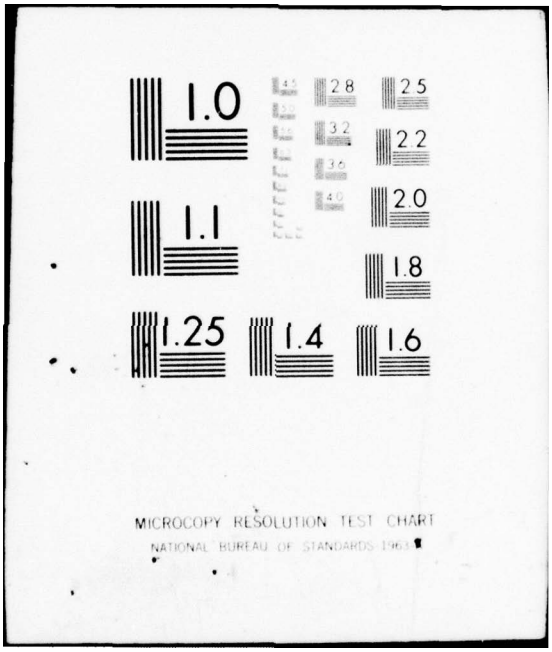
AD-A034 896 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 12/1
APPLICATION OF DIFFERENTIAL DYNAMIC PROGRAMMING TO AN AIR-TO-AI--ETC(U)
DEC 76 A H FERRARIS

UNCLASSIFIED GA/MC/76D-7 NL

1 OF 2
AD A034896


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

GA/MC/76D-7

(6) APPLICATION OF
DIFFERENTIAL DYNAMIC PROGRAMMING
TO AN AIR-TO-AIR
MISSILE GUIDANCE PROBLEM
MODELED AS A DIFFERENTIAL GAME.

(9) Masters Thesis,

(14) GA/MC/76D-7 Albert H. Ferraris
Captain USAF

(10)

(11) Dec 76

(12) 148p.

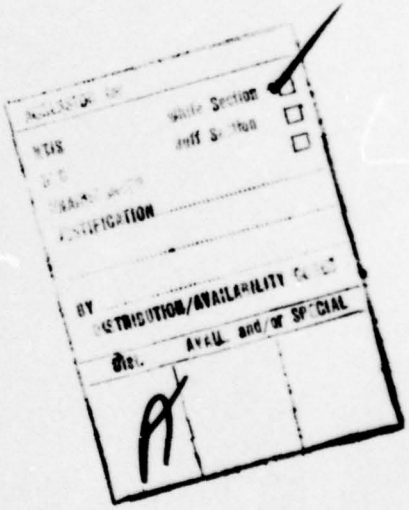
012 225

GA/MC/76D-7

APPLICATION OF DIFFERENTIAL DYNAMIC PROGRAMMING TO AN AIR-TO-AIR MISSILE GUIDANCE PROBLEM MODELED AS A DIFFERENTIAL GAME

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science



by

Albert H. Ferraris, B.S.
Captain USAF
Graduate Astronautical Engineering

December 1976

Preface

Closed-loop guidance laws resulting from differential game models are seldom realized. Several approximations to the closed-loop law, based upon updating a reference open-loop trajectory, have been postulated. This thesis represents the results of my attempt to apply a differential dynamic programming scheme with a new convergence control parameter technique to an air-to-air missile intercept problem using nonlinear dynamics.

I wish to gratefully acknowledge the assistance of Major Gerald M. Anderson for both his classroom presentations on optimal control and his interest in this thesis. In addition, I wish to express a great deal of gratitude to my wife, Terry, who not only put up with me during this task, but also typed the entire thing. Most of all, I want to thank my son Michael, who with one of his smiles was able to lift me from my periods of gloomy frustration.

Albert H. Ferraris

Contents

List of Figures

List of Tables

Abstract

An intercept problem between an air-to-air missile and an aircraft is modeled as a zero sum, free final time differential game which includes nonlinear dynamics and a payoff related to the kill probability. Previous research has shown that the currently used guidance scheme, proportional navigation, is nonoptimal in this type of problem formulation and a higher kill probability is possible with a guidance law based upon a differential game theory.

A differential dynamic programming method is applied to the intercept problem in the search for a real-time feedback solution. A convergence control procedure is introduced in an attempt to enhance the convergence of the typically long-time solution methods. The closed-loop guidance law which results is compared to both proportional navigation and some exact open-loop solutions by means of an off-line simulation on a CDC 6600 computer.

The method does not yield a real-time solution for this problem and does not give improvement over a proportional navigation scheme.

APPLICATION OF DIFFERENTIAL DYNAMIC PROGRAMMING TO AN AIR-TO-AIR MISSILE GUIDANCE PROBLEM MODELED AS A DIFFERENTIAL GAME

I. Introduction

Background

Proportional navigation, whereby a pursuer is guided toward a target at a rate proportional to the measured rate of rotation of the pursuer-target line-of-sight, is the principal guidance law currently in use with most air-to-air missiles (Ref 10). It has been shown that proportional navigation is optimum for problems using linear dynamics and non-maneuvering targets (Ref 5:287-288). Several attempts (Refs 1, 2, 3, 4) have been made to devise closed-loop optimal control laws using nonlinear dynamics, which offer an alternative to proportional navigation if formulated in a closed-loop feedback strategy. One example (Ref 4) requires that the evader's future control strategy be known, but does not allow the evader to take advantage of the pursuer's limitations in predicting the controls.

The theory of differential games (Ref 6) provides a more realistic modeling of the pursuit-evasion problem. The evader's natural desire to escape, and the ability to convert poor pursuer strategy into an advantage for the evader, can be included in the guidance philosophy. Correspondingly, any nonoptimal play by the evader would result in a more favorable condition for the pursuer.

Optimal open-loop controls can be found for the problem through the solution of a two point boundary value problem which arises from the application of optimization conditions (Ref 5: 212-246). Since these controls are open-loop, they do not allow the combatants to capitalize on each other's errors. Near optimal feedback strategies based upon a linearization about the nominal trajectory (resulting from the open-loop controls) which is periodically updated have been proposed (Ref 1, 2, 3). They provide some real-time, near optimal controls; however, the nominal saddle-point solution is required for the linearization and the updating must be accomplished often enough to keep the assumed linearization valid. This represents an enormous investment in computational time and storage space when applied to problems which include nonlinear dynamics and realistic maneuvers.

A comparison between proportional navigation and differential game guidance (Ref 11) where nonlinear dynamics and target maneuverability are allowed, conclusively proves that proportional navigation is not optimal. An off-line computer simulation (Ref 11: 94-100) was used to solve the problem but a real-time application was not realized. The potential gains involved make the search for a real-time implementation worthwhile.

Statement of the Problem

An intercept problem between a heat seeking, air-to-air missile and an aircraft (Ref 11), modeled as a zero-sum,

free final time, differential game between two intelligent combatants, forms the model for this thesis. A differential dynamic programming algorithm (Ref 8) is used to obtain a closed-loop solution for the problem. The aim is to test this algorithm for the possibility of obtaining a real-time guidance law to be used on a short duration (typically less than six seconds) air intercept problem by periodically updating a computed control history. Each combatant can change his control at updating points to capitalize on deficiencies in the adversary's strategy.

A convergence control procedure (Ref 8) was included in an attempt to accommodate convergence problems associated with the inclusion of nonlinear dynamics and to aid in keeping the computational time to a minimum, while not significantly reducing the accuracy of the final solution. The major emphasis of this thesis is to seek a real-time implementation of the differential game feedback guidance law to the nonlinear model.

Overview

Chapter II discusses the mathematical aspects of differential games. The dynamic programming algorithm used in obtaining the closed-loop control strategies is explained in Chapter III, while the game scenario is presented in Chapter IV. The results obtained in the application of this algorithm, and those resulting from an application of proportional navigation to the missile-aircraft intercept problem are compared in Chapter V.

II. Differential Game Theory

Mathematical Formulation

The zero-sum differential game may consist of the state equations, some path or terminal constraints, a terminal (stopping) condition which determines when the game ends, and a payoff or cost function. The state equations which describe the motion of the two players are represented as

$$\dot{x} = f(x, u, v, t) \quad ; \quad x(t_o) = x_o \tag{2-1}$$

where x is an n-dimensional vector which represents the state of each combatant, u is the vector control of the pursuer (minimizer), and v is the vector control of the evader (maximizer). Constraints may be imposed upon the controls of the form

$$\begin{aligned} C(x_p, u) &\leq 0 \\ C(x_e, v) &\leq 0 \end{aligned} \tag{2-2}$$

where x_p and x_e represent the pursuer and evader components of the state vector. In addition, terminal constraints of the form

$$\psi\left[x(t_f), t_f\right] = 0 \tag{2-3}$$

may be included. For situations in which the final time is left free and no terminal constraints are imposed, some stopping condition must be specified, for example $\frac{d}{dt}(J)=0$.

The cost function is expressed in general as

$$J = \varphi(x(t_f), t_f) + \int_{t_o}^{t_f} L(x,u,v,t)\,dt \qquad (2\text{-}4)$$

The cost is a numerical measure for determining the outcome of the game and for evaluating the effectiveness of a particular selected strategy. The game is termed zero-sum because there is a single payoff and one player's gain is the other player's loss. The pursuer's goal is to minimize the cost, J, while the evader strives to maximize it. This forms the basis upon which each player selects his controls. The objective of the game is to determine optimal control strategies, u* and v*, such that

$$J(u^*,v) \leq J(u^*,v^*) \leq J(u,v^*) \qquad (2\text{-}5)$$

If the pair u* and v* can be found, it is termed a saddle point of the game.

Necessary Conditions for a Solution

The problem under consideration in this thesis is a free final time differential game without terminal constraints. A necessary condition for the saddle point solution is that the Hamiltonian, H, defined as

$$H(x,\lambda,u,v,t) = \lambda^T f + L \qquad (2\text{-}6)$$

be maximized for admissible values of v, and minimized for admissible values of u. For games in which the Hamiltonian

is separable in u and v, where

$$H = H_c(x,v) + H_p(x,u) \tag{2-7}$$

the following necessary conditions apply (Ref 5):

$$\begin{aligned} \dot{\lambda}^T &= -\frac{\partial H}{\partial x} \\ \lambda^T(t_f) &= \left.\frac{\partial \varphi}{\partial x}\right|_{t_f} \\ \frac{\partial H}{\partial u} &= 0 \\ \frac{\partial H}{\partial v} &= 0 \\ H(t_f) &= -\left.\frac{\partial \varphi}{\partial t}\right|_{t_f} \end{aligned} \tag{2-8}$$

These conditions hold if there are no control constraints. For the case where control constraints are imposed, the following conditions apply (Ref 5):

$$\begin{aligned} \dot{\lambda}^T &= -\frac{\partial H}{\partial x} - \nu \frac{\partial C}{\partial x} \\ \lambda^T(t_f) &= \left.\frac{\partial \varphi}{\partial x}\right|_{t_f} \\ \frac{\partial H}{\partial u} &= -\nu \frac{\partial C}{\partial u} \\ \frac{\partial H}{\partial v} &= -\nu \frac{\partial C}{\partial v} \\ H(t_f) &= -\left.\frac{\partial \varphi}{\partial t}\right|_{t_f} \end{aligned} \tag{2-9}$$

where λ represents the n-dimensional co-state vector and ν is the Lagrange multiplier vector which obeys the following (Ref 5: 108-109):

$$\begin{aligned} \nu &= 0 \quad \text{FOR} \quad C < 0 \\ \nu &\neq 0 \quad \text{FOR} \quad C = 0 \end{aligned} \tag{2-10}$$

The Two Point Boundary Value Problem

The application of the necessary conditions, Eqs (2-8) or (2-9), result in expressions for the saddle-point controls, u* and v*. These controls, $u^*(x, \lambda, t)$ and $v^*(x, \lambda, t)$, are substituted into the state and co-state equations to form a two point boundary value problem (TPBVP) of the form:

$$\begin{aligned} \dot{x} &= f(x, \lambda, t) \quad ; \quad x(t_o) = x_o \\ \dot{\lambda} &= g(x, \lambda, t) \quad ; \quad \lambda(t_f) = \frac{\partial \varphi}{\partial x}\Big|_{t_f} \\ H(t_f) &= -\frac{\partial \varphi}{\partial t}\Big|_{t_f} \end{aligned} \tag{2-11}$$

The solution to the TPBVP yields open-loop controls of the form

$$\begin{aligned} u(t) &= u(x_o, \lambda_o, t) \\ v(t) &= v(x_o, \lambda_o, t) \end{aligned} \tag{2-12}$$

These controls are termed open-loop because they depend only upon the initial conditions, the time, and the assumption that each player will employ the optimal

strategy. They do not provide a means for either combatant to capitalize upon nonoptimal play by the adversary. One method for determining control strategies which are able to adapt to variations in the opponent's strategy requires that the solution to the TPBVP must somehow be periodically updated based upon more current information. This philosophy is based upon the fact that optimal closed-loop and open-loop controls have the same time history and state trajectories. This idea forms the basis for the determination of closed-loop controls which are able to transfer one player's nonoptimal strategy into an advantage for the other.

III. Closed-Loop Control Strategy

The traditional open-loop solution to the TPBVP, Eqs (2-10), requires that initial values of the co-states, λ (0), be known and utilized in a forward integration. These values are difficult to arrive at, and they must be reasonably close to the optimum co-states to hope for obtaining a solution to the TPBVP (the open-loop controls). Further complications arise because the TPBVP is extremely sensitive to even small variations in the initial co-state values. Large trajectory deviations may result and convergence may be inhibited (even precluded) as indicated in Fig. 1 (Ref 9). Without an accurate TPVBP solution, closed-loop controls based upon the updating of open-loop controls is impossible.

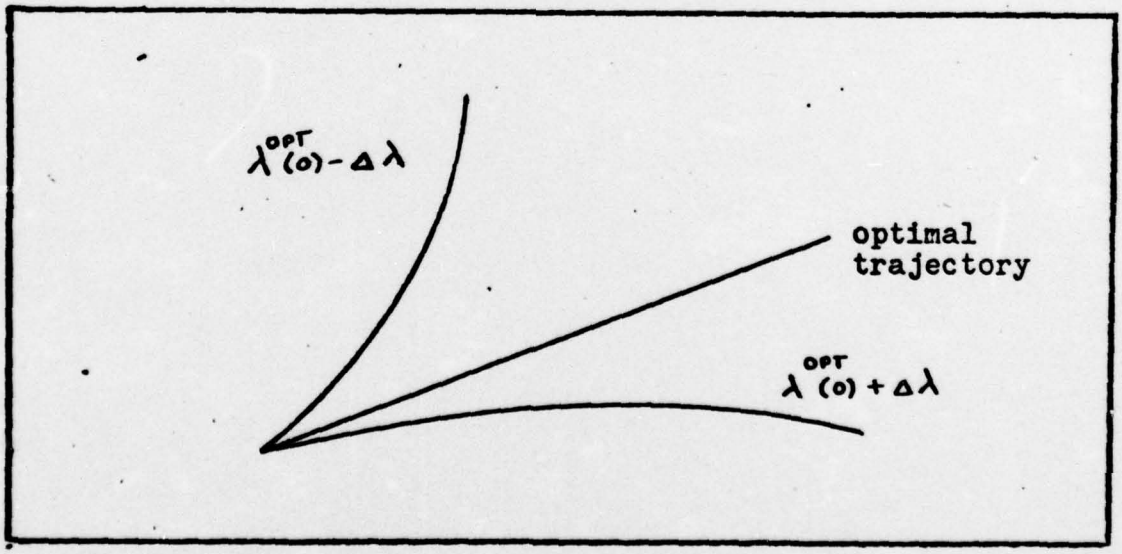


Fig. 1. Co-state Sensitivity

The Differential Dynamic Programming method (DDP) is an alternative method for obtaining closed-loop controls.

The DDP Method

The DDP method (Ref 7, 8) attempts to solve the following problem:

$$\frac{\partial J^{\circ}}{\partial t}(x,t) + H(x, J_x^{\circ}, u^*, v^*, t) = 0$$
$$J^{\circ}(x(t_f), t_f) = \varphi(x(t_f), t_f) \qquad (3\text{-}1)$$

$$H = L(x,u,v,t) + J_x^{\circ T} f(x,u,v,t)$$
$$u^*(x, J_x^{\circ}, t) = \min_u H(x, J_x^{\circ}, u, v^*, t) \qquad (3\text{-}2)$$
$$v^*(x, J_x^{\circ}, t) = \max_v H(x, J_x^{\circ}, u^*, v, t)$$

where the subscript $^{\circ}$ indicates optimality and $J_x = \frac{\partial J}{\partial x}$. The solution to Eq (3-1) yields an optimal cost, $J^{\circ}(x,t)$, which, when substituted into Eqs (3-2), results in optimal closed-loop controls u* and v*. Unfortunately, Eq (3-1) does not, in general, readily lend itself to analytical solutions. The DDP method provides a numerical tool for obtaining the solution in an iterative manner.

If Eq (3-1) is expanded to first order about the optimal trajectory, x*, the following relationships emerge:

$$-\dot{a} = H(\bar{x}, u^*, v^*, J_x, t) - H(\bar{x}, \bar{u}, \bar{v}, J_x, t) \; ; \; a(t_f) = 0$$
$$-\dot{J}_x = \frac{\partial H}{\partial x}(\bar{x}, u^*, v^*, J_x, t) + \left[\nu^T \frac{\partial C}{\partial x}(\bar{x}, u^*, v^*, t)\right]^T \qquad (3\text{-}3)$$
$$J_x(\bar{x}(t_f), t_f) = \frac{\partial \varphi}{\partial x}(\bar{x}(t_f), t_f)$$

where $\bar{u}$ and $\bar{v}$ are nominal controls which result in a

nominal trajectory $\bar{x}$, and a(t) is the difference between the optimal cost, J^o, resulting from the application of the optimal controls u* and v* in the state equations, and the cost, $\bar{J}$, obtained from using nominal controls $\bar{u}$ and $\bar{v}$.

$$a(t) = J^*(u^*, v^*) - \bar{J}(\bar{u}, \bar{v}) \qquad (3\text{-}4)$$

The predicted cost change, $a(t_o)$, can be expressed as

$$a(t_o) = a_p(t_o) + a_e(t_o) \qquad (3\text{-}5)$$

where $a_p(t_o)$ is the predicted cost change due to changes in the pursuer's control and $a_e(t_o)$ is that due to changes in the evader's control. These relationships, Eqs (3-3), are valid if $\Delta x = x^*(t) - \bar{x}(t)$ is not excessively large within the time interval remaining, due to the linearization. The derivation of Eqs (3-3) is presented in Appendix C.

The mechanization of the DDP algorithm for obtaining optimal closed-loop controls is as follows:

(a) Nominal controls, $\bar{u}$ and $\bar{v}$, are used in the state equations, Eq (2-1), which is then integrated forward in time until reaching the stopping criterion, to determine a nominal trajectory $\bar{x}(t)$ and a cost J(t) from Eq (2-3).

(b) Eqs (3-3) are integrated backward in time, using the same nominal controls, $\bar{u}$ and $\bar{v}$, with appropriate boundary conditions. At each step of the backward integration, theconditions of Eqs(3-2) are enforced to obtain new controls u*(t) and v*(t) which are stored in the computer.

(c) The new controls, u* and v*, are applied to

Eqs (2-1) and (2-3) in a forward integration as in step (a). If the change in cost, $\Delta J = J^*(u^*, v^*, t) - J^*(\bar{u}, \bar{v}, t)$, is of the same order as $a(t_o)$, u(t) and v(t) can be replaced by u*(t) and v*(t). Steps (a) through (c) are repeated until $a(t_o)$, $a_e(t_o)$ and $a_p(t_o)$ are small.

(d) When the predicted cost changes have been decreased to a prescribed small value, the computed controls are used as the optimal controls for a specified period of time (a fixed portion of the trajectory).

(e) At the end of the specified time interval, the entire sequence is begun again using the conditions at the end of the interval as the new initial conditions.

If the actual cost change, ΔJ, is not on the order of the predicted cost change, $a(t_o)$, some form of convergence control must be supplied to assure that a solution will be found. The method used in this problem is the Convergence Control Parameter method (CCP) (Ref 8).

The CCP Method

The magnitude of $\Delta x(t)$ can be restricted if the control changes between iterations, $\Delta u(t)$ and $\Delta v(t)$, are not excessively large. The idea behind CCP is to restrict the magnitude of the control changes by means of convergence parameters attached to $\Delta u(t)$ and $\Delta v(t)$.

An augmented Hamiltonian, $\tilde{H}$, is formed as follows:

$$\tilde{H} = \tilde{H}(x, u, \Delta u, v, \Delta v, J_x, P_p, P_e; t)$$

$$\tilde{H} = H(x, u, v, J_x; t) + \tfrac{1}{2} \Delta u^T P_p \Delta u - \tfrac{1}{2} \Delta v^T P_e \Delta v \qquad (3\text{-}6)$$

where P_p and P_e are the convergence parameters. These are diagonal matrices whose elements are positive values. A saddle-point to $\tilde{H}$ is sought through the DDP method.

The predicted cost change, $a(t_o)$, written as

$$a(t_o) = a_p(t_o) + a_e(t_o)$$

allows $a_p(t_o)$ to be dependent upon P_p and $a_e(t_o)$ upon P_e. An analysis of the relative magnitudes of the predicted and actual cost changes allows the penalty terms to be adjusted individually for the best convergence characteristics as explained in Appendix D.

If the predicted cost change, $a(t_o)$, is approximately equal to the actual cost change, ΔJ, the series expansion of Eq (3-1) is satisfied. The relationship between ΔJ and $a(t_o)$ can be plotted and divided into several regions to indicate the effectiveness of the selected penalty values in effecting good convergence of the solution. This is explained in detail in Appendix D.

IV. Intercept Problem

The specific problem under consideration is a pursuit-evader situation between an air-to-air missile and an aircraft. This problem is modeled as a zero-sum differential game with free final time.

Underlying Concepts

The aircraft is modeled with the stall limit, thrust, and drag dependent upon altitude and velocity. The missile is a thrust-coast air-to-air missile utilizing an infra-red seeker. Missile guidance is begun at the termination of the boost phase with drag dependent upon altitude and velocity.

The game "ground rules" are as follows:

(a) Each vehicle is represented as a point mass, maneuverable in three dimensions.

(b) All maneuvers performed are flown in a co-ordinated fashion.

(c) Gravity is represented as a constant in both magnitude and direction.

(d) Both combatants are presumed to have perfect knowledge of the state of the game at all times.

(e) With free final time and no terminal conditions to meet, the determination of when the game ends is the point where $\frac{d}{dt}(J) = 0$.

