NASA TECHNICAL NOTE

NASA TN D-6502

c. 1

NASA TN D-6502

DEVELOPMENT AND APPLICATION OF A GRADIENT METHOD FOR SOLVING DIFFERENTIAL GAMES

by David A. Roberts and Raymond C. Montgomery

Langley Research Center
Hampton, Va. 23365

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • NOVEMBER 1971

1. Report No. NASA TN D-6502	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT AND APPLICATION OF A GRADIENT METHOD FOR SOLVING DIFFERENTIAL GAMES		5. Report Date November 1971
		6. Performing Organization Code
7. Author(s) David A. Roberts and Raymond C. Montgomery		8. Performing Organization Report No. L-7770
		10. Work Unit No. 136-62-01-02
9. Performing Organization Name and Address NASA Langley Research Center Hampton, Va. 23365		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Note
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

Part of the information presented in this paper was included in a thesis entitled "The Application of Gradient Methods to Differential Games" submitted by David A. Roberts in partial fulfillment of the requirements for the degree of Master of Science, George Washington University, May 1971.

16. Abstract

A gradient technique for solving n-dimensional differential games is developed in this paper and applied to two example pursuit-evasion games: the first, a two-dimensional game similar to the homicidal chauffeur but modified to resemble an airplane-helicopter engagement, and the second, a five-dimensional game of two airplanes at constant altitude and with thrust and turning controls. In both games, the performance function to be optimized by the pursuer and evader was the distance between the evader and a given target point in front of the pursuer. The analytic solution to the first game is found and compared with the gradient solution. The analytic solution reveals that both unique and nonunique solutions exist, depending on the initial conditions selected. A comparison between the gradient results and the analytic solution shows a dependence on the nominal controls in regions where nonunique solutions exist. In the unique solution region, the results from the two methods agree very closely. The application of the gradient method to the five-dimensional two-airplane game is illustrated for one set of initial conditions. These results are also shown to be dependent on the nominal controls selected and indicate that these initial conditions are in a region of nonunique solutions.

17. Key Words (Suggested by Author(s)) Differential games; Gradient method	18. Distribution Statement Unclassified – Unlimited

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 46	22. Price* $3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

DEVELOPMENT AND APPLICATION OF A GRADIENT METHOD FOR SOLVING DIFFERENTIAL GAMES*

By David A. Roberts and Raymond C. Montgomery
Langley Research Center

SUMMARY

A gradient technique for solving n-dimensional differential games is developed in this paper and applied to two example games. The method requires selecting nominal controls for both players and then improving these controls iteratively until the optimal controls are found. An iteration scheme is recommended, consisting of a minimization phase employing only steepest descent followed by a phase employing alternately steepest descent and steepest ascent. The gradient method has been applied to two pursuit-evasion games: the first, a two-dimensional game similar to the homicidal chauffeur but modified to resemble an airplane-helicopter engagement; and the second, a five-dimensional game of two airplanes at constant altitude and with thrust and turning controls. In both games, the performance function to be optimized by the pursuer and evader was the distance between the evader and a given target point in front of the pursuer.

The analytic solution to the first game is found and compared with the gradient solution. The outcome of the gradient method is strongly dependent on the nominal controls selected in some cases. However, the analytic solution reveals the existence of both unique and nonunique solutions, depending on the initial conditions selected. A comparison between the gradient results and the analytic solution shows that the dependence on the nominal controls occurs only in regions where nonunique solutions exist. In the unique solution region, the results from the two methods agree very closely.

The application of the gradient method to the five-dimensional two-airplane game is illustrated for one set of initial conditions. These results are also shown to be dependent on the nominal controls selected and indicate that these initial conditions are in a region of nonunique solutions.

*Part of the information presented in this paper was included in a thesis entitled "The Application of Gradient Methods to Differential Games" submitted by David A. Roberts in partial fulfillment of the requirements for the degree of Master of Science, George Washington University, May 1971.

INTRODUCTION

Analytic methods useful in studying pursuit-evasion are principally derived from game theory. A game can be considered as a multiple-decision process in which at least two players take part and, in some manner, have conflicting objectives. A game has three basic features: a starting condition, a set of rules governing the evolution of the play, and a termination condition. In a pursuit-evasion game involving two airplanes, the airplane equations of motion and the controls available to the pilots define the rules of the game. A game in which these rules are stated as differential equations is called a "differential game." Obviously, each player wishes to select his controls to "play" in the best manner, that is to achieve some objective. In a differential game, this objective or "payoff" is usually formed as a quantitative measure of how well the game is played. Note that the pursuer could have a payoff different from that of the evader. The objective then of each player is to select the controls that optimize his payoff.

To determine the optimal strategies for the players, one must select a rationale for their actions. An example is to select a set of evasive maneuvers, either open-loop or closed-loop, and to determine pursuer maneuvers that lead to optimal pursuer payoff. But the difficulty with this approach is that the optimal pursuer maneuvers depend strongly on the evasive maneuvers selected. Another improved rationale is to assume that each player selects a strategy based on receiving optimal payoff for himself while assuming that his opponent will accordingly select a strategy to optimize his payoff. If, furthermore, the payoff for the two opponents always adds identically to zero, the game is called a zero-sum differential game because one player attempts to maximize the payoff while the other attempts to minimize the same quantity P.

The purpose of this paper is to develop and apply a gradient method for solving zero-sum differential games. Present solution methods rely heavily on indirect or analytic methods similar to those originated by Isaacs. (See ref. 1.) However, these methods are currently restricted to problems where the governing equations are linear or contain simple nonlinearities and to problems of low state dimensions. Direct methods that have proven to be successful in optimization theory (ref. 2) have not generally been used to solve games. However, Taylor (ref. 3) has applied Balakrishnan's Epsilon technique to a pursuit-evasion problem. His method has been applied to problems in which the payoff is the time required to reach a specified position regardless of orientation. Recently, Baron et al. (ref. 4) have developed a direct method which is global in nature (that is, the solution is found for all initial conditions). However, the computer storage required to attain this global nature limits the method's applicability. The gradient method described in this report numerically determines the optimal controls by iterating on some nominal set of controls. Since this method is not limited by the restrictions discussed, it may readily be applied to more realistic differential games. To insure convergence, this method is limited to games in which either the minimum or the maximum of the payoff can first be found for a fixed action on the part of the other player. Then, once this optimization problem is solved, both players can be permitted to optimize to find the min-max or max-min solution. For large dimensional problems, the saddle-point solution may not exist; that is, the min-max solution and the max-min solution may not be the same. The iteration scheme as presented in the appendix will insure convergence only to a max-min solution. However, the solutions presented in this paper are believed to be also saddle-point solutions. The terms "saddle point," "min-max," and "max-min" will then be used interchangeably to describe the solutions with the understanding that they may not be the same in more complicated problems.

SYMBOLS

a_n	normal acceleration, g
C	terminal manifold
C_1, C_2	constant weightings in payoff
$C_{D,i}$	induced drag coefficient
$C_{D,o}$	drag coefficient at zero angle of attack
f	function for equations of motion
g	gravitational acceleration constant, 9.8 m/sec^2
H	Hamiltonian defined in equations (20)
I	$n \times n$ identity matrix
J(x)	scalar termination function
K_u	pursuer gradient gain
K_v	evader gradient gain
m	airplane mass, kg

n	dimension of system in state space
P	scalar payoff
R	range of target point, m
R_c	minimum turning radius, m
r	distance from target point, m
S	airplane wing area, m^2
S_2,S_3	parameters of playing space
T	airplane thrust, N
t	forward time, independent variable
$\vec{u}$	pursuer control vector, $m \times 1$
V	airplane velocity, m/sec; Value of game, min-max payoff
V_i	Value derivatives, $\frac{\partial V}{\partial x_i}$
$\vec{v}$	evader control vector, $m \times 1$
w	pursuer's angular velocity
w_1	pursuer's velocity, m/sec
w_2	evader's velocity, m/sec
X,Y	axes for inertial coordinates
$\vec{x}$	state vector, $m \times 1$
θ	angular orientation of players
ρ	atmospheric density, kg/m^3

σ	independent variable of steepest descent
$\Delta\sigma$	increment in σ
τ	variable of integration, or retrograde time, sec
Φ	$n \times n$ matrix defined by equation (15)
ϕ	pursuer's turning control
Ψ	$n \times n$ transition matrix defined by equation (11)
ψ	velocity angle control for evader, rad
$\nabla_x f, A$	$n \times n$ gradient matrix, $A_{ij} \equiv \left(\frac{\partial f_i}{\partial x_j}\right)$
$\nabla_u f, B$	$n \times m$ pursuer control gradient matrix, $B_{ij} \equiv \left(\frac{\partial f_i}{\partial u_j}\right)$
$\nabla_v f, C$	$n \times m$ evader control gradient matrix, $C_{ij} \equiv \left(\frac{\partial f_i}{\partial v_j}\right)$
$\nabla_x^T J$	$1 \times n$ row vector whose ith element is $\frac{\partial J}{\partial x_i}$
$\nabla_x^T P$	$1 \times n$ row vector whose ith element is $\frac{\partial P}{\partial x_i}$

Subscripts:

A	airplane A
B	airplane B
f	evaluation at terminal time
i,j	ith and jth components
nom	nominal controls or path
o	initial value

Superscripts:

T transpose of vector or matrix

* optimal controls

Vectors are denoted by arrows; quantities without arrows are scalar. A dot over a symbol denotes differentiation with respect to time. The symbol $\triangleq$ means "is defined as" and the symbol $\equiv$ denotes identically equal terms.

PROBLEM STATEMENT

Generally, for any two airplanes the equations governing the evolution of an engagement involve a set of quantities $\vec{x}$ – referred to as the state – which describe the positions and velocities of the vehicles, and a set of quantities $\vec{u}$ and $\vec{v}$ – referred to as the controls – which describe the way in which the individual aircraft may affect the evolution of the engagement.

The first step in determining the relative superiority of two opposing airplanes is to decide, for each, how much value will be attached to being in a certain state – a certain position relative to the opponent. Indeed, if the method of gradients is to be used, the weighting should be determined so that a single number P is associated with each state $\vec{x}$. This function $P(\vec{x})$ is referred to as the "payoff" and is ordered so that $P(\vec{x}) = 0$ corresponds to ideal capture and $P(\vec{x}) \to \infty$ corresponds to escape for the evader. The numerical value of the function P for an airplane (designated A) thus provides a measure of how "far" the airplane (designated B) is from being placed in an unfavorable situation. This ordering of the situation is arbitrary. In a general game, the cost structure $P_A(\vec{x})$ for airplane A may differ from that of airplane B $P_B(\vec{x})$. The analysis presented in this report is restricted to cases where $P_A(\vec{x}) = P(\vec{x}) = -P_B(\vec{x})$, that is, to zero-sum differential games.

To illustrate these concepts, consider a simplified model of pursuit-evasion, which is restricted to maneuvers in a horizontal plane. This particular model has eight state variables made up of the absolute position coordinates and heading of both vehicles and control variables such as the accelerations normal to and along the flight path of each vehicle. Figure 1 illustrates the geometry and shows the state variables for each airplane. A typical example of the payoff for airplane A is presented in figure 2. In this figure the coordinates represent the relative position of airplane B in the horizontal plane with the x_1-axis directed along the velocity of airplane A. Contours for constant values of the function $P(x)$ are indicated in the figure. For this example, the most favorable situation for the pursuer was for the evader to be at the pursuer's target point where $P = 0$.

To provide an analytic statement of the problem considered in this study, the governing differential equations of motion are given in the form

$$\dot{\vec{x}} = \vec{f}\left(\vec{x}, \vec{u}, \vec{v}\right) \tag{1}$$

where

$\vec{x}$ an $n \times 1$ state vector

$\vec{u}$ an $m \times 1$ control vector used by pursuer

$\vec{v}$ an $m \times 1$ control vector used by evader

Admissible engagements will be those which satisfy a given initial condition

$$\vec{x}\left(t_o\right) = \vec{x}_o \tag{2}$$

a given termination condition

$$J\left(\vec{x}\left(t_f\right)\right) = 0 \tag{3}$$

and the governing differential equations (1) for a given set of controls $\vec{u}(t)$ and $\vec{v}(t)$ defined on the time interval $t_o \leqq t \leqq t_f$. The payoff of an admissible engagement is

$$P\left(\vec{u}, \vec{v}\right) = P\left(\vec{x}\left(t_f\right)\right) \tag{4}$$

Admissible control functions $\vec{u}$ and $\vec{v}$ which satisfy equations (1) to (4) are said to be of class C. The problem considered in this paper is to find a set of admissible controls $\left(\vec{u}^* \text{ and } \vec{v}^*\right)$ satisfying

$$\left.\begin{aligned} \vec{u}^*\left(\vec{v}\right) &= \underset{\left(\vec{u}, \vec{v}\right) \epsilon C}{\arg\min} P\left(\vec{u}, \vec{v}\right) \\ \vec{v}^* &= \underset{\left(\vec{u}^*, \vec{v}\right) \epsilon C}{\arg\max} P\left(\vec{u}^*\left(\vec{v}\right), \vec{v}\right) \end{aligned}\right\} \tag{5}$$

where arg min and arg max indicate the functions that minimize or maximize the payoff, respectively. The max-min solution to the differential game is termed $P(\vec{u}^*,\vec{v}^*)$.

GENERAL THEORY

The basic approach of the gradient method in optimization theory is described in references 2 and 5. The same general concepts are used in this paper to obtain solutions of differential games (that is, finding controls $\vec{u}^*$ and $\vec{v}^*$ which satisfy equations (5)). Figure 3 illustrates the gradient approach to solving this problem. The basic idea is to start with a nominal engagement between opponents and to modify this engagement iteratively to reach the desired solution. The procedure for modifying the engagements from one iteration to another is based on linearized variational equations.

To obtain the gradient formulas for modifying the controls, it is convenient to describe each step in the iterative process by an additional independent variable σ which may be thought of as an iteration number. At $\sigma = 0$, the state and control functions are those of the nominal engagement and as $\sigma \to \infty$, the state and control functions should approach the optimal solution sought. Then

$$\vec{u}(t,\sigma)\Big|_{\sigma=0} \equiv \vec{u}_{nom}(t)$$

$$\vec{v}(t,\sigma)\Big|_{\sigma=0} \equiv \vec{v}_{nom}(t)$$

and the optimal controls correspond to

$$\vec{u}^*(t) \equiv \vec{u}\left(t,\sigma_f\right)$$

$$\vec{v}^*(t) \equiv \vec{v}\left(t,\sigma_f\right)$$

The iterative process for a typical control as a function of t and σ is depicted in figure 4. The nominal control is at the intersection of the $\vec{u},t$ plane and the σ-axis. Notice that as σ increases, $\vec{u}(t,\sigma)$ gradually converges to a function considered to be the optimal. Note also that the final time is not necessarily constant but is instead a function $t_f(\sigma)$.

