AD/A-003 387

PROCEEDINGS OF THE SYMPOSIUM ON AIR FORCE APPLICATIONS OF MODERN CONTROL THEORY HELD AT AIR FORCE INSTITUTE OF TECHNOLOGY, WRIGHT-PATTERSON AIR FORCE BASE, OHIO ON 9-11 JULY 1974

AEROSPACE RESEARCH LABORATORIES

PREPARED FOR
FRANK J. SEILER RESEARCH LABORATORY

JULY 1974

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Organizations or individuals receiving reports via Aerospace Research Laboratories automatic mailing lists should refer to the ARL number of the report received when corresponding about change of address or cancellation. Such changes should be directed to the specific laboratory orginating the report. Do not return this copy; retain or destroy.

Reports are not stocked by the Aerospace Research Laboratories. Copies may be obtained from:

National Technical Information Services
Clearinghouse
Springfield, VA 22151

This report has been reviewed and cleared for open publication and public release by the appropriate Office of Information in accordance with AFR 190-12 and DODD 5230.0. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service.

AIR FORCE/56780/15 November 1974 -- 350

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A003387

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARL 74-1060	2 GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PROCEEDINGS OF THE SYMPOSIUM ON AIR FORCE APPLICATIONS OF MODERN CONTROL THEORY		5. TYPE OF REPORT & PERIOD COVERED Symposium Proceedings 9-11 July 1974 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David R. Audley, 1Lt Kenneth D. Herring, Major		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Frank J. Seiler Research Laboratories (NH) and Aerospace Research Laboratories (LB)		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS DoD Element 61102F 70710406
11. CONTROLLING OFFICE NAME AND ADDRESS Aerospace Research Laboratories Building 450 - Area B Wright-Patterson AFB, Ohio 45433		12. REPORT DATE 13. NUMBER OF PAGES 502
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)

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

Approved for public release; distribution unlimited.

18. SUPPLEMENTARY NOTES

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

Control Theory	Identification	Navigation	Fire Control
System Theory	Modeling	Guidance	Laser Weapons
Estimation	Stochastic Control	Weapons Delivery	Conventional Weapons
Filtering	Attitude Determination	Testing	Air-to-Air Missiles
Optimal Control	Flight Control	Inertal Navigation	Human Operator

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

The purpose of this symposium was to bring together DoD personnel involved in the application of modern control theory to Air Force Systems. "Modern Control Theory" was defined by the sponsors to include at least the following technical areas: estimation theory, linear and nonlinear system theory, control theory (including classical design and optimal techniques), modeling, identification, time series analysis, and the numerical methods in these areas. Through the formal paper presentations and informal exchange of ideas, the sponsors' objectives were i) to exhibit the current state-of-the-art in the application of modern control

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

theory to Air Force systems and to review some successful case histories. ii) to indicate some current limitations of the theory which prevent its effective application in certain situations. This is of special interest in helping to define areas of needed research. iii) to solicit new areas of possible application through the active participation of diverse Air Force organizations. In addition to two invited lectures, there were thirty-three contributed papers arranged into five sessions. These papers represent a broad spectrum of Air Force interests, but all are commonly bound to modern control theory. The symposium was attended by over one hundred thirty scientists and engineers. Most of the attendees were Air Force personnel, however, several were from other DoD activities. A list of attendees is included in this volume The principals of the organization of this conference were Major kenneth D. Herring (FJSRL), Symposium Manager and 1Lt David R. Audley, (ARL), Technical Program Manager. They were strongly aided by the following: Lt Col Edward J. Bauman (USAFA), Capt Richard M. Potter (AFIT), Capt Kenneth A. Myers (ARL), Mr. Richard M. Reeves AFAL), Capt Robert B. Asher (FJSRL), and Capt. Randall V. Gressand (AFFDL).

19. Key Words (continued)

Orbit Determination
Pointing and Tracking
Satellite Control
Numerical Analysis
Solar Electric Propulsion
Terrain Following
Calibration

PROCEEDINGS OF THE SYMPOSIUM ON AIR FORCE APPLICATIONS OF MODERN CONTROL THEORY

COSPONSORED BY:

FRANK J. SEILER RESEARCH LABORATORY
AEROSPACE RESEARCH LABORATORIES

HELD AT:

AIR FORCE INSTITUTE OF TECHNOLOGY
WRIGHT–PATTERSON AIR FORCE BASE, OHIO

9–11 JULY 1974

Approved for public release; distribution unlimited.

AEROSPACE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
WRIGHT–PATTERSON AIR FORCE BASE, OHIO 45433

INTRODUCTION

The purpose of this symposium was to bring together DoD personnel involved in the application of modern control theory to Air Force Systems. "Modern Control Theory" was defined by the sponsors to include at least the following technical areas: estimation theory, linear and nonlinear system theory, control theory (including classical design and optimal techniques), modeling, identification, time series analysis, and the numerical methods in these areas.

Through the formal paper presentations and informal exchange of ideas, the sponsors' objectives were

i) to exhibit the current state-of-the-art in the application of modern control theory to Air Force systems and to review some successful case histories.

ii) to indicate some current limitations of the theory which prevent its effective application in certain situations. This is of special interest in helping to define areas of needed research.

iii) to solicit new areas of possible application through the active participation of diverse Air Force organizations.

In addition to two invited lectures, there were thirty-three contributed papers arranged into five sessions. These papers represent a broad spectrum of Air Force interests, but all are commonly bound to modern control theory.

The symposium was attended by over one hundred thirty scientists and engineers. Most of the attendees were Air Force personnel, however, several were from other DoD activities. A list of attendees is included in this volume.

The principals of the organization of this conference were

Major Kenneth D. Herring (FJSRL), Symposium Manager

1Lt David R. Audley (ARL), Technical Program Manager

They were strongly aided by the following:

Lt Col Edward J. Bauman (USAFA)

Capt Richard M. Potter (AFIT)

Capt Kenneth A. Myers (ARL)

Mr. Richard M. Reeves (AFAL)

Capt Robert B. Asher (FJSRL)

Capt Randall V. Gressand (AFFDL)

TABLE OF CONTENTS

TECHNICAL PROGRAM

Symposium on Air Force Applications of Modern Control Theory

9-11 July 1974

Sponsored by the Aerospace Research Laboratories (ARL) and
the Frank J. Seiler Research Laboratory (FJSRL) at
Wright-Patterson AFB, Ohio

All presentations will be made in the auditorium of
the AFIT Engineering School (Bldg 640)

---

9 July 1974 - Tuesday

0900 hrs — Opening Remarks and Introduction of Keynote Speaker
Major Kenneth D. Herring, FJSRL, Symposium Manager

0910 hrs — Keynote Address
Colonel Bernard S. Morgan
Commander, AF Cambridge Research Laboratory

Session on Fire Control Systems
Chairman: Lt Col Edward J. Bauman, USAF Academy (USAFA)

1000 hrs — A Ballistic Trajectory Algorithm for Digital Fire Control
Arthur A. Duke, Naval Weapons Center (NWC)

1030 hrs — Non-Director Gunsight Response
Ralph S. Bryan, A.F. Avionics Laboratory (AFAL)

1100 hrs — Digital Control of a Trainable Gun
Lt Col Edward J. Bauman and Capt Jack C. Henry, USAFA

1130 hrs — Kalman Filtering for Precision Pointing and Tracking
Capt Robert B. Asher, FJSRL and Lt David H. Watjen, AFAL

1200 hrs — Lunch

1330 hrs — The Generation of Air-to-Air Missile Launch Envelopes
John Hledik and John C. Shepherd, NWC

1400 hrs — Air-to-Air Missile Launch Boundary Fire Control Techniques
Jon Ooten and Capt Al Chatmon, Aeronautical Systems Division (ASD)

1430 hrs — The Analysis and Simulation of a Fire Control Technique for Launching Air-to-Air Missiles
D. H. de Does, Space and Missile Test Center (SAMTEC) and T. E. Moriarty, Air Force Institute of Technology (AFIT)

Session on Modeling, Identification and Numerical Methods
Chairman: Capt Richard M. Potter, AFIT

1500 hrs — Feasibility Study of Autoregressive-Moving Average (p,q) Models Applied to Natural Phenomena
Capt Thomas S. Lee, Space and Missile Systems Organization (SAMSO)

1530 hrs — In-Flight Pilot Describing Functions Using a Newton-Raphson Parameter Identification Program
Daniel L. Kugel, AF Flight Dynamics Laboratory (AFFDL)

1600 hrs — Stochastic Response of an Idealized Aircraft to Atmospheric Turbulent Gusts
Jon Lee, ARL

1630 hrs — A Method of Constructing Minimal Linear State-Variable Models based on a General Class of System Representations
1Lt David R. Audley, ARL

---

10 July 1974 - Wednesday

0900 hrs — Distinguished Guest Lecture
Colonel Charles A. Scolatti, USAF-Ret.
Commander, AF Flight Dynamics Laboratory (February 1971 - May 1974)

Session on Modeling, Identification and Numerical Methods (Cont.)
Chairman: Capt Richard M. Potter, AFIT

1000 hrs — Fly and Fight: Predicting Piloted Performance in Air-to-Air Combat
Capt Thomas R. Harvey, AF Satellite Control Facility (AFSCF) and Major James D. Dillow, AFIT

1030 hrs — Human Operator Identification Using Least Squares, Maximum Likelihood and Power Spectral Density Techniques
Major John C. Durrett, USAFA and 2Lt Donald R. McMonagle, Air Training Command (ATC)

1100 hrs — Algebraic Representation of Sensitivity Operators for Linear Constant Coefficient Systems
Capt J. Gary Reid, AFAL

1130 hrs — Tuneable Integration for Air Force Applications
Capt Marc L. Sabin, RJSRL

1200 hrs — Lunch

Session on Aerospace Vehicle Control
Chairman: Capt Kenneth A. Myers, ARL

1330 hrs — A Control System Design Method for a System having a Plant with Variable Parameters
C. H. Houpis, AFIT

1400 hrs — Multi-Surface Flight Control Systems Designed using Quadratic Optimal Control
Robert D. Poyneer, AFIT

1430 hrs — Differential Thrust Controllers for Taxiing an Air Cushion Landing System Aircraft
Capt R. V. Gressang, AFFDL

1500 hrs — The Effect of High Thrust on Optimal Fighter Aircraft Turning Maneuvers
Major Gerald M. Anderson, AFIT

1530 hrs — Dynamic Missile Synthesis
1Lt Michael L. Mumford, AF Armament Laboratory (AFATL)

1600 hrs — Precision Attitude Determination for the Defense Meteorlogical Satellite Program (DMSP) Block 5D Satellite
Major Kenneth E. Nidiffer and 1Lt David A. Nichols, SAMSO

1630 hrs — Precision Attitude Control for the DMSP Block 5D Satellite
1Lt Peter Reischl, SAMSO

11 July 1974 - Thrusday

Classified Session of Air Force Applications of Modern Control Theory (U)
Chairman: Richard M. Reeves, AFAL

Notice: All presentations in this session are classified SECRET. Entry will be controlled. Those wishing to attend any part of this session must have evidence (See AFR 205-1, paragraph 7-109) of clearance level on file by 10 July 1974 with 1Lt David R. Audley, ARL/LB, WPAFB, Ohio 45433.

0900 hrs — Application of Estimation Theory of an AF Pointing and Tracking System (U)
1Lt James E. Negro, AF Weapons Laboratory (AFWL)

0930 hrs — Applications of Modern Control Theory to the Integration of Radar/Inertial Navigation Systems for all Weather Navigation and Weapons Delivery (U)
Richard M. Reeves, AFAL

1000 hrs — Improved Guidance Law Evaluation (U)
James E. Gibson, AFATL

1030 hrs — Minimal Error Trajectories in Real Time (U)
Eugene J. Pollock, AF Special Weapons Center (AFSWC)

Session on Filtering Theory
Chairman: Capt Robert B. Asher, FJSRL

1115 hrs — Application of Extended Kalman Filtering to a Calibration of an Inertial Navigation System
Lt Col Frazier J. Hellings, 6585th Test Group

1145 hrs — Filtering Theory-Methods and Applications to the Orbit Determination Problem for Near-Earth Satellites
Capt Kenneth A. Myers, ARL

1215 hrs — Lunch

1400 hrs — Analysis of Mean Square Estimation Error, Variance, and Bias in Reduced Order Filters
Capt Robert B. Asher, FJSRL and Jesse C. Ryles, AFAL

1430 hrs — Sequential Estimation of Random Thrust Errors for Solar Electric Propulsion Spacecraft
Major Thomas J. Eller, USAFA

1500 hrs	System - Level Inertial Measurement Unit (IMU) Calibration Richard A. Kitzerow, AFAL
1530 hrs	Modifications to the Iterated-Extended Kalman Filter That Can Improve Performance Capt Ronald E. Janosko, FJSRL
1600 hrs	Information Processing for an Aircraft Terrain Following System Major James E. Funk, AFIT

A BALLISTIC TRAJECTORY ALGORITHM FOR DIGITAL FIRE CONTROL

by
Arthur A. Duke
Naval Weapons Center
China Lake, California

ABSTRACT. A method of numerical intergration of the equations of unguided air-to-surface weapons is developed. This algorithm, suitable for real-time solution by airborne digital computers, yields accurate trajectory parameters, provides great flexibility in release condition and weapon type, and minimizes computer memory requirements.

The algorithm is extended for applications including helicopter and antiaircraft fire control.

INTRODUCTION

Since the first airborne digital computers were introduced into operational Navy attack aircraft some 15 years ago, a basic computational requirement imposed upon these computers has been the prediction of weapon range and time-of-fall, given sensor-supplied release conditions. Because of the nature of the problem, each aircraft system developed in this interim period (e.g., A-6A, ILAAS, A-7E, etc.) has had concurrently developed with it a "new" set of weapon ballistic equations for this air-to-ground fire control use. While all the "sets" of weapon ballistic equations are developed from the same basic physical considerations, their final formulation is subject only to the ingenuity of the mathematician, and often the equation sets bear no resemblance to each other. Each of these sets of equations has to be "proofed" through extensive developmental flight tests, and each has its own idiosyncrasies.

As the processing speed and capacity of these airborne digital computers has increased, a direct method of solution of the ballistic weapon trajectories has become more and more attractive. This method, numerical integration, has heretofore required excessive time at the processing speeds available to meet the near real-time requirements of airborne weapon delivery systems.

This paper presents a method of numerical integration of ballistic weapon trajectories that is suitable for current operational airborne digital computers in that it is fast, accurate, flexible, and efficient. It is hoped that this algorithm, or modifications thereof, will become the standard method of trajectory computation for future airborne digital weapon delivery systems.

BACKGROUND

BALLISTIC PROJECTILES

Most aircraft unguided air-to-ground weapons can be described as ballistic projectiles. That is, the only forces acting on them after release from the aircraft are gravity and aerodynamic drag. Bullets, streamlined bombs, drogued (retarded) bombs, cluster munitions, and unguided rockets (after burnout) are all ballistic projectiles. Guided weapons and weapons developing lift are not ballistic projectiles.

To successfully release a ballistic weapon from an aircraft so that it impacts at a desired point requires a measurement of--or prediction concerning--the following quantities:

1. Position of the target relative to the aircraft
2. Velocity of the aircraft in the air mass
3. Direction and magnitude of gravity
4. Velocity of the aircraft relative to the ground | (or total relative velocity)
5. Velocity of the target relative to the ground |
6. An *a priori* prediction of the weapon trajectory (for example, horizontal range, time of flight)

given known initial (release) conditions in the air mass. This includes assumptions concerning the structure of the air mass containing the trajectory and of the ballistic (drag) characteristics of the weapon. The first five of these quantities are normally supplied in current operational airborne weapon delivery systems by a variety of aircraft sensors, e.g., inertial platforms, air data sensors, radar or laser rangers, target-tracking devices, etc. While the accuracy with which a ballistic weapon can be delivered against a target depends greatly upon the accuracy of this sensor-supplied information, equally important to the problem is the *a priori* prediction of the weapon's trajectory based upon the sensor-supplied instantaneous weapon release parameters.

BASIC DIFFERENTIAL EQUATIONS

The development of an effective weapons release system is inherently dependent upon obtaining solutions of the equations for the motion of a projectile within the atmosphere. This is, generally, a difficult mathematical problem which has not been solved completely. The major difficulty stems from the nonlinearities introduced by the atmospheric effects on a falling weapon.

In choosing the mathematical model, two considerations have been kept in mind. The major objective of the mathematical analysis is to yield the weapon impact point. The main effect of this is that the weapon mass can be assumed to be a point mass. Also, the choice of the model is dictated by the need to evaluate results against some standard. Since the armed forces published range tables for various weapons, the model is chosen to conform as closely as possible to the model used for these tables.

The equations of motion are developed assuming the projectile is a point mass acted on only by the force of gravity and the retardation forces due to air resistance. The trajectory can be restricted to a plane by ignoring cross track effects such as winds. For practical applications, the effect of winds can be accounted for in a straight-forward manner.

The assumptions adopted are summarized below:

1. The Earth is flat and nonrotating.
2. The gravitational attraction is constant.
3. The projectile is a point mass.
4. The projectile is not powered and has a constant mass. (Rocket applications are discussed in detail in Ref. 1 and 2.)

Under these assumptions, the differential equations of motion (Ref. 3) have the following form.

$$\frac{d^2X}{dt^2} + H\frac{dX}{dt} = 0 \qquad (1)$$

$$\frac{d^2Y}{dt^2} + H\frac{dY}{dt} + G = 0$$

where X, Y, t, and G denote downrange, altitude, time and gravitational attraction, respectively. The coefficient H, which is the drag function, is given by

$$H = \frac{\rho}{W} d^2 \frac{\pi}{8} C_D V \qquad (2)$$

where ρ is the atmospheric mass density, W is the bomb mass, d is the bomb diameter, C_D is the weapon coeffient of drag, and V is the velocity in air mass. The atmospheric density ρ is given as a function of altitude which is fitted to measured values of atmospheric density. C_D is empirically derived and given in tabular form as a function of Mach number. A constant gravitational acceleration of 32.174 ft/sec^2 is quite adequate for most bombing applications and helps simplify the differential equations.

The above differential equations are not analytically integrable, if an accurate model of H is used, because an accurate model would render them extremely nonlinear.

The two second-order, differential equations given in Eq. 1 will now be rewritten as four first-order, differential equations. This is done to get the differential equations in a form that is more suitable to the integration process used. The new variables V_x and V_y are defined by

$$\frac{dX}{dt} = V_x \tag{3}$$

$$\frac{dY}{dt} = V_y$$

Substituting the above expressions in Eq. 1 results in

$$\frac{dV_x}{dt} = -H\,V_x$$

$$\frac{dV_y}{dt} = -H\,V_y - G \tag{4}$$

The four first-order, differential Eq. 3 and 4, are the desired equations with time as the independent variable. The Runge-Kutta integration formulas provide a step-by-step method of finding dependent variable values at given intervals of the independent variable. This is discussed in detail in Ref. 1. Figure 1 shows the salient features in the x-y plane.

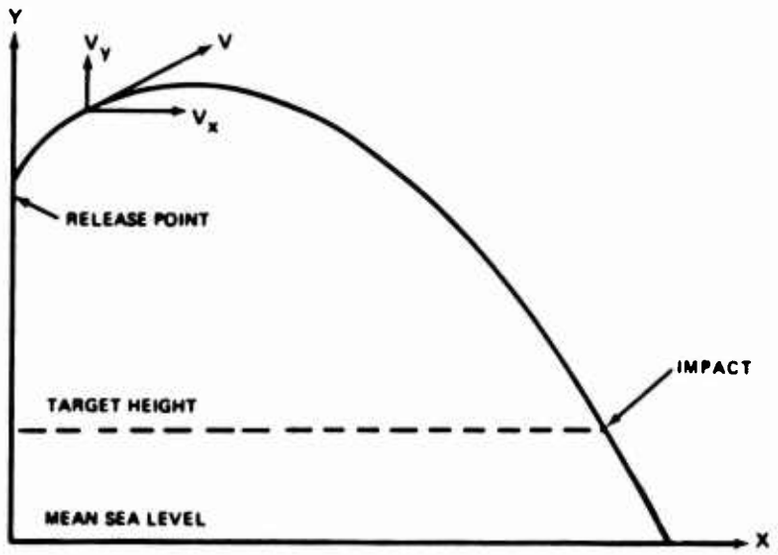


FIG. 1. Trajectory in the x-y Plane.

BALLISTIC ALGORITHM

The investigations leading to this paper and Ref. 1 were motivated by a need for a flexible, general purpose algorithm, suitable for near real-time solution by current airborne digital fire control computers, that would provide accurate weapon trajectory parameters for unguided air-to-surface weapon delivery.

The algorithm developed (see Fig. 2) uses a time-base, second-order, Runge-Kutta numerical integration process with a fixed number of time (integration) steps. Good accuracy is maintained for all weapons for all modes of delivery using 10 integration steps. To provide greater accuracy, or to expand the release envelope, more steps could be used at the expense of increased computation time.

The computer logic developed for implementing the weapon delivery algorithm into an airborne digital computer includes:

1. Starting procedure
2. Repetitive computations
3. Specification of the integration interval in each of the above cases
4. Monoriting of the state of the trajectory computations, i.e., whether the computation of a given trajectory or part of a trajectory is completed.

The algorithm logic is very general and can be used for dive, toss, loft, and over-the-shoulder weapon delivery. The logic handles the cases where target altitude is above or below the present aircraft altitude at weapon release. The algorithm is stable at any altitude, velocity, dive angle, and pullup maneuver the present A-7 type aircraft is capable of for any weapon in the inventory.

Aerospace companies and military installations have successfully coded the algorithm for different airborne digital computers. Using the algorithm, computer words can be reduced by half and computational time required reduced by up to 65 percent over conventional approaches used for airborne fire control.

BALLISTIC ALGORITHM EXTENDED

For fire control applications not requiring large ballistic lead-angles, the instantaneous velocity, V, in Eq. 2 can be approximated by the velocity component, V_r, along the slant range and thereby decreasing the computational time 30% and computer storage requirement 20% for calculating a weapon trajectory. This only requires making a change in coordinate systems, namely, letting the R-axis be along the slant range and S-axis be perpendicular to R as shown in Fig. 3. Thus Eq. 3 and 4 become

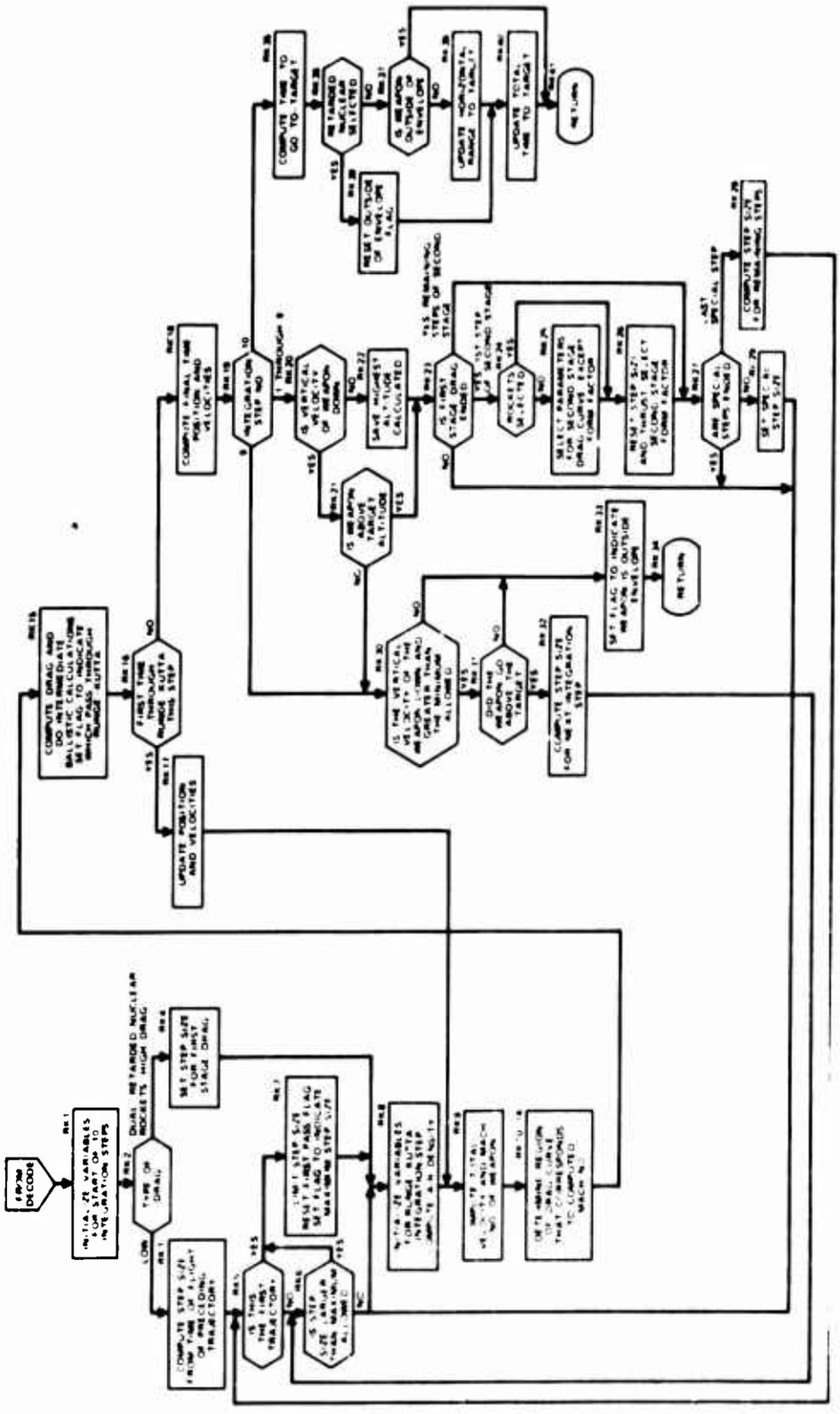


FIG. 2. Descriptive Flow Diagram of Algorithm

$$\frac{dR}{dt} = V_r$$

$$\frac{dS}{dt} = V_s \qquad (5)$$

$$\frac{dV_r}{dt} = -HV_r + G \sin \sigma$$

$$\frac{dV_s}{dt} = -H V_s - G \cos \sigma$$

For helicopter and antiaircraft fire control applications, it is advantageous to use a rotated coordinate system as in Fig. 3 since slant range to target and harp angle are normally known more accurately than target altitude and downrange to target. Ref. 2 extends the ballistic algorithm of Ref. 1 by using Eq. 5 and a convergence technique on slant range instead of target altitude to compute the ballistic lead angle and time-of-flight.

Aerospace companies and military installations have also successfully coded the extended algorithm. The logic of the extended algorithm (see Fig. 4) plus weapon parameter storage requires a core size of approximately 250 computer words. There is no restriction to the mode of weapon delivery.

CONCLUSIONS

The technique described in this paper can be used to predict the impact point and time-of-flight of any unguided weapon currently in the inventory. There is no restriction in the basic algorithm as to mode of weapon delivery. This wide applicability stems directly from the fact that an accurate representation of the differential equations of motion is numerically integrated to near-perfect (one fourth the weapon dispersion or less) accuracy by the algorithm. Furthermore, the algorithm has been extended by the author or others to handle the following problems:

1. Variable wind profile
2. Nonstandard air density
3. Altitude fuzed weapons
4. SHRIKE and other similar guided weapons
5. Applications involving helicopter, antiaircraft, and shipboard fire control

In summary, the algorithm can provide accuracy, flexibility, and efficiency in fire control applications.

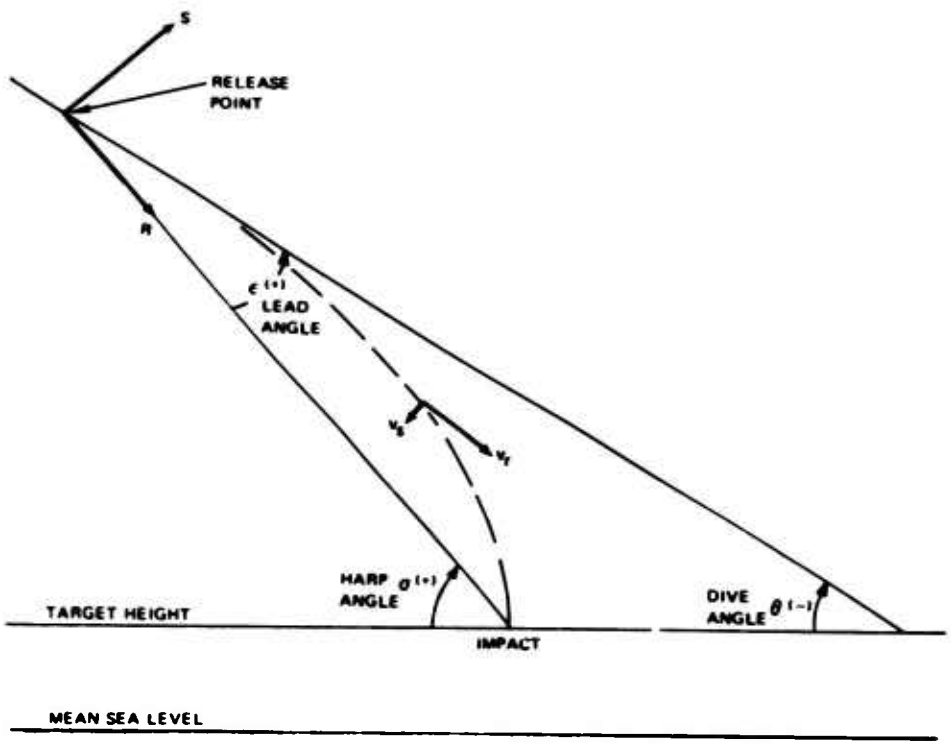


FIG. 3. Trajectory in the R-S Plane.

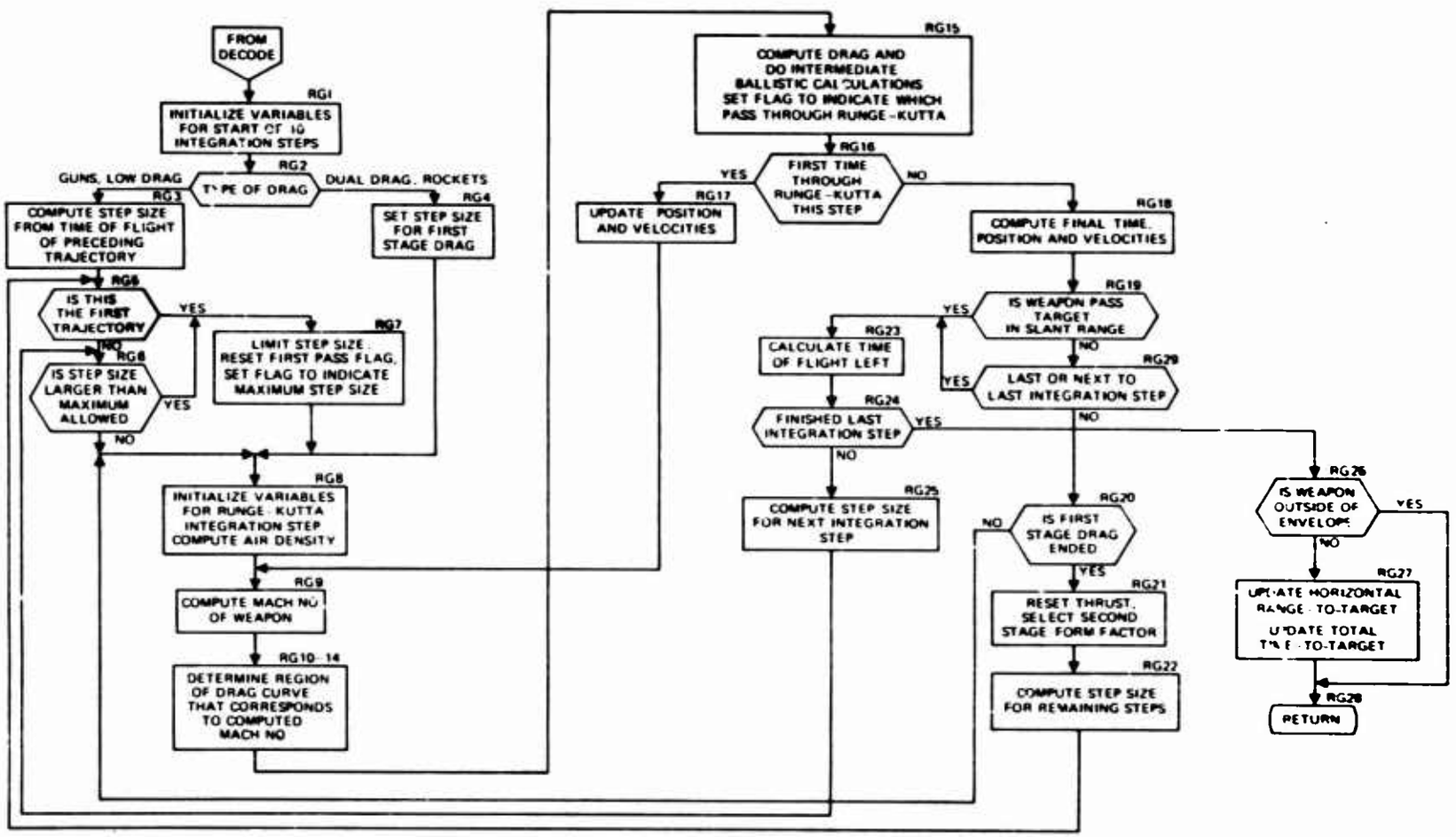


FIG. 4. Descriptive Flow Diagram of Ballistic Algorithm Extended for Helicopter and Antiaircraft Fire Control Applications.

REFERENCES

1. Naval Weapons Center. *A Ballistic Trajectory Algorithm for Digital Airborne Fire Control* (U), by Arthur A. Duke, Thomas H. Brown, Kenneth W. Burke, and Richard B. Seeley. China Lake, California, NWC, September 1972 (NWC TP 5416).

2. ------ *A Ballistic Trajectory Algorithm for Digital Fire Control* (U), by Arthur A. Duke. China Lake, California, NWC, November 1972 (NWC TN 4070-36-72).

3. Naval Ordnance Test Station. *Ballistic Handbook* (U), by R. B. Seeley and R. D. Cole. China Lake, California, NOTS, August 1965 (NOTS TP 3902).

NON-DIRECTOR

GUNSIGHT RESPONSE

MR. RALPH BRYAN

AIR FORCE AVIONICS LABORATORY

9 JULY 1974

Abstract

Background material is given on the aerial gunnery problem and the history of aerial gunfire control systems. The differential equations for the displayed lead angle are developed, with primary emphasis on the kinematic portion of the lead angle. Transfer functions describing lead angle response and inertial reticle response are developed for a Lead Computing Optical Sight and a "tracer" sight. The transfer functions are evaluated qualitatively in terms of two criteria, controllability and accuracy. Bode plots and computed and simulated time histories are used to illustrate typical results. Alternate dynamics are considered and evaluated qualitatively using the two criteria already mentioned.

I. Introduction

Air-to-air gunnery requires the attacker aircraft to aim his gun ahead of the target so that the bullet trajectory and target trajectory converge toward the same point in space. The time-of-flight is generally short by comparison with air-to-ground gunnery, usually from one-half to one and one-half seconds. The result is that, for air-to-air gunnery, the kinematic (target prediction) portion of the lead angle is much larger than the ballistic lead angle (the angle by which the bullet trajectory deviates from a straight line). Thus, while the ballistic portion of lead is not negligible, it is the kinematic portion which requires more engineering care.

Because of the shortness of the time-of-flight, and the difficulty of obtaining higher order measurements, the kinematic lead computation is not carried out to higher than second order. Even the second order measurement, target acceleration, is difficult to obtain without exceptionally good measurements. The obvious hardware with which to measure the target state is a target range and angle tracker. Due to the short time-of-flight, the linear term associated with the line-of-sight angular rate has a larger magnitude than the acceleration term. The analysis in this paper will deal only with angular rate term.

While much effort in modern control theory is directed toward estimating target states with a range and angle tracker, gunfire control systems using such estimates, called "director" systems, are not operationally used at present, nor are they forecast for near-term operational use. Non-director systems display a lead angle to the pilot based upon ownship measurements and target range (if available), without using measured angular rates from

an angle tracker. Of course, the lead angle displayed to the pilot cannot be the same as for a director system. This implies that the pilot must effectively estimate the difference between the correct lead angle and the actual display. Thus, the manner in which the non-director gunsight processes ownship information, and the manner in which the pilot estimates proper lead are highly coupled, not only from the conceptual standpoint (how well he understands the display), but also from the dynamic standpoint. The dynamic coupling is due to the use, in some manner, of ownship angular rates to compute lead angle. The result is that the pilot can control the reticle position only through the aircraft dynamics. The difficulty arises from the fact that he is simultaneously estimating his aiming error and controlling the aircraft.

Two classes of non-director systems exist today. These are the Lead Computing Optical Sight (LCOS), and the "tracer" sight. Conceptually, the LCOS estimates line-of-sight rates based upon the angular rates of the displayed reticle, and uses these rates to compute a lead angle. Conceptually, the tracer system displays the present location, in present aircraft coordinates, of previously "fired" ordnance, whether or not the gun was actually being fired. Hence, the name "tracer" accurately describes the sight concept, and a conceptual understanding of the sight is not difficult.

An understanding of the sight dynamics necessitates an understanding of the differential equations equivalent to the sight performance. Tracer sights are not mechanized on the basis of lead angle differential equations, but rather of ballistic differential equations operating separately from another differential equation set which integrates the necessary aircraft angular rates. Thus it was necessary, for this paper, to develop a set of lead angle differential equations.

Both of these sight concepts are being flight-tested in the Comparative Gunsight Evaluation Program. This program is attempting to compare state-of-the-art gunsight mechanizations in an F-106 flight test program at Tyndall AFB, Florida. In later phases of the test, alternate gunsight mechanizations will be tested. Thus, it is desireable to obtain a better analytic understanding of non-director gunsights, so that these alternate mechanizations can be intelligently designed.

II. Lead Computing Optical Sight Mathematical Development

The mathematical development of the Lead Computing Optical Sight equations is usually made upon the basis of a predicted miss distance, necessitating a prediction of target motion. The primary quantity required for this estimate is the angular rate of the line-of-sight (LOS). Since this information is not directly available to the airborne computer from a target tracker (due to noise and acquisition difficulty) the line-of-sight rate is estimated from the pilot's attempts to track the target. The way this is done is shown below. The primary term in the lead angle equation is the linear kinematic lead,

$$\underline{\lambda} \approx -\underline{\omega}_{LOS} t_f \tag{1}$$

where $\underline{\lambda}$ is the vector lead angle, $\underline{\omega}_{LOS}$ is the line-of-sight angular rate, and t_f is the bullet time-of-flight. Again, smaller and less dynamic terms have been deleted for clarity. In order to implement this equation without a LOS tracker, it is assumed that the pilot always has the reticle on the target, i.e., the pilot is performing perfect tracking. This leads to the replacement of the angular rate with the inertial angular rate of the reticle. This is formed using the difference of aircraft rates and lead angle rates. The angular rate of the reticle, $\underline{\omega}_R$, replaces $\underline{\omega}_{LOS}$

$$\underline{\omega}_{LOS} = \underline{\omega}_R \cdot \tag{2}$$

But then it is true that

$$\underline{\omega}_R = \underline{\omega}_A + \dot{\underline{\lambda}}^I \tag{3}$$

where $\underline{\omega}_A$ is the vector composed of aircraft inertial rates (p, q, r), and

$\dot{\underline{\lambda}}^I$ is the vector composed of inertial rates of reticle position relative to the aircraft. In order to obtain the conventional lead angle rates, which are relative rates of the reticle position, an additional substitution is required. The inertial lead angle rate is given by

$$\dot{\underline{\lambda}}^I = \dot{\underline{\lambda}} + \underline{\omega}_A \times \underline{\lambda} \tag{4}$$

where λ is the lead angle in aircraft coordinates. Substituting this expression for $\dot{\underline{\lambda}}^I$ in eq (3) leads to the final equation,

$$\underline{\Lambda} = -(\underline{\omega}_A + \dot{\underline{\lambda}} + \underline{\omega}_A \times \underline{\lambda})t_f$$

This can be rearranged to provide the desired differential equation,

$$\dot{\underline{\lambda}} = -\underline{\omega}_A - \frac{\underline{\lambda}}{t_f} - \underline{\omega}_A \times \underline{\lambda}$$

Separating the vector equation into its components yields:

$$\dot{\lambda}_r = -p - \frac{\lambda_r}{t_f} - q\lambda_a + r\lambda_e \tag{5}$$

$$\dot{\lambda}_e = -q - \frac{\lambda_e}{t_f} - r\lambda_r + p\lambda_a \tag{6}$$

$$\dot{\lambda}_a = -r - \frac{\lambda_a}{t_f} - p\lambda_e + q\lambda_r \tag{7}$$

At this point, λ_r, the reticle roll angle, can be assumed zero. The reason for this is that when the LCOS is implemented with a two degree-of-freedom gyroscope, only the λ_e and λ_a states can appear, due to the absence of a third degree-of-freedom. This results in the following set of equations:

$$\dot{\lambda}_e = -q - \frac{\lambda_e}{t_f} + p\lambda_a \tag{8}$$

$$\dot{\lambda}_a = -r - \frac{\lambda_a}{t_f} - p\lambda_e \qquad (9)$$

This result can also be obtained without making any assumptions concerning LCOS hardware. Returning to the original lead angle equation,

$$\underline{\lambda} = -\omega_{LOS} t_f \ ,$$

a rotation, λ_r, about the line-of-sight, corresponds to a line-of-sight roll rate. This term is of no significance, since the linear prediction term is derived from the equation

$$\underline{V}_T^I = \underline{V}_A^I + \dot{\underline{R}}^{LOS} + \omega^{LOS} \times \underline{R} \quad , \qquad (10)$$

where the target velocity is shown to be the sum the attacker velocity and the inertial relative velocity. Any component of $\underline{\omega}^{LOS}$ along $\underline{R}$ does not affect the cross product. Therefore, λ_r can be assumed zero.

Due to the difficulty of tracking with such a mechanization, it is customary to add lead angle rate feedback to the lead angle computation. Thus, the lead angle equation (1) becomes

$$\underline{\lambda} = -(\underline{\omega}_{LOS} + \sigma\dot{\underline{\lambda}}^I)t_f$$

Substituting for $\dot{\underline{\lambda}}^I$ as before results in a more general differential equation, in which sight damping can be varied parametrically. The final equations become

$$\lambda_a = -\frac{r}{1+\sigma} - \frac{\lambda_a}{(1+\sigma)t_f} - p\lambda_e \qquad (11)$$

$$\dot{\lambda}_e = -\frac{q}{1+\sigma} - \frac{\lambda_e}{(1+\sigma)t_f} + p\lambda_a \qquad (12)$$

If the time-of-flight is relatively constant, then the differential equations are very nearly linear, but have terms with a time-varying coefficient, $p(t)$, multiplying the lead angle components. To compute a general solution to perturbations about nominal conditions, it is reasonable to assume that λ_e and λ_a are approximately constant in these two terms only. The matrix formulation for this small perturbation model follows.

$$\begin{bmatrix} \dot{\lambda}_a \\ \dot{\lambda}_e \end{bmatrix} = \begin{bmatrix} \frac{-1}{(1+\sigma)t_f} & 0 \\ 0 & \frac{-1}{(1+\sigma)t_f} \end{bmatrix} \begin{bmatrix} \lambda_a \\ \lambda_e \end{bmatrix} + \begin{bmatrix} -\lambda_{e_o} & 0 & \frac{-1}{1+\sigma} \\ \lambda_{a_o} & \frac{-1}{1+\sigma} & 0 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} \quad (13)$$

Using Laplace transforms, the solution is

$$\begin{bmatrix} \lambda_a(s) \\ \lambda_e(s) \end{bmatrix} = \begin{bmatrix} \frac{-(1+\sigma)t_f\lambda_{e_o}}{(1+\sigma)t_f s + 1} & 0 & \frac{-t_f}{(1+\sigma)t_f s + 1} \\ \frac{(1+\sigma)t_f\lambda_{a_o}}{(1+\sigma)t_f s + 1} & \frac{-t_f}{(1+\sigma)t_f s + 1} & 0 \end{bmatrix} \begin{bmatrix} p(s) \\ q(s) \\ r(s) \end{bmatrix} \quad (14)$$

Features of these equations worth noting are

1. that every transfer function depends significantly upon time-of-flight, t_f,
2. that the settling time increases with t_f and σ, and
3. that roll coupling also depends upon the lead angle component in the other axis.

In order to reduce the settling time of the sight, it would appear that reducing

σ toward -1 would help. However, other considerations, developed in a later section, shall limit the success of this approach.

III. "Tracer" Sight Lead Angles

Computing the relative position of previously fired ordnance requires air data information and the vertical direction to compute the ballistic trajectory, and sufficient aircraft information to compute its flight path, and orientation. As was done for the LCOS, the computation here is simplified in order to show more clearly the dynamic effects of pilot inputs, i. e., aircraft angular rates. Therefore, it is assumed that the bullet follows a straight trajectory along the aircraft x-axis at time-of-fire, and that changes in relative orientation are due only to the aircraft rotation during the time-of-flight.

Let $\underline{R}_B$ be the bullet relative position vector for the bullet at range R_B in front of the aircraft. This range will be assumed constant for the time period of interest. At the time-of-fire this vector can be expressed as

$$\underline{R}_B = R_B \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{15}$$

in aircraft coordinates. It will remain constant in inertial coordinates, while the aircraft maneuvers for one time-of-flight, t_f. Then, the round is to be displayed. Let $T^*(t)$ be the transformation from inertial coordinates to body coordinates at time t. To obtain a single transformation, T(t), which will transform the vector $\underline{R}_B$ for a bullet fired at time $(t - t_f)$ to aircraft

coordinates at time t, the matrix product $T^*(t)T^{*\prime}(t - t_f)$ is formed. (Prime denotes transpose). The transpose can be used due to the orthogonality of T^*. A differential equation for the matrix T(t) can now be obtained

$$T(t) = T^*(t)T^{*\prime}(t - t_f) \tag{16}$$

$$\dot{T}(t) = \dot{T}^*(t)T^{*\prime}(t - t_f) + T^*(t)\dot{T}^{*\prime}(t - t_f) \tag{17}$$

Expanding $\dot{T}^*$,

$$\dot{T}^*(t) = \omega(t)T^*(t) \tag{18}$$

$$\dot{T}^{*\prime}(t) = T^{*\prime}(t)\omega'(t) \tag{19}$$

and substituting into the original equation,

$$T(t) = \omega(t)T^*(t)T^{*\prime}(t - t_f) + T^*(t)T^{*\prime}(t - t_f)\omega'(t - t_f) \tag{20}$$

$$\dot{T}(t) = \omega(t)T(t) + T(t)\omega'(t - t_f) \tag{21}$$

where

$$\omega = \begin{bmatrix} 0 & r & -q \\ -r & 0 & p \\ q & -p & 0 \end{bmatrix} \tag{22}$$

is formed from the aircraft angular rates, p, q, r. Since $\omega' = -\omega$,

$$\dot{T}(t) = \omega(t)T(t) - T(t)\omega(t - t_f) \tag{23}$$

The inverse transformation, T'(t), can be expressed as a three Euler-angle transformation, rotating about aircraft present z-axis through an azimuth

angle, λ_a; about the resultant y-axis through an elevation angle, λ_e; and finally about the resulting x-axis through a roll angle, λ_r. In terms of the angles λ_a, λ_e, and λ_r, this transformation is

$$T'(t) = \begin{bmatrix} c\lambda_e c\lambda_a & -c\lambda_r s\lambda_a + s\lambda_r s\lambda_e c\lambda_a & s\lambda_r s\lambda_a + c\lambda_r c\lambda_a s\lambda_e \\ c\lambda_e s\lambda_a & c\lambda_r c\lambda_a + s\lambda_r s\lambda_e s\lambda_a & -s\lambda_r c\lambda_a + c\lambda_r s\lambda_a s\lambda_e \\ -s\lambda_e & s\lambda_r c\lambda_e & c\lambda_e c\lambda_r \end{bmatrix} \quad (24)$$

where $c\lambda_e = \cos \lambda_e$, $s\lambda_e = \sin \lambda_e$, etc. λ_a and λ_e are the azimuth and elevation angles that would be displayed on an aircraft Heads-Up-Display (HUD).

The differential equation for T(t) is linear with time-varying coefficients. The coefficients are controlled by the pilot, and the output of interest is T_{21} and T_{31}. In order to simplify the problem enough so that the important features are illustrated, several assumptions will be made. First, it is assumed that the angles λ_a, λ_e, and λ_r are small enough to use small angle approximations. While this may not be desireable in an actual gunsight computation, this will not significantly affect an analysis of the dynamics. Secondly, it is assumed that perturbations will be made starting from a steady-state condition where the roll rate is small. With these assumptions, the equations for $\dot{T}_{21}$, $\dot{T}_{31}$, and $\dot{T}_{32}$ can be simplified to

$$-\dot{T}_{21} \approx \dot{\lambda}_a(t) = -r(t) - p(t)\lambda_e + r(t - t_f) + q(t - t_f)\lambda_r$$

$$\dot{T}_{31} \approx \dot{\lambda}_e(t) = q(t) + p(t)\lambda_a + q(t - t_f) - r(t - t_f)\lambda_r$$

$$-\dot{T}_{32} \approx \dot{\lambda}_r(t) = -q(t)\lambda_a - p(t) + r(t - t_f)\lambda_e + p(t - t_f)$$

An analytic solution to this set of equations is possible under some assumptions. However, any analytic solution to the problem when roll rates

are present is too unwieldy to be very enlightening. An analysis of this situation should use a numerical solution to the equation. It is possible, though, to describe analytically the solution when rates are present in one axis only. For example, consider

$$\dot{\lambda}_e(t) = - q(t) + q(t - t_f) \tag{25}$$

for which the solution is

$$\lambda_e(s) = \frac{q(s)\ (e^{-t_f s} - 1)}{s} \tag{26}$$

This equation would be valid for a pitching maneuver. The important feature to note is the time delay, in place of the time lag evident in the LCOS mechanization.

IV. Inertial Reticle Response

While the lead angle response equations are useful for determining steady state values and settling time characteristics, it does not give much insight into control aspects of the problem. To do this, it is necessary to describe the inertial rates of the reticle. The aircraft body rates (p, q, r) are inertial rates resolved into aircraft body axes. The body rate response to control inputs is a function of the aircraft and flight control system. Thus, the inertial reticle rates response to body rate inputs can be called fire control system response.

In order to derive a general expression for inertial reticle response, consider the reticle as a unit vector along the reticle axis. Let

$$\underline{P}^R = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

in reticle coordinates. Using the transformation, T(t), developed in the "tracer" section, this can be transformed to aircraft body axes.

$$\underline{P}^A = T'(t)\underline{P}^R = \begin{bmatrix} c\lambda_e c\lambda_a \\ c\lambda_e s\lambda_a \\ -s\lambda_a \end{bmatrix} \tag{27}$$

Using the transformation $T^*(t)$, which transforms from inertial to body axes, $\underline{P}^A$ can be transformed to inertial coordinates.

$$\underline{P}^I(t) = T^{*\prime}(t)\underline{P}^A(t) \tag{28}$$

Taking the time derivative of $P^I(t)$, and using the fact that

$$T^{*\prime}(t) = T^{*\prime}(t)\omega'(t), \tag{29}$$

$$\frac{d}{dt}\left[T^{*\prime}(t)\underline{P}^A(t)\right] = \dot{T}^{*\prime}(t)\underline{P}^A(t) + T^{*\prime}(t)\dot{\underline{P}}^A(t)$$

$$= T^{*\prime}(t)\omega'(t)\underline{P}^A(t) + \dot{T}^{*\prime}(t)\underline{P}^A(t) \tag{30}$$

Finally, transforming back to aircraft coordinates,

$$T^*(t)\ \frac{d}{dt}\left[T^{*\prime}(t)\underline{P}^A(t)\right] = \omega'(t)\underline{P}^A(t) + \dot{\underline{P}}^A(t) \tag{31}$$

Using small angle approximations for $\underline{P}^A$, substituting, and ignoring the first component (which is of no interest, since it corresponds to the length of the unit vector),

the azimuth and elevation inertial reticle rates are

$$\omega_{R_a} \approx r + p\lambda_e + \dot{\lambda}_a \tag{32}$$

$$\omega_{R_e} \approx q - p\lambda_a + \dot{\lambda}_e \tag{33}$$

where the sign of the second equation was changed when rectangular components were replaced by the approximate angular rates. These rates have the same sense as the aircraft body rates, i.e., positive right and up. They represent the two components (in aircraft coordinates) of the inertial reticle rate.

These results can now be applied to the LCOS lead angle differential equations, developed in the previous section,

$$\dot{\lambda}_a = -\frac{r}{1+\sigma} - \frac{\lambda_a}{(1+\sigma)t_f} - p\lambda_e \tag{34}$$

$$\dot{\lambda}_e = -\frac{q}{1+\sigma} - \frac{\lambda_e}{(1+\sigma)t_f} + p\lambda_a \tag{34}$$

Substituting in the equations for reticle rate, equations 32, 33

$$\omega_{R_a} = \frac{\sigma}{1+\sigma} r - \frac{\lambda_a}{(1+\sigma)t_f} \tag{36}$$

$$\omega_{R_e} = \frac{\sigma}{1+\sigma} q - \frac{\lambda_e}{(1+\sigma)t_f} \tag{37}$$

Substituting for $\lambda_a(s)$ and $\lambda_e(s)$, equation (14)

$$\omega_{R_a}(s) = \frac{\sigma t_f s + 1}{(1+\sigma)t_f s + 1} r(s) + \frac{\lambda_{e_o}}{(1+\sigma)t_f s + 1} p(s) \tag{38}$$

$$\omega_{R_e}(s) = \frac{\sigma t_f s + 1}{(1+\sigma)t_f s + 1} q(s) - \frac{\lambda_{a_o}}{(1+\sigma)t_f s + 1} p(s) \tag{39}$$

Qualitative Analysis

The transfer function for $\omega_{R_a}(s)/r(s)$ and $\omega_{R_e}(s)/q(s)$ is a lead-lag combination with a time constant proportional to time-of-flight. For good tracking, one would expect that increasing σ would result in a quickened sight response at higher control frequencies, although the larger pole would reduce the response slightly at frequencies near $\omega = 1/(1 + \sigma)t_f$. Returning to the lead angle response, however, it is noted that increasing σ increases the "settling" time of the sight, resulting in a longer time before the pilot is sure of a gun solution. The result is that increasing σ gives better response to "quick" control inputs, at the expense of poorer response in helping to estimate actual aim error.

There is no lead term in reticle response to roll rate inputs. This is natural, because there is no change in gun orientation due to a roll rate about the gun axis. Therefore, the actual aim error is slower to change. A detailed analysis of this term would depend upon the actual roll axis of the aircraft. However, an equation of this type could be used in such an analysis. For example, if an aircraft would roll about an axis below the gun, the angular rate would be a combination of "p", and "r". Thus, part of this "roll" rate input would appear in the first term of the equation, resulting a "lead" term in response to this "roll" input. The opposite case, one in which the roll axis is above the gun, would result in a negative "r" in the first term. This is sometimes known as the "pendulum" effect, since the reticle "swings" away from the target.

Application to "Tracer"

Substituting the differential equations for "tracer" lead angles into the

expression for inertial reticle response yields the following for the single-axis case:

$$\omega_{R_e} = q(t - t_f) \tag{40}$$

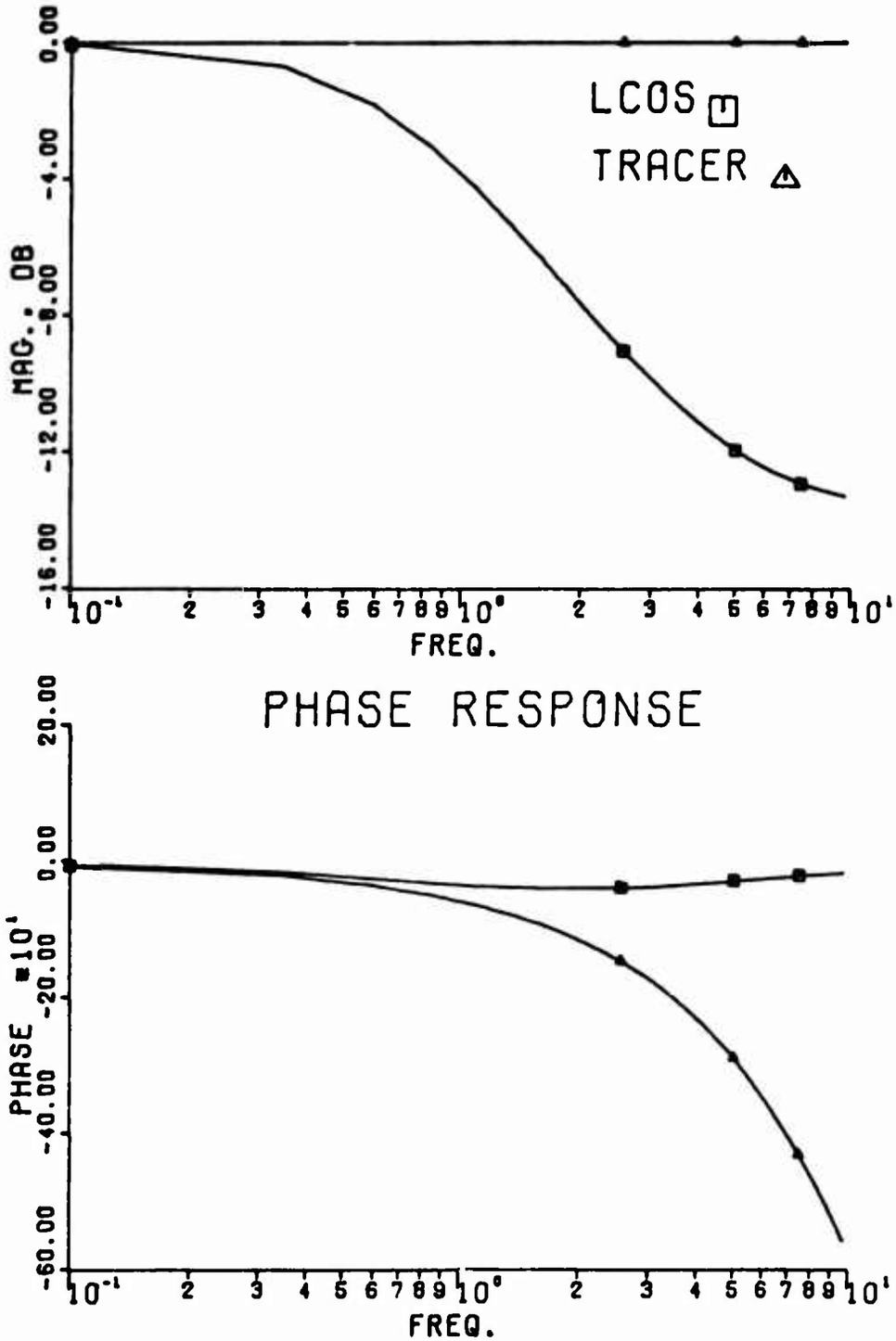
The Laplace transform is

$$\omega_{R_e}(s) = q(s)e^{-t_f s} \tag{41}$$

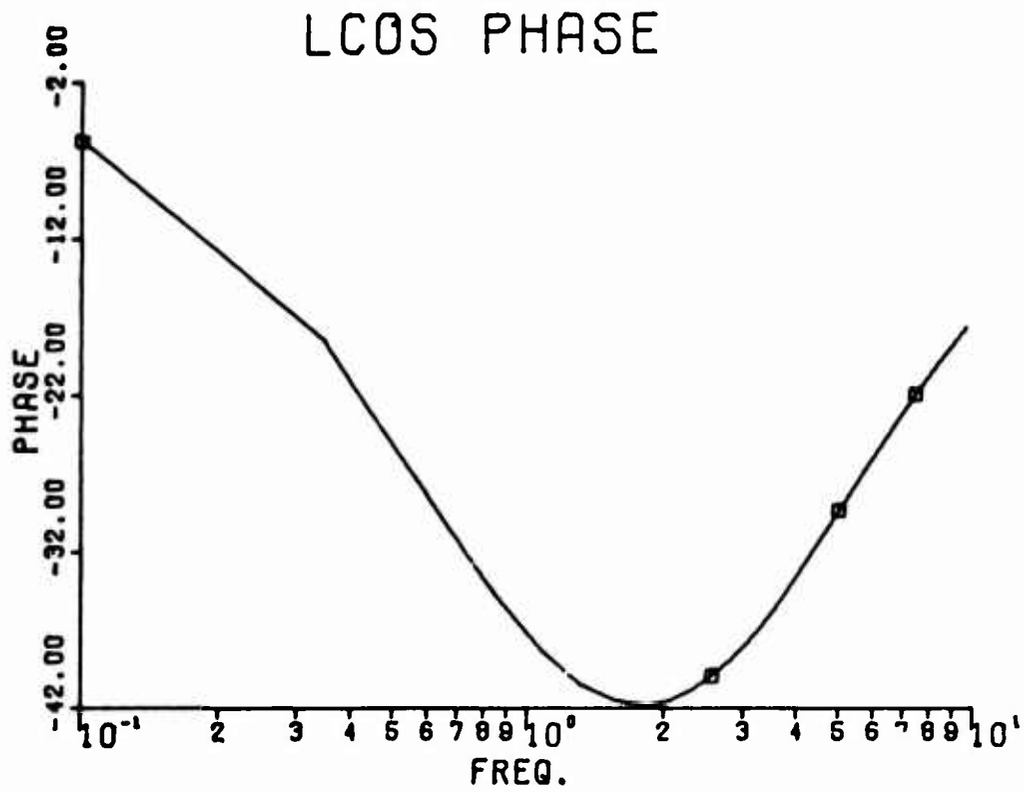
Qualitative Analysis

The transfer functions for a "tracer" involve a pure time delay, as opposed to the lead-lag characteristic of the LCOS. At this point, it is possible to clarify some of the confusion that exists concerning terms used to describe the sights. The tracer sight is referred to as an "undamped historic tracer", while the LCOS is sometimes referred to as a "damped predictive tracer". Sometimes these terms are construed to mean that the LCOS has a slower response due to the damping term which the "tracer" does not have. On the other hand, sometimes the LCOS is claimed to be predictive, in some mysterious manner. The truth is that they both have phase lag at control frequencies at or above $\omega_c = 1/t_f$. The claim that an LCOS is predictive is due to the fact that it does not have a pure delay term, but rather a lag term. The claim that a "tracer" sight is undamped only means that it does not have the lag term associated with the LCOS.

One observation to be made from the transfer functions for the "tracer" is that the amplitude response of the sight is unity for all frequencies. Another observation concerning the tracer sight is the fact that the phase



Inertial Reticle Response
LQGS σ= .25
Freq: .1 to 10. rad/sec



LCOS Phase Response -Expanded Scale

lag is equal to $\omega_c t_f$, where ω_c is the frequency of the rate input. This phase lag increases without bound at high control frequencies. Thus it is possible, for "quick" inputs, to be operating at an unstable point, where the phase lag exceeds 180^o with unity magnitude response. This fact is responsible for "wet noodle" performance of the tracer sight in dynamic situations.

IV. Alternate Dynamics

In this section, alternate gunsight dynamics are evaluated using the lead angle response and inertial reticle response developed in previous sections. Two different values of the parameter σ, used in the LCOS equations, are qualitatively evaluated. Then, a second order damping function is considered. In order to show just the principles involved, only the single axis problem will be studied.

One possible value of σ which could be used is $\sigma = -1$. Substituting this into the lead angle differential equations (14) provides:

$$\lambda_a(s) = - t_f r(s) \tag{40}$$

$$\lambda_e(s) = - t_f q(s) \tag{41}$$

We see that there is zero settling time. This is equivalent to assuming that the line-of-sight angular rate equals the aircraft (not the reticle) angular rate. The problem with this value of σ is evident upon substituting into the inertial reticle rate equations (38, 39):

$$\omega_{R_a}(s) = (1 - t_f s) r(s) + \lambda_{e_o} p(s) \tag{42}$$

$$\omega_{R_e}(s) = (1 - t_f s) q(s) - \lambda_{a_o} p(s) \tag{43}$$

The results are a zero in the right half "s" plane, an unstable situation where the reticle responds in a disturbing way to aircraft angular acceleration.

Another possible value of σ is $\sigma = -1/2$. This is an interesting possibility because upon substituting into equations (38, 39):

$$\omega_{R_a}(s) = \frac{1 - 1/2\ t_f s}{1 + 1/2 t_f s}\ r(s) + \frac{\lambda_{e_o}}{1/2 t_f s + 1}\ p(s) \tag{44}$$

$$\omega_{R_e}(s) = \frac{1 - 1/2 t_f s}{1 + 1/2\ t_f s}\ q(s) - \frac{\lambda_{a_o}}{1/2 t_f s + 1}\ p(s) \tag{45}$$

The resulting transfer functions are identical to the Pade approximation to the time delay. This approximation can be expressed as

$$e^{-t_f s} \approx \frac{1 - t_f s/2}{1 + t_f s/2} \tag{46}$$

But then, this is approximately the same transfer functions as for the tracer sight, which has the time delay function exactly. The Pade approximation is exact for magnitude response, but phase response does not match that of the time delay for frequencies above approximately 3 rad/sec, where the phase lag error is 60°.

Again, the right hand plane zero is destabilizing, though not as much as for $\sigma = -1$. The phase response of this function approaches 180° at high frequency, which could more easily become unstable than with a positive value of σ, where phase lag approaches 0° at high frequency (see graph for $\sigma = .25$). Aside from the destabilizing effect of $\sigma = -1/2$, this low value for σ does help shorten the settling time of the lead angle, as in equations (14).

Returning to equations (14), it is noted that the damping term is first order. It is possible to conceive a second order damping term for which the

general single axis equation (ignoring roll rate), would be of the form,

$$\lambda(s) = \frac{- t_f \omega_A(s)}{1 + as + bs^2} \tag{47}$$

Using this equation with equation (32) yields the inertial reticle rate.

$$\omega_R(s) = \frac{1 + (a - t_f)s + bs^2}{1 + as + bs^2} \omega_A(s) \tag{48}$$

This is yet a more generalized sight response than was obtained for the LCOS with the single parameter, σ (equation (38)). As in the case of the LCOS equations with $\sigma = - 1/2$, a proper choice of parameters can yield an approximation to tracer gunsight dynamics. If the parameters "a" and "b" are chosen such that

$$a = t_f/2 \tag{49}$$

$$b = t_f^2/12 \tag{50}$$

then the following equation results:

$$\omega_R(s) = \frac{1 - t_f s/2 + (t_f s)^2/12}{1 + t_f s/2 + (t_f s)^2/12} \tag{51}$$

This is a higher order, and more accurate approximation to the time delay than the Padé approximation (equation (46)). This approximation extends the phase response accuracy to 5 rad/sec.

V. Conclusions

1. A frequency domain analysis should be done on any gunsight modifications that affect the dynamics of the gunsight. There are many mechanizations which appear feasible until such an analysis is done.

2. The relationship and similarity between the tracer and the LCOS dynamics is better understood by studying frequency response.

3. By a proper choice of damping functions and parameters, it is possible to approximate a tracer response, at least for the single-axis case.

4. By a more extended analysis, which would include all three axes, the roll rate response could be studied, possibly in conjunction with the approximate aircraft transfer functions and pilot transfer functions (reference 2).

5. An analysis technique could be developed for the response to aircraft acceleration, similar to this analysis technique for angular rate input.

References

1. General Electric Company, "Air-to-Air Gunnery Improvement for F-4E per FSCP-34", August 1970.
2. Harvey, Captain Thomas R., "Application of An Optimal Control Pilot Model to Air-to-Air Combat", Thesis, March 1974.

ABSTRACT

DIGITAL CONTROL OF A TRAINABLE GUN

by

Edward J. Bauman, Lt Col, USAF*
Jack C. Henry, Capt, USAF*

This paper discusses the use of a mini-computer to generate accurate control commands which are used to position a movable 105 millimeter howitzer on-board the AC-130H gunship. The complete system is an airborne open loop, digital-analog control system. This hybrid combination consists of two rather sophisticated subsystems that were separately proven to have outstanding performance in previous applications. The trainable gun subsystem is an analog closed loop control system, designed and built by Delco Electronics Corporation. The AC-130E fire control subsystem uses an IBM 4 PI, model TC-2. This paper discusses the subsystems separately and then describes how these subsystems were made compatible by initial design and refinements resulting from actual flight testing. The complete system has produced excellent firing accuracies. Also included in this paper is a discussion of future applications of the trainable gun concept.

*Air Force Academy
Colorado Springs, Colorado

DIGITAL CONTROL OF A TRAINABLE GUN

Edward J. Bauman, Lt Colonel
Jack C. Henry, Captain
Department of Astronautics and Computer Science
U. S. Air Force Academy, Colorado

INTRODUCTION

The Department of Astronautics and Computer Science at the U.S. Air Force Academy, as part of the program of research and consultation, supplied assistance to the Gunship System Program Office (SPO) and was requested to conduct the principle investigation of error sources and their possible correction. Early in the research it was determined that the pilot was a major error source, since he was required to control the attitude with an exceedingly high degree of precision in order to accurately aim the fixed side-mounted gun.

One approach to eliminating the pilot's error was to replace him with an autopilot.[1] The simplest and easiest approach, however, to eliminating the pilot error is to have the pilot fly near nominal attack geometry and allow the gun to move the aim toward ("train to") the designated target. An indication of the ease and simplicity of the trainable gun approach was the time required for complete development. Less than one year elapsed between the time the approach was accepted by the Gunship SPO and actual combat deployment! The speed of completion and the success achieved was largely a function of the subsystems used. Each subsystem had been separately proven to have outstanding performance in previous applications. Delco Electronic Corp was awarded a contract to position the 105mm Howitzer in both azimuth and elevation to within certain specified accuracies. IBM was awarded the contract to provide the gun position command signal. The development, implementation, results, and future applications of this idea are discussed in this paper.

TRAINABLE GUN CONCEPT

In the gunship fire control computation, all known anomalies in the system are accounted for. The pilot is the major error source in the system, having difficulties in aligning the computed impact point (CIP) and the aim point to within

a few milli-radians. At 10,000 feet slant range, one milli-radian amounts to a 10 foot miss.

The concept of the trainable gun is to have the pilot fly near nominal attack geometry and have the gun train to the designated target. Under this system the pilot simply keeps the aim point and CIP within a specified region on the HUD as shown in Figure 1. The signal representing the difference between the CIP and the aim point is used to correct the gun position. The effectiveness of this concept depends almost entirely upon the accuracy of the CIP calculation and the accuracy of positioning the gun. The CIP calculation has been demonstrated to be accurate in the "fixed-gun" gunship operations. The gun positioning problem was solved by Delco Electronics on the U.S. Army main battle tank. The accuracy of that system was demonstrated to be sufficient for gunship operation.

FIRE CONTROL COMPUTER SUBSYSTEM

On the AC-130H an IBM 4 Pi Model TC-2 digital computer is used to compute the CIP, the aim point, and the commanded gun angles. The most complicated and time consuming computation is the CIP. In this calculation are (1) a Kalman estimator to determine gun-sensor misalignments and the effect of winds on the projectile; (2) equations to determine accurate fire control correction angles. The aim point is a series of axis transformations using sensor orientation angles and aircraft attitude. The commanded gun angle is the sum of the present gun position and the difference between the aim point and the CIP. This commanded gun angle is converted to an analog signal by the digital-to-analog (D/A) converter within the TC-2.

TRAINABLE GUN SUBSYSTEM

The trainable gun subsystem consists primarily of control transformers, an electronic control unit, servo valves, hydraulic actuators, gun mount and gun. A synchro control transformer is mounted on both the azimuth and elevation gun rotation axes. The commanded gun position is fed to these synchros from the D/A converter. The gun position error signals are fed to the electronic control unit. The control unit demodulates and filters the synchro signals.

Angular acceleration feedback using differential hydraulic pressure and forward compensation for gun bending modes are used in each control axis. When the synchro position error signal is less than 0.5 mr in each axis, a ready to fire signal is sent back to the digital computer. If this signal is not present, gun fire is inhibited. This means the only feedback to the computer from the gun is simply a go/no-go signal and is not used in any way to influence the command computation.

The hydraulic drive system for the gun can produce a maximum acceleration of 2300 deg/sec^2 and a maximum velocity of 6 deg/sec. Considering the computational rates discussed in the next section, this velocity rate is quite adequate to keep the gun aligned to the commanded position during times when the gun is allowed to be fired.

RESULTS OF AIRBORNE TESTS

Reduced data from the initial airborne testing clearly showed that the major computation rate of the digital computer was insufficient to position the gun during rapid aircraft maneuvers. For example, if aircraft attitude rates were as much as 5 degrees per second, then calculations performed only once every 0.2 seconds allowed the CIP to lag the aimline by one degree. Of course, with this large an error, a computer coincidence check would inhibit fire. Therefore, until the aircraft attitude rates were nulled by the pilot, the gun could not be fired which increased the time spent in the target area. Consequently, the computation of the CIP was changed to the minor computation rate. The error lag was considerably reduced, at the cost of much more computation time. Now for a rate of 5 degrees per second, the error lag would be only 0.25 degrees.

The digital computational rate thus becomes a direct limitation on the tolerable aircraft attitude rates. For example, if the maximum acceptable error lag is 1/20 degree, then the maximum allowable aircraft rate is one degree per second if computations are done at the minor rate. The pilot's main task now, is not to precisely align the CIP and aimline, but to null aircraft rates to

less than 1 degree per second. If the computational rate was increased, higher tolerable aircraft rates would be possible. However, not enough computer time is available to increase the computational rate and be able to complete all the computations each cycle.

FUTURE APPLICATION

Fighter pilots have the same problem as gunship pilots in positioning their aircraft attitude accurately and steadily. Additionally, fighter pilots in air-to-air combat often find themselves in positions where they can not quite pull the necessary lead for a hit. This could be due to the enemy's defensive tactics or possibly a superior maneuvering aircraft. Obviously, a trainable gun on the fighter aircraft could eliminate, or certainly alleviate, these problems. The big question is whether trainable gun technology from the gunship can be logically extended to fighters. The answer is yes; however, there is one other variable associated with air-to-air combat --- lead. Present lead computing fire control systems are inadequate for trainable gun application. Improved sensor/tracker technology is needed to acquire and track targets accurately over a whole dynamic range.

CONCLUSIONS

The use of an airborne digital computer allowed the computation of gun position signals at a rate which could compensate for moderate aircraft attitude rates. The hydraulic control system moved the trainable gun at more than sufficient accelerations and rates to follow the generated commands. The net result was essentially uncoupling the pilot's task of flying the circular trajectory from the task of aiming the gun. Confirmation that the system was improved, as much as predicted by the test results, came from actual combat missions performed in Southeast Asia. A logical and feasible extension of gunship trainable gun technology is application to fighter aircraft.

REFERENCES

1. Parkinson, B.W., Kruczynski, L.R., and Wynne, M.W., "Sight Line Autopilot: A New Concept in Air Weapons," AIAA Paper No. 71-960.

KALMAN FILTERING FOR PRECISION POINTING AND TRACKING APPLICATIONS

BY

CAPT ROBERT B. ASHER

F. J. SEILER RESEARCH LABORATORY

USAFA, CO

AND

1/Lt DAVID H. WATJEN

AIR FORCE AVIONICS LABORATORY

W-PAFB, OH

ABSTRACT

Many aerospace problems require that information to precisely locate an object relative to a moving vehicle be established. This report considers the application of an extended Kalman filter to obtain minimum variance estimates of certain required physical variables in order to accomplish precision pointing and tracking. The work considered is formulated in a general framework and then applied to a specific situation. The model necessary for the Kalman filter is generated in detail. Included in the model, are equations for pertinent error sources. If the dominant error sources (relative to the application) are not modeled, then the filter may have poor performance and may in fact diverge. Equations based on covariance analysis that relate the performance of a suboptimal filter to the true state model are developed and presented. Six filter configurations are analyzed using these covariance analysis techniques. The results show the comparison between the filter's own predicted performance and the filter's performance as predicted by the truth model.

I. Introduction

Many aerospace problems require that information to precisely locate a particular object relative to a moving vehicle be established. The object may be stationary relative to the earth, or it may in general be moving relative to inertial space. (It is assumed that for the time intervals under consideration that the earth coordinate frame is an inertial frame. The relaxation of this assumption causes little problem.) Once the particular object that is to be tracked is within the field of view of the particular tracker being used, i.e. the acquisition problem has been solved, then the necessary information must be established both to control the tracker to keep a continuous, precise track of the object and to obtain the necessary information in order to enable the vehicle to perform its required function. The accurate pointing of the tracker will allow one to obtain the necessary information as to the location of the object relative to the tracker and as to the velocity of the object relative to the tracker.

In the missile guidance problem one needs to have the line of sight to the target precisely defined. This is necessary to be able to apply a control acceleration normal to the line of sight, to obtain a measurement of range/range rate relative to the target for guidance command generation, and to obtain a measurement for the angular velocity of the line of sight for guidance command generation. In the gun fire control problem the lead prediction is dependent on the angular velocity and angular acceleration of the line of sight, and on the range and range rate to the target. An application of the philosophy in air to ground weapon delivery includes the angle rate bombing system.

In all the above applications the basic problem is that of accurately pointing a sensor to track an object. This report considers the application of an extended Kalman filter to obtain minimum variance estimates of certain required physical variables in order to accomplish the function in the pointing and tracking problem. The work considered is formulated in a general framework, and may be readily applied to the several application areas mentioned.

The model necessary for the Kalman filter is generated in detail. All apparent possible error sources are modeled. The reason is that many state of the art applications of the pointing and tracking philosophy require an extremely accurate estimate. If the dominant error sources (relative to the application) are not modeled, then the filter will at best have poor performance and may in fact diverge. The modeling leads to a high order state model. In order to obtain a reasonable real-time implementation a suboptimal filter of lower dimensionality must be used. The results are studied using covariance analysis techniques as in [12,13]. This is used to establish the design of and the performance bounds for both a radar sensor and an infrared sensor.

Work in the area of Kalman filtering for pointing and tracking includes work by Fitts [6, 7, 8, 9]. Filtering is accomplished relative to an inertial reference system while this report considers filtering in the line of sight coordinate system. Also, the modeling of the error sources was not considered in detail. Pearson [10] considers a general study but does not include the methodology for high accuracy modeling. Also Landau [5] considers filtering in the line of sight coordinate system but does not model the error sources. References [1], [2], and [11] contain preliminary work that this paper is based upon. None of the above references consider the problem of performance evaluation. The modeling of error sources is important since exclusion of pertinent error sources may cause the filter design

to diverge when applied in the system during actual operation.

Problem Statement

In order to obtain the location of the target relative to the tracker and the velocity of the target, one must have the following information: the elevation angle and the azimuth angle of the line of sight; angular velocity relative to an inertial frame; the range and range rate; and the target acceleration. Since the tracker control problem may be considered as a position control problem one must have an estimate of the angular misalignment to command the tracker to null the error. Also, the estimates of the angular velocity components may be used to rate aid the tracker system. The target acceleration is necessary because it is coupled with the estimation equations for the above variables and is desireable for fire control purposes.

Measurements of the angular error are available in the tracker system. However, the measurements are in general noisy in that certain physical phenomenon, for example scintillation noise for radar systems, act to corrupt them. There is significant power at low frequencies for several different noise phenomenon found in radar, electro-optical, and infrared trackers. Thus, one cannot build the tracker system to reject the noise. Also, a low pass characteristic will lead to undesireable dynamic response characteristics in the tracker system. The Kalman filter may be used to obtain a minimum variance estimate of the required physical variables.

Figure one depicts the kinematics and physical geometry at the tracker. In this figure, the (x,y,z) coordinate system is fixed to the vehicle and oriented along the body frame coordinate system. Figure two depicts the kinematics of the tracker relative to the vehicle. Figure three depicts the line of sight tracker geometry in detail. The nomenclature used in the figures are defined as follows:

a_T is the acceleration of the tracker origin relative to inertial space

v_T is the velocity of the tracker origin relative to inertial space

r is the range to the target measured from the tracker

$\dot{r}$ is the range rate

ε_T is the elevation angle of the tracker

η_T is the azimuth angle of the tracker

$\delta\varepsilon$ is the angular misalignment between the tracker centerline and the line of sight in elevation

$\delta\eta$ is the angular misalignment between the tracker centerline and the line of sight in azimuth

a_{Tar} is the acceleration of the target relative to inertial space

v_{Tar} is the velocity of the target relative to inertial space

R is the vector from the center of mass to the tracker origin

ω_L is the angular velocity of the line of sight relative to inertial space

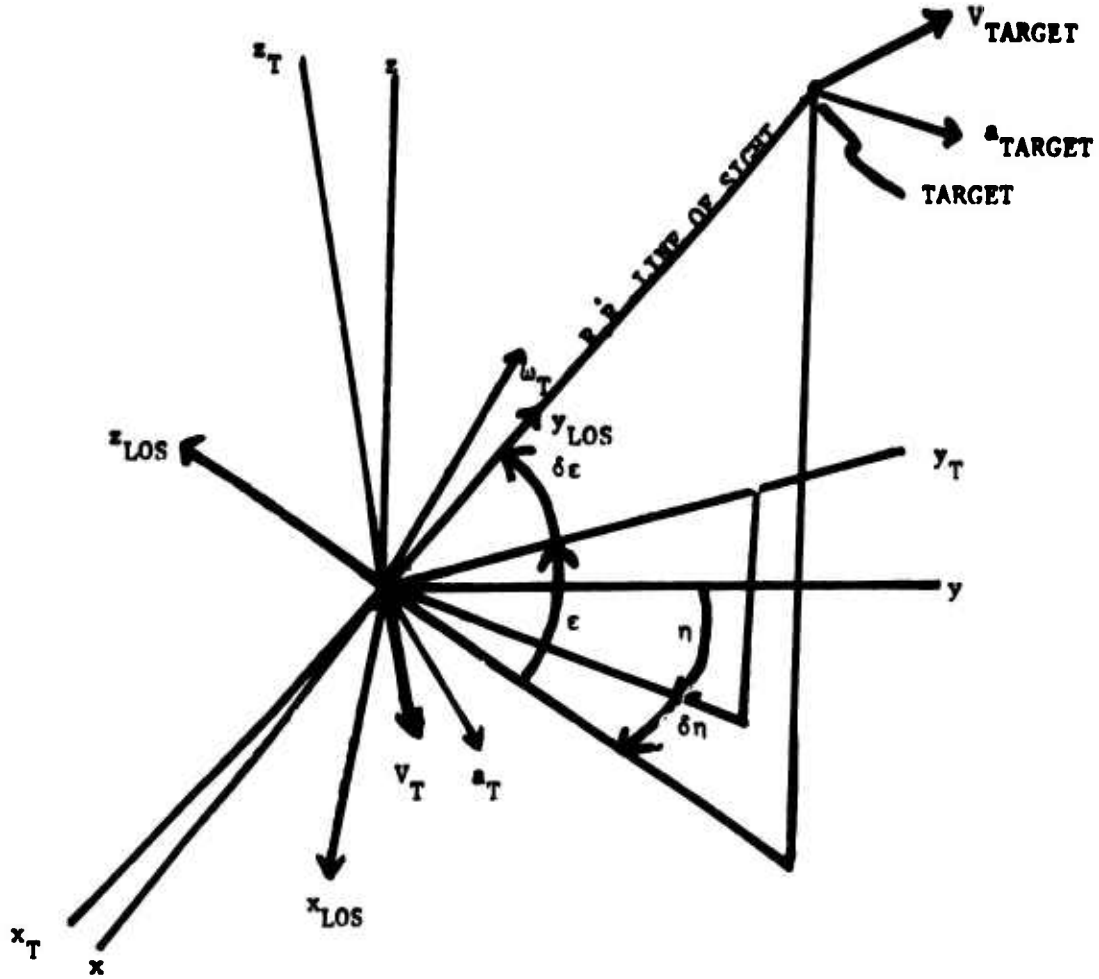


Figure 1: Tracker Geometry and Kinematics

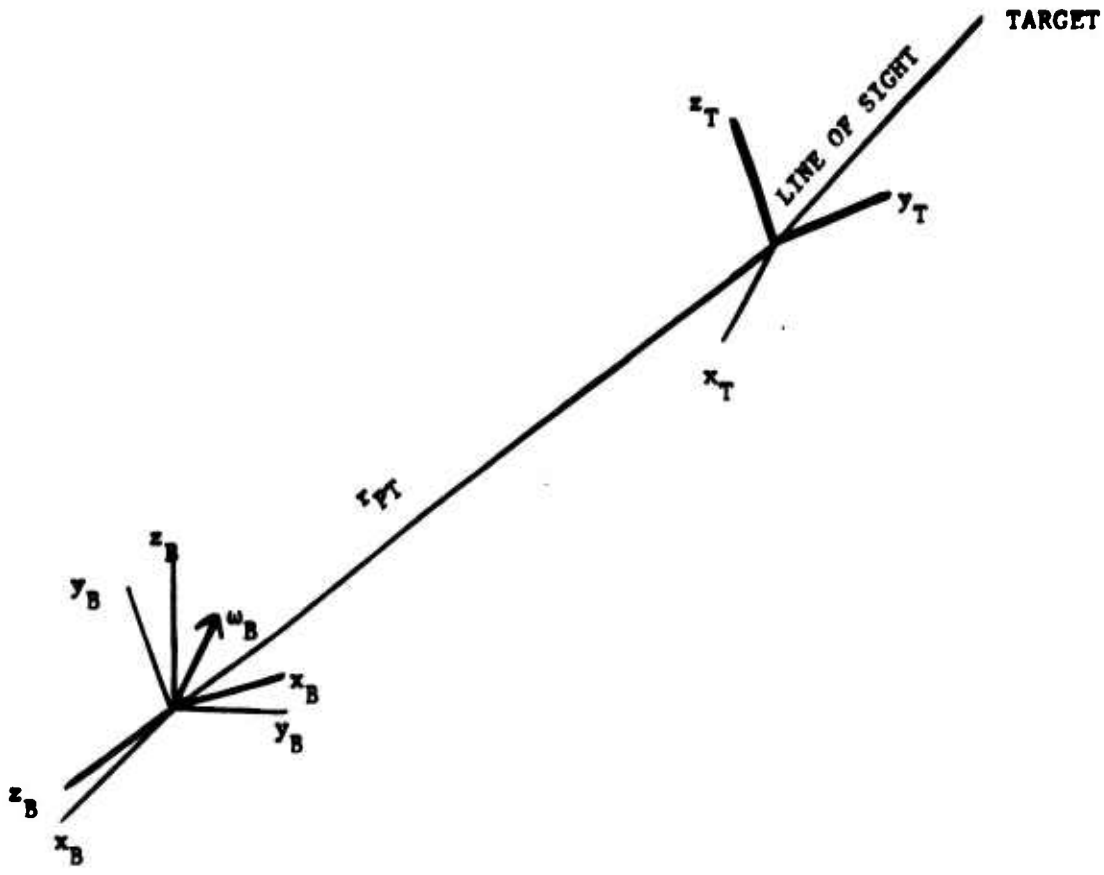


Figure 2: Aircraft/Tracker Geometry and Kinematics

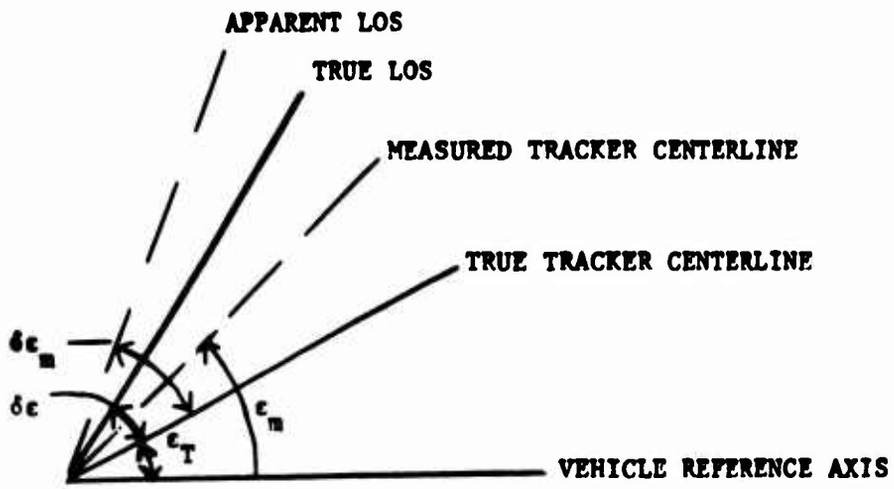


Figure 3: Angular Measurements

ω_T is the angular velocity of the tracker relative to inertial space

II. Optimal State and Measurement Equations

In this section the equations giving the true state model are given. These equations consist of both the algebraic measurement equations relating the states to the physical measurements available and the differential equations giving the time evolution of the system state. The physical error sources are modeled in detail and augmented to the state vector. This will allow one to study the effects of suboptimality by the implementation of a suboptimal filter.

As was discussed in the previous section estimates of the elevation and azimuth angles between the line of sight and the tracker centerline, the line of sight angular velocity, the range/range rate, and the target acceleration are required. Differential equations for these physical variables are required. Also, the algebraic measurement equations consisting of measurements of range, elevation and azimuth misalignment angles, and a pseudo measurement of the angular velocity of the line of sight obtained by using the tracker rate gyro outputs are required. The next subsection will give the measurement equations for the optimal system.

Measurement Equations

The measurement to be utilized includes that of a pseudo-measurement of the angular velocity of the line of sight. This measurement consists of using the tracker rate gyro outputs as a measurement of the angular velocity of the line of sight. If the tracker is very stable relative to the line of sight, then this measurement will give an accurate measure of the true angular velocity.

The angular velocity of the tracker measured by rate gyros will be in error because of gyro errors. Also, errors due to misalignment of the gyros will induce error into the measurements. Thus, the measured angular velocity is given as:

$$\omega_{T_i} = \omega_{TT_i} + b_g + c_g + \sum_{\ell=1}^{3} k_\ell a_\ell + \sum_{\ell=1}^{3}\sum_{j=1}^{3} k_{j\ell} a_\ell a_j + \delta k_s \omega_{TT_i} + [\Delta C_g^T \omega_{TT}]_i + \xi_i \quad , \; i = \varepsilon, \eta, r \tag{1}$$

where

ω_{TT_i} is the true tracker angular velocity

ω_{T_i} is the measured tracker angular velocity and i denotes the particular measurement, ε denotes elevation gyro, η denotes azimuth gyro, and r denotes roll gyro

b_g denotes a bias error term in rad/unit time

c_g denotes a random component of gyro drift in rad/unit time

k_i are the coefficients of the g-sensitive mass unbalance terms in rad/unit time -g

k_{ij} are the coefficients of the g^2-sensitive anisoelastic drift terms in rad/unit time $-g^2$

a_i are the along and cross axis acceleration terms

δk_s is a torque scale factor error factor

ω_{TT} is the vector consisting of the components of the true tracker angular velocity

ΔC_g^T is an error angle transformation matrix relating the angular velocity in gyro coordinates to tracker coordinates and $[\ (\cdot)\ ]_i$ denotes the i-th row of the vector $(\cdot)$

The error angle transformation matrix may be written as reference [3]. The magnitude of the gyro errors are dependent on the particular gyro used. The effect of the errors on the estimation error will be dependent on the magnitude of these errors and of course on the particular application of the results. The bias terms, b, are the uncompensated random bias occurring due to miscalibration. The random drift is due to any time correlated bias drifts. The bias drift and random drift may be incorporated as an exponentially correlated sequence, i.e., autocorrelation function

$$R_{RD}(T) = \sigma^2 \exp\{-\beta|T|\} \tag{2}$$

where

β is the correlation time of the process

σ is the rms value of the process

Thus, one may write the following stochastic differential equation relating the time evaluation of the random drift as

$$\dot{c}_g = -\beta c_g + \sqrt{2\beta}\,\sigma\, u. \tag{3}$$

The bias term may be included in the above drift model by appropriately choosing the initial condition.

Since in general the constants, k_i, k_{ij}, are not known exactly (if known exactly they could be calibrated out), one must model the coefficients as random constants with differential equations.

$$\dot{b}_{gm_i} = o,\ i = \varepsilon, \eta, r \tag{4}$$

$$\dot{b}_{gA_{ij}} = o,\ i, j = \varepsilon, \eta, r,\ i \neq j,$$

where $b_{gm_i} \triangleq k_i$ and $b_{gA_{ij}} \triangleq k_{ij}$, and the variance of the constants and, thus, the rms value is known. Similarly, the scale factor error may be modeled as a random constant with differential equation

$$\dot{b}_{gsf_i} = o,\ i = \varepsilon, \eta, r \tag{5}$$

and known variance. The error angles defined in the error angle transformation matrix may be considered to be random constants with differential equations

$$\dot{b}_{gmA_{ij}} = 0 \quad , i , j = x, y, z, \quad i \neq j , \tag{6}$$

where $b_{gmA_{ij}} \overset{\Delta}{=} \theta_{ij}$, and the variance known. Thus, one may augment the state vector with the above bias terms and first order Markov terms. In general the acceleration measurement and its associated error sources may be written as

$$\begin{aligned} a_{T_i} = {} & k_1 a_{TT_i} + \delta k_1 a_{TT_i} + b_a + c_a + \\ & k_2 a_{TT_i}^2 + k_3 a_{TT_i}^3 + k_{12} a_{T_j} + k_{13} a_{T_k} + \\ & [C_a^{ac} \; a_{TT}]_i \quad , i = x, y, z \\ & \qquad\qquad j, k \neq i \\ & \qquad\qquad j \neq k \end{aligned} \tag{7}$$

where

a_{TT_i} is the true acceleration of the tracker origin

a_{T_i} is the measured acceleration of the tracker origin and i denotes the particular acceleration component in the accelerometer coordinate frame

δk_1 is an accelerometer scale factor error

b_a is a bias error term

c_a denotes a random component of acceleration error, i.e., accelerometer drift

k_i are the nonlinear calibration coefficients

k_{ij} are the crosstrack acceleration errors

a_{TT} is the vector consisting of the components of the true acceleration

C_a^{ac} is an error angle transformation matrix relating the acceleration in actual accelerometer coordinates to nominal accelerometer coordinates which gives the axes for transformation to tracker coordinates

and $[(\cdot)]_i$ denotes the i-th row of the vector $(\cdot)$. One may note from the discussion above that in general the accelerometer measurements may not be in the tracker coordinate frame. One may easily transform the measurements from the accelerometer frame (i.e., taken to be the nominal, computer frame location of the accelerometer) to the required tracker frame.

Similarly to the problem of random gyro drift one may formulate the accelerometer drift term as a first order Markov process with state equation

$$\dot{c}_a = -\beta c_a + \sqrt{2\beta}\,\sigma\, u \tag{8}$$

Also, one may model the scale factor, δk, as a random constant, i.e.,

$$\dot{b}_{ASF} = 0\ ,\ i = x, y, z. \tag{9}$$

The coefficients k_i, k_{ij}, may be modeled as random constants, and the nonzero elements of the error angle matrix may be modeled as random constants.

At this point it is appropriate to note that if very long term performance is required, then it may be appropriate to include higher order terms in the random components in both gyro and accelerometer drift. This may include nonstationary effects in the statistics.

Other measurements necessary for both tracker control and obtaining the location of the target relative to the tracker are the angular misalignments $\delta\varepsilon$ and $\delta\eta$ (also referred to as the boresight error). The measurements may be assumed given on the following form

$$y_{\delta\varepsilon} = k_1\,(\delta\varepsilon_T + s_{\delta\varepsilon}) + \delta k_1 \delta\varepsilon_T + b_{\delta\varepsilon} + \omega_{\delta\varepsilon} \tag{10}$$

$$y_{\delta\eta} = k_1\,(\delta\eta_T + s_{\delta\eta}) + \delta k_1 \delta\eta_T + b_{\delta\eta} + \omega_{\delta\eta}$$

where

$y_{\delta\varepsilon}$, $y_{\delta\eta}$ are the measurements of the elevation and azimuth misalignments respectively

k_1 is a scale factor

δk_1 is a scale factor error

$s_{\delta\varepsilon}$, $s_{\delta\eta}$ are random components of the measurement-assumed to be a first order Markov process

$\delta\varepsilon_T$, $\delta\eta_T$ are the true misalignment angles

$b_{\delta\varepsilon}$, $b_{\delta\eta}$ are bias errors in the measurement

$\omega_{\delta\varepsilon}$, $\omega_{\delta\eta}$ are residual errors assumed to be represented by white noise with known variance.

The random components $s_{\delta\varepsilon}$ and $s_{\delta\eta}$ are represented as

$$\dot{s}_{\delta\varepsilon} = -\beta_\varepsilon s_{\delta\varepsilon} + \sqrt{2\beta}_\varepsilon\,\sigma_\varepsilon\, u \tag{11}$$

$$\dot{s}_{\delta\eta} = -\beta_\eta s_{\delta\eta} + \sqrt{2\beta}_\eta\,\sigma_\eta\, u$$

where the variables correspond to previously defined variables. The scintillation error components are measured relative to an inertial frame. However, as the tracker rolls the scintillation error rate will couple with the tracker roll rate. This may be taken into consideration of the following equation

$$\dot{s}_{\delta/I} = \dot{s}_{\delta/T} + w_T \times s_\delta \tag{12}$$

This yields the following components noting that only two angles are necessary to determine the location of the line of sight (where $s_{\delta\varepsilon}$ and $s_{\delta\eta}$ are the scintillation components relative to the tracker frame)

$$\dot{s}_{\delta\varepsilon} = -\beta_\varepsilon s_{\delta\varepsilon} - w_{Tr} s_{\delta\eta} + \sqrt{2\beta_\varepsilon}\sigma_\varepsilon u$$

$$\dot{s}_{\delta\eta} = -\beta_\eta s_{\delta\eta} + w_{Tr} s_{\delta\varepsilon} + \sqrt{2\beta_\eta}\sigma_\eta u.$$

The scale factor error is modeled as a first order Markov process, i.e.,

$$\dot{c}_{SF} = -\beta c_{SF} + \sqrt{2\beta}\,\sigma\ u \tag{13}$$

where, again, the variables correspond to previously defined variables. The term representing any calibration bias is modeled as a random constant. Any residual errors are assumed incorporated into the white noise components of measurement noise.

In order to estimate range and range rate one must have a filter. It is assumed that range measurements are available and are given as

$$y_r = k_r(r_T + c_r) + \delta k_r r_T + b_r + \omega_r \tag{14}$$

where

k_r is a scale factor

δk_r is a scale factor error

y_r is the range measurement

r_T is is the true range

c_r is a random component of range error-assumed to be a first order Markov process

b_r is a range bias error

ω_r is a white noise error representing all other residual error sources with known variances.

The random component of range error is modeled as a first order Markov. The scale factor error is assumed to be modeled as a first order Markov. The bias error is assumed to be a random constant. If range rate is available from an independent source, then this may be modeled in similar manner.

State Equations

In this section the necessary differential equations representing the time evolution of the state of the system are derived. These equations are necessary for the filter development.

The first equations derived are the line of sight/tracker pointing error equations. The estimate of the pointing error is necessary for tracker control. Any error in this estimate will cause a direct error in the tracker pointing problem. Thus, the estimate must be as accurate as possible. The two coordinate systems to be considered are shown in figure (1). It may be shown that the basic equations of concern for the elevation and azimuth angles are

$$\begin{aligned} \dot{\delta\varepsilon} &= \omega_{T\varepsilon} - \omega_{L\varepsilon} - \omega_{T_r}\delta\eta \\ \dot{\delta\eta} &= \omega_{T_\eta} - \omega_{L_\eta} + \omega_{T_r}\delta\varepsilon \end{aligned} \tag{15}$$

The tracker angular velocity differs from the line of sight angular velocity due to several reasons. The tracker system is of finite bandwidth and, thus, cannot follow high frequency target motions. Inherent tracker disturbances cause inaccuracies in the control of the tracker. Measurements of the error angles are not exact because of inherent noise corruption. The fact that the control system works on an error signal prohibits the system to exactly follow the line of sight. If rate aiding is used, then gyro noise will feed back into the plant causing plant disturbances. The differential equations will be in error for several reasons: the gyro measured angular velocity is corrupted by gyro noise; the line of sight angular velocity is not exactly known; and in case of a tracker looking through a radome one may have radome refraction errors. The radome refraction error perturbs the differential equations (equation (15)) as follows:

$$\begin{aligned} \dot{\delta\varepsilon} &= \omega_{T_\varepsilon} - \omega_{L_\varepsilon} + \frac{\partial\varepsilon}{\partial E}\dot{E} + \frac{\partial\varepsilon}{\partial A}\dot{A} - \omega_{T_r}\delta\eta \\ \dot{\delta\eta} &= \omega_{T_\eta} - \omega_{L_\eta} + \frac{\partial\eta}{\partial E}\dot{E} + \frac{\partial\eta}{\partial A}\dot{A} + \omega_{T_r}\delta\varepsilon \end{aligned} \tag{16}$$

where the terms

$$\frac{\partial(\cdot)}{\partial E}\dot{E} + \frac{\partial(\cdot)}{\partial A}\dot{A} \tag{17}$$

are due to an erroneous line of sight rate due to radome refraction errors. The variables are defined as follows:

E is the elevation look angle

A is the azimuth look angle

$\dot{E}$ is the time derivative of the look angle due to relative motion between the tracker and the aircraft in elevation

$\dot{A}$ is the time derivative of the look angle due to relative motion between the tracker and the aircraft in azimuth

$\partial(\cdot)/\partial E$ and $\partial(\cdot)/\partial A$ are radome error slopes

The perturbational equations with the inclusion of all error sources are given in the next section.

The next equations are those representing the time evolution of the angular velocity of the line of sight. The angular velocity of the line of sight is necessary to rate aid the tracker system. Also, the estimate is required to establish the target velocity, and because the pointing error equations are coupled to this estimate. Thus, in order to obtain an accurate pointing error estimate one must have an accurate estimate of the line of sight angular velocity. It may be shown that the required equation for $\omega_{L\varepsilon}$ is

$$\dot{\omega}_{L\varepsilon} = \frac{1}{r}\{a_{z_{rel}} - \delta\varepsilon\, a_{y_{rel}} - 2\dot{r}\,\omega_{L_\varepsilon}\} - \omega_{Tr}\,\omega_{L\eta} \tag{18}$$

with a similar equation for ω_{L_η}. If one assumes $\delta\varepsilon, \delta\eta \to 0$, then this equation becomes

$$\dot{\omega}_{L_\varepsilon} = \frac{1}{r}\{a_{z_{rel}} - 2\dot{r}\omega_{L_\varepsilon}\} - \omega_{T_r}\,\omega_{L_\eta} \tag{19}$$

Thus, the components of the acceleration equation in the y direction can be written as

$$\dot{V}_{TARGET/Ty} = a_{TARGETy} - a_{Ty} - 2\,(\omega_{T_\eta} V_{xTARGET} - \omega_{T_\varepsilon} V_{zTARGET})$$

$$-(\omega_{T_r}\omega_{T_\eta} r_z - \omega_{T_\eta}^2\, r_y - \omega_{T_\varepsilon}^2\, r_y + \omega_{T_\varepsilon}\omega_{T_r} r_x) - (\dot{\omega}_{T_\eta} r_x - \dot{\omega}_{T_\varepsilon} r_z) \tag{20}$$

with similar equations in the x and z directions.
If the approximation that $\delta\varepsilon = \delta\eta = 0$ is made then

$$\dot{V}_r = a_{yTARGET} - a_{yT} + (\omega_{T_\eta}^2 + \omega_{T_\varepsilon}^2)\, r. \tag{21}$$

The reason that the term

$$-2\,(\omega_{T_\eta} V_{xTARGET} - \omega_{T_\varepsilon} V_{zTARGET}) \tag{22}$$

drops out of the equations is due to the fact that if the y-axis is always pointing at the target, the only component of velocity seen by this coordinate system is along the y-axis, i.e., the range rate.

The target acceleration relative to inertial space is necessary in order to obtain estimates of the angular velocity and angular acceleration of the line of sight, the range, and the range rate. However, target acceleration is not directly measureable. Thus, estimates of the target acceleration may be obtained only by physical insight into the problem and by the use of advanced estimation techniques.

The method is that proposed by Singer, et al, whereby the target acceleration is assumed modeled by exponentially correlated noise. The theory behind this approach is that the target has the capability to maneuver between two acceleration limits. For example, in a dogfight situation the upper acceleration limits may be $\pm$ 8 g's. There is an assumed probability of P_{max} that the target will accelerate at the maximum g capability. Similarly, there may be a finite probability that the target will not maneuver. This could be reflected as P_0. The assumption is made that between the limits of maximum acceleration that the probability density is uniformly distributed with a discrete high probability of no maneuver. This theory is somewhat developed in reference [4]. In order to relate the theory to the particular problem at hand one must consider the environment in which the tracking is to take place. The air to air fire control problem in many instances will relate to a target which is aware of his environment. Thus, it is most likely that he will be operating with a near zero probability of no maneuver and with a near zero probability of obtaining his maximum acceleration limits. He will use an acceleration near his maximum with high probability where the maximum acceleration will be denoted as A_{max}. As an example, the probability density function will be assumed to be as in Figure 4.

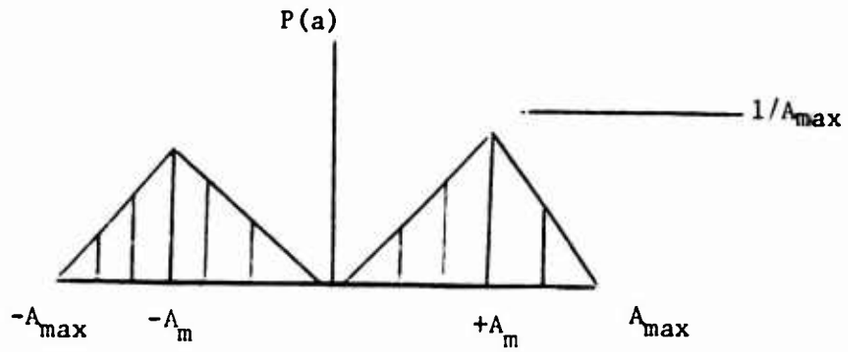


Figure 4: Maneuver Probability Density Function

This differs greatly than that suggested in reference [4]. The acceleration A_m is the acceleration at which the vehicle may maneuver frequently. This may be taken to be 3-4 g's. This allows the vehicle to have maneuverability without excessive velocity bleedoff.

The shaping filter defining the acceleration becomes

$$\dot{a}_{TARGET} = -\beta a_{TARGET} + \sqrt{2\beta}\ \sigma u \qquad (23)$$

where u is zero mean white noise with unity variance. Singer [4] suggests that an approximate value for the correlation constant $\beta = \frac{1}{20}$ for a vehicle undergoing evasive maneuvers. This implies that the correlation between acceleration levels is small and, thus, the target is maneuvering rapidly.

III. Truth State Equations

The variational equations are derived in reference [12]. An assumption as to the error sources corrupting the roll rate gyro will be made in order to limit the complexity of the filter equations. The error sources corrupting the roll rate gyro will be given as

$$\delta\omega_{Tr} = \epsilon_{gr} + b_{gs}F_r + [\omega C_y^T \omega_{TT_t}]. \tag{24}$$

The assumption implies that the acceleration nonlinearities are negligible. The addition of the required bias states to study this problem would increase the state dimensionality from seventy-four to eighty-three. As the errors multiply small terms (i.e., the errors in angle misalignment and errors in angular rates of the line of sight) the approximation is reasonable. The system matrix is as shown in figure 5. The pertinent system submatrices are shown in Appendix A.

IV. Applications

The truth model generated within was used to study two important weapon delivery problem. The first problem was that of a radar directed fire control problem similar to that of the F-4E. The results are given in reference [12]. The second problem was that of the augmented tracker system which was an infrared system for fire control. The results are given in reference [13].

	LINE OF SIGHT ANGULAR VELOCITY		RANGE/RANGE RATE		ANGULAR MISALIGNMENT		TARGET ACCELERATIONS	ANGLE TRACK SCINTILLATION	RANGE SCINTILLATION	GYRO DRIFT	ACCELEROMETER DRIFT	ANGLE MEASUREMENT SCALE FACTOR	GYRO MASS UNBALANCE	GYRO ANISOELASTICITY	ACCELEROMETER SCALE FACTOR	GYRO MISALIGNMENTS	ACCELEROMETER MISALIGNMENTS	ACCELEROMETER [illegible]	RANGE BIAS	ANGLE TRACK BIAS	GYRO SCALE FA...
$\dot{\theta}_{LOS}$ / $\dot{\epsilon}_{LOS}$	F_{11}		F_{12}		F_{13}		F_{14}	0	0	F_{15}	F_{16}	0	F_{17}	F_{18}	F_{19}	F_{110}	F_{111}	F_{112}	0	0	F_{113}
$\dot{R}$ / $\dot{r}$	0		F_{22}		0		F_{24}	0	0	F_{25}	F_{26}	0	F_{27}	F_{28}	F_{29}	F_{210}	F_{211}	F_{212}	0	0	F_{213}
$\dot{\alpha}_{LOS}$ / $\dot{\alpha}_{LOS}$	-I		0		F_{33}		0		0	F_{35}	0	0	F_{37}	F_{38}	0	F_{310}	0	0	0	0	F_{311}
$\dot{a}_{TAR}$	0	0	0	0	0	0	F_{44}	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\dot{ss}$	0	0	0	0	0	0	0	F_{55}	0	0	0	0	0	0	0	0	0	0	0	0	0
$\dot{ss}_r$	0	0	0	0	0	0	0	0	F_{66}	0	0	0	0	0	0	0	0	0	0	0	0
$\dot{c}_y$	0	0	0	0	0	0	0	0	0	F_{77}	0	0	0	0	0	0	0	0	0	0	0
$\dot{c}_a$	0	0	0	0	0	0	0	0	0	0	F_{88}	0	0	0	0	0	0	0	0	0	0
$\dot{s}_{SF}$	0	0	0	0	0	0	0	0	0	0	0	F_{99}	0	0	0	0	0	0	0	0	0
$\dot{b}_{R_{II}}$																					
$\dot{b}_{R_A}$																					
$\dot{b}_{AS}$																					
$\dot{b}_{ASA}$																					
$\dot{b}_{ASSS}$																					
$\dot{b}_r$																					
$\dot{b}_{AT}$																					
$\dot{b}_{RSF}$																					

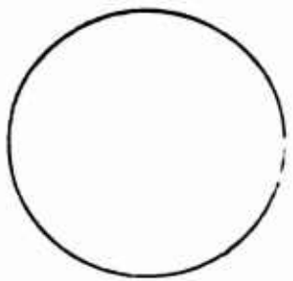


Figure 5: System Matrix

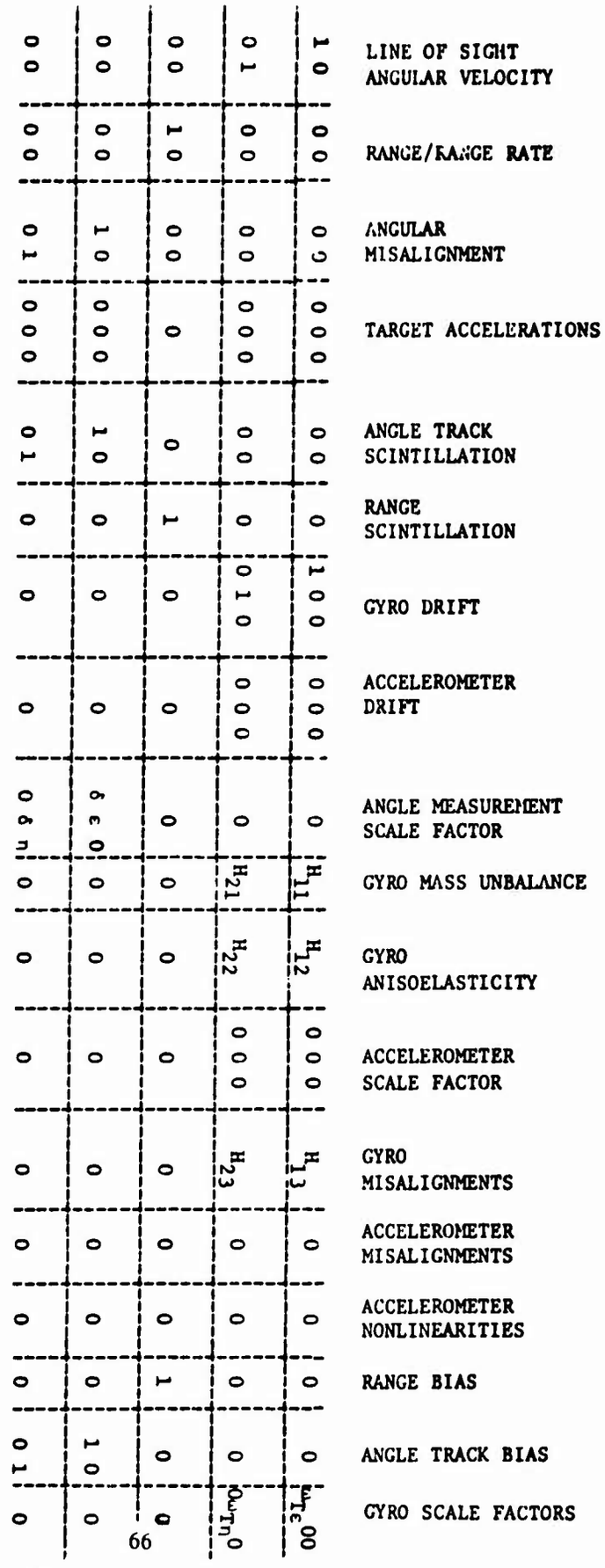


Figure 6: Measurement Matrix

Appendix A

$$F_{11} = \begin{Bmatrix} -2\dot{r}/r & -\omega_{T_r} \\ \omega_{T_r} & -2\dot{r}/r \end{Bmatrix} \qquad F_{12} = \begin{Bmatrix} f_{11} & -2\omega_{L_\varepsilon}/r \\ f_{21} & -2\omega_{L_\eta}/r \end{Bmatrix}$$

where

$$f_{11} = -\frac{1}{r^2}\left(a_{z_{TARGET}} - a_{z_T} - \delta\varepsilon(a_{y_{TARGET}} - a_{y_T}) - 2\dot{r}\omega_{L\varepsilon}\right)$$

$$f_{21} = -\frac{1}{r^2}\left(\delta\eta(a_{y_{TARGET}} - a_{y_T}) - a_{x_{TARGET}} + a_{x_T} - 2\dot{r}\omega_{L\eta}\right)$$

$$F_{13} = \begin{Bmatrix} f_{11} & 0 \\ 0 & f_{22} \end{Bmatrix}$$

$$f_{11} = -\frac{1}{r}\left(a_{y_{TARGET}} - a_{y_T}\right)$$

$$f_{22} = \frac{1}{r}\left(a_{y_{TARGET}} - a_{y_T}\right)$$

$$F_{14} = \begin{Bmatrix} 0 & \frac{-\delta\varepsilon}{r} & \frac{1}{r} \\ \frac{-1}{r} & \frac{\delta\eta}{r} & 0 \end{Bmatrix}$$

$$F_{15} = \begin{Bmatrix} 0 & 0 & -\omega_{L_\eta} \\ 0 & 0 & \omega_{L_\varepsilon} \end{Bmatrix}$$

$$F_{16} = \begin{Bmatrix} 0 & \frac{\delta\varepsilon}{r} & \frac{-1}{r} \\ \frac{1}{r} & \frac{-\delta\eta}{r} & 0 \end{Bmatrix}$$

$$F_{22} = \begin{Bmatrix} 0 & 1 \\ f_{21} & 0 \end{Bmatrix}$$

where

$$f_{21} = \omega_{T_\varepsilon}^2 + \omega_{T_\eta}^2$$

$$F_{18} = \left\{\begin{array}{ccc|ccc|ccc} 0 & 0 & -\omega_{L_\eta} a_x^2 & 0 & 0 & -\omega_{L_\eta} a_y^2 & 0 & 0 & -\omega_{L_\eta} a_z^2 \\ 0 & 0 & \omega_{L_\varepsilon} a_x^2 & 0 & 0 & \omega_{L_\varepsilon} a_y^2 & 0 & 0 & \omega_{L_\varepsilon} a_z^2 \\ 0 & 0 & -\omega_{L_\eta} a_x a_y & 0 & 0 & -\omega_{L_\eta} a_x a_z & 0 & 0 & -\omega_{L_\eta} a_y a_z \\ 0 & 0 & \omega_{L_\varepsilon} a_x a_y & 0 & 0 & \omega_{L_\varepsilon} a_x a_z & 0 & 0 & \omega_{L_\varepsilon} a_y a_z \end{array}\right\}$$

$$F_{28} = \left\{\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 2r\omega_{T_\varepsilon} a_x^2 & 2r\omega_{T_\eta} a_x^2 & 0 & 2r\omega_{T_\varepsilon} a_y^2 & 2r\omega_{T_\eta} a_y^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 2r\omega_{T_\varepsilon} a_z^2 & 2r\omega_{T_\eta} a_z^2 & 0 & 2r\omega_{T_\varepsilon} a_x a_y & 2r\omega_{T_\eta} a_x a_y & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 2r\omega_{T_\varepsilon} a_x a_z & 2r\omega_{T_\eta} a_x a_z & 0 & 2r\omega_{T_\varepsilon} a_y a_z & 2r\omega_{T_\eta} a_y a_z & 0 \end{array}\right\}$$

$$F_{38} = \left\{\begin{array}{ccc|ccc|ccc} a_x^2 & 0 & -\delta\eta a_x^2 & a_y^2 & 0 & -\delta\eta a_y^2 & a_z^2 & 0 & -\delta\eta a_z^2 \\ 0 & a_x^2 & \delta\varepsilon a_x^2 & 0 & a_y^2 & \delta\varepsilon a_y^2 & 0 & a_z^2 & \delta\varepsilon a_z^2 \\ a_x a_y & 0 & -\delta\eta a_x a_y & a_x a_z & 0 & -\delta\eta a_x a_z & & & \\ 0 & a_x a_y & \delta\varepsilon a_x a_y & 0 & a_x a_z & \delta\varepsilon a_x a_z & & & \\ a_y a_z & 0 & -\delta\eta a_y a_z & & & & & & \\ 0 & a_y a_z & \delta\varepsilon a_y a_z & & & & & & \end{array}\right\}$$

$$F_{24} = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{Bmatrix}$$

$$F_{25} = \begin{Bmatrix} 0 & 0 & 0 \\ 2r\omega_{T_\varepsilon} & 2r\omega_{T_\eta} & 0 \end{Bmatrix}$$

$$F_{26} = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \end{Bmatrix}$$

$$F_{33} = \begin{Bmatrix} 0 & -\omega_{T_r} \\ \omega_{T_r} & 0 \end{Bmatrix}$$

$$F_{35} = \begin{Bmatrix} 1 & 0 & -\delta\eta \\ 0 & 1 & \delta\varepsilon \end{Bmatrix}$$

$$F_{44} = \begin{bmatrix} -\beta_x & \omega_{T\eta} & -\omega_{Tr} \\ -\omega_{T\eta} & -\beta_y & \omega_{T\varepsilon} \\ \omega_{Tr} & -\omega_{T\varepsilon} & -\beta_z \end{bmatrix}$$

$$F_{17} = \begin{Bmatrix} 0 & 0 & -\omega_{L_\eta} a_x & 0 & 0 & -\omega_{L_\eta} a_y & 0 & 0 & -\omega_{L_\eta} a_z \\ 0 & 0 & \omega_{L_\varepsilon} a_x & 0 & 0 & \omega_{L_\varepsilon} a_y & 0 & 0 & \omega_{L_\varepsilon} a_z \end{Bmatrix}$$

$$F_{27} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 2r\omega_{T_\varepsilon} a_x & 2r\omega_{T_\eta} a_x & 0 & 2r\omega_{T_\varepsilon} a_y & 2r\omega_{T_\eta} a_y & 0 \\ 0 & 0 & 0 & & & \\ 2r\omega_{T_\varepsilon} a_z & 2r\omega_{T_\eta} a_z & 0 & & & \end{Bmatrix}$$

$$F_{37} = \begin{Bmatrix} a_x & 0 & -\delta\eta a_x & a_y & 0 & -\delta\eta a_y & a_z & 0 & -\delta\eta a_z \\ 0 & a_x & \delta\varepsilon a_x & 0 & a_y & \delta\varepsilon a_y & 0 & a_z & \delta\varepsilon a_z \end{Bmatrix}$$

$$F_{19} = \begin{pmatrix} 0 & \frac{\delta\epsilon}{r} a_y & \frac{-1}{r} a_z \\ \frac{1}{r} a_x & \frac{-\delta\eta}{r} a_y & 0 \end{pmatrix} \qquad F_{29} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -a_y & 0 \end{pmatrix}$$

$$F_{110} = \begin{pmatrix} 0 & 0 & 0 & 0 & \omega_{L_\eta}\omega_{T_\epsilon} & -\omega_{L_\eta}\omega_{T_\eta} \\ 0 & 0 & 0 & 0 & -\omega_{L_\epsilon}\omega_{T_\epsilon} & \omega_{T_\eta}\omega_{L_\epsilon} \end{pmatrix}$$

$$F_{210} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -2r\omega_{T_\epsilon}\omega_{T_\eta} & 2r\omega_{T_\epsilon}\omega_{T_r} & 2r\omega_{T_\eta}\omega_{T_\epsilon} & -2r\omega_{T_\eta}\omega_{T_r} & 0 & 0 \end{pmatrix}$$

$$F_{310} = \begin{pmatrix} -\omega_{T_\eta} & \omega_{T_r} & 0 & 0 & \delta\eta\omega_{T_\epsilon} & -\delta\eta\omega_{T_\eta} \\ 0 & 0 & \omega_{T_\epsilon} & -\omega_{T_r} & -\delta\epsilon\omega_{T_\epsilon} & \delta\epsilon\omega_{T_\eta} \end{pmatrix}$$

$$F_{111} = \begin{pmatrix} 0 & 0 & \frac{\delta\epsilon}{r} a_x & \frac{-\delta\epsilon}{r} a_z & \frac{1}{r} a_x & -\frac{1}{r} a_y \\ -\frac{1}{r} a_y & \frac{1}{r} a_z & \frac{-\delta\eta}{r} a_x & \frac{\delta\eta}{r} a_z & 0 & 0 \end{pmatrix}$$

$$F_{55} = \begin{pmatrix} -\beta_4 & 0 \\ 0 & -\beta_5 \end{pmatrix} \qquad F_{66} = \begin{pmatrix} -\beta_6 \end{pmatrix}$$

$$F_{77} = \begin{pmatrix} -\beta_7 & 0 & 0 \\ 0 & -\beta_8 & 0 \\ 0 & 0 & -\beta_9 \end{pmatrix} \qquad F_{88} = \begin{pmatrix} -\beta_{10} & 0 & 0 \\ 0 & -\beta_{11} & 0 \\ 0 & 0 & -\beta_{12} \end{pmatrix}$$

$$F_{99} = \begin{pmatrix} -\beta_{13} & 0 \\ 0 & -\beta_{14} \end{pmatrix} \qquad F_{113} = \begin{pmatrix} 0 & 0 & -\omega_{L_\eta}\omega_{T_r} \\ 0 & 0 & \omega_{L_\varepsilon}\omega_{T_r} \end{pmatrix}$$

$$F_{213} = \begin{pmatrix} 0 & 0 & 0 \\ 2r\omega_{T_\varepsilon}^2 & 2r\omega_{T_\eta}^2 & 0 \end{pmatrix} \qquad F_{211} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_x & a_z & 0 & 0 \end{pmatrix}$$

$$F_{112} = \begin{pmatrix} 0 & 0 & \frac{\delta\varepsilon}{r}a_y^2 & \frac{\delta\varepsilon}{r}a_y^3 & -\frac{1}{r}a_z^2 & -\frac{1}{r}a_z^3 \\ \frac{1}{r}a_x^2 & \frac{1}{r}a_x^3 & -\frac{\delta\eta}{r}a_y^2 & -\frac{\delta\eta}{r}a_y^3 & 0 & 0 \end{pmatrix}$$

$$F_{212} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_y^2 & -a_y^3 & 0 & 0 \end{pmatrix}$$

$$F_{311} = \begin{pmatrix} \omega_{T_\varepsilon} & 0 & -\delta\eta\omega_{T_r} \\ 0 & \omega_{T_\eta} & \delta\varepsilon\omega_{T_r} \end{pmatrix}$$

The measurement matrix is given as in figure 7 where the submatrices are

$$H_{11} = \begin{bmatrix} a_x & a_y & a_z & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$H_{21} = \begin{bmatrix} 0 & 0 & 0 & a_x & a_y & a_z & 0 & 0 & 0 \end{bmatrix}$$

$$H_{12} = \begin{bmatrix} a_xa_y & a_ya_z & a_xa_z & a_x^2 & a_y^2 & a_z^2 & 0 \cdots 0 \end{bmatrix}$$

$$H_{22} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & a_xa_y & a_ya_z & a_xa_z \end{bmatrix}$$

$$H_{13} = \{-w_{Tr} \quad w_{T\eta} \quad 0 \quad 0 \quad 0 \quad 0\}$$

$$\{a_x^2 \quad a_y^2 \quad a_z^2 \quad 0 \cdots 0\ ]$$

$$H_{23} = \{0 \quad 0 \quad w_{T\varepsilon} \quad -w_{T\eta} \quad 0 \quad 0\}$$

The 0-matrix is the identity matrix. The G matrix is shown in figure 8. The R matrix is given in figure 9.

References

1. Asher, Robert B. Precision Pointing and Tracking for Weapon Delivery Interim Report. Air Force Avionics Laboratory (AFAL) Wright-Patterson AFB, Ohio, June 1972.

2. Asher, Robert B. Precision Pointing and Tracking for Weapon Delivery Interim Report. AFAL, Wright-Patterson AFB, Ohio, July 1972.

3. Britting, Kenneth R. Inertial Navigation Systems Analysis. New York: John Wiley & Sons, Inc. (1971).

4. Singer, R. A. "Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets". IEEE Transactions on AES, Vol AES-6, No. 4 (July 1970).

5. Landau, M. I. "Angle Tracking Input to Multimode Guidance Study". Hughes Aircraft Co. Memo.

6. Fitts, J. M. "HAT Aided Tracking Analysis". Hughes Aircraft Co. Report, pp. 71-215 (June 1971).

7. Fitts, J. M. "A Strapdown Approach to HAT Aided Tracking". Hughes Aircraft Co. Report, pp. 71-351 (September 1971).

8. Fitts, J. M. "HAT Position and Rate Aided Tracking Analysis". Hughes Aircraft Co. Report, pp. 71-471 (December 1971).

9. Fitts, J. M. "Aided Tracking versus Applied to High Accuracy Pointing Systems". IEEE Transactions on AES, Vol AES-9, No. 3.

10. Pearson, J. B. "Basic Studies in Airborne Radar Tracking Systems". (Ph.D. dissertation, UCLA, 1970).

11. Asher, R. B. and R. M. Reeves. "Performance Evaluation of Suboptimal Filters". AFAL/NVS-1 Technical Note (January 1973).

12. Asher, R. B. and D. H. Watjen. "Kalman Filtering for Precision Pointing and Tracking Applications". AFAL TR (Unpublished).

13. Watjen, D. H. and R. B. Asher. "Covariance Analysis of a Proposed Kalman Filter for an Airborne Tracking System". AFAL/NVE-1 Technical Note (Unpublished).

The Generation of Air-to-Air Missile Launch Envelopes

by

J. Hledik
and

J. C. Shepherd
Naval Weapons Center
China Lake, California

1. INTRODUCTION

During the complete life cycle of an air-to-air weapon from inception through use by the services, one important measure of system performance is launch range capability. Launch range capability is normally best described by the launch envelope, that is, the area surrounding a launch aircraft within which a missile may guide to the target with an acceptable miss distance.

During the initial, or feasibility, phase of a new system design the launch envelopes are generated and compared with a given theoretical ideal. At this point the engineer is primarily concerned with determining whether or not it is feasible to design and build a system which has a pre-determined desired or ideal range capability. Generally the system model during this period is relatively simple and borrows heavily from information gathered from existing systems. This is also the time when a very critical comparison should be made with the performance of existing systems to determine the desirability of proceeding with the design. Launch envelopes can be and are a valuable aid to this decision making process.

Once the system has been designed, a continuing need exists throughout the life cycle of the weapon for critical evaluation of its performance. This should include not only the evaluation of such items as reliability etc., but also the capability of the system to meet known or hypothesized enemy threats. For this use the generation of launch envelopes generally incorporates a much more sophisticated system model. A wider range of input parameters describing the tactical relationship between the target and launch aircraft is necessary to completely describe the performance. The result is a requirement for the creation of very large numbers of launch envelopes.

Another use for a version of the launch envelope occurs during the design and development testing of a new system. At this time, the launch envelope, now with possible missile component failures included as a parameter, will be used not only to determine desirable test conditions, but also by the range safety engineer as a criterion for the safety of all personnel and real estate involved in the actual air firing tests.

The ultimate use, however, of a launch envelope is to describe the system capability to the user, that is, the man whose life may depend upon the system mounted on his aircraft. At this point the launch envelope is used to develop the system firing doctrine, or tactics, and

to establish the decision criteria such as minimum-maximum range, which must be incorporated within the fire control computer. The transition from the laboratory generated launch envelopes to something which is usable in combat, probably presents one the larger problems in the total weapon system design. The launch zone capability of any given system is dependent upon a large number of parameters including launch aircraft speed, altitude and attitude, missile seeker capability, target size, speed and altitude, etc. As a result, an extremely large amount of data is necessary to completely describe the launch range capability of a weapon system. However, such a large amount of data cannot possibly be assimilated and remembered by a pilot, especially under a tense combat situation. The solution to this problem is not an easy task but it is really not the subject of this paper.

2.0 METHODOLOGY

During the design and development of such air-to-air weapon systems as the SIDEWINDER, AGILE, and SPARROW AIM-7F as well as various air-to-ground missiles as the SHRIKE, CONDOR, and others, the Naval Weapons Center (NWC), has been heavily involved in the generation of missile system launch envelopes. The methods used for generating these envelopes have evolved from very elementary techniques using slow computers with manual controls to the currently used semi-automatic methods utilizing large, general purpose high speed computing equipment under digital control.

2.1 GENERAL REQUIREMENTS

Regardless of the system under investigation, the generation of launch envelopes to describe system performance, requires the same basic information. First a system model must be developed and verified. The model may vary in complexity from a two dimensional mathematical simulation with linearized parameters, to a complete three dimensional, six degree-of-freedom representation including all the non-linear parameters which can be mathematically represented.

Once a model has been constructed, the criteria which are to be used to determine the envelope must be established. These criteria involve both initial and terminal conditions which must be satisfied in order for a missile to successfully intercept a target. Included are miss distance, closing velocity, seeker acquisition range, signal-to-noise ratio, and launcher and launch aircraft characteristics.

Finally, the target characteristics must be provided to the launch envelope generator. For the case of the air-to-air envelopes, these parameters are primarily speed, maneuverability, altitude, expected aspect angle and target range at launch. However, in some applications target size and signature characteristics must be added.

Once the model has been derived and verified, the criteria for success or failure determined, and the initial launch and target conditions provided as inputs, the generation of the launch envelopes involves exercising the model with a systematic variation of all possible combinations of input parameters. For any given set of tactical parameters describing the target-launch relationship as discussed above, the range at which a failure occurs must be determined. For each range one complete solution of the model equations as a function of time must be obtained. As each solution of the terminal conditions is obtained, it is compared against the success criteria. If this comparison shows the flight to be a success the range must be incremented and another solution obtained. The procedure is repeated until a failure occurs. At this time, one point on a launch envelope is determined for one set of input conditions.

It is immediately obvious that the generation of launch envelopes requires the aid of relatively large scale computing devices. At the Naval Weapons Center both digital and analog computers have been utilized for this purpose. However, because of the large number of solutions required, the complexity of most models, the subsequent high cost, and the long time involved, it has been found that launch envelopes are most efficiently generated using a high speed analog computer, operated in the repetitive operation mode. Repetitive operation means that the computer is automatically cycled through its computational modes of operate (problem solution), hold and reset. Initial parameter conditions can be automatically set or changed during the reset mode so that a logical sequence of runs is obtained without the use of manual control.

The analog computers which are in use at the Naval Weapons Center are all electronic with solid-state components, and can be operated in the repetitive mode at speeds such that a 10 second missile flight can be simulated in 1/10 second. That is 100 times faster than real time. Time scaling of the computer is automatic, and the solution rate can be switched from real time to 10 or 100 times real time with a simple selector switch. The particular equipment available consists of six consoles of Electronics Associates Incorporated, Model 7800, general purpose analog computers. Each console has 108 operational amplifiers, 132 attenuators, 36 multipliers, 6 double input resolvers, plus a compliment of switches, limiters, function generators and other analog devices. All analog devices are terminated at the main analog patch panel. An extensive compliment of patchable logic is also available through a separate patch panel. Manual controls and readout devices for all components are also available for computer control, bookkeeping or mathematical calculations as needed. The laboratory is arranged so that each console may be operated independently or with any or all of the other consoles. Using the SIDEWINDER missile as an example, a single console of this equipment is adequate for a single plane, three degree-of-freedom simulation. A six degree-of-freedom simulation generally requires two consoles, while a more complex missile system such as the SPARROW AIM-7F requires three consoles, unless considerable simplification is made in the modeling.

The model most commonly used for the generation of launch envelopes is a single plane three degree-of-freedom model requiring minimum computer equipment. The model is programmed for computer solution using standard

analog computer techniques. However, since the solution time will occur at 100 times real time, extreme care must be taken in modeling all time or frequency dependent functions (such as transfer functions of servo loops) so that their characteristics will be automatically time scaled for the desired high speed operation.

The number of individual trajectories needed to describe one launch envelope dictates the use of not only the high speed solution capability, but also some form of automation. With the use of the programmable logic components, or the available digital computers, the necessary control functions to automatically generate a complete launch envelope can be pre-programmed. All criteria for a successful flight, as well as the desired flight parameter variation are previously stored within the computer. At the end of any given "run" or simulated flight, the terminal conditions are compared with the success criteria. If the flight was successful, the launch range is incremented and another trajectory solution obtained. This process is automatically repeated until a failure occurs which determines the launch boundary limit point. The particular initial range which resulted in the flight that failed is stored, and a new launch aspect angle is introduced and the entire process is repeated. When the full range of aspect angles about the launch point has been exhausted, the stored information is sufficient to complete a launch envelope such as seen in Figure 1. At this time new envelope parameters such as missile launch speed, target maneuver, etc., can be either manually or automatically inputed and a new launch envelope generated.

Using this process, a combination of high speed solution and automation, one entire launch envelope can be generated in approximately 3 to 5 minutes. One of the most significant advantages of this technique is that a much more extensive investigation of the envelope is possible. With the much slower simulation methods used previously, only points near the boundary could be examined because of the long time and high cost involved. Holes in the envelope (i.e. areas within an envelope boundary where the system will be unsuccessful) due to unexpected characteristics of the system are now immediately obvious, where before they were either not completely described or missed entirely. Finally a much larger variety of conditions including target speed and altitude, launch speed, and target maneuver can be examined, where as only a very restricted number of conditions were possible with slower techniques.

2.2 ADVANCED REQUIREMENTS

Launch envelopes as described above have always been the traditional method of examining system performance and for the comparison of competitive systems. However, to adequately evaluate a system, a very large number of envelopes must be generated in an attempt to investigate all possible target engagements as was described above. Also, if one wishes to compare various systems, launch envelopes for identical parameters must be generated for each system under investigation. To overcome this problem and again in an attempt to provide economy of funding and time, Monte Carlo methods have recently begun to be used at NWC.

The Monte Carlo method is a statistical approach of randomly introducing the various engagement parameters to the simulation model according to a predetermined probability distribution. With this random distribution of initial conditions a large sample of simulated flights are made. In each case, success or failure is recorded. This stored information is sufficient to determine the necessary statistics to assign a probability of success of the system at a given condition and also the confidence level associated with that probability. This method becomes practical only with the use of modern high speed analog computers and special purpose mini-digital computers. The advantages of the Monte Carlo method are at once obvious. Probabilities of successful guidance for different systems in identical environments may be generated for an almost instant comparison of systems effectiveness. Every envelope for all feasible engagment parameters need not be generated to describe a system.

The Monte Carlo technique has been used extensively at the NWC for the determination of many of the AGILE system requirements and characteristics. Detailed examination of missile parameters versus performance has thus been possible in great detail.

A variation of the Monte Carlo technique has been applied to an investigation of missile parameters other than launch parameters during an evaluation of the SIDEWINDER missile performance. Expected use of this capability will be a detailed examination of the effects on system performance of the interaction of various subsystem design tolerances.

3.0 CONCLUSIONS

The use of modern high speed computers, repetitive operation, and automatic control has made it possible to reduce the time required to generate launch envelopes from hours down to a few minutes. However, the number of launch envelopes generated to describe a system has not been decreased, in fact in many cases it has increased. The use of a statistical approach, such as Monte Carlo methods, has the possibility as a probabilistic distribution. However, at the Naval Weapons Center, this method needs to be further investigated for generation of launch envelopes.

The problem of presenting the data to the combat pilot in a short, concise and usable form also still remains. It doesn't appear likely that the amount of data to describe system performance will significantly reduced since the number of possible engagement situations is still huge for any given system. However, recent developments in the field of small computers with relatively large capability, make it possible to develop airborne fire-control computers which may be able to assimilate this data and present it in a usable form to the pilot in a real-time situation. This seems even more likely if the original data is of a statistical nature such as that obtained from Monte Carlo methods.

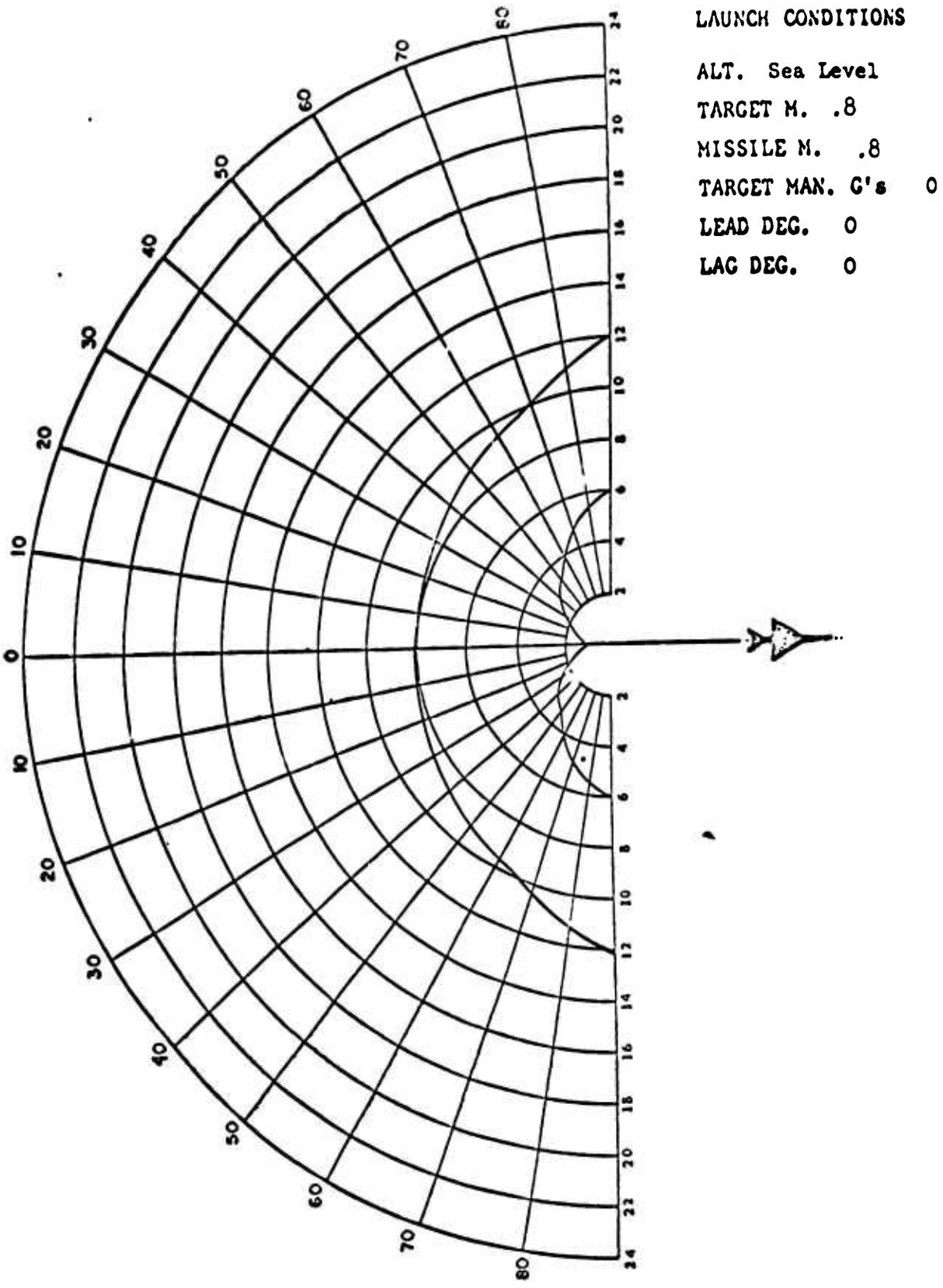


Figure 1. Typical Launch Envelope.

