the early stages only for as long as the modal responses last since the use of a test input is undesirable. The short interval of linear independence results from the fact that the closed-loop response during the early stages is excited only by the initial conditions of the game and $\bar{A}_o$ has been assumed to be a stable matrix.

In order to evaluate the effect of this identification time constraint and to determine which of the identification techniques exhibits better convergence characteristics both learning models were simulated using a digital computer. The example incomplete information game with the Evader presumed to be the ignorant player was utilized. For this example, $\underline{x}_p$, $\underline{x}_e$, $\underline{e}$, $A^o_{ppo}$, and $A^o_{peo}$ are scalars.

Although the fourth order Runge-Kutta integration routine was again used, it was discovered that an extremely short integration step size was required in order for digital simulation of the learning model techniques to be stable. The required step size was found to be inversely dependent on the parameter adjustment gain being utilized. The integration step size was determined for each value of gain by initializing the parameters at their correct values and requiring the simulation to run for a suitable period of time without diverging.

It was also discovered that if only one of the two pursuer parameters was considered to be unknown, this parameter could be rapidly identified with the identification time becoming shorter as the adjusted gain was increased. However, when both parameters were presumed unknown, the linear independence time interval was not sufficient to allow the identification to converge. Furthermore, increasing the adjustment gain appeared to worsen the response, possibly because of associated decrease in integration accuracy. The results were equally discouraging for both learning models.

Lion [17] discussed the above effect resulting from increased gain for the single input/output equation error learning model. He showed that in order to be able to arbitrarily increase the adjustment gain and thereby decrease identification time, the number of equation error states must equal the number of unknown parameters. This implies that for the general multivariable system with n states and $m \leq n$ inputs, $(n + m)$ error vectors of n dimension must be used. This approach is outlined in Appendix E for the equation error learning model. As can be seen, this expansion vastly complicates the mechanization of the learning model identification technique, particularily for the present case where initial conditions must be taken into account.

For the previous example, there are only two unknown parameters. Consequently only one additional error equation had to be added to the equation error learning model simulation in order to apply the above technique. Although some improvement was noted, the simulation results were again discouraging. Because integrators were used for the two additional state variable filters required to mechanize the additional equation error, the second error quickly exceeded the first. Therefore, the parameter adjustment rate became primarily dependent on the second error rather than the first. Since the states of the game are linearly independent for only a short period of time, the adjustment gain would have to be extremely high, resulting in extremely rapid response of the learning model in order for the method to be effective. The difference between these results and those obtained by Lion appears to be due primarily to the fact that the natural response of the second order differential game is being utilized instead of the system response to a multi-frequency or a noisy test input. The integration error associated with the digital simulation when the adjustment gains are high may also have some affect on these results.

It should also be noted that the response error and equation error models may be combined to provide an improved learning model identification technique. An original development of this combined response/equation error learning model is outlined in Appendix F. Both the response error model and the combined error model may also be extended in the same fashion as the equation error model in order to increase the dimension of the error vector.

It is obvious from the above discussion that parameter identification via learning models is highly dependent on (1) the response of the unknown system to initial conditions, (2) the characteristics of the input, and (3) the gain setting of the parameter adjustment laws. Even if suitable response is obtained for one set of unknown parameter values, a significant amount of simulation would be required to ensure that the response is adequate for the expected range of the unknown parameters. Consequently, a more sure method of parameter identification than that afforded by the learning model technique is desirable for the incomplete information game.

3. Discrete Time Identification

The learning model parameter identification technique was originally applied in order to avoid the differentiation and matrix inversion required by (65). However, the matrix inversion can easily be handled by digital computation if a discrete time equivalent of (65) is constructed. Of course, computational difficulties can still arise if the elements of the augmented state vector are not clearly linearly independent. Assuming that sufficient samples can be obtained before the modal response to initial conditions dies out, application of discrete time parameter identification to the incomplete information differential game merits investigation.

Assume that the states of the differential game are sampled at times $t_1$ and $t_2$ where $t_2 > t_1$ . Since the Pursuer's closed-loop system (62) is time-invariant during the early stages of the incomplete information game being investigated, the Pursuer's states at time $t_2$ are given by

$$\underline{x}_p(t_2) = e^{A^o_{ppo}\Delta T}\underline{x}_p(t_1) + e^{A^o_{ppo}t_2}\int_{t_1}^{t_2} e^{-A^o_{ppo}t} A^o_{peo}\underline{x}_e(t)\,dt \qquad (76)$$

where the time interval between samples is

$$\Delta T = t_2 - t_1 \quad . \qquad (77)$$

If the time interval $\Delta T$ is sufficiently short, $\underline{x}_e(t)$ may be considered a constant so that (76) may be rewritten as

$$\underline{x}_p(t_2) = e^{A^o_{ppo}\Delta T}\underline{x}_p(t_1) + e^{A^o_{ppo}\Delta T}\int_0^{\Delta T} e^{-A^o_{ppo}\alpha}\,d\alpha\; A^o_{peo}\underline{x}_e(t_1) \quad . \qquad (78)$$

Since the series expansion for $\exp[-A^o_{ppo}\alpha]$ converges uniformly for $-\infty < \alpha < \infty$, the order of integration and summation may be interchanged and the integration indicated in (78) may be carried out term by term. As a result, $\underline{x}_p(t_2)$ may be written symbolically as

$$\underline{x}_p(t_2) = e^{A^o_{ppo}\Delta T}\underline{x}_p(t_1) + [e^{A^o_{ppo}\Delta T} - I]\, A^{o\;-1}_{ppo} A^o_{peo}\underline{x}_e(t_1) \quad , \qquad (79)$$

since $[e^{A^o_{ppo}\Delta T} - I]\, A^{o\;-1}_{ppo}$ exists even if $A^{o\;-1}_{ppo}$ is singular. Assuming that $\Delta T$ is also sufficiently short so that

$$e^{A^o_{ppo}\Delta T} \approx I + A^o_{ppo}\Delta T \quad , \tag{80}$$

equation (79) becomes

$$\underline{x}_p(t_2) = \Phi^o_{ppo}\underline{x}_p(t_1) + \Phi^o_{peo}\underline{x}_e(t_1) \tag{81}$$

where

$$\Phi^o_{ppo} = I + A^o_{ppo}\Delta T \tag{82}$$

and

$$\Phi^o_{peo} = A^o_{peo}\Delta T \quad . \tag{83}$$

From (82) and (83) one sees that if $\Phi^o_{ppo}$ and $\Phi^o_{peo}$ can be identified, then the closed-loop system parameters can also be computed as

$$A^o_{ppo} = \frac{1}{\Delta T}\left[\Phi^o_{ppo} - I\right] \tag{84}$$

and

$$A^o_{peo} = \frac{1}{\Delta T}\Phi^o_{peo} \quad . \tag{85}$$

If the states of the game are sampled at equally spaced intervals, the n'th sample time becomes $n\Delta T$ and (81) may be written as

$$\underline{x}_p(n+1) = \Phi^o_{ppo}\underline{x}_p(n) + \Phi^o_{peo}\underline{x}_e(n) \quad . \tag{86}$$

Adjoining 2n successive samples, one obtains

$$\left[\underline{x}_p(2n) \ldots \underline{x}_p(2)\ \underline{x}_p(1)\right] = \left[\Phi^o_{ppo}\ \ \Phi^o_{peo}\right]\left[\underline{x}(2n-1) \ldots \underline{x}(1)\ \underline{x}(0)\right] \tag{87}$$

where $\underline{x}$ is the augmented state vector defined by (5). Therefore, one equation for estimation of $\Phi^o_{ppo}$ and $\Phi^o_{peo}$ is given by

$$[\hat{\Phi}^{o}_{ppo} \;\; \hat{\Phi}^{o}_{peo}] = [\underline{x}_p(2n)\ldots \underline{x}_p(2)\; \underline{x}_p(1)]\,[\underline{x}(2n-1)\ldots \underline{x}(1)\;\; \underline{x}(0)]^{-1} \; . \tag{88}$$

The above identification equation is the discrete time analog of the continuous time identification equation given by (65). The need to construct 2n derivative vectors has been replaced by the requirement for 2n equally spaced samples. The condition for identifiability using the above discrete time identification equation is essentially the same as that continuous time identification via (65), namely, that the states of the augmented state vector $\underline{x}$ must be linearly independent over the time interval $[0, 2n\Delta T]$, i.e., they must span the 2n dimensional space.

The discrete time identification equation given by (88) is not free of computational difficulties, however, since from (82) $\Phi^{o}_{ppo} \rightarrow I$ as $\Delta T \rightarrow 0$. Therefore, from (84) one sees that a small error in $\Phi^{o}_{ppo}$ can result in a large error in $A^{o}_{ppo}$ if $\Delta T$ is made very short.

Although (88) appears to avoid the noisy derivative process associated with (65), measurement noise does have an effect on the accuracy of the estimates for $\Phi^{o}_{ppo}$ and $\Phi^{o}_{peo}$, which in turn can seriously affect the accuracy of the estimates for $A^{o}_{ppo}$ and $A^{o}_{peo}$ given by (84) and (85). Even so, the estimates obtained using (88) are optimal in the minimum mean square error sense. This is easily shown with the aid of the following definitions:

$$\Phi^{o}_{po} = [\Phi^{o}_{ppo} \;\; \Phi^{o}_{peo}] \tag{89}$$

$$X_p(2n) = [\underline{x}_p(2n)\ldots \underline{x}_p(2)\; \underline{x}_p(1)] \tag{90}$$

$$X(2n-1) = [\underline{x}(2n-1) \ldots \underline{x}(1)\ \underline{x}(0)] \quad . \tag{91}$$

Since the estimate of $X_p(2n)$ is given by

$$\hat{X}_p(2n) = \hat{\Phi}^o_{po} X(2n-1) \quad . \tag{92}$$

the error in the estimate of $\Phi^o_{po}$ is reflected by the error in the estimate of $X_p$ which may be written as

$$\begin{aligned} E_p(2n) &= \hat{X}_p(2n) - X_p(2n) \\ &= \hat{\Phi}^o_{po} X(2n-1) - X_p(2n) \quad . \end{aligned} \tag{93}$$

Defining the cost function as

$$\begin{aligned} J_E &= \tfrac{1}{2}\, tr(E_p E_p^T) \\ &= \tfrac{1}{2}\, tr(\ [\hat{\Phi}^o_{po} X(2n-1) - X_p(2n)]\ [X^T(2n-1)\hat{\Phi}^{o\ T}_{po} - X_p^T(2n)]\ ), \end{aligned} \tag{94}$$

one sees that the minimum mean square error estimate $\hat{\Phi}^o_{po}$ must satisfy

$$\frac{\partial J_E}{\partial \hat{\Phi}^o_{po}} = 0 \quad . \tag{95}$$

Performing the indicated operation, one obtains

$$tr(\ [\hat{\Phi}^o_{po} X(2n-1) - X_p(2n)]\ X^T(2n-1)\ ) = 0 \quad . \tag{96}$$

Note that (88), which in terms of the previous definitions may be written as

$$\hat{\Phi}^o_{po} = X_p(2n)\ X^{-1}(2n-1) \quad , \tag{97}$$

satisfies (96) and is, therefore, the minimum mean square error estimate for $\Phi^o_{po}$.

From (96) one sees that the minimum mean square error estimate for $\Phi^{o}_{po}$ may also be written as

$$\hat{\Phi}^{o}_{po} = [X_p(2n)\ X^T(2n-1)]\ M(2n-1) \tag{98}$$

where

$$M(2n-1) = [X(2n-1)\ X^T(2n-1)]^{-1} \quad . \tag{99}$$

Although this parameter identification equation requires two more matrix multiplications, one also notes that it is no longer necessary that X be a square matrix. Therefore, (98) offers a method by which the mean square estimate of $\Phi^{o}_{po}$ can be based on an expanded number of equally spaced discrete time samples of $\underline{x}_p$ and $\underline{x}_e$.

Consider as an example the addition of one more sample occurring at time $(2n+1)\Delta T$. The discrete time estimate may then be written as

$$\hat{\Phi}^{o}_{po}(2n+1) = [\underline{x}_p(2n+1)\quad X_p(2n)]\ [\underline{x}(2n)\quad X(2n-1)]^T M(2n) \tag{100}$$

where

$$M(2n) = [\ [\underline{x}(2n)\quad X(2n-1)]\ [\underline{x}(2n)\quad X(2n-1)]^T]^{-1} \tag{101}$$

If equations of the form given by (100) and (101) are applied, all data must be saved and multiplication must be performed on matrices of ever increasing dimension. However, Lee [16] has shown that recursive equations can be obtained for the single input/output system. An original development of the recursive equations for the multivariable system is given in Appendix G. The resulting recursive equation for the minimum mean square error estimate at time $(m+1)\Delta T$ is

$$\hat{\Phi}^{o}_{po}(m+1) = \hat{\Phi}^{o}_{po}(m) \tag{102}$$

$$+ [\underline{x}_p(m+1) - \hat{\Phi}^{o}_{po}(m)\underline{x}(m)]\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

where the recursive coefficient equation is given by

$$M(m) = M(m-1) - M(m-1)\underline{x}(m)[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1) \tag{103}$$

Note that the indicated matrix inversion is actually just division by a scalar so that the matrix inversion originally indicated by (101) has also been avoided by the recursive equations.

The recursive discrete time identification equation (103) is quite similar to that given by Ho [12], one difference being that a computed matrix gain is utilized in place of an unspecified scalar gain. Thus Ho's technique [12] is a discrete time response error learning model which, if applied to the incomplete information differential game, would seem to be as susceptible to problems similar to those previously discussed for continuous time learning models.

The recursive discrete time parameter identification equation given by (102) and its recursive coefficient equation given by (103) can be initialized in several ways as shown in Appendix G. In the present application, it is desirable to make use of the limiting estimate for the Pursuer's open-loop parameters, i.e., $\hat{F}_p$ and $\hat{G}_p$ given by (58) and (59), respectively. Using (43) and (44) in (1) and (2), along with (39), the Evader's optimistic estimate of the incomplete information differential game becomes

$$\begin{bmatrix} \dot{\hat{\underline{x}}}_p \\ \dot{\hat{\underline{x}}}_e \end{bmatrix} = \begin{bmatrix} [\hat{F}_p - \hat{G}_p R_p^{-1} \hat{G}_p^T P] & \hat{G}_p R_p^{-1} \hat{G}_p^T P \\ - G_e R_e^{-1} G_e^T P & [F_e + G_e R_e^{-1} G_e^T P] \end{bmatrix} \begin{bmatrix} \hat{\underline{x}}_p \\ \hat{\underline{x}}_e \end{bmatrix} \tag{104}$$

or

$$\begin{bmatrix} \dot{\hat{\underline{x}}}_p \\ \dot{\hat{\underline{x}}}_e \end{bmatrix} = \begin{bmatrix} \hat{A}^o_{pp} & \hat{A}^o_{pe} \\ A^*_{ep} & A^*_{ee} \end{bmatrix} \begin{bmatrix} \hat{\underline{x}}_p \\ \hat{\underline{x}}_e \end{bmatrix} \quad . \tag{105}$$

Thus, the limiting estimates for the Pursuer's closed-loop parameters during the early stages of the incomplete information game become

$$\begin{aligned} \hat{A}^o_{ppo} &= \hat{F}_p - \hat{G}_p R_p^{-1} \hat{G}_p^T P_o \\ &= F_e - G_e R_e^{-1} G_e^T P_o \end{aligned} \tag{106}$$

and

$$\begin{aligned} \hat{A}^o_{peo} &= \hat{G}_p R_p^{-1} \hat{G}_p^T P_o \\ &= G_e R_e^{-1} G_e^T P_o \quad . \end{aligned} \tag{107}$$

The limiting estimates for $\Phi^o_{ppo}$ and $\Phi^o_{peo}$ may now be computed using (82) and (83), i.e.,

$$\hat{\Phi}^o_{ppo}(0) = I + \hat{A}^o_{ppo} \Delta T \tag{108}$$

and

$$\hat{\Phi}^o_{peo}(0) = \hat{A}^o_{peo} \Delta T \quad . \tag{109}$$

In order to initialize the recursive identification equation (102) at $t = 0$ using the limiting estimates $\hat{F}_p$ and $\hat{G}_p$ and have the identification begin with the first sample, $M(-1)$ must also be estimated. As shown in Appendix G, this may be accomplished in several different ways.

From Appendix G, the best overall choice appears to be simply let

$$M(-1) = a\ I \tag{110}$$

where $a \to \infty$. The result is an asymptotically convergent solution for M and $\hat{\Phi}^{o}_{po}$ which begins with the first sample and seems to be relatively insensitive to measurement noise.

## C. Open-Loop System Parameter Identification

It was shown in the previous section that the smart player's optimal closed-loop system parameters can be identified via a recursive discrete time parameter identification technique provided they are time-invariant. As shown by (22), these parameters are time-invariant whenever the Riccati matrix is time-invariant. The fact that during the early stages of the game the Riccati matrix is time-invariant if $\bar{A}_o$ is stable and also satisfies the nonlinear algebraic matrix given by (26) can be utilized to estimate the smart player's open-loop system parameters from the identified closed-loop parameters.

Equation (26) may be rewritten in terms of the estimated and known matrices as

$$\begin{bmatrix} \hat{P}_{ppo} & \hat{P}_{peo} \\ \hat{P}_{epo} & \hat{P}_{eeo} \end{bmatrix} \begin{bmatrix} \hat{F}_p & 0 \\ 0 & F_e \end{bmatrix} + \begin{bmatrix} \hat{F}^T_p & 0 \\ 0 & F^T_e \end{bmatrix} \begin{bmatrix} \hat{P}_{ppo} & \hat{P}_{peo} \\ \hat{P}_{epo} & \hat{P}_{eeo} \end{bmatrix} + \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix}$$
$$- \begin{bmatrix} \hat{P}_{ppo} & \hat{P}_{peo} \\ \hat{P}_{epo} & \hat{P}_{eeo} \end{bmatrix} \begin{bmatrix} \hat{G}_p R^{-1}_p \hat{G}^T_p & 0 \\ 0 & -G_e R^{-1}_e G^T_e \end{bmatrix} \begin{bmatrix} \hat{P}_{ppo} & \hat{P}_{peo} \\ \hat{P}_{epo} & \hat{P}_{eeo} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{111}$$

where the circumflex is used to distinguish estimated or identified matrices from known matrices. Performing the operations indicated in (111) and noting that

$$\hat{A}^{o}_{ppo} = \hat{F}_{p} - \hat{G}_{p}R_{p}^{-1}\hat{G}_{p}^{T}P_{ppo} \tag{112}$$

and

$$\hat{A}^{o}_{peo} = -\hat{G}_{p}R_{p}^{-1}\hat{G}_{p}^{T}P_{peo} \tag{113}$$

yields the following three equations:

$$\hat{P}_{ppo}\hat{A}^{o}_{ppo} + \hat{A}^{o}_{ppo}\hat{P}_{ppo} + \hat{P}_{ppo}\hat{G}_{p}R_{p}^{-1}\hat{G}_{p}^{T}P_{ppo} + \hat{P}_{ep}^{T}G_{e}R_{e}^{-1}G_{e}^{T}\hat{P}_{ep} + Q = [0] \tag{114}$$

$$\hat{P}_{epo}\hat{A}^{o}_{ppo} + F_{e}^{T}\hat{P}_{epo} + \hat{P}_{eeo}G_{e}R_{e}^{-1}G_{e}^{T}\hat{P}_{epo} - Q = [0] \tag{115}$$

$$\hat{P}_{epo}\hat{A}^{o}_{peo} + \hat{P}_{eeo}F_{e} + F_{e}^{T}\hat{P}_{eeo} + \hat{P}_{eeo}G_{e}R_{e}^{-1}G_{e}^{T}\hat{P}_{eeo} + Q = [0] \quad . \tag{116}$$

Equations (115) and (116) may be solved simultaneously for $\hat{P}_{epo}$. Using the fact that $\hat{P}_{peo} = \hat{P}_{epo}^{T}$ and assuming that $\hat{P}_{peo}$ is nonsingular, from (113) one obtains

$$\hat{G}_{p}R_{p}^{-1}\hat{G}_{p}^{T} = -\hat{A}^{o}_{peo}\hat{P}_{peo}^{-1} \quad . \tag{117}$$

The estimate for $G_p$ may now be obtained from (117). Since $\hat{G}_{p}R_{p}^{-1}\hat{G}_{p}^{T}$ is now known, equation (114) may be solved for $\hat{P}_{ppo}$. Rewriting (112) one obtains

$$\hat{F}_{p} = \hat{A}^{o}_{ppo} + \hat{G}_{p}R_{p}^{-1}\hat{G}_{p}^{T}\hat{P}_{ppo}. \tag{118}$$

Thus, estimates of the smart player's open-loop system parameters may be calculated from the identified closed-loop parameters.