Vehicle Models

Aircraft. The aircraft model is based upon the F4. The stall limit and thrust variations with velocity and

altitude are represented as polynomials with the maximum throttle setting used throughout. The equations of motion are

$$
\begin{aligned}
\dot{x} &= V\cos\gamma\cos\sigma \\
\dot{y} &= V\cos\gamma\sin\sigma \\
\dot{z} &= V\sin\gamma \\
\dot{V} &= \left[T\cos\alpha - D\right]\frac{1}{m} - g\sin\gamma \\
\dot{\gamma} &= \left[L + T\sin\alpha\right]\frac{\cos\mu}{mV} - \frac{g\cos\gamma}{V} \\
\dot{\sigma} &= \left[L + T\sin\alpha\right]\frac{\sin\mu}{mV\cos\gamma}
\end{aligned}
\qquad (4\text{-}1)
$$

where the variables are defined as follows:

x = distance in the north direction

y = distance in the west direction

z = altitude

V = magnitude of the velocity

γ = angle between the velocity vector and the local horizon

σ = angle between the projection of the velocity vector in the x-y plane and the x-axis

α = angle of attack (defined as the angle between the thrust and velocity vectors)

μ = bank angle

D = force due to drag

L = lift (perpendicular to velocity vector)

g = acceleration due to gravity

T = thrust along the aircraft centerline

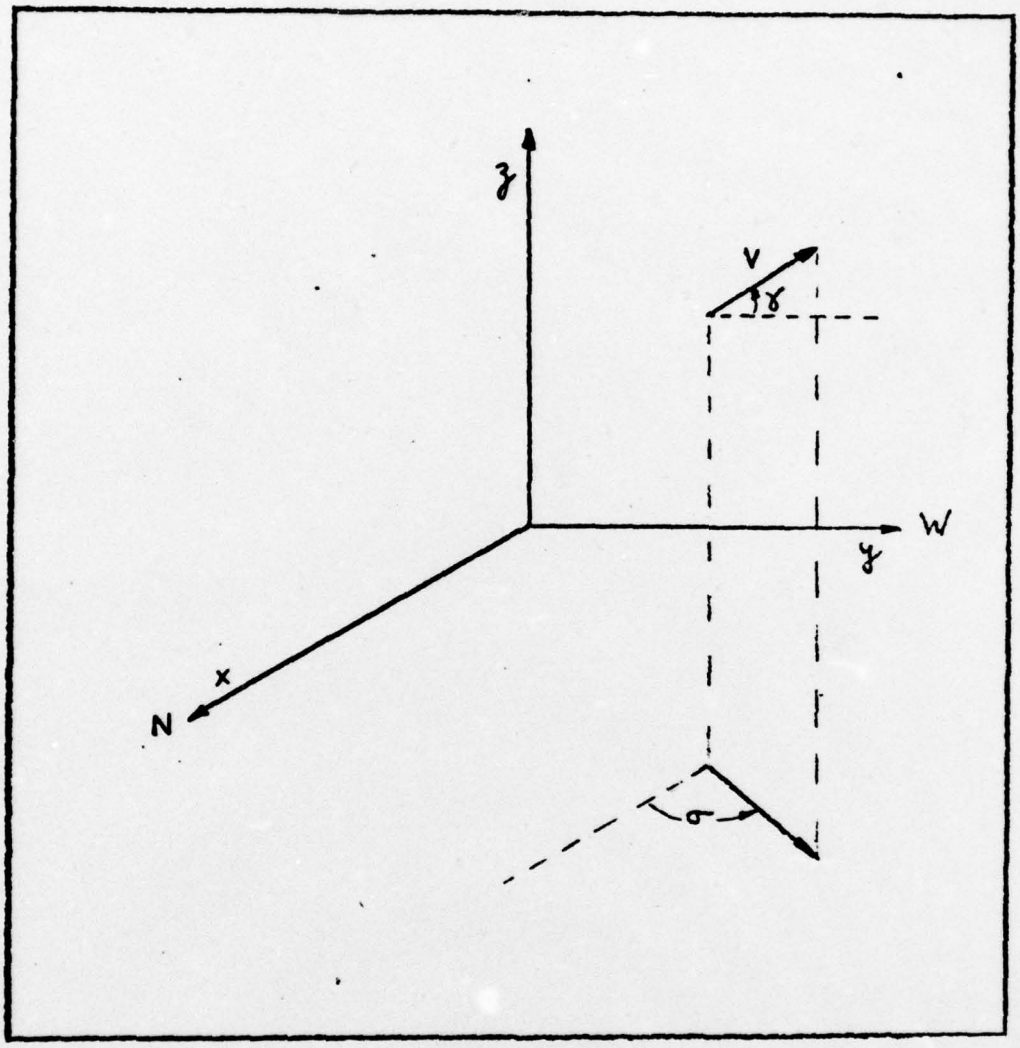


Fig. 2. State Variable Depiction

The controls are the bank angle and the load factor. No constraints are imposed upon the bank angle; however, the load factor is constrained aerodynamically as a function of altitude and airspeed, and limited structurally to six g's. Expressed mathematically, the constraints are

written as follows:

$$C_1(n, \gamma, V) \geq 0$$
$$C_2(n) \geq 0 \qquad (4\text{-}2)$$

where C_1 is the aerodynamic constraint, C_2 is the structural limit, and n is the load factor.

The force due to drag is

$$D = C_D Q S \qquad (4\text{-}3)$$

where the following relationships define the variables:

$$C_D = C_{D_o} + k_1 C_L^2$$
$$Q = \tfrac{1}{2} \rho_o e^{-\beta \gamma} V^2 \qquad (4\text{-}4)$$

S = reference area

Numerical values for the models are listed in Appendix E.

Missile. The same equations of motion, Eqs (4-1), are used in the missile model; however, the missile is considered in the coast phase only so that the thrust, T, is zero. No aerodynamic constraint is imposed upon the missile load factor. The structural limit is set at fifteen g's with the corresponding constraint relation as follows:

$$C_3(n) \geq 0 \qquad (4\text{-}5)$$

Selection of the Cost Function

The selection of a suitable cost function is probably the most subjective portion of the differential game problem formulation. The cost function used in this problem is

$$J = AR^2 - B\cos(\psi_f) + Ct_f \qquad (4\text{-}6)$$

where R represents the range between the aircraft and the missile, ψ_f represents the angle between the velocity vectors of the two players (track crossing angle), t_f is the final time (time at which $\frac{d}{dt}(J) = 0$), and A, B, and C are suitably selected weighting factors.

The selection of the "R^2" term is based upon the fact that the probability of kill, P_k, is primarily determined by the miss distance in an inversely proportional manner (Ref 9). The "R^2" term heavily penalizes the pursuer for failing to close to within a small final range, resulting in the miss distance being the predominant measure of warhead effectiveness.

Some consideration must be given to the fuzing and explosive pattern of the warhead. The "destructive fragments" radiate outward from the explosion. Clearly, a proximity fuze will be most apt to detonate the warhead within the lethal range if the flight paths of the missile and aircraft are closely aligned. The penalty associated with the track crossing angle ($-B\cos\psi_f$) reflects this consideration. A small value of ψ_f results in a small penalty for the

pursuer while a head-on attack results in the maximum penalty, reflecting the difficulties in fuzing the warhead for very large closure rates.

The penalty attached to the final time is added to preclude non-unique solutions when the missile is able to accomplish the intercept. This term is not significant unless the missile is able to reduce the final cost to a very small number.

The weighting factors are picked so that the range predominates in the cost function until a distance of ten feet is reached. The track crossing angle is significant within the ten foot range while the "t_f" term becomes significant as $\dot{r}$ and R approach zero. The following weighting values are selected to reflect this:

$$A = 1$$
$$B = 10$$
$$C = 1/6$$

Application of Differential Game Theory

The DDP Equations, Eqs (3-3), are applied to the problem which allow the adjoint equations to be found as partial derivatives of the Hamiltonain. For the problem under consideration, the Hamiltonian is

$$\tilde{H} = J_x^T f - \frac{1}{2} \Delta v^T P_e \Delta v + \frac{1}{2} \Delta u^T P_p \Delta u \qquad (4\text{-}7)$$

The particular formulation of this problem allows the Equation for $\dot{\lambda}^T$ from Chapter II and the $\dot{J}_x$ equation from Chapter III to be related as

$$\dot{J}_x^T = \dot{\lambda}^T = -\frac{\partial \tilde{H}}{\partial x} \tag{4-8}$$

This allows Eq (4-7) to be written as

$$\begin{aligned}\tilde{H} = {} & \lambda_x \dot{x} + \lambda_y \dot{y} + \lambda_z \dot{z} + \lambda_v \dot{v} \\ & + \lambda_\gamma \dot{\gamma} + \lambda_\sigma \dot{\sigma} - \frac{1}{2} \Delta v^T P_c \Delta v \\ & + \frac{1}{2} \Delta u^T P_p \Delta u\end{aligned} \tag{4-9}$$

Employing Eqs (4-1) results in the following:

$$\dot{\lambda}_x = 0$$

$$\dot{\lambda}_y = 0$$

$$\dot{\lambda}_z = -\frac{\lambda_v}{m}\left[\frac{\partial T}{\partial z}\cos\alpha - T\sin\alpha\frac{\partial \alpha}{\partial z} - \frac{\partial D}{\partial z}\right]$$

$$-\left[\frac{\lambda_\gamma \cos\mu}{mv} + \frac{\lambda_\sigma \sin\mu}{mv\cos\gamma}\right] \times$$

$$\left[\frac{\partial L}{\partial z} + \frac{\partial T}{\partial z}\sin\alpha + T\cos\alpha\frac{\partial \alpha}{\partial z}\right]$$

$$\dot{\lambda}_v = -\lambda_x \cos\gamma\cos\sigma - \lambda_y \cos\gamma \sin\sigma - \lambda_z \sin\gamma$$

$$-\frac{\lambda_v}{m}\left[\frac{\partial T}{\partial v}\cos\alpha - T\sin\alpha\frac{\partial \alpha}{\partial v} - \frac{\partial D}{\partial v}\right] + \frac{\lambda_\gamma \dot{\gamma}}{v} + \frac{\lambda_\sigma \dot{\sigma}}{v}$$

$$-\left[\frac{\lambda_\gamma \cos\mu}{mv} + \frac{\lambda_\sigma \sin\mu}{mv\cos\gamma}\right] X$$

$$\left[\frac{\partial L}{\partial v} + \frac{\partial T}{\partial v}\sin\alpha + T\cos\alpha\frac{\partial \alpha}{\partial v}\right]$$

(4-10)

$$\dot{\lambda}_\gamma = \lambda_x v\cos\sigma\sin\gamma + \lambda_y v\sin\gamma\sin\sigma$$

$$-\lambda_z v\cos\gamma + \lambda_v g\cos\gamma - \frac{\lambda_\gamma g\sin\gamma}{v}$$

$$-\lambda_\gamma \dot{\sigma}\tan\gamma$$

$$\dot{\lambda}_\sigma = \lambda_x v\cos\gamma\sin\sigma - \lambda_y v\cos\gamma\cos\sigma$$

where

$$L = nw$$

$$T_{max} = F_1 + F_2 g + F_3 v$$

$$D = \frac{1}{2}\rho_0 e^{-\beta\gamma} v^2 s\left(C_{D_0} - k_1 C_L^2\right) \qquad (4\text{-}11)$$

$$\alpha = k_2 C_L$$

$$C_L = \frac{n W}{\frac{1}{2} \rho_0 e^{-\beta \gamma} V^2 S}$$

and the partial derivatives are

$$\frac{\partial L}{\partial \gamma} = 0 \quad ; \quad \frac{\partial L}{\partial V} = 0$$

$$\frac{\partial T}{\partial \gamma} = F_2 \quad ; \quad \frac{\partial T}{\partial V} = F_3$$

$$\frac{\partial \alpha}{\partial \gamma} = k_2 \frac{\partial C_L}{\partial \gamma} \quad ; \quad \frac{\partial \alpha}{\partial V} = k_2 \frac{\partial C_L}{\partial V} \qquad (4\text{-}12)$$

$$\frac{\partial C_L}{\partial \gamma} = \beta C_L \quad ; \quad \frac{\partial C_L}{\partial V} = -\frac{2 C_L}{V}$$

$$\frac{\partial D}{\partial \gamma} = -\beta \left[\frac{1}{2} \rho_0 e^{-\beta \gamma} V^2 S \left(C_{D_0} - k_1 C_L^2 \right) \right]$$

$$\frac{\partial D}{\partial V} = \rho_0 e^{-\beta \gamma} V S \left(C_{D_0} - k_1 C_L^2 \right)$$

Adjustments to Eqs (4-9) must be made to compensate for the load factor constraints, Eqs (4-2) and Eqs (4-5). These equations may be written as

$$C_1(n, \gamma, V) = n'(\gamma, V) - n$$

$$C_2(n) = 6 - n \qquad (4\text{-}13)$$

$$C_3(n) = 15 - n$$

where n' is the aerodynamic (or lift coefficient) limit upon the load factor of the aircraft

$$n'(\gamma, V) = a_1 + (a_2 - a_3 \gamma)(V - a_4) \tag{4-14}$$

For the occasions when the load factor is on the constraint boundary, the changes to Eqs (4-10) are

$$\begin{aligned} \dot{\lambda}_\gamma &= -\frac{\partial H}{\partial \gamma} - \nu \frac{\partial C_1}{\partial \gamma} \\ \dot{\lambda}_V &= -\frac{\partial H}{\partial V} - \nu \frac{\partial C_1}{\partial V} \\ \nu &= \frac{-\partial H/\partial n}{\partial C_1/\partial n} = \frac{\partial H}{\partial n} \end{aligned} \tag{4-15}$$

This gives the following relationships for constrained load factors

$$\begin{aligned} \dot{\lambda}_\gamma &= -\frac{\partial H}{\partial \gamma} - \frac{\partial H}{\partial n}\frac{\partial C_1}{\partial \gamma} \\ \dot{\lambda}_V &= -\frac{\partial H}{\partial V} - \frac{\partial H}{\partial n}\frac{\partial C_1}{\partial V} \end{aligned} \tag{4-16}$$

and for unconstrained situations (where $\nu = 0$) the stated relations hold:

$$\begin{aligned} \dot{\lambda}_\gamma &= -\frac{\partial H}{\partial \gamma} \\ \dot{\lambda}_V &= -\frac{\partial H}{\partial V} \end{aligned}$$

Optimal Control Solution

A requirement for an optimal control is that the augmented Hamiltonian, $\tilde{H}$, be minimized for the pursuer and maximized for the evader. The augmented Hamiltonian can be written in separate form as

$$\tilde{H} = \tilde{H}_e + \tilde{H}_p$$

$$\tilde{H} = H_e - P_{e_1} \frac{\Delta\mu_e^2}{2} - P_{e_2} \frac{\Delta n_e^2}{2} + H_p + P_{p_1} \frac{\Delta\mu_p^2}{2} + P_{p_2} \frac{\Delta n_p^2}{2} \qquad (4\text{-}17)$$

The first order condition for the evader's unconstrained optimal bank angle is

$$\frac{\partial \tilde{H}_e}{\partial(\Delta\mu_e)} = 0 \qquad (4\text{-}18)$$

where

$$\tilde{H}_e = \lambda_\gamma \left[\frac{L + T\sin\alpha}{mV} \cos(\mu_e + \Delta\mu_e)\right] + \lambda_\sigma \left[\frac{L + T\sin\alpha}{mV\cos\gamma} \sin(\mu_e + \Delta\mu_e)\right] - P_{e_1} \frac{\Delta\mu_e^2}{2} \qquad (4\text{-}19)$$

The application of Eq (4-18) with appropriate small angle approximations results in

$$\Delta\mu_e = \frac{\frac{L+T\sin\alpha}{mV}\left[-\lambda_\gamma \sin\mu_e + \frac{\lambda_\sigma \cos\mu_e}{\cos\gamma}\right]}{\left[\frac{L+T\sin\alpha}{mV}\left\{\lambda_\gamma \cos\mu_e + \frac{\lambda_\sigma \sin\mu_e}{\cos\gamma}\right\} + P_{e_1}\right]} \qquad (4\text{-}20)$$

The second order condition

$$\frac{\partial^2 \tilde{H}_e}{\partial(\Delta\mu_e)^2} \leq 0 \qquad (4\text{-}21)$$

gives

$$-\frac{L+T\sin\alpha}{mV}\left[\lambda_\gamma \cos\mu_e + \frac{\lambda_\sigma \sin\mu_e}{\cos\gamma}\right] \leq P_{e_1} \qquad (4\text{-}22)$$

The selection of P_e must ensure that Eq (4-22) is satisfied.

A similar approach results in the following relationships for computing $\Delta\mu_p$:

$$\Delta\mu_p = \frac{\frac{L}{mV}\left[-\lambda_\gamma \sin\mu_p + \lambda_\sigma \frac{\cos\mu_p}{\cos\gamma}\right]}{\frac{L}{mV}\left[\lambda_\gamma \cos\mu_p + \lambda_\sigma \frac{\sin\mu_p}{\cos\gamma}\right] - P_{p_1}} \qquad (4\text{-}23)$$

$$\frac{L}{mv}\left[\lambda_\gamma \cos\mu_p + \frac{\lambda_\sigma \sin\mu_p}{\cos\gamma}\right] \leq P_{p_1} \qquad (4\text{-}24)$$

The constrained load factor is found with an analogous approach. For the evader, the separated terms of the augmented Hamiltonian are

$$\tilde{H}_e = \frac{\lambda_v}{m}\left[T\cos\alpha - D\right] + \frac{\lambda_\gamma}{mv}\left[L + T\sin\alpha\right]\cos\mu_e + \frac{\lambda_\sigma}{mv\cos\gamma}\left[L+T\sin\alpha\right]\sin\mu_e - P_{e_2}\frac{\Delta n_e^2}{2} \qquad (4\text{-}25)$$

The application of the first order necessary conditions gives

$$\frac{\partial \tilde{H}_e}{\partial(\Delta n)} = \frac{\lambda_v}{m}\left[-T\sin\alpha\frac{\partial\alpha}{\partial(\Delta n)} - \frac{\partial D}{\partial(\Delta n)}\right] + \frac{\lambda_\gamma\cos\mu_e}{mv}\left[\frac{\partial L}{\partial(\Delta n)} + T\cos\alpha\frac{\partial\alpha}{\partial(\Delta n)}\right] + \frac{\lambda_\sigma\sin\mu_e}{mv\cos\gamma}\left[\frac{\partial L}{\partial(\Delta n)} + T\cos\alpha\frac{\partial\alpha}{\partial(\Delta n)}\right] - P_{e_2}\Delta n = 0 \qquad (4\text{-}26)$$

where the partial derivatives are evaluated from Eqs (4-11).

Solving for Δn_e and assuming small angles, α, results in

$$\Delta n_e = \left\{ \frac{-\lambda_{v_e}}{m} \left[\frac{T k_2^2 \omega^2}{Q^2 s^2} + \frac{2 k_1 \omega}{Q s} \right] n_e \right.$$

$$\left. + \frac{\omega}{mv} \left[\lambda_{\gamma_e} \cos\mu_e + \frac{\lambda_{\sigma_e} \sin\mu_e}{\cos\gamma_e} \right] \left[1 + \frac{T k_2}{Q s} \right] \right\} X \qquad (4\text{-}27)$$

$$\left[\frac{\lambda_{v_e}}{m} \frac{T k_2^2 \omega^2 n_e}{Q^2 s^2} + P_{e_2} \right]^{-1}$$

The second order conditions, Eq (4-21), are applied to give

$$-\frac{\lambda_{v_e}}{m} \left[\frac{T k_2^2 \omega^2}{Q^2 s^2} + \frac{2 k_1 \omega^2}{Q s} \right] \leq P_{e_2} \qquad (4\text{-}28)$$

The same approach is used to find Δn_p with the following results:

$$\Delta n_p = \frac{\frac{-\lambda_{v_p}}{m} \frac{2 k_1 \omega^2 n_p}{Q s} + \frac{\omega}{mv} \left[\lambda_{\gamma_p} \cos\mu_p + \frac{\lambda_{\sigma_p} \sin\mu_p}{\cos\gamma_p} \right]}{\left[\frac{\lambda_{v_p}}{m} \frac{2 k_1 \omega^2}{Q s} - P_{p_2} \right]} \qquad (4\text{-}29)$$

$$\frac{\lambda_{v_p}}{m} \frac{2 k_1 \omega^2}{Q s} \leq P_{p_2} \qquad (4\text{-}30)$$

Equation (4-28) is automatically satisfied if $\lambda_{v_e} > 0$ and an interior load factor is possible in these circumstances.

Should the computed load factor, $n_e + \Delta n_e$, exceed the constraints, Eqs (4-13), the load factor is set at maximum. An interior control for the pursuer's load factor is possible if $\lambda_{v_p} < 0$, from Eq (4-30), and the load factor, $n_p + \Delta n_p$, is handled in exactly the same manner as for the evader in consideration of the pursuer's constraint. A selection of zero values for the penalty functions reduces the problem to the general differential game situation (Ref 11: 19-24).

Proportional Navigation Guidance

A proportional navigation scheme was employed against the DDP "guided" evader to rate the performance of differential game guidance. The proportional navigation pursuer was also allowed to use perfect information to determine the rate of change of the line-of-sight between the two vehicles. Two angles were used to determine the line-of-sight (LOS) as indicated in Fig 3 (Ref 11:25). Relationships for the angles are

$$\tan \theta = \frac{y}{x} \tag{4-31}$$

$$\tan \psi = \frac{z^2}{\sqrt{x^2 + y^2}} \tag{4-32}$$

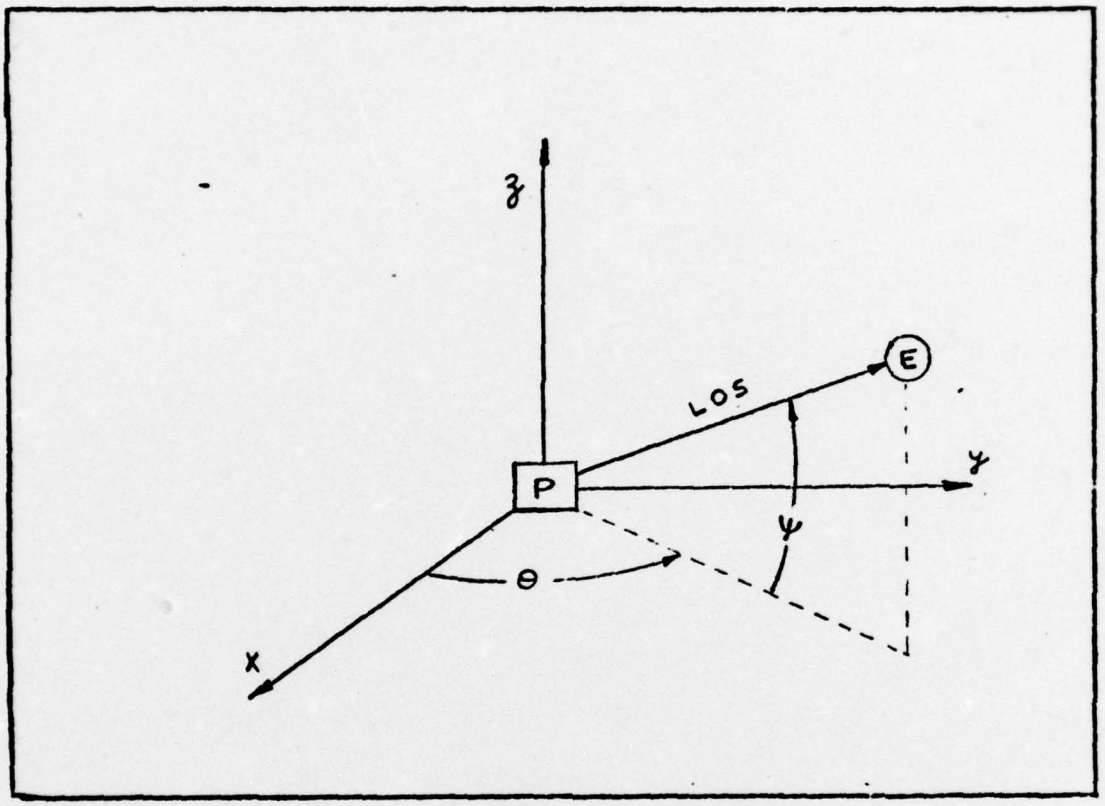


Fig. 3. Line-of-Sight Angle Determination

The time derivatives are

$$\dot{\psi} = \frac{1}{1+\frac{z^2}{(x^2+y^2)}}\left[\frac{\dot{z}}{\sqrt{x^2+y^2}} - \frac{z(x\dot{x}-y\dot{y})}{\sqrt{(x^2+y^2)^3}}\right] \tag{4-33}$$

$$\dot{\theta} = \frac{1}{1+\frac{y^2}{x^2}}\left[\frac{\dot{y}}{x} - \frac{y\dot{x}}{x^2}\right]$$

The idea is to have $\dot{\gamma}_p$ and $\dot{\sigma}_p$ be some multiple of $\dot{\psi}$ and $\dot{\theta}$:

$$\dot{\gamma}_p = RP_1 \dot{\psi} = \frac{L}{mV} \cos\mu_p - \frac{g \cos\gamma_p}{V}$$

$$\dot{\sigma}_p = RP_2 \dot{\theta} = \frac{L \sin\mu_p}{mV \cos\gamma_p} \qquad (4\text{-}34)$$

Solving yields

$$\tan\mu_p = \frac{RP_2 \dot{\theta}}{RP_1 \dot{\psi} + \frac{g \cos\gamma_p}{V}} \qquad (4\text{-}35)$$

$$L = nW = \frac{RP_2 \dot{\theta} mV \cos\gamma_p}{\sin\mu_p}$$

If the calculated value of n exceeded the structural limit, n was set to fifteen.

