

AD-758 145

PROCEEDINGS OF THE 1972 ARMY NUMERICAL  
ANALYSIS CONFERENCE SPONSORED BY THE  
ARMY MATHEMATICS STEERING COMMITTEE  
HELD AT BIOMEDICAL LABORATORY, EDGEWOOD  
ARSENAL, MARYLAND ON 20-21 APRIL 1972

Army Research Office  
Durham, North Carolina

October 1972

DISTRIBUTED BY:

**NTIS**

National Technical Information Service  
U. S. DEPARTMENT OF COMMERCE  
5285 Port Royal Road, Springfield Va. 22151

# PROCEEDINGS OF THE 1972 ARMY NUMERICAL ANALYSIS CONFERENCE

AD 758145



DDC  
RECEIVED  
APR 11 1973  
C

Approved for public release; distribution unlimited.  
The findings in this report are not to be construed  
as an official Department of the Army position, un-  
less so designated by other authorized documents.

Reproduced by  
NATIONAL TECHNICAL  
INFORMATION SERVICE  
U S Department of Commerce  
Springfield VA 22151

SPONSORED BY  
THE ARMY MATHEMATICS STEERING COMMITTEE ON BEHALF OF

THE OFFICE OF  
THE CHIEF OF RESEARCH AND DEVELOPMENT

735

U. S. Army Research Office-Durham

Report 72-3

October 1972

PROCEEDINGS OF THE 1972 ARMY NUMERICAL

ANALYSIS CONFERENCE

Sponsored by the Army Mathematics Steering Committee

Host

Biomedical Laboratory

Edgewood Arsenal

Maryland

Approved for public release; distribution unlimited. The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

U. S. Army Research Office-Durham  
Box CM, Duke Station  
Durham, North Carolina

## FOREWORD

Thirteen years ago the Office of Ordnance Research (now the Army Research Office-Durham) organized an OOR Liaason Group on Computers. Two meetings of this group were held, one in 1959 and the other in 1960, to exchange information of interest to managers of ordnance computers. The Army Mathematics Steering Committee decided that these meetings should be revived on an Army-wide basis, and asked Dr. John H. Giese, Chairman of its subcommittee on Numerical Analysis and Digital Computers, to draw up a format for, and take charge of, the new series of conferences. Dr. Giese thought these meetings should "establish a way to exchange ideas on the Army's desires, capabilities, and interest in the field of 'other-than business' applications of computers"; and they should provide the AMSC and ARO with information on the Army's needs for computers, requirements for assistance in research and numerical analysis and other kinds of mathematics. He also suggested that the title of the conferences should be the "ARO Working Group on Computers". Two meetings, one in 1962 and the other in 1964, were held under this title. Starting in 1965 these conferences have been held yearly under the title "Army Numerical Analysis Conference".

Dr. Giese has served as chairman of all these conferences. Members of the subcommittee on Numerical Analysis and Digital Computers have assisted him on some of the planning details of the meetings. However, most of the responsibilities of the arrangements were in his hands. Thanks to his continuing efforts, all of the meetings have been held at a high scientific level. Speakers and attendees at these conferences



would like to show their appreciation for all of your efforts, John, by dedicating these Proceedings to you. They are sorry that you will no longer be serving as chairman, but they do feel you have done more than your share of work in promoting these conferences. We certainly hope you will continue to participate in future conferences in this area.

The theme of the 1972 Army Numerical Analysis Conference was Systems Identification. This meeting was held on 20-21 April 1972 at the Biomedical Laboratory at Edgewood Arsenal, Maryland. Dr. William J. Sacco served as Chairman on Local Arrangements. All those in attendance are indebted to him for a well-planned conference and for selection of some of the invited speakers.

The Army Mathematics Steering Committee, the sponsor of those conferences, has asked that these Proceedings be issued to Army scientists and to others interested in the science and application of numerical analysis to applied problems. Members of this committee would like to extend their thanks to the speakers for their interesting papers, and to the chairmen and all others who participated in the conduction of this meeting.

## TABLE OF CONTENTS

Title	Page
Foreword . . . . .	iii
Table of Contents. . . . .	v
Program. . . . .	vii
Initial Value Methods for Nonlinear Boundary Value Problems and Integral Equations Robert Kalaba . . . . .	1
Optoelectronic Computational Techniques for Fast Picture Processing Jenny Bramley . . . . .	15
An Algorithm for Rapidly Parsing Automatically Digitized Drawings Charles M. Williams . . . . .	27
Evaluation of the Roots of Cross-Product Bessel Functions Shih-Chi Chu and Philip D. Benzkofer. . . . .	43
Application of Pattern Recognition to Shock-Trauma Studies W. S. Copes, A. R. Cowley C. Masaitis, W. J. Sacco and A. V. Milholland. . . . .	53
Aerodynamic Parameter Identification in Ballistic Range Tests Gary T. Chapman . . . . .	73
The WSMR Best Estimate of Trajectory - An Overview William S. Agee and Robert H. Turner. . . . .	111
A Computer Program to Investigate Exo-Atmospheric Engagements of Interceptors and Re-Entry Vehicles LTC M. L. Roberson, CPT C. Van Nostrand and W. A. Barbieri. . . . .	129

5734

<b>The Wind Tunnel Free Flight Testing Technique</b>	
A. S. Platou . . . . .	155
<b>Some Aspects of Regularization and Approximation of Solutions of Ill-Posed Operator Equations</b>	
M. Z. Nashed . . . . .	163
<b>Kinetics Analysis by Digital Parameter Estimation</b>	
Y. Bard and L. Lapidus . . . . .	183
<b>List of Attendees . . . . .</b>	<b>229</b>

A G E N D A

1972 ARMY NUMERICAL ANALYSIS CONFERENCE  
Biomedical Laboratory, Edgewood Arsenal, Maryland

Thursday, 20 April 1972

- 0830-0900 REGISTRATION - Biomedical Laboratory, Building 3100
- 0900-0920 OPENING OF THE CONFERENCE - Room 14, Building 3100
- Calling of Conference to Order  
Dr. William J. Sacco, Chairman on Local Arrangements,  
Biomedical Laboratory, Edgewood Arsenal, Maryland
- Welcoming Remarks  
Colonel J. R. Blair, Director of the Biomedical Laboratory,  
Edgewood Arsenal, Maryland
- 0920-1020 GENERAL SESSION 1 - Room 14
- Chairman: Dr. F. Heinmets, Biophysics Group, Pioneering  
Research Laboratory, US Army Natick Laboratories,  
Natick, Massachusetts
- SYSTEM IDENTIFICATION FOR PHYSIOLOGICAL PROCESSES  
Dr. Robert Kalaba, University of Southern California,  
Department of Electrical Engineering, University Park,  
Los Angeles, California
- 1020-1050 BREAK
- 1050-1200 TECHNICAL SESSION 1 - Room 14
- Chairman: Joseph S. Tyler, Biomedical Laboratory, Edgewood  
Arsenal, Maryland
- OPTOELECTRONIC COMPUTATIONAL TECHNIQUES FOR FAST PICTURE  
PROCESSING  
Jenny Bramley, Geographic Information Systems Division,  
US Army Engineer Topographic Laboratories, Fort Belvoir,  
Virginia
- AN ALGORITHM FOR RAPIDLY PARSING AUTOMATICALLY DIGITIZED  
DRAWINGS  
C. M. Williams, Visicon Inc., State College, Pennsylvania  
Representing: Directorate of Management Information  
Systems, HQ AMC
- 1200-1300 LUNCH

1300-1430 TECHNICAL SESSION 2 - Room 14

Chairman: Alan S. Galbraith, Mathematics Division, US Army  
Research Office-Durham, Durham, North Carolina

THE DYNAMIC STABILITY OF STIFFENED CYLINDRICAL SHELLS UNDER  
AXIAL IMPULSE

C. Lakshmikantham and Tien-yu Tsui, Army Materials and  
Mechanics Research Center, Watertown, Massachusetts

NONSTEADY FREE CONVECTION AND RADIATION AROUND GUN TUBES

Rao V. S. Yalamanchili and Steve Bostwick, Research  
Directorate, Weapons Laboratory at Rock Island, US Army  
Weapons Command, Rock Island, Illinois

EVALUATION AND APPLICATION OF THE ROOTS OF CROSS-PRODUCT  
BESSEL FUNCTIONS

Shih-Chi Chu and Philip D. Benzkofer, Research, Development  
and Engineering Directorate, US Army Weapons Command,  
Rock Island, Illinois

1430-1440 SHORT BREAK

1440-1530 TECHNICAL SESSION 2 (Continued)

APPLICATION OF PATTERN RECOGNITION TO TRAUMA STUDIES

W. Copes, AMSAA, Aberdeen Proving Ground, Maryland;  
A. Cowley, University of Maryland Hospital; C. Masaitis,  
Ballistics Research Laboratory, Aberdeen Proving Ground,  
Maryland; A. Milholland and W. Sacco, Biomedical Laboratory,  
Edgewood Arsenal, Maryland

1530-1600 BREAK

1600-1700 GENERAL SESSION 2 - Room 14

Chairman: Dr. Louis B. Rall, Mathematics Research Center,  
The University of Wisconsin, Madison, Wisconsin

AERODYNAMIC PARAMETER IDENTIFICATION IN BALLISTIC RANGE TESTING

Dr. G. T. Chapman, Ames Research Center, NASA, Moffett Field,  
California

\*\*\*\*\*

Friday, 21 April 1972

0830-1010 TECHNICAL SESSION 3 - Room 14

Chairman: Norman P. Coleman, Systems Research Division,  
Headquarters, U. S. Army Weapons Command, Rock  
Island, Illinois

0830-1010 TECHNICAL SESSION 3 (Continued)

THE WSMR BEST ESTIMATE OF TRAJECTORY - AN OVERVIEW

William S. Agee and Robert H. Turner, Analysis and Computation  
Division, National Range Operations Directorate, White Sands  
Missile Range, New Mexico

A COMPUTER PROGRAM TO INVESTIGATE EXO-ATMOSPHERIC ENGAGEMENTS  
OF INTERCEPTORS AND RE-ENTRY VEHICLES

LTC M. L. Roberson, CPT C. Van Nostrand and W. A. Barbieri,  
Development Branch ACS/Studies and Analysis, HQ, US Air  
Force, Washington, D. C.

SYSTEM PARAMETER ESTIMATION FOR NON-LINEAR SYSTEMS

Richard L. Moore, US Army Weapons Command, Rock Island  
Arsenal, Rock Island, Illinois

1010-1040 BREAK

1040-1200 TECHNICAL SESSION 4 - Room 14

Chairman: CPT Clarence W. Kitchens, Jr., Ballistics  
Research Laboratories, US Army Aberdeen  
Research and Development Center, Aberdeen  
Proving Ground, Maryland

REDUCTION OF FREE FLIGHT WIND TUNNEL DATA

A. S. Platou, Ballistics Research Laboratory, Aberdeen  
Research and Development Center, Aberdeen Proving Ground,  
Maryland

SOME ASPECTS OF REGULARIZATION AND APPROXIMATION OF  
PSEUDOSOLUTIONS OF ILL-POSED OPERATOR EQUATIONS

M. Z. Nashed, Mathematics Research Center, University  
of Wisconsin, Madison, Wisconsin

1200-1300 LUNCH

1300-1400 GENERAL SESSION 3 - Room 14

Chairman: Major A. A. Mullin, Office, Chief of Research and  
Development, Mathematics Branch, Department of the  
Army, Washington, D. C.

PARAMETER ESTIMATION TECHNIQUES IN CHEMICAL REACTION SYSTEMS

Professor Leon Lapidus, Chemical Engineering Department,  
Princeton University, Princeton, New Jersey

# INITIAL VALUE METHODS FOR NONLINEAR BOUNDARY VALUE PROBLEMS AND INTEGRAL EQUATIONS

Robert Kalaba  
Biomedical Engineering Program  
Department of Electrical Engineering  
University of Southern California  
Los Angeles, California

## SUMMARY

A technique has been developed for transforming nonlinear boundary value problems and integral equations into Cauchy systems. This provides an analytical approach to nonlinear problems which is different from the usual successive approximation and series expansion schemes. It is also significant computationally, for modern analog and digital computing machines can solve initial value problems with considerable speed and accuracy. There are implications for stochastic nonlinear equations.

Applications of this new approach in biology, physics, and engineering, both analytically and computationally, are sketched.

## 1. Introduction

During the early 1950's I recognized that computing machines would be able to solve large systems of nonlinear ordinary differential equations, provided that a complete set of initial conditions is known. The study of some physical systems leads directly to such initial value problems; the study of others does not. Clearly, an important task would be the transformation of integral equations and boundary value problems into initial value problems to take advantage of this great new computational ability. There were two early hints that this could be done: Ambarzumian showed that the reflecting properties of a slab could be found without first determining the entire internal field [1], and Davidenko [2] showed that nonlinear transcendental equations could be reduced to initial value problems.

In a long series of papers [3-8] my colleagues and I have shown how to transform many important integral equations and boundary value problems of applied mathematics into Cauchy systems. These ideas have been productive both computationally and analytically. For the most part these systematic earlier considerations have been for linear systems, though there have been some exceptions [9].

In recent months we have found general methods for converting nonlinear boundary value problems and nonlinear integral equations into Cauchy systems. No use of the usual successive approximation or series expansion techniques is made. Let us now take up a special case to indicate the approach. Then in § 3 we cover numerical aspects. Next stochastic equations and nonlinear integral equations are treated. Initial value problems in one parameter are transformed into an initial value problem in another in §6. A broad program of applications in biology, physics and engineering is presented in §7.

---

The remainder of this paper has been reproduced photographically from the author's manuscript.



## 2. A Nonlinear Boundary Value Problem [10]

To illustrate the reduction of a nonlinear boundary value problem to a Cauchy system we consider the problem

$$(1) \quad \ddot{u}(t) = \lambda f(u(t)), \quad 0 \leq t \leq 1,$$

$$(2) \quad u(0) = u(1) = 0,$$

and assume that a unique solution exists for  $0 \leq \lambda \leq \Lambda$ . As usual, the dots over a variable indicate differentiation with respect to  $t$ . Since the solution  $u$  is a function of  $\lambda$ , as well as  $t$ , we shall write

$$(3) \quad u = u(t, \lambda), \quad 0 \leq t \leq 1, \\ 0 \leq \lambda \leq \Lambda.$$

Equations (1) and (2) become, in this expanded notation,

$$(4) \quad \ddot{u}(t, \lambda) = \lambda f(u(t, \lambda)), \quad 0 \leq t \leq 1,$$

$$(5) \quad u(0, \lambda) = u(1, \lambda) = 0.$$

Assuming appropriate differentiability properties we find that the function  $u_\lambda$  satisfies the linear boundary value problem

$$(6) \quad [u_\lambda(t, \lambda)]'' = f(u(t, \lambda)) + \lambda f'(u(t, \lambda))u_\lambda(t, \lambda), \\ 0 \leq t \leq 1,$$

$$(7) \quad u_\lambda(0, \lambda) = u_\lambda(1, \lambda) = 0,$$

where, as usual, the subscript denotes a partial derivative with respect to  $\lambda$ .

To solve equations (6) and (7) for  $u_\lambda$  consider the function  $w$ , the solution of the linear problem



$$(8) \quad \ddot{w}(t, \lambda) = g(t, \lambda) + \lambda f'(u(t, \lambda)) w(t, \lambda) ,$$

$$0 \leq t \leq 1 ,$$

$$(9) \quad w(0, \lambda) = w(1, \lambda) = 0 .$$

In terms of an appropriate Green's function  $G$ , for an arbitrary forcing function  $g$  the function  $w$  is given as

$$(10) \quad w(t, \lambda) = \int_0^1 G(t, y', \lambda) g(y', \lambda) dy' ,$$

$$0 \leq t \leq 1 ,$$

$$0 \leq \lambda \leq \Lambda .$$

It follows that the function  $u_\lambda$  may be represented in the form

$$(11) \quad u_\lambda(t, \lambda) = \int_0^1 G(t, y', \lambda) f(u(y', \lambda)) dy' ,$$

$$0 \leq t \leq 1 ,$$

$$0 \leq \lambda \leq \Lambda .$$

This is viewed as a differential equation for the function  $u$ , the independent variable being  $\lambda$ . The initial condition at  $\lambda = 0$  is

$$(12) \quad u(t, 0) = 0, \quad 0 \leq t \leq 1 ,$$

according to equations (4) and (5).

Next we obtain a differential equation and an initial condition for the Green's function  $G$ . From equation (10) we notice that

$$(13) \quad w_\lambda(t, \lambda) = \int_0^1 G_\lambda(t, y', \lambda) g(y', \lambda) dy'$$

$$+ \int_0^1 G(t, y', \lambda) g_\lambda(y', \lambda) dy' .$$

On the other hand, we obtain a two point boundary value problem for the function  $w_\lambda$  from equations (8) and (9). It is

$$\begin{aligned}
 (14) \quad [w_\lambda(t, \lambda)]'' &= g_\lambda(t, \lambda) + f'(u(t, \lambda))w(t, \lambda) \\
 &+ \lambda f''(u(t, \lambda)) u_\lambda(t, \lambda) w(t, \lambda) \\
 &+ \lambda f'(u(t, \lambda))w_\lambda(t, \lambda) \ ,
 \end{aligned}$$

$$(15) \quad w_\lambda(0, \lambda) = w_\lambda(1, \lambda) = 0 \ .$$

According to equations (8), (9) and (10) the solution of equations (14) and (15) is

$$\begin{aligned}
 (16) \quad w_\lambda(t, \lambda) &= \int_0^1 G(t, y', \lambda) [g_\lambda(y', \lambda) + (f'(u(y', \lambda)) \\
 &+ \lambda f''(u(y', \lambda)) u_\lambda(y', \lambda)) w(y', \lambda)] dy' \ .
 \end{aligned}$$

It is now convenient to introduce the auxiliary variable M,

$$\begin{aligned}
 (17) \quad M(t, \lambda) &= f'(u(t, \lambda)) + \lambda f''(u(t, \lambda))u_\lambda(t, \lambda) \\
 &= f'(u(t, \lambda)) + \lambda f''(u(t, \lambda)) \int_0^1 G(t, y', \lambda) f(u(y', \lambda)) dy' \ , \\
 &0 \leq t \leq 1 \ , \\
 &0 \leq \lambda \leq \Lambda \ .
 \end{aligned}$$

Equation (16) then becomes

$$\begin{aligned}
 (18) \quad w_\lambda(t, \lambda) &= \int_0^1 G(t, y', \lambda) g_\lambda(y', \lambda) dy' \\
 &+ \int_0^1 G(t, y', \lambda) M(y', \lambda) w(y', \lambda) dy' \\
 &= \int_0^1 G(t, y', \lambda) g_\lambda(y', \lambda) dy' \\
 &+ \int_0^1 G(t, y', \lambda) M(y', \lambda) \int_0^1 G(y', y, \lambda) g(y, \lambda) dy dy' \ .
 \end{aligned}$$

In view of the two representations for the function  $w_\lambda$  in equations (13) and (18) and the arbitrariness of the function  $g$ , we see that

$$(19) \quad G_\lambda(t, y, \lambda) = \int_0^1 G(t, y', \lambda) M(y', \lambda) G(y', y, \lambda) dy' ,$$

$$0 \leq t, y \leq 1 ,$$

$$0 \leq \lambda \leq \Lambda ,$$

where the function  $M$  is given in equation (17). The initial condition on the Green's function  $G$  at  $\lambda = 0$  is known to be

$$(20) \quad G(t, y, 0) = \begin{cases} y(t-1), & 0 \leq y \leq t , \\ t(y-1), & t \leq y \leq 1 . \end{cases}$$

The desired Cauchy system for the functions  $u$  and  $G$  consists of the differential equations in equations (11) and (19), the auxiliary relation (17) for the variable  $M$ , and the initial conditions in equations (12) and (20).

It is a straightforward matter to establish that a solution of the Cauchy system provides a solution of the original two point boundary value problem.

### 3. Numerical Aspects

Based on much previous experience [3, 4, 8] we believe that the method of lines [12] provides an effective approach to the numerical solution of the Cauchy system just given. The basic idea is to approximate the integrals on the interval  $(0, 1)$  by means of a quadrature formula. In that way the differential-integral equations are approximated by a system of ordinary differential equations for which the independent variable is  $\lambda$ . Since a complete set of initial values for  $u$  and  $G$  known, at  $\lambda = 0$ , the original boundary value problem is reduced to a system of ordinary differential equations subject to known initial conditions. Modern digital, analog and hybrid computers are well-

suited for this task. We routinely integrate systems of order  $10^3$  or so in the year 1972.

Let us use the approximation

$$(1) \quad \int_0^1 f(y') dy' \cong \sum_{i=1}^N f(r_i) w_i .$$

Then equation (11) of the previous section is approximated by the ordinary differential equations

$$(2) \quad \frac{du_i(\lambda)}{d\lambda} = \sum_{j=1}^N G_{ij}(\lambda) f(u_j(\lambda)) w_j ,$$

$$i = 1, 2, \dots, M ,$$

where

$$(3) \quad u_i(\lambda) = u(r_i, \lambda)$$

and

$$(4) \quad G_{ij}(\lambda) = G(r_i, r_j, \lambda) ,$$

$$i, j = 1, 2, \dots, N .$$

Equation (19) becomes

$$(5) \quad \frac{dG_{ij}(\lambda)}{d\lambda} = \sum_{m=1}^N G_{im}(\lambda) M_m(\lambda) G_{mj}(\lambda) w_m ,$$

$$i, j = 1, 2, \dots, N ,$$

in an obvious notation. Thus there are  $N^2 + N$  ordinary differential equations with evident initial conditions from equations (22) and (20). In addition the analogue of equation (17) is required.

We have done [10] trial computations with  $f(u) = \exp(u)$ . We approximated the integrals by using Simpson's rule with twenty intervals. The

resulting system of ordinary differential equations was integrated for  $0 \leq \lambda \leq 1$  and gave accuracy to within one part in five thousand. This demonstrates the computational feasibility of the method. In view of the known discontinuity in the derivative of  $G(t, y, \lambda)$  with respect to  $y$  at  $y = t$ , it would be desirable to find ways to make the computation as efficient as possible and to compare it against other standard methods such as quasilinearization [13]. Where the parameter study in  $\lambda$  is required, the efficiency of the proposed method is beyond dispute. Even if the solution of the nonlinear boundary value problem is desired for only  $\lambda = \Lambda$ , the proposed method is interesting analytically and possibly numerically, for no solving of linear algebraic equations is required.

#### 4. Nonlinear Stochastic Equations and Other Matters

In the previous section we have indicated how to produce numerically the function  $u(t, \lambda)$ ,  $0 \leq t \leq 1$ , for  $0 \leq \lambda \leq \Lambda$ , where  $u$  is the solution of the nonlinear two-point boundary value problem in equations (1) and (2). There are at least three advantages in being able to produce the function  $u$  for all of these values of  $\lambda$ . In the first place, it automatically provides a "parameter study" which is often required in engineering and biological applications.

Secondly, it provides a way of treating stochastic nonlinear boundary value problems. First determine  $u = u(t, \lambda)$  as above for  $0 \leq \lambda \leq \Lambda$ . Then suppose that  $\lambda$  is a random variable having the probability density function  $p = p(\lambda)$  for  $0 \leq \lambda \leq \Lambda$ . Let the  $m^{\text{th}}$  moment of  $u(t)$  be denoted by  $M_m(t)$ ,  $0 \leq t \leq 1$ . Then we have

$$(1) \quad M_m(t) = \int_0^{\Lambda} u^m(t, \lambda) p(\lambda) d\lambda, \quad 0 \leq t \leq 1,$$

which is readily evaluated numerically.