The kinematic equations (eqs. (1)) can be rewritten in the form

$$\frac{\partial \vec{x}}{\partial t}(t,\sigma) = \vec{f}\left(\vec{x}(t,\sigma),\vec{u}(t,\sigma),\vec{v}(t,\sigma)\right) \tag{6}$$

with equation (2)

$$\vec{x}\left(t_o,\sigma\right) \equiv \vec{x}_o \tag{7}$$

For a given $\vec{u}(t,\sigma)$ and $\vec{v}(t,\sigma)$, equations (6) and (7) may be used to determine $\vec{x}(t,\sigma)$. The termination time $t_f(\sigma)$ is determined from equation (3) written as

$$J\left(\vec{x}\left(t_f,\sigma\right)\right) = 0 \tag{8}$$

If the controls $\vec{u}$ and $\vec{v}$ are modified according to the linearized relations

$$\vec{u}(t, \sigma+\Delta\sigma) = \vec{u}(t,\sigma) + \frac{\partial\vec{u}}{\partial\sigma}(t,\sigma)\ \Delta\sigma$$

$$\vec{v}(t, \sigma+\Delta\sigma) = \vec{v}(t,\sigma) + \frac{\partial\vec{v}}{\partial\sigma}(t,\sigma)\ \Delta\sigma$$

the variation in $\vec{x}$ will approximately satisfy

$$\vec{x}(t, \sigma+\Delta\sigma) \approx \vec{x}(t,\sigma) + \frac{\partial\vec{x}}{\partial\sigma}(t,\sigma)\ \Delta\sigma$$

The function $\frac{\partial\vec{x}}{\partial\sigma}(t,\sigma)$ can be obtained by formally differentiating equation (6) with respect to σ and interchanging the order of the t and σ differentiation to obtain the linear equation

$$\frac{\partial}{\partial t}\frac{\partial\vec{x}}{\partial\sigma} = A(t,\sigma)\frac{\partial\vec{x}}{\partial\sigma} + B(t,\sigma)\frac{\partial\vec{u}}{\partial\sigma} + C(t,\sigma)\frac{\partial\vec{v}}{\partial\sigma} \tag{9}$$

where

$$A(t,\sigma) \triangleq \nabla_x f\left(\vec{x}(t,\sigma), \vec{u}(t,\sigma), \vec{v}(t,\sigma)\right)$$

$$B(t,\sigma) \triangleq \nabla_u f\left(\vec{x}(t,\sigma), \vec{u}(t,\sigma), \vec{v}(t,\sigma)\right)$$

$$C(t,\sigma) \triangleq \nabla_v f\left(\vec{x}(t,\sigma), \vec{u}(t,\sigma), \vec{v}(t,\sigma)\right)$$

Equation (9) is a linear differential equation which can be solved for $\frac{\partial\vec{x}}{\partial\sigma}$ in terms of $\frac{\partial\vec{u}}{\partial\sigma}$ and $\frac{\partial\vec{v}}{\partial\sigma}$. Since these last two variables are needed to update the controls, the solution to equation (9) will be used to find expressions for $\frac{\partial\vec{u}}{\partial\sigma}$ and $\frac{\partial\vec{v}}{\partial\sigma}$ that will generate the optimal controls. It is known that the solution to equation (9) can be written as (see ref. 6)

$$\frac{\partial\vec{x}}{\partial\sigma}(t,\sigma) = \int_{t_o}^{t} \Psi^T(\tau,t,\sigma)\ B(\tau,\sigma)\frac{\partial\vec{u}}{\partial\sigma}(\tau,\sigma)\ d\tau$$

$$+ \int_{t_o}^{t} \Psi^T(\tau,t,\sigma)\ C(\tau,\sigma)\frac{\partial\vec{v}}{\partial\sigma}(\tau,\sigma)\ d\tau + \Psi^T\left(t_o,t,\sigma\right)\frac{\partial\vec{x}}{\partial\sigma}\left(t_o,\sigma\right) \tag{10}$$

where $\Psi(\tau,t,\sigma)$ is a matrix defined by the adjoint differential equation

$$\frac{\partial}{\partial t} \Psi\left(t,t_o,\sigma\right) = -A^T(t,\sigma)\ \Psi\left(t,t_o,\sigma\right) \tag{11}$$

and with the additional properties that

$$\Psi\left(t_o,t_o,\sigma\right) = I$$

$$\Psi\left(t,t_o,\sigma\right) = \Psi^{-1}\left(t_o,t,\sigma\right)$$

$$\Psi\left(t,t_f,\sigma\right) = \Psi\left(t,t_o,\sigma\right)\ \Psi\left(t_o,t_f,\sigma\right)$$

These latter properties will permit the transition matrix to be integrated numerically and simultaneously with the equations of motion in forward time.

The initial condition of equation (7) will be constant from iteration to iteration. Hence, the initial condition on $\frac{\partial \vec{x}}{\partial \sigma}$ needed to integrate equation (9) is

$$\frac{\partial \vec{x}}{\partial \sigma}\left(t_o,\sigma\right) \equiv 0$$

Since a terminal payoff is being considered, the variation in the state must be evaluated at $t = t_f$. The total variation in the terminal state $\vec{x}_f \triangleq \vec{x}\left(t_f(\sigma),\sigma\right)$ can be calculated from the expression

$$\frac{d\vec{x}_f}{d\sigma} = \frac{\partial \vec{x}}{\partial \sigma}\left(t_f,\sigma\right) + \frac{\partial \vec{x}}{\partial t}\left(t_f,\sigma\right)\frac{dt_f}{d\sigma}$$

which results in (from eqs. (6) and (10))

$$\frac{d\vec{x}_f}{d\sigma} = \int_{t_o}^{t_f} \Psi^T\left(\tau,t_f,\sigma\right)\left[B(\tau,\sigma)\ \frac{\partial \vec{u}}{\partial \sigma}(\tau,\sigma) + C(\tau,\sigma)\ \frac{\partial \vec{v}}{\partial \sigma}(\tau,\sigma)\right] d\tau$$

$$+\ \vec{f}\left(\vec{x}\left(t_f,\sigma\right),\vec{u}\left(t_f,\sigma\right),\vec{v}\left(t_f,\sigma\right)\right)\frac{dt_f}{d\sigma} \tag{12}$$

If a fixed duration game is being considered, $\frac{dt_f}{d\sigma}$ will be zero. However, for variable termination time, $\frac{dt_f}{d\sigma}$ must be evaluated from the termination condition (eq. (8)).

$$J\left(\vec{x}\left(t_f,\sigma\right)\right) = 0$$

Differentiation of equation (8) yields

$$\frac{dJ}{d\sigma} = \nabla_x^T J\left(\vec{x}_f\right)\frac{d\vec{x}_f}{d\sigma} = 0 \tag{13}$$

The term $\frac{dt_f}{d\sigma}$ may now be eliminated between equations (12) and (13) and an expression for $\frac{d\vec{x}_f}{d\sigma}$ can be obtained independent of $\frac{dt_f}{d\sigma}$. This result is

$$\frac{d\vec{x}_f}{d\sigma} = \Phi\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\left\{\int_{t_o}^{t_f} \Psi^T\left(\tau,t_f,\sigma\right)\left[B(\tau,\sigma)\frac{\partial\vec{u}}{\partial\sigma}(\tau,\sigma) + C(\tau,\sigma)\frac{\partial\vec{v}}{\partial\sigma}(\tau,\sigma)\right] d\tau\right\} \tag{14}$$

where

$$\Phi\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right) = I - \frac{\vec{f}\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\nabla_x^T J\left(\vec{x}_f\right)}{\nabla_x^T J\left(\vec{x}_f\right) f\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)} \tag{15}$$

where

$$\vec{u}_f \triangleq \vec{u}\left(t_f(\sigma),\sigma\right)$$

$$\vec{v}_f \triangleq \vec{v}\left(t_f(\sigma),\sigma\right)$$

The total derivative of the payoff P with respect to σ can be determined by using equation (14) and noting that

$$\frac{dP}{d\sigma}\left(\vec{x}_f\right) = \nabla_x^T P\left(\vec{x}_f\right)\frac{d\vec{x}_f}{d\sigma}$$

Hence,

$$\frac{dP}{d\sigma} = \nabla_x^T P\left(\vec{x}_f\right)\Phi\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\left\{\int_{t_o}^{t_f} \Psi^T\left(\tau,t_f,\sigma\right)\left[B(\tau,\sigma)\frac{\partial\vec{u}}{\partial\sigma}(\tau,\sigma) + C(\tau,\sigma)\frac{\partial\vec{v}}{\partial\sigma}(\tau,\sigma)\right] d\tau\right\} \tag{16}$$

Equation (16) gives the "slope" of the payoff with respect to the iteration variable σ. It is possible to choose expressions for $\frac{\partial\vec{u}}{\partial\sigma}$ and $\frac{\partial\vec{v}}{\partial\sigma}$ from equation (16) that yield the most negative and most positive values of $\frac{dP}{d\sigma}$. From reference 2, the steepest descent direction (used for minimizing the payoff with respect to the control $\vec{u}$) is

$$-B^T(t,\sigma)\ \Psi\left(t,t_f,\sigma\right)\ \Phi^T\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\ \nabla_x P\left(\vec{x}_f\right)$$

and the steepest ascent direction (used for maximizing the payoff with respect to the control $\vec{v}$) is

$$+C^T(t,\sigma)\ \Psi\left(t,t_f,\sigma\right)\ \Phi^T\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\ \nabla_x P\left(\vec{x}_f\right)$$

The control variations are now chosen as

$$\left.\begin{aligned} \frac{\partial \vec{u}}{\partial \sigma} &= -K_u B^T(t,\sigma)\ \Psi\left(t,t_f,\sigma\right)\ \Phi^T\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\ \nabla_x P\left(\vec{x}_f\right) \\ \frac{\partial \vec{v}}{\partial \sigma} &= +K_v C^T(t,\sigma)\ \Psi\left(t,t_f,\sigma\right)\ \Phi^T\left(\vec{x}_f,\vec{u}_f,\vec{v}_f\right)\ \nabla_x P\left(\vec{x}_f\right) \end{aligned}\right\} \tag{17}$$

where K_u and K_v are positive scalar constants whose selection is explained in the appendix. The control modifications, where the controls are now updated by using the first-order approximation, are of the form

$$\vec{u}(t,\ \sigma+\Delta\sigma) = \vec{u}(t,\sigma) + \frac{\partial \vec{u}}{\partial \sigma}\Delta\sigma$$

$$\vec{v}(t,\ \sigma+\Delta\sigma) = \vec{v}(t,\sigma) + \frac{\partial \vec{v}}{\partial \sigma}\Delta\sigma$$

The iterative process is continued until $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$ approach zero and no improvement is observed in the payoff. Notice that saturation may occur if limits are placed on $\vec{u}(t,\sigma)$ and $\vec{v}(t,\sigma)$. In that case the system will be converged when the controls saturate, even though the control derivatives do not approach zero. Details of the iteration sequence to insure convergence are found in the appendix, as are suggestions for handling the variable terminal time.

APPLICATION TO A SIMPLIFIED DIFFERENTIAL GAME

Analytic Solution

Early in the present study it became evident that the sheer size of problems involving two airplanes inhibited a clear understanding of certain problem areas. Such problems as the type of solution or whether the gradient method would converge to a solution could not be readily investigated in the high-dimensional states. Also it became apparent that the gradient method could be checked by a comparison with some analytic results. For these reasons, a simplified game of lower dimensions was solved; however, the game still embodied the basic concepts of the higher dimensional games.

The geometry of the simplified game is shown in figure 5 for an inertial coordinate system. The pursuer A moves with constant velocity w_1 and controls his radius of turn. The evader B moving with a different constant velocity w_2 controls his velocity direction ψ and is free to change it instantaneously. The equations of motion are

$$\left.\begin{aligned}
\dot{x}_A &= w_1 \sin \theta \\
\dot{y}_A &= w_1 \cos \theta \\
\dot{\theta} &= \frac{w_1}{R_c} \phi \\
\dot{x}_B &= w_2 \cos (\theta - \psi) \\
\dot{y}_B &= w_2 \sin (\theta - \psi)
\end{aligned}\right\} \tag{18}$$

To reduce the order of the system, a relative coordinate system is defined as illustrated in figure 6. The origin of the relative system is centered in and moving with the pursuer so that the pursuer's velocity is always along the positive x_1-axis. Notice that two state variables are now adequate to describe the system completely. The equations of motion are the same as those given by Isaacs in reference 1 for the homicidal chauffuer game

$$\left.\begin{aligned}
\dot{x}_1 &= w\phi x_2 + w_2 \cos \psi - w_1 \\
\dot{x}_2 &= -w\phi x_1 + w_2 \sin \psi
\end{aligned}\right\} \tag{19}$$

where

$$w = \frac{w_1}{R_c}$$

the pursuer's turning rate control is

$$-1 \leqq \phi \leqq +1$$

and the evader's direction control is

$$-\pi \leqq \psi \leqq \pi$$

The payoff will be terminal and is taken as the distance r from the target point (R,0) at the terminal time. Only one pass will be permitted, the game terminating at the closest approach to the target point for each run. It is further stipulated that $\dot{r} < 0$ throughout the region of play. To simplify computation, the payoff may be redefined as $P = \frac{r^2}{2}$, and $J = \dot{P} = 0$ can be taken as the termination condition. The problem now is to determine controls ϕ and ψ such that $P\left(\vec{x}(t_f)\right)$ is a minimum with respect to ϕ and a maximum with respect to ψ with $\dot{r} \leqq 0$ and with t_f determined by the satisfaction of

$$J = \left[-w\phi R x_2 + (x_1 - R) w_2 \cos\psi + w_2 x_2 \sin\psi - w_1(x_1 - R)\right]_f = 0$$

The analytical treatment follows Isaacs' work in reference 1. Isaacs first sets up the "main equation" similar to the Hamilton-Jacobi equation of the calculus of variations

$$\left.\begin{aligned} &\min_{\phi}\max_{\psi} H = 0 \\ &H = \sum_{i=1}^{n} V_i \dot{x}_i \end{aligned}\right\} \tag{20}$$

where $V_i = \frac{\partial V}{\partial x_i}$ and V is the Value of the game defined as $V = \min\limits_{\phi}\max\limits_{\psi} P$. Expansion of equations (20) yields

$$\min_{\phi}\max_{\psi}\left[w\phi\left(V_1 x_2 - V_2 x_1\right) + w_2\left(V_1 \cos\psi + V_2 \sin\psi\right) - V_1 w_1\right] = 0 \tag{21}$$

Optimal controls Φ^* and ψ^* are selected as

$$\Phi^* = -\operatorname{sgn} A$$

where

$$\left.\begin{aligned} &A = V_1 x_2 - V_2 x_1 \\ &\sin\psi^* = \frac{V_2}{\rho} \\ &\cos\psi^* = \frac{V_1}{\rho} \end{aligned}\right\} \tag{22}$$

where

$$\rho = \sqrt{V_1{}^2 + V_2{}^2}$$

Substituting the optimal controls into equation (21) produces the second form of the main equation

$$H = w\phi^*A + w_2\rho - w_1V_1 = 0$$

Introducing next the concept of retrograde time τ such that $\frac{d}{d\tau} = -\frac{d}{dt}$ and using the equation $\frac{dV_i}{d\tau} = \frac{\partial H}{\partial x_i}$, the retrograde path equations (RPE) become

$$\left.\begin{aligned} \frac{dx_1}{d\tau} &= -w\phi^*x_2 - w_2 \cos\psi^* + w_1 \\ \frac{dx_2}{d\tau} &= w\phi^*x_1 - w_2 \sin\psi^* \\ \frac{dV_1}{d\tau} &= -w\phi^*V_2 \\ \frac{dV_2}{d\tau} &= w\phi^*V_1 \end{aligned}\right\} \tag{23}$$

In order to integrate the RPE, the terminal conditions must be evaluated. The terminal manifold here is the locus of points where $J = 0$ for both players playing optimally. If parameters S_3 and S_2 are introduced such that

$$\left.\begin{aligned} x_1 &= R + S_3 \cos S_2 \\ x_2 &= S_3 \sin S_2 \end{aligned}\right\} \tag{24}$$

then S_2 and S_3 may be found from

$$\min_{\phi}\max_{\psi}\left\{S_3\left[w_2 \cos\left(S_2 - \psi\right) - w_1 \cos S_2 - Rw\phi \sin S_2\right]\right\} = 0$$

Choosing $\psi^* = S_2$ and $\phi^* = \mathrm{sgn}\,(\sin S_2)$ leads to $S_3 = 0$ and

$$S_2 = \sin^{-1}\left(\frac{w_2}{\sqrt{w_1^2 + R^2w^2}}\right) - \theta$$

where

$$\theta = \sin^{-1}\left(\frac{w_1}{\sqrt{w_1^2 + R^2w^2}}\right) = \cos^{-1}\left(\frac{Rw}{\sqrt{w_1^2 + R^2w^2}}\right)$$

Notice that S_2 is independent of S_3 and constant for a given problem since w_2, w_1, w, and R are specified constants. Since S_2 is constant, the terminal manifold is a line originating at the target point $(S_3 = 0)$ and extending outward to infinity on both sides of the x_1-axis; therefore, the solution need be found only for the right half plane.