To contend with situations where $\dot{\theta}$ is small, Eq (4-34) gives

$$n = \frac{V}{g \cos\mu_p} \left[RP_1 \dot{\psi} + \frac{g \cos\gamma_p}{V} \right] \qquad (4\text{-}36)$$

V. Results

The results obtained are derived from the closed-loop application of the DDP algorithm as explained in Appendix B and Chapter III, and from the proportional navigation scheme of Chapter IV against an evader employing the DDP algorithm in a closed-loop fashion. Several TPBVP solutions were available (Ref 11) and these allowed the algorithm to be compared to a known reference as a check of accuracy.

DDP Closed-Loop Application

The intercept problem was attempted first without convergence control. It was found that large control changes, $\Delta u(t)$ and $\Delta v(t)$, resulted causing trajectory changes, $\Delta x(t)$, to be too large, thus violating the linearization of Eq (3-1) and not allowing a solution to be found. Reducing the problem to extremely favorable initial pursuer positions did not help the situation, and the CCP technique was included in the algorithm.

The closed-loop guidance scheme used in this problem solution utilizes five iterations to determine the control strategies and the combatants' applied controls are updated at .5 second intervals. The integration routine uses a dt of .02 seconds. This formulation does not result in a real-time guidance philosophy with the use of current computer technology and is certainly not applicable to a small, air-to-air missile; however, it does obviate the traditional, long execution time, non-linear TPBVP

solution methods (Ref 11). The resulting trajectories appear to follow reasonable logic. The aircraft tries to cross the path of the missile and normally descends to gain a better turning rate and take advantage of the higher thrust available. This result is consistent with normal air-to-air evasion tactics and parallels the findings of reference 11. For purposes of this study, a kill is defined as a pass within ten feet of the aircraft at any value of track crossing angle, Ψ , or a pass within fifteen feet at values below 45°.

Case 1. The first situation considers an attack from directly behind the aircraft (six o'clock position). The altitude for both combatants is 33,000 feet and the initial controls for both are zero bank angle and one "g" (straight and level). The DDP algorithm is applied in a closed-loop manner as indicated (five iterations between control updates with .5 second updating intervals). The results are that the aircraft attempts a straight-ahead (no bank) maximum "g" climb. This makes sense if consideration is given to the speed and load factor advantage of the missile. By attempting a straight-ahead climb, the aircraft tries to capitalize on the zero-thrust situation of the missile; however, a kill is scored in all situations until the missile is initially positioned far enough behind the aircraft so that its airspeed is depleted before the intercept can be completed.

This test case confirms what experience has proven

in air-to-air combat; if an evader is unaware of an attack upon him and is not maneuvering at the time missile guidance begins, he will most likely be destroyed.

Case 2. This situation is initially the same relative position as in Case 1; however, the initial controls are changed. The pursuer again employs zero bank and one "g" but the evader flies a fixed 90° bank and three "g's" throughout. The DDP algorithm is used as the pursuer's guidance law and the evader flies the initial controls with no updating.

The first application used an integration step size of .02 seconds, an updating interval of 1.5 seconds, and 15 iterations between updates. The result is a 237 foot miss at the end of the game.

A second attempt with DDP guidance with a one second updating interval and ten iterations results in a 13 foot miss at $\phi = 11.5^{\circ}$, a kill.

Finally, a third application using five iterations and a .5 second updating interval gives a four foot miss at $\phi = 11.5^{\circ}$, within the lethal envelope.

This analysis shows that an improvement is realized by taking a smaller updating interval and retaining enough iterations (in this case five) to derive "near optimal" controls without undue computational time. In the large updating interval case, the "near-optimal" controls derived are not close enough to the optimal and, when applied for long periods (one second and more), cause poor results.

In the explanation of the DDP method (Chapter III) it was stated that the predicted cost changes, $a_e(t_o)$ and $a_p(t_o)$, should be small. In the 1.5 second updating application, the magnitudes of a_e and a_p after five iterations are on the order of 50,000 and decrease to 5000 by the end of the game. For the one second updating, they are 40,000 after five iterations and decrease to 200. While with a .5 second interval, the predicted changes after the iterations are 32,000 but decrease to 10^{-3}.

It can be concluded that forcing the evader to use a particular control and allowing the pursuer to update with the DDP algorithm on short time intervals will result in a successful intercept; however, this is not achieved in real-time, even with the CDC 6600 computer.

Case 3. In this simulation, both pursuer and evader use the DDP algorithm to update and compute control histories. Although a real-time solution is not achieved, the computations are based upon an integration step size of .02 seconds, a .5 second updating period, and 5 iterations between updates. Three particular situations will be presented in detail as representative.

The first situation is that of Table XI of Appendix A. The selected initial controls place the aircraft in a 90° bank, two "g" turn into the missile. The pursuer's initial control is a zero bank, one "g" path aimed ahead of the aircraft. This control was selected to reflect the fact that the pursuer does not have any idea what the aircraft

will do. It is very similar to the way a hunter would aim at a duck, some slight lead based on the present control of the evader, and does not give any inherent advantage to the missile.

The missile closes to within 15 feet at the intercept point with $\xi = 40^{\circ}$ and the resulting flight path is depicted in Figure 4. Table I lists the relative values of actual versus predicted cost changes, $\Delta J/a(t_o)$. The individual predicted cost changes start at magnitudes on the order of 100,000 and slowly decrease with successive iterations until they reach .5 at the end of the game.

Table I

Convergence Characteristics

Time (Secs)	$\Delta J/a(t_o)$	Predicted Final Miss Distance (ft)
.5	.13	646
1.0	.85	492
1.5	.91	288
2.0	.85	129
2.5	.77	50
3.0	.67	19
3.5	.71	15

A second data set, Table XII, is run in the same manner but the initial evader control is now 90^{o} bank and two "g's" away from the pursuer. This problem terminates in an 11 foot miss at $\psi = 40^{o}$. Table II depicts the convergence characteristics of this problem. The values of $a_e(t_o)$ and $a_p(t_o)$ start at 42,000 and decrease to 10^{-2}. This trajectory is shown in Figure 6.

Table II

Convergence Characteristics

Time (Secs)	$\Delta J/a(t_o)$	Predicted Final Miss Distances (ft)
.5	.42	663
1.0	.91	544
1.5	.92	305
2.0	.86	142
2.5	.79	61
3.0	.68	29
3.5	.96	16
4.0	.67	12

Finally, Table XIII is used to begin an intercept problem. In this case, the aircraft is in a 30^{o} dive using 90^{o} bank and two "g's" into the pursuer. This helps the aircraft since it is gaining both airspeed and control authority, and the miss distance is 578 feet with the convergence characteristics presented in Table III. The individual cost change values begin at 350,000 and decrease to 10^{-1}. The trajectory is depicted in Figure 8.

Table III

Convergence Characteristics

Time (Secs)	$\Delta J/a(t_o)$	Predicted Final Miss Distances (ft)
.5	1.54	1227
1.0	1.00	1467
1.5	1.11	1483
2.0	1.25	1489
2.5	.95	1487
3.0	.99	1468
3.5	.99	1413
4.0	.99	1295
4.5	.99	1077
5.0	.90	760
5.5	.87	613
6.0	.70	581
6.5	.54	579

The bank angle histories for the three cases are in Tables IV, V, and VI.

Table IV

Bank Angle History

Time (Secs)	Pursuer Bank (Deg)	Evader Bank (Deg)
.5	42	32
1.0	81	82
1.5	74	64
2.0	82	75
2.5	86	78
3.0	81	84
3.5	84	86

Table V

Bank Angle History

Time (Secs)	Pursuer Bank (Deg)	Evader Band (Deg)
.5	41	50
1.0	86	81
1.5	78	70
2.0	81	74
2.5	83	75
3.0	85	77
3.5	86	76
4.0	87	77

Table VI

Bank Angle History

Time (Secs)	Pursuer Bank (Deg)	Evader Bank (Deg)
.5	76	82
1.0	55	30
1.5	52	25
2.0	51	24
2.5	50	22
3.0	50	22
3.5	51	22
4.0	51	25
4.5	54	32
5.0	58	44
5.5	62	55
6.0	65	62
6.5	66	63

It can be seen that there are no significant discontinuities in the updating of the controls; however, it is also apparent that the controls remain nonoptimal for a good portion of the intercept. If more iterations are used between control updates, better controls would be derived but at the expense of a real-time implementation. In general, five iterations is not sufficient for determining near-optimal controls when initial control "guesses " are not close to the optimal.

The final stage of the intercept problem is the most sensitive and rapid, large control changes are demanded of

the missile reflecting the missile's desire to line up the flight paths (null $\dot{r}$) and eliminate the terminal miss distance while state values are rapidly changing. To overcome this situation, anytime the time remaining is less than one full updating period, the missile uses the last control selected.

The load factor selected by the guidance scheme was largely determined by the penalty values, P_{e_2} and P_{p_2}. Interior controls result for the missile for a large portion of the flight while the aircraft normally reaches its maximum load factor first. The values of P_e and P_p must be picked large enough to ensure convergence; however, large values also cause slow convergence as reflected in the values of $a_e(t_o)$ and $a_p(t_o)$ previously discussed.

Proportional Navigation Guided Pursuer

The pursuit-evasion problem was re-solved in a closed-loop fashion with the evader employing DDP derived controls and the pursuer relying upon proportional navigation to compare the DDP performance. The same initial states and controls were used (Tables XI - XXII). The proportionality constants, RP_1 and RP_2, are set at ten. This represents a performance level which exceeds that obtainable with current technology; however, the selection of ten is used to reflect future capabilities of improved proportional navigation methods.

Each proportional navigation trajectory follows the DDP trajectory in Appendix A. In each situation, the

missile is able to close to lethal range. The aircraft once again attempts to cross the pursuer's path, but the missile is able to keep inside of the turn. This is due to the selection of ten as a proportionality constant and the fact that the evader is using the DDP guidance law which is not optimal.

Both combatants use the maximum load factor throughout which allows the missile more control authority, but at a great expense in drag. The more unfavorable the initial position, the larger the final range. This situation is expected and reflects missile launches made from outside of the "firing envelope" where the drag penalty defeats the missile.

Tables VII, VIII, and IX present the bank angles resulting from the conditions of Tables XI, XII, and XIII with the trajectories depicted in Figures 5, 7, and 9.

Table VII

Bank Angle History

Time (Secs)	Pursuer Bank (Deg)	Evader Bank (Deg)
.5	182	87
1.0	72	87
1.5	184	89
2.0	7	89
2.5	180	88
3.0	178	89
3.5	13	90
4.0	154	90

Table VIII

Bank Angle History

Time (Secs)	Pursuer Bank (Deg)	Evader Bank (Deg)
.5	-61	-59
1.0	-21	-59
1.5	136	-59
2.0	-40	-59
2.5	-44	-59
3.0	143	-60
3.5	130	-60
4.0	-71	-60
4.5	-63	-59

Table IX

Bank Angle History

Time (Secs)	Pursuer Bank (Deg)	Evader Bank (Deg)
.5	-51	20
1.0	88	14
1.5	26	18
2.0	19	16
2.5	52	19
3.0	61	22
3.5	96	29
4.0	104	40
4.5	100	37
5.0	110	39
5.5	112	46
6.0	53	28

Open-Loop Comparison

The exact TPBVP solution for three intercept problems, yielding open-loop controls, were obtained (Ref 11). These situations are listed in Tables XX - XXII. The two guidance schemes previously discussed (DDP and proportional navigation) were applied to the problems; however, the initial controls for both players are the saddle-point controls from the TPBVP solution. Again, five iterations between .5 second updates are used with the resulting trajectories displayed in Figures 22 - 30.

For the situation of Table XX, the open-loop controls give a terminal miss of 20 feet while the DDP method results in a 177 foot miss and the proportional navigation scheme results in a 126 foot miss. The control comparison is listed in Table X. The predicted cost changes, $a_e(t_o)$ and $a_p(t_o)$, start at 240,000 and decrease to 60.

From this analysis, it is determined that the DDP method can converge to a near-optimal solution if nominal controls are selected which are very close to the optimal since the trajectories (Figures 22 - 30) are very close in all cases. In actual applications this is very difficult to do since there is no sound basis for arriving at good initial control guesses without long computer solutions (Ref 11).

Table X

Control Comparison

Time (Secs)	Optimal Pursuer (Deg)	DDP Pursuer (Deg)	Prop Nav Pursuer (Deg)
.5	25	21	16
1.0	25	22	19
1.5	25	23	22
2.0	25	23	25
2.5	25	24	28
3.0	25	24	31
3.5	29	23	34
4.0	35	22	38
4.5	39	20	44
5.0	45	20	55
5.5	53	29	72
6.0	75	-	40

VI. Conclusions and Recommendations

Conclusions

The DDP method has been applied to a non-linear, differential game modeled air-to-air intercept problem. A variety of initial positions has been examined as indicated in Chapter V and Appendix A, and a real-time, closed-loop implementation of the guidance law was not realized. The trajectory deviations which arise from large computed control changes, Δu and Δv, are too large and force the adoption of the CCP technique. Although this assures convergence, the computation time is not suitable for real-time applications.

The restriction to five iterations does not allow near-optimal controls to be found due to the slow convergence characteristics as evidenced by the large values of $a_e(t_o)$ and $a_p(t_o)$ given in Chapter V. It is concluded that any real-time guidance algorithm (for a short duration air-to-air missile) which hopes to achieve an increase in kill probability through the use of differential game theory will have to sacrifice some of the realism involved through the inclusion of nonlinear dynamics.

The algorithm demonstrates a strong reliance on the initial controls. The closer the initial controls to the optimal, the better the DDP method solutions (more near-optimal). The possibility of storing several near-optimal control histories aboard a small missile is doubtful, and merely guessing one does not lead to good results.

The tactics which arise through this simulation in a closed-loop manner are very similar to those which are used in air-to-air combat situations. The use of DDP, therefore, may be worthwhile in an off-line computer simulation to test selected tactics.

Finally, it is concluded that although the traditional iterative (neighboring extremal) TPBVP solution requirement of needing very good "guesses" for the initial co-states is eliminated with the DDP method, a strong dependence on initial conditions remains. The solution is very problem-dependent in that in high altitude intercepts (above 30,000 feet) the aircraft is penalized due to lower thrust and a reduction in the maximum load factor. At lower altitudes, the aircraft's improved performance and the nonoptimality resulting from the five iteration restriction allow the aircraft to defeat the missile as shown in Figures 10 and 18.

It has been determined (Chapter V) that the DDP method applied to the aircraft improves the evasion (increases the final range) over the selection of a particular selected strategy which employs no updating method. The final set of trajectories, Figures 22 - 30, show that the solution will converge to the optimal, given sufficient time. This "long-time" requirement precludes the real-time implementation.

Recommendations

It is recommended that further study in this area be directed toward simplification of the dynamics. A

reduction in the complexity of the models, while not completely relinquishing the basic characteristics, may allow some of the potential gains inherent in the differential game philosophy to be achieved.

Beyond the model simplification, an extension of the intercept problem to include the case of multiple missile launches may provide some useful tactical information. Presently, the results confirm accepted logic for air-to-air counter-maneuvers. Perhaps a new or revised tactic may be discovered through simulation of different encounters.

Bibliography

1. Anderson, G. M. "A Transition Matrix Method for Generating Near Optimal Closed-Loop Solutions to Nonlinear Zero-Sum Differential Games." Preprint of International Federation of Automatic Control - International Federation of Operational Research Symposium on Optimization Methods, Applied Aspects, Varna, Bulgaria, October 8 - 11, 1974, pp 27-35.

2. -----. "A Near Optimal Closed-Loop Solution Method for Nonsingular Zero-Sum Differential Games." Journal of Optimization Theory and Applications, 13:303-318 (March 1974).

3. -----. "A Real-Time Closed-Loop Solution Method for a Class of Nonlinear Differential Games." IEEE Transactions Automatic Control, 10:576-577 (August 1972).

4. Asher, R. B. and J. P. Matuszewski, "Optimal Guidance with Maneuvering Targets." Journal of Spacecraft and Rockets, 11:204-206 (March 1974).

5. Bryson, A. E. and Y. C. Ho, Applied Optimal Control. Washington, D. C.: The Halstead Press, 1975.

6. Isaacs, R. Differential Games. Huntington, New York: Robert E. Krieger Publishing Co., 1975.

7. Jacobson, D. H. and D. Q. Mayne, Differential Dynamic Programming. New York, New York: American Elsevier Publishing Company, Inc., 1970.

8. Jarmarck, B. A New Convergence Control Technique in Differential Dynamic Programming. Trita - Reg - 7502. Stockholm, Sweden: The Royal Institute of Technology, December 1975.

9. -----. Near-Optimal Closed-Loop Strategy for Aerial Combat Games. Trita - Reg - 7602. Stockholm, Sweden: The Royal Institute of Technology, March 1976.

10. Murtaugh, S. A. and H. E. Criel, "Fundamentals of Proportional Navigation." IEEE Spectrum, Vol 3, No. 12:75-85 (December 1966).

11. Poulter, R. A. "Differential Game Guidance Versus Proportional Navigation for an Air-To-Air Missile", Masters Thesis, Wright Patterson Air Force Base, Ohio: Air Force Institute of Technology, December 1975.

Appendix A

Trajectory Analysis

The following graphs depict the three-dimensional flight paths and ground tracks of the missile and aircraft resulting from a closed-loop application of the guidance scheme. In Figures 4 - 21, the aircraft uses the DDP method while the missile uses either the DDP algorithm or proportional navigation as indicated on each graph. Figures 22 - 30 are the exact TPBVP solution comparisons with Figures 22, 25 and 28 depicting the "exact" optimal solutions. The initial conditions for each set of graphs are given in the preceding tables, Tables XI - XXII.

Table XI

Initial Conditions

State	Evader	Pursuer
x ft	5000	-1000
y ft	5000	6000
z ft	33160	33160
v ft/sec	706	2219
γ rads	0.0	-.01
σ rads	.524	.01
Control		
n g's	2	1
μ rad	1.57	0.0