Thirdly, a uniform approach to system identification problems is provided. Suppose that  $b_1, b_2, \dots, b_R$  are observed values of  $u(t_i)$ , and we wish

to select  $\lambda$  so that we minimize

$$(2) \quad S(\lambda) = \sum_{i=1}^R [u(t_i, \lambda) - b_i]^2 ,$$

where equations (1) and (2) of 2 hold.

Transforming the boundary value problem into a Cauchy system, as we have explained, puts the problem in a form for which much is known [13]. It also makes possible the use of gradient techniques for effecting the minimization.

### 5. Nonlinear Integral Equations [14]

A broad and important class of nonlinear integral equations has the form

$$(1) \quad u(t) = g(t, \lambda) + \lambda \int_0^1 k(t, y, \lambda, u(y)) dy ,$$

$$0 \leq t \leq 1 .$$

The parameter  $\lambda$  may lie in an interval  $(0, \Lambda)$ . To emphasize the dependence of the unknown function  $u$  upon the parameter  $\lambda$ , as well as upon the variable  $t$ , we shall write

$$(2) \quad u(t, \lambda) = g(t, \lambda) + \lambda \int_0^1 k(t, y, \lambda, u(y, \lambda)) dy ,$$

$$0 \leq t \leq 1 ,$$

$$0 \leq \lambda \leq \Lambda .$$

Then, under suitable regularity properties, it is possible to demonstrate the equivalence between the nonlinear integral equation, Eq. (2), and the Cauchy system for  $u$  and the auxiliary function  $K$

$$(3) \quad u_\lambda(t, \lambda) = \Psi(t, \lambda) + \lambda \int_0^1 K(t, y', \lambda) \Psi(y', \lambda) dy' ,$$

$$(4) \quad K_\lambda(t, y, \lambda) = Q(t, y, \lambda) + \lambda \int_0^1 K(t, y', \lambda) Q(y', y, \lambda) dy' ,$$

$$0 \leq t, y \leq 1 ,$$

$$0 \leq \lambda \leq \Lambda ,$$

$$(5) \quad u(t, 0) = g(t, 0) ,$$

$$(6) \quad K(t, y, 0) = k_u(t, y, 0, g(y, 0)) ,$$

$$0 \leq t, y \leq 1 ,$$

where the functions  $\Psi$  and  $Q$  are certain functionals on  $u$  and  $K$ .

Preliminary numerical experiments in which  $g(t, \lambda) = 1 - (\lambda/2)t$  and  $k(t, y, \lambda, u) = tyu^2$  have shown the feasibility of the method. That  $\lambda = 3/2$  is a bifurcation point is obtained effortlessly, for the auxiliary function  $K$  becomes infinite there. How to continue the solution through such a point is a matter of great interest.

## 6. Initial Value Problems

Consider the Cauchy system

$$(1) \quad \dot{x}(t, \lambda) = f(x(t, \lambda), \lambda), \quad 0 \leq t \leq T$$

$$(2) \quad x(0, \lambda) = c,$$

where  $\lambda$  is a parameter lying in the interval  $(0, \Lambda)$ . Frequently we desire a parameter study in  $\lambda$  of the solution of the equations (1) and (2). One procedure, of course, is to solve the system as an initial value problem in  $t$  for various values of  $\lambda$ . However, there is an alternative: transform the system (1) and (2) into a Cauchy system in which  $\lambda$  becomes the time-like variable. Such a system is

$$(3) \quad x_\lambda(t, \lambda) = M(t, \lambda)$$

$$(4) \quad M(t, \lambda) = \int_0^t [\varphi(t, \lambda)/\varphi(y, \lambda)] f_\lambda(x(y, \lambda), \lambda) dy$$

$$(5) \quad x(t, 0) = g(t)$$

$$(6) \quad \varphi_\lambda(t, \lambda) = \int_0^t [\varphi(t, \lambda)/\varphi(y, \lambda)] \Psi(y, \lambda) dy$$

$$(7) \quad \Psi(t, \lambda) = f_{xx}(x(t, \lambda), \lambda)M(t, \lambda) \\ + f_{x\lambda}(x(t, \lambda), \lambda)\varphi(t, \lambda)$$

$$(8) \quad \varphi(t, 0) = h(t),$$

$$0 \leq \lambda \leq \Lambda,$$

$$0 \leq t \leq T.$$

The initial conditions at  $\lambda = 0$  in equations (5) and (8) are obtained by integrating equations (1) and (2) for  $\lambda = 0$  and by integrating the system

$$(9) \quad \dot{\varphi}(t, 0) = f_x(x(t, 0), 0)\varphi(t, 0), \quad 0 \leq t \leq T,$$

$$(10) \quad \varphi(0, 0) = 1.$$

Under some circumstances it might be preferable numerically to solve the system (3) - (10) rather than the system (1) and (2). This remains to be investigated.

There are similar discussions for the conversion of initial value problems for partial differential equations into Cauchy systems in which a selected parameter becomes the time-like variable.

## 7. Applications

I have in mind applications in physics, engineering, and biology. Electromagnetic theory and radiative transfer should be investigated. One



of the principal nonlinear integral equations of radiative transfer is

$$(1) \quad \varphi(\eta) = 1 + \frac{\lambda}{2} \eta \varphi(\eta) \int_0^1 \frac{\varphi(\xi)}{\eta + \xi} d\xi,$$

$$0 \leq \eta \leq 1,$$

$$0 \leq \lambda \leq 1.$$

A start on the study of this nonlinear integral equation using initial value methods has been made [15]; in fact, successful calculations have been performed, but much remains to be investigated. The behavior of the solution of the associated Cauchy system near  $\lambda = 1$ , a bifurcation point, is interesting.

The theory of optimal filtering, detection, and control abounds with integral equations, many of which are nonlinear [16]. These should be studied with emphasis on filtering of physiological data, as well as communication and radar signals. The detection of arrhythmias in coronary patients is a possibility.

The theory of thin shells of revolution [17] depends upon solving nonlinear systems of coupled integral equations. Here we have to derive the appropriate Cauchy systems and then do test calculations. This is an extension of our work on the linear integral equations of elasticity theory [18]. Applications to biomechanics should be stressed, especially the study of trauma due to a blow to the head.

Nonlinear integral equations are used to describe lateral inhibition in neural systems [19]. The dependent variable is a function of several spatial variables. Computational solution using initial value methods is a challenge.

Nonlinear boundary value problems and integral equations abound in the study of fluid and electrolyte transport in physiological systems [20, 21, 22].

Here the biological interpretation of the Cauchy system will be particularly interesting.

A nonlinear differential equation with nonlinear boundary conditions is treated in [23].

#### 8. Discussion

In the previous pages I have adumbrated a uniform approach to nonlinear boundary value problems and integral equations. I feel that it will become as effective for nonlinear problems, in this age of computing machines, as the eigenfunction expansion technique was for linear problems in pre-computer days. It possesses the great merits of being simple in concept, broad in application and effective in computation.

### References

1. V. Ambarzumian, "O Passeyanii Sveta Atmosferami Planet", Astronomicheskii Zhurnal, v. 19, no. 5 (1942), pp. 30-41.
2. D. Davidenko, "Ob odnom novom metode chislennovo resheniya sistem nelineinykh uravnenii", Doklady Akademii Nank SSR, v. 87, n. 4 (1953), pp. 601-602.
3. H. Kagiwada and R. Kalaba, "An Initial Value Method for Fredholm Integral Equations of Convolution Type", International J. Computer Mathematics, v. 2, n. 2 (Apr. 1968), pp. 143-155.
4. H. Kagiwada and R. Kalaba, "A New Initial Value Method for Internal Intensities in Radiative Transfer", The Astrophysical Journal, v. 147, no. 1 (Jan 1967), pp. 301-309.
5. H. Kagiwada and R. Kalaba, "Verification of the Invariant Imbedding Method for Certain Fredholm Integral Equations", J. Math. Anal. and Applic., v. 23, no. 3 (Sept. 1968), pp. 540-550.
6. H. Kagiwada, R. Kalaba, and A. Schumitsky, "Differential Systems for Eigenvalues of Fredholm Integral Equations", J. Math. Anal. Applic., v. 23, n. 2 (Aug. 1968), pp. 227-234.
7. R. Kalaba and E. Zagustin, "Reduction of Fredholm Integral Equations to Cauchy Systems", J. Franklin Institute, v. 293, n. 4 (Apr. 1972), pp. 277-283.
8. H. Kagiwada and R. Kalaba, "An Initial Value Theory for Fredholm Integral Equations with Semidegenerate Kernels", J. Assoc. Integral Mach., v. 17, n. 3 (July 1970), pp. 412-419.
9. H. Kagiwada and R. Kalaba, "Derivation and Validation of an Initial-Value Method for Certain Nonlinear Two-Point Boundary-Value Problems", J. of Opt. Theory and Applic., v. 2, no. 6 (1968), pp. 378-385.
10. R. Huss, H. Kagiwada and R. Kalaba, "Reduction of a Family of Nonlinear Boundary Value Problems to a Cauchy System", to appear in J. Applic. Analysis.
11. R. Huss, H. Kagiwada and R. Kalaba, "A Cauchy System for the Green's Function and the Solution of a Two-Point Boundary Value Problem", Journal of the Franklin Institute, v. 291, no. 3 (1971), pp. 159-168.
12. S. Mikhailin and K. Smolitskiy, Approximate Methods for Solution of Differential and Integral Equations, American Elsevier, New York, 1967.
13. R. Kalaba, "On Nonlinear Differential Equations, the Maximum Operation and Monotone Convergence", J. Math. and Mech., v. 8 (1959), pp. 519-574.

14. H. Kagiwada, R. Kalaba and C. -C. Yang, "Reduction of a Class of Nonlinear Integral Equations to a Cauchy System", Journal of Mathematical Physics, v. 13, n. 2 (Feb. 1972), pp. 228-231.
15. J. Buell, A. Fymat and R. Kalaba, "An Initial Value Method for the Ambarzumian Integral Equation", Journal of Quantitative Spectroscopy and Radiative Transfer, v. 12, no. 5 (May 1972), pp. 769-776.
16. H. Van Trees, Detection, Estimation and Modulation Theory, Wiley, New York, 1968.
17. A. A. Berezovskii, "Integro-Differentsial'nye Uravneniya Nelineinoi Teorii Pologikh Tonkikh Obolochek", "Ukranskii Matematicheskii Zhurnal, v. 12, no. 4, 1960.
18. R. Kalaba and E. Ruspini, "Invariant Imbedding and Potential Theory", Int. J. Eng. Sci., v. 7, no. 10 (Oct. 1969), pp. 1091-1101.
19. H. Landahl, "Mathematical Models in the Behavior of the Central Nervous System", a chapter in the book Mathematical Problems in the Biological Sciences, Amer. Math. Soc., Providence, 1962.
20. J. Diamond and W. Bossert, "Standing Gradient Osmotic Flow", Journal of General Physiology, v. 50 (1967), pp. 2061-2083.
21. H. Kutchai, J. Jacquez and F. Mather, "Nonequilibrium Facilitated Oxygen Transport in Hemoglobin Solution", Biophysical Journal, v. 10, no. 1 (1970), pp. 38-54.
22. D. Marsh, R. Kalman and H. Howard, "The Theory of Urine Formation in Water Diuresis with Implications for Antidiuresis", Bull. of Math. Biophysics, v. 29 (1967), pp. 67-89.
23. R. Kalaba and E. Zagustin, "Reduction of a Nonlinear Two Point Boundary Value Problem with Nonlinear Boundary Conditions to a Cauchy System", submitted for publication.

## OPTOELECTRONIC COMPUTATIONAL TECHNIQUES FOR FAST PICTURE PROCESSING

JENNY BRAMLEY

Geographic Information Systems Division  
U. S. Army Engineer Topographic Laboratories  
Fort Belvoir, Virginia 22060

In an earlier report<sup>1</sup>, I have shown how the use of analog TV-type techniques can greatly reduce both the cost of operation and the time required for the processing of pictorial information. The inherent limitation on the precision of a TV-type computer is not serious in this case since there is a comparable limitation on the accuracy of the experimentally obtained picture data. That earlier procedure was strictly sequential. The relatively high processing speed was the result of the elimination of drawbacks inherent in the digital computer, namely the delays in the input and output functions and the need for piecewise operation due to the inadequacy of the memory for the amount of data being processed.

Further consideration of possible analog approaches indicated that the addition of state-of-the-art (though not necessarily off-the-shelf) optoelectronic devices allows an increase in the previous operating speed by three orders of magnitude. Thus this report makes the earlier one obsolete. The speed-up is due to the use of parallel or near-parallel processing techniques.

To keep this presentation within a reasonable length, I restrict the discussion to the following four operations:

- (1) convolution, (2) Fourier transforms, (3) filtering,  
 (4) algorithms.

(1) The basis of the convolution approach, as well as of the other operations, is the imparting of information to a light beam by passing it through a transparency. Consider two transparencies A and B with fiducial marks to determine registry and relative orientation. Let resolution element  $(x, y)$  of A have an absorption  $\alpha_1 = -\log f_1(x, y) > 0$  (assuming that the function  $f_1(x, y) < 1$ ). The intensity  $U$  of a beam of parallel light passing through  $(x, y)$  is transformed into  $Uf_1(x, y)$ . Similarly, if the element of B superimposed on element  $(x, y)$  of A is specified by  $(s-x, t-y)$  and has an absorption  $\alpha_2 = -\log f_2(s-x, t-y)$ , then after passing through the set, the light beam has an intensity  $Uf_1(x, y)f_2(s-x, t-y)$ . The total amount of light transmitted through the set AB (usually measured by a photomultiplier) is:

$$I(s, t) = \int_S \int U f_1(x, y) f_2(s-x, t-y) dx dy$$

The integration is performed over the area  $S$  being investigated. The function  $I(s, t)$  is the convolution integral.

In principle, the change in the relative positions of A and B, which would give rise to different values of  $s$  and  $t$ , could be achieved by mechanical means. However, for an operation of interest in picture processing, this brute force approach is grossly inadequate from the point of view of achievable speed and the specification of relative orientation. I propose replacing it with the

following optoelectronic arrangement:

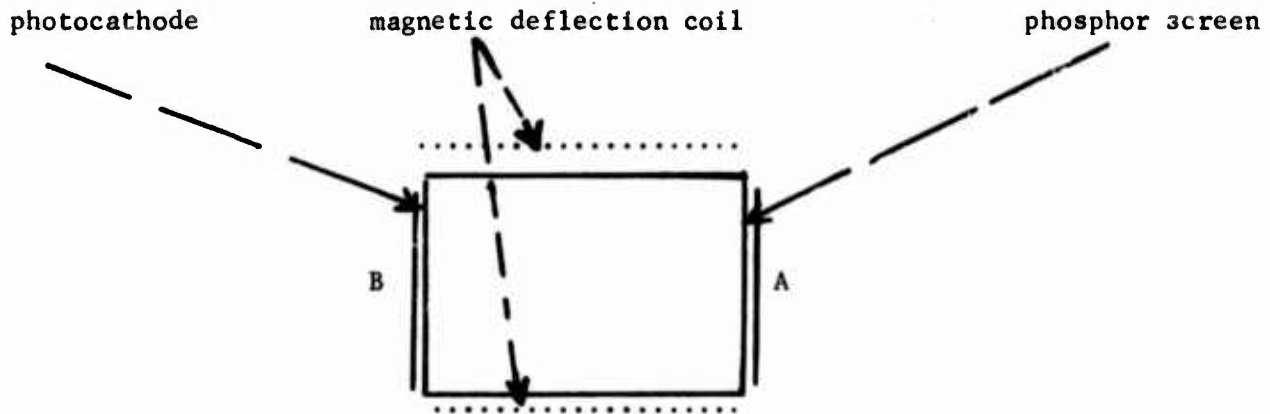


FIGURE 1

Image Converter Tube

My "central processor" is a magnetically focused image converter tube (schematized in Fig. 1) with flat fiber optics plates at input and output. This is an off-the-shelf item. Transparency A is mounted directly in front of the faceplate, and transparency B, which is several times smaller than A--to permit its correlations with different portions of A--is mounted directly in front of the photocathode. This arrangement eliminates the use of lenses with attendant loss of light and provides for compactness. Transparency B is illuminated and forms the input on the photocathode. A deflection coil, such as used in the Goodyear Correlatron, permits scanning the luminescent output image across the faceplate in any desired pattern. At each position, the light from the luminescent image of B passes through A and is picked up by a photomultiplier in an arrangement



that is conventional for flying spot scanners. The position of A is specified by fiducial marks, that of B is determined by the current passing through the deflection coil. The output of the photomultiplier can be recorded or displayed in any conventional manner to specify the occurrence and magnitude of the maximum.

The determination of each correlation value can proceed at rates standard for a flying spot scanner, no matter how large the area being correlated. The limiting speed factor is the phosphor decay time. In the case of a P16 phosphor screen, a high degree of accuracy may call for a rate somewhat slower than conventional TV time per deflection, e.g.,  $\frac{1}{4}$   $\mu$ sec per correlation. Allowing for retrace times, this provides more than 2000 correlations per millisecond. The main time delay arises in changing pictures because rigorous alignment of picture A is critical. Since picture B is displayed on an image converter tube, the controls are electronic and more easily achieved. However, with suitable fiducial marks, the alignment of picture A can probably be automated to require only a fraction of a second.

(2) The Fourier transform of a picture can be treated as a special case of correlation, based on the following mathematics:

Let  $g(n, m)$  ( $n, m=1, \dots, N$ ) represent the intensity at every point  $(n, m)$  of a picture  $N \times N$  elements considered as a matrix. The first index numbers the rows and the second one the columns. Take the cosine transform as an example. The coefficients in it are defined



by

$$C(n, m) = \sum_{\mu=1}^N g(n, \mu) \cos 2\pi m \mu / N,$$

or in terms of a positive processing function  $f_c^{(m)}(\mu) = 1 + \cos 2\pi m \mu / N$

$$C(n, m) = \sum_{\mu=1}^N g(n, \mu) f_c^{(m)}(\mu) - \sum_{\mu=1}^N g(n, \mu) \quad (1)$$

The first term on the right-hand side of Eq. (1) is a convolution--expressed as a finite sum. Using an image converter tube and an optical processing plate, all the multiplications and additions in Eq. (1) can be performed in parallel. The processing plate is a transparency of  $N+1$  parallel strips, each strip having the width of a resolution element of the picture and a length  $N$  times the width. The transmittivities of the successive strips are  $k f_c^{(0)}$ ,  $k f_c^{(1)}$ , ...,  $k f_c^{(N)}$ ,

$$k, k(1 + \cos 2\pi \mu / N), k(1 + \cos 4\pi \mu / N), \dots, 2k \quad k < 0.5$$

where  $\mu$  is the running index, which assumes all integer values from 1 to  $N$ . The quantity  $k$  is a constant of the plate. While the preparation of such a plate entails time and expense, it is a onetime operation.

To obtain the coefficients  $C(n, m)$  of the transform, line  $n$  of the picture is projected successively on the  $N+1$  transmissive strips. The light is then focussed on a photomultiplier as in the correlation operation. In suitable units, the total amount of light transmitted through strips 1, 2, and  $m+1$  is

$$k \sum_{\mu=1}^N g(n, \mu) \equiv g_n, \quad k \sum_{\mu=1}^N f_c^{(1)}(\mu) g(n, \mu), \dots, \quad (2)$$

$$k \sum_{\mu=1}^N f_c^{(m)}(\mu) g(n, \mu)$$

respectively.

Hence to obtain the coefficients  $C(n, m)$  of the Fourier transform for any value of  $m$  (and a given line  $n$ ), we store  $g_n$  and take the difference between the photomultiplier output after illumination of the line  $m+1$  and the stored signal  $g_n$ . The proportionality coefficient  $k$  can be determined by calibration.

This scheme allows each transform coefficient to be obtained in a single step instead of requiring  $N$  successive multiplications as in the case of a conventional serial computer. To implement it, we use the same arrangement as for convolution. The processing plate is picture A, while picture B is a transparency of  $N \times N$  resolution elements. ( $N \sim 1000$ ). It is illuminated one horizontal resolution line at a time and is imaged on the photocathode of the image converter tube. By means of magnetic deflection, it is placed successively in front of each strip. The light transmitted represents the operations in Eqs. (2). To minimize errors due to phosphor

persistence, about 0.2  $\mu$ sec should be allowed per deflection. But even at this "slow" speed, we obtain 5 Fourier transform coefficients per microsecond, and the one-dimensional transform of a 1000  $\times$  1000 picture requires only about 0.2 sec. Parallel processing for Fourier coefficients eliminates the need for any algorithms of the Cooley-Tukey type.

(3) As far as filtering is concerned, a number of factors must be considered. As a rule, the objective is to try a number of filters with the same picture. Therefore, in the arrangement described for convolution, the filter stands for picture B, ahead of picture A which is to be filtered. The parallel output can be used for direct viewing, or photographed (with all the processing delays involved), or it can be recorded on the cathodochromic screen of an additional image converter tube. (This also takes time.) For all other uses, a sequential output is essential.

If the filter is available in sequential form on video tape, transcribed, e. g., from the output of a digital computer or of a flying spot scanner, it is presented on the screen of a cathode ray tube rather than of an image converter tube. The light emitted by B and transmitted through A is picked up by a photomultiplier separately for every resolution element and is recorded in conventional fashion.

$$H(n, m) = G(n, m)S(m)$$

The problem is to find  $h$ , the inverse transform of  $H$ . The derivation shown appears to be self-explanatory

$$G(n, m) = \sum_{\mu=1}^N g(n, \mu) W^{\mu m} \quad W = e^{2\pi i/N} \quad W^N = 1$$

$$\text{If } \mu < p, \quad \sum_{m=1}^N S(m) W^{-m(p-\mu)} = s(p-\mu)$$

$$\text{if } \mu \geq p \quad \sum_{m=1}^N S(m) W^{-m(p-\mu)} = \sum_{m=1}^N S(m) W^{-m(N+p-\mu)}$$

$$= s(N+p-\mu) \text{ since } N+p-\mu > 0.$$

$$h(n, p) = \sum_{m=1}^N H(n, m) W^{-mp} = \sum_{m=1}^N G(n, m) S(m) W^{-mp}$$

$$= \sum_{\mu=1}^N g(n, \mu) \sum_{m=1}^N S(m) W^{-m(p-\mu)}$$

$$= \sum_{\mu=1}^{p-1} g(n, \mu) s(p-\mu) + \sum_{\mu=p}^N g(n, \mu) s(N+p-\mu)$$

To see more clearly what is involved in implementing this approach we write a few of the coefficients  $h(n, 0)$  explicitly

$$h(n, 1) = g(n, 1)s(N)+g(n, 2)s(N-1)+g(n, 3)s(N-2)+\dots+g(n, N)s(1)$$

$$h(n, 2) = g(n, 1)s(1)+g(n, 2)s(N)+g(n, 3)s(N-1)+ \dots+g(n, N)s(2)$$

$$h(n, 3) = g(n, 1)s(2)+g(n, 2)s(1)+g(n, 3)s(N)+\dots+g(n, N)s(3)$$

$$\dots\dots\dots$$

$$h(n, N) = g(n, 1)s(N-1)+g(n, 2)s(N-2)+g(n, 3)s(N-3)+\dots+g(n, N)s(N)$$

As the index  $p$  changes by one unit, the transformed filter coefficients  $s(m)$  are translated cyclically by one unit; i.e.:

$$\begin{array}{cccccc} s(N) & s(N-1) & \dots & s(3) & s(2) & s(1) \\ s(N-1) & s(N-2) & \dots & s(2) & s(1) & s(N) \\ s(N-2) & s(N-3) & \dots & s(1) & s(N) & s(N-1) \end{array}$$

.....