Figure 7 shows the terminal lines for various ratios of w_2 to w_1, when $R/R_c = 0.6$. The lowest line, where $w_2/w_1 = 0$, corresponds to the case of an immobile evader. As w_2 increases, the play becomes more favorable to the evader. The terminal line moves upward reducing the region in which pursuer can capture evader. At $w_2 = w_1$, the evader can force $\dot{r}$ to zero by flying directly away from the pursuer; hence, termination can also occur on the x_1-axis. These two terminal lines then converge to one line again at

$$w_2 = \sqrt{w_1{}^2 + R^2 w^2}$$

For

$$w_2 > \sqrt{w_1{}^2 + R^2 w^2}$$

no solution exists; thus, the evader can now everywhere force $\dot{r} > 0$ and escape.

It has been observed that the game terminates whenever the evader is forced across the terminal line. But it may well be that the optimal paths actually terminate on only a part of the terminal line. The part on which termination may occur is called the "usable part of the terminal manifold." The boundary of the usable part will delineate between the two regions in which one player or the other has control over termination.

Consider now the situation when the evader is an infinitesimal distance above the terminal line C. The evader wishes to force termination to prevent his payoff from worsening further and hence wishes to force the component of $\dot{\vec{x}}$ normal to C to be downward or negative. Likewise, the pursuer wishes to prevent termination so that his payoff will improve. The boundary of the usable part will then be just the point where the normal component to the terminal line is zero under optimal play. Expressed mathematically, this condition becomes

$$\min_{\psi} \max_{\phi} \left(\dot{x}_1 \sin S_2 - \dot{x}_2 \cos S_2\right) = 0$$

Notice the reversal of the player's ordinary minimizing and maximizing roles. Substituting the equations of motion and solving for S_3 yields

$$S_3 = \frac{w_2 + w_1 \sin S_2 - wR \cos S_2}{w} \tag{25}$$

Equation (25) defines the boundary of the usable part as a point on each of the terminal lines. Figure 8 shows this boundary as a function of w_2 for each of the terminal lines presented in figure 7. The terminal line is usable for any value of S_3 less than the value given in equation (25) and nonusable for any larger value of S_3. The optimal path equations (eqs. (23)) may then be integrated back through the playing space from points on the usable part of the terminal line.

Before integrating the equations, the variables and optimal controls must be evaluated on the terminal line. The Value $\left(\min_{\phi} \max_{\psi} P\right)$ of the game is given by S_3. From Isaacs (ref. 1, ch. 4)

$$\frac{\partial V}{\partial S_K} = \sum_{i}^{n} \frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial S_K} = \sum_{i}^{n} V_i \frac{\partial x_i}{\partial S_K} \qquad (K = 2,3) \tag{26}$$

where x_i is given as a function of S_2 and S_3 in equations (24). Solving the two equations generated by equation (26) for V_1 and V_2 yields

$$V_1 = \cos S_2$$
$$V_2 = \sin S_2$$

on the terminal manifold. With V_1 and V_2 determined, the optimal controls may also be found as

$$\psi^* = S_2$$
$$\phi^* = \mathrm{sgn}\,(\sin S_2)$$

The retrograde equations and initial (retrograde) conditions now become

$$\left.\begin{aligned}
\frac{dx_1}{d\tau} &= -wx_2 - w_2 \frac{V_1}{\rho} + w_1 && x_1(0) = R + S_3 \cos S_2 \\
\frac{dx_2}{d\tau} &= wx_1 - w_2 \frac{V_2}{\rho} && x_2(0) = S_3 \sin S_2 \\
\frac{dV_1}{d\tau} &= -wV_2 && V_1(0) = \cos S_2 \\
\frac{dV_2}{d\tau} &= wV_1 && V_2(0) = \sin S_2
\end{aligned}\right\} \tag{27}$$

These equations may be integrated to yield

$$\left.\begin{aligned}
V_1 &= \cos\left(S_2 + w\tau\right) \\
V_2 &= \sin\left(S_2 + w\tau\right) \\
x_1 &= R \cos w\tau + \left(S_3 - w_2\tau\right) \cos\left(S_2 + w\tau\right) + \frac{w_1}{w} \sin w\tau \\
x_2 &= R \sin w\tau + \left(S_3 - w_2\tau\right) \sin\left(S_2 + w\tau\right) + \frac{w_1}{w} (1 - \cos w\tau)
\end{aligned}\right\} \qquad (28)$$

The complete solution is shown in figure 9 for the case $w_2 = 0.2$, $w_1 = 0.1$, $R = 3$, and $w = 0.2$. Remember that since the coordinate system is centered in the pursuer and moving with it, the optimal paths show the motion of the evader relative to the pursuer. The analytic solution divides the relative space into three areas with each area representing a type of solution for starting points within that area. The shaded region is a termination zone where the player that desires termination may force it to occur. In this zone the evader can always force $\dot{r} > 0$; hence, the game will terminate instantly and no optimal paths will result. The terminal line which separates the shaded region from the rest of the playing space was calculated from the expression

$$\min_{\phi} \max_{\psi} J = 0$$

and the points in the shaded region then satisfy

$$\min_{\phi} \max_{\psi} J > 0$$

The unshaded region of the playing space can be divided into two areas: one where nonunique solutions occur and one where unique solutions occur. Immediately above the terminal line is an area in which unique optimal solution for both players can be found. The paths represent optimal paths for any set of initial conditions in the region. These optimal paths will terminate on the crosshatched area representing the usable part of the terminal manifold. For example, for a set of initial conditions at point A, the solution is for the state to follow the path indicated until the game ends at point B. The pursuer is required to use a saturated control (full turn) but the evader is, nevertheless, able to prevent the payoff from reaching zero. The evader chooses his optimal strategy in accordance with equations (22) and (28). Notice that the evader will be pointed directly away from the target point along the terminal line at termination. If the evader would fly any other strategy, the pursuer would gain. Likewise, if the pursuer does not use saturated controls, the evader will gain.

The boundary of the unique region is the unique solution along which the payoff will be zero. That is, the pursuer can only force the evader to the target point by turning full right. For points above this path, the pursuer no longer needs to use saturated controls to obtain zero payoff; hence, the solutions become nonunique. The payoff for all initial conditions in this region is the same, namely, zero. Any trajectory that brings the evader to the target point is optimal. Because the initial conditions lie within the region that contains nonunique solutions, it is possible to arrive at the same terminal state by many different paths. Consequently, different nominal paths will give different converged solutions but will yield the same payoff.

Gradient Solution to the Simplified Differential Game

The gradient method was applied to the problem just solved by using the procedure described in the appendix. These results are shown in figures 10, 11, and 12.

Each calculation of the gradients $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$ is considered to be one iteration; however, each iteration may have from one to ten minor cycles. (See the appendix.) As a result, one must be very careful in comparing iteration times and number of iterations for different cases. In figure 10 an initial condition is chosen in the unique trajectory region. The nominal path, an intermediate nonoptimal path, and the final converged solution are presented. A comparison with figure 9 shows the extremely good agreement with the optimal path found analytically. The method required 41 iterations to converge and averaged some 9 seconds per iteration using a Control Data 6600 computer.

Figures 11 and 12 illustrate the solutions in the nonunique region for different nominal paths. The gradient method required about 5 seconds per iteration for these cases. Figure 11(a) shows the nominal path (straight flight for both players) and iterated paths leading to a converged solution after seven iterations. Since this converged path is known to be in the nonunique region, a new game with unique solutions can be formed for this region. The payoff for this game is the time to reach the target point. The evader still maximizes and the pursuer minimizes this payoff with termination remaining $\dot{r} = 0$. This time-optimal game may be solved by using the same approach presented in the general theory above. One simply increases the dimension of the state vector by one with x_{n+1} being time and $\dot{x}_{n+1} = 1$. A typical solution for this time and position optimal game is shown in figure 11(b) by using the optimal solution from figure 11(a) as the starting point and iterating for 64 iterations. Notice that the pursuer now initiates a full turn followed by straight flight to bring the evader down the x_1-axis to the target point. This type of maneuver is found to be the solution to the time-optimal game in reference 1.

Figure 12 shows the results when a different nominal trajectory is used with the same initial conditions as in figure 11. In figure 12(a), the nominal path results from straight flight by the evader ($\psi = 0$) and a full right turn by the pursuer ($\phi = +1$). The resulting position optimal path after eight iterations is quite different from that in figure 11(a) although the payoff is the same, a fact which further verifies the nonuniqueness of this region. But reverting to the time-optimal game should yield unique results regardless of the nominal path chosen. A comparison of figures 11(b) and 12(b) show reasonable agreement in the solutions. For both nominal paths the time-optimal strategy is shown to be a full turn followed by straight flight to the target point.

Application to a Two-Airplane Game

As another example of the application of the gradient method, the results are presented for the two-airplane problem discussed in the section entitled "Problem Statement." The equations of motion are written for an engagement at constant altitude. To simplify computation, a coordinate system centered in and moving with the attacker is used, the x_1 coordinate being measured along the pursuer's velocity vector.

The relative position equations are given by

$$\left.\begin{aligned} \dot{x}_1 &= \frac{u_1 a_{n,A} g x_2}{V_A} - V_A + V_B \cos \theta_B \\ \dot{x}_2 &= \frac{-u_1 a_{n,A} g x_1}{V_A} + V_A \sin \theta_B \end{aligned}\right\} \tag{29}$$

The relative angle equation is

$$\dot{\theta}_B = \frac{g v_1 a_{n,B}}{V_B} - \frac{g u_1 a_{n,A}}{V_A} \tag{30}$$

The speed equations for the two airplanes are

$$\left.\begin{aligned} \dot{V}_A &= \frac{T_A}{m_A} u_2 - \left(\frac{\rho_A S_A}{2m_A}\left(C_{D,o}\right)_A + \frac{\rho_A S_A}{2m_A}\left(C_{D,i}\right)_A {u_1}^2\right) {V_A}^2 \\ \dot{V}_B &= \frac{T_B}{m_B} v_2 - \left(\frac{\rho_B S_B}{2m_B}\left(C_{D,o}\right)_B + \frac{\rho_B S_B}{2m_B}\left(C_{D,i}\right)_B {v_1}^2\right) {V_B}^2 \end{aligned}\right\} \tag{31}$$

where $C_{D,o}$ is the drag coefficient at zero angle of attack and $C_{D,i}$ is a coefficient representing induced drag resulting from the turn.

Since $a_{n,A}$ and $a_{n,B}$ are maximum turning accelerations and T_A and T_B are maximum thrusts, each airplane has two controls with the following constraints:

$$\text{(Full left turn)}\quad -1 \leqq u_1, v_1 \leqq +1 \quad \text{(Full right turn) turning control}$$

$$0 \leqq u_2, v_2 \leqq +1 \quad \text{(Full thrust) thrust control}$$

The aerodynamic variables ρ, S, $C_{D,o}$, $C_{D,i}$, T, and a_n are chosen for a particular flight condition and then held constant throughout the engagement. For the results presented here, the airplanes are considered to have the same aerodynamic characteristics – those of a fighter airplane operating at 9144 meters altitude. The drag characteristics are $\dfrac{C_{D,o}\rho S}{2m} = 4.968 \times 10^{-7}/\text{meter}$ and $\dfrac{C_{D,i}\rho S}{2m} = 8.29 \times 10^{-7}/\text{meter}$. Maximum thrust per unit mass is 0.03581 m/sec^2 and maximum turning acceleration is 29.44 m/sec^2 (limited to a 3g turn). To prevent a standoff situation, airplane A is given an initial velocity advantage over airplane B. These initial conditions are $V_A = 230$ m/sec, $V_B = 190.5$ m/sec, $\theta_B = 90^{\circ}$, $x_2 = 4115$ meters, and $x_1 = 7315$ meters.

To formulate the problem as a game, let one player A be the pursuer and the other player B, the evader. The payoff, as in the previous game, will be the distance from a target point $(x_2=0, x_1=R)$ directly in front of the pursuer and where $R = 915$ meters. The pursuer will attempt to minimize the distance and force $P = 0$ while the evader tries to maximize the distance. For this game the contours of constant payoff are ellipses centered around the pursuer's target point as illustrated in figure 2. Thus, $P = C_1 x_2^2 + C_2(x_1 - R)^2$ where the constants were chosen as $C_1 = 16.8 \times 10^{-6}$ and $C_2 = 10.76 \times 10^{-6}$ representing the boundaries of a capture zone for a typical fighter airplane. As in the previous example, only one pass will be considered, the game terminating at the first minimum of the payoff.

The gradient-method results for a lateral pursuit case using two different nominal controls are shown in figures 13 and 14. For visualization purposes the results are presented in absolute coordinates with the paths marked at 5-second intervals. The dashed lines indicate the nominal trajectories for each airplane whereas the solid lines are the converged optimal paths. The nominal controls in figure 13 shows the pursuer and evader flying straight at intermediate thrust and with the pursuer turning right after 20 seconds. The converged path after 28 iterations has the evader turning full right at minimum thrust and the pursuer using a combination of turns at minimum thrust to capture. Figure 14 shows the same pursuer nominal with the evader nominally turning left and then proceeding in straight flight. After 47 iterations the converged trajectory shows the evader using a full left turn to meet the pursuer head on. Both airplanes again use minimum thrust.

These two results are considered examples of nonunique solutions since widely differing strategies yield the same payoff. The payoff for figure 13 is 6.15×10^{-4} and for figure 14 is 2.81×10^{-4}. When compared with an initial payoff of 725.7, these two solutions are extremely close. Notice that since the payoff was independent of the relative angle, sizable angular differences could occur without changing the payoff. In both of these solutions, the converged evader strategy is a maximum turn at minimum thrust. Nevertheless, the pursuer's velocity advantage enables him to close in on the evader to the desired range.