DDP EVASION-DDP PURSUIT

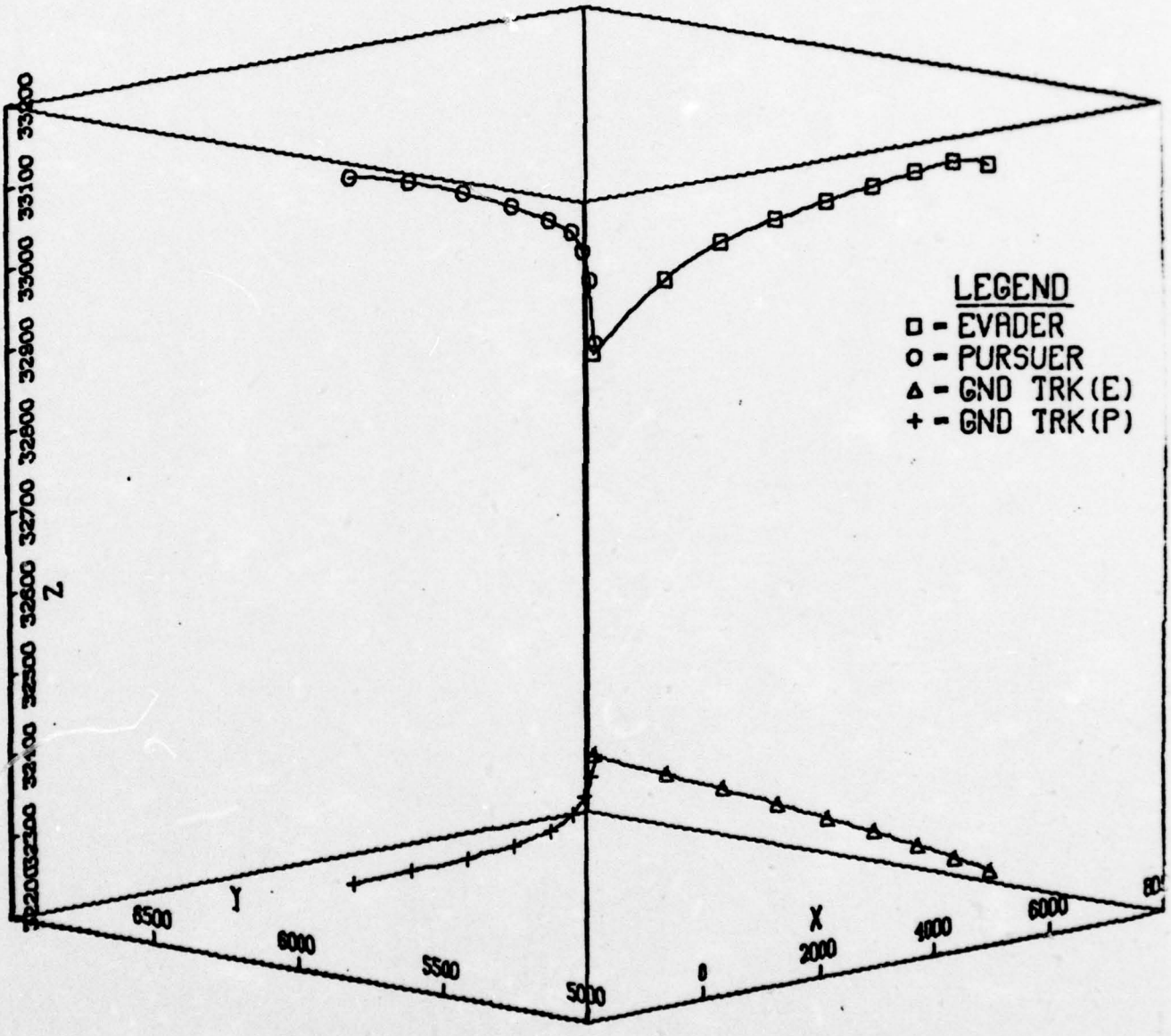


Fig. 4

DDP EVASION-PROP NAV PURSUIT

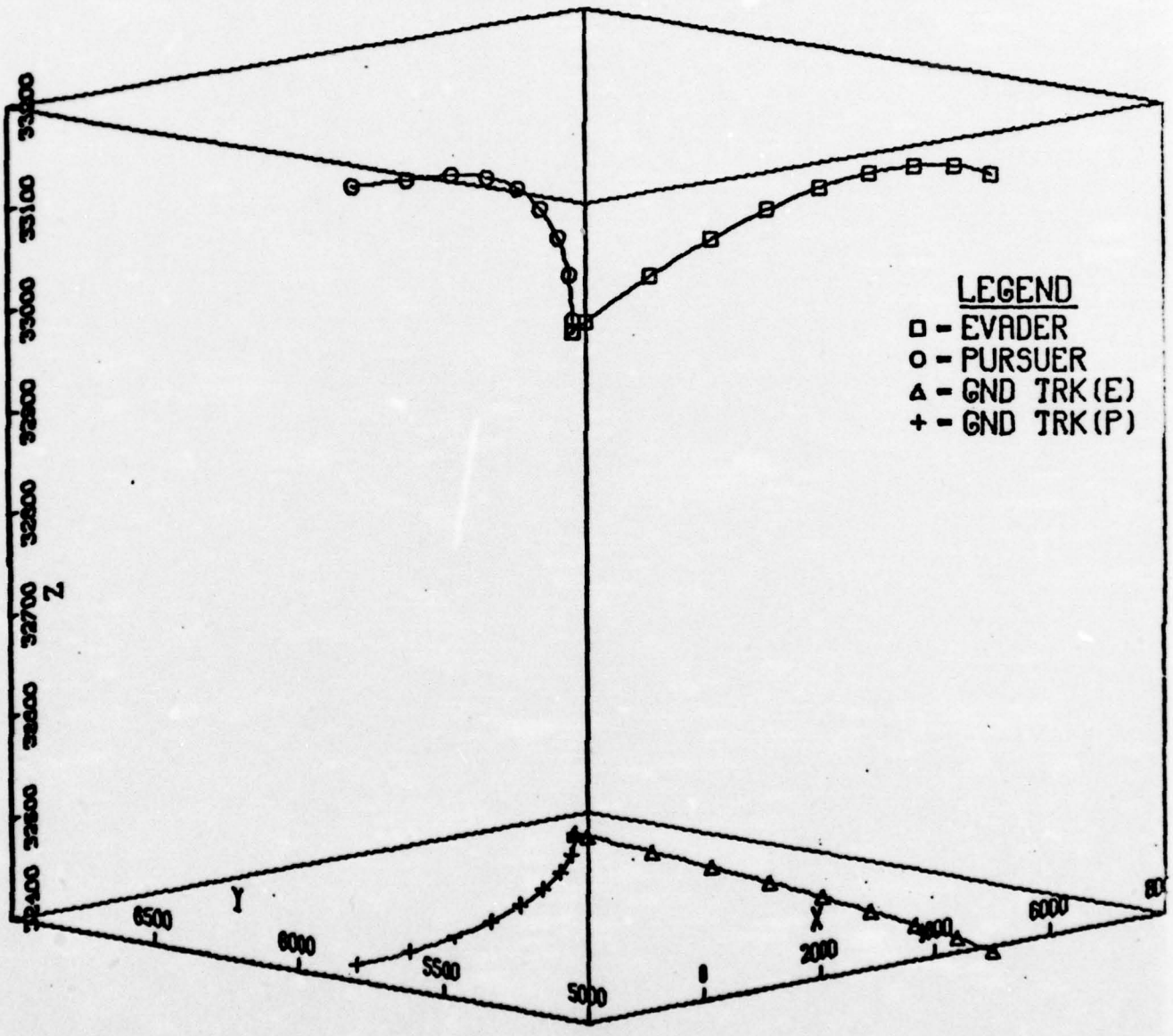


Fig. 5

Table XII

Initial Conditions

State	Evader	Pursuer
x ft	10000	4000
y ft	10000	9000
z ft	33160	33160
v ft/sec	706	2219
γ rads	0.0	-.01
σ rads	-.524	-.007
Control		
n g's	2	1
μ rad	1.57	0.0

DDP EVASION-DDP PURSUIT

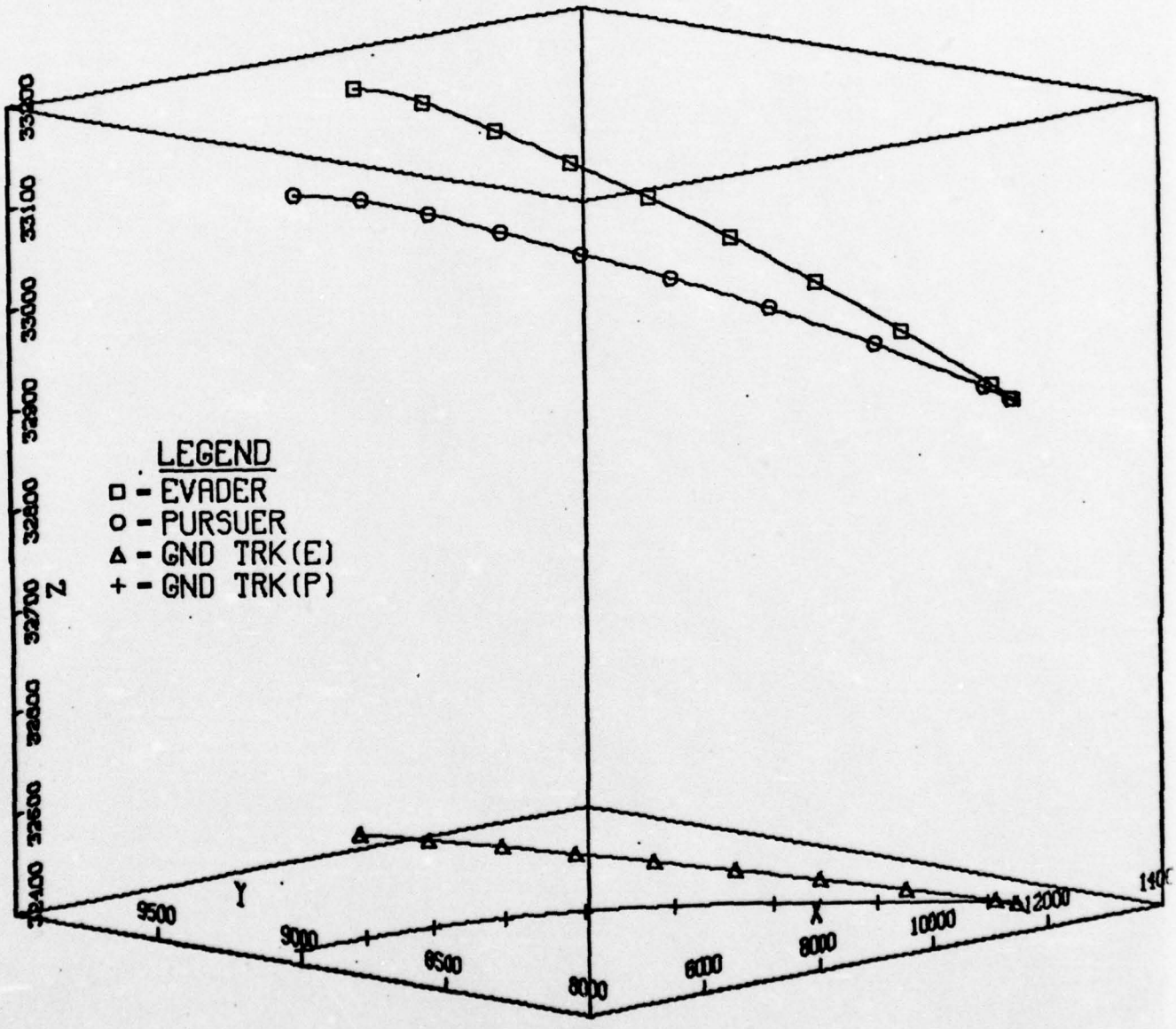


Fig. 6

DDP EVASION-PROP NAV PURSUIT

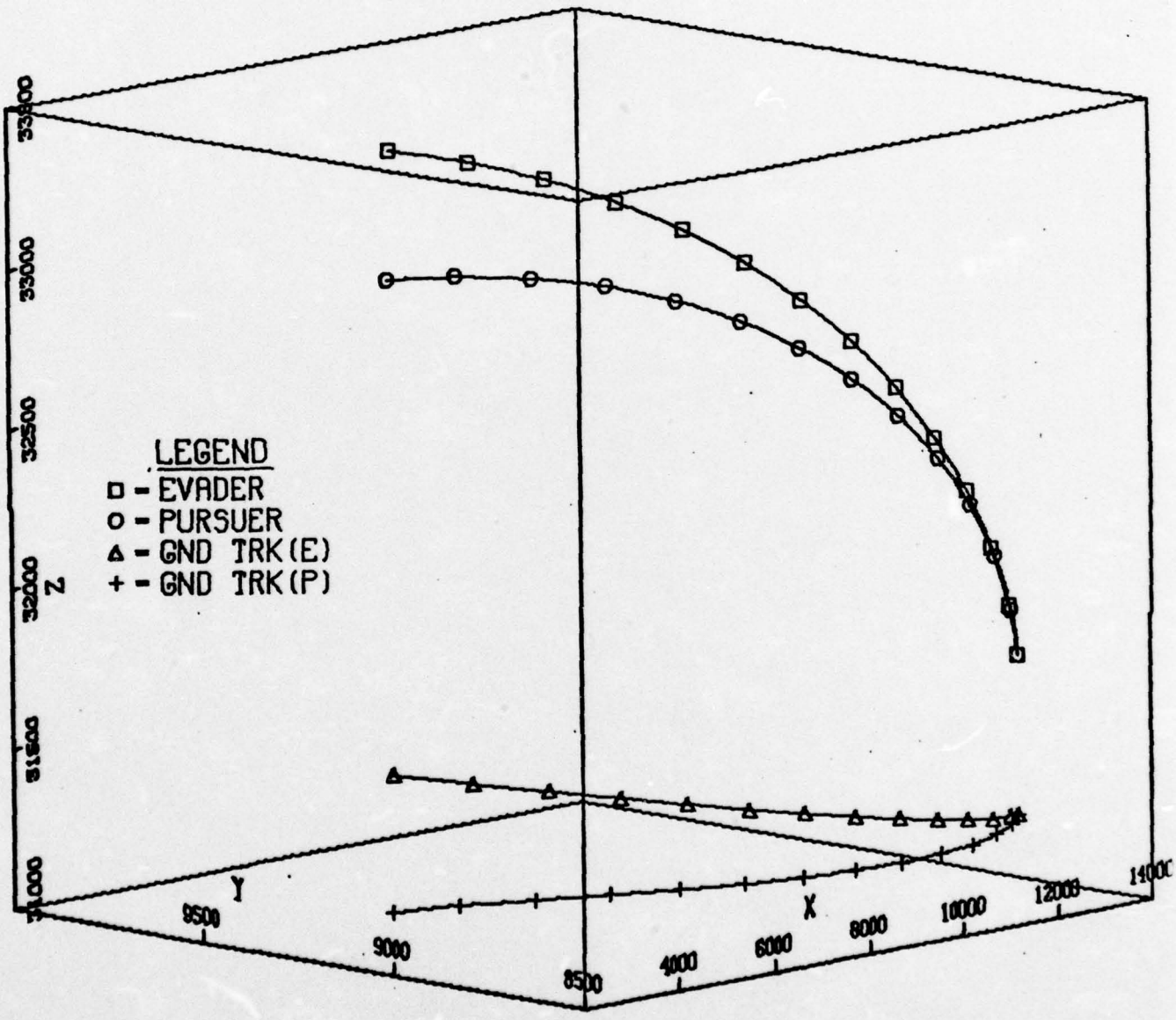


Fig. 7

Table XIII

Initial Conditions

State	Evader	Pursuer
x ft	10000	2000
y ft	10000	9000
z ft	33000	34000
v ft/sec	706	2219
γ rads	-.524	-.3
σ rads	-.524	.01
Control		
n g's	2	1
μ rads	1.57	0.0

DDP EVASION-DDP PURSUIT

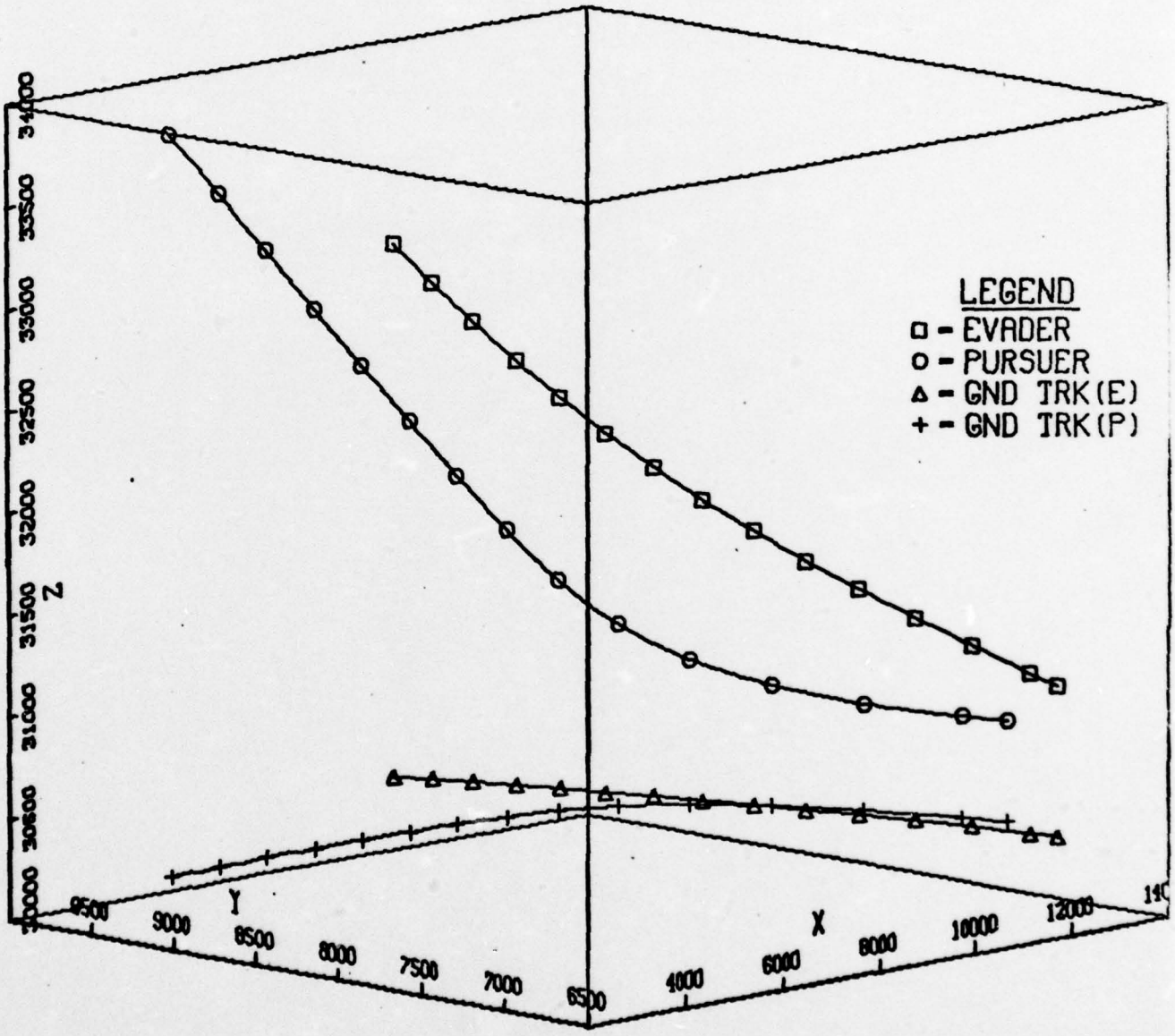


Fig. 8

DDP EVASION-PROP NAV PURSUIT

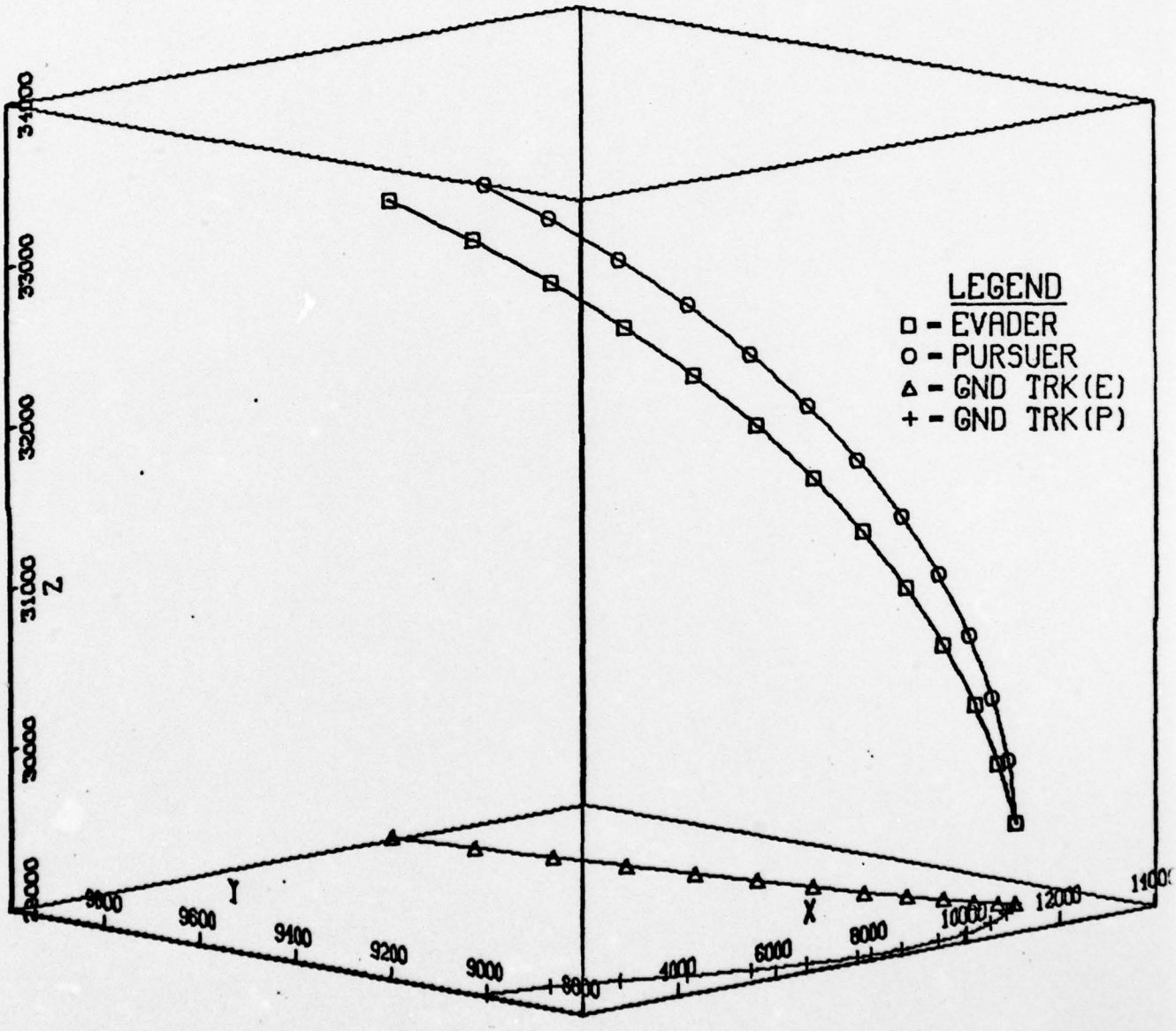


Fig. 9

Table XIV

Initial Conditions

State	Evader	Pursuer
x ft	5000	0.0
y ft	5000	6000
z ft	15000	15000
v ft/sec	770	2065
γ rads	-.1	-.02
σ rads	.524	.01
Control		
n g's	2	1
μ rads	1.05	0.0