PAPER

on

AIR-TO-AIR MISSILE LAUNCH BOUNDARY FIRE CONTROL TECHNIQUES

11 July 1974

PREPARED BY:

RONALD T. OOTEN
ASD/ENASA

and

CAPTAIN ALBERT I. CHATMON
ASD/EN

AERONAUTICAL SYSTEMS DIVISION
WRIGHT-PATTERSON AFB, OHIO 45433

ABSTRACT

Present fire control systems do not provide the capability to fully utilize the capabilities of existing or projected air-to-air missiles. This deficiency exists because of the inaccuracies present in the missile launch boundary modeling in the fire control computers. The deficiency manifests itself in one of two fashions: the missile is launched outside its envelope, thus missing the target, or conservative "rules-of-thumb" are applied, which result in missed launch opportunities. The launch boundary algorithms used in most air-to-air fire control computers are optimized on some favorable launch conditions, and, hence, become more inaccurate as the encounter deviates from these conditions. Such deviations are the rule, rather than the exception in the air combat maneuvering (ACM) environment. The dynamics of the encounter make it difficult, if not impossible, to attain the "optimal" launch conditions. Three improved launch boundary modeling techniques have been developed which alleviate these problems. These techniques are presented and discussed in this paper.

SUMMARY

Basically, the aircraft weapon delivery (fire control) system is composed of: the pilot, avionics/sensors, flight/engine control systems, aircraft, and armament. Data from Southeast Asia reports, Air Force/Navy air combat maneuver flight tests and evaluation programs, and Air Force/Navy simulation studies show that operational weapon delivery control systems have many deficiencies that can be (are being) corrected via judicial use of modern technology. One of the key problem areas illustrated in the data is the inaccuracy of air-to-air missile (A/AM) launch boundary modeling present in the avionics/sensor subsystem of present weapon delivery systems. This deficiency has resulted in missiles being launched out of conservative "rules-of-thumb," thus missing launch opportunities. This paper addresses three modern techniques that are being used/developed to greatly increase the accuracy.

Actual A/AM launch boundaries are a function of missile capability, fighter/target velocities, fighter/target altitudes, target maneuver at and after launch, fighter/missile separation trajectory, and end-game geometry. Air Force operational A/AM launch modeling equations are based for the best missile launch conditions (lead collision and/or pursuit "down-the-tail-pipe" launch). They become more and more inaccurate as the launch aircraft deviates from the "best launch condition". In an air combat environment obtaining and holding this "best condition" long enough for a launch may be impossible. The three modeling techniques that will be discussed in this paper eliminate the deterioration of the modeling accuracy caused by deviating from the "best launch condition" by incorporating the effects of all the independent variables with the exception of target maneuver after launch and/or end-game geometry.

The first modeling technique, developed by ASD, is an empirical technique based on the Taylor series expansion using partial derivatives. This technique greatly improves the inaccuracy inherent in the Taylor series as one deviates from the baseline by interpolating between stored data points to obtain the single-variable deviation effect.

The second technique, developed at McDonnell-Douglas Corporation, St. Louis, Missouri, is a closed form solution of the intercept problem obtained by first assuming that the missile travels a "straight line" lead collision trajectory to the impact point, dividing the

problem into two stages - boost and glide. The closed form solution is then modified to account for missile capability, altitude effects, target maneuver at launch and fighter/missile separation trajectory.

The third technique, developed at the Naval Missile Center, Pt Mugu, California, is a simplified simulation of the missile. The simplification is obtained by assuming: (1) that the seeker and autopilot subsystem of the missile can be approximated by linear, first order lag transfer functions, and (2) that the inputs into the subsystems are constant over an integration iteration.

This paper presents background information to show limitations of present oprational modeling techniques, and how launch boundary models are obtained; discusses and compares three modern modeling techniques, and presents areas where future improvements are required.

Additional information on the mechanization can be obtained from the working papers on "AIM-7F Maximum Launch Range Equation Development", McDonnell-Douglas Corporation.

Additional information on FAS can be obtained from TP-72-7, "A Simple Analytical Model of an Air-to-Air Guided Missile for Rapid Flight Simulation," 17 February 1972, Naval Missile Center, Point Mugu, California.

AIR-TO-AIR MISSILE LAUNCH BOUNDARY FIRE CONTROL TECHNIQUES

Tactical deployment of current Air Force air-to-air missiles in recent air combat situations has resulted in a major discrepancy between the theoretical or predicted probability of kill and that actually realized in the combat arena. Based upon recent data, it was evident that the missile/fire control system continued to be the major factor affecting unsatisfactory missile firings.

Figure 1 relates the overall control system of a fighter aircraft. The diagram shows the interaction between the flight/engine controls and the aircraft dynamics and the pilot. The pilot then interacts with the avionics which is also influenced by the aircraft dyanmics data and target data to expend the armament to the target. The avionics package consists of several functions, as shown in Figure 2, to include the inertial navigation system, a stores management system, sensors (i.e., radar, electro-optical, infrared, laser, etc.), a lead computing gyro to calculate the lead angle for steering and launching air-to-air missiles, and a central computer to interface with all the functions.

This paper will be directed toward the computational algorithms for solving the fire control problem and computing the missile launch boundaries. The probability of satisfactory missile intercept is affected by several parameters to include fighter and target velocities, fighter and target altitudes, target acceleration, lead/lag angle and aspect angle. As shown in Figure 3, the definition for the aspect angle (τ) is the angle between the velocity vector of the target and the line-of-sight between the fighter and target measured in a counter-clockwise direction from the target velocity vector. For this convention, τ is 0° for the tail aspect and 180° for a head-on attack. The lead/lag angle (λ) is the angle between the line-of-sight of the fighter and target and the velocity vector of the fighter.

Figures 4 through 7 illustrate how the variation of some of the parameters, mentioned above, affect the aerodynamic capability of an air-to-air missile. Figure 4 shows the capability of the missile, as a function of aspect angle. The position of the target is in the center of the coordinate system. The oval-shaped line segment, noted as Rmax, is the maximum range the fighter aircraft can

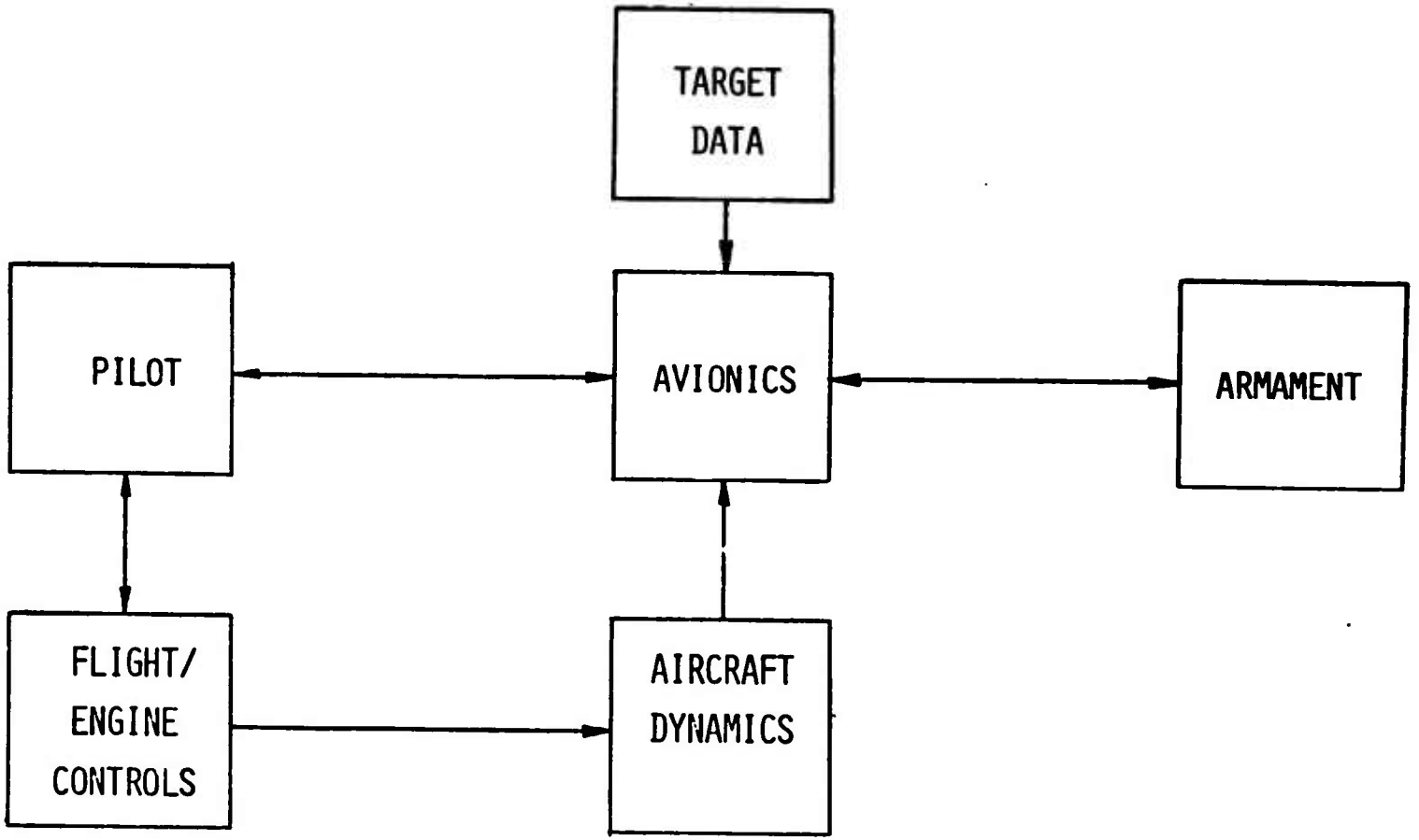


FIGURE 1. GENERAL AIRCRAFT CONTROL SYSTEM

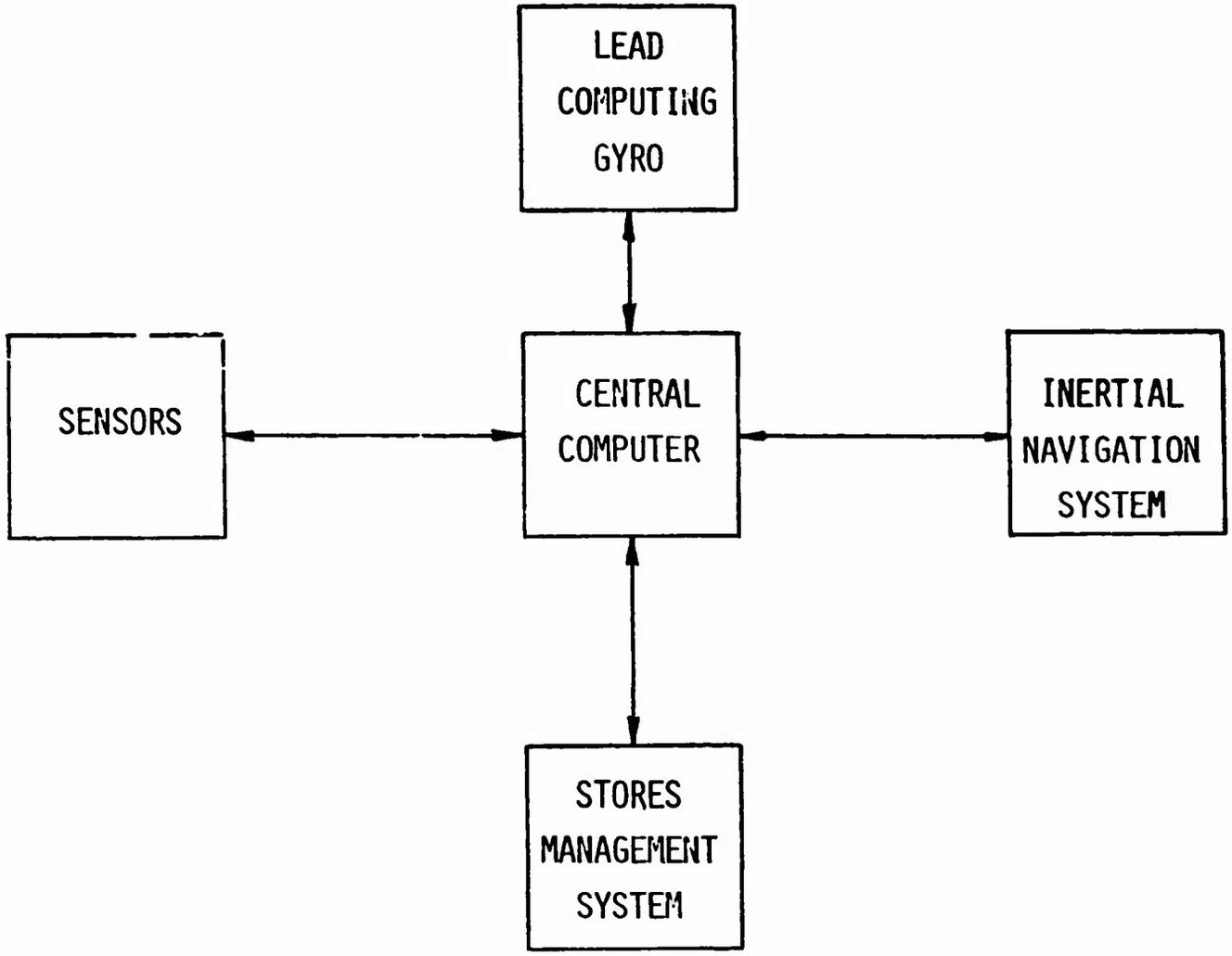


FIGURE 2. AVIONICS FIRE CONTROL SYSTEM

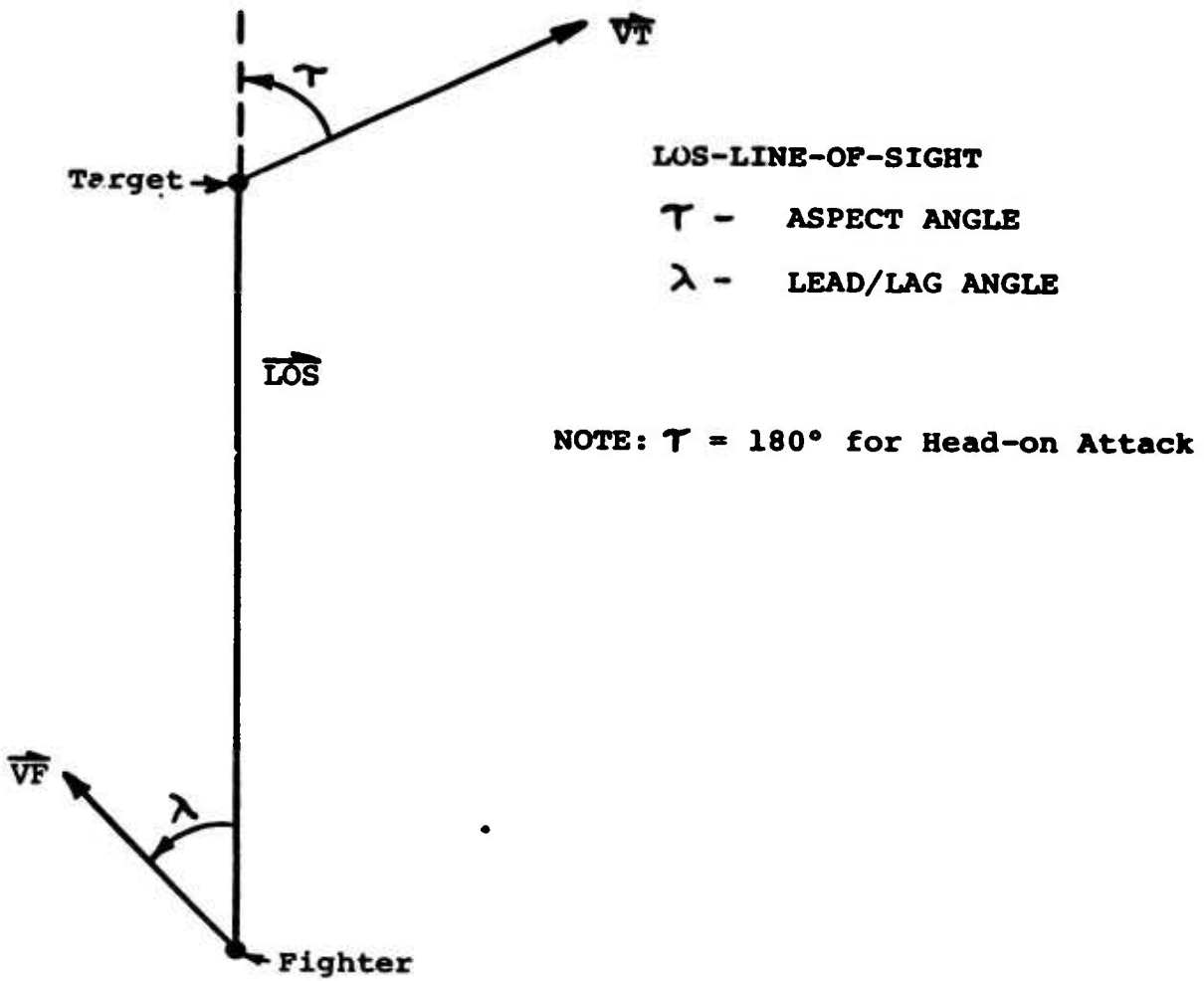


FIGURE 3. ANGLE DEFINITIONS

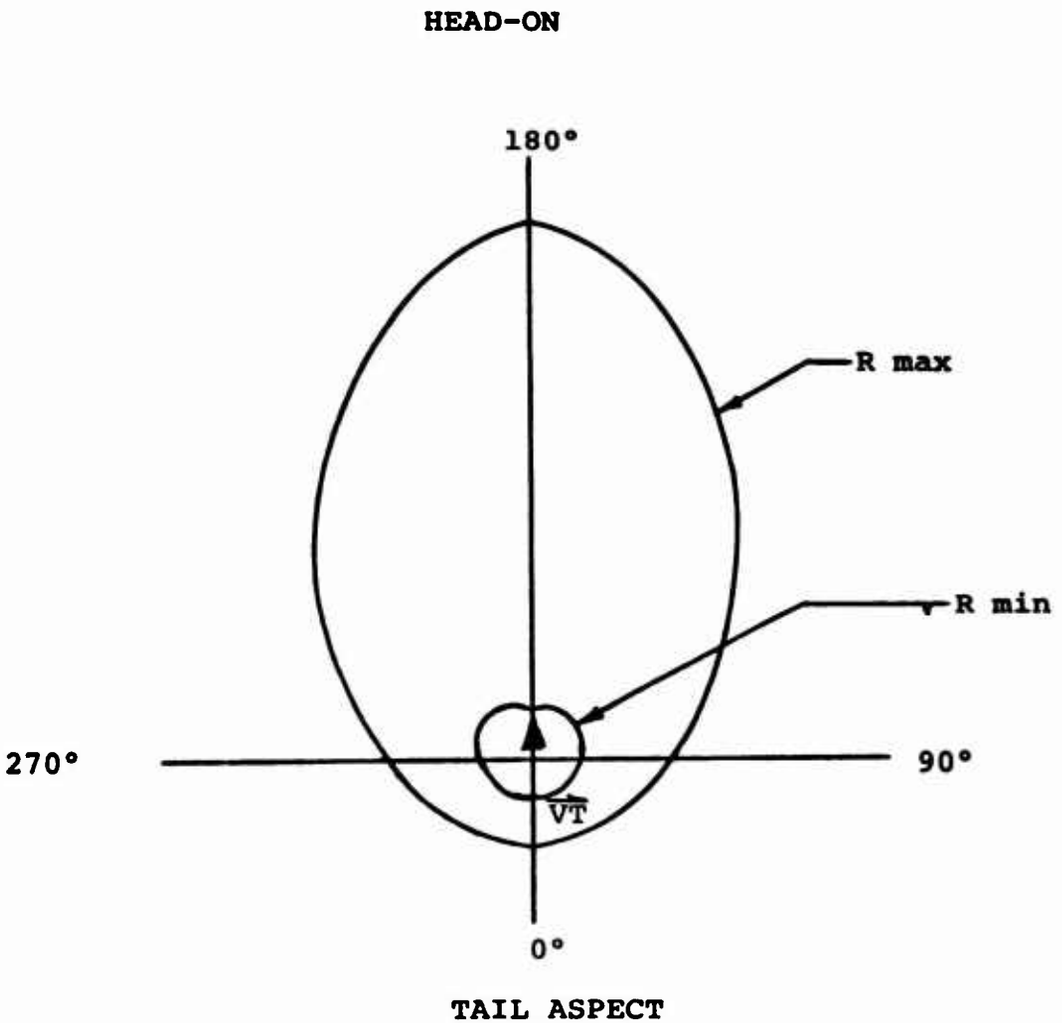
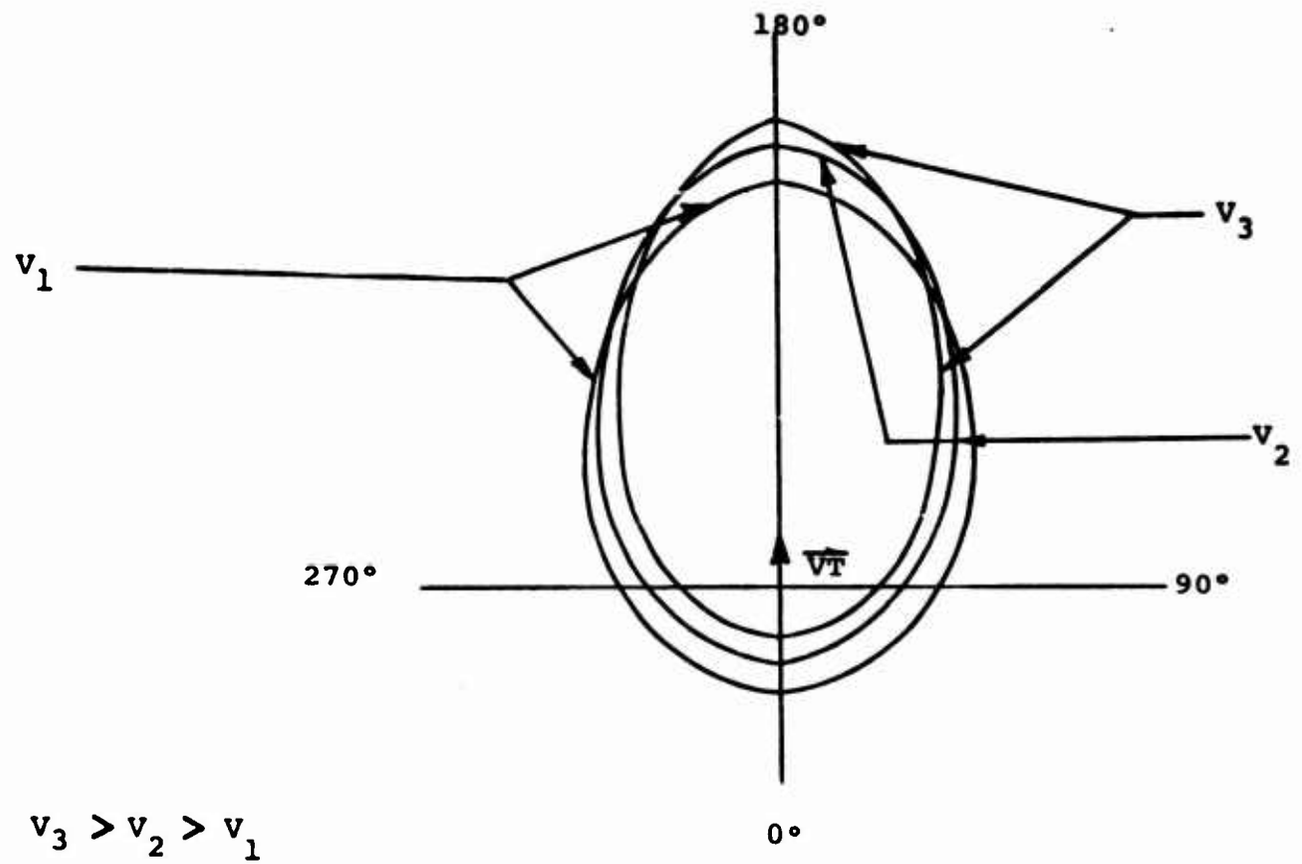


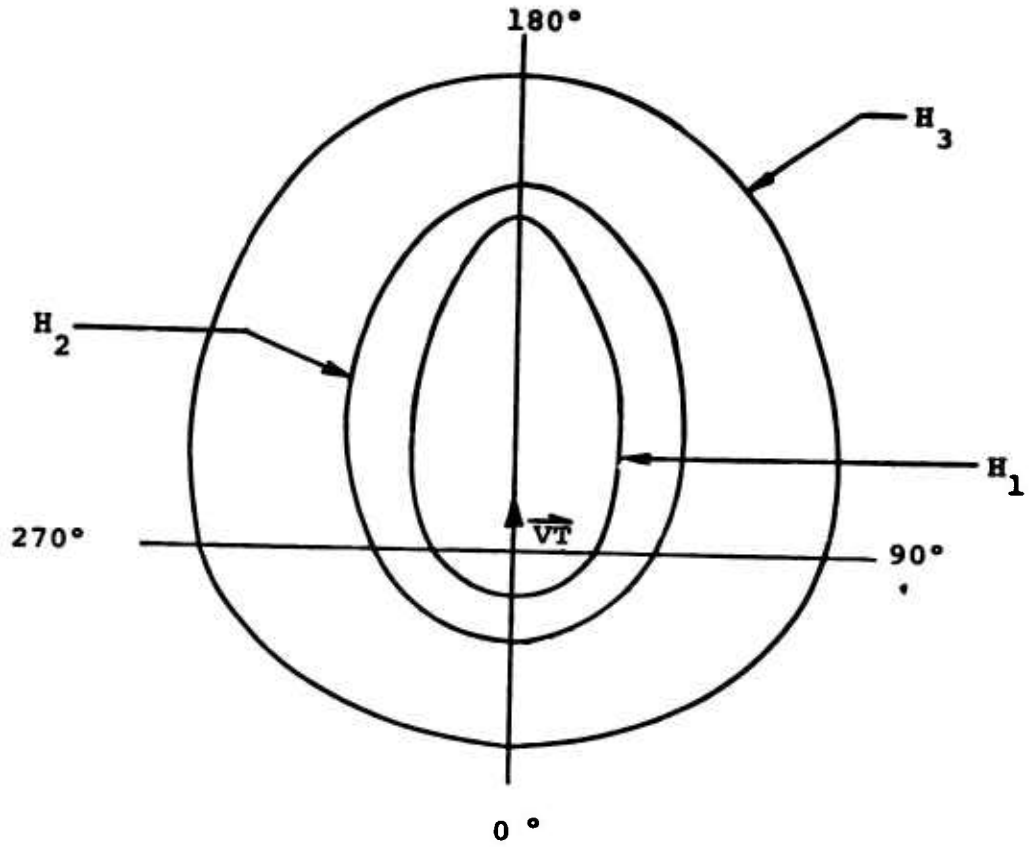
FIGURE 4. Typical Missile Capability as a Function of Aspect Angle



$V_3 > V_2 > V_1$

VF = VT

FIGURE 5. EFFECTS ON MAXIMUM LAUNCH RANGE DUE TO INCREASES OF VELOCITY F[illegible] THREE CO-SPEED ENCOUNTERS



$H_3 > H_2 > H_1$

HF = HT

FIGURE 6. EFFECTS ON MAXIMUM RANGE DUE TO INCREASES OF ALTITUDE FOR THREE CO-ALTITUDE ENCOUNTERS

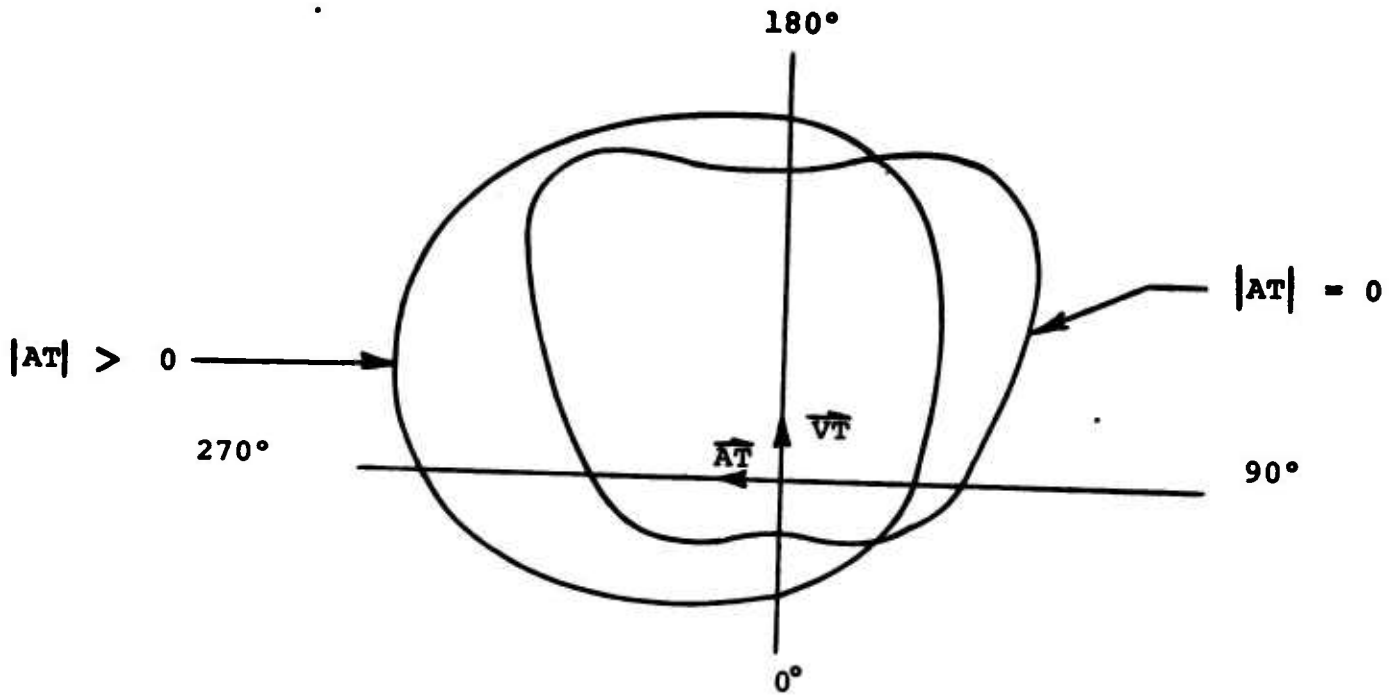


FIGURE 7. TARGET MANEUVER DISTORTS THE ZONE IN THE DIRECTION OF THE TURN

launch a missile and have a probability of killing the target. The Rmin line segment is the minimum range to launch a missile and still have a probability of intercepting the target. This figure clearly illustrates the effect aspect angle has on the capability of the missile with the majority of the aerodynamic capability existing in the forward aspect area. The higher closing velocity of the missile to the target is the cause of the increased range of launch opportunities in the forward aspect region.

Figure 5 shows the maximum capability of the missile for three co-speed encounters with velocity V_3 being greater than V_2 and V_2 greater than V_1. The increase in velocity decreases the maximum range for the beam and tail aspects, but increases the maximum range for the forward aspects. The increased velocity increases the drag of the missile which causes the closing velocity of the missile to decrease. The reduction in the closing velocity then limits the range from which the missile can be launched and successfully intercept the target. However, in the forward aspect maximum range increases because the increase in closing velocity, due to the increased velocities, over compensates for the decrease in closing velocity due to the increased drag on the missile.

The effects of changes in altitude are illustrated in Figure 6 for three co-altitude encounters. An increase in altitude reduces the drag of the missile which causes the closing velocity to increase, allowing the maximum launch range to increase.

As shown in Figure 7, the effect of target acceleration distorts the launch zone in the direction of the turn. This illustration shows the effect on the minimum range. The distortion of the maximum range boundary is also in the direction of the turn, but does not have as much effect as displayed on the minimum range.

The previous illustrations present a perspective of what is involved when attempting to define the computational algorithms required to display the missile capability to the pilot with a high degree of accuracy. Approaches that have been taken to solve the missile fire control problem are an empirical solution and a deterministic solution. The empirical method takes the known physical results and through experimental observations, attempts to define mathematical equations that will produce the expected results. In contrast to the

empirical method, the deterministic solution applies the scientific principles associated with the physical system to compute the end results. One of the most common empirical solutions employed in the airborne system is the stored zones technique. One version of this technique has been developed by Westinghouse Electric Corporation, Baltimore, Maryland, and has been implemented in the AWG-10A Fire Control System in the Navy F4J aircraft. Another version, developed by ASD/ENA, has not been employed in an airborne application. Deterministic approaches to the fire control solution includes the intercept triangle technique developed by McDonnell-Douglas, which performs the air-to-air functions in the F-15 fire control system, and the fast airborne simulation (FAS) technique developed by the Naval Missile Center, which has not been implemented in any production aircraft fire control system. Characteristics of each of the techniques will be brought out in the following paragraphs.

The stored zones technique is based upon a first order Taylor series expansion using partial derivations shown in Equation 1.

$$\begin{aligned} \text{Rmax}\ (\gamma) = \text{Rbase} &+ \frac{\partial\, \text{Rmax}}{\partial\, V_T}\,\Delta V_T + \frac{\partial\, \text{Rmax}}{\partial\, V_T}\,\Delta V_F \\ &+ \frac{\partial\, \text{Rmax}}{\partial\, G_T}\,\Delta G_T + \frac{\partial\, \text{Rmax}}{\partial\, \lambda}\,\Delta\lambda + \\ &+ \frac{\partial\, \text{Rmax}}{\partial\, H_T}\,\Delta H_T + \frac{\partial\, \text{Rmax}}{\partial\, H_T}\,\Delta H_T \end{aligned} \qquad (1)$$

Where Rmax (γ) is the maximum aerodynamic range for a launch encounter at a given aspect angle (γ), Rbase is the maximum launch range for the baseline encounter and the remaining terms are the partial derivatives of the maximum range with respect to each of the parameters, multiplied by the deviation of that parameter from the baseline. The variables include target velocity (V_T), fighter velocity (V_F), target acceleration (G_T), lead/lag angle (λ), target altitude (H_T) and fighter altitude (H_F). The accuracy of this technique is degraded with an increase in the non-linearity of the aerodynamic range as a function of the parameters. To counter this problem, additional data points can be stored other than the baseline, and by interpolating between the stored data points to obtain a single-variable deviation from the

baseline, the product term for each parameter in equation (1) can be replaced with the single-variable deviation as shown in the following equation:

$$\begin{aligned} \text{Rmax}(\tau) = \text{Rbase} &+ \Delta\text{Rmax}(V_T) + \Delta\text{Rmax}(V_F) \\ &+ \Delta\text{Rmax}(G_T) + \Delta\text{Rmax}(\lambda) \\ &+ \Delta\text{Rmax}(H_F) + \Delta\text{Rmax}(H_T) \end{aligned} \quad (2)$$

where,

$$\Delta\text{Rmax}(V_T) = \text{Rmax}(V_T) - \text{Rbase} \quad (3)$$

$$\Delta\text{Rmax}(V_F) = \text{Rmax}(V_F) - \text{Rbase} \quad (4)$$

$$\Delta\text{Rmax}(G_T) = \text{Rmax}(G_T) - \text{Rbase} \quad (5)$$

$$\Delta\text{Rmax}(\lambda) = \text{Rmax}(\lambda) - \text{Rbase} \quad (6)$$

$$\Delta\text{Rmax}(H_F) = \text{Rmax}(H_F) - \text{Rbase} \quad (7)$$

$$\Delta\text{Rmax}(H_T) = \text{Rmax}(H_T) - \text{Rbase} \quad (8)$$

Substituting equations (3) through (8) into equation (2) and simplifying, the maximum range for an aspect angle (τ) can be expressed as:

$$\begin{aligned} \text{Rmax}(\tau) = \text{Rmax}(V_T) &+ \text{Rmax}(V_F) + \text{Rmax}(G_T) \\ &+ \text{Rmax}(\lambda) + \text{Rmax}(H_F) + \text{Rmax}(H_T) \\ &- 5\ \text{Rbase} \end{aligned} \quad (9)$$

Inherent in this technique is the increase in inaccuracy as the conditions deviate from the baseline. This inaccuracy can be greatly improved by establishing other baselines. As an example, suppose the baseline for altitude is 15,000 feet and at high altitude encounters, the system accuracy is degraded extensively.

A second baseline would then be stored in the computer to reduce the error for the high altitude encounters. This technique can be accomplished in approximately 1000 sixteen-bit words for one missile and requires about 180,000 equivalent add operations per second to perform the mathematical computations and present ques to the pilot. Assuming a load of four missiles, the required computer memory is 850 sixteen-bit words/missile.

The intercept triangle technique is based on a simplified missile model which predicts the missile flight history (velocity, distance, time) as a function of launch velocity together with initial and terminal altitude. The model was developed on the basis of two highly simplified differential equations relating velocity and time for the powered and non-powered flight, as shown in equations (10) and (11), respectively.

$$\frac{dv}{dt} = A - DV^2 \qquad (10)$$

$$\frac{dv}{dt} = DV^2 \qquad (11)$$

The coefficients A and D are treated as constants for a particular engagement, but have different values for different engagements. The parameter A is considered to be an average thrust-to-mass ratio and functionally related to altitude through atmospheric pressure. The quantity DV^2 is interpreted to be a drag-to-mass ratio with D being proportional to atmospheric density which decreases in a roughly exponential fashion with increasing altitude. It is also proportional to the zero lift drag coefficient which depends on missile velocity.

The powered flight expression defined by equation (10) can be integrated using elementary methods. After rearrangement of the terms, the solution is:

$$V = \frac{V_o + (A/D)^{1/2} \tanh \left[A^{1/2} D^{1/2} (t-t_o)\right]}{1 + (D/A)^{1/2} V_o \tanh \left[A^{1/2} D^{1/2} (t-t_o)\right]} \qquad (12)$$

Integrating equation (12), an expression for the distance can be obtained which is

$$S - S_o = \frac{1}{D} \ln \left\{ \cosh \left[A^{1/2} D^{1/2} (t - t_o) \right] + (D/A)^{1/2} V_o \sinh \left[A^{1/2} D^{1/2} (t - t_o) \right] \right\} \quad (13)$$

where t_o, V_o, and S_o are initial time, velocity, and distance, respectively, and t, V, and S are the corresponding quantities at some later point along the trajectory. Making approximations for the hyperbolic functions and the natural logarithm which were derived on the basis of Taylor expansions with the expected range of the function arguments taken into account, expressions for the velocity and distance then become

$$V = \frac{V_o + A (t - t_o)}{1 + DV_o (t - t_o)} \quad (14)$$

$$S = (1/2) A (t - t_o)^2 + \frac{V_o (t - t_o)}{1 + DV_o (t - t_o)/2} \quad (15)$$

The non-powered portion of the missile trajectory is governed by equation (11) and is integrable as follows to obtain the velocity,

$$V = \frac{1}{D^2 V_o (t - t_o)} \quad (16)$$

Integrating equation (16), distance can be obtained through the following expression,

$$S = \frac{1}{D} \ln \left[\frac{V_o}{V} \right] + S_o \quad (17)$$

For the above equations, t_o, V_o, and S_o represent the conditions at the end of the powered flight and t, V, and S are the conditions at the end of the glide portion of the trajectory.

The above equations represent a missile model which predicts the relationship among missile velocity, distance, and time. With the model, a maximum launch range can be obtained by representing the factors limiting this range as minimum missile velocity requirements at intercept. These factors are the maximum relative range, gimbal limiting, closing velocity maneuver capability, and maximum time-of-flight. The maximum of the velocities associated with these factors is the required intercept velocity. From the intercept velocity, the missile distance and time-of-flight can be determined. The time-of-flight, along with the threat velocity, are then used to obtain the threat distance. Knowing the distance the target will travel and the missile distance, as shown in Figure 8, the maximum range can be determined using the following equation:

$$Rmax = Ds \cos(\tau) + \left[Dm^2 - \left[Ds \sin(\tau)\right]^2\right]^{1/2} \quad (18)$$

where Ds is the threat distance, τ is the aspect angle and Dm is the missile distance. If the quantity $\left[Ds \sin(\tau)\right]^2$ is greater than Dm^2, a launch opportunity does not exist. This technique requires about 1000 sixteen bit words for one missile and approximately 180,000 equivalent add operations per second to perform the calculations and present the resulting missile capability to the pilot at the necessary update rate. For four missiles, the average required computer memory is 750 sixteen bit words per missile.

The Fast Airborne Simulation (FAS) approach to an airborne computer estimation of missile capability uses current radar data on the target and own ship data as inputs to rapidly estimate launch opportunities on the basis of hit or miss results from the missile simulation. The inherent problem associated with this solution is one of computational time within the time-sharing environment of the airborne fire control computer. The time lag between processing data and presenting a cue to the pilot must be kept short or else the dynamic nature of the air combat environment will radically alter the state of the affairs and the cues will be unreliable.

$$Rmax = Ds \cos(\tau) + \left[Dm^2 - \left[Ds \sin(\tau) \right]^2 \right]^{1/2}$$

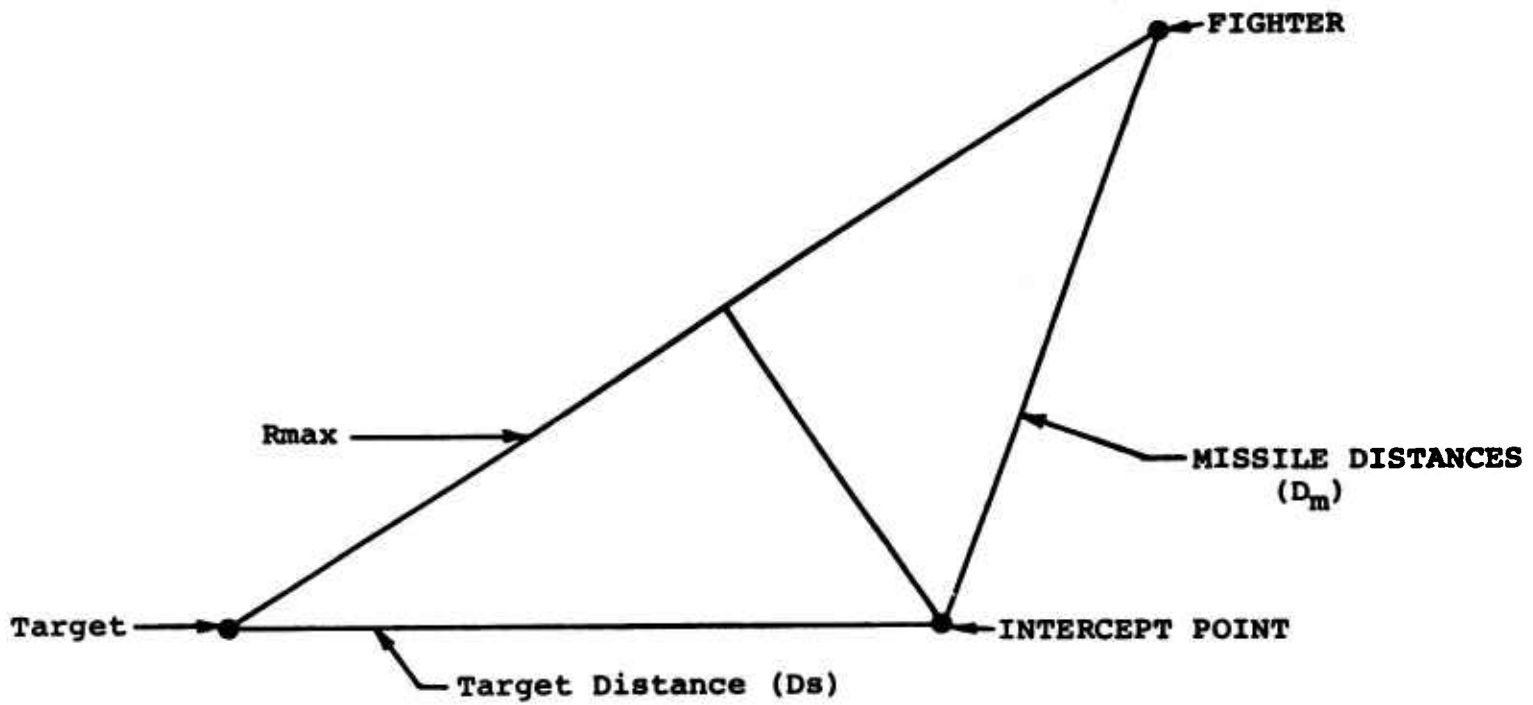


FIGURE 8. STRAIGHT LINE TRIANGLE GEOMETRY FOR OBTAINING SOLUTION FOR Rmax

Four basic assumptions are made to simplify the missile simulation. First, the effect of gravity is neglected. This divorces the simulation from an earth-referenced coordinate system. The coordinate system used is the instantaneous radar coordinate system in which the initial target data and own ship data are defined. Consequently, the simulation uses no axes transformations on trigonometric functions. Second, the angle of attack is neglected. Thus, the missile body axis and the missile velocity vector are assumed coincident. Third, the responses of the major missile subsystem and the aerodynamic response of the missile can be approximated by linear first order differential equations. Fourth, all kinematical quantities can be treated as constants during an integration time step. Thus, all differential equations are assumed to have both constant inputs and fixed time constants during an integration time step.

To start the computational process, the following variables at the time of missile launch (T = 0) must be initialized: $\overrightarrow{PT}$, $\overrightarrow{VT}$, $\overrightarrow{PM}$, $\overrightarrow{VM}$, $\overrightarrow{U}$, $\overrightarrow{WS}$, $\overrightarrow{WF}$, $\overrightarrow{AN}$, MM. All the vectors must be defined in a common right-handed orthogonal basis. This basis will then be the working basis for the computations. It is immaterial what basis is used so long as it remains fixed during the computations. When a vector quantity, say $\overrightarrow{VM}$, appears without the arrow as a vector designator; i.e., as VM, the magnitude of the vector is implied.

At its computational cycle rate, an airborne computer provides a kinematic description of the supposedly current state of affairs of the air combat engagement; i.e., the position, velocity, and acceleration vectors of the target and the velocity and acceleration vectors of the launch aircraft, all in the radar coordination system. The position vector of the launch aircraft in the radar coordinate system is the zero vector; however, altitude of the launch aircraft is available.

To predict the position of the launch aircraft, it is assumed that the acceleration, $\overrightarrow{AF}$, is a turning acceleration. The associated turning rate, $\overrightarrow{WF}$, is given by:

$$\overrightarrow{WF} = (\overrightarrow{VF} \times \overrightarrow{AF})/VF^2 \qquad (19)$$

where $\vec{VF}$ is the current velocity of the launch aircraft. The position vector of the launch aircraft 2 seconds ahead, which is the initial position of the missile $\vec{PM}$, is given by a truncated Taylor series as:

$$\vec{PM} = \vec{PF} + 2\vec{VF} + 2\vec{AF} + \frac{4}{3}(\vec{WF} \times \vec{AF}) \qquad (20)$$

The velocity of the launch aircraft 2 seconds ahead, which is the initial velocity of the missile, $\vec{VM}$, is also given by a truncated Taylor series as:

$$\vec{VM} = \vec{VF} + 2\vec{AF} + 2(\vec{WF} \times \vec{AF}) \qquad (21)$$

The position and velocity vectors of the target are predicted ahead, assuming that the target acceleration, $\vec{AT}$, is a turning acceleration. The turning rate of the target, $\vec{WT}$, is given by:

$$\vec{WT} = (\vec{VT} \times \vec{AT})/VT^2 \qquad (22)$$

where $\vec{VT}$ is the current velocity of the target.

Although assumed zero in predicting the position and velocity vectors of the target at missile launch, the tangential component of the target acceleration, ATT, is used to update the target trajectory during missile flight. Both ATT, given by:

$$ATT = (\vec{VT} \cdot \vec{AT})/VT \qquad (23)$$

and $\vec{WT}$ are assumed constant in the process of updating the target trajectory.

The predicted position and velocity vectors of the target at missile launch are given by truncated Taylor series in the form of replacement equations; i.e., the same variable appears on both sides of the equal sign with the old value of the variable on the right and the new value on the left. The predicted position of the target is:

$$\vec{PT} = \vec{PT} + 2\vec{VT} + 2\vec{AT} + \frac{4}{3}\vec{WT} \times \vec{AT} \qquad (24)$$

and the predicted velocity of the target is:

$$\vec{VT} = \vec{VT} + 2\vec{AT} + 2\vec{WT} \times \vec{AT} \qquad (25)$$

where, again, the prediction is done 2 seconds ahead. Clearly, in all of these prediction equations, the coefficients would change if the prediction time were varied from 2 seconds.

With the conditions at launch predicted, a check is made to see if missile gimbal angle limitations are exceeded. If the limits are exceeded, a miss is declared. The relative range vector at launch is:

$$\vec{R} = \vec{PT} - \vec{PM} \qquad (26)$$

measured from the missile to the target. The off-boresight angle (λ) measured from the missile velocity vector is:

$$= \text{ARCCOS}(\vec{VM}/VM \cdot \vec{R}/R) \qquad (27)$$

If λ is equal to or greater than the gimbal limits, a miss is immediately declared.

The computations, as detailed here, are performed cyclically until the criteria of either a hit or a miss is satisfied at which point the result is declared as a cue to the pilot, the simulation is terminated, and new kinematical data from the air combat is processed.

The current altitude of the launch aircraft is used to compute the physical properties of air required in the simulation; i.e., the air density and the speed of sound.

The relative range vector, $\vec{R}$, from the missile to the target is given by:

$$\vec{R} = \vec{PT} - \vec{PM} \qquad (28)$$

and the velocity of the target relative to the missile, $\overrightarrow{VTM}$, is also given by:

$$\overrightarrow{VTM} = \overrightarrow{VT} - \overrightarrow{VM} \tag{29}$$

From these two vectors, the range rate, RDT, a scalar, is computed as:

$$RDT = (\overrightarrow{R} : \overrightarrow{VTM})/R \tag{30}$$

and the line-of-sight rate, $\overrightarrow{LOSR}$, is obtained as:

$$\overrightarrow{LOSR} = (\overrightarrow{R} \times \overrightarrow{VTM})/R^2 \tag{31}$$

A test is provided at this point to adjust the computational time step and reduce its size as the missile approaches the target. This ensures a finer computation of miss distance at intercept. The test criteria is:

$$DTEST = \frac{R}{2\ |RDT|} \tag{32}$$

if DTEST < DT, then DT = Max. (DTEST, 0.005). Thus, the smallest possible time step is 0.005 second.

As the program goes through a computational cycle, a series of tests are conducted to determine whether or not a hit is possible. These tests involve the line-of-sight rate, gimbal angle limits, closing speed, range-to-arm, missile guidance, and miss distance. These tests are conducted on the basis of criteria stored in the program as missile parameters.

The missile seeker system measures the line-of-sight rate. However, the seeker is assumed to respond as a linear first order system and the output of the seeker is obtained from the differential equation:

$$\overrightarrow{LOSR} = TAUS\ (d\overrightarrow{WS}/dT) + \overrightarrow{WS} \tag{33}$$

where TAUS is the time constant of the seeker. If we assume that the line-of-sight rate is constant over the computational time step, we can represent the solution to this differential equation by a replacement equation which updates the seeker rate each computational time step given by:

$$\vec{WS} = e^{-DT/TAUS}\,\vec{WS} + (1 - e^{-DT/TAUS})\,\vec{LOSR} \qquad (34)$$

The total commanded acceleration to the missile is the vector sum of an acceleration proportional to the seeker measurement of the line-of-sight rate and an acceleration proportional to the off-boresight angle. However, physical constraints limit the commanded acceleration to a certain magnitude.

The missile responds to the total commanded acceleration as an input signed. The aerodynamic response of the missile is obtained as the solution to a linear first order differential equation with the normal acceleration as the dependent variable shown in the following equation,

$$\vec{AC} = TAUA\ (d\vec{AN}/dT) + \vec{AN} \qquad (35)$$

where $\vec{AC}$ is the commanded acceleration vector, TAUA is the aerodynamic time constant of the airframe and $\vec{AN}$ is the normal acceleration vector. The time constant is a function of mach number, altitude, and time. Again, assuming $\vec{AC}$ and TAUA are both constant over the computational time step, the solution to the differential equation is given in the form of a replacement equation for $\vec{AN}$ as:

$$\vec{AN} = e^{-DT/TAUA}\,\vec{AN} + (1 - e^{-DT/TAUA})\,\vec{AC} \qquad (36)$$

The missile has an acceleration along its velocity vector due to thrust and drag. This acceleration is given by the equation:

$$F = (TH/MM) - (QS/MM)\ CDO - (MM \cdot AN^2)/QS \cdot CDL \qquad (37)$$

where TH is the thrust of the motor, MM is the missile mass, QS is the dynamic pressure times a characteristic area, CDO is the zero-lift drag coefficient, and CDL is

the zero-induced drag coefficient. Both CDO and CDL are a function of the mach number, and thrust and mass are a function of the time-of-flight of the missile.

The missile acceleration vector, $\overrightarrow{AM}$, is given by the vector sum:

$$\overrightarrow{AM} = \overrightarrow{AN} + F\ (\overrightarrow{VM}/VM) \tag{38}$$

With this acceleration and the missile velocity, the position vector of the missile is updated in the computational time step, DT, by the replacement equation:

$$\overrightarrow{PM} = \overrightarrow{PM} + DT \cdot \overrightarrow{VM} + DT^2 \cdot \overrightarrow{AM}/2 \tag{39}$$

The missile velocity vector is updated in the computational time step, DT, in a slightly different fashion. An intermediate velocity vector, VMI, is first computed as:

$$\overrightarrow{VMI} = \overrightarrow{VM} + DT \cdot \overrightarrow{AM} \tag{40}$$

Then, the missile velocity vector is updated by the equation:

$$\overrightarrow{VM} = (\overrightarrow{VMI} \cdot \overrightarrow{VM}/VM) \cdot (\overrightarrow{VMI}/VMI) \tag{41}$$

The reason for this procedure is to reduce the effect of both large normal accelerations and large computational time steps. A situation where the normal acceleration does not just turn the velocity vector, but also causes a change in its magnitude.

With the values of ATT and $\overrightarrow{WT}$ determined during the initialization of the simulation, the acceleration of the target is computed as:

$$\overrightarrow{AT} = (ATT \cdot \overrightarrow{VT}/VT) + (\overrightarrow{WT} \times \overrightarrow{VT}) \tag{42}$$

The position vector of the target is updated by the equation:

$$\overrightarrow{PT} = \overrightarrow{PT} + DT \cdot \overrightarrow{VT} + DT^2 \cdot \overrightarrow{AT}/2 \qquad (43)$$

and the velocity vector of the target is updated by the equation:

$$\overrightarrow{VT} = \overrightarrow{VT} + DT \cdot \overrightarrow{AT} \qquad (44)$$

Finally, the value of time, T, is updated by the equation:

$$T = T + DT \qquad (45)$$

and the cycle is complete. A new computational cycle begins with the newly computed values of T, $\overrightarrow{PT}$, $\overrightarrow{PM}$, and $\overrightarrow{VM}$. The tests are made for a hit or miss and a cue given to the pilot if either are satisfied; otherwise, the computational cycle is completed again. Ultimately, a hit or miss will be declared.

This technique requires about 2000 sixteen-bit words for one missile and approximately 250,000 equivalent add operations per second to perform the appropriate calculations and present the update cues to the pilot. For four missiles, the average required computer memory is 1000 sixteen-bit words per missile.

Based upon a recent ASD/ENAS study, the recommended technique for missile fire control is the intercept triangle because of accuracy and less computational time required. The study revealed that the accuracy of the empirical solution was on the order of 75 per cent, whereas the deterministic was between 85-90 per cent probability of giving the correct missile capability to the pilot.

Other areas which need future improvement include: (1) systems engineering to provide present and future program offices with assistance in areas such as software development, hardware integration/interface associated with designing avionics systems/subsystems; (2) non-target data zones - to find out if there is any cost effective means to launch air-to-air missile effectively without radar lock-on; (3) no-escape zones for certain profiles, it is possible for the target to execute optimum maneuvers and escape the missile intercept in the end-game geometry. It may be desirable to display to the pilot a launch range such that the target would be unable to evade the missile after launch; (4) lethality boundaries - even though maximum and minimum ranges are displayed to the pilot, he has no idea how lethal the range is. It might be desirable to inform

the pilot of the probability of intercepting the target for a given launch encounter.

THE ANALYSIS AND SIMULATION OF A FIRE CONTROL TECHNIQUE FOR LAUNCHING AIR-TO-AIR MISSILES

by

D. H. deDOES
SPACE AND MISSILE TEST CENTER
VANDENBERG AFB, CALIFORNIA

T. E. MORIARTY
AIR FORCE INSTITUTE OF TECHNOLOGY
WRIGHT-PATTERSON AFB, OHIO

G. M. BERGEMAN
AIR FORCE SPECIAL WEAPONS CENTER
HOLLOMAN AFB, NEW MEXICO

R. D. GREEN
AIR FORCE ARMAMENT LABORATORY
EGLIN AFB, FLORIDA

ABSTRACT

A fire control technique which predicts and displays a missile's maximum off-boresight capability is analyzed by simulation. The fire control computer considers such factors as the target's acceleration and range and the missile's performance. The philosophy of this approach is to reduce both the hardware and on-line computational requirements while incorporating pilot judgment as part of the solution to the fire control problem.

To test the system's performance, both the target and attacking aircraft have been modeled in three dimensions and programmed for an EAI 8400 computer. Two flight stations are provided which contain control sticks for pitch and roll; the attacker's station is equipped with a firing trigger for launching air-to-air missiles. Both pilots observe their relative positions and accelerations as well as the in-flight missile on the face of a large cathode ray tube. The display is similar to that which would be observed by the attacking pilot through his heads-up display during actual combat.

In one implementation, the maximum lag angle is estimated using a first-order Taylor series expansion about one of 32 selected nominal points within the envelope. The range limit is predicted using time-of-flight estimates. The feasibility of this concept is tested, and it is concluded that the maximum off-boresight angle and the maximum range boundary can be satisfactorily predicted using the described technique and that reasonably small on-board digital storage and computational capabilities are required for implementation.

PART I

INTRODUCTION

The highly dynamic environment of air-to-air combat places a substantial workload on an attacking pilot when his primary weapon is an air-to-air missile. When firing under visual contact, it has been found that pilot generated solutions to the fire control problem (i.e. rules of thumb) have not been sufficiently accurate. However, a complete solution by a fire control computer which does not rely to some extent upon pilot judgment may be undesirable. The objective of this paper is to develop and analyze a fire control technique which relies on pilot judgment and which requires a minimum amount of onboard computation.

In particular, this study considers the problems of pursuit and weapon delivery when the target aircraft can be maintained within visual contact. Furthermore, two types of attacks are considered when the pursuer is initially located in the target's rear hemisphere. The first, when both aircraft are at nearly co-speed and performing ACM. The second, when the pursuer has a speed advantage and is performing a slashing type of attack. In the latter case the attacking pilot is faced with a time problem in that he must prevent overshooting the target before a firing opportunity occurs.

The approach of this paper is to first develop a pursuit steering law which is effective from the initial conditions of the encounter, and which results in a missile firing opportunity. Moreover, mechanization of the steering law provides: (1) an opportunity for off-boresight seeker lock on, and (2) an indication of the bounds of the missile's capability so that the first missile can be fired in nearly minimum time. The concept of capture regions is employed to determine the missile's maximum off-boresight capability. This information can then be incorporated in a helmet-mounted display or HUD.

To demonstrate the proposed technique, a computer generated, real time, combat simulation has been mechanized, Ref (1). This simulation provides the attacking pilot with the proper controls and visual cues for following the proposed steering law, achieving seeker-lock-on, and firing the missile(s). The target pilot is provided with controls and visual cues by which he can perform evasive maneuvers.

ANALYSIS OF THE FIRE CONTROL PROBLEM

In the literature one finds a number of steering laws which will effect an intercept, Refs (2), (3). These include proportional, pure pursuit, lead or lag pursuit, and those techniques applied by fighter pilots when engaged in intercept missions. The latter types are

commonly found in the radar aided, missile firing, intercepts as performed by modern fighter aircraft. Here the missile's capability is usually displayed as a steering circle on the radar scope. The pilot then attempts to satisfy the firing envelope by steering his interceptor so as to null the indicated error.* The approach taken in this paper is similar to the technique described above in that an error is presented to the pilot which he then attempts to null. However, there are basic differences between the two approaches. In the method proposed in this paper the steering error is observed on a HUD and, in addition, the fire control solution remains valid during hard maneuvering combat.

MISSILE PERFORMANCE

Missile capability is typically presented in the form of launch envelopes as shown in Figs. 1 and 2 (these plots are normally computed when the missile is launched from a pure pursuit course). A typical characteristic of launch envelopes is that they shift toward the outside of the target's turn (see Fig.2). Hence, an attacker which is turning "inside" of the target as a result of boresight tracking may be denied launch opportunities as a result of the envelope's position.

AN OPTIMUM ATTACK GEOMETRY

In view of the above discussion, a preferred pursuit technique may be that shown in Fig. 3. Here the steering law is that which drives the pursuer from some initial position in the target's rear hemisphere to a missile firing position within the dashed region. This type of pursuit may also be considered optimum in view of: (1) pursuit or lead pursuit which requires the attacker to track inside the target's turn may result in an "overshoot", (2) tracking too far outside of the target's turn provides an opportunity for the evading pilot to "reverse" his turn - this could lead to the situation described in item (1) above, and (3) for the missile design considered in this report, multiple firings at a range of 5000 to 6000 feet from within the dashed region appears to produce good results.

Obviously the attacking pilot cannot employ pure pursuit (or boresight the target) steering to achieve the desired firing position. As will be shown in the sequel, a variable lag pursuit strategy can drive the attacker along a path which terminates in a firing position

---

* This procedure is often limited to nonmaneuvering targets.

within the dashed region. Since missile launch envelopes are normally based on a pure pursuit attack, other means of defining the missile's capability must be found.

CAPTURE REGIONS FOR A MISSILE

One way to determine the performance of an interceptor is via the concept of capture regions, Ref (4). In the present study, the missile's capture region is defined by the set of initial target positions for which the missile can perform a successful intercept under specified flight conditions (see Fig. 4). For the flight condition shown, the target would be able to escape the attack if the off-boresight angle, β, were allowed to become larger than its maximum value, β_{max}. Here, β_{max} is defined by the boundary of the capture region which is in turn a function of the flight condition (i.e. of the target's range, speed, altitude, and maneuver, the attack geometry (see Fig. 6) and the attacker's speed and altitude).

Often, a discussion of capture regions involves game theory which in turn implies that the boundaries of the capture regions result from optimal strategies. For the intercept problem, the missile's strategy to attain optimum performance may be a maximum rate turn, a dash, or a combination of these maneuvers. In general, optimal steering requires that the missile have complete information as to the target's state and maneuver. In contrast, current missiles are limited in their knowledge of the target's state (the line of sight rate is convenient to measure); therefore, proportional guidance has found extensive use.

In view of the sophistication required to guide in an optimal fashion, the method employed in this study is based on the maximum off-boresight angle, β_{max}, for which a hit can be achieved by a missile which uses proportional guidance. A presentation of the missile's maximum capability to the pilot on his HUD allows him to directly estimate the amount of "turn" required to achieve a firing opportunity.

Determination of the maximum missile capability becomes a difficult task due to the large number of variables required to model both the missile and the target. For the two dimensional encounter depicted in Fig. 6, β_{max} is a function of the target's range, speed, altitude and maneuver, the angles α and β, and the attacking aircraft's speed and altitude. Ref. (1) contains an outline of the procedure used to determine β_{max}.

THE PURSUIT LAW

With an indication of missile performance available from the fire control computer, that is, the angle defined by β_{max}, the question arises as to how this information may be used by the attacking pilot. The technique proposed in this study is depicted in Fig. 5. Here the pursuer's control strategy is similar to that used for a gun attack: roll to keep the target aligned in the pitch plane and apply normal acceleration (pitch up) until the off-angle to the target, β, becomes less than β_{max}. Also, as in a gun attack, a tracking period is required in order to assure that $\dot{\beta} \approx 0$. The latter requirement is due to the assumptions made in computing β_{max}. Thus, in effect, the missile is fired as if it were an off-boresight gun. That is, a sequence of steps, as shown in Fig. 5a-d, is employed which is similar to that which would be used for a gun attack.

The type of display proposed should be compatible with pilot training for ACM. Furthermore, as with a gunsight, the pilot is always aware of the magnitude of the heading error that exists and thus can judge how much "turn" he must apply in order to satisfy the missile's requirements. How the error is reduced obviously depends upon the tactical situation. Providing the attacker has sufficient performance in comparison to the target, he can turn with the evader until a firing situation is indicated. In other situations it may be preferred to yo-yo high, yo-yo low, or high speed barrel roll in order to achieve more favorable initial conditions from which to start applying a control strategy to achieve $\beta < \beta_{max}$ and $\dot{\beta} \approx 0$.

The type of steering used to null errors can take on a variety of forms. The first technique employed in the combat simulation, Ref. (1), resembles a minimum time to launch strategy. This consists of rolling to maintain the target in the pitch plane while executing a maximum rate turn until $\beta < \beta_{max}$. This strategy is similar to the optimal feedback pursuit law as described by Lynch in Ref. (5). Lynch's solution to a three dimensional differential game, when the payoff is time to reach the weapon envelope, shows that the pursuer follows the type of strategy described above. The second method used in the simulator consists of rolling to keep the target in the pitch plane and applying sufficient G to reduce the off-angle β at an acceptable rate (the large $\dot{\beta}$ obtained from the first strategy may be unacceptable due to factors such as a large energy loss).

DEVELOPMENT OF THE SIMULATION

The attacker's view of a typical air-to-air engagement is shown in Fig. 5. The corresponding flight paths are depicted in Fig. 7. The objective of the simulation is to provide the attacking pilot with a view similar to that shown in Fig. 5, and thereby provide a means of testing the proposed solution to the fire control problem. Since it would be unfair to deny the target pilot visual information as to the attacker's position, the evader was given access to the attackers view of the encounter (this would provide the worst case for the attacker since the evader is constantly aware of any missile firing opportunities).

The simulation was mechanized on the EAI 8400 hybrid computer located in the Engineering Flight Simulation Branch of the Air Force Flight Dynamics Laboratory. The computer is used to determine the positions of the attacking and target aircraft, to simulate the sight display, and to calculate the missile trajectory.

Two flight stations are provided for the attacking and evading pilots. Both stations contain control sticks for pitch and roll; the pursuer's station also provides a firing trigger for launching air-to-air missiles. A block diagram of the hybrid simulation is shown in Fig. 8. The attacker's view of an encounter at a particular time point (see Fig. 7) as observed on the cathode ray tube (CRT) in the simulator is depicted in Fig. 9. In this figure the CRT coordinate system translates with the pursuer while maintaining a local level orientation (CRT system does not roll with the pursuer), see Fig. 10. This type of mechanization gives both pilots a horizon reference.

The target is simulated as an inverted T which represents the wings and vertical stabilizer. Coordinated flight is assumed; thus, the attacker can anticipate the evader's turns by noting the direction of his vertical stabilizer.

For the pursuer's reference, boresight is displayed as a cross centered in the lower half of the CRT (see Fig. 9).

To add realism and, in addition, to provide the target pilot with an opportunity to carry out evasive maneuvers, the missile is also displayed on the CRT (the view on the CRT is similar to that experienced by the attacking pilot while tracking the rocket motor of an air-to-air missile). A scope read out provides information on the missile's flight status, such as: velocity, range to the target, time, normal acceleration, line of sight rate, the seeker gimbal angles, and the miss distance.

The fire control solution, which is represented by β_{max}, is displayed on the CRT as a chevron (see Fig. 9). The location of the chevron is determined by the attacker's roll angle and by β_{max}. Hence

it serves two purposes by representing the attacker's wings and by indicating the missile's capability (the distance from boresight is proportional to β_{max}). Since β_{max} is a function of both the flight condition and attack geometry, the chevron can be observed to depress (as would a depressed reticle gunsight) and roll as the ACM progresses. In addition, the missile's seeker was assumed to be slaved to the vertex of the chevron. Thus, in order to achieve off-boresight seeker lock-on, the vertex must be superimposed over the target for a short period of time. To provide flexibility in achieving seeker lock-on, the chevron can be "trimmed" in the attacker's pitch plane. This mode is used when the target off-angle, β, is observed to be less than β_{max} at the beginning of an attack.

A detailed description of both the aircraft and missile models may be found in Ref. (1).

SIMULATION RESULTS

Simulations of air-to-air encounters have been conducted with the following objectives in mind:

(1) Does the solution to the fire control problem, as given by β_{max}, accurately represent the missile's performance over a wide range of flight conditions?

(2) What are the characteristics of the pursuit trajectories when steering is carried out using those techniques described in the previous section (the objective in this phase is to test the system against a programmed target, that is, a target flying straight and level or performing constant G turns)?

(3) Does the system perform adequately in a "no holds barred" dogfight?

The tests required to adequately answer the first and third objectives have not been completed as of the date of this manuscript. However, it appears that for those flight conditions investigated, the system adequately predicts the missile's performance capability-about 95% of the firings have resulted in hits. Complete results will be available in Ref. (1) and from follow on studies.

Typical behavior of the attacker when pursuing a target which is turning at a constant rate is depicted in Fig. 11. Here the attacker starts out at the target's six o'clock position at a range of 5000 feet and then allows him to turn until $\beta > \beta_{max}$. The pursuit, starting at about the 15 second point, consists of a 4 G turn until the attacking pilot is assured that $\beta < \beta_{max}$. A reduction in the attacker's G allows him

to lock the seeker and track for a short time to satisfy the $\dot{\beta} \approx 0$ criterion. The point at which the missile was fired is indicated on the trajectory in Fig. 11. Note that the lag pursuit employed by the attacker results in a pursuit path that closely follows the target's trajectory. The desirability of this type of attack was discussed previously.

An attack similar to the one discussed above is depicted in Fig. 12. Here the target is maintaining a constant 5 G horizontal turn. The pursuit strategy was to apply sufficient G so as to assure that $\beta < \beta_{max}$. However, a steady tracking situation was not achieved and the missile was fired when $\beta < \beta_{max}$. Even though a hit was achieved, this type of "snap shooting" is not recommended.

A second attack against a target which is performing a constant 5 G horizontal turn is shown in Fig. 13. Here, the attack starts at the six o'clock position with the pursuer reaching maximum G about five seconds after the target initiates the 5 G turn. Note that in contrast to the trajectories in Fig. 11, the attacker turns "inside" of the target's track. This characteristic is due to both the initial conditions for the encounter and the fact that the missile has limited off-boresight capability against a 5 G target (see Ref. 1). Therefore, when tracking under the conditions shown in Fig. 13, the chevron was depressed almost to boresight; hence, compliance with the indicated missile capability requires nearly a pure pursuit type of attack. Increasing the missile's off-boresight capability would allow an increase in the off-boresight tracking capability. In this situation, the trajectories should resemble those in Fig. 11.

A 3-D encounter of approximately 30 seconds in duration is shown in Fig. 14. Due to the nature of the aircraft models used in the simulation, and since pilot G tolerances are not a factor, both aircraft can maintain high turn rates for prolonged periods of time (they can also roll while sustaining high G flight). Hence, as illustrated in the G profiles in Fig. 15, the target uses this capability to generate large off-angles and thereby escape the missile's capture region. A reduction in the attacker's G which is evident in Fig. 15 is due to the encounter geometry which allowed the attacker to "let up", lock the seeker, then track and fire.

Several factors pertaining to ACM became very apparent during the simulation runs. The first is that it becomes increasingly difficult to boresight a target which is turning at high rates. For targets turning with a normal acceleration in the 3-4 G range, the attack can be made by turning at 4.5 to 5 G's, this reduces the off-angle at an acceptable rate and eventually drives the target into the missile's capture region. However, for targets operating above 5 G's, the attacker

must increase his load factor into the range where either pilot tolerance becomes a limit or excessive energy is being expended to sustain the high turn rates. This discussion assumes, of course, that other means such as yo-yo's, etc., are not used to improve the attack geometry. Not only is the attacker forced into the higher G range, but the time required to, say, boresight a target may increase because the misalignment between the attacker's and target's velocity vectors is likely to increase as G increases (Ref. 1).

Relating these observations to the capability of the missile which is currently modeled in the simulation (see Ref. 1), one finds that a 5 G target at a range of about 5000 feet must be nearly boresighted in order to achieve a firing opportunity. In constrast, the missile may be fired at 3 G targets up to 30 degrees off-boresight. Simulation runs using this missile have shown that it is relatively easy for the attacker to convert large off-angles to a firing opportunity when $G_T \lesssim 3$. Hence, increasing the missile's performance should substantially reduce the time required to convert a large off-angle situation to a firing opportunity in a high G environment.

CONCLUSIONS

The approach taken to solve the fire control problem for launching air-to-air missiles is that of displaying the boundary of the missile's capture region as observed in the attacker's pitch plane. It appears that this information can be presented to the pilot using a helmet mounted display or a HUD with an extended field of view to accommodate the large off-boresight angles. The pursuit strategy to be employed when using the display resembles that for a gun attack and, therefore, should be compatible with current methods for pilot training. The proposed steering law is uncomplicated and consists of rolling to keep the target in the pitch plane and turning to reduce the off angle until the target is imbedded in the missile's capture region.

Results from the simulation show that there are a number of advantages in displaying the boundary of the capture region, these include:

(1) Constant pilot awareness of the missile's capability allows him to judge how much "turn" must be achieved in order to reach a firing position,

(2) missile off-boresight capability permits a firing opportunity in nearly minimum time, and

(3) steering via the capture region boundary results in a variable lag type of pursuit which may alleviate the "overshoot" problem when attacking from short ranges.

The problem of determining a function, in terms of the system variables such as range, target acceleration, etc., which describes the boundary of the capture region has not been completely solved. The results described in this paper are for constant altitude attacks against targets flying at a constant speed. A follow on study will extend the simulations capability by incorporating more realistic aircraft and missile models and by improving the solution for the missile's maximum off-boresight capability.

In addition to an on-line application, the display of capture regions in a manned simulation may prove to be beneficial in conducting aircraft-missile performance trade-off studies. Also, pilot training programs may be enhanced by displaying the boundary of the capture region on the HUD in a combat simulator. One particular aspect of training where such an approach may be beneficial is that of reinforcing cockpit discipline. For example, a sequence of steps such as: (1) prepare the missile for launch, (2) lock the seeker, and (3) determine when a firing opportunity exists, must be carried out without omissions if a hit is to be achieved. Hence, simulators which display such information as the boundary of the capture region may very well enhance a training program for combat pilots.

REFERENCES

1. AF Institute of Technology. The Analysis and Simulation of a New Fire Control Technique for Launching Air-to-Air Missiles, by C.D. Haas, and L.D. Puckett. WPAFB, Ohio. GGC/EE/73-4, June 1973.
2. Van Nostrand, Princeton, New Jersey. Guidance, by A.S. Locke. 1955.
3. AF Flight Dynamics Laboratory. Characteristics of Pursuit Study, by M. Miller. WPAFB, Ohio. January 1971.
4. Systems Control Inc. Application of Reachable Sets Techniques to Air Combat Analysis, by M.D. Ciletti, L. Meier, and D.M. Salmon. Palo Alto, Calif. January 10, 1973.
5. AF Institute of Technology. Differential Game Barriers and Their Application in Air-to-Air Combat, by U.H.D. Lynch. WPAFB, Ohio. DS/MC/73-1, March 1973.

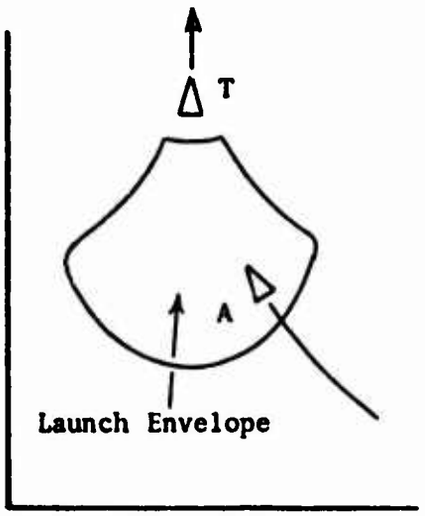


Fig. 1 Typical Launch Envelope: Target in Level Flight

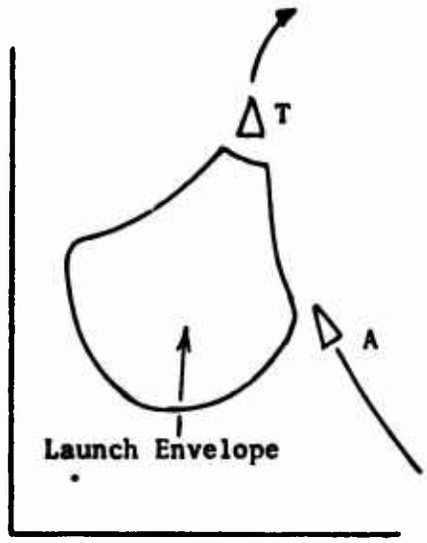


Fig. 2 Typical Launch Envelope: Target Turning

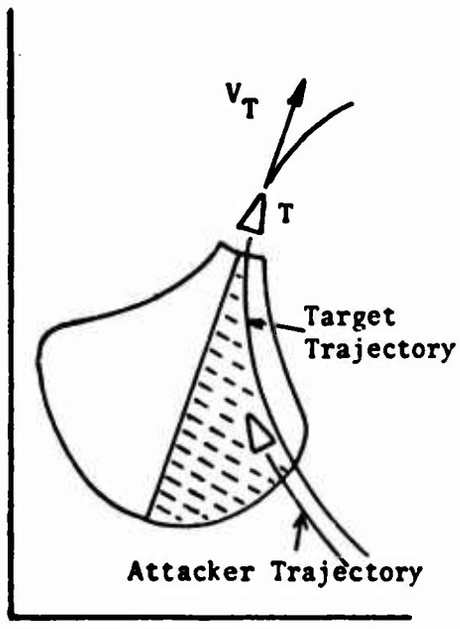


Fig. 3 Preferred Attack Position

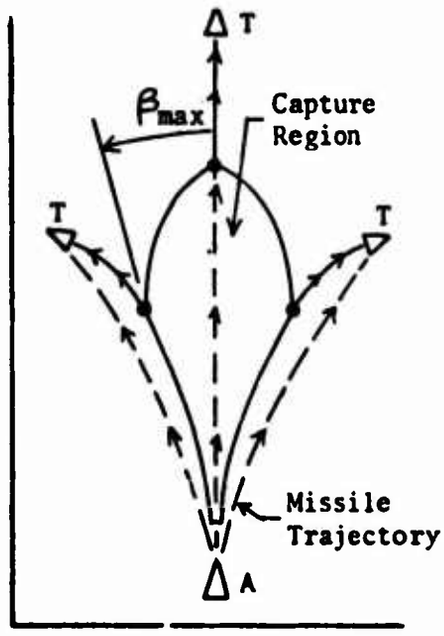
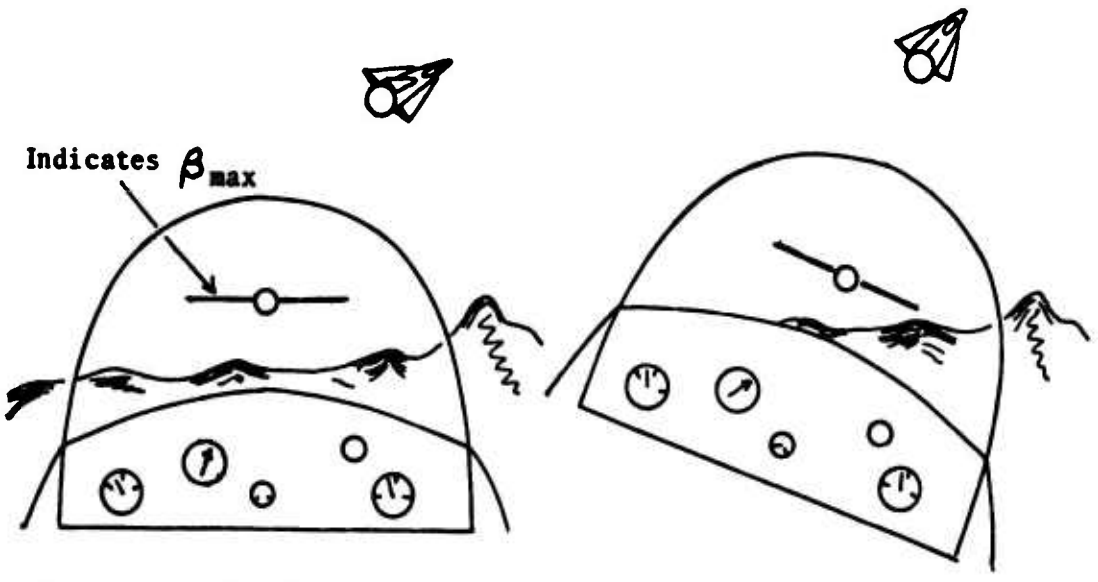
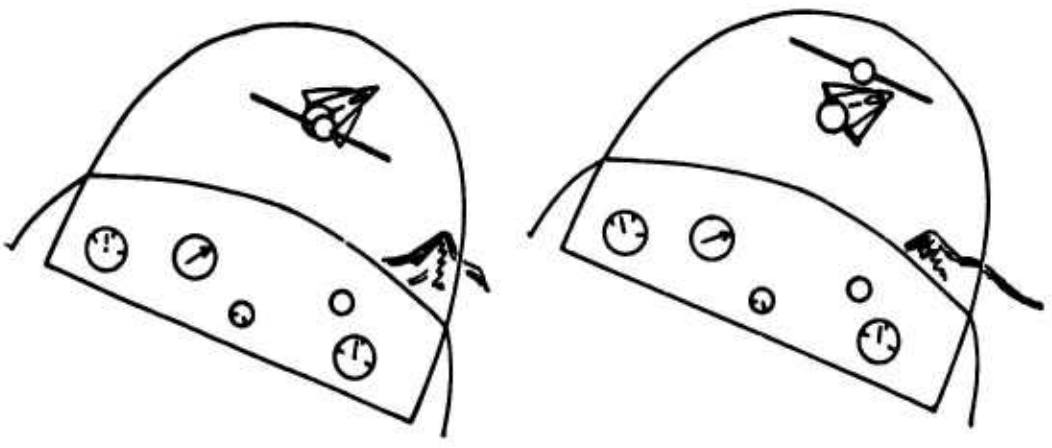


Fig. 4 Definition of Capture Region



a. Target Acquisition

b. Roll and Pitch Up



c. Lock Seeker

d. Satisfy Firing Condition and Fire

Fig. 5 Attacker's View of a Typical Encounter

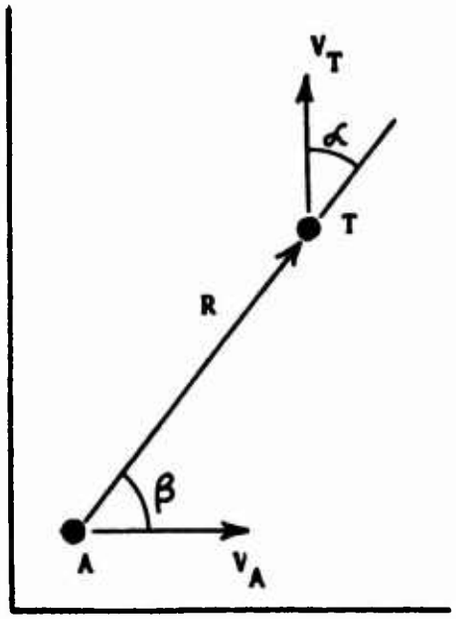


Fig. 6 Two Dimensional Attack Geometry

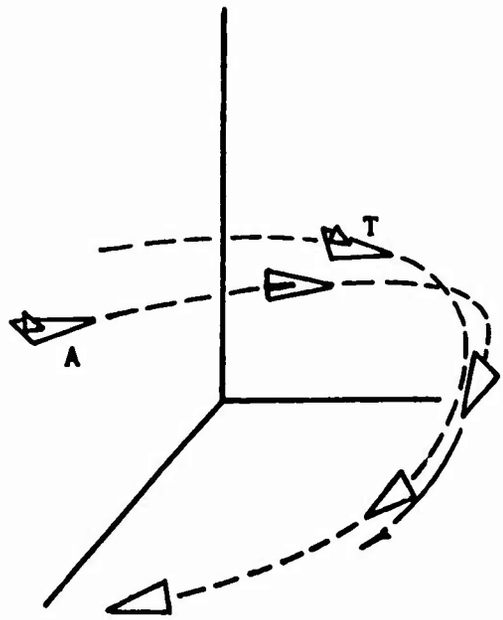


Fig. 7 Flight Paths for a Typical Encounter

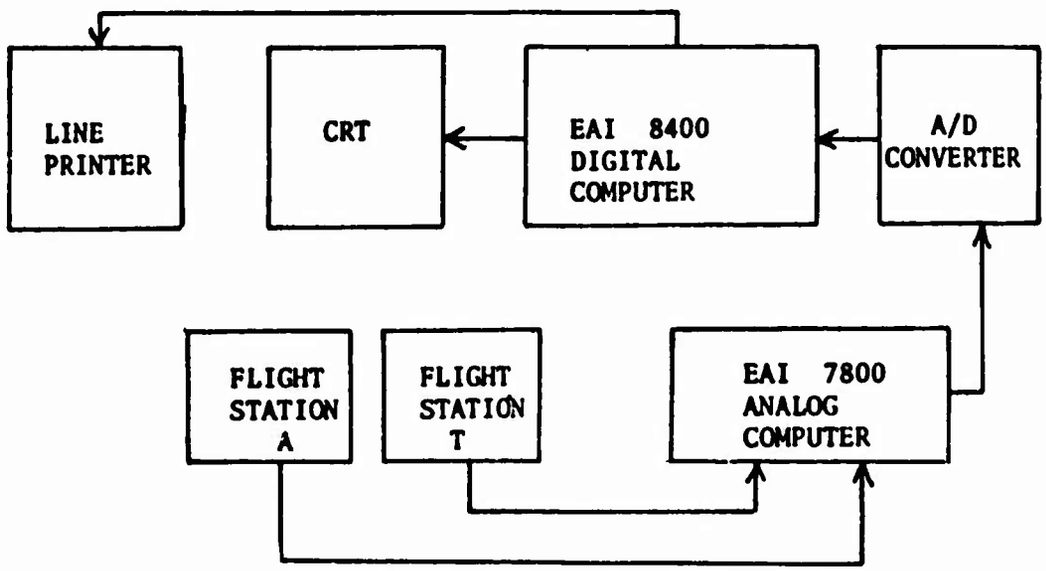
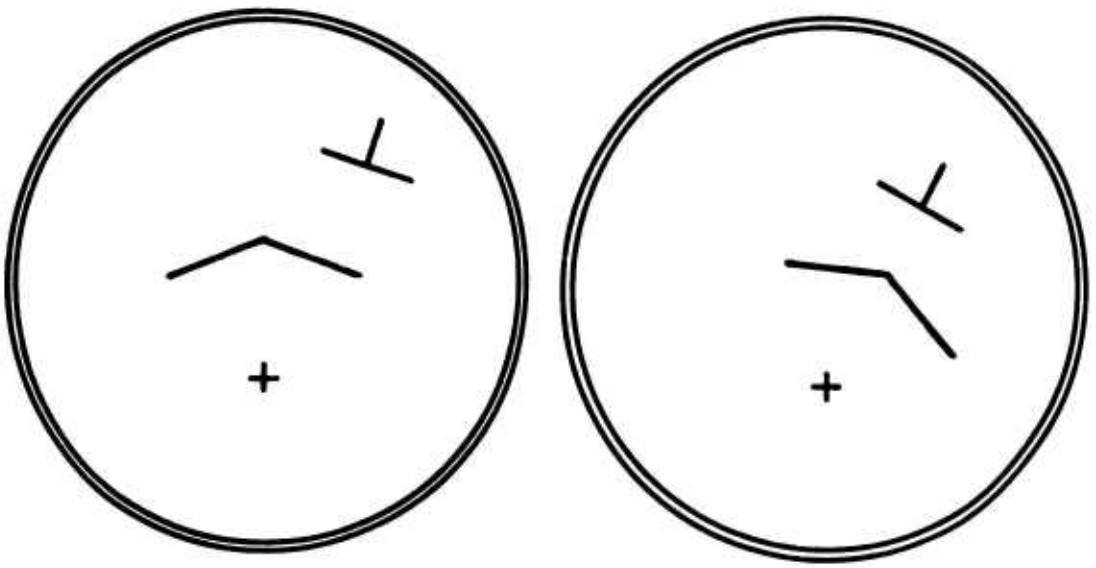
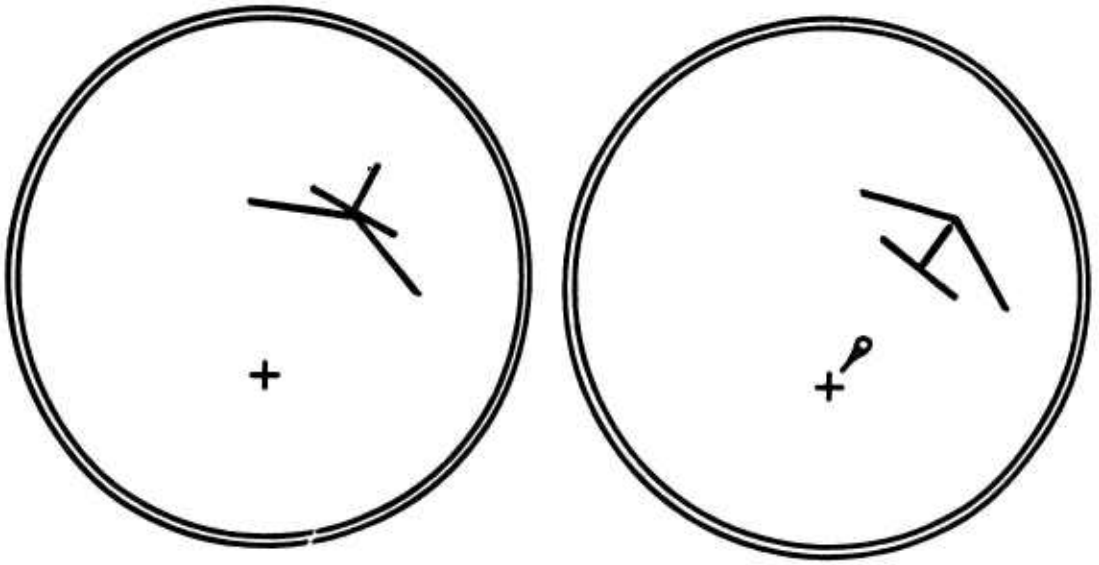


Fig. 8 Hybrid Simulation Block Diagram



a. Target Acquisition

b. Roll and Pitch Up



c. Lock Seeker

d. Satisfy Firing Condition and Fire

Fig. 9 Simulated View of a Typical Encounter

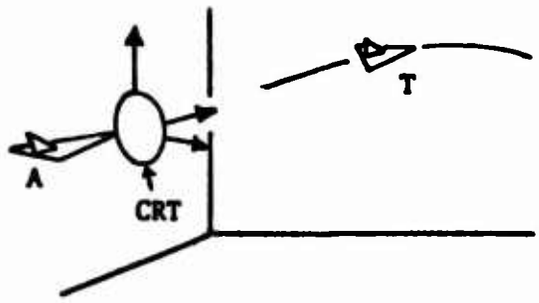


Fig. 10 Display Coordinate System

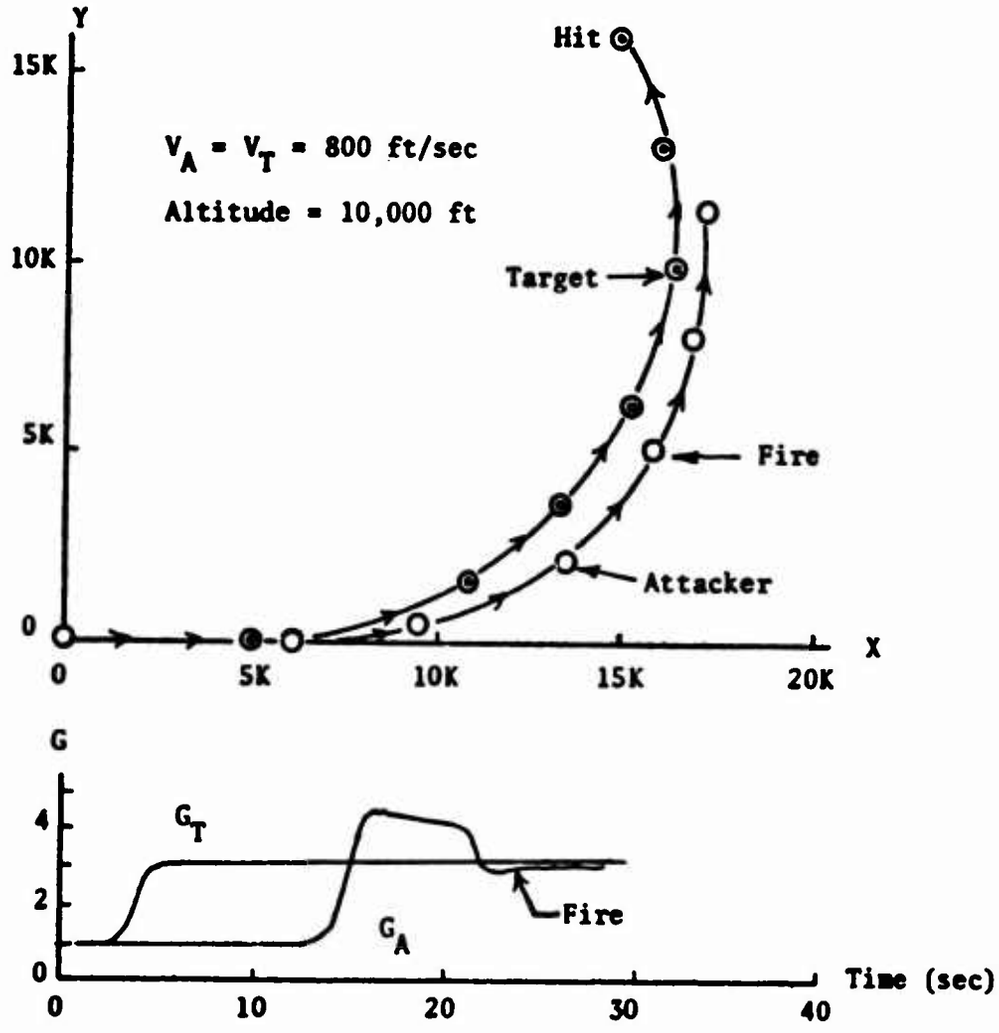


Fig. 11 Simulated Attack: Target in a 3G Horizontal Turn

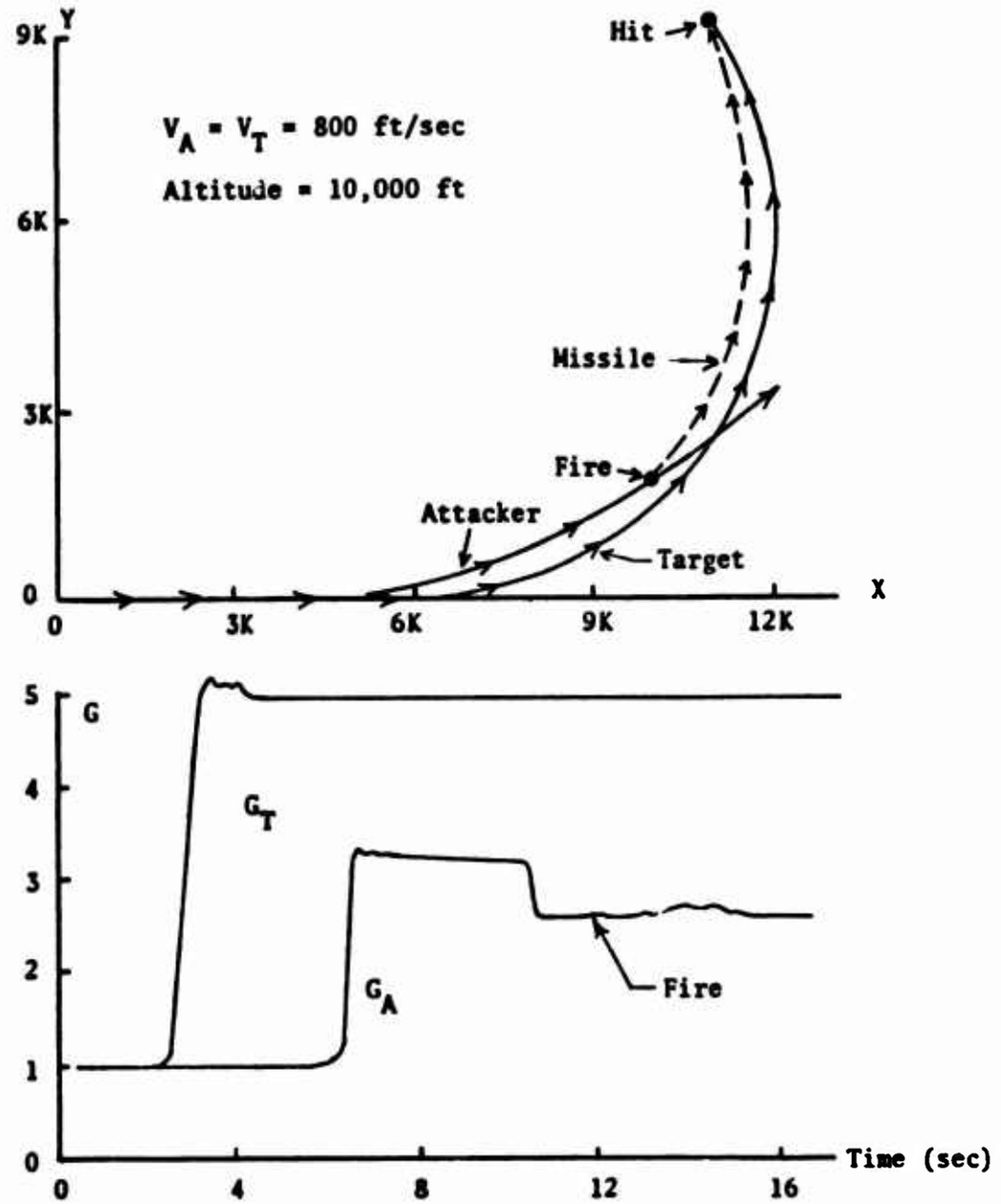


Fig. 12 Simulated Attack: Target in a 5G Horizontal Turn

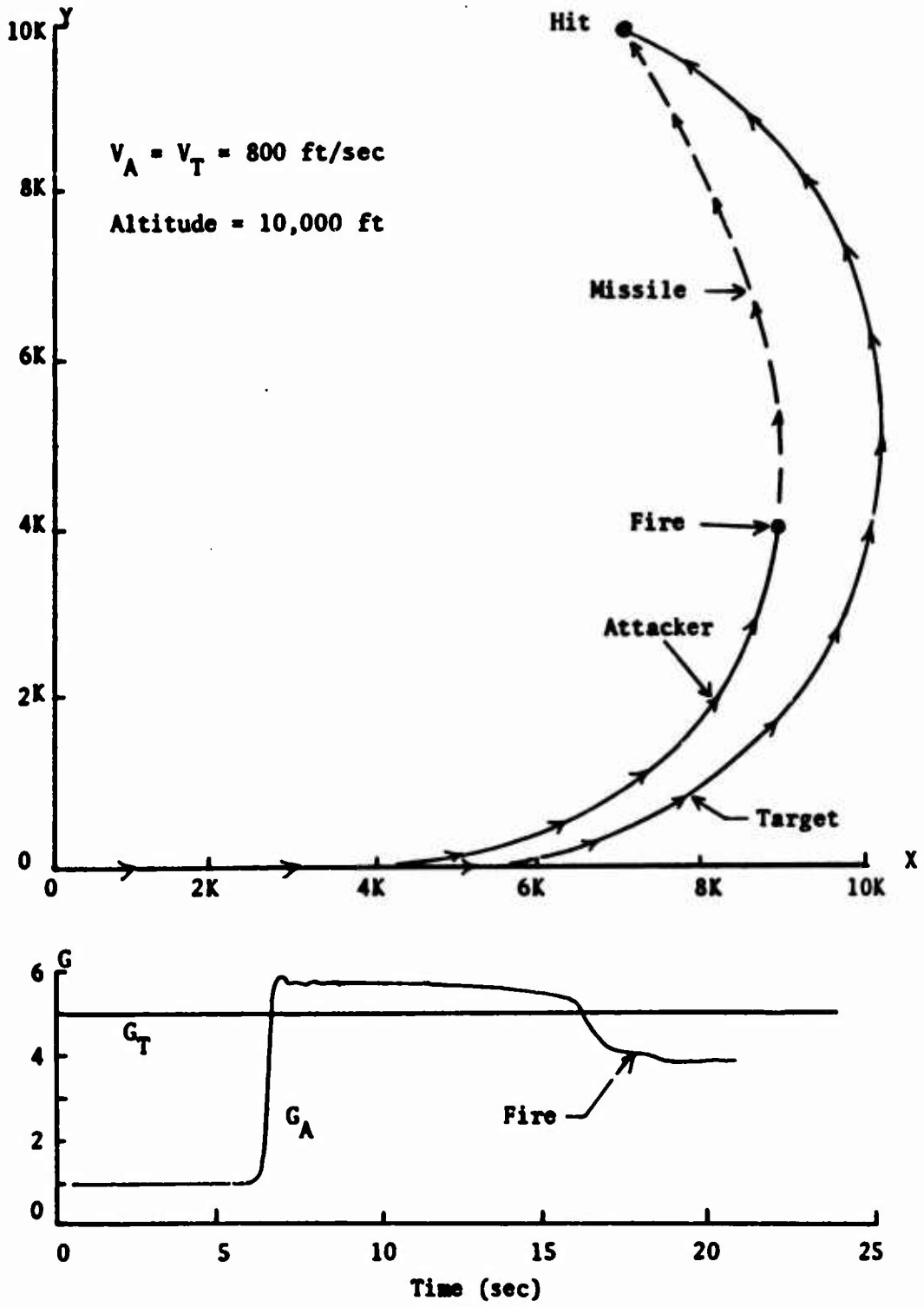


Fig. 13 Simulated Attack: Target in a 5G Horizontal Turn

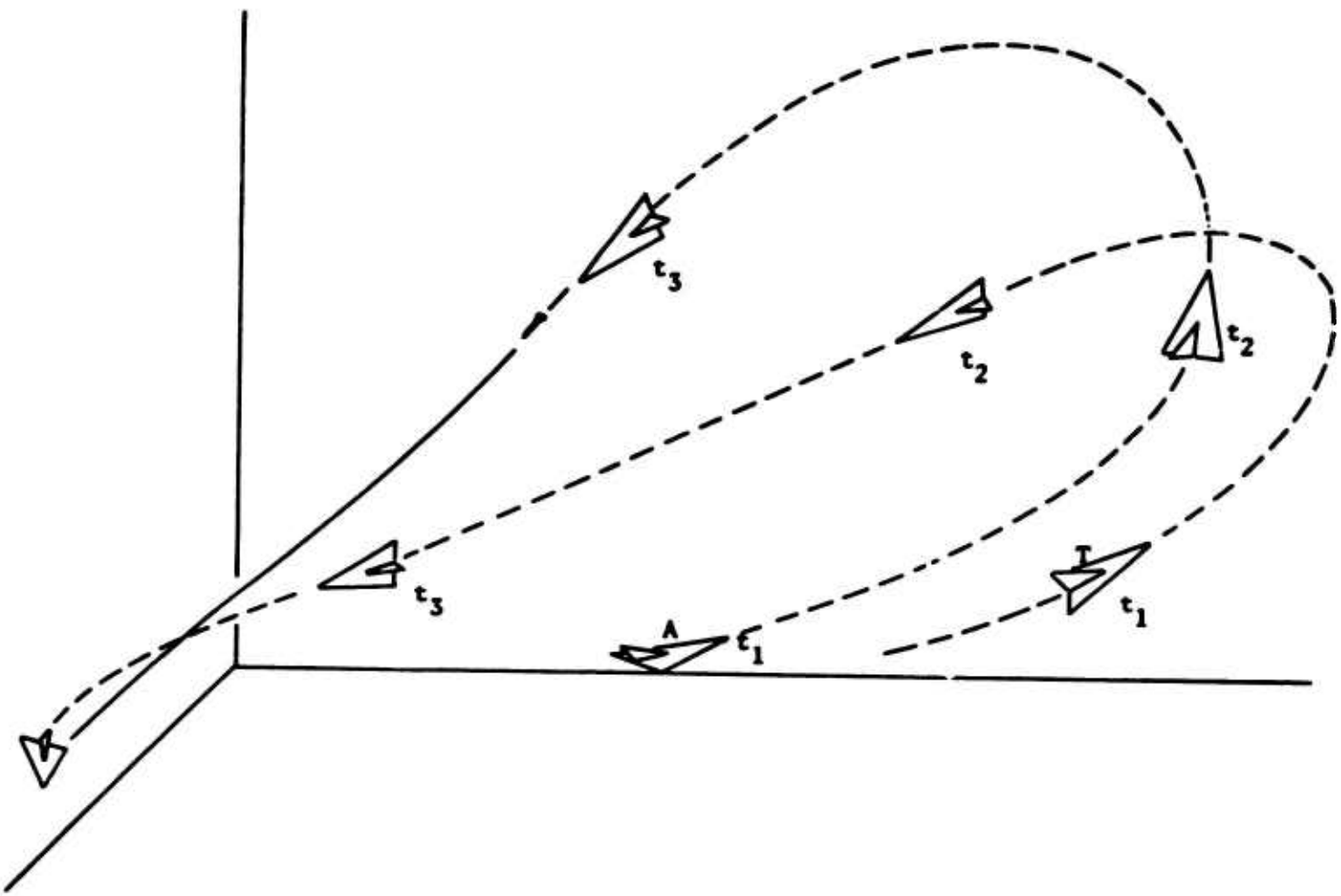


Fig. 14 Simulated Attack: Trajectories for a Three Dimensional Engagement

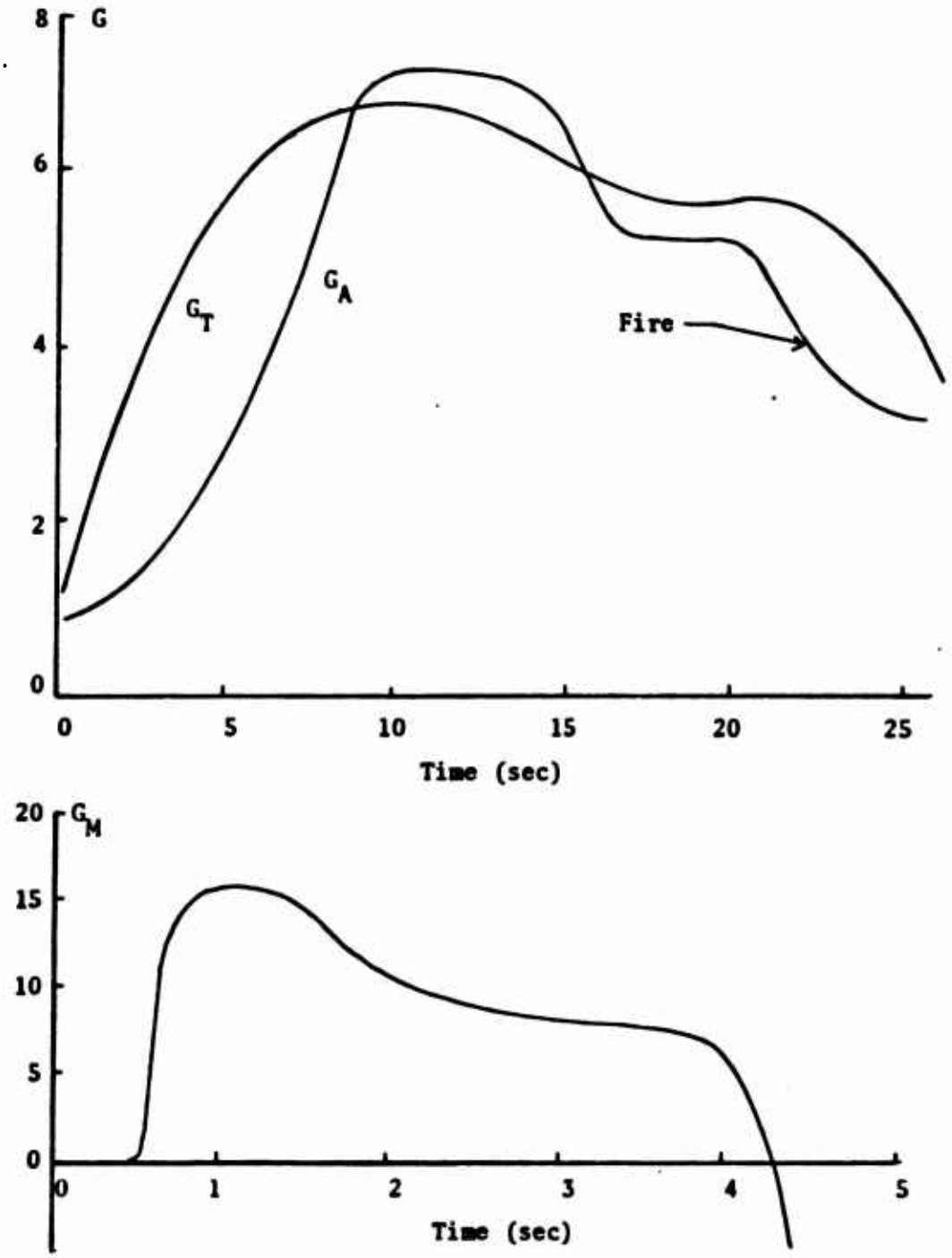


Fig. 15 Aircraft and Missile Acceleration Profiles for the Three Dimensional Engagement

PART II

INTRODUCTION

The conclusions of Part I note that a follow-on study in this same area is being conducted. It is the purpose of this part of the paper to summarize the research efforts of Green and Bergeman Ref (1), as they extended the work of Haas and Puckett, Ref (2).

There are three main areas of interest. First, the missile model and engagement model are expanded to better reflect possible launch situations. Second, the β_{max} prediction algorithm is revised to cover a larger set of launch conditions and the data base for the algorithm is expanded. Finally, the results of implementing the updated models and expanded prediction capabilities are presented.

MISSILE AND ENGAGEMENT MODEL EXPANSION

The missile trajectory analysis was expanded to the three-dimensional situation in order to include target out-of-plane effects on the missile's performance. The proportional navigation guidance system was also modeled in three dimensions. The speed and atmospheric density were modeled as functions of altitude, and the lift, drag, and maximum G's were made functions of missile Mach No. as shown in Figures 1, 2, 3, and 4. The propulsion system was modeled as a drag free boost followed by coast at constant mass until the missile slows to the speed of the target.

It was assumed that seven variables are measurable from the attacking aircraft: the range (R), the range rate ($\dot{R}$), the velocity of the attacking aircraft (V_A), the G loading of the attacking aircraft (G_A), the off-boresight of angle (β), the time rate-of-change of β ($\dot{\beta}$), and the attacker's altitude (h). By assuming (1) that the velocity of the attacker and evader are in parallel planes, (2) that β is zero, and (3) that the altitudes of the attacker and evader are the same; the velocity of the evading aircraft (V_E), the G loading of the evading aircraft (G_A), and the angle between the attacker's velocity and the evader's velocity ($\Delta\gamma$) can be calculated. These target parameters can then be used to compute trajectories and calculate the miss distances associated with missile launches.

The method used to provide an estimate of β_{max} through the heads-up display was basically that of using a first-order Taylor series expansion about various points in the operational environment. The primary measurable variables (relative to the attacker) are range (R), range rate ($\dot{R}$) attacker velocity (V_A) attacker G loading (G_A), and altitude (h); and Table 1 shows the nominal values chosen to represent the rear hemisphere engagement.

Combining the two values of 5 variables yields 32 nominal points to expand about. Table 1 also shows the variable ranges and the perturbation

increments. In keeping with the Taylor series philosophy, each of the primary variables was varied over the given range while the others were held constant, resulting in 1376 data points.

Considering the useful range of β to be + 0.9 radians, the target parameters $\Delta\gamma$, V_E, and G_E were computed for each of the 1376 data points, incrementing by 0.1 radians within the useful range. This resulted in approximately 19,000 sets of target parameters.

Finally, the maximum off-boresight angle (β_m) was determined by simulating missile trajectories for the 19,000 engagement configurations and searching for the maximum lag angle for which the miss distance was less than 10 feet. This resulted in a value of β_m for each of the data points.

To estimate the series coefficients, each set of data of β_m verses one of the primary variables was approximated using a first order least square fit. Figure 5 is an example of the form of the data and also shows the linear fit. The slope of the line is the partial derivative desired for the series expansion. There were, of course, some points at which because of range limitations the missile could not hit the target, and there the partials did not exist.

The fire control solution was implemented using 32 first order series expansions each of which had the five partials and the constant value β_{m_o} The estimate for β_{m_o} was purposely made conservative to account for unmodeled perturbations. The fire control system then utilized the stored expansions by taking the actual range, range rate, attacker velocity, attacker "G", and altitude and entering the matrix at the closest of the 32 nominal points. The difference between the actual variable values and the nominal variable values were used to predict β_m for the situation.

Since it was possible to get β_m values where the missile could not hit the target due to range limitations, an added set of calculations was performed by the fire control system. The estimated quantities:

$$\text{Average Missile Velocity} = V_{avg} = \frac{2V_A + 1790 + 575\,(1-\epsilon^{-h/3000}) + \dot{R}}{2}$$

$$\text{Apparent Range} = R^* = \frac{1 + G_A^2/60}{R}$$

$$\text{Missile G loading} = G_m = \{5.47 - 2.67\ (1 - \epsilon^{-h/3000})\}\ G_A$$

$$\text{Max Time-of-Flight} = t_m = \frac{}{.00007V_{avg}^2\ \epsilon^{-h/3000} + \frac{4478G_m^2}{V_{avg}^2\ \epsilon^{-h/3000}}}$$

$$\text{Time-to-Intercept} = t_c = \frac{R^*}{V_{avg} - V_A - R}$$

were calculated as shown above. The coefficients were obtained by repeated simulations of the missile within the allowable launch region and then adjusting the coefficients until a reasonable relationship was obtained. The fire control system evaluated the above quantities and required that t_e be less than t_m for the launch to be possible. Figure 5 shows the effect of the time-of-flight limits on certain nominal points.

The utilization of this fire control technique is as follows: first, the computer evaluates the ratio of t_e to t_m. If the ratio is greater than one, the displayed chevron flashes to indicate that the target is out of the envelope. When the ratio is less than one, the computer enters the Taylor series expansion and generates a value for β_m is displayed. The cycle is repeated at high speed until the missile is fired.

SIMULATION AND RESULTS

The attacking and evading vehicles, the missile and the fire control system were all simulated in real time on the EAI 8400/231R hybrid computer at the Air Force Flight Dynamics Laboratory. The evading pilot and the attacking pilot were both given CRT displays of the situation as shown on Figure 7. The attacking pilot could launch a missile at any time by depressing a trigger device at his station. The fire control display was cycled every 35 milliseconds.

A total of 206 simulated missile firings were accomplished, and they were evenly spread over the initial conditions shown in Table 2. During the engagements, normal accelerations up to six G's for the attacker and up to eight G's for the evader were noted. Sixty-six missiles were intentionally fired with a "NO-GO" indication (either $\beta > \beta_m$ or $t_e > t_m$), none were fired in situations that the attacking pilot considered to be obvious miss situations. The remaining 140 firings were made with a "GO" indication but near the range and angular boundaries of the capture region.

Of the 66 missiles fired with a "NO-GO" indicator, 25 broke lock prior to blind-range and missed by more than thirty five feet.

Figures 8 and 9 show the miss distance frequency distribution for the missile fired in a "GO" situation. The grouping indicates that further consideration needs to be given to the minimum range restrictions since as the time of flight goes down, the misses seem to get larger. Table 3 shows the cummulative percentages for all the launches with a categorical breakdown.

In summary, the analysis and simulations have indicated that the maximum permissible lab angle, β_m, along with a usable "in-range" indication, can be satisfactorily predicted using parameters measurable from the attacking aircraft.

Further, pilot's who have flown the simulation have indicated that this type of fire control system would be an aid to the attacking pilot even if his weapon system were restricted to boresight acquisition or launch, since there is some correlation between the magnitude of β_m and the potential of the missile to successfully manuver against the target.

Finally, the algorithm is relatively simple and has the capability of being implemented in an on board digital fire control computer since the computational requirements are not extensive.

REFERENCES

1. AF Institute of Technology. The development of a Fire Control System for Off-Boresight Launch of Air-to-Air Missiles, by R.D. Green, and G.M. Bergeman. WPAFB, Ohio. GA/GAM/EE/73A-1, December 1973.

2. AF Institute of Technology. The Analysis and Simulation of a New Fire Control Technique for Launching Air-to-Air Missiles by C.D. Haas, and L.D. Puckett. WPAFB, Ohio. GGC/EE/73-4, June 1973.

TABLE I

Variables and Nominal Values

variable	variable range		nominal values		perturbation increments
	low	high	low	high	
R (ft)	$2500 \leq R \leq 5500$	$5500 \leq R \leq 8500$	4000	7000	500
$\dot{R}$ (ft/sec)	$-500 \leq \dot{R} \leq 0$	$0 \leq \dot{R} \leq 200$	-200	+100	50
G_A (G's)	$0 \leq G \leq 3.5$	$3.5 \leq G \leq 7.0$	2	5	0.5
V_A (ft/sec)	$500 \leq V \leq 900$	$900 \leq V \leq 1400$	700	1100	50
h (ft)	$0 \leq h \leq 20,000$	$20,000 \leq h \leq 40,000$	10,000	30,000	2000

TABLE II

Parameter Combinations Used
in Simulated Missile Launches

altitude (ft)	V_A (ft/sec)	V_E (ft/sec)
5,000	1000	1000
5,000	600	600
10,000	800	800
10,000	1000	1200
15,000	1000	1000
15,000	1200	1200
25,000	1200	1200
25,000	600	700
35,000	1300	1300
35,000	600	600

TABLE III

Cummulative Percentages of Miss Distance for Several Launch Categories

miss distance (ft)	all missiles launched %	missiles launched with a "go" indication		
		total %	time-of-flight greater than 5.02 seconds %	near maximum range %
0 - 4.99	17.48	20.71	24.29	36.36
0 - 9.99	39.32	44.28	57.14	68.18
0 - 14.99	42.23	57.85	74.29	86.36
0 - 19.99	56.31	74.28	90.00	100.00
0 - 24.99	78.16	89.28	98.57	
0 - 29.99	84.47	96.43	98.57	
0 - 34.99	87.38	99.29	100.00	

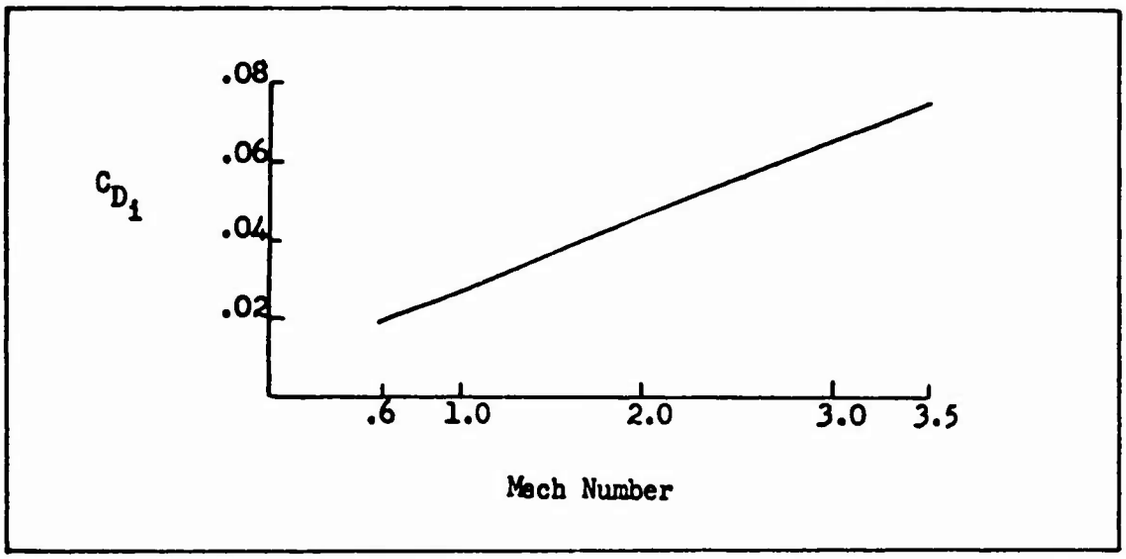


Fig. 1 Coefficient of Induced Drag

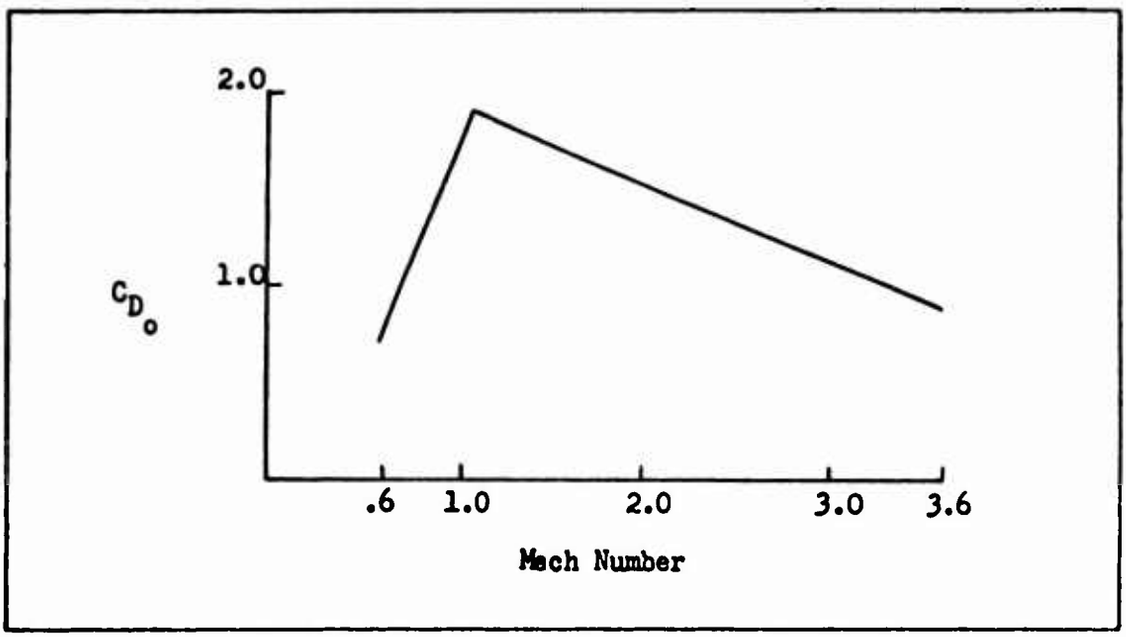


Fig. 2 Profile Drag

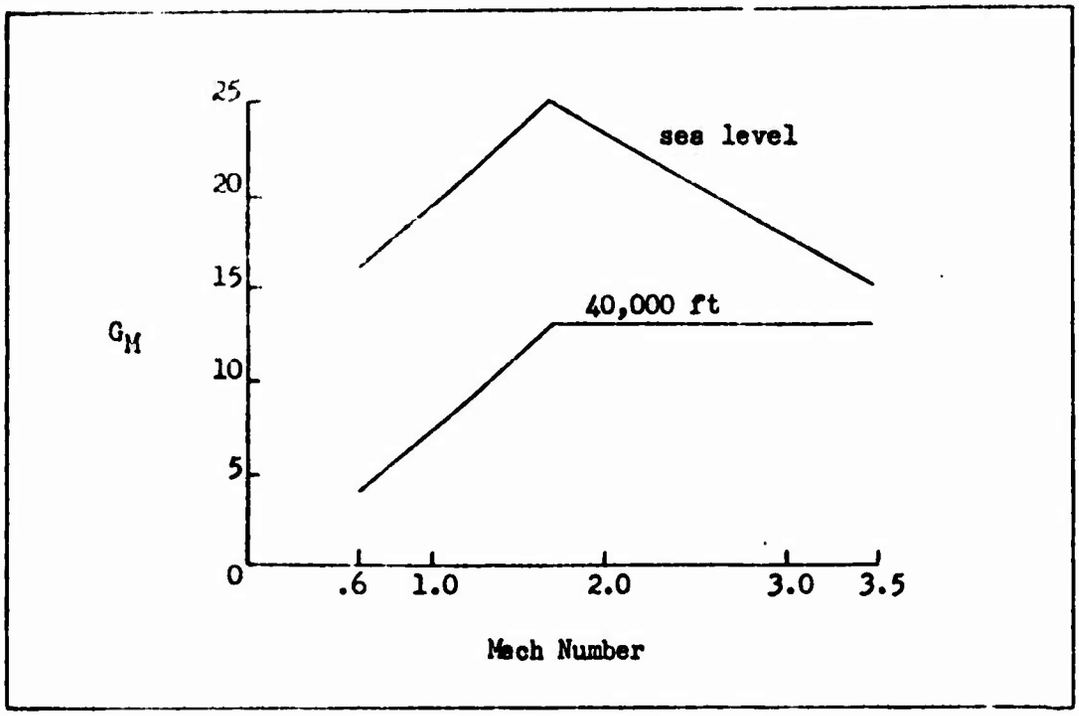


Fig. 3 G's of the Missile vs Mach Number

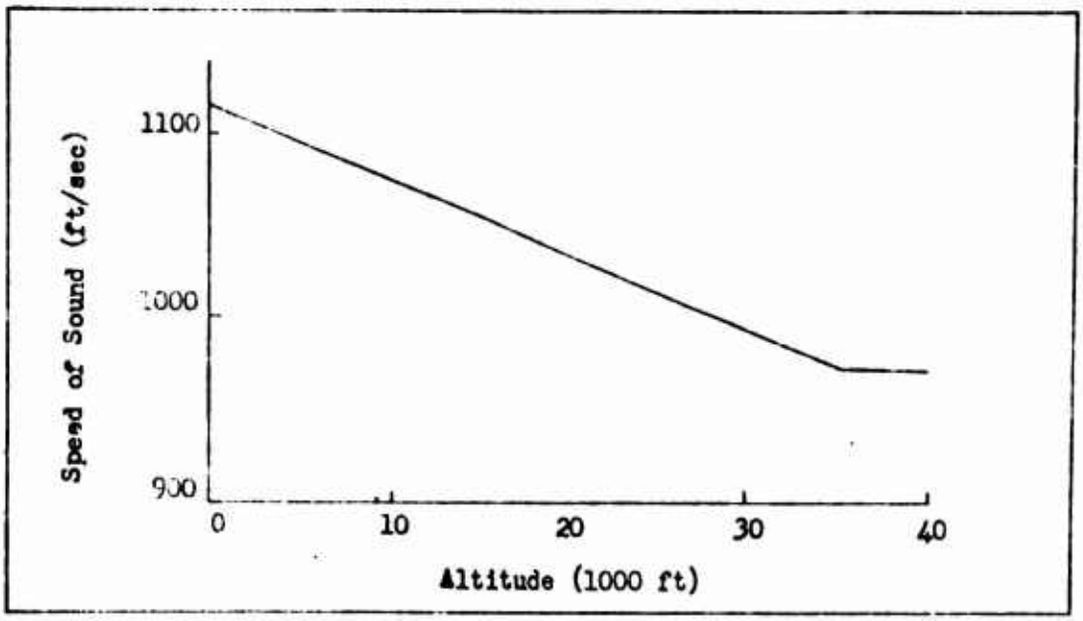


Fig. 4 Speed of Sound as a Function of Altitude

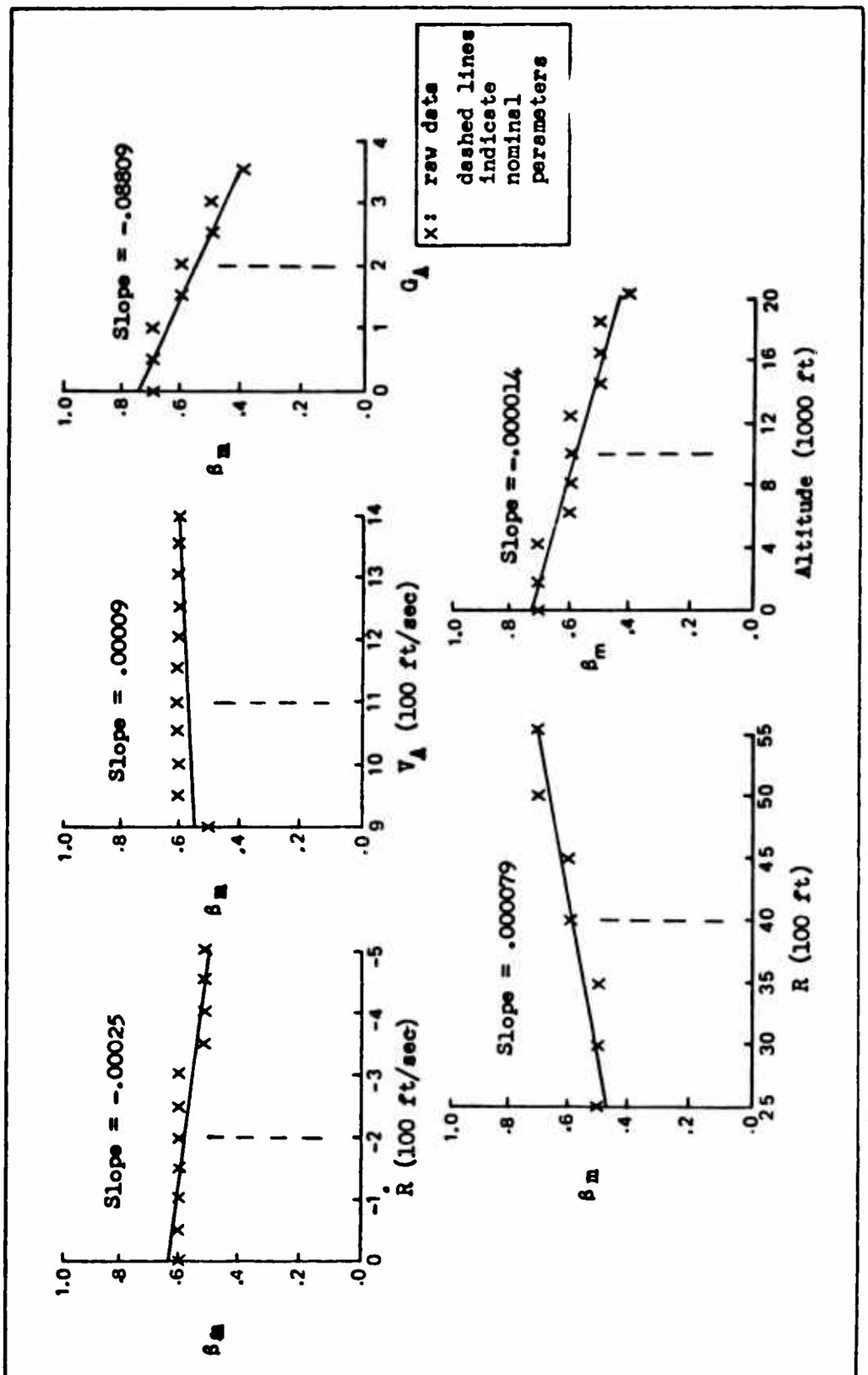


Fig. 5. β_m vs the Five Primary Variables for Nominal Point Five

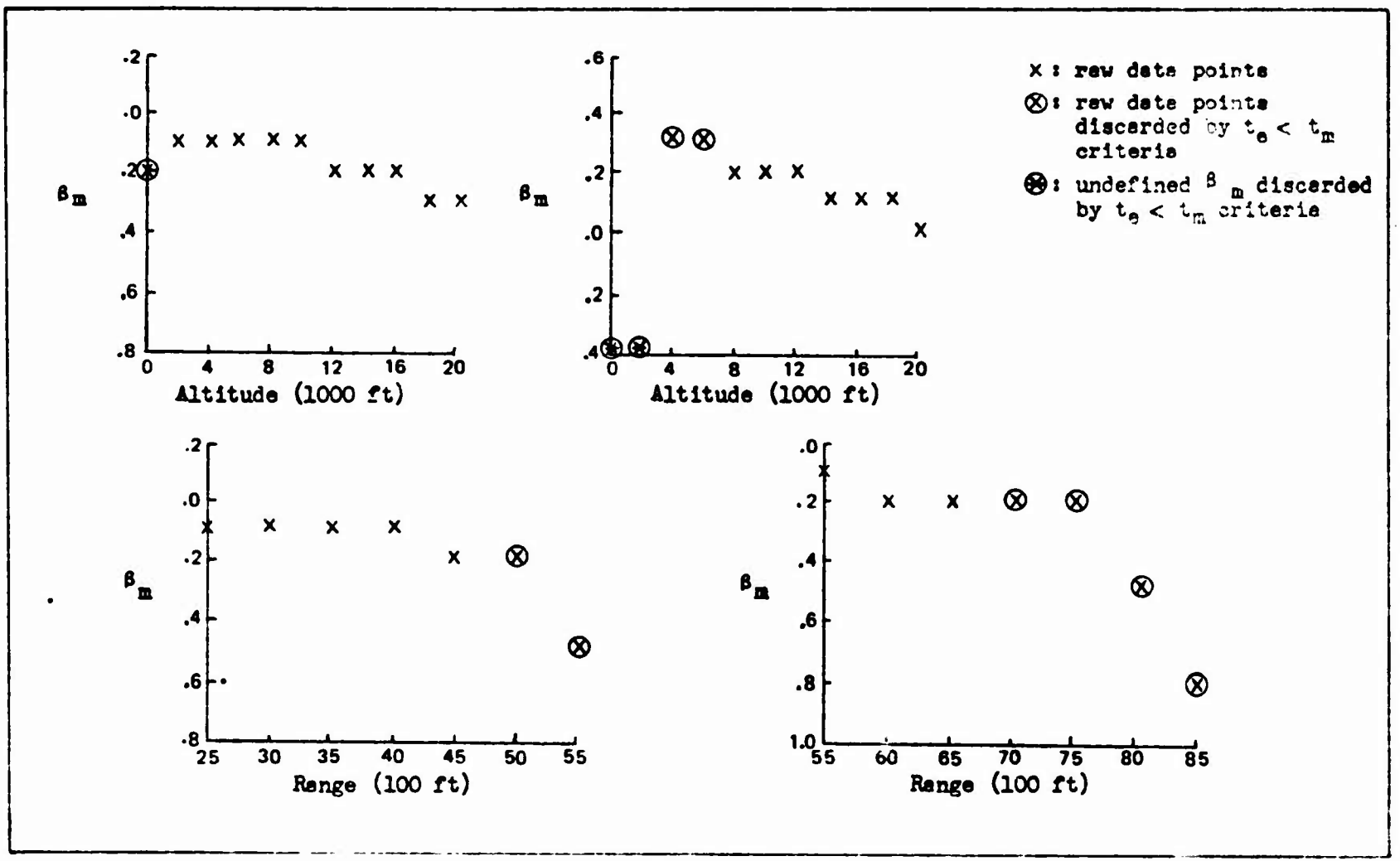


Fig. 6 . Example of $t_e < t_m$ Criteria For Some Nominal Points

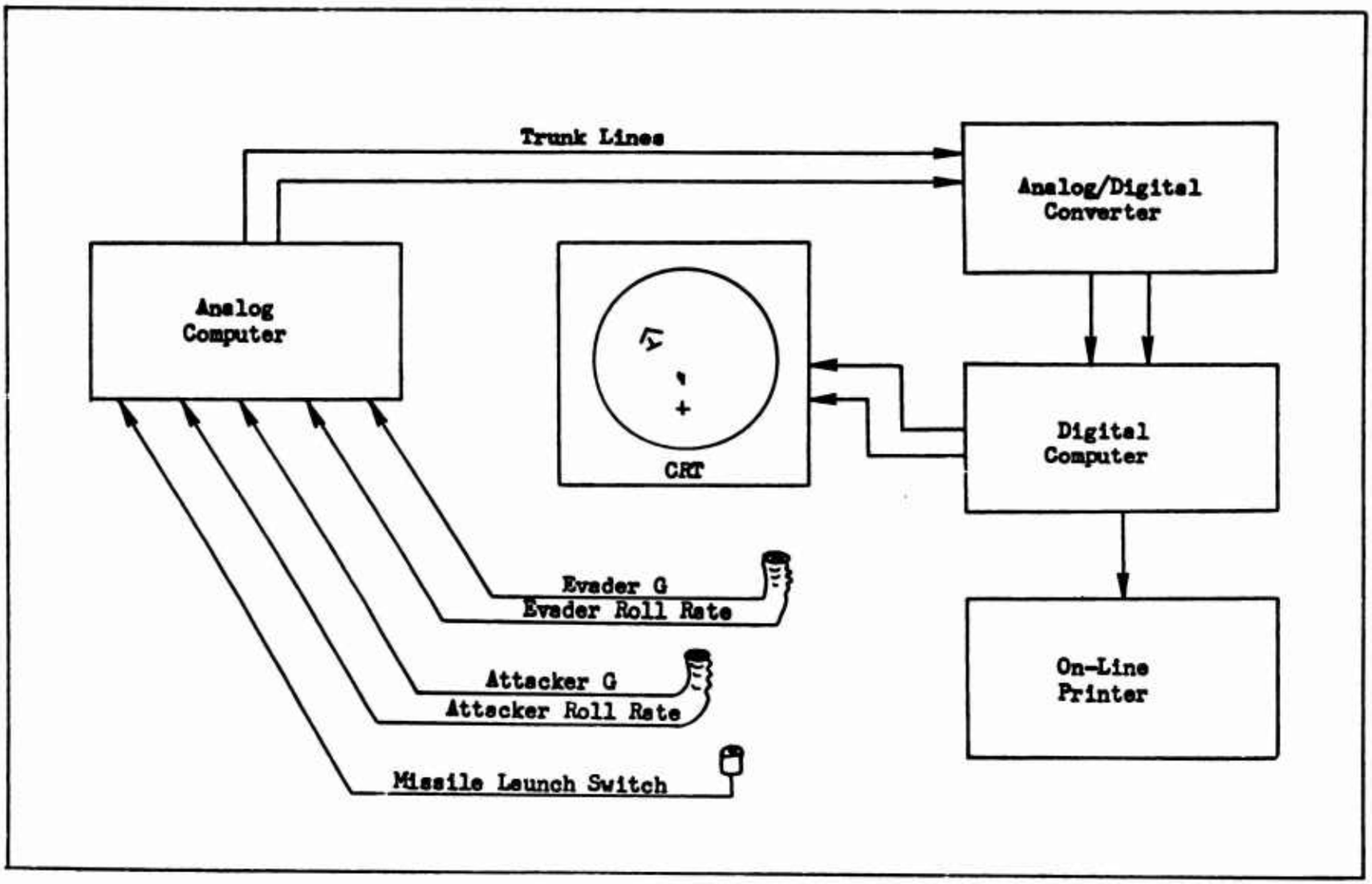


Fig. 7 Simulation Flow Signals

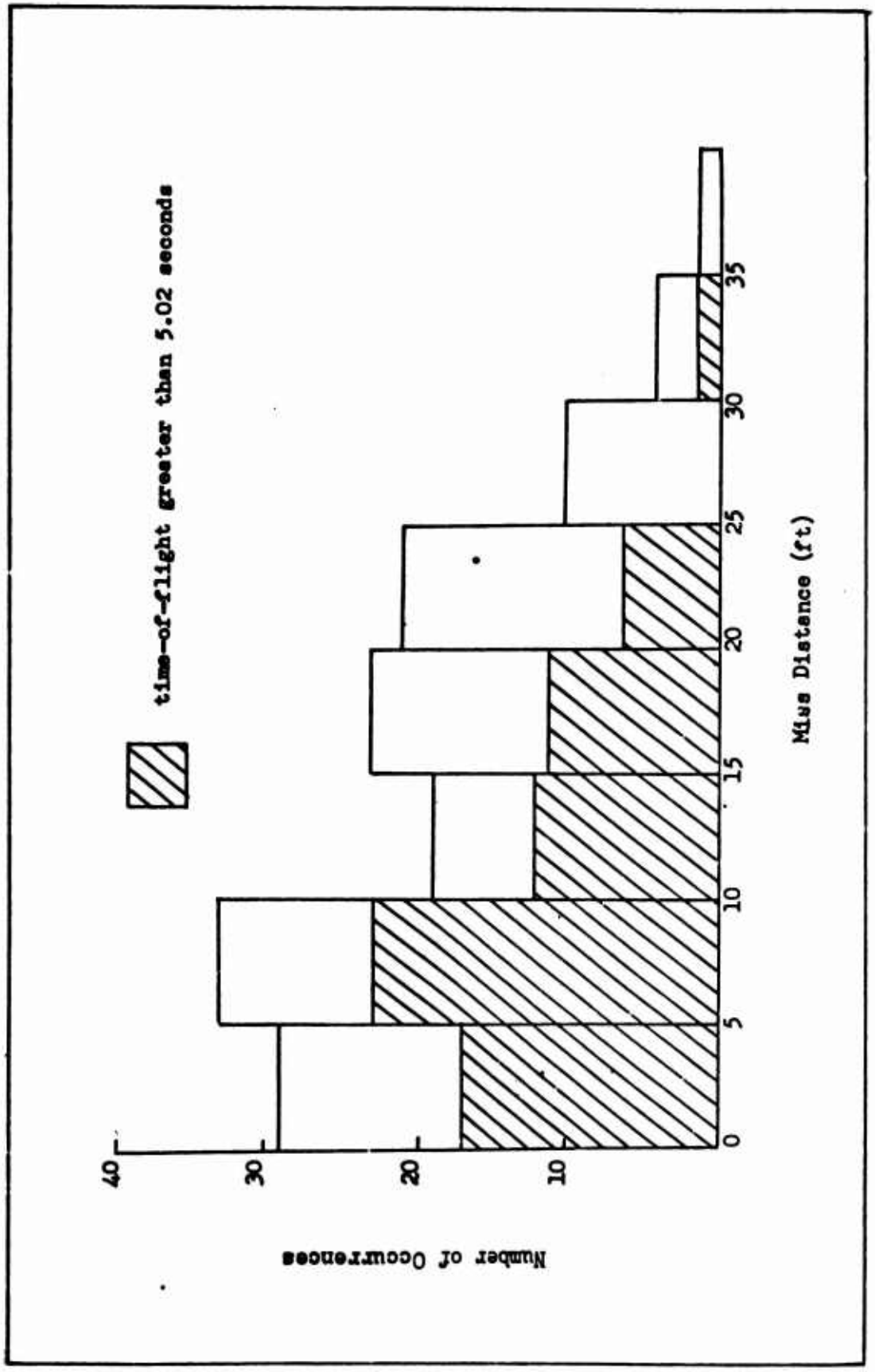


Fig.8 . Miss Distance Frequency Distribution of Missiles Fired with a "Go" Indication

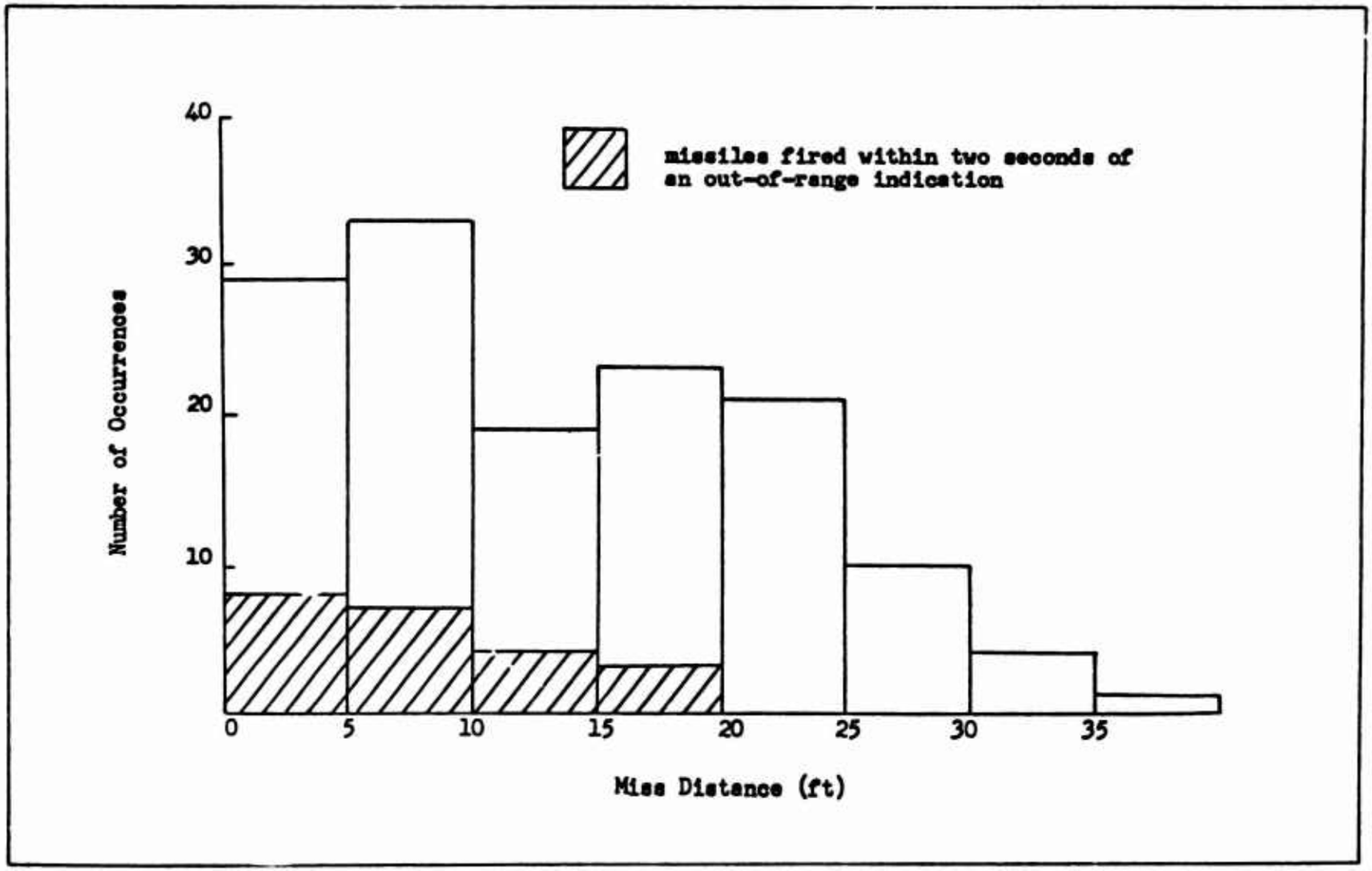


Fig. 9 . Miss Distance Frequency Distribution of Missiles Fired with a "Go" Indication

Preface

The following study was developed in order to produce a more effective alternative to the existing procedures of obtaining the exact ephemeris data of spatial bodies (missiles, satellites, etc.) under observation. The study utilized readily available pollution levels rather than classified data because the data represented many different types of random signals or waves. This data represented non-predictable natural phenomena and the only bias was inherent in the collecting devices.