To obtain the intensity  $h(n, 1)$  of the first resolution element on line  $n$  of the filtered picture, we present the first sequence of the transformed filter coefficients in luminous form on the face of a cathode ray tube and shine this filter function through line  $n$  of the original picture. The light transmitted through all the resolution elements is picked up and integrated by a photomultiplier. The operation is repeated after a one-step cyclic translation is performed on the coefficients  $s(m)$ . The only way I envision of performing this cyclic translation is to store the coefficients  $s(N), \dots, s(1)$  on a scan converter tube in a circular scan made up of  $N$  elements. The readout also follows a circular pattern, but for each successive scan line, the scanning starts at the same element where the preceding

scan terminated. Writing and reading rates for a scan converter tube can be real time or slower so that there is no problem in recording the coefficients  $s(m)$  as they are obtained by a Fourier transform from the frequency filter  $S(m)$ . As indicated above, the determination of each coefficient  $h(n, p)$  calls for the scanning of one line of the picture. This can be done in real time. Thus an entire line of a transformed picture is obtained per TV frame time. This means that starting with a  $1000 \times 1000$  element picture and a line frequency filter, we obtain a filtered picture in little over half a minute.

(4) In the digital processing of images, masking operation and statistical analysis of neighboring areas have proved very successful in extracting information. The drawback has been the length of time required to perform these operations. An equivalent approach has been tried with a special type of tube, the image storage tube, where each step of the algorithm is performed sequentially but all points in the picture are handled in parallel. In principle, this would be the ideal solution, but it was not made to work in practice except under very restrictive conditions which destroyed the usefulness of the method.

The following approach, illustrated on a very simple example, provides a means for convenient parallel handling in the vicinity of any particular point, though the different points in a picture are handled sequentially. Consider a resolution element symbolized by a point  $(0, 0)$  of a picture and assume that the processing

affects its 8 nearest neighbors as well. The intensities at that point and its vicinity constitute the array  $\{g\}$ . In the type of algorithm considered, the intensity at every point of the area is multiplied by a preselected number, positive or negative, forming the array  $\{b\}$ .

$$\begin{array}{cccc}
 g_{1-1} & g_{1 0} & g_{1 1} & b_{1-1} & b_{1 0} & b_{1 1} \\
 \{g\} = g_{0-1} & g_{0 0} & g_{0 1} & \{b\} = b_{0-1} & b_{0 0} & b_{0 1} \\
 g_{-1-1} & g_{-1 0} & g_{-1 1} & b_{-1-1} & b_{-1 0} & b_{-1 1}
 \end{array}$$

The products are added together to form

$$J = \sum_{\mu=-1}^{+1} \sum_{\gamma=-1}^{+1} g_{\mu\gamma} b_{\mu\gamma}$$

which represents a convolution of arrays  $\{g\}$  and  $\{b\}$ . If all the numbers of array  $\{b\}$  are positive, it can be used to form a transparency with the appropriate intensity values and apply the techniques described above for convolution. Otherwise, every quantity has to be measured from a bias level, which makes the new array positive. An appropriate bias function has then to be subtracted from the final result. This is a standard operation.

A more flexible approach to the same problem and one that does not require photographic registry of the algorithm on a transparency calls for a multibeam cathode ray tube. In such a tube, the entire array  $\{b\}$  is written in parallel on the screen and is then scanned

across the entire picture to be processed. Such tubes have been built by Litton Industries and by Sylvania and (based on a different principle) are under consideration by the Stanford Research Institute. Arrays of up to 5 X 7 beams have been fabricated, but a redesign is necessary to insure that the array consists of adjoining intensity squares rather than an array of discrete luminescent spots. The techniques for accomplishing this are well known.

In comparing optoelectronic and purely digital image processing, we must take into account that the former has speed on its side while the latter one has high precision. But while only minor improvements in the components could increase the precision of optoelectronic processing to the extent that the system is no longer the limiting factor, it requires a breakthrough in the speed and storage capability of a digital computer to make it competitive in these areas. If cost effectiveness is a consideration, we should concentrate on the implementation of optoelectronic system designs.

#### REFERENCES:

1. J. Bramley, Proc. 1971 Army Num. Anal. Conf., p. 11.



AN ALGORITHM FOR RAPIDLY PARSING  
AUTOMATICALLY DIGITIZED DRAWINGS \*

Charles M. Williams  
VISICON, Inc.  
State College, Pennsylvania

The potential for the breadth of applicability of computer graphics is demonstrated to some degree by the variety of the figures which appear in this paper. Such potential, however, cannot be realized until pictures and drawings can routinely be entered into computers and manipulated there with a minimum of human effort and drudgery. The purpose of this paper is to describe a working process which contributes a step toward this goal.

The drawings in this paper were all automatically digitized and transduced into a digital computer by means of the Visicon AD-1 Automatic Digitizing System shown in Figure 1. This system can process an 11" x 17" document in 58 seconds at a resolution of 100 samples per inch, or 116 seconds at 200 samples per inch. The digital output usually goes to magnetic tape, but can alternatively go directly into a computer. The digitizer itself weighs roughly 75 pounds and can fit on a desk top.

Figures 2 through 4 were reproduced directly from the digitized image onto an electrostatic (raster) plotter. A close inspection of these images will reveal that lines on them are composed of individual points. Each image is actually formed from a mosaic (raster) of black and white dots which are represented in a computer by means of ones and zeros. The reasons for digitizing such drawings are manifold. The resultant data may be transmitted over telephone lines, inserted into computer files, analyzed, or may even serve as a mechanism for controlling machinery or computer software.

The EEG of Figure 2, for example, was digitized so that a computer could make a frequency analysis for medical diagnoses. Such data can also be input into a computer with a manual digitizer or a data tablet by laboriously retracing the lines of the drawing. The drudgery of such retracing, however, results in poor human performance when attempted over extended periods of time. The work is monotonous and yet requires meticulous hand and eye coordination. One can imagine the difficulty of detecting and removing those frequencies introduced into EEG data during the tracing process by the vagaries of the human muscular motor response.

---

\*This research was funded in part by a grant from VISICON, Incorporated to Small Industries Research, The Pennsylvania State University.

The printed circuit and typed lettering shown in Figures 3 and 4 indicate the quality of the digitized image. Such information can be inserted into computer archival storage where it is accessible to interactive updating, computer analysis, or optical character recognition.

A common requirement for processing such digitized images is the ability to abstract line information from them. That is, the raster of disassociated black and white points must be collated into sets of lines strokes. As digitization can generate 40,000 data points per square inch of drawing, such line isolation and identification must be affected by a very efficient process. The graphic collation (on an IBM 360/65) which yielded the digital plots shown in Figures 5 through 10 were accomplished in times essentially proportional to the total line length on the drawing and independent of its complexity.

The digitized data for the map shown in Figure 5 required 23 seconds of computer time to collect constituent digitized points into lines, thin such lines to individual pen strokes, and generate plot commands. Figure 6 required 183 seconds of computer time and reveals the results obtained by processing a far more complex drawing. In this instance, thin lines were reproduced as individual strokes and thicker symbols were represented by their peripheries.

Figure 7 shows the results obtained by digitizing cardboard cutouts for garment patterns at 200 samples per inch. The computer processing time required to generate the plotting commands was six seconds per pattern. Such plotting commands could just as easily have been used to direct a cloth or metal cutting machine. The drawing then would have served as the direct source of instructions for a numerically controlled tool.

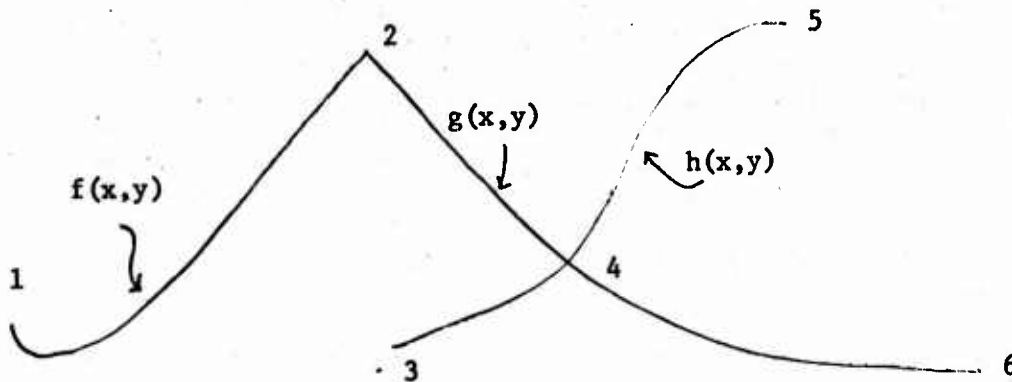
The input to this graphic collator is the raw digitized data. The output is a sequence of triples (X, Y, I):

X,Y the Cartesian coordinates in inches of a line  
point  
I an indicator which is 0 if the point is the beginning  
of a line segment and is 1 otherwise

The indicator, I, plays the role of the "pen code" commonly employed in digital plotter commands. This set of triples forms a complete geometric description of the lines which constitute the drawing. The format of this output can easily be modified to interface with a large variety of digital plotters, computer display scopes, and numerically controlled tools. In addition, this output may be further processed by the computer to yield a smoother or more compact analytic representation of the data for transmission or storage in computer files.

In many applications, computer analysis of drawings is a very important requirement. In these cases, it becomes necessary to abstract and categorize the important features of each picture. Such a process is quite similar in concept to the grammatical parsing of sentences into component parts of speech. In essence, a drawing is reduced to its component line segments to yield a network description of the original.

The structure of the process is very similar to that employed by programming language compilers on digital computers. The initial phase of graphic collation has been described and involves the collection of raw data points into component lines. Attendant to this is the process of graphical lexical analysis in which lines are segmented into component parts, and the interrelationship of separate lines is detected and noted. Basically, this involves isolating and measuring points where lines terminate, intersect, or exhibit slope discontinuities. The locations of these features along with a path description of the lines between them constitute a network description of the original drawing which can be used to regenerate or categorize it.



The drawing above, for example, can be represented by the line terminal points 1, 3, 5, and 6; the line intersection point 4; and the line slope discontinuity point 2; along with the tabular or analytic descriptions  $f$ ,  $g$ , and  $h$  which categorize the curve shapes between these points. The output from the graphical lexical analyzer for this drawing can be represented in tabular form:

<u>Initial Point</u>	<u>Terminal Point</u>	<u>Curve</u>
$(X_1, Y_1)$	$(X_2, Y_2)$	f
$(X_2, Y_2)$	$(X_4, Y_4)$	g
$(X_3, Y_3)$	$(X_4, Y_4)$	h
$(X_4, Y_4)$	$(X_5, Y_5)$	h
$(X_4, Y_4)$	$(X_6, Y_6)$	g

This network description has all the information of an interconnection matrix and contains the geometric constraints of the system as well. This description can thus serve to reproduce the drawing or provide input for specialized analysis programs. The network of Figure 8 for example, was used on an interactive graphics terminal to generate an electrical network description of the circuit in a format which could be directly input to circuit analysis programs for computer simulation or diagnostics.<sup>1</sup> The curves in this instance were represented by polynomial approximations.

The output of the graphic lexical analyzer can also serve as data for a graphic syntax analyzer. Figures 9 and 10 exemplify the results of such syntactic analyses. The lines and important features of Figure 9 were detected and isolated by the graphic collator and graphic lexical analyzer to produce a network description of the drawing. This data was then processed by a graphic syntactic analyzer specially constructed to recognize circles, squares, and rectangles. The results shown in Figure 10 are lines plotted from the appropriate words to the proper figures on the drawing. This analyzer was readily constructed as it had to work only on the fixed format data structures of the network rather than on the imperfect raw data itself. The system works on rough drawings whose allowable deviations from perfect geometric figures is specifiable by parameters.

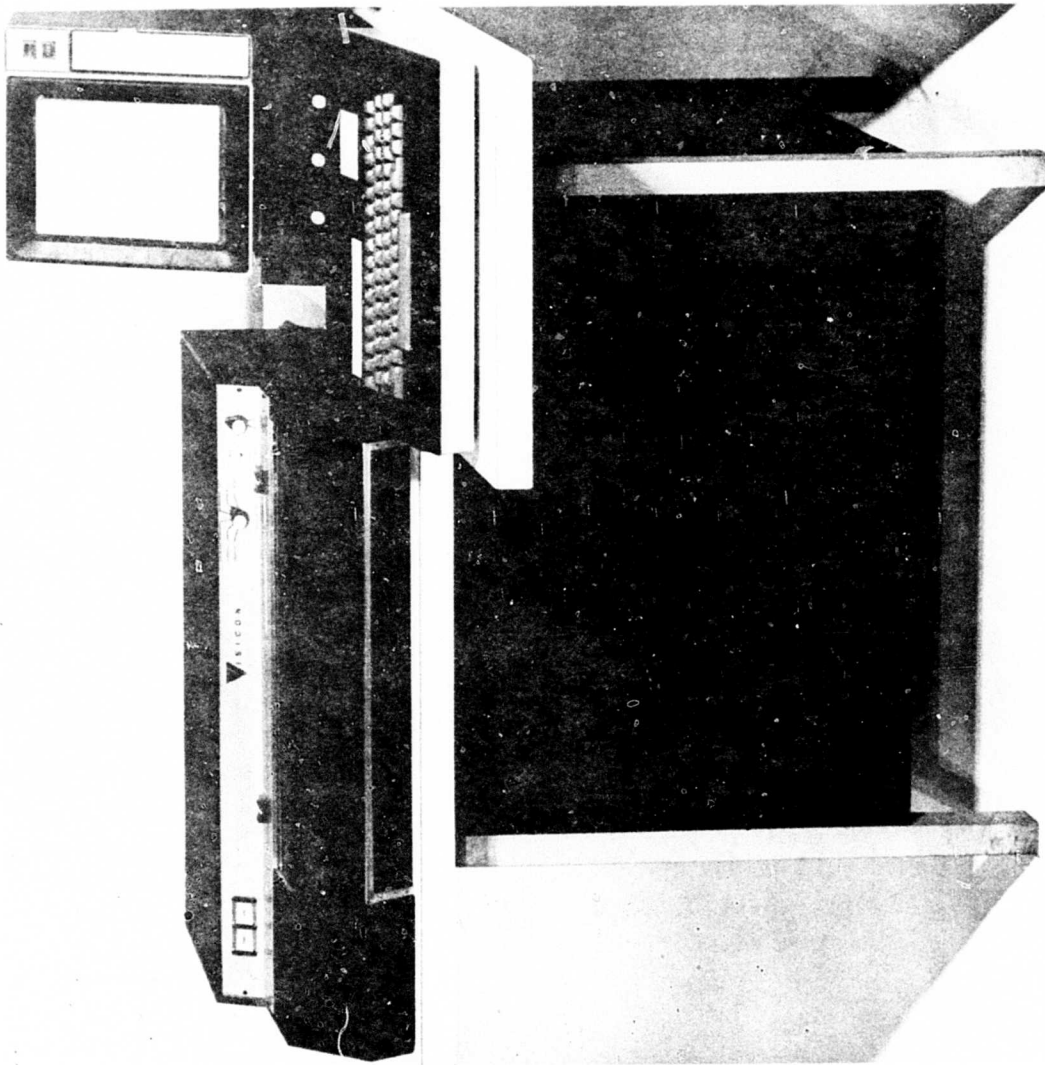
In summary, the automatic digitizing software system can be categorized as consisting of three distinct sections which perform the following functions:

1. Graphic Collator - collects and sorts digitized raster points into lines
2. Graphic Lexical Analyzer - translates lines into a descriptive network of nodal points and connecting paths
3. Graphic Syntactic Analyzer - parses and recognizes graphic symbols using the network formed by the Graphic Lexical Analyzer

---

<sup>1</sup>Joseph Novoshielski, An Interactive Program for Automated Network Description, unpublished Master's paper, Computer Science Department, The Pennsylvania State University, University Park, Pa. 16802, 1972

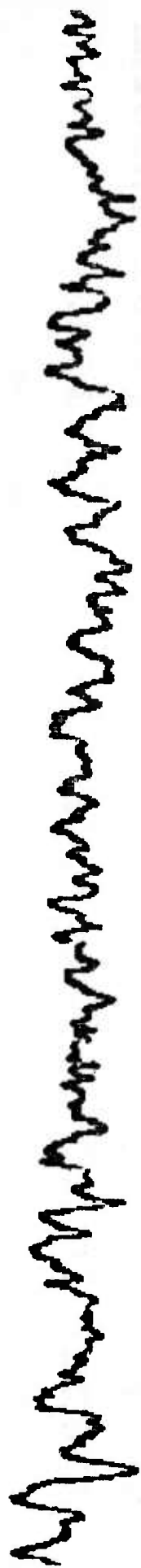
The system described here is capable of automatically and economically reducing drawings to a format which can be used as direct input to systems for computer graphics, numerically controlled tools, data transmission, network analysis, pattern recognition, or design analysis. The value of such processing is that the accuracy of the result is independent of the complexity of the original drawing and is a function only of the resolution of the digitizing operation. Humans in attempting to perform the same tasks manually will inevitably omit some data and interject inaccuracies into the rest. The detection and correction of such errors involves as much work as does the original digitizing.



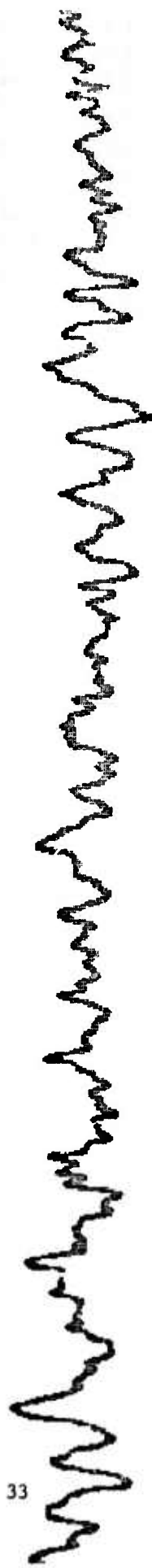
VISICON, INC.  
Model GC-3

THE VISICON AD-1 AUTOMATIC DIGITIZING SYSTEM

10-10000-1



33



Electroencephalograph digitized by a VISICON AD-1 System and reproduced on an electrostatic plotter.

FIGURE 2



the software for expansion and reformatting of the operation is completed, the reformatted raster by Versatec 1100 printer/plotter which produces the

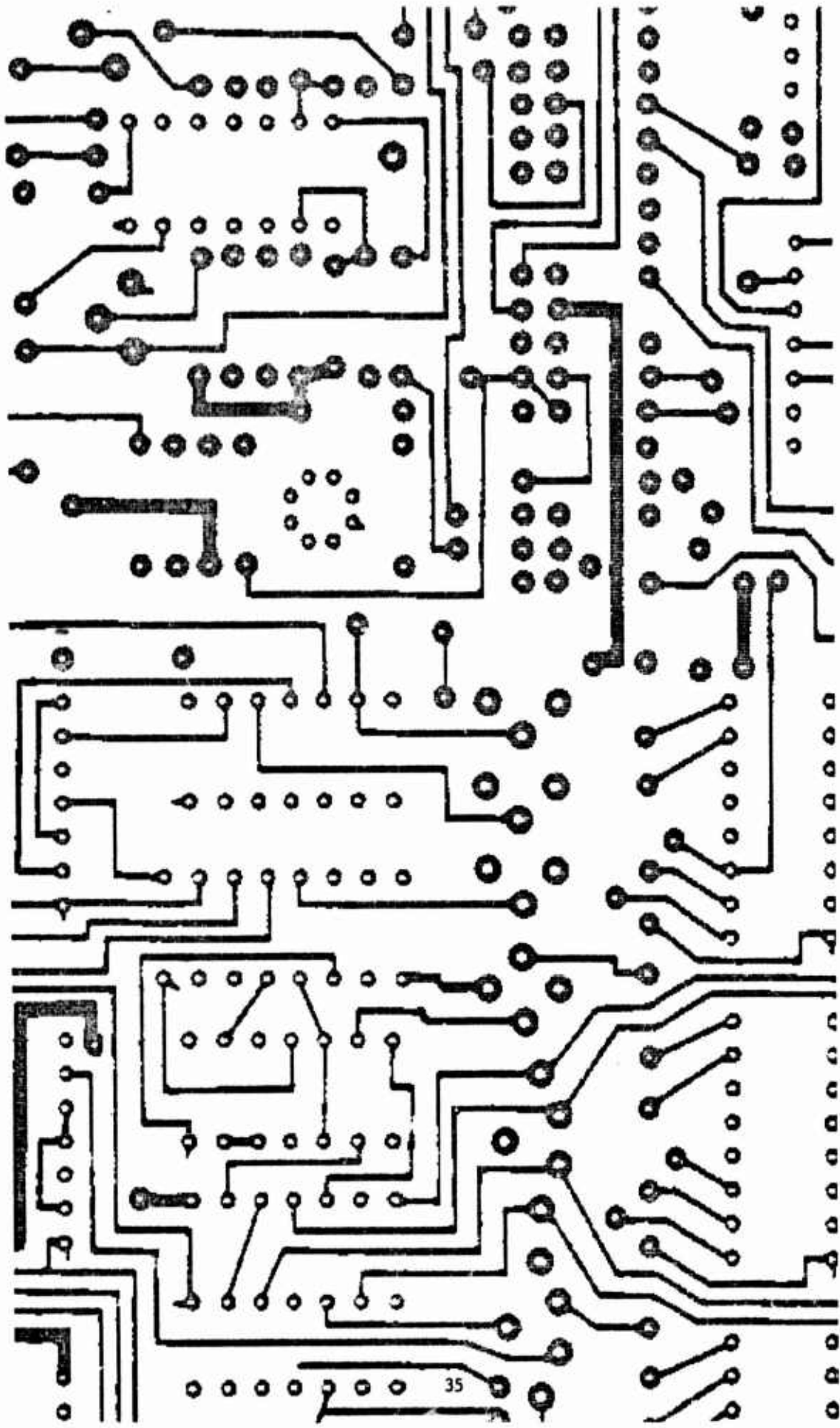
The system normally operates with a resolution of 300 dots per inch. This resolution is totally adequate for most documents and shows the quality of output which can be expected

The GC-3A digitizer is also available with a dual (200 dots per inch) output for documents which are not full size. However, the transmission times may be increased by a factor of 4 at 200 LPI as the amount of digital information is proportional to the square of the resolution.

Typed page digitized by a VISICON AD-1 System at 200 samples per inch and reproduced on an electrostatic plotter at 2 2/3 scale.

