

DIFFERENTIAL GAMES AND MANUAL CONTROL

by Sheldon Baron Electronics Research Center Cambridge, Mass.



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION . WASHINGTON, D. C. . OCTOBER 1966



NASA TN D-3659

.

DIFFERENTIAL GAMES

AND

MANUAL CONTROL

By Sheldon Baron

Electronics Research Center Cambridge, Mass.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information Springfield, Virginia 22151 – Price \$1.00

I.

|

CONTENTS

~

| | Page |
|--|------|
| ABSTRACT | . 1 |
| INTRODUCTION | . 1 |
| WHAT IS A DIFFERENTIAL GAME? | . 1 |
| A SIMPLE PURSUIT-EVASION DIFFERENTIAL GAME | . 2 |
| SIMULATION | . 6 |
| Analog Mechanization | . 6 |
| Instrument Display | . 7 |
| Controller | . 8 |
| RESULTS AND DISCUSSION | . 9 |
| CONCLUSION | . 10 |
| REFERENCES | . 11 |

•

DIFFERENTIAL GAMES AND MANUAL CONTROL

By Sheldon Baron Electronics Research Center

ABSTRACT

Variational methods are used to solve a particular pursuitevasion differential game. The problem involves the determination of optimal strategies for both the pursuer and evader. The performance measure is the miss distance at some fixed terminal time. Both pursuer and evader have limited control energy. The performance of a trained research pilot, for both single- and two-axis control tasks, is compared with that of the optimal pursuer. State vector display and "quickened" display are discussed. The results suggest that differential game problems could be quite useful in the study of manual control.

INTRODUCTION

The theory of differential games was initiated by Isaacs in 1954 (Ref. 1). It was later studied in greater detail by Fleming and Berkowitz (Refs. 2,3). Recently, Ho, Bryson, and the author applied variational techniques to solve a class of differential games (Ref. 4). In an effort to demonstrate the results of Ref. 4, a simulation of a simple pursuit-evasion differential game was conducted. As a matter of some interest, it was decided to compare the performance of a human pilot with that of an optimal pursuer. The results and some implications of this comparison are the subject of this paper. It should be emphasized that these results, from a manual control standpoint, are not extensive since the primary purpose of the research was the study of a class of differential games; nevertheless, they do suggest that differential game problems could be useful in the study of manual control.

WHAT IS A DIFFERENTIAL GAME?

A differential game problem may be stated briefly, and roughly, as follows (a rigorous, precise formulation may be found in Ref. 3):

Given the payoff

$$J(\underline{u},\underline{v}) = \phi(\underline{x}(T),T) + \int_{0}^{T} L(\underline{x},\underline{u},\underline{v},t)dt$$
(1)

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, \mathbf{v}, t) ; \mathbf{x}(t_0) = \mathbf{x}_0$$
 (2)

determine the pair of feedback control laws

$$\underline{u}^{O} = \underline{k} (\underline{x}(t), t) \quad \varepsilon \quad U \tag{4}$$

$$\mathbf{y}^{O} = \overline{\mathbf{k}} (\mathbf{x}(t)), t) \in \mathbf{V}$$
(5)

satisfying the relation

$$J(\underline{u}^{O},\underline{v}) \leq J(\underline{u}^{O},\underline{v}^{O}) \leq J(\underline{u},\underline{v}^{O})$$
(6)

for arbitrary $u \in U$, $v \in V$.

In the parlance of game theory, J is called the payoff, x the (vector) state of the game, and u and y are called (vector) strategies and are restricted to certain sets of admissible strategies, U and V, which depend, in general, on the specific problem to be solved. If strategies u° and v° can be found such that Eq. (6) is true, then they are called optimal pure strategies, and the pair (u°, v°) is called a saddle-point of J. The payoff evaluated at the saddle-point $J(u^{\circ}, v^{\circ})$ is called the Value of the game.

The class of problems in which the differential equations [Eqs. (2)] are linear and the payoff [Eq. (1)] is quadratic was solved using variational methods in Ref. 4. A special case is discussed in the next section.