It can be proven that the autoregressive-moving average models exhibit the ability to track and predict this data and the method should be appropriate when applied to classified data of less abundance.

This paper satisfies two needs. First, to supply linear models to help in solving random data sets such as the air and water pollution problems described and second, to satisfy a need for a "cook-book" approach to modeling complex missile ephemeris data.

If the reader has data points in sufficient quantity (50 or more, preferably more) that they would like to compare with a model using the autoregressive-moving average models, contact Capt Thomas S. Lee, SAMSO/MNNI, Norton AFB, CA 92409. Autovon 876-5977.

TIME SERIES MODELING OF URBAN POLLUTION LEVELS

by

Thomas S. Lee, Captain, USAF
Dr. R. Bethke, Wright State University

INTRODUCTION

In the study of pollution problems one is often confronted by pollution data observed at equispaced times at some geographic location. Questions arise as to the nature of the dynamics of a given pollutant level. For example, does the dynamic behavior of a pollutant level change with time and does it depend on geographic location? Do different pollutants exhibit different dynamic behavior? Is it possible to predict future pollution levels based on past dynamic behavior? These questions and others can only be answered if the pollution level dynamics can be adequately described or modeled.

The modeling of pollution level dynamics can be done in a variety of ways. Perhaps the most simple description is the mean and the variance of the data. While simple, this description is far from complete. The most complete model of a pollutant level would be its description as a function of all the pollution generation and transmission variables which effect it. While complete, this is an extremely formidable task. Another method which is adequately descriptive and yet not impossibly difficult will be demonstrated here, the method of autoregressive-moving average time series modeling.

The use of this time series modeling technique will be demonstrated on several 24-hour average air pollutants from Bayonne, New Jersey and several hourly water quality measures from the Great Miami River in Dayton, Ohio.

MODEL

Autoregressive-moving average time series models view the observed levels of a pollutant about its mean value as the output of a linear system. The input to the system is assumed to be normally distributed, uncorrelated random variables (normal white noise). The system configuration to be modeled is shown in Figure 1.

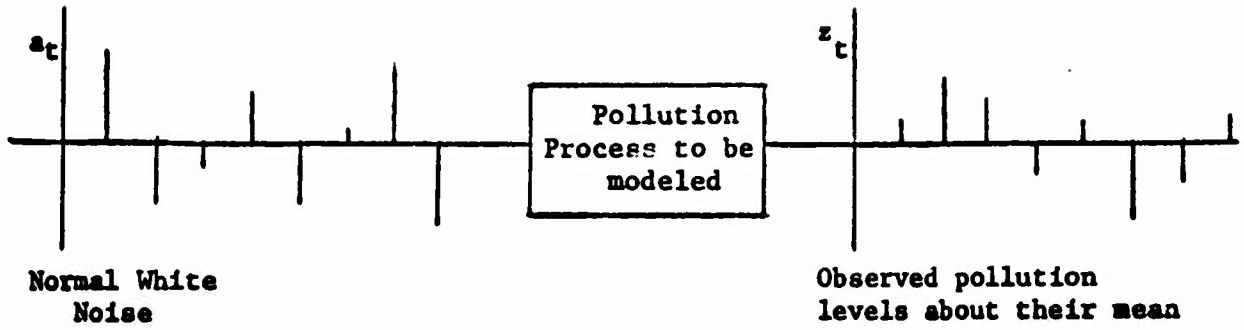


Figure 1. SYSTEM CONFIGURATION

The system shown is described by a linear difference equation, the autoregressive-moving average model:

$$z_t - \phi_1 z_{t-1} - \ldots - \phi_p z_{t-p} = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$$

where the z_t's are the observed levels of a pollutant about its mean, the a_t's are normal white noise, and the ϕ's and θ's are model parameters to be estimated.

Note that the model contains only observed values of the system output, in this case the pollution levels about their mean, and no observed values of possible system inputs, such as weather conditions, etc. The a_t's are calcuated recursively from the model and the z_t's. This makes this model an especially useful tool when the actual system inputs are unknown or very complicated and only the output levels are available.

The fitting of the model to the observed data involves three basic steps. The first is process identification, that is how many z and a terms to use in the model. The second step is estimation of the ϕ and θ parameters to best make the identified model fit the data. The last step is diagnostic checking of the fitted model to detect any shortcomings.

PROCESS IDENTIFICATION

Identification of the appropriate autoregressive-moving average model is accomplished by computation and inspection of the data autocorrelation

(normalized autocovariance) and partial autocorrelation functions. Figure 2 shows these functions computed for hourly river pH data.

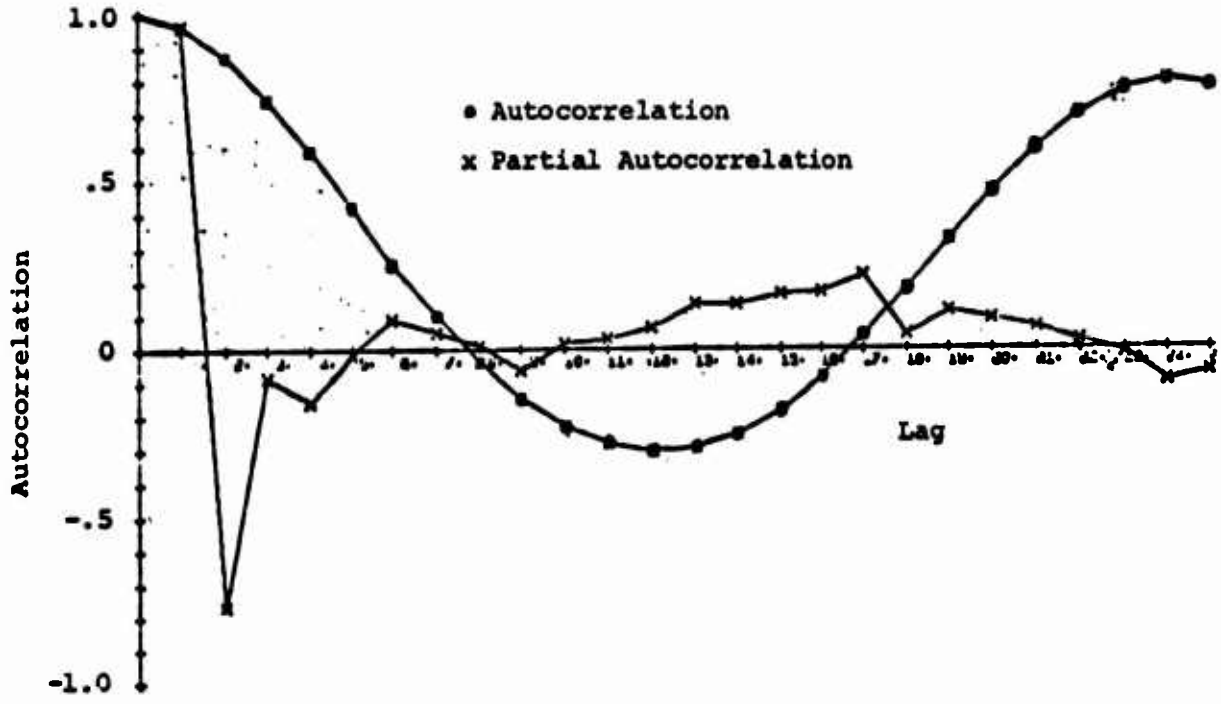


FIGURE 2. pH AUTOCORRELATION AND PARTIAL AUTOCORRELATION FUNCTIONS

Inspection of Figure 2 shows the autocorrelation function oscillates and decays while the partial autocorrelation truncates after two lags. A decaying autocorrelation indicates an autoregressive model is required to fit the data. An autoregressive (AR) model of order p has the form

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + a_t$$

where the present pollution level z_t depends on the previous p pollution levels

and the present input a_t. Truncation of the partial autocorrelation function after two lags indicates the required AR model is of order two

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

Decay of the partial autocorrelation and truncation of the autocorrelation after q lags would indicate a moving average (MA) model of order q.

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 d_{t-2} - \ldots - \theta_q a_{t-q}$$

Here the present pollution level z_t is a function of the present and the q previous system inputs.

If both the autocorrelation and partial autocorrelation functions exhibit truncation and decay the required model is autoregressive-moving average (ARMA)

$$z_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$$

Autocorrelation functions not only help identify the required model but give insight into the amount of "memory" in the data. Figure 3 shows the autocorrelation functions of daily NO and NO_2 air pollution levels.

Figure 3 shows the NO autocorrelation decays far more slowly than that of NO_2. This shows that while the present level of NO_2 is not well correlated to levels more than 2 lags (days) previous, NO is. One might say NO "remembers" what it was in the past far better than NO_2.

Autocorrelation functions are also useful in detecting and displaying periodic components in the data. Data which is strictly periodic will yield a periodic autocorrelation function having the same period. The pH autocorrelation in Figure 2 shows periodic behavior having a 24-hour lag period. This shows the pH data has a strong daily cycle and that pH levels 24 hours apart are very similar.

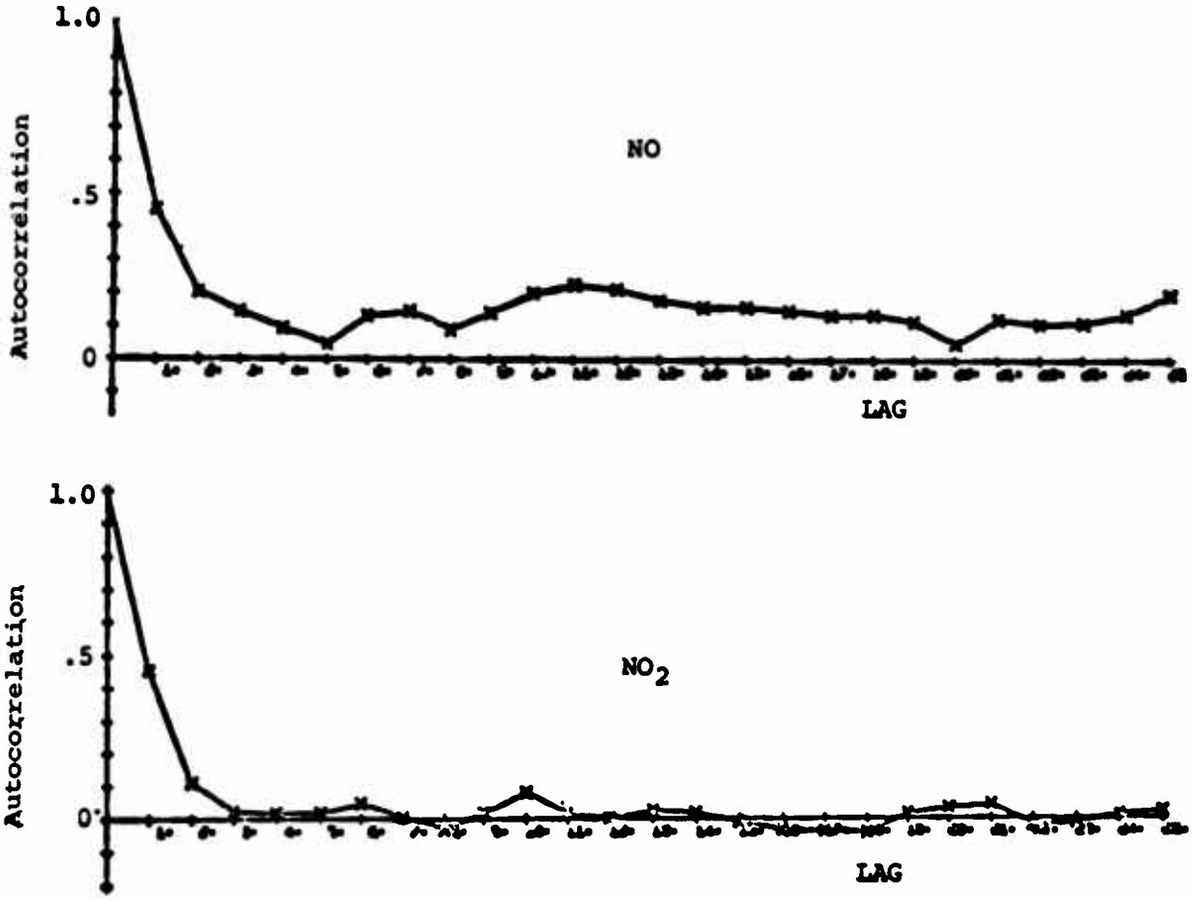


FIGURE 3. NO AND NO_2 AUTOCORRELATION

Figure 4 shows the autocorrelation of river water temperature. Note the 24-hour lag cycle which is now superimposed on a much lower frequency oscillation. This low frequency temperature oscillation is the yearly cycle and it is much more evident in temperature than in pH.

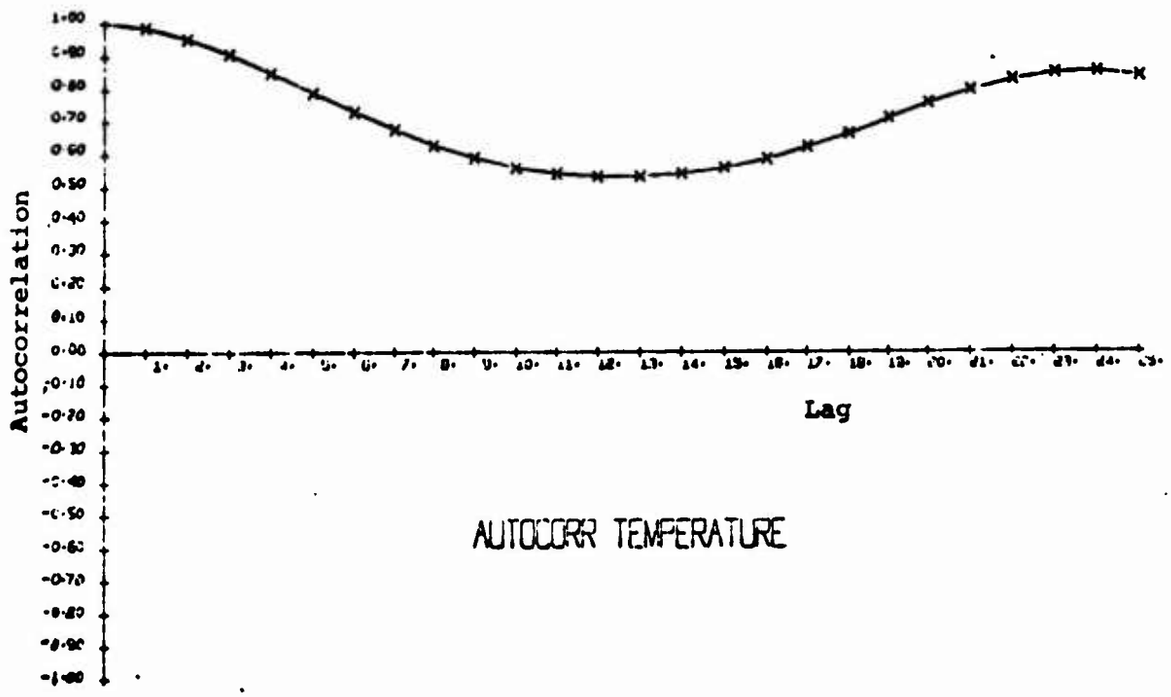


FIGURE 4. WATER TEMPERATURE AUTOCORRELATION

The data to be modeled is assumed to be stationary, that is it doesn't change its dynamic character with time. Data containing very low frequencies and hence having autocorrelation functions which decay very slowly, may over a short time appear to be drifting and nonstationary. These very low frequency trends are usually removed prior to modeling with a first difference, high pass filtering operation on the data. That is adjacent data points are subtracted from each other to form a new data set.

Figure 5 shows the autocorrelation of the temperature data after it was high pass filtered with a first difference operation. Note the low frequency yearly cycle has been filtered out and now the daily frequency predominates.

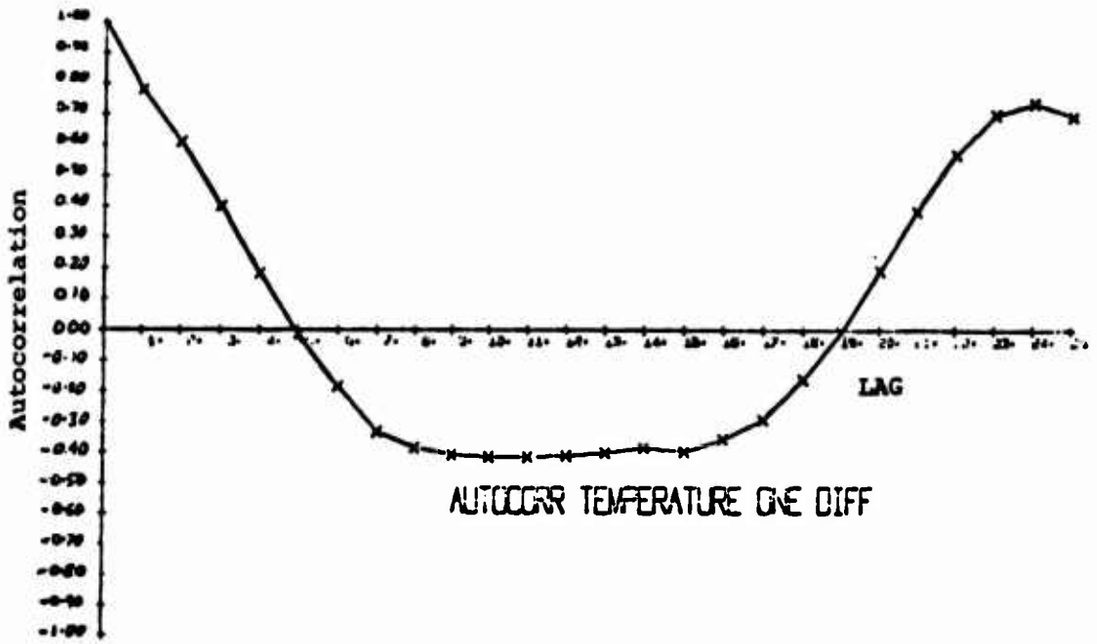


FIGURE 5. DIFFERENCED WATER TEMPERATURE AUTOCORRELATION

PARAMETER ESTIMATION

Parameters for the identified models are estimated using linear or non-linear regression, as called for by the model. Regression is accomplished by minimizing the sum of the squared errors. Errors (residuals) are the differences between actual observed z_t data and the model estimates $\hat{z}_t$. For the general ARMA case

$$\hat{z}_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$$

and the error at time t is

$$z_t - \hat{z}_t = a_t$$

For the air pollution data, stationary and normality considerations were accomodated by doing a log and a first difference operation on the data. Parameters were then estimated for the identified models. Examples of the fitted models are:

Nitrogen Dioxide (NO_2)

$$z_t = a_t - .510a_{t-1} - .363a_{t-2} - .124a_{t-3} + .016a_{t-4}$$

Nitric Oxide (NO)

$$z_t = a_t - .482a_{t-1} - .285a_{t-2} - .024a_{t-3} - .129a_{t-4}$$

Examples of fitted water quality models are:

Dissolved Oxygen

$$z_t = 1.746z_{t-1} - .818z_{t-2} + a_t$$

pH

$$z_t = 1.701z_{t-1} - .765z_{t-2} + a_t$$

DIAGNOSTIC CHECKING

Diagnostic checking of the fitted model involves several aspects. Because identification of the model to be used is not always clear cut, models similar to the most likely one may be fit as a check. Residual variance is compared among the models and the minimum variance indicates the best fit and hence the preferred model.

Another diagnostic check is to examine the errors for serial correlation. If the model fits well it should extract all the serial information from the data and leave the errors uncorrelated. This can be checked by computing an autocorrelation function of the errors and seeing if the values are near zero.

For example the error autocorrelation values for the Nitric Oxide model above are -.0023, .0007, -.0094, -.0048, -.0009,... Note the values are very close to zero and hence indicate virtually all the information in the data was put into the model. This good fit was generally found to be the case with the models fit to other pollution data.

PREDICTION

Once a model has been fit to the data it can be used to predict future data values, in this case pollution levels. Figure 6 shows the one day ahead predictions for NO using the model above and the actual observed values.

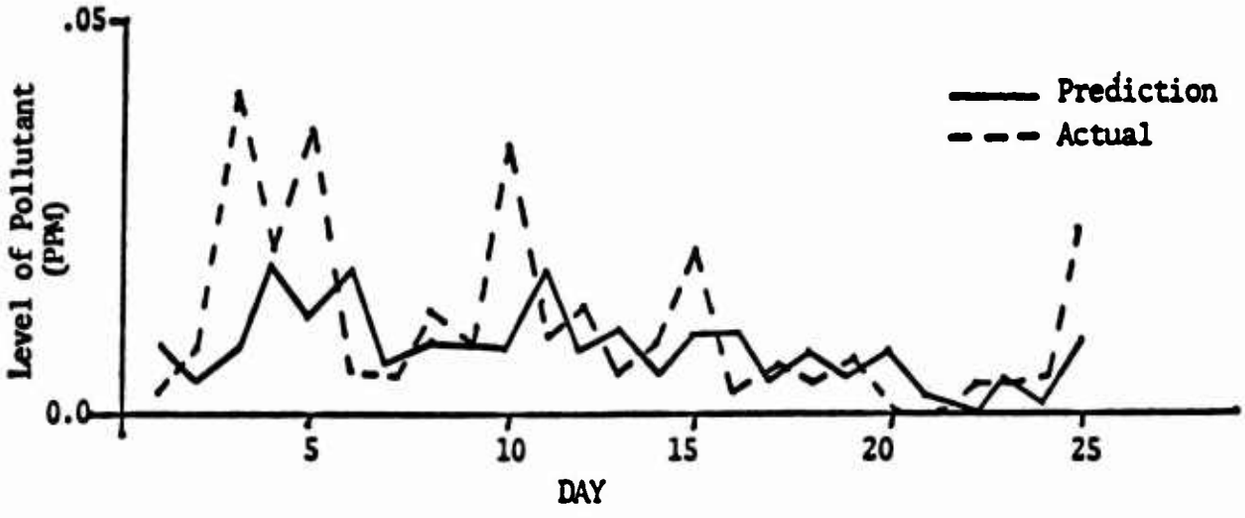


FIGURE 6. ONE DAY AHEAD NITRIC OXIDE (NO) PREDICTIONS AND ACTUAL VALUES

Note that the predictions track the actual values but do not predict perfectly. The ability to predict future values is determined by two factors; how well the model extracts the serial information from the data (fits), and how much serial information is in the data to begin with. Time series data has two components, deterministic information and random noise. If the data is highly random then no matter how well the model fits, it will not predict well. That is, one can predict deterministic trends but not random variations.

CONCLUSION

Autoregressive-moving average time series models can be used to sucessfully describe dynamic level behavior of a wide variety of pollutants. This is a fundamental step in the study of pollutant behavior and the causes of it. These models offer a compact descriptive format in which to store vast amounts of observed data. In addition the fitted models give insight into the nature of the pollutants, the problems they cause, and possible solutions.

AIR FORCE APPLICATIONS

Further applications of the technique may be in the areas of: drone guidance and control, terrain following control, navigation system integration (doppler and celestial plus pilotage), orbit prediction and detection satellite vs missile (determine mass of vehicle in orbit), anti-hunt controls on automatic systems with automatic limit adjustments, backtracking foreign missile trajectory data to obtain exact launch point.

This procedure was developed in order to arrive at a simpler and more time sensitive adaptive technique for expressing random phenomenona exiguously in predictive computer programs.

LIST OF REFERENCES

1. Stern, A.C. (1962), Air Pollution, New York: Academic Press.

2. Bibbero, R.J. (1971), "Systems Approach toward Air-Pollution Control," IEEE Spectrum, October, 20-31.

3. Lathi, B.P. (1968), Random Signals and Communication Theory, Pennsylvania: International Textbook Company.

4. Green, M.H. (1969), Technical Bulletin (A-69-1), New Jersey State Department of Health.

5. Box, G.E. and Jenkins, G.M. (1969), Time Series Analysis, Forecasting and Control, California: Holden-Day.

6. Merz, P.H., Painter, L.J., and Ryason, P.R. (1971), "Aerometric Data Analysis-Time Series Analysis and Forecast and an Atmospheric Smog Diagram," Atmospheric Environment, Vol. 6, 319-342.

7. Marquardt, D.L. (1963), "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," J. Soc. Indust. Applied Math., 2, 431-441.

8. Spiegel, M.R. (1961), Theory and Problems of Statistics, New York: Schaum Publishing Co.

9. Jenkins, G.M. and Watts, D.G. (1968), Spectral Analysis and Its Applications, California: Holden-Day.

10. Hogg, R.V. and Craig, A.T. (1970), Introduction to Mathematical Statistics, London: MacMillan.

DETERMINATION OF IN-FLIGHT PILOT PARAMETERS USING A NEWTON-RAPHSON MINIMIZATION TECHNIQUE

Daniel L. Kugel
AF Flight Dynamics Laboratory, Wright-Patterson AFB, Ohio

Abstract -- This paper describes the application of a modified Newton-Raphson parameter identification program to a post-simulation analysis of a large delta-wing aircraft similar to a Concorde supersonic transport. Pilot parameters are determined by minimizing the weighted mean square difference between the computed model responses and the measured responses of the total pilot-vehicle system. Pilot remnant is calculated using a power spectral density approach. These results are compared to presimulation analysis results obtained using an automated digital scheme and to those which were measured by an on-board analog computer. This study illustrates the utility of modern parameter identification techniques to post-simulation analysis.

I. Introduction

The mathematical modeling of pilot response in a particular task is of continuing interest in the fields of aircraft development and handling qualities evaluation. A model which can accurately represent a pilot's response is of great benefit and can be used to predict pilot rating and aircraft performance. A modeling effort such as this was applied to a recent study using the Total In-Flight Simulator (TIFS) to investigate the landing approach handling qualities of a large delta-wing aircraft. [1].

The TIFS is a variable stability research aircraft which permits the duplication of motion effects in the cockpit, as well as visual and instrument cues. Crosswinds and turbulence can be introduced electronically into the evaluation task. These signals are recorded to provide deterministic environmental disturbances which can be used later in the analysis program. All aircraft states and pilot response data are also recorded on a digital tape.

Prior to conducting this experiment a presimulation analysis of handling qualities was performed using Pitch Paper Pilot [2]. This analysis provided the predicted parameters of a pilot model for pitch control [3]. During the flight, an analog computer known as a Describing Function Analyzer (DFA) was used to measure the Bode response, (amplitude and phase) of the human operator at each frequency component of the input forcing function. The predicted and measured pilot parameters were then computed. The recorded data for the in-flight simulation provided a data base from which a technique could be developed to extract pilot model parameters from flight test data records by application of parameter identification techniques.

This paper will describe the models of the aircraft, the flight control system, and the pilot used in the post-simulation analysis. A description of the modified Newton-Raphson parameter identification routine used to extract the pilot parameters will also be discussed. The results will then be presented and compared with the results using the Describing Function Analyzer and Pitch Paper Pilot.

II. System and Disturbance Models

The total pilot-vehicle system for the approach and landing task of a large delta-wing aircraft can be represented by Figure 1.

Aircraft Dynamics

The longitudinal dynamics of the supersonic transport were programmed on the TIFS simulator using linearized, three-degree-of-freedom, small perturbation equations of motion. For parameter identification purposes the longitudinal responses were simplified by using a short period approximation to help limit the size of the overall model. This was a fairly good approximation since only a small section of data was being analyzed at any one time (40 seconds) and since approach speed was held relatively constant.

The aircraft short period linear equations used in this analysis were of the form

$$\dot{x} = Ax + Bd$$

For the short period approximation

$$x = [\theta, \alpha, q]^T$$

where

θ pitch angle (rad)

α angle of attack (rad)

q pitch rate (rad/sec)

The aircraft can be represented by the following first order linear differential equations:

$$\dot{\theta} = q$$

$$\dot{q} = (M_q + M_{\dot{\alpha}})q + (M_\alpha Z_\alpha + M_\alpha)\alpha + (M_\alpha Z_{\delta_e} + M_{\delta_e})\delta_e + (M_\alpha Z_\alpha + M_\alpha)\alpha_{gust}$$

$$\dot{\alpha} = q + Z_\alpha \alpha + Z_{\delta_e}\delta_e + Z_\alpha \alpha_{gust}$$

From the above equation it can be seen that

$$B = [0, (M_\alpha Z_\alpha + M_\alpha), Z_\alpha]^T$$

and

$$d = \text{disturbance vector} = \alpha_{gust}$$

The disturbance used was a sum of five sinusoids whose frequency content was equally spaced when plotted on a logarithmic axis [4]. The power

distribution resembled that of a Dryden spectra for an angle of attack gust, α_{gust}, of 1.272 degrees, whose associated vertical gust, Wg, was 6.0 ft/sec. The disturbance had the form

$$\alpha_{gust}(\omega) = \sum_{n=o}^{4} A_n \sin \omega_n t$$

Figure 2 shows the power distribution of this disturbance [5].

Control - Feel System

The control system for this analysis duplicated that programmed on the TIFS [1]. Coupled with the control system was a second order feel system which provided control feel to the wheel and rudder pedals. The total control-feel system can be modeled by a fourth order linear differential equation. This system is shown in Figure 3. Combining these two second order systems produces the following transfer function:

$$\frac{\delta_e}{\delta_p} = \frac{a_4}{s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4}$$

Using three dummy states, this equation can be transferred into four first order linear differential equations of the form

$$\dot{x} = Ax$$

These equations are

$$\dot{\delta}_e = \delta_{e_2}$$

$$\dot{\delta}_{e_2} = \delta_{e_3}$$

$$\dot{\delta}_{e_3} = \delta_{e_4}$$

$$\dot{\delta}_{e_4} = -a_4\delta_e - a_3\delta_{e_2} - a_2\delta_{e_3} - a_1\delta_{e_4} + a_4\delta_p$$

Human Operator

The model chosen for the human operator is a quasi-linear pilot describing function of the form

$$Y_p(s) = K_p e^{-\tau s}(T_L s + 1)\frac{K_n}{(T_L s + 1)\left[\frac{s^2}{\omega_n^2} + \frac{2\xi_n s}{\omega_n} + 1\right]}$$

[6], [7] and [8],

where the remnant term is defined to be that portion of the actual pilot's response not accounted for by the linear model.

The linear describing function, $Y_p(s)$, can be written in form

$$\dot{x} = Ax$$

by representing the pure time delay by a first order Padé approximation. The block diagram of Figure 4 represents the human operator model. Using the Padé approximation

$$e^{-\tau s} = \frac{-(s - 2/\tau)}{(s + 2/\tau)}$$

and

$$K = K_p K_n$$

then the human operator can be represented by four first order linear differential equations

$$\dot{y}_1 = \frac{1}{T_I}\theta + \frac{T_L}{T_I}q - \frac{1}{T_I}y_1$$

$$\dot{y}_2 = K\omega_n^2 y_1 - 2\xi_n\omega_n y_2 - \omega_n^2\delta'$$

$$\dot{y}_3 = (-2/\tau)y_3 + (4/\tau)\delta'$$

$$\dot{\delta}' = y_2$$

Total Pitch Tracking Model

The total pitch tracking model can be represented in the state vector form by

$$\dot{x} = Ax + Bd$$

where

$$x = (\theta, u, q, \delta_e, \delta_{e_2}, \delta_{e_3}, \delta_{e_4}, y_1, y_2, y_3, \delta')^T$$

and A is a partitioned matrix of the form given in Figure 5. The vectors B and d are the same as before.

III. Newton-Raphson Minimization

A linear system can be represented by

$$\dot{x} = Ax + Bd$$

and a set of output expressions

$$y = Fx + Gu + b$$

and

$$z = y + \eta$$

In the above equations

x	state vector
y	calculated response vector
z	measured response vector
d	disturbance vector
b	constant bias vector
η	noise vector
A	airframe/gust correlation vector
F	state transition matrix
G	gust transition vector

A cost function which is proportional to the mean square error can be represented by

$$J = \sum_{i=1}^{NM} (z_i - y_i)^T D_1 (z_i - y_i)$$

where NM is the number of samples and D_1 is a weighting matrix for the difference in response [9]. This weighing matrix should equal the inverse of the appropriate error covariance matrix [10].

Because the aircraft, feel system, and control system dynamics are known quantities (programmed on the TIFS simulator [1], only the parameters of the human operator are identified. These pilot parameters make up an unknown parameter vector, c, which relates to the system states and system responses as

$$\dot{x}(t) = f[x(t), c, u(t)]$$
$$y(t) = g[x(t), c] + \eta$$

For this experiment

$$c = [1/T_I, T_L/T_I, -\omega_n^2, K_p\omega_n^2, 2\xi_n\omega_n, 2/\tau]^T$$

The estimate for these parameters can be found by

$$\hat{c} = \text{ARG MIN } (J)$$

where ARG MIN means the vector c which minimizes the cost function J.

The calculated response vector y can be linearized with respect to the unknown parameter vector c such that

$$y_i = y_{i_o} + \nabla_c y_i (c - c_o)$$

where

y_{i_o}	nominal response calculated by using c_o
$\nabla_c y_i$	gradient of y with respect to c
c_o	nominal c vector

The optimal estimate for the unknown parameter vector is the vector c which minimizes J, and hence the mean square response error. This estimate can be found by applying the following equation iteratively to update the calculated nominal response and its gradient with respect to the vector of unknown parameters.

$$\hat{c} = c_o + \left[\sum_{i=1}^{NM} (\nabla_c y_i)^T D_1 (\nabla_c y_i)\right]^{-1} \left[\sum_{i=1}^{NM} (\nabla_c y_i)^T D_1 (z_i - y_{i_o})\right]$$

A priori estimates of the unknown parameters can be incorporated into the cost function using probability theory [9], [10]. This is done by maximizing the unconditional probability of z. The optimum parameter vector, $\hat{c}$, will result if a cost function containing the sum of the mean square response error and the mean square difference of c and its a priori value is minimized.

$$c = \underset{c}{\text{ARG MIN}} \sum_{i=1}^{NM} (z_i - y_i)^T M_1^{-1} (z_i - y_i) + (c - c_o)^T M_2^{-1} (c - c_o)$$

where

η	noise vector
c_o	nominal parameter vector
M_1	$E\{\eta\eta^T\}$
M_2	$E\{(c - c_o)(c - c_o)^T\}$

where E is the expectation operator. This occurs provided the weighting matrices used are equal to the appropriate inverse error covariance matrices. (The above information taken from [9]).

IV. Program Operation

The analysis using the Newton-Raphson method [9] was conducted on a Control Data Cyber-74 computer. The program takes 56,000 words of central memory to compile and execute. For a data record length containing 400 data points of 11 state variables, the program takes approximately 350 seconds of central processor time and 25 seconds of peripheral processor time. The program has the capability of printing out the values of the gradient and RMS error of each state and the value of the cost function at each iteration. After convergence, the program prints out the final A, B, F, and G matrices, and the bias vector b. Also printed out are the pilot parameters, an error covariance matrix of the estimated results and their approximate standard deviation. After convergence is reached, new time histories, using the estimated matrices, are calculated.

V. Results

Predictions (Paper Pilot)

The results of the presimulation analysis using Paper Pilot [11], with a pure time delay of $\tau = 0.2$ seconds and a first order neuromuscular lag of $T_I = 0.1$ second, predicted a pilot lead of $T_L = 3.89$ seconds and a pilot gain of $K_p = 0.691$. These parameters were obtained by optimizing the pilot parameters to minimize a cost functional based on the root mean square tracking error and the pilot lead [2] [3].

Frequency Techniques (DFA)

During the flight test of the TIFS simulator, the DFA calculated, on-line, the finite Fourier transform of the various system signals. The real and imaginary parts of the Fourier transform were then used in simple off-line calculations to yield system response and performance data. The resulting describing function measurements can be seen in Table 1 [3].

Newton-Raphson Method

Using the aircraft and pilot response data, which were recorded during the experiment, the Newton-Raphson minimization routine was used to extract pilot parameters based on the theory previously discussed.

During the digitization of this data, a 3-second section of data was improperly digitized midway through the run. As a result, the 100 seconds of data was divided into a two 40 second data records and each processed separately. During the identification analyses, the natural frequency and damping of the second order neuromuscular dynamics were set at ω_n = 16.5 radians/sec and ζ_n = 0.12 [6].

The identified pilot parameters in Table 2 were obtained using the Newton-Raphson technique.

It should be noted that this method can yield biased estimates if the mean of the distribution of θ is non-zero. This bias can be reduced by using the longest data records possible. The Newton-Raphson program also has the capability of estimating the bias terms of each of the states.

The above describing functions have been plotted in Bode form for comparison purposes. These plots appear in Figure 6.

Remnant

A determination of remnant was made by calculating the power spectral density [12] of the difference signal formed by subtracting the pilot output of the model from the actual recorded pilot command to the elevator. Figure 7 shows the power spectral density plots of the recorded elevator command of the pilot, the modeled elevator command, and the calculated remnant.

As can be seen from a close examination of the power spectral density functions of the pilot model, the model produces an output which has power at each of the five input frequencies. The spectral density functions of the actual recorded pilot output, however, has power at frequencies other than those of the input. These extraneous powers are defined as remnant powers. The small spikes in the power spectral density plots could result from nonlinear or nonstationary operation of the pilot or from the fact that the approach and landing tasks required the control of some side tasks and as a result was not truly a single loop compensatory task. The large spike in each of the remnant plots is attributed to "pilot pumping". Pumping is performed by the pilot to obtain dynamic information about the aircraft as it enters ground effects. The pumping frequency observed from in-flight records was always greater than 1.0 radians/sec. This oscillatory input to the elevator would show up as a pilot generated input and could therefore not be accounted for by the linear describing function. Making a sinusoidal approximation to the pilot pumping the remnant term will appear relatively flat and look more like a typical remnant spectra.

VI. Development Status

To date, the Newton-Raphson method as applied to pilot modeling has only been used with the longitudinal dynamics of a large transport aircraft simulation. By including the lateral directional dynamics and a roll pilot model, a two axis tracking situation could be modeled. Also, an analytical expression for pilot rating, such as is used in Paper Pilot [3], or in the Neal-Smith method [14], could be incorporated to give not only identified pilot parameters, but also to relate these parameters to a useful handling qualities criterion.

Finally, the Newton-Raphson method could be used in manual control situations for other than aircraft situations, such as automobile control in response to highway gusts [15], or the behavior of a helmsman steering a ship [16].

Further use and refinement of this technique could provide a valuable tool in the area of handling qualities and human operator modeling.

BIBLIOGRAPHY

[1]. Wasserman, Richard, et al, "In-Flight Simulation of Minimum Longitudinal Stability for Large Delta-Wing Transports in Landing Approach and Touchdown", Vol I and II, AFFDL-TR-72-143, Wright-Patterson AFB, Ohio, February 1973.

[2]. Anderson, R. O., Connors, Alonzo J., and Dillow, James D., "Paper Pilot Ponders Pitch", AFFDL/FGC-TM-70-1, Wright-Patterson AFB, Ohio, November 1970.

[3]. Stone, John R., "Prediction and Measurement of an In-Flight Pilot Model for Supersonic Transport", AFFDL/FGD-TM-74-22, Wright-Patterson AFB, Ohio, January 1974.

[4]. Peters, R. A. and Allen, R. W., "Operational Manual for Describing Function Analyzer", W-P-406-2, Systems Technology, Inc., Hawthorne, California, 1970.

[5]. Teper, Gary L., "An Assessment of the 'Paper Pilot' - An Analytical Approach to the Specification and Evaluation of Flying Qualities", AFFDL-TR-71-174, Wright-Patterson AFB, Ohio, June 1972.

[6]. McRuer, Duane and Krendel, Ezra, "Human Pilot Dynamics in Compensatory Systems", AFFDL-TR-65-15, Wright-Patterson AFB, Ohio, July 1965.

[7]. McRuer, Duane T., et al, "New Approaches to Human-Pilot/Vehicle Dynamics Analysis, AFFDL-TR-67-150. Wright-Patterson AFB, Ohio, February 1968.

[8]. Bekey, George A., "An Investigation of Sampled-Data Models of the Human Operator in a Control System", ASD-TR-62-36, Wright-Patterson AFB, Ohio, May 1962.

[9]. Taylor, Lawrence W., et al, "Systems Identification Using a Modified Newton-Raphson Method - A Fortran Program", NASA-TN-D-6734, Langley Research Center, May 1972.

[10]. Iliff, Kenneth W. and Taylor, Lawrence W., Jr., "Determination of Stability Derivatives from Flight Data Using a Newton-Raphson Minimization Technique", NASA-TN-D-6579, Langley Research Center, May 1972.

[11]. Stone, John R., "Paper Pilot Ponders Supersonic Transports", Ninth Annual Conference on Manual Control, Session II, Massachusetts Institute of Technology, May 1973.

[12]. Ralston, Anthony and Wilf, Herbert S., Mathematical Method for Digital Computers, Chapter 19, John Wiley and Sons, Inc., New York, 1960.

[13]. Anderson, Ronald O., "A New Approach to the Specification and Evaluation of Flying Qualities", AFFDL-TR-69-120, Wright-Patterson AFB, Ohio, June 1970.

[14]. Mayhew, David R., "A Digital Computer Program for the Calculation of Parameters Necessary to Satisfy the Closed Loop Criteria of T. P. Neal", AFFDL/FGC-TM-73-110, Wright-Patterson AFB, Ohio, August 1973.

[15]. Weir, D. H. and Hoh, R. H., "Driver/Vehicle Response and Performance in the Presence of Aerodynamic Disturbances", Fourteenth Joint Annual Control Conference, Ohio State University, June 1973.

[16]. Veldhuyzen, W., and Stassen, H. G., "Modelling the Behavior of the Helsman Steering a Ship", Ninth Annual Conference on Manual Control, Session VIII, Massachusetts Institute of Technology, May 1973.

[17]. Kugel, Daniel L., "Pilot Describing Function Estimates from In-Flight Data -- A Fortran Program", AFFDL/FGD TM-74-131, Wright-Patterson AFB, Ohio, July 1974.

PITCH TRACKING MODEL

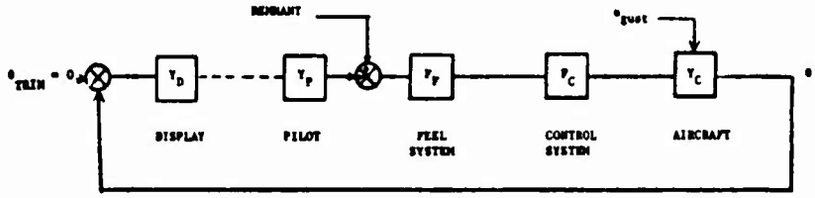


FIGURE 1

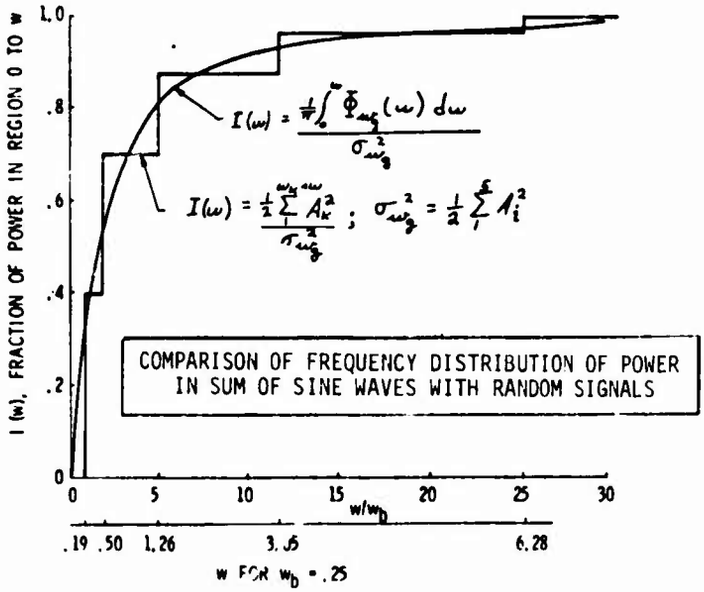


FIGURE 2

CONTROL AND FEEL SYSTEM

FIGURE 3

HUMAN OPERATOR MODEL

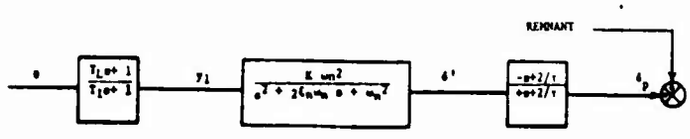


FIGURE 4

TOTAL PITCH TRACKING MODEL

$$A = \begin{array}{|c|c|c|} \hline \text{AIRCRAFT DYNAMICS} & 0 & 0 \\ \hline 0 & \text{CONTROL AND FEEL SYSTEM} & 0 \\ \hline \text{EQUALIZATION COEFFICIENTS } T_L, T_I & 0 & \text{HUMAN OPERATOR } K_p, T, \omega, f_n \\ \hline \end{array}$$

FIGURE 5

ω (rad/sec)	0.18	0.50	1.25	3.01	6.28
AMPLITUDE(dB)	-1.7	3.3	12.9	21.4	25.8
PHASE(deg)	-86.5	-150.4	-194.0	-168.6	-225.6

TABLE 1

	K_p	T_L	T_I	τ
RUN 1 T = 1 - 41 sec	0.702	1.605	0.102	0.333
RUN 2 T = 63 - 103 sec	0.475	2.528	0.104	0.336

TABLE 2

IN-FLIGHT PILOT MODEL PREDICTIONS AND MEASUREMENTS FOR AN SST IN POWER APPROACH

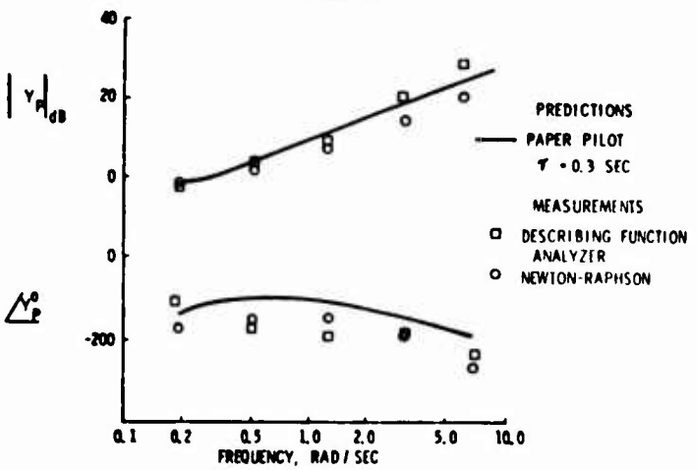


FIGURE 6

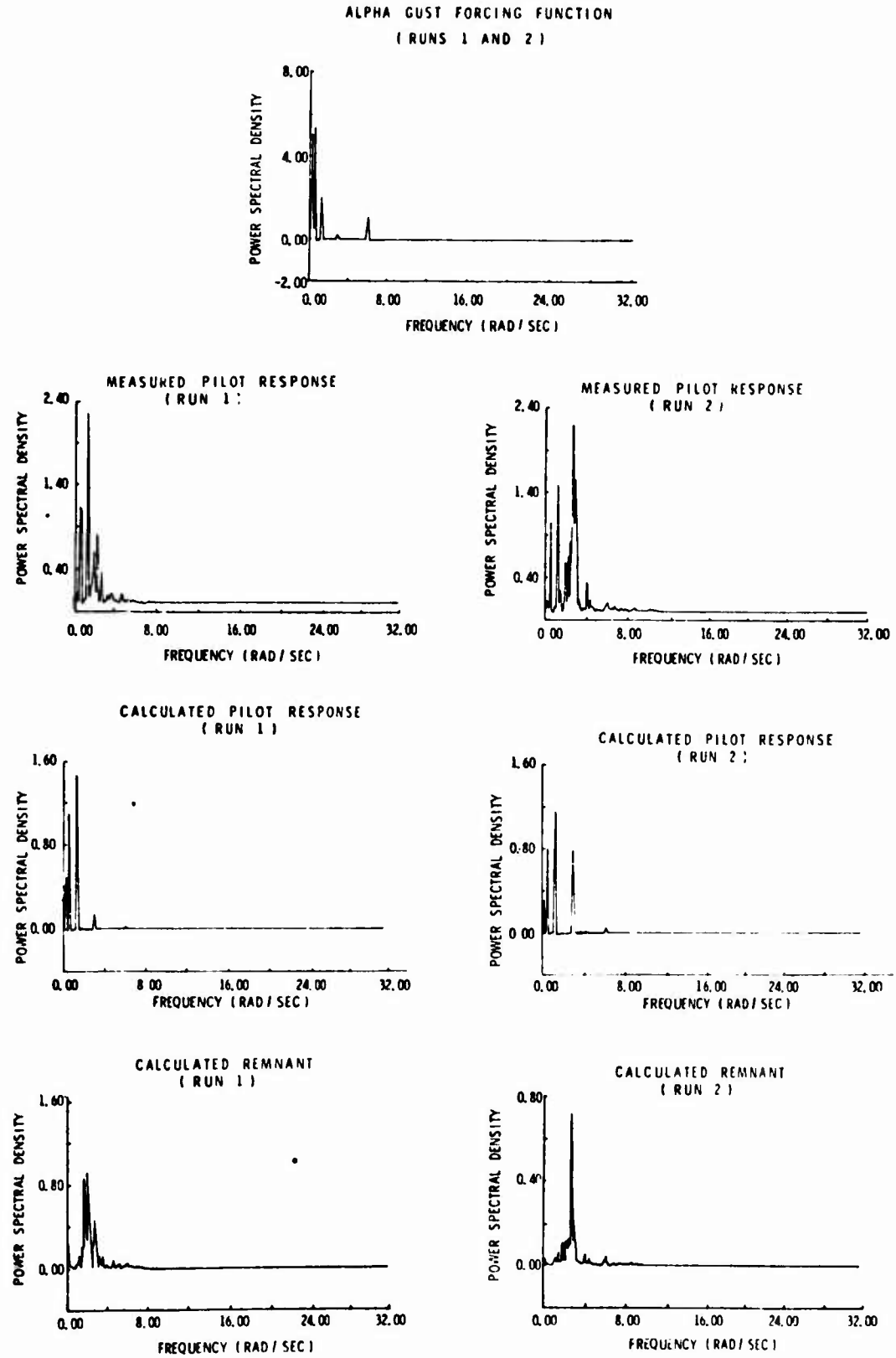

ALPHA GUST FORCING FUNCTION
(RUNS 1 AND 2)
POWER SPECTRAL DENSITY
8.00
4.00
0.00
-2.00
0.00
8.00
16.00
24.00
32.00
FREQUENCY (RAD / SEC)
MEASURED PILOT RESPONSE
(RUN 1)
2.40
1.40
0.40
MEASURED PILOT RESPONSE
(RUN 2)
CALCULATED PILOT RESPONSE
(RUN 1)
1.60
0.80
0.00
CALCULATED PILOT RESPONSE
(RUN 2)
CALCULATED REMNANT
(RUN 1)
CALCULATED REMNANT
(RUN 2)

FIGURE 7

The Stochastic Response of an Idealized Airplane to Atmospheric Turbulence*)

Jon Lee

Aerospace Research Laboratories
Wright-Patterson AFB, Ohio 45433

Abstract

The purpose of this paper is two-fold: The first is to present a new formulation of the stochastic input-response relation of linear dynamic systems to non-stationary random excitation. This formulation makes use of the Fourier-Stieltjes representation of a general random process first suggested by Priestley. The second purpose is to show how the new technique can be applied to the response of an idealized airplane encountering turbulent gusts. Since the main objective here is the illustration of the new technique, we shall not elaborate on the dynamic aspects of the aircraft response problem which are reported elsewhere.

---

*) This work was done in support of the Research Need (RN-AFFDL-04-72-2).

1. THE IDEALIZED AIRPLANE DYNAMICS

Let us idealize the plunging-mode motion of an idealized airplane with pitch neglected by the following equation [1]

$$v'(t) + a\, v(t) = -\, a\, w(t). \tag{1}$$

Here $v(t)$ denotes the vertical velocity of the airplane whose equivalent weight is supposedly massed around the center of gravity, $w(t)$ is the turbulent gust, and a is a collection of the airplane/aerodynamic parameters

$$a = \frac{\rho\, V\, S\, (dC_L/d\alpha)}{2m}, \tag{2}$$

where ρ is the air density, V is the airplane speed, S is the wing area, m is the airplane weight, and $(dC_L/d\alpha)$ is the lift curve slope (see Fig. 1).

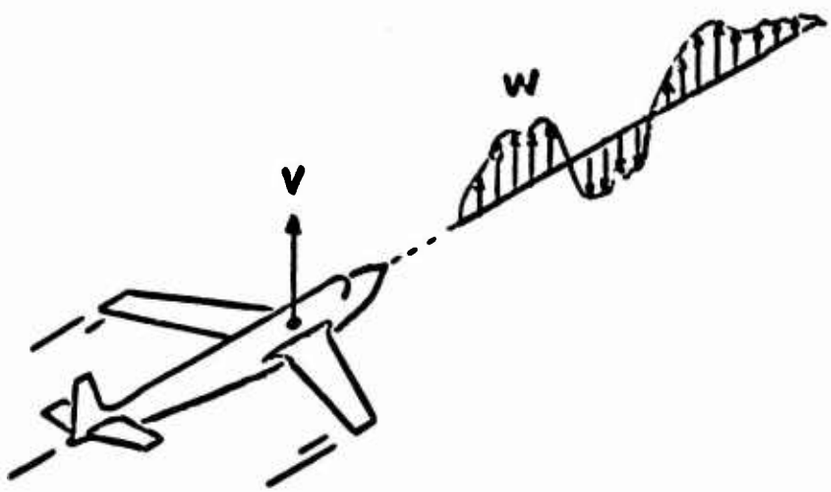


Fig. 1 Gust encounter.

The starting point is the Fourier-Stieltjes representation of a general random process first introduced by Priestley [2,3]

$$v(t) = \int_{-\infty}^{\infty} A(t,\omega)\, e^{i\omega t}\, dZ_v(\omega), \tag{3a}$$

$$w(t) = \int_{-\infty}^{\infty} B(t,\omega)\, e^{i\omega t}\, dZ_w(\omega). \tag{3b}$$

Here $A(t,\omega)$ and $B(t,\omega)$ are the yet unspecified deterministic functions, and $Z_v(\omega)$ and $Z_w(\omega)$ are random processes with orthogonal increments

$$\langle dZ_v(\omega_1) dZ_v^*(\omega_2) \rangle = 0, \quad \langle dZ_w(\omega_1) dZ_w^*(\omega_2) \rangle = 0, \text{ if } \omega_1 \neq \omega_2, \tag{4a}$$

$$\langle |dZ_v(\omega)|^2 \rangle = dF_v(\omega), \quad \langle |dZ_w(\omega)|^2 \rangle = dF_w(\omega), \tag{4b}$$

where $\langle\ \rangle$ denotes expectation. Now, introduce (3) into (1) to obtain

$$\int_{-\infty}^{\infty} [A'(t,\omega) + A(t,\omega)(i\omega + a)]\, e^{i\omega t}\, dZ_v(\omega) = -a \int_{-\infty}^{\infty} B(t,\omega)\, e^{i\omega t}\, dZ_w(\omega). \tag{5}$$

This is a stochastic relation, hence the equality must be understood in a statistical sense. Consider

$$[A' + A\,(i\omega + a)]\, dZ_v(\omega) = -a\, B(t,\omega)\, dZ_w(\omega). \tag{6}$$

By a direct substitution [4], one can argue that (6) is a unique choice satisfying (5). Then considering (6) as the equation that $A(t,\omega)$ must satisfy, the particular solution can be obtained in the form

$$A(t,\omega) = A_0(t,\omega)\ (dZ_v(\omega)/\ dZ_w(\omega)), \tag{7}$$

where

$$A_0(t,\omega) = -\,a \int_0^t e^{-(i\omega\ +\ a)(t\ -\ \tau)}\ B(\tau,\omega)\ d\tau\ . \tag{8}$$

Upon inserting (7) into (3a), we obtain

$$v(t) = \int_{-\infty}^{\infty} A_0(t,\omega)\ e^{i\omega t}\ dZ_w(\omega). \tag{9}$$

This is the general representation for v(t) subject to arbitrary turbulent gusts.

2. STATIONARY TURBULENT GUST

The situation is the simplest when the turbulent gust may be assumed stationary. We can then evaluate the $A_0(t,\omega)$ explicitly since $B(t,\omega)$ is constant (= 1). Whence,

$$A_0(t,\omega) = \frac{-\,a}{(i\omega + a)}\ [\ 1 - e^{-(i\omega\ +\ a)t}\ ]. \tag{10}$$

For the exceedance statistics, it is necessary to evaluate the 2x2 covariance matrix formed by v(t) and v'(t). After inserting (10) into (9), the differentiation with respect to t yields

$$v'(t) = \int_{-\infty}^{\infty}\ A_1(t,\omega)\ e^{i\omega t}\ dZ_w(\omega), \tag{11}$$

where

$$A_1(t,\omega) = \frac{-a}{(i\omega + a)} \left[i\omega + a\, e^{-(i\omega + a)t} \right]. \tag{12}$$

In view of (4a), we can at once write down the covariance matrix (i,j =0,1)

$$\langle v^i(t) v^{j*}(t) \rangle = \int_{-\infty}^{\infty} A_i(t,\omega)\, A_j^*(t,\omega)\, dF_w(\omega). \tag{13}$$

Here, the superscript for v(t) actually refers to the number of primes; namely, i = j = 0 corresponds to the variance of v(t), and so on. If $F_w(\omega)$ is differentiable, i.e., $dF_w(\omega) = \Phi_w(\omega)\, d\omega$, equation (13) takes a more familiar form

$$\langle v^i(t) v^{j*}(t) \rangle = \int_{-\infty}^{\infty} A_i(t,\omega) A_j^*(t,\omega)\ \Phi_w(\omega)\, d\omega \ , \tag{14}$$

where $\Phi_w(\omega)$ is the power spectral density (psd) of turbulent gusts.

Let us now explicitly compute the covariances. First, for i = j, $A_i A_i^*$ is the squared magnitude of A_i; hence, the variances $\sigma_v^2(t) = \langle v(t) v^*(t) \rangle$ and $\sigma_{v'}^2(t) = \langle v'(t) v'^*(t) \rangle$ are real

$$\sigma_v^2(t) = \int_{-\infty}^{\infty} \frac{a^2}{(a^2 + \omega^2)} \left[1 - 2\cos\omega t\, e^{-at} + e^{-2at} \right] \Phi_w(\omega)\, d\omega \ , \tag{15}$$

$$\sigma_{v'}^2(t) = \int_{-\infty}^{\infty} \frac{a^2}{(a^2 + \omega^2)} \left[\omega^2 - 2a\omega \sin\omega t\, e^{-at} + a^2\, e^{-2at} \right] \Phi_w(\omega)\, d\omega \ . \tag{16}$$

Second, $A_0 A_1^*$ has the real and imaginary parts

$$A_0(t,\omega) A_1^*(t,\omega) = \frac{a^2}{(a^2 + \omega^2)} (a\cos\omega t\, e^{-at} + \omega \sin\omega t\, e^{-at} - a\, e^{-2at}) +$$

$$+ \frac{i\,a^2}{(a^2+\omega^2)}\,(-\omega + a\sin\omega t\, e^{-at} + \omega\cos\omega t\, e^{-at}). \qquad (17)$$

The theory of stationary random processes states that the psd is even function for the real process. Since the imaginary part of (17) is odd, it integrates out to zero when introduced into (14). Hence, only the real part of (17) contributes to the covar ance

$$\langle v(t)v'(t)\rangle = \int_{-\infty}^{\infty} \frac{a^2}{(a^2+\omega^2)}\,[a\cos\omega t\, e^{-at} + \omega\sin\omega t\, e^{-at} - a\, e^{-2at}]\,\Phi_w(\omega)d\omega\,. \qquad (18)$$

This shows that the response covariance matrix is real and symmetric.

To complete the computation of $\langle v^i(t)v^j(t)\rangle$, we shall adopt the so-called Dryden's turbulent gust spectrum

$$\Phi_w(\omega) = \frac{\sigma_w^2 K}{2\pi}\,\frac{1+3K^2\omega^2}{(1+K^2\omega^2)^2}\,, \qquad (19)$$

where σ_w^2 is the gust variance. The spectrum parameter K denotes L/U, where L being the integral length scale and U the mean flow velocity. With the use of (19), the integrals of (15), (16), and (18) can be evaluated by the standard contour integration technique

$$\sigma_v^2(t) = \frac{\sigma_w^2 a^2 K}{2\pi}\,[(1+e^{-2at})\,I_1 - 2\,e^{-at}\,I_3(t)], \qquad (20)$$

$$\sigma_{v'}^2(t) = \frac{\sigma_w^2 a^2 K}{2\pi}\,[\,a^2\,e^{-2at}\,I_1 + I_2 - 2a\,e^{-at}\,I_4(t)], \qquad (21)$$

$$\langle v(t)v'(t)\rangle = \frac{\sigma_w^2 a^2 K}{2\pi}\,[-a\,e^{-2at}I_1 + a\,e^{-at}I_3(t) + e^{-at}I_4(t)]. \qquad (22)$$

Here I's are the definite integrals ($r(\omega) = (1 + 3K^2\omega^2)/(a^2+\omega^2)(1 + K^2\omega^2)^2$)

$$I_1 = \int_{-\infty}^{\infty} g(\omega)\, d\omega = \pi(1 - 3a^2K^2 + 2a^3K^3)/\, a(1 - a^2K^2)^2,$$

$$I_2 = \int_{-\infty}^{\infty} g(\omega)\omega^2 d\omega = \pi(2 - aK - 4a^2K^2 + 3a^3K^3)/K(1 - a^2K^2)^2,$$

$$I_3(t) = \int_{-\infty}^{\infty} g(\omega)\, \cos\omega t\, d\omega =$$

$$= \pi[e^{-at}(1 - 3a^2K^2) + e^{-t/K}(2a^3K^3 + at(1 - a^2K^2))]/a(1 - a^2K^2)^2$$

$$I_4(t) = \int_{-\infty}^{\infty} g(\omega)\, \omega\sin\omega t\, d\omega =$$

$$= \pi[e^{-at}(1 - 3a^2K^2) + e^{-t/K}(3a^2K^2 - 1 + t(1-a^2K^2)/K)]/(1 - a^2K^2)^2.$$

In Fig. 2, we have displayed the time developments of $\sigma_v(t)/\sigma_w$, $\sigma_{v'}(t)/\sigma_w$, and $\mu(t) = \langle v(t)v'(t)\rangle / \sigma_v(t)\, \sigma_{v'}(t)$ under a = 1.07 and K = 0.9.

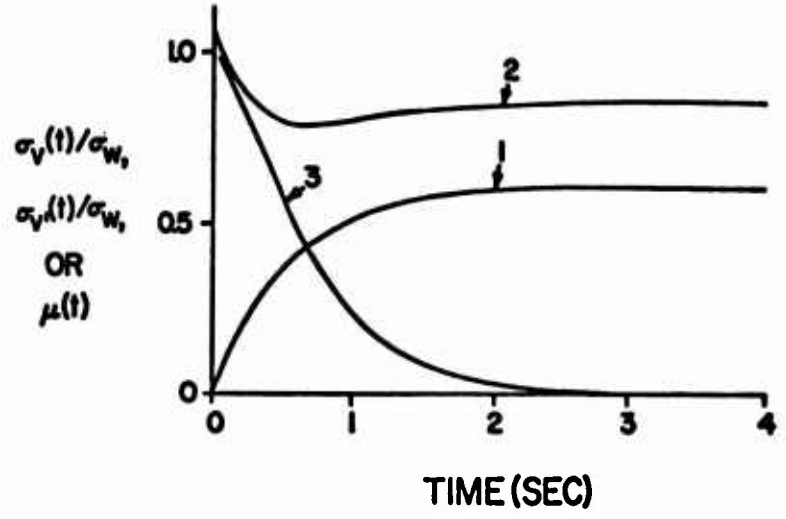


Fig. 2 Transient response covariances (a = 1.07 & K = 0.9).
(1): $\sigma_v(t)/\sigma_w$; (2): $\sigma_{v'}(t)/\sigma_w$; (3): $\mu(t)$.

3. STATIONARY INPUT-RESPONSE RELATION

After the transition period, the response variances attain the stationary values

$$\sigma_v = \left\{\frac{\sqrt{Ka(1 - 3a^2K^2 + 2a^3K^3)/2}}{1 - a^2K^2}\right\} \sigma_w, \tag{23}$$

$$\sigma_{v'} = \left\{\frac{a\sqrt{(2 - aK - 4a^2K^2 + 3a^3K^3)/2}}{1 - a^2K^2}\right\} \sigma_w. \tag{24}$$

The important dynamic conclusions are these;

(i) As a natural generalization of the gust-variance modulation [5,6], the non-stationary turbulent gust should be modeled by incorporating both the time-varying σ_w and K.

(ii) Under the Gaussianity assumption, the response exceedance of the plunging-mode motion is given by

$$N_\alpha = N_o \exp(-\alpha^2/2\,\sigma_v^2),$$

where $N_o = (\sigma_{v'}/\sigma_v)/2\pi$. Since Fig. 3 shows $(\sigma_{v'}/\sigma_v) \sim K^{-1/2}$, the non-stationary turbulent gust model of (i) can bring about response exceedance curves displaying a wide range of convexity (see Ref. [7] for the detailed discussion).

(iii) For a fixed a, the σ_v, and σ_w are related in (24) through the curly bracket which is a sole function of K. Note that we can estimate the σ_w and $\sigma_{v'}$ respectively from the gust record and the accelerometer data at the center of gravity [8]. It is therefore possible [7] to compute K indirectly

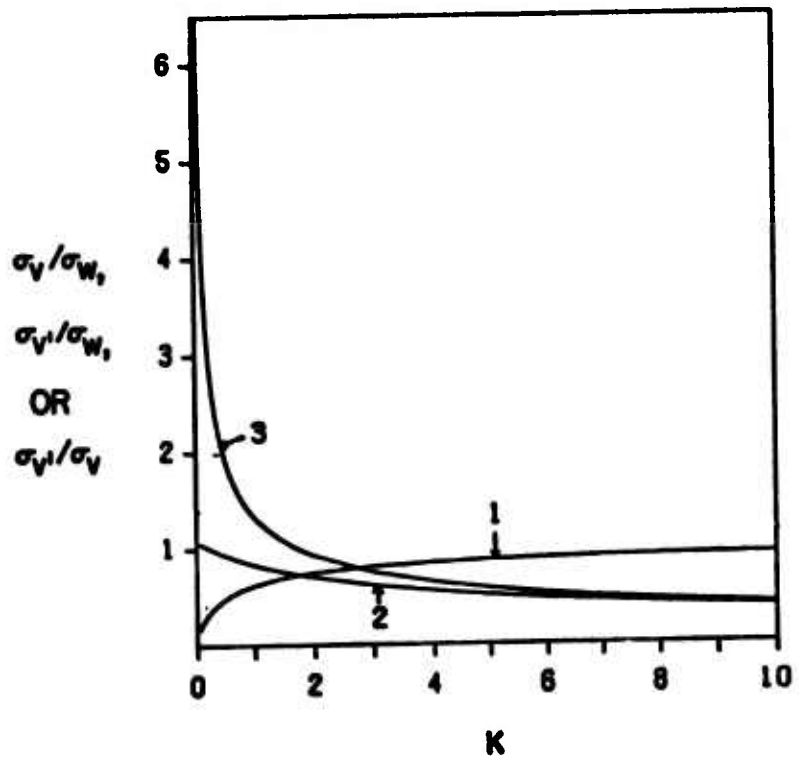


Fig. 3 Stationary input-response relation (a = 1.07).

(1): σ_v/σ_w; (2): $\sigma_{v'}/\sigma_w$; (3): $\sigma_{v'}/\sigma_v$.

from the variance data under the assumption that the airplane dynamics may be idealized by (1) and the variance measurements are truly stationary.

References

1. Bisplinghoff, R. L., Isakson, C., and O'Brien, T. F., Jour. Aero. Sci. 8, 33 (1951).
2. Priestley, M. B., Jour. Roy. Stat. Soc., B27, 204 (1965).
3. Priestley, M. B., Jour. Sound & Vibr., 6, 86 (1967).
4. Lee, J., Jour. Sound & Vibr., 34,.xxx (1974).
5. Howell, L. J. and Lin, Y. K., AIAA Jour., 9, 2201 (1971).
6. Verdon, J. M. and Steiner, R., AIAA Jour., 11, 1086 (1973).
7. Lee, J., The Plunging-Mode Response of an Idealized Airplane to Atmospheric Turbulent Gusts.
8. Crooks, W.M. and et al, Project HICAT, High Altitude Clear Air Turbulence Measurements and Meteorological Correlations, AFFDL-TR-68-127, Vol. 1, Wright-Patterson AFB, Ohio (1968).

A METHOD OF CONSTRUCTING MINIMAL LINEAR STATE-VARIABLE MODELS BASED ON A GENERAL CLASS OF SYSTEM REPRESENTATIONS*

David R. Audley, 1Lt, USAF
Applied Mathematics Research Laboratory
Aerospace Research Laboratories
Wright-Patterson AFB, Ohio 45433
(513)255-2481

Abstract The problems of partial and complete linear system realization are solved using a general class of system representations. Direct identification schemes based on different types of input/output experiments are known for many representations in the class.

I Introduction

We consider constant parameter linear systems which can be described by the state vector equation.

$$\begin{aligned} \dot{x}(t) &= Ax(t)+Bu(t), \quad t \geq 0 \\ y(t) &= Cx(t). \end{aligned} \tag{1}$$

The vectors x,u, and y are n,p, and r dimensional, respectively. We refer to (1) by the triple (A,B,C) and we say that (A,B,C) has dimension n. For our purposes the input/output description of (1) will be in terms of the $r \times p$ matrix transfer function

$$H(s) = C(sI-A)^{-1}B. \tag{2}$$

Most of the recent results concerning realization theory are based on the Markov parameter representation of H(s). [1], [2], [3]. Writing H(s) as a power series

$$H(s) = \sum_{i=1}^{\infty} Y_i s^{-i} \tag{3}$$

*Reissue of: David R. Audley & Wilson J. Rugh, "Linear System Realization Based on Data Set Representations", JHU-EE Report No. 74-6, The Johns Hopkins University, Baltimore, MD, 1974, for presentation at the Symposium on Air Force Applications of Modern Control Theory, 9-11 July 1974, Wright-Patterson AFB, Ohio.

yields the infinite sequence $\{Y_1, Y_2, \ldots\}$ of Markov parameters. From this data a minimal complete realization of $H(s)$ can be obtained. This is a least dimension triple (A,B,C) such that (2) holds. [2], [4] If we take only the first M terms, $\{Y_1, \ldots, Y_M\}$, then a minimal partial realization of order M of $H(s)$ can be obtained. This is a least dimension triple (A,B,C) such that $C(sI-A)^{-1}B$ has the specified first M Markov parameters. [2], [3]

Two disadvantages of this approach are the following. First, direct identification of the Markov parameters involves measuring the impulse response and its derivatives at $t = 0$. Aside from the well known difficulties in obtaining these values accurately, many systems cannot tolerate such severe test signals (particularly biological systems). Second, a partial realization obtained from the Markov parameters may not be useful. Such a realization tends to model the high frequency behavior of $H(s)$ and this behavior may not be of interest. Also, a partial realization of $H(s)$ can be unstable even though $H(s)$ is stable.

These difficulties can often be circumvented by using the data set representation discussed below. A wide choice of input/output experiments is available for data set identification and a partial realization reflects the nature of the information in the data set. If instabilities arise in a partial realization, a different data set representation can be chosen.

II System Realization from Data Sets

Corresponding to $H(s)$ we define a <u>data set of order M</u> as follows. According to the finite complex numbers $\lambda_1,\ldots,\lambda_m$, determine the set of $r\times p$ matrice.

$$Y_{ij} = \frac{d^j}{ds^j} H(s)\Big|_{s=\lambda_i} = H^{(j)}(\lambda_i), \quad \begin{matrix} i=1,\ldots,m \\ j=0,\ldots,k_i-1 \end{matrix} \tag{4}$$

where $\sum_{i=1}^{m} k_i = M$. We assume that each Y_{ij} is finite and that if $Y_{ij} = H^{(j)}(\lambda_i)$ is in the set, then $\bar{Y}_{ij} = H^{(j)}(\bar{\lambda}_i)$ is in the set, where overbar denotes complex conjugate. This latter assmuption is not necessary but it removes the necessity for complex arithmetic in the following development.

There are several examples of data set representations which are well known and which can be identified directly from input/output experiments. The following examples are stated for bounded input, bounded output stable systems with a single input and a single output; the extension to the multivariable case being straightforward. The time moments of the impulse response yield the values of $H(s)$ and its derivatives at $s=0$. [5], [6] It is well known that the steady-state response to $\sin(\omega t)$ yields $H(j\omega)$, $j=\sqrt{-1}$. If λ_1 and λ_2 are positive numbers, then with input $e^{-\lambda_1 t}$ and corresponding response $y(t)$,

$$H(\lambda_2) = (\lambda_1+\lambda_2)\int_0^\infty y(t)e^{-\lambda_2 t}dt. \tag{5}$$

Similar computations involving the response to $t^j e^{-\lambda_1 t}$ yield the values $H^{(j)}(\lambda_2)$. [7]

In keeping with previous terminology a triple (A,B,C) of least dimension such that $C(sI-A)^{-1}B$ satisfies a given data set of order M will be called a <u>minimal partial realization of order M</u> of $H(s)$. The following is a method for obtaining such realizations.

Suppose a data set of order M is given as in (4). We shall show how to obtain the transfer function $G(s)$ of a minimal partial realization of order M. The corresponding triple can be obtained from $G(s)$ by the usual methods.

We can write

$$G(s) = P(s)/Q(s) \tag{6}$$

where $P(s)$ is an $r \times p$ matrix polynomial, $Q(s)$ is a scalar polynomial, and deg $P(s)$ < deg $Q(s)$. Multiplying both sides of (6) by $Q(s)$ and using the known data set we obtain the set of equations

$$p^{(j)}(\lambda_i) = [G(\lambda_i)Q(\lambda_i)]^{(j)} = \sum_{k=0}^{j} \binom{j}{k} Y_{ik} Q^{(j-k)}(\lambda_i), \quad j = 0,1,\ldots,k_i-1 \quad i = 1,\ldots,m . \tag{7}$$

In [7] it is shown that solving the set of equations (7) for $P(s)$ and $Q(s)$ is equivalent to solving the matrix polynomial equation

$$P(s) = B(s)Q(s) - A(s)R(s) \tag{8}$$

where $B(s)$ is an $r \times p$ matrix polynomial with real coefficients such that degree $B(s) < M$ and

$$B^{(j)}(\lambda_i) = Y_{ij}, \quad j = 0,\ldots,k_i-1 \quad i = 1,\ldots,m. \tag{9}$$

(e.g. B(s) is an interpolating polynomial of the Full Hermite type [8]), A(s) is the scalar polynomial with real coefficients

$$A(s) = \prod_{i=1}^{m} (s-\lambda_i)^{k_i}, \tag{10}$$

and R(s) is an $r \times p$ matrix polynomial remainder term.

We now proceed to determine P(s) and Q(s) from (8). Multiplying both sides by $1/(A(s)Q(s))$ gives

$$\frac{P(s)}{A(s)Q(s)} = \frac{B(s)}{A(s)} - \frac{R(s)}{Q(s)} = o(s^{-M-1}) \tag{11}$$

where the symbol $o(s^{-M-1})$ indicates that the right side of (11) is a power series in s^{-1} with lowest power M+1. Thus B(s)/A(s) and R(s)/Q(s) written as power series in s^{-1} have identical terms through s^{-M} That is, if

$$B(s)/A(s) = V_1 s^{-1} + V_2 s^{-2} + \ldots + V_M s^{-M} + o_1(s^{-M-1}) \tag{12}$$

then

$$R(s)/Q(s) = V_1 s^{-1} + V_2 s^{-2} + \ldots + V_M s^{-M} + o_2(s^{-M-1}). \tag{13}$$

The partial sequence of Markov parameters $\{V_1, \ldots, V_M\}$ from (12) and (for instance) Ho's algorithm provides a minimal (A_v, B_v, C_v), such that $R(s)/Q(s) = C_v(sI-A_v)^{-1}B_v$ satisfies (13) if and only if there exist positive integers N and N' = M-N such that

$$\text{rank } H_{N'N} = \text{rank } H_{N'+1,N} = \text{rank } H_{N',N+1} \tag{14}$$

where H_{ij} is the $ri \times pj$ upper left corner submatrix of the Hankel matrix. [2] We know that for the partial sequence $\{V_1, \ldots, V_M\}$ there is a minimal extension for which (14) holds. Since this extension is not unique, R(s)/Q(s) is in general not unique. However, Q(s) will have minimal degree since A_v is of minimal dimension.

Having determined $R(s)/Q(s)$, we obtain a (possibly non-unique) transfer function $G(s)$ by writing (11) as

$$G(s) = P(s)/Q(s) = A(s)\left[\frac{B(s)}{A(s)} - \frac{R(s)}{Q(s)}\right]. \qquad (15)$$

This transfer function will satisfy the given data set unless so-called unattainable points arise. That is, unless for some λ_i, $P(\lambda_i) = Q(\lambda_i) = 0$. For a discussion of this phenomenon as well as a survey of the Cauchy Interpolation Problem from which our results are derived, see [9]. Suffice it to note that if the data set consists only of evaluations at points with non-negative real parts (as do all the data set examples mentioned above), and if $Q(s)$ as computed from (13) has all roots with negative real parts (the partial realization is stable), then unattainable points cannot occur.

Remark If $H(s)$ admits a minimal complete realization of dimension at most n, then this realization can be uniquely determined (up to a change of variables) from any data set of order $2n$ obtained from $H(s)$. This result follows from the fact that the rank condition (14) will be satisfied with $N' = N = n$, [2], and the fact that unattainable points cannot arise.

It should be noted that from a theoretical point of view the particular partial realization algorithm used to compute $R(s)/Q(s)$ is unimportant. There are a number of such algorithms available and computational considerations should govern the choice. [4],[10],[11]

Example Based on the data set $\{H(1/2) = 14/3, H(2) = 4/3\}$ obtained from $H(s) = (s+10)/(s+1)^2$, we will compute the transfer function of a minimal partial realization of order 2. We take

$$B(s) = \frac{-20}{9}\left(s-\frac{52}{20}\right),\quad A(s) = s^2-\frac{5}{2}s+1$$

so that

$$B(s)/A(s) = \frac{-20}{9}s^{-1}+\frac{2}{9}s^{-2}+o(s^{-3}).$$

Thus we find

$$R(s)/Q(s) = \frac{-20/9}{s+1/10}$$

and from (15)

$$G(s) = \frac{56/20}{s+1/10}\ .$$

Note that based on the Markov parameters $\{1,8\}$ of $H(s)$ we obtain

$$G(s) = \frac{1}{s-8}\ ,$$

which is unstable.

III Conclusions

Methods of obtaining realizations of a linear input/output map typically involve representing the input/output map in a particular fashion, discarding a portion of this representation, and using the remaining portion to determine a realization. In this paper we have used a "data set" representation of linear input/output maps.

For a complete realization the particular data set used is unimportant since all data sets of sufficient order will provide the unique (up to variable change) realization. However in the partial realization case the particular data set used determines the nature of the realization obtained. Thus a central question is: In what sense is the input/output map of a given partial realization an approximation to the prescribed input/output map? In some cases we have intuitive answers. For example, Markov parameters yield a "high frequency" approximation and the moment coefficient data set yields a "low frequency" approximation. For other data sets the results in [7] provide an interpretation in terms of a Hilbert space mapping.

However, these answers are not completely satisfactory. The results described here show that the structure of the problem is that of rational function interpolation. Within this structure, approximation theoretic results are needed.

Acknowledgement The authors acknowledge the contribution of Stephen L. Baumgartner concerning the occurrence of unattainable points.

IV References

1. Silverman, L.M., "Realization of linear dynamical systems," IEEE Trans. Automat. Contr., AC-16: 554-567, Dec. 1971.

2. Kalman, R.E., "On minimal partial realizations of a linear input/output map," Aspects of Network and System Theory (Guillemin Prize vol.); ed. by R.E. Kalman and N. DeClaris. Holt, Rinehart and Winston, 1971.

3. Tether, A.J., "Construction of minimal linear state-variable models from finite input/output data," IEEE Trans. Automat. Control, AC-15: 427-436, Aug. 1970.

4. Ho, B.L. and Kalman, R.E., "Effective construction of linear state-variable models from input/output functions," Proc. 3rd Ann. Allerton Conf. Circuit & System Theory, 449-459, 1965; also Regelungstech., 14: 545-548, 1966.

5. Huggins, W.H., Network Approximation in the Time Domain, AFCRL report No. E 4048, Oct. 1949.

6. Bruni, C., Isidori, A. and Ruberti, A., "A method of realization based on the moments of the impulse-response matrix" IEEE Trans. Automat. Contr., AC-14: 203-204, April 1969.

7. Audley, D.R. and Rugh, W.J., "On the H-matrix system representation," IEEE Trans. Automat. Contr., AC-18: 235-243 (Also Proc. 1972 IEEE Conf. on Dec. and Control, 66-70, 1972).

8. Davis, P.J., Interpolation and Approximation, Blaisdell, 1963.

9. Meinguet, J., "On the solubility of the Cauchy Interpolation Problem," Proc. Univ. Lancaster Symposium on Approximation Theory and Application, ed. A. Talbot, Academic Press, 1970.

10. Rissanen, J., "Recursive identification of linear systems", SIAM J. Control, Vol. 9, No. 3: 420-430, August 1971.

11. Dickinson, B.W., Morf, M., and Kailath, T., "A minimal realization algorithm for matrix sequences", IEEE Trans. Automat. Contr., AC-19: 31-38, February 1974.

APPLICATION OF AN OPTIMAL CONTROL PILOT MODEL TO AIR-TO-AIR COMBAT

Thomas R. Harvey* and James D. Dillow

Abstract

Linearized equations approximating the longitudinal dynamics of an air-to-air combat tracking task with a lead computing optical sight were simulated for three sets of aircraft dynamics, two levels of target activity, and two slight configurations. A closed loop model for the air-to-air combat task is developed using an optimal model for the pilot. The pilot model accounts for inherent human limitations such as neuromuscular dynamics, reaction time delay, and visual errors. Analytic values of performance based on the model closely match the values measured in the simulation.

Introduction

The task of accurately tracking a maneuvering target for the purpose of obtaining a kill with airborne cannon fire is one of the most difficult required of a fighter pilot. As is the case when firing any projectile at a moving target, lead for target motion must be computed and the aiming direction adjusted to compensate accordingly. Further complicating the air-to-air problem, trajectory adjustments must be made to account for projectile drag, velocity jump, and gravity drop, as well as other more minor effects. In most modern fighter aircraft, the required lead compensation is continuously computed using air data information and displayed on a heads up display. The lead information is normally presented to the pilot of the attacking aircraft in the form of a pipper that is two mils in diameter and is surrounded by a larger circle called the reticle. The pipper is depressed from the weapon line by the amount of the computed lead angle as shown in Fig. 1. When the pilot maneuvers his aircraft in such a manner as to place the pipper on the target, he has achieved the proper aiming direction to insure a kill provided the target maneuver remains constant. A sight of this type is called a lead computing optical sight (LCOS).

Two practical problems are encountered with the sight design. First, the lead computation is an approximation and there is usually a trade off between the accuracy of the computed lead and the time lag introduced by the computation time. Secondly, the continuous lead computation creates sight dynamics or motion of the pipper that may be confusing or unsuitable to the pilot. For example if the sight dynamics are lightly damped, the pilot may have to exercise special care to avoid pilot induced oscillations (PIO). The overall effect of adverse sight dynamics

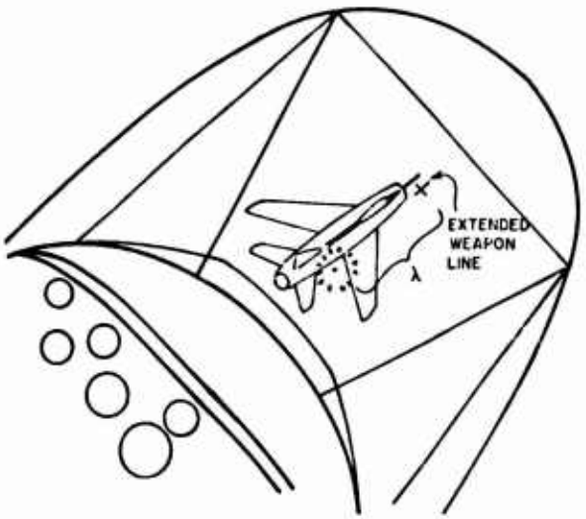


Fig. 1. Air-to-Air Tracking as Seen by Pursuing Pilot at Close Range (Approximately 750 Feet).

can be a reduction in the aiming accuracy despite the fact that the lead computation is reasonably accurate.

It would be extremely desirable to have a reliable method for analytically predicting piloted performance in an air-to-air combat situation. In this way, the trade offs involved in computing lead could be evaluated using the closed loop dynamics of the pilot-vehicle system. The sight parameters could be optimized and the suitability of the sight dynamics in the closed loop tracking task could be evaluated. Furthermore, this could be done without resorting to extensive simulation or flight tests.

A simplified model for the closed loop air-to-air combat task has been developed and validated by experimental data. This work is described in this paper. First, the piloted simulation study that was used to collect data for an air-to-air combat task is described. The equations are given and the assumptions delineated. A control theoretic optimal pilot model is used to develop an analytic model for the task. This model is briefly described and the representation of human limitations used in the model are given. A comparison of the experimental and analytic results are given as a demonstration of the model's validity.

*Currently assigned to the Air Force Systems Command, Satellite Control Facility, Sunnyvale, California.

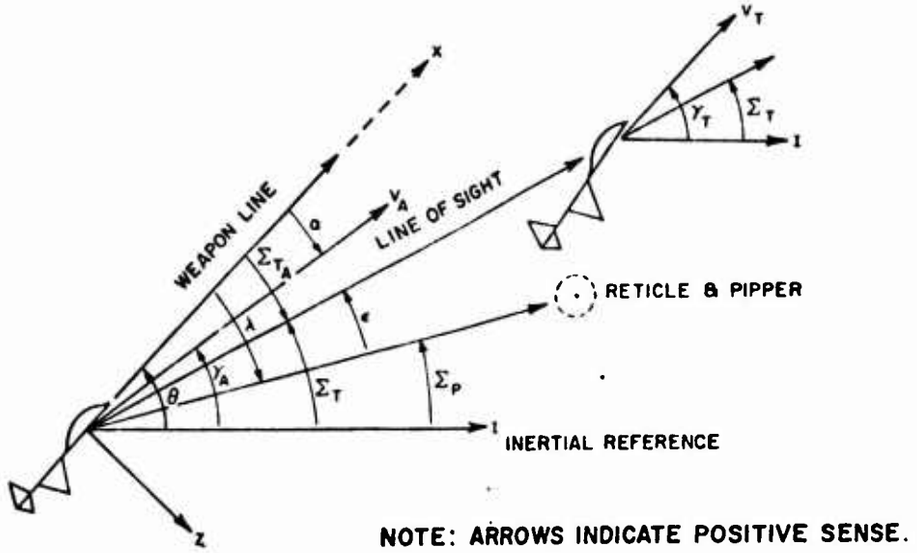


Fig. 2. Air-to-Air Tracking Geometry.

Simulation

The two dimensional, longitudinal dynamics of an air-to-air combat tracking task with a lead computing optical sight were simulated. The in plane geometry for the air-to-air tracking situation is shown in Fig. 2. The following notation is used in Fig. 2:

V_A - velocity of attacker (ft/sec)

V_T - velocity of target (ft/sec)

γ_A - attacker flight path angle (radians)

γ_T - target flight path angle (radians)

α - attacker angle of attack (radians)

θ - attacker pitch angle (radians)

Σ_{T_A} - relative line of sight angle (radians)

λ - lead angle (radians)

ϵ - error (radians)

Σ_T - inertial line of sight angle (radians)

Attacking Aircraft Equations

The equations of motion of the attacking aircraft are approximated by the linearized longitudinal short period equations of motion. These are

$$\dot{\alpha} = q + Z_\alpha \alpha + Z_\delta \delta$$

$$\dot{q} = M_q q + M_\alpha \alpha + M_{\dot{\alpha}} \dot{\alpha} + M_\delta \delta$$

$$\dot{\theta} = q$$

where

q - pitch rate (radians/sec)
δ - elevator deflection (radians)

The constants Z_α, Z_δ, M_q, etc. are stability derivatives associated with a given aircraft and flight condition.

Elevator actuator dynamics are modeled by a first order lag,

$$\dot{\delta} = -20\delta + 20K_L\delta_c$$

where δ_c is commanded elevator position and K_L is the control linkage gain. A force stick was used in the simulation and $-\delta_c = K_f F_s$ where F_s is force applied by the pilot and K_f is the force stick sensitivity. K_f was adjusted to insure the pilot had as much control authority as he wanted.

Target Equations

The target was assumed to be a fighter aircraft and the target's normal acceleration A_{nT}, was modeled by driving a second order filter with zero mean, gaussian white noise. The equations are

$$\dot{y} = -\frac{1}{\tau_T} y + \xi$$

$$\dot{A}_{nT} = -\frac{1}{\tau_T} A_{nT} + y$$

$$\dot{\gamma}_T = -.01\gamma_T - \frac{1}{V_T} A_{nT}$$

The dummy variable y is used to put the equations in first order form and ξ is the white noise input. The $-.01\gamma_T$ term in the

target flight path differential equation is an artificial feedback used to keep the steady state variance of γ_T finite. A value of $\tau_T = 3$ sec was used because it resulted in a target motion that looked good.

The inertial line of sight from the attacker to the target is given by (Ref 1)

$$\dot{\Sigma}_T = \frac{1}{D}\left[\vec{s}x\vec{V}_T - \vec{s}x\vec{V}_A\right]$$

where $\vec{s}$ is a unit vector along the line of sight and D is the distance between the attacker and the target. Assuming motion only in the plane,

$$\dot{\Sigma}_T = \frac{V_T}{D}\sin(\gamma_T - \Sigma_T) - \frac{V_A}{D}\sin(\theta - \Sigma_T - \alpha)$$

Assuming that the velocity vectors of the attacker and the target are closely aligned with the line of sight, small angle approximations can be used to arrive at

$$\dot{\Sigma}_T = \frac{V_A}{D}\alpha + \frac{V_A - V_T}{D}\Sigma_T + \frac{V_T}{D}\gamma_T - \frac{V_A}{D}\theta$$

In the simulation, the closing velocity, $V_A - V_T$, was taken to be zero and D was assumed to be constant. This restriction was made so that the simulation equations would not be time varying.

Sight Equations

The sight equation is taken from Ref 1 and a more detailed derivation is contained in Ref 2. The differential equation for lead that was used is

$$\dot{\lambda} = -\frac{1}{T_f}\lambda + q - \frac{V_A}{V_f}\frac{Z_\alpha}{2} + \frac{J_V}{T_f}\alpha$$

where

T_f - time of flight of projectile (sec)

V_f - muzzle velocity of projectile (ft/sec)

J_V - velocity jump correction angle (radians)

The velocity jump correction angle is used to compensate for the misalignment of the velocity vector of the attacker and the projectile muzzle velocity vector. The lead correction for gravity drop is a function of roll angle. Based on the assumption that the plane of flight is not rotating, the gravity drop term will only introduce a fixed bias in the equations and not affect the rms values of the states. For this reason it was not considered in the simulation.

Display

A dual beam oscilloscope was used to display the target and the sight. The display is shown in Fig. 3. The center of the oscilloscope was taken to be the extension of the aircrafts weapon line. The target was an inverted T positioned relative to the weapon line by Σ_{TA}. The sight is positioned relative to the weapon line by λ. Only the reticle portion of sight was displayed due to the limitations of the oscilloscope. In order to accomodate the large lead angles, the display was scaled for 13° per cm. The scope was positioned 20 inches in front of the pilot.

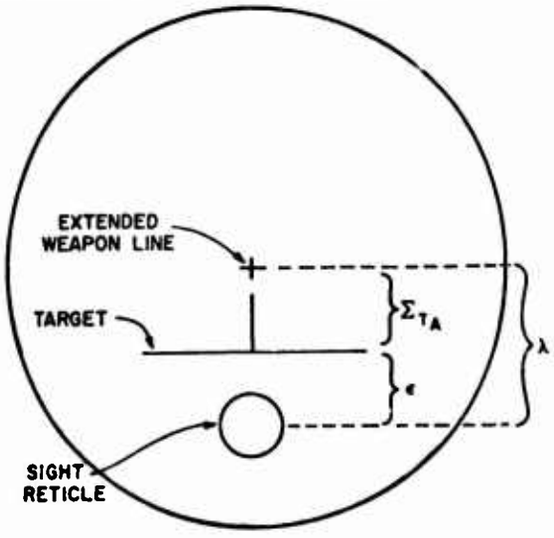


Fig. 3. Simulation Display.

Cases Simulated

Twelve cases were considered by considering all combinations of three aircraft dynamics, two ranges, and two rms normal accelerations for the target.

The three aircraft considered were F-4E, F-5, and A-7. The aircraft and corresponding flight conditions were picked to give a range of aircraft dynamics from good to bad for the air-to-air combat tracking task.

The two ranges simulated were 1000 ft and 3000 ft. In retrospect, the 3000 ft range was not very realistic as a firing range. However, the choice was fortunate in that the sight dynamics for that range were highly oscillatory and the results were useful in developing the analytic model.

The two rms acceleration levels for the target were 3.5g's and 5.0g's.

Test Conduct

Three military pilots were used in the simulation. All three pilots had air-to-air combat experience. Each pilot was thoroughly trained on each of the twelve cases prior to taking data. Each pilot made three data runs for each case.

Different target time histories were used to keep the pilot from learning the target motion. The simulation was conducted so that data was taken only during steady state tracking over a 100 second interval. The mean and the variance of the following variables were determined: elevator deflection, pitch rate, lead angle, normal acceleration of the attacker and tracking error.

Pilot Model

An analytic model was developed for the closed loop air-to-air combat tracking task that was simulated. The pilot was modeled by an optimal pilot model, the form of which was developed by Kelinman, Baron, and Levison (Ref 3). This model has been successfully applied to a number of other tasks including VTOL hover, land-approach, and anti-aircraft gun tracking (Refs 3 through 7). A block diagram of the optimal pilot model is shown in Fig. 4. The displayed variables, y, are observed by the pilot. Visual errors are accounted for by additive gaussian white noise, v_y. The noisy observation is delayed by time τ, which represents a perceptual time delay. The best estimate of the current state, $\hat{x}$, is reconstructed using an optimal estimator and predictor. The control is generated by multiplying the optimal estimate of the state by a set of optimal feedback gains. The feedback gains are optimal for the cost function

$$J(\delta_c) = \lim_{T\to\infty} E\{[M_\varepsilon \varepsilon(T)]^2 + [M_{\dot\varepsilon} \dot\varepsilon(T)]^2 + [M_{\dot\lambda} \dot\lambda(T)]^2 + [M_{\dot\delta c} \dot\delta_e(T)]^2\}$$

Gaussian white noise, v_n, is added to the control output, δ'. This noise accounts for neuromuscular noise and control errors. The pilot output, δ_c, is obtained by passing $\delta + v_n$ through a first order lag with time constant τ_N. Mathematically, the lag is a part of the optimal solution for the cost function $J(\delta_c)$ since it includes a weighting on control rate, $\dot\delta_c$. Physically, can be considered as a model of the lag in the neuromuscular system. The parameters that are used to define the specific model are described next.

Cost Functional Weightings

The weightings in the cost function are defined by the pilot's objective and subjective measures of goodness for the task. The primary objective in the air-to-air combat task is to minimize the error The pilot will also avoid excessive error. rate since it induces undesirable oscillations in the error. The first guess was $M_\varepsilon = M_{\dot\varepsilon}$, $M_{\dot\lambda} = 0$. This worked well for the 1000 ft cases, but the analytic values of rms error for the 3000 ft cases were way too low. The pilots were asked what they did different at 3000 ft than at 1000 ft. It turned out that the sight was extremely oscillatory at the 3000 ft range and the pilots had to be very careful not to cause the sight to oscillate. This subjective consideration is treated by weighting lead rate. The values used are related by

$$M_{\dot\varepsilon} = \frac{1}{4} M_\varepsilon$$

$$M_{\dot\lambda} = \frac{1}{2} M_\varepsilon$$

The value of M_ε is arbitrary.

The weighting on control rate, $M_{\dot\delta c}$, is set to a value that yields $\tau_N = .1$ sec. This value of τ_N is consistent with prior applications of the optimal pilot model (Refs 3 through 7).

Observation Noise

Error and lead are observed by the pilot. In accordance with the model ground

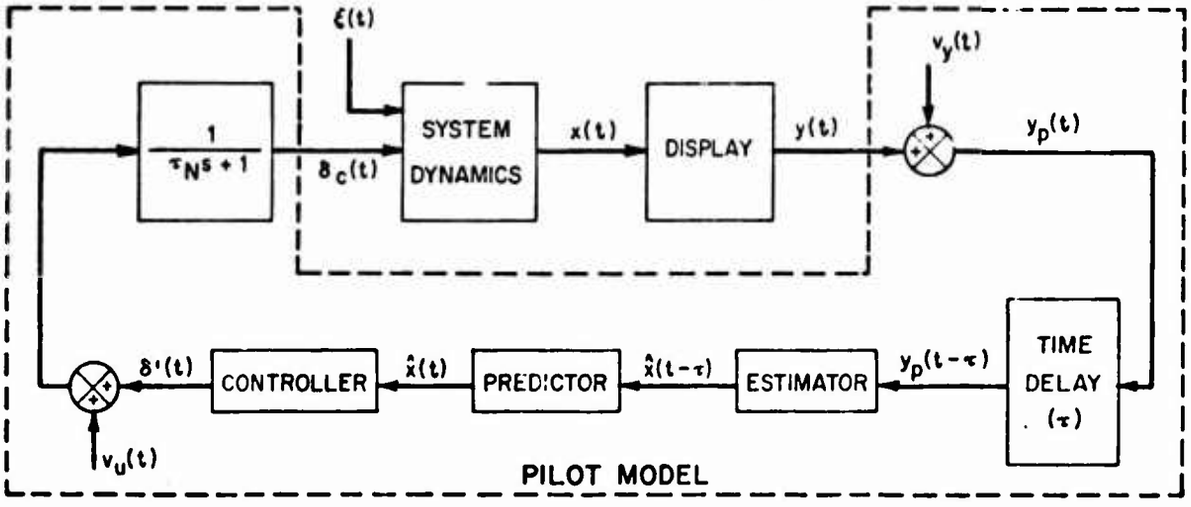


Fig. 4. Functional Schematic of the Optimal Pilot Model in the Control Loop.

rules, it is assumed that the first derivative of the displayed variable are perceived by the pilot (Ref 3). Thus the output vector is

$$y_1 = \epsilon$$

$$y_2 = \dot{\epsilon}$$

$$y_3 = \lambda$$

$$y_4 = \dot{\lambda}$$

Denoting

$$E\{v_{yi}(t)v_{yi}(s)'\} = V_{yi}\delta(t-s),\ i=1,\ldots,4$$

the observation noise is scaled to the rms magnitude of the observation by

$$V_{yi} = \pi\rho_v\sigma_{yi}^2$$

A value of $\rho_v = .01$ was used and is consistent with the value used in Refs 3 through 7. (This corresponds to a -20dB white noise power spectral density level.)

Threshold Effects

Threshold effects associated with visual acuity and the ability to perceive motion were considered in the model. These effects turn out to be significant for this simulation because of the way the display was scaled. To account for the visual thresholds, the perceived output, y_{pi}, is expressed as

$$y_{pi}(t) = f_i[y_i(t)] + v_{yi}(t)$$

where

$$f_i(y) = \begin{cases} y-a_i, & y \geq a_i \\ 0, & -a_i \leq y \leq a_i \\ y+a_i, & y \leq -a_i \end{cases}$$

and a_i is the threshold level. Describing function theory (Ref 8) is used to determine a scale factor k_i as a function of a_i so that the effect of the nonlinearity can be approximated by an increase in the observation noise as follows:

$$y_{pi}(t) = y_i(t) + k_i v_{yi}(t)$$

A typical value for the human threshold for position is .05° of visual arc (Refs 5 and 7). For the position of the pilot and the display scale factor in the simulation, this gives threshold levels of

$$a_\epsilon = a_\lambda = .65°$$

A threshold for human rate perception of .05-.1° visual arc/sec is suggested in Ref 5; however, no specific value was used. In Ref 7, a value of .18° visual arc/sec was used. A value of .18° visual arc/sec for the rate threshold was used. It was found that lower values did not result in as good agreement between the simulated and analytic values of rms error. The value of .18° visual arc/sec results in rate thresholds of

$$a_{\dot{\epsilon}} = a_{\dot{\lambda}} = 2.4°/\text{sec}$$

Time Delay

The pilot time delay was taken to be $\tau = .2$ sec. This value is consistent with Refs 3 through 7.

Neuromuscular Noise

Denoting

$$E\{v_u(t)v_u(s)\} = v_u\delta(t-s)$$

The neuromuscular noise is scaled to the rms magnitude of the pilot output by

$$V_u = \pi\rho_u\sigma_{\dot{\delta}_c}^2$$

Values for ρ_u of from .003 to .01 are suggested in Refs 3 through 7. It was found that better agreement between the simulation and analytic rms values of error was obtained with $\rho_u = .0015$. This corresponds to a white noise power density level of -28dB. There are two possible reasons that the neuromuscular noise had to be relatively small in the pilot model. First, a force stick was used in the simulation. The force stick is a very linear transducer and any errors due to stick nonlinearities are virtually eliminated. Secondly, the pilots selected stick sensitivities that were optimal and control errors should therefore be minimized.

Results

The rms values of the attacker's elevator deflection, pitch rate, lead angle, normal acceleration of the attacker, and tracking error for the model are compared to the simulation data in Figs. 5 through 9. The comparisons are made via scatter diagrams. It can be seen from Fig. 6 that the analytic values of rms pitch rate are slightly higher than the simulation values. The trend is excellent, however. The agreement between analytic values and simulation values of rms elevator deflection, lead angle, attacker normal acceleration, and tracking error is excellent.

Conclusions

The major conclusions is that the model worked! The model provided a faithful reproduction of those results that were measured in the simulation. The model was reasonably simple to develop. The analytic values of rms error and pitch rate are sensitive to the choice of weightings on state in the cost function, i.e., M_ϵ, $M_{\dot{\epsilon}}$, and $M_{\dot{\lambda}}$. The values used were established after three trials and were picked so the data matched. It may have

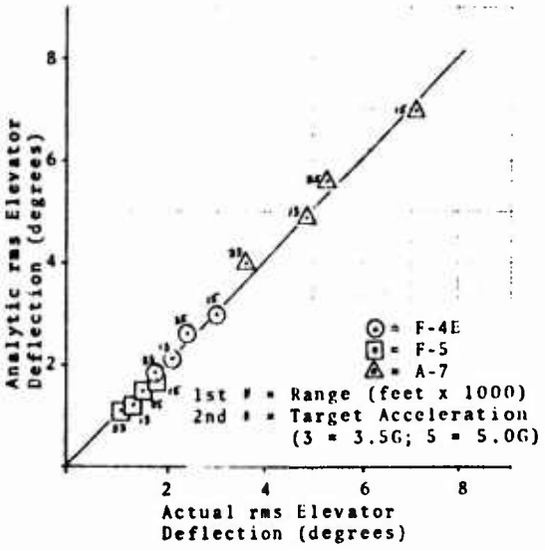


Fig. 5. Comparison of Actual and Analytic Values of rms Attacker's Elevator Deflection, δ.

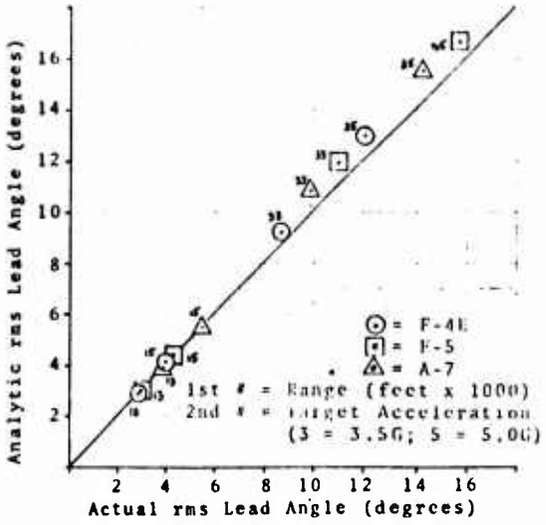


Fig. 7. Comparison of Actual and Analytic Values of rms Lead Angle, λ.

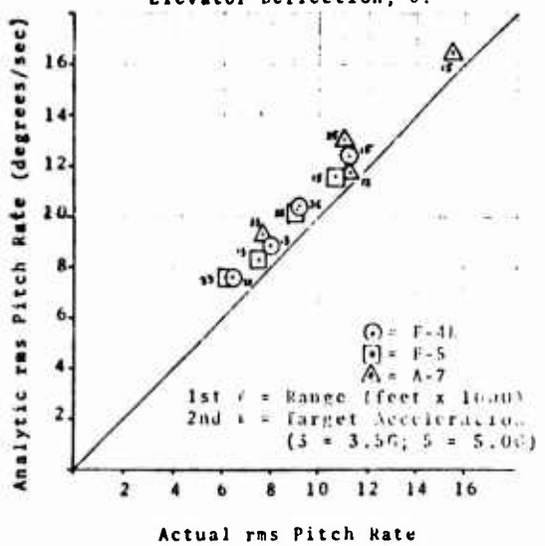


Fig. 6. Comparison of Actual and Analytic Values of rms Pitch Rate, q.

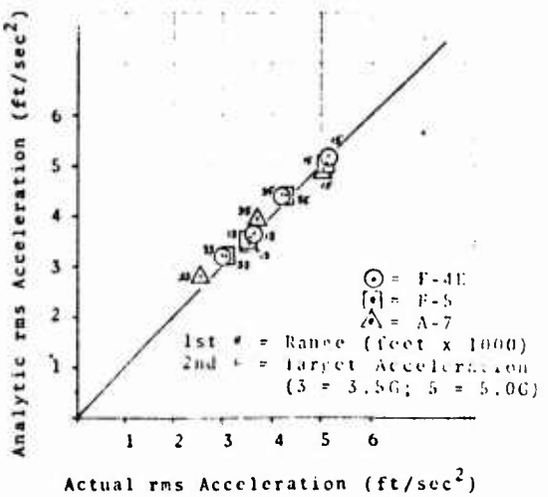


Fig. 8. Comparison of Actual and Analytic Values of rms Attacker Acceleration, A_{N_A}.

been possible to pick the values a priori by judicious questioning of the pilots. The rms value of error was also sensitive to the neuromuscular noise level. The lower value of ρ_u, corresponding to -28dB white noise power density level for the neuromuscular noise, is probably reasonable for an ideal linear control input device. A value of -25dB was used in Ref 6 and the control device was nonlinear. All other parameters used in the pilot model were basically taken from previous applications of the optimal pilot model.

It is not definite at this point if the simulation described, and hence the analytic model, is a suitable representation of the actual air-to-air combat tracking task. The pilots indicated that the dynamics in the simulation were representative of actual aircraft and sight dynamics. However, there are some tenuous assumptions (linear equations, time invariance, in plane motion, etc.) which can not be justified in an off handed manner. The next order of business is to determine if the simple model described in this paper can be

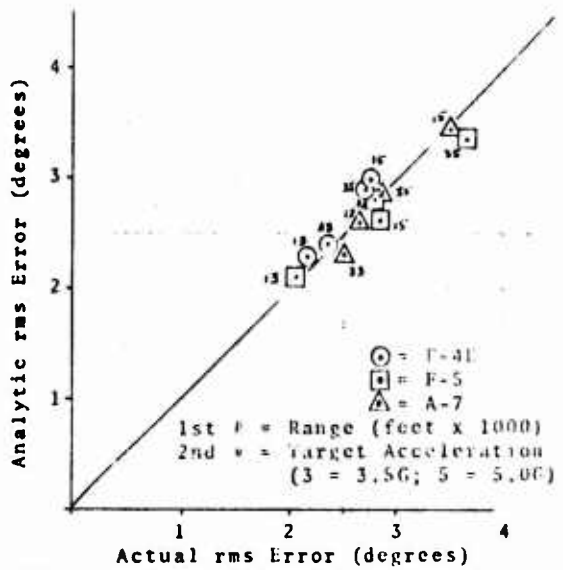


Fig. 9. Comparison of Acual and Analytic Values of rms Error, ε.

used to predict performance in the actual air-to-air combat tracking task. If not, the model can be extended to include lateral dynamics and three dimensional kinematics If it is necessary the time varying dynamics can be included in the model as in Ref 7 and even nonlinearities can be accounted for in the model as in Ref 6.

References

1. Quinlivan, R. P. Multimode Flight Control Definition Study for Precision Weapon Delivery. AFFDL-TR-71-39. Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory, June, 1971.

2. Harvey, Thomas R. Application of an Optimal Control Pilot Model to Air-to-Air Combat. AFIT Thesis, GA/MA/74M-1. Wright-Patterson Air Force Institute of Technology, March 1974.

3. Baron, S., et al. Application of Optimal Control Theory to the Prediction of Human Performance in a Complex Task. AFFDL-TR-69-81. Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory, March, 1970.

4. Kleinman, D. L., S. Baron, and W. H. Levison. "An Optimal Control Model of Human Response, Part 1," Automatica, 6:357-363 (May 1970).

5. Kleinman, D. L. and S. Baron. Analytic Evaluation of Display Requirements for Approach to Landing. NASA CR-1952. Washington: National Aeronautics and Space Administration, November, 1971.

6. Kleinman, D. L. and R. Perkins. "Modelling the Human in a Time-Varying Anti-Aircraft Tracking Loop." Submitted for publication in IEEE Transactions on Automatic Control, May 1973.

7. Levison, W. H. and D. L. Kleinman. Analysis of Pilot/System Performance in Carrier Approach. Cambridge, Massachusetts: Bolt Beranek and Newman, Inc., September, 1971.

8. Graham, Dunstan and Duane McRuer. Analysis of Nonlinear Control Systems. New York: Dover Publications, Inc., 1971.

HUMAN OPERATOR IDENTIFICATION USING LEAST SQUARES, MAXIMUM LIKELIHOOD AND POWER SPECTRAL DENSITY TECHNIQUES

Major John C. Durrett
and
2nd Lt Donald R. McMonagle*

Department of Astronautics and Computer Science
United States Air Force Academy, Colorado 80840

ABSTRACT

This paper describes the results of applying three separate system identification techniques to the identification of a math model of a human operator in a compensatory tracking task.

A least squares technique was attempted, but the results were negative because of excessive state measurement requirements. A maximum likelihood estimation program was used to process the data from the tracking experiments. However, the sensitivity of the gradient search routine was so great that another identification technique had to be used to obtain initial parameter values for the maximum likelihood method. The method used to determine starting values was to compute the power spectral density of the time histories using a fast fourier transform, compute system transfer function values at discrete frequencies, and plot the results in Bode plot form. Estimates of transfer functions from the Bode plots were then used to determine starting parameters for state variable models in the maximum likelihood identification program.

The results of the experiment indicate the advantages of concurrent usage of both time and frequency domain techniques in system identification.

I. INTRODUCTION

The human operator can be a critical part of a control task. A pilot's capability to maneuver his fighter during air-to-air combat can determine the success or failure of

*Former Academy Cadet, now in pilot training in Air Training Command.

the mission. The Department of Astronautics and Computer Science at the United States Air Force Academy has been conducting research in the development of fire control systems for air-to-ground and air-to-air weapon delivery over the past several years. During this research, analytic tools have been developed to evaluate the effectiveness of proposed fire control systems. These analytic models have been used to evaluate systems in either an open loop mode or a closed loop, fully automatic mode. However, no analytic capability exists at the Academy to aid in assessing fire control systems where a human operator is in the closed loop control mode.

In order to develop this capability, an independent study project was initiated in the Fall Semester 1973. The intent of the project was to apply various identification techniques to the identification of a math model of the human operator in compensatory tracking tasks. This would reveal the nuances of the identification schemes as well as the experimental techniques and the data acquisition process. This knowledge would then be used in future projects to develop analytic models of a pilot in an air-to-air fire control task. The major results of this study, which encompassed the fall and spring semesters of school year 1973-1974 are documented in this paper.

II. TEST SET UP

The task presented to the operator in these experiments was a single-degree-of-freedom, compensatory tracking task. The operator's control stick output was the input to a simulated controlled element. The output of the controlled element was summed with a random disturbance to form the system error. This system error was presented to the cperator on an oscilloscope as the difference between a moving horizontal line and a fixed reference mark. The operator's task was to keep the moving line aligned with the reference mark as closely as possible at all times. The operator exercised control with a pencil-type, spring-loaded controller. A symbolic diagram of the compensatory tracking task is shown in Figure 1.

The controlled element was simulated on an Applied Dynamics-4 (AD-4) Analog Computer. The signals i(t), e(t) and c(t) were sampled every 0.1 seconds and, using analog-to-digital converters, stored on a Sigma V Digital Computer in an array format. These three signals were selected in order to follow the analysis procedure of Hess, Reference 1, for

spectral analysis of the compensatory tracking task. Each test run was 120 seconds in duration. The controlled elements for the experiment were a gain (1), an integrator (1/S), and a double integrator ($1/S^2$). The noise source was a Hewlett-Packard 3722A noise generator. Cutoff frequencies used were 0.5 and 1.5 cycles per second (Hz).

All of the system identification computer programs were operational on a Burroughs (B6700) Digital Computer. To accomplish the system identification, the data were dumped onto a digital tape which was handcarried to the B6700. A translator program then manipulated the data into the B6700 digital format and stored the data in a data file on disk. At a later time, then, the identification programs could call the data off the disk file for the system analysis.

III. IDENTIFICATION PROGRAMS

Least Squares Identification: The least squares identification computer program of Reference 2 was obtained and converted to run on the B6700 computer. The program was checked out by identifying simple systems that were simulated and recorded on the AD-4/Sigma V Hybrid Computer. Essentially, the least squares program works as follows. Given a linear time-invariant system that is described by

$$\dot{x} = Fx + Gu \tag{1}$$

where x is an n-vector, u is an m-vector, $\dot{x}$ is the time derivative of the vector x, and F and G are appropriately dimensioned matrices. The least squares program accepts discrete time histories of $\dot{x}$, x and u, and then computes the elements of F and G using a least squares fit technique.

Unfortunately the least squares program was too restrictive to be used in identifying human operator dynamics. In compensatory tracking, the human operator views a single input (the error displayed on the scope) and produces a single output (control stick deflection), which means only two signals can be recorded. If the operator is to be modeled as a second order dynamic system, then the least squares program would require five signals: one for each of the two states and their time derivatives, and one for the observed input. Since these signals would not be available from a compensatory tracking task, the least squares program could not be used to identify the human operator.

Maximum Likelihood Identification:

A FORTRAN computer program that uses maximum likelihood techniques for system identification was obtained from the Air Force Flight Test Center, Reference 3. The program, written by Dr Eugene Cliff of Virginia Polytechnic Institute and State University, will identify a linear time-invariant system of the following form:

$$\dot{x} = Fx + Gu$$
$$y = Cx + Du + \eta \qquad (2)$$

For this system, x is an n-vector, u is an m-vector, y is an ℓ-vector with $m \leq n$, $\ell \leq n$, and F, G, C and D are appropriately dimensioned matrices. The measurement noise, η, is assumed Gaussian with zero mean and covariance of R. The program will accept discrete time histories of the control, u, and the outputs, y, and identify selected parameters, α, of the matrices F, G, C and D, and the initial values of the state vector, x. The critical element of the identification procedure is to maximize the probability density function of the observed measurement noise. This function is dependent on the value of the noise, R, and α at any given time.

$$f(\eta) = \frac{\exp\left[-\frac{1}{2}\eta^T R^{-1}\eta\right]}{(2\pi)^{\ell/2}|R|^{1/2}}$$

When at a maximum, the actual noise is at its highest probability of equaling the observed noise. The values of R and α which maximize the probability density function of observed noise are found using a modified Newton-Raphson iteration technique.

Compensatory tracking data was obtained and an analysis was attempted using the maximum likelihood program. However, two factors were immediately apparent. The convergence of the identification routine was very sensitive to initial parameter estimates. And the order of the math model of the human operator could not be easily established prior to the identification. These two factors combined to make it very difficult to identify the human operator using only the maximum likelihood time domain identification technique. This problem was solved, however, by using spectral analysis

techniques to estimate system order and starting values for the unknown parameters in the human operator model.

Power Spectral Density Analysis:

A FORTRAN computer program was obtained which used discrete fast fourier transform techniques to calculate power spectral densities and cross power spectral densities of the discrete time histories, Reference 4. These power spectral densities were obtained for discrete frequencies over a frequency interval which was determined by the sampling interval and the length of the data run. For the compensatory tracking experiments, the frequency range was 0.05 to 25 radians per second. By recording the error signal, u, presented to the human operator and the operator's control stick output, denoted by y, it was possible to compute the power spectral density (PSD) of the operator's input and output, Φ_{uu} and Φ_{yy}, respectively. Under the assumption that the operator behaved as a linear, time-invariant dynamic system, the PSD of the operator's input was related to his output PSD in the following manner:

$$\Phi_{yy}(\omega) = |Y_p(\omega)|^2 \quad \Phi_{uu}(\omega) \tag{3}$$

where $|Y_p(\omega)|$ is the absolute value of the operator's transfer function representation, Y_p, evaluated at the frequency ω. Equation 3 can be rearranged to give the operator's transfer function.

$$|Y_p(\omega)| = \sqrt{\Phi_{yy}(\omega)/\Phi_{uu}(\omega)} \tag{4}$$

Equation 4 was programmed and used to obtain values of $|Y_p|$ over the frequency range of interest. These values of $|Y_p|$ were plotted versus frequency in a log-log manner to obtain a Bode plot. The plotting was accomplished automatically using a Hewlett-Packard Digital Plotter.

The Bode plots of the human operator dynamics were used to estimate transfer function representations of the operator. These transfer functions were converted into a state space format to use as starting conditions for the maximum likelihood identification program.

IV. OPERATOR IDENTIFICATION

Since the intent of this paper is to show how the identification programs were applied to the identification of a math model of a human operator, only the analysis of the data from one compensatory tracking run will be presented. During this run the "controlled element" of Figure 1 was an integrator, 1/S. Figure 2 shows the results of the power spectral density and subsequent transfer function calculations. Curve (a) of the figure shows the calculated Bode plot of the integrator which is in excellent agreement with the theoretical slope of -20 db per decade (-1 on this curve). Curve (b) is a plot of the calculated magnitude of the product of the operator's transfer function and the controlled element transfer function, $|Y_p Y_c|$, versus frequency. This curve shows a crossover frequency of 5 radians per second. According to the classical human operator theory of Reference 5, the slope of $|Y_p Y_c|$ should go to -1 at the crossover frequency; which it does, sort of. Curve (c) is a plot of the magnitude of the operator's transfer function, $|Y_p|$, versus frequency. The unusual characteristic of this curve is the high frequency lead generated by the operator at 5 radians per second. The data has been thoroughly checked and the lead is present for all runs controlling 1/S and $1/S^2$ systems. A possible explanation for this is that the extremely light stick forces of the control stick allowed the injection of high frequency inputs by the operator.

A larger scale plot of the operator's transfer function is shown in Figure 3. Using the "eyeball" technique, four transfer functions were selected which most closely approximated the Bode plot of Figure 3. These transfer functions are given in Table 1. Using canonical forms of Reference 6, state space equations were obtained for the transfer functions of Table 1. These state space equations were used as initial conditions for the parameters in the maximum likelihood identification program. To account for the effect of the time delay, the discrete time histories were shifted by an appropriate amount prior to the maximum likelihood identification.

The selection of the "best fit" math model for the human operator was accomplished using the covariance of the observation noise, R, as a criterion for acceptance. Those

identification runs with the lowest covariance were considered the best fit for the recorded time histories. After the maximum likelihood identification, the state space equations were transformed into transfer function format for comparison with the initial estimates. Table 2 shows the identified transfer functions with their time delays. The two transfer functions with the lowest noise covariance have been plotted on Figure 3 to compare with the test data. The comparison appears favorable.

V. CONCLUSIONS

The least squares identification program was not applicable to human operator identification. The maximum likelihood identification program was very effective, but extremely sensitive to initial conditions. The power spectral density analysis was found to be necessary to establish starting conditions for the maximum likelihood program.

Future experiments at the Air Force Academy will concentrate on human operator identification and definition in air-to-air and air-to-ground fire control tasks. This first effort has been very successful in debugging the identification programs and in defining the necessary experimental techniques for data acquisition and processing.

REFERENCES

1. Hess, R.A., "An Introduction to Human Describing Function and Remnant Measurement In Single Loop Tracking Tasks," AFFDL/FGC-TM-72-9, Wright-Patterson AFB, Ohio, May 1972.

2. Kugel, D.L. and Pope, R.E., "A Least Squares Approach to Parameter Identification - A Fortran Program," AFFDL/FGC-TM-73-109, Wright-Patterson AFB, Ohio, 1973.

3. Cliff, E.M., "A Maximum Likelihood Method For The Identification of Dynamical System Parameters," Flight Test Technology Branch Office Memo, AFFTC, Edwards AFB, California, September 1973.

4. Koenigsberg, W.D., "Spectral Analysis of Random Signals - Techniques And Interpretation," Report E-2771, The Charles Stark Draper Laboratory, M.I.T., Cambridge, Massachusetts, June 1973.

5. McRuer, D., and Graham, D., "Human Pilot Dynamics In Compensatory Systems," AFFDL-TR-65-15, Wright-Patterson AFB, Ohio, July 1965.

6. Brogan, W.L., Modern Control Theory, Quantum Publishers, Inc., New York, 1974.

Table 1

Initial Estimated Operator Transfer Functions

$4\ e^{-\tau s}$

$\frac{5\ s\ e^{-\tau s}}{(\frac{s}{.7} + 1)}$

$4(\frac{s}{3.0} + 1)e^{-\tau s}$

$\frac{5\ s(\frac{s}{3} + 1)\ e^{-\tau s}}{(\frac{s}{.7} + 1)}$

Range of time delay, τ, was .0 to .4 seconds.

Table 2

Identified Operator Transfer Functions

$4.04\ e^{-.1s}$

$\frac{2.42(\frac{s}{1.03} + 1)\ e^{-.2s}}{(\frac{s}{2.32} + 1)}$

R = .187

R = .147

$4.31(\frac{s}{5.09} + 1)e^{-.3s}$

$\frac{30.7\ s(\frac{s}{5.57} + 1)\ e^{-.3s}}{(\frac{s}{.143} + 1)}$

R = .00517

R = .00492

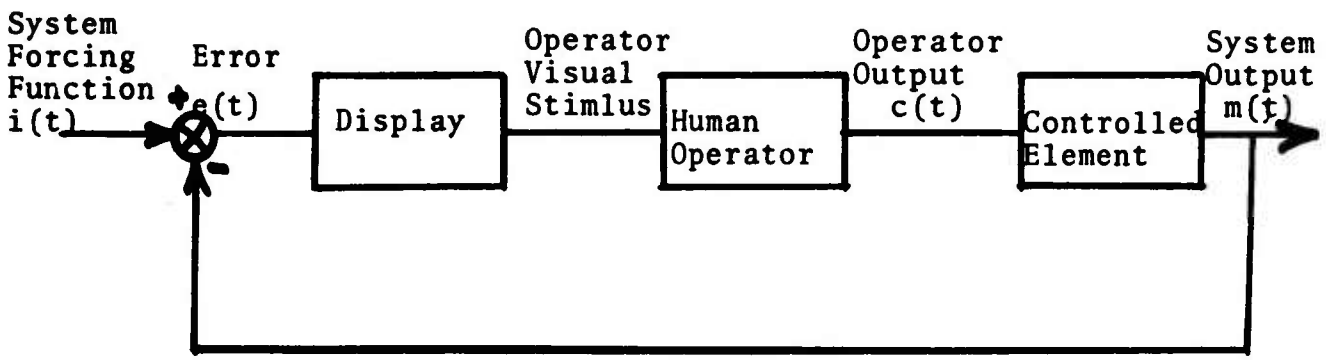


Figure 1

The Human Operator in a Compensatory Tracking Control Task

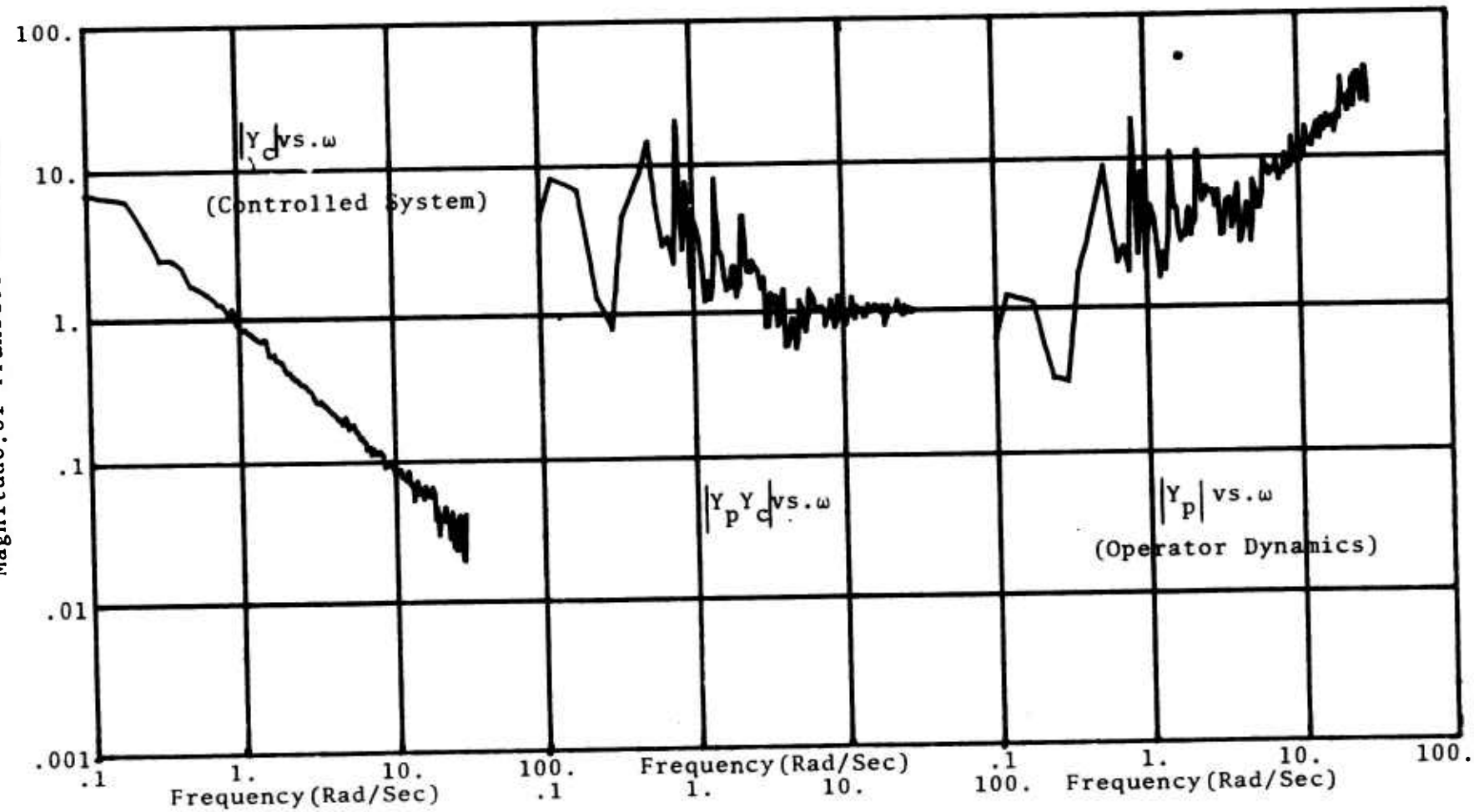


Figure 2. Spectral Analysis Results

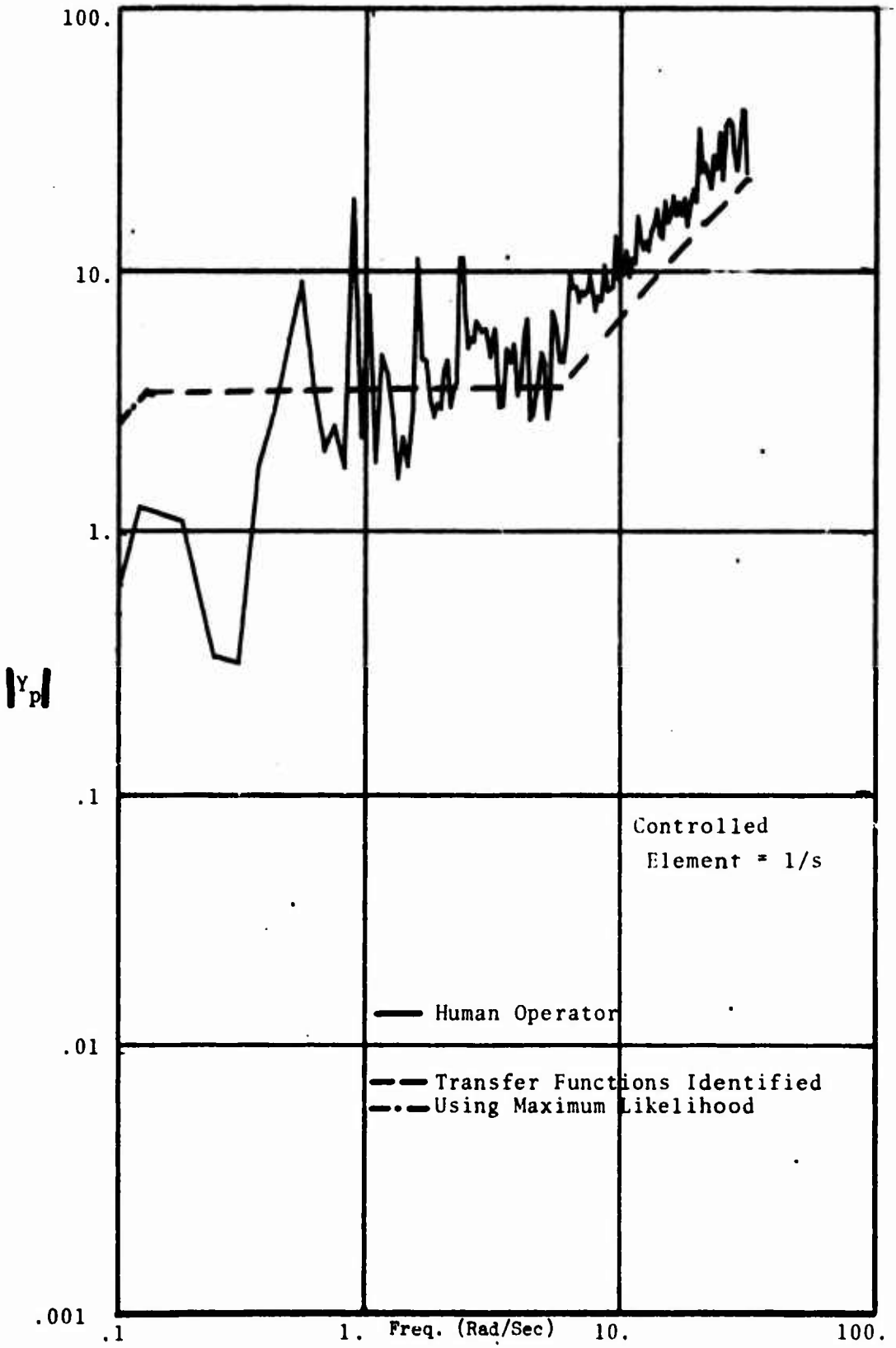


Figure 3. Human Operator Dynamics

ALGEBRAIC REPRESENTATION OF SENSITIVITY OPERATORS FOR LINEAR CONSTANT COEFFICIENT SYSTEMS

by

Captain J. Gary Reid

Air Force Avionics Laboratory

Wright Patterson Air Force Base Ohio

10 May 1974

*Captain Reid is a Ph.D candidate at the Air Force Institute of Technology, Wright-Patterson AFB, Ohio. This paper presents partial results from Capt Reid's Ph.D dissertation [1]. The author wishes to express his sincere appreciation to his dissertation committee - Lt Maybeck, Capt Asher, and Maj Dillow - for their tremendous encouragement in this research.

The author wishes to acknowledge that the views represented in this paper are his own and not necessarily those of the Air Force Institute of Technology, the Air Force Avionics Lab, or the Air Force.

Abstract

This paper presents a new approach for the efficient generation and description of parameter sensitivity operators in linear, time-invariant systems. Rather than attempting to find minimal realizations of the "sensitivity system", this paper develops the sensitivity operators directly from their defining relations as the partial derivatives of the state with respect to the unknown parameters. This approach provides a completely algebraic description of the sensitivity operators, thus enabling the entire sensitivity system to be obtained through matrix manipulations and quadrature integrations.

Perhaps more important than the computational advantages which are emphasized in this paper are the structural insights which are provided by the algebraic description of the sensitivity operators. This physical insight is further enhanced by separately treating the zero input and zero state response sensitivities.

1.0 Introduction

System parameter sensitivity plays a fundamental role in the analysis and design of engineering systems. (eg [2], [3], [4]) However, one of the primary drawbacks in using sensitivity analysis and sensitivity design techniques is often the enormous amount of computation which is required to generate the sensitivity operators. This is particularly true if the system has a large number of states and parameters. The present paper addresses an important subset of this general computational problem - namely, this paper is concerned with the efficient generation and description of sensitivity operators for the linear constant coefficient system S_{LC}

$$\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) \qquad x(o) = \phi(\theta) \tag{1}$$

$x(t)\epsilon\ R^n$ $\qquad u(t)\epsilon\ R^n$

$\theta\epsilon\ R^p$ $\qquad \phi(\theta)\epsilon\ R^n$ $\qquad t\epsilon[o,t_f]$

$A(\theta)\epsilon\ R^{nxn}$ $\qquad B(\theta)\epsilon\ R^{nxr}$

The unknown constant parameter vector, $\theta\epsilon R^p$, may appear in any of the system matrices A, B, or in the initial condition vector ϕ. The components of θ are designated θ_i, i=1, 2, ..p. The nominal parameter vector is designated θ_o. The only restriction that is placed upon S_{LC} is that the unknown matrices $A(\theta)$, $B(\theta)$, and $\phi(\theta)$ be defined, bounded, and differentiable at θ_o. For θ_o and for a prescribed control $u_o(\cdot)\ \epsilon L_2\{R^n,[o,t_f]\}$, the system S_{LC} has the unique solution [5].

$$x_o(t;\theta_o) = \Phi_o(t)\phi(\theta_o) + \int_o^t \Phi_o(t-S)B(\theta_o)\ u_o(s)ds \tag{2}$$

The state transition matrix, $\Phi_o(t)$, for the time invariant system may be determined either from the linear matrix equation

$$\dot{\Phi}_o(t) = A(\theta_o)\Phi_o(t) \qquad \Phi_o(o) = I \tag{3}$$

or from the matrix exponential

$$\phi_o(t) = e^{A(\theta_o)t} \tag{4}$$

This nominal trajectory is designated

$$(x_o, u_o, \theta_o, t) \in R^n \times R^r \times R^p \times [o, t_f]$$

Now the state sensitivity operators for the system S_{LC} are defined as follows:

$$\xi^{(i)}(t;\theta_o,o) \equiv \frac{\partial x(t;\theta)}{\partial \theta_i} \Big| (x_o,u_o,\theta_o,t) \tag{5}$$

$$= \lim_{\Delta\theta_i \to o} \frac{x(t;\theta_o + \Delta\theta_i) - x(t;\theta_o)}{\Delta\theta_i}$$

$$i = 1, 2, \ldots p$$

Notice that the sensitivity operators are, in general, functions of both the current time, t, and the initial time, t_o. However, for time invariant systems, they are only dependent upon the time difference $t-t_o$. Since t_o is selected as zero, this initial time dependency will henceforth be deleted. Also for notational convenience, the dependencies of the matrices and vectors upon θ and the subscripting of the nominal values will be dropped. It will be understood, unless otherwise noted, that all matrices, vectors, and partial derivatives are evaluated along (x_o, u_o, θ_o, t).