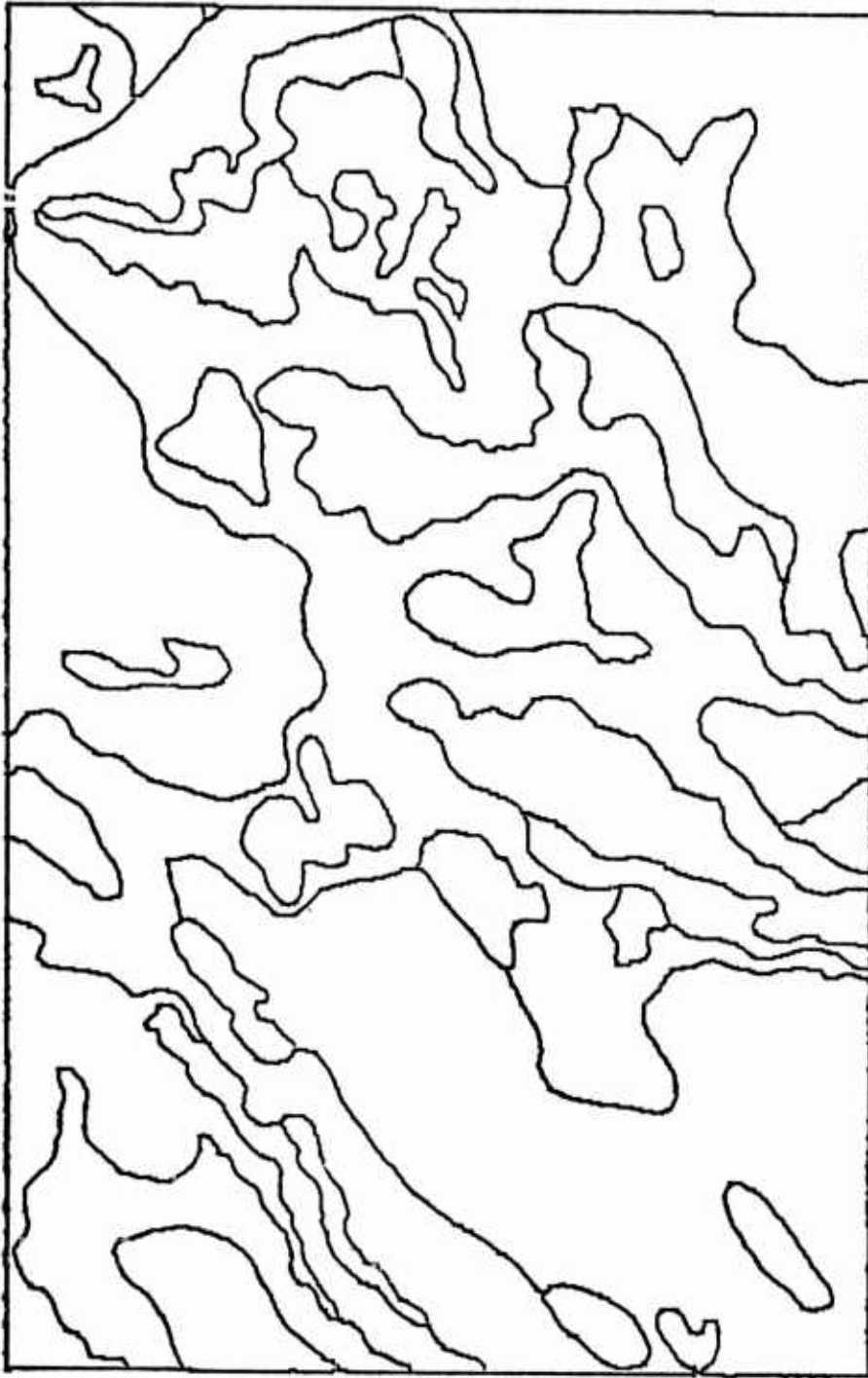
Figure 3





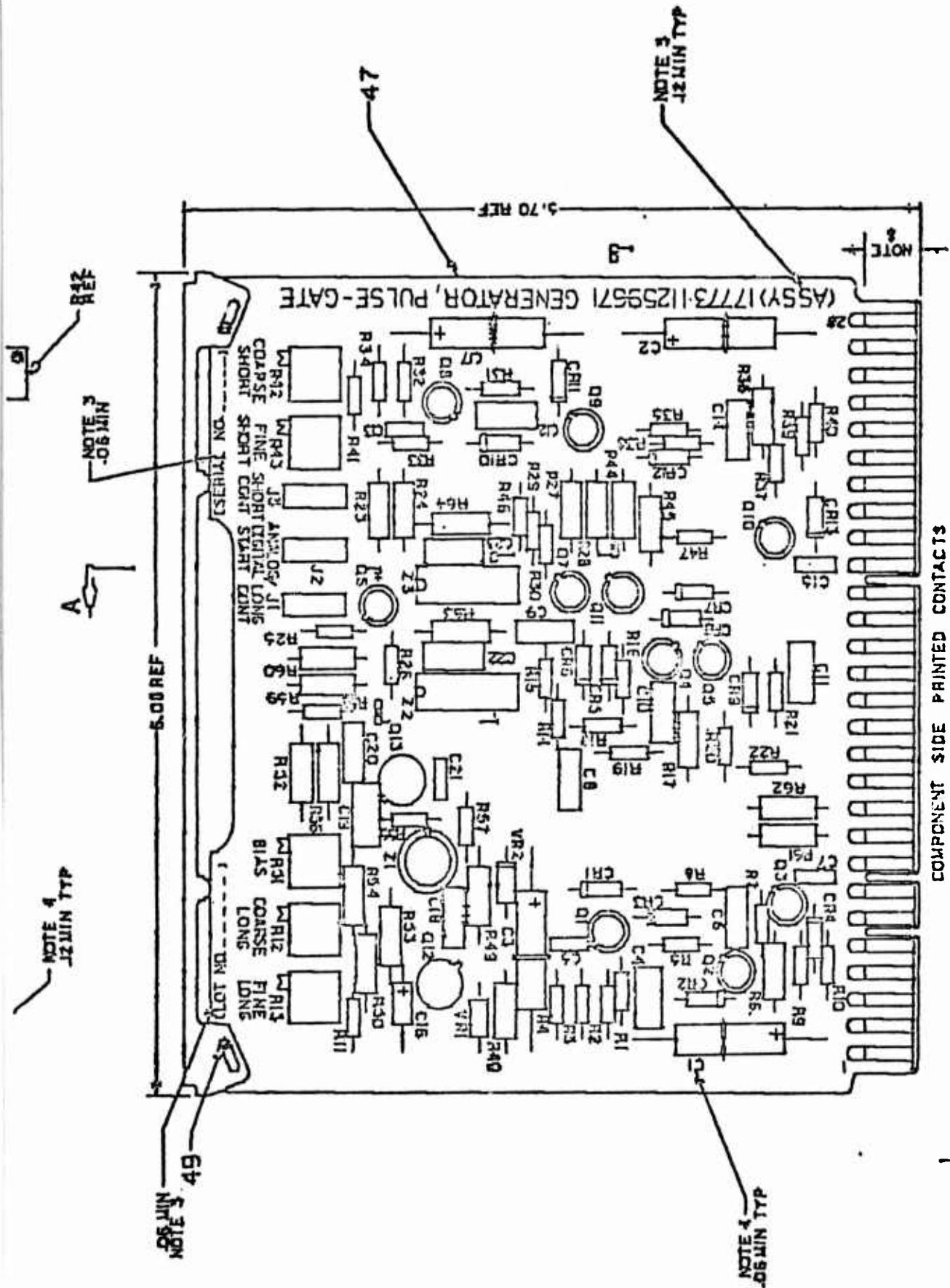
Printed circuit board digitized by a VISICON AD-1 System and reproduced on an electrostatic plotter.

Figure 4



Digital plot of a map generated by the Graphic Collator from digitized information supplied by a VISICON AD-1 System.

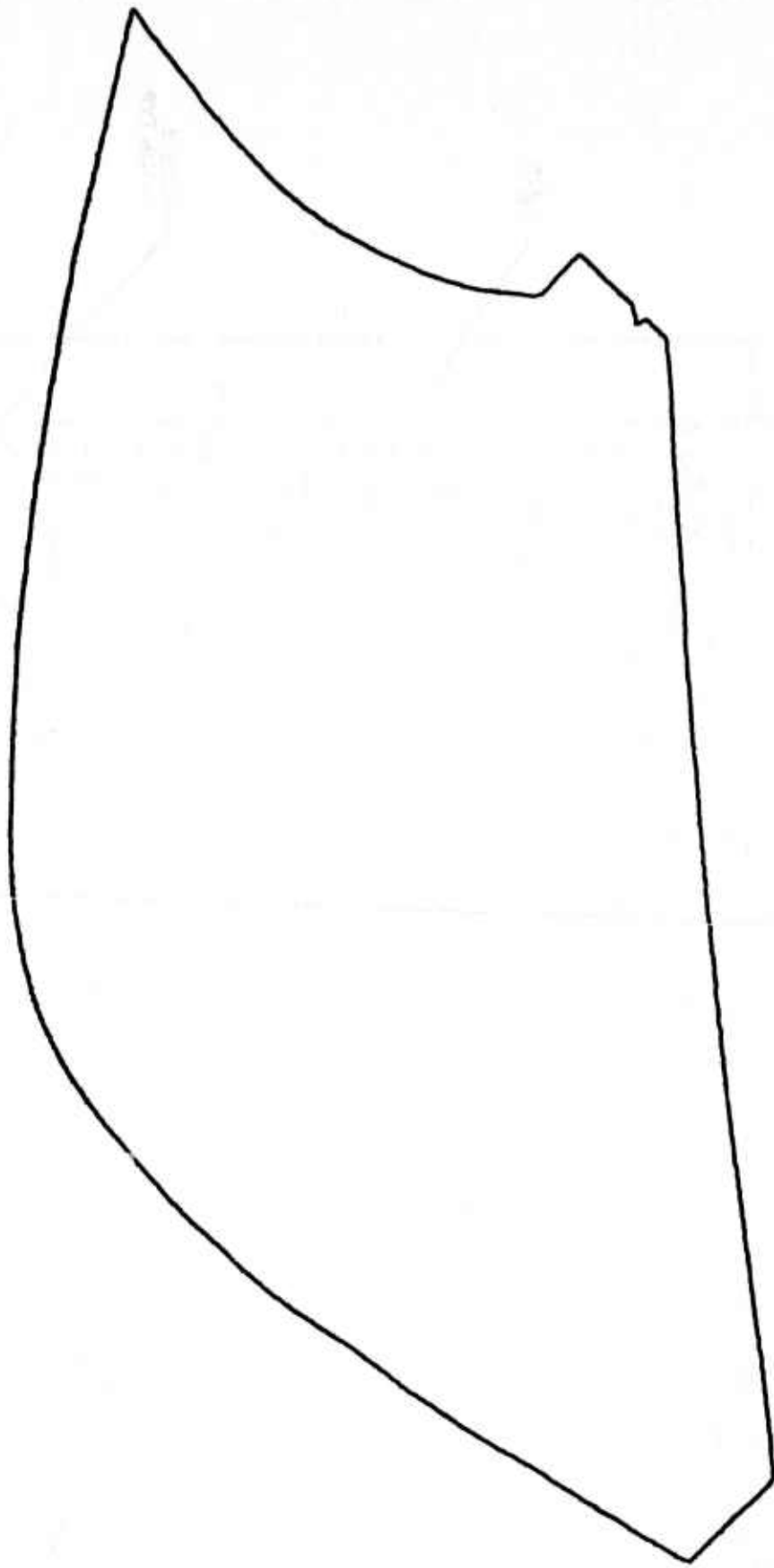
Figure 5



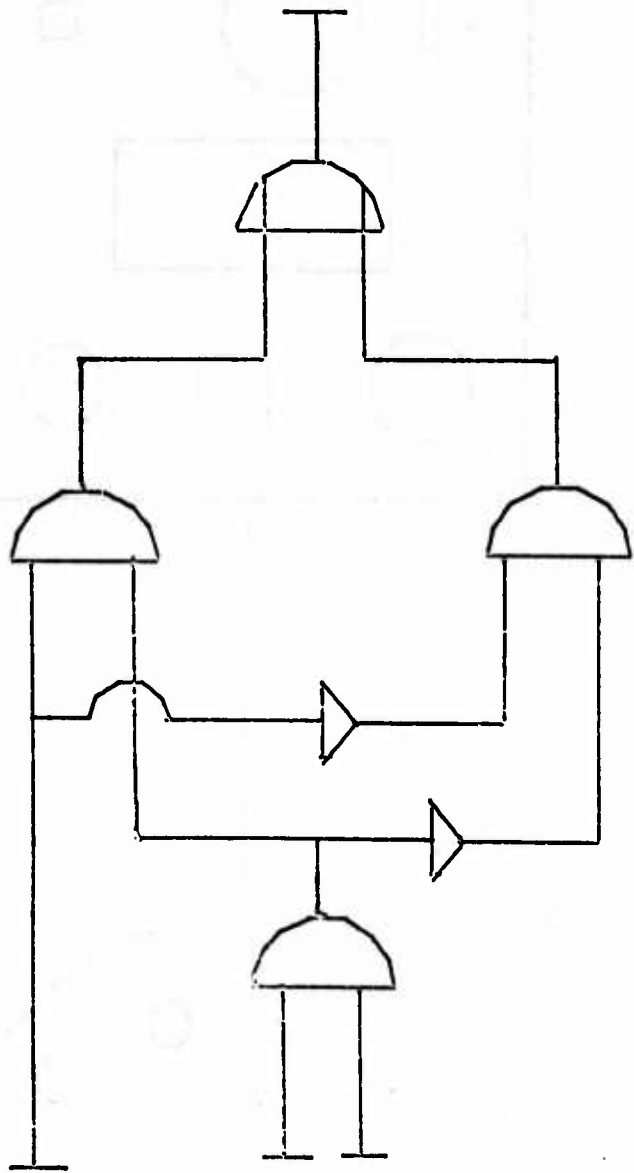
COMPONENT SIDE PRINTED CONTACTS

Digital plot of an intricate drawing generated by the Graphic Collator.

Figure 6

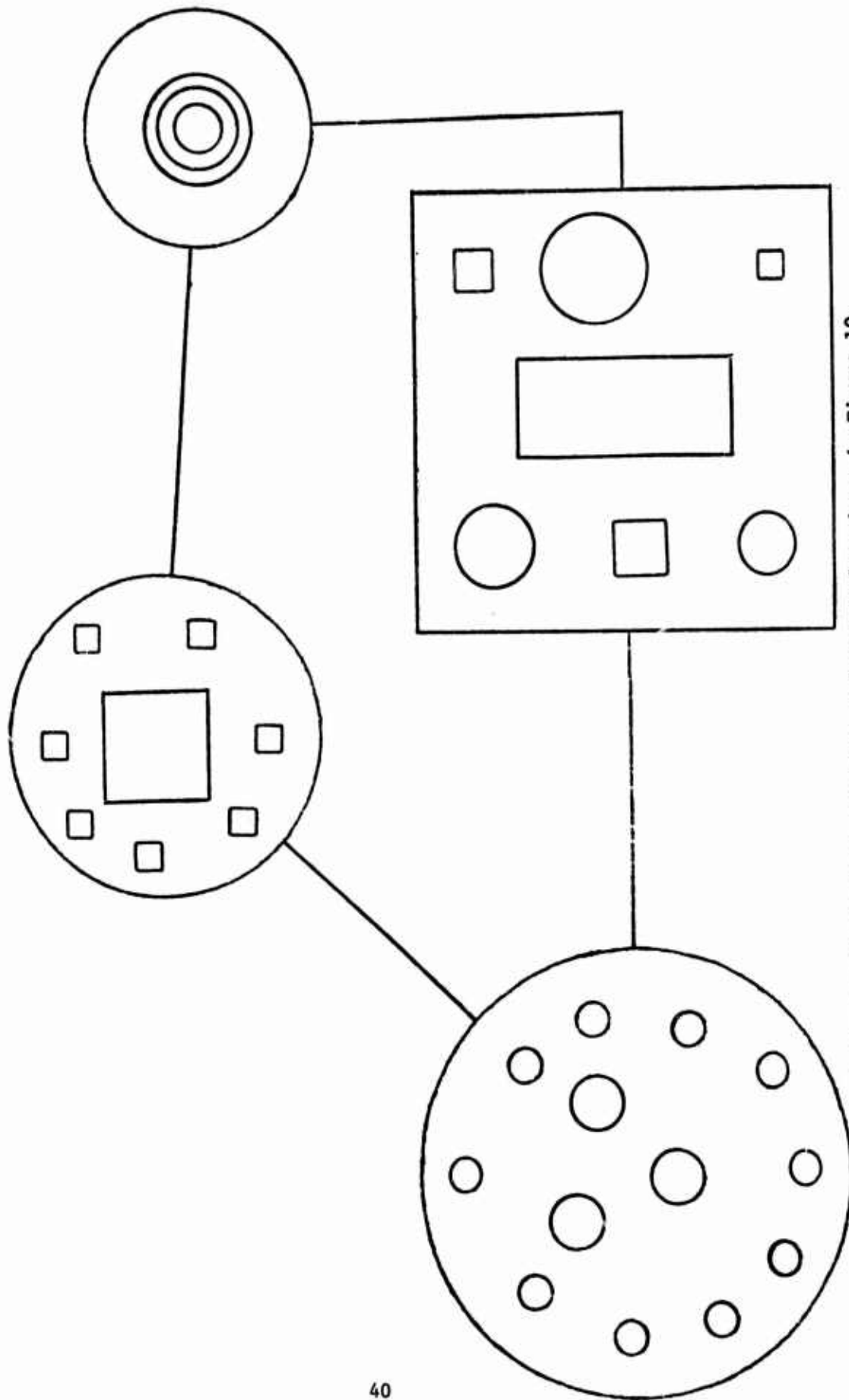


Digital plot of a garment pattern generated by the Graphic Collator.



Digital plot of electrical schematic from polynomial equations  
generated by the Graphic Lexical Analyzer from  
digitized information supplied by a VISICON AD-1 System.

Figure 8

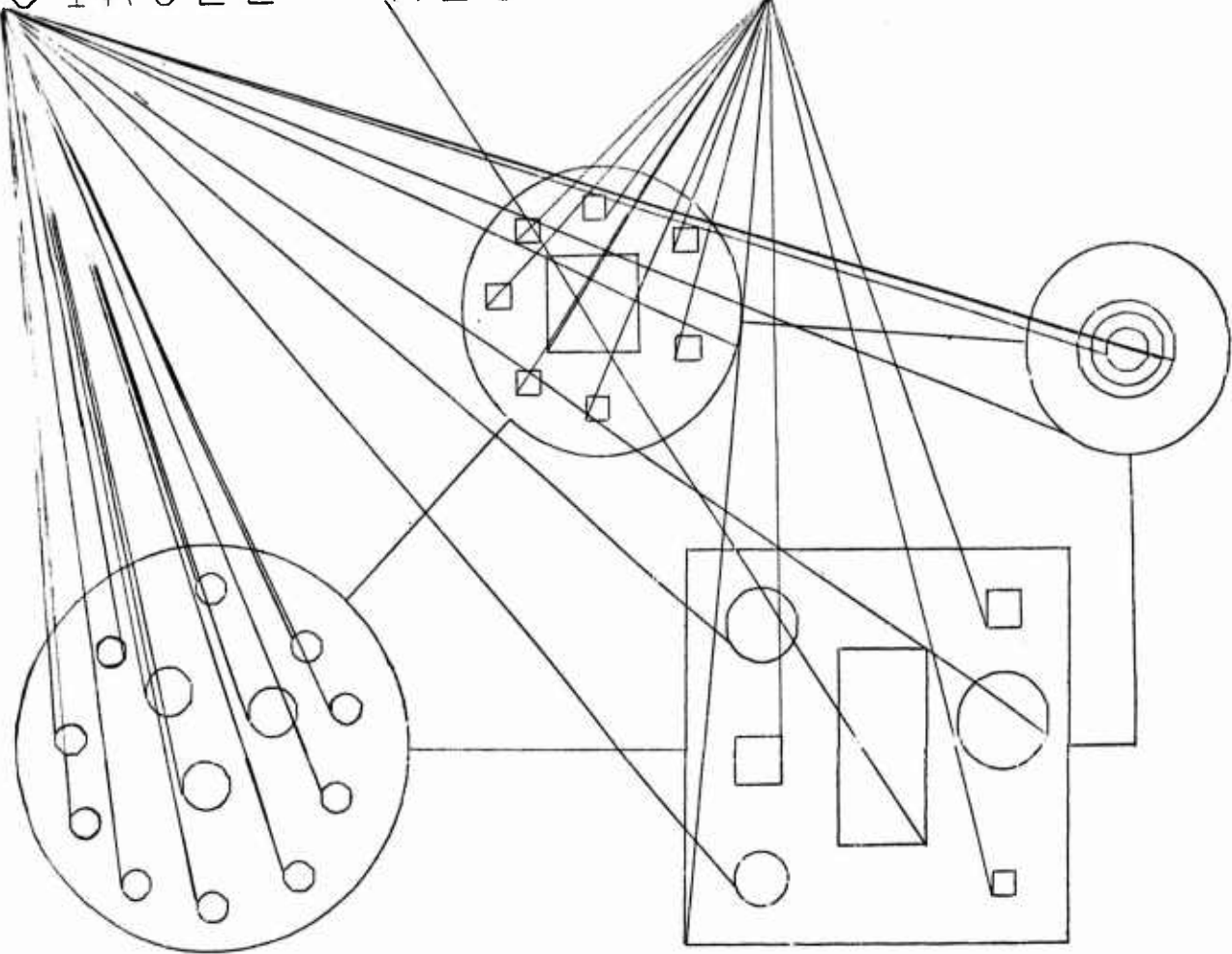


Original drawing of circles, squares, and rectangles shown in Figure 10.

Figure 9



CIRCLE      RECT      SQUARE



Drawing digitized by a VISICOM AD-1 System and replotted on a digital plotter from data generated from the Graphic Lexical Analyzer. The Graphic Syntactic Analyzer has located, identified, and measured the circles, rectangles, and squares in the picture (original drawing shown in Figure 9)

Figure 10

EVALUATION OF THE ROOTS  
OF CROSS-PRODUCT BESSEL FUNCTIONS

Shih-Chi Chu & Philip D. Benzkofer  
Research Directorate  
Weapons Laboratory at Rock Island  
U. S. Army Weapons Command  
Rock Island, Illinois

ABSTRACT

A difficulty frequently encountered in the solving of ordinary and partial differential equations in the problems of heat transfer, electricity, fluid and solid mechanics involving an annular region (such as a gun tube) subjected to Dirichlet, Neumann, or mixed type boundary conditions is that of obtaining the roots of nonlinear cross-product Bessel functions. By use of implicit iterative techniques, a digital computer program was prepared by personnel in the Research Directorate of the Weapons Laboratory at Rock Island to overcome this difficulty and thus to provide all necessary roots for cross-product Bessel functions. Tables of roots of certain particular cross-product Bessel functions with varying order are given. These tables are particularly useful for gun-tube heat transfer analysis and can be used directly by a designer for the calculation of transient temperature distribution, and thermal stresses and strains in a small or large caliber gun barrel.

---

The remainder of this article was reproduced photographically from the author's manuscript.



## 1. INTRODUCTION

Cross-product Bessel functions are frequently encountered in solving Bessel's equation in the problems of heat transfer, electricity, hydrodynamics, and mechanics. Laslett and Lewish<sup>1</sup> studied cross-product Bessel functions of the types

$$J_n(\beta a)Y_n(\beta b) - J_n(\beta b)Y_n(\beta a) = 0 \quad (1)$$

where  $J_n$ ,  $Y_n$  are Bessel functions of the first and second kinds of order  $n$ , respectively, and

$$J'_n(\beta a)Y'_n(\beta b) - J'_n(\beta b)Y'_n(\beta a) = 0 \quad (2)$$

where  $J'_n$ ,  $Y'_n$  are derivatives of the Bessel functions  $J_n$ ,  $Y_n$  respectively, with respect to their total arguments.

Kirkham<sup>2</sup> graphed various combinations of Equations 1 and 2 given above. Extensive tables of roots of Equation 2 are given by Bridge and Angrist.<sup>3</sup> Other authors have solved for the roots of different types of cross-product Bessel functions. A general form of cross-product Bessel functions that could be reduced to many specific cases would obviously be most advantageous. The purpose of this paper is to present several general cross-product Bessel functions that can be reduced to specific cases with proper selection of constants and parameters.