A SIMPLE PURSUIT-EVASION DIFFERENTIAL GAME

The kinematic equations of motion for an interceptor and target in space may be written as

$$\ddot{r}_{p} = f_{p} + a_{p} ; r_{p}(t_{o}) = r_{p_{o}} , \dot{r}_{p}(t_{o}) = \dot{r}_{p_{o}}$$

$$\vdots \\ \ddot{r}_{e} = f_{e} + a_{e} ; r_{e}(t_{o}) = r_{e_{o}} , \dot{r}_{e}(t_{o}) = \dot{r}_{e_{o}}$$
(7)

where r represents the position vector of the body in threedimensional space, f is the external force per unit mass, a is the control acceleration of the body and the subscripts "p" and "e" refer to pursuer and evader, respectively. If it is assumed

that the altitude difference between the two bodies is small and we consider their relative position, the effect of external forces may be neglected and we obtain

$$\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_{p} - \ddot{\mathbf{r}}_{e} = \underline{\mathbf{a}}_{p} - \underline{\mathbf{a}}_{e} .$$
(8)

We take as a payoff for this game

SHELL F

$$J = \frac{a^2}{2} (r(T) \cdot r(T))$$
(9)

i.e., a measure of the miss distance at some fixed terminal time T. The objective of the pursuer is to minimize the miss distance while the evader attempts to maximize it. The controls of both pursuer and evader are assumed to be constrained by the following relations:

$$\int_{t_{o}}^{T} a_{p} \cdot a_{p} dt \leq E_{p}^{2}(t_{o})$$

$$\int_{t_{o}}^{T} a_{e} \cdot a_{e} dt \leq E_{e}^{2}(t_{o}) .$$
(10)
(11)

Equations (10) and (11) may be thought of as constraints on the control energy available to the two players. It is intuitively clear, and readily proven, that under optimal play the evader will use all his energy. Similarly, it can be shown that if the pursuer has less energy than is required for capture,* or just enough energy to capture, then he, too, will use all his energy in an optimal play of the game. We shall only consider such cases so that equations (10) and (11) may be replaced by the corresponding equality constraints:

$$\int_{t_0}^{T} a_p \cdot a_p dt = E_p^2(t_0)$$

$$\int_{t_0}^{T} a_e \cdot a_e dt = E_e^2(t_0) .$$
(12)
(13)

Then the differential game problem is to determine a saddle-point (a_p^O, a_e^O) of (9) subject to the constraints (8), (12), and (13).

*Capture is defined here as r(T) = 0.

This problem may be solved by considering the payoff

$$J = \frac{a^2}{2}(\underline{r}(T) \cdot \underline{r}(T)) + \int_{t_0}^{T} \left\{ \frac{\underline{a_p}(t) \cdot \underline{a_p}(t)}{2c_p} - \frac{\underline{a_e}(t) \cdot \underline{a_e}(t)}{2c_e} \right\} dt \quad (14)$$

where c_p^{-1} and c_e^{-1} are Lagrange multipliers to be determined such that (12) and (13) are satisfied. The results of Ref. 4 may now be applied and, upon evaluating c_p and c_e , one finds that the optimal controls and the minimax miss distance are:*

$$a_{p(e)}(t) = - \frac{E_{p(e)}(t_{o})(T-t)\hat{r}(t_{o})}{(T-t_{o})^{3}} ||\hat{r}(t_{o})||$$
(15)

$$||\hat{\mathbf{r}}(\mathbf{T})|| = ||\hat{\mathbf{r}}(\mathbf{t}_{o})|| - (\mathbf{E}_{p}(\mathbf{t}_{o}) - \mathbf{E}_{e}(\mathbf{t}_{o})) \sqrt{\frac{(\mathbf{T}-\mathbf{t}_{o})^{3}}{3}}$$
 (16)

where $\hat{r}(t)$ is defined by

$$\hat{\mathbf{r}}(t) \stackrel{\Delta}{=} \mathbf{r}(t) + (\mathbf{T}-t) \dot{\mathbf{r}}(t)$$
(17)