Under the assumptions which have been made it is well known [2] that the sensitivity operators for S_{LC} may be determined from the "sensitivity system":

$$\dot{\xi}^{(i)}_{(t)} = A\, \xi^{(i)}(t) + A_{(i)}\, x(t) + B_{(i)} u(t) \tag{6}$$

$$\xi^{(i)}(o) = \phi_{(i)} \qquad i = 1, 2, \ldots p$$

Here and elsewhere the notation is that

$$A_{(i)} \equiv \frac{\partial}{\partial \theta_i} A(\theta) \Big|_{\theta = \theta_o}$$

etc.

The solution of the sensitivity system together with the solution of the system S_{LC} requires solving $(p+1)n$ coupled linear differential equations. For large

p and n this complete solution represents quite a formidable task. Therefore, a number of researchers have investigated the use of low order sensitivity models in the generation of the complete sensitivity system. [7] - [12] These approaches have the potential of reducing the number of required differential equations from $(p+1)n$ to $n(r+3)$. [12]

However, the disadvantages in the use of the low order sensitivity models are several. [12] First, one must be able to transform the nominal system to a suitable canonical form. This generally places the restriction on the nominal system that it be normal [7] or cyclic [9]. Second, once obtaining the transformed sensitivity operators using the companion relationships, one usually wishes to obtain the original system sensitivity operators. The reverse algebraic operations necessary in doing this can require computational and storage burdens which rival the original sensitivity system computations. Additionally, numerical errors in this procedure can be quite severe, particularly if the nominal system has eigenvalues which are near to instability [12] [13]. Lastly, physical interpretation of the sensitivity operators is lost when they are dealt with in the transformed coordinate system.

Although related to these previous efforts, the present paper takes a fundamentally different approach in the analysis of sensitivity operators for the system S_{LC}. Namely, rather than concentrating on the "sensitivity system" and attempting to find minimal realizations for this system, the present paper takes an operator approach by working directly with the defining relations, equation (5). This is a tractable approach, in general, for linear systems because of the ability to write the solution of a linear system explicitly in terms of fundamental matrix solutions. [5] [1] However, for the linear constant coefficient system, S_{LC}, this approach leads to exceedingly simple relizations involving merely algebraic manipulations and quadrature integrations.

But perhaps more important than the computational advantages (which are the emphasis of this paper) are the structural insights provided by the algebraic description of the sensitivity operators. This physical insight is further enhanced by defining and analyzing separate zero input and zero state sensitivity operators. Although not analyzed in this paper, these results should provide powerful design tools for sensitivity consideration in linear constant coefficient systems. Indeed, in [1] these results and this methodology form the basis for algebraic analysis of system parameter identifiability, sensitivity operator controllability, insensitivity, and generalized open loop control design for improved parameter identification.

2.0 Definition of Sensitivity Operators

In this section the description of the sensitivity operators for the system S_{LC} is initiated. Their complete description will depend upon the central theorem of this paper which is given in the next section. Although the definitions and discussion of this section are directed towards the constant coefficient system S_{LC}, it is noted that they are completely general and are valid for linear time varying systems as well. [1]

The zero input and zero state responses for the syster S_{LC} are defined as

$$x_{z.i.}(t) \equiv \phi(t)\phi \tag{7}$$

$$x_{z.s.}(t) \equiv \int_0^t \phi(t-s)Bu(s)ds \tag{8}$$

Then

$$x(t) = x_{z.i.}(t) + x_{z.s.}(t) \tag{9}$$

Now rather than dealing directly with the complete sensitivity operators as defined by equation (5) first two new quantities are defined. These are the zero input and zero state sensitivity operators

$$\xi^{(i)}_{z.i.}(t) = \frac{\partial}{\partial\theta_i} x_{z.i.}(\theta;t) \Big|_{\theta=\theta_o} \qquad i=1,2,..p \tag{10}$$

$$\xi^{(i)}_{z.s.}(t) \equiv \frac{\partial}{\partial\theta_i} x_{z.s.}(\theta;t)\Big|_{\theta=\theta_o} \qquad i=1,2,..p \tag{11}$$

Then it is obvious that the complete sensitivities are given by

$$\xi^{(i)}(t) = \xi^{(i)}_{z.i.}(t) + \xi^{(i)}_{z.s.}(t) \qquad i=1,..p \tag{12}$$

Physically the zero input and zero state sensitivity operators are the parameter sensitivities of the system transient response and the system forced response when initial conditions are zero, respectively. If the system transient response is asymptotically stable then one would expect the zero input sensitivity operators to likewise decay exponentially to zero. The zero state sensitivity operators would then become the parameter sensitivities of the system steady state response. This fact is indeed true and is easily proved as a result of Theorems 1 and 2 of this paper. [1] Also, it is noted that the zero state sensitivities are in the form of a linear operator operating on the control input, u. This particular fact enables one to apply extremely effective sensitivity operator design techniques. [1] [14]

Now the system zero input and zero state responses, equation (7) - (8), may be substituted into the definitions of the zero input and zero state sensitivity operators to yield

$$\xi^{(i)}_{z.i.}(t) = \frac{\partial}{\partial\theta_i}[\phi(t)\phi] \tag{13}$$

$$= \phi_{(i)}(t)\phi + \phi(t)\phi_{(i)}$$

$$\xi^{(i)}_{z.s.}(t) = \int_o^t \frac{\partial}{\partial\theta_i}[\phi(t-s)B]u(s)ds \tag{14}$$

$$= \int_o^t [\phi_{(i)}(t-s)B + \phi(t-s)B_{(i)}]u(s)ds$$

All of the quantities in expression (13) and (14) are easy to obtain with the exception of the partial derivative of the state transition matrix with respect to θ_i:

$$\phi_{(i)}(t) = \frac{\partial}{\partial \theta_i}\,\phi(t) = \frac{\partial}{\partial \theta_i}\,[e^{A(\theta)t}]_{\theta=\theta_o} \tag{15}$$

3.0 Computation of $\phi_{(i)}(t)$

This section considers the partial derivative of the state transition matrix with respect to the parameter component θ_i.

To begin, consider the linear homogeneous system S_H:

$$\dot{x}(t) = A(\theta)x(t) \qquad x(o) = \phi(\theta) \tag{16}$$

$$x(t)\epsilon R^n \qquad \theta\epsilon R^p$$

with unique nominal solution for $\theta_o \epsilon R^p$

$$x_o(t) = \phi_o(t,\theta_o) \tag{17}$$

The notation and assumptions are the same as for the system S_{LC}. Then the sensitivity operators for S_H (i=1,2,..p) are given uniquely by*

$$\xi^{(i)}(t) = \phi(t)\phi_{(i)} + \phi_{(i)}(t)\phi \tag{18}$$

Now consider the equivalent augmented sensitivity system

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\xi}^{(i)}(t) \end{bmatrix} = \begin{bmatrix} A & o \\ A_{(i)} & A \end{bmatrix} \begin{bmatrix} x(t) \\ \xi^{(i)}(t) \end{bmatrix} \tag{19}$$

with initial conditions

$$\begin{bmatrix} x(o) \\ \xi^{(i)}(o) \end{bmatrix} = \begin{bmatrix} \phi \\ \phi_{(i)} \end{bmatrix} \tag{20}$$

This augmented system has the unique solution

$$\begin{bmatrix} x(t) \\ \xi^{(i)}(t) \end{bmatrix} = e^{\begin{bmatrix} A & o \\ A_{(i)} & A \end{bmatrix} t} \begin{bmatrix} \phi \\ \phi_{(i)} \end{bmatrix} \tag{21}$$

*For simplicity the existence of $\phi_{(i)}(t)$ is here merely assumed. In [1] a rigorous proof of existence is given. This also provides both an integral equation and differential equation representation of $\phi_{(i)}(t)$. Although such representation may be generalized to other linear systems, they are not of particular value computation in the time invariant case and are so not included here.

However, from (17) and (18) and the uniqueness of the representations it must be true that

$$e^{\begin{bmatrix} A & o \\ A_{(i)} & A \end{bmatrix} t} = \begin{bmatrix} \phi(t) & o \\ \phi_{(i)}(t) & \phi(t) \end{bmatrix} \tag{22}$$

For convenience, define the augmented matrix by

$$\tilde{A}^{(i)} \equiv \begin{bmatrix} A & o \\ A_{(i)} & A \end{bmatrix} \tag{23}$$

Then the characteristic polynomial of $\tilde{A}^{(i)}$ is

$$\tilde{f}^{(i)}(\lambda) = f(\lambda) \cdot f(\lambda) \tag{24}$$

where $f(\lambda)$ is the characteristic polynomial for A. Now suppose that $\{\lambda_k; k=1,2,..\sigma\}$ are the set of distinct roots of $f(\lambda)=0$ and that n_k, $k=1,2,..\sigma$, are the multiplicity of each distinct eigenvalue λ_k in $f(\lambda)$. Then

$$f(\lambda) = \prod_{k=1}^{\sigma} (\lambda - \lambda_k)^{n_k} \tag{25}$$

and

$$\tilde{f}^{(i)}(\lambda) = \prod_{k=1}^{\sigma} (\lambda - \lambda_k)^{2n_k} \tag{26}$$

Futhermore, this relationship is true for each parameter component θ_i, $i=1,2,..p$. But now it is well known (eg,[15]) that

$$e^{\tilde{A}^{(i)}t} = \sum_{j=o}^{2n-1} (\tilde{A}^{(i)})^j \, a_j(t) \tag{27}$$

where the $a_j(t)$ are the unique solution of the set of 2n linear equations

$$t^{\ell} e^{\lambda_k t} = \frac{d^{(\ell)}}{d_{\lambda_k}(\ell)} \begin{bmatrix} 1 & \lambda_k & \lambda_k^2 \ldots \lambda_k^{2n-1} \end{bmatrix} \begin{bmatrix} a_o(t) \\ a_1(t) \\ \cdot \\ \cdot \\ \cdot \\ a_{2n-1}(t) \end{bmatrix} \tag{28}$$

$k = 1,2,..\sigma$

$\ell = 0,1,..2n_k-1$

But then it is easily shown that

$$(\tilde{A}^{(i)})^j = \begin{bmatrix} A^j & o \\ \frac{\partial}{\partial \theta_i}(A^j) & A^j \end{bmatrix} \tag{29}$$

And so using equation (22) we have established

Theorem 1

Let A (θ) be an nxn real matrix which is dependent upon the constant parameter vector $\theta \epsilon R^p$. Suppose that for a nominal $\theta_o \epsilon R^p$ the matrix A (θ) is differentiable with respect to each parameter component θ_i, i=1,2,..p, at θ_o, and that the characteristic polynomial of $A(\theta_o)$ is given by

$$f(\lambda) = \prod_{k=1}^{\sigma} (\lambda - \lambda_k)^{n_k} \tag{30}$$

Then

$$e^{A(\theta_o)t} = \sum_{j=o}^{2n-1} [A(\theta_o)]^j \, a_j(t) \tag{31}$$

and

$$\frac{\partial}{\partial \theta_i} [e^{A(\theta)t}]_{\theta=\theta_o} = \sum_{j=o}^{2n-1} \frac{\partial}{\partial \theta_i} [A(\theta)]^j \, a_j(t) \tag{32}$$

$i=1,2,..p$

where the scalar functions $\{a_j(t);\ j=0,1,..2n-1\}$ are uniquely determined from the 2n linear equations (28)

Comments

1. Equations (31) and (32) may be put in a minimal "basis function" representation involving the minimal polynomials of $A(\theta_o)$ and $\tilde{A}^{(i)}(\theta_o)$, respectively. By considering the adjoint matrix of $\tilde{A}^{(i)}(\theta_o)$ it is relatively easy to show that the order of the minimal polynomial of $\tilde{A}^{(i)}(\theta_o)$ is between the bounds $n_m \leq \tilde{n}_m^{(i)} \leq 2n_m$, where n_m and $\tilde{n}_m^{(i)}$ are the order of the minimal polynomials of $A(\theta_o)$ and $\tilde{A}^{(i)}(\theta_o)$ respectively. [1] [16]
2. For computer applications it is conveient to compute the partial derivatives

$$\frac{\partial}{\partial \theta_i} [A^j(\theta)]_{\theta=\theta_o} \qquad j=0,1,..2n-1$$

from the recursive relation

$$\frac{\partial}{\partial \theta_i} [A^j(\theta)]_{\theta=\theta_o} = \frac{\partial}{\partial \theta_i} [A^{j-1}(\theta)]_{\theta=\theta_o} A(\theta_o) + A^{j-1}(\theta_o) \frac{\partial}{\partial \theta_i} [A(\theta)]_{\theta=\theta_o} \tag{33}$$

4.0 Algebraic Representation of the Sensitivity Operators

In this section Theorem 1 is used to provide an algebraic description of the zero input and zero state sensitivity operators. In fact, based upon equations (13) and (14) of Section 2.0 and Theorem 1, one immediately obtains

Theorem 2

For the system S_{LC}

$$x_{z.i.}(t) = \sum_{j=o}^{2n-1} A^j \phi \, a_j(t) \tag{34}$$

$$x_{z.s.}(t) = \sum_{j=o}^{2n-1} A^j B \int_o^t a_j(t-s)u(s)ds \tag{35}$$

$$\xi^{(i)}_{z.s.}(t) = \sum_{j=o}^{2n-1} \frac{\partial}{\partial \theta_i} [A^j \phi] \, a_j(t) \tag{36}$$

$$\xi^{(i)}_{z.s.}(t) = \sum_{j=o}^{2n-1} \frac{\partial}{\partial \theta_i} [A^j B] \int_o^t a_j(t-s)u(s)ds \tag{37}$$

$i = 1,2,..p$

The scalar functions $\{a_j(t): j=o,1,....2n-1\}$ are determined from equation (28), Theorem 1.

Comments

1. Similar relations for the system sensitivities are easily obtained by inclusion of a measurement matrix, say $H(\theta)$.
2. Again for computation one would use

$$\frac{\partial}{\partial \theta_i} [A^j \phi] = \frac{\partial}{\partial \theta_i} [A^j] \phi + A^j \frac{\partial}{\partial \theta_i} [\phi] \tag{38}$$

and comment 3 of Theorem 1, etc.

3. Equations (34) - (37) give a great deal of structural and physical insight into the system sensitivity operators [1]. On the other hand, the convolution integrals which are indicated in equations (35), (37) can be quite troublesome for digital computer computations. However, through a judicious examination of the scalar functions, $a_j(t)$, these convolution integrals may be transformed

to ordinary quadrature integrals. Using our established notation it is relatively straightforward to obtain the following Lemma. [1]

Lemma 1

Let Λ be the 2nx2n non-singular generalized Vandermonde matrix

$$\Lambda \equiv \left| \frac{d^{\ell}}{d\lambda^{\ell}} \left[1 \quad \lambda_k \quad \lambda_k^2 \cdots \lambda_k^{2n-1}\right] \right| \quad \begin{matrix} k = 1,2,..\sigma \\ \ell = 0,1,..2n_k-1 \end{matrix} \tag{39}$$

Let $\alpha^{(j)} = [\alpha_{k\ell}^{(j)}]_{1 \times 2n}$; $k = 1,2,..\sigma$; $\ell = 0,1,..2n_k-1$;

be the j^{th} the row vector of Λ^{-1}. Then

$$a_j(t) = \sum_{k=1}^{\sigma} \sum_{\ell=0}^{2n_k-1} \alpha_{k\ell}^{(j)} \, t^{\ell} \, e^{\lambda_k t} \tag{40}$$

and for each component $u_q(\cdot)$, q=1,2,..r, of the control vector $u(\cdot)$

$$\int_0^t a_j(t-s)u_q(s)ds =$$

$$\sum_{k=1}^{\sigma} \sum_{\ell=0}^{2n_k-1} \sum_{\nu=0}^{\ell} (-1)^{\nu} \binom{\ell}{\nu} \alpha_{k\ell}^{(j)} \, t^{(\ell-\nu)} \, e^{\lambda_k t} \tag{41}$$

$$\int_0^t s^{\nu} \, e^{-\lambda_k s} \, u(s)ds$$

where $\binom{\ell}{\nu} = \frac{\ell!}{(\ell-\nu)!\nu!}$

Comments

1. Through the use of Theorem 1 and Lemma 1, one may obtain the complete system and sensitivity system, both state and output, zero input and zero state, through merely matrix manipulations and the solution of the 2nr quadrature integrals

$$\int_0^t s^{\ell} \, e^{\lambda_k s} \, u_q(s)ds$$

$k = 0,1,..\sigma$; $\ell = 0,1,..2n_k-1$; and $q = 1,2,..r$.

This replaces the (p+1)n linear differential equations for the "sensitivity system" approach.

2. Although the generalized Vandermonde matrix, Λ, may be of rather large dimension $(2n \times 2n)$, specialized techniques exist for obtaining its inverse. (see eg, [17])

5.0 Extension to Second Order Sensitivity Operators

In this section the approach of sections 3.0 and 4.0 are outlined to give an algebraic representation of the second order sensitivity operators. The second order sensitivity operators for the system S_{LC} are defined as

$$\xi^{(i,j)}(t) = \frac{\partial^2}{\partial\theta_i \partial\theta_j} [x(\theta;t)]_{\theta=\theta_o} \tag{55}$$

with corresponding definitions for the zero input and zero state sensitivities.

Again consider the homegeneous system S_H:

$$\dot{x}(t) = A(\theta)\, x(t) \qquad x(o) = \phi(\theta) \tag{56}$$

The matrix $A(\theta)$ is now assumed to be twice differentiable at θ_o. The first and second order sensitivity operators for this system at the nominal parameter θ_o are then given uniquely by

$$x(t) = \Phi(t)\phi \tag{57}$$

$$\xi^{(i)}(t) = \Phi_{(i)}(t)\phi + \Phi(t)\phi_{(i)} \tag{58}$$

$$\xi^{(j)}(t) = \Phi_{(j)}(t)\phi + \Phi(t)\phi_{(j)} \tag{59}$$

$$\xi^{(i,j)}(t) = \Phi_{(i,j)}(t)\phi + \Phi_{(j)}(t)\phi_{(i)} + \Phi(t)\phi_{(i,j)} + \Phi_{(i)}(t)\phi_{(j)} \tag{60}$$

As before, all matrices and partial derivatives are evaluated at θ_o and the double subscript (i,j) denotes the second partial derivative.

Next consider the equivalent augmented "sensitivity system"

$$\frac{d}{dt}\begin{bmatrix} X(t) \\ \xi^{(i)}(t) \\ \xi^{(j)}(t) \\ \xi^{(i,j)}(t) \end{bmatrix} = \begin{bmatrix} A & 0 & 0 & 0 \\ A_{(i)} & A & 0 & 0 \\ A_{(j)} & 0 & A & 0 \\ A_{(i,j)} & A_{(j)} & A_{(i)} & A \end{bmatrix} \begin{bmatrix} X(t) \\ \xi^{(i)}(t) \\ \xi^{(j)}(t) \\ \xi^{(i,j)}(t) \end{bmatrix}$$

$$\begin{bmatrix} X(o) \\ \xi^{(i)}(o) \\ \xi^{(j)}(o) \\ \xi^{(i,j)}(o) \end{bmatrix} = \begin{bmatrix} \phi \\ \phi_{(i)} \\ \phi_{(j)} \\ \phi_{(i,j)} \end{bmatrix}_{\theta=\theta_o} \tag{61}$$

As in section 3.0, by considering the unique solution of this augmented system and comparing it with equations (57 - 60), it is evident that

$$e^{\tilde{A}^{(i,j)}t} = \begin{bmatrix} \phi(t) & 0 & 0 & 0 \\ \phi_{(i)}(t) & \phi(t) & 0 & 0 \\ \phi_{(j)}(t) & 0 & \phi(t) & 0 \\ \phi_{(i,j)}(t) & \phi_{(j)}(t) & \phi_{(i)}(t) & \phi(t) \end{bmatrix} \tag{62}$$

Now suppose that $m(\lambda)$ is the minimal polynomial of A of order n_m, and that the characteristic polynomial of A is given by $f(\lambda) = m(\lambda)g(\lambda)$. Then using the adjoint relation [16, p83]

$$\text{adj}\,(\lambda I - \tilde{A}^{(i,j)})\,(\lambda I - \tilde{A}^{(i,j)}) = \det\,(\lambda I - \tilde{A}^{(i,j)})I$$

$$= (f(\lambda))^4 I \tag{63}$$

one can show that $f(\lambda)g^3(\lambda)$ is a common factor of each component of $\text{adj}(\lambda I - \tilde{A}^{(i,j)})$ for all i,j. Thus the order of the minimal polynomial of $\tilde{A}^{(i,j)}$ must be less than or equal to $3n_m \leq 3n$. Then by methods similar to those of Section 3.0 we obtain:

Theorem 3

Let $A(\theta)$ be a real nxn matrix which is twice differentiable at θ_o and has characteristic polynomial

$$f(\lambda) = \prod_{k=1}^{\sigma} (\lambda - \lambda_k)^{n_k} \tag{64}$$

Then

$$e^{At} = \sum_{j=o}^{3n-1} A^j b_j(t) \tag{65}$$

$$\frac{\partial}{\partial\theta_i} [e^{At}] = \sum_{j=o}^{3n-1} \frac{\partial}{\partial\theta_i} [A^j] b_j(t) \tag{66}$$

and

$$\frac{\partial^2}{\partial\theta_i\partial\theta_i} [e^{At}] = \sum_{j=o}^{3n-1} \frac{\partial^2}{\partial\theta_i\partial\theta_i} [A^j] b_j(t) \tag{67}$$

All matrices and partial derivatives are evaluated at $\theta=\theta_o$ and the 3n scalar function $b_j(t)$ are uniquely determined from the 3n linear equations:

$$t^{\ell} e^{\lambda kt} = \frac{d^{\ell}}{d\lambda_k^{\ell}} [1 \ \lambda_k \ \lambda_k^2 \ . \ . \ . \ \lambda_k^{3n-1}] \begin{bmatrix} b_0(t) \\ b_1(t) \\ \cdot \\ \cdot \\ \cdot \\ b_{3n-1}(t) \end{bmatrix} \tag{68}$$

$k = 1, 2, \ . \ . \ . \ \sigma$

$\ell = 0, 1, \ . \ . \ 3n_k-1$

Comments

1. Theorem 3 may now readily be applied to the definition of the second order sensitivity operators, equation (55) and etc., to provide an algebraic description of the second order sensitivity operators similar to that of Theorem 2. Thus the entire system, first, and second order sensitivity operators may be obtained through mere matrix manipulations and, now, 3nr quadrature integrals. These computations would replace the solution of $n(p + 1 + pJ)$ coupled linear differ-

ential equations which would be required for an equivalent "sensitivity system" approach.

2. These second order results should be of great significance in the use of second order numerical optimization techniques such as Newton-Raphson algorithm (eg, [4]).

This paper has provided an algebraic description of the sensitivity operators for linear constant coefficient systems. Computational advantages have been emphasized in this paper, but the methodology and the algebraic results which have been presented offer powerful design tools for many system considerations. [1]

References

[1] Reid, J. Gary, "Sensitivity Operators and Associated System Concepts for Linear Dynamic Systems", Ph.D Dissertation, Air Force Institute of Technology, Wright-Patterson AFB, Ohio 1974.

[2] Tomovic, R., and M. Vukobratovic, General Sensitivity Theory, American Elsevier, New York, 1972.

[3] Sage, A. P. and J. L. Melsa, System Identification, Academic Press, 1971.

[4] Wilde, D. J., and C. S. Beightler, Foundations of Optimization, Prentice-Hall, 1967.

[5] Brauer, F. and J. A. Nohel, Qualitative Theory of Ordinary Differential Equations, W. A. Benjamin Inc., 1969.

[6] Wilkie and Perkins, "Generation of Sensitivity Functions for Linear Systems Using Low-Order Models", IEEE Transactions on Automatic Control, Vol AC-14, No 2, April 1969.

[7] Wilkie, D. F. and W. R. Perkins, "Essential Parameters in Sensitivity Analysis", Automatica, Vol 5, P 191-197, 1969.

[8] Guardabassi, G., A. Locatelli, and S. Rinaldi, "On the Optimality of

the Wilkie-Perkins Low-Order Sensitivity Model", IEEE Transactions on Automatic Control, June 1970.

[9] Denery, D. G., "Simplification in the Computation of the Sensitivity Functions for Constant Coefficient Linear Systems", IEEE Transactions on Automatic Control, August 1971.

[10] Neumann and Sood, "Sensitivity Functions for Multi-Input Linear Time-Invariant Systems", International Journal of Control, 1971, Vol 13, No 6, 1137-1150.

[11] _____, "Sensitivity Functions for Multi-Input Linear Time-Invariant Systems - II. Minimal Order Models", Internationa Journal Control, 1972, Vol 15, No 3, 451-463.

[12] Varshney, R. K., "Reduced - Order Sensitivity Models of Linear Systems with Applications", Ph.D. Dissertation, University of Illinois, Urbana Illinois, 1972.

[13] Singer, R. A. "Selecting State Variables to Minimize Eigenvalue Sensitivity of Multivariable Systems", Automatica, Vol 5, Pergamon Press, 1969.

[14] Mehra, R. K., "Optimal Inputs for Linear System Identification", Paper 28-5, Preprints, 1972 Joint Automatic Control Conference, Stanford, Calif, August 1972.

[15] Zadeh, L. A. and C. L. Desoer, Linear System Theory, McGraw-Hill, 1963.

[16] Gantmacher, F. R., The Theory of Matrices, Vol 1, Chelsea Publishing Co., N.Y., 1959.

[17] Goknar, I. C., "Obtaining the Inverse of the Generalized Vandermonde Matrix of the Most General Type", IEEE Transactions on Automatic Control, October 1973.

TUNEABLE INTEGRATION FOR AIR FORCE APPLICATIONS

Marc L. Sabin, Captain, USAF
Research Associate
Frank J. Seiler Research Laboratory
USAF Academy, Colorado 80840

Introduction

The purpose of this paper is to introduce you to and give you a feel for a new way of looking at numerical integration of dynamic systems. In the course of this introduction, I will give you the mathematical formulation of a specific integrator which may be able to solve some of your accuracy and computation-time problems associated with numerical integration.

The subject of this paper is "Tuneable Integration," (TI), a technique originally proposed by Jon M. Smith of Software Research Corporation.[1] The precepts for the approach arise in sampled-data control theory; specifically, they are based on digital filters employing variable phase and gain to control distortion in the integrated "signal." In most classical integrators this distortion arises from the polynomial approximations used in their development, and integration step size is the only available control on the amount of distortion (classically called truncation error). The TI, on the other hand, encounters distortion due to reconstruction of the sampled signal, and it employs variable phase and gain compensation to minimize the effects of that distortion.

Though the results reported herein are from my own work, I do not extend the theory of Smith. My purpose, as indicated earlier, is to introduce you to Tuneable Integration, for I believe that this is a technique of significant potential value to the Air Force. The particular items I will discuss in this paper are as follows:

1. The Zero-Order-Hold Tuneable Integrator (ZOH-TI),
2. Application of the ZOH-TI to a damped, second-order oscillator;
3. Application of the ZOH-TI to the Air Force Academy algorithm for a digital lead-computing optical sight system.

Zero-Order Hold Tuneable Integrator

Before proceeding with the development, let me explain a basic guideline which governs a number of decisions to follow. The guideline is simplicity, and its purpose is twofold. By keeping the result as simple as possible, programming complexity and computation time are minimized. Secondly, simplicity minimizes the number of poles in the numerical integrator. Ideally we would like to have a result with only one pole, the number possessed by a real, continuous integrator. Recall that each extraneous pole induces more phase lag, and hence a greater tendency toward instability.

Now let us proceed with the development. Our objectives are to model a discrete approximation to the continuous integrator, to define the input-output transfer function, and to finally write the difference equation which represents the numerical integrator.

The basic form of the discrete approximation is presented in Figure 1.

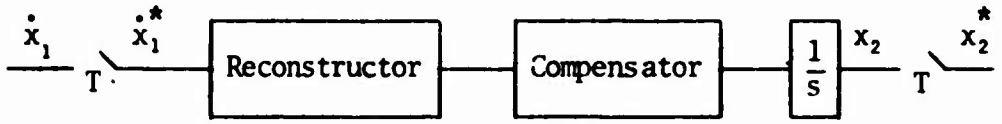


Figure 1. Basic Discrete Integrator

Referring to the figure, we have a continuous signal $\dot{x}_1$ to be integrated. Samples are taken at intervals T to give the data sequence $\dot{x}_1^*$. A continuous signal is reconstructed, compensated for phase and gain distortion, and integrated yielding the signal x_2, which is approximately the integral of $\dot{x}_1$. The second sampler, again with interval T, provides the discrete, output data train x_2^*.

The reconstructor may be chosen in many ways. Smith has looked at zero-, first- and second-order and trapezoidal holds. The higher the order of the hold the more complex the integrator becomes. This additional complexity yields one or two extra poles and zeros in the z-plane,† and the integrators yield more high frequency distortion. Thus I will restrict the discussion to the zero-order-hold reconstructor.

The purpose of the compensator is to counter the distortion of the reconstructor with gain (λ) and phase (γ) adjustments. The basic

† The poles are at $z = 0$, so that $z \triangleq e^{sT}$ results in the real part of the s-plane pole being at minus infinity.

form of the compensator is $\lambda e^{\gamma sT}$, and there are two primary considerations with regard to its application. The first involves continuous vs discrete compensation. Smith proposed a form of discrete compensation using a triangular hold and a first order expansion of $e^{\gamma sT}$. For the higher order reconstruction processes the discretely compensated integrator is simpler than the continuously compensated integrator with the same reconstructor and the same expansion of $e^{\gamma sT}$, while for the zero-order hold the result is the same as for continuous compensation. I will discuss only the case of the first-order expansion of the exponential for reasons to be mentioned shortly; and thus, since both forms of compensation give the same result for the zero order hold, I will not go into further detail regarding discrete compensation.

The second consideration pertains to the expansion of $e^{\gamma sT}$. From the definition of the z-transform, $e^{\gamma sT} = z^{\gamma}$, and for continuous variation in γ we would have non-integer powers of z in our transfer function. To prevent this, a series expansion of the exponential is desirable. The simplest expansion is the first order expansion

$$e^{\gamma sT} \approx 1 + \gamma sT \tag{1}$$

Two other expansions looked at by Smith were $1/(1-\gamma sT)$ and $(2+\gamma sT)/(2-\gamma sT)$. Each of these results in a more complex integrator with extraneous poles and is thus excluded from further discussion here.

We have now narrowed our field of consideration down to the zero-order-hold tuneable integrator with "single-zero" compensation (this being the name given to the expansion of Equation 1 by Smith), and we designate it ZOH-TI. Figure 2† depicts the ZOH-TI.

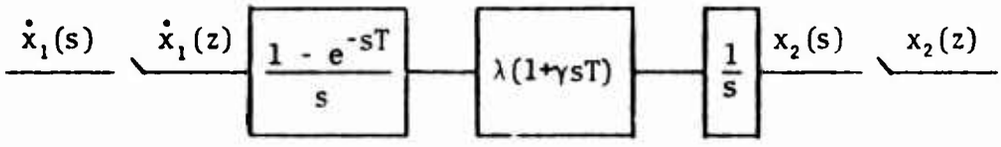


Figure 2. ZOH-TI

The z-plane transfer function for the discrete integrator is evaluated as follows:

† This figure is essentially the same as Figure 5 in Reference 1.

$$\frac{x_2(z)}{\dot{x}_1(z)} = Z\left\{\frac{1 - e^{-sT}}{s} \cdot \lambda(1+\gamma sT) \cdot \frac{1}{s}\right\}$$

$$= Z\left\{1 - e^{-sT}\right\} \cdot Z\left\{\frac{\lambda(1+\gamma sT)}{s^2}\right\}$$

$$(2) \qquad \frac{x_2(z)}{\dot{x}_1(z)} = \frac{\lambda T\,[\gamma z + (1-\gamma)]}{z - 1}$$

We see that there is a single pole at $z = 1$, so that the integrator is neutrally stable. The difference equation, which is our desired numerical integrator, is found from Equation 2 by solving for $x(z)$ in terms of $\dot{x}(z)$ and noting that z is a unit advance in the time domain; i.e.,

$$Z^{-1}\left\{zx(z)\right\} = x(t_n+T)$$

The result is

$$(3) \qquad x_n = x_{n-1} + \lambda T\left[\gamma\dot{x}_n + (1-\gamma)\dot{x}_{n-1}\right]$$

where $x_n = x(t_n)$ and n replaces the $1,2$ subscripts used earlier. Equation 3 is inherently a closed type integrator, for it involves the derivative $\dot{x}_n$; but the weighting of this derivative, call it a leading derivative, may be varied inversely to the weighting of $\dot{x}_{n-1}$, which we can call a lagging derivative. It is the ability to vary this weighting of lead and lag that allows us control of the integrator accuracy.

A very significant observation is made by Smith regarding classical numerical integration and various forms of the tuneable integrator. He states that "from the frequency-domain viewpoint, many of these [classical] integrators are actually the same integrator, differing only in the amount of phase shift of the integrand."[1] For instance, as γ takes on the values 0, $1/2$, 1, and $3/2$, Equation 3 takes the forms of the Euler, trapezoidal, rectangular and implicit Adams second order integrators.

Test Performance

Having developed the form of the ZOH-TI, we shall now apply it to the integration of a second-order oscillator's response to various

inputs. The test vehicle was, as indicated, a second-order oscillator, represented by the differential equation

$$\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x = \omega_n^2 F(t) \tag{4}$$

where ζ is the damping ratio, ω_n is the undamped natural frequency, and $F(t)$ is the forcing function. The standard specifications used were $\zeta = 0.3$ and $\omega_n = 1$ Hz. The tests examined the step response, the impulse response, and the response to sinusoidal forcing at resonance. The step response was also examined with critical damping ($\zeta = 1.0$) and a high natural frequency ($\omega_n = 5$ Hz). In all cases the results were similar, indicating an "optimal" choice† of γ in the neighborhood of 0.5. For example, Figures 3-5 demonstrate the resonance test results for three values of γ(0.0, 0.5, and 1.0), each used with a range of integration time steps (T = 0.001, 0.010, 0.050, 0.100). Clearly the case of $\gamma = 0.5$ (i.e., a trapezoidal integrator) is the best of the three shown; we will see shortly that $\gamma \equiv 0.5$ is not necessarily precisely the best choice. In Figure 5 a factor of 50 increase in the integration time step has resulted in virtually no change in the integrator output.

Before making a detailed examination of the results of the step response test, a comment is necessary concerning Smith's reported results for the same type of test. While my results indicate that $\gamma \approx 0.52$ gives the best performance, Smith obtained his best results with $\gamma = 1.0$. The two results are consistent and serve to demonstrate an important point. That point is that when modeling a system (such as in the development of an aircraft simulator) the manner in which the continuous elements of the system are discretely modeled will have a strong impact on resultant performance.

Let me explain why mine and Smith's results are consistent, and at the same time I will demonstrate the import of the manner of modeling. Figure 6 is a block diagram of the second-order system after it has been discretized.

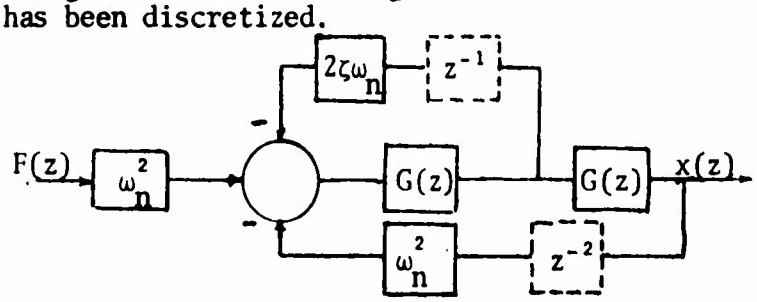


Figure 6. Discretized Second-Order System

† I will say more about choosing γ and λ later. For these tests $\lambda = 1$ was the value used.

In the figure, $G(z)$ represents the ZOH-TI transfer function given by Equation 2. There are two delays (z^{-1} and z^{-2}) indicated in the dotted blocks of the feedback loops. In my model I do not include these delays, for in the simple integration situation I have the values of $\dot{x}$ and x when I require them. Smith, however, proceeds from a different viewpoint and includes the delays† which are "characteristic of the explicit formulation of discrete systems where the forward loop must be computed before the feedback can be computed and the loop closed."[1] If the delays are included, the characteristic equation is fourth order, while if they are not it is second order. If $\gamma = 1$ in the first case, however, two of the poles move to the origin where they are cancelled by two zeros. The resultant second-order system is then very nearly the same as the one for the case of no delays. Using $\omega_n = 1$ Hz, $\zeta = 0.3$ and $T = 0.01$ seconds, we find the poles of the no-delay model with $\gamma = 1/2$ to be at $z \approx 1.00$ and .96 and for the delay model with $\gamma = 1$, at $z \approx .98 \pm j.06$. It is therefore very reasonable that my results using $\gamma \approx 0.5$ should be similar to Smith's results using $\gamma = 1.0$. It should also be clear that the manner in which the overall system is modeled will have a strong influence on the results obtained.

Let us now look in some detail at the results of the step response test. Figures 7-11 each show the system output for a given value of γ with four values of time step T. As γ increases from 0.40 to 0.60 in the sequence of figures, the ZOH-TI progresses from a condition of too much lag to one of too much lead. The "true" output in each case is generated by using a small enough step size so that changes in γ have little effect on the output. That time step was chosen to be 0.001 seconds.

In Figure 7, $\gamma = 0.40$ and we have a case of too much lag. As T increases from 0.001 to 0.100, each respective curve overshoots the previous curves, for the lag in the integrator increases with T and a greater tendency toward instability exists.

Figures 8, 9 and 10 depict the response near $\gamma = 0.52$. In all three cases, a factor of 50 increase in step size results in no discernible change in the plotted curve. A two order of magnitude increase in T results in relatively small errors. The thing to observe from these three figures is the shift from overshoot to undershoot with increasing T. This results in a gradual shift of the crossover points from left to right, as can be easily seen for the minimum near $t = 1$ sec. It is this change in the effect of increasing T from one of increasing overshoot to one of increasing undershoot that indicates the "optimal" γ for this situation is in the neighborhood of $\gamma = 0.52$.

---

† Still another formulation is possible: it results in the same set of z-plane poles, but has one less zero.

Note that we are picking the proper γ in an empirical manner by analyzing these figures. The question of what is truly optimal is not an easy one, for the change in the effect of T is gradual. One possible criterion would be the minimum RMS deviation over all integration points between the output at a given T and the so-called true output. I will say a little more concerning the choice of γ near the end of this paper.

The last figure of this sequence, Figure 11, evinces the excessive lead induced by values of γ greater than about 0.52. The undershoot is obvious, as is the shift of crossover points to the right.

To emphasize the effect of γ rather than T, Figure 12 shows the step response for a range of γ values with $T = 0.100$ seconds. The curve for $\gamma = 0.50$ was computed with $T = 0.001$ to give a "true" output history. Though the individual curves are hard to distinguish, the values shown for $t = 0.5$ sec and the general envelope of the five curves clearly demonstrate the lead/lag effect of γ.

As indicated earlier, the results of all tests were similar to those shown here. Therefore, I will not go into a detailed discussion of those results, but rather I will now discuss the application of a ZOH-TI to the integration algorithm for a lead-computing, optical gunsight.

Lead Computing Optical Sight System (LCOSS)

The LCOSS is a fire control system which computes the lead angle necessary to fire upon a target being tracked at a specified range. Background for this type of sight is discussed in the paper on "Non-Director Gunsight Response" contained elsewhere in these proceedings. Researchers at the Air Force Academy have developed some improvements to previous solutions of the LCOSS problem, and in their testing desired to evaluate various numerical integrators for usage in their version of the gunsight: the tuneable integrator was included among these.[2] McClendon's computer program was modified and further testing was conducted to obtain the results shown below.

The differential equation being integrated is that of a simple, first-order system with an external forcing function, and the quantity being integrated is the lead-angle derivative. Since the details of the equation are not important for our present purposes, let me simply say that for given flight and weapon specifications (e.g., acceleration, speed, altitude, muzzle velocity, angle of attack, etc.)

the equation is one of constant coefficients and is simply integrated for a period of time necessary for the bullet to reach the target's range. In the case of straight-and-level flight the lead angle need only compensate for gravity, and the solution may readily be determined analytically. This condition was used for the tests from which data will be shown.

The bases for integrator comparison were the accuracy of the output and the computation time required to obtain it. Since the program was tested on a B6700 computer with virtual memory, run times were not repeatable, and only a rough comparison should be made of the relative computation times. Further, none of the programming was optimized for minimum computation time; and though all of the integrators were penalized in this respect, some could be affected more than others. Again, only rough comparisons can be made from the relative computation time figures.

The equations for the five integrators used in the test are shown in Table 1. Note that Euler and rectangular integration are also included in this list by letting $\gamma = 0$ and $\gamma = 1$, respectively, for the tuneable integrator. The Simpson integrator is shown because it was included in the Academy's program, but it is unstable in a dynamic situation, the integration error growing without bound.[3]

Table 1. Numerical Integrators for $\dot{x} = f(x,t)$

Integrator	Formula
4th Order Runge-Kutta	$x_{n+1} = x_n + \frac{1}{6}\left(k_1+2k_2+2k_3+k_4\right)$ $k_1 = T\, f\left(x_n\ ,\ t_n\right)$ $k_2 = T\, f\left(x_n + \frac{1}{2}k_1,\ t_n + \frac{1}{2}T\right)$ $k_3 = T\, f\left(x_n + \frac{1}{2}k_2,\ t_n + \frac{1}{2}T\right)$ $k_4 = T\, f\left(x_n + k_3,\ t_n + T\right)$
Trapezoidal	$x_{n+1} = x_n + \frac{T}{2}\left(\dot{x}_{n+1} + \dot{x}_n\right)$
Simpson	$x_{n+2} = x_n + \frac{T}{3}\left(\dot{x}_{n+2} + 4\dot{x}_{n+1} + \dot{x}_n\right)$
3/8 Newton-Cotes	$x_{n+3} = x_n + \frac{3T}{8}\left(\dot{x}_{n+3}+3\left(\dot{x}_{n+2}+\dot{x}_{n+1}\right)+\dot{x}_n\right)$
ZOH-TI	$x_{n+1} = x_n + \lambda T\left(\gamma\dot{x}_{n+1} + (1-\gamma)\,\dot{x}_n\right)$

Two points should be noted at this time. The first is that no predictor-corrector formulations have been used. The second is that the derivative at the time being integrated to (our so-called "leading derivative") is required for all but the Runge-Kutta integrator; i.e., they are closed type integrators. The two points relate strongly to each other, for the corrector is typically a closed type integrator which uses the output of the predictor to evaluate the leading derivative. I have employed a simple iteration scheme to converge on simultaneous values of the leading derivative and output by using as an initial guess for the derivative the derivative at the last computed time step. In the ZOH-TI, for example, $\dot{x}_{n+1}$ is initially set to $\dot{x}_n$, x_{n+1} is calculated and a new $\dot{x}_{n+1}$ computed and so on. Agreement between two consecutive values of x_{n+1} to a set convergence criterion results in moving to the next time interval. For the LCOSS, testing was done to see the effect of the convergence criterion: there was little effect over a wide range. Letting the criterion vary between 10^{-4} and 10^{-2}, I found that the ZOH-TI with $\gamma = 0.60$ showed a change only after the fifth decimal digit. The reason for the low sensitivity to the convergence criterion is the smooth nature of the exponential rise in the computed lead angle. (The true solution in the case tested was given by the expression $a(1-e^{-bt})$.)

Even with a loose convergence criterion, at least one iteration must be done at every time step and this is costly in terms of computation time, as the results below will show. Thus it is desirable to consider formulating the tuneable integrator into predictor-corrector schemes. Smith has formulated a few of these; but I have not yet examined them, nor formulated others. Another possibility, which worked well for the smooth exponential involved in these tests of LCOSS, is to use an old value of the derivative and employ more lead compensation through a larger value of γ. Both of these possibilities require further examination.

The integration tests involved a five-second time of flight with time steps varied between .02 and .20 sec at increments of .02 sec. The data of Table 2 were obtained using the largest time step and loosest convergence criterion. The units of time are in sixtieths of seconds and the lead angle is given in mils. The numbers in parentheses indicate values of γ.

Immediately obvious are the large errors in the Simpson and Newton-Cotes integrator outputs. A trapezoidal integrator was used to start the Simpson technique and the large initial error grew to the value shown; an extremely accurate start-up value, artificially inserted into the algorithm, reduced the error but did not eliminate it.

Table 2. LCOSS Test Results

T = 0.20 Convergence Criterion = 0.01

Integrator	Lead Angle	Process Time
Runge-Kutta	6.455610566336	9.3
Trapezoidal	6.455610534656	11.9
Simpson	1154.3256.....	24.3
Newton-Cotes	3.770586.....	21.0
ZOH-TI(0.0)	6.455610566336	10.6
ZOH-TI(0.2)	6.455610566336	11.2
ZOH-TI(0.4)	6.455610566336	11.8
ZOH-TI(0.6)	6.455605657856	13.6
ZOH-TI(0.8)	6.452087800576	25.7

True Value = 6.455610566451

As more lead is inserted into the ZOH-TI, the derivative becomes too small too fast and the steady state value is reached at a later time. Hence we see the decrease in the outputs for the last two entries in the table. Too much lag results in attaining the steady state at an earlier time and there is no apparent impact on the computed output. This has much significance for the choice of an LCOSS integrator. If the lead angle can typically be expected to grow as a smooth exponential, and if the target range will be such that the final value of the exponential rise will be attained, then a simple Euler integrator (the same output as the ZOH-TI with $\gamma = 0$) will suffice. No leading derivative will then be required, thereby eliminating the need for any iteration and significantly reducing the computation time from the values shown in Table 2.

As noted previously, the time figures, in Table 2 should only be used for rough comparisons. With that in mind, we see that the tuneable integrator gives equivalent accuracy to the Runge-Kutta for nearly the same computation time. The number of iterations required for convergence during the early time steps penalizes the ZOH-TI severely. Since $\gamma = 0$ is equivalent to an Euler integrator, such an integrator could be programmed explicitly and would avoid the single iteration that is made by the tuneable integrator at every step.† This would essentially half the time required without affecting accuracy. A word of caution, however, is in order: at time steps larger than those investigated these results would most likely be changed somewhat.

The conclusion to be drawn from this LCOSS study is that while flexible, accurate and relatively simple, tuneable integration is

† Current programming of the ZOH-TI does not recognize that for $\gamma = 0$ $\dot{x}_{n+1}$ is not required, and hence a single iteration is made at every step when $\gamma = 0$.

not necessarily appropriate to problems where still simpler techniques will suffice.

Choosing λ and γ

In the example of the second-order oscillator which we discussed earlier, an "optimal" choice of phase compensation, γ, was found empirically by trial and error. This is typically what presently must be done when a complex system is involved. A fruitful area for investigation lies in examining potential means for determining a priori what the proper choice should be. The selection of λ is subject to the same concern, though apparently less freedom is available in its choice. This is because λ scales the steady state output and values other than unity can cause difficulties. (From the LCOSS results we can note that a value slightly greater than unity would be appropriate.) Smith has also found that λ has less influence on the ZOH-TI performance than γ, and he recommends selection initially of the latter and then the former. Also noted by Smith is the primary influence of γ on the transient response as opposed to the influence of λ on the steady state output.

For some systems it is possible to apply analytic techniques to the determination of λ and γ. Two possible methods discussed by Smith are root matching of the discrete system poles to those of the continuous system and, secondly, minimizing the RMS deviation of the output from the analytically determined response to a known forcing function. It is easy to see how system complexity can preclude the application of techniques such as these.

Conclusions

From the discussion presented in this paper, one should be able to see that there exists in tuneable integration the potential for significant reductions in programming complexity and computation time (through the ability to use larger integration time steps) by providing the user with external control of integrator phase and gain characteristics. Though it does provide a simple yet flexible and accurate technique, tuneable integration is not a panacea and may be superfluous in cases such as the LCOSS, where a simpler integrator is sufficient.

The real potential for this technique lies in applications to complex systems where numerical integration time must be minimized. Two possible areas of application are in missile range-safety tracking and in SLBM tracking and impact prediction.

A significant area for further research concerns finding a priori methods of determining the proper integrator phase and gain for complex systems.

References

1. Smith, Jon M., "Recent Developments in Numerical Integration for Digital Simulation, Digital Control, and Digital Information Systems," Recent Developments in Modern Numerical Methods, Seminar Notebook. Saint Louis: Software Research Corporation, 1973. (See also the Journal of Dynamic Systems, Measurement and Control, 96 (1974) 61-70.)

2. McClendon, Mike H., "Accuracy Research on the Digital Lead Computing Optical Sight System." Department of Astronautics and Computer Science Course Report. USAF Academy, 1974.

3. Ralston, Anthony, "Numerical Integration Methods for the Solution of Ordinary Differential Equations." Mathematical Methods for Digital Computers, Vol I, Ralston and Wilf, eds. New York: John Wiley and Sons, Inc., 1960.

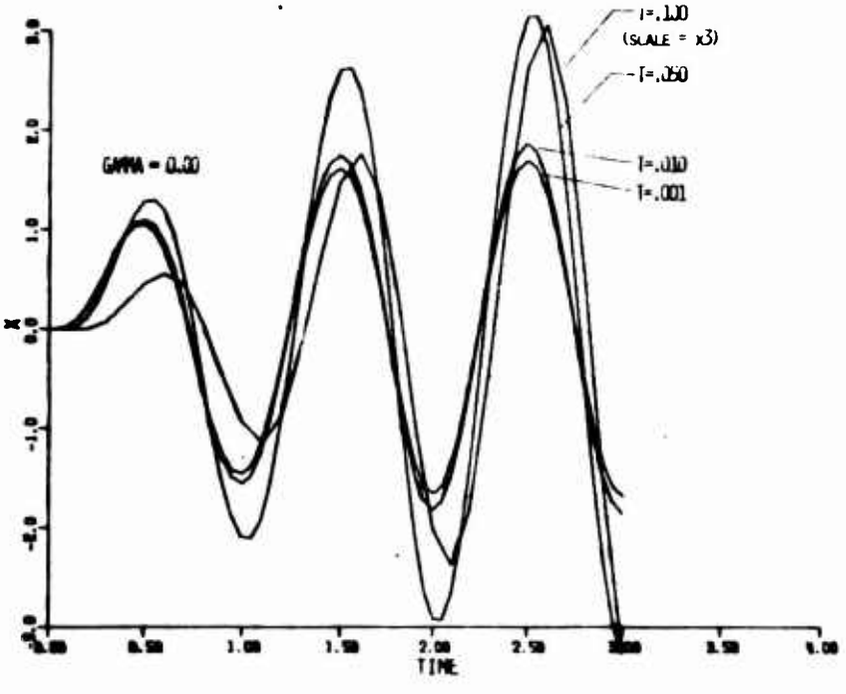


Figure 3. Response to Sinusoidal Forcing at Resonance: Excessive Lag

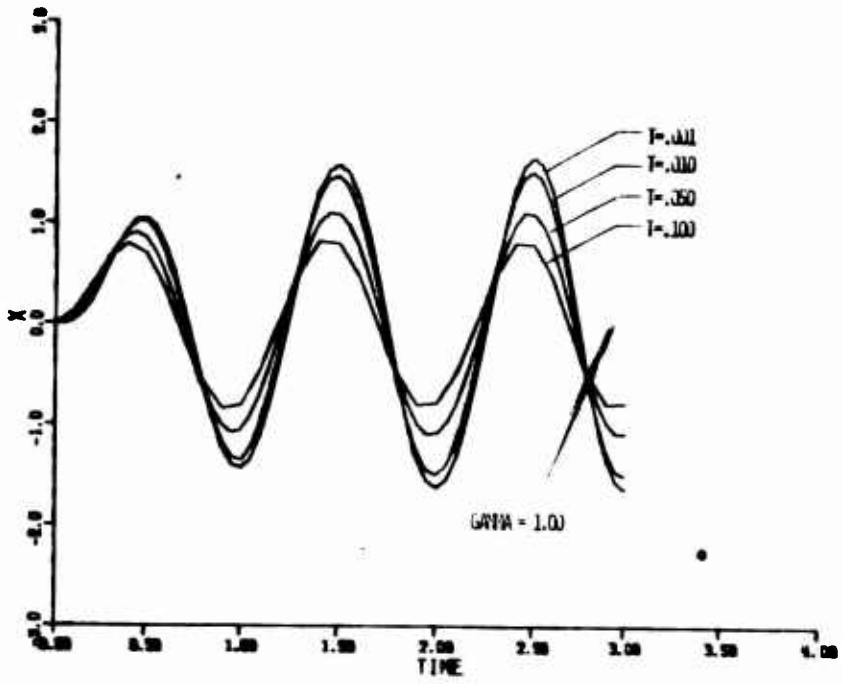


Figure 4. Response to Sinusoidal Forcing
at Resonance: Excessive Lead

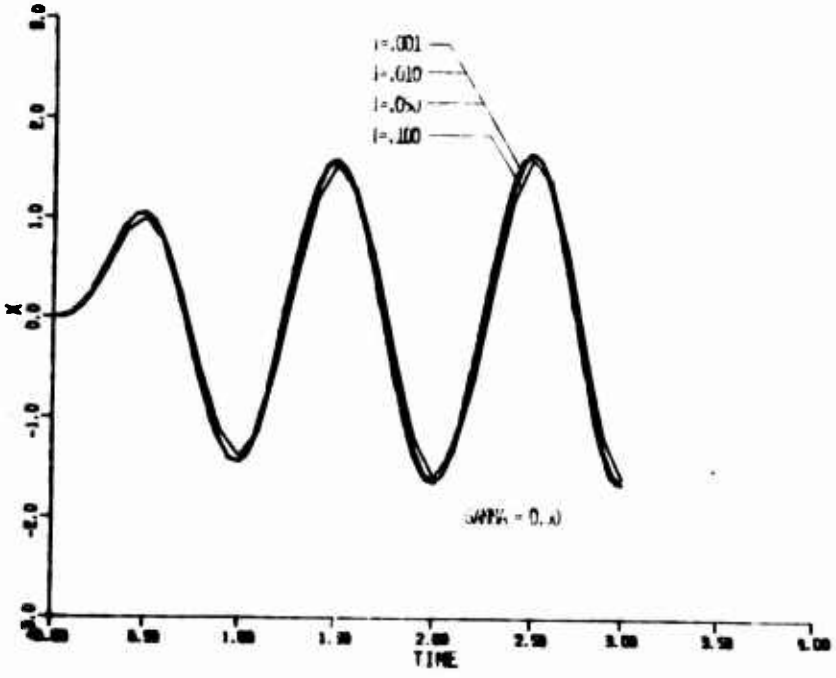


Figure 5. Response to Sinusoidal Forcing
at Resonance: Nearly Optimal Compensation

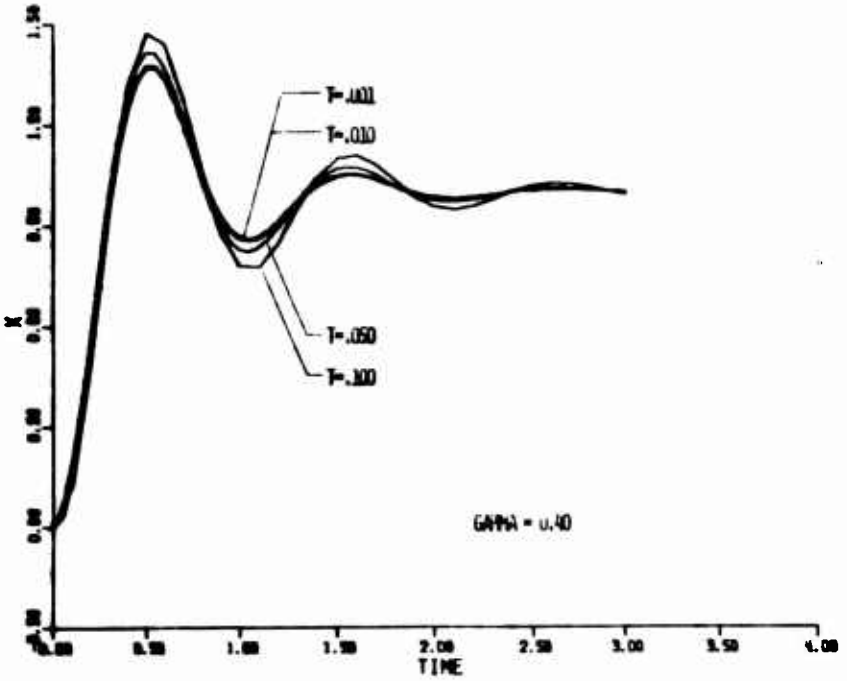


Figure 7. Step Response: Excessive Lag

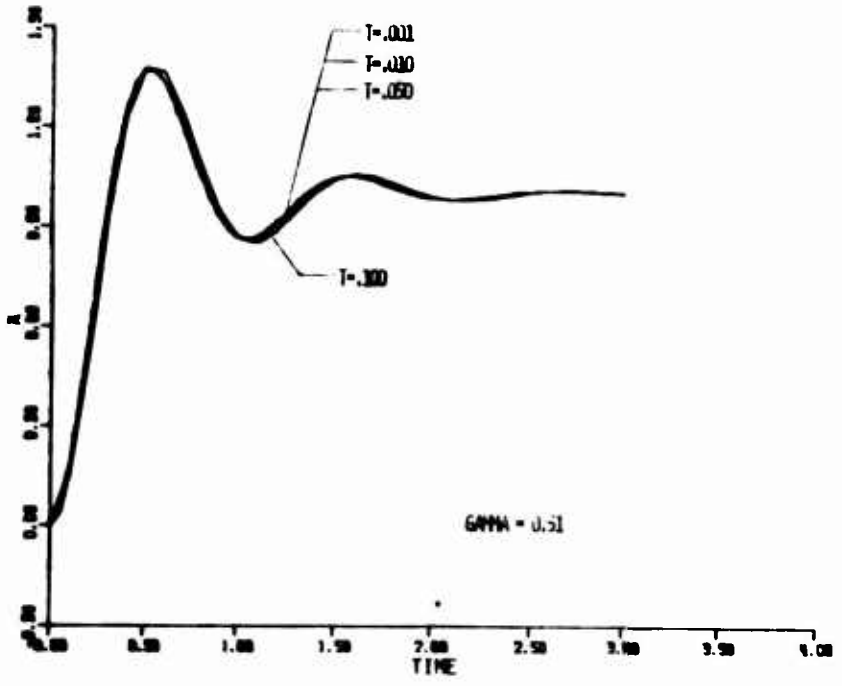


Figure 8. Step Response: Compensation
Slightly Less Than Optimal

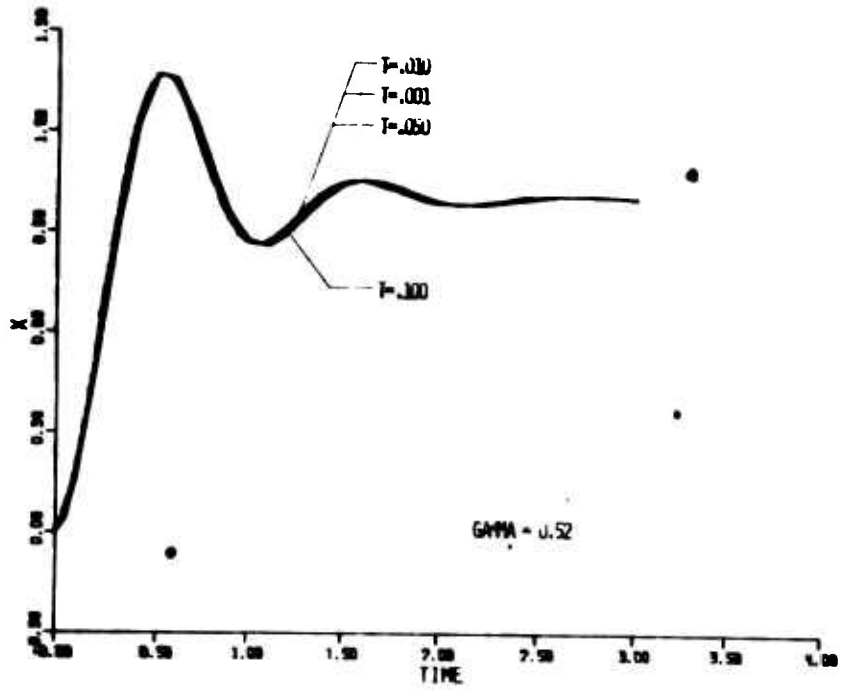


Figure 9. Step Response: Compensation Approximately Optimal

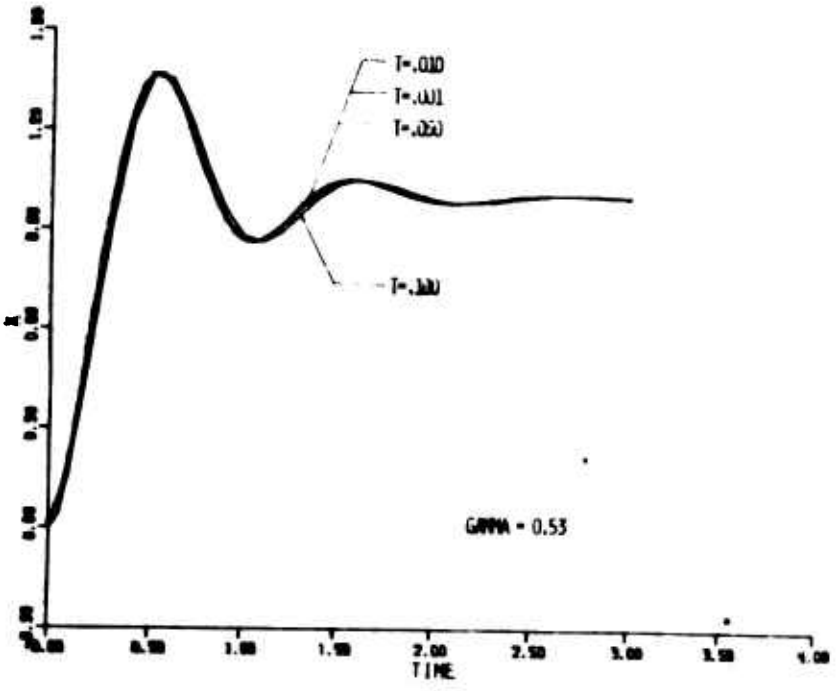


Figure 10. Step Response: Compensation Slightly More Than Optimal

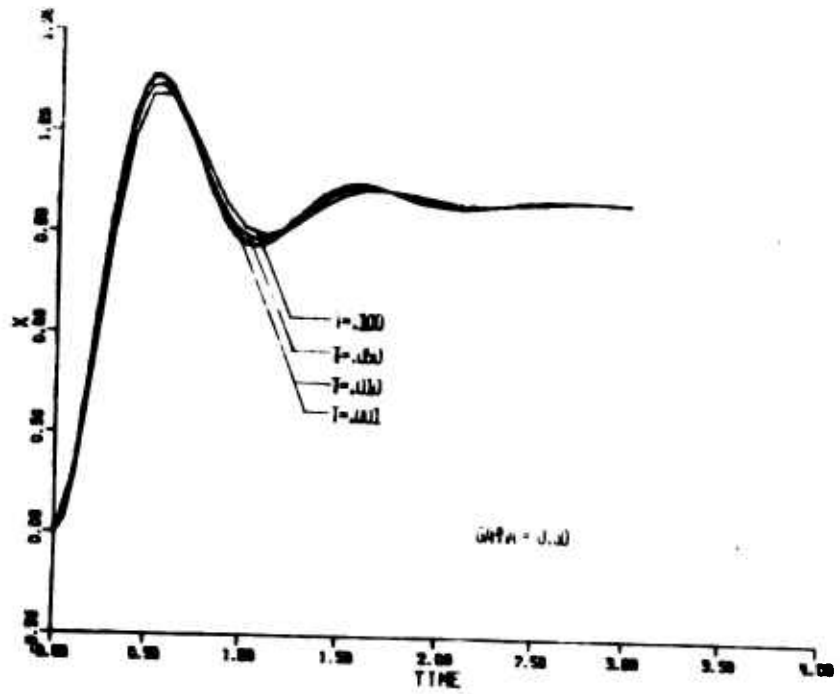
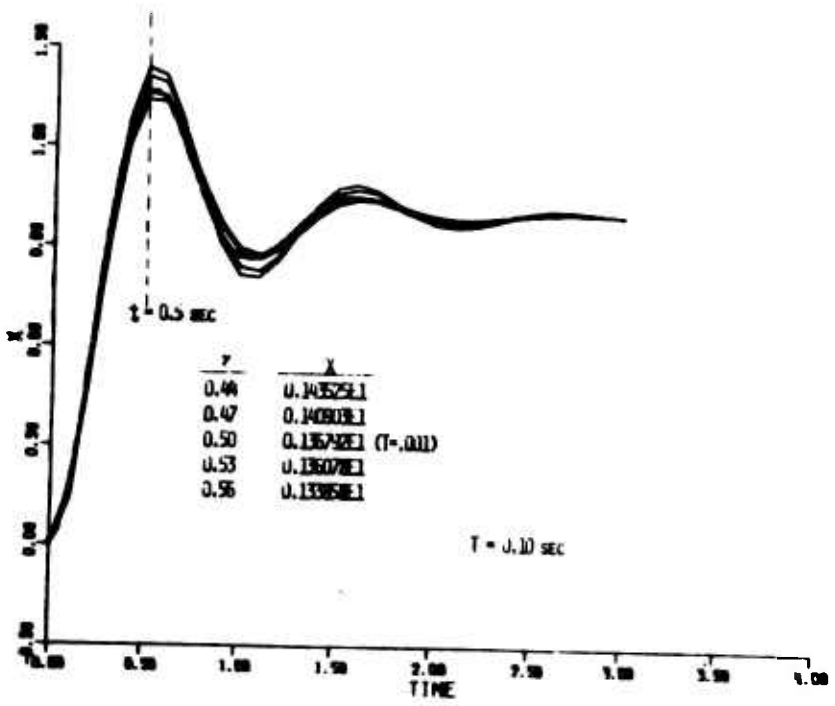


Figure 11. Step Response: Excessive Lead



T	X
0.44	0.1435254
0.47	0.1409018
0.50	0.1367928 (T=.011)
0.53	0.1360788
0.56	0.1338568

Figure 12. Step Response: Effect of Variable Lead/Lag Compensation

A CONTROL SYSTEM DESIGN METHOD FOR A SYSTEM HAVING A PLANT WITH VARIABLE PARAMETERS [10]

by

*C. H. HOUPIS

ABSTRACT

A design method is presented which permits a feedback control system with variable plant parameters to satisfy the desired performance specifications. The method utilizes state-variable feedback with constant feedback gains and satisfies the quadratic performance index (P.I.). The design application is an aircraft pitch attitude control system. The control system must be insensitive to parameter variations which occur over a range of flight conditions. The design example allowed large variations in the plant parameters while still maintaining the desired response. The gain in this design is high, but not in the sense of Horowitz's high-gain system designs [1]. Because of this relatively high gain, the disturbance response of this system is within acceptable limits. In comparison with previous control system designs [1], the design presented in this paper resulted in a simpler system configuration, i.e., reduced system complexity.

I. INTRODUCTION

A design approach [2,3,9] utilizing the quadratic P.I. is presented for designing a pitch attitude control system for an airplane that is insensitive to parameter variations over a given range of flight conditions. The

---

* Professor of Electrical Engineering, Department of Electrical Engineering, School of Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, U.S.A.

values of these parameters for six flight conditions (F.C.) (M_δ, control-surface effectiveness, and Mach number) are listed in Table 1. In order to design a feedback control system for such a plant, see Fig. 1, it is necessary to determine the parameter which has the most influence on the time response of the pitch rate, $\dot{\theta}$. When this parameter is identified and the desired response selected, the control system may be designed to satisfy the desired performance specifications. The system performance specifications are: the response to a step command must have less than 25% overshoot and must damp to at least 1/8 amplitude within one cycle; the response must rise to within 90% of the commanded value in three seconds or less; and the system must damp the response to a step disturbance input, w(t), to less than 1/4 amplitude in one cycle or if first order effects dominate, then 90% or more or the response to the disturbance should be eliminated within three seconds. Although not shown in Fig. 2, in which the state-variables are identified, w(t) is summed with $x_3(t)$ to yield the actual input into the block representing the airplane dynamics. Numerous nonoptimal designs have been made that achieve these specifications [1,4, and 5].

II. SYSTEM CONFIGURATION

The plant transfer function is given by (1), the plant state equation by (2), and the P.I. by (3).

$$G(s) = \frac{-6.67\, M_\delta K_1 \left(s + \frac{1}{T_a}\right)}{s(s + 6.67)(s^2 + 2\zeta\omega_a s + \omega_a^2)} \tag{1}$$

$$\dot{\underline{x}} = \underline{A}\,\underline{x} + \underline{B}\,u \tag{2}$$

$$\text{P.I.} = \int_0^\infty (\underline{x}^T \underline{h}\,\underline{h}^T \underline{x} + pu^2)\, dt \tag{3}$$

For this P.I., the optimal control is given by $u^o = -\underline{k}^T\underline{x}$ [6]. The configuration of the closed-loop system designed by minimizing the P.I. of (3) is shown in Fig. 2. The parameters a, b, c, and d are related to the aircraft parameters as follows:

$$a = \zeta\omega_a,\ b = \frac{1}{T_a},\ c = (\omega_a^2 - ab)/(a - b),\ \text{and}\ d = (\omega_a^2 - a^2)/M_\delta(a - b).$$

Note that the state x_2 is inaccessible.

III. DESIGN APPROACH [6,9]

Using the phase-variable approach, the design method is based upon selecting a defined dynamic equaion (DDE) for the desired optimal performance. If it were permissible to set $p = 0$ in (3), the P.I. would attain an absolute minimum of zero for $\underline{h}^T\underline{x}(t) = 0$. This minimum may be expressed as

$$h_1X_1(s) + h_2X_2(s) + \ldots + h_jX_j(s) + \ldots + h_nX_n(s) = (h_1 + h_2s + \ldots + h_ns^{n-1})X_1(s) = 0 \tag{4}$$

because, $sX_{j-1}(s) = X_j(s)$ for $j = 1,2,\ldots,n$. Note that the vector $\underline{h}$ is related to the standard matrix $\underline{Q}$ as follows:

$$\underline{Q} = \underline{h}\,\underline{h}^T$$

For a solution to the optimal control problem, then

$$(h_1 + h_2s + \ldots + h_ns^{n-1}) = 0 \tag{5}$$

may be taken as the DDE of the system. The limiting form of the closed-loop system as $p \to 0$ is the model described by (5). For other values of p there is a mismatch. For the system design presented in this paper $p = 1$. For the aircraft under consideration, where $n = 4$, it is desirable to place a root, α, of the DDE, coincident with the plant zero to reduce dependence of the system time response on the zero. If the remaining two roots of the DDE are a complex-conjugate pair, then the DDE is

$$(s + \alpha)(s^2 + 2\zeta_N\omega_Ns + \omega_N^2) = 0$$

or

$$\alpha + \left(\frac{\omega_N + 2\zeta_N\alpha}{\omega_N}\right)s + \left(\frac{\alpha + 2\zeta_N\omega_N}{\omega_N^2}\right)s^2 + \left(\frac{1}{\omega_N^2}\right)s^3 = 0 \tag{6}$$

The necessary constraint on h_1, for zero steady-state error with a step input, [7] is $h_1 = \alpha$. Thus, comparing (5) and (6), the weighting vector for a fourth-

order plant with one zero, becomes [3]

$$\underline{h}^T = \left[\alpha, (1 + \frac{2\zeta_N}{\omega_N}\alpha), (\frac{\alpha}{\omega_N^2} + \frac{2\zeta_N}{\omega_N}), \frac{1}{\omega_N^2} \right] \tag{7}$$

Since a fixed-gain feedback solution is specified, then the values of the feedback coefficients must be determined on the basis of a nominal F.C. For the airplane dynamics of Fig. 2, the variable parameter, M_δ, has the most influence on the system's response. The nominal F.C. chosen is F.C. #1 for which M_δ has its minimum value [2,3]. The control ratio for the system of Fig. 2 is

$$\frac{Y(s)}{R(s)} = \frac{6.67\, M_\delta^* K_1 (s + 1/T_a^*)}{s^4 + (6.972 + K_1 k_4^*)s^3 + (4.29 + K_1 k_4^* b_3)s^2 + (15.21 + K_1 k_4^* b_2)s + K_1 k_4^* b_1} \tag{8}$$

where

$$b_1 = 6.67 \left[\omega_a^2 \left(1 - \frac{k_3^*}{k_4^*}\right) + \left(\frac{M_\delta^*}{k_4^*}\right)\beta\left(\frac{1}{T_a^*}\right) \right] \tag{9}$$

$$b_2 = 6.67 \left[\left(1 - \frac{k_3^*}{k_4^*}\right)(2\zeta\omega_a)^* + \frac{M_\delta^*}{k_4^*}\left(\beta + A\,\frac{1}{T_a^*}\right) \right] \tag{10}$$

$$b_3 = 6.67 \left(1 + \frac{M_\delta^* A - k_3^*}{k_4^*}\right) + (2\zeta\omega_a)^* \tag{11}$$

and where the asterisks denote the nominal values for F.C. #1. The terms A and β are given by

$$A = (-k_2^*/M_\delta^*) \tag{12}$$

$$\beta = k_1^* - (a^*/M_\delta^*)\, k_2^* \tag{13}$$

An analysis of (8) reveals that zero steady-state error with a step input [2,9] for all flight conditions, occurs when the following conditions are satisfied:

$$k_3^* = k_4^* \quad \text{and} \quad \beta^* = 1 \tag{14}$$

The k's of (14) are the values determined by use of the physical variable representation of Fig. 2. Based upon the criterion established by Mirmak [3] the minimum value of K_1 required for an optimal solution is

$$(K_1)_{min} = \xi\omega'\omega_N^2/6.67M_\delta^* \tag{15}$$

where ω' represents the maximum corner frequency of $\Gamma(j\omega) = \underline{h}^T\underline{\Phi}(j\omega)\ \underline{B}$, for all F.C.'s, $\underline{\Phi}(s)\ \underline{A}\ (sI - \underline{A})^{-1}$, and ω_c represents the 0-db cross-over point of the Bode plot of $\Gamma(j\omega)$. ξ represents the value that yields the desired separation between ω' and $\omega_c\ (\geq \xi\omega')$. If $\underline{k}_p^*$ (the designation of the feedback coefficients for phase-variable representation) is to be determined by the Bode diagram, the solution of the algebraic Riccati equation, utilizing the Kalman equation

$$|1 + \underline{k}_p^T\ \underline{\Phi}(s)\underline{B}|^2 = 1 + |\underline{h}^T\ \underline{\Phi}(s)\ \underline{B}|^2/\rho \tag{16}$$

then the values of ξ and ω_c must be selected.

$\underline{k}_p$ is determined by the root-square-locus solution of the Riccati equation utilizing (16) [6,9]. Equation (16), utilizing the spectral factorization technique, can be rewritten as follows:

$$[1 + G(s)H_{eq}(s)]^+\ [1 + G(-s)H_{eq}(-s)]^- = [1 + \Gamma(s)\Gamma(-s)]^+\ [1 + \Gamma(s)\Gamma(-s)]^- \tag{17}$$

where $G(s)H_{eq}(s) = \underline{k}_p^T\ \underline{\Phi}(s)\ \underline{B}$, $\Gamma(s) = \underline{h}^T\underline{\Phi}(s)\underline{B}$. The (+) superscript indicates factors with roots in the left-half plane, and the (-) superscript denotes factors with roots in the right-half plane. For a given value of static loop sensitivity and $\underline{h}^T$ the factors on the right hand side of (17) are known. Thus, the following equation is used to solve for $\underline{k}_p^*$

$$[1 + G(s)H_{eq}(s)]^+ = [1 + \Gamma(s)\Gamma(-s)]^+ \tag{18}$$

A number of defined dynamic equations are synthesized that it is felt will yield the desired system performance for all F.C.'s. From numerical experiments, utilizing a digital computer program that solves for $\underline{k}_p^*$ from (18) and yields $y(t)\big|_{w(t)=0}$ for $r(t) = 1$, and $y(t)\big|_{r(t)=0}$ for $w(t) = 1$, for each synthesized DDE, a DDE (6) is determined that does yield the desired performance specifications for all F.C.'s. Since (16) is expressed in phase-variables, a linear transformation is used to determine $\underline{k}^*$ for the system of Fig. 2. The values for the final design are $K_1 = 800$, $\xi = 2.8$, $\omega_N = 7.5$, $\zeta_N = 0.7$, $\omega' = 7.5$, $\alpha^* = 0.0356$, and $M_\delta^* = 0.22$ which yield.

$$\underline{h}^T = [0.0356,\ 1.00665,\ 0.18730,\ 0.01778] \tag{19}$$

$$(\underline{k}^*)^T = [0.90526,\ -0.04868,\ 0.03102,\ 0.3123] \tag{20}$$

This $\underline{k}^*$ resulted in satisfactory system transient performance for all F.C.'s. Zero steady-state error was only achieved for F.C. #1.

The numerator polynomial of $1 + G(s)H_{eq}(s)$ is

$$s^4 + (6.972 + K_1k_4^*)s^3 + (4.29 + K_1k_4^*b_3)s^2 + (15.21 + K_1k_4^*b_2)s + K_1k_4^*b_1 \qquad (21)$$

where b_1, b_2, and b_3 are functions of the k_i^*'s. Inserting the values from (20) into (13) yields $\beta = 0.93867$. Thus, in order to obtain the normalized value of $\beta^* = 1$, in order to achieve zero steady-state error for all F.C.'s, the k_i^*'s of (20) are divided by this value of β. The values for b_1, b_2, and b_3 remain unchanged by this normalization of β. In order to achieve the condition of $k_3^* = k_4^*$, their normalized values of 0.0330466 and 0.0332704, respectively, are approximated by $k_3^* = 0.033$. Thus the normalized $\underline{k}_N^*$ is

$$\underline{k}_N^* = [0.9644, -0.05186, 0.033, 0.033] \qquad (22)$$

Since $(k_{4_N}^* - k_4^*)/k_4^*$ represents approximately a 6.5% change in k_4^*, the utilization of $k_{4_N}^*$ is not detrimental to achieving the desired performance specifications, i.e., the coefficients of (21) are not appreciably affected. The computer program incorporated the normalization of β to obtain all time responses, for r(t) and w(t), for all F.C.'s.

Table 2 lists the poles and zeros of (8) for all six F.C.'s. In all cases there are two real poles and a complex-conjugate pair. In every case one of the real poles essentially coincides with the plant zero. This was a desired result of the design: to minimize the effect of the plant zero on y(t). For all cases, except for $F.C._N$, the remaining real pole is the dominant pole whereas for $F.C._N$ the complex-conjugate pair are the dominant poles. Note that the values of all real dominant poles and the real part of the dominant poles for F.C. #1 do not vary significantly (-4.23 to -5.125). This is a desired result of the design, i.e., the system meets the desired performance specifications under fixed-gain feedback for all flight conditions.

IV. SUMMARY

The implementation of the final design values is shown in Fig. 3 where (a) states x_3 and x_4 are fed back through a common feedback unit to ensure the condition of $k_{3_N}^* = k_{4_N}^*$, and (b) the feedback signal $(x_2k_{2_N}^*)$ is obtained from the output state, x_1. This requires a proportional plus rate feedback

unit. The system of Fig. 3 yields a sub-optimal performance for all F.C.'s. In comparison to the system designed by Ray [2] a significant gain reduction is achieved, a large decrease in response time, and a zero steady-state error is achieved for all F.C.'s. The settling time, t_s [$\pm$ 2% of the steady-state value of y(t)] for F.C. #1 is in excess of 15 seconds; for all other F.C.'s t_s < 15 sec. The design technique presented in this paper illustrates a method of obtaining a fixed-gain feedback system that satisfies the desired performance specifications over the entire range of plant parameter variations.

V. REFERENCES

1. LaBounty, R. H. and C. H. Houpis, "Root Locus Analysis of a High-gain Linear System with Variable Coefficients; Application of Horowitz's Method. IEEE Transactions on Automatic Control, AC-11, 255-263 (April 1966).

2. Ray, R. A. A State-variable Design Approach for a High-performance Aerospace Vehicle Pitch Orientation System with Variable Coefficients, Thesis, Wright-Patterson AFB, Ohio: Air Force Institute of Technology, March 1973.

3. Mirmak, E. V. Some Techniques for Optimal Linear Regulator Design to Satisfy Conventional Figures of Merit, Thesis, Wright-Patterson AFB, Ohio: Air Force Institute of Technology, December 1972.

4. Beale, R. E. and F. J. Hellings, Comparisons of High-Gain-Linear and Self-adaptive Flight Control Systems for a Typical Re-entry Vehicle, Thesis, Wright-Patterson AFB, Ohio: Air Force Institute of Technology, May 1963.

5. Parry, I. S. and C. H. Houpis, "A Parameter Identification Self-adaptive Control System," IEEE Transactions on Automatic Control, AC-15, 462-468 (August 1970).

6. Schultz, D. G. and J. L. Melsa, State Functions and Linear Control Systems, New York: McGraw-Hill Book Co., 1967.

7. Houpis, C. H. The Relationship Between Conventional Control Theory Figures of Merit and Performance Indices in Optimal Control Theory, Doctoral Dissertation, University of Wyoming, 1971.

8. D'Azzo, J. J. and C. H. Houpis, Feedback Control System Analysis and Synthesis (Second Edition), New York: McGraw-Hill Book Co., 1966.

9. D'Azzo, J. J. and C. H. Houpis, Linear Control System Analysis: Conventional and Modern, New York: McGraw-Hill Book Co., Tentative Publishing Date, April 1975.

TABLE I

System Parameters

F.C.	M_δ	$1/T_a$	$2\zeta\omega_a$	ω_a^2	Mach No
1	0.22	0.0356	0.302	2.28	0.2
2	16.29	1.163	2.226	6.51	0.6
3	52.95	2.070	4.980	56.10	1.2
4	20.86	0.325	0.652	18.71	6.0
5	2.24	0.0366	0.0792	3.68	4.0
6	0.70	0.00794	0.0165	0.65	6.0

TABLE II

Poles and Zeros of (8) for the Final Design

F.C.	Dominant Pole(s)	Non-Dominant Poles	Zero, $s = -1/T_a$
1 (Nominal)	-5.125 ± j5.0228	-23.0865, -0.03515	-0.0356
2	-4.2550	-14.939 ± j142.14, -1.16306	-1.163
3	-4.2386	-15.871 ± j257.76, -2.069	-2.070
4	-4.2574	-14.57 ± j161.127, -0.32464	-0.325
5	-4.4382	-14.337 ± j49.93, -0.03652	-0.0366
6	-5.0468	-14.016 ± j23.335, -0.007931	-0.00794

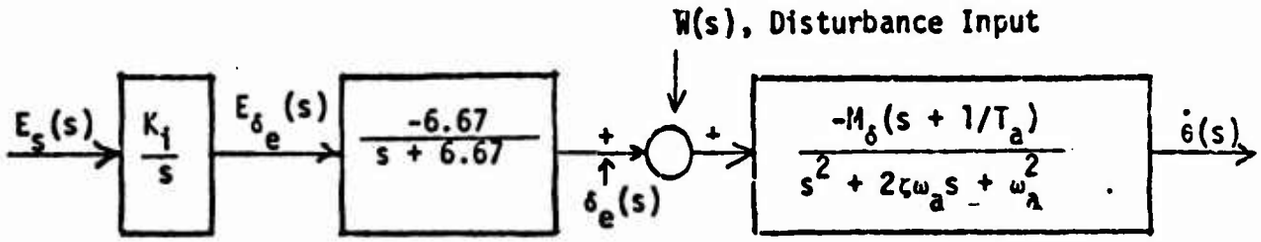


Fig. 1. Block Diagram of Basic Plant

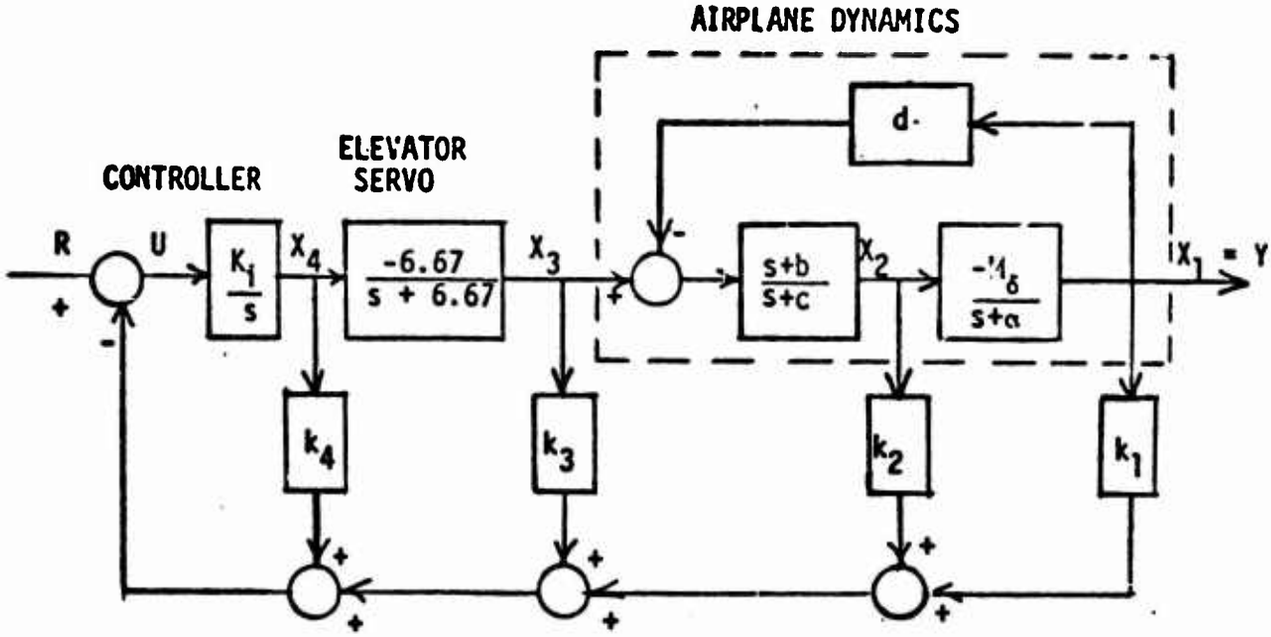


FIGURE 2

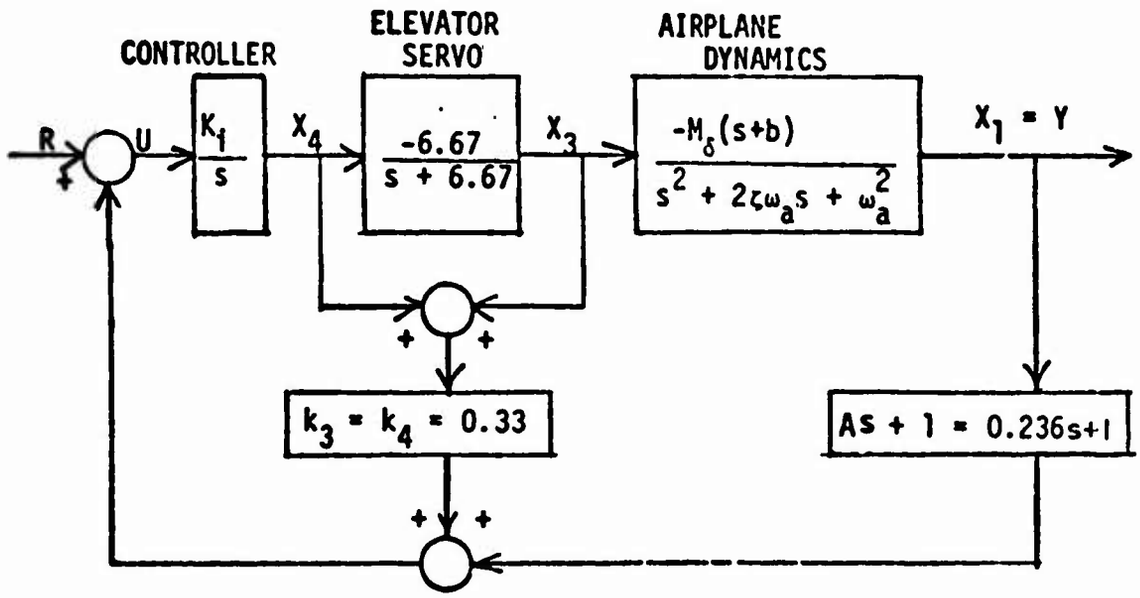


FIGURE 3

MULTI-SURFACE FLIGHT CONTROL SYSTEMS DESIGNED USING QUADRATIC OPTIMAL CONTROL

Robert D. Poyneer
AF Flight Dynamics Laboratory
Wright-Patterson Air Force Base, Ohio

ABSTRACT

Multiple-input, multiple-output flight control systems were designed using quadratic optimal control programs. The flight control systems incorporated Control Configured Vehicle (CCV) concepts of: (1) Ride Control, (2) Maneuver Load Control, and (3) Load Alleviation. The flight control systems also impacted: (1) Empennage acceleration, (2) Short-period damping, (3) Dutch Roll damping, (4) flexible mode damping, and (5) flying qualities. The first of four programs converts equations from generalized coordinate format into state equation format and deletes selected bending modes. The second program optimally designs the control laws for a constant linear system using states as inputs. The third program optimally designs the control laws using measurements as inputs. The fourth program generates root loci information to determine stability margins. The quadratic optimal control programs provided adequate designs for both the longitudinal and lateral-directional axes flight control system for the CCV B-52.

INTRODUCTION

In February 1972, it was decided that the CCV concepts previously developed for implementation as separate subsystems should be combined to provide optimum performance through a single system controlling multiple surfaces. The goal of the program was to develop the more advanced computational techniques required to perform quantitative analyses of multi-function CCV systems. Work started on 1 July 1972 and the contractor finished 1 October 1973.

The techniques leading up to this program were reported on in References 1 and 2. Reference 1 described a practical controller design procedure for aircraft over their entire flight envelope using quadratic optimal control technology. The resulting design was an optimal gain schedule with some gains variable. Reference 2 showed how the techniques described in Reference 1 were applied to the C-5A aircraft. The applicability of the CCV concepts of Relaxed Static Stability and Maneuver Load Control were determined.

The B-52 CCV Program validated that five advanced flight control concepts were operationally possible, and that they significantly improved the airplanes characteristics. The CCV Program for the B-52

is described in Reference 3. The new control concepts that were flight validated were augmented stability, flutter mode control, maneuver load control, ride control, and fatigue reduction. All of these concepts were designed separately, and then tested for compatibility.

The augmented stability system allows the center of gravity of the airplane to be moved aft (to the neutral point and beyond) while pitch stability is maintained by the use of active controls. The size of the horizontal tail can be reduced because this surface no longer needs to supply a large down load pitching moment for static stability. On the CCV B-52, (the test bed) the c.g. was moved aft using the fuel distribution system to get to the neutral stability point. Adequate handling qualities were maintained, even though the center of gravity was shifted from 29.% Mean Aerodynamic Chord (MAC) to 42% MAC.

Flutter Mode Control uses control surfaces on the wing to add damping or stiffness to selected wing structural bending modes. On a new airplane the structure of the wing could be made less rigid (resulting in less weight) and still maintain the same flutter placard speed. On the CCV B-52, the flutter placard speed was increased using control surfaces to add damping to the existing wing. The B-52 was successfully flight tested to 10 knots above the projected bare airframe flutter speed.

Maneuver Load Control uses control surfaces outboard on the wing to reduce lift and control surfaces inboard on the wing to garner more lift, so as to move the wing center of lift inboard. With the center of lift more inboard, the moment arm for stress on the wing root is decreased. With a decreased wing root stress, the size and weight of the wing root structure could be reduced. On the CCV B-52, it was shown that the wing root bending moment was reduced more than 30% compared to bare airframe value.

Ride Control System reduces the acceleration at the pilot station due to turbulence through the use of active canards. The canards help to increase the stiffness of the forward fuselage. On a new airplane the structural stiffness of the forward fuselage could be reduced (saving weight) with adequate crew comfort maintained by the active canards. On the CCV B-52, the canards reduced the acceleration at the pilot station by 30% in both the vertical and lateral axes.

Fatigue Reduction uses active control surfaces to reduce the stress levels along the fuselage and wing to increase the working structural life of the airplane. On the CCV B-52, the damage rates along the fuselage and wing in the longitudinal axis were reduced.

The MSS (Multi-Surface System) was designed to perform as well as the separately designed CCV concepts, at a single design condition. The concepts of Flutter Mode Control and Augmented Stability were not included in the design study because they require multiple design points to verify performance.

Results of the analysis of the lateral-directional axis MSS was reported on in Reference 4. This paper will give the results of the lateral-directional and longitudinal axes of the MSS, as well as more detail of the optimal control programs.

DIGITAL COMPUTER PROGRAMS

In designing the Multi-Surface Systems four digital computer programs were used (See Figure 1). Three programs (MODEL, DIAK, and FFOC) helped generate the quadratic optimal control solutions. The fourth program (VALUE) provided root locus information for gain and phase stability margins.

The B-52 airplane equations of motion are normally in second-order differential equation format (see Reference 5).

$$[As^2 + B\,s + C]\,q = 0$$

where A = second order coefficients
B = first " "
C = zero " "
s = LaPlace operator
q = generalized coordinate vector

MODEL

In the longitudinal axis, MODEL program reduced the number of flexible mode from 27 to 6. Even though 21 modes were dropped, their steady-state effects are included as residuals. In the lateral-directional axis, the MODEL program reduced flexible mode displacements from 27 to 5.

The MODEL Program converts the second order differential equations into state equation format. The program generates the F, G_1, G_2, M, H, and D matrices which are the matrices of the state equations (see Figure 2) which is the format required by DIAK and FFOC.

$$\dot{x} = Fx + G_1 u + G_2 \eta$$

$$y = Mx$$

$$r = Hx + Du$$

where:
x = generalized state vector
u = control input vector
η = disturbance input vector
y = measurement vector
r = response vector

A description of the MODEL, DIAK and FFOC programs is presented in Reference 6.

DIAK

The DIAK program generates optimal solutions for the state equations and the quadratic cost function. The feedback gains, which are defined by the program, relate the inputs (surface deflections) to the outputs (states). The optimal feedback gains are generated using the Ricatti equation.

The quadratic cost function is

$$J = \int_0^\infty r^T Q\, r\, dt$$

where Q is the quadratic weighting matrix. The optimal control laws are

$$u = Kx$$

where K is a matrix of gains determined by solving an algebraic Ricatti equation, on an iterative basis. The algebraic Ricatti equation is of the form

$$0 = \bar{A}^T P + P \bar{A} + \bar{Q} - P \bar{E} P$$

where

$$\bar{A} = F - G_1 (D^T QD)^{-1} D^T QH$$

$$\bar{Q} = H^T QH - H^T QD (D^T QD)^{-1} D^T QH$$

$$\bar{E} = G_1 D^T QD)^{-1} G_1^T$$

and

$$K = -(D^T QD)^{-1} (D^T QH + G_1^T P)$$

Other outputs from the DIAK program are: (1) covariance matrices, (2) cross-correlation matrices, (3) row-correlation matrices, (4) row-sum correlation matrices, (5) rms controls, (6) rms measurements, (7) rms responses, and (8) eigenvalues. DIAK also computes and plots time responses for states or measurements for initial conditions or step inputs.

FFOC

The FFOC Program generates quadratic optimal control solutions where the inputs to the feedback gains are measurements. The control law is

$$u = K^* Mx$$

where $K^* = KM^{-1}$ = fixed form gains matrix

M^{-1} = inverse of Measurement Matrix

The weighting on the states determined in the DIAK program are used as the starting weights in FFOC. The measurements are generally accelerometers and rate gyros. Any filtering thought necessary has to be included as part of the sensor, since all the program can do is generate optimal gains.

The program can start any of three ways: (1) full state gains (2) zero gains or (3) a set of stable measurement gains. If a set of measurement gains are used as the starting position, these gains were usually the results of a previous FFOC computer run.

A general procedure is to generate a set of measurement gains using all sensors (measurements) and surfaces. For this analysis the full measurement solution had 50 gains (10 sensors and 5 surfaces). Since 50 feedback gains is not very manageable, the number of gains was reduced, using engineering judgment, until significant degradation in performance occurred. The restricted feedback gain solutions were much more useful.

VALUE

A frequency-domain computer program, VALUE, was used to generate root locus printouts which could be used to determine phase and gain margins. A description of the VALUE program may be found in Reference 7.

IMPLEMENTATION

The equations of motion of the B-52 were obtained from The Boeing Company. The equations were in second order differential equation format. Added to the airplane equations were the Dryden wind model, the response equations and the sensor equations. The MODEL program will do two basic things; first, it will present the above equations in the state equation format; and second, it will reduce the number of flexible modes in the simulation. The "residual" effects of the bending modes that are eliminated are included in the flexible modes that are retained.

The outputs of the MODEL program are the following matrices: F (system), G_1 (control input), G_2 (noise input), H (state response), D (control response), and M (measurement).

The output matrices of the MODEL program form the input matrices of the DIAK program.

Added to these matrices is the Q (weighting) matrix. To suppress a response, the weight Q_{ii} is increased. The original selection of weights is based on the experience of the engineer. If the state is a response it can be undamped by weighting the state rate or model-following responses. Weights on the controls restrict the amount of displacement or rates generated by the surface actuators. To emphasize certain frequencies that a sensor picks up, the sensor may be modified by adding a bandpass filter to it. The use of off-diagonal elements to the Q matrix can help eliminate crosstalk between surfaces.

Model-following is used to place airplane roots to a desirable location. The model may be placed in the control law or the performance index (rate model-following). The optimal control program DIAK picks

gains which will minimize the difference between the controlled system ($F + G_1 K$) and the model (F_m).

The output of the DIAK program are 185 gains (37 states and 5 surfaces). The weights determined here can be used for the initial measurement solution using the FFOC program. Normally all the states are not observable, and for practical systems only observables can be used as inputs. The FFOC program generates an optimal solution using observables (measurements) as inputs to the flight control system. The function of the DIAK program in this study is to provide a matrix of weights (Q) for the FFOC program.

To achieve a solution from the FFOC program, alternate paths may be chosen. The initial gains (between measurements and surfaces) may be zero or the full state gains from the DIAK program may be used, where the measurement gains are calculated according to

$$K^* (\lambda) = K^1 (1) + \lambda K^2$$

as λ varies from 1 to 0. The weights and filters (on sensors) can be varied to obtain the best full measurement solution. The full measurement solution has 50 gains (10 measurements and 5 surfaces.)

The next step is to simplify the full-measurement solution to a restricted feedback solution. In this process, the number of gains, weights, and filtering may be varied to obtain the best solution for the fewest number of gains. It is a trial-and-error process to reduce the number of gains and still maintain adequate performance.

The FFOC program will provide the rms responses for wind turbulence and surface rms disturbances. To determine the performance of the MSS for step inputs of the surfaces, or wind, the DIAK program is used to plot up time responses. To ascertain gain and phase stability margins, VALUE program was used to give root locus printouts. The best solution is a combination of outputs from: (1) FFOC program restricted feedback rms responses, (2) DIAK time responses, and (3) VALUE gain and phase stability margins.

MULTI-SURFACE SYSTEM

The optimal control design study was for a B-52 modified as shown in Figure 3. The airplane flight condition is: (1) altitude of 2000 feet, (2) Mach of 0.517, (3) and a gross weight of 270,000 lbs.

The study is divided into two parts, a longitudinal-axis MSS and a lateral-directional-axis MSS. The longitudinal axis MSS will use five sets of surfaces: (1) elevator, (2) inboard aileron, (3) outboard aileron, (4) flaperon, and (5) horizontal canards. The lateral-directional axis MSS will also use five sets of surfaces: (1) rudder, (2) flaperon, (3) inboard aileron, (4) outboard aileron, and (5) vertical canard.

LATERAL-DIRECTIONAL AXIS

The MSS for the lateral-directional axis has the following goals: reduce pilot station lateral acceleration by 30%, not degrade the lateral acceleration at the c.g. (BS860) by more than 5%, reduce the tail acceleration, in the lateral axis, by 30%, provide adequate Dutch Roll damping and frequency, reduce fuselage stresses by 10% for wind turbulence input, keep the flexible mode (1st, 2nd, 6th, 9th, and 10th) frequencies within 10% of their open loop values, not reduce the flexible mode damping by more than 10% and maintain the surface deflections and rates less than saturation. The goals of stress reduction and flexible mode damping and frequency, were added the goals from the CCV Program.

The responses that are weighted in the cost functional (J) are: (1) nine lateral stresses at fuselage, wing, and vertical tail, (2) three side accelerations along the fuselage, (3) five washed-out sensors, (4) two wing vertical accelerations (5) five surface deflection rates, (6) eight model-following errors, and (7) five surface deflections.

The full-state solution generated by the DIAK program had 185 gains (37 states and 5 surfaces). The performance of the full-state solution is shown in Table I. Of the 31 goals set forth, only one was not met, the frequency of the sixth flexible mode was 23.1% higher than the open loop value (goal was not more than 10% change).

The full-measurement solution generated by the FFOC program had 50 gains (10 measurements and 5 surfaces). The performance of the full-measurement solution is shown in Table I. Of the 31 goals, only four were not met; 2nd, 6th and 10th mode damping were less than 90% of the open loop values, and 2nd mode frequency was more than 10% different than open loop.

The restricted feedback (MSS) solution generated by the FFOC program had 10 gains (5 measurements and 4 surfaces). The performance of the restricted feedback solution is shown in Table I. Of the 37 goals, only two were not met. The damping of the 6th flexible mode was 16.1 % less than the open loop value (goal was not more than 10% loss in damping). The phase margin of the vertical canard loop was $\pm$45 deg. instead of $\pm$60 deg. Neither of these discrepancies is critical, so it can be said that the lateral-directional axis MSS is adequate.

A functional block diagram of the lateral-directional axis MSS is shown in Figure 4.

LONGITUDINAL AXIS

The MSS for the longitudinal axis of the B-52 had the following goals: reduce pilot station vertical acceleration by 30%, not degrade more than 5% of the vertical acceleration at BS860, reduce tail (BS1655) acceleration in the vertical direction by 30%, provide adequate short period damping and frequency, reduce the wing root vertical stress by 30% for an elevator input, reduce fuselage and wing stresses by 10% for wind turbulence input, keep the

TABLE I PERFORMANCE OF LATERAL-DIRECTIONAL AXIS MSS

PARAMETER	DISTURBANCE	UNITS	FULL STATE	FULL MEASU.	MSS	GOAL
NY_{172} Reduction	RMS Wind	%	64.2	34.8	43.1	>30.
NY_{860} "	"	%	58.8	40.0	51.0	> -5.
NY_{1655} "	"	%	76.8	42.8	38.3	> 30.
Stress BS475 Reduction	"	%	53.6	11.0	36.5	> 10.
" BS760 "	"	%	71.5	43.6	47.7	> 10.
" BS1412 "	"	%	78.5	47.8	51.0	> 10.
" WBL222 "		%	58.3	50.1	51.2	> 10.
" WBL820 "	"	%	60.4	33.2	64.5	> 10.
" WBL974 "	"	%	74.0	52.6	66.4	> 10.
Rudder Displ.	RMS Wind/Rud/Ail*	Deg	1.8	1.4	3.0	< 19.
" Rate	"	Deg/Sec	8.4	6.1	23.8	< 80.
Flaperon Displ.	"	Deg	0.7	0.2	0.8	< 20.
" Rate	"	Deg/Sec	3.7	1.8	2.0	< 80.
Inboard Aileron Displ	"	Deg	0.8	1.6	2.1	< 17.
" Rate	"	Deg/Sec	5.4	14.6	30.5	< 80.
Outboard Aileron Displ	"	Deg.	1.0	1.4	-	< 20.
" Rate	"	Deg/Sec	5.0	11.1	-	< 80.
Vertical Canard Displ	"	Deg.	0.9	3.1	2.8	< 10.
" Rate	"	Deg/Sec	5.3	12.3	29.7	< 80.
Dutch Roll Damping	RMS Wind & Rudder and Elevator		.527	.884	.71	> .19
Dutch Roll (W_n)(ζ)	"	Rad/Sec	.67	1.15	.87	> .35
Mode 1 Damping	"		.183	.181	.083	> .071
" Frequency	"	Rad/Sec	9.64	8.92	8.8	8.7-10.7
Mode 2 Damping	"		.362	.099	.159	> .138
" Frequency	"	Rad/Sec	12.54	9.52	11.6	10.7-13.2
Mode 6 Damping	"		.496	.156	.193	> .207
" Frequency	"	Rad/Sec	22.24	17.46	17.7	16.2-19.9
Mode 9 Damping	"		.101	.074	.054	> .030
" Frequency	"	Rad/Sec	27.73	22.18	21.1	20.6-25.7
Mode 10 Damping	"		.108	.020	.065	> .031
" Frequency	"	Rad/Sec	29.88	30.31	28.8	27.2-33.4
Rudder Gain Margin	-	$\pm$db	-	-	11.	> 6.
" Phase "	-	$\pm$Deg	-	-	80.	> 60.
Flaperon Gain "	-	$\pm$db	-	-	7.	> 6.
" Phase "	-	$\pm$Deg	-	-	180.	> 60.
Inboard Aileron Gain	-	$\pm$db	-	-	23.	> 6.
" " Phase	-	$\pm$Deg	-	-	60.	> 60.
Vertical Canard Gain	-	$\pm$db	-	-	43.	> 6.
" " Phase	-	$\pm$Deg	-	-	45.	> 60.

* Worst case of 3 disturbances: rms wind, rms rudder, or rms aileron.

flexible mode (1st, 2nd, 5th, 7th, 8th, and 12th) frequencies within 10% of their open loop values, provide good pitch rate performance for an elevator step input, maintain surface deflections and rates below saturation, and not reduce the flexible mode damping by more than 10%.

The responses that are weighted in the cost functional (J) are: eight longitudinal stresses at fuselage, wing, and horizontal stabilizer locations, three vertical accelerations along the fuselage, longitudinal stress rates at eight locations, five surface rates, eight model-following errors, and five surface deflections. For the given weights, inputs, and outputs, the quadratic optimal control program gives the best set of feedback gains.

The restricted-feedback solution generated by the FFOC program had 19 gains (7 measurements and 4 surfaces). The performance of the restricted feedback solution is shown in Table II. Of the 41 goals, only 5 were not met. Vertical stress at BS1412 increased 28.5% instead of decreasing by at least 10%. Eighth mode structural damping decreased to 0.020, rather than staying above the goal of 0.043. The twelfth mode frequency was 31.5 rad/sec whereas the goal was between 34.2 and 41.9 rad/sec. Flaperon phase margins is $\pm$39.3 deg. instead of $\pm$60 deg. Horizontal canard phase margin was $\pm$31.9 degs. instead of $\pm$60 degrees. The most important of the five discrepancies was the reduction in damping of the eighth structural mode. The only way to increase the eighth mode damping was to either increase empanage acceleration or dramatically slow down the pitch rate response for an elevator step.

On the whole the MSS does a very fine job (meeting 36 out of 41 goals) but it can't do everything. The performance of the longitudinal axis MSS is adequate.

A functional block diagram of the longitudinal axis MSS (restricted feedback) is shown in Figure 5.

LIMITATIONS

Expensive

For a computer problem that has 37 responses in its quadratic cost functional, the computer time on a CDC 6600 required for a DIAK run is about 150 seconds. The computer time required for a FFOC run is about 900 seconds. The FFOC program runs until the ratio of costs J_{i+1}/J_i is between 0.99 and 1.00, and if the solution is not converging very fast, the run time is large. The FFOC program is used many more times than the DIAK program.

TABLE II PERFORMANCE OF LONGITUDINAL AXIS MSS

PARAMETER	DISTURBANCE	UNITS	MSS	GOAL
Nz_{172} reduction	RMS Wind	%	35.6	> 30
Nz_{860} "	"	%	25.6	> -5.
Nz_{1655} "	"	%	32.5	> 30.
Stress BS 475 Reduction	"	%	15.6	> 10.
" BS 760 "	"	%	16.4	> 10.
" BS 1412 "	"	%	-28.5	> 10.
" WBL 222 "	"	%	18.8	> 10.
" WBL 974 "	"	%	20.1	> 10.
" HS 56 "	"	%	36.2	> 10.
" WBL 222 "	RMS Elev.	%	29.7	> 30.
Pitch Rate (90% of Steady State)	Step Elev.	Sec	.35	> .75
Elevator Displ	RMS Wind/Elev*	Deg	5.0	< 19.
" Rate	"	Deg/Sec	42.9	< 80.
Flaperon Displ	"	Deg	3.6	< 20.
" Rate	"	Deg/Sec	19.7	< 80.
Outboard Aileron Displ	"	Deg	1.1	< 20.
" " Rate	"	Deg/Sec	24.6	< 80.
Horizontal Canard Displ	"	Deg	1.3	< 10.
" " Rate	"	Deg/Sec	25.3	< 80.
Short Period Damping	RMS Wind & Elev		.789	> .583
" " Frequency	"	Rad/Sec	1.71	1.6-10.0
Mode 1 Damping	"		.164	> .129
" Frequency	"	Rad/Sec	6.1	5.5-6.8
Mode 2 Damping	"		.189	> .106
" Frequency	"	Rad/Sec	12.0	11.3-13.9
Mode 5 Damping	"		.064	> .028
" Frequency	"	Rad/Sec	14.4	13.5-16.6
Mode 7 Damping	"		.081	> .012
" Frequency	"	Rad/Sec	18.4	17.6-21.5
Mode 8 Damping	"		.020	> .043
" Frequency	"	Rad/Sec	19.8	17.3-21.2
Mode 12 Damping	"		.095	> .027
" Frequency	"	Rad/Sec	31.5	32.2-41.9
Elevator Gain Margin		±db	15.	> 6.
" Phase "		±Deg	69.6	> 60.
Flaperon Gain "		±db	10.	> 6.
" Phase "		±Deg	39.3	> 60.
Outboard Aileron Gain Margin		±db	10.	> 6.
" " Phase "		±Deg	180.	> 60.
Horizontal Canard Gain Margin		±db	20.	> 6.
" " Phase "		±Deg	31.9	> 60.

* Worst case of 2 disturbances: rms wind, or rms elevator.

Coupled Equations

In order to increase the damping of the 8th structural mode in the longitudinal axis, the weighting of the 8th mode was increased from 0.4×10^{-6}. to 0.5×10^{-3}. The result was that the damping increased on the 2nd, 5th, 7th, 12th and 8th modes, and the acceleration reduction at BS1655 decreased below the goal of 30%. Since the equations are coupled, a certain amount of trial and error engineering is necessary to produce the desired results.

Lack of Versatility

The programs output the best set of feedback gains for the given set of inputs (sensors, surfaces and weights). Washouts or lag filters can be included as part of either the input or the output. For a given set of weights the solution may come out unstable. The engineer's experience must be used to change weights (or inputs, or outputs) to get a stable solution. The program does not know how or what to modify to get a stable solution.

Stability Margins

The FFOC program has no provision to provide a predetermined gain or phase margin. Reference 8 suggests a way to guarantee the degree of stability by using a performance index such as:

$$J = \int_0^\infty e^{2\alpha t} \, (u^T R u + x^T Q x) \, dt$$

Reference 9 shows that modal control theory can be used to position the characteristic roots which produce the desired damping.

SUMMARY

This paper described at set of quadratic optimal control programs (MODEL ,DIAK, and FFOC) that were used to design a Multi-Surface Systems (MSS) for the B-52 Bomber. The MSS is a flight control system in the longitudinal or lateral-directional axes which can use up to ten sensors and five flight control surfaces to enhance the aircraft's performance.

The quadratic optimal control programs were developed by Honeywell Inc., for use by the Air Force in designing multi-input multi-output automatic feedback control systems. The MODEL program converts the second order differential equations of motion into state equation format. It also reduces the number of flexible modes and residualizes them. The DIAK program generates an optimal solution for the flight control system using the states of the system as inputs (sensors). The FFOC program generates an optimal solution for the flight control system using the measurements as inputs.

The lateral-directional axis MSS was designed to reduce acceleration and stresses, and to maintain adequate handling qualities. The longitudinal axis MSS was designed to reduce acceleration and stresses, reduce wing root bending moments for maneuvers, load alleviation, and maintain adequate handling qualities.

The MSS will be reported on by a Technical Memo (AFFDL-TM-74-138 FGB). The MSS will not be flight tested since the test bed B-52 has been mothballed.

The main limitations of the quadratic optimal control programs are that the program use up a lot of computer time; the network filtering design has to be done on a trial-and-error basis; the programs can only optimize for the given set of inputs; the stability margin cannot be guaranteed; the engineer is still a vital part of the design loop; and the equations are still coupled.

For very simple systems, the use of these optimal control programs would mean extra work and complexity that is not needed. For extremely complicated or large systems, the core memory required by the programs and computation time needed, limit the program's use.

RECOMMENDATIONS

For complex systems, an optimal control program should be used instead of classical techniques. If the equations are predominately uncoupled then the DIAK and FFOC computer programs should be used.

The FFOC digital computer program should be "optimized" for the particular application. Eliminate general purpose options that will not be used. The core memory requirement should be reduced. Some of the optional features should be eliminated when it is not expected that they will be used.

Some way of assuring a stability margin should be included in the program.

REFERENCES

1. Van Dierendonck, A.J., "Design Method for Fully Augmented Systems for Variable Flight Conditions," Air Force Flight Dynamics Laboratory Technical Report AFFDL-TR-71-152, Wright-Patterson Air Force Base, Ohio, January 1972.

2. Stone, C.R., Ward, M.D., Harvey, C.A., Ebsen, M.E., McBride, E.E., Holenbeck, W.W., "Studies of the Compatibility of Relaxed Static Stability and Maneuver Load Control to C-5A-Type Aircraft," Vols, I and II, Air Force Flight Dynamics Laboratory Technical Report AFFDL-TR-72-38, Wright-Patterson Air Force Base, Ohio June 1972.

3. Kass, G.J., Johannes, R.P.,"B-52 Control Configured Vehicles Program," AIAA 4th Aircraft Design, Flight Test, and Operations Meeting, AIAA Paper 72-747, Los Angeles, CA, August 7-9, 1972.

4. Poyneer, R.D., "Multi-Surface System for the CCV B-52", AIAA 12th Aerospace Sciences Meeting, AIAA Paper 74-126, Washington, D.C. January 31, 1974.

5. Stockdale, C.R.,Poyneer, R.D., "CCV B-52 Data Base Report", Air Force Flight Dynamics Laboratory Technical Memo, AFFDL-TM-72-1-FG-2, Wright-Patterson Air Force Base, Ohio, May 1972.

6. Van Dierendonck, A.J.,Hartman, G.L., "Quadratic Methodology/Vol I., A Short Course on the Application of Optimal Control Theory to the Design of Practical Control Systems/Vol II, Documentation of Computer Programs" Honeywell Report No. F0161-FR, Minneapolis, Minnesota, October 1973.

7. Nelson, G.D., "Digital Handbook", Section 4.2 Honeywell Aerospace Division, Minneapolis, Minnesota, November 1967.

8. Anderson, B.D.,Moore, J.B., "Linear Optimal Control" Chapter 4, Prentice Hall Inc., Englewood Cliffs, N.J., 1971.

9. Porter, B., "Modal Control Theory-Modal Analysis & Modal Synthesis" Belgian Institute for Automatic Control, Brussels, Belgium, February 1974.

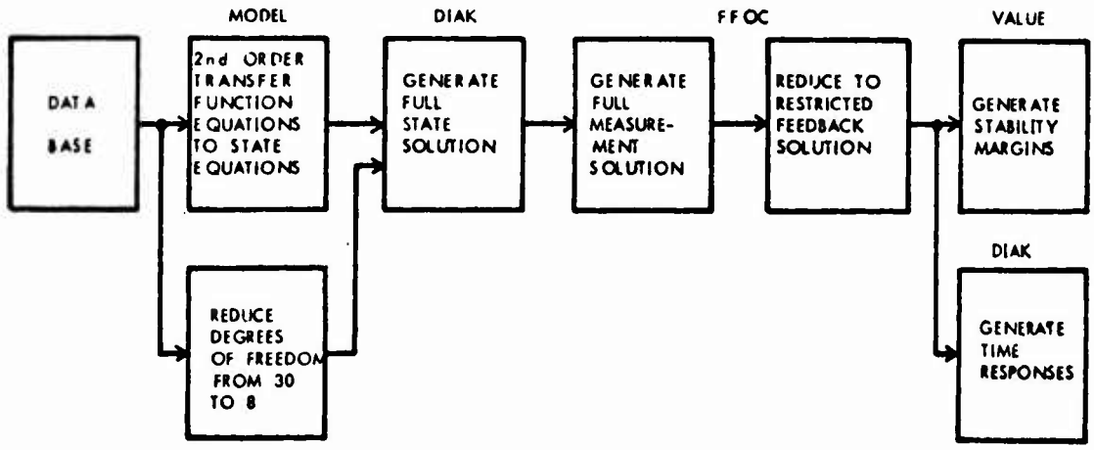


Figure 1 Design Procedure

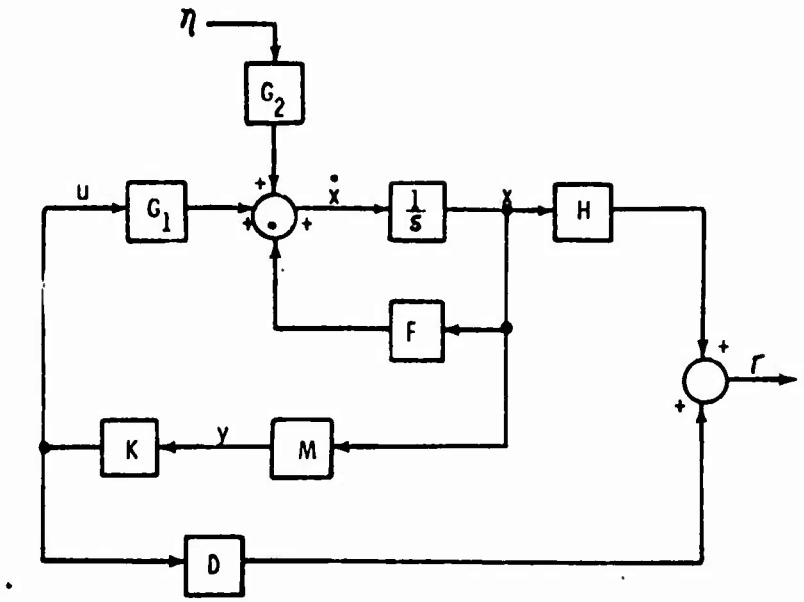


Figure 2 Optimal Control System

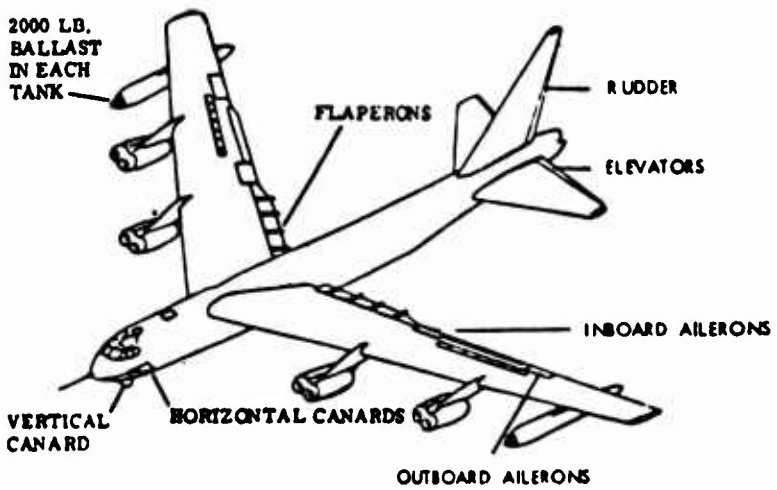


Figure 3 B-52 MSS Vehicle Configuration

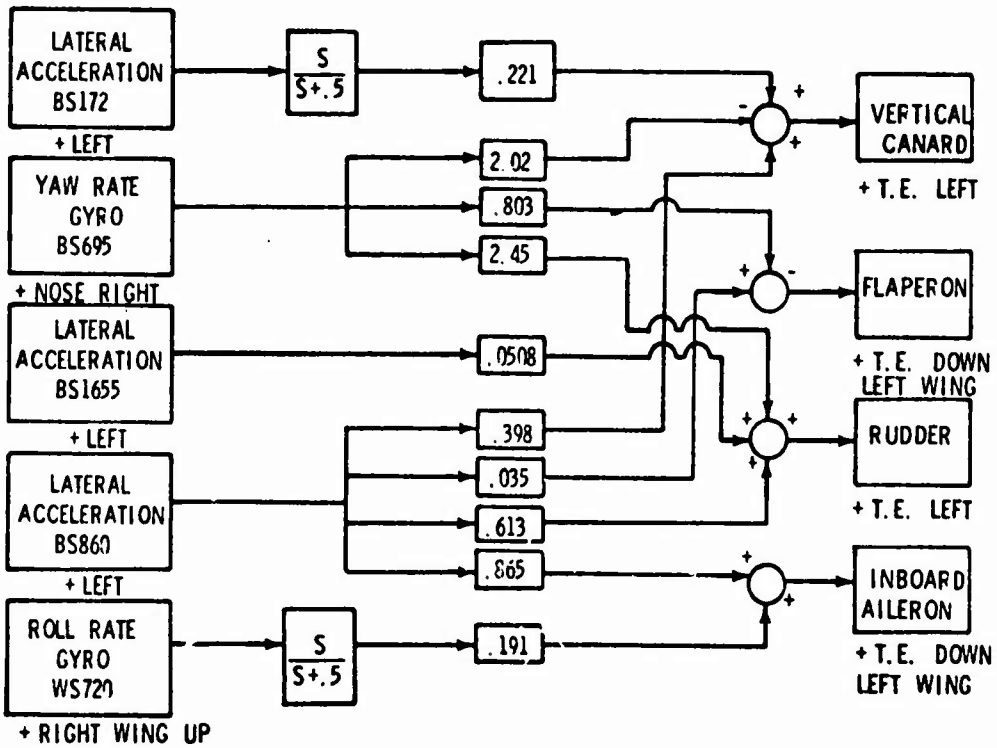


Figure 4 Lateral MSS

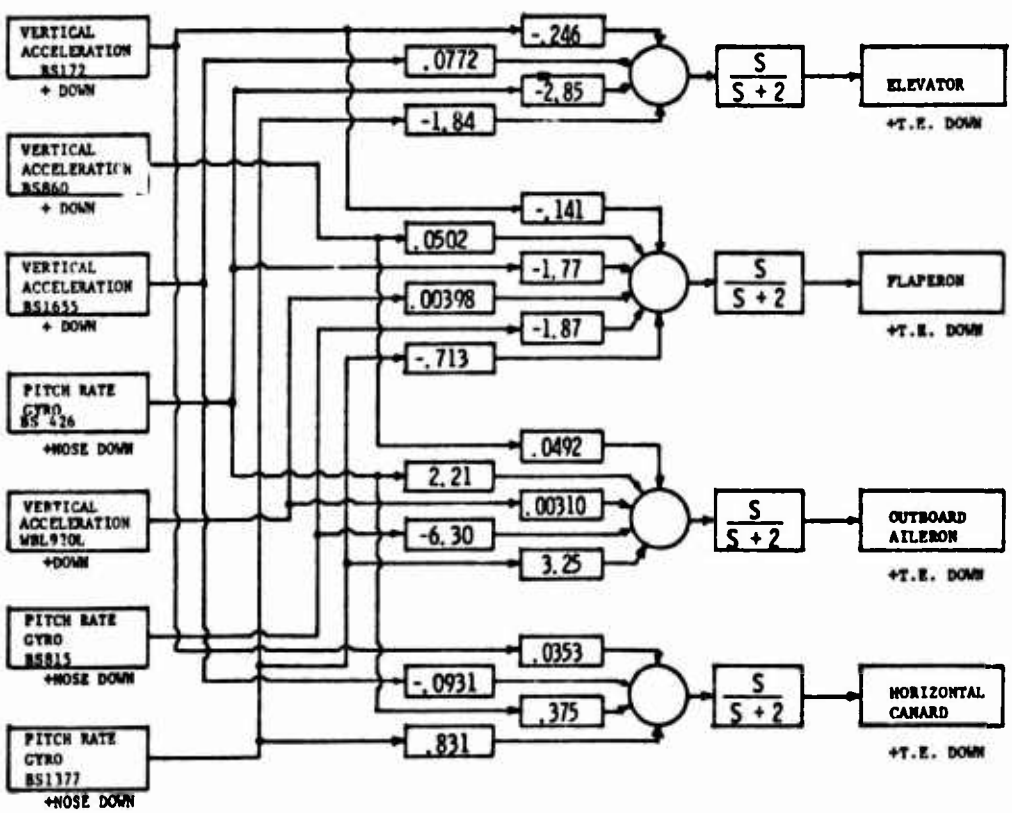


Figure 5 Longitudinal MSS

DIFFERENTIAL THRUST CONTROLLERS FOR TAXIING AN AIR CUSHION LANDING SYSTEM AIRCRAFT

By Randall V. Gressang, Captain, USAF
Air Force Flight Dynamics Laboratory (AFFDL)
Wright-Patterson Air Force Base, Ohio

ABSTRACT

The problem of improving the yaw control of a slowly taxiing Air Cushion Landing System (ACLS) aircraft by means of a linear feedback controller is posed. A simplified model of the ACLS is formulated, and a control system for the ACLS is derived using this model, pseudo observers to reconstruct missing observations, and the history vector approximation to calculate state feedback control laws. Two alternate controllers are also derived by simplifying this control law. The responses of these controllers for an initial yaw angle misalignment are calculated and compared with the control free response and the response of a rate feedback controller. The controllers derived in this paper compare favorably with both the no control and rate feedback system responses.

I. INTRODUCTION AND PROBLEM STATEMENT

An Air Cushion Landing System [1][2] offers an attractive method of freeing an aircraft from dependence on prepared runways, while achieving a weight reduction compared to conventional wheel landing gear. The Air Cushion Landing System is formed by attaching an elastic trunk to the underside of the fuselage, the trunk appearing like a life raft attached to the bottom of the fuselage. Air is pumped into the trunk, and exhausts through holes in the bottom of the trunk. This forms an air pocket underneath the aircraft, similar to an air bearing, which supports the aircraft weight and suspends the aircraft in a manner insensitive to the condition of the ground surface. The insensitive suspension allows takeoffs and landings from unprepared sites. After takeoff, the elastic trunk is depressurized and pulled up against the underside of the fuselage to minimize drag.

A vehicle supported on an air pocket lacks the side forces associated with ground contact. Without these side forces, the aircraft may exhibit poor weather cocking and yaw damping characteristics [2][3]. For an aircraft taxiing at low speed, aerodynamic controls are ineffective due to the small dynamic pressure, therefore brakes or thrusting devices must be used to overcome directional control difficulties. On a multi-engine aircraft, a powerful thrusting effect for directional control can be obtained by differential thrusting of the engines.

The current test aircraft considered in this paper is a DeHavilland Buffalo Aircraft (CC-115 or XC-8A) modified to incorporate an Air Cushion Landing System. The problem then is to design an active control system to assist the pilot in controlling the heading of the aircraft while taxiing at low speed, using only differential thrust for directional control. Thus directional control by means of brakes or side force generators is not considered. The differential thrust control is to be obtained by varying engine RPM rather than propeller pitch, so that the results will be indicative of results for an aircraft of similar size with turbojet engines.

The control system is designed to act in parallel with the pilot. Since maintaining a desired heading in the presence of disturbances is a tracking or regulator task, a pilot model [4] is used to represent the pilot. Using the pilot model, a differential delay equation is derived for the pilot-vehicle system. The control system design is split into two parts, the design of a state feedback control law yielding acceptable response using an approximate method of Hanson [5], and the design of an observer for the differential delay equation to reconstruct missing observations, using a method due to Gressang [6]. After the control system is designed, several simplifications of the control system are considered and the responses of all of the systems are compared with the no control response and the response for a rate feedback controller.

II. MATHEMATICAL MODEL OF AIRCRAFT AND PILOT

In this section a simplified mathematical model of the aircraft and pilot is derived. This model takes maximum advantage of the task limitation to low speed taxiing, and is used for the control design and evaluation of the following sections.

Table I, taken from Kurylowich [2], lists parameters of the CC-115. The yawing moment equation of the aircraft is given by [2]

$$N = - I_{xz}\dot{P} + I_{zz}\dot{R} - I_{xx}QP + I_{xz}QR + I_{yy}PQ$$

where P, Q, and R are roll, pitch, and yaw rates in body axes, and N is the yawing moment in body axes. Neglecting the products of angular rates as being of second order, and neglecting I_{xz} compared to I_{zz} (as I_{xz} is only about 5% of I_{zz}), the yawing moment equation uncouples from the other aircraft equations of motion. The yawing moment equation thus becomes

$$N = I_{zz}\dot{R}$$

The moment N is generated by aerodynamic, control surface, air cushion, and differential thrust components. As the task considered in this paper is limited to low ground speeds, the aerodynamic and control surface forces are neglected. This approximation is valid below about 30 knots [1:151]. Since the air cushion suspension system does not directly contribute a moment about the yaw axis [1:34 -114], only the differential thrust term is kept in the yawing moment equation.

The differential thrust term is assumed to arise from the pilot correcting for yaw error using the engine throttles. This term is represented by the cascade of two transfer functions, one for the engine response, and one for a human engaged in a closed loop compensatory tracking task. The engine transfer function is taken as

$$G_E(s) = \frac{510e^{-.5s}}{1 + 2s}$$

based upon limited data about thrust versus time following a reverse thrust command [1:156].

A crossover model is used for the pilot transfer function [4][7][8]. The model is that for k/s^2 plant dynamics with the gain, lead, and lag chosen to stabilize the closed loop system. The resulting pilot transfer function is

$$G_p(s) = 10\ (1 + 5s)\ e^{-.5s}$$

A block diagram of the complete system model used to design the control system is given in Figure 1. In this block diagram, it is assumed that any external control helping the pilot must also act through the engine as the pilot does. If the only measurement available is yaw angle, this block diagram is equivalent to the following differential delay equation:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -.5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -.005013 & -.02507 & 0 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \\ x_3(t-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ .0005013 \end{bmatrix} u(t-.5)$$

and the measurement equation

$$y(t) = [1 \quad 0 \quad 0] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

III. CONTROL DESIGN

The control system is designed using a technique which restricts the control to being a piecewise constant functional of the state of the differential delay equation, but otherwise places no apriori restrictions upon the control system structure. The design proceeds in two stages, the first stage being to reconstruct missing observations, the second stage being to design a state feedback regulator.

The missing observations are reconstructed using a system called a pseudo observer. Reference 6 defines a pseudo observer, gives sufficient conditions to construct a pseudo observer, and indicates an algebraic procedure that can be used to design a pseudo observer. The algebraic procedure of Reference 6 is used to design a pseudo observer for the ACLS System, under the assumption that the only measurement available is the yaw angle. It should be noted that a pseudo observer cannot be designed if the only measurement is the yaw angle rate, because with this measurement the system is not observable, and the state cannot be determined uniquely from the measurement.

For the ACLS System, the pseudo observer given by the design procedure is actually an observer, and is specified by the equations

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 6 \\ 56 \end{bmatrix} y(t) + \begin{bmatrix} .04513 \\ -.09527 \end{bmatrix} y(t-1) + \begin{bmatrix} .0005013 \\ -.0005013 \end{bmatrix} u(t-.5)$$

$$\begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5.5 & .5 & .5 \\ 5.25 & 1.75 & .75 \end{bmatrix} \begin{bmatrix} y(t) \\ z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -.02507 & 0 & 0 \end{bmatrix} \begin{bmatrix} y(t-1) \\ z_1(t-1) \\ z_2(t-1) \end{bmatrix}$$

The first set of these equations gives the dynamics of the observer. The second set of equations furnishes an approximation to the state of the differential delay system, which becomes exact as t approaches infinity.

The state feedback regulator is designed using the history vector approximation used by Hanson [5]. This method is chosen for simplicity, and because it did not involve consideration of non-standard partial differential equations. The method provides a piecewise constant control, and proceeds as follows.

First, the differential delay equation representing the system to be controlled is converted to an integral equation. Then using the assumption that the control inputs are piecewise constant, this integral equation is transformed to an infinite dimensional state discrete time system. Approximating the integrals in the preceding step by the trapezoidal rule, a finite dimensional discrete time system is constructed which approximates the infinite dimensional discrete time system. This finite dimensional system is called the history vector system, and has the form

$$h(k + 1) = Ah(k) + Bv(k)$$

where h(k) is the history vector, and v(k) is the control input sequence.

A quadratic loss discrete time regulator problem [9] is now posed for the history vector system. Controllability of the history vector system implies that the regulator problem will provide a unique gain matrix K, which can be used to specify the next control value in the sequence of control values of the piecewise constant state feedback control.

For the ACLS System, the control is held constant over .5 second intervals. The history vector for this system is of dimension 10, and given by

$$h(k) = \begin{bmatrix} x_1(.5k) \\ x_2(.5k) \\ x_3(.5k) \\ x_1\{.5(k-1)\} \\ x_2\{.5(k-1)\} \\ x_3\{.5(k-1)\} \\ x_1\{.5(k-2)\} \\ x_2\{.5(k-2)\} \\ x_3\{.5(k-2)\} \\ u\{.5(k-1)\} \end{bmatrix}$$

Table II gives the A and B matrices of the history vector system. It should be noted that even though the dimension of the history vector system is much larger than the original number of differential delay equations, the matrices in the history vector system are sparse.

The weighting matrices Q and R of the discrete time regulator problem are chosen by trial and error, the object being to find a Q and an R such that the controlled system exhibits good recovery from an initial heading offset.

$$Q = \begin{bmatrix} 0 & & & & & & & & & \\ & 1. \times 10^4 & & & & & & & & \\ & & 1. \times 10^4 & & & & 0 & & & \\ & & & 0 & & & & & & \\ & & & & 0 & & & & & \\ & & & & & 0 & & & & \\ & & & & & & 0 & & & \\ & & 0 & & & & & 0 & & \\ & & & & & & & & 0 & \\ & & & & & & & & & 0 \end{bmatrix}$$

and

$$R = .5$$

were found to yield satisfactory results for the ACLS System. The feedback gain matrix corresponding to this Q and R was determined by iterating the matrix Riccati equation until a steady state was reached, which required 175 iterations.

The calculations were performed on a CDC 6600 computer. The feedback gain matrix determined by this procedure is

$$K = \begin{bmatrix} .722 \\ 126.8 \\ 219.1 \\ -.522 \\ -2.61 \\ 0. \\ -.275 \\ -1.37 \\ 0. \\ .103 \end{bmatrix}$$

During the trials with various different Q and R, it was noted that K was not very sensitive to even large changes in Q and R.

Given K, the next value of the control is then specified by

$$u_{next} = .772x_1(t) + 126.8x_2(t) + 219.1x_3(t)$$

$$-.522x_1(t-.5) - 2.61x_2(t-.5) - .275x_1(t-1)$$

$$-.137x_2(t-1) + .103\, u_{present}$$

This formula completely specifies the piecewise constant state feedback control law.

The closed loop controller for the system with incomplete measurements is now formed by replacing $x_1(t)$ by $y(t)$ (the measurement), $x_2(t)$ by $z_1(t)$, and $x_3(t)$ by $z_2(t)$ in the expression for the next value of the control ($z_1(t)$ and $z_2(t)$ are the pseudo observer reconstructions of $x_2(t)$ and $x_3(t)$). The specification of the feedback controller for the ACLS System is completed by requiring that during initial start up, the control input to the ACLS be held at zero until the observer part of the controller has operated for 3 seconds. This delay before turning on the control input is required to prevent large transients arising from incorrect control inputs that result from initial observer errors. The 3 second time interval is chosen as it is greater than 3 time constants of the largest time constant in the pseudo observer, yet is short compared to the characteristic response time of the open loop ACLS System.

IV. SIMPLIFIED CONTROLS

As only two elements of the gain matrix have significant magnitudes, and these elements are associated with undelayed states, it is possible to design a simple continuous time controller from the sampled controller of the previous section. The simplified control is developed by simplifying both the state feedback control law and the pseudo observer, and then cascading them to form the closed loop controller.

The feedback control law is simplified by modifying the gain matrix. The new gain matrix

$$K = \begin{bmatrix} 0 \\ 100 \\ 200 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

is determined from the old gain matrix by rounding off all of its elements to one significant figure. As .5 second is much smaller than the response time of the closed loop system, the control law is taken to be continuous rather than piecewise constant.

Requiring the pilot to generate a 1 + 5s lead time taxes his capabilities. However, by adding 50 to the K_2 position, the pilot feedback loop can be closed by the pilot transfer function

$$Gp(s) = 10e^{-.5s}$$

which corresponds to a relaxed pilot. Therefore the gain matrix is augmented to

$$K = \begin{bmatrix} 0 \\ 150 \\ 200 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and the lead term dropped from the pilot model.

Since all of the delayed terms in the pseudo observer are small, the pseudo observer is simplified by dropping all delayed terms.

The simplified controller for heading measurements is then specified by the equations

$$\dot{z}_1 = -2z_1 + 6y$$

$$\dot{z}_2 = -4z_2 - 56y$$

$$\bar{x}_2 = 5.5y + .5z_1 + .5z_2$$

$$\bar{x}_3 = 5.25y + 1.75z_1 + .75z_2$$

$$u = 150\bar{x}_2 + 200\bar{x}_3$$

Note that in this controller, yaw angle is never fed back (except by the pilot), so that a yaw angle set point need not be specified. This overcomes what could be a weakness in implementing the controller of the previous section. The 3 second delay between starting the observer and closing the controller loop is retained in this system.

A second simplified controller is derived under the assumption that the yaw angle rate is measured instead of the yaw angle. This possibility may arise, as it may be easier in the ACLS (using a rate gyro) to measure the angular rate than to measure the yaw angle.

If angular rate is measured instead of the angle, the only part of the controller that must be changed is the observer section. An observer or pseudo observer cannot be designed if only angular rate is measured. This is due to the impossibility of reconstructing $x_1(t)$ and to the presence of the $-.005013\ x_1(t-1)$ term in the $\dot{x}_3$ equation. Since $x_1(t)$ is not required for the simplified feedback gain matrix, neglecting delayed terms allows a conventional observer to be designed for $x_3(t)$. Using this observer, the simplified controller for yaw angle rate measurement is specified by

$$\dot{z} = -3.5z + 7y$$

$$\bar{x}_3 = 1.5z - 3y$$

$$u = 150y + 200\bar{x}_3$$

Note that here the observation y is $x_2(t)$, the yaw rate, instead of $x_1(t)$. The controller is continuous time, and incorporates the 3 second delay in starting control action.

It should be noted that both of the simplified controllers developed in this section have the same form as conventional analog autopilots, with the exception of the 3 second delay in starting control action. Thus both of the simplified controllers could be synthesized using conventional procedures.

IV. EVALUATION OF THE CONTROLLERS

The various control laws were evaluated by calculating the time response of the ACLS System to an initial yaw misalignment of one unit. The time responses were computed by integrating the differential delay equations using a modified fourth order Runge Kutta Algorithm. These calculations were used solely for comparisons between the various controllers for the same task, and were not intended or interpreted as transients to be expected from the ACLS System with a pilot. Interpreting the results only as solutions of various sets of differential delay equations, the assumptions underlying the pilot model are not violated. Plots of the time response trajectories and time response specifications determined from the trajectories were used to compare the various controllers. The plots are given as Figures 2 through 6, and the time response specifications are presented in Table III.

Figure 2 gives the response of the ACLS System without a control to aid the pilot, and the pilot generating the (1 + 5s) lead term. Figure 3 gives the plot for the piecewise constant control, with the pilot supplying the (1 + 5s)

term, while Figures 4 and 5 are for the simplified controllers, with the pilot not supplying the lead term. The last figure, Figure 6, shows the response of a system where the pilot is generating the (1 + 5s) lead term, and is being assisted by rate feedback (the control being of the form $u = 150x_2(t)$). It is assumed in the rate feedback system that the rate can be measured directly. This system is included so a comparison can be made with a simple straight-forward control scheme.

Examining the plots, it is seen that the uncontrolled response is oscillatory with considerable overshoot, and is lightly damped. The addition of any of the four controllers effectively damps the oscillation, with the pure rate feedback going so far as to overdamp the system. The piecewise continuous, simplified angle, and simplified angle rate controls all give similar, almost critically damped responses, with small overshoot. Compared to them, the pure rate feedback system seems sluggish. Depending upon the available measurements, either the simplified angle or simplified angle rate controllers should provide an effective, easily realized control system.

VI. CONCLUSIONS AND RECOMMENDATIONS

For the ACLS System in the simplified form considered in this paper, it is shown that the combination pseudo observer/history vector control design procedure yields an effective controller. Furthermore, this controller can be used as a basis from which to form simplified controllers with comparable performance. All of these controllers compared favorably with a simple rate feedback controller.

Three recommendations can be made for further investigations along the lines of this paper. The first recommendation is to incorporate a gust model into the system, to see if random disturbances seriously degrade the performance of the control systems. Secondly, the controllers should be tried with more complex models of the ACLS dynamics, so as to check into the possibly that the model used to design the controller was simplified too much. Third and finally, the controllers should be tried in simulations of the ACLS, in order to check the validity of using the pilot models in the way they were used, and to obtain pilot opinions as to whether the controllers improve the taxiing capabilities of the ACLS.

REFERENCES

[1] Coles. Air Cushion Landing System CC-115 Aircraft. AFFDL-TR-72-4, Part I, May 1972.