D. Simulation

The digital simulation for the example incomplete information game described in Chapter III was modified in order to include identification of the Pursuer's open-loop and closed-loop parameters, following the previous assumption that the Evader is the ignorant player. The recursive discrete time parameter identification technique discussed above in Section B-3 and derived in Appendix G was utilized to estimate the Pursuer's closed-loop parameters. The recursive identification equations were initialized as discussed in Section B-3, which is the third initialization technique proposed in Appendix G. The Pursuer's open-loop parameters were calculated from the closed-loop parameters using the technique discussed in Section C above.

In order to allow sampling of the state variables every 0.001 second, the integration step size of the fourth order Runge-Kutta integration scheme utilized by the digital simulation was also reduced to 0.001 second. The effect of this change in integration step size on the results of the simulation was negligible.

The simulation showed that for the example incomplete information game the Pursuer's closed-loop parameters can be readily identified. The resulting identification is almost identical to that shown in Figure G-3 of Appendix G. As shown there, the parameter $A^o_{peo}$ is identified at $t = 0.001$ second, while identification of the parameter $A^o_{ppo}$ is essentially complete by $t = 0.050$ second.

The resulting open-loop identification is shown in Figure 6. The jump seen at $t = 0.001$ second results from the fact that $A^o_{peo}$ is

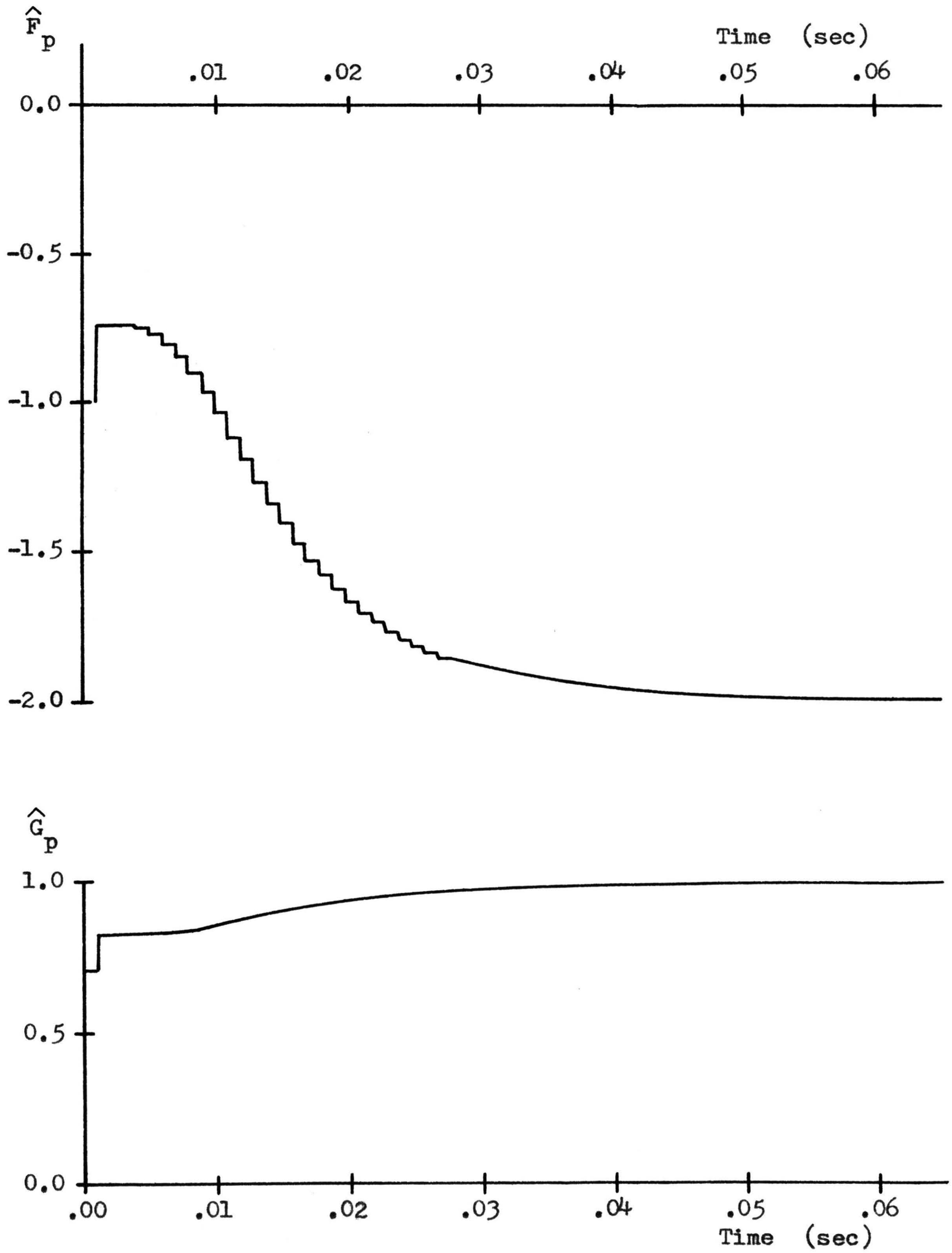



Figure 6. Identification of Pursuer's Open-Loop Parameters

identified immediately, while the estimate $\hat{A}^{o}_{ppo}$ has not changed appreciably. The smooth response after the 0.001 second point results from the smooth convergence of $\hat{A}^{o}_{ppo}$.

As shown in Figure 6, the estimates for $F_p$ and $G_p$ remain at their respective optimistic limiting estimates, as obtained from (58) and (59), for only one sample period. Note also that since the closed-loop parameters $A^{o}_{ppo}$ and $A^{o}_{peo}$ are not identified at the same rate, the estimate for $F_p$ actually exceeds the optimistic limiting estimate, as given by (58), for the first 0.010 second of the example incomplete information game.

# V. ASYMPTOTICALLY OPTIMAL ADAPTIVE GAMES

## A. Adaptive Strategies

As was shown in Chapter IV, the Pursuer's open-loop system given by (1) can be identified if: (1) the Pursuer plays his optimal feedback control strategy, (2) the $\bar{A}_o$ matrix for the game given by (28) is stable, (3) the remaining playing time, i.e., $[t-t_f]$ is long, and (4) the initial conditions excite all of the modes of the closed-loop incomplete information game. The net effect of conditions (1), (2), and (3) above is that the Pursuer's optimal feedback control gains and, therefore, the optimal closed-loop matrices $A^o_{pp}$ and $A^o_{pe}$ are approximately time-invariant.

A review of Chapter IV shows further, that the only effect that the Evader's closed-loop control gains have on the identification of the Pursuer's open-loop system parameters is associated with the response of the closed-loop game to initial conditions. Therefore, the Evader's feedback control gains are allowed to vary as desired as long as all modes of the game remain excited by the initial conditions.

Based on the above discussion, it is obvious that the Evader's feedback control gains can be made adaptive by using the estimate for the Pursuer's open-loop system parameters to solve the matrix Riccati equation for the game, i.e.,

$$\begin{bmatrix} \dot{\hat{P}}_{pp} & \dot{\hat{P}}_{pe} \\ \dot{\hat{P}}_{ep} & \dot{\hat{P}}_{ee} \end{bmatrix} = - \begin{bmatrix} \hat{P}_{pp} & \hat{P}_{pe} \\ \hat{P}_{ep} & \hat{P}_{ee} \end{bmatrix} \begin{bmatrix} \hat{F}_p & 0 \\ 0 & F_e \end{bmatrix} - \begin{bmatrix} \hat{F}_p^T & 0 \\ 0 & F_e^T \end{bmatrix} \begin{bmatrix} \hat{P}_{pp} & \hat{P}_{pe} \\ \hat{P}_{ep} & \hat{P}_{ee} \end{bmatrix}$$

$$- \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} + \begin{bmatrix} \hat{P}_{pp} & \hat{P}_{pe} \\ \hat{P}_{ep} & \hat{P}_{ee} \end{bmatrix} \begin{bmatrix} \hat{G}_p R_p^{-1} \hat{G}_p^T & 0 \\ 0 & -G_e R_e^{-1} G_e^T \end{bmatrix} \begin{bmatrix} \hat{P}_{pp} & \hat{P}_{pe} \\ \hat{P}_{ep} & \hat{P}_{ee} \end{bmatrix} \tag{119}$$

The suboptimal feedback control strategy given by (44) is then replaced by a feedback control strategy of the form given by (20), i.e.,

$$\underline{u}_e(t) = R_e^{-1} G_e^T \hat{P}_{ep}(t)\, \underline{x}_p(t) + R_e^{-1} G_e^T \hat{P}_{ee}(t)\, \underline{x}_e(t) \quad . \tag{120}$$

It is obvious that $\hat{P}_{ep} \rightarrow P_{ep}$ and $\hat{P}_{ee} \rightarrow P_{ee}$ and, therefore, $\underline{u}_e \rightarrow \underline{u}_e^o$ as $\hat{F}_p \rightarrow F_p$ and $\hat{G}_p \rightarrow G_p$. Thus, the proposed adaptive feedback control strategy is asymptotically optimal.

B. Simulation

The digital simulation described in Chapter IV was modified in order to include the asymptotically optimal adaptive feedback control strategy proposed above. The Evader utilized the estimates for the Pursuer's open-loop parameters, i.e., $\hat{F}_p$ and $\hat{G}_p$, throughout the game. However, the estimate for $F_p$ was not allowed to exceed the optimistic limiting estimate obtained from (58). Furthermore, the Pursuer's open-loop parameters were assumed to be accurately identified at 0.1 second play of the example game. These estimates were then utilized by the Evader for the remaining 4.9 seconds play of the game.

An integration step size of 0.001 second was used for the first 0.1 second of play while parameter identification was taking place. Only a 0.1 second integration step size was required for the remaining 4.9 seconds of the example adaptive game.

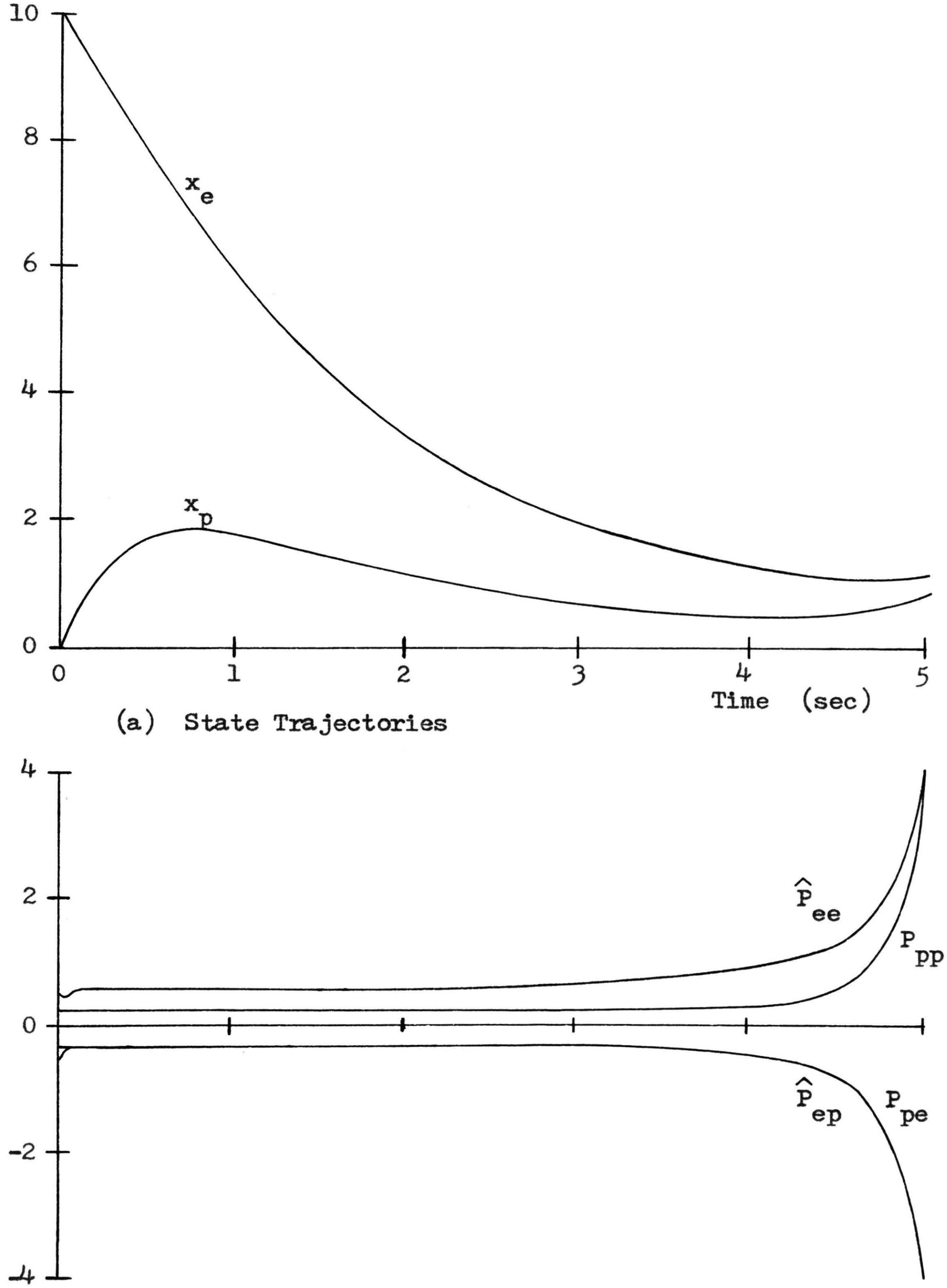
The results of the adaptive game are compared with those of the incomplete information game and of the perfect information game in Table II. As shown, the use of the asymptotically optimal adaptive control improves, at least from the Evader's point of view, both the performance index and the terminal miss distance.

The trajectories for the states of the example adaptive game are shown in Figure 7a. As seen by comparison to Figure 2a, the trajectories for the adaptive game are almost identical to those for the complete information game. Likewise, the Evader's solution to the matrix Riccati equation, which is based on the estimates $\hat{F}_p$ and $\hat{G}_p$, is nearly the same as that for the perfect information game after the initial 0.05 second adaptive transient, as can be seen by comparison of Figures 7b and 2b.

Table II.

Adaptive Game Performance Comparison

| Ignorant Player | Evader | Evader | Neither |
|---|---|---|---|
| Adaptive Control | No | Yes | —— |
| Performance Index (J) | 26.1389 | 26.3634 | 26.3672 |
| Terminal Miss $[x_e(5)-x_p(5)]$ | 0.2321 | 0.3386 | 0.3382 |




(a) State Trajectories

(b) Riccati Equation Solution

Figure 7. Example Adaptive Game - Evader Ignorant

# VI. CONCLUSIONS

A two player zero-sum differential game has been investigated for the case in which one of the players has incomplete a priori knowledge of the parameters of his opponent's dynamic system. The motion of each player is governed by a decoupled linear time-invariant dynamic system. The players, who are termed the Pursuer and the Evader, simultaneously try to minimize and maximize a quadratic time-invariant performance index by selection of deterministic feedback control strategies.

As has been shown herein, the incomplete system parameter information game is playable since the ignorant player can make limiting estimates for his opponent's system parameters from the relative controllability condition for the game. While play of the incomplete information game is possible, the performance, at least from the ignorant player's point of view, is suboptimal.

The response of the smart player to initial conditions and to the ignorant player's motion inherently contains information on the smart player's dynamic system. In this dissertation, parameter identification has been applied in the particular case in which the smart player's optimal closed-loop feedback control gains become time-invariant during the early stages of the game. As has been shown, the accessibility of the player's states allows the ignorant player to identify the smart player's optimal closed-loop parameters. The smart player's open-loop parameters can then be estimated from the identified closed-loop parameters by using the algebraic matrix Riccati equation for the game. If the ignorant player utilizes these estimated open-loop parameters in the optimal control law, the result is an asymptotically optimal adaptive feedback control strategy.

All of the results described above have also been verified by digital simulation for an example differential game.

In attempting to identify the smart player's optimal closed-loop system parameters, both continuous and discrete time identification techniques were applied. The continuous time techniques included the multivariable response error learning model developed by Pazdera and others [21], [28] ; an original multivariable extension of the single input/output equation error learning model developed by Lion [17] ; an original generalization of the multivariable equation error learning model; and an original combination of the response and equation error learning models. The results of the application of continuous time identification were disappointing, at least for the example incomplete system parameter information game. However, all of these continuous time learning model identification techniques are deserving of further investigation, both in the present application and in more normal system identification problems.

The discrete time parameter identification applied to the incomplete system parameter information game was an original multivariable extension of the single input/output identification technique developed by Lee [16]. This technique, which results in a minimum mean square error estimate, yielded good results for the example incomplete system parameter information game.

While important new results have been obtained for the incomplete system parameter information game, these results apply to a fairly specific differential game, namely, one in which (1) the dynamic systems are dimensionally identical, (2) the states are directly measurable, and (3) the optimal feedback control gains are approximately constant if the remaining playing time is sufficiently long. Generalization of the previous results to include dynamic systems of arbitrary dimensions whose states are observable, but not directly measurable would be worthwhile as would generalization to allow identification and adaptive control while the optimal feedback control gains are time-varying. Investigation of the incomplete information game in which both players lack a priori knowledge of their opponent's system parameters is also of interest.