Optimal Strategies (feedback control laws) may be obtained from Eq. (15) by letting $t_0 = t$. The result is

$$a_{p(e)}(t) = -\frac{E_{p(e)}(t)}{\sqrt{\frac{T-t}{3}}} \frac{\hat{r}(t)}{||\hat{r}(t)||}$$
 (15')

with a corresponding minimax miss distance

$$||\mathbf{r}(\mathbf{T})|| = ||\mathbf{r}(\mathbf{t})|| - (\mathbf{E}_{p}(\mathbf{t}) - \mathbf{E}_{e}(\mathbf{t})) \sqrt{\frac{(\mathbf{T}-\mathbf{t})^{3}}{3}}$$
 (16')

*Norm notation is used to denote the length of a vector, i.e., $||y|| = (y \cdot y)^{1/2}$

E (t) and E (t) are just the energies remaining at time t and are calculated from

$$E_{p(e)}^{2}(t) = E_{p(e)}^{2}(t_{o}) - \int_{t_{o}}^{t} a_{p(e)}(t) \cdot a_{p(e)}(t) dt.$$
(18)

The vector $\mathbf{r}(t)$ will be called the predicted miss. Given a relative position and velocity, $\mathbf{r}(t)$ and $\mathbf{\dot{r}}(t)$, at time t, then $\hat{\mathbf{r}}(t)$ is the relative position which would be obtained at time $\mathbf{\tilde{r}}(t)$ is the relative position which would be obtained at time $\mathbf{\tilde{T}}$ if no further control were applied by either pursuer or evader. The quantity in the brackets in Eq. (15') is just a unit vector in the direction of the predicted miss. Hence, the minimax controls are applied in the opposite direction of the predicted miss and have magnitude depending only on the energy remaining and the time to go. (Note that as a result of the minus sign in Equation (8), the evader's control is actually a positive feedback in the system, as one would expect.) From Eq. (16') we see that the minimum pursuit energy required for capture, under optimal play, is:

$$E_{p}(t) = \frac{\left| \left| \hat{r}(t) \right| \right|}{\sqrt{\left(\frac{T-t}{3}\right)^{3}}} + E_{e}(t) .$$
 (19)

An interesting special case of the above results is the following: Let the pursuer and the target be on a nominal collision course with range R and closing velocity $V_C = R/(T-t)$. Let r_X represent the lateral deviation from the collision course (Figure 1) and let the pursuer have just enough energy to capture at time T. Then, for small lateral deviations $r_X = R\sigma$ and the optimal pursuit strategy according to Eq. (15') is:

$$a_{p_{\mathbf{X}}}(t) = -\frac{E_{p}(t)}{\sqrt{(\frac{T-t}{3})}} \frac{\nabla_{c} \sigma (T-t)^{2}}{||\tilde{r}(t)||} .$$
(20)

Substituting for $||\hat{r}(t)||$ from Eq. (19) yields:

$$a_{p_{x}}(t) = -\frac{3V_{c}\sigma}{1 - E_{e}(t)/E_{p}(t)}$$
 (21)

which is simply proportional navigation with an effective navigation constant which depends on the energies of both players.

Note that although the above problem has been interpreted as

a pursuit-evasion game, one could also interpret it, simply, as a problem of controlling, in a specified manner, several externally disturbed double integrator plants. The disturbance in this case is not a random disturbance, but rather, the worst possible disturbance in a class of admissible disturbances. In this context it should be pointed out that, by a relatively straightforward extension of the results and techniques of Ref. 4, one can obtain a solution to the problem with a payoff:

$$J = \frac{a^{2}}{2} (\underline{r}(T) \cdot \underline{r}(T)) + \int_{t_{o}}^{T} \underbrace{\frac{\underline{r}(t) \cdot \underline{r}(t)}{2}}_{t_{o}} + \frac{\underline{a}_{p}(t) \cdot \underline{a}_{p}(t)}{2c_{p}} - \frac{\underline{a}_{e}(t) \cdot \underline{a}_{e}(t)}{2c_{e}} \overset{a}{\overset{(22)}{\overset{$$

However, in the present investigation, the problem with payoff given by Eq. (22) is not considered, since this payoff did not seem to be consistent with the pursuit-evasion interpretation.