## 2. EVALUATION OF THE BESSEL FUNCTIONS

To solve the various cross-product Bessel functions, one must solve the individual functions  $J_n$  and  $Y_n$ . The first kind of Bessel function,  $J_n(x)$ , is evaluated by use of the recurrence relation<sup>4</sup>

$$F_{n+1}(x) + F_{n-1}(x) = (2n/\alpha)F_n(x) \quad (3)$$

Then the desired Bessel function is given by

$$J_n(x) = F_n(x)/\alpha \quad (4)$$

where

$$\alpha = F_0(x) + 2 \sum_{m=1}^{M-2} F_{2m}(x) \quad (5)$$

M is initialized at  $M_0$ , where  $M_0$  is the greater of  $M_A$  and  $M_B$ , where

$$M_A = \begin{cases} [x + 6]; & x < 5 \\ [1.4x + 60/x]; & x \geq 5 \end{cases} \quad (6)$$

and

$$M_B = [n + x/4 + 2] \quad (7)$$

$F_{M-2}, F_{M-3}, \dots, F_2, F_1, F_0$  are evaluated by use of Equation 3 with  $F_M = 0$  and  $F_{M-1} = 10^{-30}$ . Incrementing M by 3 and again evaluating  $J_n$ , and if

$$\left| J_n(x)_M - J_n(x)_{M+3} \right| \leq \delta \left| J_n(x)_{M+3} \right| \quad (8)$$

is satisfied, then  $J_n$  is within the required accuracy range determined by a given  $\delta > 0$ . If this condition is not met, M is again incremented by 3, and the same procedure is used until the desired accuracy is obtained, with a default value for M given by

$$M_{\max} = \begin{cases} [20 + 10x - x^2/3]; & x \leq 15 \\ [90 + x/2]; & x > 15 \end{cases} \quad (9)$$

Similarly, the recurrence relationship for  $Y_n$  is given by

$$Y_{n+1}(x) = (2n/x)Y_n(x) - Y_{n-1}(x) \quad (10)$$

For  $x > 4$ , the  $Y_0$  and  $Y_1$  Bessel functions are given by the asymptotic relationships

$$Y_0(x) = \sqrt{\frac{2}{\pi x}} (P_0(x)\sin(x-\pi/4) + Q_0(x)\cos(x-\pi/4)) \quad (11)$$

$$Y_1(x) = \sqrt{\frac{2}{\pi x}} (-P_1(x)\cos(x-\pi/4) + Q_1(x)\sin(x-\pi/4)) \quad (12)$$

where  $P_0(x)$ ,  $P_1(x)$ ,  $Q_0(x)$  and  $Q_1(x)$  are defined<sup>5</sup>, but are too lengthy to include in this discussion.

For  $x \leq 4$ ,

$$Y_0(x) = \frac{2}{\pi} \sum_{m=0}^{15} (-1)^m (x/2)^{2m} \frac{1}{(m!)^2} [\text{Log}(x/2) + \gamma - H_m] \quad (13)$$

where

$$H_m = \begin{cases} \sum_{r=1}^m \frac{1}{r}; & m \geq 1 \\ 0 & ; m = 0 \end{cases} \quad (14)$$

and

$$\gamma = \text{Euler's Constant} = .5772156649$$

and

$$Y_1(x) = -\frac{2}{\pi x} + \frac{2}{\pi} \sum_{m=1}^{16} (-1)^{m+1} (x/2)^{2m+1} \frac{1}{m!(m-1)!} \quad (15)$$
$$[\text{Log}(x/2) + \gamma - H_m + \frac{1}{2m}]$$

Then the  $Y_n$ 's for  $n > 1$  can be obtained from Equation 10 for any value of  $x$ . The two subroutines for the solving of Bessel functions are available at the University of Iowa Computer Center in their subroutine package; both subroutines were entirely workable in all cases.

### 3. DISCUSSION OF THE NUMERICAL METHOD

#### A. NEWTON-RAPHSON METHOD

The calculation of the roots of cross-product Bessel functions was accomplished mainly by use of the efficient Newton-Raphson method. The iteration technique for the Newton Method is defined by

$$X_{n+1} = X_n - F(X_n)/F'(X_n); \quad n = 0, 1, 2, \dots \quad (16)$$

The graph of the function  $F$  is approximated by its tangent at the point  $X_n$ , that is,  $F(X)$  is replaced by

$$F(X_n) + (X - X_n)F'(X_n) \quad (17)$$

The first problem encountered with this method was that of finding a good first estimate for  $X$ . Since the shape of the cross-product Bessel functions is not generally known, this first estimate is difficult to make. The problem is avoided by the checking of function values for increasing  $X$

until a sign change in the function, say  $F$ , is detected. At this point, evaluate  $F'$ , the first derivative of  $F$  with respect to its total argument, and then evaluate Equation 16 for a new  $X$ . Then, using this new  $X$ , evaluate  $F$  and  $F'$ , checking the value of  $F$  to determine whether this value is within the desired range of accuracy for a root. If the accuracy requirements are not met, determine a new  $X$  from Equation 16 and repeat the process given above until the root has been found. When the first root has been determined, the initial  $X$  for finding the second root will be  $\text{root}(1)$  plus an arbitrary, small estimate for  $X$ . The identical procedure as followed in finding the first root is used to find the second root. After the first two roots have been obtained, the increment size for  $X$  is set at  $X = \text{root}(2) - \text{root}(1)$ , since the roots are generally periodic. From this point, all roots past the first two are found by use of the Newton-Raphson method only. Thus, the change of sign test is no longer necessary. The convergence is guaranteed with Newton's method; specifically, the convergence is said to be quadratically convergent. That is, for each iteration of  $X$ , the number of correct decimal places in the root are doubled.

#### B. BISECTOR METHOD

If the slope or shape of the cross-product Bessel function is relatively flat, convergence may be on a previous or later root. Thus, another method is necessary to solve this type of function. For lack of a more suitable name, the Bisector method is chosen. The first two roots can generally be found by the Newton method since the sign change is generally used in conjunction with it. Therefore, suppose that the first two roots have been determined and that with the Newton method the third root is not found. Then, return to the second root, reset the increment size of  $X$  to the increment used to find the first two roots, and add this increment to the second root. Thus, the first estimate to find the third root is the second root plus the small increment of  $X$ . Then, use the sign change technique and check the function value  $F$  with increasing  $X$  for a sign change. When the sign change is detected, decrement  $X$  by one increment of  $X$ , then bisect the increment size of  $X$ . Now proceed again with the sign change technique. When the next change is detected, again decrement  $X$  one increment of  $X$ , bisect the increment size, and proceed with the sign change test. Repeated iteration of this process will give us the desired roots. This method was quite workable in all cases tested, the only disadvantage of this method was that the speed of convergence is slower than the Newton method.

Thus, with the computer program developed to solve for the roots of cross-product Bessel functions, any number of roots can be found for virtually any type of cross-product Bessel function. The specific cases can then be fitted into the following general equation forms:

$$F(x) = [C_1 J_n(bx) + C_2 x J_{n+1}(bx)][C_3 Y_n(ax) + C_4 x Y_{n+1}(ax)] \\ - [C_5 J_n(ax) + C_6 x J_{n+1}(ax)][C_7 Y_n(bx) + C_8 x Y_{n+1}(bx)] \quad (18)$$

and

$$F(x) = [C_1 J'_n(bx) + C_2 x J'_{n+1}(bx)][C_3 Y'_n(ax) + C_4 x Y'_{n+1}(ax)] \\ - [C_5 J'_n(ax) + C_6 x J'_{n+1}(ax)][C_7 Y'_n(bx) + C_8 x Y'_{n+1}(bx)] \quad (19)$$

where  $C_1, C_2, \dots, C_8$  are constants or parameters, and  $J_n, Y_n, J'_n$  and  $Y'_n$  are as defined previously. If the cross-product Bessel function cannot be fitted into the general form of Equations 18 and 19, one could simply supply a new  $F$  and  $F'$ . Several functions have been tabulated to illustrate how one can determine any number of roots for any order  $n$ . Each table gives but only partially because of lack of space, the magnitude of  $n$  and the number of roots.

#### 4. ROOTS OF SPECIFIC CROSS-PRODUCT BESSEL FUNCTIONS

Suppose one sets  $C_1 = C_5 = -1, C_3 = C_7 = 1,$  and  $C_2 = C_4 = C_6 = C_8 = 0$  in Equation 18. One obtains identically Equation 1. Since this specific cross-product Bessel function is commonly found in various fields, the roots of Equation 1 are tabulated for order  $n = 0$  through  $n = 10$  in Table I.

Another common type of cross-product Bessel function that has identical coefficients as assigned above, but applied to Equation 19 is tabulated in Table II.

Thus, with the use of the computer program developed in this study, one can find any number of roots of a given cross-product Bessel function.

N	Root(1)	Root(2)	Root(3)	Root(4)	Root(5)	Root(6)	Root(7)	Root(8)	Root(9)	Root(10)
0	3.123	6.273	9.418	12.561	15.704	18.846	21.988	25.130	28.272	31.414
1	3.197	6.312	9.444	12.581	15.720	18.860	22.000	25.140	28.281	31.422
2	3.407	6.428	9.523	12.640	15.767	18.899	22.034	25.170	28.307	31.446
3	3.729	6.616	9.652	12.738	15.846	18.965	22.090	25.220	28.352	31.485
4	4.133	6.871	9.831	12.875	15.956	19.057	22.169	25.289	28.413	31.541
5	4.595	7.187	10.056	13.048	16.096	19.175	22.271	25.378	28.492	31.612
6	5.094	7.555	10.326	13.257	16.267	19.318	22.394	25.486	28.589	31.699
7	5.618	7.969	10.637	13.501	16.466	19.486	22.539	25.613	28.702	31.802
8	6.156	8.421	10.986	13.778	16.693	19.678	22.705	25.760	28.833	31.920
9	6.701	8.904	11.370	14.086	16.947	19.894	22.892	25.925	28.980	32.053
10	7.251	9.412	11.786	14.423	17.228	20.133	23.100	26.108	29.144	32.201

TABLE I - ROOTS OF  $J_n(x)Y_n(2x) - J_n(2x)Y_n(x) = 0$

N	Root(1)	Root(2)	Root(3)	Root(4)	Root(5)	Root(6)	Root(7)	Root(8)	Root(9)	Root(10)
0	3.197	6.312	9.444	12.581	15.720	18.860	22.000	25.140	28.281	31.422
1	0.677	3.282	6.353	9.471	12.601	15.736	18.873	22.011	25.150	28.390
2	1.341	3.531	6.475	9.552	12.661	15.784	18.913	22.045	25.180	28.316
3	1.979	3.920	6.674	9.684	12.761	15.863	18.979	22.102	25.230	28.361
4	2.588	4.418	6.946	9.868	12.899	15.974	19.072	22.181	25.299	28.422
5	3.169	4.993	7.287	10.100	13.075	16.116	19.190	22.283	25.388	28.502
6	3.731	5.613	7.691	10.379	13.288	16.288	19.334	22.407	25.497	28.598
7	4.279	6.255	8.154	10.702	13.536	16.489	19.503	22.553	25.625	28.712
8	4.819	6.898	8.670	11.068	13.819	16.719	19.696	22.720	25.772	28.843
9	5.354	7.532	9.234	11.475	14.134	16.976	19.914	22.908	25.937	28.991
10	5.884	8.152	9.834	11.920	14.481	17.261	20.155	23.117	26.121	29.156

TABLE II - ROOTS OF  $J'_n(x)Y'_n(2x) - J'_n(2x)Y'_n(x) = 0$



## LITERATURE CITED

1. Laslett, L. J., and Lewish, W., "Evaluation of the Zeroes of Cross-Product Bessel Functions," Iowa State University, Ames, Iowa, 1960, Ames Laboratory Report IS-189.
2. Kirkham, D., "Graphs and Formulas for Zeroes of Cross-Product Bessel Functions," Iowa Agricultural Experiment Station, Ames, Iowa, Project 998, Journal Paper J-2998.
3. Bridge, J. F., and Angrist, S. W., "An Extended Table of Roots of  $J'_n(x)Y'_n(\beta x) - J'_n(\beta x)Y'_n(x) = 0$ ," Department of Mechanical Engineering, Ohio State University, 1961.
4. Goldstein, H., and Thaler, R. M., "Recurrence Techniques for the Calculation of Bessel Functions," M.T.A.C., V. 13, pp. 102-108.
5. Stegun, I. A., and Abramowitz, M., "Generation of Bessel Functions on High Speed Computers," M.T.A.C., V. 11, 1957, pp. 255-257.



APPLICATION OF PATTERN RECOGNITION TO SHOCK-TRAUMA STUDIES

W. S. Copes  
US Army Materiel Systems Analysis Agency  
Aberdeen Proving Ground, Maryland

A. R. Cowley  
University of Maryland Hospital  
Baltimore, Maryland

C. Masaitis  
Ballistic Research Laboratories  
Aberdeen Proving Ground, Maryland

W. J. Sacco  
Biomedical Laboratory  
Edgewood Arsenal, Maryland

A. V. Milholland  
Biomedical Laboratory  
Edgewood Arsenal, Maryland

**ABSTRACT.** This paper reports the results of a study to determine pattern vectors (Profiles) composed of physiological and biochemical measurements which reflect the severity of injury to traumatized individuals. Profiles, selected by clinicians at the University of Maryland Center for the Study of Trauma, were obtained from the Center data bank and subjected to pattern analyses using OLPARS, an on-line pattern analysis and recognition system, located at Rome Air Development Center. Prognosis regions were delineated in the Eigenvector Plane and the Discriminant Plane. The time courses of individual patients were plotted in the Eigenvector Plane.

**INTRODUCTION.** Shock is usually associated with severe injury to the soft tissues, the skeleton, and to specific organs. Tissue injury, hemorrhage, and pain cause a multidimensional and widespread body response to injury which can involve every organ system within the body. Moreover, the responses are interconnected in a very complex way.

The Center for the Study of Trauma at the University of Maryland Hospital was established to study the effect of inadequate tissue perfusion induced by injury at the organ and tissue level by assessing physiological and biochemical responses.

In support of this objective, several analyses using pattern recognition techniques have been conducted by clinicians and researchers from the Center, together with analysts from the Biomedical Laboratory (Edgewood Arsenal, Maryland) and the Army Materiel Systems Analysis Agency (Aberdeen Proving Ground, Maryland).

A data bank at the Center contains clinical, cardiovascular, metabolic, and therapeutic data on over a thousand patients.

The initial step in the study was to determine pattern vectors composed of physiological and biochemical measurements which reflect the severity of a patient's traumatic state, that is, to determine prognosis regions of the pattern space and to analyze the time course of patients as a function of therapy.

A candidate pattern profile, selected by the clinicians, was subjected to pattern analysis routines, using OLPARS (an On-Line Pattern Recognition System belonging to Rome Air Development Center), in an effort to find some structure in the data and to delineate various prognosis regions. The pattern profile consisted of 12 measurements, so in this instance, the condition of each patient was characterized by a 12-dimensional vector  $X = (x_1, x_2, \dots, x_{12})$ . From the data bank profiles were retrieved on 140 patients, 70 of whom ultimately recovered, and 70 of whom ultimately died in the Center. The methods and results of the analyses will now be described.

## METHODS

### Patient Sample and Data

Included in this study were initial and final measurements from 140 patients, 70 of whom died, and 70 of whom survived. For each patient, the first measurement on each variable was called the initial value and the last measurement before death or discharge, the final value. The set of 12 measurements used in this study\* were systolic blood pressure (SBP), diastolic blood pressure (DBP), hemoglobin (Hgl), hematocrit (Hmt), serum fibrinogen (Fib), serum sodium (Na), serum potassium (K), serum chloride (Cl), serum osmolality (Osm), blood urea nitrogen (BUN), glucose (Gl), and serum creatinine (Cr).

The set of measurements will be called a profile or a patten vector. Throughout the study the vectors were considered as belonging to classes A, B, L, and D where:

A is the set of 70 vectors composed of final measurements on the surviving patients;

---

\*Results of a study of organ system profiles will be reported in another paper.

B is the set of 70 vectors composed of final measurements on the dying patients;  
 L is the set of 70 vectors composed of initial measurements of the surviving patients; and  
 D is the set of 70 vectors composed of initial measurements of the dying patients.

Subsequent to the start of the study nine vectors from Class A were discarded. These vectors came from patients who were still seriously ill when they left the Center

### Analyses

In this study Class A was considered to be the control set. Indeed, the vectors in Class A are the final measurements of patients who recovered and should be close (and in this case are) to normal values. The means and standard deviations for the measurements in Class A are given in Table 1.

All of the data were normalized with respect to the vectors of Class A. More specifically let the pattern vector  $X = (x_1, \dots, x_{12})$ . Let  $x_i^A$  and  $S_i^A$  be the mean and standard deviation of  $x_i$  with respect to Class A. Then each X vector was transformed into a new X vector whose component  $x_i^{\text{new}}$  was defined by the relationship

$$x_i^{\text{new}} = \frac{x_i^{\text{old}} - x_i^A}{S_i^A}$$

The normalized mean values for each class and the total data set are given in Table 2.

### Profile Analyses

The normalized vectors from all four classes were subjected to a detailed structure analysis. Several mathematical transformations including Non-Linear Mapping, Eigenvector Plane Mapping, and Discriminant Plane Mapping were used in an effort to uncover some revealing aspects of the inherent data structure. For this data set the latter two mappings appeared to provide the better settings for delineating prognosis regions. The Eigenvector Plane Mapping projects the original data onto a plane which "best" fits the data in the linear least squares sense. Minimizing the sum of the squared distance from a two-dimensional subspace of the original space requires the solution for the eigenvectors of the lumped data covariance matrix. The eigenvector plane is defined by the two eigenvectors  $E_1, E_2$  corresponding to the two largest eigenvalues of the

matrix. The Discriminant Plane Mapping projects two-class data onto a plane which enhances discrimination between data classes. The plane is defined by the discriminant vector  $D_1$  (which is the direction which maximizes the projected between class scatter relative to the sum of projected within class scatter) and a vector  $D_2$ , where  $D_2$  is that direction orthogonal to  $D_1$  which maximizes the projected between class scatter relative to the sum of the projected within class scatter. For a more complete description of these techniques see Reference 1. The eigenvectors  $E_1$  and  $E_2$  are given in Table 3. The discriminant vectors for the four classes taken two at a time (namely, A-B, A-L, A-D, B-L, B-D, L-D) are given in Table 4. All of the eigenvectors and discriminant vectors are unit vectors. Therefore, the mapping of a vector  $X$  to a point  $(y_1, y_2)$  in the Eigenvector Plane, say, is given by the scalar products  $y_1 = E_1 \cdot X$ ,  $y_2 = E_2 \cdot X$ . The Eigenvector Plane Plot is given in Figure 1. Figure 2 gives a count of the number of vectors from each class which are located in each square of the grid appearing in Figure 1. The Discriminant Plane Plots for pairs A-B and A-D are given in Figures 3 and 4.

#### Measurement Reduction

In solving a pattern classification problem it is desirable to use the minimum number of measurements to achieve a satisfactory solution. The OLPARS system provides two functionals (Discriminant Measure and Confusion Probability) for ranking the measurements.

The functionals were used to evaluate the discriminatory power of each individual measurement. The measurements which ranked high with respect to both functionals were systolic blood pressure, hematocrit, fibrinogen, potassium, osmolality, and creatinine. The original profile was reduced to these six measurements. Structure analyses were repeated on the reduced vectors. Little loss in discriminating among various data classes was noted by comparing the Discriminant Plane Plots. For example, Figure 5 is the Discriminant variables. The reader may compare this plot with Figure 3, the corresponding plot based on 12 variables.

#### Delineation of Good and Poor Prognosis Regions and Patient Trajectories

Figure 6 gives the Eigenvector Plane Plot for the six-dimensional profile. This plot was used to delineate prognosis regions, with good (G) and poor (P) regions as indicated. Region G contains 46 vectors from Class A, 2 vectors from B, 31 vectors from L, and 14 vectors from D. The regions labeled with a P contain 14 vectors from A, 68 from B, 39 from L, and 56 from D.

The Eigenvector Plane is also being used to observe the time courses of patients. Indeed, let  $X_1, X_2, \dots$ , be a time-ordered sequence of pattern vectors obtained from a patient. Let  $Y_1, Y_2, \dots$ , be the corresponding points in the Eigenvector Plane. We shall call this latter sequence a trajectory. From the trajectory we can observe a patient's current state and the rate of change of the state. From a family of trajectories we may compute transition probabilities from region to region as functions of therapy. Figures 7 to 10 show sample trajectories of daily profiles of various patients some who survived (designated S) and some who expired (designated D). Each trajectory is labeled  $1, 2, \dots$ , for Day<sub>1</sub>, Day<sub>2</sub>, ..., respectively.

SUMMARY AND OBSERVATIONS. A pattern profile consisting of 12 physiological and biochemical measurements was used to reflect the severity of a patient's traumatic state. Prognosis regions were delineated in the Eigenvector Plane and Discriminant Plane based on initial and final measurements from 140 patients, 70 of whom ultimately recovered and 70 of whom died in the Center. The original profile was reduced to 6 variables with little apparent alteration of prognosis regions. The Eigenvector Plane Plot based on the 6-dimensional profile was used to exhibit sample trajectories of several patients.

The Discriminant Plane may be used as well for plotting trajectories. Indeed, we are currently converting the data from over 1,000 patients into trajectories in the Eigenvector Plane and the Discriminant Plane. In addition, we are computing trajectories based on "distance" from normality using original variables for various profiles. Some of this work is directed toward isolating a small but efficient set of measurements which can be used as a basis for initiating and evaluating therapies. We call this set an "Action Profile." Clinicians are polled to determine their preferences. Since each clinician usually has his own priorities, we would like to establish the smallest set which is sufficient to regulate, say, 90% of patients suffering from trauma.

#### REFERENCES

1. J. W. Sammon, "Interactive pattern analysis and classification," IEEE Transactions on Computers, Volume C-19, Number 7 (1960).
2. J. Siegal, R. Goldwyn, and M. Friedman, "Pattern and process in the evolution of human septic shock," Surgery, Volume 70, Number 2 (1971).