BIBLIOGRAPHY


1. B. D. O. Anderson and J. B. Moore, Linear Optimal Control, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.

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

3. R. D. Behn and Y. C. Ho, "Characteristics of Stochastic Pursuit-Evasion Games," 1969 IFAC Preprint.

4. R. D. Behn and Y. C. Ho, "On a Class of Linear Stochastic Differential Games," IEEE Transactions on Automatic Control, Vol. AC-13, No. 3, pp. 227-240, June, 1968.

5. A. Blaquiere, F. Gerard, and G. Leitmann, Quantitative and Qualitative Games, Academic Press, New York, New York, 1969.

6. A. E. Bryson and Y. C. Ho, Applied Optimal Control, Blaisdell Publishing Co., Waltham, Massachusetts, 1969.

7. M. Dresher, A. W. Tucker, and P. Wolfe, Contributions to the Theory of Games, Vol. III, Princeton University Press, Princeton, New Jersey, 1957.

8. L. E. Elsgolc, Calculus of Variations, Addison-Weslay Publishing Co., Inc., Reading, Massachusetts, 1962.

9. R. W. Hamming, Numerical Methods for Scientists and Engineers, McGraw-Hill Book Co., Inc., New York, New York, 1962.

10. Y. C. Ho,"Toward Generalized Control Theory," IEEE Transactions on Automatic Control, Vol. AC-14, No. 6, pp. 753-754, December, 1969.

11. Y. C. Ho and A. W. Starr, "Nonzero-Sum Differential Games," Journal on Optimization and Applications, March, 1969.

12. Y. C. Ho and B. H. Whalen, "An Approach to the Identification and Control of Linear Dynamic Systems with Unknown Parameters," IEEE Transactions on Automatic Control, Vol. AC-8, No. 3, pp. 255-256, July, 1963.

13. R. Isaacs, Differential Games, John Wiley & Sons, Inc., New York, New York, 1965.

14. H. W. Kuhn and A. W. Tucker, Contributions to the Theory of Games, Vol. I, Princeton University Press, Princeton, New Jersey, 1950.

15. H. W. Kuhn and A. W. Tucker, Contributions to the Theory of Games, Vol. II, Princeton University Press, Princeton, New Jersey, 1953.

16. R. C. K. Lee, Optimal Estimation, Identification, and Control, M.I.T. Press, Cambridge, Massachusetts, 1964.

17. P. M. Lion, "Rapid Identification of Linear and Nonlinear Systems," Proceedings of the 1966 Joint Automatic Control Conference, pp. 605-615, August, 1966.

18. A. L. Nelson, K. W. Folley, and M. Coral, Differential Equations, D.C. Heath and Co., Boston, Massachusetts, 1960.

19. J. J. O'Donnell, "Asymptotic Solution of the Matrix Riccati Equation of Optimal Control, "Proceedings Fourth Annual Allerton Conference on Circuits and System Theory, pp. 577-586, October, 1966.

20. K. Ogata, State Space Analysis of Control Systems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1967.

21. J. S. Pazdera and H. J. Pottinger, "Linear System Identification Via Liapunov Design Techniques," Proceedings of the Joint Automatic Control Conference, pp. 795-801, August, 1969.

22. W. A. Porter, "On the Matrix Riccati Equation," IEEE Transactions on Automatic Control, Vol. AC-12, No. 6, pp. 746-749, December, 1967.

23. I. B. Rhodes, "On Nonzero-sum Differential Games with Quadratic Cost Functionals," Proceedings of the First International Conference on the Theory and Applications of Differential Games, University of Massachusetts, pp. IV-1-6, September, 1969.

24. I. B. Rhodes and D. G. Luenberger, "Differential Games with Imperfect State Information," IEEE Transactions on Automatic Control, Vol. AC-14, No. 1, pp. 29-38, February, 1969.

25. I. B. Rhodes and D. G. Luenberger, "Stochastic Differential Games with Constrained State Estimators," IEEE Transactions on Automatic Control, Vol. AC-14, No. 5, pp. 476-481, October, 1969.

26. A. P. Sage and J. L. Melsa, System Identification, Academic Press, New York, New York, 1971.

27. H. F. Spence, "Identification of Linear Systems With Delay Via a Learning Model," Ph. D. Dissertation, University of Missouri-Rolla, 1971.

28. H. F. Spence and J. S. Pazdera, "Identification of Linear Systems With Delay Via a Learning Model," American Society of Mechanical Engineers Paper No. 72 - WA/AUT-13, November, 1972.

29. A. W. Tucker and R. D. Luce, Contributions to the Theory of Games, Vol. IV, Princeton University Press, Princeton, New Jersey, 1959.

30. D. R. Vaughan, "A Negative Exponential Solution for the Matrix Riccati Equation," IEEE Transactions on Automatic Control, Vol. AC-14, No. 1, pp. 72-75, February, 1969.

31. J. Von Neumann, "Zur Theorie der Gesellschaftsspiele," Mathematische Annalen 100, pp. 259-320, 1928.

32. J. Von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, New Jersey, 1953.

33. R. Weinstock, Calculus of Variations, McGraw-Hill Book Co., New York, New York, 1952.

34. W. W. Willman, "Formal Solutions for a Class of Stochastic Pursuit-Evasion Games," Harvard University, Cambridge, Massachusetts.

# VITA

John D. Corrigan was born on March 5, 1942 in St. Charles, Missouri where he also received his primary and secondary education.

He entered the University of Missouri, School of Mines and Metallurgy, in September, 1960. During his undergraduate career, he supplemented his education with engineering experience as a Co-operative Engineering Student with McDonnell Aircraft Corporation, St. Louis, Missouri. He received a Bachelor of Science Degree in Electrical Engineering from the University of Missouri at Rolla in May, 1965. Following graduation, he joined the McDonnell Aircraft Company as an Associate Engineer in the Electronic Systems Engineering Department.

Mr. Corrigan enrolled in the Graduate School of the University of Missouri at Rolla in September, 1965 and received a Master of Science in Electrical Engineering from the University in May, 1966. He held a National Defense Education Act, Title IV, Fellowship for the period from September, 1965 through August, 1966.

Mr. Corrigan is presently employed as a Technical Specialist in the Electronic Systems Technology Department of McDonnell Aircraft Company, Division of McDonnell Douglas Corporation, St. Louis, Missouri.

Mr. Corrigan is a member of Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu, and the Institute of Electrical and Electronics Engineers.

# APPENDIX A

## SUFFICIENCY CONDITIONS FOR THE DIFFERENTIAL GAME OPTIMAL SOLUTION

The extremal problem presented by the zero sum differential game defined in Chapter II is easily solved via the calculus of variations [8], [33]. However, since the differential game must satisfy the saddle point condition (4) of Chapter II, it is a minimax problem rather than either a maximum or a minimum problem. In order for a deterministic (pure) solution to exist, the minimax solution must be equivalent to the maximin solution, i.e., the optimal solution must be independent of the order of selection of control strategies. The minimax and maximin solutions are equivalent if the extremal problem is separable into one-sided minimum and maximum problems having the same solutions as obtained for the original two-sided problem.

In the following sections, the above separability condition is shown to hold for the differential game defined in Chapter II. The necessary and sufficient conditions for the optimal solution of this game are also derived.

### A-1. The Two-Sided Extremal Problem

The two-sided extremal problem for the differential game is defined by the performance index

$$J(\underline{u}) = \tfrac{1}{2}\int_0^{t_f} (\underline{x}^T\overline{Q}\ \underline{x} + \underline{u}^T\overline{R}\ \underline{u})\ dt + \tfrac{1}{2}\ \underline{x}^T(t_f)\ \overline{S}\ \underline{x}\ (t_f) \tag{A-1}$$

subject to the differential equation constraint

$$\dot{\underline{x}} = \overline{F}\ \underline{x} + \overline{G}\ \underline{u} \quad ; \quad \underline{x}(0) = \underline{x}_o \tag{A-2}$$

where $\bar{Q}$, $\bar{R}$, $\bar{S}$, $\bar{F}$, and $\bar{G}$ are defined by (9) and (10) of Chapter II. Adjoining the differential constraint to the performance index via the Lagrange undetermined multiplier vector $\underline{y}$ yields the augmented performance index

$$\bar{J}(\underline{u}) = \int_0^{t_f} \left(\tfrac{1}{2}\, \underline{x}^T \bar{Q}\, \underline{x} + \tfrac{1}{2}\, \underline{u}^T \bar{R}\, \underline{u} + \underline{y}^T [\bar{F}\, \underline{x} + \bar{G}\, \underline{u} - \dot{\underline{x}}]\right) dt + \tfrac{1}{2}\, \underline{x}^T(t_f)\, \bar{S}\, \underline{x}\,(t_f) \quad . \tag{A-3}$$

Integrating by parts, (A-3) becomes

$$\bar{J}(\underline{u}) = \int_0^{t_f} \left(\tfrac{1}{2}\, \underline{x}^T \bar{Q}\, \underline{x} + \tfrac{1}{2}\, \underline{u}^T \bar{R}\, \underline{u} + \underline{y}^T [\bar{F}\, \underline{x} + \bar{G}\, \underline{u}] + \dot{\underline{y}}^T \underline{x}\right) dt + [\tfrac{1}{2}\, \underline{x}^T(t_f)\, \bar{S} - \underline{y}^T(t_f)]\, \underline{x}(t_f) + \underline{y}^T(0)\, \underline{x}(0) \; . \tag{A-4}$$

The first order necessary conditions for an extremum are obtained from the first variation of the augmented performance index, which is given by

$$\delta\bar{J}(\underline{u}) = \int_0^{t_f} \left( [\underline{u}^T \bar{R} + \underline{y}^T \bar{G}]\, \delta\underline{u} + [\underline{x}^T \bar{Q} + \underline{y}^T \bar{F} + \dot{\underline{y}}^T]\, \delta\underline{x}\right) dt + [\underline{x}^T(t_f)\, \bar{S} - \underline{y}^T(t_f)]\, \delta\underline{x}(t_f) + \underline{y}^T(0)\, \delta\underline{x}(0) \; . \tag{A-5}$$

Note that $\delta\underline{x}(0) = \underline{0}$ since the initial conditions are specified.

The first variation given by (A-5) must vanish along the extremal path for arbitrary values of $\delta\underline{u}$ and $\delta\underline{x}$. Therefore, the first order necessary conditions for an extremum are given by

$$\underline{u} = -\, \bar{R}^{-1} \bar{G}^T \underline{y} \quad , \tag{A-6}$$

$$\dot{\underline{y}} = -\bar{Q}\,\underline{x} - \bar{F}^T\underline{y} \quad , \tag{A-7}$$

and

$$\underline{y}(t_f) = \bar{S}\,\underline{x}(t_f) \quad . \tag{A-8}$$

Substituting (A-6) into (A-2) yields

$$\dot{\underline{x}} = \bar{F}\,\underline{x} - \bar{H}\,\underline{y} \quad ; \quad \underline{x}(0) = \underline{x}_o \quad , \tag{A-9}$$

where

$$\bar{H} = \bar{G}\,\bar{R}^{-1}\bar{G}^T \quad . \tag{A-10}$$

Equations (A-7), (A-8), and (A-9) form a two point boundary value problem (TPBVP) which may be written in state equation form as

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{y}} \end{bmatrix} = \begin{bmatrix} \bar{F} & -\bar{H} \\ -\bar{Q} & -\bar{F}^T \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} \quad ; \quad \begin{matrix} \underline{x}(0) = \underline{x}_o \\ \underline{y}(T) = \bar{S}\,\underline{x}(t_f) \quad . \end{matrix} \tag{A-11}$$

As a result of (A-11), the Lagrange multiplier vector $\underline{y}$ is also often termed the costate vector.

The above TPBVP may be solved by defining

$$\underline{y}(t) = \bar{P}(t)\,\underline{x}(t) \quad . \tag{A-12}$$

Differentiating (A-12), one obtains

$$\dot{\underline{y}} = \dot{\bar{P}}\,\underline{x} + \bar{P}\,\dot{\underline{x}} \tag{A-13}$$

which, after substitution of (A-7), (A-9), and (A-12), becomes

$$[\dot{\bar{P}} + \bar{Q} + \bar{P}\,\bar{F} + \bar{F}^T\bar{P} - \bar{P}\,\bar{H}\,\bar{P}]\,\underline{x} = 0 \quad . \tag{A-14}$$

Also from (A-8) and (A-12), one gets

$$[\bar{P}(t_f) - \bar{S}]\,\underline{x}(t_f) = 0 \quad . \tag{A-15}$$

Since $\underline{x}(0)$ is arbitrary, $\underline{x}(t) \neq 0$ is the general case. Therefore, (A-14) is true for all $t \in [0, t_f]$ if and only if

$$\dot{\overline{P}} = -\overline{Q} - \overline{P}\,\overline{F} - \overline{F}^T\overline{P} + \overline{P}\,\overline{H}\,\overline{P} \quad ; \quad \overline{P}(t_f) = \overline{S} \quad . \tag{A-16}$$

Equation (A-16) is the nonlinear matrix Riccati equation. Note that since $\overline{H}$, $\overline{Q}$, $\overline{S}$ are symmetric matrices, $\overline{P}$ is also symmetric.

Substituting (A-12) into (A-6), the optimal control vector, $\underline{u}^o$, for the game is given by

$$\underline{u}^o(t) = -\overline{R}^{-1}\overline{G}^T\overline{P}(t)\,\underline{x}(t) \tag{A-17}$$

where $\overline{P}(t)$ is obtained by solution of the matrix Riccati equation. Performing the operations indicated in (A-17), the control strategies for the Pursuer and the Evader respectively are given by

$$\underline{u}_p^o(t) = -\left[R_p^{-1}G_p^T \quad 0\right]\overline{P}(t)\,\underline{x}(t) \tag{A-18}$$

and

$$\underline{u}_e^o(t) = +\left[0 \quad R_e^{-1}G_e^T\right]\overline{P}(t)\,\underline{x}(t) \quad . \tag{A-19}$$

A-2. Separability Condition

In order for (A-18) and (A-19) to be deterministic optimal control strategies, they must also satisfy the saddle point condition

$$J(\underline{u}_p^o, \underline{u}_e) \leq J(\underline{u}_p^o, \underline{u}_e^o) \leq J(\underline{u}_p, \underline{u}_e^o) \tag{A-20}$$

i.e., (A-18) and (A-19) must be both the minimax and the maximin solutions of the two-sided extremal problem. The minimax and maximin solutions are equivalent if the two separate one-sided extremal problems have the same solutions as those obtained for the original two-sided problem.

In order to show that the differential game defined in Chapter II is separable, consider the Pursuer's one-sided minimum problem wherein the Evader utilizes (A-19) as a closed-loop feedback control strategy. For this case, the extremal problem is defined by the performance index

$$J_p(\underline{u}_p) = \frac{1}{2}\int_0^{t_f} \left(\underline{u}_p^T R_p \underline{u}_p - \underline{x}^T \overline{P} \begin{bmatrix} 0 \\ G_e R_e^{-1} \end{bmatrix} R_e \begin{bmatrix} 0 & R_e^{-1} G_e^T \end{bmatrix} \overline{P}\, \underline{x}\right) dt$$

$$+ \frac{1}{2}\int_0^{t_f} (\underline{x}^T \overline{Q}\, \underline{x})\, dt + \frac{1}{2}\, \underline{x}^T(t_f)\, \overline{S}\, \underline{x}(t_f) \qquad \text{(A-21)}$$

subject to the differential equation constraint given by

$$\dot{\underline{x}} = \overline{F}\, \underline{x} + \begin{bmatrix} G_p & 0 \\ 0 & G_e \end{bmatrix} \begin{bmatrix} \underline{u}_p \\ \begin{bmatrix} 0 & R_e^{-1} G_e^T \end{bmatrix} \overline{P}\, \underline{x} \end{bmatrix} . \qquad \text{(A-22)}$$

Adjoining this differential constraint to (A-21) via the undetermined multiplier vector $\underline{y}$ yields the augmented performance index

$$\overline{J}_p(\underline{u}_p) = \frac{1}{2}\int_0^{t_f} \left(\underline{x}^T \overline{Q}\, \underline{x} + \underline{u}_p^T R_p \underline{u}_p + \underline{x}^T \overline{P} \begin{bmatrix} 0 & 0 \\ 0 & -H_e \end{bmatrix} \overline{P}\, \underline{x}\right) dt$$

$$+ \int_0^{t_f} \left(\underline{y}^T\left[\overline{F}\, \underline{x} + \begin{bmatrix} G_p \underline{u}_p \\ \begin{bmatrix} 0 & H_e \end{bmatrix} \overline{P}\, \underline{x} \end{bmatrix} - \dot{\underline{x}}\right]\right) dt \qquad \text{(A-23)}$$

$$+ \frac{1}{2}\, \underline{x}^T(t_f)\, \overline{S}\, \underline{x}(t_f)$$

where

$$H_e = G_e R_e G_e^T \quad , \tag{A-24}$$

Integrating by parts, (A-23) becomes

$$\bar{J}_p(\underline{u}_p) = \frac{1}{2}\int_0^{t_f} (\underline{x}^T\bar{Q}\ \underline{x} + \underline{u}_p^T R_p \underline{u}_p + \underline{x}^T\bar{P} \begin{bmatrix} 0 & 0 \\ 0 & -H_e \end{bmatrix} \bar{P}\ \underline{x})\ dt$$

$$+ \int_0^{t_f} (\underline{y}^T\bar{F}\ \underline{x} + \underline{y}^T \begin{bmatrix} G_p \underline{u}_p \\ [0 \quad H_e]\ \bar{P}\ \underline{x} \end{bmatrix} + \dot{\underline{y}}^T\underline{x})\ dt \tag{A-25}$$

$$+ [\tfrac{1}{2}\ \underline{x}^T(t_f)\ \bar{S} - \underline{y}^T(t_f)]\ \underline{x}(t_f) + \underline{y}^T(0)\underline{x}(0)\ .$$

The first order necessary conditions for the one-sided minimum problem are obtained from the first variation of (A-25), which is given by

$$\delta\,\bar{J}_p(\underline{u}_p) = \int_0^{t_f} ([\underline{u}_p^T R_p + \underline{y}^T \begin{bmatrix} G_p \\ 0 \end{bmatrix}]\ \delta\underline{u}_p + [\underline{x}^T\bar{Q} + \underline{x}^T\bar{P} \begin{bmatrix} 0 & 0 \\ 0 & -H_e \end{bmatrix}\ \bar{P}$$

$$+ \underline{y}^T\bar{F} + \underline{y}^T \begin{bmatrix} 0 & 0 \\ [0 & H_e] \end{bmatrix} \bar{P} + \dot{\underline{y}}^T]\ \delta\underline{x})\ dt \tag{A-26}$$

$$+ [\underline{x}^T(t_f)\ \bar{S} - \underline{y}^T(t_f)]\ \delta\underline{x}(t_f) + \underline{y}^T(0)\ \delta\underline{x}(0)$$

where $\delta\underline{x}(0) = 0$ since the initial conditions are specified.