SIMULATION

Analog Mechanization

Two particular cases of the above problem were simulated on an analog computer. Both cases involved planar motion. However, in the first case, two-axis control was required, whereas, in the second case, which corresponds to the proportional navigation situation described above, only single-axis control was necessary. The values for the initial conditions were selected for convenience and do not necessarily correspond to any realistic situation. The initial values, along with the minimax miss distance for each case, are given in Table I.

An interesting development occurred in attempting to mechanize the optimal solution on the analog computer. In the first attempt at accomplishing this task, Eqs. (8), (15'), and (18) were mechanized directly. The results of this mechanization deviated considerably from the analytically obtained optimal solution. The difficulty, from an analog mechanization standpoint, is apparent upon examination of Table I. If the y-component of the miss distance is considered, one sees that the analog computer will encounter resolution difficulties; when the computer is scaled to accommodate the initial miss distance of 2832 feet, the terminal miss distance, 2.4 feet, is represented by a voltage in the noise range of the computer. The effects of the limited resolution in this formulation were most pronounced in the terminal phase of the solution. While feedback control might normally be expected to reduce these errors, it must be remembered that the evader's

control introduces a positive feedback loop. It was, in fact, demonstrated by generating the open-loop optimal trajectories that the feedback loop aggravated the problem.

Further examination of Eqs. (8) and (15'), along with Table I, indicates the means for overcoming the resolution difficulties. The procedure is to formulate the problem directly in terms of the predicted miss,* which, for the y-component, has a much smaller dynamic range. The differential equation for the predicted miss is simply:

 $\dot{\tilde{r}} = (a_{p} - a_{e}) (T-t)$ (23)

With this formulation the instantaneous position and velocity are calculated as open-loop outputs for display purposes only; the actual problem solution involves variables which present no resolution difficulties. The results of the analog mechanization for this formulation were in excellent agreement with the analytic results.

Instrument Display

The prime consideration in designing the display for this study was that the pilot must be presented all the information necessary to generate the optimal pursuit strategy. Secondary considerations were that the display should be easy to read and reasonably realistic. The resulting display is shown in the photograph presented as Figure 2.

The scope at the top of the panel presented the relative position of the evader (the pursuer is located at the origin). A scale change was programmed to improve resolution when the pursuer closed to within 50 feet in the y-direction and/or 20 feet in the x-direction. A light situated below the scope indicated the appropriate scale.

The vertical instrument at the center of the panel displays predicted y-miss and instantaneous closing (relative) velocity (note that the instantaneous closing velocity and the predicted closing velocity are identical for this problem). The horizontal instruments present the same information for the x-components. The circular meters on the right and left of the panel provide, respectively, pursuit energy remaining and time-to-go. The circular meter at the bottom of the panel indicates evasive energy

^{*}It is interesting to note that parallel theoretical work associated with Ref. 4 led to the conclusion that the results for the general problem were most simply and meaningfully stated in terms of the predicted miss.

remaining. The display of evasive energy was included in accordance with the ground rule stated above, viz., that all information required for generation of the optimal pursuit strategy would be displayed. In the derivation of the optimal pursuit strategy, it was assumed that the evader's energy was known; thus, although the final expression for the pursuit strategy [Eq.(15')] does not depend explicitly on E_e , it was decided to display this information.