The one-pass restriction ($\dot{P} \leqq 0$) in the games considered in this paper is more important than may appear at first glance. In the simplified game solved earlier, this restriction generated the line on which termination could occur. The effects there could readily be seen in that points where a solution might not occur were eliminated. However, in this larger problem the effect is not so easily visualized. One effect is to prevent any type of swerve maneuver such as in the homicidal chauffeur problem of reference 1 in which the pursuer initially turns away and then swings back to capture. The airplanes here are forced to continually close in on one another, or nearly so. Depending, of course, on the initial conditions, one can reasonably argue that the pursuer can always gain if this one-pass restriction is lifted. However, the authors believe this restraint is not unrealistic but often representative of actual pursuit evasion.

CONCLUDING REMARKS

A gradient method has been presented in this paper which may be applied to solving general zero-sum differential games. This method, a first step in developing a computational capability in game theory, is applicable to nonlinear multidimensional game problems representing realistic combat between two airplanes. Problems of this magnitude previously could not be solved. The technique requires selecting nominal controls which are then improved iteratively by a scheme consisting of a minimization phase followed by a minimizing-maximizing (min-max) phase.

The analytic solution to a simple differential game which is analogous to the particular aerial combat problem studied has been presented to give insight into the nature of the higher dimensional solution. This analytic solution revealed that different nominal engagements often produce different final solutions for the same initial conditions. This condition occurs because there is a region of space in which nonunique answers exist. Computational results have been presented that illustrate this behavior for the problem solved analytically. The method was then extended to a larger more realistic problem of two airplanes at constant altitude which also was found to exhibit this nonunique feature.

The question of types of solutions to large-dimensional games is a difficult one. When no analytic solution is available, it is often difficult to determine whether solutions are min-max, max-min, saddle points, nonunique, or perhaps local solutions. Nevertheless, any method which gives a solution to these higher order problems could be a valuable tool for studying such things as the effects of aircraft performance parameters on the outcome of an aerial combat engagement.

Langley Research Center,
National Aeronautics and Space Administration,
Hampton, Va., September 13, 1971.

APPENDIX

COMPUTATIONAL ASPECTS

The theory described in the section entitled "General Theory" has been implemented on the Control Data 6600 digital computer at the Langley Research Center in a FORTRAN program. Figure 15 shows a diagram of the steps presented below. The method starts with a nominal control table and then improves this estimate by updating the control table to move in the steepest descent-ascent direction. The method has evolved into the following format:

(a) Select a nominal control table as a function of time for each player. (Use of the digital computer requires the control function to be discretized. Intervals of 1.0 and 0.5 second have generally been used.)

(b) Integrate the equations of motion from the given initial conditions by using the control tables from step (a). Integrate simultaneously the transition matrix with initial conditions $\Psi(t_o,t_o,\sigma) = I$ and store the products $B^T(\tau,\sigma)\,\Psi(\tau,t_o,\sigma)$ and $C^T(\tau,\sigma)\,\Psi(\tau,t_o,\sigma)$ for each control table time point. The integration is stopped when the termination condition is satisfied.

(c) Calculate the control derivatives from equation (17) and update the controls at each point of the control table.

(d) Determine the best step size by changing K_u for each K_v to keep the payoff a minimum for each evader control change. Only the equations of motion will have to be integrated unless the termination time increases, in which case the transition matrix will also have to be integrated for the extended part of the table. To save computer time a large integration step size may be used here.

(e) Once the final control step size is found, update the control table and repeat steps (b) to (e) until convergence is evident. Convergence will be characterized by first, a negligible change in the payoff from iteration to iteration, and second, the vanishing of the control derivatives $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$. An automated criterion can be constructed by taking some suitable time average of $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$, for example, the root-mean-square value, and stopping the iterations whenever the measure falls below some specified number.

The first consideration in applying the formulas to a game involves the selection of an admissible nominal engagement. It is important to select nominal engagements in the vicinity of the global optimal solution. But this task becomes increasingly difficult as the game becomes more complicated since intuition and experience fail. In high-dimensional games and games with a complicated payoff, the payoff function may possess several local optimal solutions. In gradient methods the iterations follow the direction of the "slope" of the payoff; therefore, the controls will converge to the saddle point in whose neighborhood the iterations are initiated. A solution is to repeat the iterative sequence with various nominal strategies to search for the global optimal. The problem is complicated even more since games with terminal payoffs usually possess regions in which nonunique solutions exist. Since the payoff is a function of the terminal states alone, the players are only concerned with the final value of the payoff function and not with the paths used to arrive at this terminal state. In this case different nominals will give different optimal trajectories but will converge to the same payoff. A solution again is to try different nominals and compare solutions to search for the nonunique region. Another alternative is to change to a time-optimal game by including a function of time in the payoff; this change results in unique optimal trajectories.

One of the first questions asked about any iterative procedure is that of convergence. Under what conditions is convergence guaranteed? In applying the present gradient technique to differential games, the iteration sequence is found to be of extreme importance in assuring convergence. For purposes of explanation, consider a scalar problem where a function $P(u,v)$ is to be minimized with respect to u and maximized with respect to v. Figure 16 shows contours of constant P plotted in the u,v plane. The intersection of the two curves $\frac{\partial P}{\partial u} = 0$ and $\frac{\partial P}{\partial v} = 0$ represents the desired solution in function space.

The iteration scheme proceeds as follows: First, select nominal controls (u_0,v_0) and minimize P with respect to u keeping v at v_0 until $\frac{\partial P}{\partial u} = 0$ is reached; second, permit v to be updated but then minimize P with respect to u for the new v. This procedure will force the iteration to follow the $\frac{\partial P}{\partial u} = 0$ curve to the desired solution.

Alternate iteration schemes may be used successfully in certain instances. For example, a sequence alternating minimization and maximization might be used. For the problem considered here, the solution to the minimization phase is known to exist from physical considerations for any $\vec{v}$. However, the solution to a maximization phase at constant $\vec{u}$ does not generally exist. Thus, the sequence shown in figure 16 is the one which appears to be applicable to multidimensional problems since P can be chosen so that a definite minimum exists. But this sequence can guarantee only a max-min solution. A saddle-point solution may occur but is not insured.

In the functional problems considered in this report, K_u and K_v are selected to follow the iteration scheme illustrated in figure 16. Generally, K_u and K_v are selected initially so that the control derivatives $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$ are less than one-tenth of the maximum control range. During the minimization phase, K_v is set to zero and iterations are performed only on the $\vec{u}$ control. During this phase, K_u is adjusted to give the largest change in payoff for each evaluation of the gradients. Upon entrance to the min-max phase, both $\vec{u}$ and $\vec{v}$ are updated by using the initial K_u and K_v. Then K_u is adjusted to give the minimum of the payoff by using the new updated $\vec{v}$ control. That is, for each step in the $\vec{u},\vec{v}$ plane, the updating of $\vec{u}$ is adjusted to force the payoff back to the minimum curve where $\frac{\partial P}{\partial u} = 0$.

As indicated in step (d) of the computing procedures, the final time may vary significantly from iteration to iteration. This variable time format is handled through an extrapolation and correction procedure in the variable-step-size routine. The procedure is illustrated in figure 17. Upon entrance to the routine, begin an iteration sequence to determine the best step size (that is, the best multiple or fraction of K_u) as discussed above. For $\Delta\sigma = 0$ let the terminal time be designated by $t_{f,0}$. By using the updated control ($\Delta\sigma = 0$), the equations of motion (1) are now integrated, extrapolation on the control being used if necessary. The integration is stopped when the termination condition is reached at some time $t_{f,1}$. If $t_{f,1} > t_{f,0}$, then as soon as time exceeds $t_{f,0}$ the transition matrix equations are integrated, and the control variations $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$ are calculated for the extended part of the table. The control variations $\frac{\partial \vec{u}}{\partial \sigma}$ and $\frac{\partial \vec{v}}{\partial \sigma}$ are then added as a correction to the extrapolated control values when the controls are updated. Because the time extensions between cycles are usually not large, this procedure helps to minimize any errors due to extrapolation.

REFERENCES

1. Isaacs, Rufus: Differential Games. John Wiley & Sons, Inc., c.1965.
2. Bryson, A. E.; and Denham, W. F.: A Steepest-Ascent Method for Solving Optimum Programming Problems. Trans. ASME, Ser. E.: J. Appl. Mech., vol. 29, no. 2, June 1962, pp. 247-257.
3. Taylor, Lawrence W., Jr.; and Rees, Neville W.: Application of the Epsilon Technique to a Realistic Optimal Pursuit-Evasion Problem. Paper presented at Fourth Hawaii International Conference on System Sciences (Honolulu, Hawaii), Jan. 1971.
4. Baron, Sheldon; Chu, Kai-Ching; Ho, Yu-Chi; and Kleinman, David L.: A New Approach to Aerial Combat Games. NASA CR-1626, 1970.
5. Leitmann, George, ed.: Optimization Techniques. Academic Press, Inc., 1962.
6. Ogata, Katsuhiko: State Space Analysis of Control Systems. Prentice-Hall, Inc., c.1967.

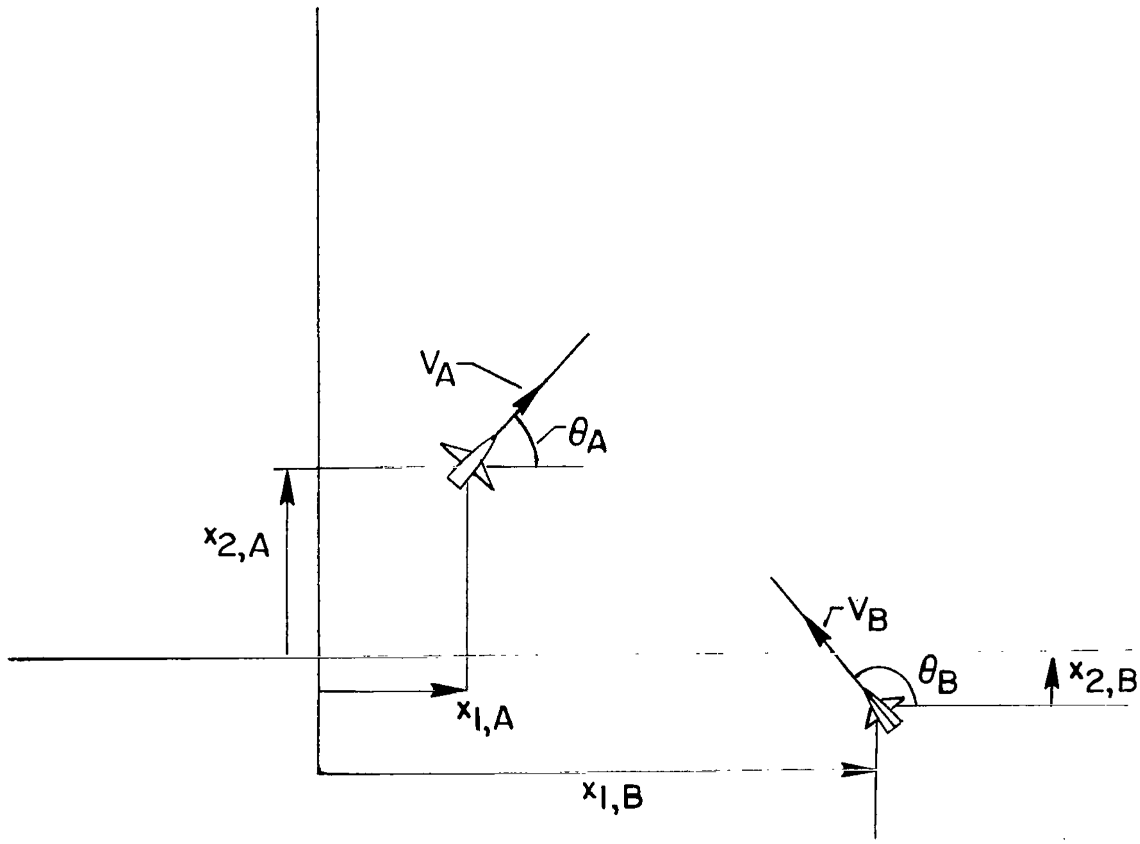


Figure 1.- Geometry of a simplified pursuit-evasion model.

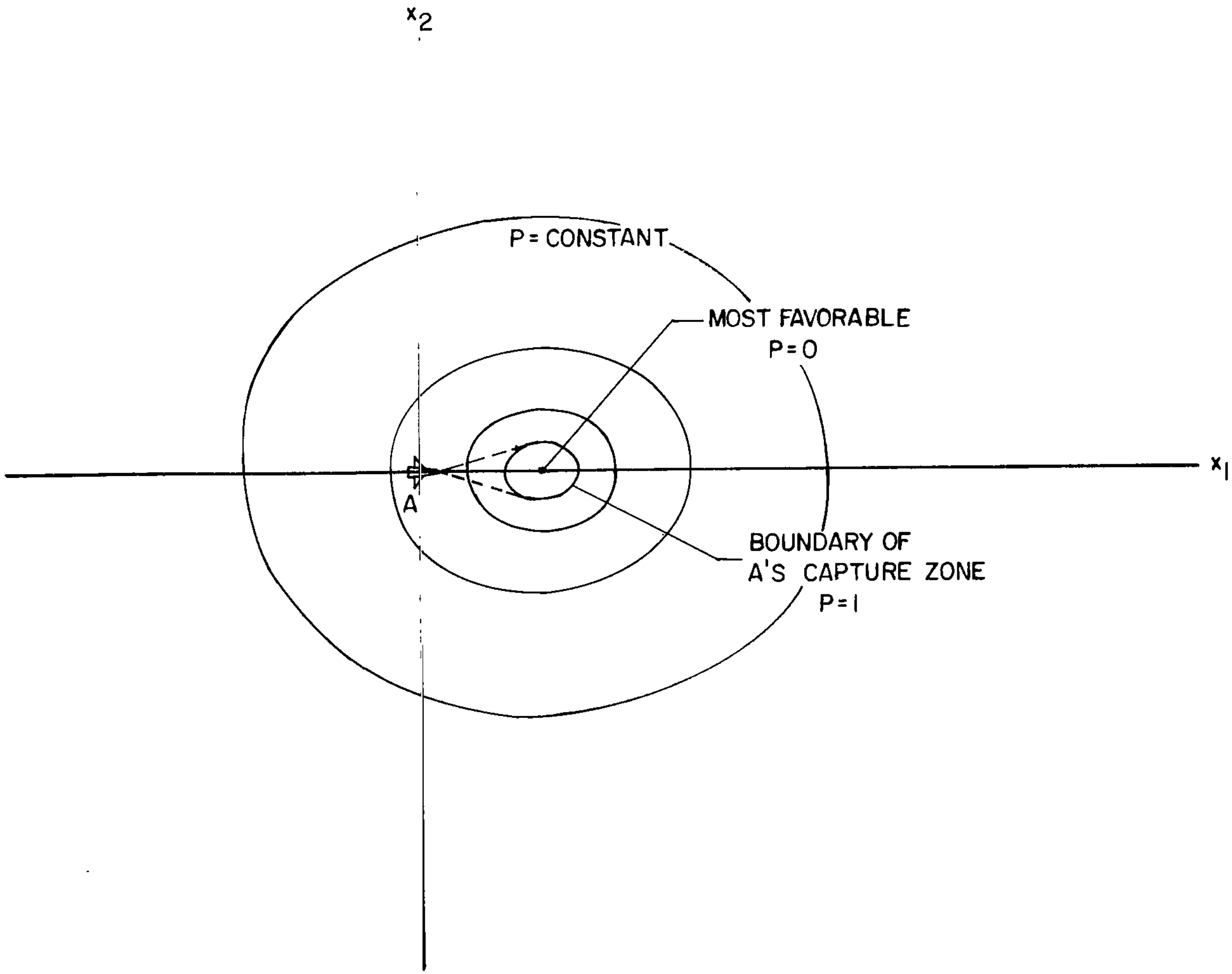


Figure 2.- An example ordering of the airplane positions for $\theta_A = \theta_B$ and $V_A = V_B$.