[2] Kurylowich. A Digital Computer Flight Simulation of an ACLS Vehicle. Proc. 1st Conf. on Air Cushion Landing Systems, Miami Beach, Florida, 12 - 14 December 1972.

[3] Tolley. "Hovercraft Operating in Theory and Practice". Hovering Craft and Hydrofoil, 13, 2:12-16. (November 1973)

[4] McRuer and Jex. "A Review of Quasi-Linear Pilot Models". IEEE Trans. Human Factors in Eng., HFE-8, 3:231 - 249. (September 1967)

[5] Hanson. Linear Difference Games with Delays. PH.D. Dissertation, University of Illinois, Champaign, Illinois, 1970.

[6] Gressang. Observers for Infinite Dimensional Linear Systems. PH.D. Dissertation, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, 1974.

[7] McRuer, Graham, Krendel, Reisener. Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15.

[8] Halstenberg and Steketee. Application of Pilot Describing Function to Handling Qualities. GDC-ERR-1452, General Dynamics Corporation (December 1969).

[9] Kwakernaak and Sivan. Linear Optimal Control Systems. John Wiley, New York, 1972.

TABLE I. CC-115 AIRCRAFT DATA [2]

Engines	2 x GE/T64-14
	2 x UCAL ST6-73 (For ACLS)
Propellers	Hamilton Standard 63-E60-15
	3 bladed, 14.5 ft. diameter
Wing Area	945 Sq. Ft.
Wing Span	96 Ft.
Wing Reference Chord, $\bar{C}$	10.3 Ft.
Weights	Max. Takeoff 41,000 lb.
	Max. Landing 39,000 lb.

41,000 lb. A/C (Inertias in Slugs-Ft^2)

C.G. Position	26.5% $\bar{C}$	41.5% $\bar{C}$
I_{xx}	301263.0	301263.0
I_{yy}	244977.0	266987.0
I_{zz}	486428.0	508642.0
I_{xz}	29852.0	24192.0

TABLE II. A and B FOR HISTORY VECTOR SYSTEM

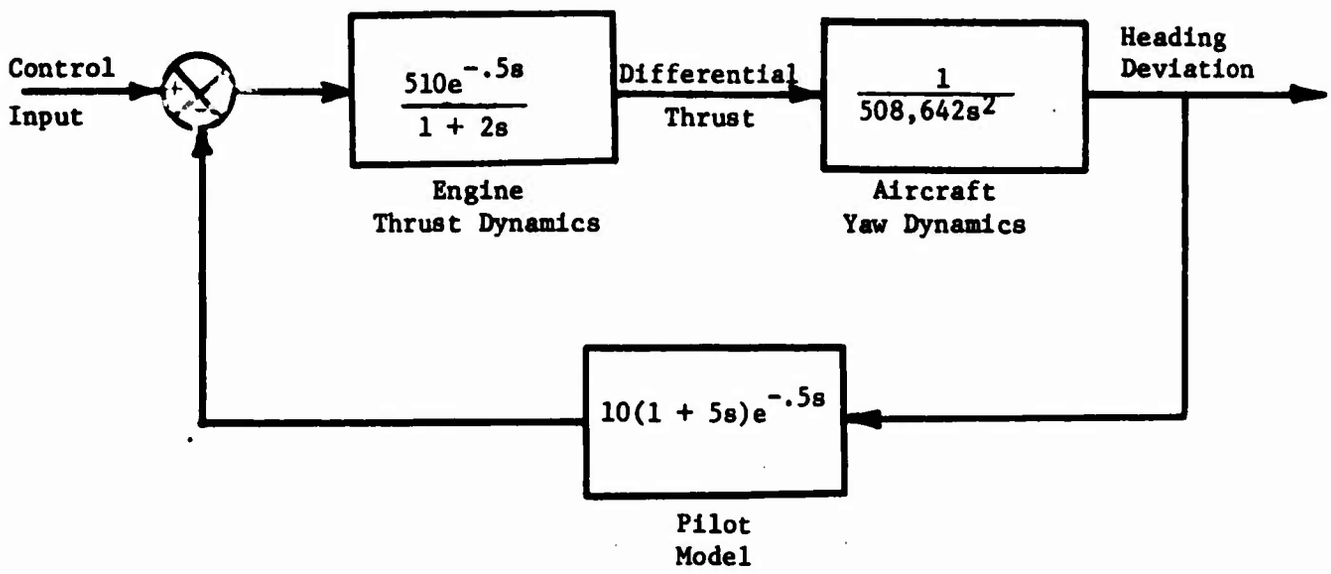
$$A = \begin{bmatrix} 1 & .5 & .1152 & 0 & 0 & 0 & -1.444_{\times 10^{-4}} & -7.220_{\times 10^{-4}} & .0 & 9.820_{\times 10^{-6}} \\ 0 & 1 & .4424 & 0 & 0 & 0 & -5.544_{\times 10^{-4}} & -2.773_{\times 10^{-3}} & 0 & 4.436_{\times 10^{-4}} \\ 0 & 0 & .7788 & -1.253_{\times 10^{-3}} & -6.268_{\times 10^{-3}} & 0 & -9.760_{\times 10^{-4}} & -4.881_{\times 10^{-3}} & 0 & 2.218_{\times 10^{-4}} \\ 1 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

TABLE III. TIME RESPONSE CHARACTERISTICS

Control System	Maximum Overshoot	Time to Maximum Overshoot	Time to First Zero Error	Settling Time To 5%	Frequency of Oscillation of the Transient
No Control	72.7%	32.5 Seconds	18.3 Seconds	> 50 Seconds	.017 Hertz
Piecewise Continuous Control	5.0%	39.5 Seconds	30.3 Seconds	41 Seconds	< .025 Hertz
Simplified Control, Angle Measurement	6.0%	39.5 Seconds	29.1 Seconds	45.5 Seconds	< .024 Hertz
Simplified Control, Angle Rate Measurement	2.6%	35 Seconds	32.2 Seconds	26.3 Seconds	.09 Hertz
Rate Feedback	No Overshoot	No Overshoot	> 50 Seconds	46.8 Seconds	No Oscillation

FIGURE 1: ACLS YAW CONTROL BLOCK DIAGRAM



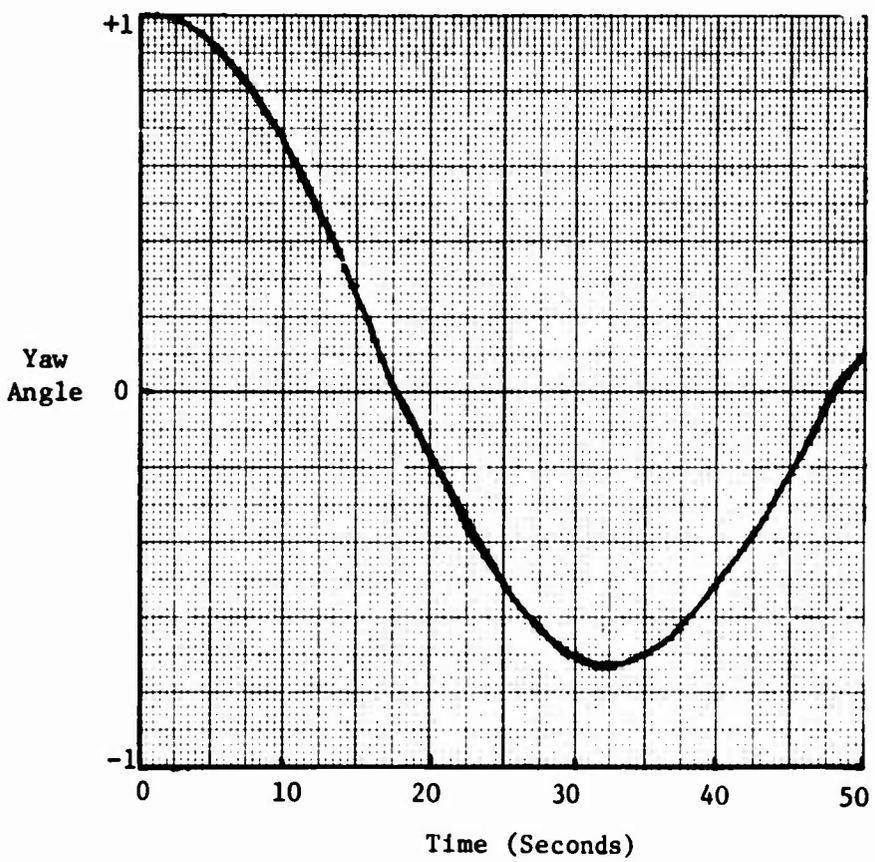


FIGURE 2: NO CONTROL TIME RESPONSE

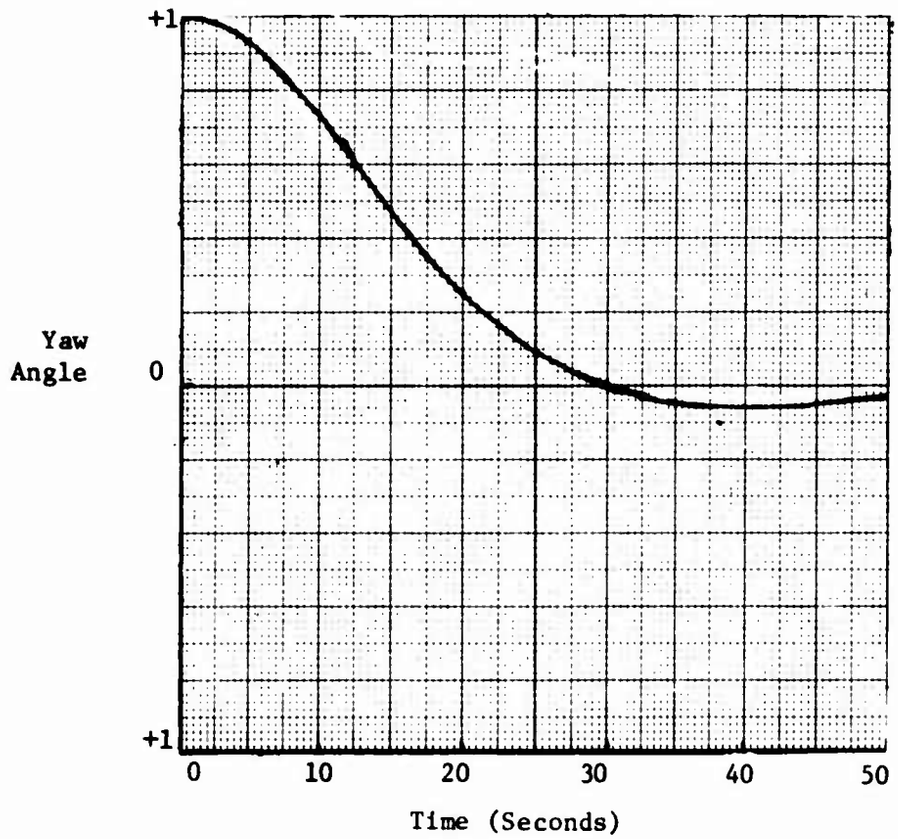


FIGURE 3: PIECEWISE CONSTANT CONTROLLER TIME RESPONSE

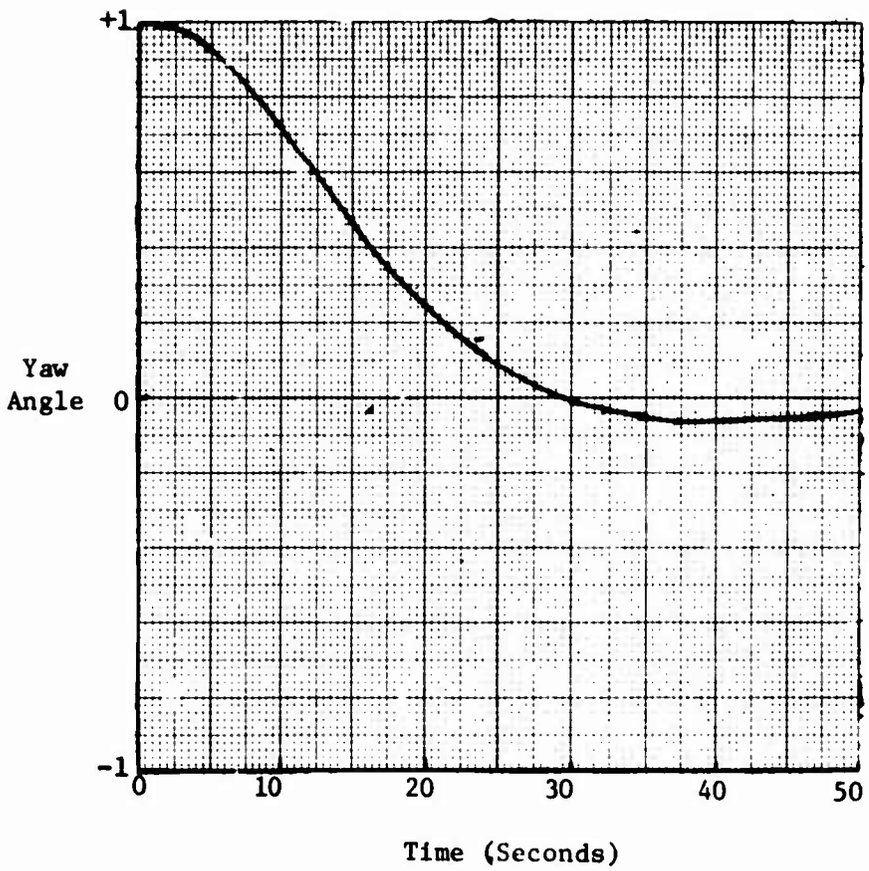


FIGURE 4: SIMPLIFIED CONTROLLER, ANGLE MEASUREMENT TIME RESPONSE

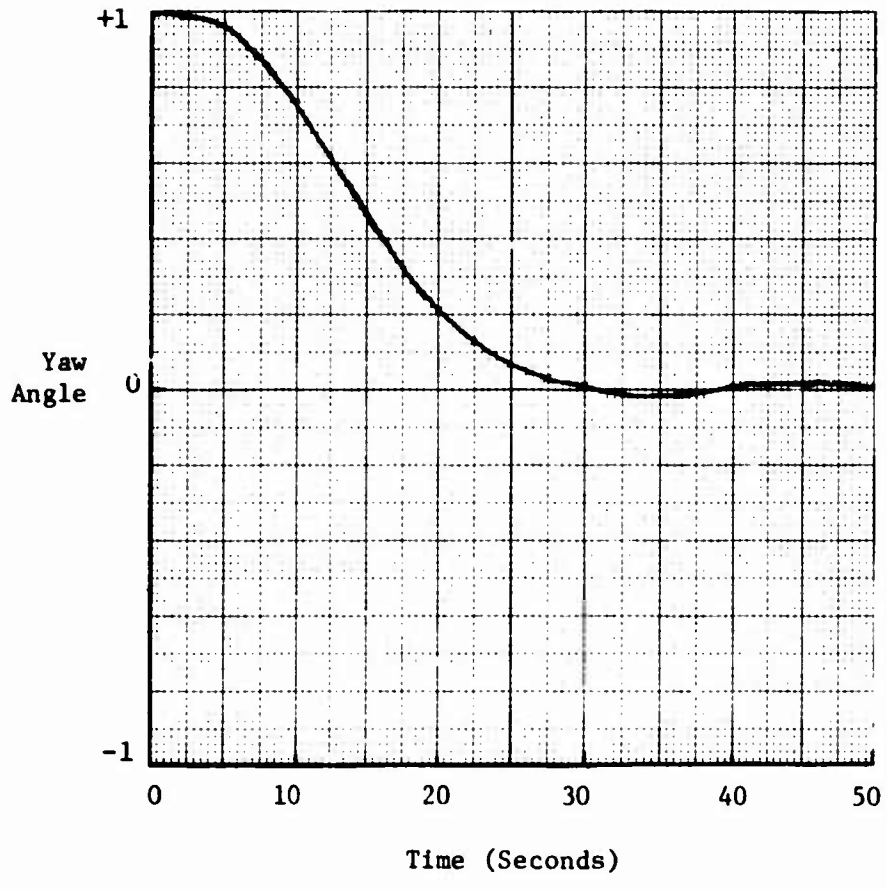


FIGURE 5: SIMPLIFIED CONTROLLER, ANGLE RATE MEASUREMENT TIME RESPONSE

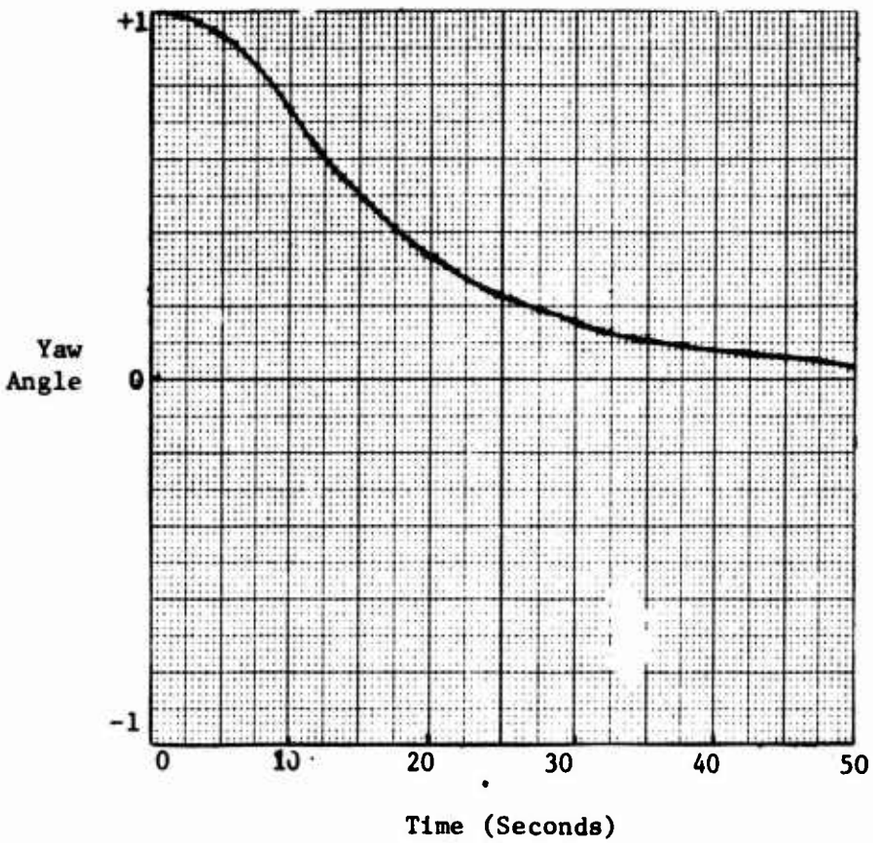


FIGURE 6: RATE FEEDBACK TIME RESPONSE

THE EFFECT OF HIGH THRUST ON OPTIMAL FIGHTER AIRCRAFT TURNING MANEUVERS

by

Gerald M. Anderson
Associate Professor, Department of Mechanics
AF Institute of Technology (AFIT/ENB)
Wright-Patterson AFB, Ohio 45433

ABSTRACT

Many fighter aircraft combat maneuvers require sustained periods of maximum turn rate. This maximum turn rate occurs when the aircraft flies at its corner velocity, the velocity at which the limits on the maximum lift coefficient and the maximum normal acceleration load factor constraints are met simultaneously. Current fighter aircraft have insufficient thrust to overcome the induced drag due to lift so that sustained maximum turn rate flight at the corner velocity is not possible. However, some future fighter aircraft may have sufficient thrust to fly this type of arc. To investigate optimal turning maneuvers, the necessary conditions for an optimal control can be applied. Unfortunately, these necessary conditions degenerate on sustained maximum turn rate corner velocity arcs. In this paper the necessary conditions are modified so that they can be applied to these corner velocity arcs. Some minimum time-to-turn trajectories are then presented to illustrate the effects of high aircraft thrust.

1. INTRODUCTION

Many fighter aircraft combat maneuvers require sustained periods of maximum turn rate flight for which maximum lift coefficient C_L is required. Below the corner velocity C_L is limited by the aerodynamics of the aircraft (the C_L limit). At velocities above the corner velocity C_L is limited by the maximum allowable normal acceleration of the aircraft due to structural or human factor constraints (the load factor limit). The fastest maximum turn rate occurs when the aircraft files at its corner velocity where both of the limits on C_L are met simultaneously. With current fighter aircraft the maximum available thrust is insufficient to overcome the induced drag due to lift so that sustained maximum turn rate flight at the corner velocity is not possible. Future fighter aircraft may, however, have sufficient thrust to allow sustained maximum turn rate flight at the corner velocity. Flight under these conditions will generally require a thrust less than maximum to maintain flight at the corner velocity. In investigating optimal turning maneuvers of high thrust aircraft, the usual necessary conditions for an optimal control breakdown on sustained maximum turn rate flights at the corner velocity.

In this paper these necessary conditions are modified so that they can be applied to corner velocity arcs. The solution to a minimum time-to-turn problem is then given to illustrate some features of optimal turning trajectories with high thrust aircraft.

2. PROBLEM STATEMENT AND NECESSARY CONDITIONS

The "point mass" aircraft equations of motion are

$$\dot{X} = V\cos\psi\cos\gamma,\ \dot{Y} = V\sin\psi\cos\gamma,\ \dot{Z} = V\sin\gamma,$$

$$\dot{V} = g[(T - D)/W - \sin\gamma],\ \dot{\psi} = gL\sin\phi/VW\cos\gamma, \tag{1}$$

$$\dot{\gamma} = g(L\cos\phi/W - \cos\gamma)/V$$

where X and Y are horizontal position coordinates, Z is altitude, V is velocity, ψ is the heading angle, γ is the flight path angle, W is the aircraft weight (which is assumed to be constant here), T is the thrust, D is drag, L is lift, ϕ is the bank angle, and g is the acceleration of gravity. Lift L has the form

$$L = (1/2)\rho(Z)AV^2C_L \tag{2}$$

where $\rho(Z)$ is the air density, A is a reference area and C_L is the lift coefficient. The drag D is assumed to be a function of V, Z and C_L, and increases monotonically with C_L for fixed V and Z.

The controls are the lift coefficient C_L, the bank angle, ϕ, and the thrust T. The thrust must satisfy the constraint

$$T_{min} \leq T \leq T_{max} \tag{3}$$

There are three constraints on C_L. First there is an upper limit of C_{Lmax} due to the aerodynamics of the aircraft. There is also a lower limit of zero. These two constraints can be summarized as

$$0 \leq C_L \leq C_{Lmax} \tag{4}$$

There is, in addition, an upper limit on the maximum acceleration of the aircraft normal to the velocity vector which can be expressed as

$$(1/2)\rho AV^2C_L/W \leq n_{max} \tag{5}$$

where n_{max} is the maximum allowable normal acceleration in g's.

Now consider a general class of optimal aircraft trajectory problems. Assuming a Meyer form for the problem, the necessary conditions can be formally applied as follows [1]. The Hamiltonian is

$$
\begin{aligned}
H = {} & V\,(\lambda_x \cos\gamma \cos\psi + \lambda_y \cos\gamma \sin\psi + \lambda_z \sin\gamma) \\
& + \lambda_v g\,[(T-D)/W - \sin\gamma] \\
& + \lambda_\psi g \rho A V\, C_L \sin\phi/2W \cos\gamma \\
& + \lambda_\gamma g\,(\rho A V\, C_L \cos\phi/2W - \cos\gamma/V) \\
& + \mu_1\,(C_L - C_{Lmax}) + \mu_2\,(\rho A V^2 C_L/2 - n_{max}\, W)
\end{aligned}
\tag{6}
$$

where the λ_i are costate variables and the multipliers μ_1 and μ_2 associated with the C_{Lmax} and n_{max} inequality constraints are defined as $\mu_1 \neq 0$ only if

$$C_L - C_{Lmax} = 0 \tag{7}$$

and $\mu_2 \neq 0$ only if

$$\rho A V^2 C_L/2 - n_{max} W = 0 \tag{8}$$

Additional multipliers adjoining the other inequality control constraints to H could be introduced, but they are not needed in this analysis.

The costate differential equations are

$$\dot{\lambda}_x = \dot{\lambda}_y = 0 \tag{9}$$

$$
\begin{aligned}
\dot{\lambda}_z = {} & \lambda_v g(\partial D/\partial Z)/W - \lambda_\psi g(\partial\rho/\partial Z) A V C_L \sin\phi/2W \cos\gamma \\
& - \lambda_\gamma g(\partial\rho/\partial Z) A V C_L \cos\phi/2W - \mu_2(\partial\rho/\partial Z) A V^2 C_L/2
\end{aligned}
\tag{10}
$$

$$
\begin{aligned}
\dot{\lambda}_v = {} & -(\lambda_x \cos\gamma\cos\psi + \lambda_y \cos\gamma \sin\psi + \lambda_z \sin\gamma) \\
& + \lambda_v g(\partial D/\partial V)/W - \lambda_\psi g\rho A C_L \sin\phi/2W\cos\gamma \\
& - \lambda_\gamma g(\rho A C_L \cos\phi/2W + \cos\gamma/V^2) - \mu_2 \rho A V C_L
\end{aligned}
\tag{11}
$$

$$\dot{\lambda}_\psi = V(\lambda_x \cos\gamma \sin\psi - \lambda_y \cos\gamma \cos\psi) \tag{12}$$

$$
\begin{aligned}
\dot{\lambda}_\gamma = {} & V(\lambda_x \sin\gamma \cos\psi + \lambda_y \sin\gamma \sin\psi - \lambda_z \cos\gamma) \\
& + \lambda_v g \cos\gamma - \lambda_\psi g \rho A V C_L \sin\phi \sin\gamma/2W \cos^2\gamma \\
& + \lambda_\gamma g \sin\gamma/V
\end{aligned}
\tag{13}
$$

The control variables must minimize H. Thus the optimal bank angle ϕ is given by

$$
\begin{aligned}
\sin\phi &= -(\lambda\psi/\cos\gamma)/[(\lambda\psi/\cos\gamma)^2 + \lambda\gamma^2]^{1/2} \\
\cos\phi &= -\lambda_\gamma/[(\gamma_\psi/\cos\gamma)^2 + \lambda_\gamma^2]^{1/2}
\end{aligned}
\tag{14}
$$

If the optimal lift coefficient is interior to the inequality constraints (4) and (5), C_L is found from the solution to the equations

$$\partial H/\partial C_L = -\lambda_v g(\partial D/\partial C_L)/W - g\rho AV[(\lambda_\psi/\cos\gamma)^2 + \lambda_\gamma{}^2]^{1/2}/2W=0 \quad (15)$$

and

$$\partial^2 H/\partial C_L{}^2 = -\lambda_v g(\partial^2 D/\partial C_L{}^2)/W > 0 \quad (16)$$

where Equations (14) have been substituted for $\sin\phi$ and $\cos\phi$. If inequality (16) is not satisfied or if the solution to Equation (15) does not satisfy the inequality constraints on C_L, then the optimal C_L is given by Equation (8), $C_L = C_{Lmax}$, or $C_L = 0$, depending on which value minimizes H and satisfies all the constraints. The optimal thrust is given by

$$T = T_{max} \text{ if } \lambda_v < 0$$

$$T = T_{min} \text{ if } \lambda_v > 0 \quad (17)$$

if both the C_{Lmax} and n_{max} constraints are not both satisfied with equality simultaneously.

If $C_L = C_{Lmax}$ and C_L is not on the n_{max} limit, the multiplier μ_1 is found from $\partial H/\partial C_L = 0$ to be

$$\mu_1 = \lambda_v g(\partial D/\partial C_L)/W + g\rho AV[(\lambda_\psi/\cos\gamma)^2 + \lambda\gamma^2]^{1/2}/2W \quad (18)$$

Recall that $\mu_2=0$ in this case. If $C_L < C_{Lmax}$ and Equation (8) is satisfied by C_L, $\mu_1 = 0$ and μ_2 is found from $\partial H/\partial C_L = 0$ to be

$$\mu_2 = 2\{\lambda_v g(\partial D/\partial C_L) + g\rho AV[(\lambda\psi/\cos\gamma)^2 + \lambda\gamma^2]^{1/2}/2\}/\rho AV^2 W \quad (19)$$

Now consider the effect of C_L being on both the C_{Lmax} and n_{max} limits simultaneously. The condition $\partial H/\partial C_L = 0$ gives

$$\partial H/\partial C_L = -\lambda_v g(\partial D/\partial C_L)/W - g\rho AV[(\lambda\psi/\cos\gamma)^2 + \lambda\gamma^2]^{1/2}/2W + \mu_1 + \mu_2\rho AV^2/2 = 0 \quad (20)$$

This is the only equation in which μ_1 and μ_2 appears. Therefore, no unique solution for these quantities exist. Since a unique solution for μ_2 is required to integrate the costate equations, it is apparent that these necessary conditions breakdown on a corner velocity arc where the C_{Lmax} and n_{max} limits are simultaneously satisfied

3. MODIFICATIONS TO NECESSARY CONDITIONS REQUIRED FOR CORNER VELOCITY ARCS

If both the C_{Lmax} and n_{max} limits are met simultaneously (i.e., a corner velocity arc) C_L can be eliminated from Equations (7) and (8) to obtain the following state equality constraint that must hold along such an arc:

$$\rho(Z)V^2 - 2Wn_{max}/AC_{Lmax} = 0 \tag{21}$$

Note that this equation can be solved for the corner velocity as a function of altitude Z. Now the methods from section 3.4 of Reference [1] pertaining to problems with state equality constraints can be applied to corner velocity arcs. Differentiation of Equation (21) gives

$$(\partial\rho/\partial Z)V^2\dot{Z} + 2\rho V\dot{V} = V\{(\partial\rho/\partial Z)V^2 \sin\gamma + 2\rho g[(T-D)/W-\sin\gamma]\} = 0 \tag{22}$$

or, since $V > 0$,

$$(\partial\rho/\partial Z)V^2 \sin\gamma + 2\rho g[(T-D)/W-\sin\gamma] = 0 \tag{23}$$

This equation contains the thrust T explicitly, thereby allowing us to solve for the value of T required to sustain maximum turn rate flight at the corner velocity.

$$T_c = T = D + W\sin\gamma - (\partial\rho/\partial Z)V^2 \sin\gamma/2\rho g \tag{24}$$

Since T must always satisfy inequality (3), it is easily seen that if $T_c > T_{max}$, sustained maximum turn rate flight at the corner velocity is not possible, as is the case with current fighter aircraft. Equation (23) is a state dependent control equality constraint that must hold along a corner velocity arc and Equation (21) can be considered to be an interior point state constraint that must hold at the start (and end) of a corner velocity arc.

To modify the necessary conditions to handle arcs on which the new control constraint given by Equation (23) is satisfied, we can introduce a new multiplier μ_3 associated with this constraint and a new Hamiltonian $\bar{H}$ defined by

$$\bar{H} = H + \mu_3\{(\partial\rho/\partial Z)V^2 \sin\gamma + 2\rho g[(T-D)/W-\sin\gamma] \tag{25}$$

where H is given by Equation (6). The definitions of μ_1, μ_2 and μ_3 for this modified problem are $\mu_1 \neq 0$, $\mu_2 = \mu_3 = 0$ only if $C_L = C_{Lmax}$ and Equation (8) is not satisfied; $\mu_2 \neq 0$, $\mu_1 = \mu_3 = 0$ only if $C_L < C_{Lmax}$ and Equation (8) is satisfied; and $\mu_3 \neq 0$, $\mu_1 = \mu_2 = 0$ only if $C_L = C_{Lmax}$ Equation (8) is satisfied, and, therefore, Equation (23) is satisfied.

The introduction of μ_3 now requires the modification of the λ_z, λ_v and λ_γ costate differential equations. These revised equations are

$$\begin{aligned}\dot{\lambda}_z = {} & \lambda_v g(\partial D/\partial Z)/W - \lambda_\psi g(\partial\rho/\partial Z)AVC_L \sin\phi/2W\cos\gamma \\ & - \lambda_\gamma g(\partial\rho/\partial Z)AVC_L\cos\phi/2W - \mu_2(\partial\rho/\partial Z)AV^2C_L/2 \\ & - \mu_3\{(\partial^2\rho/\partial Z^2)V^2\sin\gamma + 2g(\partial\rho/\partial Z)[(T-D)/W - \sin\gamma] \\ & - 2\rho g(\partial D/\partial Z)/W\}\end{aligned} \tag{26}$$

$$\begin{aligned}\dot{\lambda}_v = {} & -(\lambda_x\cos\gamma\cos\psi + \lambda_y\cos\gamma\sin\psi + \lambda_z\sin\gamma) \\ & + \lambda_v g(\partial D/\partial V)/W - \lambda_\psi g\rho AC_L\sin\phi/2W\cos\gamma \\ & - \lambda_\gamma g(\rho AC_L\cos\phi/2W + \cos\gamma/V^2) - \mu_2\rho AVC_L \\ & - \mu_3[2(\partial\rho/\partial Z)V\sin\gamma - 2\rho g(\partial D/\partial V)/W]\end{aligned} \tag{27}$$

$$\begin{aligned}\lambda_\gamma = {} & V(\lambda_x\sin\gamma\cos\psi + \lambda_y\sin\gamma\sin\psi - \lambda_z\cos\gamma) \\ & + \lambda_v g\cos\gamma - \lambda_\psi g\rho AVC_L\sin\phi\sin\gamma/2W\cos^2\gamma \\ & + \lambda_\gamma g\sin\gamma/V - \mu_3[(\partial\rho/\partial Z)V^2\cos\gamma - 2\rho g\cos\gamma]\end{aligned} \tag{28}$$

These equations replace Equations (10), (11), and (13) respectively. Note that when the aircraft is not on a corner velocity arc, $\mu_3=0$ so that Equations (26-28) reduce to the same form as Equations (10), (11), and (13).

If $\mu_2\neq 0$, it is again given by Equation (19). On a corner velocity arc where $\mu_3\neq 0$, μ_3 is found from $\partial H/\partial T$ to be

$$\mu_3 = -\lambda_v/2\rho \tag{29}$$

Due to the interior point constraint, given by Equation (21), that must be satisfied at the start of a corner velocity arc, the costate variables λ_z and λ_v may be discontinuous at this point. (The other costate variables and the Hamiltonian are continuous here since Equation (21) is an explicit function only of the states Z and V, and not X, Y, ψ, γ, nor time t.) If a corner velocity arc starts at time t_1, the discontinuities in λ_z and λ_v are given by

$$\lambda_z(t_1^-) = \lambda_z(t_1^+) + \pi V^2(\partial\rho/\partial Z) \tag{30}$$

$$\lambda_v(t_1^-) = \lambda_v(t_1^+) + 2\pi\rho V \tag{31}$$

where π is a scalar multiplier. These two costate variables are continuous at the termination of a corner velocity arc. (It is also possible to let these discontinuities occur at the termination of the corner velocity arc, with λ_z and λ_v now being continuous at the beginning of the arc. This is because there are nonunique solutions for λ_z and λ_v along such an arc [2].)

The application of these modified necessary conditions results in a multi-point boundary-value problem that must be solved to find candidates for the optimal trajectory. To summarize this multi-point boundary-value problem, the state and costate differential equations are given by Equations (1), (9), (12), and (26-28), the controls by Equations (14), (15-17), and (24), and the multipliers μ_2 and μ_3 by Equations (19) and (29), respectively, when they are nonzero. The initial states are usually all specified. At the beginning of a maximum turn rate corner velocity arc, Equations (21), (31), and (32) must be satisfied, and all the states, the costates λ_x, λ_y, λ_ψ, and λ_γ, and the Hamiltonian are continuous. At the termination of an optimal corner velocity arc Equation (21) must again be satisfied, and all the states, costates and Hamiltonian are continuous. At the final time any specified terminal conditions and the resulting transversality conditions must be satisfied. The multiplier π in Equations (31) and (32) must be chosen to ensure that all these conditions are satisfied.

4. A MINIMUM TIME-TO-TURN PROBLEM

To illustrate the characteristics of optimal aircraft trajectories with corner velocity arcs, here we consider the problem of finding the trajectory of a high thrust aircraft that results in the minimum time-to-turn through a specified heading angle. All the initial states are given and only the final heading angle ψ_f is specified. All the other final states are free. The aircraft parameters assumed for this problem are W = 42,000 lbs, A = 430 ft^2, C_{Lmax} = 1, n_{max} = 5g's, T_{max} = 65,270 lbs and T_{min} = -6,527 lbs. The expression for drag is

$$D = (1/2)\rho AV^2(C_{Do} + kC_L^2) \tag{32}$$

where the zero lift drag coefficient is C_{Do} = .04 and the induced drag factor is k = .2. With the exception of T_{max}, which is very high, all of these parameters are typical of current fighter aircraft. The negative value of T_{min} is attained through the use of speed brakes. The exponential air density expression is, in slugs/ft^3,

$$\rho = .0023769\ e^{(-Z/23,800)} \tag{33}$$

where Z is in feet. The payoff is time or

$$J = t_f \tag{34}$$

The expressions for the state equations, the costate equations, the control variables, and the multipliers μ_2 and μ_3 given previously are valid here. Application of the transversality conditions yields

$$\lambda_x(t_f) = \lambda_y(t_f) = \lambda_z(t_f) = \lambda_v(t_f) = \lambda\gamma(t_f) = 0 \tag{35}$$

$$\bar{H}(t_f) = -1 \tag{36}$$

With $\lambda\gamma(t_f) = 0$ and $\lambda_\psi(t_f) \neq 0$, in general, Equation (14) gives at t_f

$$\cos\phi = 0 \ , \quad \sin\phi = \pm 1 \tag{37}$$

where the positive sign is chosen if $\lambda_\psi(t_f) < 0$ and the negative sign if $\lambda_\psi(t_f) > 0$. A combination of Equations (6,25, and 35-37) yields the following expression for $\lambda_\psi(t_f)$

$$\lambda_\psi(t_f) = \mp 2W \cos\gamma/g\rho AVC_L\Big|_{t_f} \tag{38}$$

In the specific examples that follow the upper signs are always chosen in Equations (37) and (38).

When a maximum turn rate arc at the corner velocity is part of an optimal trajectory for this problem, the trajectory terminates on this arc. Thus we have a three-point boundary-value problem to be solved here. In order to generate optimal trajectories from this three-point boundary-value problem, solutions were generated by backward integration from assumed terminal conditions with the transversality conditions satisfied. Switching from the corner velocity arc to a $T=T_{max}$ or $T=T_{min}$ arc was then forced at some time t_1. The multiplier π must then be found so that the Hamiltonian is continuous at t_1. Equating the expressions for the Hamiltonian at t_1^- and t_1^+, and using Equations (30) and (31), the following expression for π is obtained:

$$\pi = g\lambda_v^+ (T^+-T^-)/[2V\rho g(T^- - D - W\sin\gamma) + WV^3(\partial\rho/\partial Z)\sin\gamma] \tag{39}$$

where the superscript (+) indicates that the quantity is evaluated at t_1^+ on the corner velocity arc and (-) indicates that the quantity is evaluated at t_1^- on the T_{max} or T_{min} arc. Note that the only quantity evaluated at t_1^- is the thrust T. For this problem it turns out that the same value of π is obtained for both $T^- = T_{max}$ and $T^- = T_{min}$. This value of π always results in $\lambda_v(t_1^-) = 0$. The value of T^- is then determined by

$$T = T_{max} \text{ if } \lambda_V > 0$$
$$T = T_{min} \text{ if } \lambda_V < 0 \qquad (40)$$

The backward integration is then stopped at some arbitrary initial time.

Figure 1 shows the solution to a typical problem with a thrusting sequence $\{T_{max}, T_c\}$. The assumed values of t_1 and t_f are $t_1 = 5$ sec and $t_f = 10$ sec. The initial states (at t=0) are X = -3506 ft, Y = 5457 ft, Z = 9380 ft, V = 769 ft/sec, $\dot{\psi}$ = -2.05 radians, and γ = .39 radian. The resulting final states at $t = t_f = 10$ sec are X = Y = 0, Z = 10,000 ft, V = 791 ft/sec (the corner velocity at 10,000 ft altitude), ψ = 0, and γ = -.20 radian. The lift coefficient is on the C_{Lmax} limit throughout the flight. Some trends can be noted from Fig. 1. The velocity starts below the corner velocity with a positive flight path angle γ. In order to quicken the acceleration of the aircraft to the corner velocity, the bank angle, ϕ, is initially set at about 96° so that a component of lift acts to decrease γ and, hence, increase $\dot{V}$. Once the corner velocity is reached, the thrust is reduced to sustain flight at this condition. This corner velocity thrust decreases between t_1 and t_f because the flight path angle γ decreases from .08 radian, where a component of gravity tends to reduce $\dot{V}$, to -.20, where a component of gravity tends to increase $\dot{V}$ allowing a corresponding reduction in thrust.

Figure 2 presents similar results for a typical problem with a $\{T_{min}, T_c\}$ thrust sequence. Again $t_1 = 5$ sec and $t_f = 10$ sec. The initial states are X = -3387 ft, Y = 6028 ft, Z = 8292 ft, V = 10,037 ft/sec, ψ = -1.94 radians, and γ = .20 radian. The resulting final states are the same as for the previous problem except that $\gamma(t_f) = .1$ radian. (It is the final value of γ that determines the optimal thrust sequence in these problems.) The lift coefficient is on the load factor limit throughout this trajectory. Note that $T = T_{min}$ is initially required to reduce the velocity to the corner velocity. This deceleration is aided by using a bank angle less than 90° to increase γ which, in turn, further reduces $\dot{V}$.

Some general trends can be noted on the choice of control required in these minimum time-to-turn problems. The maximum allowable value of C_L is always used. The initial thrust and the initial bank angle are determined by the initial velocity, V(0). If $V(0) > V_c$, the corner velocity, $T(0) = T_{min}$ and ϕ is set slightly less than 90° to further

reduce V. Of $V(0) < V_c$, $T(0) = T_{max}$ and ϕ is set slightly greater than 90° to decrease γ and thereby further increase V. Once the corner velocity is attained, the thrust is adjusted to sustain flight at this condition and ϕ smoothly approaches 90° as the desired terminal heading is approached.

5. CONCLUSIONS

Sustained maximum turn rate arcs at the corner velocity will be of considerable importance for future high thrust fighter aircraft when the fastest possible turn is required. In the context of optimal turning maneuvers for these aircraft, the possible presence of corner velocity arcs increase the complexity of the solution to these problems in that a multi-point boundary value problem must now be solved rather than a two-point boundary value problem required for current fighter aircraft. However, by investigating general trends in the solution to these multi-point boundary value problems, it may be possible to generate near optimal "rules of thumb", as was done with the examples in Section 4, to aid the pilot in flying near optimal maneuvers.

6. REFERENCES

1. Bryson, A. E. and Ho,Y. C., Applied Optimal Control, Ginn and Co., Waltham, Mass., 1969.

2. Bryson, A. E.; Denham, W. F.; and Dreyfus, S. E., "Optimal Programming Problems with Inequality Constraints," AIAA Journal, Vol. 1 (1963), 2544.

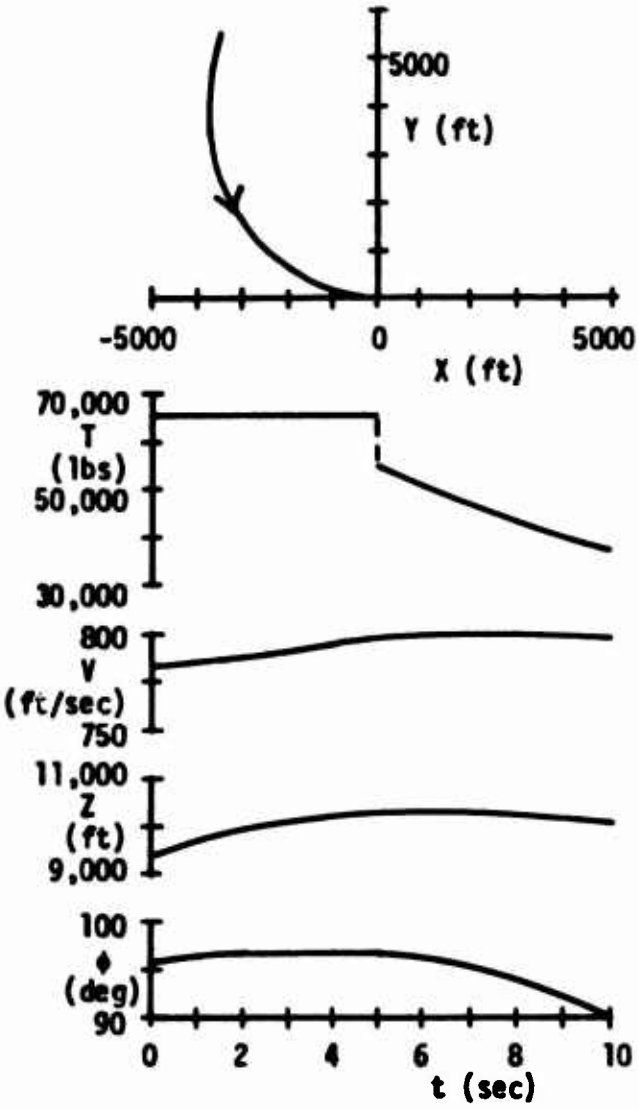


Figure 1. Results for minimum-time-to-turn problem with thrust sequence $\{T_{max}, T_c\}$.

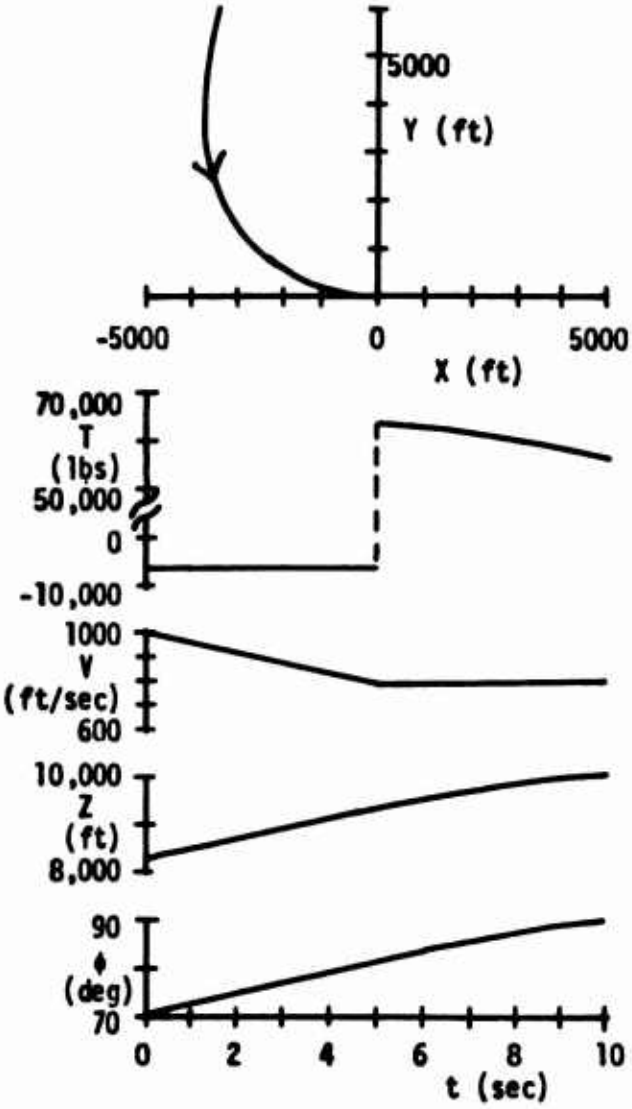


Figure 2. Results for minimum-time-to-turn problem with thrust sequence $\{T_{min}, T_c\}$.

DYNAMIC MISSILE SYNTHESIS

1 Lt Michael L. Mumford
AF Armament Laboratory
Eglin AFB, Florida

1. Introduction

Recognizing a specific lack in the area of evaluation of the aircraft/weapons system for fighter aircraft/missile systems, the Analysis Division of the Armament Laboratory initiated a study designed to develop a tool for evaluation of the total weapons system. This work resulted in the Missile and Aircraft System Effectiveness (MASE) computer model, allowing the combined evaluation of air-to-air missiles and fighter aircraft in terms of kill probability from the onset of the weapon delivery maneuver.

To date, the primary emphasis in missile performance evaluation has been on the operation of the missile alone, including the terminal fuze/warhead/target interaction. This approach has involved the detailed simulation of the missile subsystems, through which the terminal lethality characteristics of the missile have been mapped back to the time of launch as kill probability (PK) envelopes. Such an approach has thus evaluated the missile without consideration of the characteristics of the launch platform.

Past aircraft performance evaluations, even when coupled with missile flight simulations, have generally been oriented toward gun firings. Also, the ones which include missile models usually generate miss distances, which do not show a direct correlation with PK information due to the complex nature of the end game interaction.

This study was generated to combine the aircraft/missile systems so as to answer such questions as early launch versus press-on for a higher PK. Other unknowns that could be resolved include what are the effects of changes in the PK envelope on launch opportunity? What benefits accrue from an off-boresight capability? What missile or aircraft changes are necessary to improve the overall missile/aircraft effectiveness in air-to-air combat? To answer these questions, an accurate model of the dynamics prior to missile launch is required. For such a model, the objective is to identify meaningful weapon delivery trajectories which can serve as a basis for the total system evaluation. Since the primary objective of air-to-air combat is to kill the opposing aircraft, it was decided that the meaningful trajectories for this analysis are those that result in maximizing the attainable PK at launch. A sample trajectory illustrating this idea is shown in Figure 1.

To start this project, computer models based on classical optimal control theory were developed at the Armament Laboratory. Having as an objective function the maximization of PK through a steepest-descent technique, these programs suffered the usual difficulties of nominal control selection and very uneconomic run times. Also, in this approach perfect information was assumed, i.e., it was effectively assumed that the attacker a priori knows the future target maneuver history, yielding an open-loop solution. However, in real world aerial combat each pilot maneuvers based on what his opponent is doing in the present. Thus, a feedback control solution, i.e., a closed-loop control solution, is

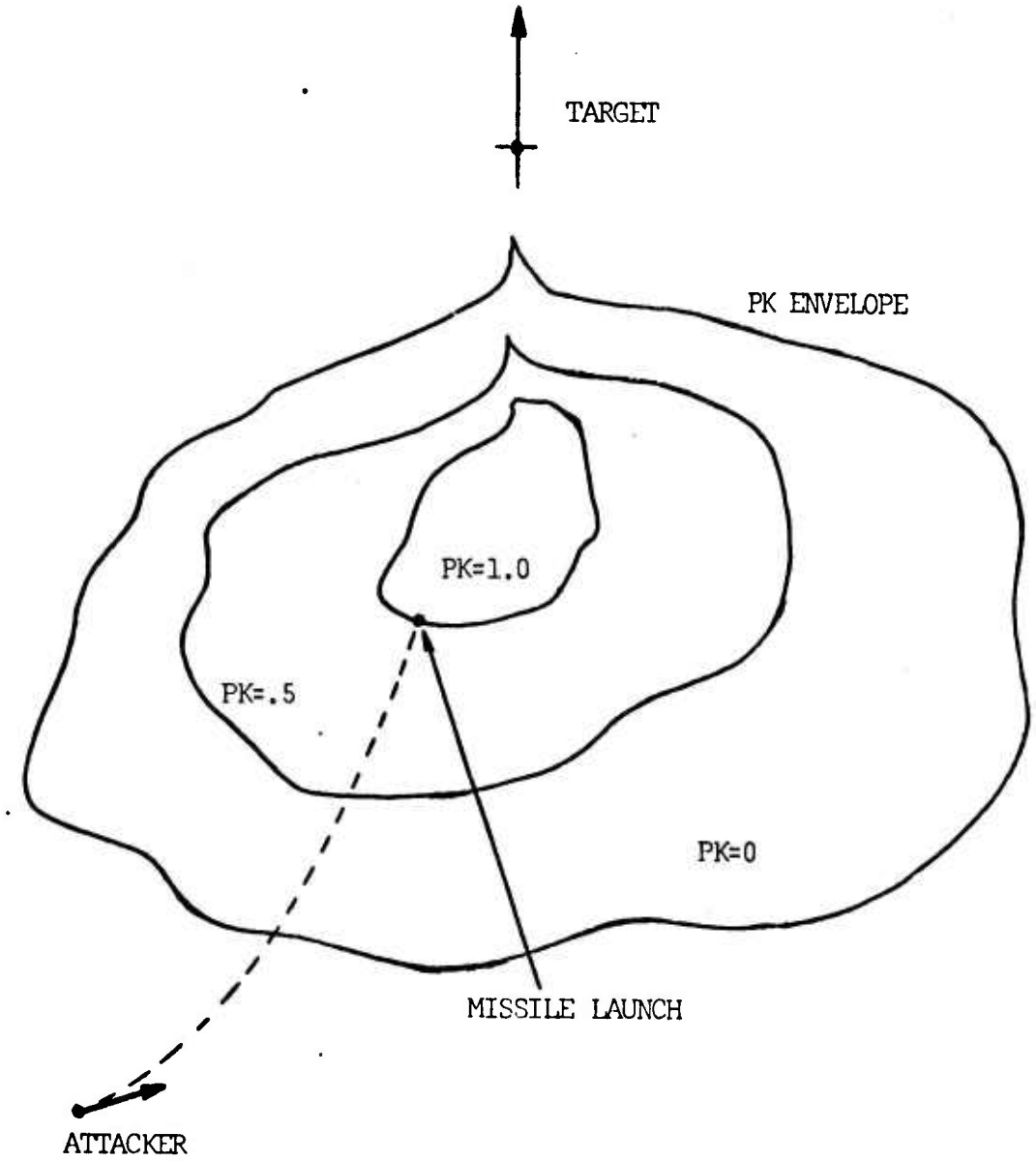


FIGURE 1. SAMPLE TRAJECTORY FOR MAXIMIZING PK

required for a realistic simulation. To circumvent the problems with the classical optimal control approach, a contract was let to Dynamics Research Corporation, Wilmington MA, to apply their already-demonstrated expertise in singular perturbation theory to this problem. The rationale for turning to this source is that singular perturbation theory yields an approximate closed-loop control solution which circumvents the problems associated with the steepest-descent solution.

2. Feedback Control Law Generation

For the purposes of this study an optimal weapon delivery strategy was defined as the control that should be applied, as a function of the current state of the system, to maximize the probability of kill. Optimal control theory states that such a control law does indeed exist. This statement is part of the classical Hamilton-Jacobi theory and is basically Bellman's principle of optimality (1). Finding the optimal feedback control law amounts to finding a solution to the Hamilton-Jacobi-Bellman partial differential equation. Unfortunately there is no known analytical technique for finding this solution for nonlinear problems.

However, if one is willing to sacrifice some precision, these difficulties may be circumvented by recourse to singular perturbation techniques. An illuminating, although slightly misleading, method of considering these techniques is to regard them as reduced order modeling techniques. Singular perturbation theory allows one to take a reduced order (free stream) model and account for the faster or transient dynamics

in a separate (boundary layer) analysis. The composite solution (free stream + boundary layers) is a better approximation to the original problem than the solution of the reduced order problem. In the following paragraphs the free stream and boundary layer solutions will be outlined. It is important to bear in mind that the controls derived using the reduced order models are used in the MASE program to drive an engagement simulation. The simulation consists of the full system of equations needed to realistically describe the motion of the target and attacker so that the resulting trajectories will be accurate and realizable.

Free Stream Solution

The reduced order, free stream model contains the horizontal plane relative position rate, and the energy rate dynamics. Turning dynamics are associated with the faster transients in the model and are therefore ignored in the free stream. These dynamics will be accounted for in a subsequent boundary layer analysis.

The free stream solution consists of a dash along a constant heading in the horizontal plane (see Figure 2). This optimal heading $\bar{\phi}$ is found so that the attacker aircraft will intercept the PK envelope center (projected into the horizontal plane). The target motion is projected using the horizontal components of target velocity (V_T) and target turn rate.

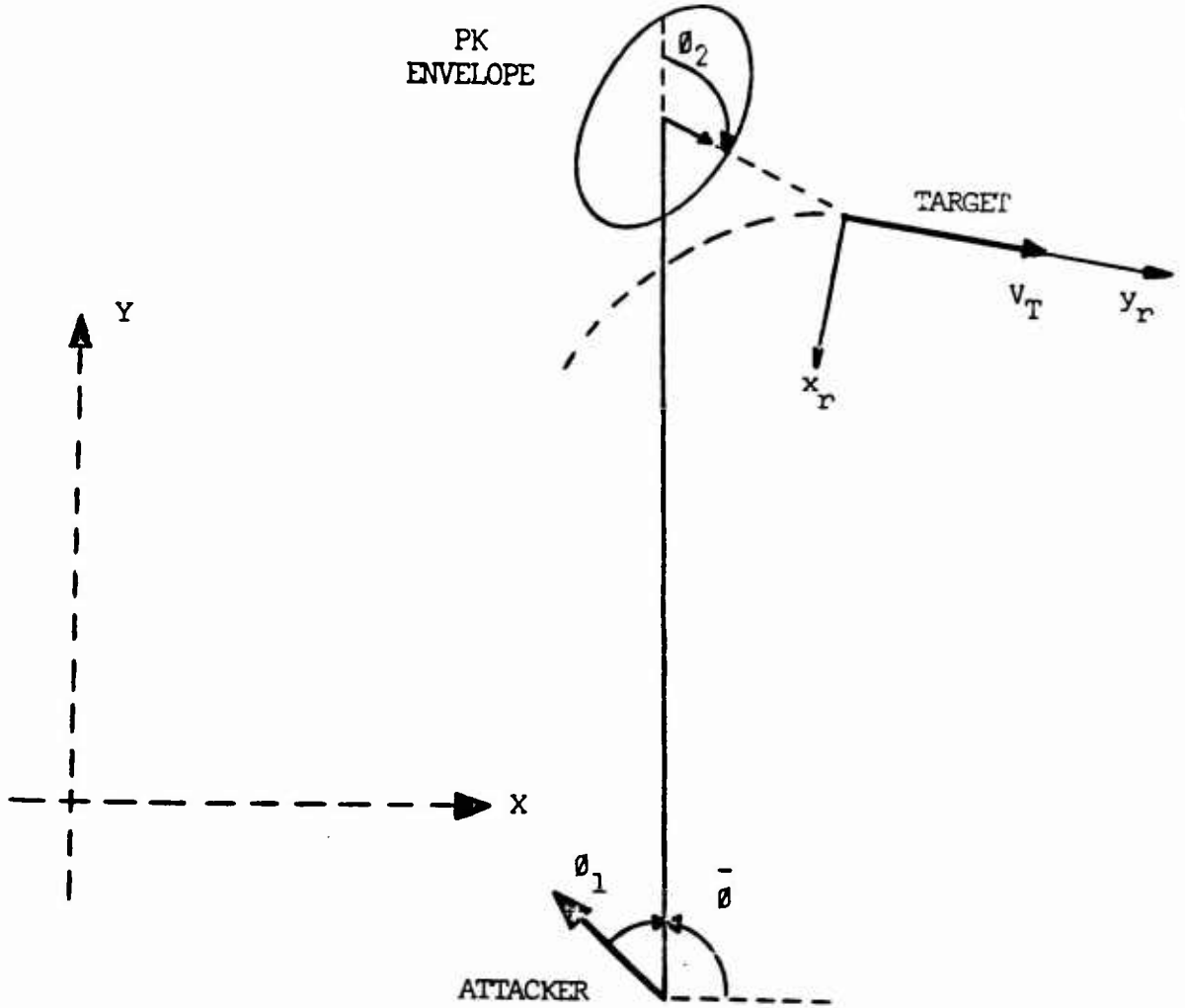


FIGURE 2. FREE STREAM DASH SOLUTION

Heading Boundary Layer

The heading transients ϕ_1 and ϕ_2 ignored in the free stream are accounted for in a separate analysis, the heading boundary layer. This allows a realistic modification of the transients ϕ_1 and ϕ_2, which are assumed instantaneous in the free stream. Analysis of the control equations lead to a feedback law of the form

$$\dot{\phi} = \pm k_{\phi} (\bar{\phi} - \phi),$$

where the plus sign applies to the initial boundary layer where the attacker approaches $\bar{\phi}$, and the minus applies to the terminal boundary layer where the attacker must turn away from $\bar{\phi}$ to satisfy the terminal boresight requirements (lead or lag at launch). The coefficient k_{ϕ} comes, through the singular perturbation theory, from the adjoint equations governing the optimality of the solution. It results in the attacker achieving an optimum balance between energy rate ($\dot{E}$) and turn rate ($\dot{\phi}$) for the current state of the engagement. Also resulting from the heading boundary layer is an optimum altitude ($\bar{h}$), which is the optimum altitude for the energy rate-turn rate balance.

Altitude and Flight-Path Angle Boundary Layers

The determination of an optimal altitude to be achieved in the heading boundary layer requires an altitude and a flight-path angle boundary layer to model transitions in the altitude (h) and flight-path angle (γ). These transients are again exponential in form and are governed by

the equation

$$\dot{h} = V \sin \gamma = \pm k_h (\bar{h} - h),$$

where $\bar{h}$ is the optimal altitude from the heading boundary layer and k_h is computed using the adjoint equations as was k_ϕ: Again, the plus sign applies to the initial boundary layer and the minus sign to the terminal boundary layer.

Since $\bar{\gamma}$ is determined by the altitude boundary layer, a flight-path angle boundary layer is required to model transients from the initial γ to the optimal γ and from the optimal γ to a terminal value. The controls in this boundary layer are the load factor and the bank angle of the attacker, which are the control variables used by the point-mass aircraft model (see Figure 3).

Summarizing, singular perturbation theory allows a systematic method of compensating for the approximations made in reduced order modeling techniques through separate boundary layer analyses. In deriving the weapon delivery maneuvers for maximizing PK, the states ignored in the reduced order, free stream solution are individually accounted for in separate heading, altitude, and flight path angle boundary layer analyses. The resulting optimal control law is obtained in feedback form.

It is important to note here that the free-stream solution and boundary layers are updated at each integration step. Thus a blending of the characteristics of each part is obtained, which allows the approximate solutions generated by this method to show excellent correlation with optimal solutions generated by the steepest-descent technique

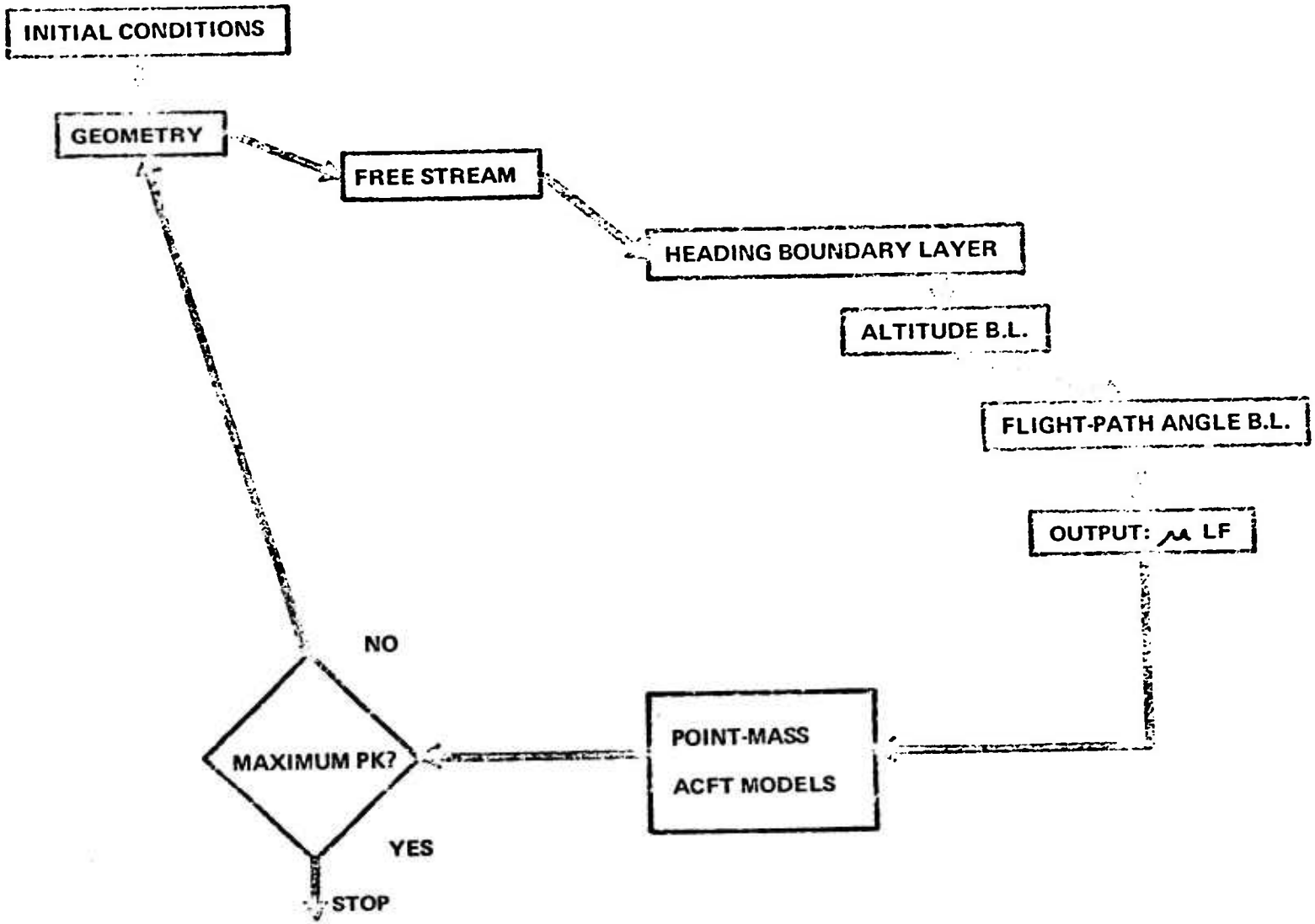


FIGURE 3. BLOCK DIAGRAM OF THE DMS SIMULATION

(see Figure 4). This singular perturbation solution is an improvement on the steepest-descent technique since no nominal control need be selected and since this is not an iterative technique, which yields a tremendous savings in computer time. Figure 5 shows some sample results for different flight conditions.

3. Conclusions

Optimal, three-dimensional attacker control laws were developed for maximizing PK against a three-dimensional maneuvering target. These feedback control laws were used to develop the Missile and Aircraft System Effectiveness (MASE) computer program. MASE is a three degree-of-freedom simulation program which generates the attacker control law and integrates the equations of motion for the attacker and target to determine the highest attainable PK for the specified initial conditions. This program can be used to map the missile PK envelopes backwards in time to the start of the weapon delivery maneuver. The resulting PK envelopes reflect the combined capabilities of both missile and aircraft, and they can be used to determine the impact that missile design changes or aircraft design changes have on the outcome of an engagement.

The major problem with the present MASE program is that it is only programmed for a predetermined target maneuver. Nothing in the MASE methodology precludes the use of a target performing countering maneuvers, and a search is being made to find a suitable target driver. Two

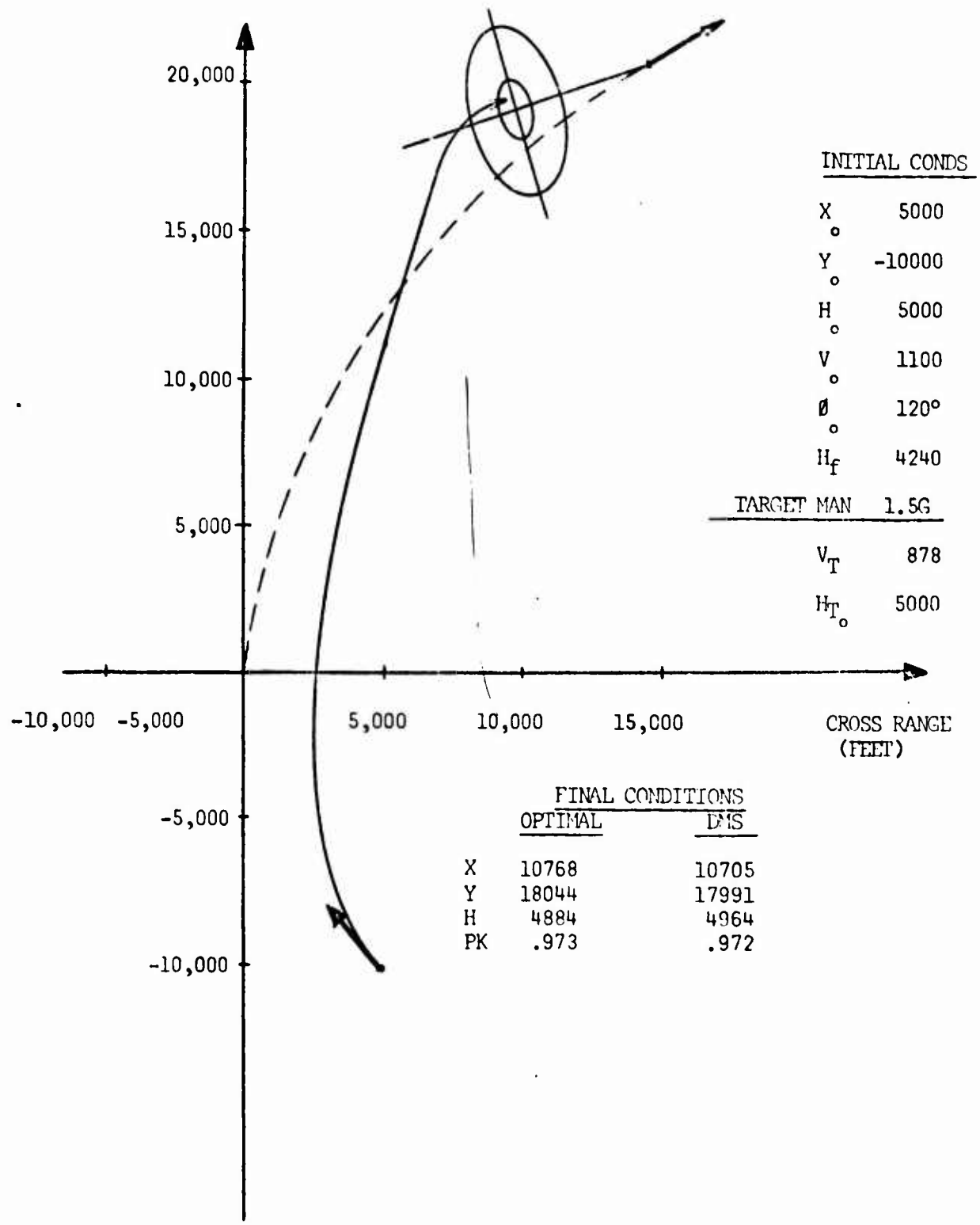


FIGURE 4. COMPARISON OF DMS AND OPTIMAL TRAJECTORIES

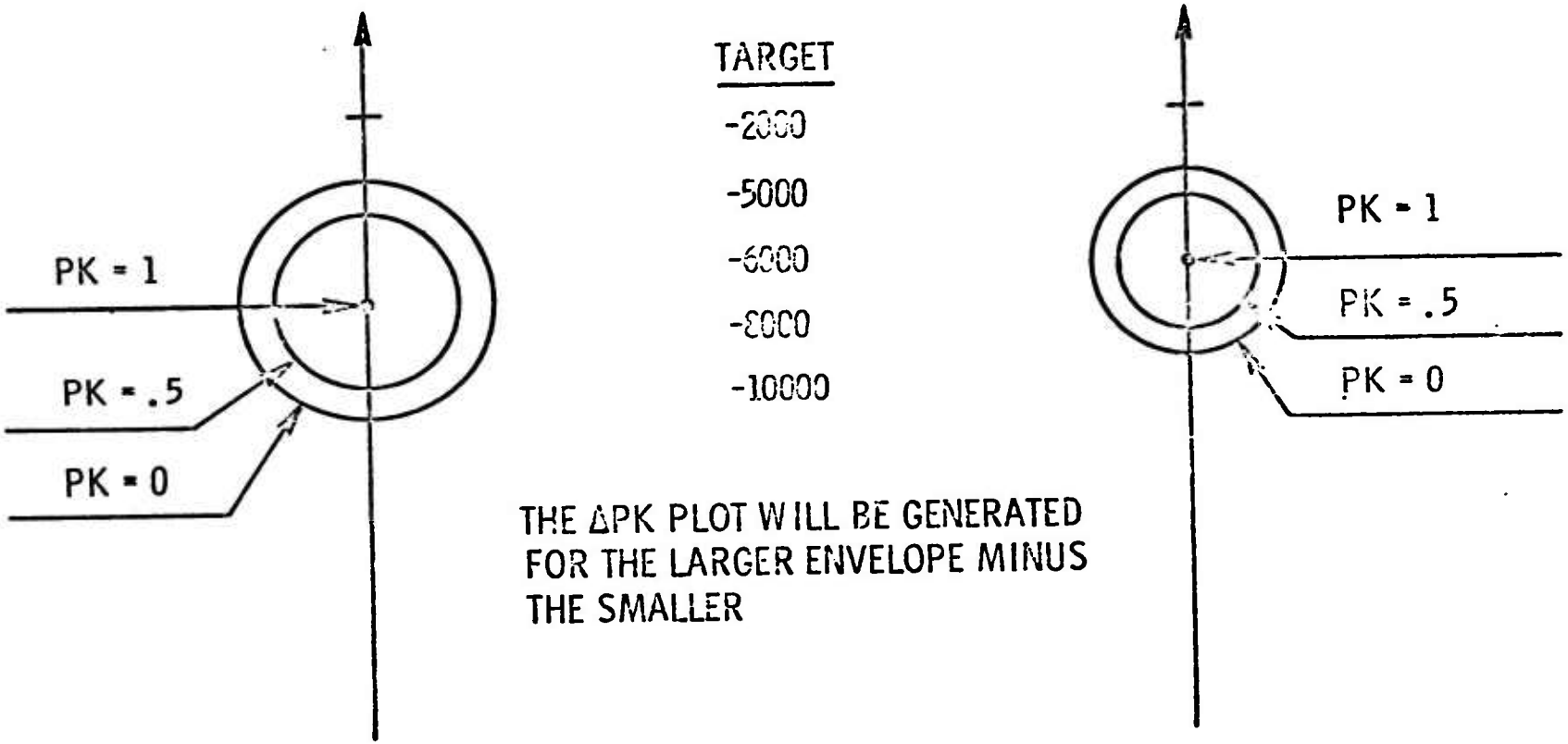


FIGURE 5a. SAMPLE ENVELOPES FOR A ΔPK COMPARISON

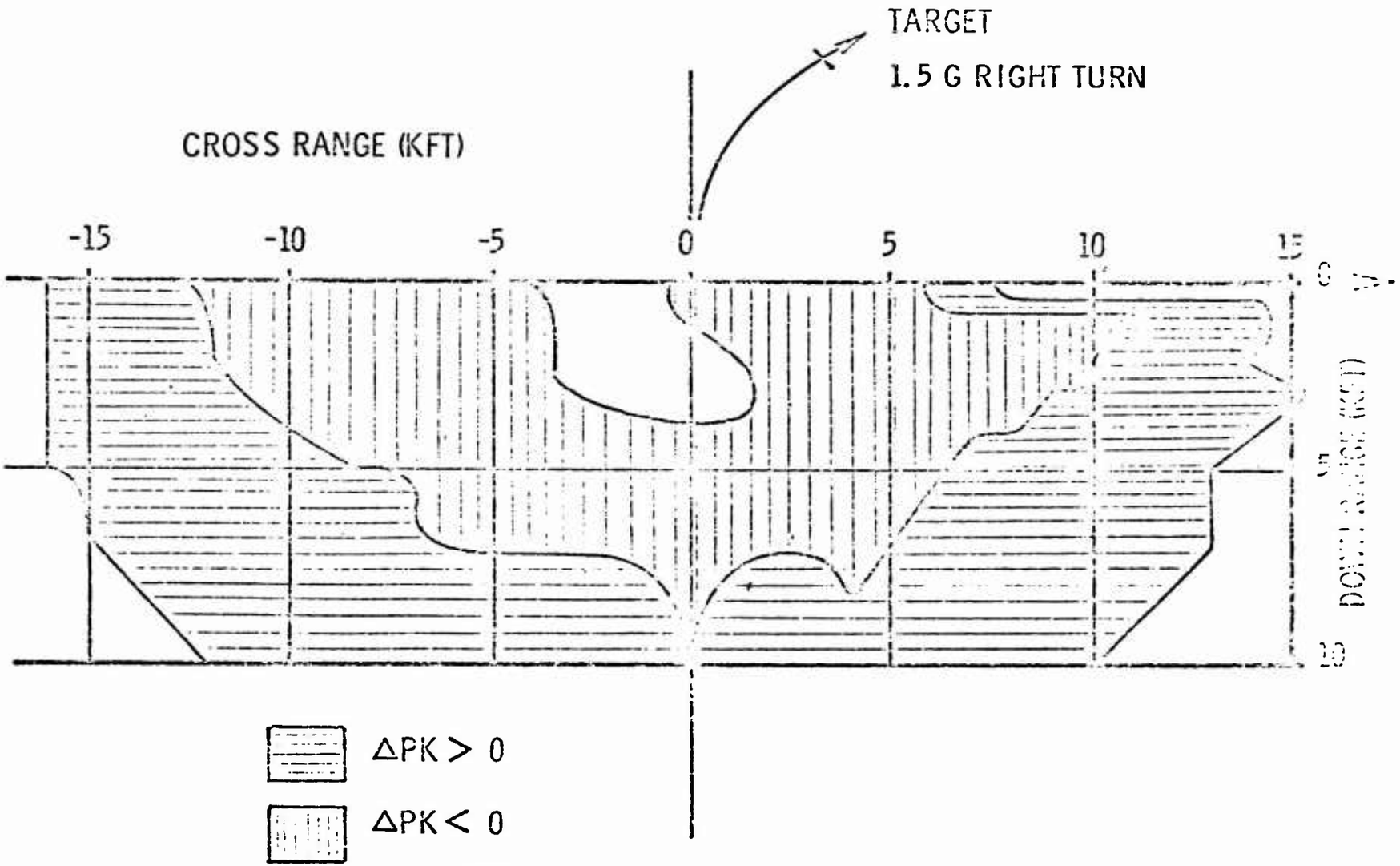


FIGURE 5b. SAMPLE ΔPK COMPARISON

candidates for such a target controller are another MASE program, and a reasonable well-known air battle simulator such as TACTICS II.

It is felt that the only realistic missile performance evaluations may be made using a countering target, so such evaluations are awaiting the programming of the target maneuver logic. When the target is programmed, one of the first applications of the MASE program is intended to be an evaluation of the benefits of an off-boresight missile capability. Another application of the MASE program is through a modification which is currently in use in the Improved Guidance Law Evaluation (reported elsewhere in this symposium) to generate feedback missile control laws. Other applications anticipated include the evaluation of thrust vector control, all-aspect capabilities, and new seeker technologies.

Title: Precision Attitude Determination for the DMSP Block 5D Satellites

Authors: Major Kenneth E. Nidiffer and 1st Lt David A. Nichols

I. SYSTEM DESCRIPTION

The Primary Attitude Determination System (PADS) for the DMSP Block 5D satellite consists of hardware and software designed to provide precise payload attitude determination with respect to a geodetic reference frame. The PADS hardware also provides attitude and acceleration inputs to the ascent guidance software for launch vehicle control.

The PADS hardware consists of two subsystems: the Inertial Measurement Unit (IMU) and the Celestial Sensor Assembly (CSA). The Primary Attitude Software (PAS) operates within the spacecraft Central Processing Unit (CPU) to process and convert IMU and CSA data for attitude determination. PAS computes the satellite attitude information on the basis of gyro and star sensor outputs. The majority of the computational logic (PAS contains less than 5000 16-bit words) deals with the processing of star sensor data. PADS uses the satellite fixed (strapdown) star sensor in a mode in which it marks the time of transit of a star crossing the sensor's field of view. To provide for continuous attitude information from intermittent star sightings, satellite rate (from gyros) is integrated. These intermittent star transits provide periodic attitude fixes and are used to estimate and correct for gyro drifts. The process is implemented with Kalman recursive filtering in the spacecraft CPU.

At the time of Block 5D conceptual design, the only practical spacecraft attitude determination systems which could satisfy DMSP pointing requirements for imagery collection were based on Space Precision Attitude Reference System (SPARS) technology. Consequently, PADS for Block 5D uses SPARS technology and will be the first operational SPAR type system to fly.

The design approach used in PADS has been to utilize a stellar augmented inertial attitude determination system evolved from SPARS. PADS uses software residing in an on-board general purpose C/MOS computer to operate on attitude and rate outputs from hardware to minimize attitude error.

PADS hardware includes three strapdown single degree of freedom gas-bearing gyros (and one skewed gyro for backup) for short term atttitude reference and one strapdown star sensor to provide the data necessary to compensate for gyro drift.

PADS is supported by ground tracking and telemetry to establish orbit parameters and requires a communications uplink for ephemeris and inertial reference data. Ephemeris and updated star catalog are uplinked daily to allow PAS to compute the local geodetic vertical PAS outputs, to the spacecraft telemetry system, significant system health data for evaluation on the ground.

II. PADS DESIGN

This section describes, in general, the PADS operation concept

used to meet DMSP performance requirements. The general requirements imposed on PADS can be outlined in six steps as follows:

(1) To utilize the outputs of a strapdown, rebalanced orthogonal set of gyros and one backup replacement, if required, (skewed to the orthogonal set) as inputs to an on-board digital computer attitude integration algorithm.

(2) To determine, using this algorithm, the satellite attitude Euler angles relating a payload fixed frame to a geoid centered rotating reference frame.

(3) To periodically obtain an independent measurement of satellite attitude with respect to inertial space from the time at which an identifiable star transited a detector slit in the GSA.

(4) At the time of this measurement, to compute in the CPU a dot product which is a measure of attitude error at the known epoch. Numerous measurements of transits compared to known stars are required to completely determine the attitude error.

(5) To solve in the CPU the software algorithms of a discrete recursive Kalman estimator, to provide a weighting or gain vector that multiplies the dot product error measurement above.

(6) To generate at accepted star transits, using the Kalman weighting vector, corrections to the instantaneous satellite Euler angle attitudes and also corrections to the IMU outputs for gyro rate bias errors.

The objective of this concept is to compute a precise estimate of payload attitude errors with respect to a coordinate frame computed from ephemeris estimates which have been provided as an input to PADS. This precise attitude error estimate is then provided to the Attitude Control System of the satellite.

III. PAS DESIGN

The PADS software consists of the Primary Attitude Software (PAS) which operates within the CPU to process IMU and CSA data and determines the precise attitude of the Primary Sensor payload relative to a geodetic local vertical reference frame.

PAS provides attitude updates for the primary sensor payload every 500 milliseconds in the orbit mode. To accomplish this, PAS utilizes an executive system to provide logical control of the program, including gyro data processing to accomplish required 2 hertz functions, and star transit processing to accomplish state corrections. The primary functions performed by PAS include:

(1) Input IMU data (gyros and star sensor) and select gyro configuration.

(2) Derive rate, compensate for drift, and compute attitude direction cosines.

(3) Compute ellipsoid earth model and transform direction cosines to geodetic frame.

(4) Process star transit data for star identification and state update.

(5) Mechanize Kalman filter utilizing star sensor data to perform state corrections.

(6) Provide those operations necessary to control program flow, and process interrupts, and interface with other software programs.

IV. IMU DESIGN

The PADS IMU senses satellite body rates, integrates the sensed rate over a fixed time interval and presents the resulting incremental attitudes to the spacecraft CPU via the Controls Interface Unit (CIU).

Body rates are sensed by miniature rate-integrating, gas-bearing, single degree of freedom, Northrup Model GI - K7G gyros. Three gyros are configured with their sensitive axes in a nominally orthogonal frame. A fourth gyro (incorporated for reliability purposes) is oriented with its sensitive axis skewed at equal angles to the sensitive axes of the other three gyros.

Each gyro is operated continuously in a dedicated analog rebalance loop. Requirements for rate sensing for ascent and orbital operations differ; therefore, loop parameters affecting range and frequency response are changed upon transfer between ascent and orbit modes in the IMU. Loop rebalance current is sampled across precision resistors, providing an accurate d.c. voltage readout proportional to sensed body rate.

Each analog rate signal is integrated with a precision electronic integrator and used by a digitizer to produce a variable frequency pulse train, each pulse of which represents the sensing of a precise value of integrated rate; thus, the pulses represent attitude increments. These attitude pulses are accumulated by the gyro channel for a 100 millisecond period and the net count is provided to the CPU.

In addition to the net count data, status information concerning the gyros and associated electronics is provided to the CPU. All output data available to the CPU is presented serially to be shifted out under control of the CPU.

V. CSA DESIGN

The Celestial Sensor Assembly (CSA) is comprised of a celestial sensor (CS), a CS Sun Shield (CS-SS), and a light tight sleeve between the CS and CS-SS. The function of the CS is to provide output pulses at a repeatable delay time, as stars transit known angular positions relative to the satellite. The CS consists of sensor optics, detector, and processing electronics. The CS-SS, which attenuates the off axis solar radiation to provide adequate star intensity, consists of a housing and two baffles. Two primary considerations in the design of the CS were:

(1) The CS output pulses must accurately define the time of the transit of each star relative to known axes of the telescope structure.

(2) The set of stars (population of stars) capable of giving rise to CS output pulses must be well defined within specified limits.

The "ideal" star sensor would have an output pulse delay time fixed (unchanging) for each star transit of a given CS slit. Furthermore, CS output pulses would be caused only by a known and prescribed star population. CS output pulse delay time from transit to transit will vary from the predicted value because of both random effects and systematic effects. The star population will appear to vary because of stochastic effects and effects of uncertainty in CS sensitivity.

The CS is a completely solid state device which utilizes a silicon photodetector array, on a single wafer, as the sensing element. A high resolution solid catadioptric optical system images the star field onto the detector array. The field of view of the star sensor is defined by a radial slit reticle (arranged in a spoke-like array) placed above the detector surface. Vehicle pitch rotation causes a star image to transverse slits in a nominally transverse direction. The six photodetector outputs are amplified and then processed to yield output pulses at a fixed delay after star transit.

Each of the six detectors has its own preamplifier-postamplifier signal processing circuit. The six detector assemblies are divided into two sets of three, each set powered separately.

The amplified detector outputs are applied to two threshold (level) circuits in the signal processing. The threshold sensor outputs are used to activate integrators and logic circuitry so as to generate

the delayed star transit pulse. Some of this circuitry is located physically within the IMU. A self-check feature is incorporated into the CSA. A light-emitting diode (LED) is located outside of the primary optical path but so as to simultaneously illuminate all six photodetectors. Thus, pulsing of the LED simulates star transit and permits verification of sensor operation.

VI. STAR SENSING AND PROCESSING

PADS requires intermittent star position information for updating and correcting the system estimates of vehicle attitude and rate gyro biases. PADS employs a vehicle-fixed star sensor (CS) in a mode that marks the time of transit of a star as it crosses the sensor field of view. As the image transits a slit, a current pulse is generated. The time of occurrence of the pulse is the fundamental measurement used to determine attitude.

The CS utilizes a silicon photodetector array as the sensing element and thus its sensitivity to particular stars is different from that of other types of sensors. Standard star catalogs such as the Smithsonian Astrophysical Observatory (SAO) Catalog give visual magnitudes. Therefore, correction factors based on the spectral characteristics of each star and the response characteristics of the sensor material are used to convert from visual magnitude to the effective magnitude as seen by the silicon detector. This correction ranges from +0.1 for class B0 (very blue) stars to -2.1 for class M9 (very red) stars.

In PADS, the CS is always operative and ready to accept star transits. No prediction is made at the time of transit as to when the next transit should occur or which star should be transited next. Thus, each transit poses a new star identification problem. This process is summarized as follows:

Based on vehicle attitude at the time of transit, a star search window is computed. This window is bounded by star right ascension values $\pm$ 7.25 deg about the vehicle pitch attitude (nominally the orbit angle). Stars in the on-board catalog with right ascension values within this window are considered potential candidates for the transiting star. No sorting is performed using star declination because the on-board catalog contains only those stars within an acceptable swath width.

For each star within the window, the dot product between the star line of sight vector and the slit normal vector, both based on estimated vehicle attitudes, is computed. This provides an estimate of how far the star is from the slit plane. If the system attitude errors are small, and if the transited star is included in the on-board catalog, or there is another cataloged star very close to the transited star, the minimum dot product might not correspond to the true star, and an attitude correction based on it could lead to an increase in system errors.

Therefore, a dot product tolerance is computed for each star in the window. This tolerance is the variance of the dot product,

which is a function of both the attitude uncertainties and measurement uncertainties. The dot products of all the stars in the window are compared with their tolerances (actually dot vs square root of tolerance), and if only one star has a dot less than tolerance, it is assumed that it is the transiting star, and its dot product is used in computing the state corrections. If no star has a dot less than tolerance, it is assumed that the transit indication is due to an uncataloged star or other cause, and the transit information is disregarded. If two or more stars have dots less than their tolerances, it is assumed that they are too close together to permit a unique identification, so again the transit information is disregarded.

Memory limitations in the on-board computer make it impossible to carry along a complete catalog of the stars that may be seen during the PADS mission. Consequently, on a nominally daily basis, positions (right ascensions, declinations) of those stars which may be transited during the next 24 to 36 hours are sorted out from a master ground catalog which includes the entire celestial sphere and uplinked to the spacecraft. The number of such stars is limited to 80. Although the CS FOV is not greater than 10.3 deg, the on-board catalog must contain stars within an equivalent swath width of 13.8 deg for steady-state and 15.8 deg for acquisition to allow for attitude uncertainties, vehicle motions and orbit precession between updates.

Because of this 80-star limitation, CS sensitivity can have a significant impact on PADS performance. Flight simulation studies

have shown that acceptable performance can be obtained with as few as 14 detected stars in the FOV swath (which results in about 35 transits during an orbit). With fewer stars, the intervals between transits can become long enough to adversely affect acquisition time as well as system errors during steady-state operations.

VII. TYPICAL MISSION SUMMARY

A typical DMSP flight mission requiring the use of the PADS system would be accomplished in the following phases. Prior to launch, PADS is precisely calibrated to measure the relationship of the attitude reference sensors and the primary sensor with respect to each other. As soon as practical after orbit injection, necessary information (such as PADS software, star catalog and accurate satellite ephemeris data) is uplinked to the computer memory. After the insertion of this data, the PADS attitude reference system begins operation in the acquisition mode, during which the system converges to steady-state. Results to date indicate that within two orbits, the attitude reference system is operating in the steady-state mode.

VIII. FUTURE IMPROVEMENTS

Future improvements to the attitude reference system are being appraised. Main areas of consideration are (a) the Third Generation Gyros (TGG) to improve the attitude accuracy and reduce the number of star updates (i.e. less star transits required to maintain a specified accuracy) and (b) a gyro package which would combine the attitude

sensing and attitude control function (i.e. the present system uses the gyro reference assembly for attitude determination and momentum wheels for stability and control). This combined precise gyro and control moment gyro is often referred to as the Attitude Sensing and Control Moment Gyro (ASCMG). The Third Generation Gyro designed by The Charles Stark Draper Laboratory, Inc. is the starting point for both of the above efforts (a) and (b). Effort (a) may remain close to the present TGG design; however, it is expected that the ASCMG design may depart considerably from the TGG by the time an acceptable product is developed.

ACKNOWLEDGEMENT: PADS was developed by Honeywell Incorporated, Aerospace Division, St. Petersburg, Florida under contract to RCA/AED, Hightstown, New Jersey, the spacecraft integrator (F04701-72-C-0221).

Title: Precision Attitude Control for the DMSP Block 5D Satellite[1]

Author: 1st Lt Peter Reischl
SAMSO/YDE
Los Angeles Air Force Station

The Defense Meteorological Satellite Program (DMSP) requires a highly accurate attitude control system to allow precision pointing of satellite meteorological sensors from a 450 n.m. circular, sun synchronous orbit. The attitude control system of the Block 5D Meteorological Satellite provides precise attitude control for the satellite which itself incorporates (1) a precision mounting platform for affixing sensors and other equipment requiring precise alignment, (2) an equipment support module enclosing the bulk of the electronics, (3) a reaction control equipment support structure which contains the spent third stage rocket motor and supports the ascent phase reaction control equipment, and (4) a rotating solar array. The overall weight of the satellite is about 1000 pounds.

The attutide control system in conjunction with the attitude determination system precisely earth orients the precision mounting platform. Each payload axis is required to be maintained within 3.6 AU[2] (3 sigma) of a geodetic local vertical frame. Maximum vehicular rates are required to be no more than 0.03°/second (3 sigma) per axis.

The pointing accuracy is influenced by several error sources including satellite ephemeris, attitude determination, attitude control and structual alignment. While the attitude determination system is based on Space Precision Attitude Reference System (SPARS) technology, the satellite attitude control system is of much more conventional design.

(1) Control system is being developed by RCA under Air Force Contract F 04701-72-C-0221.

(2) AU stands for Attitude Units and is used to keep this report unclassified.

Block 5D satellite attitude control is provided by a highly reliable closed-loop subsystem utilizing satellite computers, momentum exchange components such as a Momentum Wheels Assembly (MWA), and magnetic torquing coils. The errors due to attitude control have been specified to be no more than 12 arc-seconds per axis.

Attitude Control

Figure 1 depicts a functional block diagram of the attitude control subsystems, together with its main interfaces. Attitude errors based

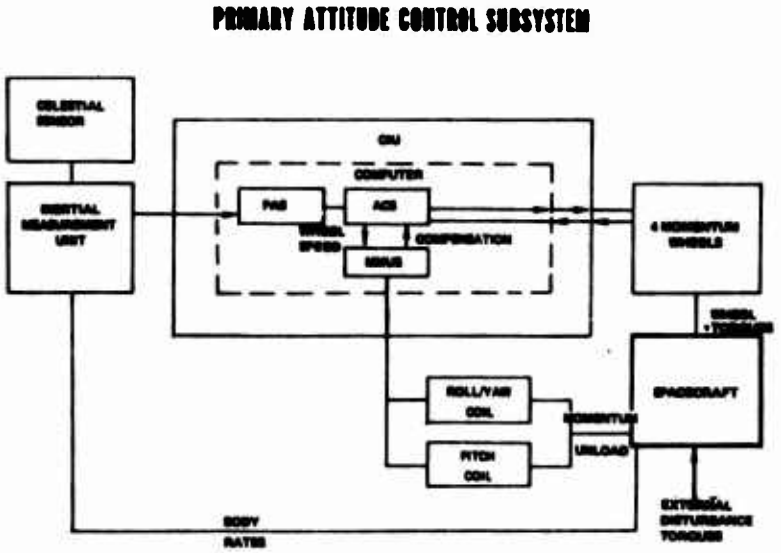


Figure 1.

on payload Euler angles are derived from gyro, celestial sensor, and ephemeris data computed with the Primary Attitude Software (PAS) of an on-board computer. This 16 bit computer has a 16 K memory, a set of 52 microinstructions, and a cycle time of about 2.5 micro-seconds. It consumes 4.2 watts of power and weighs 7 pounds.

The computed attitude errors command via the Attitude Control Software (ACS) three orthogonal momentum wheels to provide control torques for satellite stabilization and pointing. In the event of a single wheel failure, a fourth skewed wheel is provided for back-up. Each momentum wheel assembly weighs approximately 8 pounds and is

driven by a brushless DC motor. Individual wheel maximum momentum storage capacity is about 50 in lb-sec. Each wheel torquing command consists of an 8 bit word plus sign and has a resolution dither dead band of no more than 2.5 arc-seconds.

A Control Interface Unit (CIU) converts the digital torquing commands to equivalent voltage levels for input to the momentum wheel motor drive circuitry. Control updates are performed at fixed intervals of 0.5 second. The internally developed control torque of the commanded momentum wheel reacts with the spacecraft structure to correct the satellite's attitude about an axis parallel to the wheel spin axis.

Wheel speed and sense information are converted by the control interface unit into an 11-bit digital format for tramsission back to the attitude control software. In this software, the digital data is reconverted to wheel speeds for generation of individual wheel desaturation commands through the Magnetic Momentum Unloading Software (MMUS).

External disturbance torques and the rotating solar array cause an accumulation of angular momentum in the three operating reaction wheels. The accumulated momentum is dissipated via momentum unloading. Momentum unloading torques are developed at specified orbital regions through interaction of the earth's magnetic field with on-board magnetic dipoles. There are two redundant air core aluminum wire coils aboard the spacecraft which deliver nominal dipole magnitudes of 62.8 and 24.4 amp-turn-m^2 for roll/yaw and pitch, respectively.

Control Laws

For the derivation of the control laws, the 5D spacecraft was idealized to consist of a main rigid body with a large flexible appendage representing the solar array (Fig. 2).

DYNAMIC MODEL FOR SYSTEM SIMULATION (IDEALIZED 5D SPACECRAFT)

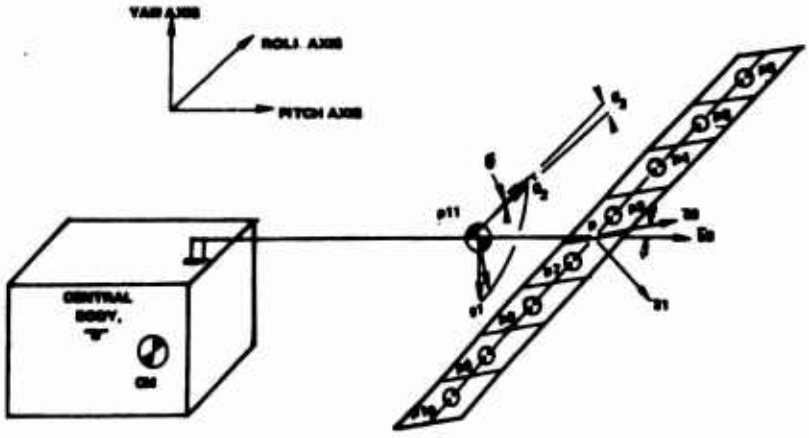


Figure 2.

The solar array was approximated by a frame structure with a discrete mass distribution. The inertia property of the eight panels of the array were modeled by the eight mass particles located at the center of each panel. The mass of the array drive motor is shown centered at P_{11}. The mass center of the entire system was assumed fixed relative to the central body B.

A basic prerequisite for the rational design of an attitude control system is a knowledge of the disturbance torques in the spacecraft environment. In the on-orbit phase, the 5D spacecraft is subjected to the following external disturbances:

a. Gravity gradient
b. Solar radiation pressure
c. Earth's magnetic field
d. Aerodynamic pressure

The torques due to residual magnetic dipoles and aerodynamic forces at a 450 n.m. orbit were considered negligible in comparison to the gravity-gradient and solar radiation pressure torques.

Disturbance torques internal to the spacecraft include:

a. Solar Array Drive (SAD) Cogging Torque
b. Momentum Wheel Assembly (MWA) Commutation Ripple
c. Three Tape Recorders (1 second on or off torque ramps)
d. Oscillating Optical Sensor at 6 Hz about roll axis; 1 percent along yaw and pitch axis
e. Other scanning sensors

The servo design of the four control loops did take the disturbance torques into consideration, as well as disturbances due to magnetic unloading and gyroscopic cross coupling.

The roll, pitch and yaw attitude loops were designed to have common segments for redundancy switching. All gyro loops and MWA's are therefore identical. A digital rolloff filter having a 6 db bandpass at approximately 2.87 Hz and sampled at 10 Hz is common to all three axes.

To simplify the attitude control design task, all equations were fully linearized with respect to a null nominal state. The resulting set of linear differential equations with constant coefficients were utilized to construct transfer functions.

The equations were changed from the time domain into the frequency domain via the Laplace transformation. The equations were manipulated and compressed into a system of four equations containing only the four attitude variables of primary interest, i.e., the body yaw, roll, pitch angles, and the relative angle between the array drive motor rotor and the housing.

To investigate the interaction between the attitude control loops caused by the highly cross-coupled nature of the spacecraft inertial and structural characteristics, the dynamic transfer function for any of the four control loops was constructed with the other three attitude control loops closed and included in the formulation. The control system characteristics were adjusted via Bode and Nyquist analysis. The Nyquist diagrams provided conclusive information concerning the spacecraft attitude stability margin.

The highly asymmetrical flexible structure imposed bandwidth limitations for stability. Modal frequencies and damping considerations constrained the achievable bandwidth.

The 0.5 second computer control updates assist in attenuating the destabilizing effects of structural frequencies, and are adequate in achieving high enough loop bandwidth to counter disturbances such as those due to tape recorder turn on transients.

Figure 3 shows the array open loop Nyquist plot with all other loops closed.

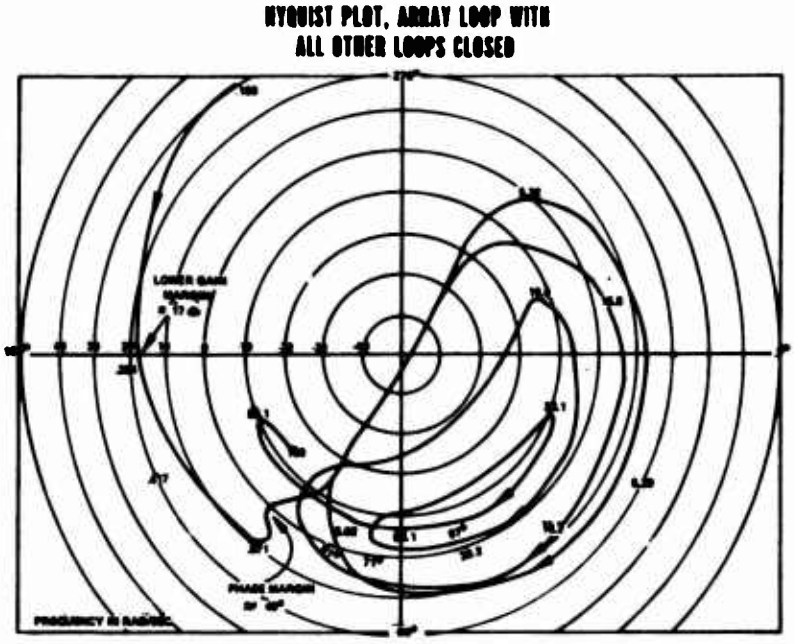


Figure 3.

The roll and yaw loop coupling effects are small but the pitch loop exerts considerable influence due to its direct coupling to the Solar Array Drive (SAD) loop. The phase margin for the Solar Array Drive (SAD) loop is 49 degrees and the lower gain margin is 17 db. The usual specifications for an acceptable design are a minimim gain margin of 6 db and a minimum phase margin of 30 degrees (Greensite, 1970).

It was not possible to design the loop gain stable and simultaneously obtain sufficient attenuation of the SAD cogging torque which manifests itself at 96 times orbital frequency. Since saturation in the array loop could cause instability, this difficulty was thoroughly studied and simulated and found to be no problem.

The pitch loop design was dependent upon the array loop parameters. Furthermore, the pitch inertia is the smallest and hence more disturbance sensitive. Consequently, compensation was added to obtain a type-3 loop with zero-error response from a steady state to a parabolic external torque input, which provided greater attenuation of the SAD motor cogging torque. The pitch loop upper and lower gain margins are 11 and 12 db, respectively. The phase margin is 49 degrees.

The roll and yaw loop compensation are identical since both axes have almost equal inertias. The major design considerations were to obtain high bandwidth to counteract tape recorder and magnetic unloading disturbance torques and obtain sufficient attenuation at higher frequencies to avoid structural modes. A high gain at low frequency is needed to minimize attitude error due to slowly varying torques (magnetic unloading, gyroscopic coupling, solar pressure). This resulted in a conditionally stable Type-2 roll and yaw loop. The upper and lower gain margins are 17 and 19 db, respectively. The phase margin is 60 degrees.

Magnetic Torquing Compensation

The accumulation of angular momentum within the spacecraft due to solar pressure, gravity gradient, and cogging torques from the rotating solar array would eventually drive the momentum wheels into saturation were it not for the magnetic momentum unloading subsystem. The roll/yaw torquing coil and pitch torquing coil provide, periodically, the necessary magnetic torques to unload angular momentum in all three operating momentum wheels. However, magnetic torquing unloading has the same effect as an abrupt error torque applied externally. This apparent error is reduced by modeling

the magnetic torque and programming the momentum wheel assembly to produce an essentially torque-balancing effect on the body. This is accomplished by proper adjustment of gains via software such that a constant output torque is superimposed on the regular control signal. The value of the compensation for a given spacecraft axis is a function of the earth's magnetic field and the dipole moment of the coil. The small variability due to the earth's magnetic field change over the torquing zone is expected to produce an acceptable small transient error.

Simulation Studies

To verify attitude control subsystem design and transient response performance, extensive digital computer simulation studies of the complete mission-mode spacecraft dynamic system were undertaken. In addition to the attitude determination algorithms, the computer program included provisions for simulating the system's transient response to internal and external disturbances, e.g., momentum unloading by magnetics. Random noise characteristics of the gyros were simulated by the separate addition of pseudo-random numbers to the computed outputs. In the simulation perfect attitude determination was assumed.

It was demonstrated that the system's peak jitter produced by the oscillating optical sensor about the roll, pitch, and yaw axes are 0.0375, 0.00431, and 0.002 degrees per second, respectively. These jitter rates are larger than those due to any other disturbance sources by at least one order of magnitude.

The effect of the solar array drive's cogging torque of 3 inch-ounces was shown to influence pitch axis performance only. The resulting pitch axis RMS attitude error is 5.21 arc-seconds, corresponding to a peak value of 7.37 arc-seconds.

Table 1 presents a detailed breakdown of the predicted Root Sum Square (RSS) peak errors due to yaw, roll, and pitch attitude control,

respectively. As shown, the predicted total errors attributable to attitude control are close to the 3 sigma error budget of 12 arc-seconds per axis. Analytical results are scheduled to be verified soon with the beginning of both three-axis and single-axis air-bearing table tests. These tests will also verify the interaction of the Primary Attitude Determination System (PADS) with the attitude control system.

Table 1. Predicted Attitude Control System Error

Error Source	Worst-Case Peak Error (Arc-Sec)	Frequency of Occurance	Remarks
1. Oscillating Optical Sensor	1.30 yaw 7.36 roll 2.69 pitch	Continuous " "	
2. Gyro Noise	0.34 yaw 1.39 roll 2.29 pitch	Continuous " "	
3. Tape Recorder Playback	3.75 yaw 6.83 roll 3.26 pitch	Random " "	 Coupled from roll/yaw
4. SAD Cogging Torque	0.90 yaw 1.40 roll 7.34 pitch	Continuous " "	Coupled from pitch Coupled from pitch
5. 1.5% SAD Rate Change	0.67 yaw 1.37 roll 3.29 pitch	Random " "	Coupled from pitch Coupled from pitch
6. RWA Cogging Torque	Negligible Negligible Negligible	Continuous " "	0.25 in-oz peak amp 0.25 in-oz peak amp
7. Secular Environmental Disturbance Torque	Negligible 1.70 roll pitch	Continuous Continuous Continuous	$0.66+10^{-4}$ in-lb (max) $2.77+10^{-3}$ in-lb (max) Type III Servo
8. Wheel Speed Reversal	3.50 yaw 2.39 roll pitch	2 per orbit 2 per orbit	 **
9. Roll/yaw Desaturation	7.09 yaw 3.44 roll 3.10 pitch	2 per orbit(max)** 2 per orbit(max) 2 per orbit(max) **	 Coupled from roll/yaw
10. Pitch (Roll) Desaturation	6.68 yaw	2 per orbit(max)	

**Cannot occur simultaneously

Summary:

Error Sources	(1-9) yaw	= 8.92 arc-sec (RSS)
Error Sources	(1-8) roll	= 11.00 arc-sec (RSS)
	(1-7+9) roll	= 11.28 arc-sec (RSS)
Error Sources	(1-9) pitch	= 10.27 arc-sec (RSS)
	(1-8+10) pitch	= 11.86 arc-sec (RSS)

References:

Greensite, Arthur L., Elements of Modern Control Theory, Spartan Books, New York, 1970.

Application of Estimation Theory

to an

AF Pointing and Tracking System

Lt J.E. Negro

INTRODUCTION

An application of modern estimation theory in an Air Force pointing and tracking system to provide control system augmentation signals is demonstrated to result in system performance that is significantly improved over that otherwise obtainable. While indeed estimation theory has been used with great success in other Air Force programs, notably, satellite orbit determination, missile trajectory measurement, target identification, aircraft parameter and system identification, and inertial navigation - this application is unique in its real-time, high data-rate requirements, a lack of accurate state extrapolation models, and a multiplicity of widely differing alternative solutions. Also, it differs from other tracking estimation problems in that information from only a single range-angle sensor is assumed available rather than from several spatially dispersed range-angle sensors as common in other applications.

The remainder of the paper is devoted to a general outline of the precision pointing and tracking problem, presentation of two tracking filter estimation algorithms and a discussion of their salient features.

PRECISION POINTING SYSTEMS

In general our objective is to accurately orientate the inertially stabilized boresight of an airborne optical system such as required for precisely aiming an optical telescope or aligning the transmitter receiver pair of an optical communication link. Pointing requirements for these narrow field of view optical applications are more exacting than corresponding radio-frequency applications. Furthermore, the mechanical optical steering mechanisms are more difficult to implement and susceptible to external disturbances than are phased-array RF steering techniques.

Pointing error sources may be catalogued into two broad areas - stabilization errors and track loop errors. Stabilization errors arise from the inability of the gyro-gimbal actuator loops to perfectly isolate the pointing line-of-sight (LOS) from disturbance phenomena. Disturbances include gyro noises and biases, gimbal kinematic torques, and for airborne pointing systems - base motion and aerodynamic windload torques. Minimization of these pointing error sources is

attained by proper gimbal configuration design as described in References 1, 2, 3 and 4.

Track loop errors are defined as those errors other than stabilization error. Dominant track loop error sources are track error sensor noise and dynamic following errors characteristic of low bandwidth control loops. These error sources are complementary in the sense that optimum track loop bandwidth (gain) is basically selected as a compromise between high loop gain dynamic following capability, on one hand, and minimum bandwidth error sensor noise transmission on the other hand. The basic estimation algorithms subsequently considered essentially provide a track loop feedforward compensation signal which obviates the loop gain compromise.

Consider the four gimbal pointing system depicted in Fig. 1 with gimbal orientation Euler angles described in Fig. 2. The four gimbals provide the degree of freedom necessary for inner gimbal base motion isolation and allow a piggy-back inner-outer gimbal arrangement having high bandwidth track loops afforded by the linear actuation possible with the limited travel inner gimbal set while maintaining full angular pointing coverage with the outer follow-up gimbals. A simplified single axis block diagram is given in Fig. 3. Fig. 4 represents a further simplified block diagram gimbal model that

provides a pointer model basis for the remainder of this paper. Observe the Type II servo loop, and hence note that a steady-state pointing error exists when following constant angular acceleration motion. Also observe in Fig. 5 that receiver motions benign in inertial space present considerable dynamics as expressed in LOS pointer coordinates. LOS angle and angle derivatives are plotted for constant inertial velocity receiver motion.

The induced dynamic following errors are approximated by

$$e(t) \simeq \frac{1}{K} \left[\ddot{\theta}_R(t) - \frac{1}{\omega_a} \dddot{\theta}_R(t) + \ldots \right] \tag{1}$$

The linear constant velocity receiver motion has a corresponding second derivative maximum value

$$|\ddot{\theta}_R|_{MAX} = \frac{3\sqrt{3}}{8} (v/R_o)^2 \tag{2}$$

V = receiver velocity

R_o = crossing range

Therefore

$$e_{MAX} \simeq \frac{1}{K} \frac{3\sqrt{3}}{8} (v/R_o)^2 \tag{3}$$

which may be unacceptably large in some applications. We, therefore, ask what control system modification might be implemented to achieve the desired accuracy.

CONTROL LAW SELECTION

A suitable control law or alternate loop compensation giving desired pointing accuracy is sought. First of all, the stabilization loop must remain unaltered so as to retain its disturbance rejection properties. Secondly, we intuitively note that the desired command for the ideal (i.e. infinite bandwidth) stabilization loop is simply the LOS target angular rate. That is, zero pointing error is maintained if the pointer is slewed at precisely the target angular rate. This same control law is also derived as the solution to the problem of finding a control which when summed with the compensated track error signal will give zero steady-state pointing error for constant angularly accelerating input motions. This control law provides acceptable pointing error; unfortunately it cannot be implemented by classical techniques since no measurement of $\dot{\theta}_R(t)$ is available. Estimation theory techniques, however, can be applied to available measurements to provide an estimate of $\dot{\theta}_R$. A zero estimated error control law has also been proposed [5]. It too, requires the estimation of unknown quantities.

ESTIMATION OF $\dot{\theta}_R$

A proliferation of tracking filters have been proposed for the tracking problem and include α-β, α-ρ-σ, Kalman, Kalman with polynomial dynamics, extended Kalman and iterated extended Kalman filters. An excellent description and comparative analysis of these filters is developed by Wishner (6) and more briefly by Baugh (7). Treatments of tracking filter estimation techniques directly applicable to the precision pointing and tracking problem are given in References 5 and 8-11. In the following, the $\dot{\theta}_R$ estimation problem is posed mathematically for the single axis, two dimensional case. Two solutions are presented and discussed.

Consider the gimbal-receiver geometry of Fig. 6. Estimates of $\dot{\theta}_R$ are to be optimally derived from noisy measurements of r, θ_R and the pointing system state variables. Two solution techniques are presented. One is a linear Kalman filter estimate of cartesian coordinate model of receiver dynamic motion states from which the control is derived. The second is a non-linear extended Kalman filter estimate of the cylindrical coordinate model of receiver dynamic motion states from which the control is directly obtained.

CARTESIAN COORDINATE ESTIMATION

With reference to Fig. 6, we describe the relative motion dynamics between the receiver-target and the gimbal platform as

$$\begin{aligned} \ddot{x} &= a_x \\ \ddot{y} &= a_y \end{aligned} \tag{4}$$

where a_x, and a_y represent the linear relative motion accelerations along each axis. Knowledge of a_x and a_y together with appropriate initial conditions completely determine x(t) and y(t). In practice a_x and a_y are unknown functions of x, y and t. For lack of any other knowledge, we assume these variables to be represented as random walk processes, and define a state variable model

$$\underset{\sim}{x} = [x \;\; \dot{x} \;\; a_x \;\; y \;\; \dot{y} \;\; a_y]^T \tag{5}$$

Thus

$$\dot{\underset{\sim}{x}} = \begin{bmatrix} \underset{\sim}{A} & \underset{\sim}{o} \\ \underset{\sim}{o} & \underset{\sim}{A} \end{bmatrix} \underset{\sim}{x} + \underset{\sim}{G}\,\underset{\sim}{\omega} \tag{6}$$

where

$$\underset{\sim}{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \underset{\sim}{G} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and where $\underset{\sim}{\omega}$ represents the white noise random walk forcing function.

Standard Kalman filter equations for the cartesian co-ordinate model are straightforward provided that we can obtain linear measurements of the states. Such measurements are not directly available but can be approximated by suitable manipulations of the measurements, as proposed by Fitts [10]. In particular, for $\Theta_p \simeq \Theta_R$ we have

$$\begin{aligned} x_m &\simeq \hat{r} \cos \Theta_p \\ y_m &\simeq \hat{r} \sin \Theta_p \end{aligned} \tag{7}$$

where $\hat{r} = \sqrt{\hat{x}^2 + \hat{y}^2}$

and where the caret ($\hat{}$) denotes estimated quantities and the subscript

'm' denotes measured (or pseudo measurement) quantities. Similarly

$$\dot{x}_m \simeq \hat{\dot{r}} \cos\theta_p - \hat{r}\,\dot{\theta}_p \sin\theta_p$$

$$\dot{y}_m \simeq \hat{\dot{r}} \sin\theta_p + \hat{r}\,\dot{\theta}_p \cos\theta_p \qquad (8)$$

$$\hat{\dot{r}} = (\hat{x}\,\hat{\dot{x}} + \hat{y}\,\hat{\dot{y}})/\hat{r}$$

and where $\omega_c \simeq \dot{\theta}_p$ is an available pointing system signal. Therefore, for each axis we have generated a set of pseudomeasurements x_m and $\dot{x}_m$ derived from pointer variables θ_p and $\dot{\theta}_p$ and estimated quantities $\hat{r}$ and $\hat{\dot{r}}$. Note that the state propagation equations and the observed measurements (with respect to the pseudo-measurements) are uncoupled between axes. Thus the Kalman equations also decouple into two identical equation sets. Note that an actual range measurement r_m should be used, when available, in the pseudomeasurements.

The discretized standard Kalman equation sets for each axis are

$$\underset{\sim}{x} = [x\ \dot{x}\ a_x]^T$$

$$\hat{\underset{\sim}{x}}_{k+1}(-) = \underset{\sim}{\Phi}\,\hat{\underset{\sim}{x}}_k(+)$$

$$\underset{\sim}{z}_k = \underset{\sim}{H}\,\underset{\sim}{x}_k + \underset{\sim}{v}_k$$

$$\hat{\underset{\sim}{x}}_k(+) = \hat{\underset{\sim}{x}}_k(-) + \underset{\sim}{K}_k(\underset{\sim}{z}_k - \underset{\sim}{H}\,\hat{\underset{\sim}{x}}_k(-))$$

$$\underset{\sim}{P}_k(-) = \underset{\sim}{\Phi}\,\underset{\sim}{P}_{k-1}(+)\,\underset{\sim}{\Phi}^T + \underset{\sim}{Q}$$

$$\underset{\sim}{P}_k(+) = (\underset{\sim}{I} - \underset{\sim}{K}_k\,\underset{\sim}{H})\,\underset{\sim}{P}_k(-)$$

$$\underset{\sim}{K}_k = \underset{\sim}{P}_k(-)\,\underset{\sim}{H}^T[\underset{\sim}{H}\,\underset{\sim}{P}_k(-)\,\underset{\sim}{H}^T + \underset{\sim}{R}]^{-1}$$

for which $\underset{\sim}{R} = E[\underset{\sim}{v}\underset{\sim}{v}^T]$, $\underset{\sim}{Q} = E[\underset{\sim}{\omega}\underset{\sim}{\omega}^T]$ and $\underset{\sim}{P}_o$ are the measurement noise covariance, plant noise covariance and initial plant state estimate covariance. $\underset{\sim}{P}_k$ denotes the estimation error covariance and $\underset{\sim}{K}_k$ denotes the Kalman gain matrix at the k^{th} process iteration.

Also

$$\underset{\sim}{\Phi} = \begin{bmatrix} 1 & \Delta T & \Delta T/2 \\ 0 & 1 & \Delta T \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underset{\sim}{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Pseudo-measurements are computed with equations 7 and 8. The desired pointer control $\dot{\theta}_R$ is computed from estimated quantities by sol ing the following equation for $\dot{\theta}_R$.

$$\dot{\hat{y}} = \frac{d}{dt}(\hat{r} \sin \hat{\theta}_R) = \dot{\hat{r}} \sin \hat{\theta}_R + \hat{r} \cos \hat{\theta}_R \dot{\hat{\theta}}_R$$

or

$$\dot{\hat{\theta}}_R = (\dot{\hat{y}} - \dot{\hat{r}} \sin \hat{\theta}_R)[\hat{r} \cos \hat{\theta}_R]^{-1}$$

DISCUSSION

The cartesian coordinate method for obtaining $\dot{\theta}_R$ estimates for use in pointing system control augmentation has been found to yield a 40 db pointing error reduction as shown in Fig. 7 for modelled receiver dynamics. Other advantages of this mechanization are that the gains can be precomputed and the same gains stored and used for each axis. Also, the linear state extrapolation equations are readily and accurately machine computed. Major objections to this estimation technique are the measurement basis and incorrect noise covariances induced by the pseudo-measurement technique. Another consequence of the pseudo-measurement technique is that at initialization, the Kalman gains are 'large' and that the filter must rely upon the pointing precision of the normal track loop to provide good measurements while later when the gains have reached their smaller steady-state values the pseudomeasurement are more likely to be in error.

POLAR COORDINATE ESTIMATION

The cartesian coordinate model previously used is equivalently expressed in polar coordinates by the dynamic equations.

$$\dot{r} = r\,\dot{\theta}^2 + a_r$$

$$\ddot{\theta} = -\frac{2\dot{r}\dot{\theta}}{r} + \frac{a_\theta}{r} \tag{11}$$

in which r and θ represent the independent polar coordinate variables while a_r and a_θ represent the receiver relative acceleration along the r and e axes respectively. As before, we use a random walk model for these quantities. The polar coordinate choice of variables allows direct estimation of the desired quantity $\dot{\theta}_R$ with available linear measurements at the cost of having to perform a non-linear estimation.

Measurements of range r_m are available as before and measurements of line of sight angle to the receiver are formulated from

$$\theta_{R_m} = \theta_p + \varepsilon \tag{12}$$

in which ε represents the pointing error measured by the tracking error

sensor.

The state vector x is defined as

$$\underset{\sim}{x} = [r\ \dot{r}\ a_r\ \Theta\ \dot{\Theta}\ a_\Theta]^T \tag{13}$$

The model dynamics are

$$\dot{\underset{\sim}{x}} = \underset{\sim}{f}(\underset{\sim}{x}) + \underset{\sim}{G}\ \underset{\sim}{\omega}$$

$$= \begin{bmatrix} x_2 \\ x_1 x_4^2 + x_3 \\ 0 \\ x_4 \\ \dfrac{-2x_2x_4 + x_6}{x_1} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \tag{14}$$

The discrete extended Kalman equations are

$$\hat{\underset{\sim}{x}}_{k+1}(-) = [\underset{\sim}{I} + \underset{\sim}{F}(\hat{\underset{\sim}{x}}_k(+))\ \Delta T]\ \hat{\underset{\sim}{x}}_k(+)$$

$$\underset{\sim}{z}_k = \underset{\sim}{H}\ \underset{\sim}{x}_k + \underset{\sim}{v}_k$$

$$\hat{\underset{\sim}{x}}_k(+) = \hat{\underset{\sim}{x}}_k(-) + \underset{\sim}{K}_k(\underset{\sim}{z}_k - \underset{\sim}{H}\ \hat{\underset{\sim}{x}}_k(-)) \qquad (15)$$

$$\underset{\sim}{P}_k(-) = [\underset{\sim}{I} + \underset{\sim}{F}(\hat{\underset{\sim}{x}}_k(+))\ \Delta T]\ \underset{\sim}{P}_k(+)[\underset{\sim}{I} + \underset{\sim}{F}(\hat{\underset{\sim}{x}}_k(+))\ \Delta T]^T + \underset{\sim}{Q}$$

$$\underset{\sim}{P}_k(+) = (\underset{\sim}{I} - \underset{\sim}{K}_k\ \underset{\sim}{H})\ \underset{\sim}{P}_k(-)$$

$$\underset{\sim}{K}_k = \underset{\sim}{P}_k(-)\ \underset{\sim}{H}^T[\underset{\sim}{H}\ \underset{\sim}{P}_k(-)\ \underset{\sim}{H}^T + \underset{\sim}{R}]^{-1}$$

$\underset{\sim}{R}$, $\underset{\sim}{Q}$, $\underset{\sim}{P}_o$, $\underset{\sim}{P}_k$ and $\underset{\sim}{K}_k$ remain as previously defined.

Also

$$\underset{\sim}{F}(\underset{\sim}{x}_k) = \frac{\sigma}{\sigma \underset{\sim}{x}_k}\ \underset{\sim}{f}(\underset{\sim}{x}_k) \qquad (16)$$

$$\underset{\sim}{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

The desired control is simply the estimated state x_4.

DISCUSSION

Our two models are identical with the exception that the random walk acceleration models differ in that the direction of the polar coordinates accelerations vary as a function of position. Thus, assuming each estimation technique equally accurate, we have in effect obtained the advantage of direct use of available measurements and direct estimation of $\dot{\theta}_R$ at the cost of the additional computational complexity of the extended Kalman filter. This filter also yields a 40 db pointing error reduction when following well-modelled receiver dynamics as shown by the pointing error history for the crossing target motion is presented in Fig. 7.

EXTENTIONS AND COMMENTS

The straightforward adaptation of these techniques to the 3-dimensional 2-axis case are presented in References 9 and 10. Spherical coordinates replace polar coordinate models. An inertial frame reference generated either by an inertial platform or a strapdown equation update of the pointer LOS is required to relate the pointing system boresight coordinates to the non-rotating inertial receiver dynamic model frame.

Other practical considerations include: 1) filter tuning (by adjustment of noise characterizing parameters) to maximize the performance tradeoff between receiver dynamic model uncertainty on one hand and sensor noise on the other, 2) inclusion of the tracking filter within the track loop to reduce pointing error sensor noise transmission through the loop and 3) adaptive filter tuning to reduce the tracking filter sensitivity to unmodelled target dynamics. These areas remain the subject of continuing research.

References

1. A.K. Rue, 'Stabilization of precision electro-optical pointing and tracking systems,' IEEE Transactions on Aerospace and Electronic Systems, Vol AES-5, pp. 805-819, Sept, 1969.

2. G.C. Comfort, "Base motion isolation of a 2-axis beam deflector" SRL-TR-72-0008, F.J. Seiler Research Laboratory, USAF Academy, Co. 80840.

3. L. Sher, 'Stabilization requirements for airborne precision pointing systems,' Proceedings of the JACC, pp. 788-789, June, 1973.

4. A.K. Rue, 'Precision stabilization systems,' Proceedings of the JACC, June, 1973.

5. R.T. Curran et al. 'Fundamental Problems of Precision Acquisition Pointing and Tracking Systems,' AFOSR Final Technical Report, November, 1973.

6. R.P. Wishner, et al. 'Status of radar tracking algorithms,' Proceedings of Symposium on Nonlinear Estimation Theory and its Applications, pp. 32-53, Sept, 1970.

7. R.A. Baugh, Computer Control of Modern Radars, RCA Press, 1973.

8. R.B. Asher and D.H. Watjen, Kalman Filtering for Precision Pointing and Tracking Applications,' AFAL Technical Report (to be published).

9. J.M. Fitts, "Aided tracking as applied to high accuracy pointing systems," IEEE Transactions on Aerospace and Electronic Systems, Vol AES-9, No. 3, pp. 350-368, May, 1973.

10. C.F. Price, "A three-dimensional pointing and tracking algorithm," Third Quarterly Progress Report, Contract F29601-73-C-0005, Air Force Weapons Laboratory, June, 1973.

11. J.M. Revoire, Precision Tracking, TG1237, The Johns Hopkins University, Applied Physics Laboratory, March, 1974.

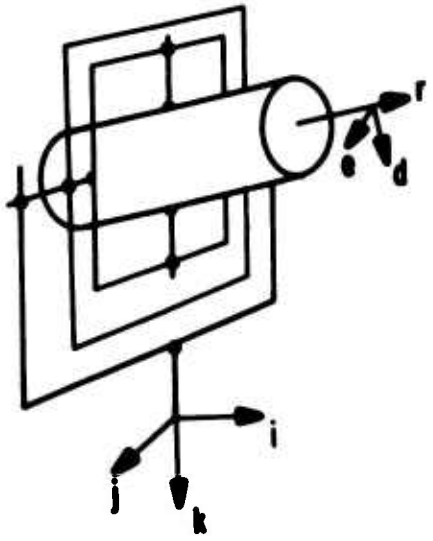


Fig. 1. A Four Gimbal Pointing System

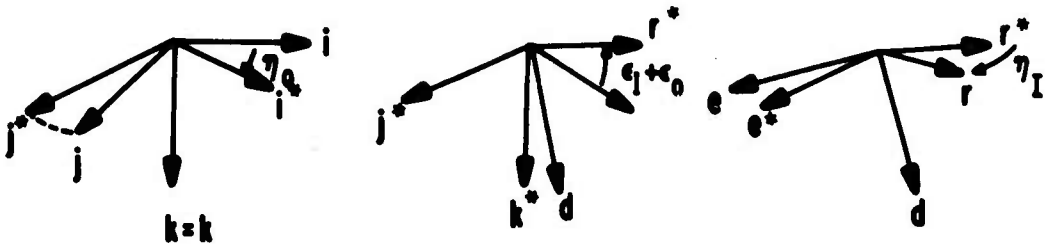


Fig. 2a. Gimbal Coordinate Transforms

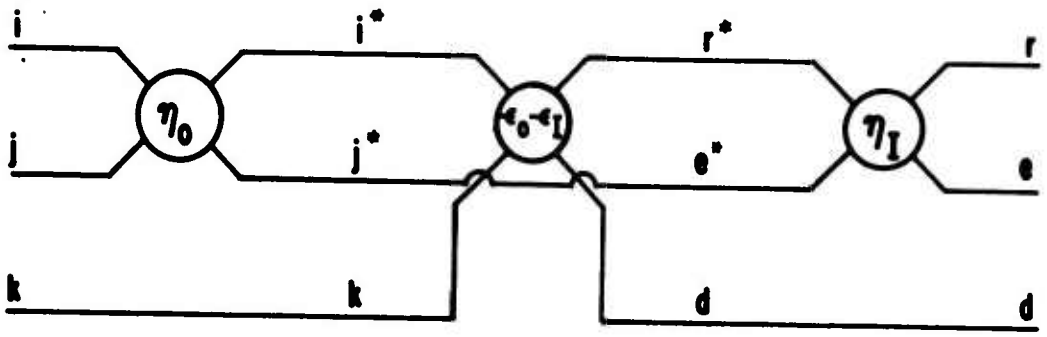
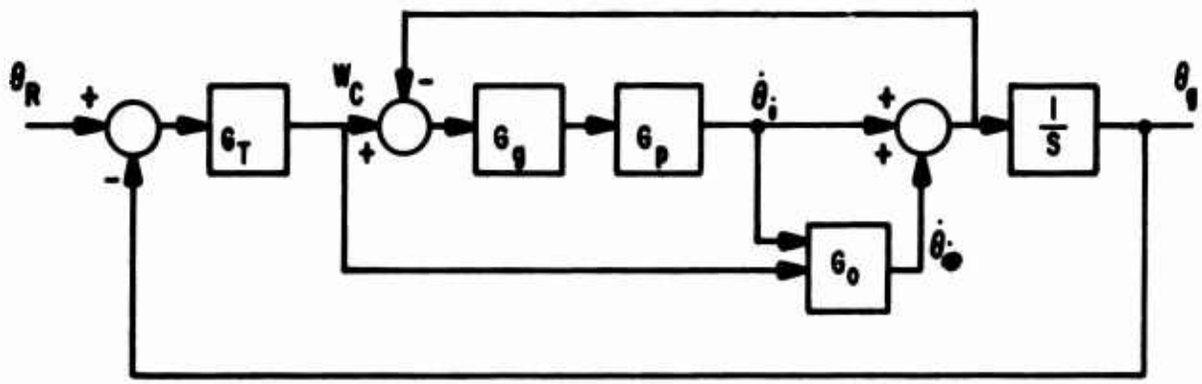


Fig. 2b. P_{10} Diagram Representation of Gimbal Coordinate Transforms



G_T Tracker transfer and compensation

G_g Gyro transfer and compensation

G_p Actuator pressure control and inner gimbal dynamics

G_o Outer gimbal rate loop

Fig. 3. A Single Axis Two Gimbal Model

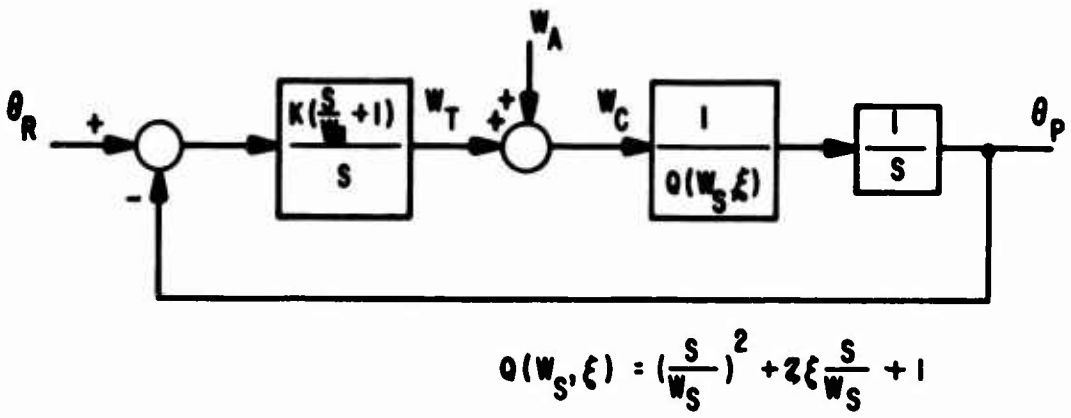


Fig. 4. A Simplified Single Axis Gimbal Model

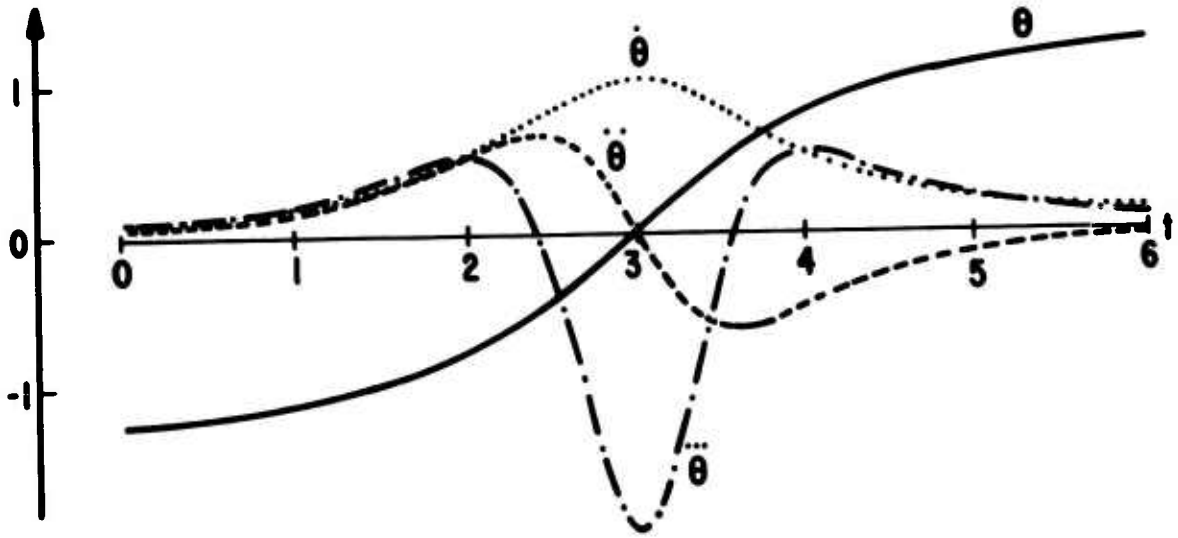


Fig 5. LOS Angle and Angle Derivatives for a Flyby Scenario

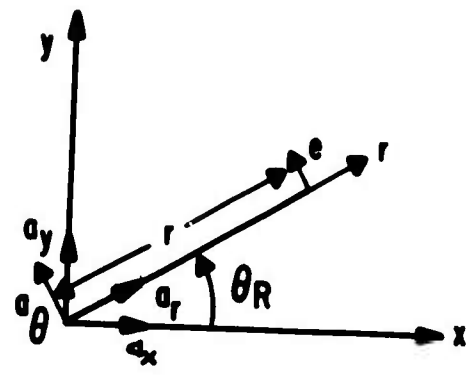


Fig. 6. Single Axis Pointing Geometry

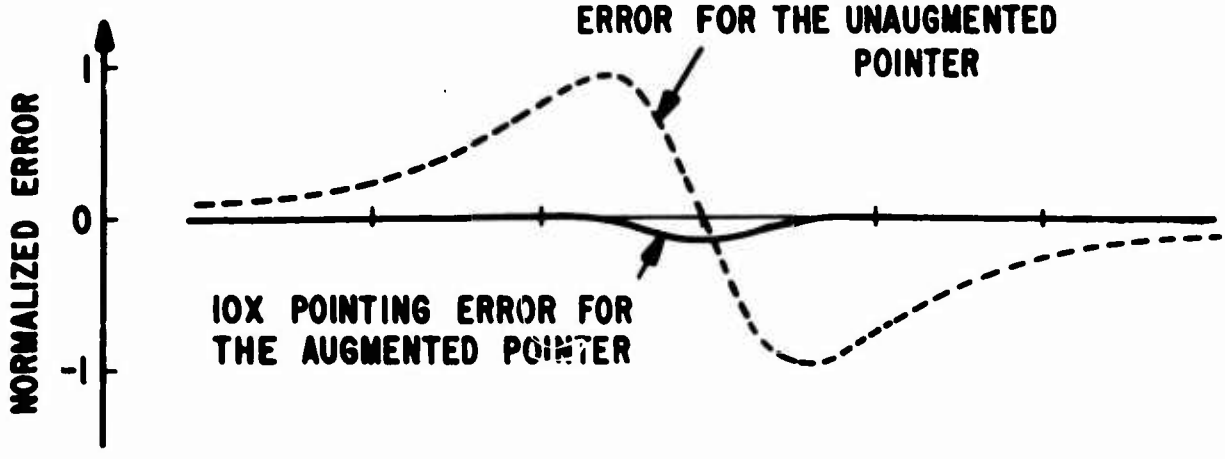


Fig 7. A Comparison of Pointing Errors with and without Control Augmentation

Applications of Modern Control Theory to the Integration of Radar/Inertial Navigation Systems for All Weather Navigation and Weapon Delivery

Richard M. Reeves
Air Force Avionics Laboratory

Recent advances in airborne radars have led to avionics system requirements which include several functions. Advnaced multimode radar systems may combine all or several of the following modes into a single radar: (1) velocity and position update of an inertial navigation system, (2) synthetic aperture radar mapping, (3) air-to-ground target acquisition and tracking, (4) air-to-air search, acquisition and tracking, (5) air-to-ground moving target detection and tracking, (6) terrain following/terrain avoidance, and (7) command guided weapon delivery.

At the present time the Air Force Avionics Laboratory is involved in several exploratory and advanced development programs, the objectives of which are to develop advanced multimode radar systems for both strategic and tactical navigation and weapon delivery. The Air-to-Ground Analysis Group of the Navigation and Weapon Delivery Division is conducting a series of in-house analysis efforts, the objectives of which are to develop cost effective techniques for integrating multimode radars with aircraft inertial navigation systems.

Two areas of particular interest are radar motion compensation using the inertial navigation system, and position and velocity update of the INS using the radar. Due to the complementary nature of the two sensors their outputs can be combined to provide an integrated multisensor system with both the long term and short term stabilities required for precision navigation and weapon delivery. In particular precision velocity and position updates are best obtained using synthetic aperture radar techniques which in turn depend upon high quality INS motion compensation signals. The radar/inertial system thus becomes closed loop in nature. Due to the dynamic real time environment within which the system must operate their integration is best accomplished using Kalman filtering techniques.

This paper will present results from several in-house studies related to the radar/inertial integration problem. Due to the classified nature of the material to be presented the contents of this paper will not be included in the conference proceedings. Additional information on these subjects can be obtained from the author or from the unclassified references listed at the end of this paper.

References

1. Duncan, D., "Combined Doppler Radar and Inertial Navigation Systems", Journal of the Institute of Navigation, Spring 1959, pp 321-327

2. Danik, B. and R. Stow, "Integrated Hybrid-inertial Aircraft Navigation Systems," Presented at the AGARD Guidance and Control Symposium, Delft, Netherlands, 23 Sep 1969

3. Fried, W. "Principles and Performance Analysis of Doppler Navigation Systems", IEE Transactions, Vol. ANE-4, No 4, Dec 1957, pp 176-196

4. Skolnik, M.I., Introduction to Radar Systems, McGraw-Hill, New York, 1962

5. Kayton, M. and W. Fried, Avionics Navigation Systems, John Wiley and Sons, New York, 1969

6. Broxmeger, C., Inertial Navigation Systems, McGraw-Hill, New York, 1964

7. Rhodes, D., Introduction to Monopulse, McGraw-Hill, New York. 1959

8. Reeves, R. and M. Needler, "The OMNIS Covariance Program - Analysis and Design of Multisensor Navigation Systems", Naval Avionics Facility Technical Report 1838, Nov 1973

9. Reeves, R. and Capt R. Asher, "Performance Evaluation of Suboptimal Filters", to be published IEEE Transactions on Aerospace and Electronic Systems

10. Reeves, R. and Robert M. Blumgold, "Radar/Inertial Navigation Using a Forward Looking Multimode Radar", Proceedings of the 1974 National Aerospace and Electronics Conference, May 13-15, 1974

11. Reeves, R. and Harold Noffke, "Motion Compensation in Synthetic Aperture Radars", Proceedings of the 1973 Joint Automatic Control Conference, June 20-22, 1973

ABSTRACT

IMPROVED GUIDANCE LAW EVALUATION

James E. Gibson
AF Armament Laboratory
Eglin AFB, Florida

Abstract: Both the Navy and the Air Force have identified operational deficiencies in US air-to-air missiles as a result of SEA combat experience. Efforts to correct these deficiencies have largely been concentrated on improvements to seeker characteristics and have led to complex, expensive and sometimes questionable solutions. This paper describes an on-going program at the Air Force armament Laboratory to explore simple, low-cost guidance alternatives through modifications to the basic proportional navigation guidance law. The effort is for a utility analysis of short-range air-to-air missile trajectory shaping which results in improved intercept geometry and increased terminal lethality for infrared homing systems. Two options are being examined: improvements which require only modifications to the missile itself and improvements which require modifications to both missile and launch aircraft. These two options will be compared and recommendations made as to which option to pursue.

SUMMARY

IMPROVED GUIDANCE LAW EVALUATION

This paper describes an on-going Air Force Armament Laboratory study to determine techniques of trajectory shaping for short-range infrared air-to-air missiles which result in improved intercept geometry and increased terminal lethality.