If the information content of the display is examined, it is seen that the display may be interpreted as a state vector display and/or a "quickened" display. In the usual fashion, the quantities r_x , r_y , \dot{r}_x , and \dot{r}_y may be considered the components of the "state" of the system and their display constitutes a "state vector display"; displaying the predicted miss, which is a linear combination of the state vector and constitutes a signal proportional to the desired control, corresponds to the so-called "quickened" display (Ref. 5). However, the distinction between the two types of display, at least for this problem, seems somewhat arbitrary. As was seen in the above discussion, and in Ref. 4, the predicted miss may be taken as the state vector of the system and, then, the distinction between state vector display and "quickened" display vanishes. In this regard, it is important to note that one can, and indeed should, include the energies remaining, E_p^2 (t) and E_e^2 (t), and the time-to-go, (T-t), in the state vector of the system. Hence, a display of acceleration command [i.e., Eq. (15')] could also be considered a "quickened" display.

The above discussion is indicative of a more general point concerning "state vector" displays. Since the state representation of a system is, in general, not unique, there often exists considerable freedom in choosing a set of state variables. Different selections will have different implications in terms of a state vector display and the proper choice of state variables could easily make the difference between a good and a poor display.

Controller

The pilot's acceleration inputs were introduced through a grip-type, two-axis side controller located at the end of the pilot's right arm rest. (Actually, the controller could have been used for three-axis inputs if such had been required.) Acceleration inputs in the x-direction are actuated by rotating the grip laterally about pivot axis located slightly below the grip; y-acceleration inputs are actuated by motions of the hand about a pivot axis passing through the wrist.

The controller had physical stops which imposed an amplitude

constraint on the pilot's control inputs. However, no amplitude constraint was imposed in deriving the optimal pursuit strategy (to impose such a constraint complicates the problem considerably). To avoid the difficulties associated with the amplitude constraint, the full-scale deflection of the controller was initially scaled to correspond to twice the maximum acceleration ever used in the optimal pursuit strategy. After some preliminary runs the scaling was changed so that full deflection yielded an acceleration which was equal to the maximum optimal acceleration. The reason for this change will be discussed below.

RESULTS AND DISCUSSION

A NASA research pilot served as a subject for the demonstrations of the pursuit-evasion game. His task was to minimize the miss-distance at the terminal time, subject to the energy constraint, i.e., he replaced the optimal pursuer. The evader's control remained an optimal evasive strategy. Some typical piloted runs for Case 1 are plotted in the $\hat{r}_{,-}\hat{r}_{,-}$ plane in Figure 3. The pilot's best performance was characterized by a miss-distance of approximately 16 feet for Case 1 and a miss-distance of approximately 15 feet for Case 2. It should be noted that the pilot made about 50 runs during the course of one afternoon. The majority of these runs were for Case 1. The following presents and discusses some of the more interesting qualitative results obtained from the study.

When the scope was used as a primary position informa-1. tion source for the pilot, his performance was quite poor. After learning to interpret the predicted miss quantities, the pilot improved his performance considerably. Part of this improvement may be due to reduced scan requirements. However, it seems clear that the major sources of improvement are improved resolution and the fact that the predicted miss represents information more pertinent to the required task than does the instantaneous relative position. It is interesting to note that the difficulties associated with the analog mechanization were indicative of the problems the pilot would encounter in trying to use instantaneous relative position information. Of more importance is the fact that the vector $(\hat{r}_x, \hat{r}_y, E_p^2(t), E_e^2(t), (T-t))$ is the minimal state representation of the system (if we insist on including the last three components as state variables), and therefore, excluding possible integrated displays, the minimal state vector display appears to be the best state vector display for this problem. The question of whether this is true in general seems worthy of further investigation.

2. The pilot tended to ignore the display of the evader's

energy. This may have been due to the fact that he was not used to having this information available. Of course, as seen from Eq. (15'), the pilot did not need the evader's energy to construct the optimal control. 3. The pilot's performance improved markedly after he was allowed to observe the optimal trajectory several times. He attributed this improvement to what he called a "pinball fix." In essence, this amounted to duplicating a correlation he noted between the energy remaining and the time-to-go for the optimal trajectory. This point is interesting from the standpoint of understanding, and possibly modeling, the pilot's learning process.