TABLE 1'

Means and Standard Deviations for  
Measurements from Class A

	<u>MEAN</u>	<u>STANDARD DEVIATION</u>
Systolic Blood Pressure	129.5	16.3
Diastolic Blood Pressure	79.14	12.7
Hemoglobin	12.24	2.03
Hematocrit	35.87	5.79
Fibrinogen	350.0	139.0
Sodium	141.5	6.37
Potassium	4.426	0.910
Chloride	100.6	8.85
Osmolality	302.2	18.4
Blood Urea Nitrogen	22.7	16.7
Glucose	128.7	59.2
Creatinine	1.40	1.38

TABLE 2  
Normalized Mean Values

	CLASS/MEASUREMENT				
	<u>Total Data Set</u>	<u>Class L</u>	<u>Class D</u>	<u>Class A</u>	<u>Class B</u>
SBP	-1.609	-1.301	-1.423	0	-3.505
DBP	-.9356	-.7051	-.9724	0	-1.990
Hgl	-.1418	.1684	.06782	0	-.7850
Hmt	-.1332	.2643	.07497	0	-.8549
Fib	-.4457	-.5592	-.4726	0	-.6957
Na	-.1461	-.1113	-.3868	0	-.06754
K	.1707	-.1953	.09207	0	.7641
Cl	-.2870	-.00326	-.3836	0	-.7243
Osm	1.107	.3724	1.435	0	2.480
BUN	1.018	.3448	1.376	0	2.220
G1	1.041	.9485	1.372	0	1.708
Cr	1.062	.4428	1.175	0	2.495



TABLE 3  
Eigenvectors

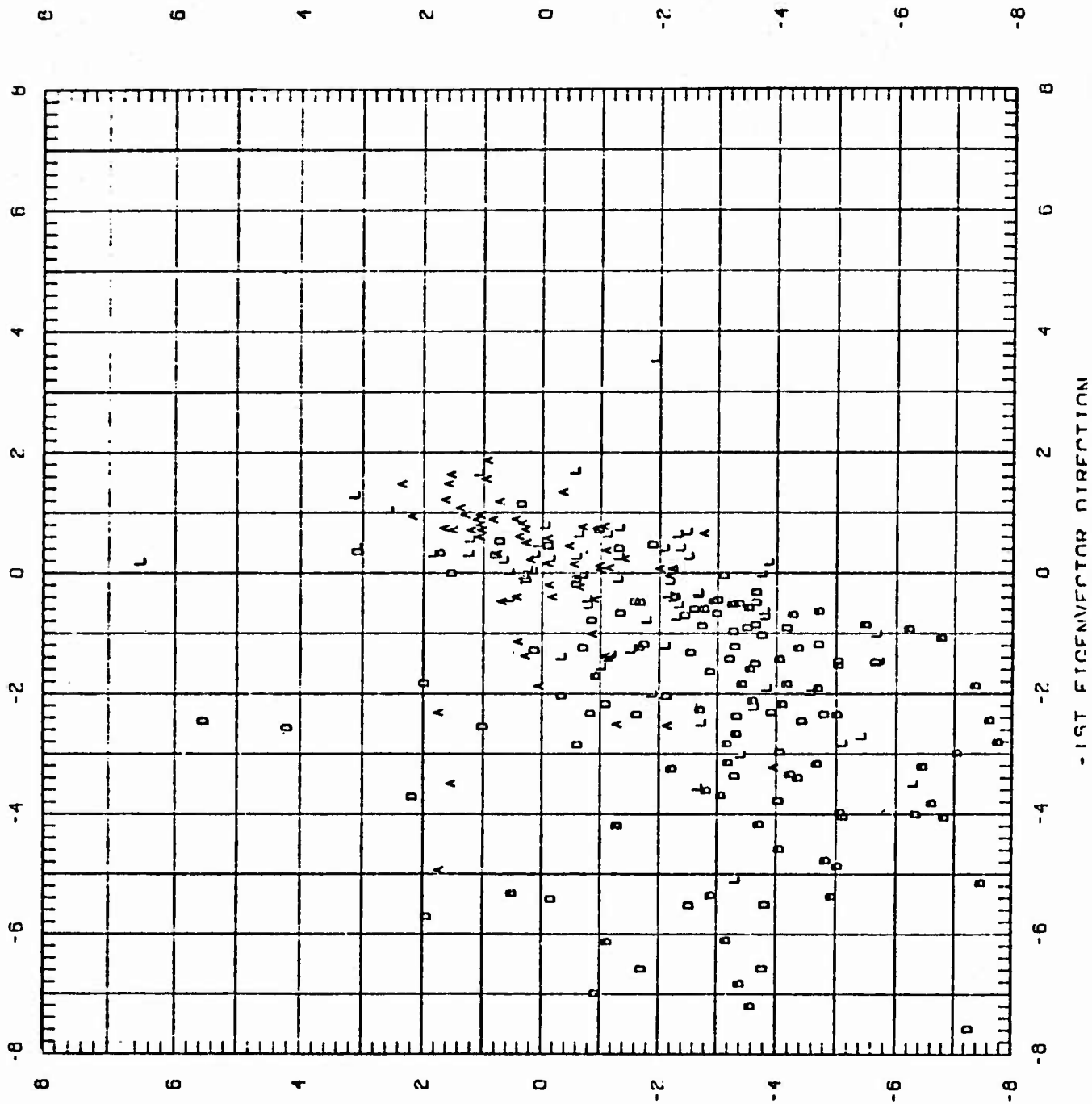
<u>MEASUREMENTS</u>	$E_1$	$E_2$
SBP	.145	.600
DBP	.0936	.595
Hg1	.0739	.121
Hmt	.0857	.0926
Fib	-.0136	.0525
Na	-.0796	-.0488
K	-.149	.0252
Cl	.0141	-.0238
Osm	-.353	-.182
BUN	-.378	-.0909
Gl	.0525	-.318
Cr	-.811	.338



TABLE 4 - Discriminant Vectors

MEASUREMENT/ CLASS PAIRS	A-B		A-L		A-D	
	<u>D<sub>1</sub></u>	<u>D<sub>2</sub></u>	<u>D<sub>1</sub></u>	<u>D<sub>2</sub></u>	<u>D<sub>1</sub></u>	<u>D<sub>2</sub></u>
SBP	.847	.121	-.364	-.203	-.431	-.0835
DBP	-.120	.115	.0396	-.136	-.123	-.387
Hg1	-.0106	.671	-.154	.671	.0268	-.167
Hmt	.198	-.648	.383	-.568	-.00573	.224
Fib	.169	.141	-.443	-.249	-.489	.366
Na	.254	-.0626	-.00577	-.0315	-.174	-.180
K	.0592	-.146	-.421	-.0591	-.141	.0716
Cl	.115	.102	-.0883	-.113	-.163	-.289
Osm	-.342	-.0798	-.0372	-.0597	.0136	.392
BUN	-.0221	-.154	.0280	-.0140	.265	.454
Gl	-.0372	-.129	.435	.257	.228	.287
Cr	-.0231	-.00698	.349	.136	.601	-.263
	B-L		B-D		L-D	
	<u>D<sub>1</sub></u>	<u>D<sub>2</sub></u>	<u>D<sub>1</sub></u>	<u>D<sub>2</sub></u>	<u>D<sub>1</sub></u>	<u>D<sub>2</sub></u>
SBP	.442	.439	.689	.434	-.0788	-.448
DBP	-.00652	-.0125	-.0835	-.0497	.319	.226
Hg1	-.382	.623	-.151	.761	-.462	.530
Hmt	.587	-.399	.605	-.430	.507	-.471
Fib	.128	.111	.224	.115	.0139	.0376
Na	.173	.128	.0733	-.00767	.204	.226
K	-.193	-.118	-.222	-.127	-.123	.0274
Cl	.333	.256	.135	.0835	.375	.260
Osm	-.330	-.353	-.0266	-.0592	-.274	-.274
BUN	-.0842	-.139	-.0608	-.0646	-.382	-.485
Gl	-.0627	-.0892	-.0111	-.0285	-.0593	-.113
Cr	-.000989	-.000203	-.0498	-.0334	-.0215	-.0380

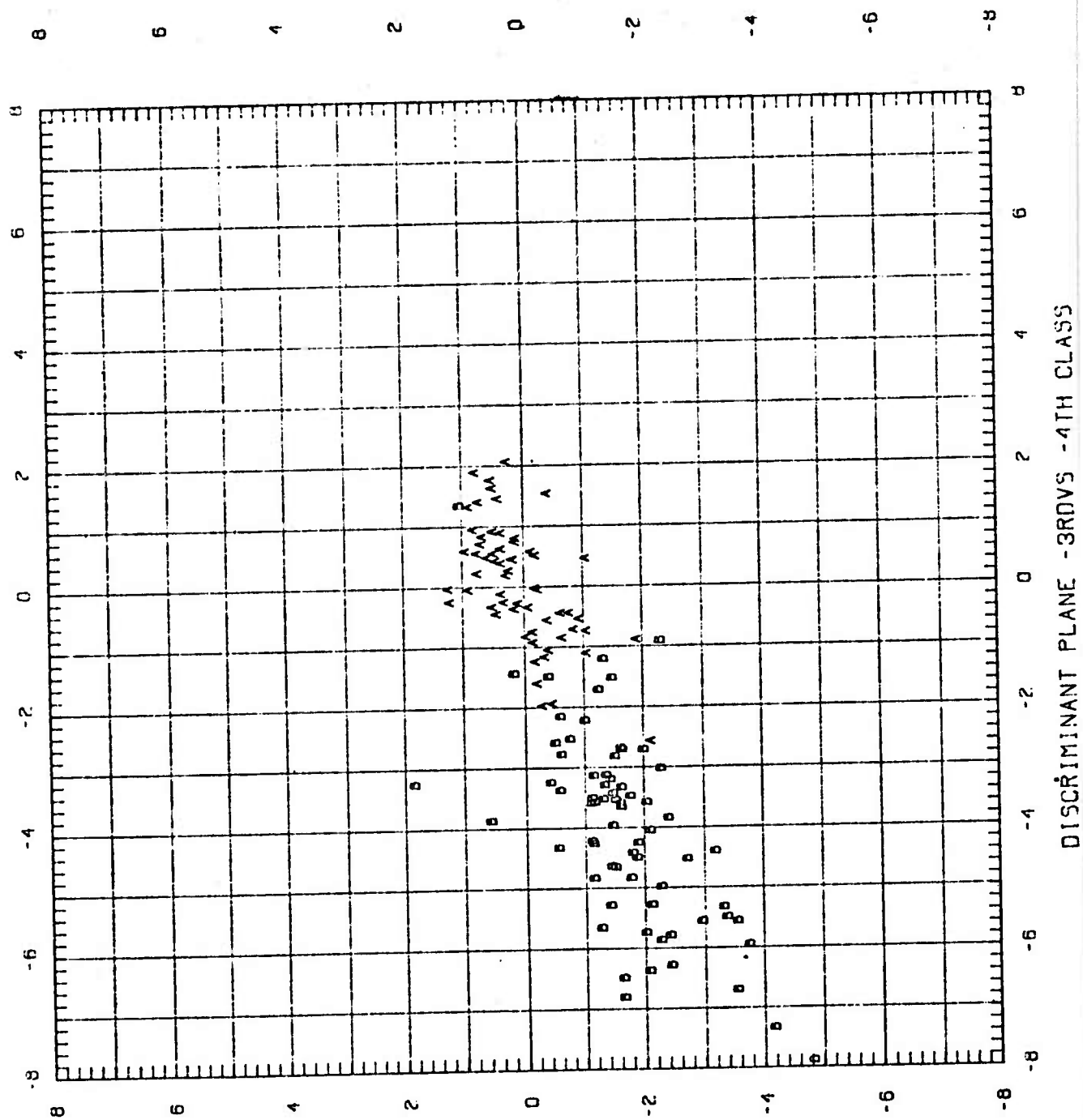
ORIG 12 DIM(SYS. DIAS. HEMO. HEMAT. FIBR. NA. K. CL. OSMO. BUN. GLUCOSE. CREAT)



- FIGURE 1 -

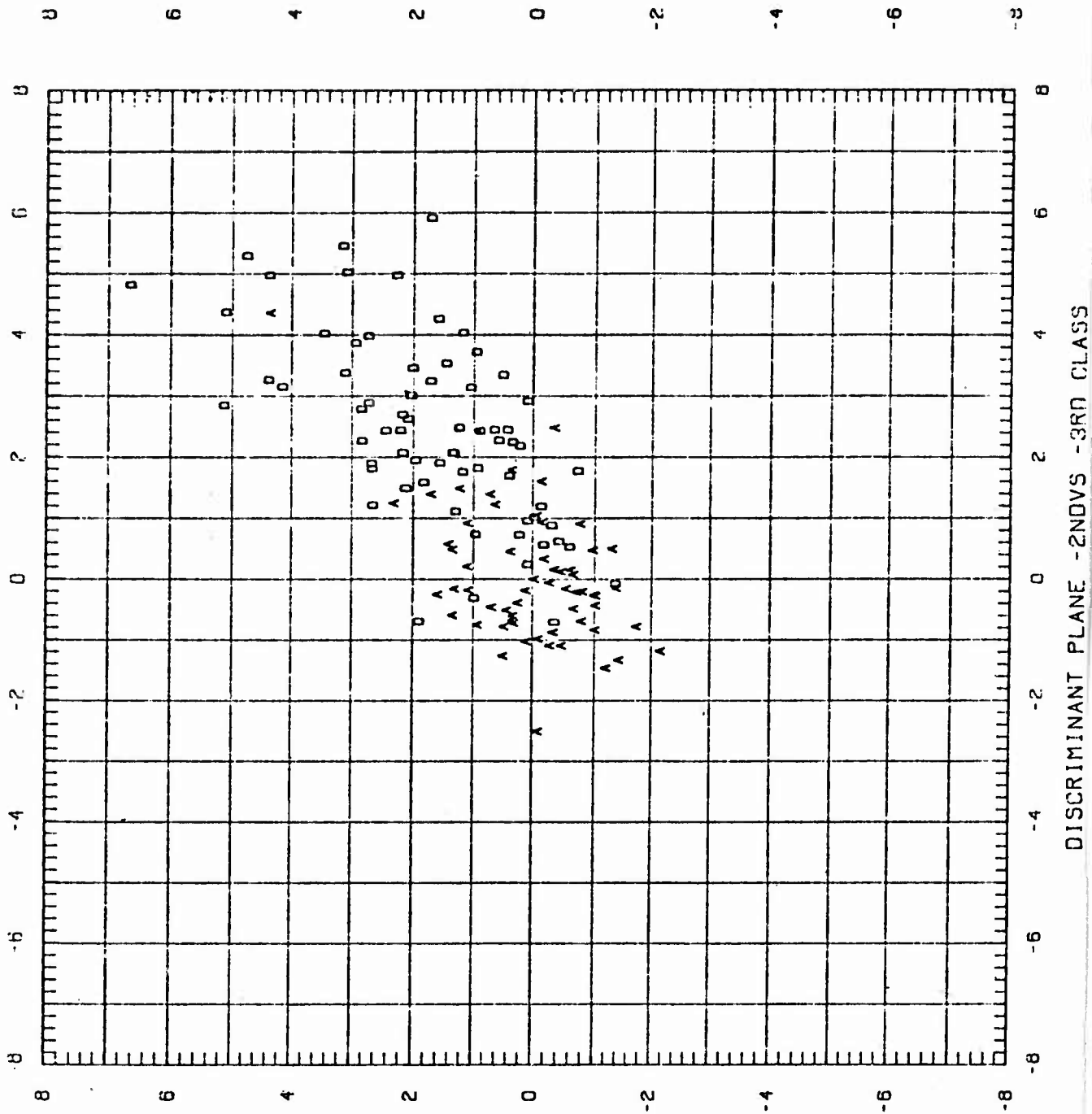


ORIG 12 DIM(SYS.DIAS,HEMO,HEMAT,FIBR,NA,K,CL,OSMC,RUN,GLUCOSE,CREAT)



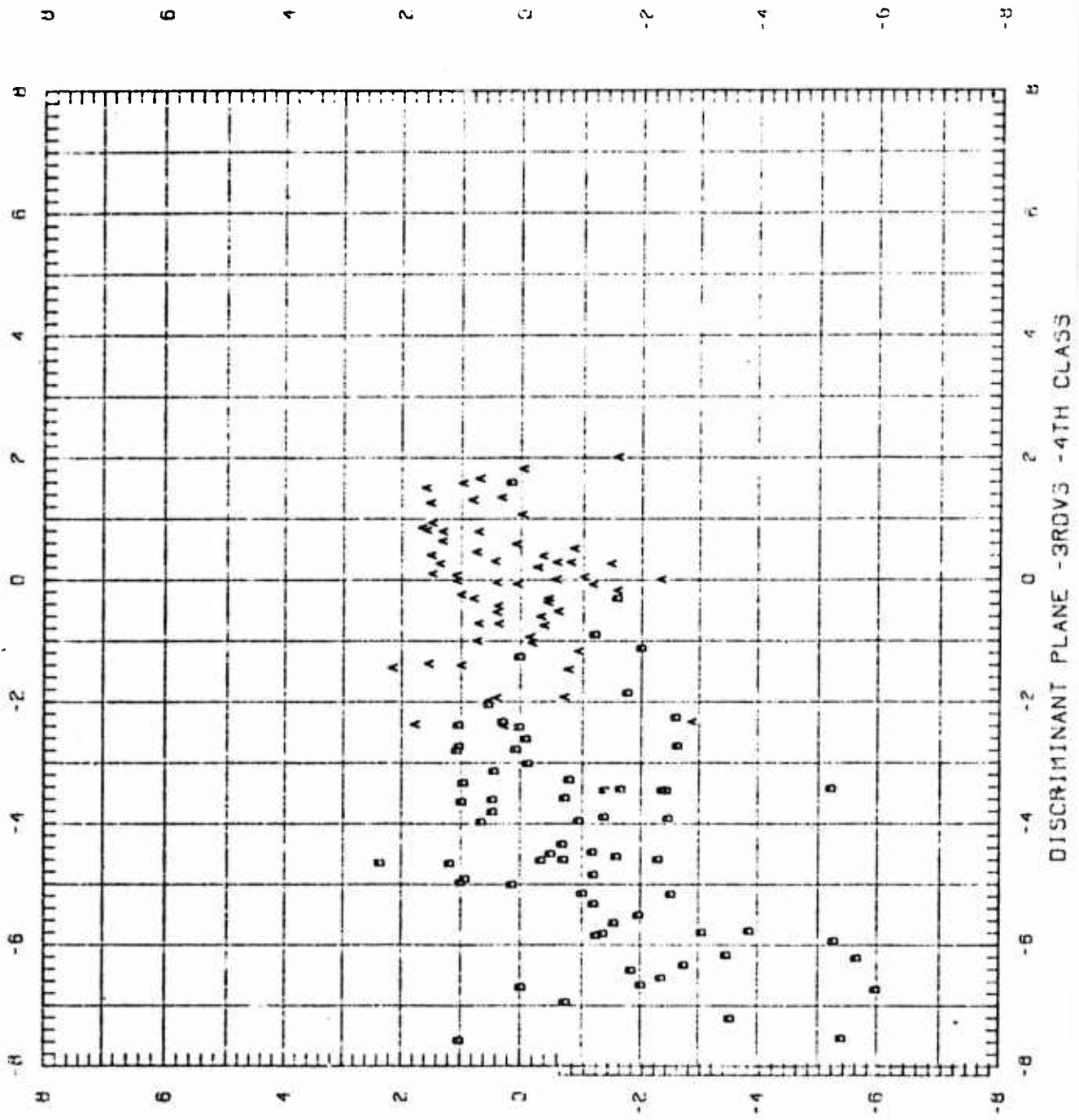
- FIGURE 3 -

ORIG 12 DIM(SYS. DIAS. HEMO. HEMAT. FIBR. NA. K. CL. OSMO. BUN. GLUCOSE. CREAT)



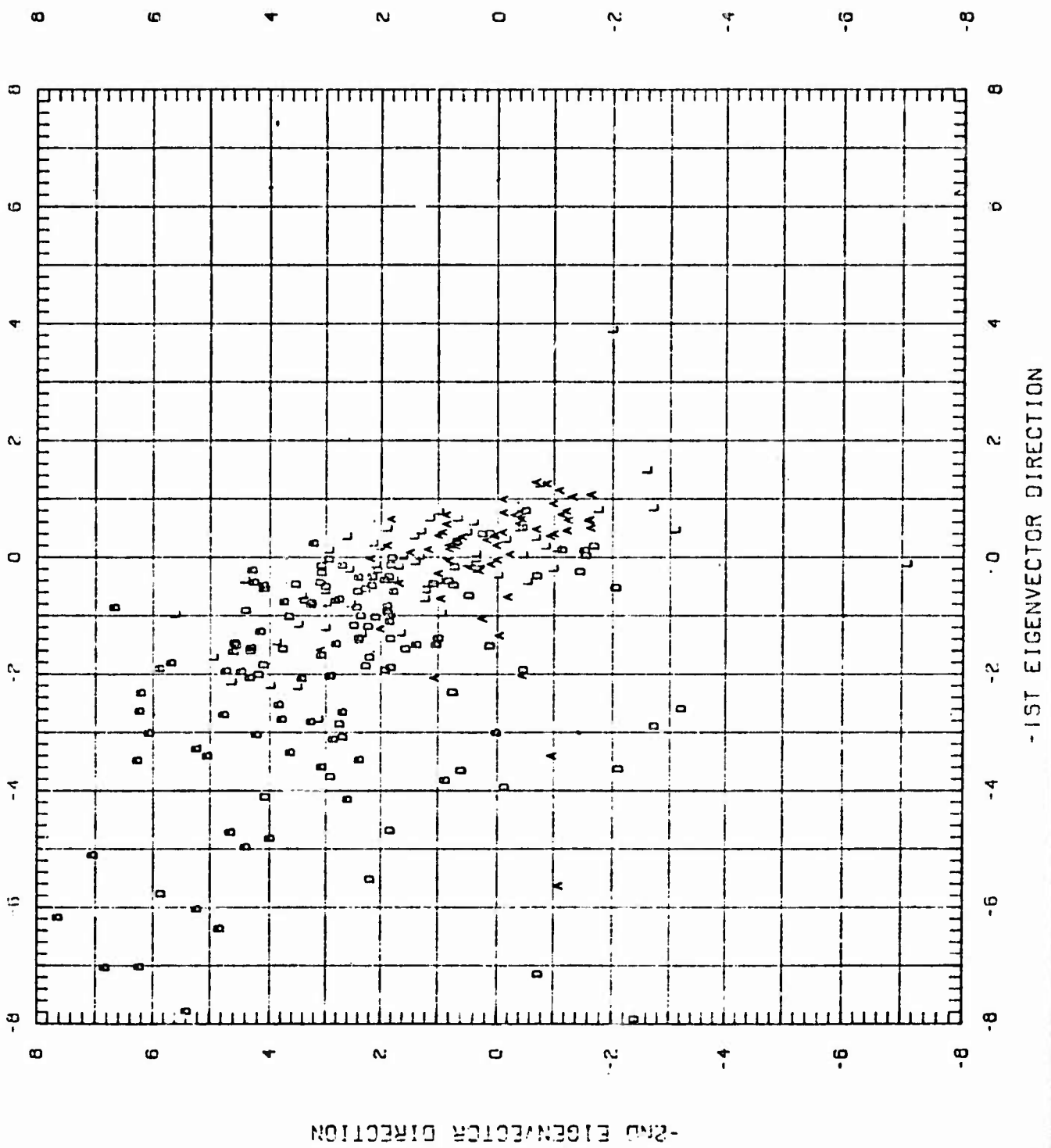
-FIGURE 4-

SIX FROM TWELVE (SYS BP, K. CREAT, OSMOLAL, FIBRIN, HEMATO)



- FIGURE 5 -

SIX FROM TWELVE (SYS BP, K.CREAT, OSMOLAL, FIBRIN, HEMATO)