The first variation given by (A-26) must vanish along the minimal path for arbitrary values of $\delta\underline{u}_p$ and $\delta\underline{x}$ . Therefore, the first order necessary conditions for a minimum are given by

$$\underline{u}_p = - R_p^{-1}[G_p^T \quad 0]\, \underline{y} \quad ; \tag{A-27}$$

$$\dot{\underline{y}} = - \overline{F}^T \underline{y} - \overline{P} \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \underline{y} - \overline{Q}\, \underline{x} + \overline{P} \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \overline{P}\, \underline{x} \quad ; \tag{A-28}$$

and

$$\underline{y}(t_f) = \overline{S}\, \underline{x}(t_f) \quad . \tag{A-29}$$

Substituting (A-27) into (A-22) yields

$$\dot{\underline{x}} = \overline{F}\, \underline{x} + \begin{bmatrix} -[H_p \quad 0]\, \underline{y} \\ [0 \quad H_e]\, \overline{P}\, \underline{x} \end{bmatrix} \quad ; \quad \underline{x}(0) = \underline{x}_o \tag{A-30}$$

where

$$H_p = G_p R_p^{-1} G_p^T \quad . \tag{A-31}$$

Equations (A-28), (A-29) and (A-30) form a TPBVP which may be solved by defining

$$\underline{y}(t) = \overline{P}_p(t)\, \underline{x}(t) \quad . \tag{A-32}$$

Differentiating (A-32) and substituting (A-28) and (A-30), followed by (A-32) yields

$$\begin{aligned} \dot{\overline{P}}_p = & - \overline{Q} - \overline{P}_p \overline{F} - \overline{F}^T \overline{P}_p - \overline{P} \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \overline{P}_p + \overline{P} \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \overline{P} \\ & + \overline{P}_p \begin{bmatrix} H_p & 0 \\ 0 & 0 \end{bmatrix} \overline{P}_p - \overline{P}_p \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \overline{P} \quad . \end{aligned} \tag{A-33}$$

Adding the identically zero quantity

$$\bar{P}_p \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \bar{P}_p - \bar{P}_p \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} \bar{P}_p$$

(A-33) may be rewritten as

$$\dot{\bar{P}}_p = -\bar{Q} - \bar{P}_p\bar{F} - \bar{F}^T\bar{P}_p + \bar{P}_p\bar{H}\,\bar{P}_p + [\bar{P}_p - \bar{P}] \begin{bmatrix} 0 & 0 \\ 0 & H_e \end{bmatrix} [\bar{P}_p - \bar{P}] \qquad \text{(A-34)}$$

where, from (A-29) and (A-32),

$$\bar{P}_p(t_f) = \bar{S} \quad . \qquad \text{(A-35)}$$

As seen by comparing (A-34) and (A-16), the only difference between the equations for $\dot{\bar{P}}$ and $\dot{\bar{P}}_p$ is the last term in the $\dot{\bar{P}}_p$ equation. However, since

$$\bar{P}_p(t_f) = \bar{S} = \bar{P}(t_f), \qquad \text{(A-36)}$$

then

$$\bar{P}_p(t) = \bar{P}(t) \qquad \text{(A-37)}$$

for all $t \in [0, t_f]$ since the last term in (A-34) can never contribute to $\dot{\bar{P}}_p$ .

Substituting (A-32) in (A-27) yields the Pursuer's optimal control strategy, $\underline{u}_p^o$, given by

$$\underline{u}_p^o = -[R_p^{-1}G_p^T \quad 0]\,\bar{P}_p\underline{x} \quad , \qquad \text{(A-38)}$$

or, using (A-37),

$$\underline{u}_p^o = -[R_p^{-1}G_p^T \quad 0]\,\bar{P}\,\underline{x} \quad , \qquad \text{(A-39)}$$

which is identical to (A-18).

As seen from (A-39) the solution to the Pursuer's one-sided minimum problem is identical to that obtained for the two-sided differential game. From the symmetry of the game, it is obvious that the solution to the Evader's one-sided maximum problem is also identical to that obtained for the two-sided differential game. Therefore, the differential game is separable and the controls given by (A-18) and (A-19) are deterministic control strategies satisfying (A-20).

The above proof is based on the assumption that the Evader utilized (A-19) as a closed-loop feedback control strategy. The Evader could, of course, utilize (A-19) as an open-loop feedback control strategy by using the solution of (A-16) together with (A-12) to obtain an analytic solution for (A-9). Since this results in the Evader following the same trajectory as for the closed-loop case, it makes no difference whether the Evader's control is open-loop or closed-loop.

## A-3. Sufficient Conditions

It has previously been shown that the deterministic control strategies $\underline{u}_p^o(t)$ and $\underline{u}_e^o(t)$ as given by (A-18) and (A-19) satisfy the first order necessary conditions for the optimal solution of the differential game defined in Chapter II. The sufficient conditions for these control strategies to be optimal may be obtained by examining the neighboring field of extremals.

Since the first variation of the augmented performance index given by (A-5) vanishes along the extremal path, the conditions for a neighboring extremal path are obtained from the second variation of the augmented performance index, which is given by

$$\delta^2\overline{J}(\underline{u}) = \frac{1}{2}\int_0^{t_f}(\delta\underline{x}^T\overline{Q}\,\delta\underline{x} + \delta\underline{u}^T\overline{R}\,\delta\underline{u})\,dt + \frac{1}{2}\,\delta\underline{x}(t_f)\,\overline{S}\,\delta\underline{x}(t_f) \qquad \text{(A-40)}$$

and the first variation of the constraint given by

$$\delta\dot{\underline{x}} = \overline{F}\,\delta\underline{x} + \overline{G}\,\delta\underline{u} \quad ; \quad \delta\underline{x}(0) = \delta\underline{x}_o \quad . \qquad \text{(A-41)}$$

The second variation given by (A-40) must vanish along the neighboring extremal path. Since (A-40) and (A-41) are identical in form to (A-1) and (A-2), respectively, it is obvious that the solution for the neighboring extremal path is given by

$$\delta\underline{u}(t) = -\,\overline{R}^{-1}\overline{G}^T\overline{P}(t)\,\delta\underline{x}(t) \qquad \text{(A-42)}$$

where $\overline{P}(t)$ satisfies (A-16).

In order to determine the sufficient conditions for the optimal solution to the differential game, consider the addition of the identically zero quantity,

$$\int_0^{t_f}(\delta\underline{x}^T\overline{P}\,[\overline{F}\,\delta\underline{x} + \overline{G}\,\delta\underline{u} - \delta\dot{\underline{x}}])\,dt$$

to (A-40). As a result, the second variation is also given by

$$\begin{aligned}\delta^2 J(\underline{u}) = {} & \frac{1}{2}\int_0^{t_f}(\delta\underline{x}^T[\overline{Q} + \overline{P}\,\overline{F} + \overline{F}^T\overline{P}]\,\delta\underline{x})\,dt \\ & + \frac{1}{2}\int_0^{t_f}(\delta\underline{x}^T\overline{P}\,\overline{G}^T\delta\underline{u} + \delta\underline{u}^T\overline{G}^T\overline{P}\,\delta\underline{x} + \delta\underline{u}^T\overline{R}\,\delta\underline{u})\,dt \\ & - \int_0^{t_f}(\delta\underline{x}^T\overline{P}\,\delta\dot{\underline{x}})\,dt + \frac{1}{2}\,\delta\underline{x}(t_f)\overline{S}\,\delta\underline{x}(t_f) \quad .\end{aligned} \qquad \text{(A-43)}$$

Note that

$$\frac{d}{dt}\left[\delta\underline{x}^T\overline{P}\,\delta\underline{x}\right] = 2\,\delta\underline{x}^T\overline{P}\,\delta\dot{\underline{x}} + \delta\underline{x}^T\dot{\overline{P}}\,\delta\underline{x} \tag{A-44}$$

so that

$$\int_0^{t_f} (d\underline{x}^T\overline{P}\,\delta\dot{\underline{x}})\,dt = \tfrac{1}{2}\,\delta\underline{x}^T(t_f)\,\overline{S}\,\delta\underline{x}(t_f) - \tfrac{1}{2}\,\delta\underline{x}(0)\,\overline{P}(0)\,\delta\underline{x}(0)$$

$$- \tfrac{1}{2}\int_0^{t_f} (\delta\underline{x}^T\dot{\overline{P}}\,\delta\underline{x})\,dt \quad . \tag{A-45}$$

Substituting (A-45) in (A-43) yields

$$\delta^2 J(\underline{u}) = \tfrac{1}{2}\int_0^{t_f} (\delta\underline{x}^T[\dot{\overline{P}} + \overline{Q} + \overline{P}\,\overline{F} + \overline{F}^T\overline{P}]\,\delta\underline{x})\,dt$$

$$+ \tfrac{1}{2}\int_0^{t_f} (\delta\underline{x}^T\overline{P}\,\overline{G}\,\delta\underline{u} + \delta\underline{u}^T\overline{G}^T\overline{P}\,\delta\underline{x} + \delta\underline{u}^T\overline{R}\,\delta\underline{u})\,dt$$

$$+ \tfrac{1}{2}\,\delta\underline{x}(0)\,\overline{P}(0)\,\delta\underline{x}(0) \quad . \tag{A-46}$$

Substituting (A-16) into (A-46) and using (A-10), the second variation may be written as

$$\delta^2\overline{J}(\underline{u}) = \tfrac{1}{2}\int_0^{t_f} (\,[\delta\underline{x}^T\overline{P}\,\overline{G}\,\overline{R}^{-1} + \delta\underline{u}^T]\,\overline{R}\,[\overline{R}^{-1}\overline{G}^T\overline{P}\,\delta\underline{x} + \delta\underline{u}])\,dt$$

$$+ \tfrac{1}{2}\,\delta\underline{x}(0)\,\overline{P}(0)\,\delta\underline{x}(0) \quad . \tag{A-47}$$

Performing the indicated matrix multiplications and assuming that the initial conditions are specified so that $\delta\underline{x}(0) = 0$ , (A-47) may be written as

$$\delta^2\bar{J}(\underline{u}_p,\underline{u}_e) = \frac{1}{2}\int_0^{t_f} \left( \left\| [R_p^{-1}G_p^T \quad 0]\ \bar{P}\ \delta\underline{x} + \delta\underline{u}_p \right\|_{R_p} \right) dt$$
$$- \frac{1}{2}\int_0^{t_f} \left( \left\| [0 \quad -R_e^{-1}G_e^T]\ \bar{P}\ \delta\underline{x} + \delta\underline{u}_e \right\|_{R_e} \right) dt \tag{A-48}$$

where $\left\| \underline{z} \right\|_R = \underline{z}^T R\ \underline{z}$ , the generalized norm of the vector $\underline{z}$ with respect to R.

From (A-48) it is evident that the solution is minimax or, equivalently, maximin if $R_p$ and $R_e$ are both positive definite. This condition is analogous to the "convexity" or "strengthened Legendre-Clebsch" of the calculus of variations.

While it has not been shown directly herein, it is also evident that $\bar{P}(t)$ must exist, i.e., be finite, for all $t \in [0,t_f]$ in order for the solution to be minimax. This condition is analogous to the "no conjugate point" or "Jacobi" condition of the calculus of variations.

## A-4. Summary

In summary, the necessary and sufficient conditions for the deterministic optimal solution of the zero sum differential game defined in Chapter II are:

(1) $\underline{u} = -\bar{R}^{-1}\bar{G}^T\bar{P}\ \underline{x}$ (A-49)

where $\bar{P}(t)$ satisfies

$$\dot{\bar{P}} = -\bar{Q} - \bar{P}\bar{F} - \bar{F}^T\bar{P} + \bar{P}\bar{H}\bar{P} \quad ; \quad \bar{P}(t_f) = \bar{S} \quad , \tag{A-50}$$

(2) $R_p$ and $R_e$ are both positive definite,

(3) $\bar{P}$ exists for $t \in [0,t_f]$ .

# APPENDIX B

## NONLINEAR MATRIX RICCATI EQUATION ANALYTIC SOLUTION

The solution to the time-invariant nonlinear matrix Riccati equation, given by (16) as

$$\dot{\overline{P}} = -\overline{P}\,\overline{F} - \overline{F}^T\overline{P} + \overline{P}\,\overline{G}\,\overline{R}^{-1}\overline{G}^T\overline{P} - \overline{Q} \quad ; \quad \overline{P}(t_f) = \overline{S} \quad , \tag{B-1}$$

may be obtained by considering the original two point boundary value problem defined by (14). In order to aid discussion the TPBVP matrix is defined as $\overline{M}$, i.e.,

$$\overline{M} = \begin{bmatrix} \overline{F} & -\overline{G}\,\overline{R}^{-1}\overline{G}^T \\ -\overline{Q} & -\overline{F}^T \end{bmatrix} \quad . \tag{B-2}$$

O'Donnell [19] has shown that the eigenvalues of $\overline{M}$ must be symmetric with respect to the imaginary axis of the complex plane and, if $\overline{A}_o$ is stable, that there are no purely imaginary eigenvalues. Utilizing the symplectic similarity property of the $\overline{M}$ matrix, O'Donnell [19], and later Vaughan [30] in a slightly different form, provide a solution for $\overline{P}(t)$ for the case in which the eigenvalues of $\overline{A}_o$, and consequently $\overline{M}$, are distinct. Subsequently, Anderson and Moore [1], in presenting the solution to the matrix Riccati equation, include the case for multiple eigenvalues.

To apply the above results to the solution to the matrix Riccati equation for the differential game (16), assume that $\overline{A}_o$ is a stable matrix and consider the similarity transformation W with the property that

$$\bar{M}\, W = W \begin{bmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{bmatrix} \tag{B-3}$$

where $\Lambda$ is the 2n x 2n Jordan canonical form with $Re(\lambda_i) < 0$, i = 1, 2,...,2n and $\lambda_i$ are the eigenvalues of $\bar{M}$. By partitioning the 4n x 4n W matrix into 2n x 2n submatrices such that

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \tag{B-4}$$

and using its symplectic characteristic, it can be shown [1] that

$$\bar{P}(t) = [W_{21} + W_{22} e^{\Lambda t_r} N\, e^{\Lambda t_r}]\, [W_{11} + W_{12} e^{\Lambda t_r} N\, e^{\Lambda t_r}]^{-1} \tag{B-5}$$

where

$$N = -\, [W_{22} - \bar{S}\, W_{12}]^{-1}\, [W_{21} - \bar{S}\, W_{11}] \tag{B-6}$$

and

$$t_r = t_f - t \quad . \tag{B-7}$$

From (B-5) it is seen that

$$\bar{P}_o = \lim_{t_r \to \infty} \bar{P}(t) = W_{21}\; W_{11}^{-1} \tag{B-8}$$

or

$$W_{21} = \bar{P}_o W_{11} \quad . \tag{B-9}$$

The fact that $W_{11}$ is nonsingular and $\bar{A}_o$ is similar to $\Lambda$ can be shown by performing the matrix operations defined by (B-3) using the partitioned W matrix (B-4) and examining the upper left submatrix which is given by

$$\bar{F}\, W_{11} - \bar{G}\, \bar{R}^{-1} \bar{G}^T W_{21} = W_{11} \Lambda \tag{B-10}$$

Substituting $W_{21}$ from (B-9), equation (B-10) may be rewritten as

$$[\bar{F} - \bar{G}\, \bar{R}^{-1} \bar{G}^T \bar{P}_o] W_{11} = W_{11} \Lambda \quad . \tag{B-11}$$

From (28) of Chapter II of this dissertation,

$$\bar{A}_o = \bar{F} - \bar{G}\, \bar{R}^{-1} \bar{G}^T \bar{P} \quad . \tag{B-12}$$

so that

$$\bar{A}_o\, W_{11} = W_{11} \Lambda \quad . \tag{B-13}$$

Thus, the Jordan canonical form $\Lambda$ is similar to $\bar{A}_o$ and, if $\bar{A}_o$ is stable, contains only eigenvalues with negative real parts. Consequently, the similarity transformation $W_{11}$ must also be nonsingular so that $\bar{P}_o$ exists.

# APPENDIX C

## RESPONSE ERROR LEARNING MODEL IDENTIFICATION TECHNIQUE

The identification technique discussed in this Appendix is the multivariable response error learning model originally developed by Pazdera and Pottinger [21] as modified by Spence [27], [28] to provide improved response characteristics. Since these references are readily available, only a brief, but descriptive, derivation is included in this Appendix.

The linear time-invariant system to be identified, given by (62), is of the form

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u} \quad ; \quad \underline{x}(0) = \underline{0} \quad , \tag{C-1}$$

where A and B are the unknown constant matrices, $\underline{x}$ is the state vector of known dimension n, and $\underline{u}$ is the control vector of known dimension m. The states of (C-1) are assumed to be accessible, i.e., directly measurable, and the system (C-1) is assumed to completely state controllable.

Consider a learning model of the form

$$\dot{\hat{\underline{x}}} = [\hat{A} + C_A\dot{\hat{A}}]\,\underline{x} + [\hat{B} + C_B\dot{\hat{B}}]\,\underline{u} - D\,\underline{e} \quad ; \quad \hat{x}(0) = \underline{x}_o \tag{C-2}$$

where $\hat{A}$ and $\hat{B}$ are the estimates for A and B, respectively, $C_A$ and $C_B$ are constant real symmetric non-negative matrices and $\underline{e}$ is the response error defined by

$$\underline{e} = \underline{x} - \hat{\underline{x}} \quad . \tag{C-3}$$

Note that (C-2) reduces to (C-1) and the identification is complete when

$$\hat{A} = A \quad , \dot{\hat{A}} = [0] \quad , \tag{C-4}$$

$$\hat{B} = B \quad , \quad \dot{\hat{B}} = [0] \quad , \tag{C-5}$$

and

$$\underline{e} = \underline{0} \tag{C-6}$$

The response error learning model identification technique is illustrated by the block diagram of Figure C-1.