DDP EVASION-DDP PURSUIT

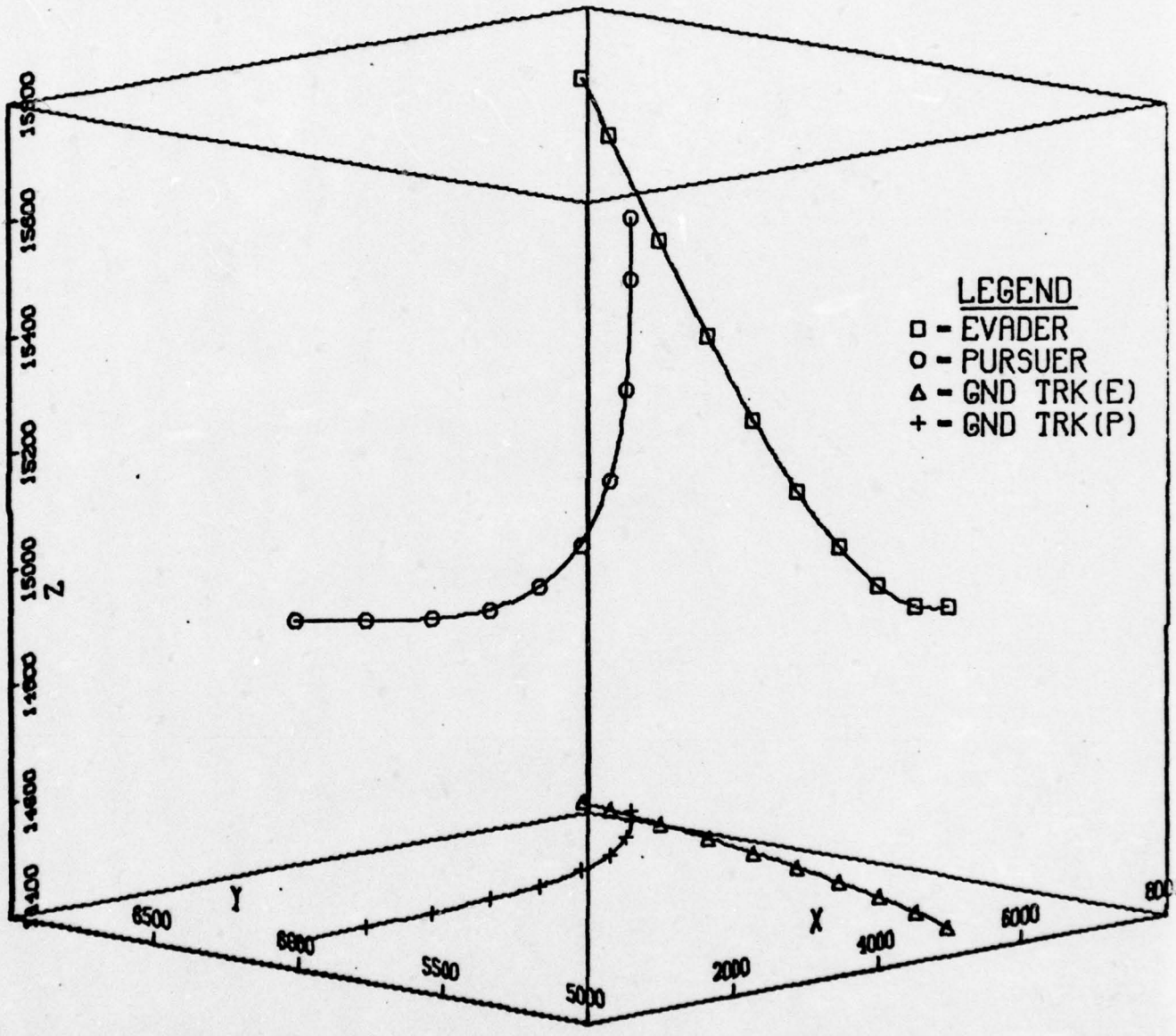


Fig. 10

DDP EVASION-PROP NAV PURSUIT

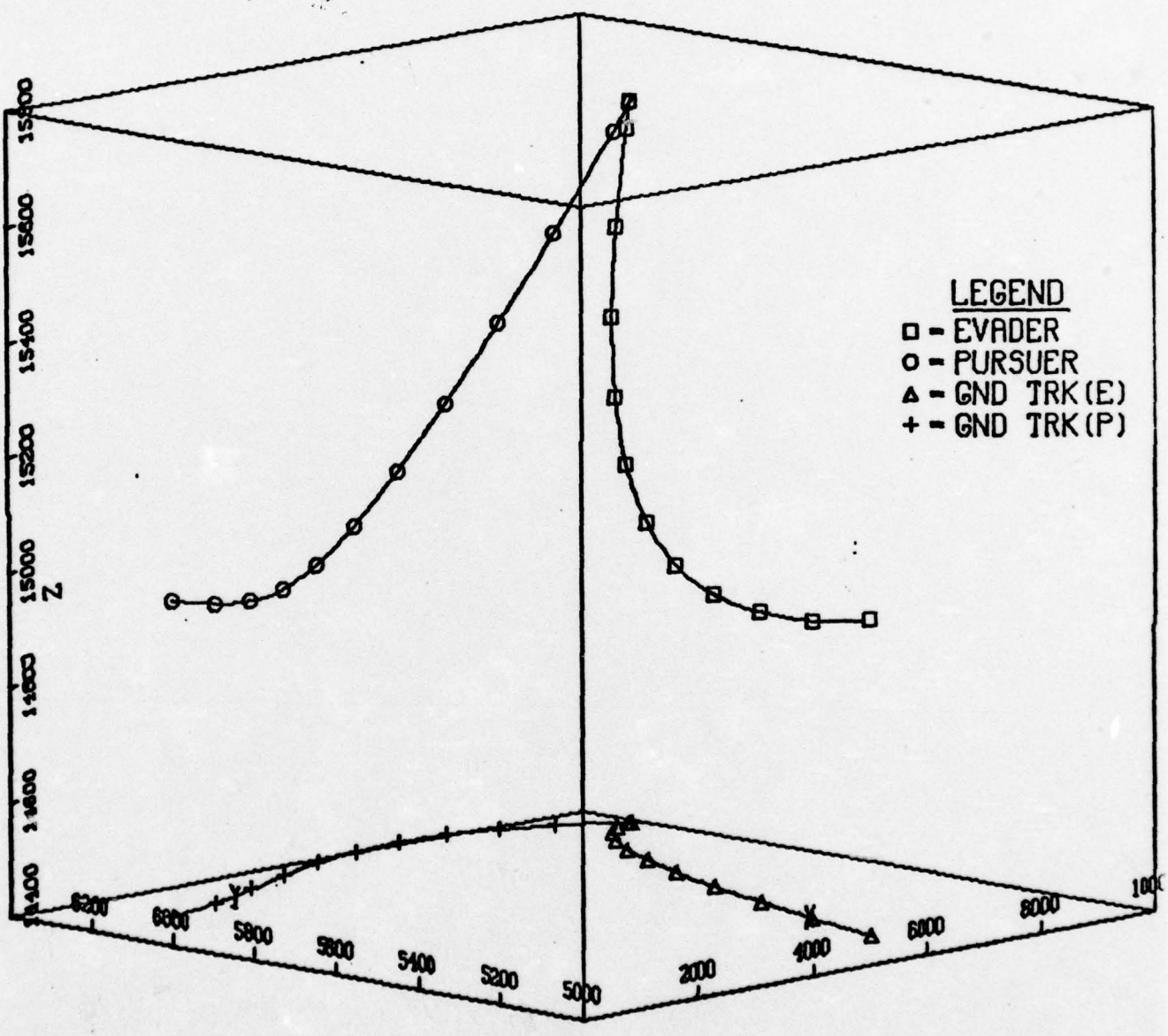


Fig. 11

Table XV

Initial Conditions

State	Evader	Pursuer
x ft	10000	12000
y ft	10000	2000
z ft	33160	33160
v ft/sec	706	2219
γ rads	0.0	-.01
σ rads	0.0	1.57
Control		
n g's	2	1
μ rads	1.57	0.0

DDP EVASION-DDP PURSUIT

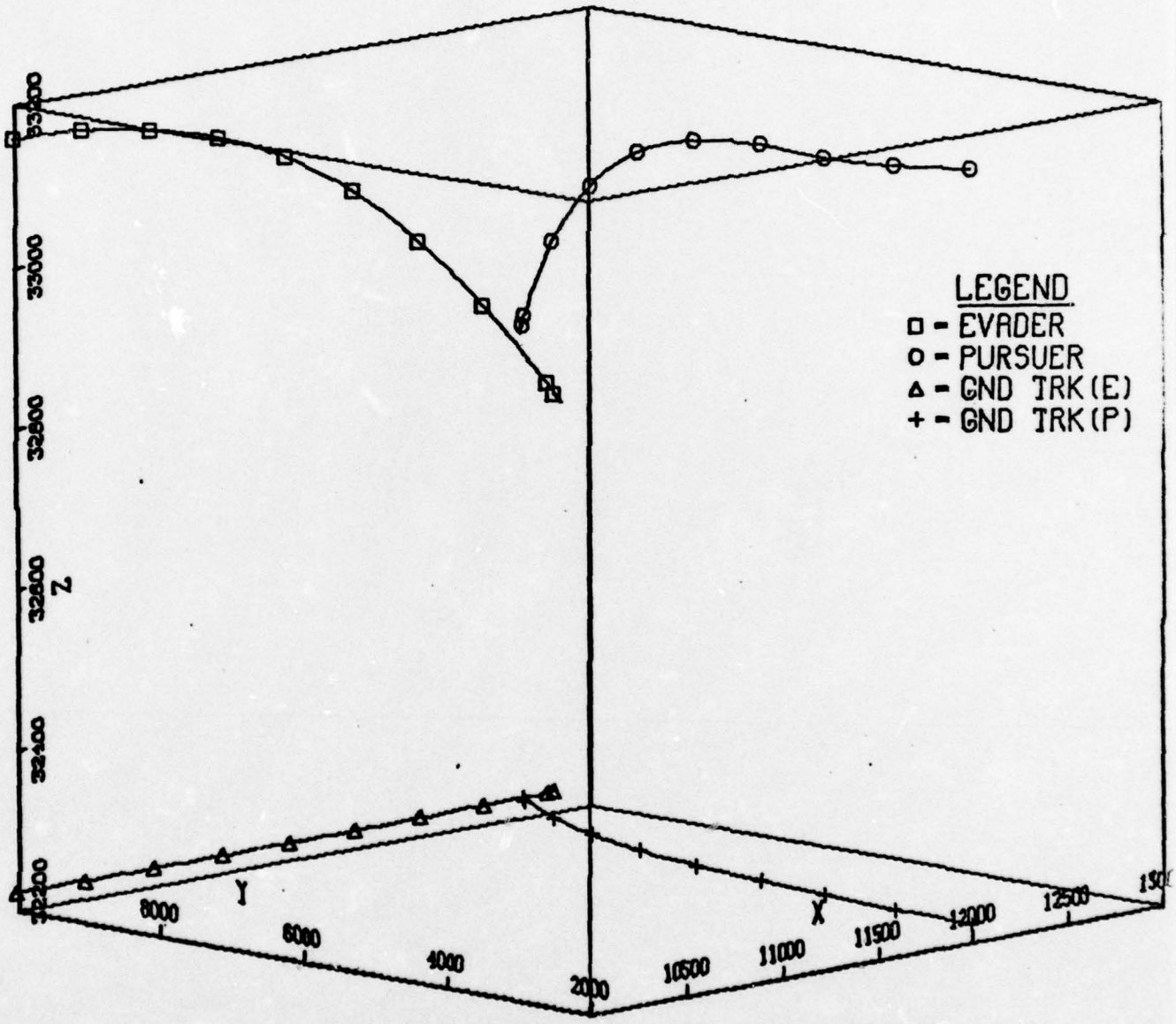


Fig. 12

DDP EVASION-PROP NAV PURSUIT

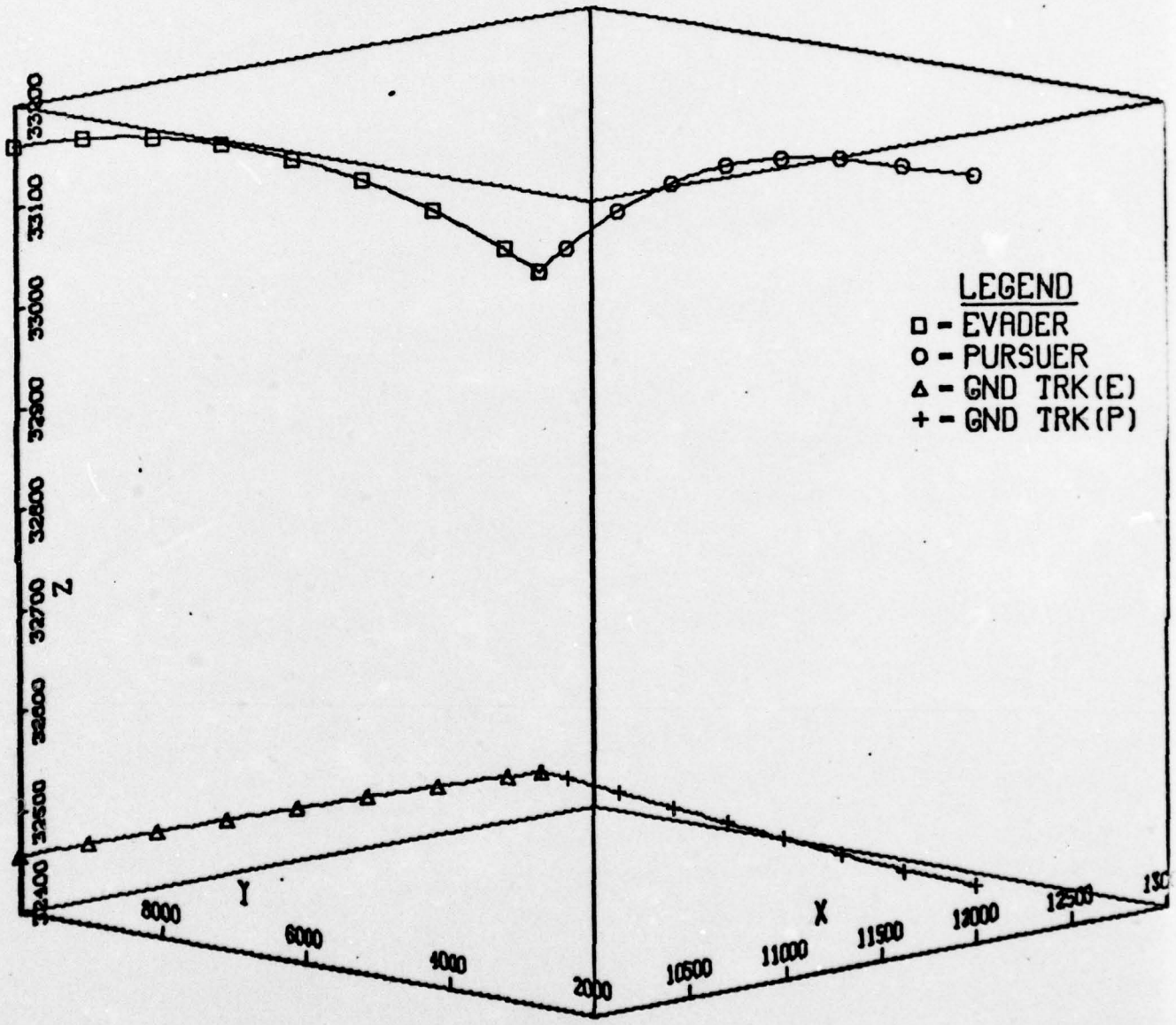


Fig. 13

Table XVI

Initial Conditions

State	Evader	Pursuer
x ft	10000	4500
y ft	10000	11000
z ft	15000	15000
v ft/sec	770	2065
γ rads	.1	.11
σ rads	.524	.01
Control		
n g's	3	1
μ rads	.524	0.0

DDP EVASION-DDP PURSUIT

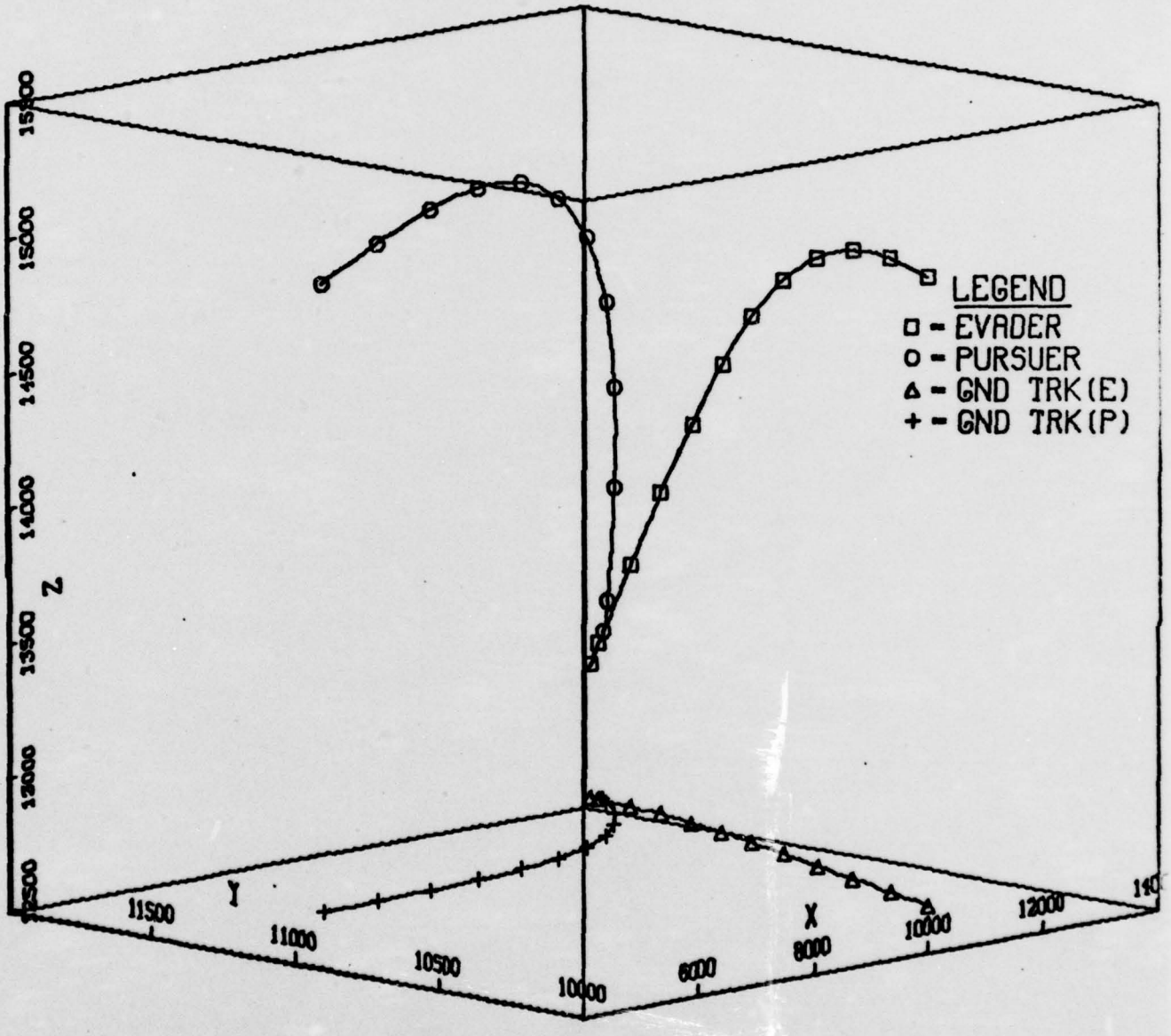


Fig. 14

DDP EVASION-PROP NAV PURSUIT

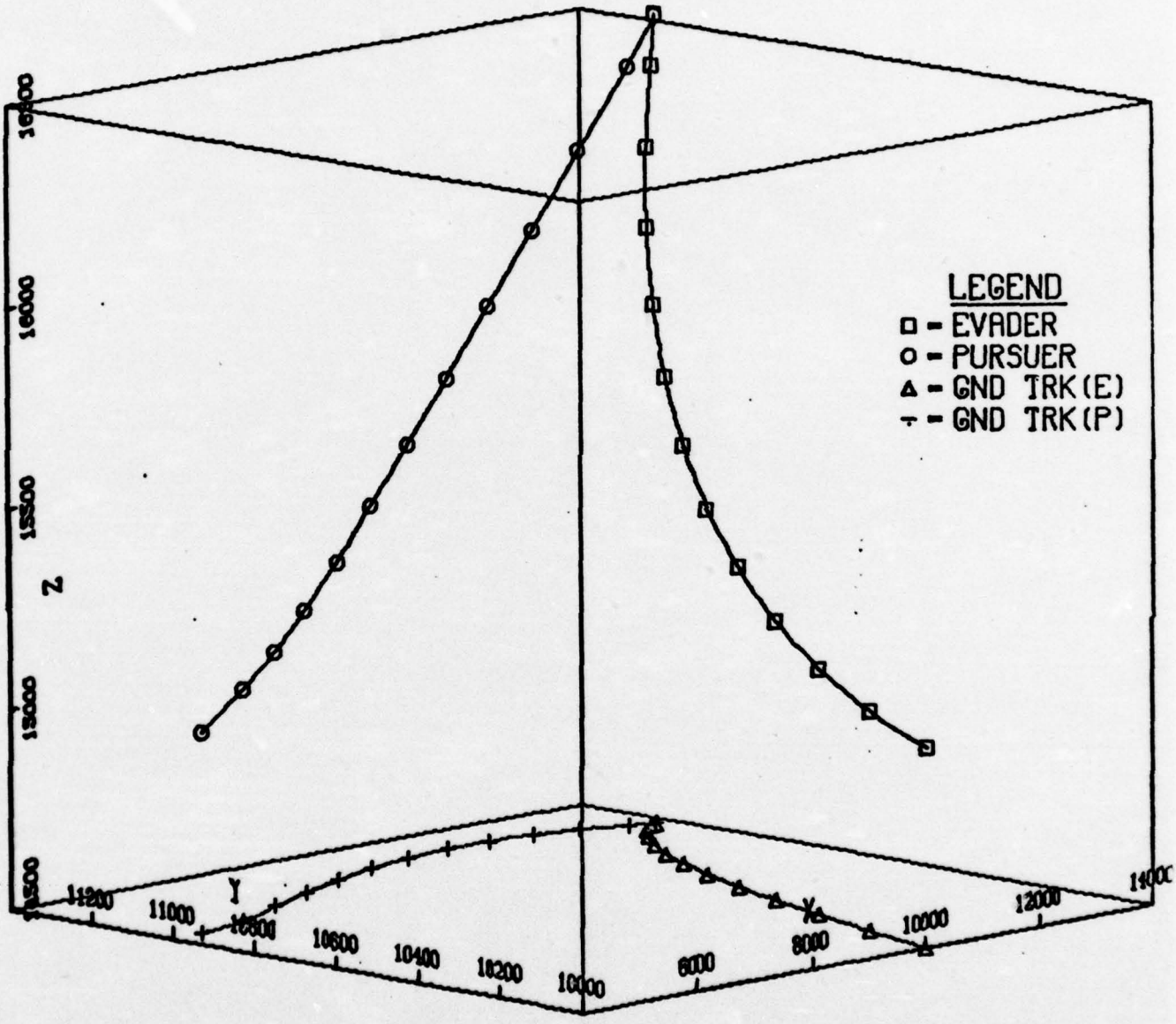


Fig. 15

Table XVII

Initial Conditions

State	Evader	Pursuer
x ft	10000	3000
y ft	10000	9000
z ft	33160	33160
v ft/sec	706	2219
γ rads	0.0	-.01
σ rads	-.524	-.001
Control		
n g's	2	1
μ rads	1.57	0.0