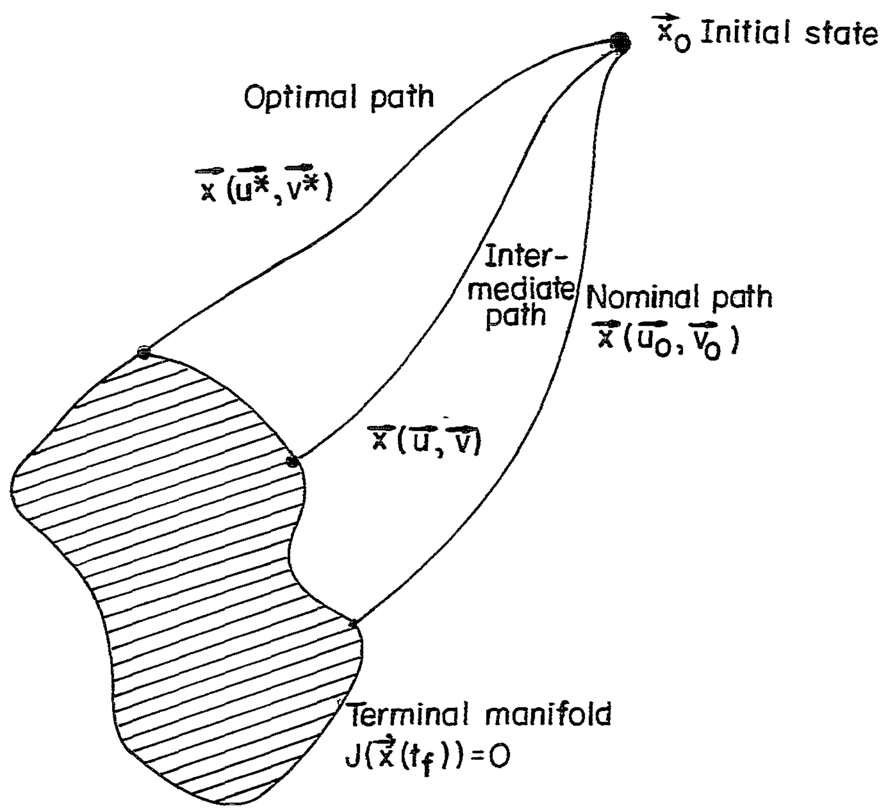


Figure 3.- Path iterations in state space.

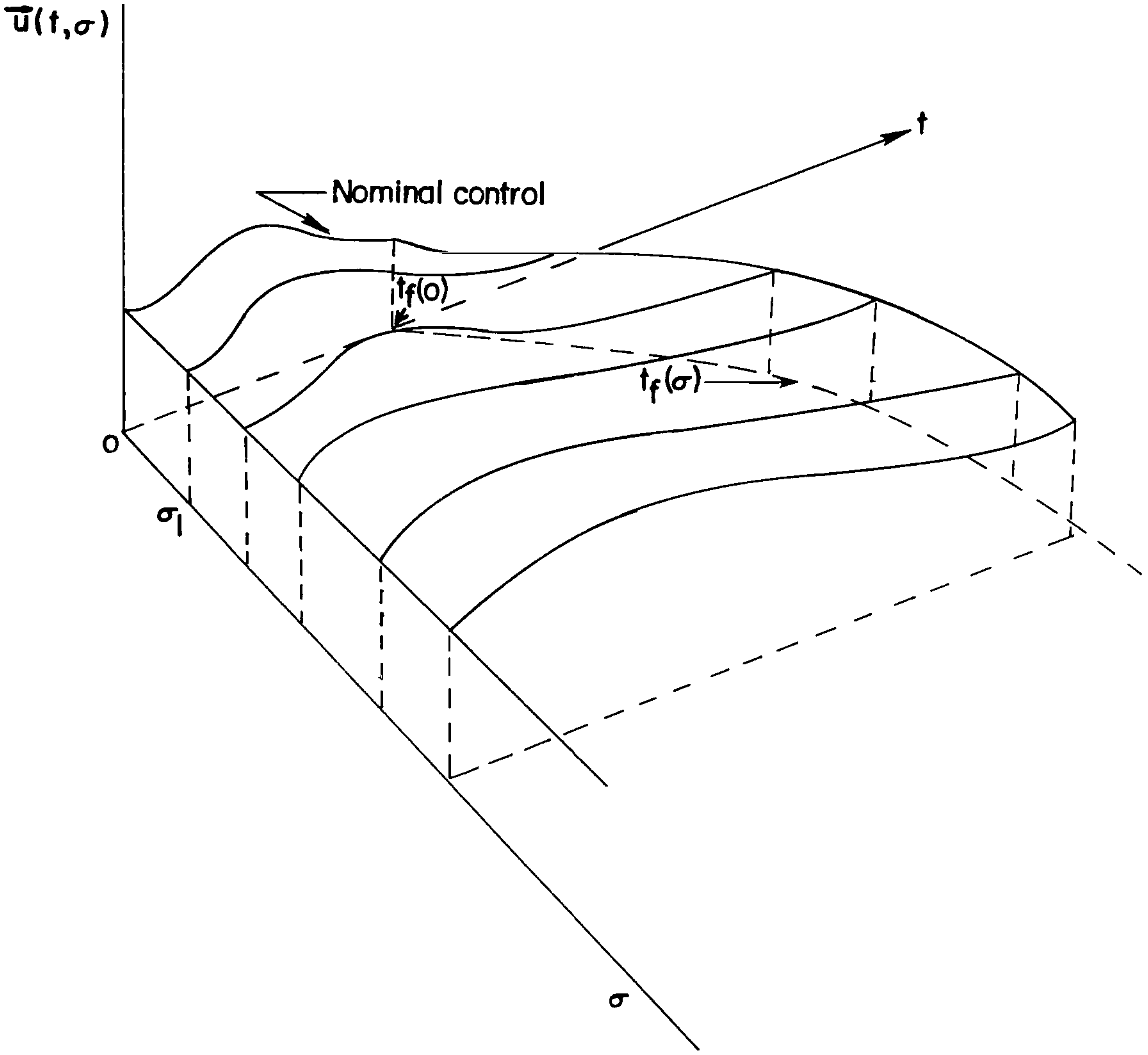


Figure 4.- Iterative process for typical control.

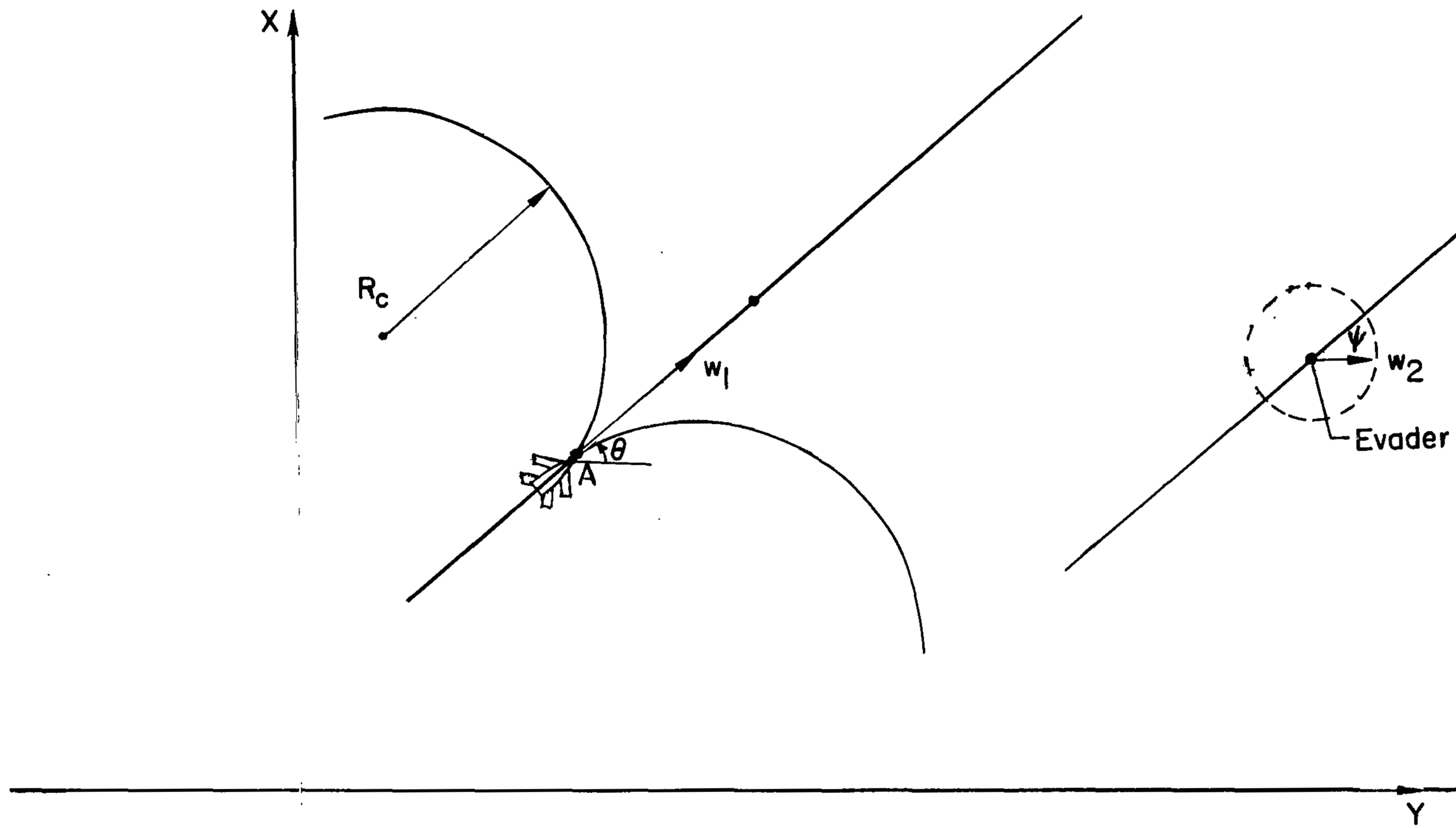


Figure 5.- Inertial coordinate geometry for the two-dimensional differential game.

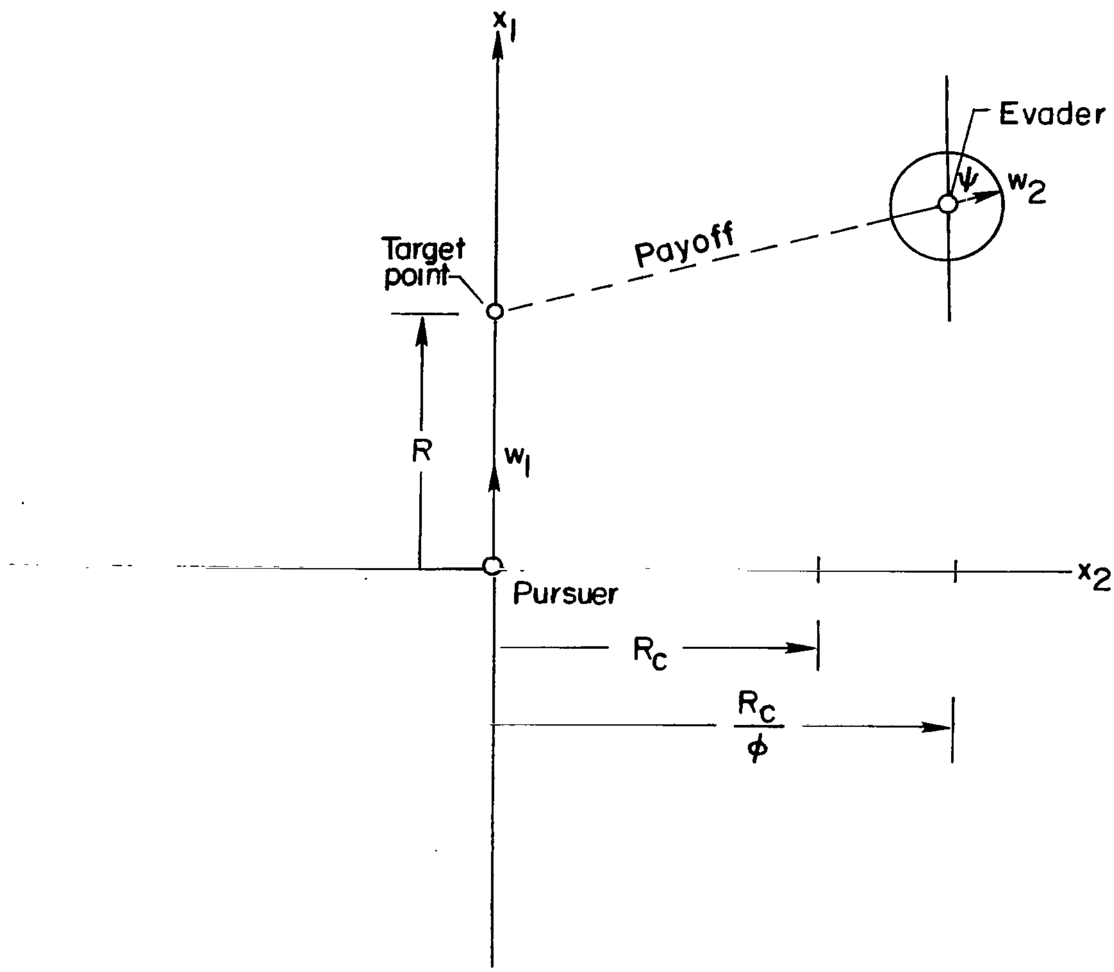


Figure 6.- Relative coordinate geometry for the two-dimensional differential game.