Classical proportional navigation results in relatively high track crossing angles (TCA's) between the missile and the target at intercept. Against afterburning targets, infrared missiles tend to track hot spots in the afterburner plume which are substantially displaced from the target tailpipe. A guidance law which minimizes terminal TCA would eliminate this deficiency since the hot spots in the plume are nearly colinear with the tailpipe for small TCA's. The optimum time delay for a fixed time delay fuze is a function of TCA. Since TCA is not a controllable variable with proportional navigation, the present fixed time delay in infrared missiles represents a suboptimal compromise over a broad range of TCA's. Again, a guidance law which minimizes terminal TCA should result in a more nearly optimal fixed time delay, with an attendant increase in missile lethality.

Candidate airframes for mechanization of new guidance laws developed under this effort are the AIM-4D-8, AIM-9E and AIM-9J; however, consideration is being given to improvements which can be implemented in new missile designs which are free of existing hardware constraints.

This analysis is being conducted considering realistic combat engagements and missile performance characteristics. Special attention is being given to the stochastic aspects of the engagement and to such features as system time delays, sensor saturation, seeker limits and missile aerodynamics. Additionally, the candidate missiles, as well as the F-4 fire control system and related missile launching equipment, are being studied to determine the suitability of mechanizing improved guidance laws in these airframes.

The effect of proposed guidance law implementations on the stability of the three candidate airframes is being examined as well as the effect of various noise sources on missile performance.

This study will identify the best guidance law/airframe combination for each of two options: improvements which only require modifications to the missile itself, and improvements which require modifications to both the missile and to the launch aircraft. These two options will then be compared with the baseline to determine which option represents the greatest improvement over the baseline.

MINIMAL ERROR TRAJECTORIES IN REAL TIME

EUGENE J. POLLOCK

Test and Evaluation Systems Program Office
Air Force Special Weapons Center
Kirtland Air Force Base, New Mexico

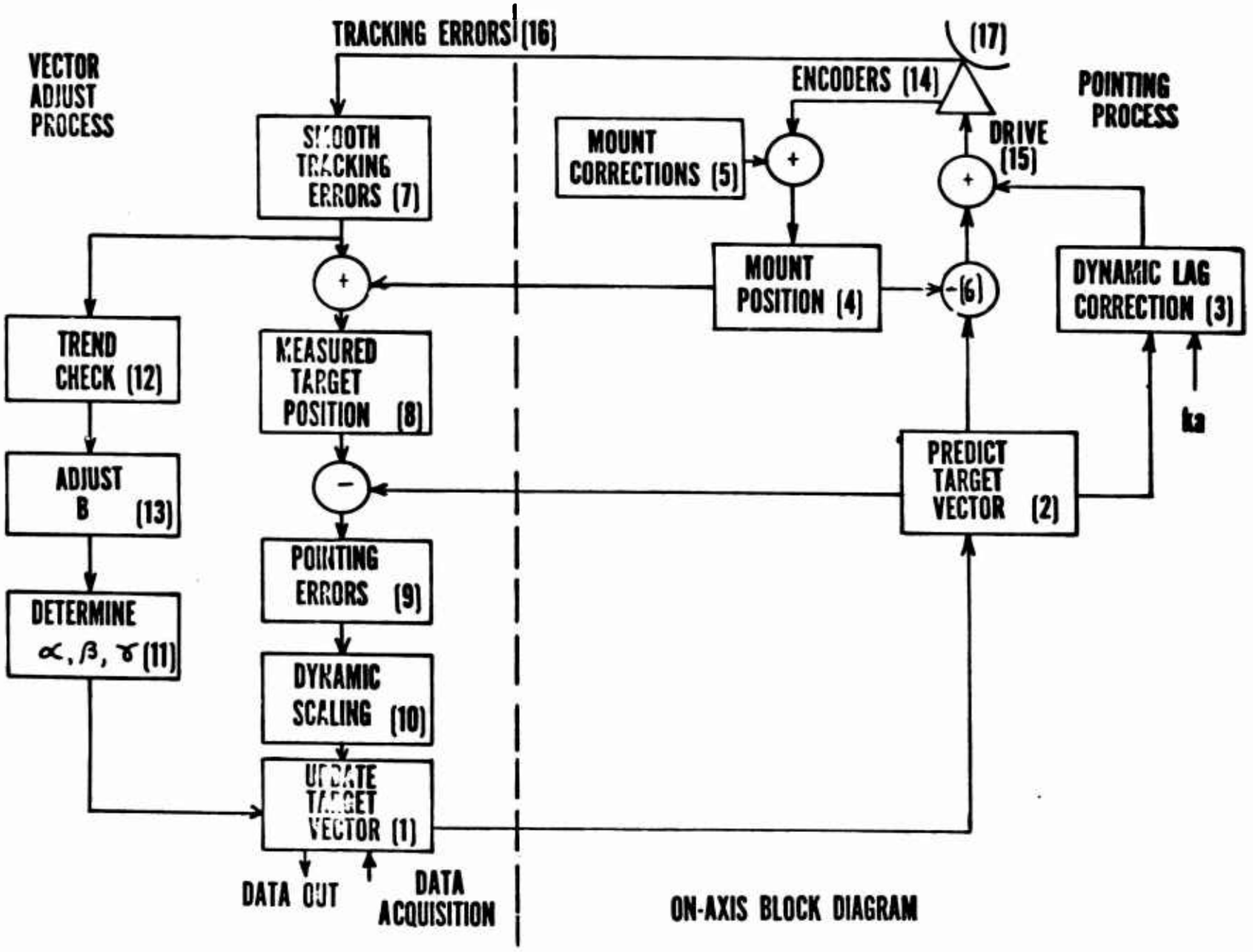
11 July 1974

1.0 INTRODUCTION:

A number of laser and very long focal length optical projects were assigned to the Range Measurements Laboratory (RML) in 1965. These projects were assigned to the RML in part because of the advertised high accuracy instrumentation on the AFETR. Although the AFETR did possess very precise instrumentation, this did not translate to accuracy except in a post flight sense. The ARPA projects all required accurate data in real time, and in addition, required accuracies beyond the resolution of the instrumentation then available. After a considerable time spent in trying, it became apparent that conventional tracking technology was the limiting factor and not the equipment.

Even if one had excellent trajectory data, one first had to learn how to use this information to accurately point remote long focal length telescopes. It was the development of the pointing technology which suggested the application to the closed loop system described in this paper.

There are errors which were not of primary concern if one used very wide apature systems, but the ARPA projects required operation with beam widths of a few arc seconds and errors of this magnitude could easily cause failure. One of these, dynamic lag posed a real dilemma. One suggested approach was to use a very high data rate to limit the lag error, but this solution was not acceptable because it also limited the range of operations to less than the minimum range to the target. A second alternate, the one successfully used, was to use a prediction process, but first one had to reduce the tracking noise.


VECTOR ADJUST PROCESS
TRACKING ERRORS (16)
(17)
ENCODERS (14)
POINTING PROCESS
DRIVE (15)
SMOOTH TRACKING ERRORS (7)
MOUNT CORRECTIONS (5)
+
+
+
MOUNT POSITION (4)
-(6)
DYNAMIC LAG CORRECTION (3)
TREND CHECK (12)
MEASURED TARGET POSITION (8)
ka
PREDICT TARGET VECTOR (2)
ADJUST B (13)
-
POINTING ERRORS (9)
DETERMINE α, β, γ (11)
DYNAMIC SCALING (10)
UPDATE TARGET VECTOR (1)
DATA OUT
DATA ACQUISITION
ON-AXIS BLOCK DIAGRAM

The happy solution was to drive from a state vector wherein noise is introduced only by the step size in a digital integration process.

Probably the most significant consideration in this process is the ability to evaluate each step in the process by direct observation of the effect or result.

With this introduction already too long, I should like to summarize by saying we learned how to accurately track only after we found out how to correctly point our instrumentation. Once we learned to point the instruments, including radars, the project was 100% successful.

2.0 SYSTEM DESCRIPTION:

This system, known as ON-AXIS has several versions, the one described here is an instrumentation radar in an aircraft tracking configuration.

2.1 Process

The system is composed of two on-line processes, one the pointing loop whose bandwidth is wide (approximately 50 Hz) in order to pass the dynamics of a close overhead pass. The pointing loop is predictively driven by a digital computer from a state vector corrected for mount and environmental anomolies. The second process is a variable, but narrower bandwidth loop known as the vector adjust process. It contains a fading memory filter operating on the apparent error as part of the state vector adjustment to the true trajectory. These two processes, separated by a dashed line, are shown in block diagram form in Figure 1. Each box in the ON-AXIS block diagram is numbered. These numbers will be referred to in the text to indicate the area of discussion. The pointing process is updated 100 times/second, while the vector update process occurs 10 times per second.

2.2 Coordinate Systems

Although not necessary for an understanding of the ON-AXIS process, some mention of the coordinate systems being used is in order.

The prediction process (2) is calculated using a geocentric coordinate system. The data collection process is accomplished in a local Cartesian system known as the boresight coordinate system. These data are transformed to the geocentric system to adjust the state vector, and then further transformed to a local reference for other computational and display purposes.

2.3 Encoders

This process does not necessarily require new equipment; rather it suggests a different method of operating instrumentation. The system does need a set of one speed angle encoders and a range machine whose resolution is at least equal to the accuracy desired, along with an accurate clock.

2.4 Time Base

The independent variable of the ON-AXIS technique is time. The system uses an internal clock to generate high frequency time good to one microsecond. Although time accuracy is excellent, special care is taken to time tag the state vector carefully. This is accomplished by entering the mount variables into computer memory at the same instant that the clock output is inserted into core.

2.5 Direct Observation

An important consideration in the ON-AXIS process is the capability for direct observation of the effect of any change to the system, rather than the evaluation of the composite data produced. A necessary part of the system is a set of equations, called mount corrections, (5) describing a dynamic "error model" used to provide apriori corrections to the drive

data. In this process, the term "error model" might be replaced by the term "coordinate transformation." All of the dynamic transformation coefficients are determined by direct and recurrsive observations using a long focal length boresight telescope mounted on the radar antenna, and coupled to a closed circuit television system and recorder. Transformation coefficients to provide direct optical observations of the stars are evaluated by pointing at stars independently selected for each transformation term. For those terms unresolved by direct optical observation of the star field, a comparison is made of the computed location to the actual location of the target in the boresight optics.

2.6 Digital Techniques

The discussion determining whether to build an optimal or sub-optimal system is generally an economic or technological one, and frequently requires a special purpose digital computer. Fortunately, the ability to build special purpose digital machines has been reduced to a relatively low cost enabling one on this project to do, in hardware, things realizable only by software a few years ago.

3.0 Acquisition (Box 1)

The determination of the initial state vector is called acquisition and is accomplished in two general ways. In one case, a state vector is given, and while not exactly correct, it is sufficient to start the update process. A specific example of this first case is the seven component vector describing the trajectory in the ORBITAL problem. One then causes the instrument to search locally about the given state vector, and starts the refinement process when acquisition occurs. Another example of this first case results when a series of time varying state vectors are furnished as driving data from some remote sensor.

In the second acquisition case only the state variables of position are given. One must again search locally for the target, and once the target is acquired, sufficient data must be collected, filtered in an expanding memory sense until a state vector of sufficient accuracy is obtained to start the prediction process.

4.0 THE POINTING LOOP:

The error sensing detector (17), is not directly part of the pointing process. In this discussion, it is a microwave monopulse system, driven (pointed) along the trajectory generated from the state vector.

4.1 Predict Target Vector

A driving bandwidth of approximately 50 Hz assures a very smooth motion of the antenna with respect to the target. The 50 Hz drive (2) is sustained by integrating the state vector forward in time 100X/sec to a new series of positions. The predicted state variables shown in Figure 2 are components of position obtained by successive integration of the second derivative. Time updated position and velocity assumes that the second derivative (acceleration) is constant for a 100 ms interval Δt.

* $X_{j+1} = X_j + \dot{X}_j\,(\Delta t) + \ddot{X}_j \frac{(\Delta t)^2}{2}$

* $\dot{X}_{j+1} = \dot{X}_j + \ddot{X}_j\,(\Delta t)$

* $\ddot{X}_{j+1} = \ddot{X}_j$

X REPRESENTS THREE STATE VARIABLE COMPONENTS

X = POSITION

$\dot{X}$ = VELOCITY

$\ddot{X}$ = ACCELERATION

j = .001 sec.

PREDICTIVE RELATIONS

Figure 2

4.2 Encoders (14)

A radar system uses four encoders to describe the mount position. One encoder is coupled directly to the azimuth turntable, another is installed on the elevation axis, and a digital counter describes the range gate location. As an example, a 20 bit encoder divides azimuth positions into one part in 10^6. These encoders are strobed 100X/sec to the computer along with time. Time is encoded using an on-site clock.

4.3 Systematic Error Mount Corrections (5)

Corrections are added to the encoder readings, reducing known systematic errors, thereby assuring that the sensor is pointed to the appropriate space coordinates at all times. Each additive term for the system is calculated on line. The effect of any added correction is observable, and will be recursively corrected in the calibration process.

4.3.1 Calibration

With accurate calibration, ON-AXIS will point and update the vector very accurately. It will, however, not produce high quality metric data if the system is not calibrated. The calibration process is one of observing how well the system points to the target of interest. The primary targets from a calibration viewpoint are the stars. The system is assumed calibrated, if and only if, it can point to approximately 20 to 40 stars randomly selected from the hemisphere of observation. Using the same integration process as described in 4.1, a state vector is employed to describe a star's motion. We point to the selected celestial objects from a catalog of approximately 600 stars, and observe

through a long focal length telescope television system how well we are calibrated. Each term in the transformation model is independently and recursively evaluated, by noting whether the system diverges or converges to the correct position as we adjust the coefficients of the model.

4.3.1.1 Encoder Bias

When sampled, the angle encoders indicate the orientation of the mount axis with respect to zero references of the encoders. In general zero is not coincident with the local zero references of north and up. The model for this transformation is a simple algebraic process representing the difference between the observed (o) and the reference.

$$\Delta A_n = A_o - A_n$$

$$\Delta E_u = E_o - E_u$$

4.3.1.2 Nonorthogonality is the mechanical misalignment of the elevation axis with respect to the azimuth axis, and α_{n_o} is the new orthogonality coefficient.

$$\Delta A_{n_o} = \alpha_{n_o} \tan E_u$$

4.3.1.3 Droop. Droop is caused by the bending of the mount under gravitational loading and α_d is the droop correction coefficient.

$$\Delta E_d = \alpha_d \cos E_u$$

4.3.1.4 Skew. Skew is the misalignment of the sensor with respect to the elevation axis, and α_s is the skew correction coefficient.

$$\Delta A_s = \alpha_s / \cos E_u$$

4.3.1.5 Mislevel is the noncoincidence of the azimuth plane of the mount to the local reference plane. The parameter $_m$ is the mislevel amplitude and ϕ_m is the phase angle of the amplitude with reference to the local zero reference. The correction ΔA_m to each sample of drive data is calculated as shown.

$$\Delta A_m = \sin (A_e - \phi_m) \, _m \tan E_e$$

$$\Delta E_m = \cos (A_e - \phi_m) \, _m$$

4.3.1.6 Environmental. In addition to the stellar evaluated terms, there are environmental influences which must be modeled. These are refraction and transit time.

4.3.1.6.1 Refraction. Refraction corrections are applied to elevation and range variables only, azimuth being unaffected. The corrections are calculated by using the widely known ray-trace integral technique. The effects of refraction are calculated in both the microwave and optical region. This process is pointing the radar system, and the telescope is mounted on the elevation axis of the radar. In order to observe optically how well the radar is pointing, one must correct the optical observation as well as the microwave.

4.3.1.6.2 Transit Time. In pointing one designates to a virtual target, i.e., that point in space where the target is expected to be. Microwave receivers are sensing reflected energy. By the time the energy reaches the receiver, the target will have traveled some distance. The time required between the time the energy hits the target and its receipt by the receivers is related to the range of the target and the velocity of propagation of the energy.

$$\Delta t = (R_e + \Delta R_r)/C$$

The corrections must be applied to position terms of the state variable and are determined from third order differential equations.

$$\Delta A_t = A_u \ \Delta t + 1/2\ A_u\ (\Delta t)^2$$

$$\Delta E_t = E_u \ \Delta t + 1/2\ E_u\ (\Delta t)^2$$

$$\Delta R_t = \quad \Delta t + 1/2\ R_u\ (\Delta t)^2$$

4.3.1.7 Dynamic Lag Correction (3). Servo systems will lag an input driving function by an amount dependent on the response of the system and the input driving function. By expanding the system error equation in a power series, the error may be expressed in terms of the input function, its derivatives and the servo error coefficients. For a positioning system, the error appears as a position error, and is called dynamic lag. This error is highly predictable allowing an appropriate addition to the input driving function so that dynamic lag does not degrade the positioning accuracy.

The servo system used in the ON-AXIS formulation is a hybrid; a type II analog servo is implemented in hardware, additional integrations are performed in the software digital drive data. A hardware type II servo in effect has velocity memory. The steady state error is described by an equation.

$$Ka = \frac{\text{accel.}}{\text{error}}$$

The Ka value of the system is determined by a calibration process using a synthetic drive matching the most difficult target trajectory one can describe. This trajectory, a near overhead pass, will be used to drive the mount at 100X/sec, with a constantly changing acceleration.

The ON-AXIS system is driven by a state vector, which contains the current estimate of target acceleration. With both the Ka and acceleration known, one may calculate that correction which must be added to prevent dynamic lag.

4.4 Mount Position (4)

The corrected encoder data now gives the position of the mount and range machine in space coordinates. Mount position data (4) are now compared with the predicted position; the difference (6) between them generate the drive data, that is converted to analog signals for the mount servos (15). At this time, the mount is being driven along the state vector trajectory calibrated both statically and dynamically. If the state vector is correct, then the target will be in the center of the field of view of the sensor, i.e., ON-AXIS.

4.4.1 There is a significant point to be emphasized here. We are pointing the antenna with very small errors, i.e., the least bit of our encoders, to absolute positions in space primarily because of the independent calibration of each term in the transformation model.

4.4.2 The corrected encoder data and the state vector drive data are compared and displayed on the television monitor as designate differences. $\dot{A}$ non-zero difference indicates that the mount is not precisely following the drive data.

5.0 TRACKING ERROR PROCESSING:

Even though one can point correctly, the state vector used to drive the mount may not accurately describe the target dynamics. The tracking

errors are processed to correct the state vector. This is accomplished by adjusting the drive data to keep the tracking error random at the boresight, making the target remain on the average in the center of the beam, i.e., ON-AXIS. Since the pointing process is very smooth, every attempt at adjustment of the state vector must be carefully and slowly done or we will introduce discontinuities. In all cases, the final smoothness is slightly disturbed to make the system accurate.

5.1 Vector Adjust

The antenna is now moving very smoothly along with the target, allowing long time integration of the tracking error to achieve accuracies not previously possible. The sensor is in this case a monopulse system operating at 160 pulses per second. The outputs of this system are radar returns forming an orthogonal set. One is up/down about the boresight, the second left/right about the boresight while the third is in/out about the target.

5.2 Noise

Although most of the tracking noise has disappeared because of the smooth pointing, these data (16) are still noisy because of atmospheric and target influences, and are averaged over 16 returns, giving 10 relatively noise free observations per second (7). To these observations are added the mount corrected encoder data (4), to form the measured target position (8). When the measured target position is compared to the predicted target position (2), one obtains a series of pointing errors (9) related to the radar, not just to the boresight.

5.3 Randomness

Literally, the ON-AXIS vector adjust process is based on the assumption that if the boresight tracking errors (16) can be kept random, then all of the useful information has been extracted. A trend check (12) is made of successive smoothed samples (7) of the tracking errors, by multiplying together the successive samples. If the resultant product is positive, then the boresight data (16) are said to be trended, otherwise the boresight data are assumed to be random. When the data are random, then the adjustment to the state vector is minimized, while if the data are trended, then the correction are more pronounced.

5.4 State Vector Corrections

The pointing errors (9) are three state variable position errors, and since we point using a state vector, one must obtain corrections for the nine state variables. This process operates as a separate loop using a fading memory polynomial filter (11), see Figure 3, that provide corrections to components of position, velocity, and acceleration.

* AIRCRAFT MODEL

* POSITION CORRECTION = $\alpha = 1-B^3$

* VELOCITY CORRECTION = $\beta = (1.5)\ (1-B)^2\ (1+B)$

* ACCELERATION CORRECTION = $\gamma = (0.5)\ (1-B)^3$

FADING MEMORY FILTER (11)

Figure 3

5.4.1 Bandwidth Adjustment

The value of the operator B determines the magnitude of the position, velocity, and acceleration correction. Operator B, or as it is generally called, B number, is always related to the randomness of the smoothed tracking errors as determined by the trend check (12). The B number may vary between .998 indicating the vector adjust bandwidth is near zero, and .850 corresponding to a bandwidth of approximately two Hz. See Figure 4, Operator B Determination (13).

As a result of randomness check, if random, the B number is increased 10X/sec by a number which varies with the B value, and target dynamics. If the data are trended, then the B number is decreased 10X/sec by .001.

A variable up/down rate is produced to meet the target maneuvering capability. To get a high B#, the state vector must be very good and, therefore, a high B# yields very small corrections. If the target aircraft maneuvers, then the update bandwidth must be rapidly increased and the B# goes down. As soon as the maneuver is over, with the B# down, it must be easy to go up, and rapidly goes to 950.

5.5 Dynamic Scaling

An editing procedure is employed which scales down rather than rejects invalid data points. The pointing errors (9) are components of position. The value of these errors are scaled in relation to the target velocity. One version, the Dynamic Ellipse Shaped Scaling (DESS), see Figure 5, uses the eccentricity of an ellipse as a function of the target velocity, and the vector direction difference between the measured target position (8) and the predicted target position to calculate the error scaling

* $B_n = B_{n-1} + \Delta B$

* WHEN (12) TRENDED $\Delta B = -.001$

* WHEN (12) RANDOM $\Delta B = \frac{(1\ B^3)\ (DESS_k)\ (.001)}{.12}$

* AS B# INCREASES - UPDATE BANDWIDTH DECREASES

.85	2.0 Hz
.90	.7
.95	.3
.98	.25
.99	.06
.999	.006

OPERATOR "B" DETERMINATION (10)

Figure 4

* DYNAMIC ELLIPSE SHAPED SCALING

* $DESS_k = \frac{L}{1 + C}$

* $L = 2 - \frac{(1 + C^2 - 2\,C\,\cos\theta)}{1 - C\,\cos\theta}$

* $C = \text{ECCENTRICITY} = \frac{\sqrt{a^2 - b^2}}{a}$

* $\theta = <$ BETWEEN MEASURED (8) AND PREDICTED (2) POSITION

* IF VELOCITY VECTOR IS LARGE THEN ECCENTRICITY (C) OF ELLIPSE IS LARGE;
  IF VELOCITY VECTOR IS SMALL, THEN C IS SMALL

DESS (15)

Figure 5

* $X_{\eta 1} = X_{\eta-1} + (\alpha)\bar{E}_{\eta}$

$\dot{X}_{\eta 1} = \dot{X}_{\eta-1} + (B)\bar{E}_{\eta}/\Delta t$

$\ddot{X}_{\eta 1} = \ddot{X}_{\eta-1} + (\gamma)\bar{E}_{\eta}/\Delta t^2$

* E = ERROR (POINTING ERRORS (9) AFTER SCALING (10)

* $X_{\eta 1}$ = POSITION STATE VARIABLE AT TIME η

* Δt = .1 SEC.

RECURSIVE RELATIONS

Figure 6

coefficient. The ellipse is used as the weighting model, becoming a circle at very low velocities and extending to a very long ellipse at high velocity. If the vector direction difference is small, i.e., along the predicted path, the full errors are used, but if large angular differences are noted, then the value is reduced by direct scaling, using the value determined by its location on the ellipse.

The value of $DESS_k$ is also used in the B# determination (10) is part of the ΔB computation when the data (12) are random.

5.6 Update Target Vector (1)

The state vector is updated by applying the recursive relations shown in Figure 6.

The result of this computation is a new set of numbers for the state vector used to point the sensor. Although this radar description shows the total process in the same instrument, it is not uncommon for the radar to output a state vector 10X/sec. These data are then used to point remote instrumentation such as telemetry, theodolites, etc. The remote instrument will then consist of the pointing loop only.

6.0 SUMMARY:

What has been described here is an alternate to the conventional method of employing instrumentation. The process enables one to achieve on-line accuracies previously only obtainable thru off-line processing. One can also argue that the off-line process is never checked, while the on-line process is continuously evaluated.

7.0 APPLICATIONS:

7.1 Past/Current Applications

7.1.1 The On-Axis technique has been applied to long focal length satellite photography. Images using focal length to 7000" have been observed.

7.1.2 Laser illumination of low orbiting satellites has been employed routinely.

7.1.3 Images have been obtained on satellites at 22K miles. Active illumination of satellites at 60K miles has been demonstrated.

7.1.4 Lasers have been continuously pointed onto aircraft through all maneuvers of aircraft for extended time periods.

7.1.5 Accurate orbit determination using a single telescope has been demonstrated.

7.1.6 Real-time determination of refraction effect of exoatmospheric bodies has been demonstrated.

7.1.7 Small practice bombs have been tracked by O/A radar from aircraft release to impact at ranges exceeding 30 miles, and the event photographed using boresight optics.

7.1.8 Aircraft launched rockets have been tracked and photographed using an O/A radar.

7.1.9 The automatic pointing of remote theodolites using an O/A radar has been demonstrated.

7.2 Future Applications

7.2.1 Coordinate bombing.

7.2.2 Airborne laser applications.

7.2.3 Fire control systems from fixed or moving platforms.

8.0 REFERENCES:

8.1 ARPA Tech Memo #211, ON-AXIS. Philosophy/Technology/ Development, published by Range Measurement Laboratory, Patrick AFB, Fla, 15 Dec 1970.

8.2 Introduction to Sequential Smoothing and Prediction, Norman Morrison, McGraw-Hill Book Company, New York.

8.3 Discrete-Time Systems, James A. Cadzow, Prentice-Hall Inc., Englewood Cliffs, New Jersey, especially pages 298-303, 272-279, 53-57, 63-66, 291.

8.4 Discrete-Time and Control Systems, James A. Cadzow and Henrich R. Martins, Prentice-Hall Inc., Englewood Cliffs, New Jersey, especially pages 27, 83, 89 and 90.

8.5 Introduction to Radar Systems, Merrill I. Skolnik, McGraw-Hill Book Co., New York.

8.6 AFSWC/TESPO Technical Memorandum No. 102, Time, 21 September 1973, Kirtland AFB NM 87117.

ABSTRACT

APPLICATION OF EXTENDED KALMAN FILTERING TO A CALIBRATION OF AN INERTIAL NAVIGATION SYSTEM

BY

FRAZIER J. HELLINGS, LT COLONEL, USAF

6585th TEST GROUP

HOLLOMAN AFB, NEW MEXICO

The paper describes a data reduction technique that obtains estimates of inertial sensor error model coefficients from a dynamic calibration of a typical Inertial Navigation System. The error model coefficients are those associated with gyros, accelerometers, and their misalignment errors that have been found by test and analysis to be the predominant sources of error affecting system accuracy. Accurate estimates of the error terms requires a positively controlled test procedure that varies the system orientation with respect to the input acceleration to provide better observation of the error coefficients in the measurement data, and a data reduction program that provides the "best" estimates of the error coefficients from noisy measurement data. This paper proposes a dynamic test procedure and a data reduction algorithm for obtaining estimates of the error coefficients from a 1-g system calibration. The results of a 17-state simulation verifies the application of the filtering algorithm to the calibration problem.

SUMMARY

The paper describes a data reduction technique that obtains estimates of inertial sensor error model coefficients from a dynamic calibration of a typical Inertial Navigation System. The error model coefficients are those associated with gyros, accelerometers, and their misalignment errors that have been found by test and analysis to be the predominant sources of error affecting system accuracy. All the error terms considered are categorized as either fixed (independent of applied acceleration), first-order (proportional to the first power of acceleration), or higher-order terms, which are proportional to the square or cube of acceleration. In the case of the higher-order terms, the error model coefficients of inertial grade sensors are from one to four orders of magnitude smaller than the fixed and first-order terms. To obtain measurable quantities of these error sources requires the application of precise acceleration inputs. Unfortunately, accurate estimates of the error terms require more than precision input accelerations. In addition one must have a positively controlled test procedure that varies the system orientation with respect to the input acceleration to provide better observation of the error coefficients in the measurement data, and a data reduction program that provides the "best" estimates of the error coefficients from noisy measurement data. This paper proposes a dynamic test

procedure and an associated data reduction method for obtaining estimates of the error model coefficients from a 1-g system calibration.

The dynamic test procedure is obtained by command torquing the platform gyros to drive the platform through a prespecified torque profile so that each axis is subjected to the applied acceleration. Since the platform motion is a combination of commanded rate, applied earth-rate, and drift due to the gyro error terms, tracking the platform motion with Euler angle rates provides the basis for the state dynamics. The measurements, obtained from the triad of platform accelerometers, are functions of the accelerometer error terms and the Euler angles which in turn contain the drift motion due to the gyro error coefficients. Therefore, the problem is a parameter estimation problem where the Euler angles and error model coefficients are designed as system states to be estimated from noisy measurements. Since both the state and measurement equations are nonlinear functions of the state variables, an Extended Kalman Filter is applied.

The results of a 17-state simulation verifies the application of the filtering algorithm to the calibration problem. The results show that significant improvement can be obtained in reducing the time required to calibrate an Inertial Navigation System. In addition, a method of applying the off-diagonal terms of the filtering error

covariance matrix provides a new dimension in evaluating filter performance with respect to system observability.

The scope of the paper is as follows:

1. Development of the system and measurement dynamic equations for a complete Inertial Navigation System Calibration Test Program. This task includes the inertial sensor error model equations, the platform rate equations, the kinematic and gravitational acceleration equations with the corresponding coordinate transformations, and finally the Euler rate equations for tracking the platform motion.
2. Development of the extended Kalman filter equations.
3. Evaluation of the data reduction technique via a 17-state variable simulation of a 1-g system calibration.

So that the error model coefficients in the simulation are representative of an actual physical system, the Carousel VB Inertial Measurement Unit developed by Delco Electronics was selected as the baseline Inertial Navigation System, since it is representative of the type of Inertial Guidance System being used in both ballistic missile and aircraft navigation applications.

The United States Air Force selected the Carousel VB for use in their Titan IIIC series of launch vehicles. The Carousel IV, which is similar to the Model VB, is presently installed in the Boeing Models 747 and 707, and McDonnell Douglas Models DC-8 and DC-10 aircraft.

For further information see:

Air Force Report No. SAMSO-TR-73-219

Aerospace Report No. TR-0073(3115)-3

APPLICATION OF EXTENDED KALMAN FILTERING TO A CALIBRATION OF AN INERTIAL NAVIGATION SYSTEM

- PROBLEM - OBTAIN ESTIMATES OF SENSOR ERROR COEFFICIENTS FROM A 1-G SYSTEM CALIBRATION

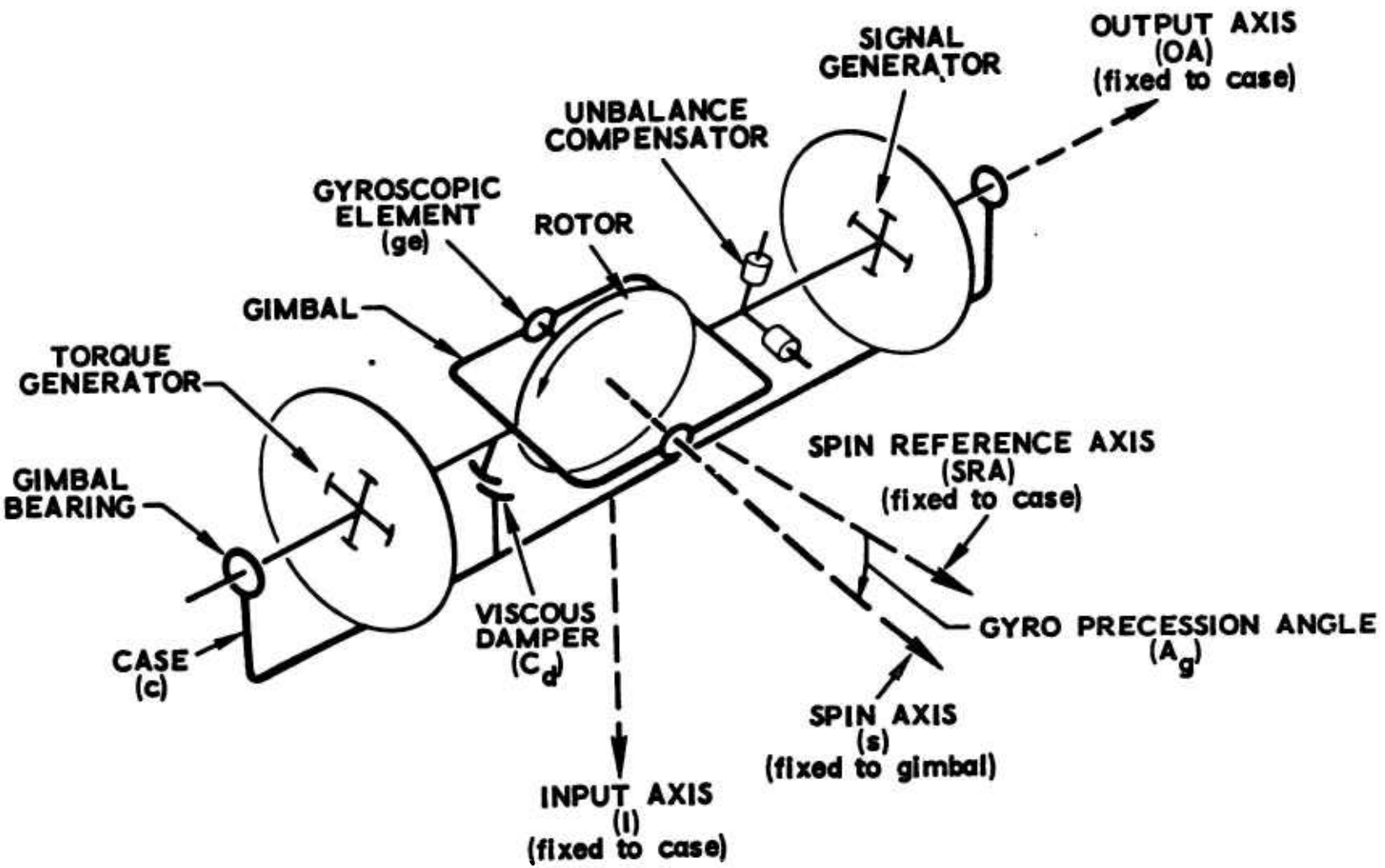
- PARAMETER ESTIMATION PROBLEM - OBTAIN STATE ESTIMATES FROM NOISY MEASUREMENTS

- APPLY DYNAMIC COMMAND TORQUE PROFILE TO OFFSET OBSERVABILITY PROBLEM

- SINCE BOTH THE STATE AND MEASUREMENT EQUATIONS ARE NONLINEAR FUNCTIONS OF THE STATE VARIABLES - EXTENDED KALMAN FILTER

OVERVIEW

- INERTIAL SENSOR PERFORMANCE MODEL EQUATIONS
- PLATFORM PERFORMANCE MODEL EQUATIONS
- APPLIED KINEMATIC AND GRAVITATIONAL ACCELERATION EQUATIONS
- EULER RATE EQUATIONS
- EXTENDED KALMAN FILTER ALGORITHM
- 17-STATE SIMULATION RESULTS AND CONCLUSIONS



Line Schematic of a Single-Degree-of-Freedom Integrating Gyro

GYRO PERFORMANCE MODEL

$$\dot{\underline{H}}^{i}_{i,ge} = \underline{M}$$

$$\underline{M} = \dot{\underline{H}}^{ge}_{s} + \dot{\underline{H}}^{ge}_{ns} + \underline{\omega}_{i,ge} \times \underline{H}_{s} + \underline{\omega}_{i,ge} \times \underline{H}_{ns}$$

$$\dot{\underline{H}}^{i}_{i,ge} = \dot{\underline{H}}^{ge}_{i,ge} + \underline{\omega}_{i,ge} \times \underline{H}_{i,ge}$$

$$\underline{M} = \dot{\underline{H}}^{ge}_{ns} + \underline{\omega}_{i,ge} \times \underline{H}_{s} + \underline{\omega}_{i,ge} \times \underline{H}_{ns}$$

$$M^{c} = C^{c}_{ge}\, \dot{H}^{ge}_{ns} + W^{ck}_{ic}\, C^{c}_{ge}\, H^{ge}_{s} + W^{ck}_{ic}\, C^{c}_{ge}\, H^{ge}_{ns}$$

$$M^c = \{M_I,\ M_{OA},\ M_{SRA}\}$$

$$\omega_{ic}^c = \{\omega_I,\ \omega_{OA},\ \omega_{SRA}\}$$

$$C_{ge}^c = \begin{bmatrix} 1 & 0 & A_g \\ 0 & 1 & 0 \\ -A_g & 0 & 1 \end{bmatrix}$$

$$\omega_x = \omega_I - A_g\,\omega_{SRA}$$

$$\omega_y = \omega_{OA} + \dot{A}_g$$

$$\omega_z = \omega_{SRA} + A_g\,\omega_I$$

$$H_s^{ge} = \{0,\ 0,\ H_s\}$$

$$\omega_{i,ge}^{ge} = \{\omega_x,\ \omega_y,\ \omega_z\}$$

$$H_{ns}^{ge} = \{I_x\,\omega_x,\ I_y\,\omega_y,\ I_z\,\omega_z\}$$

$$W_{ic}^{ck} = \begin{bmatrix} 0 & -\omega_{SRA} & \omega_{OA} \\ \omega_{SRA} & 0 & -\omega_I \\ -\omega_{OA} & \omega_I & 0 \end{bmatrix}$$

$$H_{ns}^{ge} = \begin{bmatrix} I_x(\omega_I - A_g\,\omega_{SRA}) \\ I_y(\omega_{OA} + \dot{A}_g) \\ I_z(\omega_{SRA} + A_g\,\omega_I) \end{bmatrix}$$

$$\dot{H}_{ns}^{ge} = \begin{bmatrix} I_x(\dot{\omega}_I - \dot{A}_g\,\omega_{SRA} - A_g\,\dot{\omega}_{SRA}) \\ I_y(\dot{\omega}_{OA} + \ddot{A}_g) \\ I_z(\dot{\omega}_{SRA} + \dot{A}_g\,\omega_I + A_g\,\dot{\omega}_I) \end{bmatrix}$$

$$M_{OA} = I_y \dot{\omega}_{OA} + I_y \ddot{A}_g - H_s \omega_I + H_s \omega_{SRA} A_g$$

$$- (I_z - I_x) \omega_I \omega_{SRA} - A_g (\omega_I^2 - \omega_{SRA}^2) (I_z - I_x)$$

$$M_{OA} \doteq M_{tg} - C_d \dot{A}_g + (u) M$$

$$\tau_g = I_y / C_d$$

$$\frac{C_d}{H_s} \tau_g \ddot{A}_g + \omega_O = \omega_I + \omega_{CMD} + (u)\omega_D - \omega_{SRA} A_g - \frac{C_d}{H_s} \tau_g \dot{\omega}_{OA}$$

$$+ \frac{(I_z - I_x)}{H_s} \omega_I \omega_{SRA} + \frac{A_g}{H_s} (\omega_I^2 - \omega_{SRA}^2) (I_z - I_x)$$

where

$\omega_O = \frac{C_d}{H_s} \dot{A}_g$ = output angular rate of the gyro

ω_I = applied rate about the gyro input axis

$\omega_{CMD} = \frac{M_{tg}}{H_s}$ = commanded angular rate

$(u)\omega_D = \frac{(u)M}{H_s}$ = error drift generated by internal gyro error sources

A_g = 1 arc-minute

C_d = 62,500 gm-cm^2/sec

H_s = 1 × 10^5 gm-cm^2/sec

τ_g = 2 milliseconds

ω_I = 5.4 × 10^{-4} rad/sec

ω_{SRA} = 5 × 10^{-4} rad/sec

$\dot{\omega}_{OA}$ = 3.2 × 10^{-8} rad/sec^2

$I_z - I_x$ = 10 gm-cm^2

$\omega_I + \omega_{CMD} + (u)\omega_D = 6.0 \times 10^{-4}$ rad/sec

$\omega_{SRA} A_g = 14.5 \times 10^{-8}$ rad/sec

$\frac{C_d}{H_s} \tau_g \dot{\omega}_{OA} = 4.0 \times 10^{-11}$ rad/sec

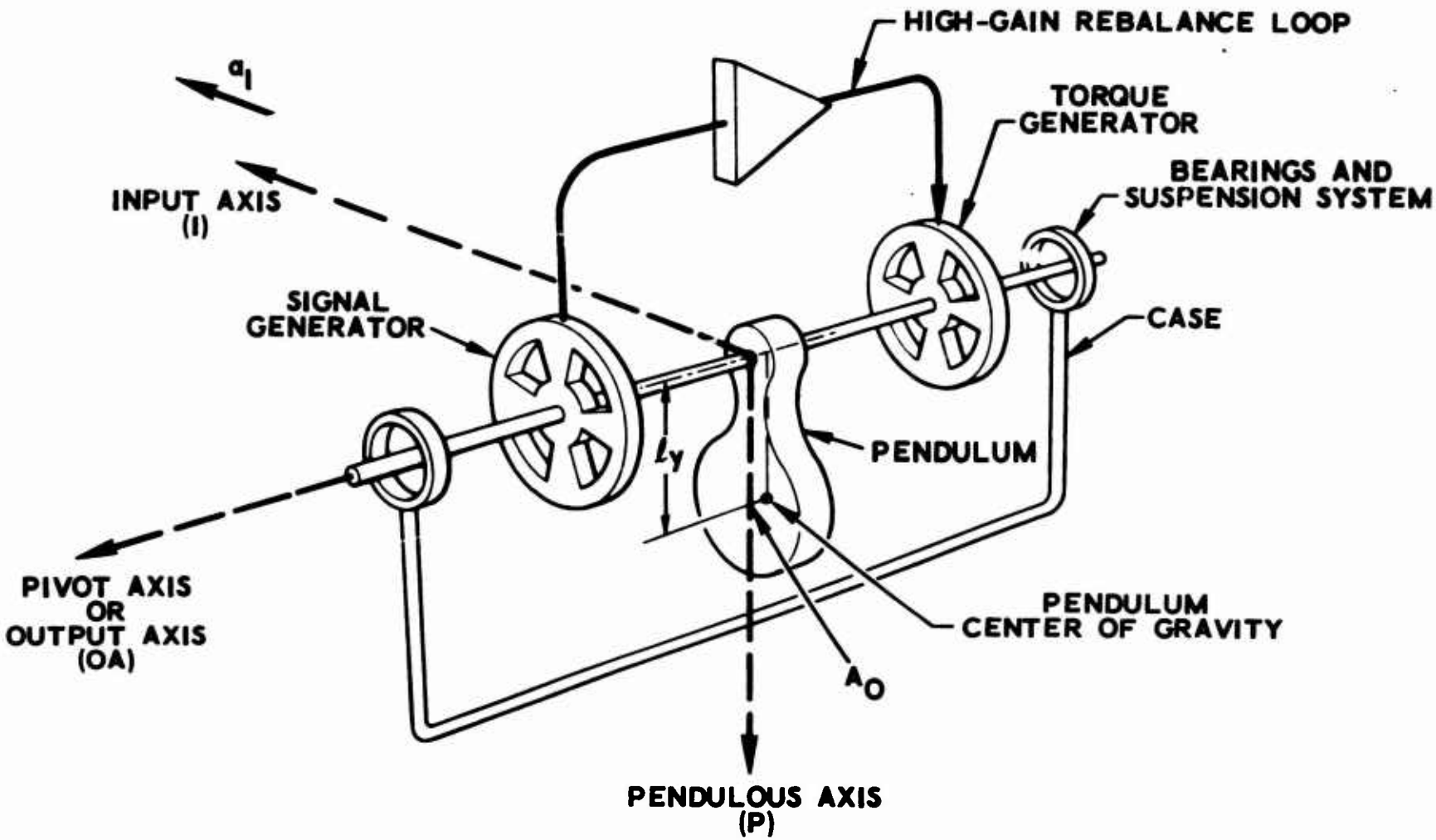
$\frac{(I_z - I_x)}{H_s} \omega_I \omega_{SRA} = 2.7 \times 10^{-11}$ rad/sec

$\frac{A_g}{H_s} (\omega_I^2 - \omega_{SRA}^2)(I_z - I_x) = 1.2 \times 10^{-15}$ rad/sec

$\omega_O = \omega_I + \omega_{CMD} + (u)\omega_D$

$$\omega_O = (D_F + D_I a_I + D_O a_O + D_S a_S + D_{II} a_I^2 + D_{OO} a_O^2 + D_{SS} a_S^2$$

$$+ D_{IO} a_I a_O + D_{IS} a_I a_S + D_{OS} a_O a_S) + \omega_I + \omega_{CMD}$$



Line Schematic of a Force-Rebalance Pendulous Accelerometer

ACCELEROMETER PERFORMANCE MODEL

$$\frac{I_P}{k}\ddot{A}_O + \frac{C_d}{k}\dot{A}_O + A_O = \frac{ml}{k}a_I + \frac{(u)M}{k}$$

$$a_{IND} = a_I + (u)a$$

$a_{IND} = \frac{k}{ml}\dot{A}_O$ = indicated output of the accelerometer

$(u)a = \frac{(u)M}{ml}$ = uncertainty in acceleration due to internal error sources

$$(u)a = K_O + (u)K_1\, a_I + K_2\, a_I^2 + K_3\, a_I^3 + K_{IP}\, a_I\, a_P$$

$$a_{IND} = K_O + K_1\, a_I + K_2\, a_I^2 + K_3\, a_I^3 + K_{IP}\, a_I\, a_P$$

$K_1 = (1 + (u)K_1)$

$(u)K_1$ = uncertainty in the accelerometer scale factor

PLATFORM PERFORMANCE MODELS

$$\underline{\omega}_{rp} = -\underline{\omega}_O + \underline{\delta\omega}_{SERVO}$$

$$\omega^p_{rp} = \begin{bmatrix} \omega_{xp} \\ \omega_{yp} \\ \omega_{zp} \end{bmatrix} = -C^p_s \left[\omega^s_{CMD} + \omega^s_D\right] - C^p_r\, \omega^r_{ir} + \omega^p_{SERVO}$$

$$\begin{bmatrix} \omega_{xp} \\ \omega_{yp} \\ \omega_{zp} \end{bmatrix} = -C^p_{SGX}\left[\omega^{SGX}_{CMD} + \omega^{SGX}_D\right] - C^p_{SGY}\left[\omega^{SGY}_{CMD} + \omega^{SGY}_D\right]$$

$$- C^p_{SGZ}\left[\omega^{SGZ}_{CMD} + \omega^{SGZ}_D\right] - C^p_r\, \omega^r_{ir} + \omega^p_{SERVO}$$

$$\begin{bmatrix}\omega_{xp}\\ \omega_{yp}\\ \omega_{zp}\end{bmatrix} = -C^{p}_{SGX}\left\{\begin{bmatrix}D_{FX}\\ 0\\ 0\end{bmatrix} + \begin{bmatrix}D_{IX} & D_{OX} & D_{SX}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{IX}\\ a_{OX}\\ a_{SX}\end{bmatrix} + \begin{bmatrix}D_{IIX} & D_{OOX} & D_{SSX}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a^2_{IX}\\ a^2_{OX}\\ a^2_{SX}\end{bmatrix}\right.$$

$$\left. + \begin{bmatrix}D_{IOX} & D_{ISX} & D_{OSX}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{IX} & a_{OX}\\ a_{IX} & a_{SX}\\ a_{OX} & a_{SX}\end{bmatrix} + \begin{bmatrix}TSF_X & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}t_{gX}\\ 0\\ 0\end{bmatrix}\right\} - C^{p}_{SGY}\left\{\begin{bmatrix}D_{FY}\\ 0\\ 0\end{bmatrix} + \begin{bmatrix}D_{IY} & D_{OY} & D_{SY}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{IY}\\ a_{OY}\\ a_{SY}\end{bmatrix}\right.$$

$$\left. + \begin{bmatrix}D_{IIY} & D_{OOY} & D_{SSY}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a^2_{IY}\\ a^2_{OY}\\ a^2_{SY}\end{bmatrix} + \begin{bmatrix}D_{IOY} & D_{ISY} & D_{OSY}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{IY} & a_{OY}\\ a_{IY} & a_{SY}\\ a_{OY} & a_{SY}\end{bmatrix} + \begin{bmatrix}0 & TSF_Y & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}0\\ t_{gY}\\ 0\end{bmatrix}\right\} - C^{p}_{SGZ}\left\{\begin{bmatrix}D_{FZ}\\ 0\\ 0\end{bmatrix}\right.$$

$$+ \begin{bmatrix}D_{IZ} & D_{OZ} & D_{SZ}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{IZ}\\ a_{OZ}\\ a_{SZ}\end{bmatrix} + \begin{bmatrix}D_{IIZ} & D_{OOZ} & D_{SSZ}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a^2_{IZ}\\ a^2_{OZ}\\ a^2_{SZ}\end{bmatrix} + \begin{bmatrix}D_{IOZ} & D_{ISZ} & D_{OSZ}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{IZ} & a_{OZ}\\ a_{IZ} & a_{SZ}\\ a_{OZ} & a_{SZ}\end{bmatrix}$$

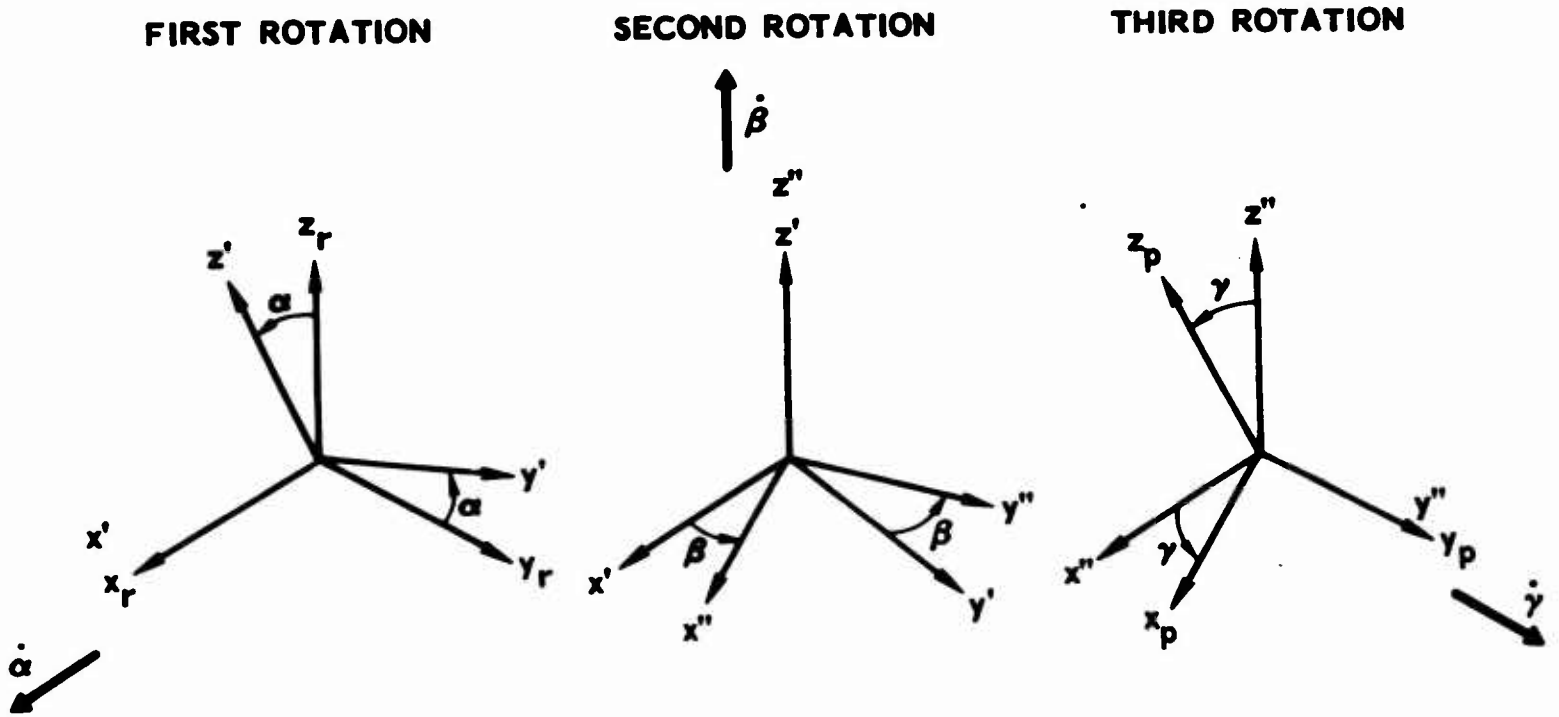
$$\left. + \begin{bmatrix}0 & 0 & TSF_Z\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}0\\ 0\\ t_{gZ}\end{bmatrix}\right\} - C^{p}_{r}\,\omega^{r}_{ir} + \omega^{p}_{SERVO}$$

$$A^s_{IND} = \begin{bmatrix} A_X \\ A_Y \\ A_Z \end{bmatrix} = \begin{bmatrix} K_{OX} \\ K_{OY} \\ K_{OZ} \end{bmatrix} + \begin{bmatrix} K_{1X} & 0 & 0 \\ 0 & K_{1Y} & 0 \\ 0 & 0 & K_{1Z} \end{bmatrix} \begin{bmatrix} a_{IAX} \\ a_{IAY} \\ a_{IAZ} \end{bmatrix} + \begin{bmatrix} K_{2X} & 0 & 0 \\ 0 & K_{2Y} & 0 \\ 0 & 0 & K_{2Z} \end{bmatrix} \begin{bmatrix} a^2_{IAX} \\ a^2_{IAY} \\ a^2_{IAZ} \end{bmatrix}$$

$$+ \begin{bmatrix} K_{3X} & 0 & 0 \\ 0 & K_{3Y} & 0 \\ 0 & 0 & K_{3Z} \end{bmatrix} \begin{bmatrix} a^3_{IAX} \\ a^3_{IAY} \\ a^3_{IAZ} \end{bmatrix} + \begin{bmatrix} K_{IPX} & 0 & 0 \\ 0 & K_{IPY} & 0 \\ 0 & 0 & K_{IPZ} \end{bmatrix} \begin{bmatrix} a_{IAX} & a_{PX} \\ a_{IAY} & a_{PY} \\ a_{IAZ} & a_{PZ} \end{bmatrix}$$

APPLIED ACCELERATION

$$A^s_{is} = C^s_p C^p_r G^e + C^s_p \left[2 w^{ek}_{ie} \, w^{pk}_{rp} \, R^p_{ps} + \dot{w}^{pk}_{rp} \, R^p_{ps} + w^{pk}_{rp} \, w^{pk}_{rp} \, R^p_{ps} \right]$$



x_r, y_r, z_r - R FRAME

x_p, y_p, z_p - INNER PLATFORM FRAME

Euler Transformation

$$C_r^p = \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix}$$

$$\dot{\alpha} = (\omega_{xp} \cos\gamma + \omega_{zp} \sin\gamma)/\cos\beta$$

$$\dot{\beta} = \omega_{zp} \cos\gamma - \omega_{xp} \sin\gamma$$

$$\dot{\gamma} = (\omega_{xp} \cos\gamma \sin\beta + \omega_{zp} \sin\gamma \sin\beta)/\cos\beta + \omega_{yp}$$

$$E_{11} = \cos\gamma \cos\beta$$

$$E_{12} = \cos\gamma \sin\beta \cos\alpha + \sin\gamma \sin\alpha$$

$$E_{13} = \cos\gamma \sin\beta \sin\alpha - \sin\gamma \cos\alpha$$

$$E_{21} = -\sin\beta$$

$$E_{22} = \cos\beta \cos\alpha$$

$$E_{23} = \cos\beta \sin\alpha$$

$$E_{31} = \sin\gamma \cos\beta$$

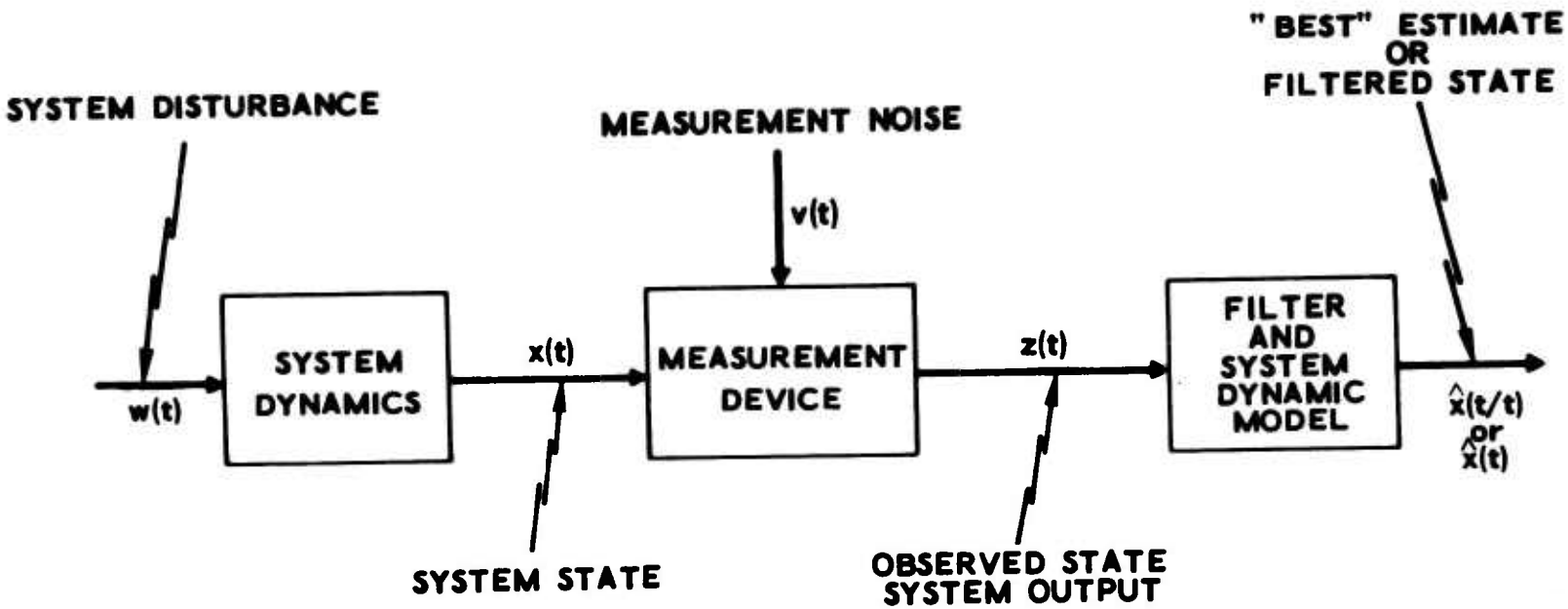
$$E_{32} = \sin\gamma \sin\beta \cos\alpha - \cos\gamma \sin\alpha$$

$$E_{33} = \sin\gamma \sin\beta \sin\alpha + \cos\gamma \cos\alpha$$

$$0 \leq \alpha < 2\pi$$

$$-\frac{\pi}{2} < \beta < \frac{\pi}{2}$$

$$0 \leq \gamma < 2\pi$$



State Estimation Problem

States of the Calibration Process

$X_1 = D_{FX}$	$X_{17} = D_{OOY}$	$X_{33} = TSF_Z$	$X_{49} = K_{3X}$
$X_2 = D_{FY}$	$X_{18} = D_{OOZ}$	$X_{34} = \eta_{x_i}$	$X_{50} = K_{3Y}$
$X_3 = D_{FZ}$	$X_{19} = D_{SSX}$	$X_{35} = \eta_{x_o}$	$X_{51} = K_{3Z}$
$X_4 = D_{IX}$	$X_{20} = D_{SSY}$	$X_{36} = \eta_{y_i}$	$X_{52} = K_{IPX}$
$X_5 = D_{IY}$	$X_{21} = D_{SSZ}$	$X_{37} = \eta_{y_o}$	$X_{53} = K_{IPY}$
$X_6 = D_{IZ}$	$X_{22} = D_{IOX}$	$X_{38} = \eta_{z_i}$	$X_{54} = K_{IPZ}$
$X_7 = D_{OX}$	$X_{23} = D_{IOY}$	$X_{39} = \eta_{z_o}$	$X_{55} = \psi_{x_i}$
$X_8 = D_{OY}$	$X_{24} = D_{IOZ}$	$X_{40} = K_{OX}$	$X_{56} = \psi_{x_o}$
$X_9 = D_{OZ}$	$X_{25} = D_{ISX}$	$X_{41} = K_{OY}$	$X_{57} = \psi_{y_i}$
$X_{10} = D_{SX}$	$X_{26} = D_{ISY}$	$X_{42} = K_{OZ}$	$X_{58} = \psi_{y_o}$
$X_{11} = D_{SY}$	$X_{27} = D_{ISZ}$	$X_{43} = K_{1X}$	$X_{59} = \psi_{z_i}$
$X_{12} = D_{SZ}$	$X_{28} = D_{OSX}$	$X_{44} = K_{1Y}$	$X_{60} = \psi_{z_o}$
$X_{13} = D_{IIX}$	$X_{29} = D_{OSY}$	$X_{45} = K_{1Z}$	$X_{61} = \alpha$
$X_{14} = D_{IIY}$	$X_{30} = D_{OSZ}$	$X_{46} = K_{2X}$	$X_{62} = \beta$
$X_{15} = D_{IIZ}$	$X_{31} = TSF_X$	$X_{47} = K_{2Y}$	$X_{63} = \gamma$
$X_{16} = D_{OOX}$	$X_{32} = TSF_Y$	$X_{48} = K_{2Z}$	

EXTENDED KALMAN FILTER EQUATIONS

$$\dot{x}_i(t) = w_i(t) \qquad [i = 1, 2, \ldots, 60]$$

$x_i(t_o)$ - is a gaussian random variable representing the initial value of the ith state with mean $\bar{x}_i(t_o) \equiv x_{o_i} = E\left\{x_i(t_o)\right\}$ and variance $\sigma_{x_{o_i}}^2 = p_{o_i} = E\left\{\left[x_i(t_o) - x_{o_i}\right]^2\right\}$

$E\ \{\ \}$ - is the expectation operator

$\{w_i(t), t \geq t_o\}$ - is a zero mean, gaussian white process (i.e., time uncorrelated) representing the ith state disturbance with covariance $c_{w_i}(t, \tau) = E\left\{w_i(t)w_i(\tau)\right\} = \sigma_{w_i}^2 \delta(t - \tau)$

$\delta(t - \tau)$ - is the Dirac delta function

$E\left\{\left[x_i(t_o) - x_{o_i}\right]\left[w_i(t)\right]\right\} = 0$ [i.e., the initial state, $x_i(t_o)$ is independent of the disturbance process $w_i(t)$]

$$\dot{x}_{61}(t) = \Big[\{\omega_{xp}(x) \cos [x_{63}(t)] + \omega_{zp}(x) \sin [x_{63}(t)]\} / \cos [x_{62}(t)]\Big] + w_{61}(t)$$

$$\dot{x}_{62}(t) = \{\omega_{zp}(x) \cos [x_{63}(t)] - \omega_{xp}(x) \sin [x_{63}(t)]\} + w_{62}(t)$$

$$\dot{x}_{63}(t) = \Big[\{\omega_{xp}(x) \cos [x_{63}(t)] \sin [x_{62}(t)] + \omega_{zp}(x) \sin [x_{63}(t)] \sin [x_{62}(t)]\} / \cos [x_{62}(t)] + \omega_{yp}(x)\Big] + w_{63}(t)$$

$$\dot{\underline{x}} = \underline{f}(\underline{x}, t) + \underline{w}(t)$$

$$\underline{a}_{IND} = \underline{h}(\underline{x}, t) + \underline{m}(t)$$

$\underline{a}_{IND}$ - is a 3-element indicated acceleration vector

$\underline{h}(\underline{x}, t)$ - is a 3-dimensional vector-valued, nonlinear, function of the state

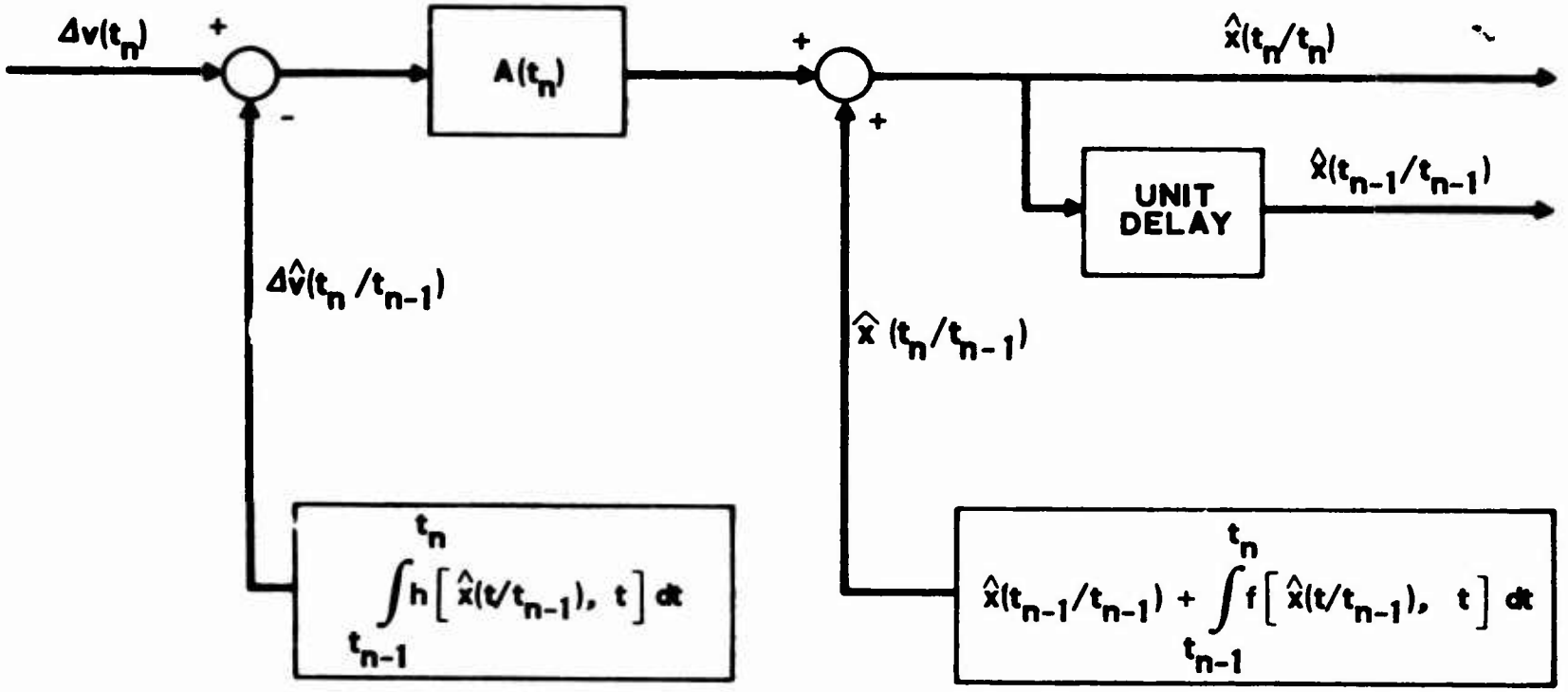
$\underline{m}(t)$ - is a 3-element measurement noise vector

$$\underline{v}(t + \Delta t) - \underline{v}(t) = \Delta \underline{v} = \underline{h}(\underline{x}, t)\Delta t + \underline{m}(t)\Delta t$$

Δt - is the time interval between measurements

Δv - is the 3 element vector velocity change over the measurement interval

$$\Delta \underline{v}(t_n) = \underline{h}[\underline{x}(t_n), t_n]\, \Delta t + \underline{m}(t_n)\Delta t$$



Schematic of the Extended Kalman Filter Algorithm

SIMULATION OF A 1-g CALIBRATION

States of the 1-g Simulation

Symbol	Description	Units	State Designation
D_{FX}	X - gyro fixed drift	meru	x_1
D_{FY}	Y - gyro fixed drift	meru	x_2
D_{FZ}	Z - gyro fixed drift	meru	x_3
D_{IX}	X - gyro input axis unbalance	meru/g	x_4
D_{IY}	Y - gyro input axis unbalance	meru/g	x_5
D_{IZ}	Z - gyro input axis unbalance	meru/g	x_6
K_{0X}	X - accelerometer bias	μg	x_7
K_{0Y}	Y - accelerometer bias	μg	x_8
K_{0Z}	Z - accelerometer bias	μg	x_9
K_{1X}	X - accelerometer scale factor	μg/g	x_{10}
K_{1Y}	Y - accelerometer scale factor	μg/g	x_{11}
K_{1Z}	Z - accelerometer scale factor	μg/g	x_{12}
η_{x_o}	X - gyro out-of-plane misalignment	arc-sec	x_{13}
ψ_{x_i}	X - accelerometer in-plane misalignment	arc-sec	x_{14}
α	Euler angle about X platform axis	arc-sec	x_{15}
β	Euler angle about Z platform axis	arc-sec	x_{16}
γ	Euler angle about Y platform axis	arc-sec	x_{17}

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{w}(t)$$

$$f_{15} = (\omega_{xp} \cos x_{17} + \omega_{zp} \sin x_{17})/\cos x_{16}$$

$$f_{16} = \omega_{zp} \cos x_{17} - \omega_{xp} \sin x_{17}$$

$$f_{17} = (\omega_{xp} \cos x_{17} \sin x_{16} + \omega_{zp} \sin x_{17} \sin x_{16})/\cos x_{16} + \omega_{yp}$$

$$\omega_{xp} = -x_1 - x_4 E_{13} + x_4 x_{13} E_{33} - TGX$$

$$- E_{12} WEC - E_{13} WES$$

$$\omega_{yp} = -x_2 - x_5 E_{23} - TGY - E_{22} WEC - E_{23} WES$$

$$\omega_{zp} = x_1 x_{13} + x_{13} x_4 E_{13} + x_{13} TGX - x_3$$

$$-x_6 E_{33} - TGZ - E_{32} WEC - E_{33} WES$$

$$h_1 = x_7(t_n) + x_{10}(t_n) E_{13}(t_n) - x_{10}(t_n) x_{14}(t_n) E_{23}(t_n)$$

$$h_2 = x_8(t_n) + x_{11}(t_n) E_{23}(t_n)$$

$$h_3 = x_9(t_n) + x_{12}(t_n) E_{33}(t_n)$$

Carousel VB 1-Sigma Error Budget

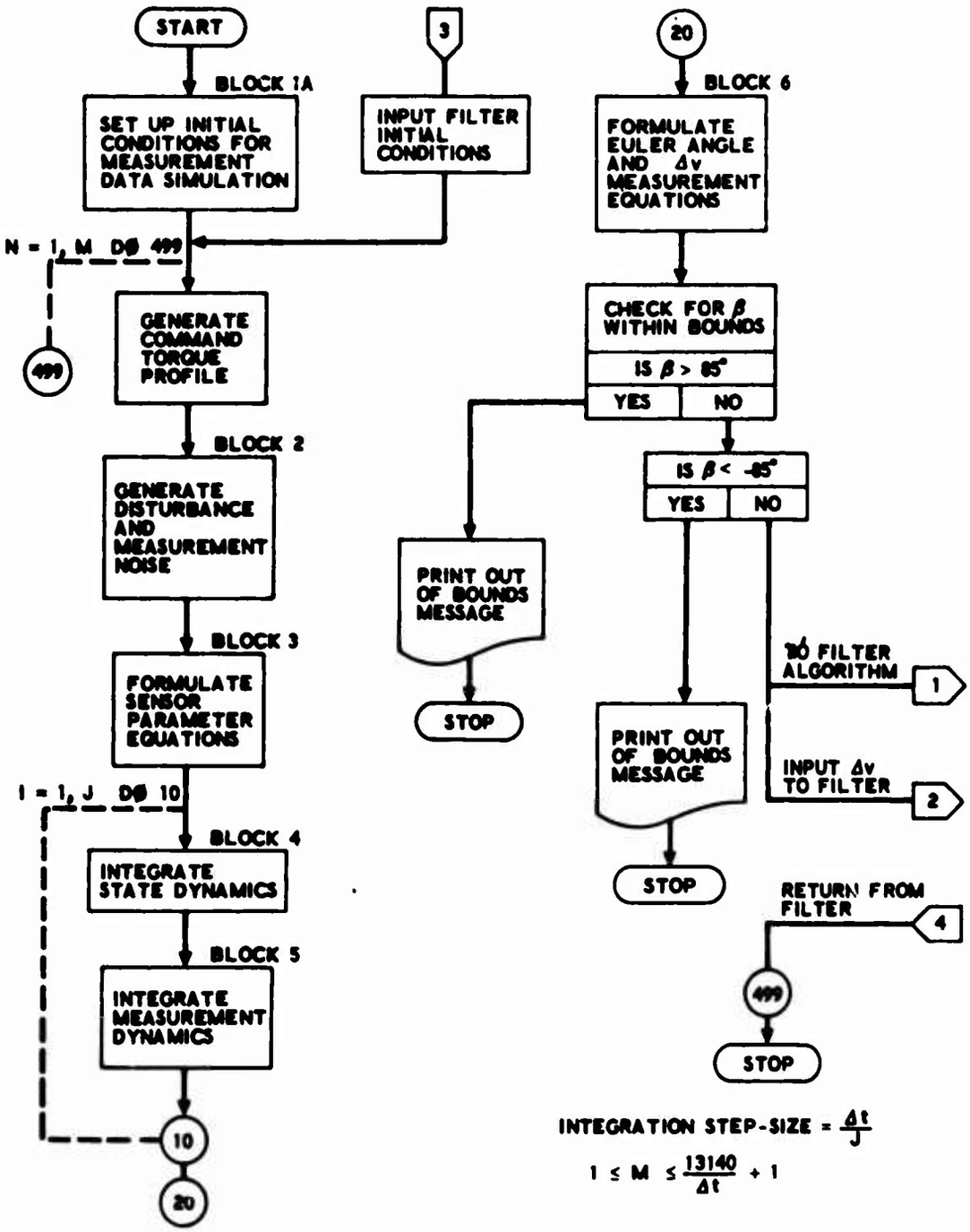
Error Sources	1-Sigma Uncertainty
Initial Attitude	
Azimuth (z - axis)	25 arc-sec
Level (x & y axes)	5 arc-sec
Accelerometer (each sensor)	
Bias	25 - 86.6 μg
Scale factor	25 - 86.6 μg/g
Input axis alignment	13 arc-sec
Gyro (each sensor)	
Fixed drift	2.0 - 6.93 meru
Unbalance drift	2.25 - 7.80 meru/g
In-axis compliance	0.208 meru/g^2
Cross-axis compliance	0.087 meru/g^2
Input axis alignment	17.3 arc-sec

Initial State Means and Standard Deviations

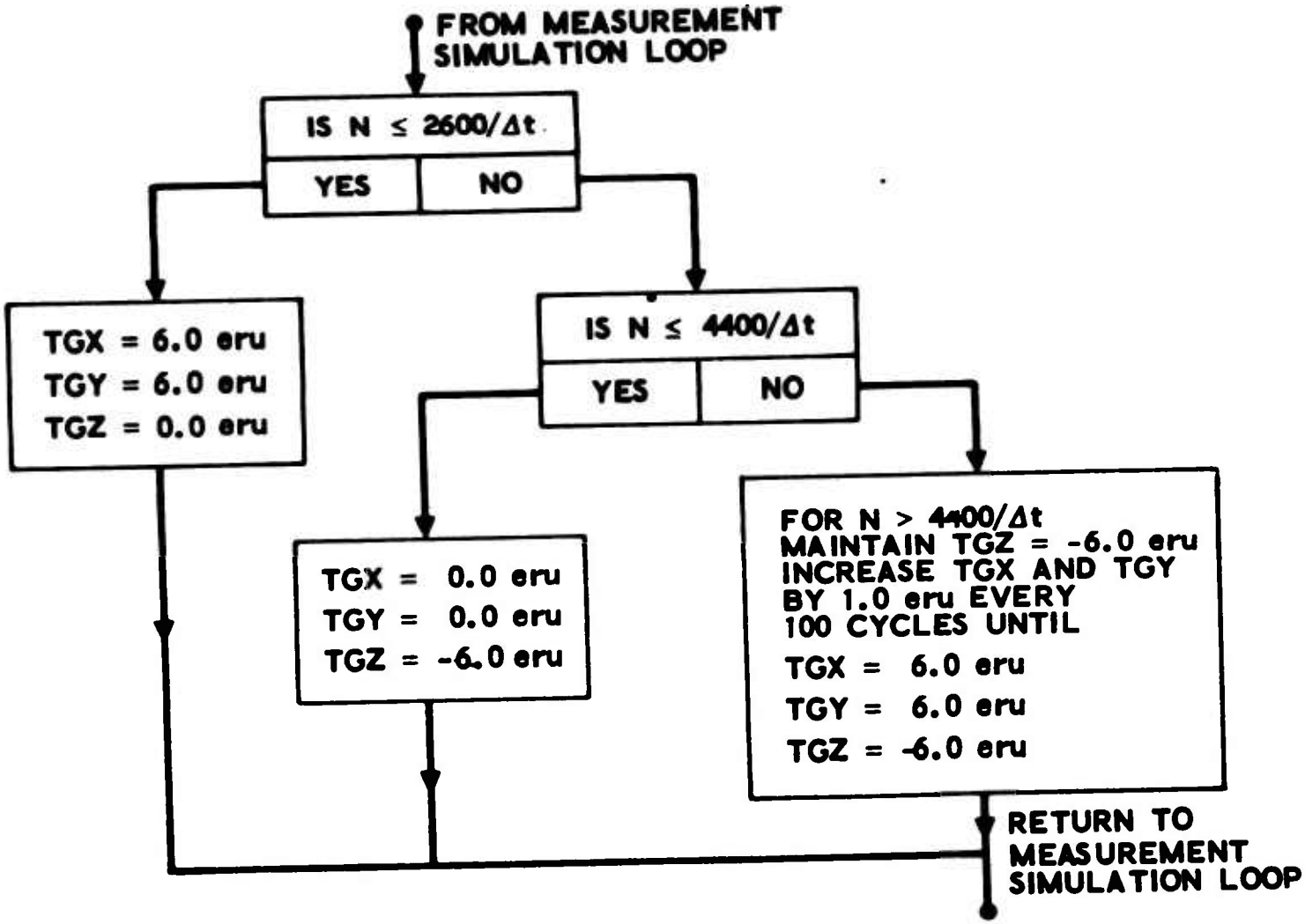
State	Symbol	Mean Value (x_o)	Standard Deviation (σ_{x_o})	Units
x_1	D_{FX}	490.0	6.0	meru
x_2	D_{FY}	520.0	5.0	meru
x_3	D_{FZ}	110.0	4.0	meru
x_4	D_{IX}	72.0	4.0	meru/g
x_5	D_{IY}	68.0	5.0	meru/g
x_6	D_{IZ}	62.0	6.0	meru/g
x_7	K_{0X}	3500.0	30.0	μg
x_8	K_{0Y}	4250.0	40.0	μg
x_9	K_{0Z}	6310.0	60.0	μg
x_{10}	K_{1X}	999840.0	30.0	μg/g
x_{11}	K_{1Y}	999900.0	27.0	μg/g
x_{12}	K_{1Z}	998900.0	70.0	μg/g
x_{13}	η_{x_o}	260.0	15.0	arc-sec
x_{14}	ψ_{x_i}	305.0	10.0	arc-sec
x_{15}	α	180.0	5.0	arc-sec
x_{16}	β	110.0	5.0	arc-sec
x_{17}	γ	120.0	5.0	arc-sec

Standard Deviations of Disturbance and Measurement Noise

Disturbance (w) No.	Standard Deviation (σ_w)	Units	Measurement Noise (m) No.	Standard Deviation (σ_m)	Units
1	0.6	meru	1	13.0	μg-sec
2	0.5	meru	2	13.0	μg-sec
3	0.4	meru	3	13.0	μg-sec
4	0.4	meru/g			
5	0.5	meru/g			
6	0.6	meru/g			
7	3.0	μg			
8	4.0	μg			
9	6.0	μg			
10	3.0	μg/g			
11	2.7	μg/g			
12	7.0	μg/g			
13	1.5	arc-sec			
14	1.0	arc-sec			
15	5.0	arc-sec			
16	5.0	arc-sec			
17	5.0	arc-sec			



Flow Diagram of Measurement Simulation Program



Command Torque Generation Block

$\hat{x}^*(N) = \hat{x}(t_n|t_{n-1})$ (17 × 1) predicted state vector

$\Delta v(N)$ - (3 × 1) measurement data vector from measurement simulation program

$\Delta\hat{v}(N)$ - (3 × 1) predicted measurement vector

$A(N)$ - (17 × 3) Kalman gain matrix

$F(N-1) = F[\hat{x}(t_{n-1}|t_{n-1})]$ (17 × 17) matrix of state dynamic partials

$H(N) = H[\hat{x}(t_n|t_{n-1})]$ (3 × 17) matrix of measurement dynamic partials

$\Phi(N, N-1) = \Phi[t_n, t_{n-1}; \hat{x}(t_{n-1}|t_{n-1})]$ (17 × 17) state transition matrix

$\Gamma(N, N-1) = \Gamma[t_n, t_{n-1}; \hat{x}(t_{n-1}|t_{n-1})]$ (17 × 17) disturbance transition matrix

$\overline{P}(N-1) = P(t_{n-1}|t_{n-1})$ (17 × 17) filtering error covariance matrix at time t_{n-1}

$\overline{P}(N) = P(t_n|t_n)$ (17 × 17) filtering error covariance matrix at time t_n

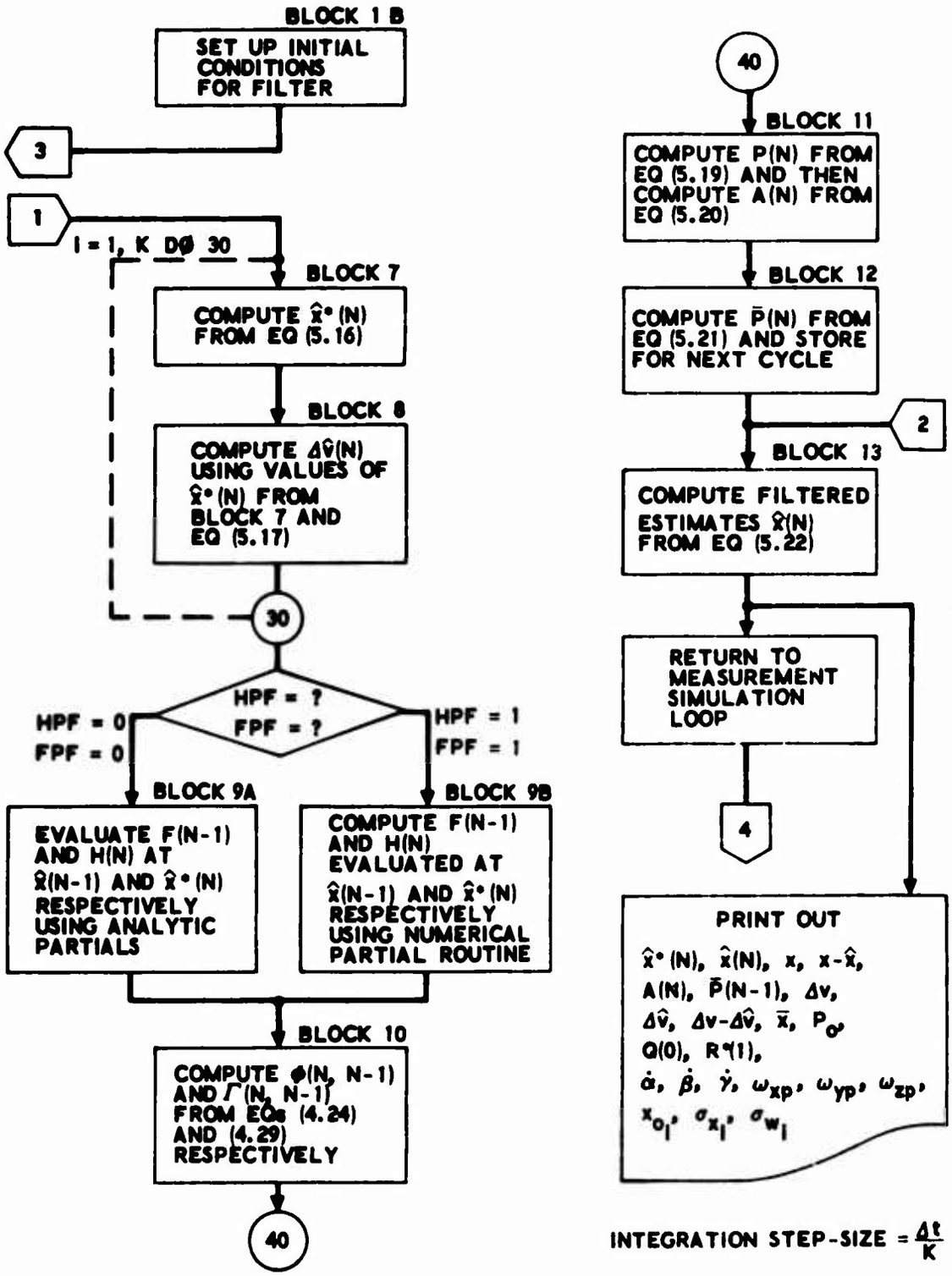
$P(N) = P(t_n|t_{n-1})$ (17 × 17) prediction error covariance matrix

$Q(N-1)$ - given (17 × 17) disturbance covariance matrix

$R^*(N)$ - given (3 × 3) measurement error covariance matrix

$\overline{x}_o$ - given mean value initial state vector (17 × 1)

P_o - given initial state covariance matrix (17 × 17)



Flow Diagram of Filter Algorithm

Values of Initial State Estimate for 1-g Simulation

State Variable	True x_o	$\bar{x}_o$ 10% of x_o	$\bar{x}_o$ 30% of x_o	$\bar{x}_o$ 60% of x_o	$\bar{x}_o$ 100% of x_o	Units
D_{FX}	490.0	441.0	343.0	784.0	0.0	meru
D_{FY}	520.0	572.0	364.0	208.0	0.0	meru
D_{FZ}	110.0	99.0	77.0	44.0	0.0	meru
D_{IX}	72.0	79.2	50.4	115.2	0.0	meru/g
D_{IY}	68.0	61.2	88.4	27.2	0.0	meru/g
D_{IZ}	62.0	55.8	80.6	24.8	0.0	meru/g
K_{0X}	3500.0	3150.0	2450.0	5600.0	0.0	μg
K_{0Y}	4250.0	4675.0	5525.0	6800.0	0.0	μg
K_{0Z}	6310.0	5679.0	8203.0	2524.0	0.0	μg
K_{1X}	999840.0	1000000.0	1000000.0	1000000.0	1000000.0	μg/g
K_{1Y}	999900.0	1000000.0	1000000.0	1000000.0	1000000.0	μg/g
K_{1Z}	998900.0	1000000.0	1000000.0	1000000.0	1000000.0	μg/g
η_{x_o}	260.0	286.0	182.0	416.0	0.0	arc-sec
ψ_{x_i}	305.0	274.0	396.0	122.0	0.0	arc-sec
α	180.0	162.0	126.0	72.0	0.0	arc-sec
β	110.0	99.0	143.0	44.0	0.0	arc-sec
γ	120.0	108.0	84.0	192.0	0.0	arc-sec

Results of 10-Percent Case for 1-g Simulation

State Variable	$\hat{x}$	σ	x	$x-\hat{x}$	Error Budget	Units
D_{FX}	496.773	0.516	497.633	0.859	2.0 - 6.93	meru
D_{FY}	513.773	0.368	514.110	0.337	2.0 - 6.93	meru
D_{FZ}	111.850	0.383	111.448	-0.402	2.0 - 6.93	meru
D_{IX}	79.111	1.730	79.310	0.200	2.25 - 7.80	meru/g
D_{IY}	71.280	1.800	71.826	0.545	2.25 - 7.80	meru/g
D_{IZ}	59.402	2.082	60.219	0.818	2.25 - 7.80	meru/g
K_{0X}	3516.932	6.335	3504.268	-12.664	25.0 - 86.6	μg
K_{0Y}	4239.670	13.400	4191.698	-47.972	25.0 - 86.6	μg
K_{0Z}	6243.403	5.198	6242.021	-1.382	25.0 - 86.6	μg
K_{1X}	999889.55	7.12	999873.51	-16.04	25.0 - 86.6	μg/g
K_{1Y}	999957.72	14.15	999905.84	-51.87	25.0 - 86.6	μg/g
K_{1Z}	998969.16	7.47	998953.54	-15.62	25.0 - 86.6	μg/g
η_{x_o}	257.454	7.443	245.356	-12.099	17.3	arc-sec
ψ_{x_i}	310.528	2.514	315.715	5.187	13.0	arc-sec
α	-609446.50	9.44	-609458.58	-12.08		arc-sec
β	-57158.61	168.98	-57091.14	67.46		arc-sec
γ	-169021.82	33.31	-169004.95	16.87		arc-sec

N = 6200 cycles

Time = 1 hr 43 min 20 sec

Accelerometer Parameter Correlation Pattern for 1-g Simulation

	K_{0X}	K_{0Y}	K_{0Z}	K_{1X}	K_{1Y}	K_{1Z}	ψ_{x_i}
K_{0X}	*	0.64	0.46	0.32	0.57	0.45	0.59
K_{0Y}	0.64	*	0.29	0.39	0.94	0.54	0.39
K_{0Z}	0.46	0.29	*	0.36	0.096	0.082	0.12
K_{1X}	0.32	0.39	0.36	*	0.57	0.22	0.75
K_{1Y}	0.57	0.94	0.096	0.57	*	0.39	0.53
K_{1Z}	0.45	0.54	0.082	0.22	0.39	*	0.12
ψ_{x_i}	0.59	0.39	0.12	0.75	0.53	0.12	*

α	ρ
0°	0.0033
-78°	0.52
-90°	0.79
-106°	0.97
-169°	0.94 (N = 6200)
-223° (43° above horizontal)	0.71

Results of 30-Percent Case for 1-g Simulation

State Variable	$\hat{x}$	σ	x	x-$\hat{x}$	Error Budget	Units
D_{FX}	496.635	0.516	497.633	0.998	2.0 - 6.93	meru
D_{FY}	513.557	0.368	514.110	0.553	2.0 - 6.93	meru
D_{FZ}	111.683	0.383	111.448	-0.235	2.0 - 6.93	meru
D_{IX}	79.253	1.727	79.310	0.058	2.25 - 7.80	meru/g
D_{IY}	71.160	1.797	71.826	0.665	2.25 - 7.80	meru/g
D_{IZ}	59.359	2.077	60.219	0.860	2.25 - 7.80	meru/g
K_{0X}	3508.775	6.335	3504.268	-4.506	25.0 - 86.6	μg
K_{0Y}	4260.383	13.408	4191.698	-68.685	25.0 - 86.6	μg
K_{0Z}	6245.022	5.198	6242.021	-3.001	25.0 - 86.6	μg
K_{1X}	999890.08	7.12	999873.51	-16.57	25.0 - 86.6	μg/g
K_{1Y}	999979.94	14.16	999905.84	-74.09	25.0 - 86.6	μg/g
K_{1Z}	998975.52	7.47	998953.54	-21.98	25.0 - 86.6	μg/g
η_{x_o}	245.112	7.445	245.356	0.244	17.3	arc-sec
ψ_{x_i}	314.863	2.515	315.715	0.852	13.0	arc-sec
α	-609440.88	9.43	-609458.58	-17.70		arc-sec
β	-57193.37	168.56	-57091.14	102.22		arc-sec
γ	-169029.41	33.24	-169004.95	24.46		arc-sec

N = 6200 cycles

Time = 1 hr 43 min 20 sec

Results of 60-Percent Case for 1-g Simulation

State Variable	$\hat{x}$	σ	x	$x-\hat{x}$	% of x	Units
D_{FX}	495.466	0.527	497.681	2.215*	0.4	meru
D_{FY}	514.717	0.283	514.109	-0.608*	0.1	meru
D_{FZ}	113.644	0.391	111.485	-2.159*	1.9	meru
D_{IX}	80.365	1.640	79.249	-1.116*	1.4	meru/g
D_{IY}	71.319	1.746	71.990	0.671*	0.9	meru/g
D_{IZ}	59.279	2.018	60.244	0.965*	1.6	meru/g
K_{0X}	3650.730	6.179	3503.438	-147.292	4.2	μg
K_{0Y}	4498.232	12.183	4192.156	-306.076	7.3	μg
K_{0Z}	6195.416	5.401	6241.252	45.836*	0.7	μg
K_{1X}	999940.03	6.85	999873.30	-66.73*	0.007	μg/g
K_{1Y}	1000209.24	12.80	999905.98	-303.26	0.03	μg/g
K_{1Z}	999062.60	7.63	998953.02	-109.59	0.01	μg/g
η_{x_o}	312.307	7.375	245.152	-67.155	27.4	arc-sec
ψ_{x_i}	278.137	2.493	315.543	37.406	11.9	arc-sec
α	-665368.96	5.28	-665437.69	-68.73		arc-sec
β	-60438.96	173.58	-60369.18	69.78		arc-sec
γ	-187580.62	15.55	-187534.86	45.76		arc-sec

N = 6600 cycles

Time = 1 hr 50 min

*Indicates $x-\hat{x}$ ($\tilde{x}$) is within error budget

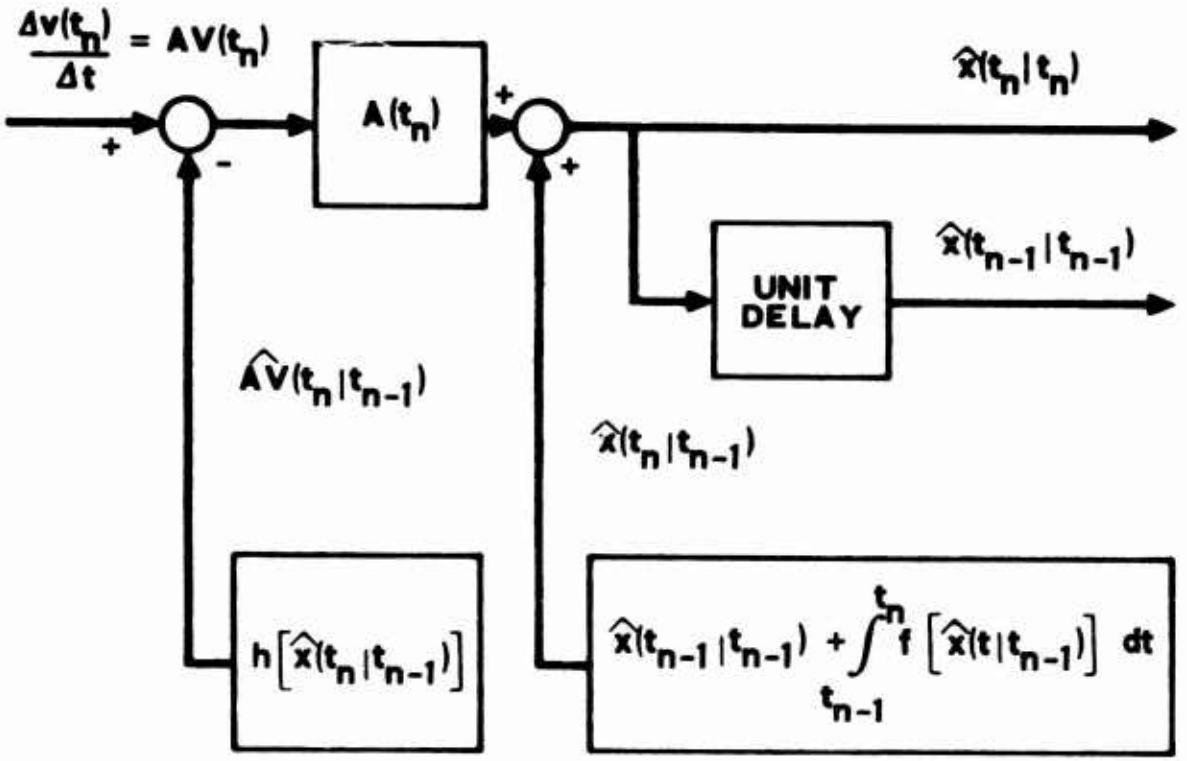
Results of 100-Percent Case for 1-g Simulation

State Variable	$\hat{x}$	σ	x	$x-\hat{x}$	% of x	Units
D_{FX}	494.927	0.442	497.508	2.580*	0.5	meru
D_{FY}	512.597	0.237	514.138	1.541*	0.3	meru
D_{FZ}	112.255	0.418	111.747	-0.508*	0.4	meru
D_{IX}	67.379	1.049	79.280	11.901	15.0	meru/g
D_{IY}	60.095	1.111	71.645	11.550	16.2	meru/g
D_{IZ}	46.901	1.219	60.549	13.648	22.5	meru/g
K_{0X}	3436.020	5.298	3503.198	67.177*	1.9	μg
K_{0Y}	4111.232	6.570	4192.535	81.302*	1.9	μg
K_{0Z}	6288.163	5.881	6242.950	-45.213*	0.7	μg
K_{1X}	999852.38	6.67	999872.76	20.38*	0.002	μg/g
K_{1Y}	999837.07	7.00	999906.32	69.25*	0.007	μg/g
K_{1Z}	998927.07	7.28	998954.23	27.16*	0.003	μg/g
η_{x_o}	187.508	7.193	245.279	57.771	23.5	arc-sec
ψ_{x_i}	330.994	2.365	315.684	-15.310	4.8	arc-sec
α	-802561.57	36.24	-802154.33	407.24		arc-sec
β	-84749.28	88.35	-83832.86	916.41		arc-sec
γ	-231740.46	89.92	-232708.19	-967.72		arc-sec

N = 7600 cycles

Time = 2 hr 6 min 40 sec

*Indicates $x-\hat{x}$ ($\tilde{x}$) is within error budget



Schematic of the Alternate Extended Kalman Filter Algorithm

Comparison of Basic and Alternate Filters for Δv Approximation

State Variable	$\tilde{x}$ of Basic Filter	$\tilde{x}$ of Alternate Filter	Units
D_{FX}	0.998	25.877	meru
D_{FY}	0.553	8.079	meru
D_{FZ}	-0.235	-10.311	meru
D_{IX}	0.058	90.223	meru/g
D_{IY}	0.665	89.070	meru/g
D_{IZ}	0.860	110.043	meru/g
K_{0X}	-4.506	-21.171*	μg
K_{0Y}	-68.685	-128.034	μg
K_{0Z}	-3.001	-7.856*	μg
K_{1X}	-16.57	-54.03*	μg/g
K_{1Y}	-74.09	-155.81	μg/g
K_{1Z}	-21.98	-18.79*	μg/g
η_{x_o}	0.244	0.608*	arc-sec
ψ_{x_i}	0.852	15.157	arc-sec
α	-17.70	-634.69	arc-sec
β	102.22	9230.82	arc-sec
γ	24.46	1825.49	arc-sec

N = 6200 cycles Time = 1 hr 43 min 20 sec

*Indicates parameters of Alternate Filter where $\tilde{x}$ is within error budget.

Results of 30-Percent Case Basic Option

State Variable	$\hat{x}$	σ	x	$\tilde{x}$	Error Budget	Units
D_{IX}	72.4021	0.0125	72.6609	0.2588	2.25 - 7.80	meru/g
K_{1X}	999802.45	0.03	999801.66	-0.79	25.0 - 86.6	$\mu g/g$
D_{OOY}	4.1600156	0.0002020	4.1599216	-0.0000940	0.208	$meru/g^2$
D_{ISZ}	2.878289	0.000609	2.869569	-0.008719	0.087	$meru/g^2$
K_{2Y}	15.65374	0.00414	15.69673	0.04299	0.1	$\mu g/g^2$
K_{3Z}	3.567258	0.000350	3.560426	-0.006832	0.01	$\mu g/g^3$
K_{IPX}	3.62399	0.00375	3.74119	0.11720	0.1	$\mu g/g^2$
α	-196257.39	0.01	-196257.50	-0.11		arc-sec
β	-56786.08	0.01	-56785.98	0.10		arc-sec
γ	-218890.68	0.01	-218891.12	-0.44		arc-sec

N = 15,000 cycles

Time = 10 min

CONCLUSIONS

- SIGNIFICANT IMPROVEMENT IN REDUCING THE TIME REQUIRED TO CALIBRATE AN INS
  - 24 TO 48 HOURS -2 HOURS
- PROVIDES CAPABILITY OF ON-LINE SYSTEM CALIBRATIONS
  - PERMITS EVALUATION OF TEST PROCEDURES, SYSTEM MODELS, AND SIGNIFICANCE OF STATE ESTIMATES WHILE SYSTEM UNDER TEST
- THE LINEARIZATION AND THEREFORE THE FILTERING PROCESS IS VALID

- APPLICATION OF OFF-DIAGONAL TERMS PROVIDES A NEW DIMENSION IN ASSESSING FILTER PERFORMANCE WITH RESPECT TO SYSTEM OBSERVABILITY

- ΔV APPROXIMATION IN KALMAN GAIN DOES NOT PREVENT FILTER CONVERGENCE

- CALIBRATION IS QUITE STABLE WITH RESPECT TO INITIAL STATE ESTIMATE

- OPTIMAL TORQUE PROFILE CAN BE CONSTRUCTED USING OFF-DIAGONAL TERMS THAT DEFINE OBSERVABILITY REGIONS WHERE SUBSETS OF MODEL COEFFICIENT ESTIMATES ARE "OPTIMAL"

"Filtering Theory Methods and Applications to the Orbit Determination Problem for Near-Earth Satellites"

by

Kenneth A. Myers, Capt, USAF*
Aerospace Research Laboratories
Wright-Patterson AFB, Ohio

Introduction

A common operational requirement in all U. S. Air Force space applications is orbit determination. Some degree of orbit determination accuracy is essential to perform those tracking or spacecraft guidance, navigation, and control functions which are required to meet particular mission objectives. Operational methods in orbit determination have progressed from deterministic methods[1], used during the early years of the space program, to least squares and differential correction techniques[2], still widely used today, to sequential filtering and smoothing schemes, the subject of much current research[3]. Well proven techniques in batch processing or even deterministic methods may suffice for some operations; however, they can be inadequate (especially for on-line operation) in many modern and sophisticated missions, where accuracy requirements are specified in terms of meters and arc-seconds. In some cases, such as precision pointing control missions and geophysical applications, the orbit determination problem may well become a limiting factor in mission success. There exists an inherent mathematical equivalence between the more conventional batch and the relatively newer sequential processing techniques, but the latter techniques usually converge more rapidly and provide more accurate estimates when used in their extended form with state noise compensation.

Higher accuracy requirements are accompanied by a need for more precise mathematical models for prediction of the spacecraft state. In conflict with this need, however, are limitations imposed not only by an abbreviated knowledge of the true forces acting on the spacecraft but also by computational considerations, such as computation time, storage, and round-off error. For instance,

*Astronautical Engineer, Applied Mathematics Research Laboratory. This research was conducted during the authors' AFIT-sponsored Ph.D. program at the University of Texas at Austin under Dr. Byron D. Tapley, Chairman of the Department of Aerospace Engineering and Engineering Mechanics. Selected portions of this paper are to be published in the AIAA Journal.

current geopotential models provide spherical harmonic coefficients up to degree 22 and order 16; sophisticated dynamical atmospheric models provide density values for drag calculations which account for seasonal-latitudinal variations, diurnal and geomagnetic effects, etc. However, the necessary storage for programs which utilize these models and the computer time required for the numerical integrations is often prohibitive. Regardless of the sophistication of the mathematical model, there will always be unknown forces acting on the spacecraft due to uncertainty in the modeled parameters, tidal forces, higher-order gravitational effects, variations in solar radiation pressure, magnetic disturbances, vehicle outgassing, and so forth ad infinitum.

The fundamental limitations of model accuracy and computational load can be offset by employing observational data and well known techniques in discrete stochastic filtering theory for accurately determining the spacecraft state[4]. This approach is made possible due to recent developments in three major areas: 1) Mathematical theories in optimal linear filtering have matured[3]. 2) The computations required in typical filtering methods can easily be accommodated with modern ground-based as well as airborne and spaceborne computers[5]. 3) High precision observations are available with current space hardware and radar instrumentation[6].

Despite the recent developments cited above, several difficulties arise with the filtering approach, as is recognized from the vast technical literature devoted to this area. Generally, these difficulties can be divided into four categories for both the state and observation processes: 1) Linearization errors; 2) Model errors; 3) Computation errors; and, 4) Unknown noise statistics. It is the intent of this study to investigate techniques which are designed to compensate for each of these filtering difficulties through computer simulation of the orbit determination problem for a typical near-earth satellite.

Description of the Simulation

Simulation results for this investigation were obtained on the CDC 6600 digital computer at the University of Texas at Austin. Observations are generated in the program from a set of true equations of motion and are corrupted by the addition of white, Gaussian noise. An extended Kalman filter is used for estimating the true spacecraft state vector based on the simulated observations. Translational equations of motion for the spacecraft are expressed in an earth-centered, inertial (ECI) coordinate system, where the X-Y plane lies

in the earth equatorial plane with the X- and Z- axes directed to the mean equinox and north pole, respectively. The basic equations of motion are integrated with an efficient Runge-Kutta-Fehlberg algorithm. The perturbing acceleration for the simulated real-world equations of motion includes the earth's gravitational field (6th-order/4th-degree geopotential model), two-body attractions of the sun and moon, and atmospheric drag. Perturbing acceleration for the filter-world equations of motion includes only the effect of earth oblateness (J_2). This provides a reasonable mismatch between the filter-world and real-world models which permits an evaluation of the filter performance in the presence of realistic dynamical model errors. Furthermore, the inclusion of J_2 in the filter-world model is necessary for effective estimation. During prediction phases, when the satellite is out of view of the tracking stations, filtered estimates are propagated with the real-world dynamical model, since a large state error growth between passes would hamper the filter performance analysis.

Radar range ρ and doppler range-rate $\dot{\rho}$ observations are made from ten land-based tracking stations at known locations (see Table 1). Standard deviations

Station Location	East Longitude (deg.)	Latitude (deg.)	Elevation (m.)	Observation Schedule (minutes)
Trinidad	298.5	11.0	245	0-8 , 89-93
Antigua	298.0	17.0	-27	0-9 , 90-94
Eglin AFB	273.8	30.5	45	10-11, 93-101
New Hampshire	288.4	42.9	211	8-16, 98-100
Cold Lake	250.0	54.5	654	15-20, 101-108
Clear	210.9	64.2	240	20-25, 107-110
Shemya	174.2	52.7	93	27-28
Carnarvon	113.7	-24.9	1	47-53
Antarctica	60.0	-66.5	100	60-67
Santiago	289.3	-33.1	727	80-85

Table 1
Tracking Station Locations and Observation Schedule

in the measurement errors are σ_ρ = 5 meters and $\sigma_{\dot{\rho}}$ = 0.005 meters/sec, respectively. Simultaneous observations are conducted at 5-second intervals. The satellite is in a low-altitude orbit with a perigee/apogee height of 160/380 kilometers, an eccentricity of 0.0167, an argument of perigee of 130°, and an inclination of 110°. An initial nominal state estimate is chosen by perturbing the true initial state by 100 meters and 1 meter/sec in each component of position and velocity, and the initial a priori covariance is a diagonal matrix with standard deviations of 1000 meters and 10 meters/sec in the position and

velocity elements, respectively. The simulation is started with the satellite positioned near its ascending node at about 50° longitude west of Greenwich.

Since the true state is available in the simulation, it is straightforward to compute actual estimation errors and to evaluate qualitative filter performance. Results are judged in terms of the true RSS position error magnitude ΔR and the ability of the estimated covariance norm N_R to predict and bound the RSS errors. Another figure of performance is provided by the "average" RSS error EP_R and covariance norm CP_R over the total estimation interval of about 110 minutes (one orbital period is about 90 minutes).

Linearization Errors - Nonlinear Filtering

The development of the extended Kalman filter (EKF) is well documented[7], and its advantages in reducing linearization errors and computational complexity are well known[8]. As illustrated in Ref. 4, this algorithm provides state estimates $\hat{X}_k$ which are unbiased to first-order (in a Taylor's series expansion) with a near-minimum (within first-order) error covariance estimate $\hat{P}_k$. The use of conventional state noise compensation (SNC), with a constant state noise covariance matrix Q_k, provides acceptable estimation results for the example problem (see Fig. 1).

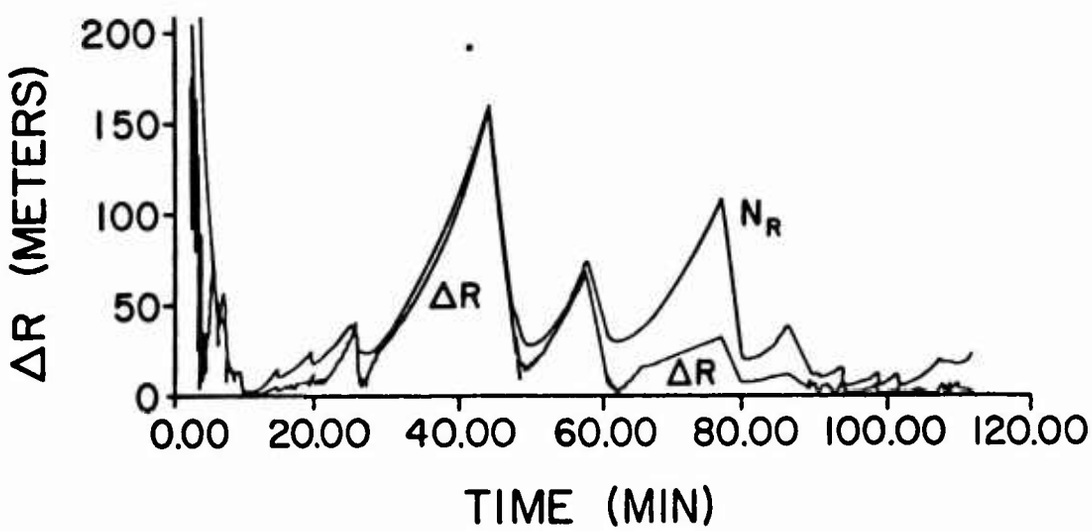


Figure 1. First-Order Filter Performance

The covariance estimates N_R provide a reliable measure of the true error ΔR, even during the prediction phases. After initial filter convergence, the estimation errors become small (on the order of one meter); during this period, the satellite passes over the North American continent, where usually two or more

tracking stations are making simultaneous observations. Toward the middle of the estimation interval, as the satellite descends into the Southern Hemisphere, fewer observations are available, and errors typically grow to 10 or 20 meters in position when only one station views the satellite. During observation coverage gaps, the errors grow to as much as 150 meters; errors are again sharply reduced when the satellite passes over North America after a full orbit. Average error and covariance performance for this case is listed first in Table 2.

Filter Algorithm	EP_R(m)	CP_R(m)
First-Order (EKF)	27.31	18.97
Second-Order	27.33	18.97
DMC	19.13	65.68
EKF (P_{0A})	Did not converge	
SRF (P_{0A})	24.89	62.56
SOSRF	27.28	19.01
EKF (Case A)	28.15	-3.90
CWLMF (Case A)	20.89	20.10
EKF (Case B)	0.216	0.0456
CWLMF (Case B)	0.181	0.1532

Table 2
Average Position Error and Covariance Performance Factors for Various Filter Algorithms

Despite the effectiveness of the above results, it is recalled that the EKF is only a first-order development; therefore, it is logical to ask what additional benefit might be obtained by extending approximations in the derivation to second-order. Research in the area of nonlinear filtering has been pursued along two general avenues: first, the stochastic Itô calculus[7] provides the optimal filter, but it is infinite dimensional and requires the solution to a complex partial differential equation; second, finite dimensional approximate filters[7,9] have been derived in which the density function is characterized by a set of "sufficient" statistics. The latter approach includes the Taylor series approximation methods which have led to several versions of the second-order filter[7,10]. Performance comparison studies of the first- and second-order discrete filters[11,12,13], indicate that results for the second-order methods are highly problem dependent and varying degrees of success can be obtained depending upon assumptions made in the development. A new approach, guided by these experimental findings, is pursued in Ref. 4. The resultant filter is identical to the conventional extended Kalman filter, except for the addition of a new term

involving second-order partials of the observation-state relation to the observation covariance matrix in the Kalman gain expression. State estimates obtained from the filter are unbiased to first-order, but the state covariance matrix is minimized to second order. Despite the incorporation of second-order effects, the estimation performance is essentially unchanged from that of the conventional EKF (see Table 2). Those differences which could be detected occurred during the early phases of the estimation, where the second-order filter tended to down-weight large observation residuals in a desirable manner. The fact that the second-order filter produced slightly larger estimation error in Table 2 is attributed to numerical inaccuracies, discussed below. As a result of these subtle differences, it is concluded that nonlinearities are not a source of difficulty in this class of orbit determination problems. However, a new algorithm has been developed which may have useful application in other problems; in fact, its usefulness has been demonstrated by Choe[11] in the Jupiter fly-by problem.

Model Errors - Dynamical Model Compensation

It is probably safe to say that dynamical model errors constitute by far the greatest detriment to effective, accurate estimation in practically every orbit determination problem; and, unfortunately, this source of difficulty is the most difficult to correct. As detailed in other studies[7,14], dynamical model errors, if left uncorrected, typically result in the well known phenomena of "filter divergence". A wide variety of techniques designed to compensate for these errors has appeared in the literature[15]. One of the most effective approaches is dynamical model compensation (DMC), due to Tapley and Ingram[16], which treats the unknown functional expression governing the neglected dynamics as an assumed, time-correlated dynamical structure. Adaptive qualities of this approach are enhanced by estimating those parameters which characterize the assumed structure. In this study, a first-order, Gauss-Markov structure is employed. As indicated in Fig. 2 and Table 2, a dramatic improvement in estimation accuracy can be obtained with the DMC approach. Further, it is evident that prediction capability has been enhanced, e.g., the large 150-m prediction error at 45 min has been reduced to about 30 m. When the estimate of the unmodeled acceleration is good, such as the period from 5 min to 30 min, the RSS position error is reduced by factors like "two" or "ten", e.g., at 18, 19, and 20 min, the standard EKF filter errors were 6.5, 6.7, and 10.9 m, whereas

the DMC errors were 2.3, 0.6, and 6.9 m, respectively. During periods of sparse observations or poor geometry, the estimate of the unmodeled acceleration can deteriorate, but degradation in the position and velocity estimate appears to be of small consequence; and, when the unmodeled acceleration estimate is good, superior filter performance can be obtained.

Computation Errors - Square Root Filtering

Computational errors stem primarily from the constraint that filtering calculations must be performed with a finite wordlength computing machine. The effects of computational errors are felt most directly by the state covariance estimate $\hat{P}$. Numerical instabilities often result in the undesirable tendency of $\hat{P}$ becoming negative definite. To combat this difficulty, several alternative formulations[3] have been proposed, of which the square root filters have been found most useful. The basic approach in the development of a square root filter is to reformulate the algorithm in terms of a square root covariance S, defined by $P \equiv SS^T$. The advantage of this approach is that the propagation and estimation of S cannot destroy the positive definiteness of P; also, greater effective precision is retained, since the order of magnitude of S is half that of P. Several algorithms[17] have been formulated for both the estimation and propagation of S, and generally, when process noise is present, it is the propagation algorithm which imposes the greater computational burden. The implementation selected in this study consists of the Andrews algorithm for square root covariance estimation and the modified Gram-Schmidt algorithm for propagation.

To a large extent, computational errors in the example problem are negligible compared to model errors, so the improvement in accuracy is so small (in the 6th significant figure) as to be inconsequential. Thus, for a well-conditioned problem, square root filtering becomes unnecessary on large wordlength machines, such as the CDC 6600. Numerical conditioning was further evaluated, however, by increasing the a priori state covariance to a value P_{0A} with standard deviations of 3×10^5 m and 7×10^3 m/sec on the diagonal position and velocity elements, respectively. This produced an immediate failure in the EKF; but, the SRF encountered no difficulty whatsoever, as shown in Table 2 and Figure 3.

It is difficult to make a fair comparison of computation time and storage required in the two filter algorithms, but it is generally agreed that square root filters require at most about 30% more time than conventional filters.

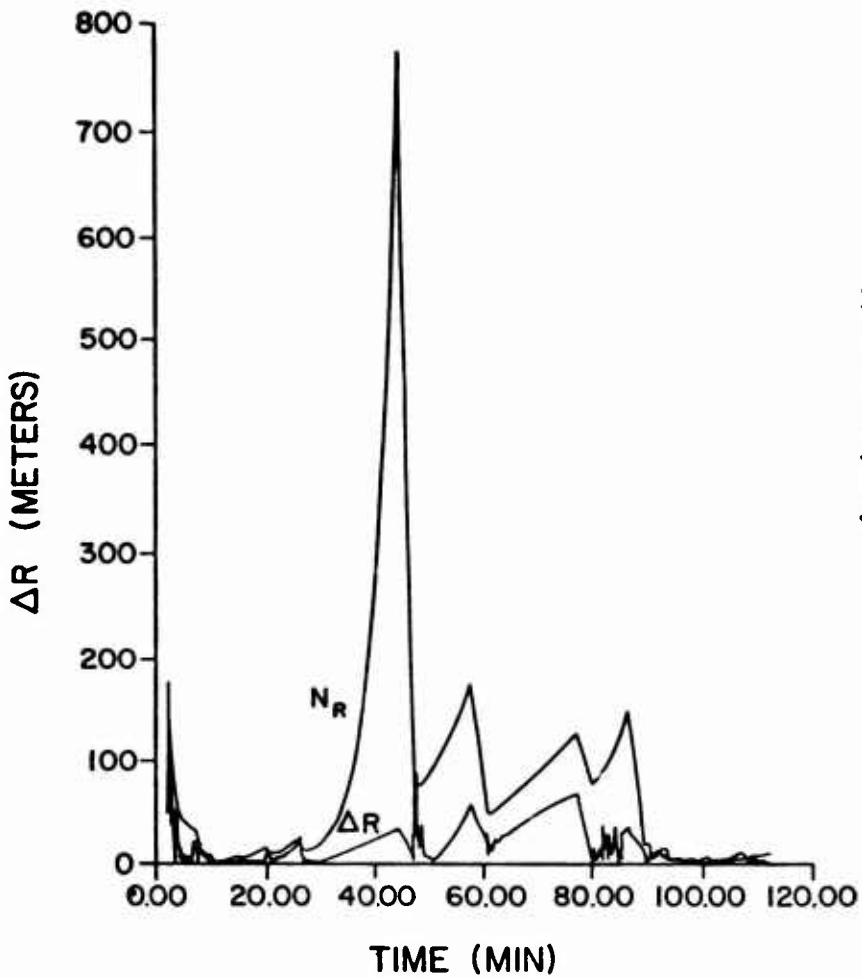


Figure 2. DMC Filter Performance

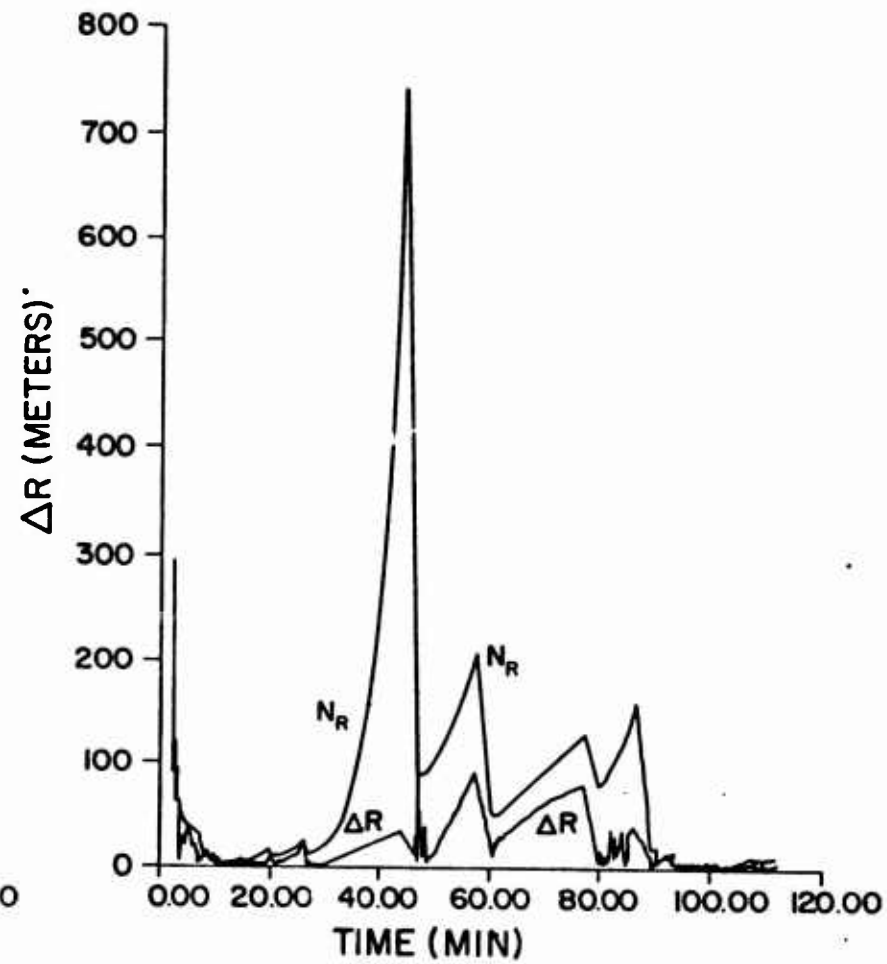


Figure 3. SRF for P_{0A}

When other computational functions are considered, the overall difference can be of minor importance as seen by the total central processor execution time in Table 3.

Filter	CP Time (sec)
First-Order (EKF)	60
DMC	127
SRF	132
SOSR	142
CWLMF	68

Table 3
Central Processor Time

It is concluded that the extra computations required in square root filtering should be well worth the effort and could even be mandatory in operational applications with short wordlength machines.

The second-order filter discussed above is in a form suitable for the square root algorithms, thereby permitting the implementation of a second-order square root filter (SOSRF). It is interesting to note in Table 2, that this algorithm provides the expected reduction in estimation error when second order terms are included. In addition, the errors are bounded slightly better by the estimated covariance. An explanation for this result is that the effects of nonlinearities, however mild, appear to be more accurately corrected with the use of a square root algorithm.

Unknown Noise Statistics - Adaptive Filtering

It is appropriate to conclude the study with an investigation of adaptive filtering, because these techniques are designed to correct the basic filter algorithm for the combined effects of all error sources. In the literature[18] this area of research is often referred to as "simultaneous estimation of the state and noise statistics". For the purposes of this study, the approach is regarded as statistical, contrasted with the structural approach used in DMC, above. Obviously, unless exact noise statistics for the state and observation processes are used in the Kalman filter algorithm, suboptimal state estimation will result. This is the typical case in actual filtering applications. In addition, these processes are usually assumed to have zero mean, but this seldom occurs due to neglected nonlinearities and model errors.

The adaptive technique developed in Ref. 4 is designed to provide estimates of the desired mean and covariance of both the observation and state noise

processes. It is based on an empirical sampling of state and observation noise samples generated directly from the basic filter algorithm. A sample space of size N is created after N filter steps, from which the desired statistics are estimated. The noise estimators could be implemented in many different ways, but in this study the straightforward assumption is made that the noise statistics remain constant over the interval of N samples; thus, the value of N must be tailored to a particular problem. A recursive, limited memory form of the algorithm, called a Complete Limited Memory Filter (CLMF), is presented in Fig. 4. The term "complete" refers to the fact that all unknowns in the problem are being estimated. In order to "ease" the estimates of the noise statistics into the basic filter, successive noise samples were multiplied by a weight factor that grows from 0 to 1 over the estimation interval, which provides a Complete Weighted Limited Memory Filter (CWLMF). Also, steps were taken to preserve the positive definite character of the observation and state noise covariances.

Performance of the standard EKF with an erroneous state noise covariance Q is compared with that for the CWLMF in Fig. 5. Both runs are started with the same value of Q; the observation noise bias and covariance R is initiated at its correct value in both runs, but it is further estimated during the CWLMF run. A sample size of N = 15 is employed. It appears from Fig. 5 and Table 2 that the CWLMF can be highly effective in providing necessary corrections to the filter. Due to small initial weight factors there is little difference in performance from the EKF until after 15 min. At this time, the errors are reduced by several meters and the CWLMF begins to correct the a priori statistics. It begins to display superior performance at about 25 min, and the large EKF prediction error at 44 min is reduced from 150 m to about 40 m; likewise, at 56 min the large 80-m error for the EKF is reduced to less than 40 m. However, CWLMF results become inferior after about 70 min; moreover, numerical difficulties were encountered at about 110 min, where the state covariance matrix developed negative diagonal elements.

A primary source of difficulty in the selected example is due to the observation coverage "gaps"; this tends to introduce erroneous noise samples at each reacquisition. Therefore, another tracking station configuration was tried with twelve stations scattered along the ground track during the first 18 minutes of the orbit. This provided simultaneous observation coverage by three to six stations at all times. As shown in Fig. 6 and Table 2, the CWLMF again provides

A Priori Data: $\hat{x}_0, \hat{P}_0, \hat{q}_0, \hat{Q}_0, \hat{r}_0, \hat{R}_0 \;\; ; \;\; k = 1$

State Propagation: $\bar{x}_k = \Phi_k \hat{x}_{k-1} + \hat{q}_{k-1}$

$$\bar{P}_k = \Phi_k \hat{P}_{k-1} \Phi_k^T + \hat{Q}_{k-1}$$

Compute Observation Noise ($k \geq \ell_r$): $r_k = y_k - H_k \bar{x}_k$, $\Gamma_k = H_k \bar{P}_k H_k^T$

$$\hat{r}_k = \hat{r}_{k-1} + \frac{1}{\ell_r} (r_k - r_{k-\ell_r})$$

$$\hat{R}_k = \hat{R}_{k-1} + \frac{1}{\ell_r - 1} \{ (r_k - \hat{r}_k)^2 - (r_{k-\ell_r} - \hat{r}_k)^2$$

$$+ \frac{1}{\ell_r} (r_k - r_{k-\ell_r})^2 + (\frac{\ell_r - 1}{\ell_r})[\Gamma_{k-\ell_r} - \Gamma_k] \}$$

Shift Noise Samples: $r_j = r_{j+1}$, $\Gamma_j = \Gamma_{j+1}$, for $j = k - \ell_r, \ldots, k - 1$

Compute Kalman Gain: $K_k = \bar{P}_k H_k^T [\Gamma_k + \hat{R}_k]^{-1}$

State Estimation: $\hat{x}_k = \bar{x}_k + K_k [r_k - \hat{r}_k]$

$$\hat{P}_k = \bar{P}_k - K_k H_k \bar{P}_k$$

Compute State Noise ($k \geq \ell_q$): $q_k = \hat{x}_k - \Phi_k \hat{x}_{k-1}$, $\Delta_k = \Phi_k \hat{P}_{k-1} \Phi_k^T - \hat{P}_k$

$$\hat{q}_k = \hat{q}_{k-1} + \frac{1}{\ell_q} (q_k - q_{k-\ell_q})$$

$$\hat{Q}_k = \hat{Q}_{k-1} + \frac{1}{\ell_q - 1} \{ (q_k - \hat{q}_k)^2 - (q_{k-\ell_q} - \hat{q}_k)^2$$

$$+ \frac{1}{\ell_q} (q_k - q_{k-\ell_q})^2 + (\frac{\ell_q - 1}{\ell_q}) [\Delta_{k-\ell_q} - \Delta_k] \}$$

Shift Noise Samples: $q_j = q_{j+1}$, $\Delta_j = \Delta_{j+1}$, for $j = k - \ell_q, \ldots, k - 1$

Reset: $k = k + 1$ and repeat

Figure 4. CLMF Algorithm

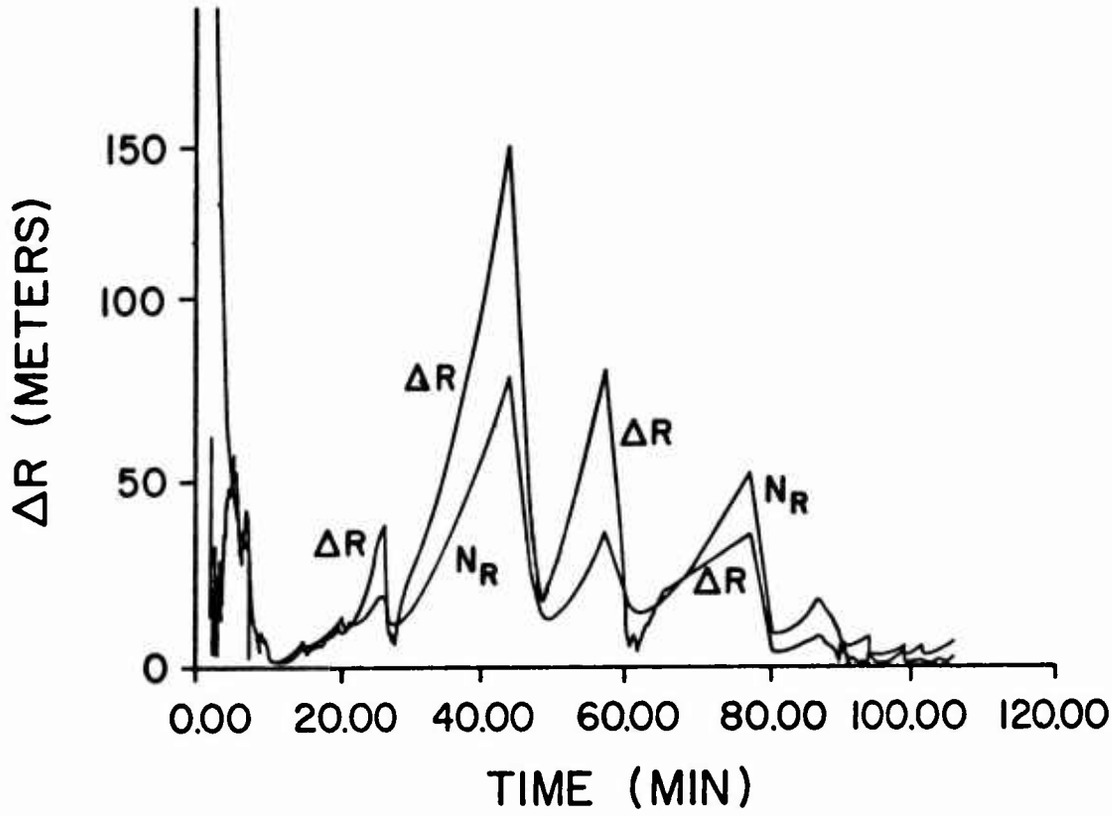


Figure 5a. EKF Results (Case A)

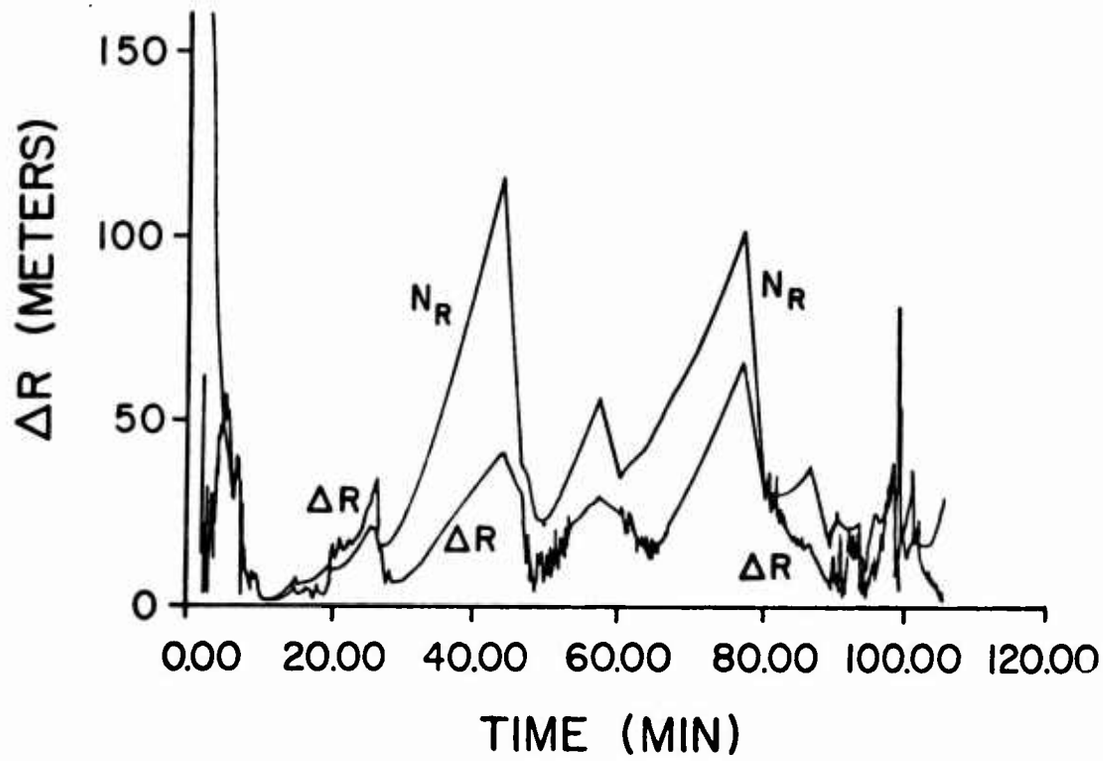


Figure 5b. CWLMF Results (Case A)

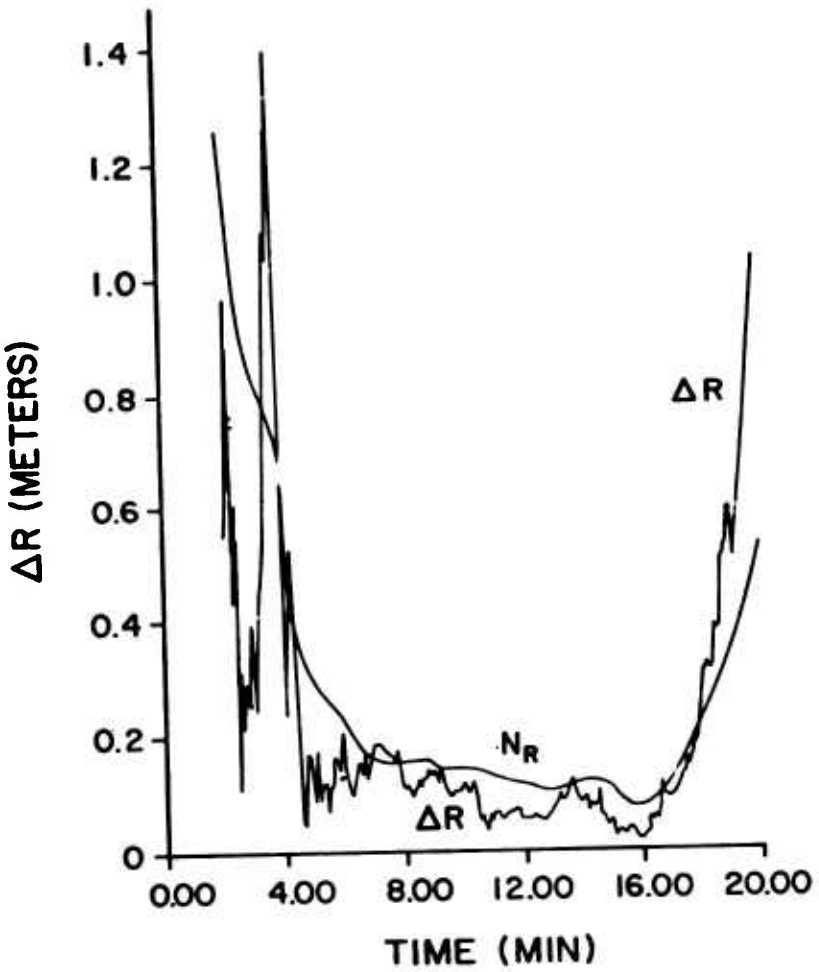


Figure 6a. EKF Results (Case B)

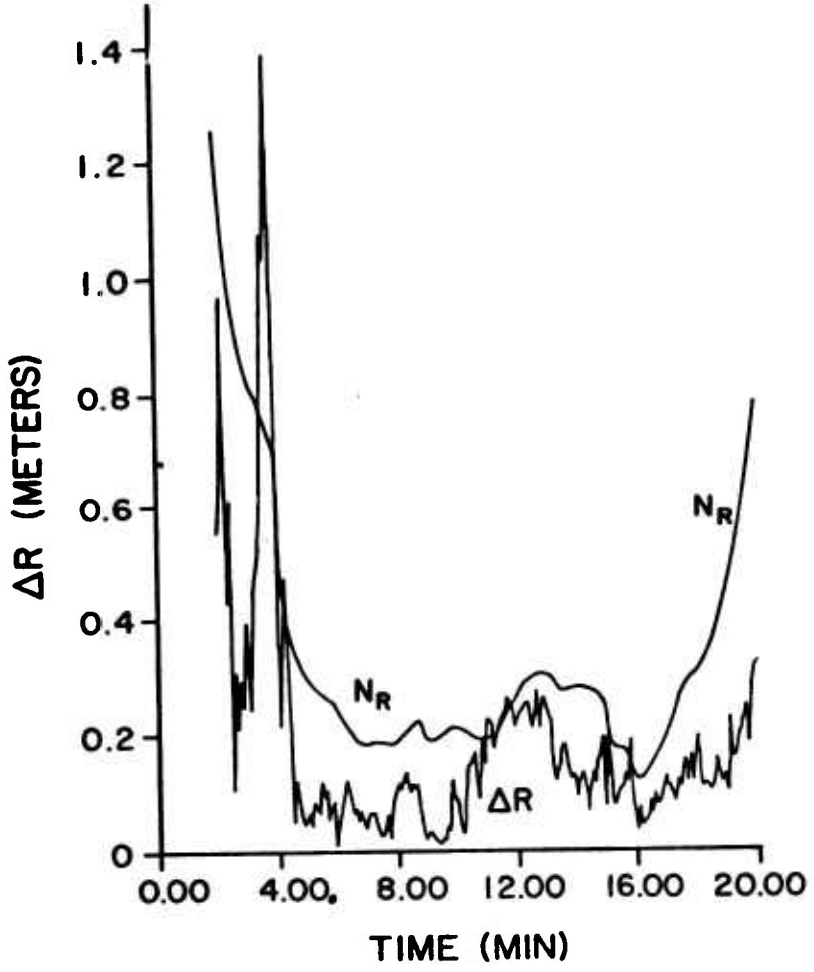


Figure 6b. CWLMF Results (Case B)

improved filter performance, and in this case no numerical difficulties were encountered.

In summary, the CWMF algorithm can improve the state estimation when a priori noise statistics are in error. Furthermore, the improvement is obtained at a modest cost, as shown in the computation times in Table 3. Some extra storage is required to accommodate and shift the noise samples, but this may be a favorable trade-off when compared with other adaptive approaches, such as those requiring state vector augmentation.

Conclusions

Based on the simulation results, it is concluded that the EKF essentially eliminates linearization errors, and second-order methods are unnecessary. The DMC algorithm provides highly effective compensation in the presence of dynamical model errors. Square root filtering imposes a minimal additional computation burden in view of the total computational requirements; this could be a very modest price to pay, especially in short wordlength computers, when the added advantages in numerical stability and accuracy are considered. Additional stability could be obtained with a second-order square root filter in other highly nonlinear problems. The new adaptive filter algorithm can improve estimation performance when the a priori statistics are in error.

References

1. Escobal, P. R., Methods of Orbit Determination, Wiley, New York, N. Y., 1965.

2. Deutsch, R., Estimation Theory, Prentice-Hall, Englewood Cliffs, N. J., 1965.

3. Leondes, C. T., (ed.), Theory and Applications of Kalman Filtering, NATO AGARDograph No. 139, Technical Editing and Reproduction, Ltd., London, Feb., 1970. (Reproduced by National Technical Information Service, Springfield, Va.).

4. Myers, K. A., "Filtering Theory Methods and Applications to the Orbit Determination Problem for Near-Earth Satellites", AMRL Report No. 1058, Nov. 1973, Applied Mechanics Research Laboratory, The University of Texas at Austin.

5. Leondes, C. T., (ed.), The Application of Digital Computers to Guidance and Control, AGARD-CP-68-70, North Atlantic Treaty Organization, Technical Editing and Reproduction, Ltd., London, Nov., 1970. (Reproduced by National Technical Information Service, Springfield, Va.).

6. GEOS-C Mission Proposal Briefing Information, Wallops Station (NASA), Wallops Island, Va., Dec. 13, 1972.

7. Jazwinski, A. H., Stochastic Processes and Filtering Theory, Academic Press, New York, N. Y., 1970.

8. Tapley, B. D., "Statistical Orbit Determination Theory", Presented at NATO Advanced Study Institute in Dynamical Astronomy, Cortina d'Ampezzo, Italy, Aug., 1972.

9. Kushner, H. J., "Approximations to Optimal Nonlinear Filters", IEEE Trans. on Auto. Cont., Vol. AC-12, No. 5, Oct., 1967, pp. 546-556.

10. Bass, R. W., V. D. Norum, and L. Schwartz, "Optimal Multichannel Nonlinear Filtering", J. Math. Anal. and Appl. 16, 1966, pp. 152-164.