DDP EVASION-DDP PURSUIT

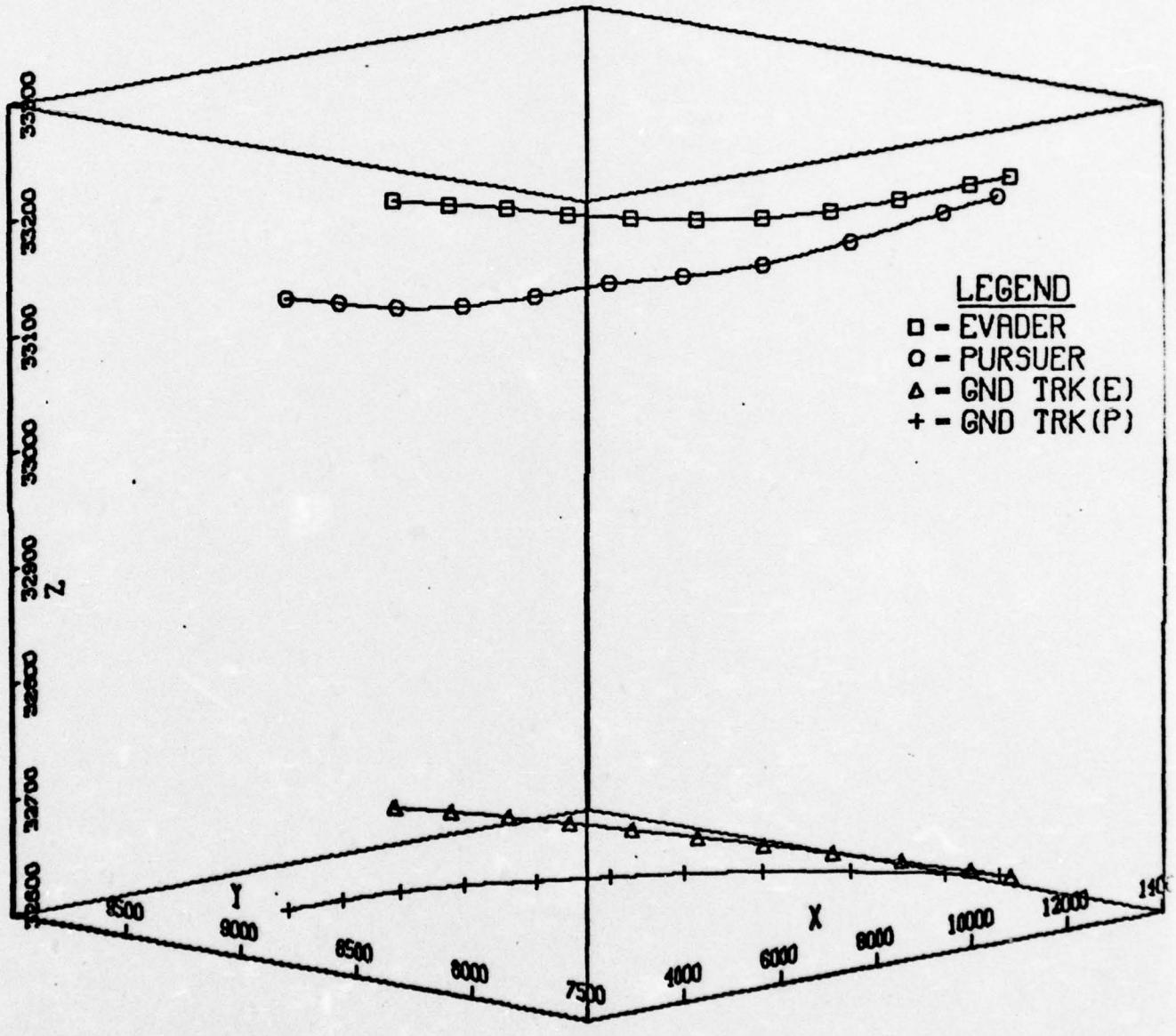


Fig. 16

DDP EVASION-PROP NAV PURSUIT

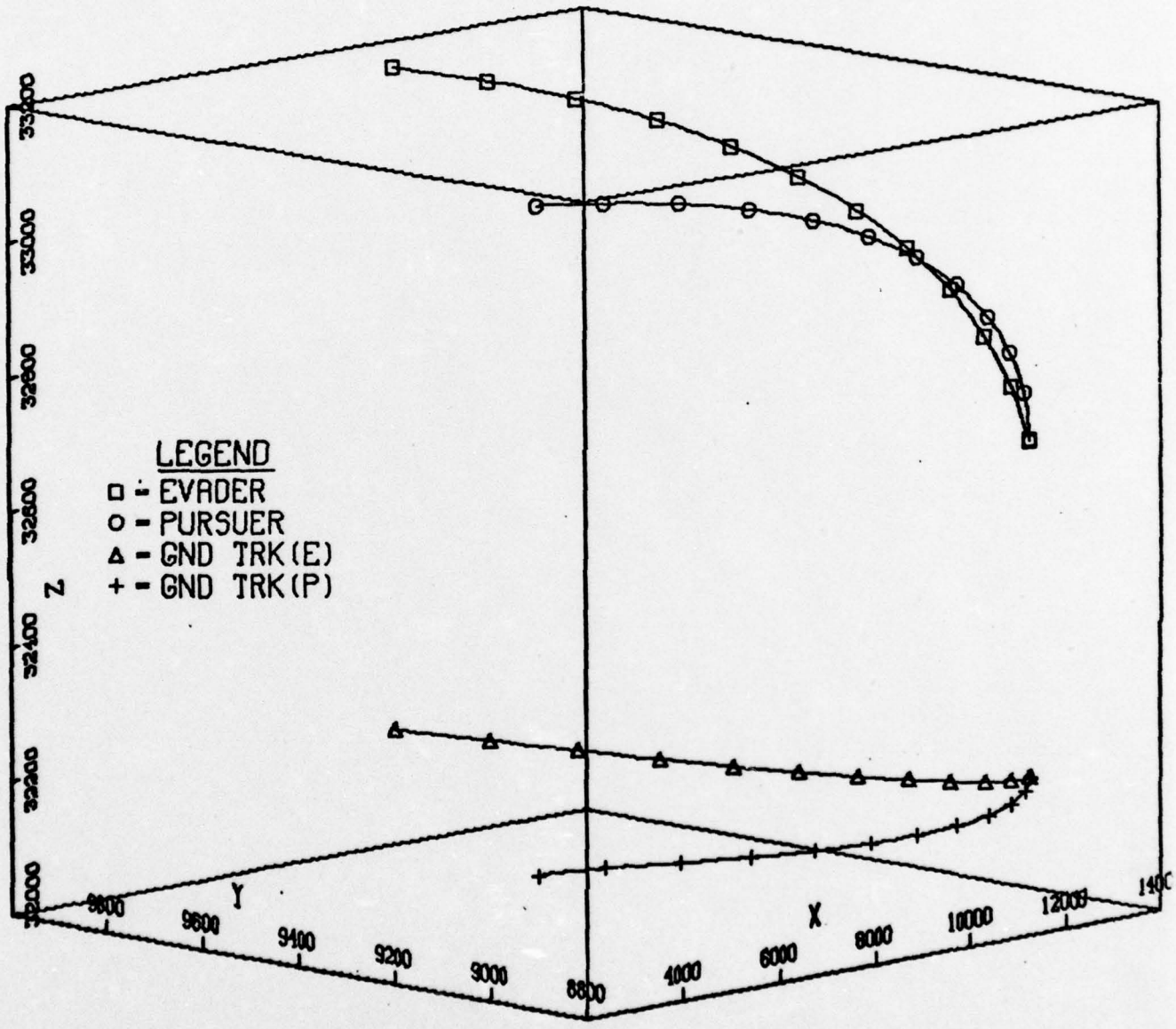


Fig. 17

Table XVIII

Initial Conditions

State	Evader	Pursuer
x ft	10000	5000
y ft	10000	9000
z ft	11000	11000
v ft/sec	850	2350
γ rads	0.0	-.01
σ rads	-.524	-.01
Control		
n g's	2	1
μ rads	1.57	0.0

DDP EVASION-DDP PURSUIT

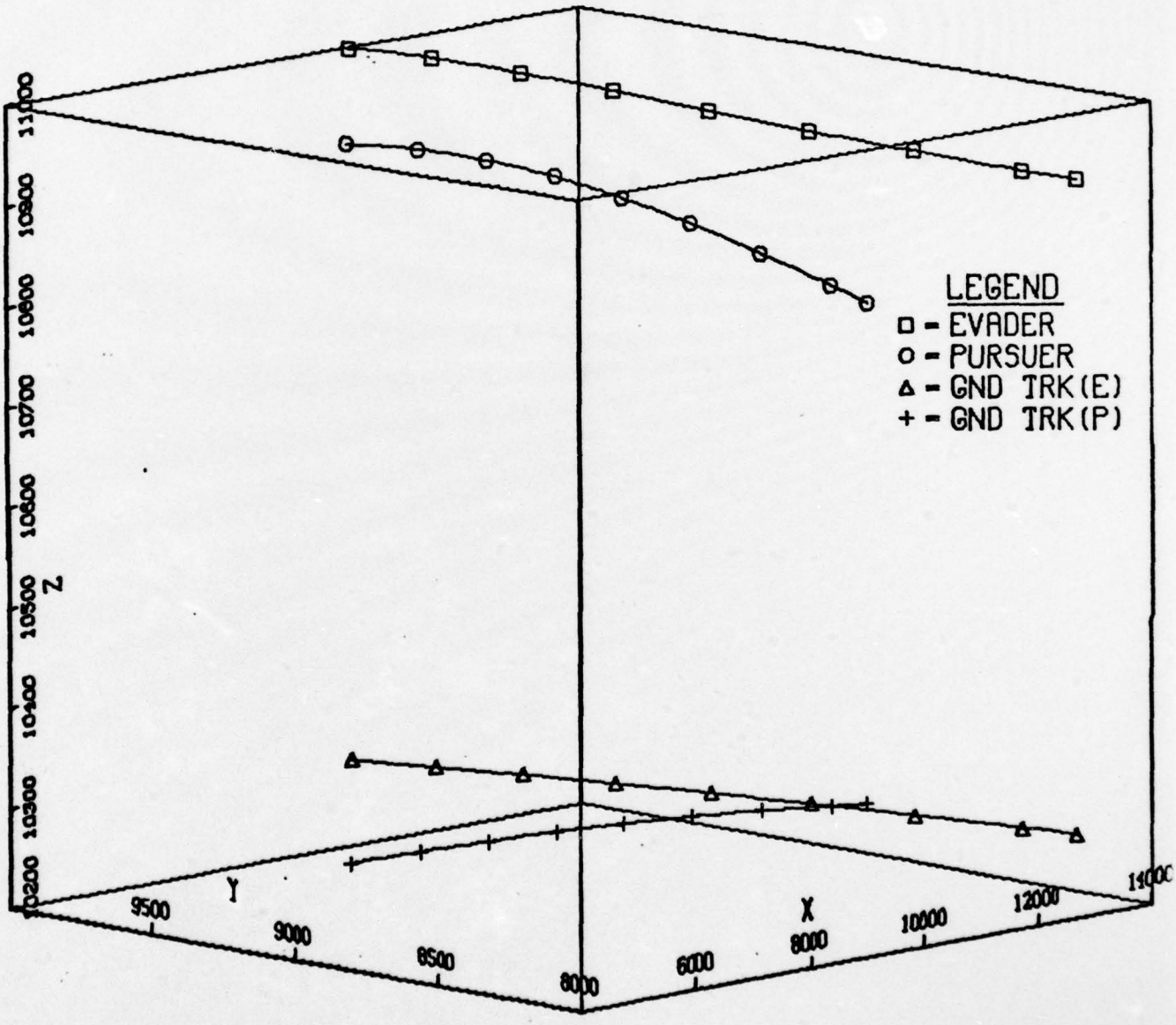


Fig. 18

DDP EVASION-PROP NAV PURSUIT

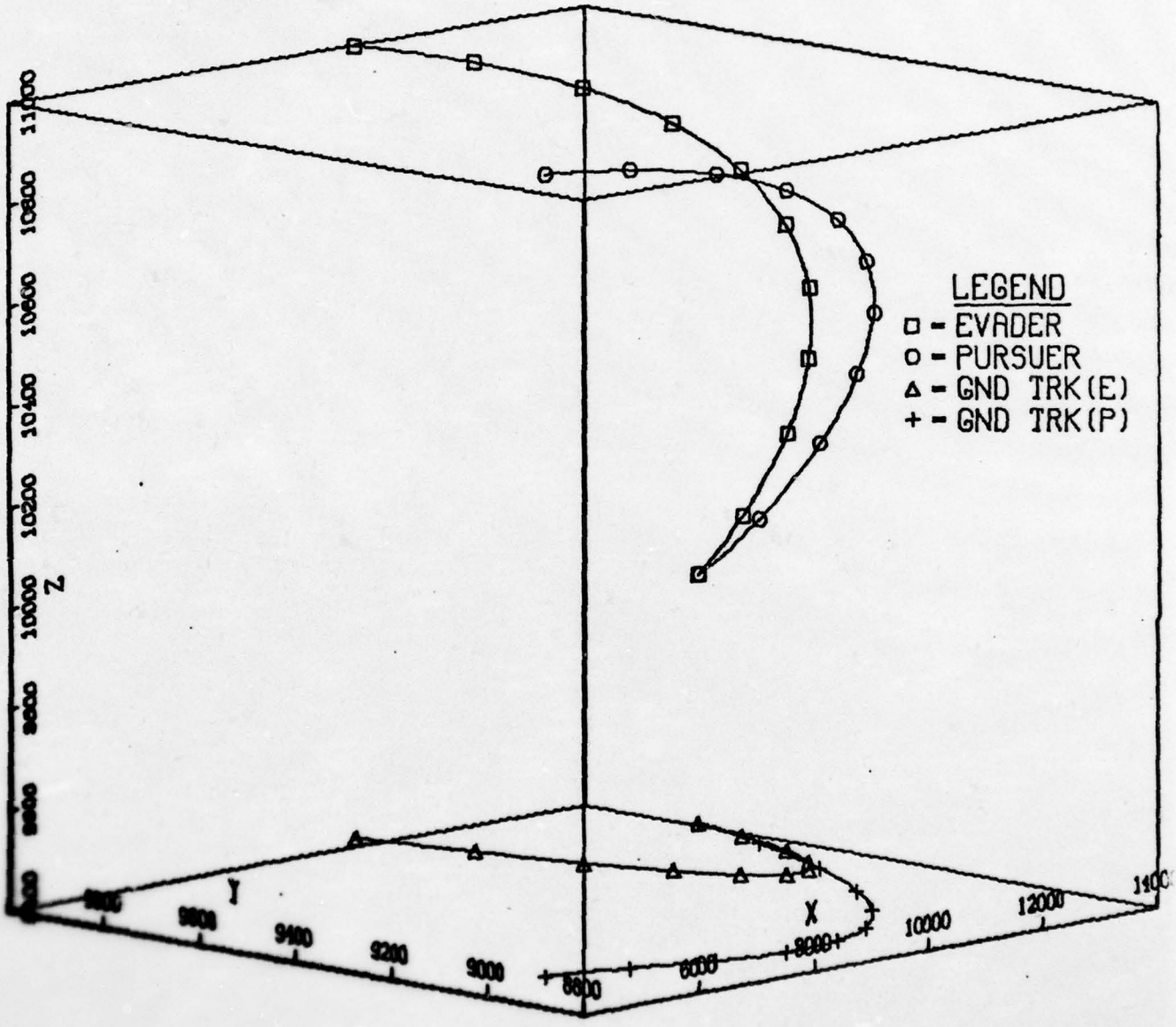


Fig. 19

Table XIX

Initial Conditions

State	Evader	Pursuer
x ft	10000	4000
y ft	10000	11000
z ft	33160	33160
v ft/sec	706	2219
γ rads	0.0	-.01
σ rads	.524	.01
Control		
n g's	3	1
μ rads	1.05	0.0

DDP EVASION-DDP PURSUIT

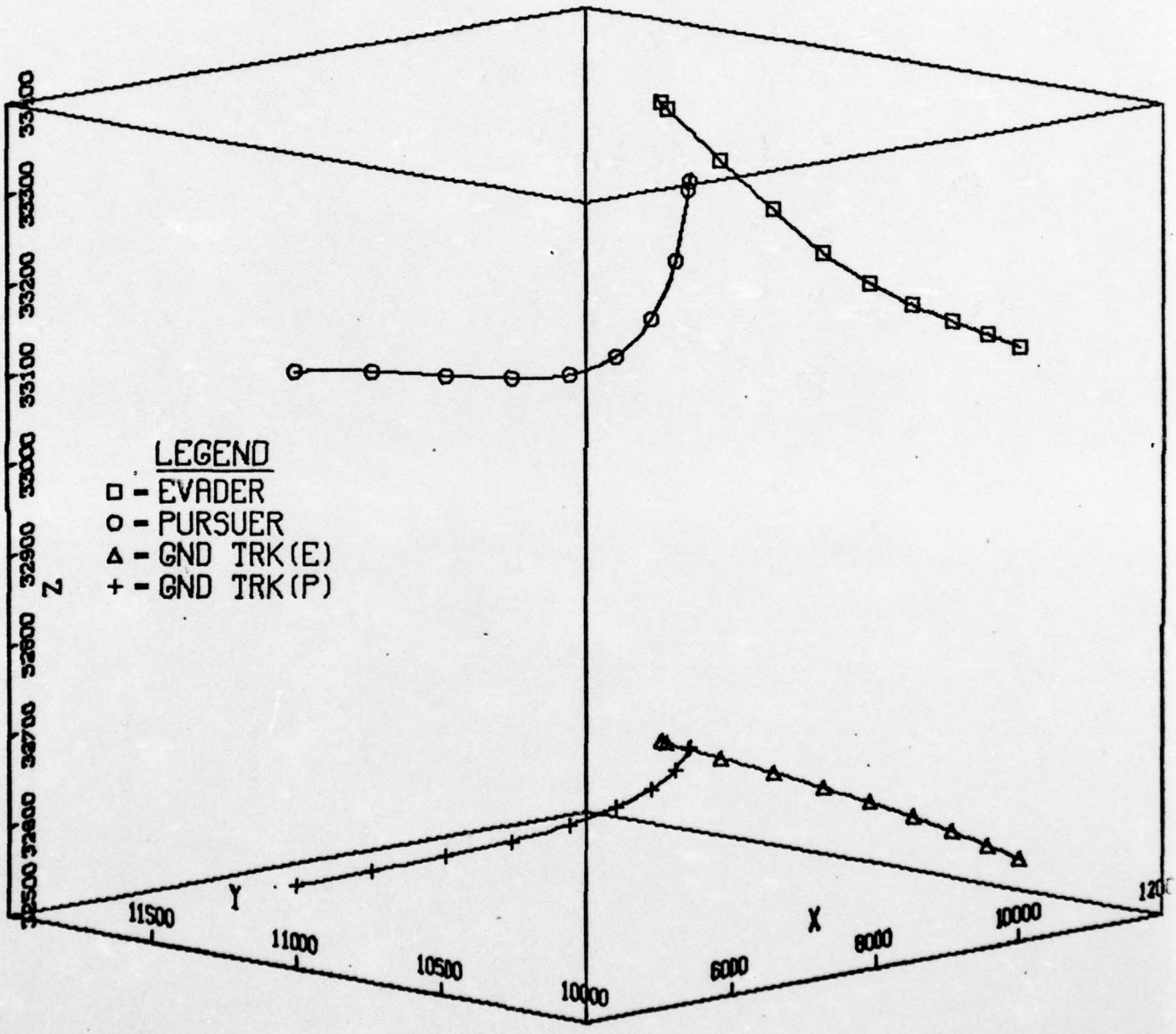


Fig. 20

DDP EVASION-PROP NAV PURSUIT

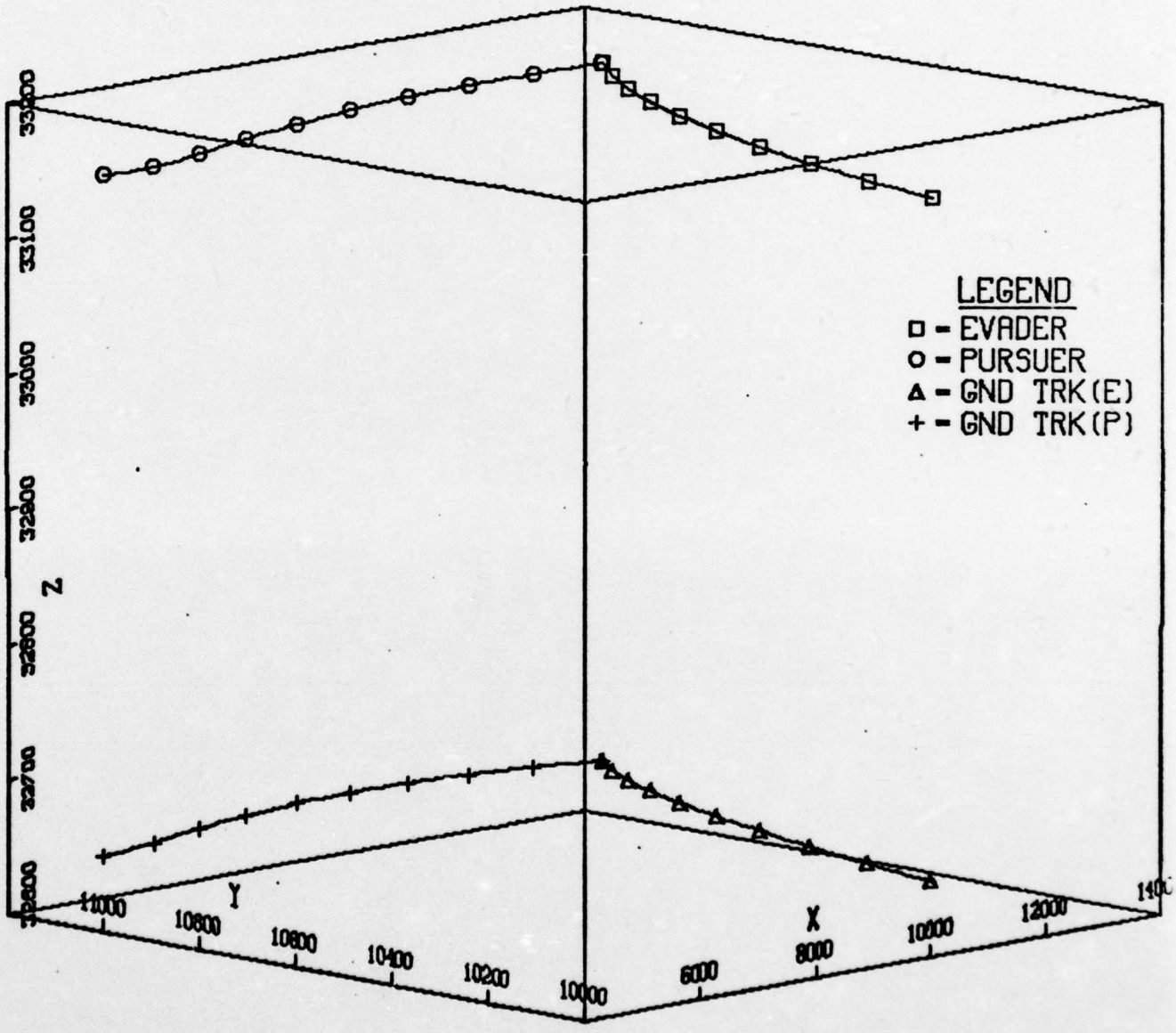


Fig. 21

Table XX

Initial Conditions

State	Evader	Pursuer
x ft	4178	10920
y ft	4435	10242
z ft	11123	7782
v ft/sec	849	2491
γ rads	-.5	-.29
σ rads	-1.2	-2.18
Control		
n g's	6	15
μ rads	1.05	.524