The response error state equation can be obtained by differentiating (C-3) and substituting (C-1) and (C-2). Performing these operations, one obtains

$$\dot{\underline{e}} = D\,\underline{e} + [\Delta_A - C_A\dot{\hat{A}}]\,\underline{x} + [\Delta_B - C_B\dot{\hat{B}}]\,\underline{u} \tag{C-7}$$

where $\Delta_A$ and $\Delta_B$ are defined as

$$\Delta_A = A - \hat{A} \tag{C-8}$$

and

$$\Delta_B = B - \hat{B} \quad . \tag{C-9}$$

In order to determine how to adjust $\hat{A}$ and $\hat{B}$ in order to achieve the desired identification, consider the Liapunov function

$$2V = tr(\Delta_A^T K_A^{-1} \Delta_A) + tr(\Delta_B^T K_B^{-1} \Delta_B) + \underline{e}^T \underline{e} \tag{C-10}$$

where $K_A$ and $K_B$ are constant real symmetric positive definite matrices. Differentiating, one obtains

$$\dot{V} = -tr(\Delta_A^T K_A^{-1} \dot{\hat{A}}) - tr(\Delta_B^T K_B^{-1} \dot{\hat{B}}) + \underline{e}^T \dot{\underline{e}} \quad . \tag{C-11}$$

Noting from the trace identity that

$$\underline{e}^T \dot{\underline{e}} = tr(\underline{e}\;\dot{\underline{e}}^T) \tag{C-12}$$

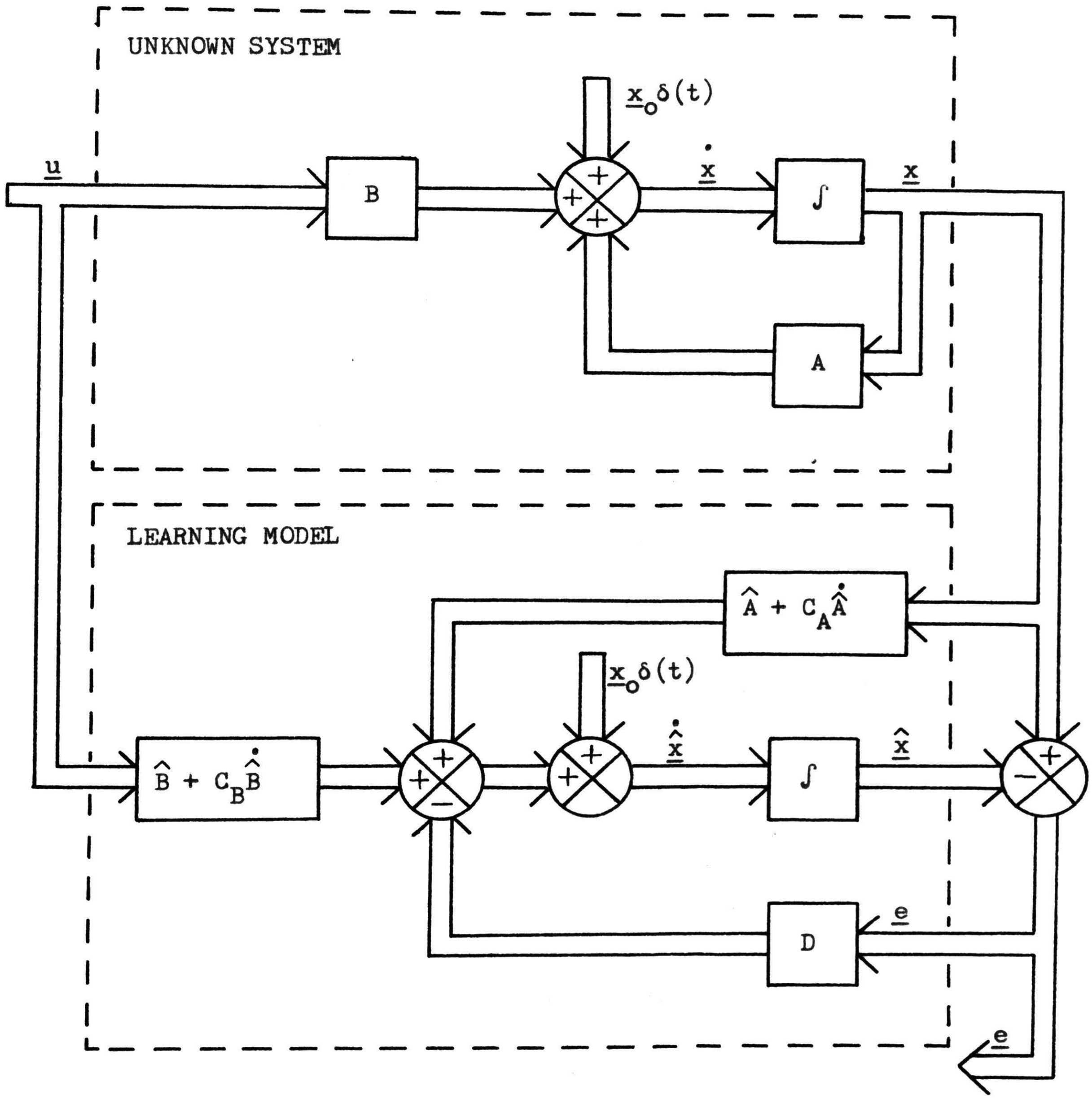



Figure C-1 Response Error Learning Model Identification Technique and substituting (C-7), (C-11) may be rewritten as

$$\dot{V} = tr(\Delta_A^T [\underline{e}\ \underline{x}^T - K_A^{-1}\dot{\hat{A}}]) + tr(\Delta_B^T [\underline{e}\ \underline{u}^T - K_B^{-1}\dot{\hat{B}}]) + \underline{e}^T D\ \underline{e} - \underline{e}^T C_A \dot{\hat{A}}\ \underline{x} - \underline{e}^T C_B \dot{\hat{B}}\ \underline{u} \quad . \tag{C-13}$$

The first two terms of (C-13) can be eliminated by setting

$$\dot{\hat{A}} = K_A \underline{e}\ \underline{x}^T \tag{C-14}$$

and

$$\dot{\hat{B}} = K_B \underline{e}\ \underline{u}^T \quad . \tag{C-15}$$

Substituting (C-14) and (C-15) into (C-13), one obtains

$$\dot{V} = \underline{e}^T D\ \underline{e} - \underline{e}^T C_A K_A^{-1} \underline{e}\ \underline{x}^T \underline{x} - \underline{e}^T C_B K_B^{-1} \underline{e}\ \underline{u}^T \underline{u} \quad . \tag{C-16}$$

Note from (C-4), (C-5), and (C-6) that if (C-14) and (C-15) are used as parameter adjustment laws, then the identification is complete when

$$\Delta_A = [0] \quad , \tag{C-17}$$

$$\Delta_B = [0] \quad , \tag{C-18}$$

and

$$\underline{e} = \underline{0} \quad . \tag{C-19}$$

Thus, from (C-10) one sees that V is positive except when the identification is complete.

Consider now the case discussed by Pazdera and Pottinger [21] in which $C_A = C_B = [0]$. For this case, $\dot{V}$ and $\dot{\underline{e}}$ are given by

$$\dot{V} = \underline{e}^T D\ \underline{e} \tag{C-20}$$

and

$$\dot{\underline{e}} = D\,\underline{e} + [\Delta_A\,\Delta_B]\begin{bmatrix}\underline{x}\\ \underline{u}\end{bmatrix} \quad . \tag{C-21}$$

Thus, $\dot{V}$ is negative definite in $\underline{e}$ if D is a stable matrix, i.e., all of the eigenvalues of D have negative real parts. Therefore, assuming that D is stable, by the second method of Liapunov $\underline{e}$ asymptotically approaches zero, i.e., $\underline{e} \to \underline{0}$ as $t \to \infty$.

The fact that $\underline{e} \to \underline{0}$ does not guarantee the identification of either A or B since one sees from (C-21) that $\underline{e} = \underline{0}$ implies only that

$$[\Delta_A\,\Delta_B]\begin{bmatrix}\underline{x}\\ \underline{u}\end{bmatrix} = \underline{0} \quad . \tag{C-22}$$

Identification of A and B is guaranteed, however, if the elements of the vector $[\underline{x}^T \;\; \underline{u}^T]^T$ are linearly independent over the required identification interval $[0, t_{id}]$ since $\underline{e} \equiv \underline{0}$ then implies that $\Delta_A \equiv [0]$ and $\Delta_B \equiv [0]$. The requirement for linear independence of the elements of $[\underline{x}^T \;\; \underline{u}^T]^T$ is the usual condition for identifiability found in the literature.

All of the above conditions remain true for the case considered by Spence [27], [28] in which $C_A > 0$ and/or $C_B > 0$. However, as seen from (C-16), the use of positive definite or even positive semi-definite matrices improves the asymptotic response of $\underline{e}$. As seen from (C-7), the use of these matrices also alters the error response of the learning model leading to an improved response in the identification of A and B in many cases.

Even though identification is guaranteed for the response error learning model, the identification time is highly dependent on the settings of $C_A$, $C_B$, $K_A$, and $K_B$, the initial condition $\underline{x}(0)$, the frequency content and amplitudes of the control signal $\underline{u}(t)$ as well as the values of A and B themselves. Furthermore, the identification time is not necessarily reduced when the adjustment law gains are increased, i.e., when the elements of the matrices $K_A$ and $K_B$ are increased in magnitude. For any combination of $\underline{x}(0)$ and $\underline{u}(t)$, there is also a combination of $C_A$, $C_B$, $K_A$, and $K_B$ which yields the minimum identification time. However, at present there is no criterion for the selection of these matrices.

# APPENDIX D

## EQUATION ERROR LEARNING MODEL IDENTIFICATION TECHNIQUE

The identification technique discussed in this Appendix is a new multivariable extension of the single input/output equation error learning model originally developed by Lion[17]. Since the rationale behind the development of the multivariable learning model is very similar to that of the single input/output learning model, for which the details are given in [17], only a brief, but descriptive derivation is included here. Unlike the work of Lion, however, the effect of non-zero initial conditions is included in this Appendix.

The linear time-invariant system to be identified, given by (62), is of the form

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u} \quad ; \quad \underline{x}(0) = \underline{x}_0 \tag{D-1}$$

where A and B are the unknown constant matrices, $\underline{x}$ is the state vector of known dimension n, and $\underline{u}$ is the control vector of known dimension m. The states of (D-1) are assumed to be accessible, i.e., directly measurable, and the system (D-1) is assumed to completely state controllable. Note also that since A and B are constant matrices, term by term integration of (D-1) yields

$$\underline{x}(t) = A\,\tilde{\underline{x}}(t) + B\,\tilde{\underline{u}}(t) + \underline{x}_0. \tag{D-2}$$

where

$$\tilde{\underline{x}}(t) = \int_0^t \underline{x}(t)\,dt + \tilde{\underline{x}}(0) \quad ; \quad \tilde{\underline{x}}(0) = \underline{0} \tag{D-3}$$

and

$$\underline{\tilde{u}}(t) = \int_0^t \underline{u}(t)\, dt + \underline{\tilde{u}}(0) \quad ; \quad \underline{\tilde{u}}(0) = \underline{0} \quad . \tag{D-4}$$

Consider a learning model of the form of (D-2), i.e.,

$$\underline{\hat{x}}(t) = \hat{A}\, \underline{\tilde{x}}(t) + \hat{B}\, \underline{\tilde{u}}(t) + \underline{x}_o \tag{D-5}$$

where $\hat{A}$ and $\hat{B}$ are the estimates for A and B, respectively, and the equation error $\underline{e}$ is defined as

$$\underline{e} = \underline{x} - \underline{\hat{x}} \quad . \tag{D-6}$$

Note that (D-5) reduces to (D-2) and the identification is complete when

$$\hat{A} = A \tag{D-7}$$

and

$$\hat{B} = B \quad . \tag{D-8}$$

The equation error learning model identification technique is illustrated by the block diagram of Figure D-1.

Substituting (D-5) into (D-6), the equation error becomes

$$\underline{e} = \underline{x} - \hat{A}\, \underline{\tilde{x}} - \hat{B}\, \underline{\tilde{u}} - \underline{x}_o \quad . \tag{D-9}$$

From (D-9) it is evident that the equation error learning model (D-5) is easily implemented and avoids the requirement that $\underline{\dot{x}}$ be accessible, which would have arisen had a learning model of the form analogous to (D-1) been utilized.

Substituting (D-2) into (D-9), the equation error becomes

$$\underline{e} = \Delta_A \underline{\tilde{x}} + \Delta_B \underline{\tilde{u}} \tag{D-10}$$

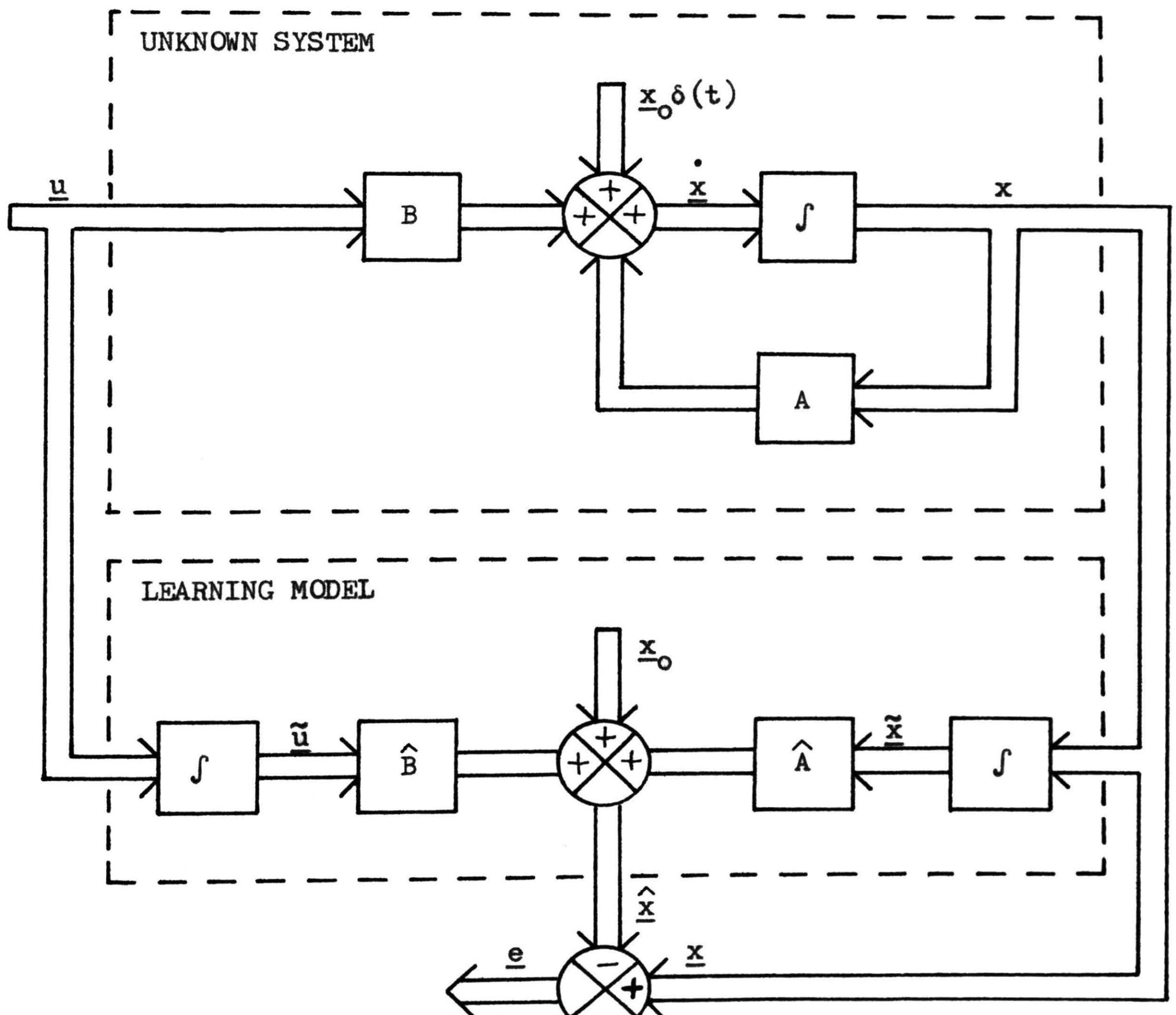



Figure D-1 Equation Error Learning Model Identification Technique where the parameter misalignment matrices $\Delta_A$ and $\Delta_B$ are defined as

$$\Delta_A = A - \hat{A} \tag{D-11}$$

and

$$\Delta_B = B - \hat{B} \quad . \tag{D-12}$$

In order to determine how to adjust A and B in order to achieve the desired identification, consider the Liapunov function

$$2V = \mathrm{tr}(\Delta_A^T K_A^{-1} \Delta_A) + \mathrm{tr}(\Delta_B^T K_B^{-1} \Delta_B) \tag{D-13}$$

where $K_A$ and $K_B$ are constant real symmetric positive definite matrices. Note that V is positive unless the identification is complete, in which case V = 0. Differentiating (D-13) one obtains

$$\dot{V} = -\,\mathrm{tr}(\Delta_A^T K_A^{-1} \dot{\hat{A}}) - \mathrm{tr}(\Delta_B^T K_B^{-1} \dot{\hat{B}}) \quad . \tag{D-14}$$

Consider now parameter adjustment laws of the same form as those used in the response error learning model identification technique (Appendix C), i.e.,

$$\dot{\hat{A}} = K_A \underline{e}\, \underline{\tilde{x}}^T \tag{D-15}$$

and

$$\dot{\hat{B}} = K_B \underline{e}\, \underline{\tilde{u}}^T \tag{D-16}$$

Substituting (D-15) and (D-16) into (D-14), one obtains

$$\dot{V} = -\,\mathrm{tr}(\Delta_A^T \underline{e}\, \underline{\tilde{x}}^T) - \mathrm{tr}(\Delta_B^T \underline{e}\, \underline{\tilde{u}}^T) \quad , \tag{D-17}$$

or, using the trace identity

$$\dot{V} = \underline{e}^T[\Delta_A\tilde{\underline{x}} + \Delta_B\tilde{\underline{u}}] \quad . \tag{D-18}$$

Substituting (D-10) into (D-18) yields

$$\dot{V} = -\underline{e}^T\underline{e} \quad . \tag{D-19}$$

Thus $\dot{V}$ is negative unless $\underline{e} = \underline{0}$, in which case $\dot{V} = 0$. Therefore, from the definition of the Liapunov function given by (D-13), one sees that the net misalignment between the system and learning model parameters must be decreasing whenever $\underline{e} \neq \underline{0}$.

From (D-10) it is evident that $\underline{e} = \underline{0}$ implies only that

$$[\Delta_A \;\; \Delta_B] \begin{bmatrix} \tilde{\underline{x}} \\ \tilde{\underline{u}} \end{bmatrix} = \underline{0} \tag{D-20}$$

and it is possible for $\underline{e}$ to be a null vector without the parameter misalignments being zero. However, if the elements of the vector $[\tilde{\underline{x}}^T \;\; \tilde{\underline{u}}^T]^T$ are linearly independent over the required identification interval $[0,t_{id}]$, then $\underline{e} \equiv 0$ implies that $\Delta_A \equiv [0]$ and $\Delta_B \equiv [0]$ and, conversely, $\Delta_A \not\equiv [0]$ and/or $\Delta_B \not\equiv [0]$ implies that $\underline{e} \not\equiv 0$. Therefore, since the Liapunov function is positive definite in the misalignments $\Delta_A$ and $\Delta_B$ and its derivative is negative definite in $\underline{e}$, then $\Delta_A$ and $\Delta_B$ must asymptotically approach a null, i.e., $\Delta_A \rightarrow [0]$ and $\Delta_B \rightarrow [0]$ as $t \rightarrow \infty$, if the elements of $[\tilde{\underline{x}}^T \;\; \tilde{\underline{u}}^T]^T$ are linearly independent.

As has been shown previously, the linear independence of the elements of the vector $[\underline{x}^T \;\; \underline{u}^T]^T$ is the sufficient condition for identifiability when using either the parameter estimation equation given by

(65) or the response error learning model identification technique discussed in Appendix C. It may be possible to prove that the linear independence of the elements of $[\underline{x}^T \ \ \underline{u}^T]^T$ implies the linear independence of the elements of $[\tilde{\underline{x}}^T \ \ \tilde{\underline{u}}^T]^T$ and is, therefore, also a sufficient condition for identification for the equation error learning model technique. However, as no proof is currently available, the above statement is at present only a hypothesis.