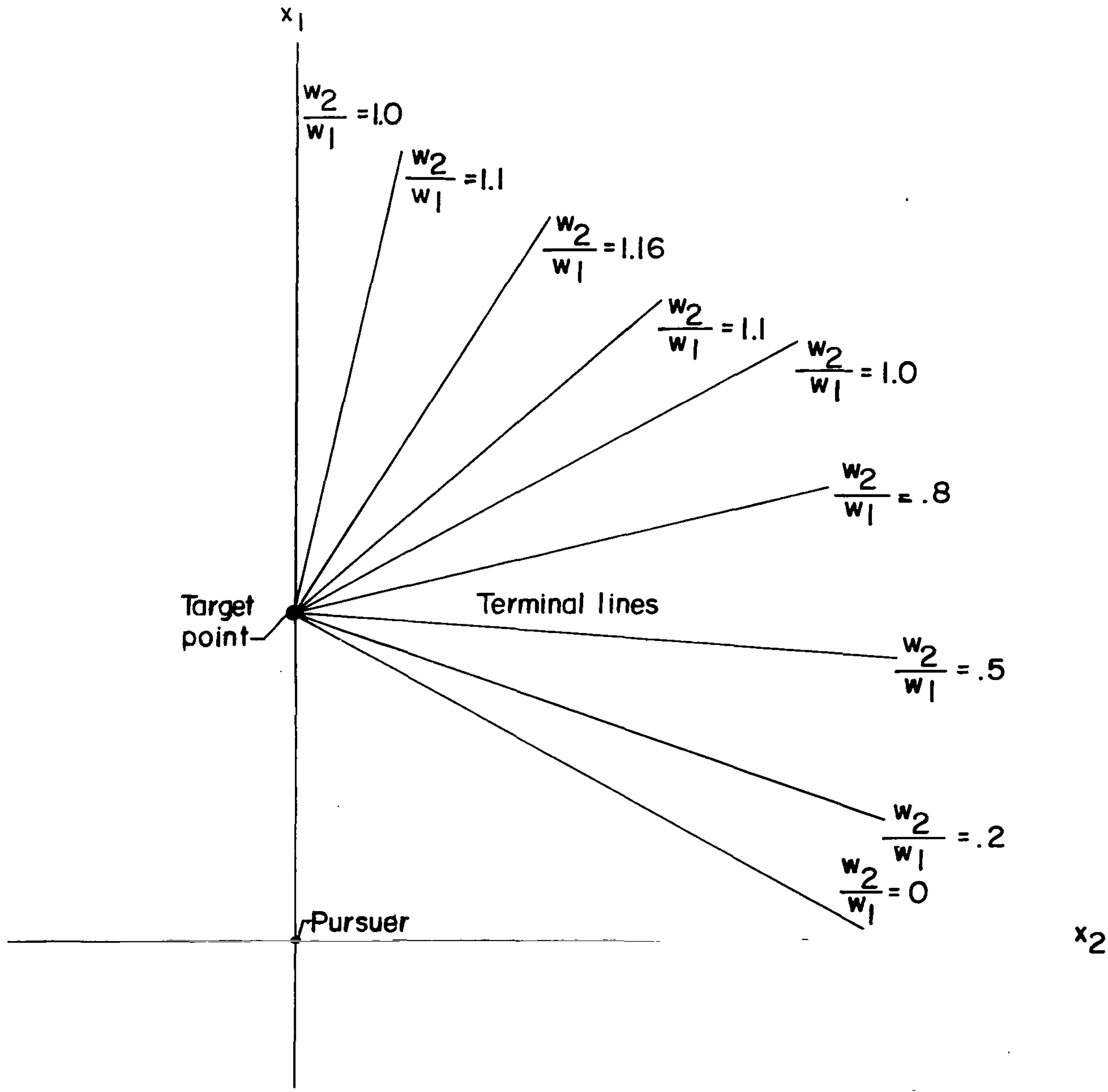


Figure 7.- Terminal lines for various velocities. $R/R_c = 0.6$.

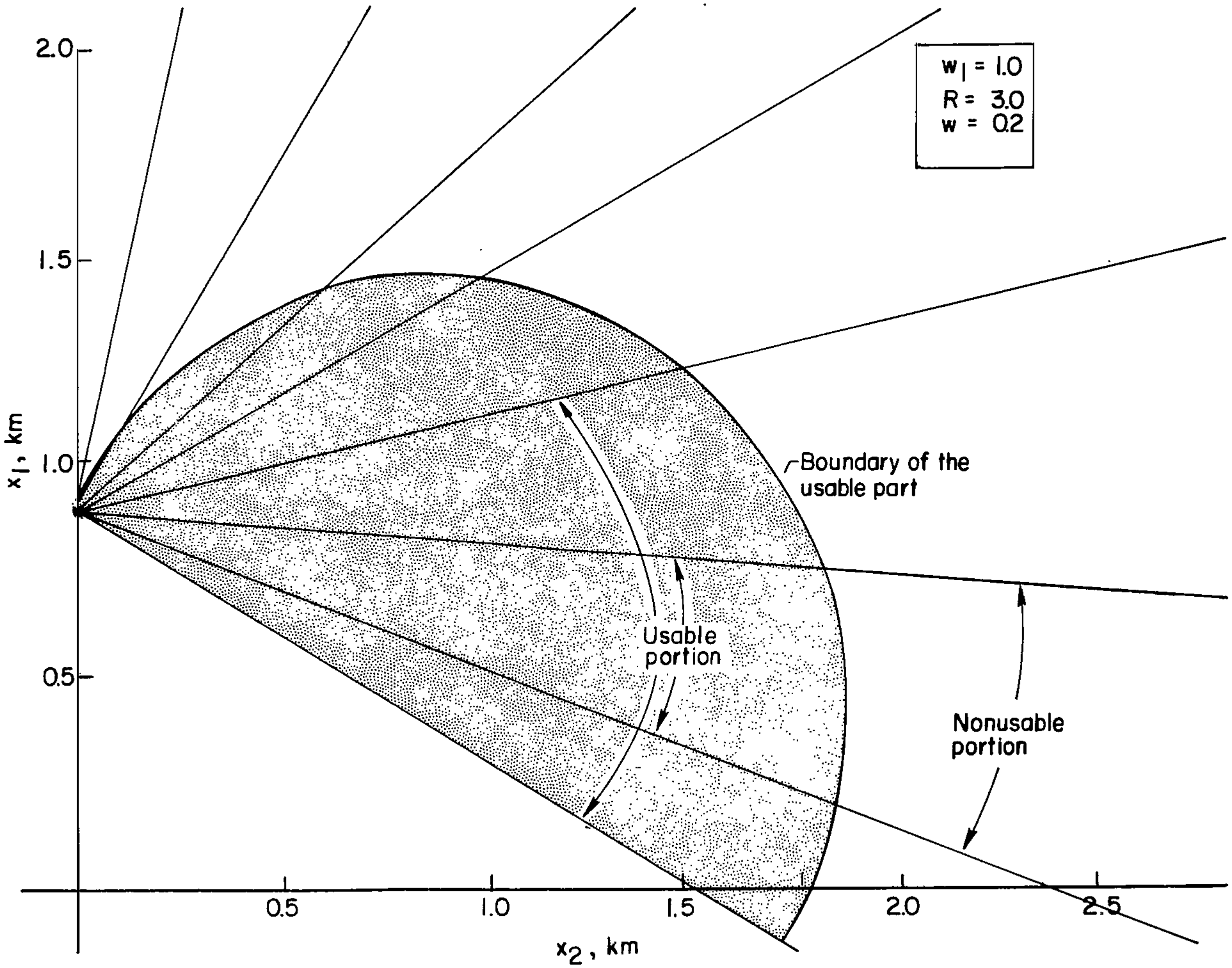


Figure 8.- Usable part of the terminal lines.

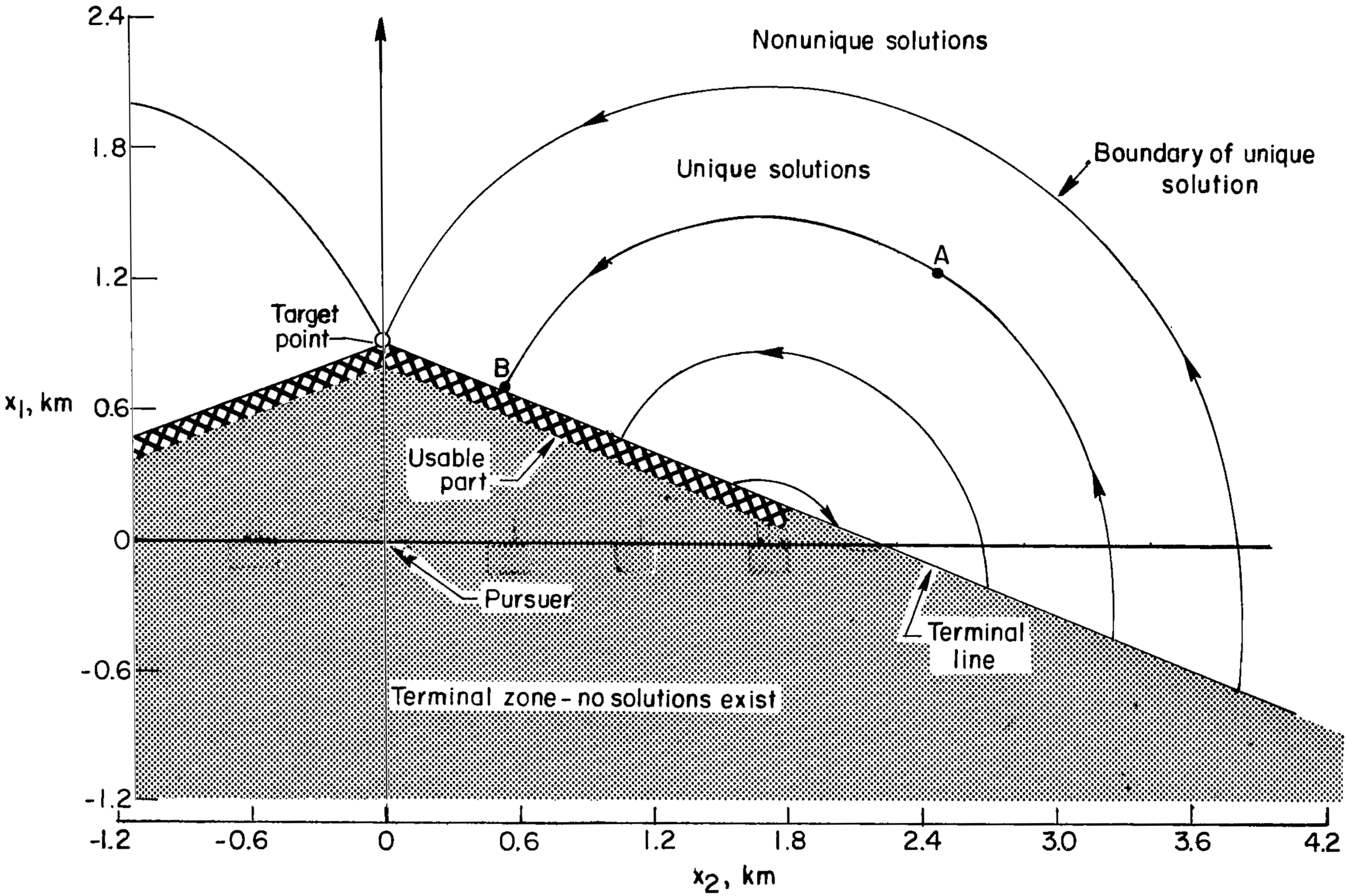


Figure 9.- Analytic solution for the two-dimensional game.