OPT EVASION-OPT PURSUIT

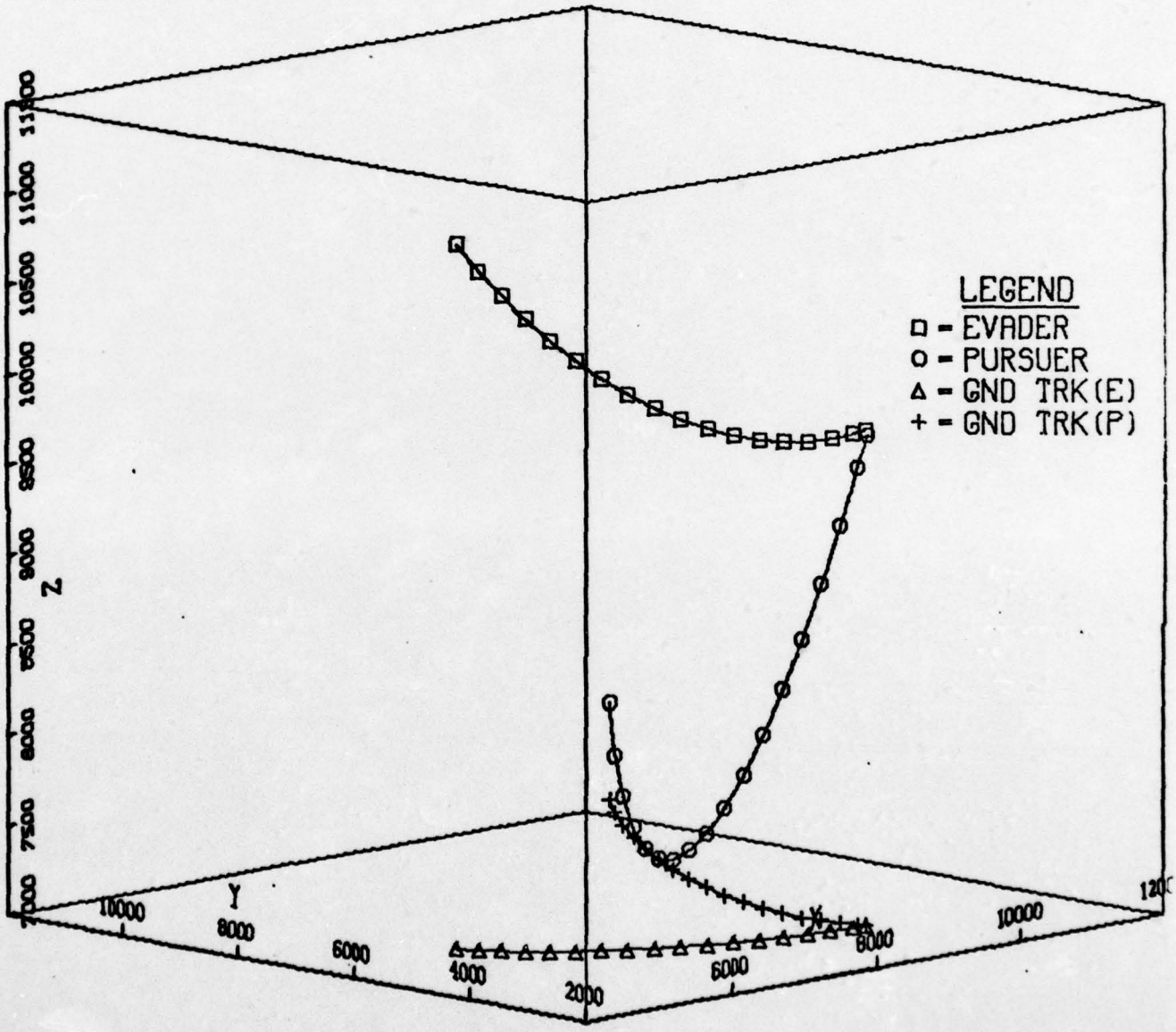


Fig. 22

DDP EVASION-DDP PURSUIT

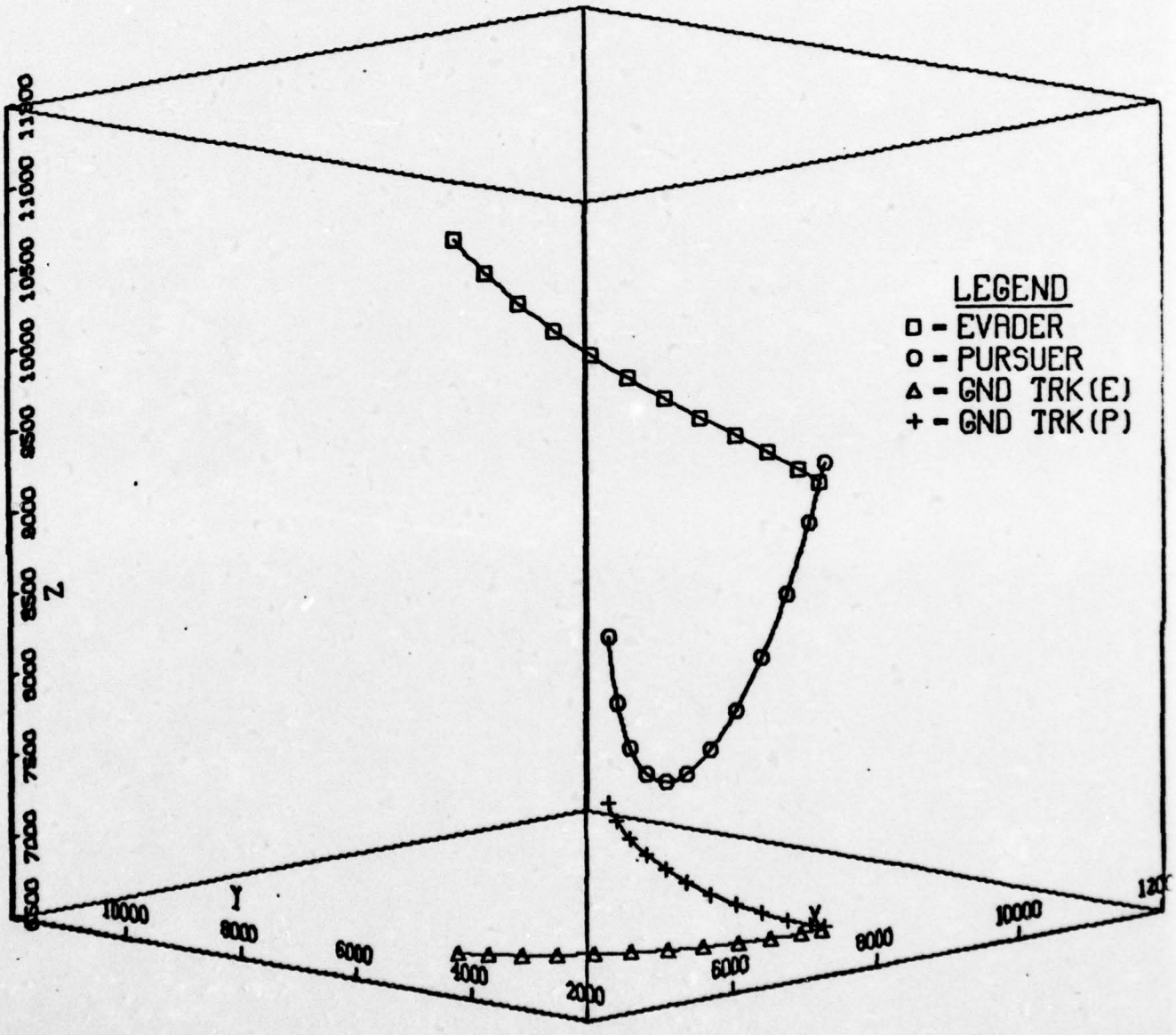


Fig. 23

DDP EVASION-PROP NAV PURSUIT

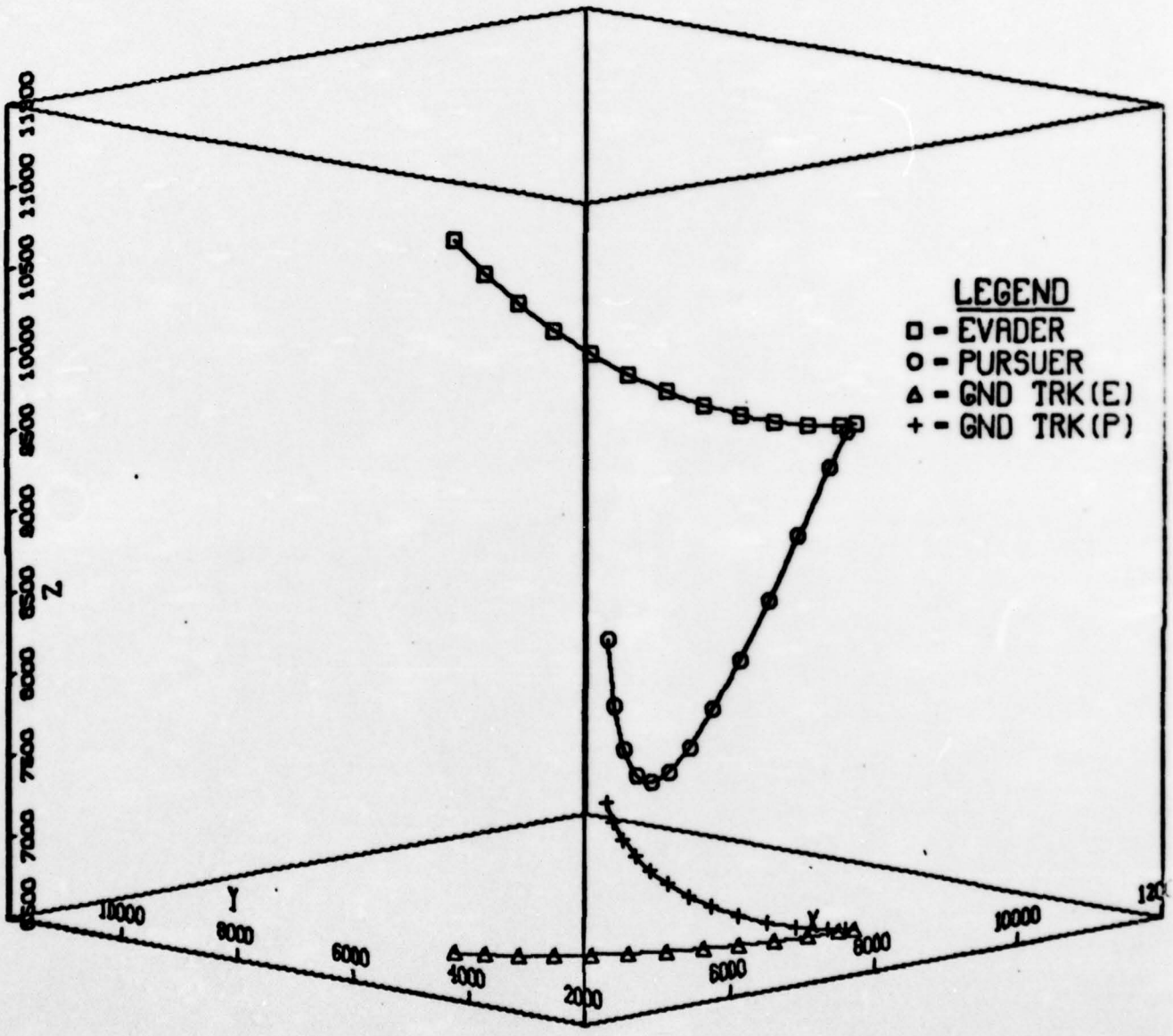


Fig. 24

Table XXI

Initial Conditions

State	Evader	Pursuer
x ft	4936	9341
y ft	3013	10358
z ft	33198	33326
v ft/sec	703	2235
γ rads	-.84	-.62
σ rads	-.64	-2.43
Control		
n g's	2	15
μ rads	0.0	1.22

OPT EVASION-OPT PURSUIT

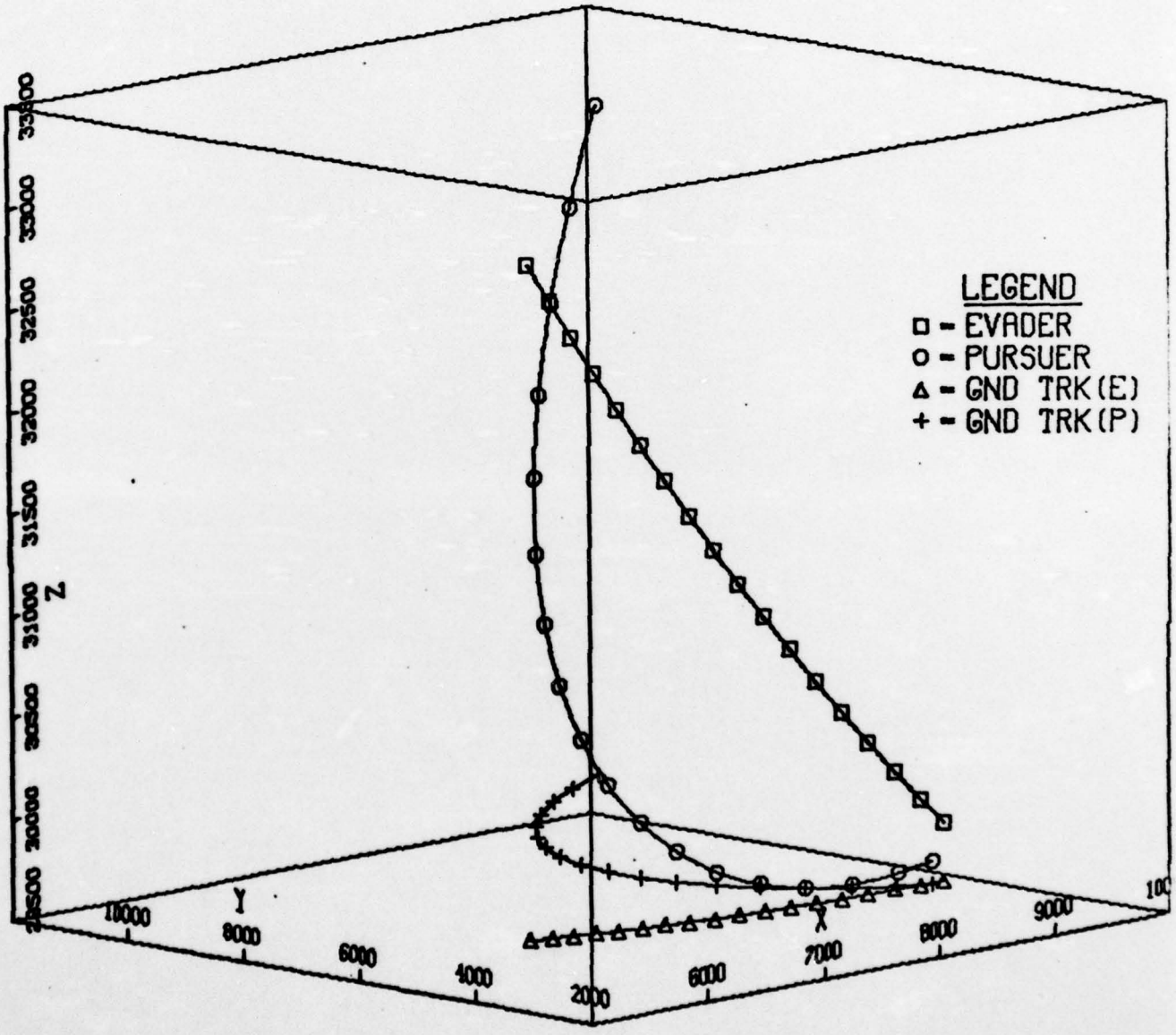


Fig. 25

DDP EVASION-DDP PURSUIT

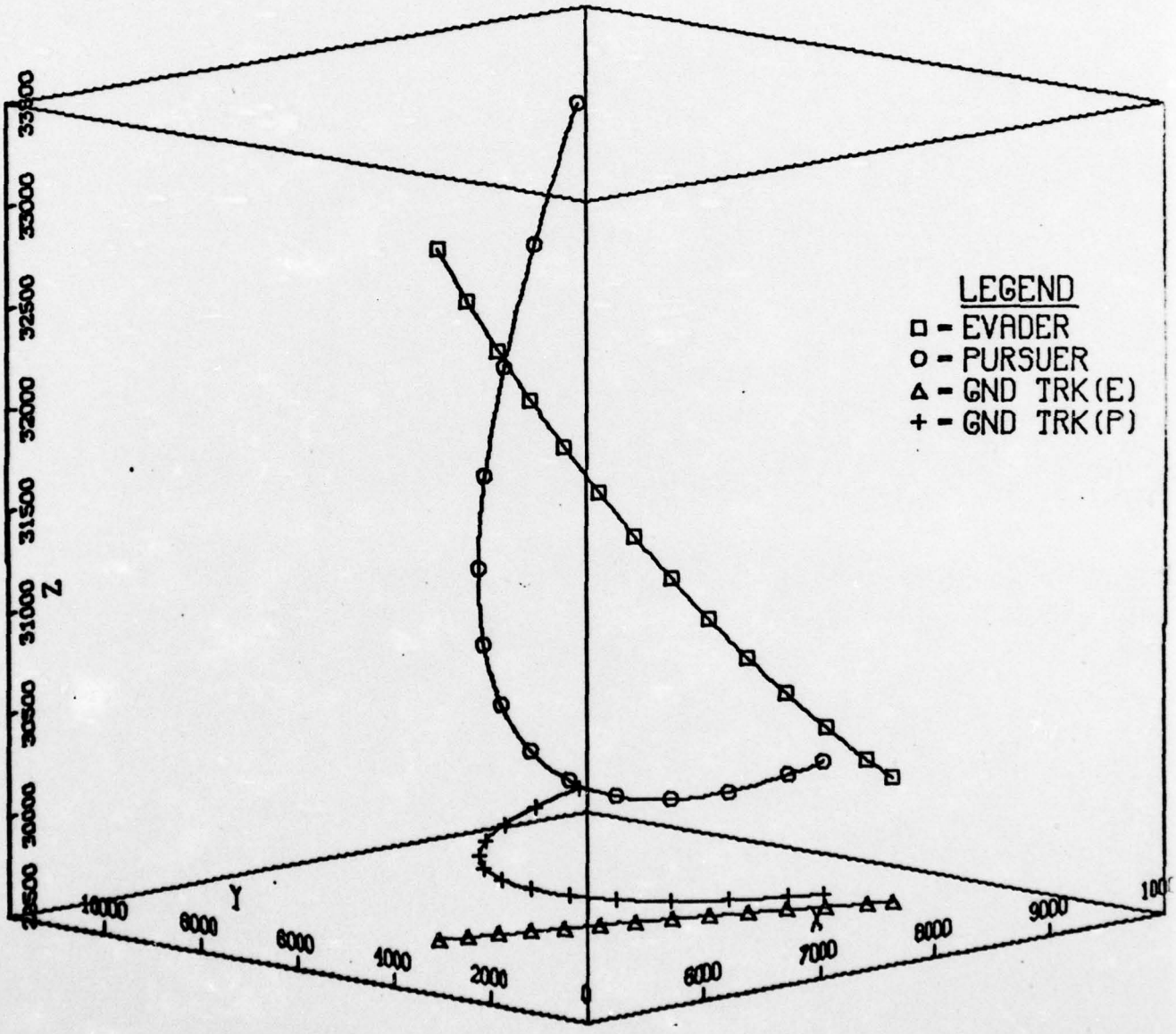


Fig. 26

DDP EVASION-PROP NAV PURSUIT

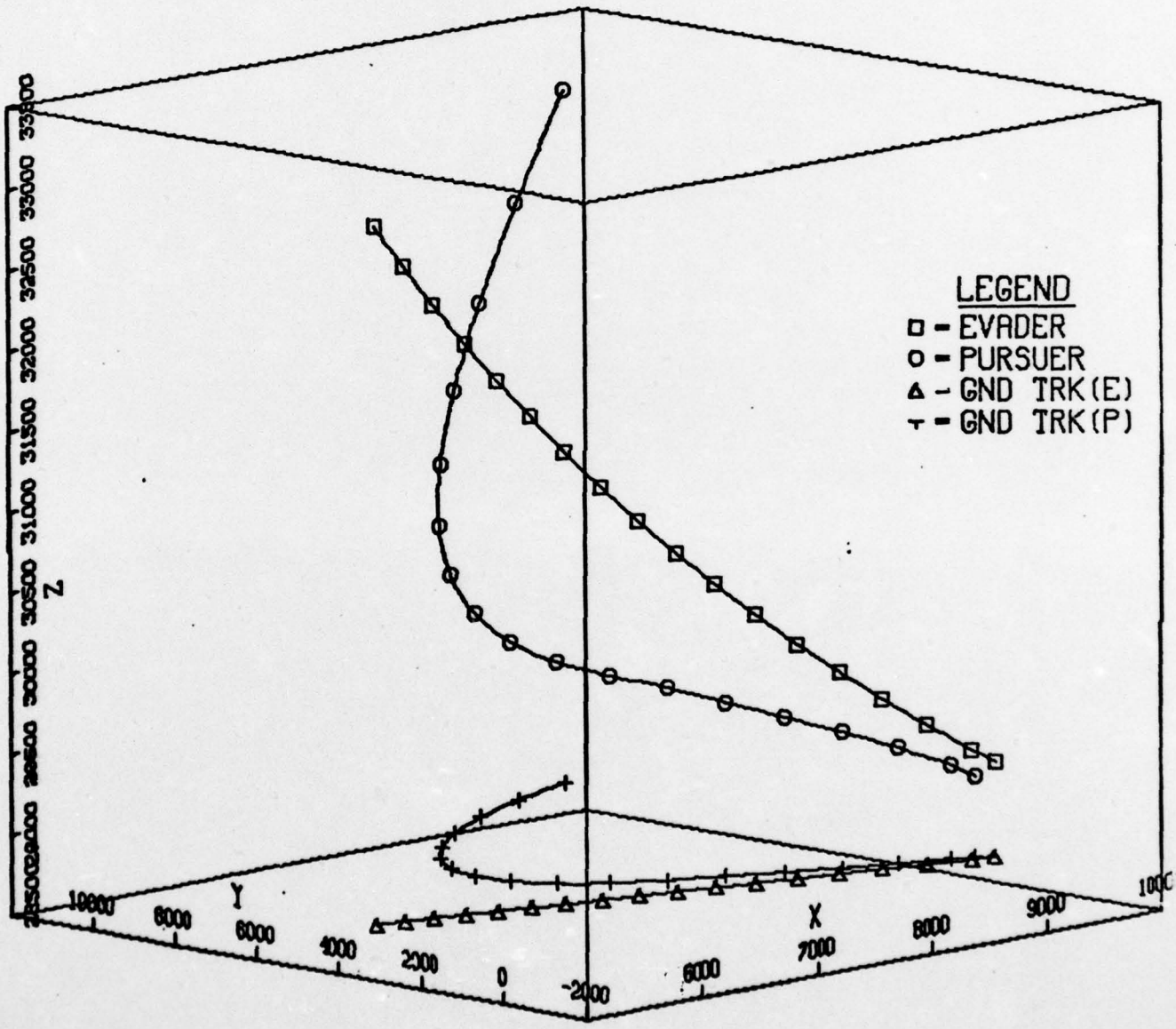


Fig. 27

Table XXII

Initial Conditions

State	Evader	Pursuer
x ft	4972	9199
y ft	2985	10260
z ft	33150	33170
v ft/sec	706	2219
γ rads	-.846	-.619
σ rads	-.632	-2.41
Control		
n g's	2	15
μ rads	0.0	1.22