4. In the preliminary runs the pilot generally started by initially commanding zero acceleration. This placed him at a disadvantage since the initial acceleration for optimal pursuit is, in fact, the maximum commanded acceleration. The pilot was informed that optimal pursuit required commanding an initial acceleration. However, so long as he was required to judge the initial acceleration required, his performance did not substantially improve. In order to minimize the effects of the initial conditions, it was decided to scale full-scale deflections of the controller to correspond to the maximum commanded optimal accelerations. The pilot then started his pursuit with the controller against the stops and his performance improved considerably.

5. The idea of "doing battle" with an intelligent adversary seemed to provide excellent motivation for the pilot. In fact, the "game" nature of the problem resulted in a large number of untrained "volunteers" for the experiment.

6. Several runs were tried with "unskilled" subjects. Their performance was, as could be expected, quite inferior to that of the pilot. As an additional cue, these subjects were displayed the optimal trajectory in $\hat{r}_x - \hat{r}_y$ coordinates and given the task of "tracking" this trajectory. It soon became apparent that a timing reference was needed for such a display to be effective. However, this approach was not pursued further.

CONCLUSION

A simple pursuit-evasion differential game has been solved by variational methods. The results of a limited investigation comparing a pilot's performance with that of an optimal pursuer indicate that differential game problems could be useful in the study of manual control. Since optimal control problems are simply one-player differential games, it is apparent that the "game" will provide at least as much information, from a manual control standpoint, as will a similar optimal control problem; the "game" has the additional advantages of providing excellent motivation for the subject pilot and allowing the study of performance subject to worst-case disturbances. In fairness, it should be noted that differential game problems will, in general, be more complicated theoretically than their optimal control counterparts.

Electronics Research Center National Aeronautics and Space Administration Cambridge, Massachusetts, August 19, 1966 125-19-01-01-25

REFERENCES

- 1. Isaacs, R.: Differential Games I, II, III, IV. RAND Corporation Research Memorandum RM-1391, 1399, 1411, 1468, 1954-56.
- Berkowitz, L.D., and Fleming, W.H.: On Differential Games with Integral Payoff. Annals of Math. Study No. 39, pp. 413-435, Princeton University, 1957.
- Berkowitz, L.D.: A Variational Approach to Differential Games, Advances in Game Theory. Annals of Math. Study No. 52, pp. 127-173, Princeton University Press, 1964.
- 4. Ho, Y.C., Bryson, A.E., and Baron, S.: Differential Games and Optimal Pursuit-Evasion Strategies. IEEE Trans. on Automatic Control, vol. AC-10, No. 4, October 1965.
- 5. Obermayer, R.W., and Muckler, F.A.: Modern Control System Theory and Human Control Functions. NASA CR-256, July 1965.

TABLE I

INITIAL CONDITIONS AND MINIMAX MISS DISTANCES FOR CASES 1 AND 2

| Case | r _x (t _o) (ft) | | | | $\frac{\text{E}_{p}^{2}(\text{t}_{o})}{(\text{ft}^{2}/\text{sec}^{3})}$ | | | $\hat{r}_{x}(t_{o})$ (ft) | r _y (t _o) (ft) | r _x (T) (ft) | - | r(T) (ft) |
|------|--|------|---|------|---|---|----|------------------------------|--|----------------------------|-----|----------------|
| 1 | 99 | 2832 | 0 | -100 | 25 | 9 | 27 | 99 | 132 | 1.8 | 2.4 | 3 |
| 2 | 162 | 2700 | 0 | -100 | 25 | 9 | 27 | 162 | 0 | 0 | 0 | 0 |