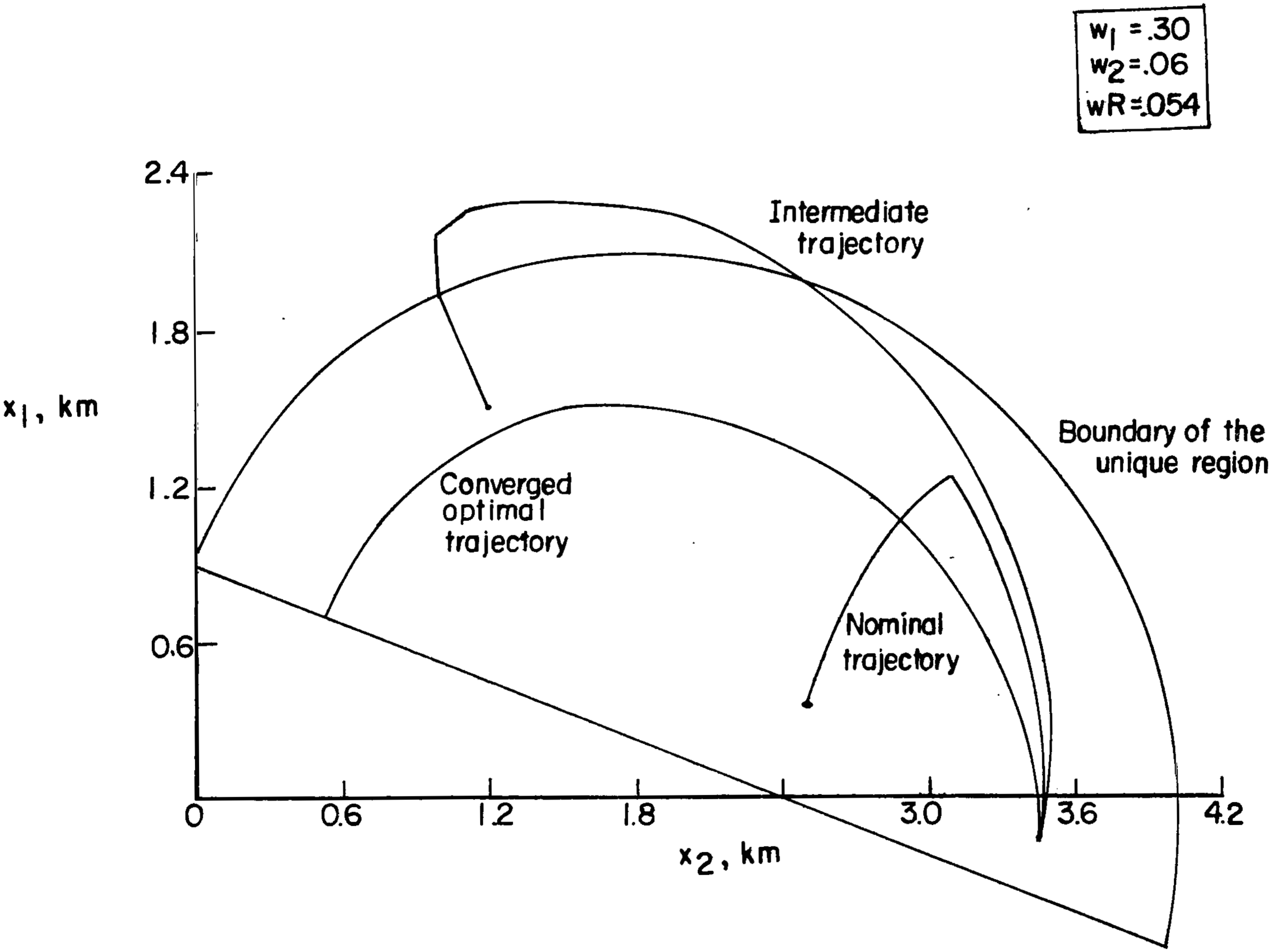
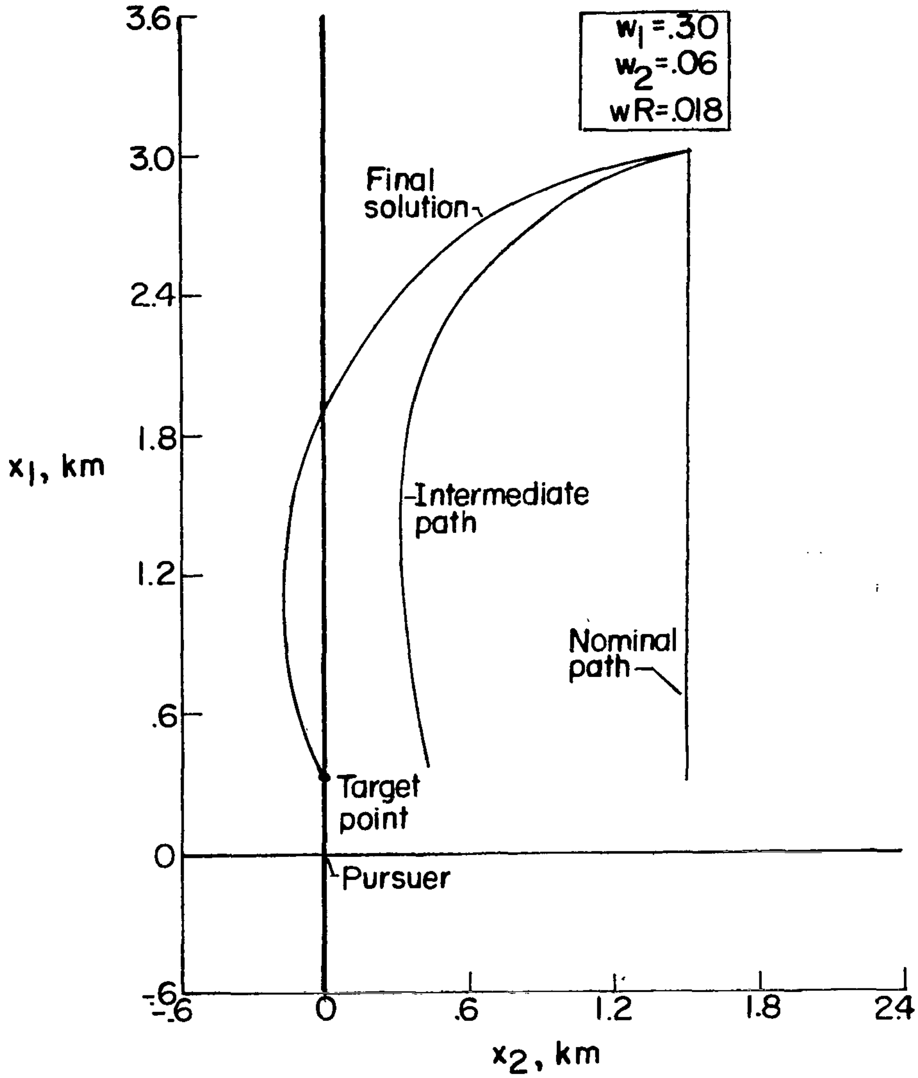
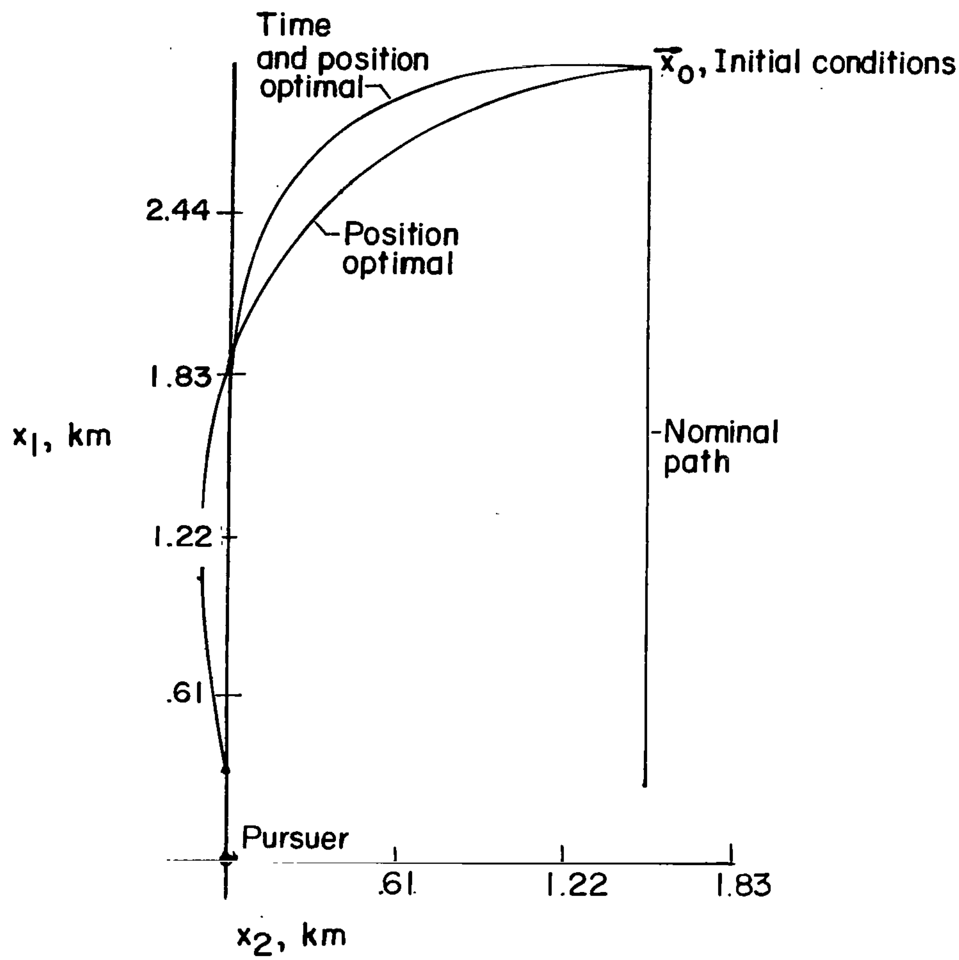


Figure 10.- Gradient method results for the unique region.



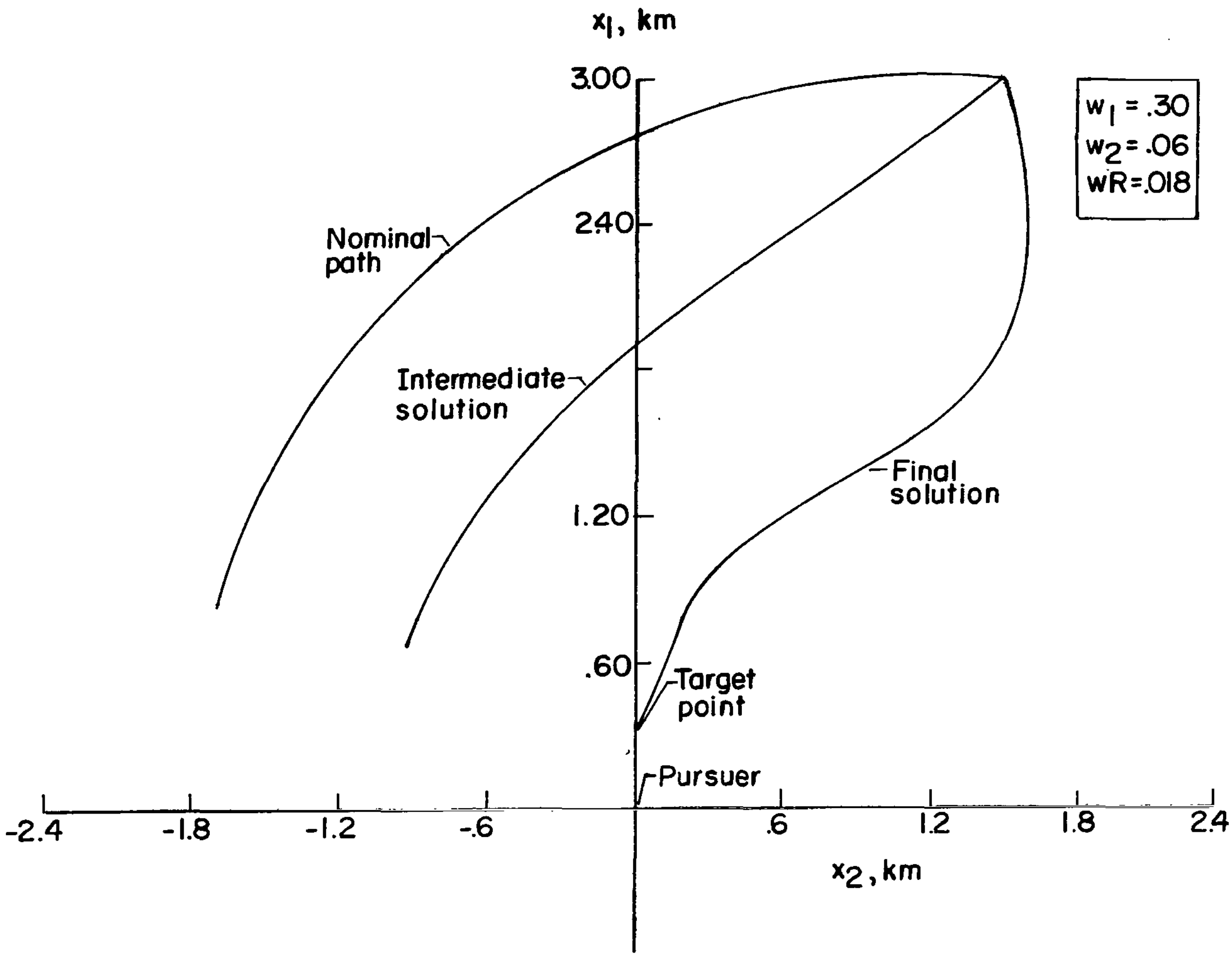
(a) Position-optimal game.

Figure 11.- Gradient solution for the nonunique region using a straight flight nominal.



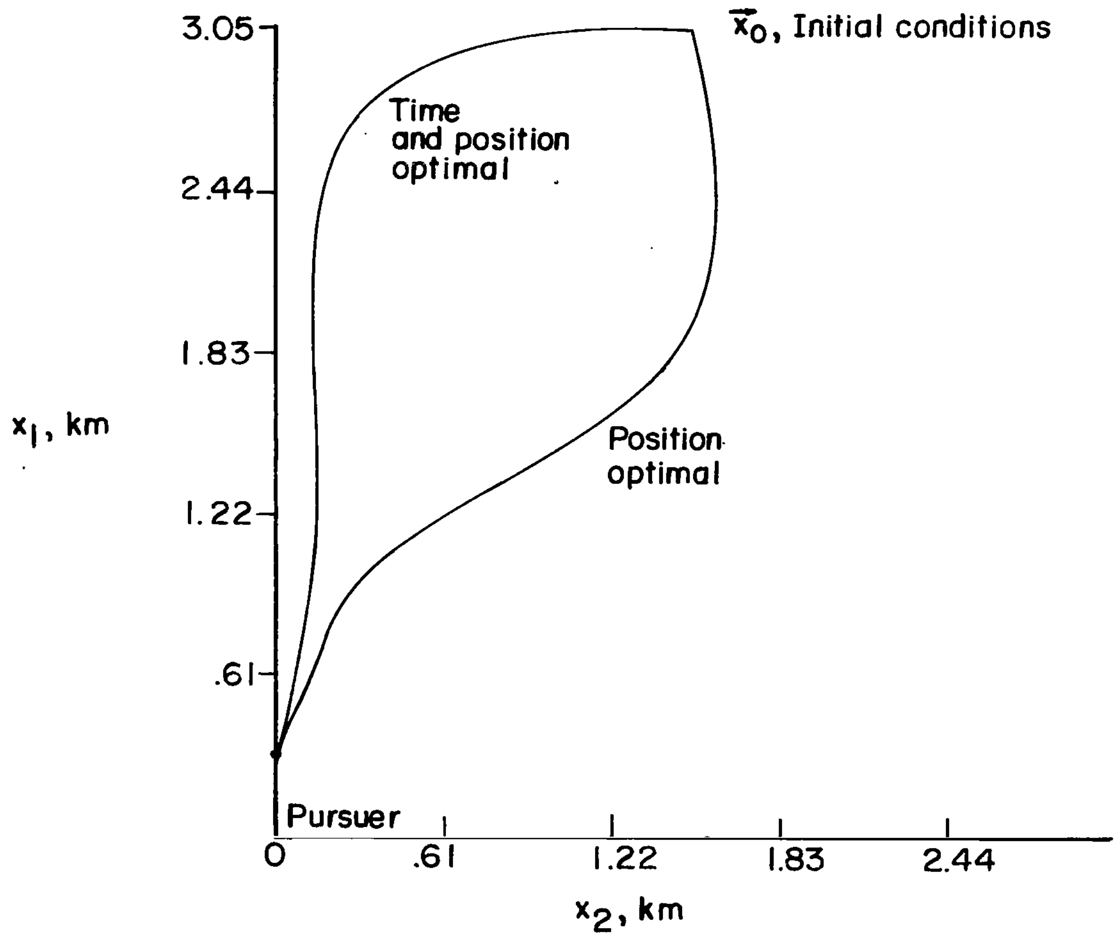
(b) Time-optimal game.

Figure 11.- Concluded.



(a) Position-optimal game.

Figure 12.- Gradient method results for the nonunique region using a turning nominal.



(b) Time-optimal game.

Figure 12.- Concluded.

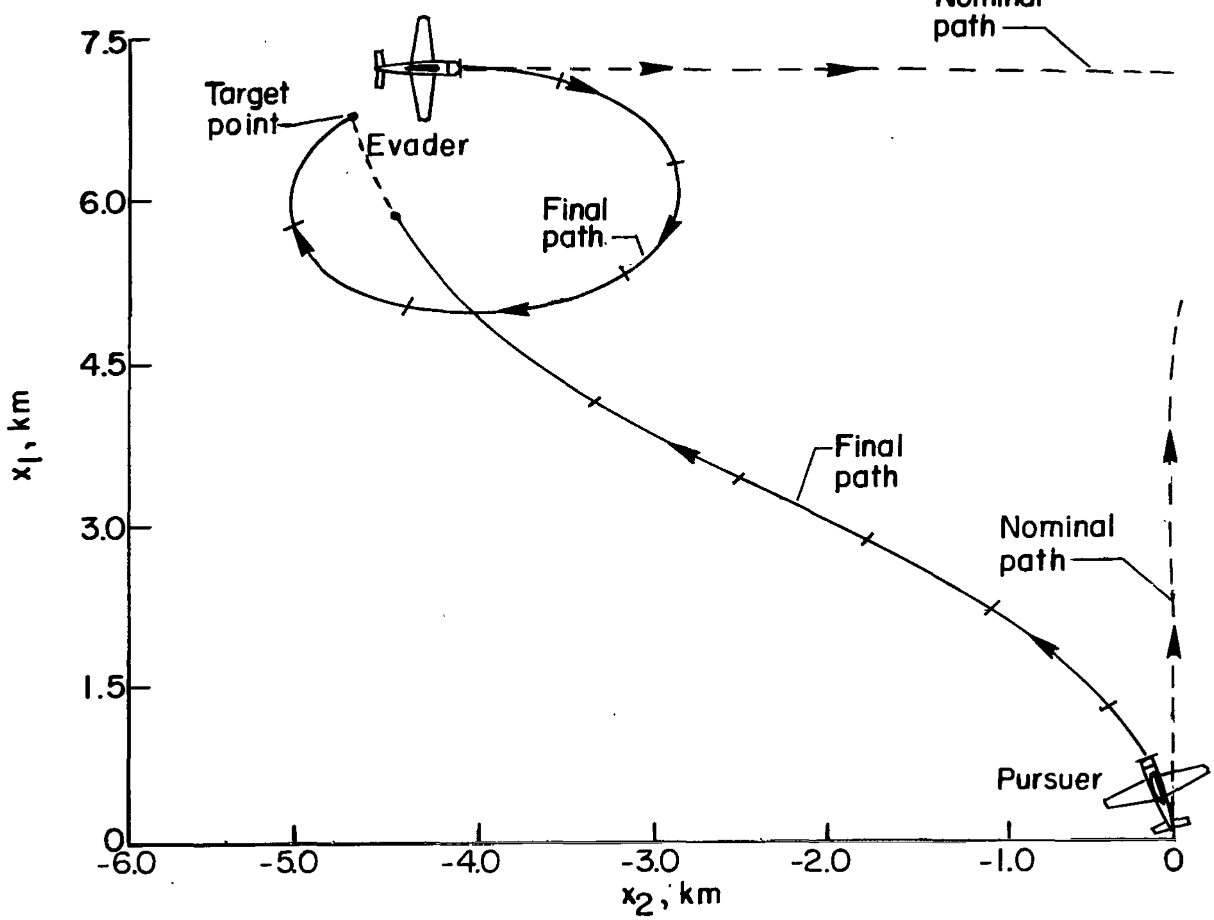


Figure 13.- Gradient results for the two-airplane game using a straight flight nominal.