OPT EVASION-OPT PURSUIT

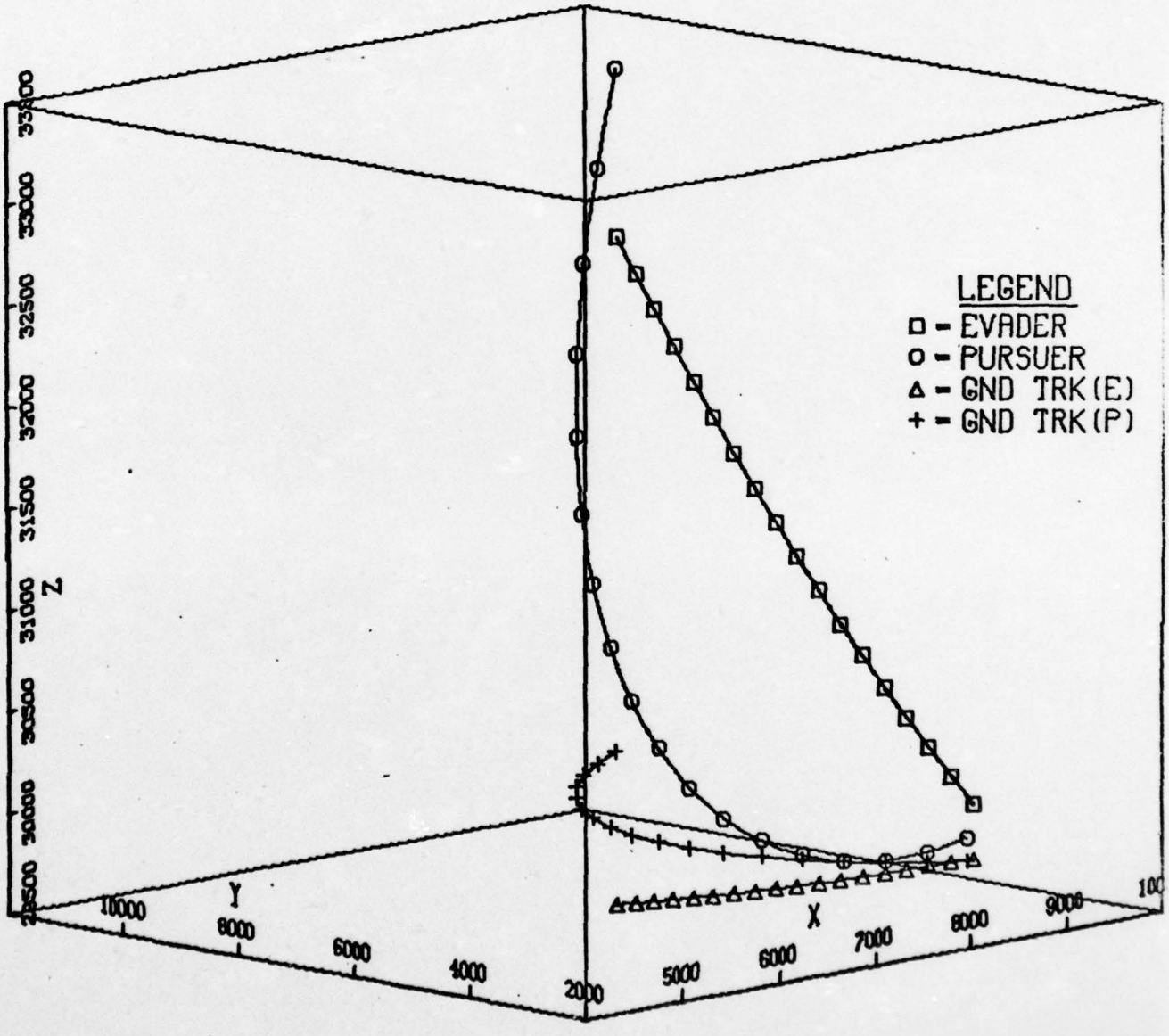
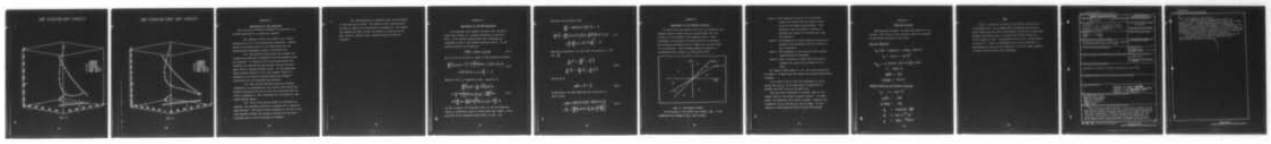


Fig. 28

AD-A034 896 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 12/1
APPLICATION OF DIFFERENTIAL)YNAMIC PROGRAMMING TO AN AIR-TO-AI--ETC(U)
DEC 76 A H FERRARIS

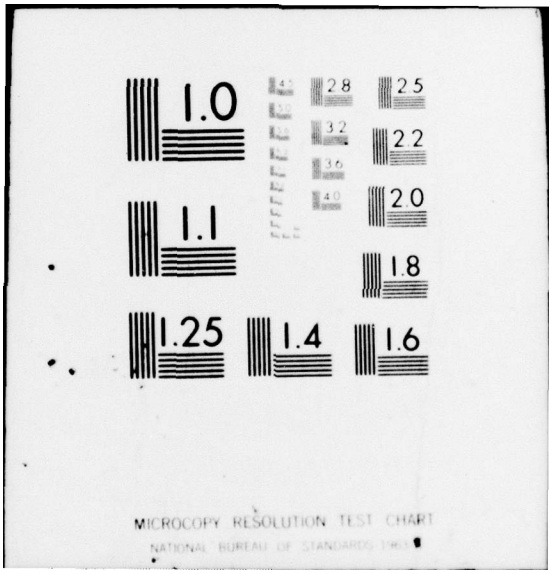
UNCLASSIFIED GA/MC/76D-7 NL

2 OF 2
AD A034896



END

DATE FILMED
2-77


1.0
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
1.1
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963

DDP EVASION-DDP PURSUIT

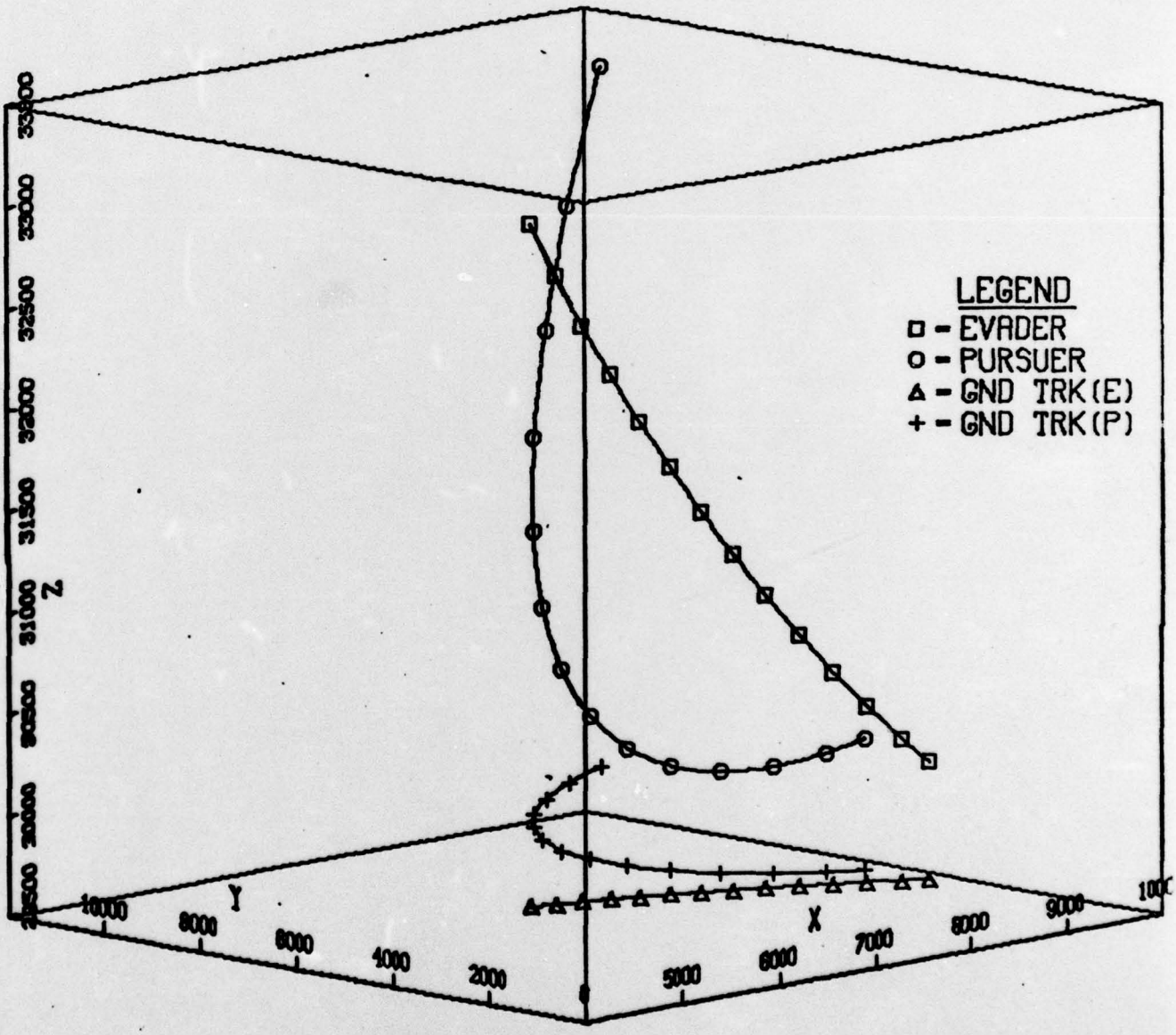


Fig. 29

DDP EVASION-PROP NAV PURSUIT

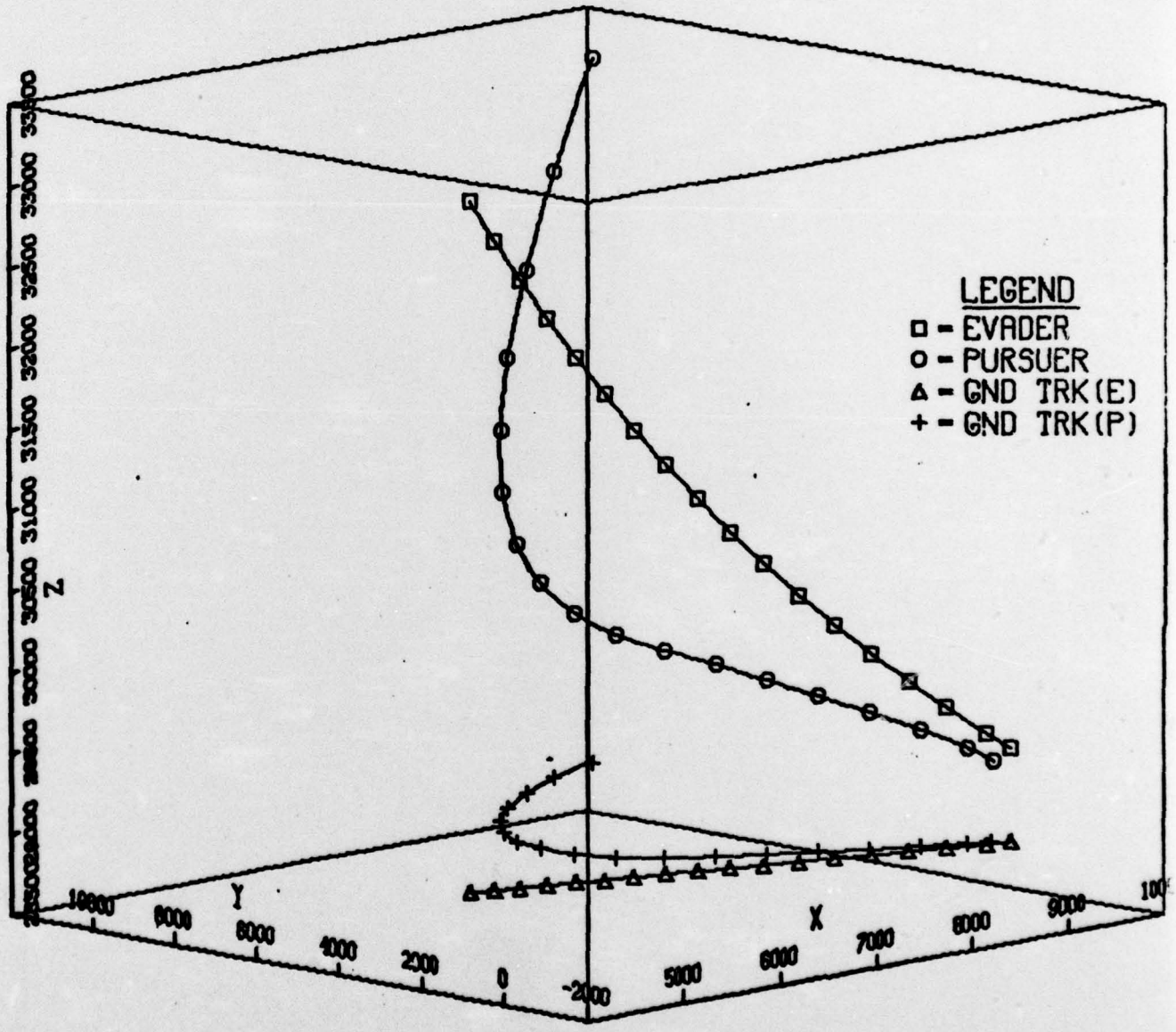


Fig. 30

Appendix B

Application of the Algorithm

The following discussion explains the application of the DDP algorithm in a closed-loop manner:

(a) Nominal controls are used in the forward integration of Eq (2-1) as previously indicated. Penalty values are selected prior to the integration and the mechanization discussed on page 11 is carried out.

(b) The predicted and actual cost changes are compared as explained in Appendix D and the penalty values are adjusted. If regions A or A' are encountered, the iteration is rejected and the nominal controls are used again with an adjusted penalty value. If the comparison falls outside regions A or A', the nominal control is replaced by the control obtained during the backward integration and the penalties are adjusted.

(c) The controls derived during the backward integration are exchanged for the controls used during the forward integration as long as the cost ratios satisfy the rules of Appendix D. The iterations are continued for some predetermined number.

(d) After the required number of iterations is reached, the resulting controls are applied for a specified time interval. The state of each combatant at the end of the time interval becomes the initial condition for the next iteration cycle, and the process is repeated.

(e) The intercept is completed when the derivative of the cost goes to zero. The state at each updating point is used to obtain the trajectories of Appendix A. The evader and pursuer are able to base the choice of controls on the more current, updated state, thereby giving closed-loop guidance.

Appendix C

Derivation of the DDP Equations

It is assumed that nominal controls, $\bar{u}(t)$ and $\bar{v}(t)$, exist which result in a nominal trajectory to Eq (2-1), $\bar{x}(t)$. It is further assumed that this trajectory is reasonably close to the optimal solution, $x^*(t)$. If the optimal solution is written as

$$x^*(t) = \bar{x}(t) + \Delta x(t) \tag{C-1}$$

Eq (3-1) can be written in terms of the nominal as follows:

$$\frac{\partial J^\circ}{\partial t}(\bar{x}+\Delta x;t) + \min_u \max_v \left[H\left(\bar{x}+\Delta x, u, v, J_x^\circ(\bar{x}+\Delta x); t\right) + \nu^T C(\bar{x}+\Delta x, u, v; t)\right] = 0 \tag{C-2}$$

Equation (C-2) is expanded in Δx, resulting in

$$\frac{\partial J^\circ}{\partial t}(\bar{x};t) + \frac{\partial}{\partial t} J_x^{\circ T}(\bar{x};t)\,\Delta x + \min_u \max_v \left[H(\bar{x},u,v,J_x^\circ;t) + \left(\frac{\partial H}{\partial x}\right)^T \Delta x + \nu^T \frac{\partial C}{\partial x}\Delta x + \left[\frac{\partial}{\partial x} J_x^\circ \Delta x\right]^T f(\bar{x},u,v;t) + R'\right] = 0 \tag{C-3}$$

If Δx is small, the remainder term, R', can be neglected since it represents terms of second order and higher. Since Eq (C-3) is not dependent upon choice of Δx, the

following relationships hold:

$$\frac{\partial J^\circ}{\partial t} + H(\bar{x}, u^*, v^*, J_x^\circ; t) = 0$$

$$\frac{\partial}{\partial t} J_x^\circ + \frac{\partial H}{\partial x}(\bar{x}, u^*, v^*, J_x^\circ; t) + \frac{\partial}{\partial x} J_x^\circ f(\bar{x}, u^*, v^*; t) \qquad \text{(C-4)}$$

$$+ \left[\nu^T \frac{\partial C}{\partial x}(\bar{x}, u^*, v^*; t)\right]^T = 0$$

Employing expressions for the total derivatives of J° and J_x°

$$\frac{d}{dt} J^\circ = \frac{\partial J^\circ}{\partial t} + J_x^{\circ T} f$$

$$\frac{d}{dt} J_x^\circ = \frac{\partial}{\partial t} J_x^\circ + \frac{\partial}{\partial x} J_x^\circ f \qquad \text{(C-5)}$$

and Eq (3-4)

$$a(t) = J^\circ - \bar{J} \qquad \text{(C-6)}$$

in Eqs (C-4), the DDP equations are determined as (Ref 8:3-5):

$$-\dot{a}(t) = H(\bar{x}, u^*, v^*; t) - H(\bar{x}, \bar{u}, \bar{v}; t) \qquad \text{(C-7)}$$

$$-\dot{J}_x^\circ = \frac{\partial H}{\partial x}(\bar{x}, u^*, v^*, J_x^\circ; t) + \left[\nu^T \frac{\partial C}{\partial x}\right]^T$$

Appendix D

Adjustment of the Penalty Functions

To derive the most benefit from the CCP technique, the penalty function values must be altered based upon the reaction of the problem to the present values. When convergence is indicated the values should be decreased. The following discussion forms a basis for the alteration of the penalties. The following figure is helpful in determining how to adjust P_e and P_p (Ref 9:30).

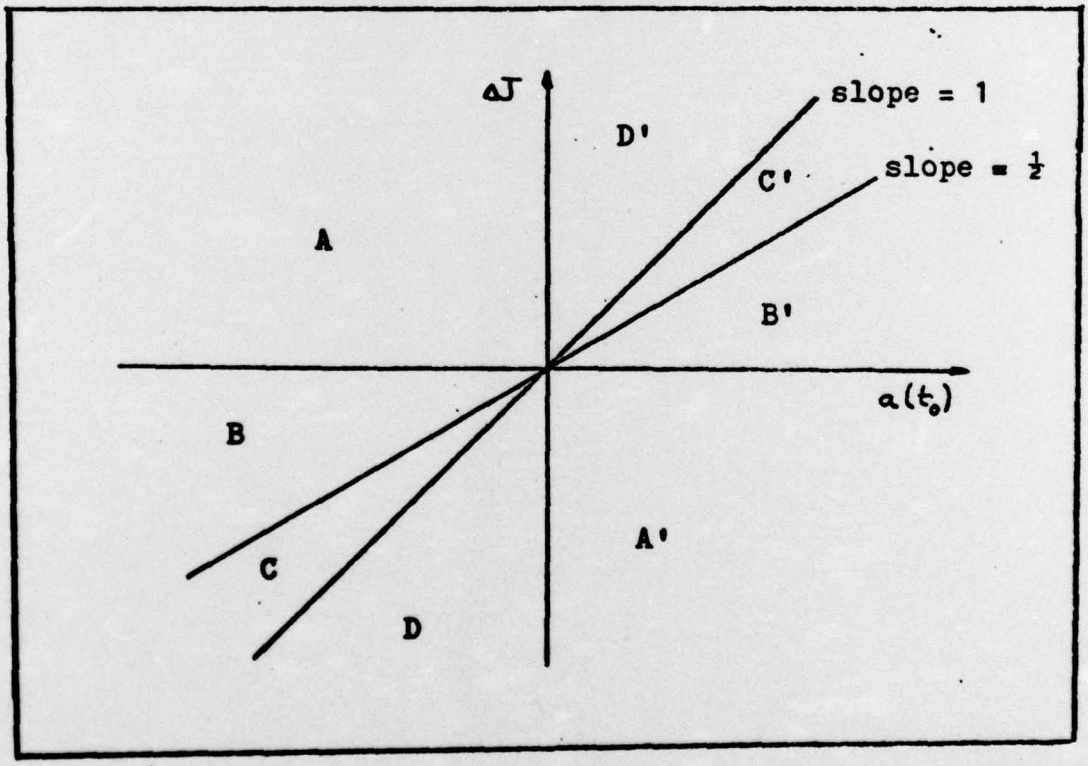


Fig. 31 Convergence Domain

If the ratio of the actual cost change, ΔJ, to the predicted cost change, $a(t_0)$, falls within:

Area A - The expansion of Eq (3-1) is not valid (Δx is too large) and the penalty values should be increased significantly. This indicates that the pursuer's penalty is dominant and should be increased more than the evader's.

Area B - The expansion is poorly satisfied (pursuer's penalty still dominates) and an increase in P_p is indicated.

Area C - The expansion is satisfied and both penalty values should be decreased.

Area D - This is similar to Area B but the evader's penalty, P_e, should be increased in this case.

For points within Areas A' - D', the rules pertaining to Areas A - D apply with the evader and pursuer philosophies reversed.

It is wise to try to keep the components of $a(t_o)$, $a_e(t_o)$ and $a_p(t_o)$, of the same order of magnitude by varying the ratio of P_p to P_e (Ref 8:31).

One area which represents a "special case" is the region close to the origin in Areas A and A'. In these areas, the predicted cost change is small; however, the components, $a_e(t_o)$ and $a_p(t_o)$, may be large. In this situation, both penalties may be reduced moderately.

Appendix E

Numerical Aspects

This Appendix presents the numerical aspects of the problem. The values of the constants used are presented to aid further research in this area.

Aircraft Equations

$$T_{max}\ (lbs) = (22346.7 - .7018\,z + 18.141\,v)$$

$$C_D = .01675 + .223\,C_L^2$$

$$\eta_{AERO} = .16 + (.01422 - .304 \times 10^{-6} z)(v - 250)$$

$$\alpha = .2565\,C_L$$

$$s(ft^2) = 530$$

$$w(lbs) = 40{,}000$$

Missile Equations and General Constants

$$C_D = .9 + .042\,C_L^2$$

$$s(ft^2) = .223$$

$$w\ (lbs) = 103$$

$$\rho_0 = .0023769\ \frac{slug}{ft^2}$$

$$\beta = 3.5 \times 10^{-5}\ ft^{-1}$$

$$g = 32.2\ ft/sec^2$$

Vita

Albert H. Ferraris was born in Morristown, New Jersey on 30 September 1947. He attended North Plainfield High School, North Plainfield, New Jersey and Stevens Institute of Technology, Hoboken, New Jersey prior to entering the United States Air Force Academy in 1966. He graduated with a B. S. in Astronautics in June, 1970. Prior to attending the Air Force Institute of Technology, he served as an instructor pilot and flight examiner at Vance Air Force Base, Enid, Oklahoma.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER GA/MC/76D-7	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) APPLICATION OF DIFFERENTIAL DYNAMIC PROGRAMMING TO AN AIR-TO-AIR MISSILE GUIDANCE PROBLEM MODELED AS A DIFFERENTIAL GAME		5. TYPE OF REPORT & PERIOD COVERED MS THESIS 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) ALBERT H. FERRARIS CAPT USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT*EN) Wright-Patterson AFB, Ohio 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE December 1976 13. NUMBER OF PAGES 96
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

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

18. SUPPLEMENTARY NOTES

Approved for public release;
IAW AFR 190-17
JERRAL F. GUESS, Captain, USAF
Director of Information

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

Differential Games
Dynamic Programming
Air-to-Air Missile
Missile Guidance
Game Theory

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

An intercept problem between an air-to-air missile and an aircraft is modeled as a zero sum, free final time differential game which includes nonlinear dynamics and a payoff related to the kill probability. Previous research has shown that the currently used guidance scheme, proportional navigation, is nonoptimal in this type of problem formulation and a higher kill probability is possible with a guidance law based

NEXT

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

upon a differential game theory.

A differential dynamic programming method is applied to the intercept problem in the search for a real-time solution. A convergence control procedure is introduced in an attempt to enhance the convergence of the typically long-time solution methods. The closed-loop guidance law which results is compared to both proportional navigation and some exact open-loop solutions by means of an off-line simulation on a CDC 6600 computer.

The method does not yield a real-time solution for this problem and does not give improvement over a proportional navigation scheme.