- FIGURE 6 -



PATIENT NO. 400034D

FIGURE 7

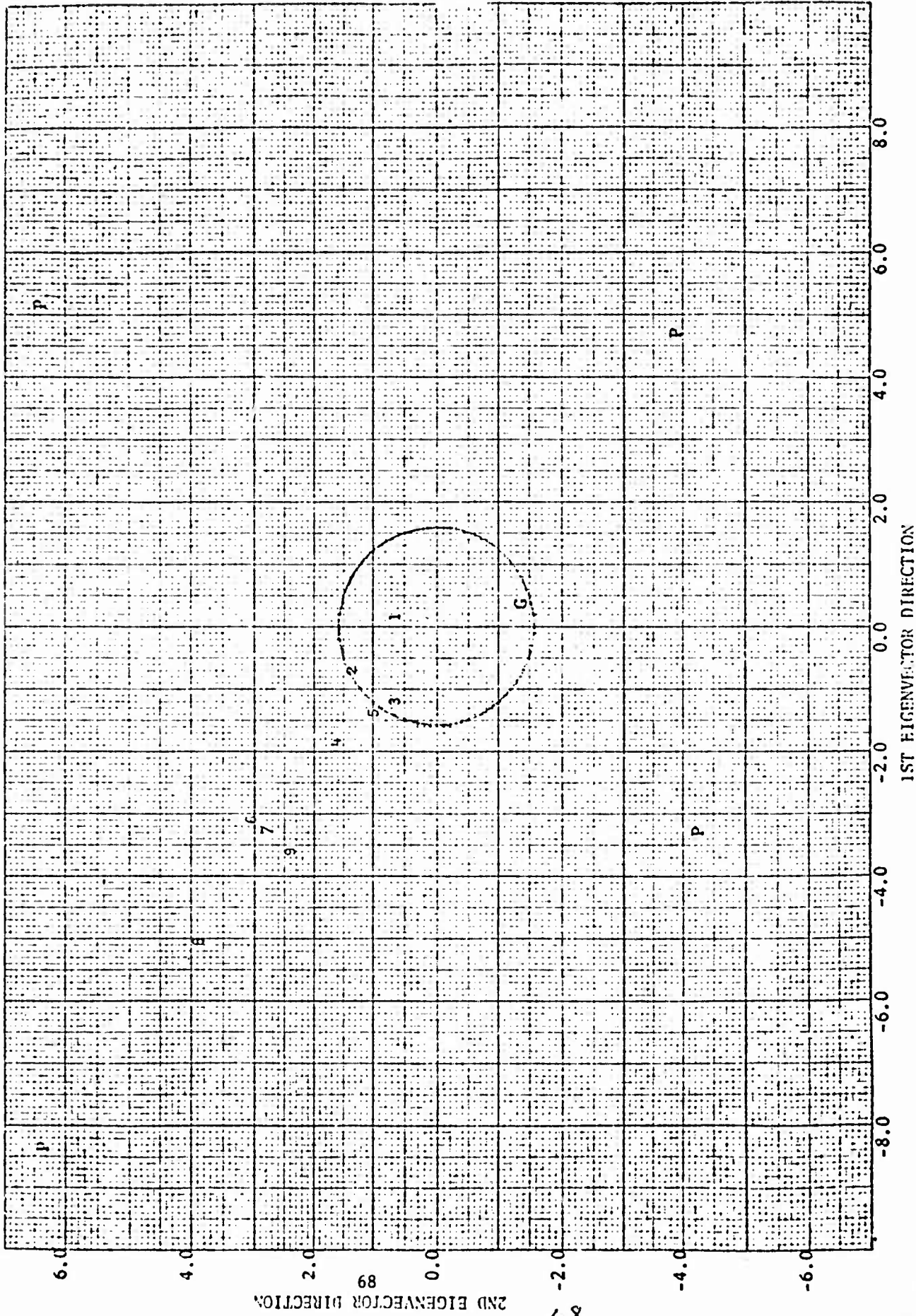
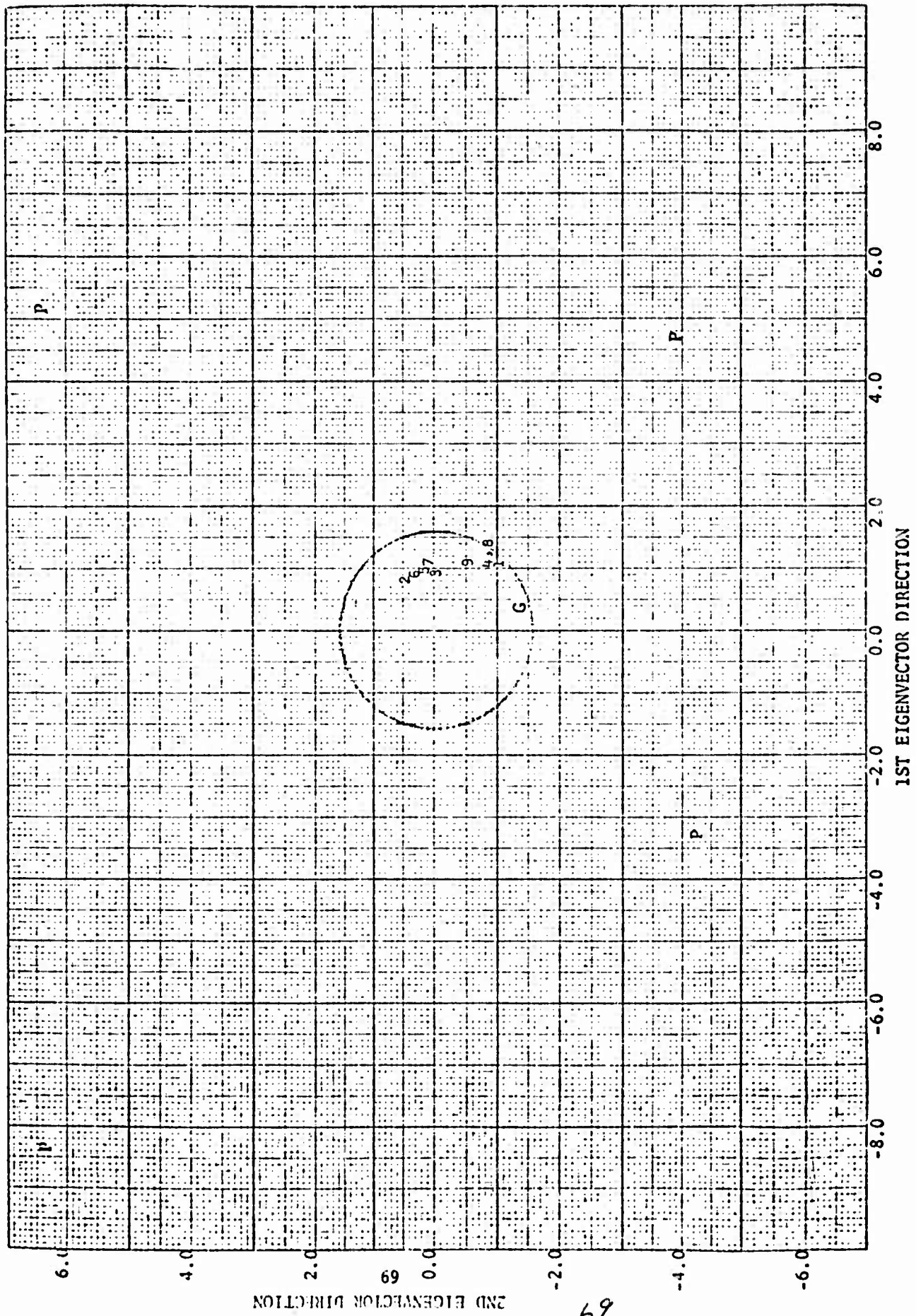


FIGURE 8



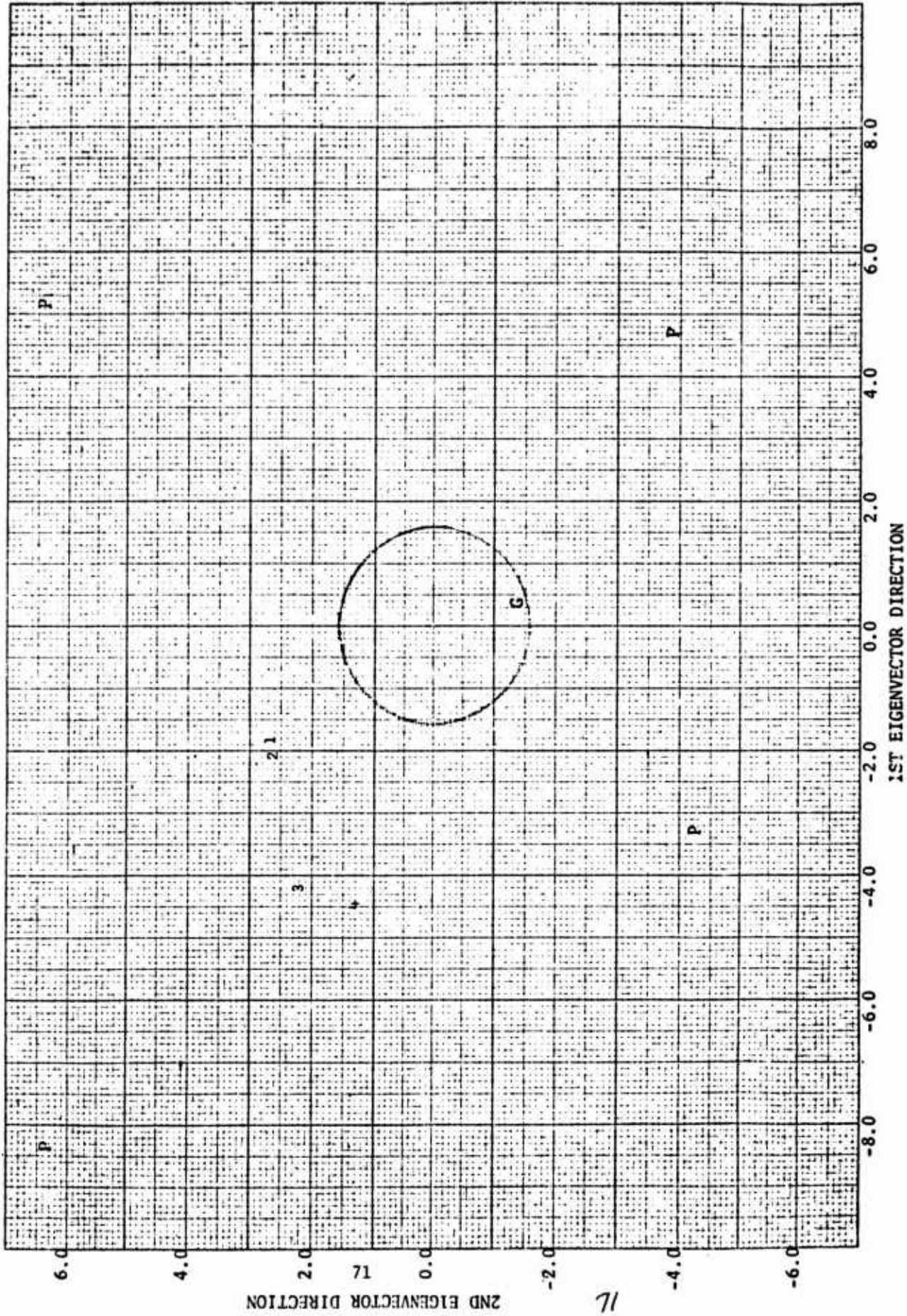
69





CST 1370 D

FIGURE 10



AERODYNAMIC PARAMETER IDENTIFICATION  
IN BALLISTIC RANGE TESTS

Gary T. Chapman

Ames Research Center, NASA  
Moffett Field, California

INTRODUCTION. In general, the problem of aerodynamic parameter identification has as its goal the determination of certain aerodynamic forces and moments that will be used to predict the dynamic behavior of some full-scale vehicle or projectile as it flies through the atmosphere. Here then the problem of parameter identification must not only be concerned with the problem of determining the aerodynamics in laboratory tests, but must also consider how those results can be applied to a geometrically similar but larger vehicle flying at some altitude above sea level. The purpose of this paper is to consider a "real" world problem and trace through the various steps that may be required to obtain the aerodynamic data needed to calculate the dynamic behavior of the vehicle. In this approach I will concern myself with both the physical world and a mathematical world. A simplified box diagram of the basic process that will be discussed is shown in Figure 1. The left-hand series of boxes represents physics, the right-hand boxes mathematics. The physical world problem is at the top left, with its corresponding mathematical counterpart on the top right. As we move down the chart to the second box, we simplify the problem until we arrive at a satisfactory experiment and corresponding mathematical model, that is depicted by the third box down. The data from the experiment are combined with the mathematical model in what is classically called the parameter identification step. From this step we get not only results that will be applied to the real problem but also information on experimental errors that can be fed back into the experiment to improve the data and information about the accuracy of the mathematical model that should be fed back to improve the mathematical model. Note also that there are dotted lines connecting the physical side of the chart to the mathematical side. This is used to indicate that the model selection process arises from the joint flow of information back and forth between laboratory and mathematical considerations.

In the following sections we will discuss briefly the pressures that cause one to move to simpler modeling (both mathematical and physical). We will then discuss in detail the interacting of the experiment and mathematical model through parameter identification. Where possible, concrete examples of all the steps will be given. In any one case all of the steps are not followed; hence a single set of data cannot be followed throughout, but rather the various ideas are illustrated by representative data. Much of the information we will describe is contained in detail in Reference 1.

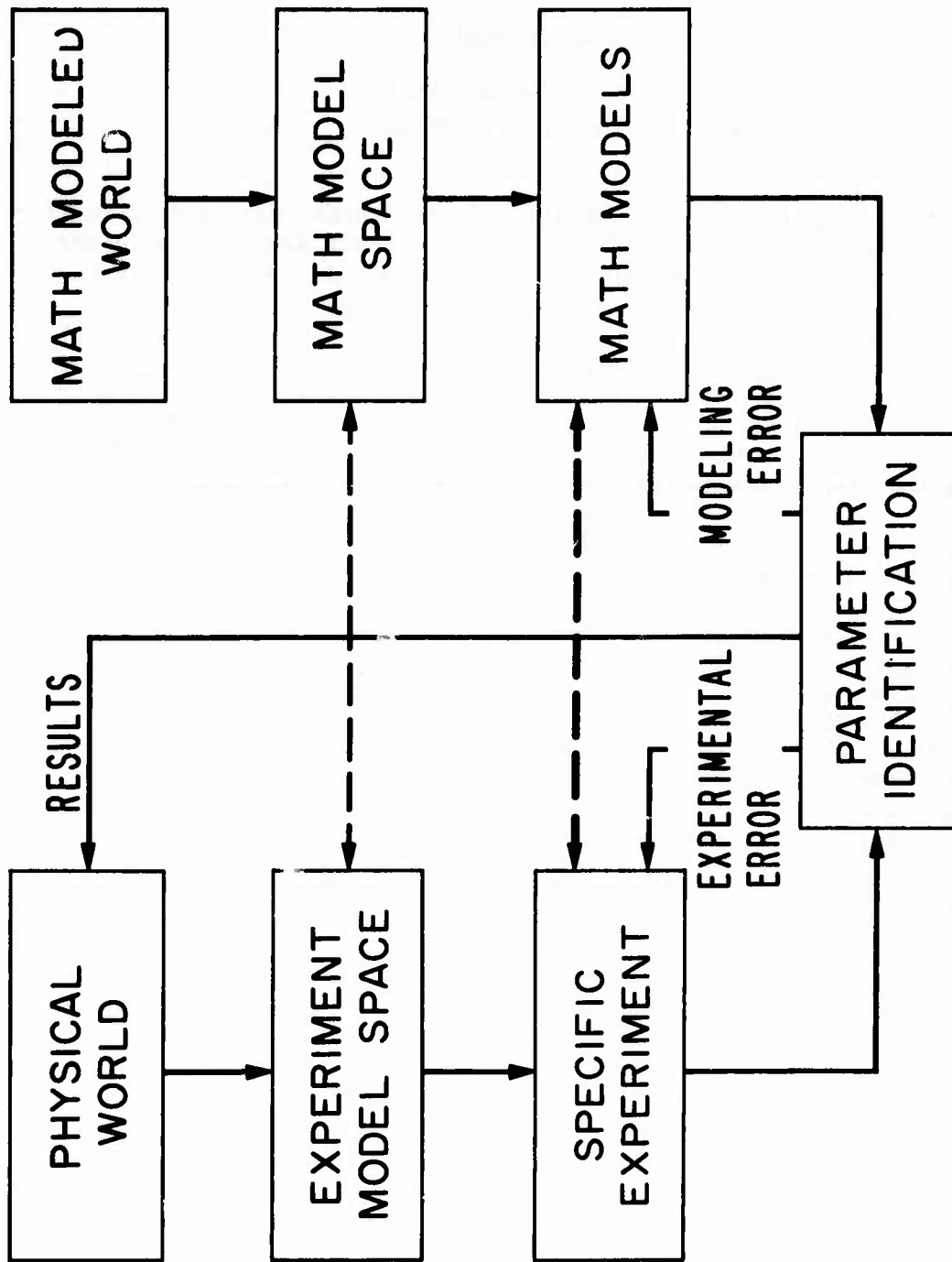


Fig. 1. Problem overview.

PROBLEM DEFINITION. The physical problem we will try to find answers to will be: What are the forces and moments that act on a full scale vehicle flying in an arbitrary atmosphere? One's first thought might be that there must be theoretical prediction techniques available to compute the aerodynamic forces and moments, and there are, but they are not sufficiently good at this time to risk a multimillion dollar vehicle on the results from them. The next question may be: Would it be possible to construct a theoretical procedure that would be accurate enough? The answer is yes in principle, but we lack the large-scale computer required to make the calculation. The ILLIAC IV computer presently being installed at Ames Research Center may be able to provide such results but, even there, we would want some supporting information particularly for turbulent or separated flows. Hence, we are left with the need to test, but not the full-scale vehicle for it is expensive. Therefore we must consider smaller-scale experiments. This step away from the real physical problem takes us into what will be referred to here as the experimental modeling space and correspondingly the mathematical modeling space.

#### Modeling Space

We referred to the step of subscale modeling as a modeling space because there still remain options as to which particular model will best meet the needs within any constraints that may exist. There may be no unique answer to the question of which modeling method is best; in fact, some of the techniques may be complementary.

The first step one might consider is to test a one-half to one-fourth scale model of the vehicle with the internal systems greatly simplified. Even here, the cost is prohibitive in all but a very few cases. Next, we consider ground-based tests such as in a wind tunnel or a ballistic range. These two approaches are not necessarily exclusive. Availability normally dictates which will be used, but one should be aware of the limitations and advantages of each because the results may be altered significantly by the choice. For example, in a wind tunnel, it is easy to measure force and moments but the presence of a model support may affect the data. The wind tunnel may also be limited in its ability to simulate the proper environment at very high speeds. In the ballistic range, on the other hand, force and moment results are difficult to obtain; they have to be inferred from the measured motion of a small model as it flies through a suitably instrumented range. The ballistic range test however can simulate the flight conditions better and there are no support effects to worry about. Figure 2 shows a typical ballistic range shadow-graph from which the motion measurements must be made. For the remainder of the work we will assume that the ballistic range has been selected.



A-33694

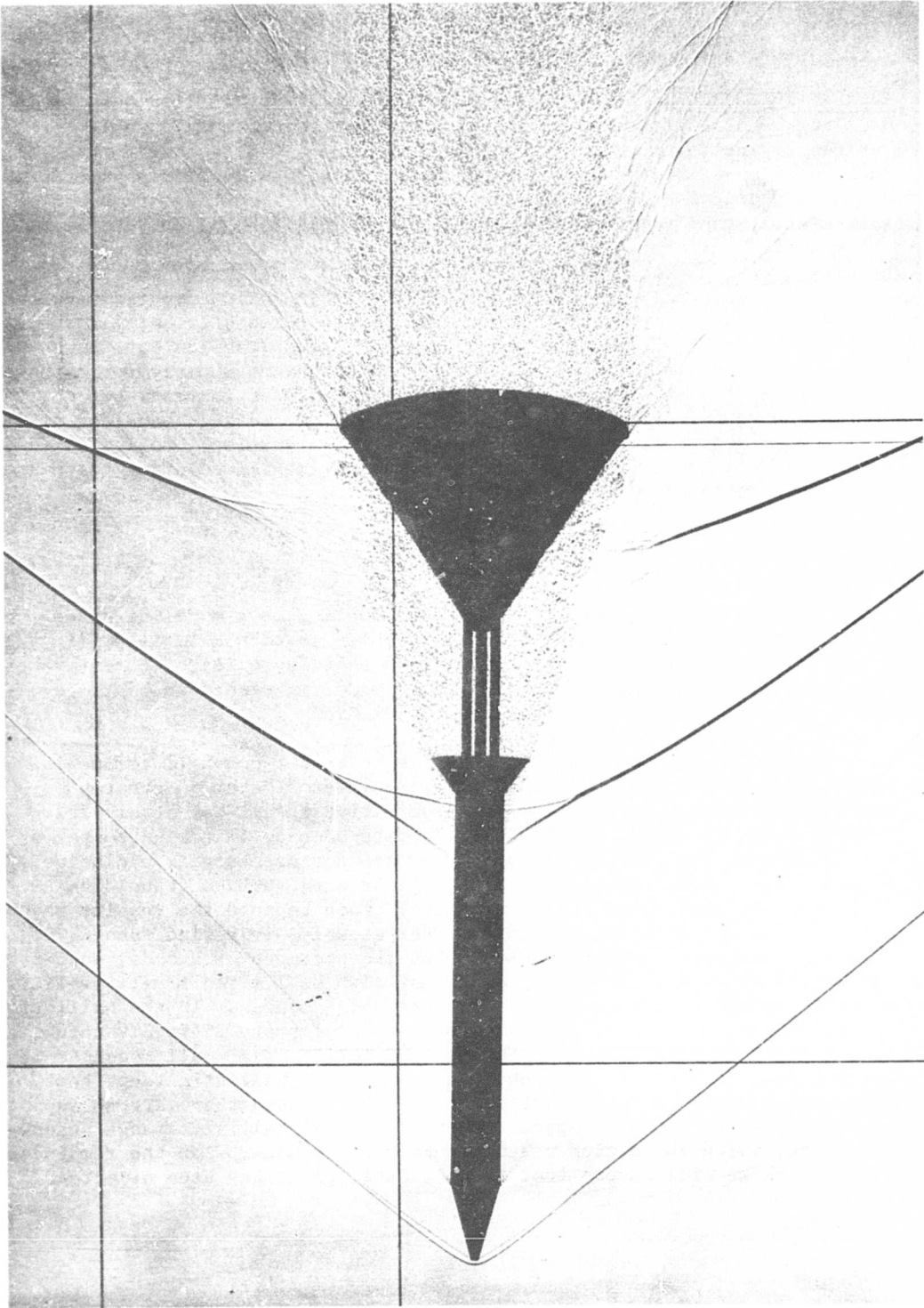


Fig. 2. Typical ballistic range shadowgraph.

## Dimensional Analysis

No matter which technique had been selected, we would be faced with the first major problem that parameter identification must deal with; that is, how are results obtained on small-scale models in a laboratory facility applicable to full-scale flight? Hence, we must understand scaling and simulation rules. We will take a brief look at this from the standpoint of dimensional analysis.<sup>2</sup> To be specific, we will look at the drag of our vehicle to see what we must simulate and how we must treat the results. The basic idea of dimensional analysis is as follows: In the functional relationship between the quantity of interest, (drag, in this case) to other important variables of the problem, such as air density  $\rho$ , flight velocity  $U$ , reference length  $L$ , vehicle configuration  $C$ , and vehicle orientation  $OR$ , the important variables must appear in combinations such as to yield the same dimensions as the quantity of interest.