Lion [17] has shown for the single input/output equation error that the state variable filters need not be integrators. This is also true for the multivariable learning model described herein, but the complexity of the learning model is, of course, increased.

Consider, for example, multiplication of the LaPlace transform of (D-1), given as

$$s\ \underline{x}(s) = A\ \underline{x}(s) + B\ \underline{u}(s) + \underline{x}_o \quad , \tag{D-21}$$

by the scalar function g(s), which yields

$$g(s)\ s\ \underline{x}(s) = g(s)\ A\ \underline{x}(s) + g(s)\ B\ \underline{u}(s) + g(s)\ \underline{x}_o \quad . \tag{D-22}$$

It is evident that (D-22) may also be written as

$$s\ \tilde{\underline{x}}(s) = A\ \tilde{\underline{x}}(s) + B\ \tilde{\underline{u}}(s) + g(s)\ \underline{x}_o \tag{D-23}$$

where

$$\tilde{\underline{x}}(s) = g(s)\ \underline{x}(s) \tag{D-24}$$

and

$$\tilde{\underline{u}}(s) = g(s)\ \underline{u}(s) \quad . \tag{D-25}$$

The inverse LaPlace transform of (D-23) yields

$$\dot{\tilde{\underline{x}}}(t) = A\ \tilde{\underline{x}}(t) + B\ \tilde{\underline{u}}(t) + g(t)\ \underline{x}_o \tag{D-26}$$

where

$$\underline{\tilde{x}}(t) = L^{-1}[\underline{\tilde{x}}(s)] \tag{D-27}$$

and

$$\underline{\tilde{u}}(t) = L^{-1}[\underline{\tilde{u}}(s)] \quad . \tag{D-28}$$

The learning model equivalent to (D-5) may now be written as

$$\underline{\hat{x}}(t) = \hat{A}\ \underline{\tilde{x}}(t) + \hat{B}\ \underline{\tilde{u}}(t) + g(t)\ \underline{x}_o \tag{D-29}$$

where A and B are the estimates for A and B, respectively, and the equation error equivalent to (D-6) is defined as

$$\underline{e} = \dot{\underline{\tilde{x}}} - \underline{\hat{x}} \quad . \tag{D-30}$$

The proof that $\hat{A} \rightarrow A$ and $\hat{B} \rightarrow B$ as $t \rightarrow \infty$ is identical to that given previously.

It has been shown that if the filtering of each element of both $\underline{x}$ and $\underline{u}$ is identical and the initial condition term in the model (D-29) reflects the impulse response of the filter, then a state variable filter other than an integrator may be used.

Even though identification is guaranteed for the equation error learning model, the identification time is highly dependent on the settings of $K_A$ and $K_B$, the initial condition $\underline{x}(0)$, the frequency content and amplitudes of the control signal $\underline{u}(t)$ as well as the values of A and B themselves. Furthermore, the identification time is not necessarily reduced when the adjustment law gains are increased, i.e., when the elements of the matrices $K_A$ and $K_B$ are increased in magnitude. For any combination of $\underline{x}(0)$ and $\underline{u}(t)$, there is also a combination of

$K_A$ and $K_B$ which yields the minimum identification time. At present, however, no criterion for the selection of these matrices can be given.

# APPENDIX E

## GENERALIZED EQUATION ERROR LEARNING MODEL IDENTIFICATION TECHNIQUE

The equation error learning model identification technique discussed in Appendix D can be further generalized by defining additional equation error vectors. The parameter adjustment laws then become functions of these additional equation error vectors. The net result is a reduction in the identification time. A brief derivation of the generalized equation error learning model is included in this Appendix. This derivation follows closely that of the Appendix D.

The linear time-invariant system to be identified is of the form

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u} \quad ; \quad \underline{x}(0) = \underline{x}_o \qquad \text{(E-1)}$$

where A and B are the unknown constant matrices, $\underline{x}$ is the state vector of known dimension n, and $\underline{u}$ is the control vector of known dimension m. The states of (E-1) are assumed to be accessible and the system is assumed to be completely state controllable.

Since A and B are constant matrices, one can write after i term by term integrations of (E-1)

$$\dot{\tilde{\underline{x}}}_i = A\,\tilde{\underline{x}}_i + B\,\tilde{\underline{u}}_i + (t^{i-1}/(i-1)!)\,\underline{x}_o \qquad \text{(E-2)}$$

where

$$\tilde{\underline{x}}_i = \int_0^t \overset{i}{\cdots} \int_0^t x(t)\,(dt)^i \quad , \qquad \text{(E-3)}$$

$$\tilde{\underline{u}}_i = \int_0^t \overset{i}{\cdots} \int_0^t u(t)\,(dt)^i \quad , \qquad \text{(E-4)}$$

and $\tilde{\underline{x}}_j(0) = \underline{0}$ and $\tilde{\underline{u}}_j(0) = \underline{0}$ , $j = 1, 2, \ldots, i$ .

Consider k learning models of the form

$$\underline{\hat{x}}_i = \hat{A}\,\underline{\tilde{x}}_i + \hat{B}\,\underline{\tilde{u}}_i + (t^{i-1}/(i-1)!)\,\underline{x}_o \tag{E-5}$$

and define the i'th equation error as

$$\underline{e}_i = \underline{\dot{\tilde{x}}}_i - \underline{\hat{x}}_i \quad . \tag{E-6}$$

Notice that (E-5) reduces to (E-2) when

$$\hat{A} = A \tag{E-7}$$

and

$$\hat{B} = B \tag{E-8}$$

and that the above model reduces to that given in Appendix D when $i = 1$ .

Substituting (E-2) into (E-5), the i'th equation error becomes

$$\underline{e}_i = \underline{\dot{\tilde{x}}}_i - \hat{A}\,\underline{\tilde{x}}_i - \hat{B}\,\underline{\tilde{u}}_i - (t^{i-1}/(i-1)!)\,\underline{x}_o \quad . \tag{E-9}$$

The generalized equation error learning model identification technique based on (E-9) is illustrated in the block diagram of Figure E-1 for $k = 2$.

Substituting (E-2) into (E-9) the i'th equation error becomes

$$\underline{e}_i = \Delta_A \underline{\tilde{x}}_i + \Delta_B \underline{\tilde{u}}_i \tag{E-10}$$

where $\Delta_A$ and $\Delta_B$ are defined as

$$\Delta_A = A - \hat{A} \tag{E-11}$$

and

$$\Delta_B = B - \hat{B} \quad . \tag{E-12}$$

In order to determine how to adjust $\hat{A}$ and $\hat{B}$ in order to achieve the

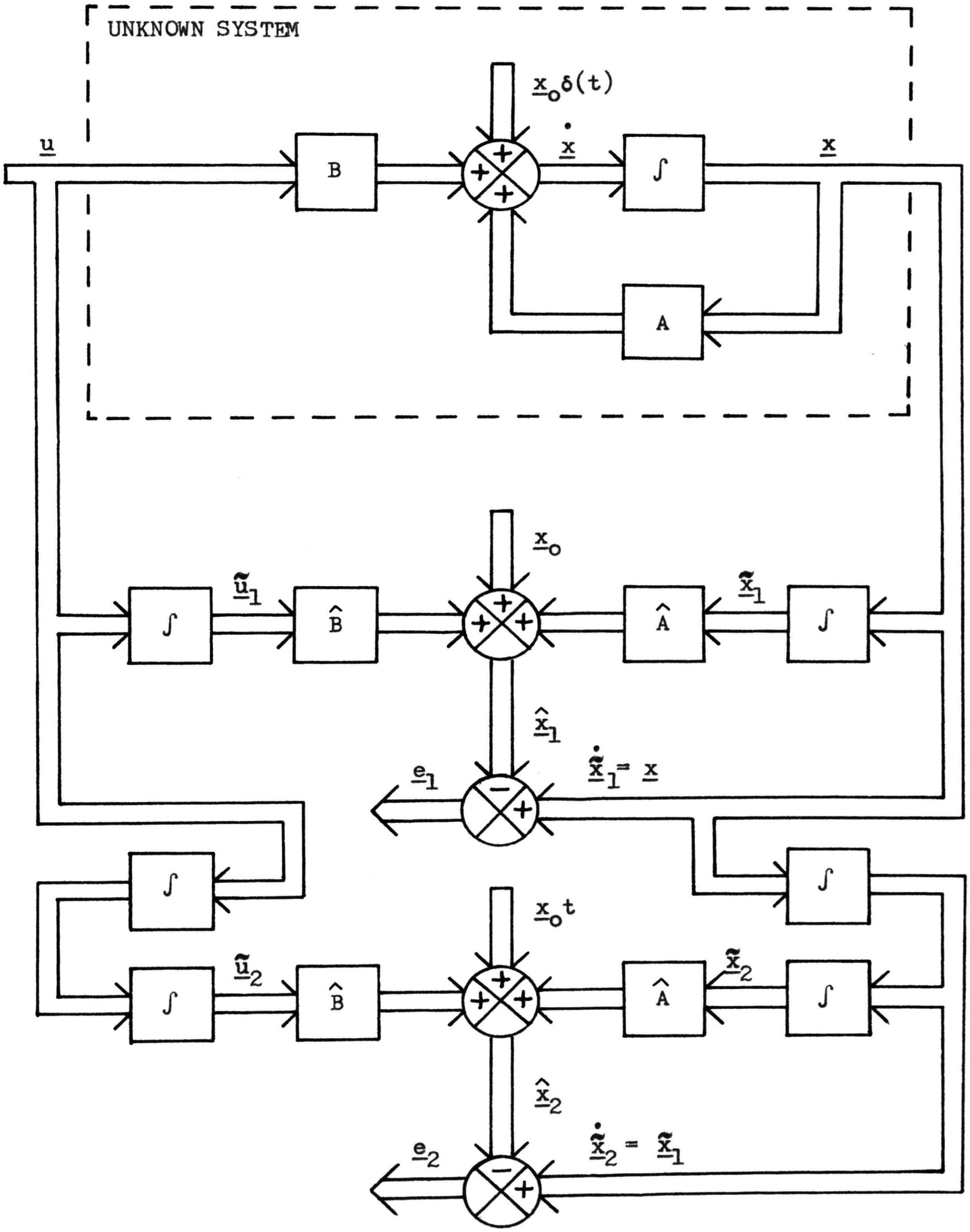



Figure E-1. Generalized Equation Error Learning Model Identification Technique For k = 2 .

desired identification, consider the Liapunov function

$$2V = tr(\Delta_A^T K_A^{-1} \Delta_A) + tr(\Delta_B^T K_B^{-1} \Delta_B) \tag{E-13}$$

where $K_A$ and $K_B$ are constant real symmetric positive definite matrices. Note that V is positive unless the identification is complete, in which case V = 0. Differentiating (E-13) one obtains

$$\dot{V} = - tr(\Delta_A^T K_A^{-1} \dot{\hat{A}}) - tr(\Delta_B^T K_B^{-1} \dot{\hat{B}}) \quad . \tag{E-14}$$

Assume now that parameter adjustment laws of the form

$$\dot{\hat{A}} = K_A \sum_{i=1}^{k} K_i \underline{e}_i \tilde{\underline{x}}_i^T \tag{E-15}$$

and

$$\dot{\hat{B}} = K_B \sum_{i=1}^{k} K_i \underline{e}_i \tilde{\underline{u}}_i^T \tag{E-16}$$

are utilized where the $K_i$ matrices are real symmetric positive definite. Substituting (E-15) and (E-16) into (E-14), one obtains

$$\dot{V} = - \sum_{i=1}^{k} [tr(\Delta_A^T K_i \underline{e}_i \tilde{\underline{x}}_i^T) + tr(\Delta_B^T K_i \underline{e}_i \tilde{\underline{u}}_i^T)] \tag{E-17}$$

or, using the trace identity,

$$\dot{V} = - \sum_{i=1}^{k} \underline{e}_i^T K_i [\Delta_A \tilde{\underline{x}}_i + \Delta_B \tilde{\underline{u}}_i] \quad . \tag{E-18}$$

Substituting (E-10) into (E-18) yields

$$\dot{V} = - \sum_{i=1}^{k} \underline{e}_i^T K_i \underline{e}_i \tag{E-19}$$

or

$$\dot{V} = -\underline{e}^T \begin{bmatrix} K_1 . & & 0 \\ & . & \\ 0 & & . K_k \end{bmatrix} \underline{e} \tag{E-20}$$

where the generalized equation error $\underline{e}$ is defined as

$$\underline{e} = [\underline{e}_1^T \ldots \underline{e}_k^T]^T \quad . \tag{E-21}$$

Thus $\dot{V}$ is negative unless $\underline{e} = \underline{0}$, in which case $\dot{V} = 0$. Therefore, from the definition of the Liapunov function given by (E-13), one sees that the net misalignment between the system and learning model parameters must be decreasing whenever $\underline{e} \neq \underline{0}$.

From (E-10) it is evident that $\underline{e} = \underline{0}$ implies only that

$$[\Delta_A \;\; \Delta_B] \begin{bmatrix} \tilde{\underline{x}}_i \\ \\ \tilde{\underline{u}}_i \end{bmatrix} = \underline{0} \tag{E-22}$$

for $i = 1, \ldots, k$ and it is therefore possible for $\underline{e}$ to be a null vector without the parameter misalignments being zero. However, if for any value of $i$ the elements of the vector $[\tilde{\underline{x}}_i^T \;\; \tilde{\underline{u}}_i^T]^T$ are linearly independent over the identification interval $[0,t_{id}]$, then $\underline{e} \equiv \underline{0}$ implies that $\Delta_A \equiv [0]$ and $\Delta_B \equiv [0]$ and, conversely, $\Delta_A \not\equiv [0]$ and/or $\Delta_B \not\equiv [0]$ implies that $\underline{e} \not\equiv \underline{0}$. Therefore, since the Liapunov function is positive definite in the parameter misalignments $\Delta_A$ and $\Delta_B$ and its derivative is negative definite in $\underline{e}$, $\Delta_A$ and $\Delta_B$ must asymptotically approach zero, i.e., $\Delta_A \rightarrow [0]$ and $\Delta_B \rightarrow [0]$ as $t \rightarrow \infty$ if the elements of $[\tilde{\underline{x}}_i^T \;\; \tilde{\underline{u}}_i^T]^T$ are linearly independent for some value of i.

The advantage of the generalized equation error learning model discussed herein as compared to the learning model discussed in Appendix D can be seen by examining the effect that the use of k learning models has on the Liapunov function and its derivative. As seen from (E-13), the Liapunov function is not affected by the number of learning models being utilized. However, recalling that the generalized model reduces to that of Appendix D for the case of $k = 1$, one sees from (E-19) that the use of $k > 1$ learning models makes the Liapunov function derivative more negative, which should increase the rate of convergence of the parameter misalignments since the Liapunov function itself has not been altered.

Of course, the increase in complexity of the learning model is not worthwhile unless each additional error vector is distinct, i;e., it does not contain either the same elements or scalar multiples of the same elements as any of the other error vectors. This is true since if each additional error vector is not distinct, then the effect is the same as could be obtained by a simple change to the parameter adjustment law gains. From (E-10) one sees that whether or not the k error vectors are distinct is directly dependent upon the characteristics of the filtered vectors, i.e., $\underline{\tilde{x}}_i$ and $\underline{\tilde{u}}_i$ as given by (E-3) and (E-4). This is, of course, the reason that a different state variable filter is defined for each value of i.

Lion [17] has shown for the single input/output identification problem that if there are p parameters to be identified, then the identification time must decrease as the adjustment law gains are increased if p equation errors are used. Lion [17] further states that for this to be true, the input signal must also have a given frequency content, namely, the number of frequencies must be greater than or equal to p/2 or the input signal must be noisy.

For the state space representation given by (E-1), there are n(n+m) unknown parameters to be identified. If the results obtained by Lion [17] for the single input/output system also hold for the multivariable input/output system discussed herein, then (n+m) learning models of the form (E-5) would be required to make the identification time decrease as the adjustment law gains are increased. As can be seen from Figure E-1, this expansion vastly complicates mechanization of the learning model identification technique. The Lion requirement on input signal frequency content appears to be equivalent to the condition that the elements of $[\underline{x}^T \quad \underline{u}^T]^T$ be linearly independent over the identification time interval.

# APPENDIX F

## COMBINED ERROR LEARNING MODEL IDENTIFICATION TECHNIQUE

The response error learning model identification technique discussed in Appendix C and the equation error learning model identification technique discussed in Appendix D can be combined into a single identification technique. The net result is a reduction in the identification time. A brief derivation of the combined error learning model is given in this Appendix. This derivation follows closely that of Appendices C and D.

The linear time-invariant system to be identified, given by (62), is of the form

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u} \quad ; \quad \underline{x}(0) = \underline{x}_o \tag{F-1}$$

where A and B are the unknown constant matrices, $\underline{x}$ is the state vector of known dimension n, and $\underline{u}$ is the control vector of known dimension m. The states of (F-1) are assumed to be accessible and the system is assumed to be completely state controllable. Notice also that since A and B are constant matrices, term by term integration of (F-1) yields

$$\underline{x}(t) = A\,\tilde{\underline{x}}(t) + B\,\tilde{\underline{u}}(t) + \underline{x}_o \tag{F-2}$$

where

$$\tilde{\underline{x}}(t) = \int_0^t \underline{x}(t)\,dt + \tilde{\underline{x}}(0) \quad ; \quad \tilde{\underline{x}}(0) = \underline{0} \tag{F-3}$$

and

$$\tilde{\underline{u}}(t) = \int_0^t \underline{u}(t)\,dt + \tilde{\underline{u}}(0) \quad ; \quad \tilde{\underline{u}}(0) = \underline{0} \tag{F-4}$$

Consider now learning models of the form

$$\dot{\hat{\underline{x}}}_1 = \hat{A}\,\underline{x} + \hat{B}\,\underline{u} - D\,\underline{e}_1 \tag{F-5}$$

and

$$\hat{x}_2 = \hat{A}\,\underline{\tilde{x}} + \hat{B}\,\underline{\tilde{u}} + \underline{x}_o \qquad \text{(F-6)}$$

where the response error is defined by

$$e_1 = \underline{x} - \underline{\hat{x}}_1 \qquad \text{(F-7)}$$

and the equation error is defined by

$$e_2 = \underline{x} - \underline{\hat{x}}_2 \quad . \qquad \text{(F-8)}$$

Note that (F-5) reduces to (F-1) and (F-6) reduces to (F-2) and the identification is complete when

$$\hat{A} = A \quad , \qquad \text{(F-9)}$$

$$\hat{B} = B \quad , \qquad \text{(F-10)}$$

and

$$\underline{e}_1 = \underline{0} \quad . \qquad \text{(F-11)}$$

The combined error learning model technique is illustrated in Figure F-1. It is evident that both the response error learning model identification technique and the equation error learning model identification technique avoid the requirement that $\underline{\dot{x}}$ be accessible.