11. Tapley, B. D., and C. Y. Choe, "Nonlinear Estimation Theory Applied to the Interplanetary Orbit Determination Problem", Third Symposium on Nonlinear Estimation Theory and its Applications, San Diego, Calif., Sept. 11-13, 1972.

12. Jazwinski, A. H., "Nonlinear Filtering with Discrete Observations", AIAA 3rd Aerospace Science Meeting, AIAA Paper No. 66-38, Jan., 1966.

13. Athans, M., R. P. Wishner, and A. Bertolini, "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements", IEEE Trans. on Auto. Cont., Oct., 1964, pp. 504-514.

14. Schlee, S. F., C. J. Standish, and N. F. Toda, "Divergence in the Kalman Filter", AIAA J., Vol. 5, No. 6, 1967, pp. 1114-1120.

15. Hagar, H., Jr., "Model Error Compensation Techniques for Linear Filtering", AMRL Rept. 1055, Applied Mechanics Research Laboratory, The University of Texas at Austin, Aug., 1973.

16. Tapley, B. D., and D. S. Ingram, "Orbit Determination in the Presence of Unmodeled Accelerations", Proc. Second Symposium on Nonlinear Estimation Theory, San Diego, Calif., Sept., 1971.

17. Kaminski, P. G., "Square Root Filtering and Smoothing for Discrete Processes", Ph.D. Dissertation, SUDAAR No. 427, Guidance and Control Laboratory, Stanford University, July, 1971.

18. Weiss, I. M., "A Survey of Discrete Kalman-Bucy Filtering with Unknown Noise Covariances", AIAA Guidance, Control, and Flight Mechanics Conf., Santa Barbara, Calif., Aug. 17-19, 1970.

Paper Presented to Symposium on
Air Force Applications of Modern
Control Theory

ANALYSIS OF BIAS, VARIANCE AND MEAN SQUARE ESTIMATION ERROR IN REDUCED ORDER FILTERS

by

Capt Robert B. Asher
Frank J. Seiler Research Laboratory
USAF Academy CO 80840

and

Dr. Jesse C. Ryles
Air Force Avionics Laboratory
Wright-Patterson AFB OH 45433

11 July 1974

I. Introduction

The Kalman filter has been utilized in many applications throughout the last decade. These applications range from chemical processing plants to aerospace navigation and guidance [1, 2, 3]. In many of these applications the implemented filter did not behave as the theoretical analysis showed as its predicted behavior. Rather than achieving a degree of optimality the filter diverged [4] causing an erroneous state estimate to be generated. This problem is due to several causes: erroneous state models including neglected biases, incomplete or erroneous knowledge of statistical models for filter derivation, nonlinearities, and roundoff and truncation errors [5, 6]. With respect to the erroneous state model problem many filter designs have been developed without due regard to additional states that arise because of error sources such as biases, first or higher order Markov processes driving the dynamic system and/or the measurement system. As is well known [7, 8] one may use state augmentation techniques in order to take these errors into account. However, this may lead to a significant and perhaps intolerable state dimensionality if a fully optimal Kalman filter were designed from this new model. For example, in a navigation system one may identify and model 61 state elements [9] and in a precision pointing and tracking problem one can identify and model 75 states [10]. However, this leads to a significant computational problem and in fact an intolerable one for a real-time flight computer. Therefore, a reduced order filter (ROF) that estimates only the required minimum number of state elements must be designed. If this is done then since all the high order state elements are coupled, one must consider the performance of the ROF against the true state model for otherwise when implemented the filter may at best yield poor performance and at worst may diverge. One must consider the sensitivity of the filter due to deliberate modeling errors introduced by the suboptimal filter structure.

The Kalman filter [4] gives the optimal minimum variance, unbiased estimate of the system state. It is shown in this paper that for a ROF one cannot in general obtain an unbiased estimator. The conditional mean of the estimation error is non-zero and, therefore, the true covariance of the estimation error is not equal to the second moment of the estimation error as implied in some previous work [11, 12]. This problem of bias is recognized in the work of several references [13, 20, 21]; however, the reduced order problem was not addressed. The work of this paper considers both bias and reduced order filters. It is shown that a reduced order filter in general will be biased. The equations for the second moment, true covariance, and bias are presented for the continuous dynamics, continuous measurement and continuous dynamics, discrete measurement ROF problem. The results include the subcases of continuous and impulsive control. Structure properties of the filter are being studied and a simulation program is being developed to evaluate this approach. These results will be reported in a separate paper.

Previous work accomplished in the area of ROF's is numerous. Aoki and Huddle [14] consider the design of an estimator for a discrete-time stochastic system with time-invariant dynamics and memory constraints. Sims [15] considers the design of an estimator based upon fixing the filter structure and optimizing certain parameters within the estimator to minimize the estimation error. Griffin and Sage [16, 17] obtain sensitivity equations for error analysis for another class of errors. Heffes [18] considers the problem for discrete-time systems. All the above referenced work, however, consider only the problem where the state and filter are of the same dimensionality. Stubberud and Wisner [19] considers the problem in which one estimates only a certain subset of the state called the primary variables; however, the structure considered is not general nor is the problem of feedback control considered.

II. Continuous Dynamics and Measurement Problem

The equations for the second moment, covariance and bias of the true estimation error are developed in this section for the continuous dynamics and continuous measurement ROF problem. The conditions for obtaining an unbiased estimate using a ROF are developed. These conditions show that in general one cannot obtain an unbiased estimate. The problem of continuous control is discussed and the necessary equation modification is given.

A. Problem Statement

Consider the time varying system

$$x_S(t) = F_S(t)x_S(t) + G_S(t)u_S(t) \tag{1}$$

where $x_S \epsilon R^{n_1}$ is the system state vector, $F_S(t)$ is an $n_1 \times n_1$ system matrix with continuous and bounded elements, $G_S(t)$ is an $n_1 \times m_1$ system matrix with continuous and bounded elements, and $u_S \epsilon R^{m_1}$ is a zero-mean, white noise vector with known variance

$$E\{u_S(t)u_S^T(\gamma)\} = Q_S(t)\delta(t - \gamma).$$

The system is observed via the measurement equation

$$y(t) = H_S(t)x_S(t) + v_S(t) \tag{2}$$

where $y \epsilon R^p$ is the measurement vector, $H_S(t)$ is a $p \times n_1$, measurement matrix with continuous and bounded elements, and $v_S \epsilon R^p$ is a zero-mean, white noise vector with known variance

$$E\{v_S(t)v_S^T(\gamma)\} = R_S(t)\delta(t - \gamma).$$

The system model (1) and measurement model (2) are assumed to be the best known model of the system structure. This representation contains all the physical states of interest and all error sources that are elements of the state vector due to the utilization of state augmentation techniques. In many application problems it is computationally infeasible to design a Kalman filter that will estimate the total system state vector. Thus, the problem is one of structuring a ROF that estimates (with minimum performance degradation) a linear transformation of the state vector, x_S. The filter vector to be estimated is given as

$$x_d(t) = Tx_S(t) \tag{3}$$

where $x_d \epsilon R^{n_2}$ are the pertinent states of interest and T is an $n_2 \times n_1$ selector matrix, $n_2 \leq n_1$. It is assumed without loss of generality that the pertinent state elements are the first n_2 elements of x_S. Therefore, the selector matrix consists of an $n_2 \times n_2$ identity matrix and an $(n_1 - n_2) \times n_2$ null matrix, i.e.,

$$T = [I : 0] \tag{4}$$

A differential equation for x_d may be found by differentiating equation (3). This yields

$$\dot{x}_d(t) = T[F_S(t)x_S(t) + G_S(t)u_S(t)]. \tag{5}$$

The reduced order filter is assumed to be of the structure

$$\hat{x}_d(t) = F_F(t)\hat{x}_d(t) + K_d(t)[y(t) - H_F(t)\hat{x}_d(t)] \tag{6}$$

where $\hat{x}_d \epsilon R^{n_2}$ is the linear estimate of x_d as determined by the filter structure in equation (6), F_F is an $n_2 \times n_2$ filter system matrix, K_d is an $n_2 \times p$ filter gain matrix, $y \epsilon R^p$ is the measurement vector, and H_F is a $p \times n_2$ measurement matrix. The filter structure is derived by use of the optimal Kalman filter equations [5] applied to the assumed (however erroneous) state model, x_F, for the state x_d, i.e.,

$$\dot{x}_F(t) = F_F(t)x_F(t) + G_F(t)u_F(t) \tag{7}$$

where F_F has continuous and bounded elements, G_F is an $n_2 \times m_2$ noise gain matrix with continuous and bounded elements, and $u_F \epsilon R^{m_2}$ is a zero mean, white noise vector with known variance

$$E\{u_F(t)u_F^T(\gamma)\} = Q_F(t)\delta(t-\gamma)$$

and to the assumed (however erroneous) measurement model for the measurements, y, i.e.,

$$y(t) = H_F(t)x_F(t) + v_F(t) \tag{8}$$

where H_F has continuous and bounded elements, and v_F is a zero-mean, white noise vector with known variance

$$E\{v_F(t)v_F^T(\gamma)\} = R_F(t)\delta(t-\gamma)$$

B. ROF Equations

The true estimation error for the states of interest is given as

$$e_d(t) = x_d(t) - \hat{x}_d(t). \tag{9}$$

A differential equation for the estimation error may be obtained by differentiating equation (9) and using equations (1), (2), (3), and (6). This yields

$$\dot{e}_d = T(F_S x_S + G_S u_S) - [F_F \hat{x}_d + K_d(H_S x_S + v_S - H_F \hat{x}_d)] \tag{10}$$

This equation may be rewritten by adding and subtracting the terms $(F_F - K_d H_F)Tx_S$. This yields

$$\dot{e}_d = (F_F - K_d H_F)e_d + (TF_S - K_d H_S - F_F T + K_d H_F T)x_S + TG_S u_S - K_d v_S \tag{11}$$

The bias properties of the estimation error are demonstrated as follows.

The requirement for a conditionally unbiased estimator is that

$$E\{[x_d(t) - \hat{x}_d(t)]|Y(t)\} = 0, \ \forall t\varepsilon[t_o, t_f] \tag{12}$$

where Y(t) is the measurement functional ($y(\gamma)$, $\gamma\varepsilon[t_o, t]$). This condition may be rewritten by use of equation (9) as

$$E\{e_d(t)|Y(t)\} = 0, \ \forall t\varepsilon[t_o, t_f]. \tag{13}$$

Equation (13) implies that

$$E\{\dot{e}_d(t)|Y(t)\} = 0, \ \forall t\varepsilon[t_o, t_f]. \tag{14}$$

Thus, equations (13), (14), and (11) imply that

$$0 = (TF_S - K_d H_S - F_F T + K_d H_F T)\hat{x}_S(t), \ \forall t\varepsilon[t_o, t_f] \tag{15}$$

where

$$\hat{x}_S = E\{x_S(t)|Y(t)\}. \tag{16}$$

However, in general x_S does not lie in the null space of the coefficient matrix throughout the interval $[t_o, t_f]$. Therefore, the condition

$$TF_S - K_d H_S - F_F T + K_d H_F T = 0 \tag{17}$$

must be satisfied $\forall t \in [t_o, t_f]$. One may partition the matrix F_S as

$$F_S = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \tag{18}$$

where F_{11} is $n_2 \times n_2$, F_{12} is $n_2 \times (n_1 - n_2)$ and the remainder of the submatrices are compatible with the dimensions of F_S. Also, the $n_2 \times n_1$ matrix $K_d H_S$ may be partitioned into two submatrices $(KH)_1$ and $(KH)_2$ of dimensions $n_2 \times n_2$ and $n_2 \times (n_1 - n_2)$ respectively, i.e.,

$$K_d H_S = [(KH)_1 : (KH)_2]. \tag{19}$$

The use of equation (4), (18), and (19) in equation (17) yields the following two conditions that must be satisfied for a conditionally unbiased estimator, i.e.,

$$F_{11} - (KH)_1 - F_F + K_d H_F = 0 \tag{20}$$

$$F_{12} - (KH)_2 = 0. \tag{21}$$

These conditions will <u>not</u> be satisfied in general and, thus, the ROF will be biased. Thus, the conditional expectation of the estimation error will be such that

$$\bar{e}_d(t) = E\{e_d(t) | Y(t)\} \neq 0 \tag{22}$$

except at most a countable number of points. Therefore, the conditional variance, $V_d(t)$ of the estimation error is given as

$$V_d(t) = E\{(e_d(t) - \bar{e}_d(t))\ (e_d(t) - \bar{e}_d(t))^T | Y(t)\} \tag{23}$$

Therefore, the variance of the estimation error is not equal to the second moment, $E\{e_d(t) e_d^T(t)\}$ as is commonly used [12] for ROF's. A differential equation for the variance of the estimation error is now derived.

One may for notational convenience define a new variable, $\delta e_d(t)$, defined as

$$\delta e_d(t) = e_d(t) - \bar{e}_d(t) \tag{24}$$

The derivative of this new variable is given as

$$\delta \dot{e}_d(t) = \dot{e}_d(t) - \dot{\bar{e}}_d(t) \tag{25}$$

The differential equation for the conditional mean of the estimation error is

$$\dot{\bar{e}}_d = (F_F - K_d H_F)\bar{e}_d + (TF_S - K_d H_S - F_F T + K_d H_F T)\hat{x}_S \tag{26}$$

The use of equation (11) and (26) in (25) yields the differential equation for δe_d as

$$\delta \dot{e}_d = (F_F - K_d H_F)\delta e_d + (TF_S - K_d H_S - F_F T + K_d H_F T)(x_S - \hat{x}_S) + TG_S u_S - K_d v_S \tag{27}$$

The term $x_S - \hat{x}_S$ is recognized as the estimation error of the optimal Kalman filter. This quantity is denoted as e_S. Thus, equation (27) may be rewritten as

$$\delta \dot{e}_d = (F_F - K_d H_F)\delta e_d + (TF_S - K_d H_S - F_F T + K_d H_F T)e_S + TG_S u_S - K_d v_S \tag{28}$$

The true error variance, $E\{\delta e_d \delta e_d^T\}$ will be denoted by, V_e, i.e.,

$$V_e = E\{\delta e_d \delta e_d^T\} \tag{29}$$

In order to develop the necessary differential equation for V_e one may define the augmented state, z_1, and the augmented noise vector, u_e as

$$z_1(t) = \begin{bmatrix} \bar{e}_d(t) \\ e_S(t) \end{bmatrix}, \quad \bar{u}(t) = \begin{bmatrix} u_S(t) \\ v_S(t) \end{bmatrix} \tag{30}$$

This yields the differential equation for z_1 as

$$\dot{z}_1(t) = \begin{bmatrix} (F_F - K_d H_F), & (TF_S - K_d H_S - F_F T + K_d H_F T) \\ 0 & , (F_S - K_S H_S \end{bmatrix} z_1 + \begin{bmatrix} TG_S & , -K_d \\ G_S & , -K_S \end{bmatrix} \bar{u} \tag{31}$$

where K_S is the system Kalman gain matrix.

This is of the form

$$\dot{z}_1 = \bar{A}z_1 + \bar{B}\bar{u} \tag{32}$$

where

$$E\{\bar{u}(t)\bar{u}^T(\gamma)\} = \bar{Q}(t)\delta(t - \gamma)$$

where

$$\bar{Q}(t) = \begin{bmatrix} Q_S(t) & 0 \\ 0 & R_S(t) \end{bmatrix} \tag{33}$$

The differential equation for the second moment of z_1, P_{z_1}, may be easily derived. This yields

$$\dot{P}_{z_1} = \bar{A}P_{z_1} + P_{z_1}\bar{A}^T + \bar{B}\bar{Q}\bar{B}^T \tag{34}$$

where $\bar{A}$, $\bar{B}$, and $\bar{Q}$ follow from the previous discussions. The second moment equation for z_1 may be partitioned as follows

$$P_{z_1}(t) = \begin{bmatrix} V_e(t) & V_{12}(t) \\ V_{21}(t) & V_S(t) \end{bmatrix} \tag{35}$$

where $V_S(t)$ is the variance associated with the optimal Kalman filter yielding $\hat{x}_S$. The components of equation (34) may be written as

$$\dot{V}_e = (F_F - K_d H_F)V_e + V_e(F_F - K_d H_F)^T + (TF_S - K_d H_S - F_F T + K_d H_F T)V_{12}^T$$

$$+ V_{12}(TF_S - K_d H_S - F_F T + K_d H_F T)^T + TG_S Q_S G_S^T T^T + K_d R_S K_d^T \quad (36)$$

$$\dot{V}_{12} = (F_F - K_d H_F)V_{12} + (TF_S - K_d H_S - F_F T + K_d H_F T)V_S +$$

$$V_{12}(F_S - K_S H_S)^T + TG_S Q_S G_S^T + K_d R_S K_S^T \quad (37)$$

$$\dot{V}_S = (F_S - K_S H_S)V_S + V_S(F_S - K_S H_S)^T + G_S Q_S G_S^T + K_S R_S K_S^T \quad (38)$$

since

$$V_{21} = V_{12}^T \quad (39)$$

where

$$K_S = V_S H_S^T R_S^{-1} \quad (40)$$

and

$$V_e(t_o) = V_e^o;\ V_{12}(t_o) = V_{12}^o;\ \text{and}\ V_S(t_o) = V_S^o$$

One may note that in the calculation of V_e one must compute the optimal Kalman filter variance, V_S. This quantity is useful as information about the performance of the optimal system filter.

The differential equation for the second moment of another matrix is derived next to determine the remaining ROF properties. Define the augmented state z_2 as

$$z_2(t) = \begin{bmatrix} e_d(t) \\ x_S(t) \end{bmatrix} \quad (41)$$

Thus,

$$\dot{z}_2(t) = \begin{bmatrix} (F_F - K_d H_F), (TF_S - K_d H_S - F_F T + K_d H_F T) \\ 0 \qquad , \qquad F_S \end{bmatrix} z_2 +$$

$$\begin{bmatrix} TG_S & -K_d \\ G_S & 0 \end{bmatrix} \begin{bmatrix} u_S \\ v_S \end{bmatrix} \quad (42)$$

or

$$\dot{z}_2 = \bar{A}z_2 + \bar{B}\bar{u} \quad (43)$$

Proceeding as before, the following differential equations are obtained.

$$\dot{P}_e = (F_F - K_dH_F)P_e + (TF_S - K_dH_S - F_FT + K_dH_FT)P_{12}^T + P_e(F_F - K_dH_F)^T$$
$$+ P_{12}(TF_S - K_dH_S - F_FT + K_dH_FT)^T + TG_SQ_SG_S^TT^T + K_dR_SK_d^T \quad (44)$$

$$\dot{P}_{12} = (F_F - K_dH_F)P_{12} + (TF_S - K_dH_S - F_FT + K_dH_FT)P_S +$$
$$P_{12}F_S^T + TG_SQ_SG_S^T \quad (45)$$

$$\dot{P}_S = F_SP_S + P_SF_S^T + G_SQ_SG_S^T \quad (46)$$

where

$$P_{21} = P_{12}^T$$

and

$$P_e(t_o) = P_e^o,\ P_{12}(t_o) = P_{12}^o \text{ and } P_S(t_o) = P_S^o$$

One may readily note that P_S is the propagation of the second moment when no measurements are used to update the state.

The bias matrix, $\bar{e}_d\bar{e}_d^T$, may be computed from the knowledge of V_e and P_e, i.e.,

$$\bar{e}_d\bar{e}_d^T = P_e - V_e \quad (47)$$

Thus, equations (36, 37, 38, 44, 45, 46, 47) yield all the relevant statistical properties of the ROF.

C. ROF with Continuous Control

The problem of covariance analysis of systems with continuous control is now considered. The control is accomplished by using the estimated state as a feedback control signal to control the system state. This may be modeled by including a feedback signal in equation (1). This yields the modified equation

$$\dot{x}_S(t) = F_S(t)x_S(t) + C_S(t)\hat{x}_d(t) + G_S(t)u_S(t) \quad (48)$$

where $C_S(t)$ is the system control matrix. The corresponding filter equation (6) is modified as

$$\dot{\hat{x}}_d(t) = [F_F(t) + C_F(t)]\hat{x}_d(t) + K_d[y(t) - H_F(t)\hat{x}_d(t)] \quad (49)$$

The differential equation for the true estimation error is given as

$$\dot{e}_d = (F_F - K_dH_F)e_d + (TF_S - K_dH_S - F_FT + K_dH_FT)x_S + TG_Su_S$$
$$- K_dv_S + TC_Sv_d \quad (50)$$
$$= (F_F - K_dH_F + C_F - TC_S)e_d + (TF_SK_dH_S - F_FT + K_dH_FT + TC_ST$$
$$- C_FT)x_S + TG_Su_S - K_dv_S$$

It may be easily shown that equation (36), (37) and (38) may be modified as

$$\dot{V}_e = (F_F - K_dH_F + C_F - TC_S)V_e + V_e(F_F - K_dH_F + C_F - TC_S)^T + (TF_S$$

$$-K_dH_S - F_FT + K_dH_FT + TC_ST - C_FT)V_{12}^T + V_{12}(TF_S - K_dH_S - F_FT \quad (51)$$

$$+ K_dH_FT + TC_ST - C_FT)^T + TG_SQ_SG_S^TT^T + K_dR_SK_d^T$$

$$\dot{V}_{12} = (F_F - K_dH_F + C_F - TC_S)V_{12} + (TF_S - K_dH_S - F_FT + K_dH_FT + TC_ST$$

$$- C_FT)V_S + V_{12}(F_S - K_SH_S)^T + TG_SQ_SG_S^T + K_dR_SK_S^T \quad (52)$$

$$\dot{V}_S = (F_S - K_SH_S)V_S + V_S(F_S - K_SG_S)^T + G_SQ_SG_S^T + K_SR_SK_S^T \quad (53)$$

with known initial conditions.

The second moment equations may be modified as

$$\dot{P}_e = (F_F - K_dH_F + C_F - TC_S)P_e + (TF_S - K_dH_S - F_FT + K_dH_FT + TC_ST$$

$$- C_FT)P_{12}^T + P_e(F_F - K_dH_F + C_F - TC_S)^T + P_{12}(TF_S - K_dH_S - F_FT \quad (54)$$

$$+ K_dH_FT + TC_ST - C_FT)^T + TG_SQ_SG_S^TT^T + K_dR_SK_d^T$$

$$\dot{P}_{12} = (F_F - K_dH_F + C_F - TC_S)P_{12} + (TF_S - K_dH_S - F_FT + K_dH_FT + TC_ST$$

$$- C_FT)P_S + P_{12}F_S^T + TG_SQ_SG_S^T \quad (55)$$

$$\dot{P}_S = F_SP_S + P_SF_S^T + G_SQ_SG_S^T \quad (56)$$

with the known initial conditions and the bias equation remains that of equation (47). One may note an interesting feature of the above variance and second moment equations. This is that unlike the optimal variance equations since the control matrices feedback into the variance and second moment equations. Thus, it is possible to control the variance and second moment by judicious choice of the control matrices. Therefore, separation between control and estimation does not generally occur in ROF problems.

The quantities V_e, P_e, and e_d yield the statistical information for the ROF. In order to determine the effect of an unestimated error source on the filter estimation error one may consider the correlation matrices V_{12} and P_{12} where

$$V_{12} \triangleq E\{\delta e_d e_S^T\} \quad (57)$$

and

$$P_{12} \triangleq E\{e_d x_S^T\} \quad (58)$$

and are generated via the solutions to equations (37) and (45) respectively.

Appendix A expands the consideration of the ROF problem to the case of continuous dynamics, discrete measurements with continuous, impulsive and continuous plus impulsive control.

Appendix A. Continuous Dynamics, Discrete Measurement Problem

The equations giving the second moment, covariance, and bias of the true estimation error are developed herein for the continuous dynamics and discrete measurement problem. The problem of continuous, impulsive, and continuous plus impulsive control is discussed and the necessary equation modifications are given. Consider the dynamic equation given in (1). The system is observed at discrete time intervals t_k denoted by subscript "k". The measurement equation is

$$y(k) = H_S(k)x_S(k) + v_S(k) \tag{A-1}$$

where all other definitions are obvious from Section II and where

$$E\{V_S(k)V_S^T(j)\} = R(k)\delta(k - j).$$

An estimate of the states as defined in equation (3) is required. The filter design is based upon a reduced order structure given by the measurement equation

$$y(k) = H_F(k)x_F(k) + v_F(k) \tag{A-2}$$

where

$$E\{v_F(k)v_F^T(j)\} = R_F(k)\delta(k - j)$$

and the filter equations, between measurements

$$\dot{\hat{x}}_d = F_F(t)\hat{x}_d \tag{A-3}$$

and at a measurement

$$\hat{x}_d^+(k) = \hat{x}_d^-(k) + K_d(k)\{y(k) - H_F(k)\hat{x}_d(k)\} \tag{A-4}$$

where the gain $K_d(k)$ is calculated by using the optimal filtering equations given in [4] for the dynamics

$$\dot{x}_F(t) = F_F(t)x_F(t) + G_F(t)u_F(t) \tag{A-5}$$

where

$$E\{u_F(t)u_F^T(\gamma)\} = Q_F(t)\delta(t - \gamma)$$

and the measurement equation as given in A-2. The estimation error is defined in equation (9).

Between measurements the estimation error is given by

$$\dot{e}_d = F_F e_d + (TF_S - F_F T)x_S + TG_S u_S \tag{A-6}$$

The equation for the estimation error at a measurement is

$$e_d^+ = (I - K_d H_F) e_d^- + (K_d H_F T - K_d H_S) x_S - K_d v_S \qquad \text{(A-7)}$$

The equation for the bias is derived from the conditional expectation of e_d^+ as follows.

$$E\{e_d^+ | Y^k\} = (I - K_d H_F) E\{e_d^- | Y^k\} + (K_d H_F T - K_d H_S) E\{x_S | Y^k\}$$

$$= (I - K_d H_F) E\{e_d^- | Y^k\} + (K_d H_F T - K_d H_S) \hat{x}_S^+$$

Now,

$$E\{e_d^- | Y^k\} = E\{T x_S - \hat{x}_d^- | Y^k\}$$

$$= T\hat{x}_S^+ - E\{E\{x_d | Y^{k-1}\} | Y^k\}$$

$$= T\hat{x}_S^+ - \hat{x}_d^-$$

Thus,

$$\bar{e}_d^+ = (I - K_d H_F)(T\hat{x}_S^+ - \hat{x}_d^-) + (K_d H_F T - K_d H_S)\tilde{x}_S^+ \qquad \text{(A-8)}$$

$$\bar{e}_d^+ = (I - K_d H_F) e_d^- + (TK_S H_S - K_d H_S K_S H_S) e_S^- + (K_d H_F T - K_d H_S)\hat{x}_S^- \qquad \text{(A-9)}$$

$$+ (TK_S - K_d H_S K_S) v_S$$

One may define the new variables

$$\delta_e^+ = e_d^+ - \bar{e}_d^+ \qquad \text{(A-10)}$$

$$\delta_e^- = e_d^- - \bar{e}_d^- \qquad \text{(A-11)}$$

The equation for δe_d may be written as

$$\delta e_d^+ = (I - K_d H_F)\delta e_d^- + (K_d H_F T - K_d H_S) e_S^- - K_d V_S - (T - K_d H_S) K_S H_S e_S^-$$

$$- (T - K_d H_S) K_S v_S,$$

or

$$\delta e_d^+ = (I - K_d H_F)\delta e_d^- + (K_d H_F T - K_d H_S - TK_S H_S + K_d H_S K_S H_S) e_S^- \qquad \text{(A-12)}$$

$$+ (K_d H_S K_S - TK_S - K_d) v_S$$

In order to develop the equations for the variance propagation between measurements one may first define an augmented state z_3 where

$$z_3^+ = \begin{bmatrix} \delta e_d^+ \\ \\ e_S^+ \end{bmatrix} \tag{A-13}$$

Therefore,

$$z_3^+ = \begin{bmatrix} (I - K_d H_F), (K_d H_F T - K_d H_S - T K_S H_S + K_d H_S K_S H_S) \\ \\ 0 \quad , \quad (I - K_S H_S) \end{bmatrix} \bar{z}_3^+$$

$$\begin{bmatrix} K_d H_S K_S - T K_S - K_d \\ \\ -K_S \end{bmatrix} v_S \tag{A-14}$$

This is of the form

$$z_3^+ = \bar{A}\, \bar{z}_3 + \bar{B}\, v_S \tag{A-15}$$

The second moment of z_3 may be written as

$$P_{z_3}^+ = E\{z_3^+ z_3^{+T}\} \tag{A-16}$$

then

$$P_{z_3}^+ = \bar{A}\, P_{z_3}^-\, \bar{A}^T + \bar{B}\, R_S \bar{B}^T \tag{A-17}$$

The second moment matrix, $P_{z_3}^+$, may be partitioned as

$$P_{z_3}^+ = \begin{bmatrix} V_e^+ & V_{12}^+ \\ \\ V_{21}^+ & V_S^+ \end{bmatrix} \tag{A-18}$$

where V_e^+ is the required variance across a measurement the variance of the optimal Kalman filter. This yields the component equations

$$V_e^+ = (I - K_d H_F) V_e^- (I - K_d H_F)^T + (I - K_d H_F) V_{12}^- (K_d H_F T - K_d H_S - T K_S H_S + K_d H_S K_S H_S)^T + (K_d H_F T - K_d H_S - T K_S H_S + K_d H_S K_S H_S) V_{12}^{-T} (I - K_d H_F)^T$$

$$+ (K_d H_F T - K_d H_S - T K_S H_S + K_d H_S K_S H_S) V_S^- (K_d H_F T - K_d H_S - T K_S H_S + K_d H_S K_S H_S)^T + (K_d H_S K_S - T K_S - K_d) R_S (K_d H_S K_S - T K_S - K_d)^T \quad \text{(A-19)}$$

$$V_{12}^+ = (I - K_d H_F) V_{12}^- (I - K_S H_S)^T + (K_d H_F T - K_d H_S - T K_S H_S + K_d H_S K_S H_S) V_S^- (I - K_S H_S)^T - (K_d H_S K_S - T K_S - K_d) R_S K_S^T \quad \text{(A-20)}$$

$$V_S^+ = (I - K_S H_S) V_S^- (I - K_S H_S)^T + K_S R_S K_S^T \quad \text{(A-21)}$$

The variance between measurements may be easily derived from equations (36), (37) and (38) since K_d and K_S equal zero between measurements. This yields

$$\dot{V}_e = F_F V_e + V_e F_F^T + (T F_S - F_F T) V_{12}^T + V_{12} (T F_S - F_F T)^T + T G_S Q_S G_S^T T^T \quad \text{(A-22)}$$

$$\dot{V}_{12} = F_F V_{12} + (T F_S - F_F T) V_S + V_{12} F_S^T + T G_S Q_S G_S^T \quad \text{(A-23)}$$

$$\dot{V}_s = F_S V_S + V_S F_S^T + G_S Q_S G_S^T \quad \text{(A-24)}$$

with the known initial conditions.

The equation for the second moment across a measurement may be easily derived by use of the augmented state z_4, i.e.,

$$z_4^+ = \begin{bmatrix} e_d^+ \\ x_S^+ \end{bmatrix} \quad \text{(A-25)}$$

Thus,

$$z_4^+ = \begin{bmatrix} (I - K_d H_F), (K_d H_F T - K_d H_S) \\ 0 \qquad\qquad I \end{bmatrix} z_4^- + \begin{bmatrix} -K_d \\ 0 \end{bmatrix} v_S \quad \text{(A-26)}$$

and is of the form

$$z_4^+ = \bar{A}\, z_4^- + \bar{B}\, v_S \quad \text{(A-27)}$$

This yields

$$P_{z_4}^+ = \bar{A}\, P_{z_4} \bar{A}^T + \bar{B}\, R_S \bar{B}^T \quad \text{(A-28)}$$

$$P^+_{z_4} = \begin{bmatrix} P_e & P_{12} \\ P_{21} & P_S \end{bmatrix}$$

The component equations are

$$P_e^+ = (I - K_dH_F)P_e^-(I - K_dH_F)^T + (I - K_dH_F)P_{12}^-(K_dH_FT - K_dH_S)^T + (K_dH_FT - K_dH_S)P_{12}^{-T}(I - K_dH_F)^T + K_dR_SK_d^T \quad \text{(A-29)}$$

$$P_{12}^+ = (I - K_dH_F)P_{12}^- + (K_dH_FT - K_dH_S)P_S(K_dH_FT - K_dH_S)^T \quad \text{(A-30)}$$

where P_S is the value of the second moment of x_S at the update time as generated by the equations between measurements (the measurements do not affect P_S). The equations between measurements may be easily derived from equations (44), (45), and (46) since K_d and K_S equal zero between measurements.

$$\dot{P}_e = F_FP_e + (TF_S - F_FT)P_{12}^T + P_eF_F^T + P_{12}(TF_S - F_FT)^T + TG_SQ_SG_S^TT^T \quad \text{(A-31)}$$

$$\dot{P}_{12} = F_FP_{12} + (TF_S - F_FT)P_S + P_{12}F_S^T + TG_SQ_SG_S^T \quad \text{(A-32)}$$

$$\dot{P}_S = F_SP_S + P_SF_S^T + G_SQ_SG_S^T \quad \text{(A-33)}$$

with the known conditions.

The equations for continuous control between measurements may be easily derived from equations (51) through (56) by setting K_d and K_S equal to zero between measurements. This also takes care of the continuous control plus impulsive control problem between measurements as the impulsive control only takes place at a measurement. Thus, the equations between measurements for the variance are given as

$$\dot{V}_e = (F_F + C_F - TC_S)V_e + V_e(F_F + C_F - TC_S)^T + (TF_S - F_FT + TC_ST - C_FT)V_{12}^T + V_{12}(TF_S - F_FT + TC_ST - C_FT)^T + TG_SQ_SG_S^TT^T \quad \text{(A-34)}$$

$$\dot{V}_{12} = (F_F + C_F - TC_S)V_{12} + (TF_S - F_FT + TC_ST - C_FT)V_S + V_{12}F_S^T + TG_SQ_SG_S^T \quad \text{(A-35)}$$

$$\dot{V}_S = F_SV_S + V_SF_S^T + G_SQ_SG_S^T \quad \text{(A-36)}$$

The equations for the second moment may be derived by setting $K = K_d = 0$ in equations (54), (55) and (56). The bias may be calculate as before.

The impulsive control problem will now be solved. At a measurement the system state is modified by use of an impulsive control matrix ∇_S, i.e.,

$$x_S^+ = x_S^- + \nabla_S \hat{x}_d^+ \tag{A-37}$$

The filter update equation at the time of control application is given by

$$\hat{x}_d^{c+} = (I + \nabla_F)\hat{x}_d^+ \tag{A-38}$$

where

$$\hat{x}_c^+ = \hat{x}_d^- + K_d(y - H_F\hat{x}_d^-) \tag{A-39}$$

and $\hat{x}_d^-$ represents $\hat{x}_d$ at the time t- after the impulsive control has been applied. Thus, at a measurement and after an impulsive control update

$$\begin{aligned}
e_d^+ &= Tx_S^+ - \hat{x}^{c+} \\
&= T(x_S^- + \nabla_S\hat{x}_d^+) - (I - \nabla_F)[\hat{x}_d^- + K_d(H_S x_S^- + v_S - H_F\hat{x}_d^-)] \\
&= Tx_S^- + T\nabla_S[\hat{x}_d^- + K_d(H_S x_S + v_S - H_F\hat{x}_d^-)] - (I + \nabla_F)[\hat{x}_d^- + K_d \\
&\cdot (H_S x_S^- + v_S - H_F\hat{x}_d^-)] \\
&= [T + T\nabla_S K_d H_S - (I + \nabla_F)K_d H_S]x_S^- + [T\nabla_S - T\nabla_S K_d H_F - (I + \nabla_F) \\
&+ (I + \nabla_F)K_d H_F]\hat{x}_d^- + [T\nabla_S K_d - (I + \nabla_F)K_d]v_S \\
&= [-T\nabla_S + T\nabla_S K_d H_F + (I + \nabla_F) - (I + \nabla_F)K_d H_F](Tx_S^- - \hat{x}_d^-) \\
&+ [T\nabla_S K_d H_S - (I + \nabla_F)K_d H_S + T\nabla_S T - T\nabla_S K_d H_F T - \nabla_F T + (I \\
&+ \nabla_F)K_d H_F T]x_S^- + [T\nabla_S K_d - (I + \nabla_F)K_d]v_S \\
&= [T\nabla_S K_d H_F - T\nabla_S + (I + \nabla_F) - (I + \nabla_F)K_d H_F]e_d^- + [T\nabla_S K_d(H_S \\
&- H_F T) + (I + \nabla_F)(K_d H_F T - K_d H_S) + T\nabla_S T - \nabla_F T]x_S^- + [T\nabla_S K_d \\
&- (I + \nabla_F)K_d]v_S \\
&= m_1 e_d^- + m_2 x_S^- + m_3 v_S
\end{aligned} \tag{A-40}$$

where the definitions of the m_i's are obvious.

The equation for the bias will now be derived. The conditional expectation of e_d^+ may be taken. This yields

$$E\{e_d^+|Y^k\} = m_1 E\{e_d^-|Y^k\} + m_2 E\{x_S|Y^k\}$$

$$= m_1 E\{e_d|Y^k\} + m_2 \hat{x}_S^+ \tag{A-41}$$

where $\hat{x}_S^+$ is the optimal estimate of the state x_S given the measurement sequence Y^k but prior to any impulsive control action. Now,

$$E\{e_d^-|Y^k\} = E\{Tx_S^- - \hat{x}_d^-|Y^k\}$$

$$= T\hat{x}_S^+ - E\{E\{x_d|Y^{k-1}\}|Y^k\} \tag{A-42}$$

$$= T\hat{x}_S^+ - \hat{x}_d^-$$

Since

$$\hat{x}_S^+ = \hat{x}_S^- + K_S(y - H_S\hat{x}_S^-) \tag{A-43}$$

Then one may rewrite equation (A-42) as

$$E\{e_d^-|Y^k\} = T[\hat{x}_S^- + K_S H_S e_S^- + K_S v_S] - \hat{x}_d^-$$

$$= T\hat{x}_S^- - \hat{x}_d^- + TK_S H_S e_S^- + TK_S v_S \tag{A-44}$$

where

$$e_S^- = x_S - \hat{x}_S^- \tag{A-45}$$

By definition

$$\bar{e}_d^- = E\{Tx_S^- - \hat{x}_d^-|Y^{k-1}\} \tag{A-46}$$

$$= T\hat{x}_S^- - \hat{x}_d^-$$

Thus,

$$E\{e_d^-|Y^k\} = \bar{e}_d^- + TK_S H_S e_S^- + TK_S v_S \tag{A-47}$$

This may be used to rewrite equation (A-45) as

$$\bar{e}_d^+ = m_1\bar{e}_d^- + m_1 TK_S H_S e_S^- + m_1 TK_S v_S + m_2\hat{x}_S^+$$

$$= m_1\bar{e}_d^- + m_1 TK_S H_S e_S^- + m_1 TK_S v_S + m_2\hat{x}_S^- + m_2 K_S v_S + m_2 K_S H_S e_S^- \tag{A-48}$$

$$= m_1 e_d^- + (m_1 TK_S H_S + m_2 K_S H_S) e_S^- + m_2 \hat{x}_S + (m_1 TK_S + m_2 K_S) v_S$$

Now by definition

$$\delta e_d^+ = e_d^+ - \bar{e}_d^+$$

$$= m_1 \delta e_d^- + (m_2 - m_1 TK_S H_S - m_2 K_S H_S) e_S^- + (m_3 - m_1 TK_S \quad \text{(A-49)}$$

$$- m_2 K_S) v_S$$

where e_S satisfies the equation

$$e_S^+ = (I - K_S H_S) e_S^- - K_S v_S \quad \text{(A-50)}$$

One may define the augmented state

$$z_5^+ = \begin{bmatrix} \delta e_d^+ \\ \\ \\ e_S^+ \end{bmatrix} \quad \text{(A-51)}$$

This yields the state equation

$$z_5^+ = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} z_5^- + \begin{bmatrix} B_1 \\ \\ B_2 \end{bmatrix} v_S \quad \text{(A-52)}$$

$$= \bar{A} z_5^- + \bar{B} v_S$$

where

$$A_{11} = m_1$$

$$A_{12} = (m_2 - m_1 TK_S H_S - m_2 K_S H_S)$$

$$A_{22} = (I - K_S H_S)$$

$$B_1 = m_3 - m_1 TK_S - m_2 K_S$$

and

$$B_2 = -K_S$$

The second moment of z_5 may be written as

$$P_{z_5}^{+} = \bar{A}\, P_{z_5}\, \bar{A}^T + \bar{B}\, R_S \bar{B}^T \tag{A-53}$$

where

$$P_{z_5} = \begin{bmatrix} \bar{V}_e & \bar{V}_{12} \\ V_{21} & V_S \end{bmatrix}$$

The component equations may be written as

$$V_e^{+} = A_{11}\bar{V}_e^{-}A_{11}^T + A_{12}\bar{V}_{12}^{-T}A_{11}^{+} + A_{11}\bar{V}_{12}^{-}A_{12}^T + A_{12}\bar{V}_S^{-}A_{12}^T + B_1 R_S B_1^T \tag{A-54}$$

$$V_{12}^{+} = A_{11}\bar{V}_{12}^{-}A_{22}^T + A_{12}\bar{V}_S^{-}A_{22}^T + B_1 R_S B_2^T \tag{A-55}$$

$$V_S^{+} = A_{22}\bar{V}_S^{-}A_{22}^T + B_2 R_S B_2^T \tag{A-56}$$

where V_e is the desired true variance and V_S is the optimal filter variance. The variance between measurements propagates according to equations (A-34), (A-35) and (A-36).

In order to derive the equation for the second moment across a measurement one may consider the following augmented state

$$z_6^{+} = \begin{bmatrix} e_d^{+} \\ x_S^{+} \end{bmatrix}$$

The equation for e_d^{+} is given by (A-40). In order to derive the equation for x_S^{+} in terms of x_S^{-} and e_d^{-} consider equations (A-37) and (A-39). They may be used as follows

$$\begin{aligned} x_S^{+} &= x_S^{-} + \nabla_S x_d^{+} \\ &= x_S^{-} + \nabla_S[\hat{x}_d^{-} + K_d(H_S x_S + v_S - H_F \hat{x}_d^{-})] \\ &= (I + \nabla_S K_d H_S)x_S^{-} + (\nabla_S - \nabla_S K_d H_F)\hat{x}_d^{-} + \nabla_S K_d v_S \\ &= (\nabla_S K_d H_F - \nabla_S)(Tx_S^{-} - \hat{x}_d^{-}) + (I + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T)x_S^{-} + \nabla_S K_d v_S \\ &= (\nabla_S K_d H_F - \nabla_S)e_S^{-} + (I + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T)x_S^{-} + \nabla_S K_d v_S \end{aligned}$$

Thus,

$$z_6^+ = \begin{bmatrix} m_1 & , & m_2 \\ (\nabla_S K_d H_F - \nabla_S, & & (I + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T) \end{bmatrix} z_6^- + \begin{bmatrix} m_3 \\ \nabla_S K_d \end{bmatrix} v_S$$

The second moment equations may be easily found as

$$P_e^+ = m_1 P_e^- m_1^T + m_2 P_{12}^{-T} m_1^T + m_1 P_{12}^- m_2^T + m_2 P_S^- m_2^T + m_3 R_S m_3^T \quad \text{(A-57)}$$

$$\begin{aligned} P_{12}^+ = {} & m_1 P_e^- (I + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T)^T + m_2 P_S^- (I + \nabla_S K_d H_S \\ & - \nabla_S K_d H_F T + \nabla_S T)^T + m_1 P_e^- (\nabla_S K_d H_F - \nabla_S)^T + m_2 P_{12}^{-T} (\nabla_S K_d H_F \\ & - \nabla_S)^T + m_3 R_S (\nabla_S K_d)^T \end{aligned} \quad \text{(A-58)}$$

$$\begin{aligned} P_S^+ = {} & (\nabla_S K_d H_F - \nabla_S) P_e^- (\nabla_S K_d H_F - \nabla_S)^T + (\nabla_S K_d H_F - \nabla_S) P_{12}^- (I \\ & + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T)^T + (I + \nabla_S K_d H_S - \nabla_S K_d H_F T \\ & + \nabla_S T) P_{12}^{-T} (\nabla_S K_d H_F - \nabla_S)^T + (I + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T) P_S^- (I \\ & + \nabla_S K_d H_S - \nabla_S K_d H_F T + \nabla_S T)^T + \nabla_S K_d R_S (\nabla_S K_d)^T \end{aligned} \quad \text{(A-59)}$$

The second moment between measurements propagates according to equations (54), (55), and (56) by setting $K_d = K_S = 0$. The bias may be calculated as before.

References

1. Crawford, B. S., J. Dunn, and A. Sutherland, Jr., "CIRIS Design Evaluation", TASC Tech Rpt 213-1, 1970.

2. Wu, T. K., "Application of Kalman Filtering to the Tactical Aircraft Navigation System", Proceedings of the First NWC Symposium on the Application of Control Theory to Modern Weapons Systems, China Lake, Calif., 1973.

3. Leondes, C. T., "Theory and Applications of Kalman Filtering", AGARDograph No. 139, 1970.

4. Jazwinski, A. H., Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

5. Gelb, A., Applied Optimal Estimation, MIT Press, Mass., 1973.

6. Fitzgerald, R. J., "Divergence of the Kalman Filter", IEEE Trans on A. C., Vol AC - 16, no 6, Dec 1971.

7. Aoki, M., Optimization of Stochastic Systems, Academic Press, New York, 1967.

8. Kalman, R. E., "New Methods in Wiener Filtering Theory", in Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability, John Wiley and Sons Inc., New York, 1963.

9. Nash, R. Jr., K. Roy and J. Kasper, Jr., "GEANS Error Minimization Study", TASC Tech Rpt 172-3, 1970.

10. Asher, R. B. and D. Watjen, "Kalman Filtering for Precision Pointing and Tracking", AFAL Tech Rpt to be published, 1974.

11. D'Appolito, J. A., "The Evaluation of Kalman Filter Designs for Multisensor Integrated Navigation Systems", TASC, AFAL-TR-70-271, 1971.

12. Duiven, E. M., "Suboptimal Linear Filtering", AIAA J. of Spacecraft and Rockets, Vol 11, No 3, 1974.

13. Warren, A. W., "Bias and Covariance Formula for Filter Error Analysis and Model Sensitivity", Proceedings of the 1973 Joint Automatic Control Conference, Columbus, Ohio.

14. Aoki, M. and J. R. Huddle, "Estimation of the State Vector of a Linear Stochastic System with a Constrained Estimator", IEEE Trans AC, August 1957.

15. Sims, C. S. and J. L. Melsa, "Specific Optimal Estimation:, IEEE Trans AC, April 1967.

16. Griffin, R. E. and A. P. Sage, "Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms", IEEE Trans AC, AS - 13, 1968.

17. Griffin, R. E. and A. P. Sage, "Sensitivity Analysis of Discrete Filtering and Smoothing Algorithms", AIAA Journal, Volume 7, 1969.

18. Heffes, H., "The Effect of Erroneous Models on Kalman Filter Response", IEEE Trans AC, July 1966.

19. Stubberud, A. R. and D. A. Wismer, "Suboptimal Kalman Filter Techniques", in Theory and Applications of Kalman Filtering by C. T. Leondes.

20. Brown, R. J., and A. P. Sage, "Analysis of Modeling and Bias Errors in Discrete - Time State Estimation", IEEE Trans on AES, Vol AES - 7, No 2, 1971.

21. Brown, R. J. and A. P. Sage, "Error Analysis of Modeling and Bias Errors in Continuous Time State Estimation", Automatica, Vol 7, 1971.

SEQUENTIAL ESTIMATION OF RANDOM THRUST ERRORS FOR SOLAR ELECTRIC PROPULSION SPACECRAFT

by

Major Thomas J. Eller

-Department of Astronautics and Computer Science-
USAF Academy, Colorado

ABSTRACT

Adaptability of Dynamic Model Compensation (DMC) algorithms for compensating for the effects of dynamical model errors due to random anomalies in the continuous-low-thrust propulsion system of a spacecraft on a 152-day trajectory from Earth to the asteroid Eros is investigated through digital stochastic simulation. Onboard star tracker and Earth-based range-rate observations are used in an extended sequential filter to estimate state and acceleration vectors. Limiting accuracies obtainable using four Gauss-Markov-process error models are determined.

INTRODUCTION

Filtering techniques have been used with good results for determining the orbits of near-earth, lunar, and ballistic interplanetary spacecraft, both in simulations and in actual operations. Because of the extremely long distances involved, high quality sensors and sophisticated computations are required for the orbit determination of interplanetary spacecraft. For continuously thrusting solar-electric propulsion (SEP) spacecraft, errors due to unmodeled forces in the thrust present a fundamental limitation on the navigation accuracy. These random, unmodeled variations in the thrust--which represent the dominant source of error--are three orders of magnitude larger than the other purturbations that also affect ballistic spacecraft. In 1970 Rourke and Jordan concluded that conventional least-squares filtering methods were not adequate for SEP missions.[1] In 1973 McDannell concluded that because of the large errors, even the idealized sequential filter is unsatisfactory.[2]

Since the thrust is applied continuously, even small anomalies will eventually cause significant estimation errors. The effect of these errors as well as those due to linearization assumptions can be compensated to some extent by using the extended form of a sequential filter. Filter divergence due to model errors can be delayed by the addition of a state noise covariance matrix to the equation used to propagate the state error covariance matrix. One particular approach to compensating for errors in the mathematical model was introduced to statistical orbit determination by Ingram and Tapley (1971) and is now known as Dynamic Model Compensation (DMC).[3]

DMC consists of assuming that the dynamical model is a linear combination of the known (modeled) and unknown (error) dynamics and is modeled in two parts:

1. The model for the known dynamics is intended to be the best available mathematical representation of the physical process involved--subject to computational realizability and efficiency considerations.

2. The model for the unknown dynamics is made by loosely specifying a model whose form and parameters are then estimated along with the system state. This model is not intended to mathematically emulate complicated

physical processes in the same manner as they may be suspected to occur, but to merely produce a similar effect using a simple form that changes under the influence of a priori observation residuals, hence the name Dynamic Model Compensation. The concepts of modeling, filtering, errors, and error compensation are shown in Figure 1.

PROBLEM STATEMENT

The dynamical system considered is a solar electric spacecraft assumed to be influenced by random errors in the thrust. A 152-day mission from heliocentric injection at the Earth's sphere of influence to a fly-by of the asteroid Eros at 1.45 AU is simulated. The projection of the design trajectory on the eccliptic plane is shown as a solid line labeled LOW THRUST in Figure 2. The dotted trajectory is the elliptical path of a non-thrusting (BALLISTIC) vehicle with the same initial conditions. This trajectory is shown as a basis for comparison with the powered trajectory. Also shown are astronomical symbols to identify the orbits of Earth, Mars, and Eros.

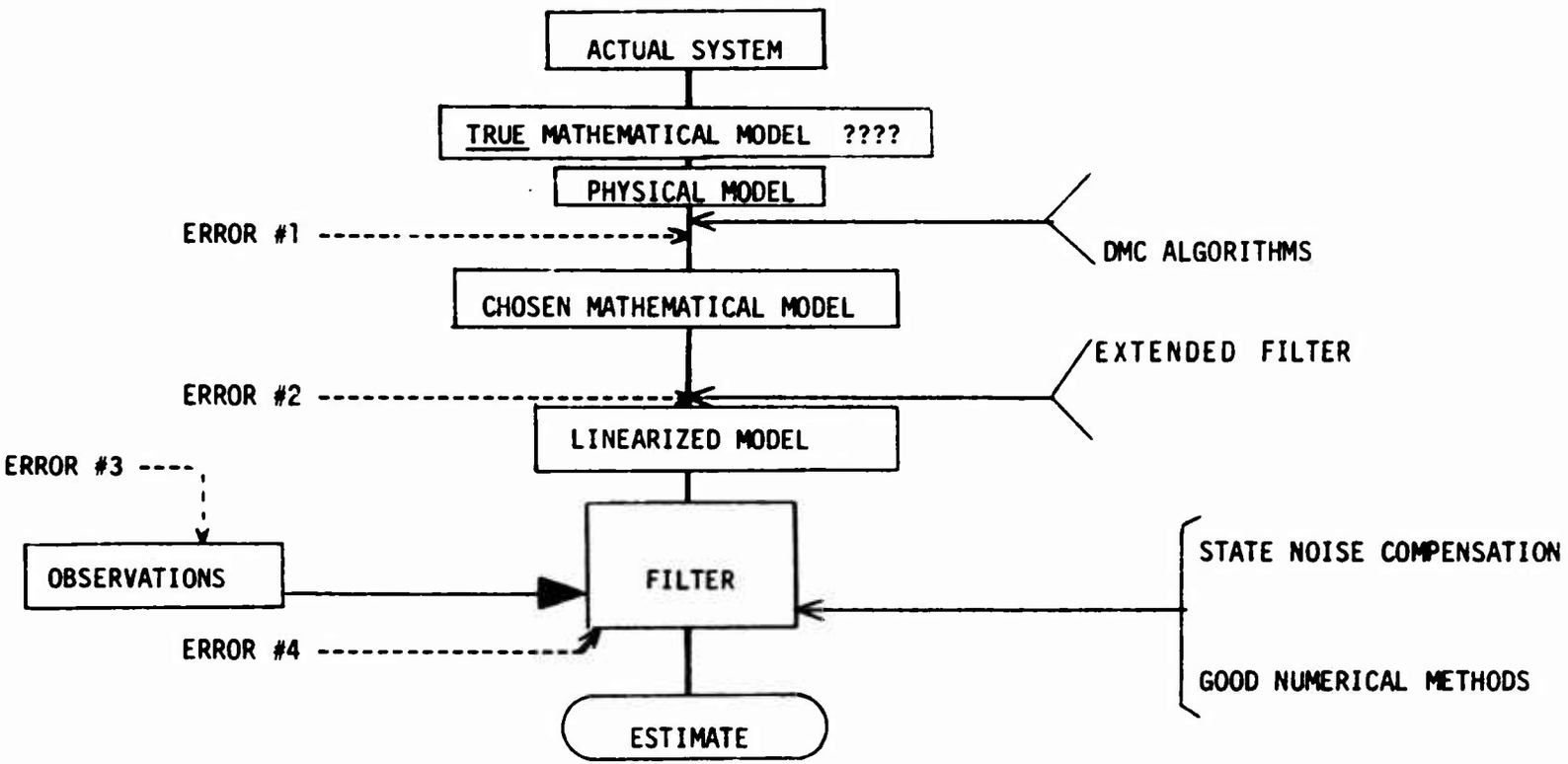
If the only gravitational force considered is due to the Sun, the equations of motion for the point mass vehicle are:

$$\dot{r} = v$$
$$\dot{v} = -\frac{\mu}{|r|^3} r + T \qquad (1)$$

where, as shown in Figure 3, r is a 3-vector of heliocentric position components, X, Y, Z; v is a 3-vector of heliocentric velocity components $\dot{X}$, $\dot{Y}$, $\dot{Z}$; $|r|$ is the magnitude of r; μ is the gravitational parameter of the sun; and T is the thrust acceleration produced by the solar electric engine. The thrust contains errors due to beam voltage and current variations, grid warpage, deadband control errors, etc. The heliocentric components of T are related to the vehicle centered, orbit frame components by

$$T = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = R \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \qquad (2)$$

where ψ is the heliocentric orientation angle as shown in Figure 3. The two reference frames are oriented such that the Z and z axes are parallel. The X



TYPES OF ERRORS

1. Errors due to inaccurate or incomplete modeling
2. Errors due to linearization assumptions.
3. Errors which occur in the process of observing
4. Errors introduced in the computation

FIGURE 1 CONCEPTS OF MODELING, ERRORS, AND ERROR COMPENSATION

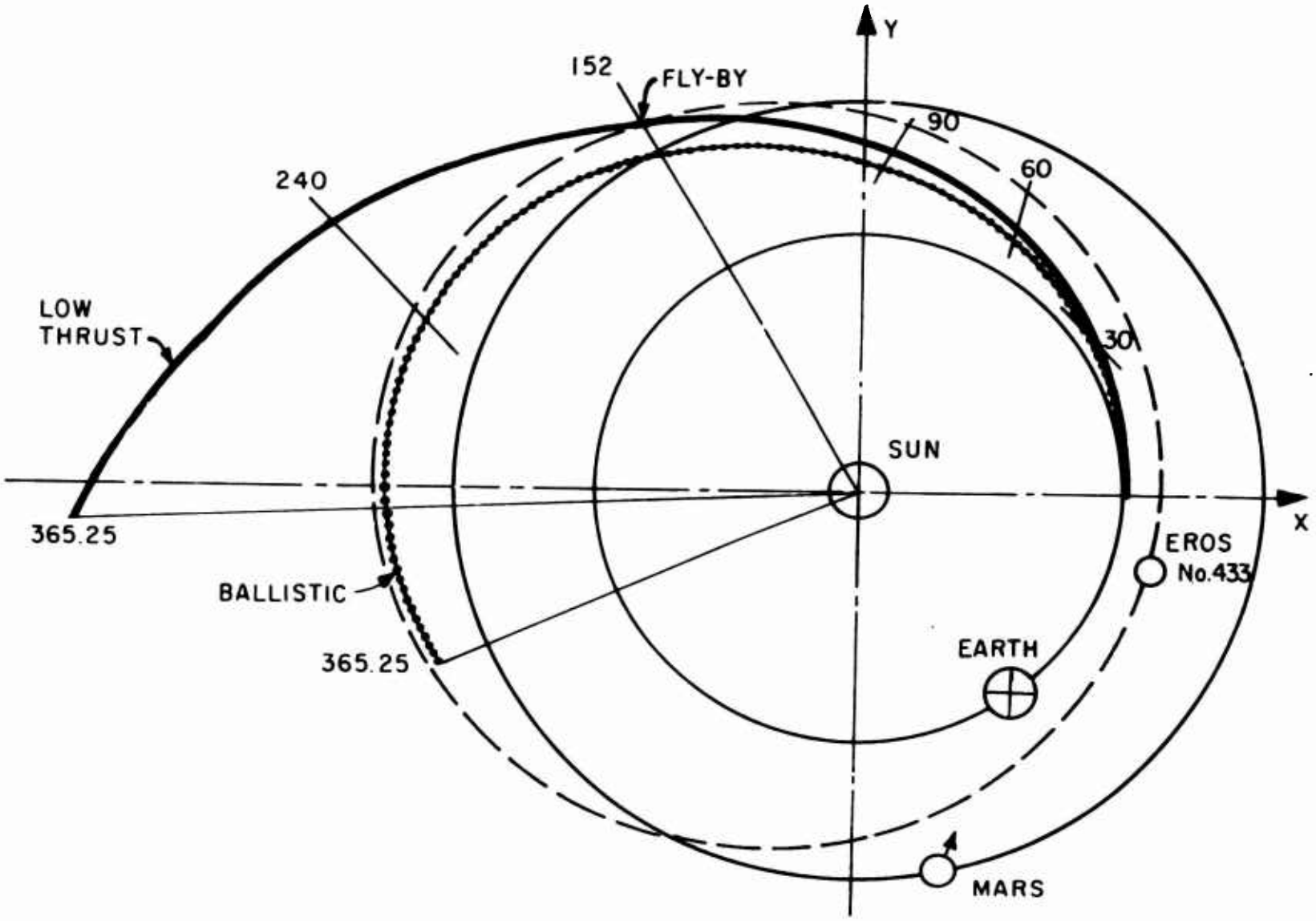


Figure 2 Mission Trajectories

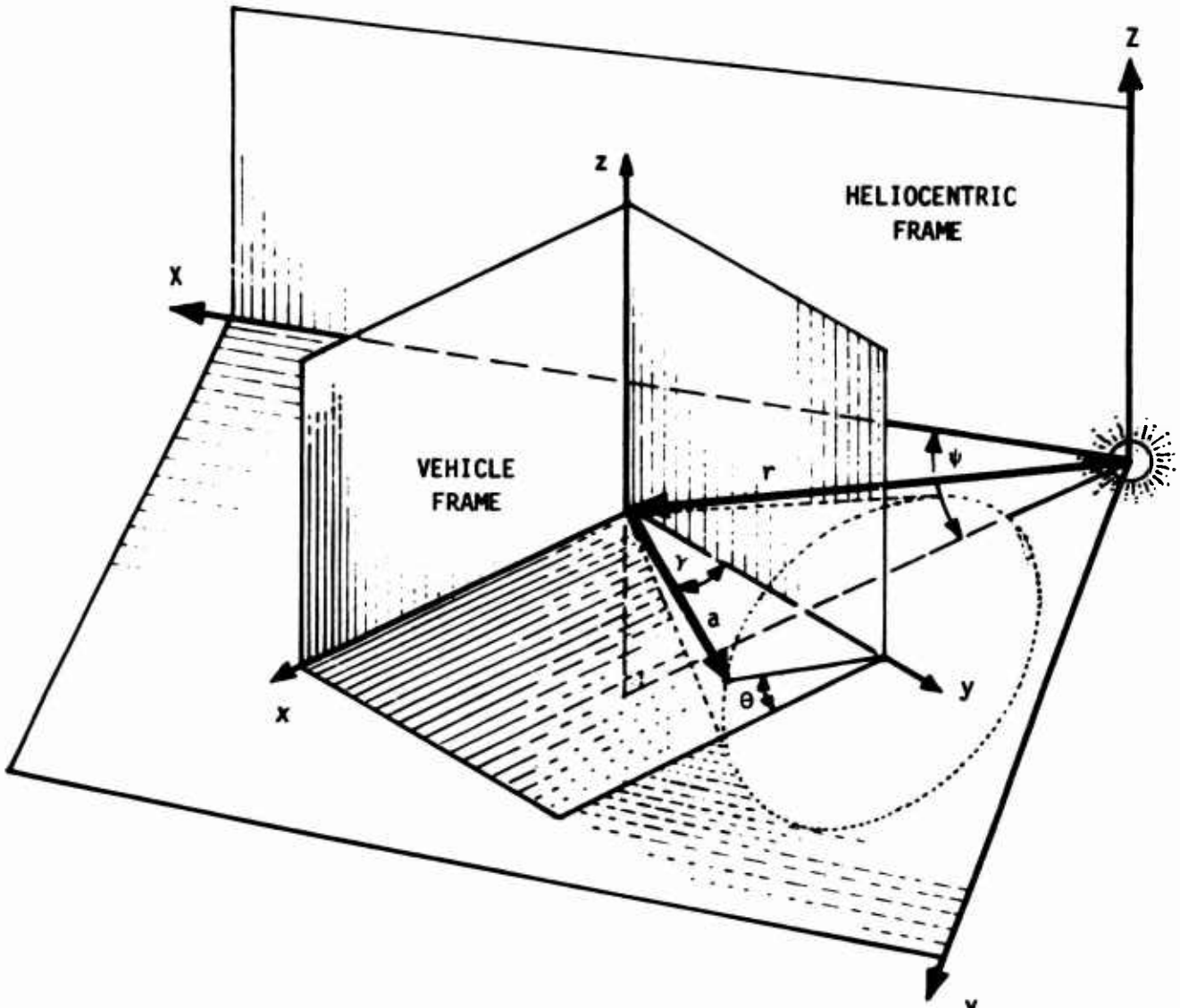


Figure 3 Reference Frames

and x axes form the angle ψ, with $\cos\psi = X/(X^2 + Y^2)^{\frac{1}{2}}$, and $\sin\psi = Y/(X^2 + Y^2)^{\frac{1}{2}}$. Thus the x axis of the vehicle frame is always along the eccliptic plane projection of the heliocentric position vector of the spacecraft.

The true thrust acceleration, T, can be resolved conveniently into its orbital frame components in terms of its magnitude a, and the clock and cone angles θ and γ as shown in Figure 3. Thus,

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = a \begin{bmatrix} \sin\gamma \cos\theta \\ \cos\gamma \\ \sin\gamma \sin\theta \end{bmatrix} \tag{3}$$

$$a = a^* + \delta a \tag{4}$$

where a^* is the magnitude of the design thrust acceleration and δa is its associated error. It is seen from Equation (3) that if γ and δa are zero, the true thrust is the same as the design thrust.

The acceleration error magnitude is simulated using

$$\delta a = \delta a_0 \sin\omega t + u_a \tag{5}$$

where δa_0 and ω are constants and u_a is a random variable with the statistics

$$E[u_a] = 0, \qquad E[u_a^2] = \sigma_a^2 \tag{6}$$

The relationship between successive instantaneous values of the pointing angles γ and θ are assumed to be characterized by both time correlated and purely random aspects.

Two types of observations are used in the orbit determination process. These are range-rate, measured by radar from the Earth, and the angle formed by the lines-of-sight from the vehicle to the Earth and to a specified navigation star.

FILTER ALGORITHM

The sequential linear filter has been applied to nonlinear problems by linearizing about a fixed reference solution. In this form, large initial estimation errors propagate through time and, if the state deviations become

large, violating our linearity assumptions, we soon have filter divergence. A technique intended to reduce these adverse effects consists of reinitializing the reference trajectory after processing each set of observations. This relinearization about the most current best estimate of the state is the basis of the formulation used here known as the Extended Sequential Filter (see Jazwinski, 1970, p. 272).[4]

ERROR COMPENSATION MODELS

The dynamical model used for the estimation process is essentially the same as Equation 1, except that we assume that the thrust acceleration can be separated into modeled and error components (see the discussion of DMC). Thus

$$T = T^* + m(t) \tag{7}$$

where $m(t)$ is a 3-vector of unknown thrust acceleration error components. This vector, $m(t)$, is referred to as the unmodeled acceleration since it represents all accelerations not accounted for in the mathematical model used to describe the motion of the low thrust vehicle.

The errors $m(t)$ are approximated by $\tilde{R}\varepsilon(t)$, where $\tilde{R}$ is the coordinate transformation matrix and $\varepsilon(t)$ satisfies one of several possible first-order or second-order differential equations. The elements of $\varepsilon(t)$ are the three orthogonal components of the acceleration error in the orbital reference frame. The values of $\varepsilon(t)$ and any unknown parameters in the differential equations which describe $\varepsilon(t)$ are part of the state vector and are estimated simultaneously with the position and velocity components. Four models are used as approximations to $m(t)$ as described below.

The filter dynamical model is expressed mathematically as

$$\begin{aligned} \dot{r} &= v \\ \dot{v} &= -\frac{\mu}{|r|^3} r + T^* + \tilde{R}\varepsilon \end{aligned} \tag{8}$$

where r and v are the heliocentric position and velocity, respectively, μ is the gravitational parameter for the Sun,

$$T^* = \tilde{R} \begin{bmatrix} 0 \\ a^* \\ 0 \end{bmatrix}, \quad \tilde{R} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{bmatrix}, \quad (9)$$

where ψ is the heliocentric orientation angle (see Section 2.1), and ε is governed by one of the following differential equations:

Model 0. Q-matrix compensation for process noise

$$\varepsilon = u \quad (10)$$

where u is a random vector with statistics given by Equation 12 and the state vector is $X^T = [r^T \; v^T]$. This model is based on the assumption that the thrust errors are purely random process noise; thus, an arbitrary constant state noise covariance matrix is used to maintain a positive definite state error covariance.

Model 1A. A first-order Gauss-Markov process

$$\dot{\varepsilon} = u_\varepsilon \quad (11)$$

where u_ε is a random 3-vector with a priori statistics

$$E[u] = 0 \qquad E[u(t)u(\tau)] = q(t)\delta(t-\tau) \quad (12)$$

The state vector is $X^T = [r^T \; v^T \; \varepsilon^T]$.

Model 1B. A first-order Gauss-Markov process

$$\begin{aligned} \dot{\varepsilon} &= -\alpha\varepsilon + u_\varepsilon \\ \dot{\alpha} &= u_\alpha \end{aligned} \quad (13)$$

where α is a diagonal matrix and u_ε and u_α are random with the same generic expression for their statistics as Equation 12 above. The state vector is $X^T = [r^T \; v^T \; \varepsilon^T \; \alpha^T]$.

Model 2A. A second-order Gauss-Markov process

$$\begin{aligned} \dot{\varepsilon} &= \eta \\ \dot{\eta} &= u_\eta \end{aligned} \quad (14)$$

where u_η is random with statistics given by Equation 12. The state vector is $X^T = [r^T \; v^T \; \varepsilon^T \; \eta^T]$.

Model 2B. A second-order Gauss-Markov process (undamped harmonic oscillator)

$$\begin{aligned} \dot{\varepsilon} &= \eta \\ \dot{\eta} &= -\beta\varepsilon + u_\eta \\ \dot{\beta} &= u_\beta \end{aligned} \tag{15}$$

where β is a diagonal matrix and u_η and u_β are random statistics given by Equation 12. The state vector is $X^T = [r^T\ v^T\ \varepsilon^T\ \eta^T\ \beta^T]$.

In each of the models above, for each equation that is driven by random noise, an arbitrary noise covariance sub-matrix q is added to the corresponding differential equation governing the state error covariance. The full state noise co-variance matrix is symbolized by Q.

A characterization of the solution of each of these models as an approximation to time correlated and random thrust errors is shown in Figure 4. These characterizations are as follows:

Model 1A - horizontal straight line

Model 1B - exponential curve

Model 2A - non-zero slope straight line

Model 2B - sinusoidal curve

Since the values of the acceleration as well as those of the parameters in the equations are reset by the extended sequential filter after each observation is processed, the particular solution is valid only until the next observation epoch, when it may change again.

SIMULATION

Figure 5 is a conceptual flow diagram of the computation. The right side of the figure, labeled Simulation, is the real-world for this study. It represents the Sun, the spacecraft, the observation stations moving with the Earth, all hardware associated with spacecraft navigation and tracking, and all errors that arise outside the Estimator. All of this knowledge enters the Estimator at one point--as Observed Observations--which is the Estimator's only contact with the world, whether it be an actual or simulated world.

As presented above, the true thrust as simulated for this investigation may be expressed in the vehicle reference frame as

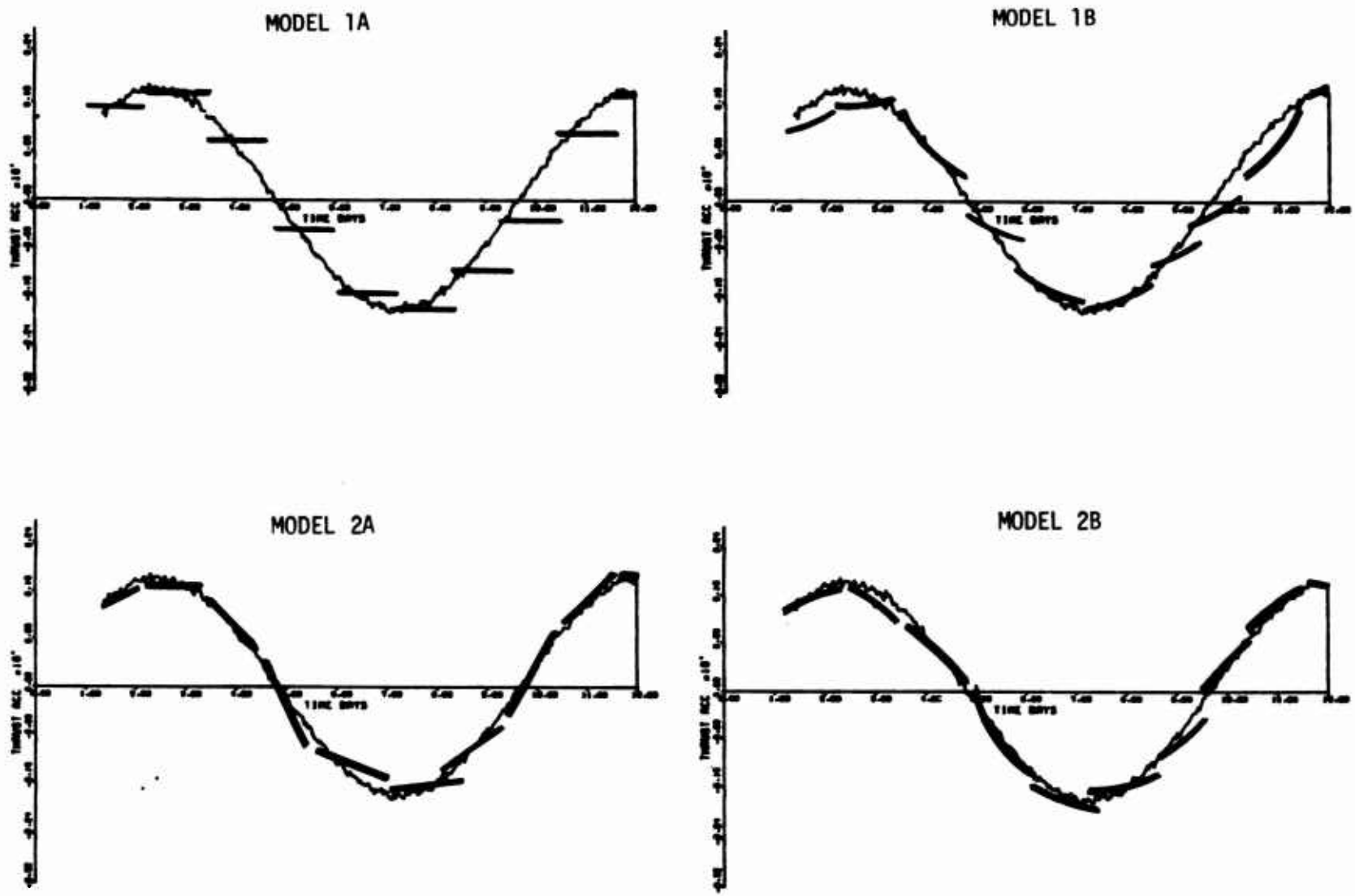


Figure 4 Error Compensation Model Characterization

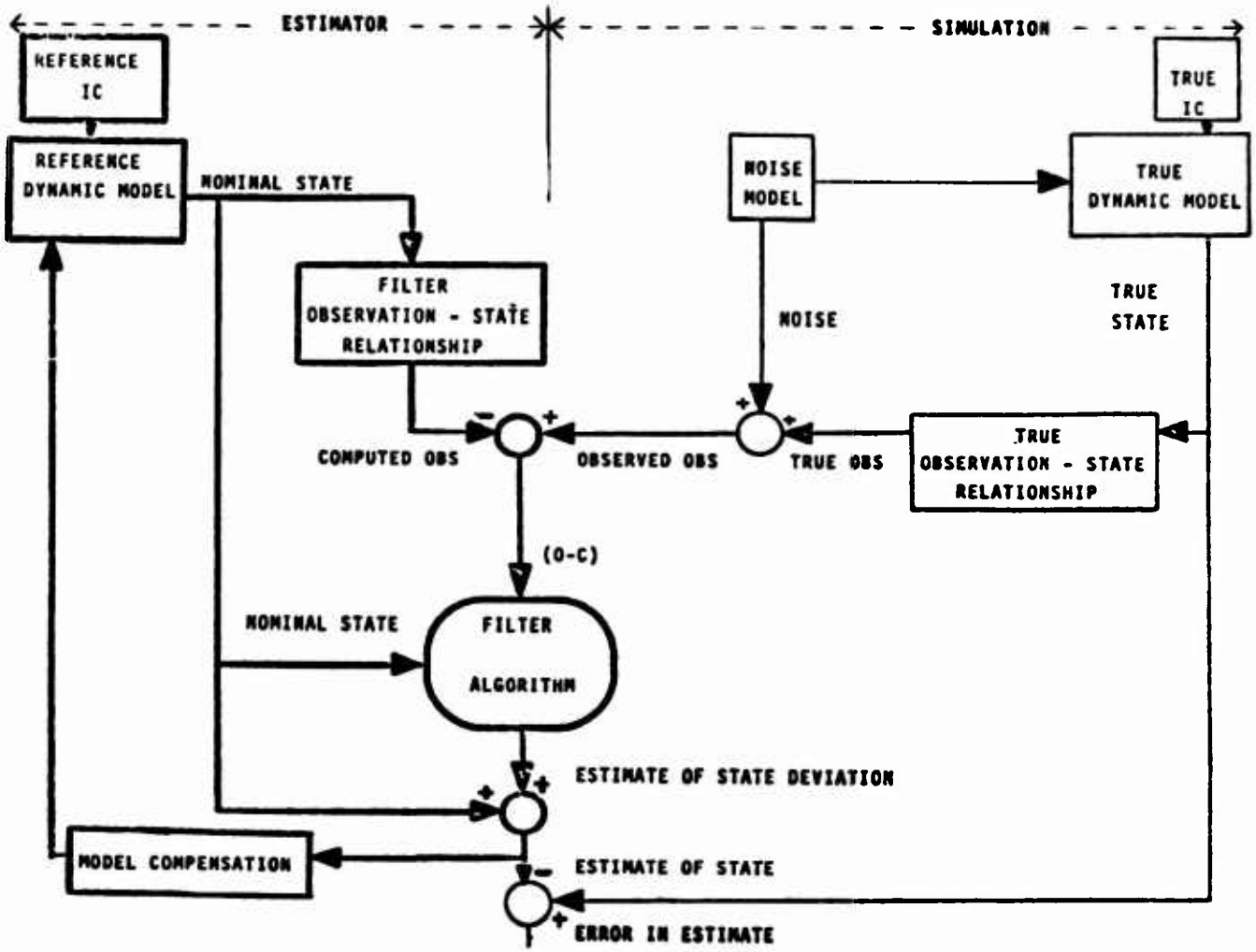


Figure 5 Conceptual Flow Diagram of Computation

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = (a^* + \delta a_o \sin \omega t + u_a) \begin{bmatrix} \sin \gamma \cos \theta \\ \cos \\ \sin \gamma \sin \theta \end{bmatrix} \qquad (16)$$

In simulating the thrust acceleration error, the instantaneous values of the pointing angles, γ and θ, are assumed to be related as shown in Figure 3. During the simulation, random numbers are introduced into a, θ, γ, s, and ϕ as described in Reference 5. A separate random number generator is used for each random distribution specified in the problem. Typical thrust error simulation is illustrated by Figure 6.

As an indication of the nature of the thrust error components produced by the simulation, a series of rough approximations can be made that results in the following representation of the type of unknown errors that must be approximated by the $\epsilon(t)$ of Equation 8.

$$\epsilon_x = (0.35 + .02\sin .6t + .002\sin 144t) \qquad (17)$$
$$(.02\sin 6t + .001\sin 144t)(\cos 3t + .001\sin 144t)$$

RESULTS

References 6 and 7 give results for this problem where only the magnitude of the thrust error is estimated. The simulations discussed here include estimates of all three components of the thrust error.

Each figure below shows the error norm and the square root of the trace of the appropriate portion of the state error covariance matrix for 160 days of simulation using three tracking stations with an observation interval of 50 minutes. Model 0 uses only a state noise covariance matrix, Q, to compensate for modeling errors. Figure 7 shows the results of a 35-day run of Model 0 with a bad choice of values for the state noise covariance matrix. Figure 8 shows the results of a 160-day run of the same case, but with an improved Q matrix. Note that the Position estimate is improved significantly and, for the latter case, that the error is bounded by the covariance. Figure 9 shows the results for Model 2B. Note the improvement in the velocity estimate. Figure 10 shows the simulated and estimated thrust errors. Note that for the y component (along the nominal thrust direction) the estimate can hardly be distinguished from the actual thrust. Table 1 compares the

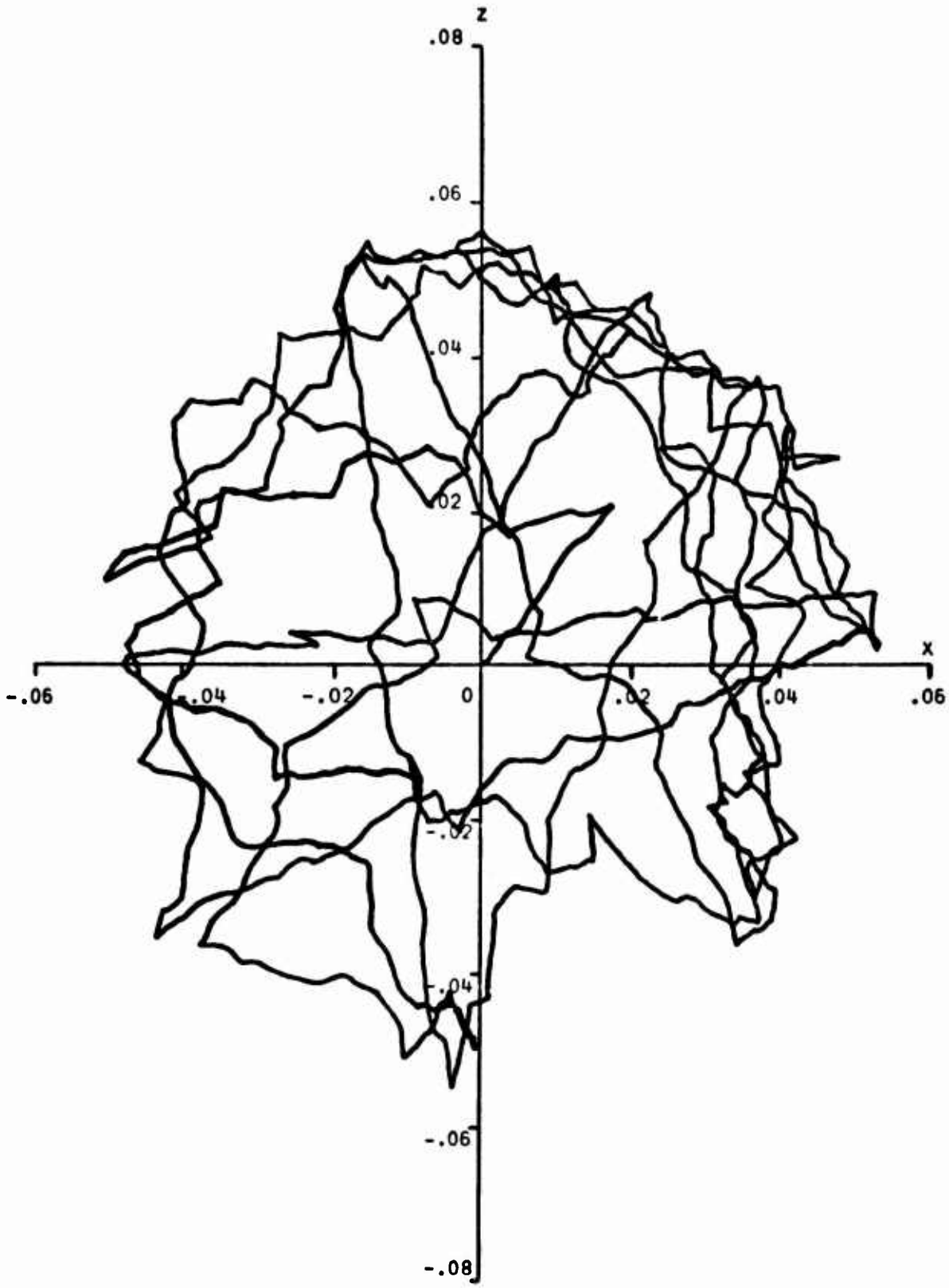
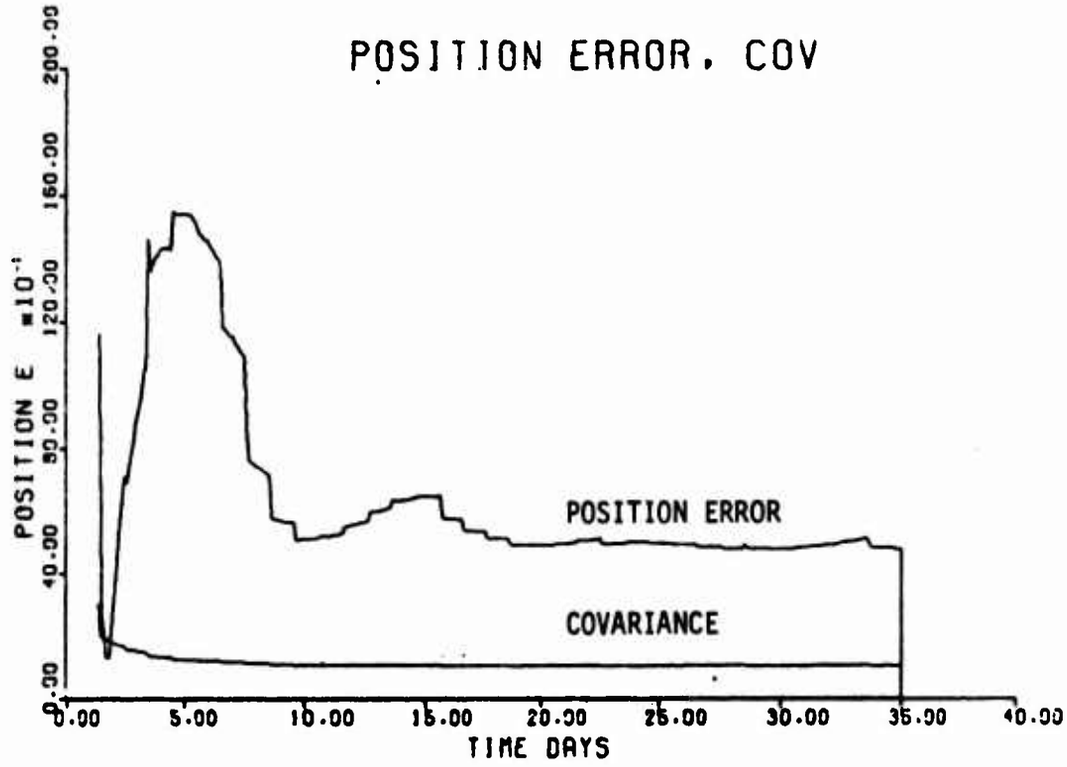


Figure 6 Simulated Acceleration Errors in x-z Plane for 35 Days

(10^{-4} m/sec^2)



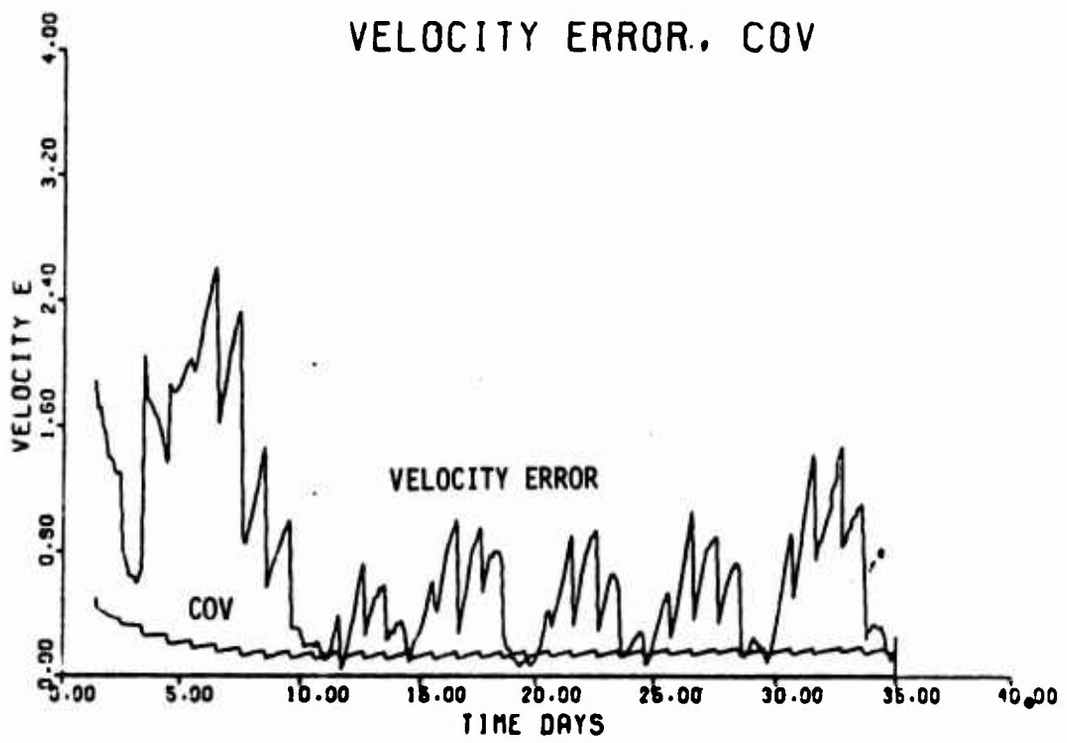
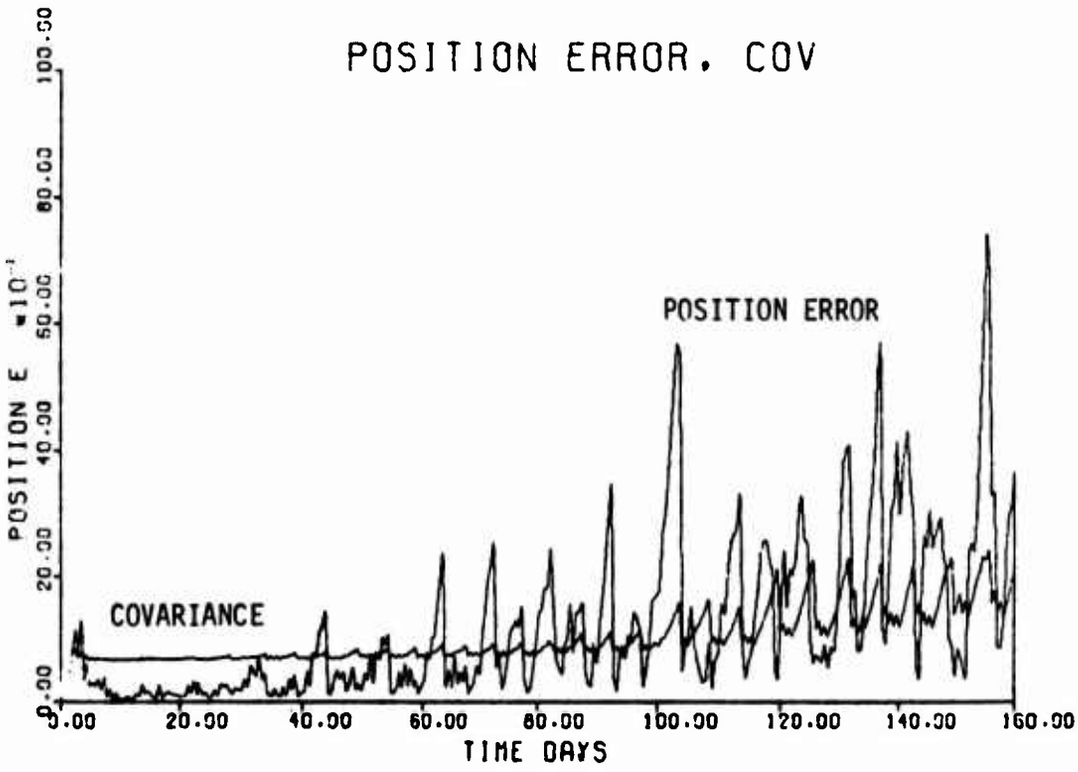


Figure 7 Model 0, Not Well Tuned



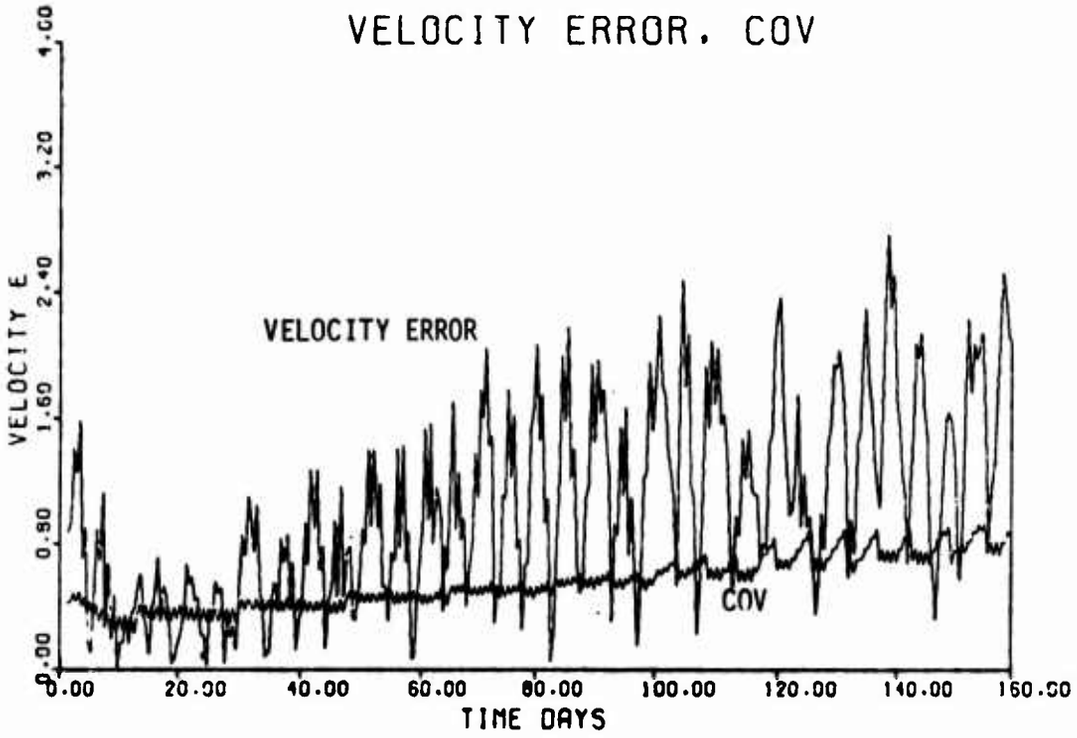
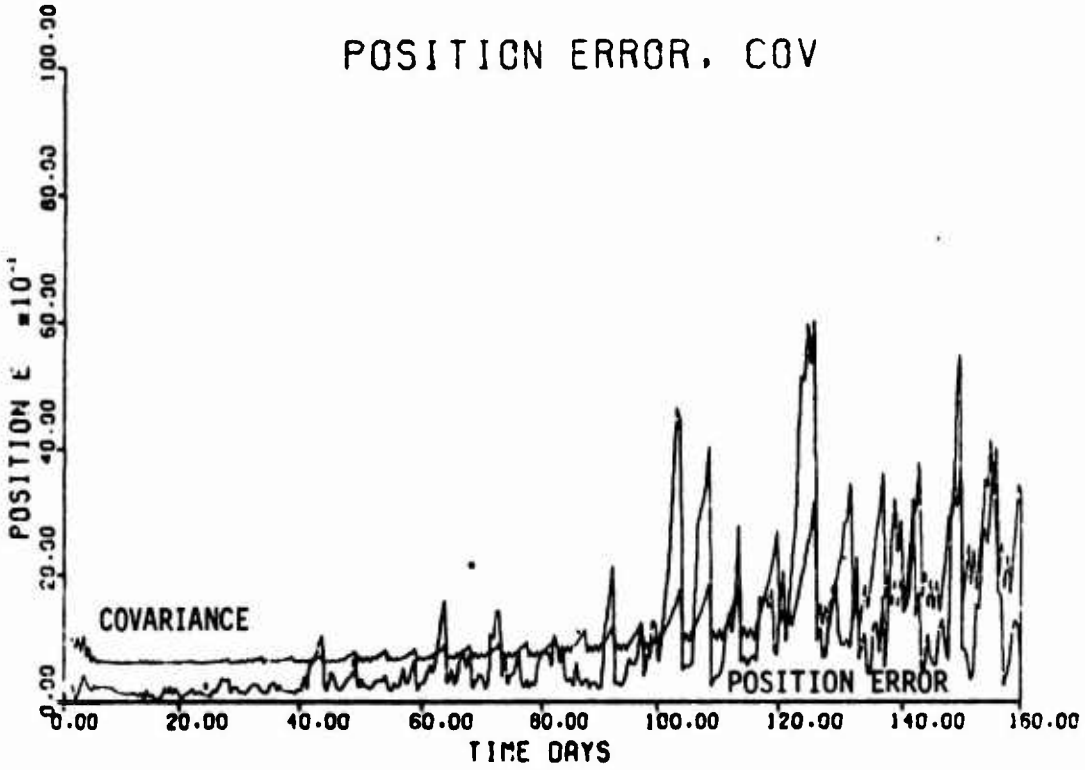


Figure 8 Model 0, Well Tuned



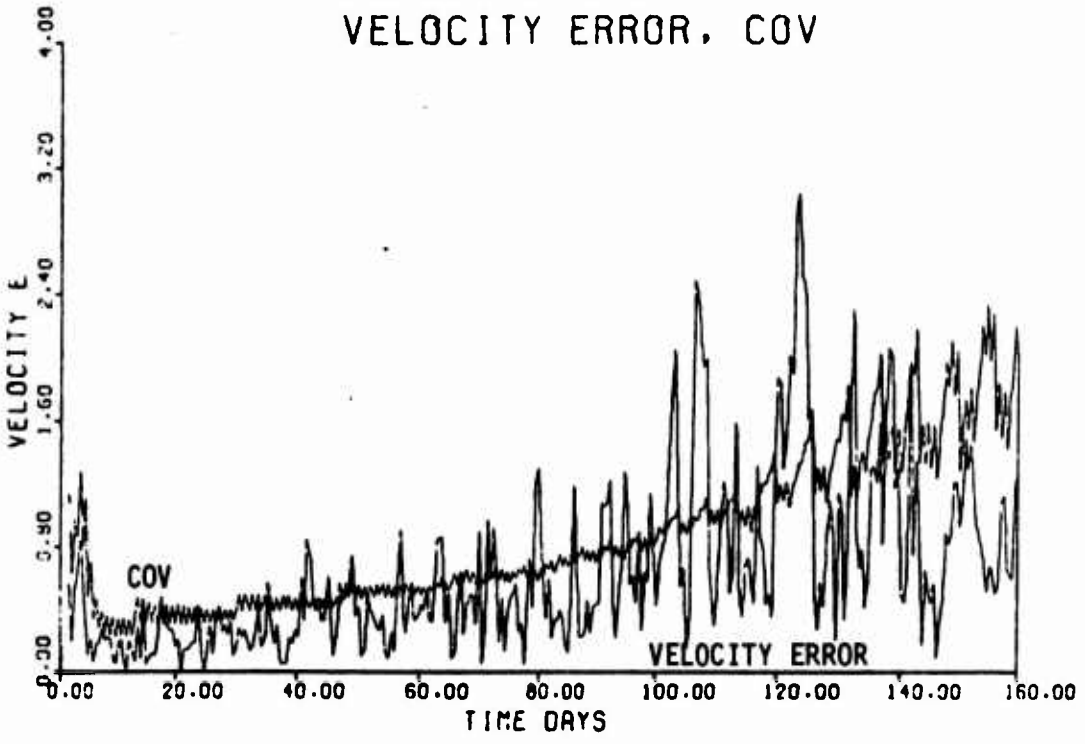


Figure 9 Position and Velocity, Model 2B

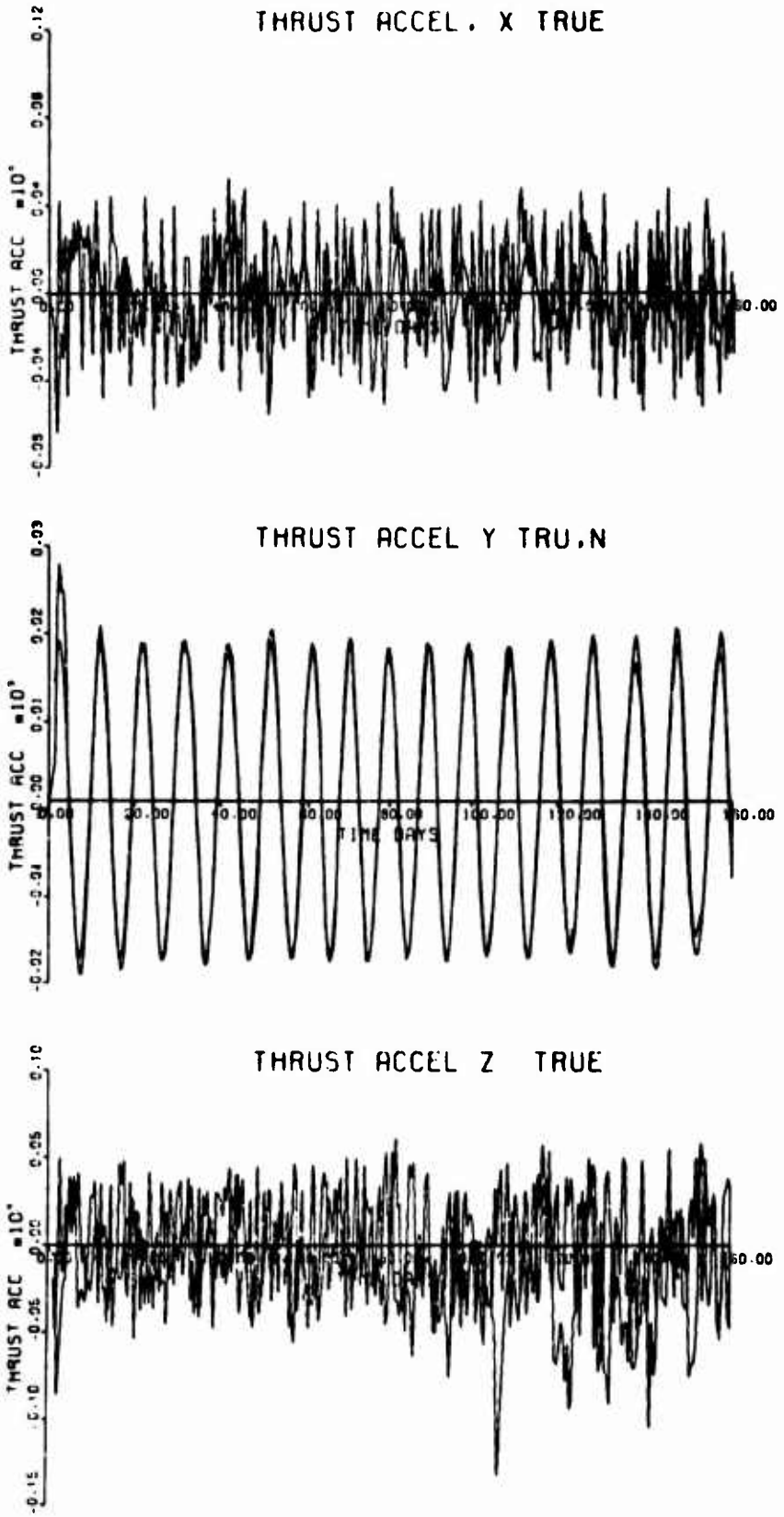


Figure 10 Simulated and Estimated Thrust Errors, Model 2B

MODEL	POSITION ERROR (km) Average	Peak	VELOCITY ERROR (m/sec) Average	THRUST ERROR (m/sec^2) x 10^6 x	y	z
0	116.98	450	1.009			
1A	123.70	680	1.123	2.571	6.974	4.141
1B	148.01	700	1.600	7.239	11.92	12.23
2A	173.09	1600	2.458	11.57	24.92	26.30
2B	86.17	400	0.613	2.247	1.183	3.653

TABLE 1. Filter Performance Data

performance of the various models. Clearly, for the particular runs shown, Model 2B was the best. However, just as indicated for Model 0 above, a better choice of Q matrix might significantly improve the results for the other models. For a detailed analysis of these and other results, see reference 5.

CONCLUSIONS

For the continuous low-thrust navigation problem considered in this investigation it has been found that if sufficient data is available to allow the estimation of the additional quantities in the enlarged state vector, then the use of DMC Gauss-Markov models is superior to Q-matrix compensation for the purpose of estimating the vehicle state, and yields an estimate of the accelerations acting on the space vehicle as well.

Furthermore, the second-order model, an undamped harmonic oscillator, gives the best results. However, it is apparent that poor choices for the state noise covariance matrices can result in larger estimation errors than the differences in estimation accuracy between the models tested.

SYSTEM-LEVEL IMU CALIBRATION

Richard A. Kitzerow

Air Force Avionics Laboratory
Analysis and Evaluation Branch AFAL/NVE-3
Wright-Patterson Air Force Base, Ohio 45433

Abstract

The United States Air Force Avionics Laboratory at Wright-Patterson Air Force Base, Ohio, is presently conducting inhouse advanced and exploratory development in the area of system-level inertial measurement unit (IMU) calibration techniques.

This paper describes an optimal filter IMU error analysis methodology developed for simultaneous fault detection, fault isolation, calibration and alignment utilizing only system-level IMU velocity and gimbal synchro angle measurements. The analysis methodology is applicable to gimballed and strapdown IMUs in either a fixed or perturbed base environment. IMU error parameter estimation can be accomplished either on-line in real time or off-line using recorded IMU outputs.

Simulation results obtained from a CDC-6600 computer simulation of a 57 state Kalman filter calibrating a gimballed KT-70 IMU indicate that the technique is conceptually more comprehensive, accurate, and will require less IMU specialized test equipment and test time than conventional calibration methods. Due to the conceptual advantages and the minimal specialized AGE requirements, the technique appears to be a likely candidate for intermediate and depot level IMU maintenance and calibration.

I. INTRODUCTION

Current IMU fault detection, fault isolation and calibration methods employ costly specialized AGE and either component or system-level tests to calibrate only a fraction of the significant IMU error sources. The costs of the time consuming testing, specialized test equipment and other maintenance function items presently constitute the major portion of IMU cost-of-ownership. The excessive test time and AGE requirements of the current methods also prohibit the AFAL/NVE-3, and other organizations involved in inertial navigation software development and validation, from accurately establishing IMU status immediately prior to conducting a software validation test.

Recognizing the need to reduce weapon system avionics cost-of-ownership and to enhance it's software validation capability, AFAL/NVE-3 initiated in July 1973 an inhouse exploratory development program to develop a generalized IMU check-out, calibration and alignment methodology. Another objective of the program is to reduce IMU cost-of-ownership by simplifying flight-line, intermediate and depot level IMU check-out and calibration while reducing the associated specialized AGE requirements.

Methodology versatility objectives include on-line and off-line data reduction, gimballed and strapdown IMU mechanizations and fixed and perturbed base applications. Performance goals are to provide a comprehensive calibration, maximum accuracy and to minimize IMU test time. It is desirable to estimate each significant IMU error source to an accuracy such that the navigational accuracy achievable with the IMU would be limited only by uncompensatable random disturbances. Ground rules are that the IMU not be disassembled, only easily accessible system-level IMU measurements be used, and that the solution be implementable on a general purpose digital computer.

II. Calibration Methodology

The performance and versatility goals dictate that the methodology be based on only the essential functions an IMU performs. These functions are realization of a reference platform coordinate frame and measurement of the instantaneous specific force vector. Noting that fault detection, fault isolation, calibration and alignment are error estimation processes, the IMU error analysis methodology formulated can be understood with the following rationale. If the orientation of the platform of an error free IMU was initially exactly known with respect to the physical earth rate and gravity vectors, then both the orientation of the platform with respect to these vectors and the platform coordinatized specific force vector could be deterministically calculated as a function of time. Any deviations between the ideal time dependent platform orientation and accelerometer outputs and those of a real IMU must be caused by either IMU error sources or an initial platform misalignment. This fact is true regardless of initial platform orientation and can be employed for the derivation of system-level IMU error equations relating measurable IMU performance parameters to IMU and platform misalignment angle error quantities.

Application of the generalized methodology to a fixed-base gimballed IMU can be illustrated with the IMU line schematic of Figure 1. The fundamental function of the gyros, gimbals and servo feedback control system is to maintain a reference platform coordinate frame relative to an inertial frame. If the gyro triad of an ideal IMU were commanded to precess in equality with the earth's angular rotation vector, then the angular velocity of the platform with respect to an earth fixed reference frame would be identically zero. This ideal IMU performance would result in a fixed angular orientation of the gimbals relative to each other and constant

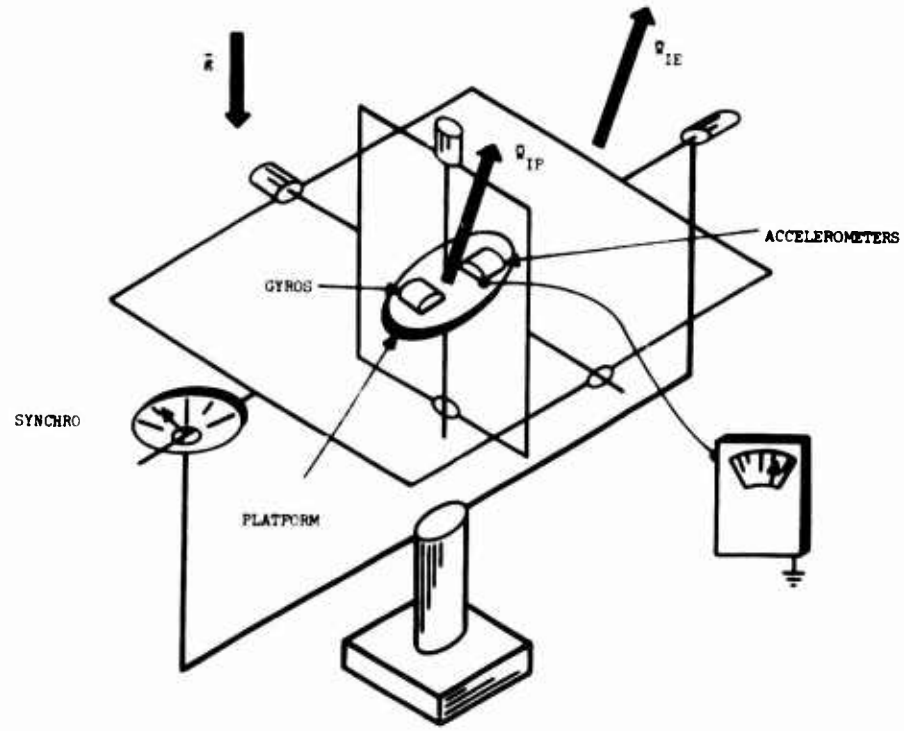


FIGURE (1) GIMBALLED IMU LINE SCHEMATIC

system-level synchro gimbal angle and accelerometer measurement outputs. The exact angular orientation of the platform frame with respect to the earth rate and gravity vectors must be known for a priori computation of the required gyro precession rate torquing magnitudes. In the real world, inexact initial platform attitude information and IMU error sources preclude the above ideal IMU performance. Synchro gimbal angle and accelerometer measurements exhibit a time dependency.

III. Error Equation Development

The essential task involved in developing the error equations is the derivation of expressions relating system-level IMU synchro gimbal angle and velocity measurements to IMU and platform-to-navigation frame misalignment angle error quantities. A prerequisite for the derivation is an IMU error model explicitly containing each IMU error source's functional dependency. To achieve the stated performance goals, the error model must contain all IMU error sources that are significant relative to the gyro and accelerometer inertial instrument random disturbances.

A Kearfott KT-70 gimballed IMU was chosen for the methodology application largely because it is utilized in the AFAL/NVE-4 Mobile Evaluation Laboratory as part of the reference navigation system and requires periodic status validation and calibration. Reference (1) contains a suitable error model for the IMU.

Expressions for the time rates of change of the IMU gimbal angles can be derived by noting that the gimbals rotate in such a manner as to isolate the platform from the IMU case angular velocity. The angular velocity vector of the platform frame relative to an inertial frame coordinatized in platform frame coordinates is given as:

$$\bar{W}^{P}_{I,P} = C^{P}_{g}\bar{W}^{g}_{I,g}$$

C^{P}_{g} is a nonorthogonal error transformation matrix to account for the physical misalignments of the gyro input axes from the platform reference frame.

$$C^{P}_{g} = \begin{bmatrix} 1 & -\beta_{xz} & \beta_{xy} \\ \beta_{yz} & 1 & -\beta_{yx} \\ -\beta_{zy} & \beta_{zx} & 1 \end{bmatrix}$$

$\bar{W}^{g}_{I,g}$ is the angular velocity vector of the gyro frame relative to an inertial frame coordinatized in gyro frame coordinates.

$\bar{W}_{I,g}$ is composed of two basic vectors. The first represents the angular velocity at which the gyros are commanded to precess relative to inertial space $\bar{W}_{I,C}$. The second represents gyro frame angular velocity relative to inertial space due to all gyro imperfections and is commonly referred to as gyro drift $\bar{W}_{I,D}$.

$$\bar{W}^{g}_{I,g} = \bar{W}^{g}_{I,C} + \bar{W}^{g}_{I,D}$$

To maintain the gyro frame stabilized at approximately a constant attitude with respect to the navigation frame, it is necessary to torque the gyros corresponding to the best a priori information of the coordinatization of the earth rate vector in the gyro frame.

$$\bar{W}^{g}_{I,C} = [K_F]\overline{PRF} = C^{P}_{PG}C^{PG}_{RG}C^{RG}_{CASE}\bar{W}^{N}_{I,N}$$

$[K_F]$ is a 3 × 3 diagonal matrix of the loaded gyro scale factors. The true gyro scale factors $[K_{FT}]$ differ from the loaded ones by an error $[\Delta K_F]$. $\overline{PRF}$ is a 3 state vector of the gyro torquing pulse repetition frequencies. $\bar{W}^{N}_{I,N}$ is the angular velocity vector of the navigation frame relative to inertial space coordinatized in the navigation frame.

$$\bar{W}^{N}_{I,N} = \begin{bmatrix} \Omega C(L) \\ 0 \\ \Omega S(L) \end{bmatrix}$$

Ω is the magnitude of the earth rate vector. L is the latitude. C^{RG}_{CASE} is the IMU case to roll (outer) gimbal frame coordinate transformation matrix.

$$C^{RG}_{CASE} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & C(\theta_x) & S(\theta_x) \\ 0 & -S(\theta_x) & C(\theta_x) \end{bmatrix}$$

C^{PG}_{RG} is the roll gimbal to pitch (middle) gimbal frame coordinate transformation matrix.

$$C^{PG}_{RG} = \begin{bmatrix} C(\theta_y) & 0 & -S(\theta_y) \\ 0 & 1 & 0 \\ S(\theta_y) & 0 & C(\theta_y) \end{bmatrix}$$

C^{P}_{PG} is the pitch gimbal to azimuth (inner) gimbal or platform frame coordinate transformation matrix.

$$C^{P}_{PG} = \begin{bmatrix} C(\theta_z) & S(\theta_z) & 0 \\ -S(\theta_z) & C(\theta_z) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

θ_x, θ_y and θ_z are the nominal synchro gimbal angle readings which are obtained by subsequent rotations about the roll, pitch and azimuth gimbal axes.

The gyro drift vector is an error quantity composed of all gyro error sources relative to and including the random disturbances.

$$\bar{W}^{g}_{I,D} = \bar{W}_B + \bar{W}_A + \bar{W}_M + \bar{W}_O + \bar{W}_{RL} + \bar{W}_{RS}$$

$\bar{W}_B$ is the constant gyro bias drift vector. $\bar{W}_A$ is the platform-to-case attitude dependent gyro bias drift vector. $\bar{W}_M$ is the gyro mass unbalance drift vector:

$$\bar{W}_M = \begin{bmatrix} -M_{Sx} & 0 & M_{Ix} \\ 0 & -M_{Sy} & M_{Iy} \\ 0 & -M_{Iz} & M_{Sz} \end{bmatrix} \bar{f}^{g}_{s}$$

M_{Sx}, M_{Sy} and M_{Sz} are the gyro spin axes mass unbalance coefficients. M_{Ix}, M_{Iy} and M_{Iz} are the gyro input axes mass unbalance coefficients. $\bar{f}^{g}_{s}$ is the specific force coordinatized in the gyro frame:

$$\bar{f}^{g}_{s} = C^{g}_{P} C^{P}_{PG} C^{PG}_{RG} C^{RG}_{CASE} C^{CASE}_{N} \bar{f}^{N}_{s}$$

C^{CASE}_{N} is an orthogonal small angle navigation to IMU case frame coordinate transformation matrix.

$$C^{CASE}_{N} = \begin{bmatrix} 1 & \eta_z & -\eta_y \\ -\eta_z & 1 & \eta_x \\ \eta_y & -\eta_x & 1 \end{bmatrix}$$

$$\bar{f}^{N}_{s} \simeq \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix}$$

g is the magnitude of the earth's gravity vector. $\bar{W}_O$ is the gyro anisoelastic drift vector:

$$\bar{W}_O = \begin{bmatrix} K_{Ox} \cdot f^{g}_{sy} \cdot f^{g}_{sz} \\ K_{Oy} \cdot f^{g}_{sx} \cdot f^{g}_{sz} \\ K_{Oz} \cdot f^{g}_{sx} \cdot f^{g}_{sy} \end{bmatrix}$$

K_{Ox}, K_{Oy} and K_{Oz} are the gyro anisoelastic coefficients. $\bar{W}_{RL}$ is the long correlation time exponentially correlated gyro random drift vector. $\bar{W}_{RS}$ is the short correlation time exponentially correlated gyro random drift vector.

$$\dot{\bar{W}}_R(t) = \frac{-\bar{W}_R(t)}{\tau} + \sigma\sqrt{\frac{2}{\tau}}\,\bar{N}(t)$$

τ is the correlation time. σ is the standard deviation. $N(t)$ is unity variance white noise.

The above equations for the IMU error sources, IMU case-to-navigation frame misalignment angles and gyro frame commanded angular velocity permit calculation of the angular velocity of the platform frame relative to an inertial frame. The IMU gimbals and servo-control mechanism function such that this angular velocity is isolated from that of the IMU case. For the fixed base situation, IMU case angular velocity is simply that of the rotating earth. To achieve this isolation, the gimbals must rotate relative to one another in a manner dictated by the following equation:

$$\bar{W}^P_{I,P} = \bar{W}^P_{PG,P} + C^P_{PG}[\bar{W}^{PG}_{RG,PG} + C^{PG}_{RG}(\bar{W}^{RG}_{CASE,RG} + C^{RG}_{CASE}C^{CASE}_{N}\bar{W}^{N}_{I,N})]$$

$\bar{W}^{RG}_{CASE,RG}$ is the angular velocity vector of the roll gimbal frame relative to the IMU case coordinatized in the roll gimbal frame.

$$\bar{W}^{RG}_{CASE,RG} = \begin{bmatrix} \dot{\theta}_{gx} \\ 0 \\ 0 \end{bmatrix}$$

$\bar{W}^{PG}_{RG,PG}$ is the angular velocity vector of the pitch gimbal frame relative to the roll gimbal frame coordinatized in the pitch gimbal frame.

$$\bar{W}^{PG}_{RG,PG} = \begin{bmatrix} 0 \\ \dot{\theta}_{gy} \\ 0 \end{bmatrix}$$

$\bar{W}^P_{PG,P}$ is the angular velocity vector of the platform frame relative to the pitch gimbal frame coordinatized in the platform frame.

$$\bar{W}^P_{PG,P} = \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_{gz} \end{bmatrix}$$

$\dot{\theta}_{gx}$, $\dot{\theta}_{gy}$ and $\dot{\theta}_{gz}$ are the time rates of change of the roll, pitch and azimuth gimbal angles. C^{RG}_{CASE}, C^{PG}_{RG} and C^P_{PG} are the three roll, pitch and azimuth gimbal frame coordinate transformation matrices as defined previously except the trigonometric function arguments are now the θ_{gx}, θ_{gy} and θ_{gz} gimbal angles.

The roll, pitch and azimuth synchros provide a measurement of the corresponding gimbal angles. The absolute accuracy of the KT-70 IMU synchro devices is 6 arc minutes. However, for sufficiently small deviations about any nominal gimbal angle, the synchros are capable of 5 arc second precision. To fully utilize the synchro angle measurement information, each gimbal angle is subdivided into three components:

$$\theta_g(t) = \theta + \Delta\theta + d\theta(t)$$

$\theta_g(t)$ is the total time varing gimbal angle. θ represents the nominal synchro

gimbal angle used in orientating the platform relative to the IMU case (considered fixed after initial platform orientation). $\Delta\theta$ represents a 6 arc minute bias angular error incurred between θ and the actual gimbal orientation due to synchro measurement accuracy limitations. $d\theta(t)$ represents the differential time varying gimbal angle component measurable to 5 arc second accuracy by utilizing the inherent precision of the synchros.

The foregoing system of equations can be manipulated to yield non-linear stochastic differential equations which relate IMU error sources and platform-to-navigation frame misalignment angles to system-level synchro gimbal angle measurable quantities. The non-linear nature of the equations preclude a rigorous solution and for this reason a first order linearization of the equations is assumed to be an accurate approximation. This is a major assumption and will be justified in the sequel. The manipulation and linearization of the system of equations by the use of partial derivatives is beyond the scope of this paper. The simplified linear stochastic differential equation that results from this process is given below for the differential azimuth gimbal angle. Similar equations result for the roll and pitch differential gimbal angles.

$$\dot{d\theta z} = -T(\theta_y)C(\theta_z)[W_{Bx}+W_{Ax}+W_{RSx}+W_{RLx}]$$

$$+T(\theta_y)S(\theta_z)[W_{By}+W_{Ay}+W_{RSy}+W_{RLy}]$$

$$+[W_{Bz}+W_{Az}+W_{RSz}+W_{RLz}]+g_zM_{Sz}-g_yM_{Iz}$$

$$+T(\theta_y)C(\theta_z)[g_xM_{Sx}-g_zM_{Ix}]$$

$$+T(\theta_y)S(\theta_z)[-g_yM_{Sy}+g_zM_{Iy}]$$

$$-T(\theta_y)C(\theta_z)g_yg_zK_{Ox}+g_xg_yK_{Oz}$$

$$+T(\theta_y)S(\theta_z)g_xg_zK_{Oy}-T(\theta_y)C(\theta_z)PRF_x\Delta K_{Fx}$$

$$+T(\theta_y)S(\theta_z)PRF_y\Delta K_{Fy}+PRF_z\Delta K_{Fz}$$

$$+T(\theta_y)C(\theta_z)[W_{Cy}\beta_{xz}-W_{Cz}\beta_{xy}]$$

$$+T(\theta_y)S(\theta_z)[W_{Cx}\beta_{yz}-W_{Cz}\beta_{yx}]$$

$$-W_{Cx}\beta_{zy}+W_{Cy}\beta_{zx}$$

$$+W_{IN}[S(\theta_x)/C(\theta_y)][S(L)\eta_x-C(L)\eta_z]$$

$$+W_{IN}[C(\theta_x)/C(\theta_y)][-C(L)\eta_y+S(L)(\Delta\theta_x+d\theta_x)]$$

$$-\{[C(\theta_z)W_{Cx}-S(\theta_z)W_{Cy}]/C^2(\theta_y)$$

$$-W_{IN}T(\theta_y)C(\theta_x)S(L)/C(\theta_y)\}[\Delta\theta_y+d\theta_y]$$

$$+T(\theta_y)[S(\theta_z)W_{Cx}+C(\theta_z)W_{Cy}][\Delta\theta_z+d\theta_z]$$

$S(\cdot)$, $C(\cdot)$ and $T(\cdot)$ represent the trigonometric sine, cosine and tangent functions.

$$g_x = g[C(\theta_x)S(\theta_y)C(\theta_z)-S(\theta_x)S(\theta_z)]$$

$$g_y = -g[C(\theta_x)S(\theta_y)S(\theta_z)+S(\theta_x)C(\theta_z)]$$

$$g_z = -gC(\theta_x)C(\theta_y)$$

$$W_{Cx} = \Omega\{C(L)C(\theta_z)C(\theta_y)$$

$$-S(L)[C(\theta_z)S(\theta_y)C(\theta_x)-S(\theta_z)S(\theta_x)]\}$$

$$W_{Cy} = \Omega\{-C(L)S(\theta_z)C(\theta_y)+S(L)[S(\theta_z)S(\theta_y)$$

$$C(\theta_x)+C(\theta_z)S(\theta_x)]\}$$

$$W_{Cz} = \Omega[C(L)S(\theta_y)+S(L)C(\theta_y)C(\theta_x)]$$

Derivation of the equations relating system-level IMU velocity measurements to IMU error sources and platform-to-navigation frame misalignment angles follows directly from the specific force equation coordinatized in the accelerometer frame. The navigation frame is considered fixed to the earth.

$$\bar{f}_s^A = \bar{g}^A - (p_N^2 \bar{R}_{NA})^A - (2\bar{\omega}_{IN} \times p_N \bar{R}_{NA})^A$$

$\bar{R}_{NA}$ is the position vector from the navigation frame origin to the accelerometer test mass. p_N is the derivative with respect to time taken relative to the navigation frame.

For the fixed base situation being considered, the last two terms on the right hand side of this equation represent tangential, centripital and coriolis acceleration specific force components produced by a finite accelerometer test mass displacement from the origin of the accelerometer frame and both an angular acceleration and velocity of this frame relative to the earth fixed navigation frame. These components can, however, be shown to be insignificant relative to the inertial instrument random disturbances for the KT-70 IMU. Under these conditions, the accelerometers provide a measure of the earth's gravity vector coordinatized in the accelerometer frame.

$$\bar{f}_s^A = C_P^A C_{PG}^P C_{RG}^{PG} C_{CASE}^{RG} C_N^{CASE} \bar{g}^N$$

C_P^A is the platform-to-accelerometer frame non-orthogonal transformation matrix to account for physical misalignments of the accelerometer sensitive axes from the platform axes.

$$C_P^A = \begin{bmatrix} 1 & \gamma_{xz} & \gamma_{xy} \\ -\gamma_{yz} & 1 & \gamma_{yx} \\ \gamma_{zy} & \gamma_{zx} & 1 \end{bmatrix}$$

In addition to ideally providing a measure of specific force, the accelerometer outputs are corrupted by bias and random noise errors. For the KT-70 IMU, the accelerometer output signals are integrated in capacitive reset integrators which index velocity registers with delta velocity pulses. The accumulated velocity pulses are subsequently multiplied by computer loaded scale factors to obtain total velocity in the appropriate units. If velocity quantization effects are treated as errors in the system-level IMU velocity measurements, then the accelerometer output signals can be used to define an artificial time derivative velocity vector modeled as follows:

$$\dot{\bar{V}} = [K_A]\{\bar{f}_s^A + \bar{A}_B + \bar{a}_R\}$$

$[K_A]$ is a 3×3 diagonal matrix of the loaded accelerometer scale factors. The true scale factors $[K_{AT}]$ differ from the loaded ones by an error $[\Delta K_A]$.

$\bar{A}_B$ is the accelerometer bias vector.

$\bar{a}_R$ is the exponentially correlated accelerometer random noise vector.

$$\dot{\bar{a}}_R(t) = \frac{-\bar{a}_R(t)}{\tau} + \sigma\sqrt{\frac{2}{\tau}}\,\bar{N}(t)$$

In contrast to the gimbal angles time rates of change, the accelerometer outputs do not have an error free value of zero. A zero mean accelerometer error vector equation can, however, be defined as follows:

$$\dot{\Delta\bar{V}} = \dot{\bar{V}} - C_{PG}^P C_{RG}^{PG} C_{CASE}^{RG} \bar{g}^N$$

All terms are as previously defined with the gimbal frame coordinate transformations being computed as functions of the initial θ_x, θ_y and θ_z synchro angle settings.

The above system of equations can be manipulated and linearized to yield linear stochastic differential equations for the accelerometer output signals. The resultant error equation for the x-axis accelerometer is given below for illustration.

$$\dot{\Delta V}_x = -g[C(\theta_z)S(\theta_y)S(\theta_x)+S(\theta_z)C(\theta_x)]\eta_x + g[C(\theta_z)C(\theta_y)]\eta_y$$

$$-g[S(\theta_z)C(\theta_x)+C(\theta_z)S(\theta_y)S(\theta_x)][\Delta\theta_x+d\theta_x]$$

$$+gC(\theta_z)C(\theta_y)C(\theta_x)[\Delta\theta_y+d\theta_y]$$

$$-g[S(\theta_z)S(\theta_y)C(\theta_x)+C(\theta_z)S(\theta_x)][\Delta\theta_z+d\theta_z]$$

$$+g[C(\theta_z)S(\theta_y)C(\theta_x)-S(\theta_z)S(\theta_x)]\Delta K_{Ax}$$

$$-g[S(\theta_z)S(\theta_y)C(\theta_x)+C(\theta_z)S(\theta_x)]\gamma_{xz}$$

$$-gC(\theta_y)C(\theta_x)\gamma_{xy}+A_{Bx}+a_{Rx}$$

IV. Separation of IMU Error Sources

The preceding derivation has resulted in six equations relating the six IMU measurable quantities to roughly fifty error sources. Since a set of six simultaneous stochastic differential equations cannot be solved for fifty unknowns, a means of functionally separating the error quantities had to be formulated. In view of the stochastic nature of the equations and IMU measurement noise, stochastic estimation of the error quantities is the best solution possible.

The conceptual approach to error source separation is the generation of new equations by angularly orientating the platform at various attitudes relative to the earth rate and gravity vectors. This approach follows directly from the observation that the measurable IMU parameters are functions of both the IMU error sources and the coordinatization of the above vectors in the platform and inertial instrument reference frames.

In developing this approach, it is convenient to group the error quantities into three groups. The first group consists of the time independent IMU error sources which are modeled as random biases and commonly denoted as calibration parameters. The second group consists of those angular misalignment errors associated with platform-to-navigation frame alignment. The third group consists of the stochastic process modeling the inertial instrument random disturbances.

If the error sources associated with the second and third group are omitted and IMU measurement noise neglected, then, the set of six stochastic differential equations can be seen to reduce to a system of ordinary differential equations. This system of ordinary differential equations can be indefinitely expanded by positioning the platform at new attitudes. A deterministic solution to the system of equations can be obtained by application of Laplace transforms and utilizing the initial and final IMU measurements for each attitude. Table 1 illustrates how the IMU calibration parameters become separable as a function of the cumulative number of attitudes utilized. The specific attitudes shown were determined by a combination judgement and trial-and-error process.

The first order effect of introducing the platform-to-navigation frame alignment errors of the second group into the solution for the separation of the IMU calibration parameters of the first group can be understood with reference to Figure 2. For a small platform-to-navigation frame misalignment, the effect is precisely the classical problem of separating west gyro drift from the west platform axis commanded angular velocity error due to the azimuth misalignment angle ψ. In order to null the angular velocity of an ideal IMU platform with respect to the navigation frame, the gyros must be commanded to precess at exactly the magnitude of the earth rate vector coordinatized in the gyro frame. Since ψ is a zero mean random variable whose magnitude is not known a priori, the west platform axis would not be commanded to precess. For this condition, the angular velocity of the west platform axis with respect to the navigation axis is approximately given by:

$$W_{N,P_y} \simeq W_{D_y} + W_{IE}C(L)\psi$$

TABLE 1

Separation of IMU Calibration Parameters

	Cumulative Platform Attitudes						
Nominal Gimbal Angle	1	2	3	4	5	6	7
θ_x-Roll	0	π	$\pi/2$	$\pi/2$	$\pi/2$	$\pi/2$	$\pi/4$
θ_y-Pitch	0	0	0	0	0	$\pi/4$	$-\pi/4$
θ_z-Azimuth	0	0	0	π	$\pi/2$	0	0
Calibration Parameter							
A_{Bx}		X					
A_{By}		X					
A_{Bz}		X					
ΔK_{Ax}					X		
ΔK_{Ay}			X				
ΔK_{Az}		X					
γ_{xz}			X				
γ_{xy}		X					
γ_{yz}					X		
γ_{yx}		X					
γ_{zy}					X		
γ_{zx}			X				
W_{Bx}				X			
W_{By}				X			
W_{Bz}				X			
M_{Sx}					X		
M_{Ix}						X	
M_{Sy}					X		
M_{Iy}						X	
M_{Sz}						X	
M_{Iz}					X		
ΔK_{Fx}				X			
ΔK_{Fy}					X		
ΔK_{Fz}						X	
β_{xz}				X			
β_{xy}						X	
β_{yz}				X			
β_{yx}						X	
β_{zy}				X			
β_{zx}					X		
K_{Ox}							X
K_{Oy}							X
K_{Oz}							X

Given the standard deviation of the azimuth misalignment angle σ_ψ and a perfect measurement of W_{N,P_y}, the standard deviation σ_W of the W_{Dy} estimation error is given as:

$$\sigma_W = \sigma_\psi W_{IE} C(L)$$

This inability to separate the west platform commanded angular velocity error from the composite west gyro drift does not permit separation of all the IMU calibration errors as predicted in Table 1. The error source separation problem can, however, be overcome by noting that the north and up platform axes do not encounter this first order limitation due to a different commanded angular velocity attitude dependency. The solution is to add additional platform attitudes until all IMU calibration parameters become separable. It was found that by adding three new attitudes, similar to the first three in Table 1 except with the x and y platform axes interchanged, the calibration parameters became separable.

Consideration of IMU measurement noise and inertial instrument random disturbances of the third group of IMU error sources changes the problem description from one of ordinary differential equations and a deterministic solution to stochastic differential equations requiring some form of stochastic estimation if optimal results are to be obtained. Separation of some of the IMU calibration parameters from the random disturbances will have to be made on the basis of time rather than platform attitude since it is the only difference in functional dependency exhibited.

V. Kalman Filter

Selection of the technique for estimating the IMU error sources was made in consideration of three major criteria. (1) The technique should be

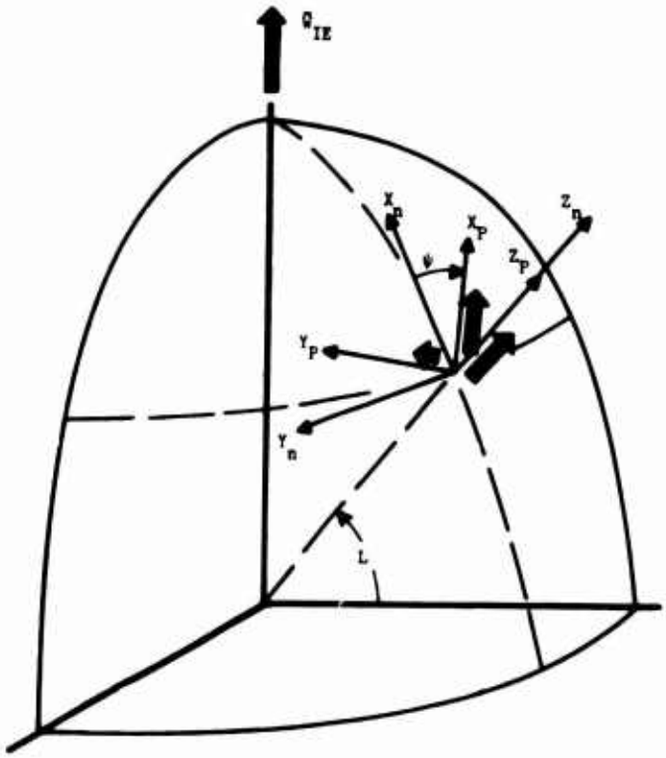


FIGURE (2) PLATFORM AZIMUTH MISALIGNMENT GEOMETRY

optimal in the minimum variance sense. (2) The technique should provide a statistical measure of the accuracy to which the IMU error sources can be estimated. And (3) the technique should provide a convenient means of validating the first order linearization and other simplifying assumptions. While many optimal stochastic estimation techniques are available, the Kalman filter algorithm with it's associated covariance matrix was considered to be the most applicable approach. Conversely, the formulated problem satisfies the linearity, a priori system dynamics and Gauss-Markov stochastic process modeling restrictions under which the Kalman filter algorithm was derived.

The specific Kalman filter equations mechanized correspond to the continuous system-discrete measurement case and are given in reference 2. Due to the rather large dimensionality, the various matrices could not be pictorially illustrated here. Formulation of the developed equations in state variable vector and matrix notation is, however, a straight-forward procedure. The resultant 57 element state vector is mechanized as:

$X^T = [W_{Bx}, W_{By}, W_{Bz},$
$W_{Ax}, W_{Ay}, W_{Az},$
$W_{RSx}, W_{RSy}, W_{RSz},$ (Disturbances)
$W_{RLx}, W_{RLy}, W_{RLz},$ (Disturbances)
$M_{Sx}, M_{Sy}, M_{Sz},$
$M_{Ix}, M_{Iy}, M_{Iz},$
$K_{Ox}, K_{Oy}, K_{Oz},$
$\Delta K_{Fx}, \Delta K_{Fy}, \Delta K_{Fz},$
$\beta_{xy}, \beta_{xz}, \beta_{yx},$
$\beta_{yz}, \beta_{zx}, \beta_{zy}$
$\eta_x, \eta_y, \eta_z,$
$\Delta\theta_x, \Delta\theta_y, \Delta\theta_z,$
$d\theta_x, d\theta_y, d\theta_z,$ (Measurements)
$A_{Bx}, A_{By}, A_{Bz},$
$\Delta K_{Ax}, \Delta K_{Ay}, \Delta K_{Az},$
$a_{Rx}, a_{Ry}, a_{Rz},$ (Disturbances)
$\gamma_{xy}, \gamma_{xz}, \gamma_{yx},$
$\gamma_{yz}, \gamma_{zx}, \gamma_{zy},$
$\Delta V_x, \Delta V_y, \Delta V_z]$ (Measurements)

The six measured states represent error quantities as computed by subtracting the nominal values from the absolute system-level IMU velocity and synchro gimbal angle measurements. The great majority of the non-zero 57 x 57 system matrix (F) elements are located in the F matrix rows corresponding to these measurement states. Specific F matrix elements are readily identified from the previous stochastic differential equations. The nine disturbance states correspond to those of the inertial instruments. The F matrix contains the negative reciprocals of the correlation times as diagonal elements in the corresponding rows. The remaining state vector elements are considered as IMU calibration parameters with the exception of the three states modeling IMU case-to-navigation frame misalignment angles. These remaining states are modeled as random biases and do not contribute any non-zero F matrix elements. The F matrix elements are independent of time but do change in magnitude each time the platform frame is re-orientated due to the nominal synchro angle dependency.

The 57 x 9 system noise distribution matrix (G) contains nine non-zero elements located in the rows corresponding to the disturbance states. These non-zero elements are readily identified from the nine stochastic differential equations modeling the inertial instrument disturbances. The 57 x 6 measurement matrix (H) contains six non-zero elements of unity magnitude in the columns corresponding to the rows of the measurement states. The measurement (z) and measurement error (v) vectors contain six states. The first three elements of the v vector are zero mean Gaussian-White random

variables with 5 arc second standard deviations and model the differential synchro gimbal angle measurement noise. The last three elements have similar properties and a standard deviation representative of that of the accelerometer quantization process and model the artifical velocity measurement noise.

The 57 x 57 covariance matrix (P) is initially a diagonal matrix composed largely of the IMU error source variances as given in reference 1. The initial variances of the measurement states are identically zero due to the manner in which these states were defined. The initial variances of the IMU case-to-navigation frame misalignment angles depend on the laboratory test table used and the manner in which the alignment is performed. A partial P matrix re-initialization is required each time the platform frame is re-orientated due to the platform-to-IMU case attitude dependent random variables and the measurement states.

In addition to those errors sources included in the Kalman filter state vector, it is desirable to compute optimal estimates and estimation errors for several composite error quantities. Inertial navigation system parameters such as total gyro and platform drift and platform and accelerometer-to-navigation frame misalignment angles provide additional insight to the IMU calibration and alignment methodology's capability to initialize an IMU for the subsequent navigation function. Optimal estimates for these system parameters are computed as functions of the sum of appropriate state vector element's optimal estimates. Optimal estimation errors for these composite error quantities are computed by use of the expectation operator and appropriate elements of the covariance matrix.

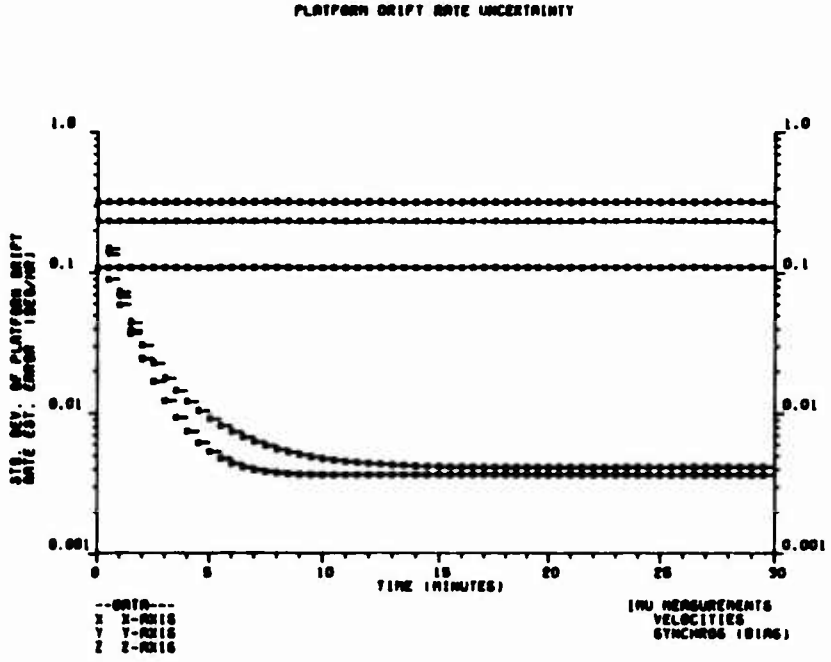
VI. Simulation Results

A CDC-6600 computer simulation of the KT-70 IMU, Kalman filter and overall calibration procedure has been developed in consideration of three major objectives: (1) conceptual validation of the calibration and alignment methodology. (2) Analytic verification that the stated performance goal can be achieved. And (3) preparation of a computer software specification and laboratory test procedure for the ensuing engineering development program.

In order to obtain the highest level of confidence possible by the use of computer simulations, the concept validation philosophy of demonstrating strict convergence of the Kalman filter optimal estimates to the error sources of a simulated KT-70 IMU was adopted. Strict sense convergence implies that the ensemble average of the simulated estimation errors conform to the zero mean Gaussian processes having time dependent variances given by the filter's covariance matrix.