-

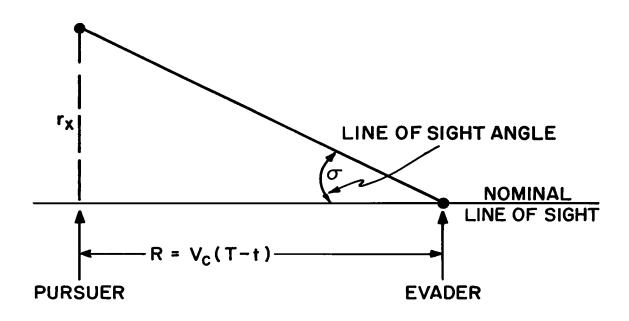
-

7

٠

1

.



|

Figure 1.-Geometry of proportional navigation

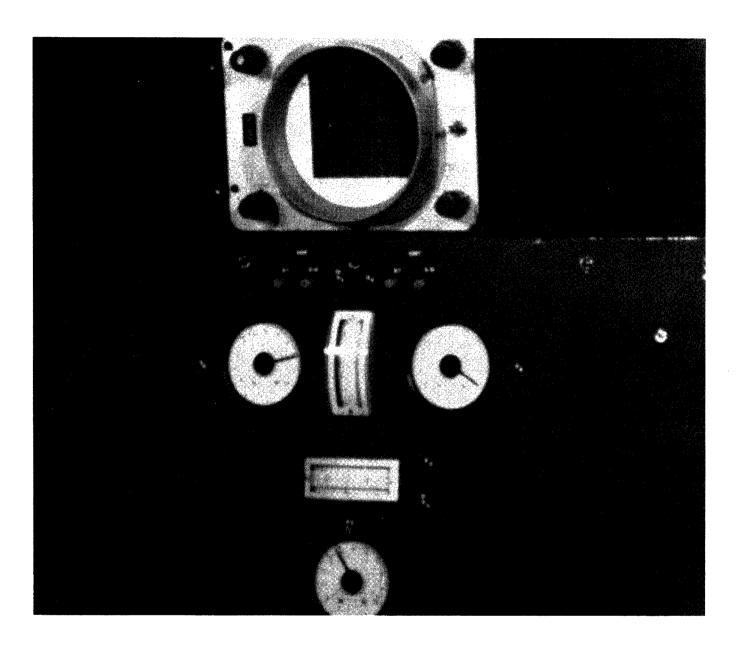


Figure 2.-Instrument display panel

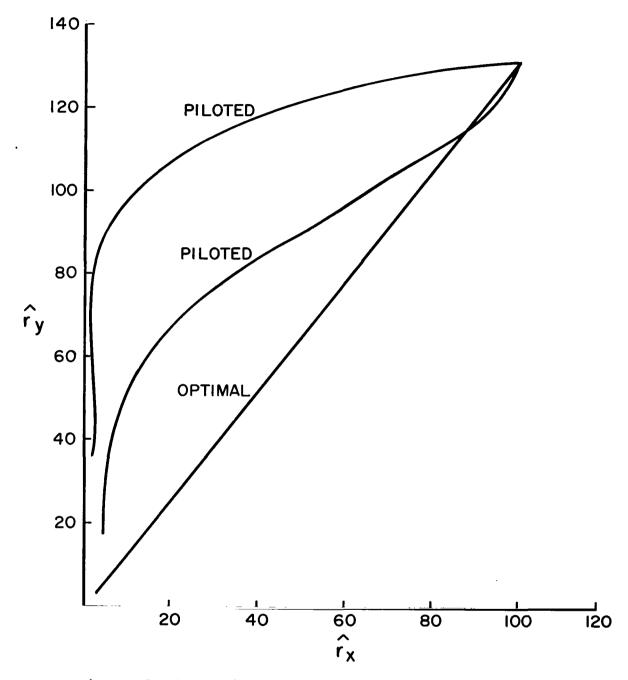


Figure 3.-Comparison of typical piloted trajectories with optimal trajectory for case 1

NASA-Langley, 1966

I.

Į

"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

14

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: Technical information generated in connection with a NASA contract or grant and released under NASA auspices.

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

TECHNICAL REPRINTS: Information derived from NASA activities and initially published in the form of journal articles.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities but not necessarily reporting the results of individual NASA-programmed scientific efforts. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

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

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546