The response error state equation is obtained by differentiating (F-7) and substituting (F-1) and (F-5). The equation error is obtained by substituting (F-2) and (F-6) into (F-8). Performing these operations, one obtains

$$\underline{\dot{e}}_1 = D\,\underline{e}_1 + \Delta_A\underline{x} + \Delta_B\underline{u} \qquad \text{(F-12)}$$

and

$$\underline{e}_2 = \Delta_A\underline{\tilde{x}} + \Delta_B\underline{\tilde{u}} \quad . \qquad \text{(F-13)}$$

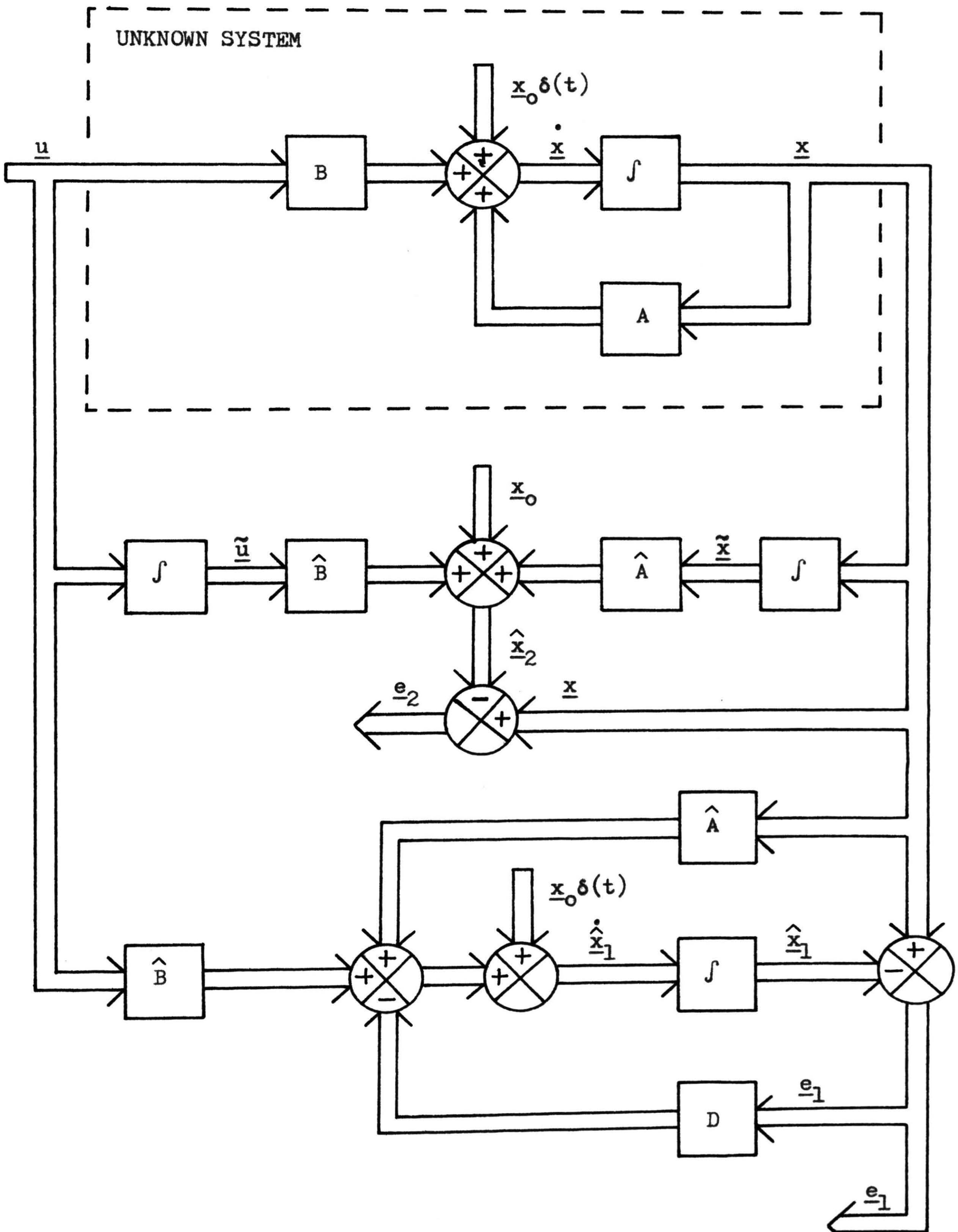



Figure F-1. Combined Error Learning Model Identification Technique where $\Delta_A$ and $\Delta_B$ are defined as

$$\Delta_A = A - \hat{A} \tag{F-14}$$

and

$$\Delta_B = B - \hat{B} \quad . \tag{F-15}$$

In order to determine how to adjust $\hat{A}$ and $\hat{B}$ in order to achieve the desired identification, consider the Liapunov function

$$2V = tr(\Delta_A^T K_A^{-1} \Delta_A) + tr(\Delta_B^T K_B^{-1} \Delta_B) + \underline{e}_1^T \underline{e}_1 \tag{F-16}$$

where $K_A$ and $K_B$ are constant real symmetric positive definite matrices. Note that V is positive unless the identification is complete, in which case $V = 0$. Differentiating (F-16) one obtains

$$\dot{V} = - tr(\Delta_A^T K_A^{-1} \dot{\hat{A}}) - tr(\Delta_B^T K_B^{-1} \dot{\hat{B}}) + \underline{e}_1^T \dot{\underline{e}}_1 \quad . \tag{F-17}$$

Noting that

$$\underline{e}_1^T \dot{\underline{e}}_1 = \underline{e}_1^T D \, \underline{e}_1 + \underline{e}_1^T \Delta_A \underline{x} + \underline{e}_1^T \Delta_B \underline{u} \tag{F-18}$$

and using the trace identity, (F-17) may be rewritten as

$$\dot{V} = \underline{e}_1^T D \, \underline{e}_1 - tr(\Delta_A^T [K_A^{-1} \dot{\hat{A}} - \underline{e}_1 \underline{x}^T]) - tr(\Delta_B^T [K_B^{-1} \dot{\hat{B}} - \underline{e}_1 \underline{u}^T]) \quad . \tag{F-19}$$

Assume now that parameter adjustment laws of the form

$$\dot{\hat{A}} = K_A [\underline{e}_1 \underline{x}^T + K \, \underline{e}_2 \tilde{\underline{x}}^T] \tag{F-20}$$

and

$$\dot{\hat{B}} = K_B[\underline{e}_1\underline{u}^T + K\ \underline{e}_2\tilde{\underline{u}}^T] \tag{F-21}$$

are utilized, where K is a constant real symmetric positive definite matrix. Substituting (F-20) and (F-21) into (F-19), one obtains

$$\dot{V} = \underline{e}_1^T D\ \underline{e}_1 - \underline{e}_2^T K[\Delta_A \tilde{\underline{x}} + \Delta_B \tilde{\underline{u}}] \quad . \tag{F-22}$$

Substituting (F-13), one gets

$$\dot{V} = \underline{e}_1^T D\ \underline{e}_1 - \underline{e}_2^T K\ \underline{e}_2 \tag{F-23}$$

or

$$\dot{V} = \underline{e}^T \begin{bmatrix} D & 0 \\ 0 & -K \end{bmatrix} \underline{e} \tag{F-24}$$

where the combined error vector $\underline{e}$ is defined as

$$\underline{e} = [\underline{e}_1^T \quad \underline{e}_2^T]^T \quad . \tag{F-25}$$

Thus, $\dot{V}$ is negative definite in $\underline{e}$ if D is a negative definite matrix. Therefore, from the definition of the Liapunov function given by (F-16) one sees that the net parameter misalignment, as well as the error $\underline{e}_1$, must be decreasing whenever $\underline{e} \neq \underline{0}$ .

From (F-12) and (F-13) it is evident that $\underline{e} = \underline{0}$ implies only that

$$\begin{bmatrix} \Delta_A & 0 & \Delta_B & 0 \\ 0 & \Delta_A & 0 & \Delta_B \end{bmatrix} \begin{bmatrix} \underline{x} \\ \tilde{\underline{x}} \\ \underline{u} \\ \tilde{\underline{u}} \end{bmatrix} = \underline{0} \tag{F-26}$$

and it is possible for $\underline{e}$ to be a null vector without the parameter misalignments being zero. However, if the elements of either $[\underline{x}^T \ \ \underline{u}^T]^T$ or $[\tilde{\underline{x}}^T \ \ \tilde{\underline{u}}^T]^T$ are linearly independent over the identification interval $[0, t_{id}]$, then $\underline{e} \equiv \underline{0}$ implies that $\Delta_A \equiv [0]$ and $\Delta_B \equiv [0]$ and, conversely, $\Delta_A \not\equiv [0]$ and/or $\Delta_B \not\equiv [0]$ implies that $\underline{e} \not\equiv \underline{0}$. Therefore, since the Liapunov function is positive definite in the parameter misalignments $\Delta_A$ and $\Delta_B$ and the response error $\underline{e}_1$ and the Liapunov derivative is negative definite in $\underline{e}$, $\Delta_A$ and $\Delta_B$ must asymptotically approach zero, i.e., $\Delta_A \rightarrow [0]$ and $\Delta_B \rightarrow [0]$ as $t \rightarrow \infty$, if either of the vectors $[\underline{x}^T \ \ \underline{u}^T]^T$ or $[\tilde{\underline{x}}^T \ \ \tilde{\underline{u}}^T]^T$ has linearly independent elements.

The advantage of the combined error learning model discussed herein as compared to the response error learning model discussed in Appendix C can be seen by examining the effect that the addition of the equation error learning model has on the Liapunov function and its derivative. As seen from (F-16), the Liapunov function is not affected by the addition of the equation error learning model. The derivative of the Liapunov function has, however, been made more negative by the addition of the equation error learning model, which should increase the rate of convergence of the parameter misalignments since the Liapunov function itself has not been altered.

The increase in learning model complexity due to the addition of the equation error learning model is not worthwhile unless the two error vectors are distinct, i.e., they do not contain either the same elements or scalar multiples of the same elements. This is true since if the equation error is not distinct from the response error, then the effect is the same as could be obtained by a simple change to the negative definite matrix D. Since the response error is governed by (F-12), while the equation error (F-13) is directly dependent upon the characteristics of $\tilde{\underline{x}}$ and $\tilde{\underline{u}}$, the choice of integrators for the state variable filters should allow the two error vectors to be distinct.

# APPENDIX G

## DISCRETE TIME PARAMETER IDENTIFICATION

### G-1. Single Point Estimation

The discrete time-invariant system whose parameters are to be identified, as given by (86), is

$$\underline{x}_p(m+1) = \Phi^o_{ppo}\underline{x}_p(m) + \Phi^o_{peo}\underline{x}_e(m) \tag{G-1}$$

where $\underline{x}_p$ and $\underline{x}_e$ are the n dimensional state vectors of the Pursuer and the Evader, respectively, and the m'th sample occurs at $t = m\Delta T$. Defining $\underline{x}$ and $\Phi^o_{po}$ as

$$\underline{x} = [\underline{x}_p^T \quad \underline{x}_e^T] \tag{G-2}$$

and

$$\Phi^o_{po} = [\Phi^o_{ppo} \quad \Phi^o_{peo}] \tag{G-3}$$

the system equation may be written in more convenient form as

$$\underline{x}_p(m+1) = \Phi^o_{po}\underline{x}(m) \quad . \tag{G-4}$$

The parameter identification equation for $\Phi^o_{po}$ is derived in Chapter IV based on the initial condition plus 2n additional equally spaced samples. However, any 2n+1 successive equally spaced samples are sufficient. Therefore, the minimum mean square error estimate for $\Phi^o_{po}$ at sample time $m\Delta T$, based on the last 2n+1 samples, may be written from (98) and (99) as

$$\hat{\Phi}^o_{po}(m) = [X_p(m) \; X^T(m-1)] \; M(m-1) \tag{G-5}$$

where

$$M(m-1) = [X(m-1)\ X^T(m-1)]^{-1} \tag{G-6}$$

and $X_p(m)$ and $X(m-1)$ are defined as

$$X_p(m) = [\underline{x}_p(m) \ \ldots \ \underline{x}_p(m-2n+2) \quad \underline{x}_p(m-2n+1)] \tag{G-7}$$

and

$$X(m-1) = [\underline{x}(m-1) \ \ldots \ \underline{x}(m-2n+1) \quad \underline{x}(m-2n)] \quad . \tag{G-8}$$

Although $X(m-1)$ as defined by (G-8) is a square matrix since it includes 2n successive samples of $\underline{x}$, this restriction is not inherent in the identification equations given by (G-5) and (G-6). Thus, more than 2n samples of $\underline{x}_p$ and $\underline{x}$ could be included in $X_p$ and $X$, respectively.

If the measurements of $\underline{x}_p$ and $\underline{x}_e$ are perfect, then inclusion of additional samples in the identification equations is superfluous. However, depending on the magnitude of the states of $\underline{x}_p$ and $\underline{x}_e$ as well as the sampling interval, even small measurement errors can cause significant errors in the estimate for $\Phi^o_{po}$ when based on only 2n+1 samples. Obviously, the minimum mean square error estimate can be improved through the use of additional samples. Since measurement error is inherent in any practical system, the estimate should always include as many samples as practical.

Consider the addition of one additional sample at time $(m+1)\Delta T$. The discrete time estimate may now be written as

$$\hat{\Phi}^o_{po}(m+1) = [\underline{x}_p(m+1) \quad X_p(m)]\ [\underline{x}(m) \quad X(m-1)]^T M(m) \tag{G-9}$$

where

$$M(m) = [\,[\underline{x}(m) \quad X(m-1)]\,[\underline{x}(m) \quad X(m-1)]^T]^{-1} \quad . \tag{G-10}$$

## G-2. Recursive Estimation

One usually desires to make a new estimate of $\Phi^o_{po}$ for each new set of samples. However, if equations of the form of (G-9) and (G-10) are employed, then in order to make successive estimates: (1) all samples must be saved, (2) matrix multiplication must be performed for each sample using matrices of ever expanding dimension, and (3) matrix inversion must be performed for each sample. Luckily, a multivariable recursive estimation equation, which eliminates all of the above requirements, can be derived in a manner similar to that utilized by Lee [16] for single input/output systems identification problem.

By performing the indicated operations, equations (G-9) and (G-10) may be rewritten as

$$\hat{\Phi}^o_{po}(m+1) = [\underline{x}_p(m+1)\ \underline{x}^T(m) + X_p(m)\ X^T(m-1)]\ M(m) \tag{G-11}$$

and

$$M(m) = [\underline{x}(m)\ \underline{x}^T(m) + X(m-1)\ X^T(m-1)]^{-1} \quad . \tag{G-12}$$

Using (G-6) in (G-12), M(m) may be rewritten in terms of M(m-1) as

$$M(m) = [\underline{x}(m)\ \underline{x}^T(m) + M^{-1}(m-1)]^{-1} \tag{G-13}$$

or, applying the matrix inversion lemma,

$$\begin{aligned} M(m) &= M(m-1) \\ &\quad - M(m-1)\ \underline{x}(m)\ [1 + \underline{x}^T(m)\ M(m-1)\ \underline{x}(m)]^{-1}\underline{x}^T(m)\ M(m-1) \ . \end{aligned} \tag{G-14}$$

Substituting M(m) into (G-11) one gets

$$\hat{\Phi}^{o}_{po}(m+1) = \underline{x}_p(m+1)\underline{x}^T(m)M(m-1) \tag{G-15}$$

$$- \underline{x}_p(m+1)\underline{x}^T(m)M(m-1)\underline{x}(m)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$+ X_p(m)X^T(m-1)M(m-1)$$

$$- X_p(m)X^T(m-1)M(m-1)\underline{x}(m)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$+ \underline{x}_p(m+1)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$- \underline{x}_p(m+1)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1) \quad .$$

Rearranging terms and using (G-5), equation (G-14) becomes

$$\hat{\Phi}^{o}_{po}(m+1) = \hat{\Phi}^{o}_{po}(m) \tag{G-16}$$

$$+ \underline{x}_p(m+1)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$- \hat{\Phi}^{o}_{po}(m)\underline{x}(m)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$+ \underline{x}_p(m+1)\underline{x}^T(m)M(m-1)$$

$$- \underline{x}_p(m+1)\underline{x}^T(m)M(m-1)\underline{x}(m)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$- \underline{x}_p(m+1)\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

or, gathering terms,

$$\hat{\Phi}^{o}_{po}(m+1) = \hat{\Phi}^{o}_{po}(m) \tag{G-17}$$

$$+ [\underline{x}_p(m+1) - \hat{\Phi}^{o}_{po}(m)\underline{x}(m)]\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$+ \underline{x}_p(m+1)\underline{x}^T(m)M(m-1)$$

$$- \underline{x}_p(m+1)\,[\underline{x}^T(m)M(m-1)\underline{x}(m) + 1]\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

or, finally,

$$\hat{\Phi}^{o}_{po}(m+1) = \hat{\Phi}^{o}_{po}(m) \tag{G-18}$$

$$+ [\underline{x}_p(m+1) - \hat{\Phi}^{o}_{po}(m)\underline{x}(m)]\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

$$+ \underline{x}_p(m+1)\underline{x}^T(m)M(m-1) - \underline{x}_p(m+1)\underline{x}^T(m)M(m-1)$$

Therefore, the recursive discrete time minimum mean square error identification equation is

$$\hat{\Phi}^{o}_{po}(m+1) = \hat{\Phi}^{o}_{po}(m) \tag{G-19}$$

$$+ [\underline{x}_p(m+1) - \hat{\Phi}^{o}_{po}(m)\underline{x}(m)]\,[1 + \underline{x}^T(m)M(m-1)\underline{x}(m)]^{-1}\underline{x}^T(m)M(m-1)$$

where the recursive equation for $M(m-1)$ is given by (G-14). The indicated matrix inversion is actually just division by a scalar so that the original matrix inversion has been avoided.

## G-3. Initialization of the Recursive Estimation Process

In order to utilize the previously derived recursive identification equation (G-19) and its recursive coefficient equation (G-14), $\hat{\Phi}^{o}_{po}$ and $M$ must be initialized. This initialization can be accomplished using the following techniques.

The first and most obvious initialization technique utilizes (G-6) at time $(2n-1)\Delta T$ to compute $M(2n-1)$ and (G-5) at time $2n\Delta T$ to compute $\hat{\Phi}^{o}_{po}(2n)$. As shown by example later, the estimate $\hat{\Phi}^{o}_{po}(2n)$ computed from (G-5) can be very much in error if the initial measurements are noisy. However, no a priori information is required or utilized in this technique.

The second initialization technique also utilizes (G-6) to compute M(2n-1). However, for the incomplete information differential game, there is sufficient a priori information available from the limiting estimates of the opponent's system parameters to allow computation of an initial value of $\hat{\Phi}^o_{po}$ at $t = 0$. Therefore, this initial estimate can be used as the estimate at $t = 2n\Delta T$. As shown by example later, this initialization results in slower parameter identification, but the estimates are also much less sensitive to measurement noise.