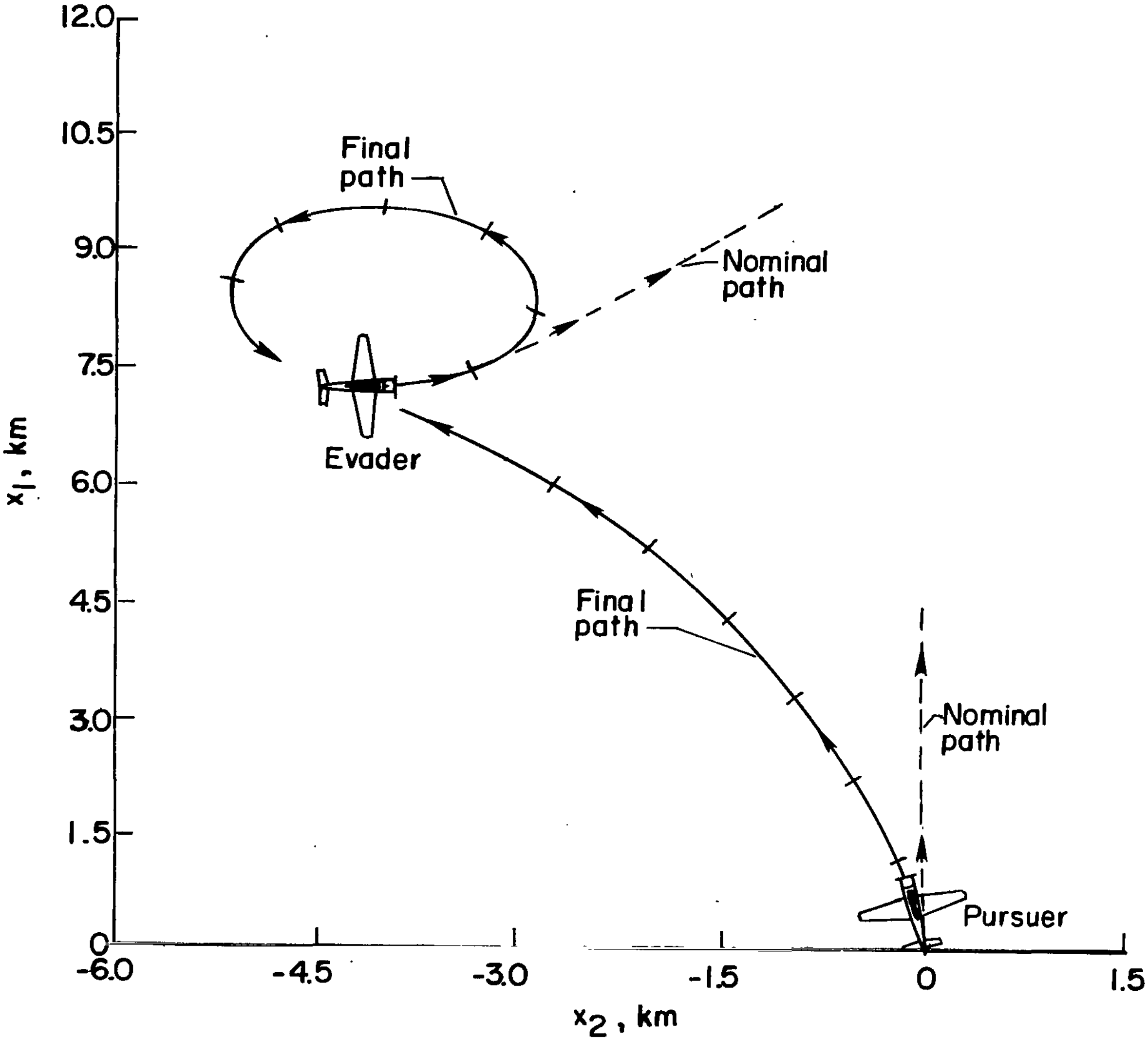


Figure 14.- Gradient results for the two-airplane game using a turning evader nominal.

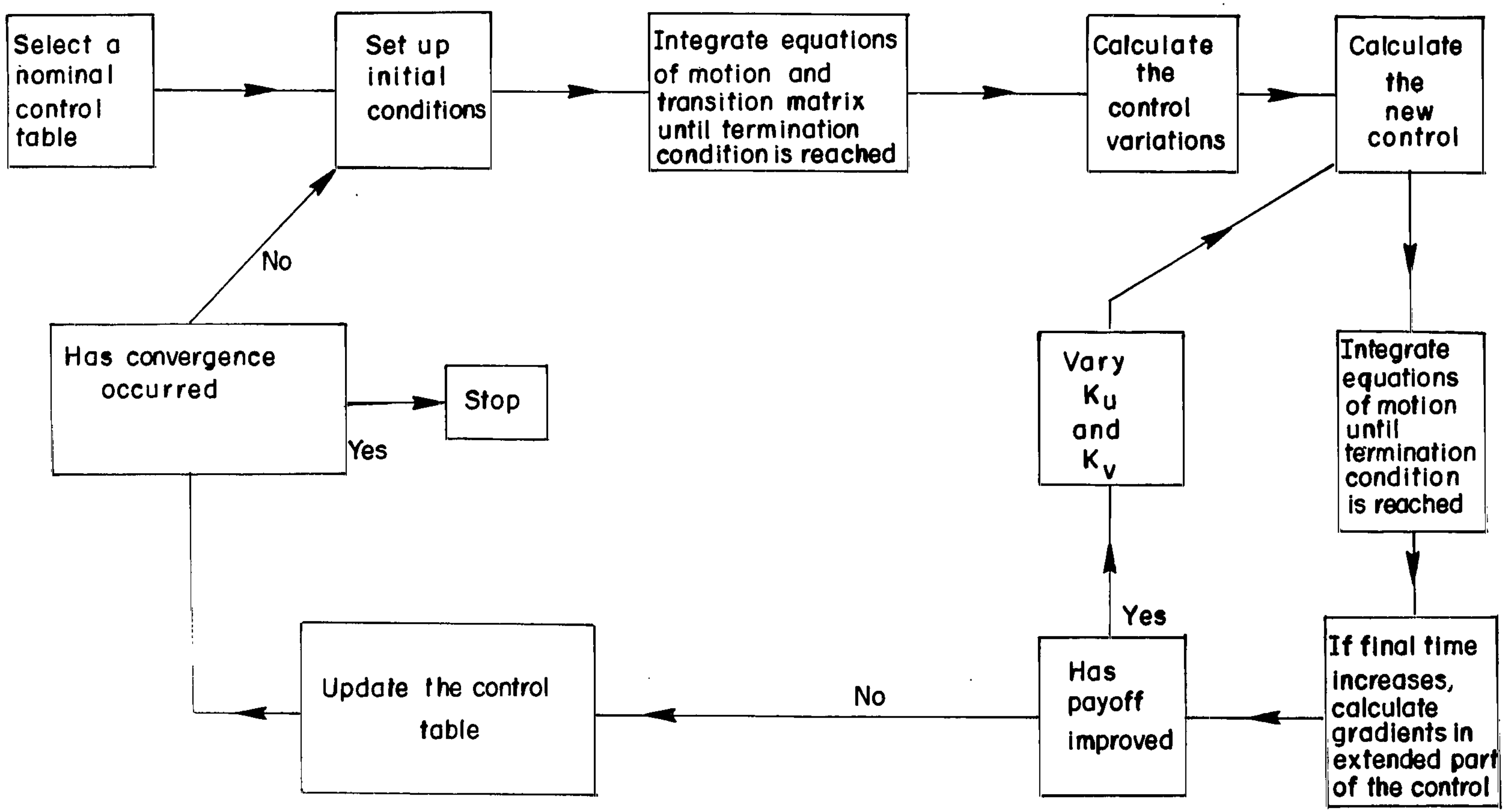


Figure 15.- Flow chart of the computing steps.

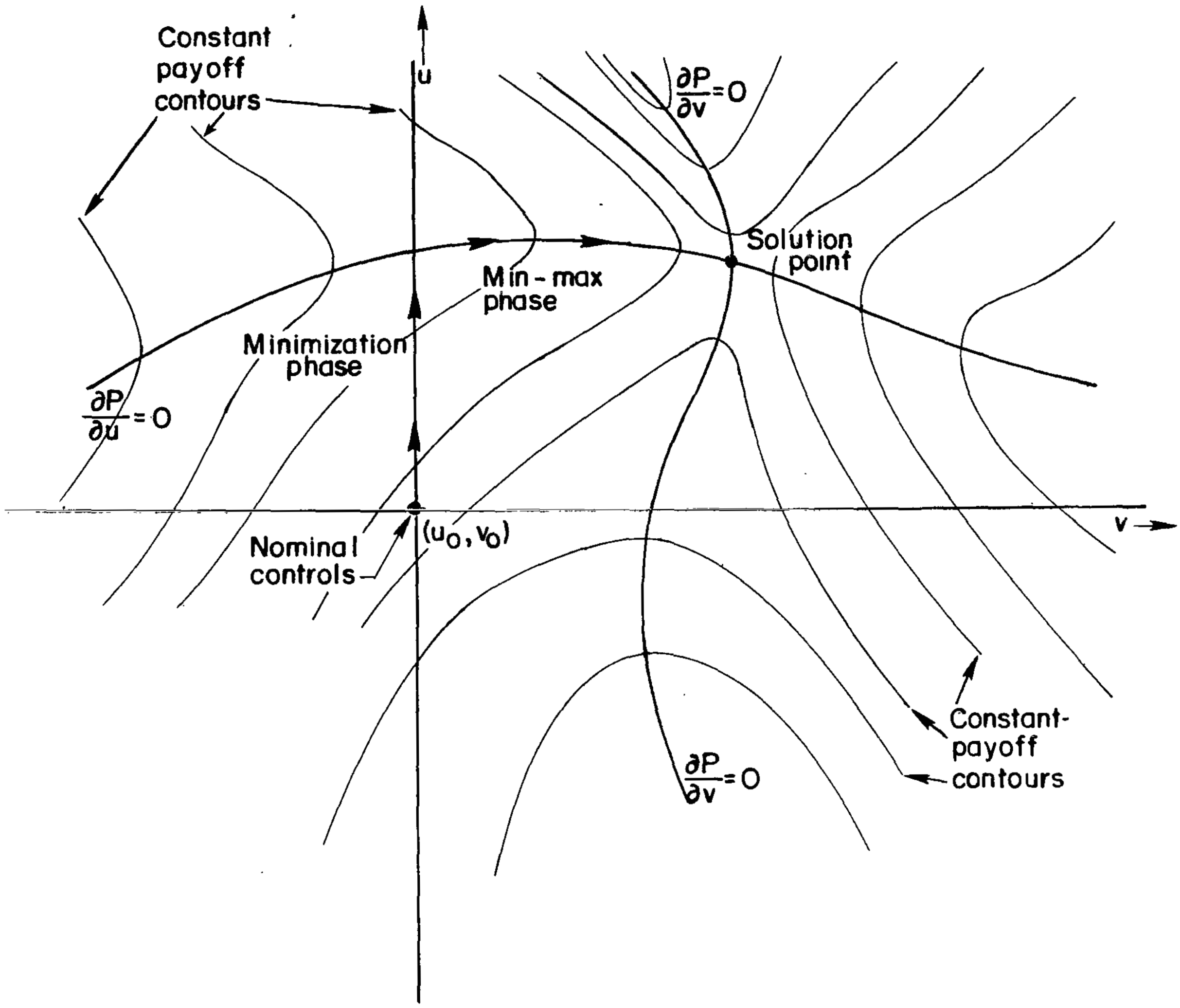


Figure 16.- Convergence procedure for scalar controls.

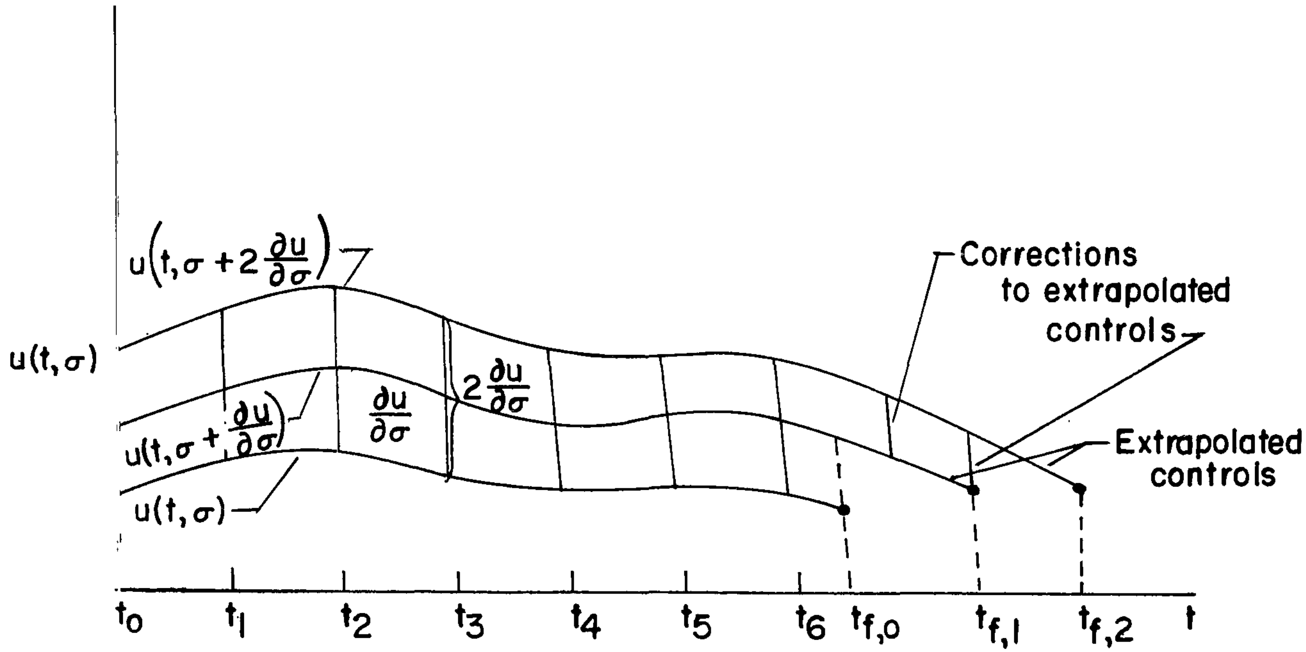


Figure 17.- Extrapolation and correction procedure for variable final time.

NASA-Langley, 1971 — 19 L-7770

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546