In keeping with the concept validation philosophy, the developed KT-70 IMU simulation contains all error sources given in reference 1 with the exception of the inertial instrument warm-up transients. A warm-up period of fifteen minutes is sufficient for these transients to decay to insignificant levels and is assumed to have preceded the start of the calibration test. The IMU simulation is based on the whole number, non-linear equations developed above and utilizes random number generators to simulate the random bias and stochastic process error sources. Digital gyro torquing and accelerometer pulse quantization effects are included. The adequacy to which the reference 1 error sources model the real KT-70 IMUs will be addressed in the engineering development program.

Figure 3 illustrates the filter's capability to estimate IMU platform drift and artificial velocity errors in terms of the standard deviation of the estimation error. Platform drift is a composite error source containing up to sixteen error sources per axis for the platform and navigation frames



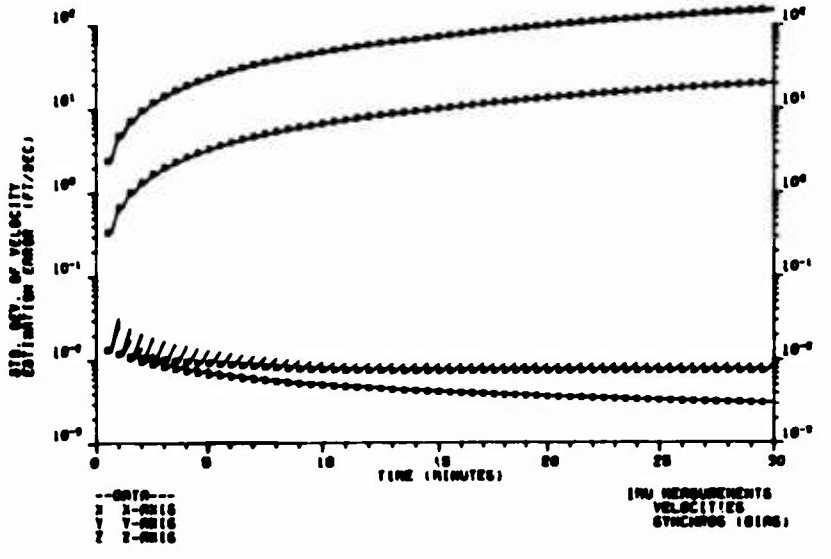


FIGURE (3) PLATFORM DRIFT AND VELOCITY UNCERTAINTY

approximately aligned situation shown. Similar remarks apply to the velocity errors. Estimation error standard deviations for these parameters are computed separately from the filter calculations but utilize covariance matrix elements as data.

Each graph contains two sets of data, corresponding to the absence and presence of IMU measurements. The symbols indicate the time instants the six simultaneous IMU measurements are taken and are uniformly spaced at thirty second intervals. The plots between the measurement points indicate how the estimation uncertainty propagates in time. The primary information conveyed by these graphs is that the filter is capable of estimating composite system-level IMU performance errors to an accuracy limited by the instrument random disturbances in a five to ten minute IMU test. The length of the test required to achieve this level of accuracy is heavily influenced by the measurement errors and the magnitude of the disturbances.

While the filter can accurately estimate the composite error sources from a single platform orientation test, it cannot separate or observe the individual IMU error sources. The filter has, however, heavily correlated certain groups of error sources from an initially uncorrelated condition. To calibrate the IMU, the error sources must be decorrelated and accurately observed. The means by which this is accomplished is by the use of numerous platform orientations as discussed previously.

Figure 4 is a graphic illustration of the separation process. The upper figure shows an increasing filter capability to estimate the y gyro bias drift component as additional platform orientations are introduced. Platform attitudes #1 thru #5 and #10 and #11 correspond to those of Table 1. Attitudes #7 thru #9 are the additional attitudes introduced to permit error source separation in the presence of the west gyro commanded angular velocity error due to the azimuth platform-to-navigation frame misalignment. Attitudes #6 and #12 are identical to #1 and illustrate that repetition of the same attitude does not improve error source observability. These attitudes are synomous with the navigation frame axes of north, west and up.

The IMU is tested for five minutes per attitude as suggested by the estimation accuracy versus measurement time characteristic of Figure 3. Total IMU test time would be somewhat longer than the one hour measurement time due to the one to two minute platform slew time required to periodically re-orientate the platform. The sequence of platform attitudes shown was selected such that the most significant IMU error sources are calibrated first and allow the option of omitting latter test attitudes in favor of reducing test time.

The upper and lower bounds superimposed about the simulated y gyro bias represent three standard deviations of the optimal estimation error stochastic process. For a Gaussian process, three standard deviations imply that, on an ensemble average basis, over 99% of the optimal estimates will fall between these bounds. Since all 120 optimal estimates are seen to fall within the three standard deviation bounds, the criterion is met. Satisfaction of this criterion is considered as justification for the first order equation linearizations, the simplifying assumptions and use of the Kalman filter as the optimal stochastic estimation technique.

The lower graph in Figure 4 shows the optimal estimation errors for all three gyro bias drift components. This representation contains the same estimation accuracy information as the upper figure but is plotted as one standard deviation on an absolute scale. This data is obtained by taking the square root of the appropriate diagonal element of the updated covariance matrix.

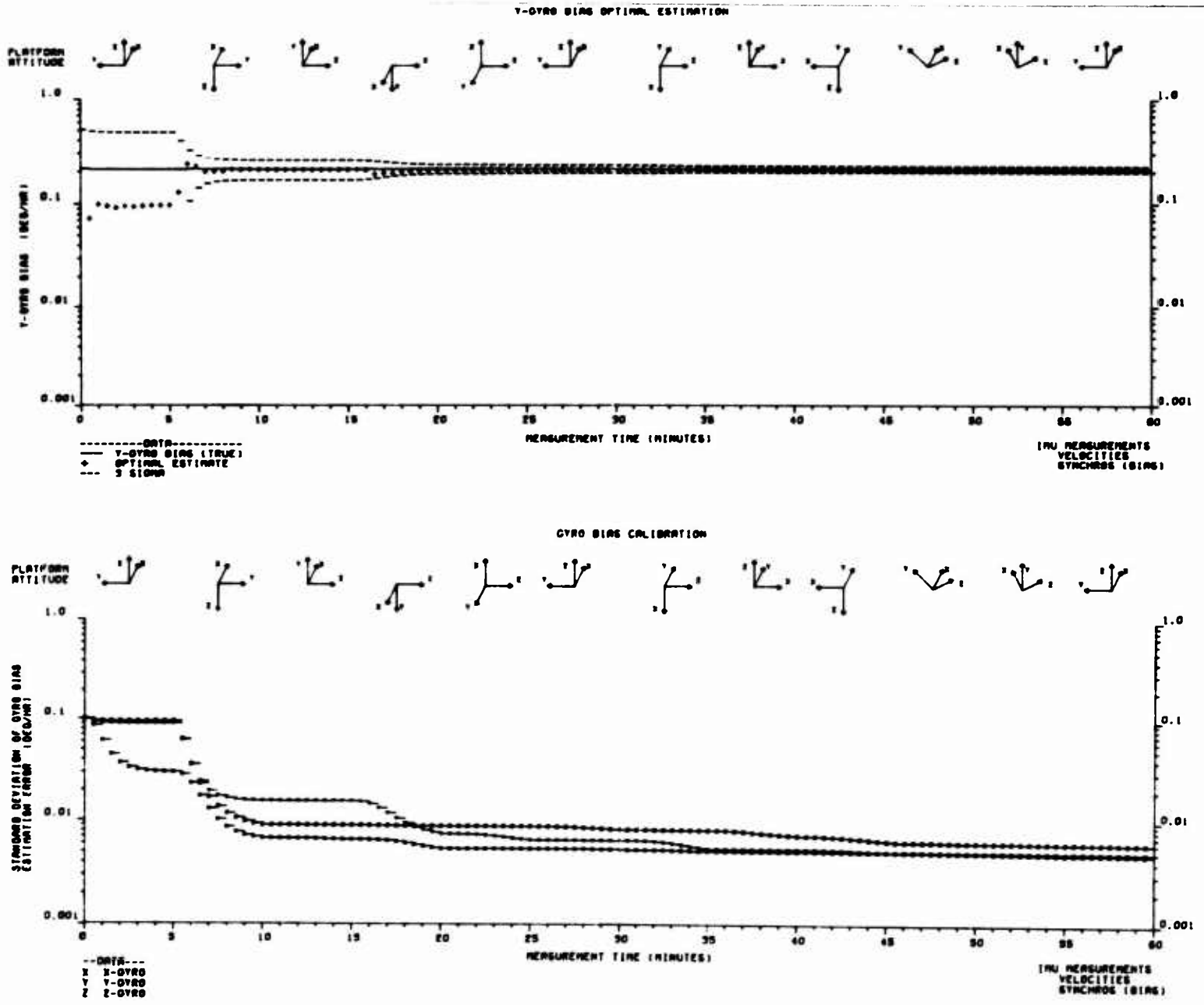


FIGURE (4) GYRO BIAS CALIBRATION

The upper graph in Figure 5 corresponds to the lower one in Figure 4 except the synchro gimbal angle measurements are not utilized. After the one hour test, level axis gyro bias calibration accuracy is about the same as before but the vertical gyro bias drift component has not been completely separated from other IMU error sources. The implication of this graph is that system-level IMU calibration appears possible by utilizing only velocity measurements. Additional platform orientations will, however, be necessary for separation of the error sources.

The lower graph in Figure 5 shows the estimation accuracy for the random variable defined as the sum of the gyro bias and long correlation time random gyro drift disturbance components. The purpose of this graph is to illustrate the IMU calibration accuracy limitations imposed by the gyro disturbances. The two exponentially correlated disturbances per gyro have 1 and 3600 second correlation times and 0.003 and 0.005 degrees per hour rates respectively. Reference to the lower graph in Figure 4 shows that the long correlation time disturbance imposes an approximate lower bound on gyro bias calibration accuracy. The reason is that time is the only difference in the functional dependency of these two error sources and accurate measurements taken over roughly a correlation time period are required to separate them. This limitation is illustrated in the lower graph of Figure 5 by the ability of the filter to estimate the sum of these two error sources to an accuracy now limited by the short correlation time disturbance.

Gyro spin axis mass unbalance drift calibration is shown in Figure 6. The mass unbalance drift is a function of both the magnitude and the orientation of the gyro relative to the specific force vector and drifts higher than those indicated will be experienced during maneuvering flight conditions. The mass unbalance drift rather than the mass unbalance coefficient estimation error is shown to emphasize the relative significance of this error source and the calibration accuracy limitations imposed by the gyro disturbances. It should be noted that the initial uncertainty corresponds to an operational system and not to that of a totally uncalibrated or raw IMU.

The lower graph in Figure 6 illustrates the well known result that platform drift about the vertical axis produces only higher order effects in the level axes accelerometer outputs. To calibrate the spin axis mass unbalance drift by use of only velocity measurements, additional platform orientations similar to #10 would have to be used such that the mass unbalance drift produces first order effects in the accelerometer outputs.

The filter's capability to estimate total gyro drift is shown in Figure 7. Total gyro drift has been defined as the sum of the disturbances, constant and attitude dependent bias, input and spin axis mass unbalance, and anisoelastic drift components. The simulated x gyro drift is plotted with straight line segments connecting the 30 second intervals and does not show the fine grain structure due to the disturbances. Filter convergence is illustrated for one sigma upper and lower bounds which, implies that 68% of the optimal estimates fall within these boundaries. Optimal estimation errors appear consistent with this criterion.

In certain military aided navigation situations, it is conceptually desirable to accomplish continuous in-flight IMU alignment and calibration to reduce weapon system reaction time and enhance self-contained navigation accuracy in the event the aids are lost. Continuous in-flight IMU calibration is also desirable since it provides a means for IMU fault detection and performance monitoring. The conceptual requirements for accomplishing these functions are shown by the total gyro calibration

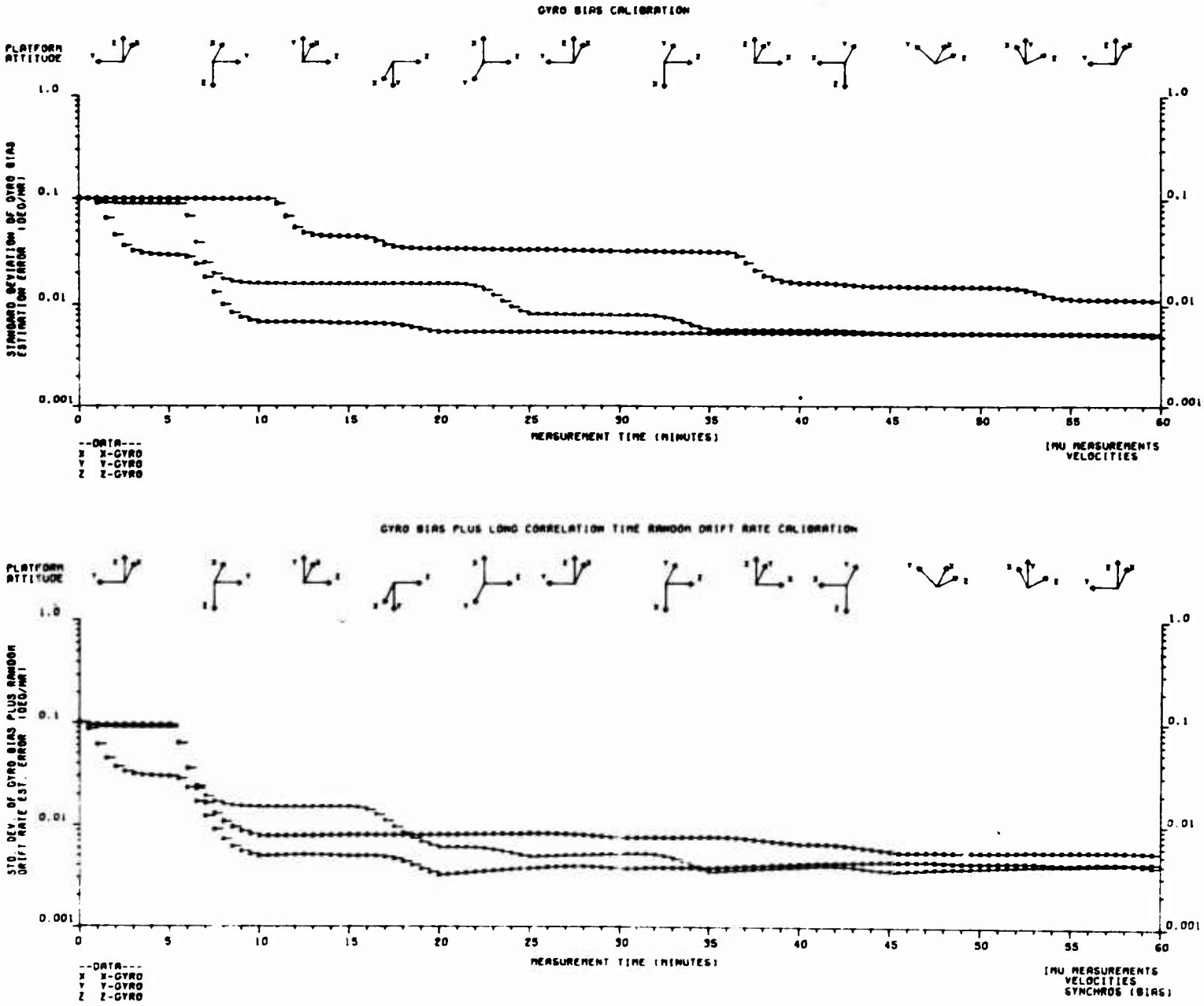


FIGURE (5) GYRO BIAS PLUS LONG CORRELATION TIME RANDOM DRIFT CALIBRATION

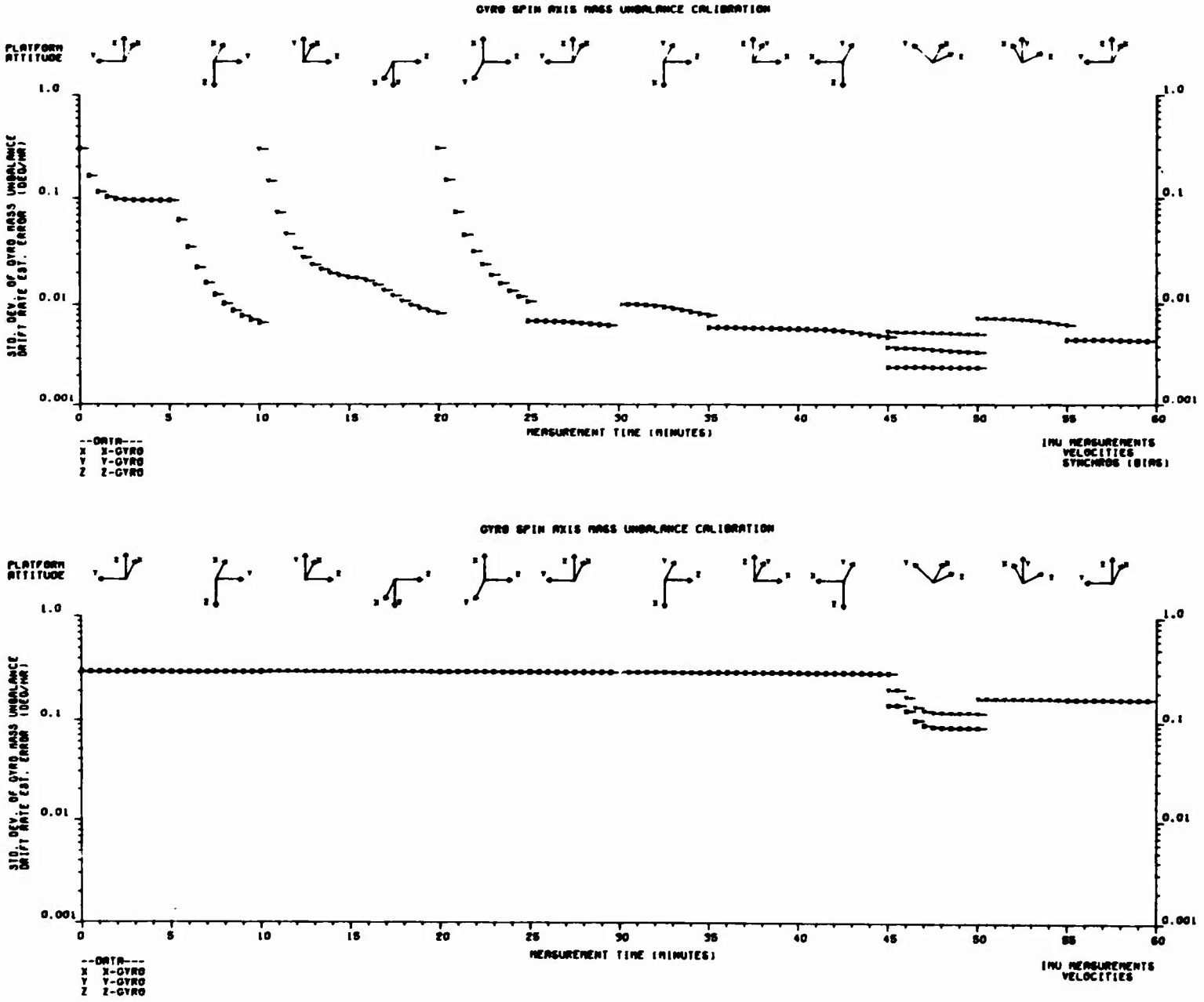


FIGURE (6) GYRO SPIN AXIS MASS UNBALANCE CALIBRATION

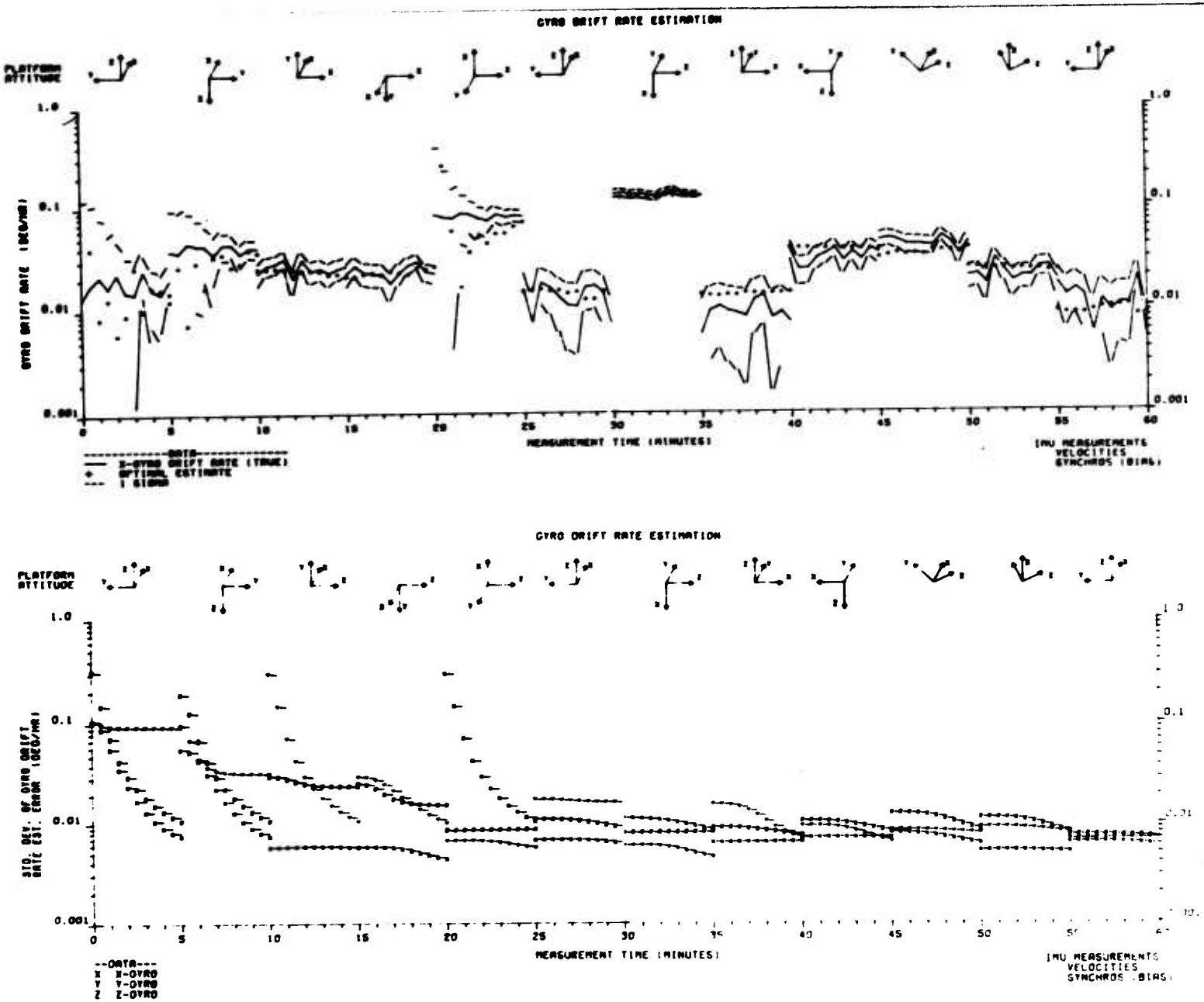


FIGURE (7) GYRO DRIFT ESTIMATION

characteristics of Figure 7. Accurate external measurements of vehicle velocity or position by OMEGA, LORAN or satellite navigation systems would be essential. The other essential requirement would be to periodically alter the relative orientation of the gyro frame with respect to the specific force vector. In-flight, these re-orientations could conceively be accomplished by aircraft maneuvers rather than changing the platform's attitude.

Gyro input-to-platform axes misalignment angle calibration accuracy data is shown in Figure 8. For the KT-70 IMU, these six error sources do not produce a significant platform drift relative to the gyro disturbances which precludes the filter from reducing the uncertainty associated with them. The lower graph of this figure was obtained by omitting the synchro bias errors and effectively results in very accurate platform attitude measurements. Attitude measurements of this accuracy are highly improbable with synchros but analytically demonstrate that absolute calibration accuracy is largely a function of the disturbances rather than the measurement accuracy.

An IMU cost-effectiveness consideration can be noted in relation to Figure 8. Due to the relatively insignificant effect of the manufacturing related gyro misalignment error sources, the production specifications for the hardware could be relaxed resulting in lower production costs. Specification relaxization to the point where the misalignment angles become significant error sources is also permissable since the IMU calibration error sources can be software compensated for in the navigation computer. Considering all manufacturing related IMU error sources, a cost-effective navigation system design could be obtained by trading-off the IMU hardware production costs against delta computer and software compensation costs.

The inherent fault detection and fault isolation capability of the IMU error analysis methodology is illustrated in Figure 9. The upper graph shows that the nominal computer loaded gyro scale factor errors do not contribute significant platform drift components. The lower graph shows the filter's performance against a simulated IMU containing a x gyro scale factor error an order-of-magnitude larger than assumed in the initial covariance matrix.

The optimal estimates for two simulation runs are plotted in the lower graph. The plus symbols correspond to the first run and an initial covariance matrix based on the nominal error data. The filter begins to estimate an abnormally large error, optimal estimates compared with the estimation error standard deviation, during the forth platform attitude test as predicted by the error source separation data of Table 1. The optimal estimates do not, however, strictly converge to the simulated error in accordance with the three sigma bounds. This divergence is due to the non-conformance of the initial covariance matrix with the simulated IMU. Having accomplished fault detection, successful filter fault isolation requires that the optimal estimates for all the other error sources converge to those simulated. This convergence did indeed occur during the simulation runs.

The second simulation run was made to verify that filter convergence could be obtained by increasing the initial covariance matrix in accordance with the x gyro scale factor optimal estimates obtained from the first run. This capability is desirable since it gives the user the option of software compensating the error rather than repairing the IMU. The zero symbols represent second run optimal estimates and show that final convergence is consistent with the estimation uncertainty obtained from the first run. This result is another indication that calibration accuracy is essentially determined by the disturbances and insensitive to the initial covariance matrix. The two step process of accurately estimating abnormally large errors can idealy be reduced to a single

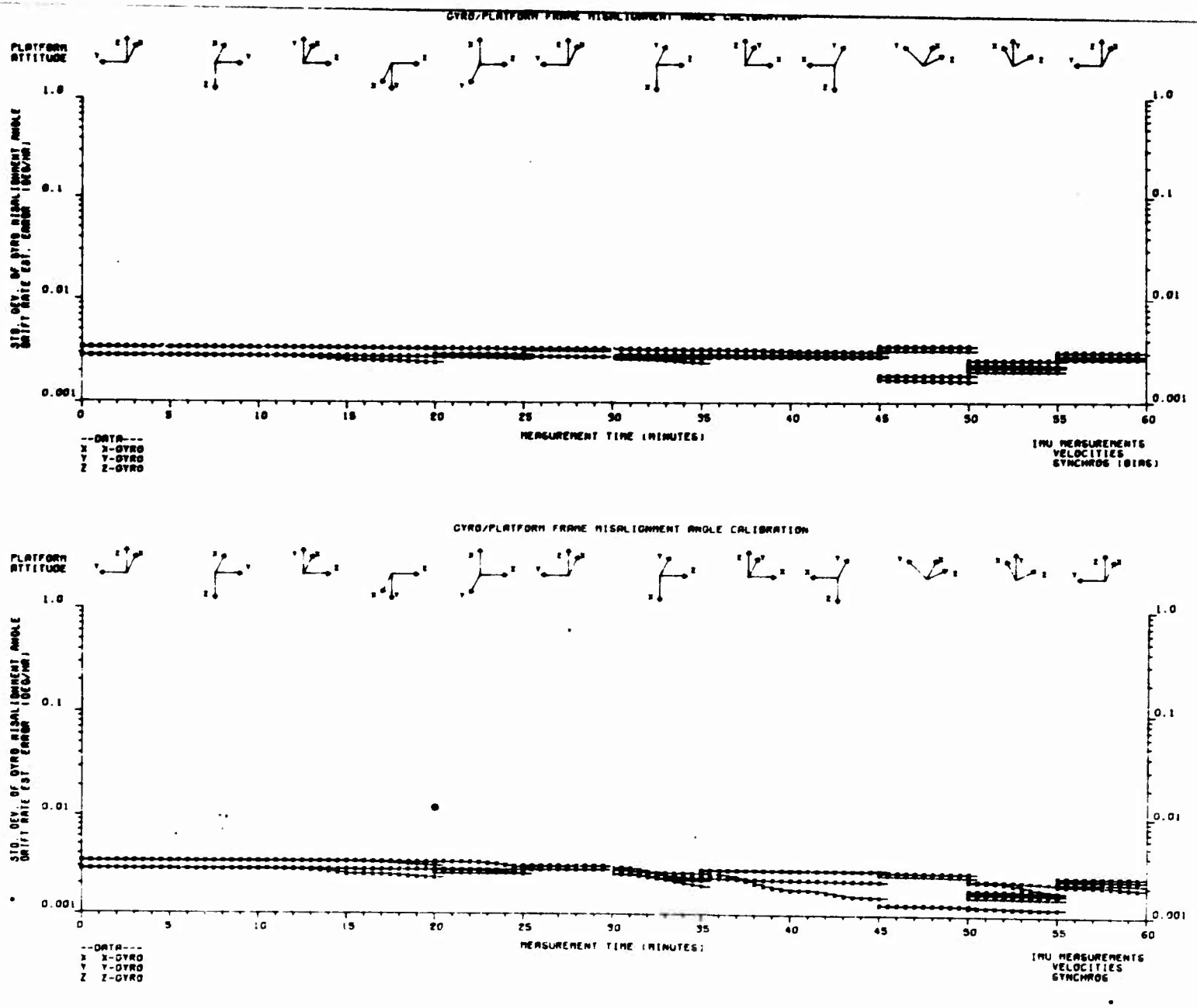
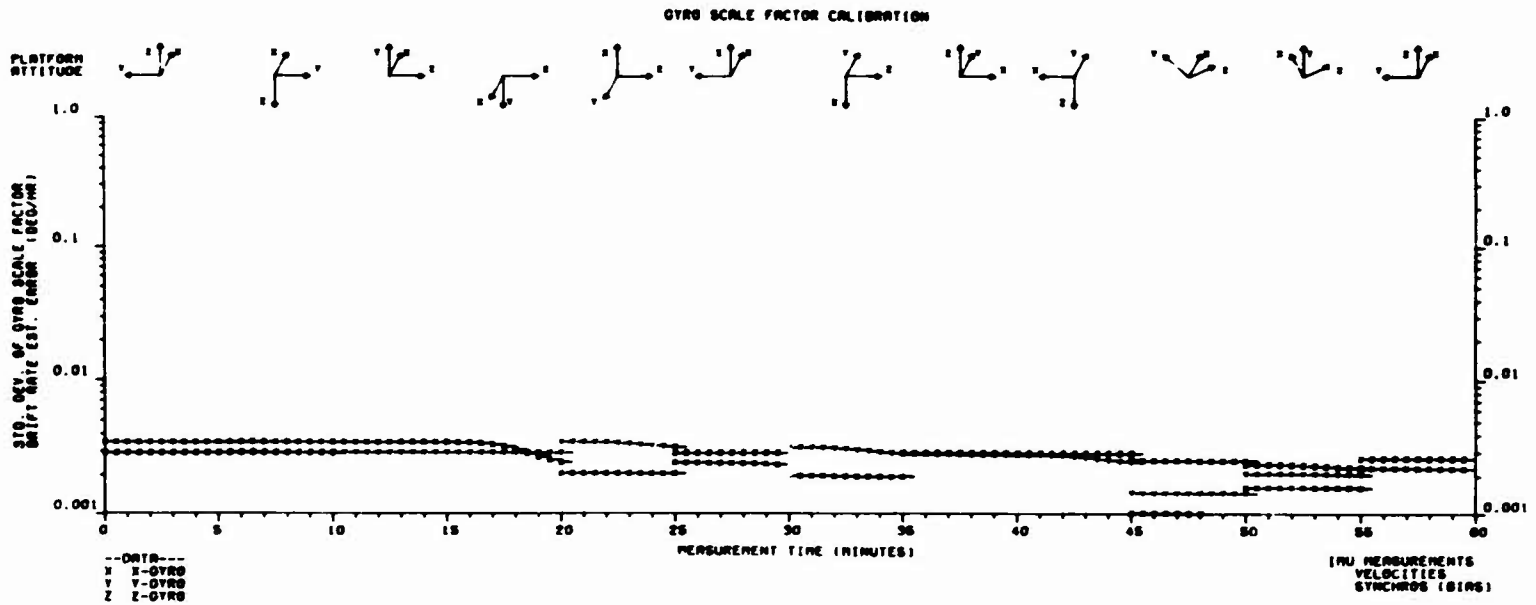
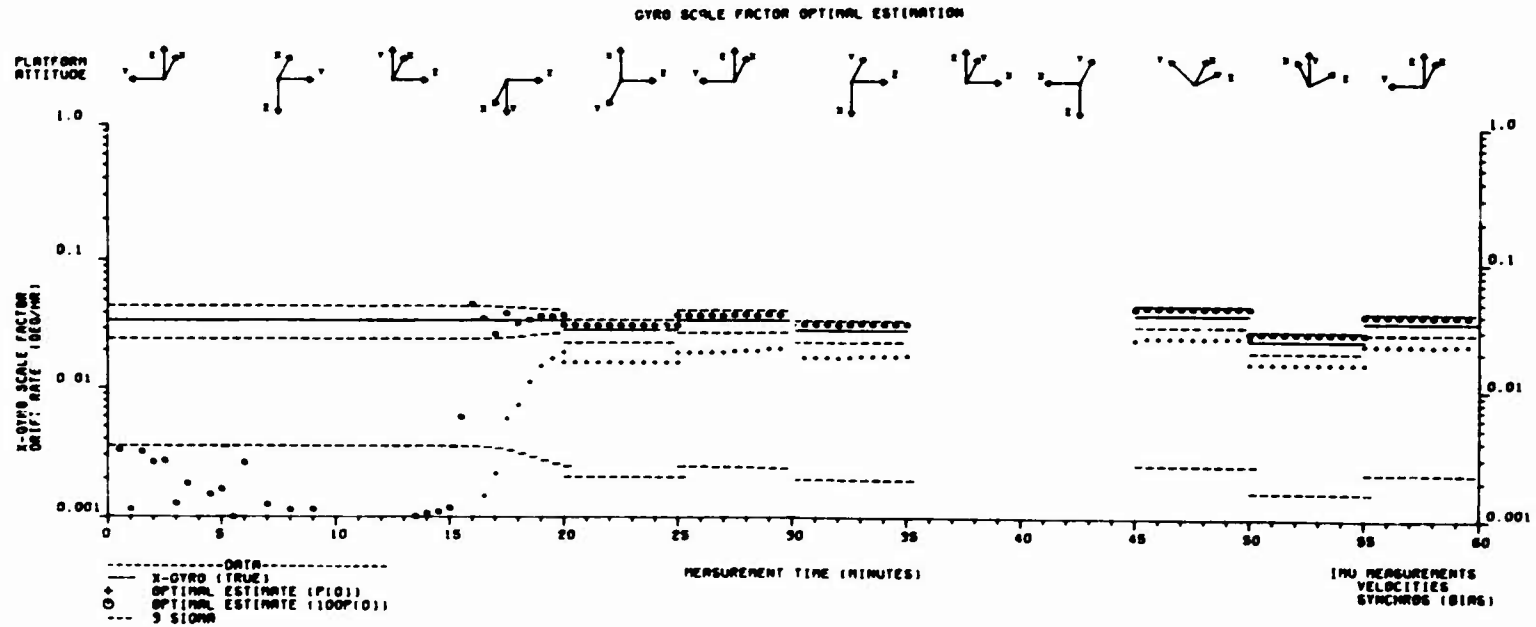


FIGURE (8) GYRO-TO-PLATFORM FRAME MISALIGNMENT ANGLE CALIBRATION


GYRO SCALE FACTOR CALIBRATION
PLATFORM ATTITUDE
1.0
0.1
0.01
0.001
0
5
10
15
20
25
30
35
40
45
50
55
60
MEASUREMENT TIME (MINUTES)
--DATA--
X X-GYRO
Y Y-GYRO
Z Z-GYRO
IMU MEASUREMENTS
VELOCITIES
GYROBIAS (BIAS)


GYRO SCALE FACTOR OPTIMAL ESTIMATION
PLATFORM ATTITUDE
1.0
0.1
0.01
0.001
0
5
10
15
20
25
30
35
40
45
50
55
60
MEASUREMENT TIME (MINUTES)
--------DATA--------
X-GYRO (TRUE)
OPTIMAL ESTIMATE (P(0))
OPTIMAL ESTIMATE (100P(0))
3 SIGMA
IMU MEASUREMENTS
VELOCITIES
GYROBIAS (BIAS)

test by increasing the initial covariance matrix in accordance with the expected IMU fault magnitudes.

Figure 10 provides an illustration of the process required to separate west or y axis platform drift due to a commanded angular velocity error from all the other error sources as discussed under separation of IMU error sources. Platform gyro dependent drift contains all platform drift compenents, shown in Figure 3, except those originating from the platform-to-navigation frame misalignment angles.

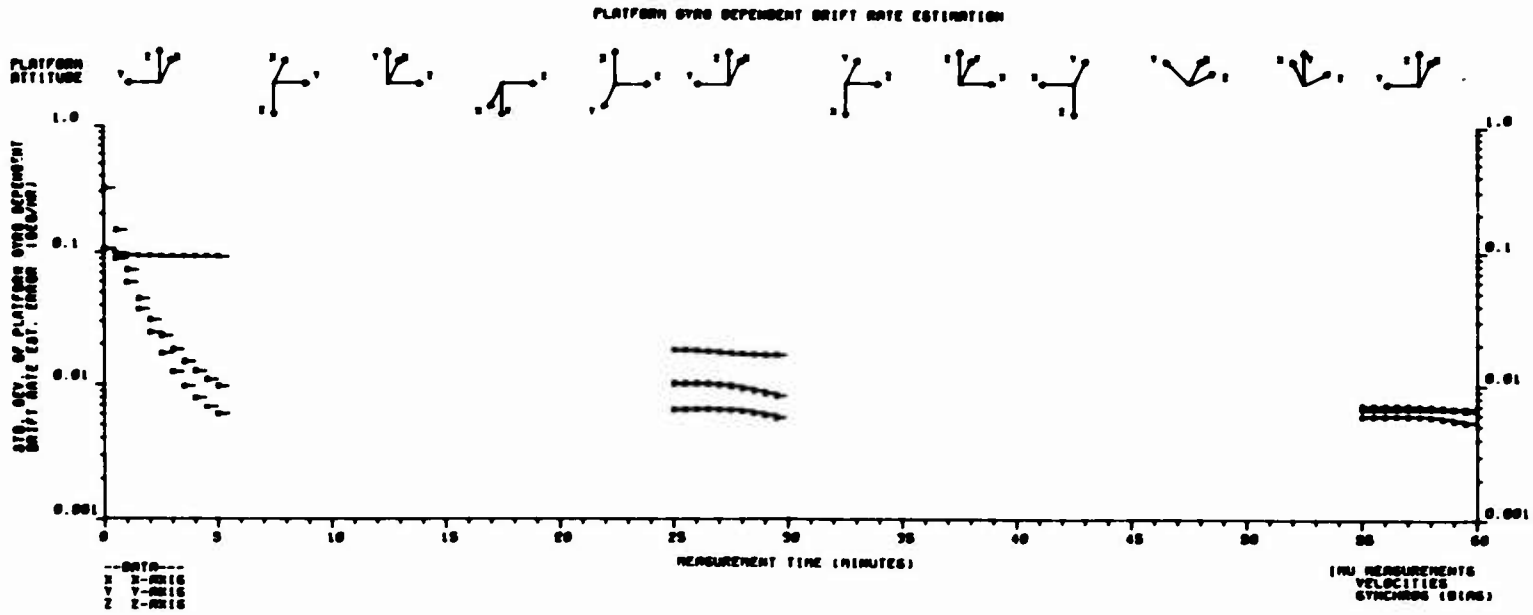
Accelerometer bias and scale factor calibration accuracy data is shown in Figure 11. Separation of these error sources occurs after an accelerometer has been orientated up and down in navigation frame coordinates. Separation occurs in this sequence, rather than that predicted by the group 1 errors in Table 1, because the group 2 platform-to-navigation frame angle bias errors do not produce a first order effect in the vertical accelerometer's output.

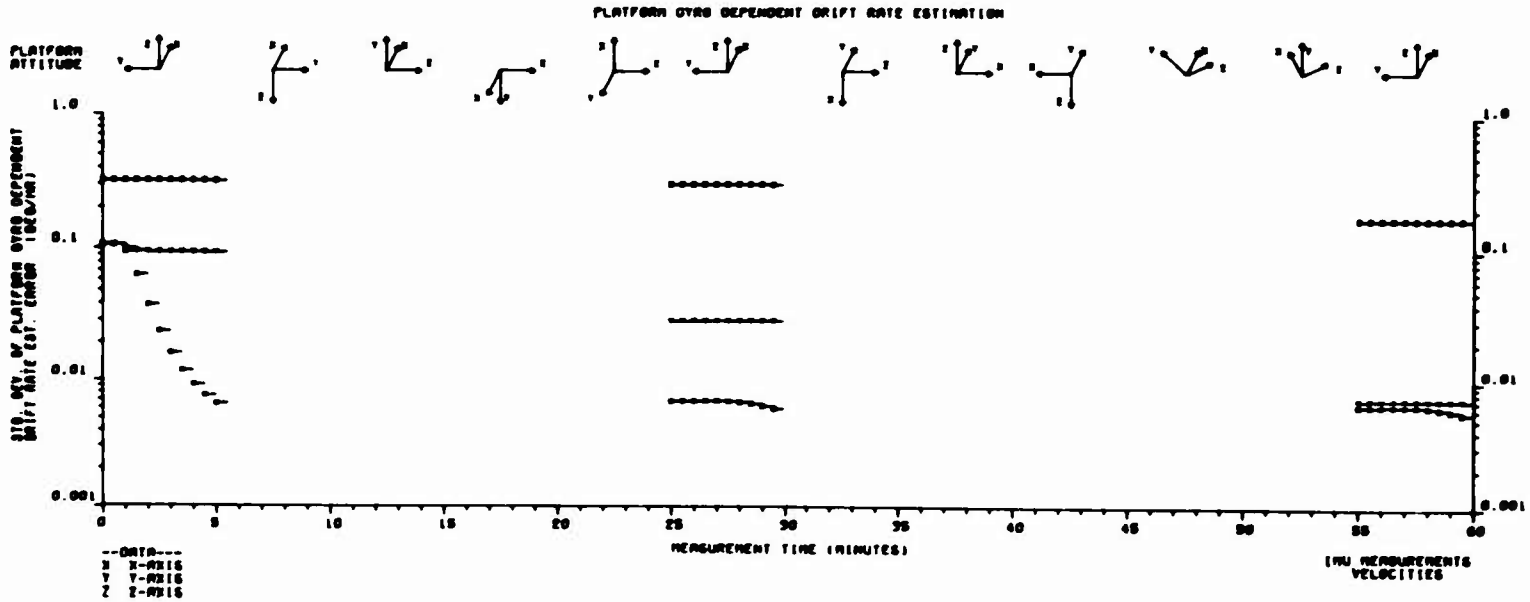
Failure of the linearized vertical accelerometer output equations to exhibit any first order angular dependency results in marginal filter convergence for the bias and scale factor error sources separated by this effect. The amount of filter divergence is a function of the accelerometer's sensitive axis angular tilt from the gravity vector. Simulation results indicate that the bias errors can be estimated to roughly one micro-g, rather than a fraction of a micro-g indicated by Figure 11, for the anticipated calibration tests. While techniques such as closed loop platform control, extended Kalman filter algorithm and inclusion of second order effects in the propagation of the optimal estimates can be utilized to improve filter convergence, these techniques complicate the data reduction process and are considered unnecessary in view of the above accuracy.

Calibration accuracy of the two small angular misalignments of the x accelerometer's sensitive axis from the platform frame is shown in Figure 12. To achieve accurate navigation, the major criterion is alignment of the accelerometer's electrical null axis to the navigation axis. The filter can estimate these composite misalignment angles with a high degree of accuracy, as will be shown in the sequel, but cannot effectively separate the components of these composite angles. Introduction of new synchro gimbal angle bias errors each time the platform is re-orientated precludes this separation process. The lower graph in this figure illustrates that separation of these angle error components is possible if a highly accurate platform attitude measurement system was used.

The purpose of Figures 13 thru 15 is to illustrate the estimation accuracy achievable for the platform-to-navigation frame component and composite misalignment angles of the group 2 error sources. The IMU case-to-navigation frame misalignment and synchro angle bias errors are shown in Figure 13. The upper graph shows that the vertical IMU case axis can be aligned about as accurately as the level axes. The lower graph shows that the six arc minute platform-to-case dependent synchro angle bias error uncertainty can be calibrated to one and a half arc minutes. It should be noted that the marginal accelerometer bias error filter convergence problem can be eliminated by compensating for these random bias errors on succeeding calibration tests.

The platform-to-navigation frame alignment accuracy is shown in Figure 14. These are composite error sources composed of the IMU case-to-navigation frame, synchro angle bias and differential synchro angle errors. These are commonly denoted as platform tilt angles and are only computed when the platform frame is approximately aligned to the navigation frame. One and one an a half arc minutes are the level and vertical axes estimation accuracies respectively.


PLATFORM GYRO DEPENDENT DRIFT RATE ESTIMATION
PLATFORM ATTITUDE
1.0
0.1
0.01
0.001
0
5
10
15
20
25
30
35
40
45
50
55
60
MEASUREMENT TIME (MINUTES)
--DATA--
X X-AXIS
Y Y-AXIS
Z Z-AXIS
IMU MEASUREMENTS
VELOCITIES
GYROBIAS (BIAS)


PLATFORM GYRO DEPENDENT DRIFT RATE ESTIMATION
PLATFORM ATTITUDE
1.0
0.1
0.01
0.001
0
5
10
15
20
25
30
35
40
45
50
55
60
MEASUREMENT TIME (MINUTES)
--DATA--
X X-AXIS
Y Y-AXIS
Z Z-AXIS
IMU MEASUREMENTS
VELOCITIES

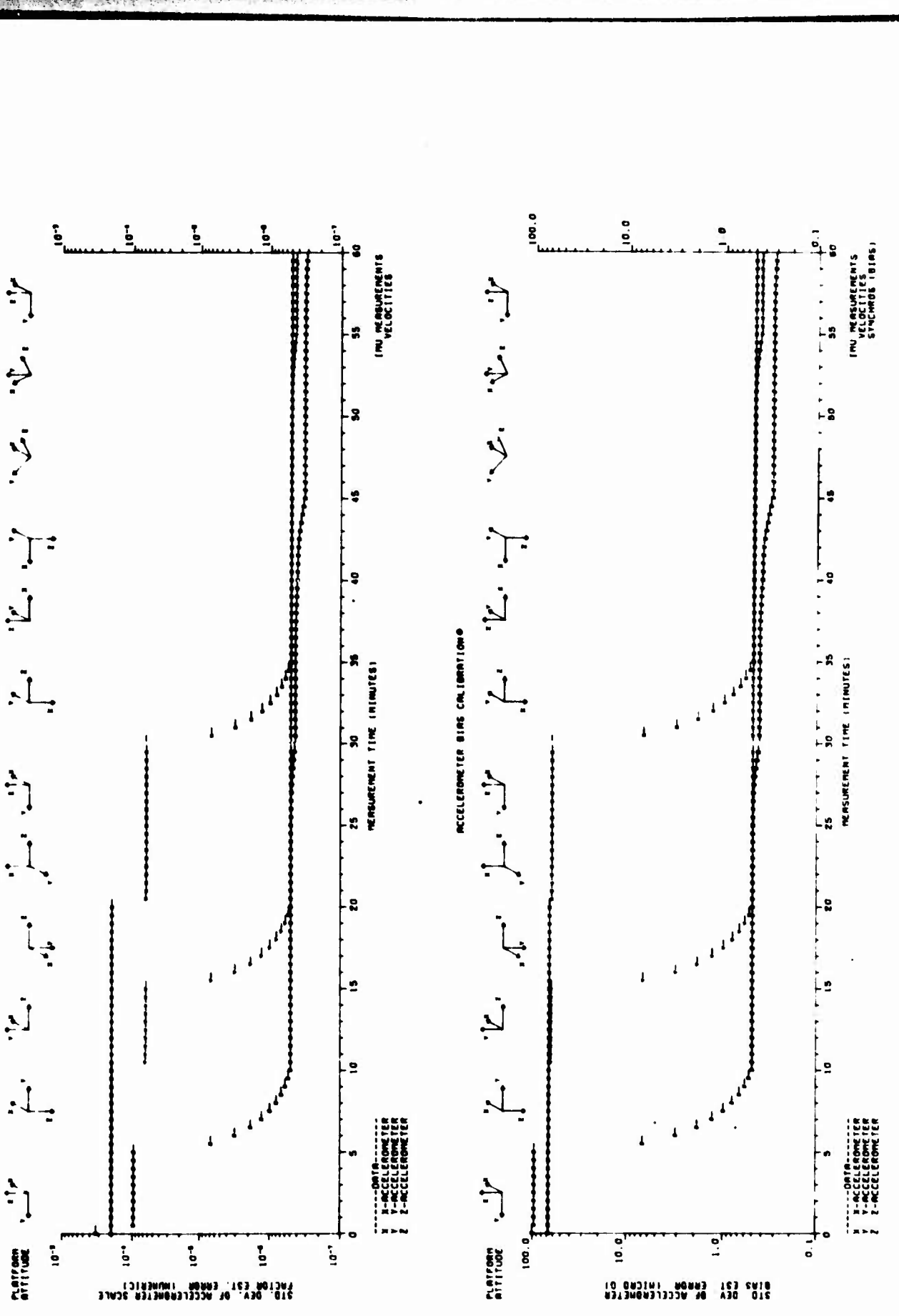
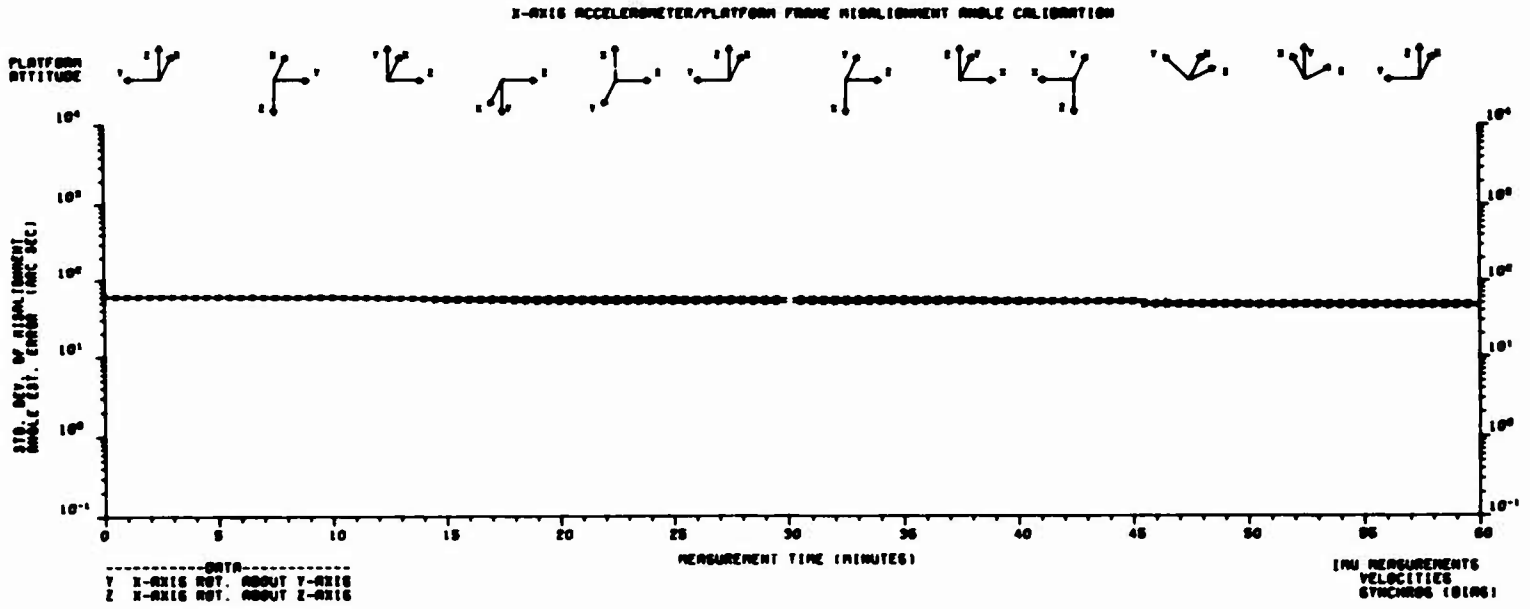


FIGURE (11) ACCELEROMETER BIAS AND SCALE FACTOR CALIBRATION



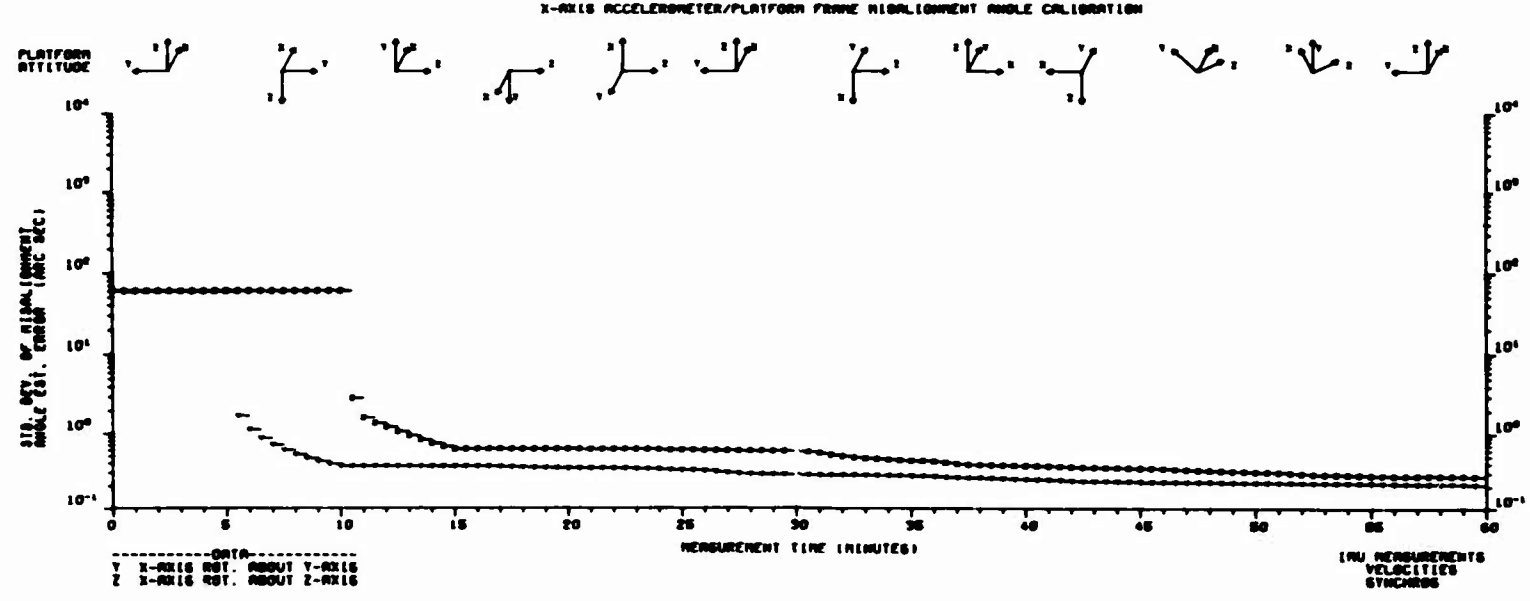


FIGURE (12) ACCELEROMETER-TO-PLATFORM FRAME MISALIGNMENT ANGLE CALIBRATION

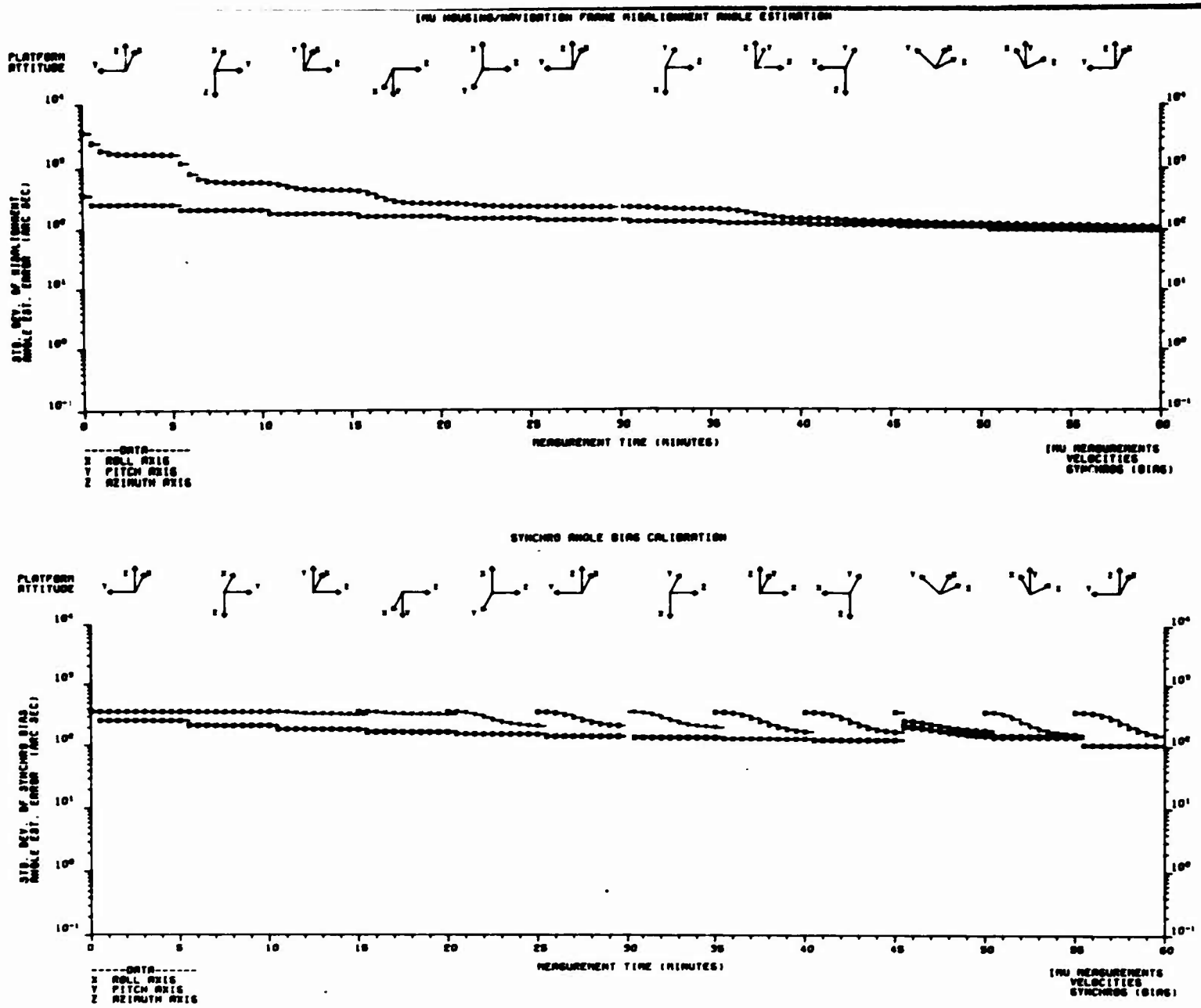


FIGURE (13) GROUP 2 ANGLE ERROR COMPONENT ESTIMATION

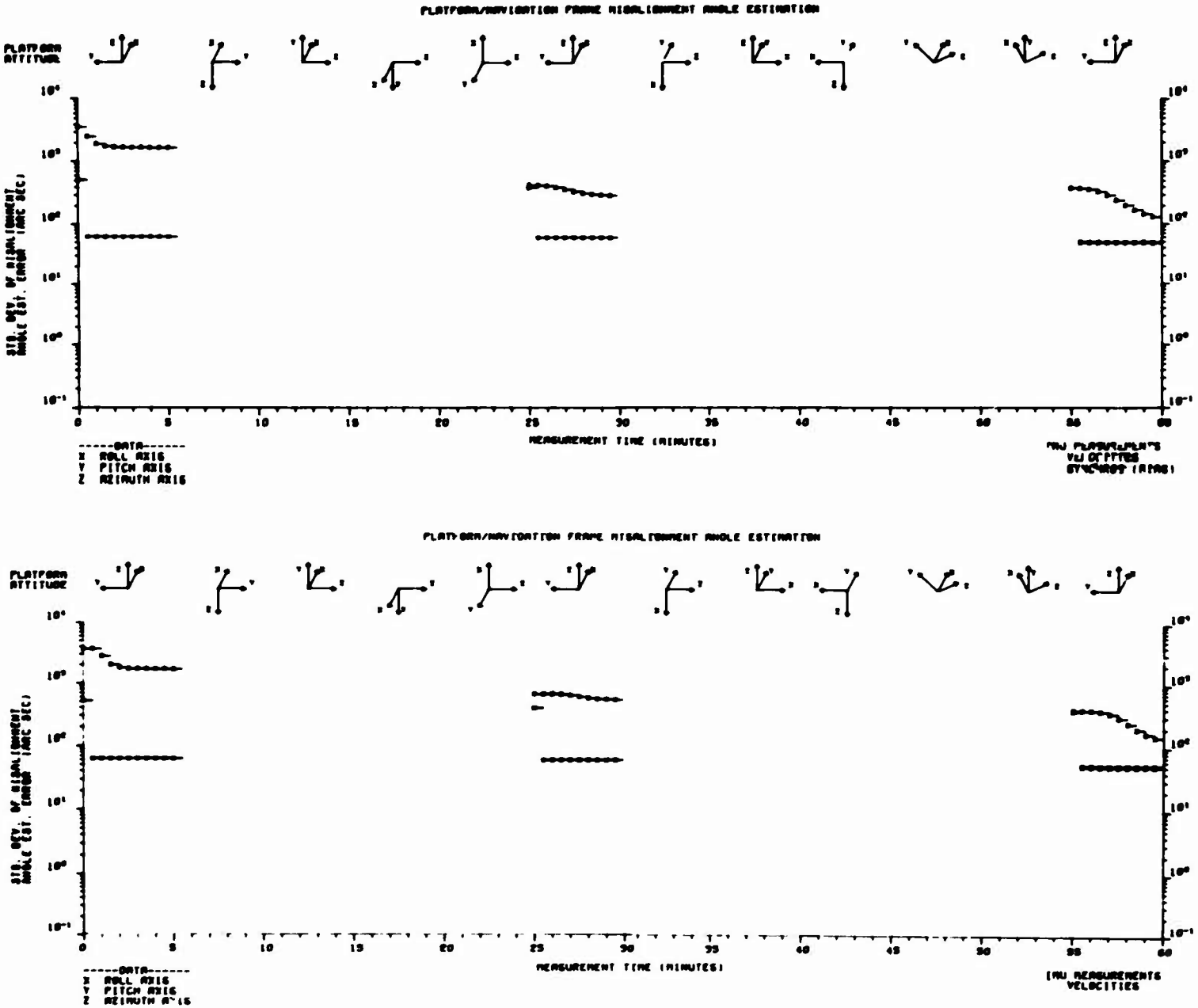


FIGURE (14) PLATFORM-TO-NAVIGATION FRAME MISALIGNMENT ANGLE ERROR ESTIMATION

Accurate estimation of the accelerometer electrical null axis-to-navigation frame misalignment angles is the major alignment criterion for navigation. The non-orthogonal transformation matrix includes six random variables as elements. The six misalignment angles are composite error sources composed of the platform tilt angles plus the accelerometer-to-platform frame and the pseudo accelerometer electrical bias angle errors. The misalignment angles can be separated into two similar groups. The estimation accuracy for one group is shown in Figure 15. The x symbol represents the non-orthogonality error of a level axis accelerometer with respect to the gravity vector. The y symbol represents the non-parallel error of the vertical accelerometer with respect to the gravity vector. The z symbol represents a level axis accelerometer azimuth misalignment error with respect to the navigation frame.

The pseudo accelerometer electrical bias angle error components were ommitted from the first platform test attitude to illustrate the accelerometer tilts that occurs when aligning with uncalibrated accelerometers. Comparison of the upper and lower graphs show that the synchro angle measurements do not have a significant effect on alignment accuracy. The final alignment accuracies are essentially limited by the inertial instrument random disturbances.

The specific IMU calibration and alignment parameters selected to illustrate the separation process were chosen from the eighty three component and composite error sources analyzed on the basis of emphasizing the methodology's salient performance characteristics. Simulation results obtained for the other error sources conform to the established performance characteristics. The adequacy of the error source estimation accuracies is examined in the next section.

VII. Verification of Performance Goal

The purpose of this section is to conceptually demonstrate that the formulated methodology is capable of satisfying the stated performance goal. Estimation of each significant IMU error source to an accuracy such that the navigational accuracy achievable with the IMU is essentially limited by uncompensatable system disturbances can be illustrated by running the simulation in a pseudo pure inertial navigation mode. The limitations of the computer program permit only a static open-loop navigation simulation. Dynamic navigation effects of Schuler tuning, gravity perturbations, specific force vector variations, etc. are not included.

A cause and effect type validation philosophy was assumed and is illustrated in Figure 16. The upper left graph shows the platform drift uncertainty associated with an ideally calibrated and aligned IMU. This configuration was simulated by making all initial covariance matrix elements, except those corresponding to the disturbance states, identically zero. The simulation begins with the platform exactly aligned with IMU measurements and filter updating occurring for the first five minutes. The twenty five minute pure inertial navigation performance is simulated by omitting the IMU measurements and propagating the covariance matrix. The upper right graph was obtained in the same manner except the group 2 platform-to-navigation frame misalignment angle error sources were included.

The lower left graph was obtained by retaining all the error sources and calibrating the IMU prior to the five minute alignment sequence. The twelfth calibration test sequence was utilized for the initial five minute alignment period. The lower right

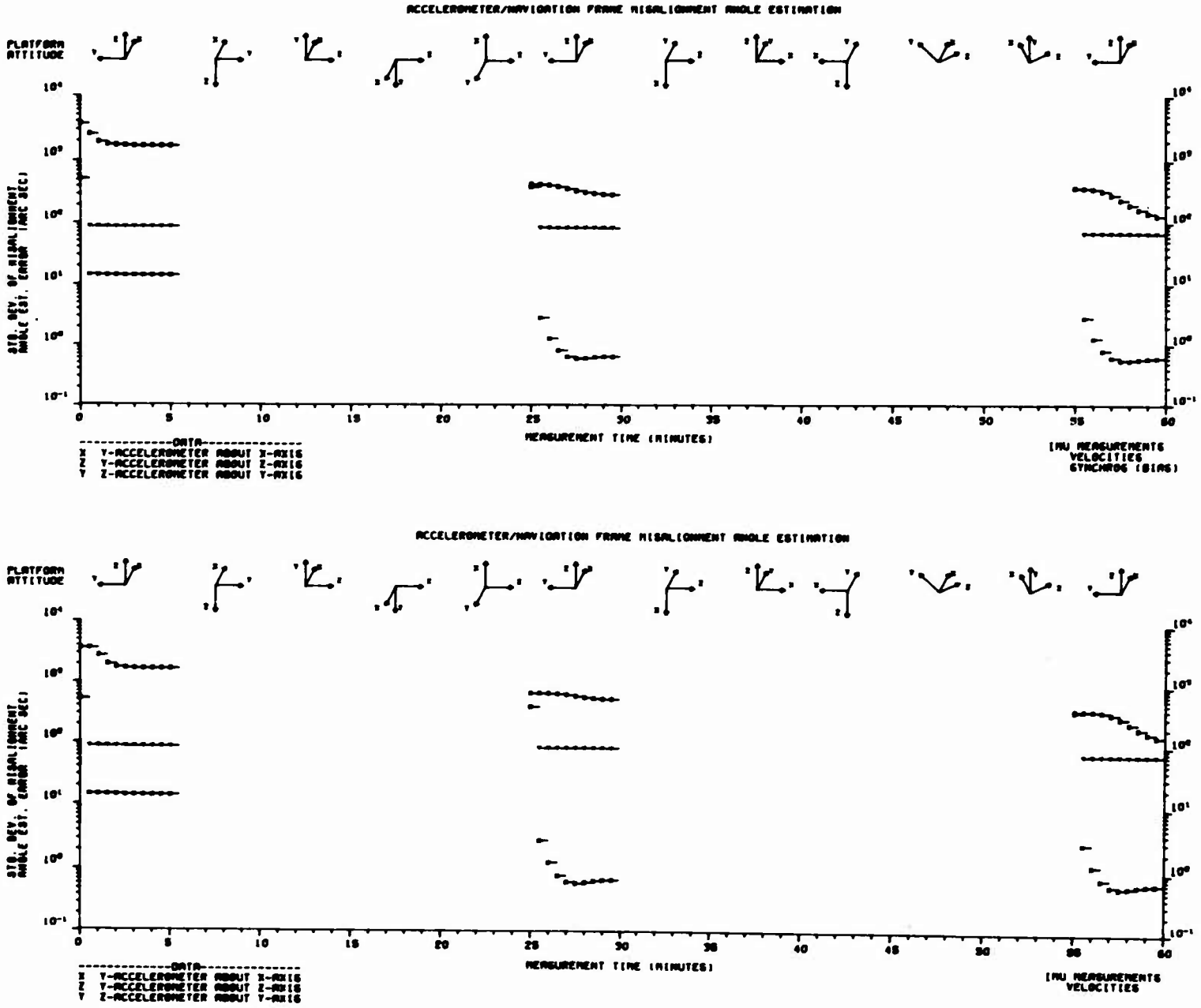


FIGURE (15) ACCELEROMETER-TO-NAVIGATION FRAME MISALIGNMENT ANGLE ESTIMATION

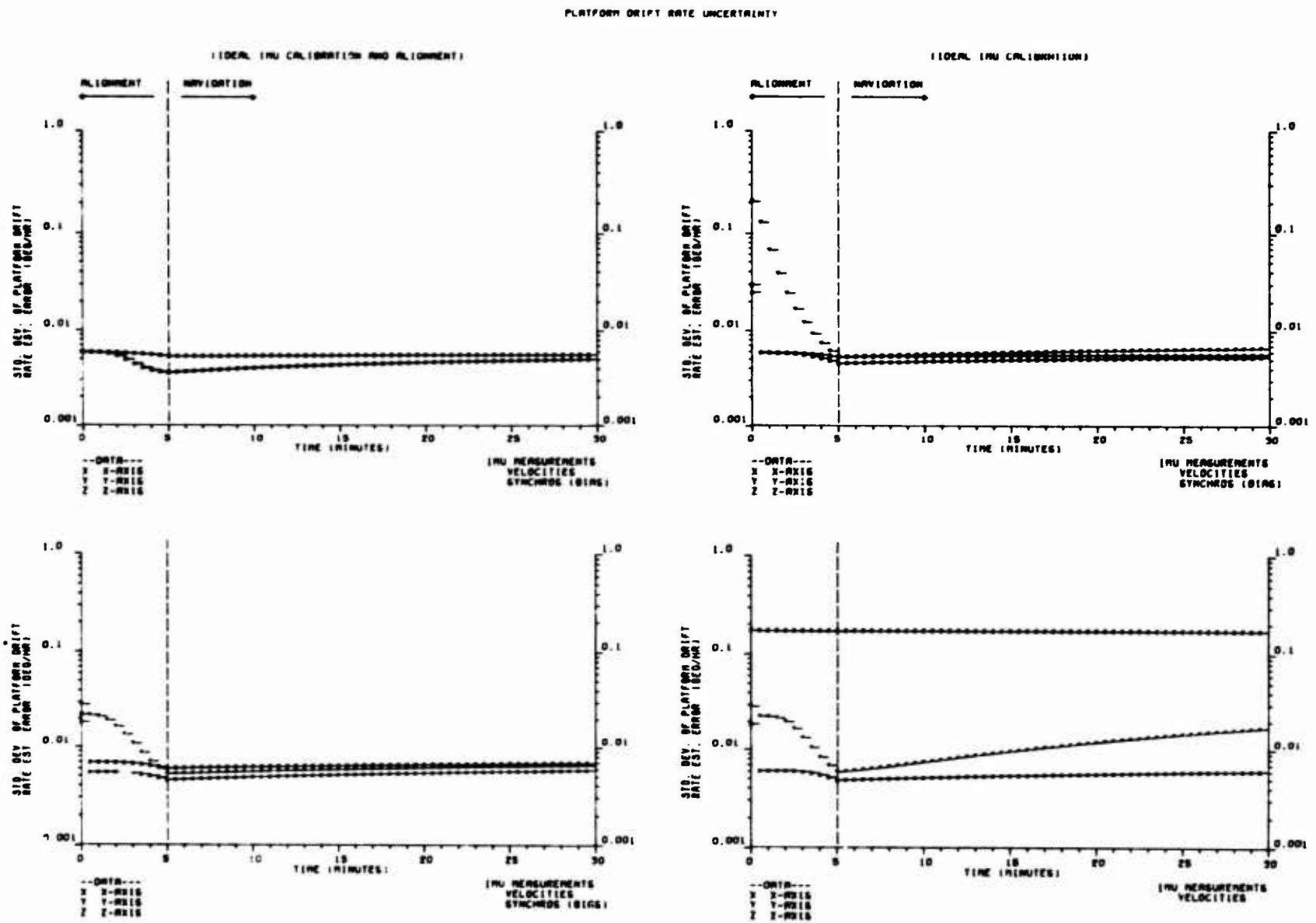


FIGURE (16) PLATFORM DRIFT UNCERTAINTY

graph was obtained in a similar mahner except the synchro angle measurements were not utilized.

Comparison of the first three graphs illustrate that platform drift uncertainty after twenty five minutes of navigation is roughly equivalent for each situation. The fourth graph shows the difficulty in estimating platform drift about the vertical axis from level axis accelerometer measurements. Vertical axis platform drift results in an increased azimuth platform misalignment angle, which, produces a larger west platform commanded angular velocity error component.

The manner in which the platform drifts integrate to produce accelerometer-to-navigation frame angular misalignments is shown in Figure 17. The x and y symbols represent the non-orthogonality errors of the level axes accelerometers with respect to the gravity vector. The z symbol represents the platform aximuth misalignment. The first graph shows the rapidity with which the gyro disturbances produce the misalignments. The level axis alignment uncertainty initially increases but begins to decrease after the filter has a knowledge of the corresponding platform drifts.

The second graph illustrates the problem of functionally separating the azimuth platform misalignment angle from the composite west platform drift. This is reflected by both the large azimuth angle uncertainty and the rather high rate at which the x accelerometer alignment uncertainty increases between measurements. Comparison of the second and third graphs shows no significant differences in the filter's capability to estimate the misalignments for the perfectly calibrated IMU as opposed to the calibrated one.

Figure 18 illustrates the erroneous velocity indicated by the stationary IMU. The open-loop navigation accuracy data shown approximately corresponds to that of a closed-loop Schuler tuned system since both of these systems exhibit similar performance characteristics over the first quarter cycle of the 84 minute Schuler oscillation period. This data is also similar to that resulting from a nominal Kalman filter IMU alignment technique. The nominal techniques being employed do not, however, include all the significant IMU error sources or utilize vertical axis accelerometer and syncho angle information in the estimation process.

Comparison of the first two graphs indicates that the five minute alignment time is not quite long enough for the filter to effectively remove the velocity error components associated with the misalignment angles. The second and third graphs again show equivalent level axis performance for an IMU containing no calibration errors as opposed to the calibrated IMU. The vertical axis errors are of little practical concern in cruise navigation systems since this channel is inherently unstable and must be damped with altitude information derived from an external source.

The filter's inability to accurately estimate vertical axis platform drift, without the precise azimuth synchro angle measurement, it reflected in the fourth graph. Comparison of the third and fourth graphs suggests that both the alignment time and subsequent navigation errors can be reduced by incorporating a more effective means of determinating vertical axis drift in the nominal Kalman filter alignment techniques.

In summary, Figures 16 thru 18 illustrate that the estimation uncertainty associated with platform drift is the major factor determing navigation accuracy for the KT-70 IMU type system. This uncertainty can be reduced to a level approaching the theoretical minimum, imposed by the gyro random disturbances, by use of the postulated calibration methodology.

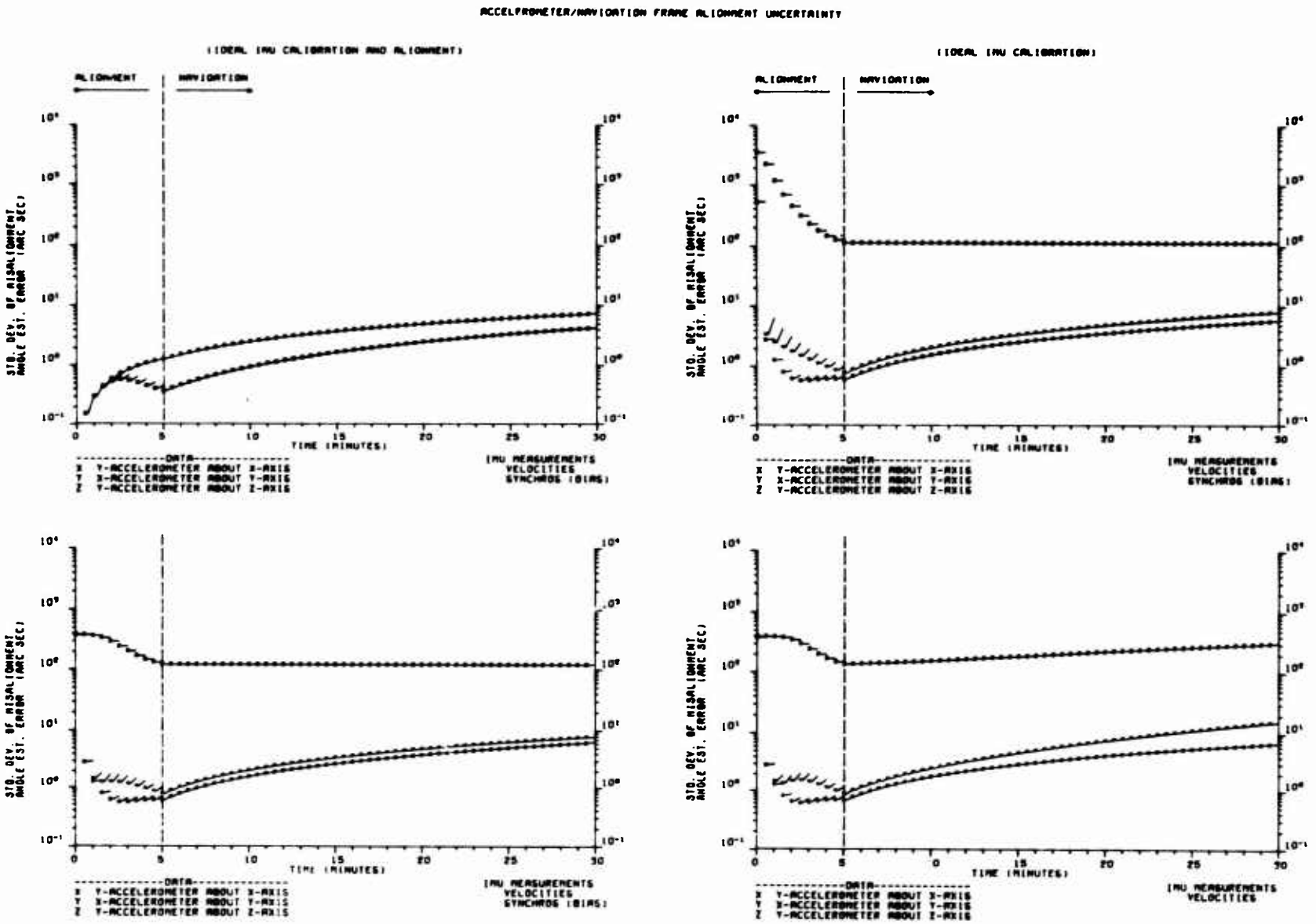


FIGURE (17) ACCELEROMETER-TO-NAVIGATION FRAME ALIGNMENT UNCERTAINTY

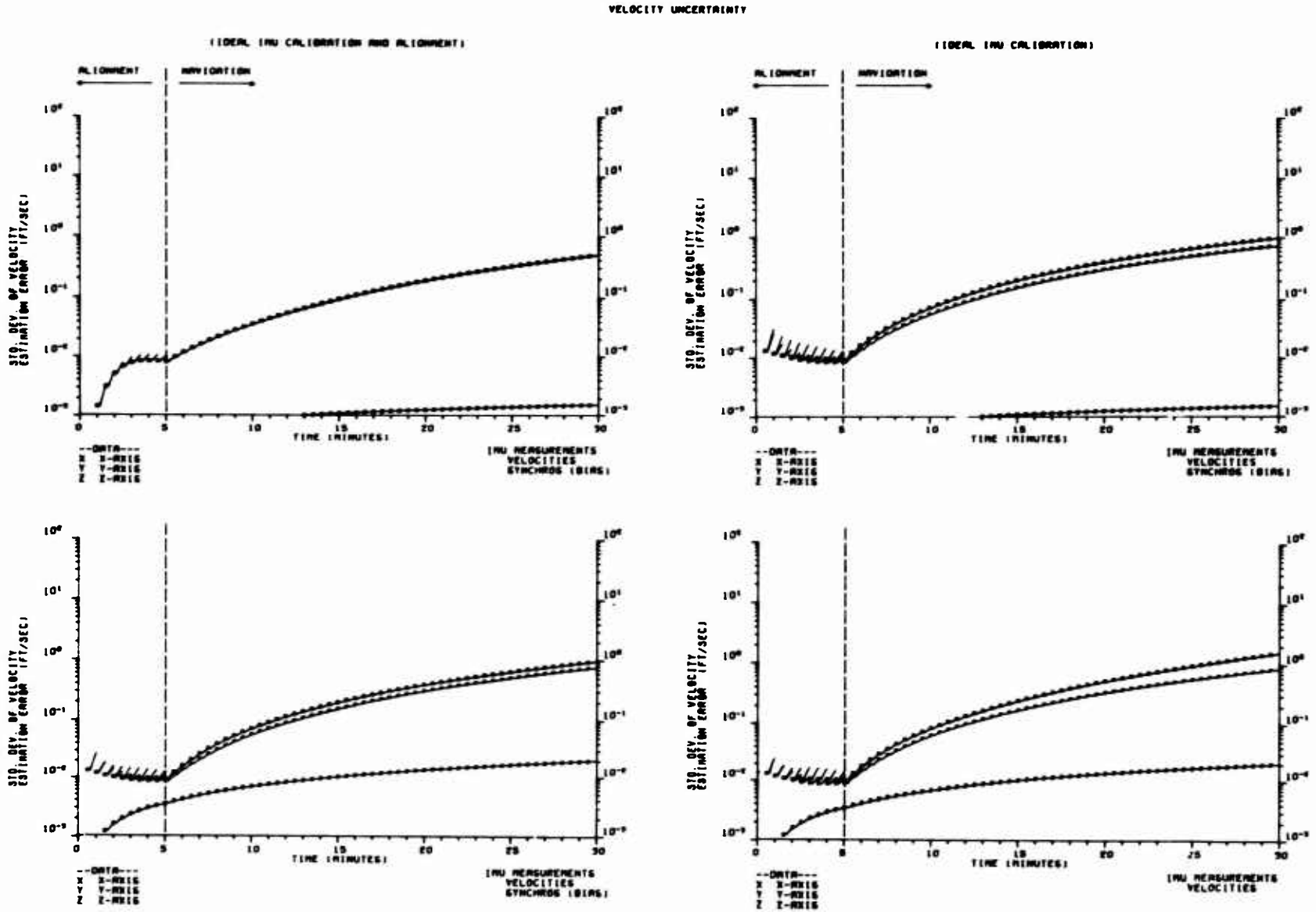


FIGURE (18) VELOCITY UNCERTAINTY

VIII. Software Concepts

Digital computer computational, storage and associated programming requirements for implementation of the IMU calibration and alignment methodology vary greatly in accordance with the application and software concept. All software concepts must, however, satisfy three functional requirements: (1) the platform must be commanded to precess at earth rate during each IMU test sequence. (2) System level IMU measurements must be recorded. And (3) IMU error source estimates must be computed. The first two requirements could be met in essentially the same manner for all concepts with software complexity being a function of the desired automation. The manner of accomplishing the third function distinguishes the concepts.

Three approaches to estimating the IMU error sources are outlined in the sequel to illustrate the software variations possible. These approaches cover the spectrum of expected methodology applications. Variations of the outlined approaches are also possible.

Depot and intermediate level IMU fault detection, fault isolation and calibration applications might utilize an off-line parameter estimation software approach. The Kalman filter mechanization and IMU test procedure formulated purposely omitted real time closed-loop platform control in favor of the off-line software flexibility. The IMU test would be automatically conducted by a small general purpose digital computer with IMU measurements being recorded on a magnetic tape. This test equipment could also be used with other IMUs by slight modification of the software.

The recorded IMU measurements would serve as input data to a comprehensive computer program hosted on a large digital computer. This program would contain extensive fault detection and fault isolation logic in addition to the Kalman filter algorithm. The program could be written in FORTRAN or another higher order programming language to reduce developmental costs and permit easy modifications.

Laboratory and field IMU status evaluation and calibration applications might utilize the small GP test computer and stored Kalman gain matrices to accomplish on-line in real time parameter estimation. This technique is possible because the filter's covariance matrix can be propagated and updated independently of the IMU measurements. This allows prior computation of the covariance and related Kalman gain matrices on a large digital computer. The precomputed Kalman gain matrices, corresponding to specified IMU measurement times, would be stored on a magnetic tape and read by the small test computer as required during the test.

Figure 19 shows the estimation accuracy possible with this technique when only a single stored Kalman gain matrix is used per platform attitude. The final 0.007 degree per hour level axis gyro bias calibration accuracy compares favorably with the 0.005 degree per hour accuracy obtained with the ten filter updates per attitude as shown in Figure 4. Estimation of the error sources is well within the small test computer's capability and is accomplished by multiplying the measurement vector by the Kalman gain matrix after each IMU measurement and propagating the optimal estimates between measurements. The software program could be written in either assembly or a higher order language depending on the test computer's capabilities.

Navigation reference system applications could utilize a reduced dimensionality Kalman filter algorithm for real time error source estimation and platform alignment. This would permit the subsequent aided-navigation function to be performed by the calibration and

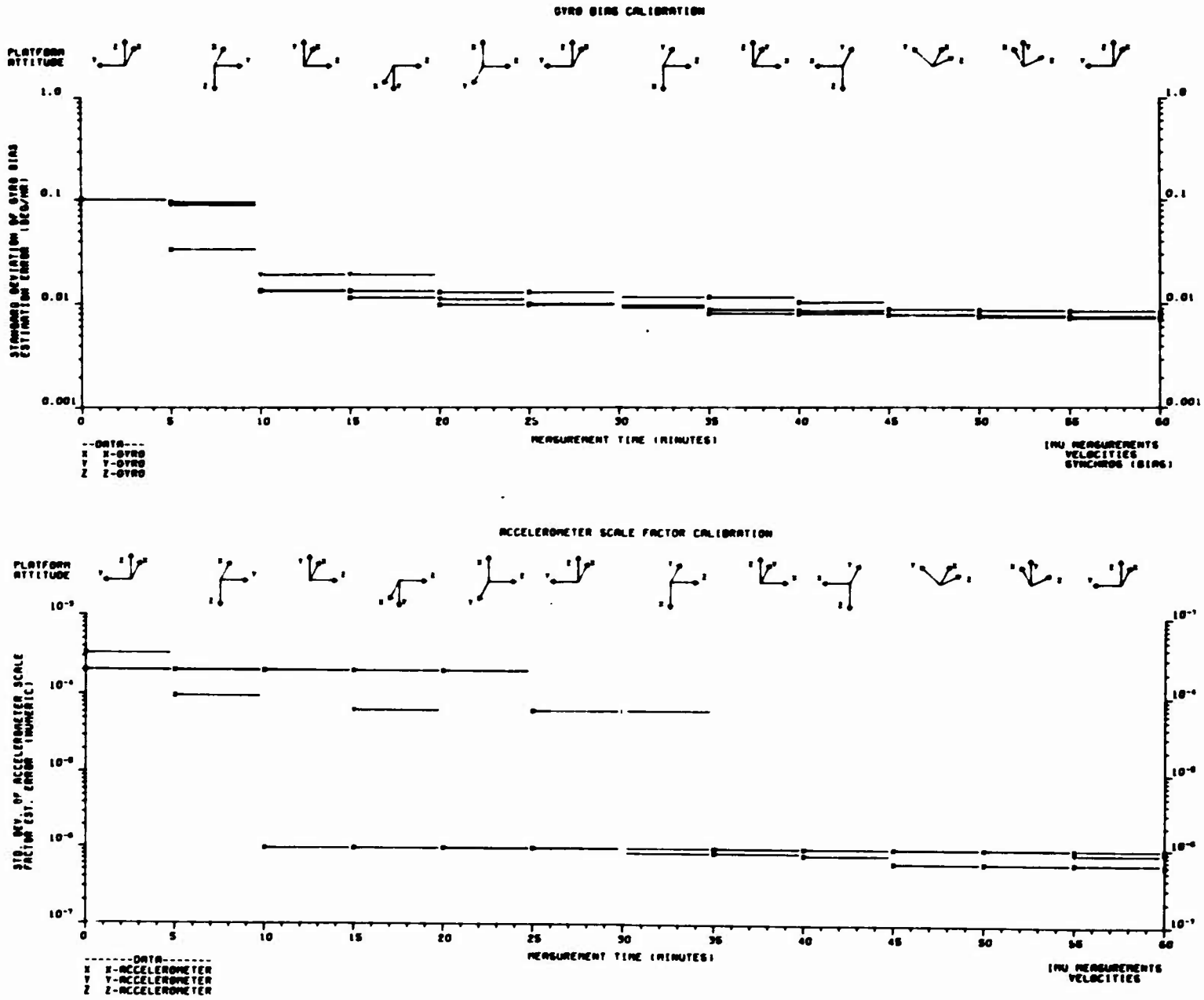


FIGURE (19) GYRO BIAS AND ACCELEROMETER SCALE FACTOR CALIBRATION

alignment filter by suitable algorithm modifications at the start of the navigation phase.

The filter's dimensionality can be reduced by software compensation of those IMU calibration parameters whose magnitudes are relatively fixed in time. These error sources would be periodically calibrated using either of the above approaches. Error states modeled by the filter would include only those calibration parameters having correlation times of less than several weeks or exhibit significant turn-on to turn-on variations.

If reference system IMU calibration and alignment is to be accomplished on a mobile base, whose perturbating motions introduce IMU measurement outputs significant relative to the inertial instrument noise levels, additional filter states will be required to model the perturbation errors. These perturbation errors can become the major factor limiting error source estimation accuracy. The total number of filter states implemented will depend to a large extent on the special purpose navigation computer resources available.

IX. Conclusions

A generalized IMU error analysis methodology has been formulated for simultaneous fault detection, fault isolation, calibration and alignment. Methodology application to a gimballed IMU results in a system of stochastic differential equations that can be manipulated into a form consistent with the Kalman filter algorithm. Simulation results indicate that optimal estimation of the IMU calibration and alignment error sources, to an accuracy limited essentially by inertial instrument random disturbances, can be accomplished in approximately one hour of open-loop testing utilizing only system-level IMU measurements. IMU fault detection and fault isolation can be accomplished by comparing the magnitudes of the recovered error sources with their a priori statistical distributions.

A small digital computer is required to conduct the IMU tests. On-line parameter estimation can be accomplished with a minimum of software by utilizing stored Kalman gains. Parameter estimation flexibility and a comprehensive fault detection and isolation capability can readily be achieved by using recorded IMU measurements in conjunction with an off-line higher order language program hosted on a large general purpose digital computer. On-line real time filter operation appears possible for those navigation reference system applications where the calibration and alignment filter would also be utilized for the subsequent aided navigation phase.

References

(1) Platform Model for the A-7D/E Inertial Measurement Set, Kearfott Systems Division, Kearfott Group, General Precision Systems Inc, Wayne, New Jersey, Doc. No. 6K032D025C, 1968, Unclassified.

(2) Stochastic Processes and Filtering Theory, Andrew H. Jazwinski, Academic Press, New York, 1970.

MODIFICATIONS TO THE ITERATED-EXTENDED KALMAN FILTER THAT CAN IMPROVE PERFORMANCE

Ronald E. Janosko, Captain, USAF
Research Associate
Frank J. Seiler Research Laboratory
USAF Academy, Colorado 80840

ABSTRACT

In any nonlinear estimation problem modifications can often be made to the standard iterated extended Kalman filter that will in some sense improve the performance of the filter. These modifications are generally tailored to treat the specific nonlinearities of the problem being studied. Often these changes to the filter are not reported in sufficient detail to make them useful to other investigators. Thus researchers are often led down blind alleys because others have not reported their methods completely or given attempts that may have failed.

Several mutations to the standard filter were attempted for a problem that is quadratic in the measurement nonlinearity. All attempts, both successes and failures, are reported in this paper. These changes include updating the state error covariance matrix more often, artificially changing the assumed measurement error covariance matrix, using a variable step gain in the filter, and various combinations of the above. The results of numerous computer simulations are included for all modifications attempted.

INTRODUCTION

The extended Kalman filter attempts to account for measurement and/or state transition nonlinearities by using a first order expansion of the nonlinear equations about some nominal value of the state to produce equations that are linear in the states By continually iterating about an updated state estimate at each measurement until a prespecified criterion is met, some of the convergence properties of this filter can be improved.

This paper presents some further changes to the filter that can affect the convergence properties of the filter. No claim is made that the modifications to be presented will improve filter performance in all situations. However, it will be demonstrated that these changes

do help the estimator for the particular problem studied. The modifications listed in this paper are changes to the form of the filter as well as to how some of the filter parameters are selected. These mutations were determined through computer simulations and are not analytically derived.

Before presenting the modifications, the basic Kalman filter equations will be presented so as to indicate the notation used and the problem simulated. The situation studied is restricted to a discrete problem with constant states and a nonlinear measurement.(1)

That is

$$\bar{x}_{i+1} = \hat{x}_i \ , \tag{1}$$

and

$$\bar{z}_i = \bar{h}(x_i) + \bar{v}_i \ . \tag{2}$$

In the above x_i is an n dimensional state vector at the ith measurement time, z_i is the corresponding m dimensional measurement vector, $\bar{h}(x_i)$ is a nonlinear measurement function relating the n states to the m measurements and $\bar{v}_i$ is a random noise with $E(\bar{v}_i \bar{v}_j^T) = R_i \, \delta_{ij}$ where R_i is the covariance matrix of the measurement noise at time i. The quantity $\bar{x}_{i+1}$ is the estimate of x_i before the $(i+1)$st measurement, and $\hat{x}_{i+1}$ is the estimate after the $(i+1)$st measurement.

The extended Kalman filter obtains an estimate of x_{i+1} based on the $\bar{z}_{i+1}$ via the relation

$$\hat{x}_{i+1} = \bar{x}_{i+1} + K_{i+1}\left(\bar{z}_{i+1} - \bar{h}(\bar{x}_{i+1})\right). \tag{3}$$

In the above equations

$$K_{i+1} = P_i H_{i+1}^T \left(H_{i+1} P_i H_{i+1}^T + R_{i+1}\right)^{-1} \ , \tag{4}$$

and

$$P_{i+1} = \left(I - K_{i+1} H_{i+1}\right) P_i \left(I - K_{i+1} H_{i+1}\right)^T + K_{i+1} R_{i+1} K_{i+1}^T \ . \tag{5}$$

The quantity K_{i+1} is the Kalman gain, P_{i+1} is the state error covariance matrix. Note that because the transition matrix is the identify matrix the P matrix does not have to be updated between

measurements. The elements of H_{i+1} are defined by

$$H_{i+1}^{\ell j} = \left.\frac{\partial h^{\ell}}{\partial x^{j}}\right|_{\substack{\text{evaluated at} \\ x = \bar{x}_{i+1} = \hat{x}_i}} \qquad \begin{matrix} \ell = 1,\ldots m \\ j = 1,\ldots n \end{matrix} \quad . \tag{6}$$

The superscripts ℓ and j are used to indicate elements of the vector or matrix.

In the iterated extended Kalman filter the same measurements are repeatedly used in updating the estimate of the state obtained from Equation (3) except that the H_{i+1} and K_{i+1} matrices are reevaluated after each new state estimate.[2] This procedure continues until the difference between successive updates is below a prespecified criterion. At that point the P_i matrix is updated by means of Equation (5) with the latest H_{i+1} and K_{i+1} values. The filter is then ready to process a new set of measurements.

Helpful though this technique is in dealing with nonlinear problems, it does not solve all of the associated pitfalls. In general, two problems often remain. The first is common to almost all forms of the Kalman filter in that as time progresses the Kalman gain becomes extremely small. Should a change now occur in the states, the filter is unable to track it accurately. The second is more limited to nonlinear applications. In this area, the change in the state estimate given by the Kalman filter, called $\Delta\hat{x}_i$ and defined as

$$\Delta\hat{x}_i = K_{i+1}\left(\bar{z}_{i+1} - \bar{h}(\hat{x}_i)\right) , \tag{7}$$

may not be useful for highly nonlinear problems. This is because the step taken may be too large and the linearization used in extending the filter may no longer be valid.

The modifications to be presented will attempt to alleviate the above problems.

SIMULATION EXAMPLE

All modifications were attempted on one particular example. This problem is quadratic in the measurement nonlinearity and is given below

$$z_i^{\ell} = x_i^T x_i - 2x_i^T s_i^{\ell} + v \qquad \ell = 1,2,\ldots m \quad . \tag{8}$$

In the above x_i is a three dimensional constant state vector, z_i^{ℓ} is the ℓth element of an m element measurement vector, and s_i^{ℓ} is a time varying, but known, quantity. The matrix R_i is diagonal with the value 2×10^2 km^4 on all nonzero elements. For reference the magnitude of z_i^{ℓ} is approximately 1.6×10^7 km^2. The initial state error is about 78 km with the true state, x_i, and the initial approximation, $\hat{x}_0$, as given below

$$x_i = \begin{pmatrix} 1683.50 \\ 2871.61 \\ 506.34 \end{pmatrix} \text{km} \ , \qquad \hat{x}_0 = \begin{pmatrix} 1633.81 \\ 2911.18 \\ 461.08 \end{pmatrix} \text{km} \ . \tag{9}$$

The measurements were simulated using the true states and then corrupted by a random number generator so as to have the above variance. No claim of a Monte Carlo simulation is made because of the extensive number of simulations that would have been needed to place any confidence in the results.[3] The technique used for the simulations is mentioned later in the paper.

NUMBER OF MEASUREMENTS STUDY

Because we can essentially have any number of measurements in any one data set, it was at first reasoned that processing as many measurements as possible simultaneously would give a better and faster state estimate. That is, taking four measurements per data set and processing all four as a set should give a better result than taking and processing only one measurement per set. However, it was found that this is not necessarily the case. In general there is a point where improvement in state estimate no longer increases significantly with an increase in the number of measurements processed as a set.[4] The above fact is clearly depicted in Figure 1. In this figure the rss state error after 100 data sets are processed is plotted as a function of the number of measurements processed in a set. It is noted that after 3 measurements per set there is not much improvement in the final error.

In this simulation the measurement error covariance matrix used in the filter equations was given the value 2×10^2 km^4 on all elements of the diagonal and zero elsewhere. That is, the filter was using the actual value for the variance of the measurement error.

Because the final rss state error was still larger than desired, regardless of the number of measurements processed, it was decided to vary another filter parameter. This time the value used on the diagonal of the measurement error covariance matrix used in the filter equations was varied. The measurements were, however, generated to have the same variance as in the previous case.

Figure 2 shows the results of this study. In this figure the rss state error after the processing of 100 data sets is plotted as a function of the number of measurements taken per set for three values of the measurement error covariance matrix. It is noted in this figure that the point beyond which increasing the number of measurements processed no longer significantly decreases the rss state error is a function of the value used for the measurement covariance.

It appears from this figure that increasing the error covariance matrix improves the response for a low number of measurements but deteriorates it as the number of measurements increases.

Because this performance was still not acceptable it was decided to try to modify the form of the filter. This modification is given below.

FIRST MODIFICATION

The attempt to change the form of the filter at first seemed very logical. In this first modification it was decided to try to update the P_{i+1} matrix every time the state vector was updated, that is at every iteration and not just once per measurement set. It was felt that if the state estimate was being updated, the confidence in that update should also be changed.

The result of using this change is given in Table 1. In this table the number of iterations used for 50 data sets is listed for both the standard and the modified filter. The data is given as a function of the stopping criteria in the state iterations. It is noted that this modification does produce a filter that requires a lower number of iterations. The error in both filters was about equal. In this set of simulations the value used for R_i was 2×10^2 km^4.

Again it was decided to vary the value used for R_i. The results of that exercise are given below.

ERROR COVARIANCE STUDY

Because this was a simulation study the actual error covariance could be calculated. Varying the value used for the error covariance in the filter equations from its true value indicated that lying to the filter can improve the convergence properties at least for the particular problem studied.

Figure 3 shows the results of varying the estimate of the measurement error covariance for processing 4 measurements per set. In this figure the rss state error after 100 data sets is plotted as a function of the value used for the diagonal elements of the measurement error covaraince matrix.

From this figure we note that in general using a smaller value for the error estimate will yield better results than using a larger value. This however is contradicted by the unusual behavior at about 2×10^{12} km^4. Also it is seen that too small a value can cause the filter to diverge.

The response of the system starting from a new initial condition was also studied. In this case the initial rss state error was 234.56 km. Also this necessitated a new initial state error covariance matrix to be calculated. The rss error after 100 data sets for various measurement error covariances for this case is shown in Figure 4. Of note is the sharp decrease in error at 2×10^{7} km^4 almost indicative of some form of resonance behavior. Also it is noted that increasing the estimate of the error covariance matrix above its true value will improve the response whereas decreasing it will first deteriorate and then improve the response. None of the above behavior can be adequately explained at the present time.

In addition to this extreme sensitivity to the R_i value, this filter had another larger drawback. In this filter the value obtained for the P_{i+1} matrix was extremely optimistic and generally of no value. Because of the above difficulties and the filter's unexplainable behavior it was decided to abandon this type of modification and look for another approach to improve the filter performance. The mutation listed below is to an otherwise standard filter.

SECOND MODIFICATION

Because the value chosen for the R_i matrix did seem to control the response characteristics of the filter it was decided to see what role this value played in the filter equations. Much has been written

about the part that the R matrix has in relationship to filter divergence; it is treated somewhat differently than what follows.[5],[6] To gain some insight into the role that it plays in this particular problem assume that the R_i matrix can be written as a constant times the identity matrix. We now note that this constant then determines the relative weights given the $H_{i+1}P_iH_{i+1}^T$ term and the R_{i+1} term in Equation (4). Further because this is a scalar weight it seems reasonable to assume that it might affect the value of K_i in the same manner as a scalar multiplier. Strictly speaking this is not the case although there does appear to be good reason to assume that the Kalman gain can be controlled by a scalar multiplier.

The above gives us reason to rewrite Equation (3) as

$$\hat{x}_{i+1} = \bar{x}_{i+1} + \alpha_i K_{i+1}\left(\bar{z}_{i+1} - \bar{h}(\bar{x}_{i+1})\right) , \qquad (10)$$

where α_i is a scalar multiplier greater than zero. Thus if α_i is small it minimizes the step size taken to aid in highly nonlinear problems. On the other hand a large α_i can compensate for the gain becoming too small. The problem that remains is how to choose an α_i that is optimal in some sense. It should be noted that $\alpha_i = 1$ yields a standard Kalman estimate.

First let us consider possible criteria under which α_i can be chosen, along with their obvious advantages and/or disadvantages. In general most criteria fall into two broad classes: those that consider some type of average value for a data set, and those that treat only individual components. Most often we desire some minimum within these classes.

Under the first heading the following are possible criteria:

[1] minimize : average measurement error

[2] minimize : weighted average measurement error

[3] minimize : average absolute measurement error

[4] minimize : rms measurement error.

The characteristic common to all of these is that the measure thus obtained is not overly biased by one excessively large error if all of the other errors are small.

The disadvantage of [1] is that large positive errors can cancel out large negative errors for various values of α_i. Thus this

criterion may show no change as α_i is varied or it may even decrease when in fact large errors do exist. This same situation is possible with criterion [2] with the additional drawback that the weights have to be properly determined in order to make this measure of much value.

Criteria [3] and [4] eliminate this positive-negative cancellation problem. However, it is still possible for both of these measures to be decreasing while some of the individual errors are increasing. This will happen, of course, if the remaining errors are decreasing faster than those that are increasing. This phenomenon is due to the averaging property of these measures. Most often this property is an advantage, though at times it can be a disadvantage. This is particularly true if it is desirable to keep all individual error components minimal.

To eliminate this problem, the criterion can be based on individual component behavior. A common example of this is the technique to be presented in this paper. That is the criterion:

minimize : maximum measurement error

will be used. This measure allows a good index of performance and also stops divergence of any one individual error. In effect, all error components will now be kept within some upper bound.

This criterion is satisfied by finding an α_i, denoted by $\tilde{\alpha}_i$, that minimizes the maximum deviation between each z^{ℓ}_{i+1} and $h^{\ell}(\hat{x})$ for $\ell = 1,2,\ldots m$ as the state estimate moves from $\bar{x}_{i+1}$ along the direction $\Delta\hat{x}_i$ given by the Kalman filter.

The minimization can be shown diagrammatically in Figure 5. In this figure, $\varepsilon^1_{i+1}(\alpha_i)$, $\varepsilon^2_{i+1}(\alpha_i)$ and $\varepsilon^3_{i+1}(\alpha_i)$ are components of the measurement error vector. The function $f_{i+1}(\alpha_i)$ is seen to be the least upper bound of the momentarily largest error component. It should be noted that the largest error can be any element of the error vector and more importantly the element number can change as α_i is varied. The optimal step, $\tilde{\alpha}_i$, is also shown in this figure.

It cannot be claimed that the algorithm used to solve for the optimal is the most efficient possible, however, it is felt that it is rather complete.

Further, as will be shown later, the reduction achievable in the number of iterations required for each data set by using the variable step process more than compensates for any inefficiencies. This algorithm is extremely straightforward considering it is used to minimize a function that can have a discontinuity.

The algorithm used will not be presented in this paper but is contained elsewhere.[7]

It should be noted that this approach might be sensitive to measurement noise. To see this possibility assume that we have a scalar measurement and a scalar state to be estimated. In this case the algorithm will solve for the state which exactly corresponds to the measurement and disregards the fact that noise is present. However, with a large number of components in the measurement vector $\bar{z}_{i+1}$ this should not be much of a problem. This is because we would expect the noise to be better distributed with a large number of samples. The example presented in this paper has $m > n$, that is, more measurements than states. The sensitivity of this criterion to the relative size of m to n is to be investigated at a later date.

Table 2 shows the results of using the standard extended iterated Kalman filter and the variable step modification for the problem studied. In this table it is noted that number of iterations required to process the data from 50 sets is significantly less for the modified filter. It is also seen that the modified filter is less sensitive to the stopping criterion used to terminate the iterations than is the standard filter. This later fact allows the use of a tighter convergence criterion without an increase in the number of iterations required.

The fact that both forms of the filter give nearly the same estimates for the states is shown in Figure 6. In this figure the magnitude of the state error vector is plotted for both forms of the filter as a function of the number of data sets processed. It is clear that both filters exhibit the same basic response characteristics with the possible exception that the modified filter is not quite as smooth as the standard form.

The same filter characteristics are shown in Table 3 when a step change is made in the true state at the 26th data set. The step was such as to give a magnitude change in the state vector of 10.34 km. Here the advantage of using the variable step is even more evident. It is noted that the standard filter shown hit the maximum number of iterations allowable per orbit at least once during the 50 data sets. This maximum was arbitrarily set at 500 to control the amount of computer time required in case the filter failed to converge for any data set.

Figure 7 reinforces the fact that the variable step filter can track a step change in the states. The magnitude of the error vector for both the fixed state and the step change simulations using the modified filter are drawn on this figure. It is seen that there is little difference in the error for both cases. This was not the situation with the standard filter.

Table 4 shows the effect of varying the value of R_i. It is noted that the response is fairly independent of the R_i value. The rss error associated with each case was approximately the same except for the one situation noted as having a large error. This was not the case for the standard filter.

It should also be mentioned that the value obtained for the P_{i+1} matrix was consistent with the actual state error for the variable step filter.

REMARKS AND CONCLUSIONS

The numerical simulations for this problem were performed on a Burroughs B6700 computer and programmed in the FORTRAN IV language.

Instead of using Monte Carlo techniques most simulations were run for two cases, one adding and the other subtracting the random component to the true measurement. No noticeable differences were noted between these two cases. In all of the simulations the random number generator was started at the same initial value, thus all sets of random numbers were identical. This eliminated the possibility that one filter simulation would have a "better" set of measurements than another and thus bias the results.

At present the results are limited to the particular problem studied although it is felt that other problems may exhibit a similar phenomenon. It is hoped that the presentation of these results will stimulate analytical investigations whose goal will be the determination of the correct filter parameters to use for a variety of estimation problems.

REFERENCES

1. Bryson, A. E., and Ho, Y.C., Applied Optimal Control, Blaisdell Publishing Company, Waltham, Mass., 1969.

2. Jazwinski, A.H., Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

3. Bucy, R.S., Hecht, C., and Senne, K.D., "An Engineer's Guide to Building Nonlinear Filters, Vol. I," SRL-TR-72-0004, Frank J. Seiler Research Laboratory, Colorado, May 1972.

4. Janosko, R.E., "Preliminary Investigation into the Initial Response of the Iterated Extended Kalman Filter," Proceedings of the Fourth Symposium on Nonlinear Estimation Theory and its Applications, Western Periodicals Co., North Hollywood, California, 1973.

5. Denham, W.F. and Pines, S., "Sequential Estimation when Measurement Function Nonlinearity is Comparable to Measurement Error," AIAA/ION G&C Conference Proceedings, August 1965, pp. 174-180.

6. Fitzgerald, R.J., "Divergence of the Kalman Filter," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, pp. 736-747, December 1971.

7. Janosko, R.E., "Variable-Step Algorithm to Aid Convergence of the Iterated Extended Kalman Filter for Use in a Navigation Satellite Problem," AIAA Mechanics and Control of Flight Conference, 5-9 August 1974.

FIGURES

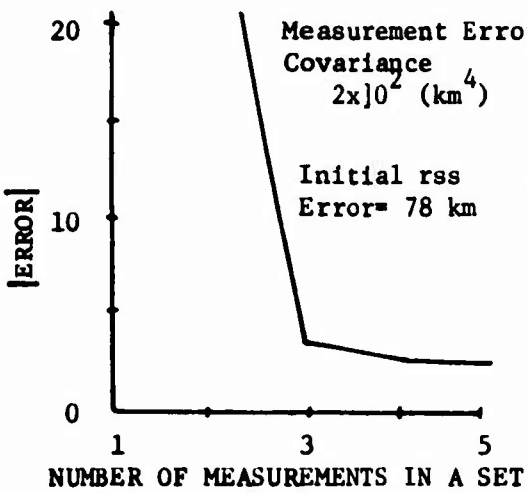


FIGURE 1

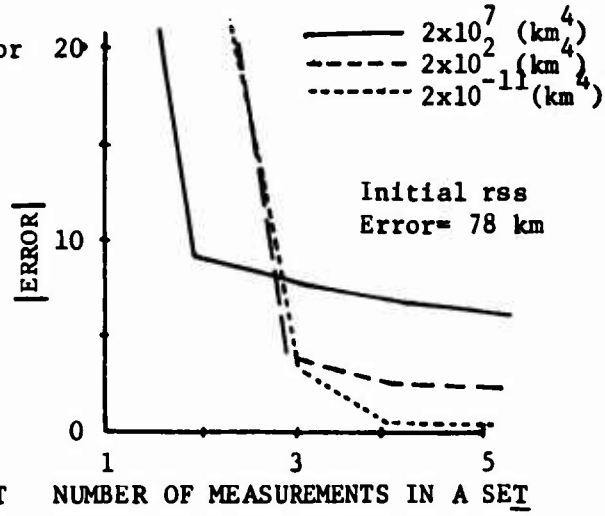
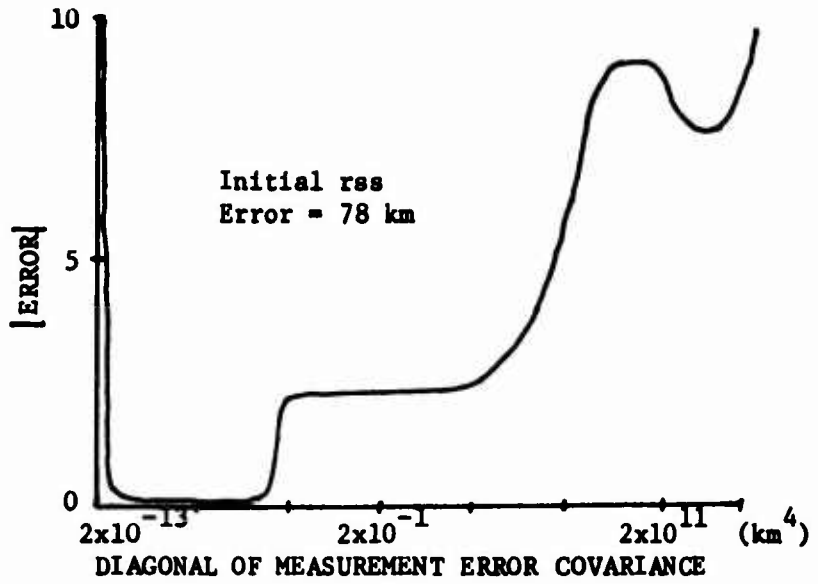


FIGURE 2

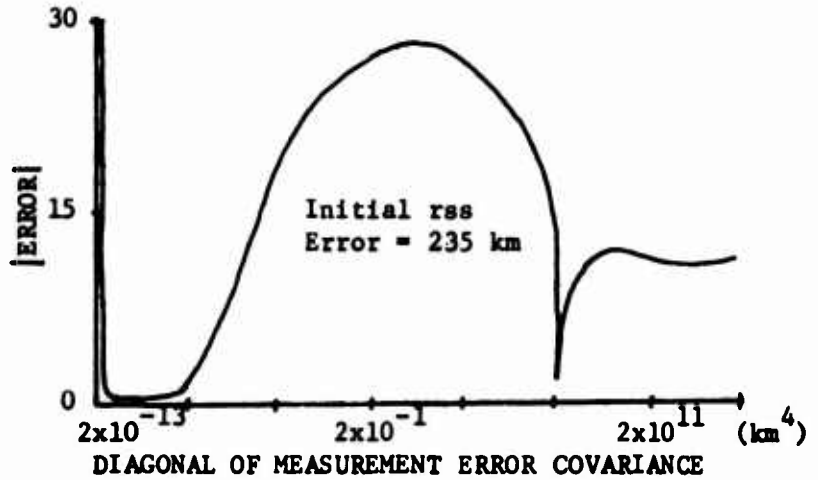
ITERATION STOPPING CRITERIA	NUMBER OF ITERATIONS (50 DATA SETS)	
	STANDARD FILTER	MODIFIED FILTER
1×10^{-7}	4662	3121
1×10^{-6}	1843	684
1×10^{-5}	130	202

Table 1.



DIAGONAL OF MEASUREMENT ERROR COVARIANCE

FIGURE 3



DIAGONAL OF MEASUREMENT ERROR COVARIANCE

FIGURE 4

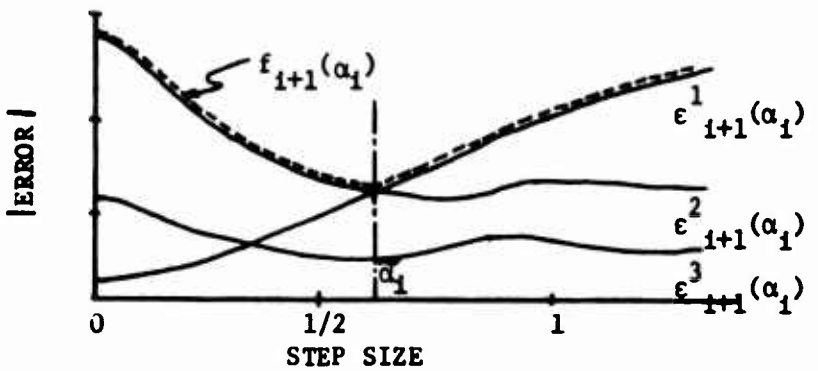


FIGURE 5

ITERATION STOPPING CRITERIA	TOTAL ITERATIONS FOR 50 DATA SETS	
	STANDARD STEP	VARIABLE STEP
1×10^{-7}	4662	127
1×10^{-6}	1843	127
1×10^{-5}	130	113

Table 2.

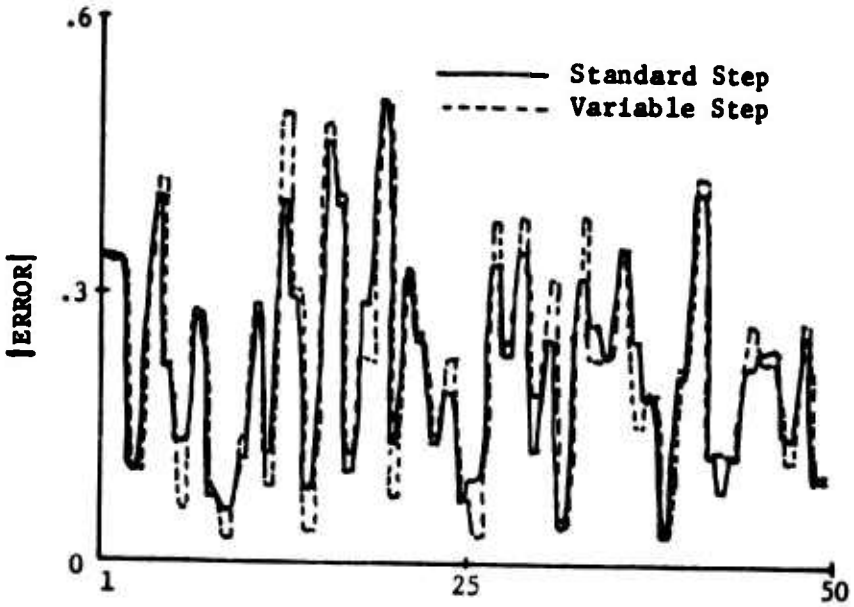
ITERATION STOPPING CRITERIA	TOTAL ITERATIONS FOR 50 DATA SETS	
	STANDARD STEP	VARIABLE STEP
1×10^{-7}	5732	158
1×10^{-6}	2657	158
1×10^{-5}	859	141

Table 3.

ITERATION STOPPING CRITERIA	NUMBER OF ITERATIONS (50 DATA SETS)		
	$R = 2 \times 10^{9}$	$R = 2 \times 10^{2}$	$R = 2 \times 10^{-4}$
1×10^{-7}	102	127	128
1×10^{-6}	100	127	127
1×10^{-5}	63*	113	113

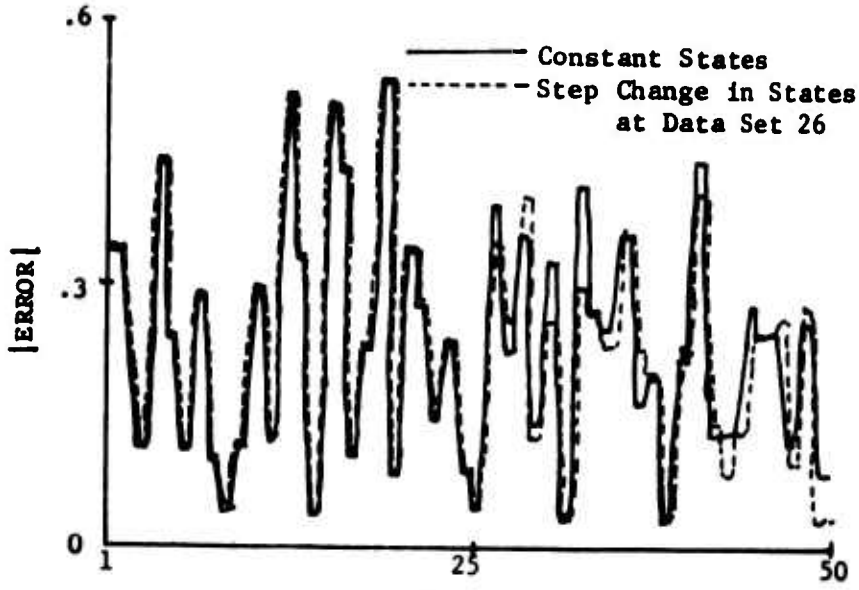
* Large Errors Present

Table 4.



NUMBER OF DATA SETS

FIGURE 6



NUMBER OF DATA SETS

FIGURE 7

INFORMATION PROCESSING FOR AN AIRCRAFT TERRAIN-FOLLOWING SYSTEM

Maj. James E. Funk

The purpose of this study was to investigate and improve the information processing in an aircraft terrain system. The input to the "system" was specific terrain data points, while the output was the aircraft flight path, from a non-linear digital simulation. Specifically, the question: "Is it feasible to incorporate an optimal-path-determination scheme into a real-time terrain-follower?" was to be answered. The information processing was divided into four phases:

1) terrain model construction from discrete data points,
2) "clearance path" determination for a specified minimum clearance distance above the terrain,
3) "optimal" or "ideal" flight path determination considering aircraft acceleration constraints and the the desire to fly as close to the terrain as possible without descending below the "clearance path", and
4) design of a tracking system, or proper inputs to the aircraft control system, to make the aircraft fly very near the determined "ideal path".

The source of the terrain data points is not specified for this study. Usually a forward-looking radar system would generate the points, but conceivably other sources are also possible-such as satellite acquired data. The preprocessing of any radar data to obtain the terrain points is beyond the scope of this study. Certainly it is a vital link in the total system, but the stress here is on deciding what path the aircraft should fly assuming reliable terrain information is available. Thus, effects such as "shaddowing", which occurs because the forward-looking radar cannot see the back sides of hills, are not considered directly. However, any terrain following system using a forward-looking radar will be subject to the same practical limitations. So the concern here is for the best response to the available terrain data.

The design concept stressed in study was to consider the physical system and problem with as many of its actual constraints and limitations as possible. Then the mathematical problem was formulated with those in mind, so that the solution to the mathematical problem would apply directly to the physical problem. This is somewhat opposed to the more usual approach of setting up a mathematical problem which can be readily solved and then attempting to adapt its solution to a physical problem with many more limitations.

Current operational terrain-following systems generate pitch commands to the aircraft control system based on essentially a single critical point for terrain ahead of the aircraft. The method of determining the critical point differs with different systems. In this study the attempt is to use all available terrain points insofar as possible. This concept agrees with that used in a system proposed and tested by General Electric [1,2]; however, the information processing is considerably different. One would expect systems based on all terrain points to achieve a greater degree of "closeness" to the terrain than those based on a single point. Results indicate that this is true While this study does not include comparisons of different types of systems, the ideal paths obtained are such that there could probably be little, if any, significant improvement in the aircraft flight path as obtained by any other system. One significant place for improvement in the overall system would be the addition of direct-lift to the aircraft. This could substantially improve the task of tracking the ideal path.

A continuous terrain model or curve was desired for clearance computations. The slant distance to the terrain was considered a more reasonable criteria for safety, rather than just the vertical clearance distance. A cubic spline was selected to represent a smooth terrain curve. The cubic spline consists of piecewise-cubic-polynomial segments there are joined at each of the "KNOTS" between two segments such that the first and second deriviatives of the two segments are equal, respectively, at that point. The knot positions were the terrain data points, so that the terrain curve passed through each data point, as illustrated in Fig. 1.

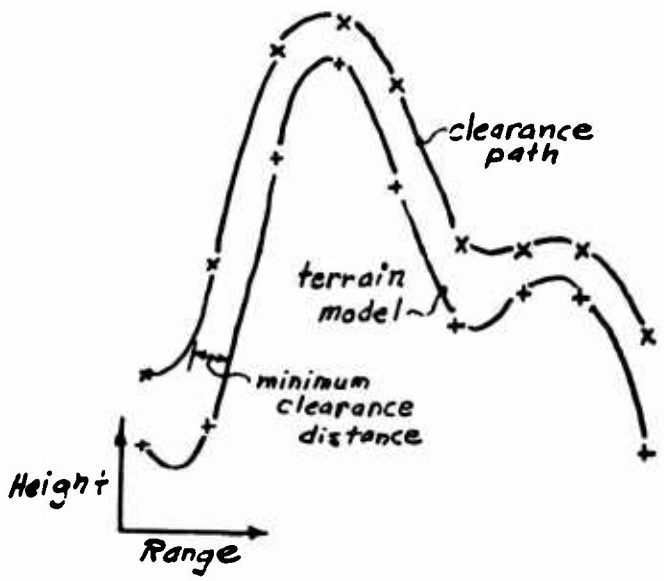


Fig. 1

The clearance path was also selected as a cubic spline, fitted through points vertically above the terrain data points. The clearance points were calculated by an iterative routine to allow slant clearance equal to the minimum clearance distance. This path is also illustrated in Fig 1.

The attempt to include as many physical constraints as necessary in the mathematical formulation created a more difficult optimization problem. The primary and most difficult constraint was that the optimal flight path should not be lower at any point than the specified minimum-clearance distance above the terrain. Restrictions on the aircraft normal-acceleration are also placed on all terrain following systems [3]. This is primarily for pilot comfort, but can also be performance and structural limitations on the aircraft. The final constraint added to the formulation was on the behavior of command signals to the aircraft control system. They should be reasonably smooth and somewhat limited in frequency, so that the aircraft will respond well.

The clearance-distance constraint appears mathematically as a state-variable inequality constraint (SVIC). There has not been very much success with general optimization algorithms that attempt to apply the SVIC's "directly". Most direct methods require some assumptions of how often the constraints are active or inactive [4, 5]. The greatest success with SVIC's has been in treating them indirectly with penalty functions. [6]. That approach was orginally tried in this study. A Davon-Rank-One parameter optimization scheme with an exterior penalty function was used with reasonable results. After some revisions to the optimization problem it was found that it could be stated as a quadratic programming problem. That currently appears to be the most promising approach, although no results from that are available yet. Direct methods for the quadratic programming problem with inequality constraints are readily available [7].

The normal acceleration constraints were approximated by "control constraints" which can be applied directly by most optimization methods. The control variable for the ideal path determination was second derivative of the altitude with respect to range— referred to hereafter as "curvature". The class of allowable control variables was purposely restricted to achieve smoothness and frequency limitations in the ideal path, since it should be a path which the aircraft can follow. The class of controls was selected as linear splines (continuous piecewise-linear functions of range), illustrated in Fig 2. Upper and lower bounds on the control values tend to limit aircraft accelerations. And the combination of these bounds with limits on the number of linear segments in the control produces frequency restrictions, as well. Double integration of this constrained curvature produces a path with continuous first and second derivatives with respect to range and with limited frequency content.

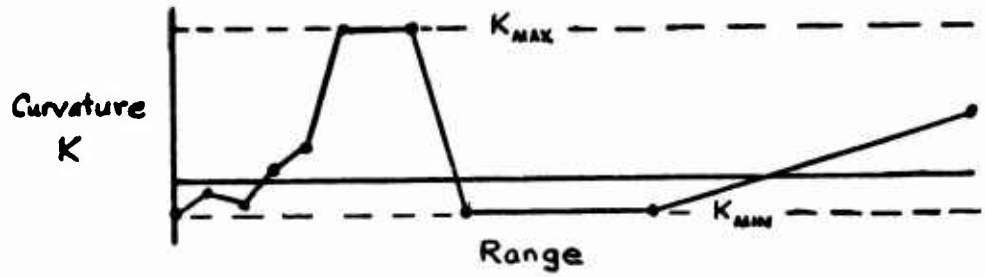


Fig. 2

The choice of a linear-spline control variable reduces the dimensionality of the control problem from a function space search to a parameter space search. The ideal path model was chosen to be as simple as possible and still contain the most significant restrictions associated with the physical problem. The equations of motion were simply:

(1) $h'(R) = s(R)$

(2) $s'(R) = K(R)$

where h is the height,

s is the slope,

and R is the range.

The restrictions on the curvature and height were, for all R:

(3) $K_{min} \leq K(R) \leq K_{max}$

(4) $h(R) \geq c(R)$, the clearance path value

The performance, or measure of closeness, was based on the elevation, or error, above the clearance path.

(5) $e(R) = h(R) - c(R)$

Using the linear spline control form, eqs. 1 and 2 can be integrated in closed form to yield an expression for e that is an explicit function of the K_i.

Originally an integral cost function was used, but this was subsequently simplified to a summation of the effects at discrete points along the terrain, R_n. J of eq. 6 was to be minimized, by optimal selection of the $K_i = K(R_i)$ parameters.

$$(6) \qquad J = \sum_{n=1}^{S} (Q_n e_n^2 + L_n e_n) + M$$

The quadratic term penalized the peak errors more heavily while the last term allowed for the possibility of a min-max criteria with the added constraint:

$$(7) \qquad e_n = e(R) \leq m \quad \text{For all n.}$$

When constraints (3) and (4) are applied only at the R_n points the optimization problem reduces to a quadratic programming problem. Since e_n is a linear combination of the K_i values, J becomes quadratic in the K_i parameters. The spacing of the control parameters (R_i) was not the same as for the performance summation (R_n). It was found the number of parameters could be effectively reduced by using a fine spacing near the aircraft where only the low frequency terrain components are of concern.

A typical ideal path is shown in Fig 3 for a fairly rough terrain segment. The curvature bounds corresponded to +3g and -1.5g accerelations. The path was generated by repetitive optimizations - optimizing a path segment, or frame, 24000 ft ahead of the aircraft, position then repeating the optimization for a new 24000 ft frame after the aircraft position moved 4000 feet down range. The control parameter values were at ranges of 1000, 2000, 3000, 4000, 6000, 8000, 14000, and 24000 feet from the start of each frame.

The tracking problem was evaluated using a non-linear digital simulation of an F-4C aircraft with its SAS (Stability Augmentation System) included. The SAS provided attitude response improvement for the aircraft, but additional control commands were necessary to obtain desireable positional response of the aircraft. The only linearizations used in the simulation were the aerodynamic force coefficients and a portion of the control-surface-actuator model. However, control surface deflection limits were considered. The aerodynamic force coefficients were used because non-linear aerodynamic data was not available.

Initially, the tracking systems considered maintained a constant aircraft thrust and allowed the speed of the aircraft to vary during the maneuvers. This approach was used because extensive throttle cycling usually reduces engine life. But tracking ability suffered severe degradation for rough terrain and "rough ride" (relatively high accelerations).

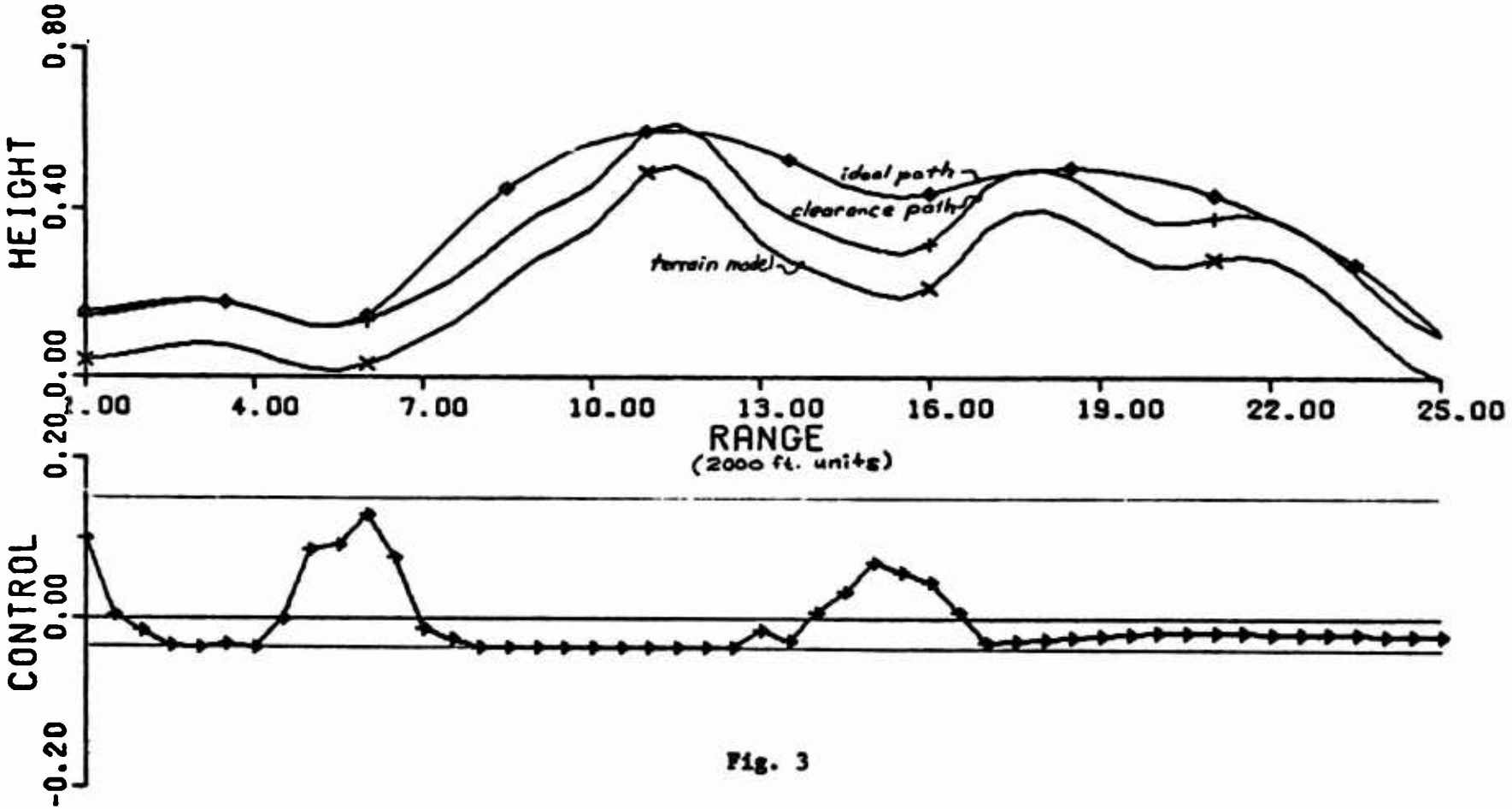

HEIGHT
0.80
0.40
0.00
ideal path
clearance path
terrain model
1.00
4.00
7.00
10.00
13.00
16.00
19.00
22.00
25.00
RANGE
(2000 ft. units)
CONTROL
0.20
0.00
-0.20

Fig. 3

The tracking control system under current investigation uses a command signal which is the sum of signals proportional to the errors in the heights, slopes, and curvatures. The errors are determined by subtracting the ideal path values from the simulated aircraft values. This gives a control system which is predictive as well as a feedback system. As was expected the tracker does not do well in a open loop mode of operation. The tracking system portion of the study is not yet complete. Convention linear control frequency response techniques which were tried have not proven helpful in this tracking problem because the non-linear effects are significant, especially for rough terrains.

From the work that has been completed it appears that a real-time, optimal-path system is feasible for terrain following, though addition verification of this must still be completed, particularly in the tracking phase. The penalty-function FORTRAN optimization routine can run in approximately real time on the CDC 6600 (5 seconds per frame). The quadratic-programming optimization should reduce computation time, which is the significant factor in adapting the scheme to an airborne computer. Program size and reduced word length would not appear to be significant problems in adaption to an airborne computer.

REFERENCES

1. Quinlivan, Richard P., "Terrain Following Controller Flight Test, 1st Tech. Rpt.", AFFDL-TR-67-114, Part I, June 1967.

2. Freedman, R.J., "Terrain Following Controller Flight Test Summary Report", AFFDL-TR-67-114, Part IV, Sept 1967.

3. Handbook for Low-Altitude Penetration Terrain-Following System Criteria, Boeing No 81205, Boeing, Wichita, (to be published).

4. Denham, W.F. & A.E. Bryson, "Optimal Programming Problems with Inequality Constraints, II: Solution by Steepest-Ascent."

5. Hennig, George R., "A Numerical Method for the Solution of Optimal Control Problems with Bounded State Variables." Dissertation, AF Institute of Technology, WPAFB, Ohio DS/EE/71-1, Feb 1971.

6. Beltrami, Edward J., An Algorithmic Approach to Nonlinear Analysis and Optimization, Academic Press, N.Y., 1970.

7. Ravinduan, A., "Computational Routine for Quadratic & Linear Programming Problems.", Comm. of A.C.M., Vol. 15, No. 9, pp 818-820, Sept 72.

AIR FORCE FLIGHT DYNAMICS LABORATORY

Finley Barfield
Robert Denaro, 1Lt
Randall V. Gressang, Capt
John E. Houtz
Robert Huber
Daniel L. Kugel
Bohdan G. Kunciw, 1Lt
George C. Perley
R. D. Poyneer
Robert Quaglieri
E. Rachovitsky
James E. Rader, Lt
Charles A. Scolatti, Col
David J. South, Major
Robert J. Woodcock

AIR FORCE AVIONICS LABORATORY

Ralph S. Bryan
Arthur A. Duke
Joe Hoffman
Richard A. Kitzerow
R. A. Manske
Stanton H. Musick
Robert A. Phelps
Richard Reeves
J. Gary Reid, Capt
Jesse C. Ryles
R. E. Salters, Capt
John Spicer
Raymond P. Wasky
D. H. Watjen, Lt
Thomas K. Wu

AIR FORCE AIR PROPULSION LABORATORY

Jon Dell

AIR FORCE ARMAMENT TEST LABORATORY

James E. Gibson
Michael L. Mumford, Lt

AIR FORCE CAMBRIDGE RESEARCH LABORATORY

Bernard S. Morgan, Col

AEROSPACE RESEARCH LABORATORY

David R. Audley, Lt
Jon Lee
Kenneth A. Myers, Capt
Robert W. Milling, Col

AIR FORCE WEAPONS LABORATORY

Craig E. Miller, Capt
James Negro, Lt

AEROSPACE MEDICAL RESEARCH LABORATORY

Mark W. Cannon
J. W. Carl, Capt
C. B. Harrah, Major
P. D. Houck, Lt
Lee A. Kjesbo, Lt
Don Monk
John A. Neff, Capt

AIR FORCE MATERIALS LABORATORY

Thomas J. Mallets, Major

AIR FORCE INSTITUTE OF TECHNOLOGY

Gerald M. Anderson, Major
Richard L. Bloom, Capt
D. W. Breuer
John Cunningham, Capt
John J. Dazzo
Jim Dillow, Major
James E. Funk, Major
Constantine H. Houpis
Peter G. Jenkins
Jerry D. Kendrick, Major
Leonard R. Kruczynski, Capt
Gary B. Lamont
Eric K. Lindberg, Capt
Robert A. K. Mitchell
Paul M. J. Moore
Thomas E. Moriarity, Capt
Jerry L. Mrozek, Capt
David N. Olmstead
Richard M. Potter, Capt
Robert D. Powell, Capt
Eugene A. Smith, Capt
Vernon P. Saxon, Capt
G. W. Titmas, Capt
Joseph Widhalm, Capt

AERONAUTICAL SYSTEMS DIVISION

James L. Blair
Albert I. Chatmon, Capt
Andrew Ceranowicz
Dan E. Flory
Felix O. Grover
Alan I. Kirschbaum, Lt
Edward A. Martin
Henry J. McGlynn, Capt
Eric B. Nelson, Lt Col
Richard W. O'Dell
Ronald T. Ooten
Jon Roll
Wylie E. Rudhman
Stanley A. Sneegasm, Lt
Donald M. Sovine
Marc S. Weinberg, Lt

AIR FORCE SPECIAL WEAPONS CENTER

George F. Forman
David J. Kempi, Major
E. J. Pollock
Erik M. Stolle, Capt

FRANK J. SEILER RESEARCH LABORATORY

Robert R. Asher, Capt
Kenneth D. Herring, Major
Ronald E. Janosko, Capt
Marc L. Sabin, Capt

FOREIGN TECHNOLOGY DIVISION

Mel Baker
Kenneth F. Bornhorst
Walter B. Boley
John E. Brothers
John E. Carles
Mike Custodio
David W. Diggle, Capt
I. W. Garnto
Mike Gural
Philip F. Honsberger
Lou H. Jackson, Lt
William Kovacs, II
Delmar D. Notestine
David Richards
James E. Ruth
Thomas J. Smith, Major
Joe Swingle

NAVAL WEAPONS CENTER

J. Hledik
John Shepherd

OFFICE OF NAVAL RESEARCH

Stuart L. Brodsky

6585th TEST GROUP

Austin L. Foote, Major
Frazier J. Hellings, Lt Col
Dittmar Kittler, Capt

SPACE AND MISSILE SYSTEMS ORGANIZATION

James A. Davis, Lt
W. F. Galbraith
Thomas S. Lee, Capt
David A. Nichols, Lt
Peter Reischl, Lt

UNITED STATES AIR FORCE ACADEMY

Edward J. Bauman, Lt Col
Daniel E. Busse, Capt
Harold L. Daniel, Capt
John C. Durrett, Major
Thomas J. Eller, Major
Ronald P. Fuchs, Capt
Jack C. Henry, Capt
Anthony L. Leatham, Major
Max I. Miller, Jr., Major

UNITED STATES AIR FORCE SCHOOL OF AERO MEDICINE

Harry M. Hughes

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Walter J. Rabe, Lt Col

*U.S.Government Printing Office: 1974 — 657-015/214