The third initialization technique, which avoids the matrix inversion required by (G-6), is discussed by Lee [16] for the single input/output system identification problem. As applied to the multivariable system identification problem being considered herein, this technique consists of setting

$$M(-1) = a\, I \tag{G-20}$$

where $a \rightarrow \infty$, in order to initialize (G-14) and using the a priori value of $\hat{\Phi}^o_{po}(0)$ to initialize (G-19). The result, as shown by example later, is an asymptotically convergent solution for M and $\hat{\Phi}^o_{po}$ which begins with the first rather than the (2n+1) sample.

A fourth initialization technique exists for the incomplete information differential game since an initial estimate for

$$\Phi = \begin{bmatrix} \Phi^o_{ppo} & \Phi^o_{peo} \\ \Phi^*_{epo} & \Phi^*_{eeo} \end{bmatrix} \tag{G-21}$$

can be made at time $t = 0$. This a priori information can be used to initialize the recursive coefficient equation (G-14) by computing

$$M(-1) = [X(-1)\ X^T(-1)]^{-1} \tag{G-22}$$

where

$$X(-1) = [\hat{\Phi}^{-1}\underline{x}(0) \quad \hat{\Phi}^{-2}\underline{x}(0) \quad \ldots \quad \hat{\Phi}^{-2n}\underline{x}(0)] \quad . \tag{G-23}$$

The initial value of $\hat{\Phi}^{o}_{po}(0)$ is again utilized in order to initialize (G-19). Consequently, estimation of $\Phi^{o}_{po}$ again starts with the first sample. However, as shown by (G-22), one matrix inversion is still required.

## G-4. Simulation

Digital simulation was used to examine the response characteristics of the recursive discrete time identification technique when applied to the continuous linear time-invariant system given by

$$\begin{bmatrix} \dot{\underline{x}}_p \\ \dot{\underline{x}}_e \end{bmatrix} = \begin{bmatrix} -5/2 & 2/3 \\ -1/2 & -1/2 \end{bmatrix} \begin{bmatrix} \underline{x}_p \\ \underline{x}_e \end{bmatrix} \quad ; \quad \begin{bmatrix} \underline{x}_p(0) \\ \underline{x}_e(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \end{bmatrix} \quad . \tag{G-24}$$

A fourth order Runge-Kutta integration routine with a fixed 0.001 second step size was used in order to maintain accuracy. Note that (G-24) closely approximates the early stages of the example incomplete information game, discussed in Chapter III, in which the Evader is the ignorant player.

The recursive identification technique was simulated for the case of both perfect and noisy measurements using the four initialization techniques previously discussed. The initial values of $\hat{\Phi}^{o}_{ppo}(0)$ and $\hat{\Phi}^{o}_{peo}(0)$ were calculated using (82) and (83) of Chapter IV, where the a priori limiting estimates for $A^{o}_{ppo}$ and $A^{o}_{peo}$ are given as

$$\hat{A}^{o}_{ppo} = -3/2 \tag{G-25}$$

and

$$\hat{A}^{o}_{peo} = 1/2 \quad . \tag{G-26}$$

The simulation results using the four initialization techniques are shown in Figures G-1, G-2, G-3, and G-4, respectively. As shown, the first initialization technique yields the fastest identification, but the estimates are very sensitive at the start to even minor measurement errors. The third initialization technique results in a slightly slower parameter identification, but is relatively insensitive to measurement noise. The second and fourth initialization techniques are quite similar in that they both result in a large initial transient and the identification is much slower than for either the first or third technique. From these simulation results, the third initialization technique, which avoids the requirement for matrix inversion, is the best overall choice, at least for systems and initial conditions similar to those of the simulation example.

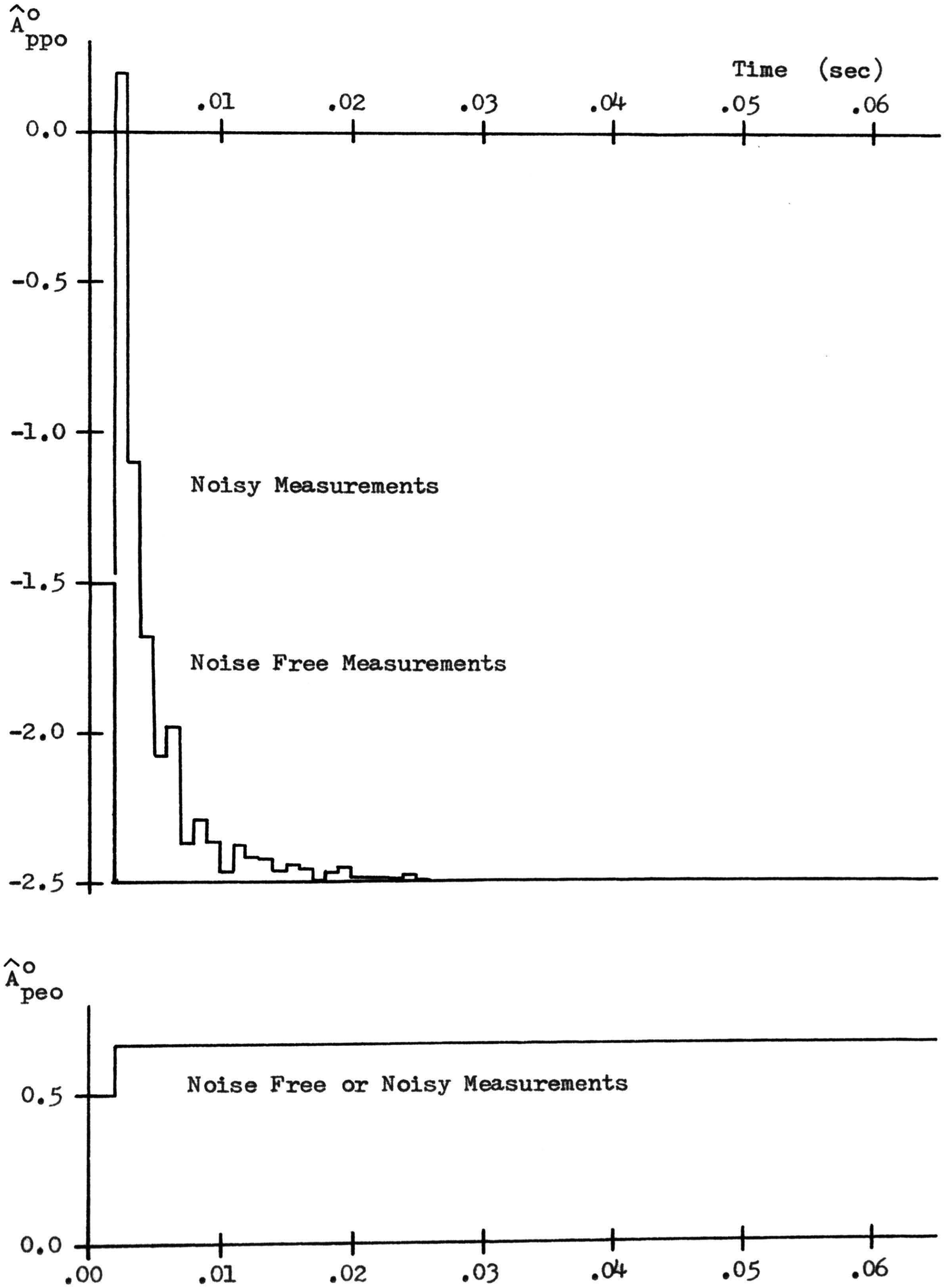



Figure G-1. Discrete Time Parameter Identification - First Initialization Technique

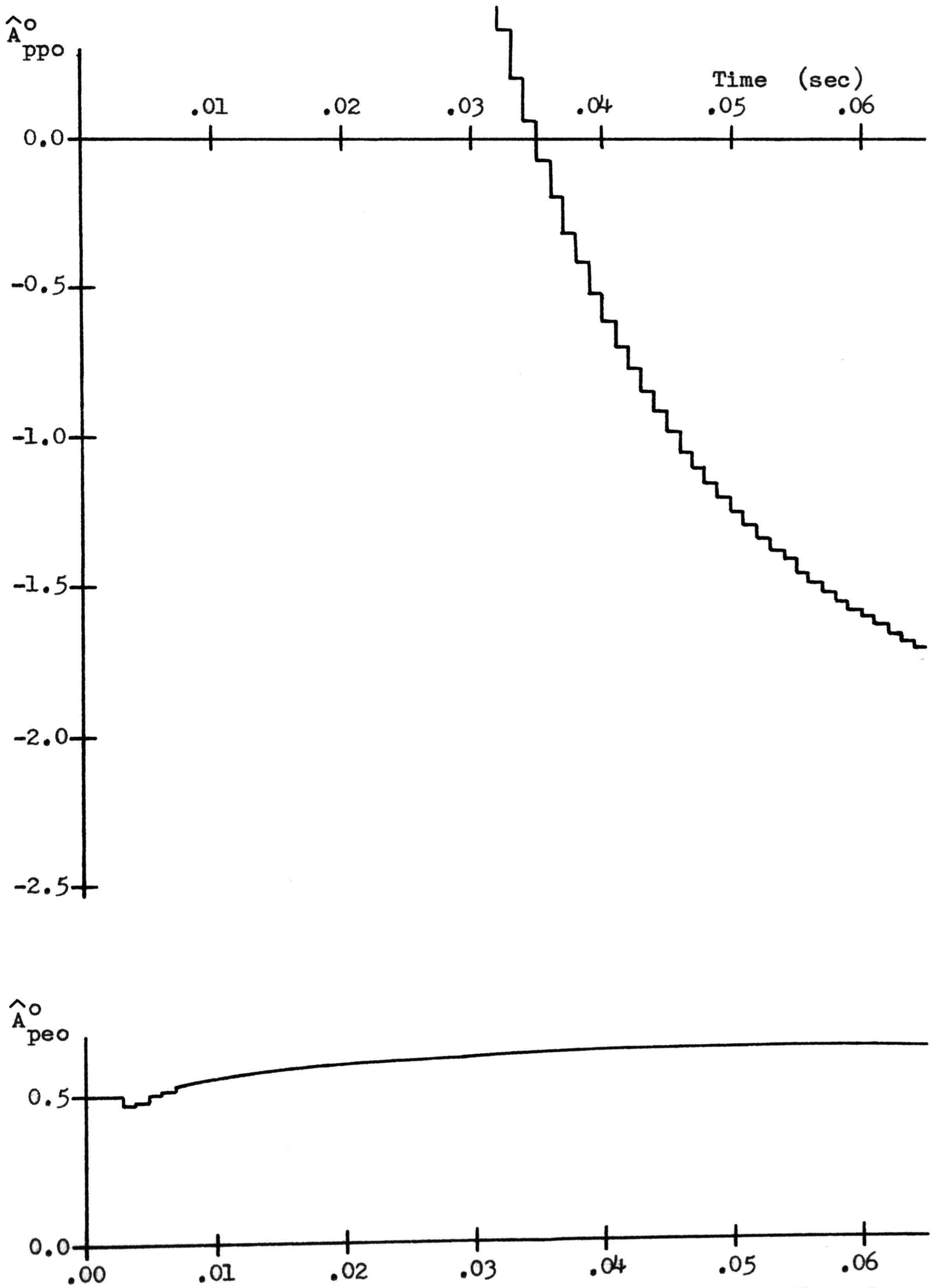



Figure G-2. Discrete Time Parameter Identification - Second Initialization Technique

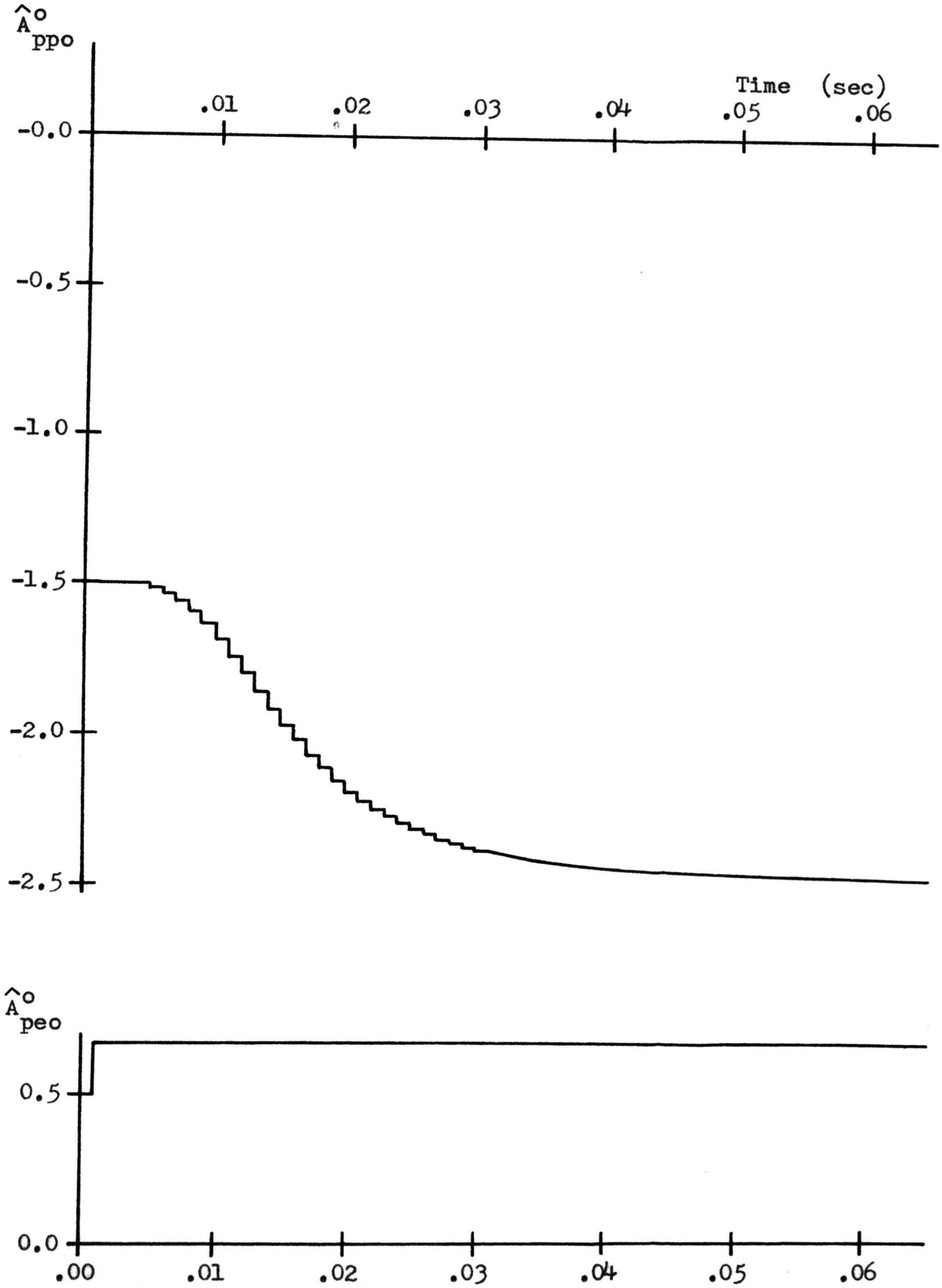



**Figure G-3. Discrete Time Parameter Identification - Third Initialization Technique**

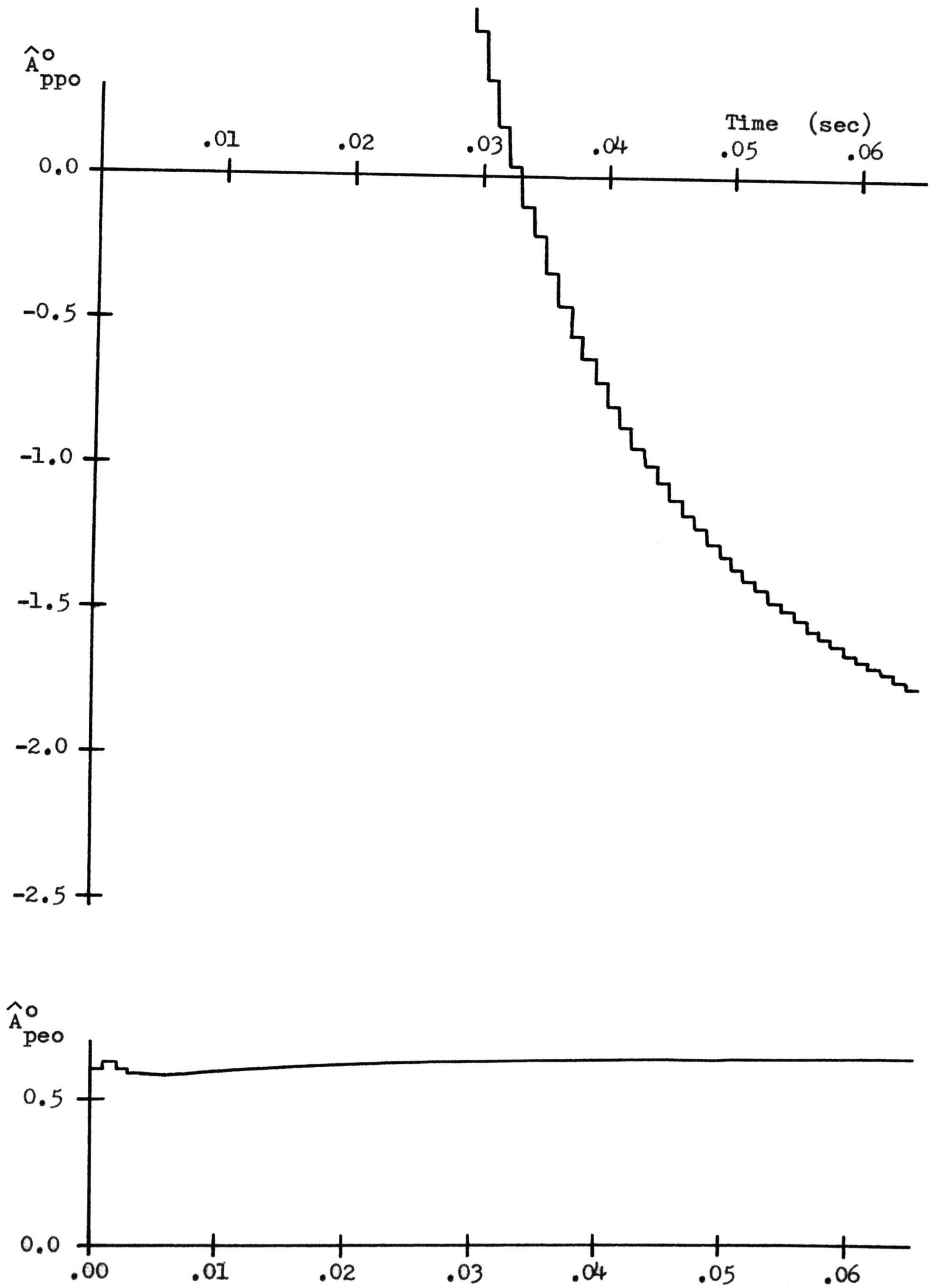



Figure G-4. Discrete Time Parameter Identification - Fourth Initialization Technique