Equation (1)

$$\text{Drag} = D(\rho, U, L; C, OR) \quad (1)$$

shows such a functional relationship, and Equation (2)

$$C_D = \frac{D}{(1/2)\rho U^2 L^2} = C_D(C, OR) \quad (2)$$

shows the particular grouping of  $\rho$ ,  $U$ , and  $L$  that yield the dimensions of drag. In this case they have been used to nondimensionalize the drag and hence produce a drag coefficient that in principle depends only on the remaining two nondimensional quantities  $C$  and  $OR$ . (The  $1/2$

appearing in the denominator arises historically because the quantity  $(1/2)\rho U^2$  is a term that appears in simple fluid mechanic problems and is referred to as the dynamic pressure.) The dependence shown in Equation (2) is found not to be sufficient in practice. The fluid viscosity ( $\mu$ ) and speed of sound ( $a$ ) have also been found to be important. When these two quantities are incorporated we get

$$\text{Drag} = D(\rho, U, L, \mu, a; C, OR). \quad (3)$$

Since we have already established a group of terms which had the same units as the drag, the remaining variables must form nondimensional groups. When we do this, we get the results shown in Equation (4), where the drag coefficient  $C_D$  is shown to be a function of the configuration, the orientation and two parameters; one called the Reynolds number  $Re$ , and the other the Mach number  $M$ .

$$C_D = \frac{D}{(1/2)\rho U^2 A} = C_D(M, Re; C, OR). \quad (4)$$

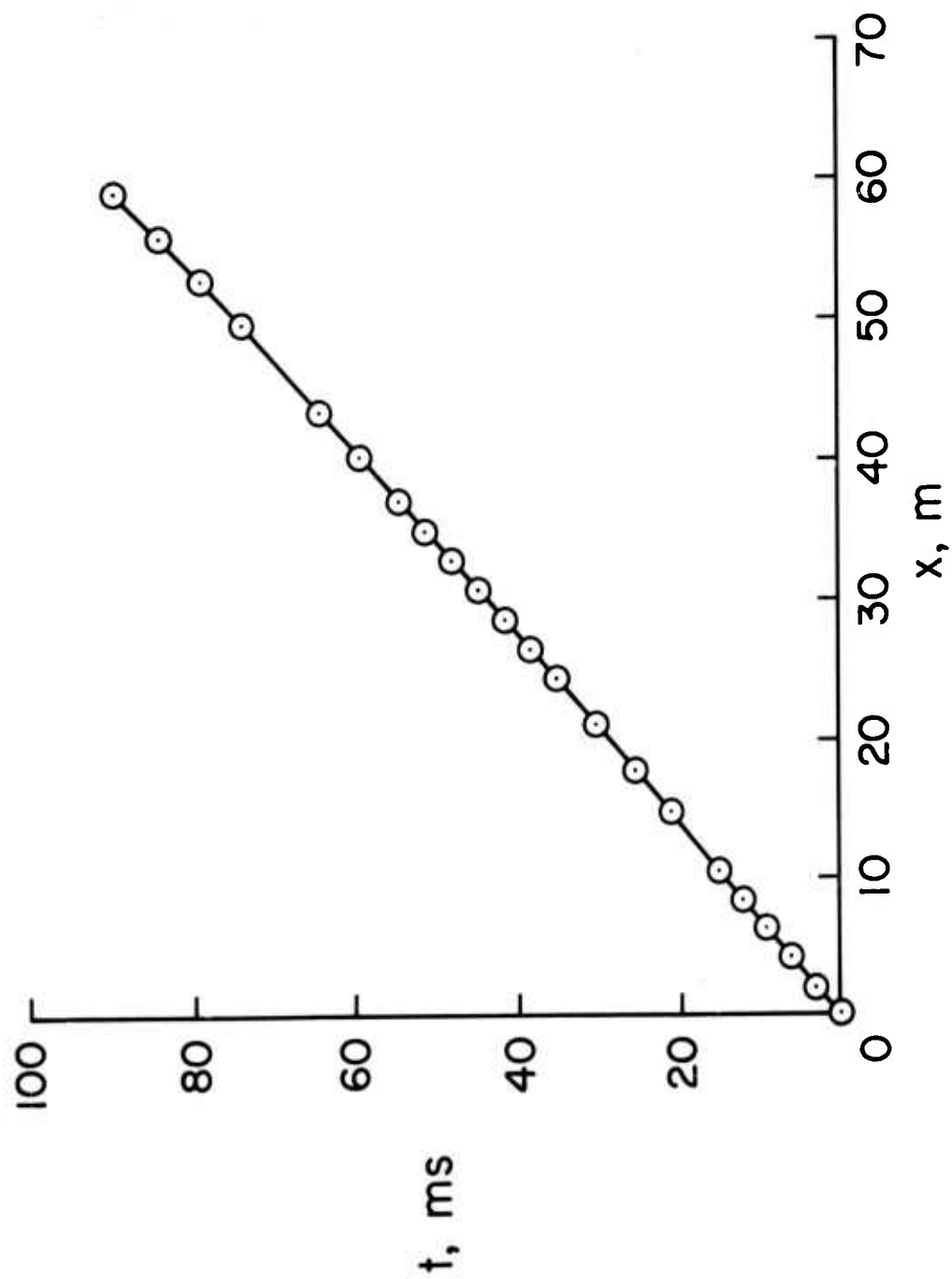
Where  $M = U/a$  and  $Re = \rho UL/\mu$ . The Reynolds number is an indication of the viscous drag effects, and the Mach number, the compressibility effects brought on by flying at speeds other than those much less than the speed of sound. Note also that a reference area  $A$  has been used in place of  $L^2$ . This last functional relationship [Equation (4)] tells us that as long as our laboratory test is of a model with the same configuration and at the same orientation as the full-scale vehicle and flying at the same Mach number and Reynolds number, then the drag coefficients for the two cases are the same and we can use laboratory tests to establish full-scale results. All other aerodynamic forces and moments must be treated in a similar manner. This then is the basis for going from laboratory tests to full scale; it has limitations that are not considered here but they will not influence our discussion (see Ref. 2 for more details). It might seem that we have not considered the mathematical modeling here but it has been implicit that both experience and theory are used to determine what variables are important.

#### Experiment

We will now consider the ballistic range test in sufficient detail to obtain a better understanding of the physical and mathematical models that we will be considering in the classical parameter identification step which terminates the formal job of obtaining the aerodynamic parameters for our full-scale vehicle.

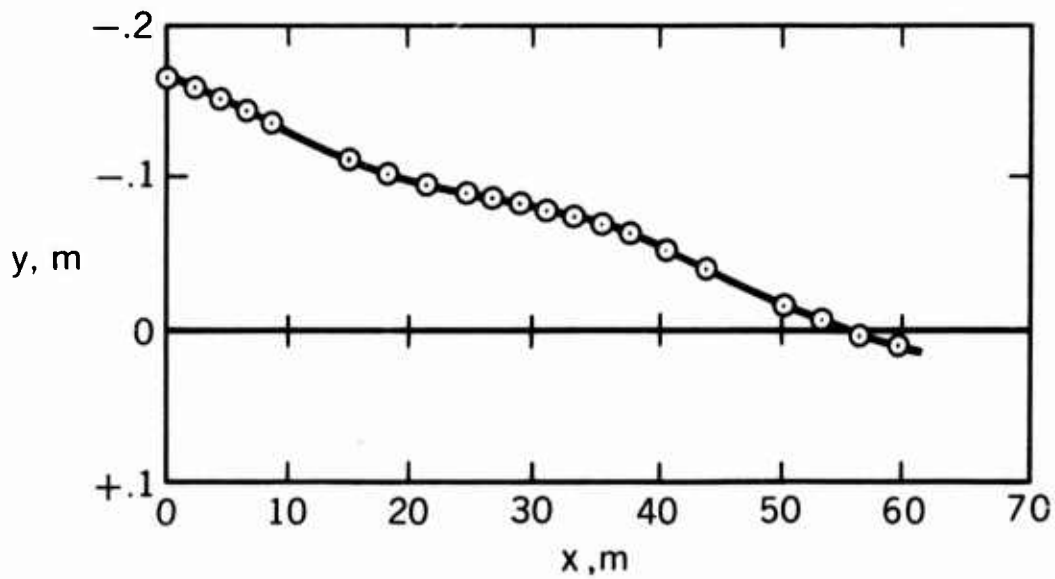
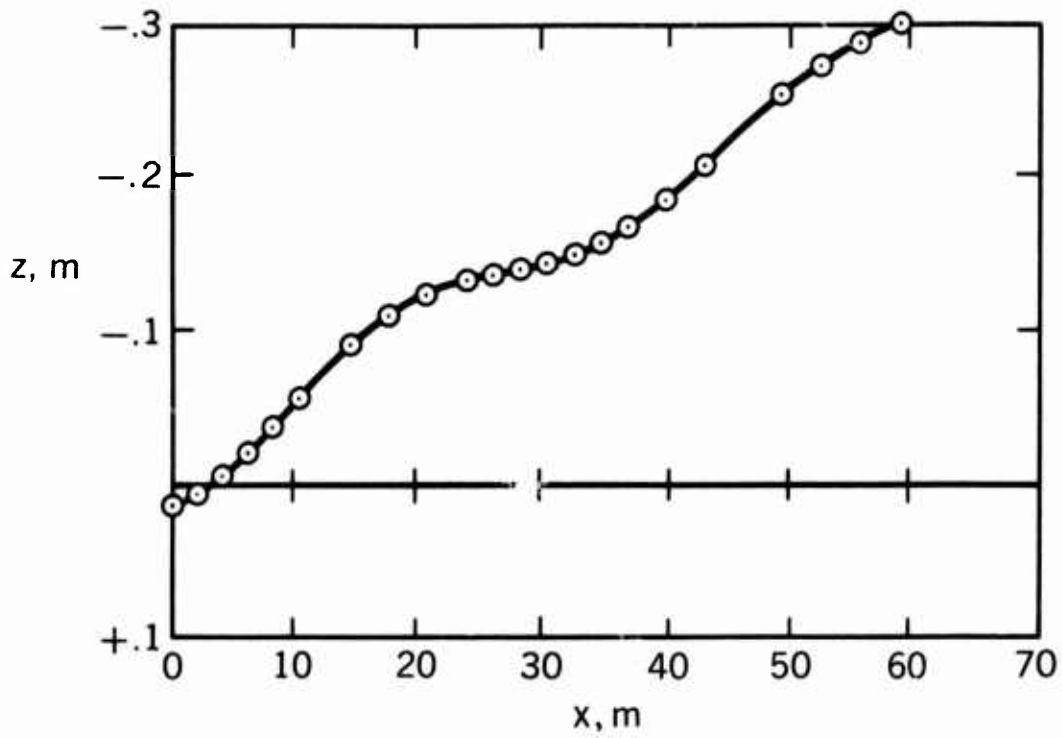
The ballistic range consists of a gun from which a small-scale model of the vehicle is launched into an instrument range. The instrumentation consists of a series of orthogonal shadowgraph stations at which spark shadowgraphs are taken. In addition, the time at which each shadowgraph is taken is recorded. An example of a shadowgraph of this type was shown in Figure 2. From the shadowgraphs one obtains the position and orientation of the model as a function of time or distance. A set of data from an actual flight is shown in Figure 3(a), (b), (c). The distance  $x$  is along the range and the distances  $y$  and  $z$  make up an orthogonal set;  $z$  is positive downward in the vertical direction. The angle of yaw is  $\psi$  and the angle of pitch,  $\theta$ . The roll angle  $\phi$  is not shown but was essentially constant.

Here the first thing we notice is that we are dealing with a problem involving six degrees of freedom. Hence the mathematical modeling problem could be very difficult. Second, although it is not obvious at this point, there is experimental error present in the data; that is, the data points define some exact trajectory with error superimposed. These errors should mostly be statistical with zero mean but, as we will see later, this is not always so.



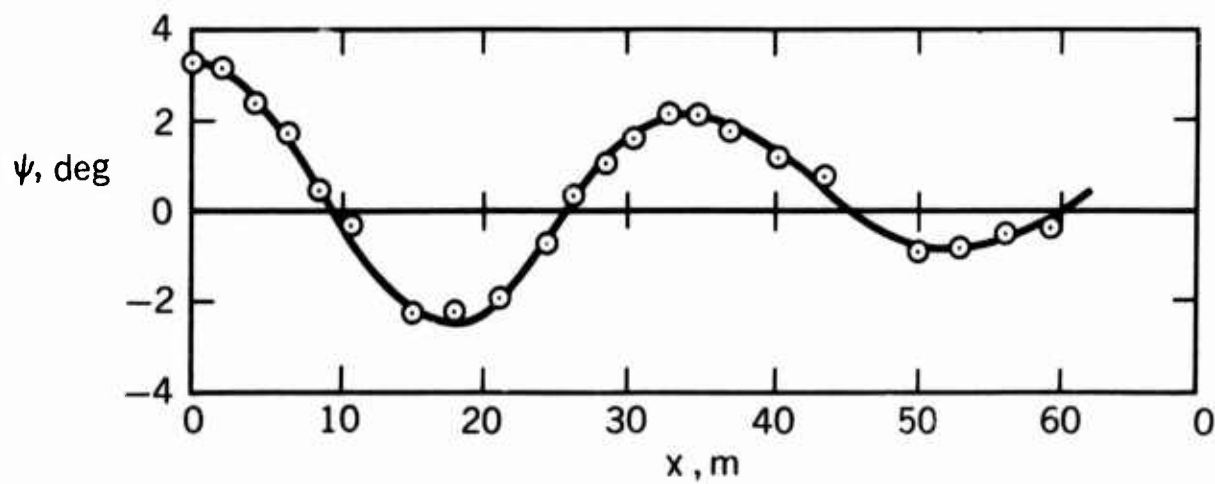
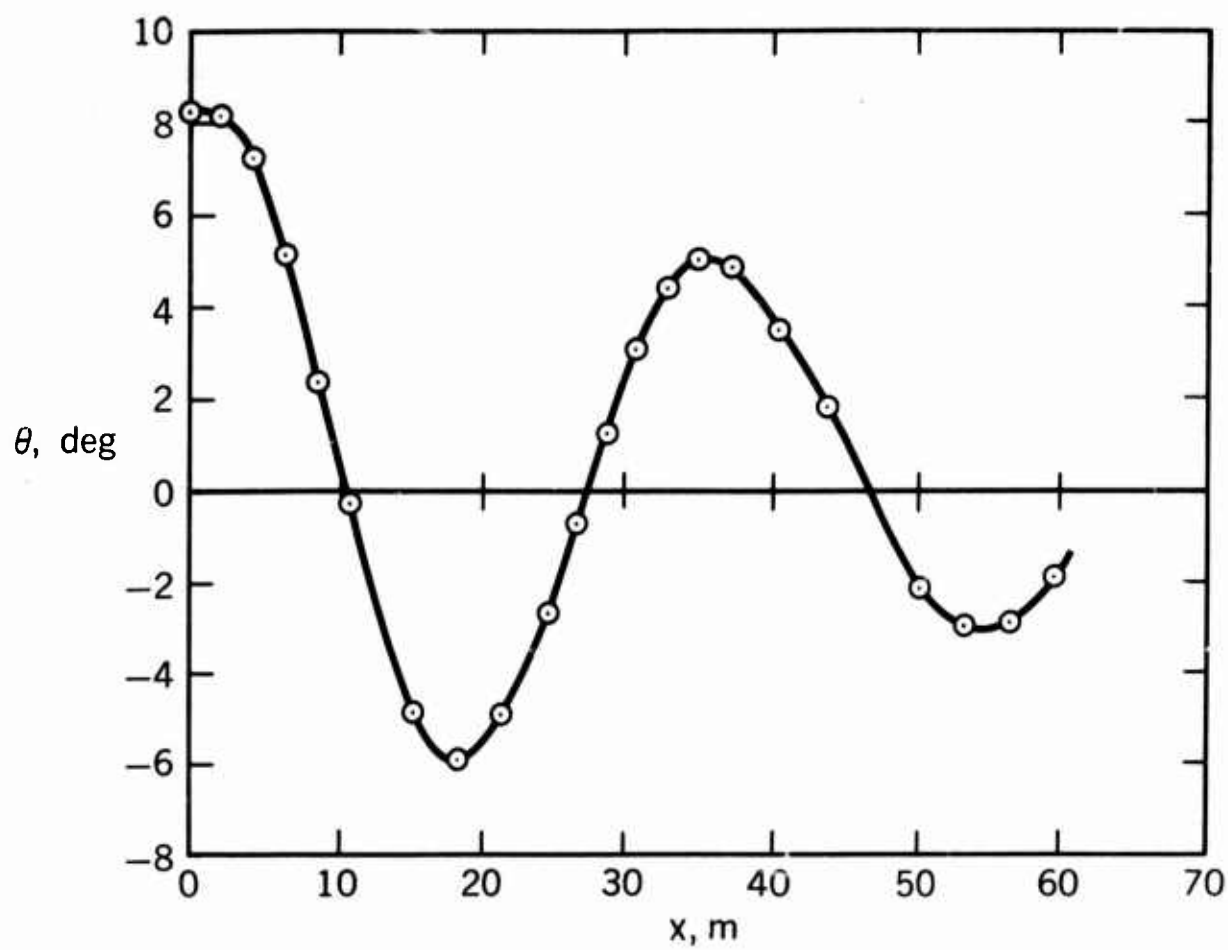
(a) Time versus distance measurements.

Fig. 3. Typical ballistic range measurements.



(b) Swerve measurements.

Fig. 3.- Continued.



(c) Angular orientation measurements.

Fig. 3.- Concluded.

The mathematical modeling of this six degree of freedom dynamic system consists of two sets of three second-order differential equations, one set for linear momentum, Equation (5) and one set for angular momentum, Equation (6).

$$m \frac{d^2 x}{dt^2} = F_x \quad (5a)$$

$$m \frac{d^2 y}{dt^2} = F_y \quad (5b)$$

$$m \frac{d^2 z}{dt^2} = F_z \quad (5c)$$

$$I_x \dot{p} - qr(I_y - I_z) = M_\ell \quad (6a)$$

$$I_y \dot{q} - pr(I_z - I_x) = M_m \quad (6b)$$

$$I_z \dot{r} - pq(I_x - I_y) = M_n \quad (6c)$$

where  $F_x$ ,  $F_y$ , and  $F_z$  represent the forces (aerodynamic and gravity) that on a body of mass  $m$ , and  $M_\ell$ ,  $M_m$ ,  $M_n$  are the aerodynamic moments that act on the body about principal axes that have moments of inertia  $I_x$ ,  $I_y$ , and  $I_z$ , respectively. Note that the angular velocities  $p$ ,  $q$ , and  $r$  are related to the pitch, yaw, and roll angles (see Ref. 1).

These two sets of equations would be decoupled except for the aerodynamic forces and moments that occur. It has been found from inspection and experience with data that the coupling in some directions is weak or can be accounted for in an after-the-fact manner. This weak coupling can and should be exploited in setting up the mathematical models for parameter identification.

The equation of motion in the  $x$  direction is completely decoupled from the remainder of the equations if  $F_x$  is independent of orientation. Even if it does depend on orientation, Seiff and Wilkins<sup>3</sup> have shown that the dependence can be accounted for in an after-the-fact manner. This will be considered in detail later. Next a major decoupling is brought about by transforming the remaining equations from time as the independent variable to distance along the direction of flight  $x$ . When this has been done, the equations of angular momentum depend only weakly on the equations of linear momentum but the coupling in the opposite direction is strong.



The weak coupling can be accounted for in an iterative manner or neglected entirely in many cases. Furthermore within the set of angular momentum equations there is some weak coupling that can be exploited. This latter coupling depends on the set of angles that are used to describe the motion.

The aerodynamic forces and moments are unknown functions of angular orientations and angular rates. This is the information that we are trying to find. It is also the area where much modeling needs to be done. At present most modeling consists of polynomial expressions in terms of the angles and angular rates. Two approaches have been used here. The classical approach<sup>4</sup> uses conventional concepts of angle of attack and angle of sideslip. The other uses a resultant angle approach.<sup>5</sup> The latter includes nonlinear terms in a simpler and often more natural way.

PARAMETER IDENTIFICATION. We now take up the classical parameter identification step, namely the bringing together of experimental data and a mathematical model to deduce some aerodynamic parameters. In this step we must fit our equations to the measured motions and determine the initial conditions and aerodynamic parameters that give the "best fit". "Best fit" here is meant in the least square sense. Hence our starting point is the sum of the squares of the difference between discrete points of an experimentally determined function  $f_{e_i}$  and the calculated function  $f_{c_i}$ . This is

written as

$$SSR = \sum_{i=1}^N \left( f_{e_i} - f_{c_i} \right)^2 \quad (7)$$

where  $i$  is the index of the measuring station and  $N$  is the number of these stations. We will now consider three cases of increasing complexity of the function  $f_c$ .

#### Simple Case - Drag

For a projectile fired horizontally in a ballistic range, the gravity term can be dropped and the equation for drag is

$$\frac{d^2x}{dt^2} = -KC_D \left( \frac{dx}{dt} \right)^2 \quad (8)$$

where  $K = \rho A/2m$ . The solution to Equation (8) for the case of constant  $C_D$  is

$$t = \left( t_0 - \frac{1}{v_0 KC_D} \right) + \frac{1}{v_0 KC_D} e^{KC_D x}. \quad (9)$$

Equation (9) cannot be used for  $f_{c_1}$  in Equation (7) to give a normal closed form least squares solution because  $C_D$  appears in a transcendental manner. Hence we must consider some other alternative and an iterative application of the conventional procedure will suffice. To start this we need a first approximation for  $C_D$ . This can be obtained by noting from experience that  $KC_D$  is normally small resulting in the following approximation for Equation (9).

$$t = t_0 + \frac{x}{V_0} + \frac{KC_D}{2V_0} x^2. \quad (10)$$

Since this is a polynomial in  $x$ , the straightforward least squares procedure readily leads to a value of  $C_D$ . This value is normally within a few percent of the value obtained using Equation (9). To obtain the value of  $C_D$  consistent with Equation (9), this approximate value of  $C_D$  is substituted into Equation (9) and the standard least square procedure is used to determine the leading constant and the multiplier on the exponential (these appear in a linear manner). The value of SSR is then calculated;  $C_D$  is now changed by a small amount and the process repeated. This entire procedure is repeated, always looking for the value of  $C_D$  that yields the minimum value of SSR. This search can be carried out in many ways. A simple and fast way is to repeat the process for three values of  $C_D$ , fit a parabola to the values of  $C_D$  and SSR and determine a minimum. This value of  $C_D$  can be the starting place to repeat the process, with smaller changes in  $C_D$ , if more accuracy is required.

#### Variable Drag Coefficient

In most cases of interest drag depends on the angular orientation. Seiff and Wilkins<sup>3</sup> have shown that for axis-symmetric bodies when the drag coefficient depends on the resultant angle of attack,  $\sigma$ , as

$$C_D = C_{D_0} + C_{D_2} \sigma^2 \quad (11)$$

the drag coefficient determined as though it were constant is the correct value at the root mean square angle-of-attack for that flight. If this is so, values of  $C_D$  determined assuming constant  $C_D$  when plotted as a function of the mean squared angle of attack should result in a straight line whose intercept is  $C_{D_0}$  and whose slope is  $C_{D_2}$ . Two examples of such

data are shown in Figure 4.<sup>6</sup> Here we see that we do indeed obtain a straight line. Note also that in this case the intercept is in good agreement with theoretical values.

#### Data Comparison

To show that data obtained in ground facilities are applicable to a full-scale flight we have made a comparison of wind tunnel, ballistic range, and full-scale data. The full-scale flight was for a 55° blunted cone approximately one meter in diameter.<sup>7</sup> Onboard accelerometer measurements were combined with meteorological measurements of density to determine the drag coefficients. The ballistic range data were obtained by Robert Sammonds and Robert Kruse of Ames Research Center. The wind tunnel results are compiled for several sources.<sup>8</sup> Comparison of these data are shown in Figure 5. Note  $C_D$  is plotted versus Mach number but since the full-scale flight experienced a particular Reynolds number and angle of attack at each Mach number, the wind tunnel and ballistic range results are for the appropriate Reynolds number and angle of attack. It can be seen that while the agreement is not perfect, it is within 2-4% for most of the range of conditions. This is well within the accuracy of the full-scale measurements.

#### Transcendental or Nonlinear Case - Linear Stability

The next most complicated parameter identification problem arises when some of the parameters occur in a transcendental or nonlinear manner. An important example of this in ballistic range testing is the solution of the equations of angular momentum when linear aerodynamics and constant roll rate have been assumed for an axisymmetric body with small asymmetries. The equation of motion for that case is.

$$\xi'' + A\xi' + B\xi = Ce^{ipx} \quad (12)$$

where  $\xi = \beta + i\alpha$ ,  $\beta$  is the angle of sideslip, and  $\alpha$  is the angle of attack. The solution to Equation (12) is

$$\xi = K_1 e^{(\eta_1 + i\omega_1)x} + K_2 e^{(\eta_2 - i\omega_2)x} + K_3 e^{ipx} \quad (13)$$

which is called the tricyclic equation because it is equivalent to three rotating vectors in the  $\alpha - \beta$  plane. This equation was first derived by Nicholaides.<sup>9</sup> Here the  $\eta$ 's and  $\omega$ 's are related to the important aerodynamic parameters, damping and pitching moments, respectively. Equation (13) can be written in terms of components as

$V \approx 5 \text{ km/sec}$      $M \approx 15$      $Re_1 \approx 0.4 \times 10^6$

	THEORY		EXPERIMENT	
	NO ABLATION	ABLATION	NO ABLATION	ABLATION
WAVE DRAG (KOPAL)	0.065	SAME		
SKIN FRICTION	.039	0.010		
BASE DRAG	.007	SAME		
VISCOUS INTERACTION	.004	.006		
<b>TOTAL DRAG (<math>\alpha=0^\circ</math>)</b>	<b>0.115</b>	<b>0.088</b>	<b>0.112</b>	<b>0.086</b>

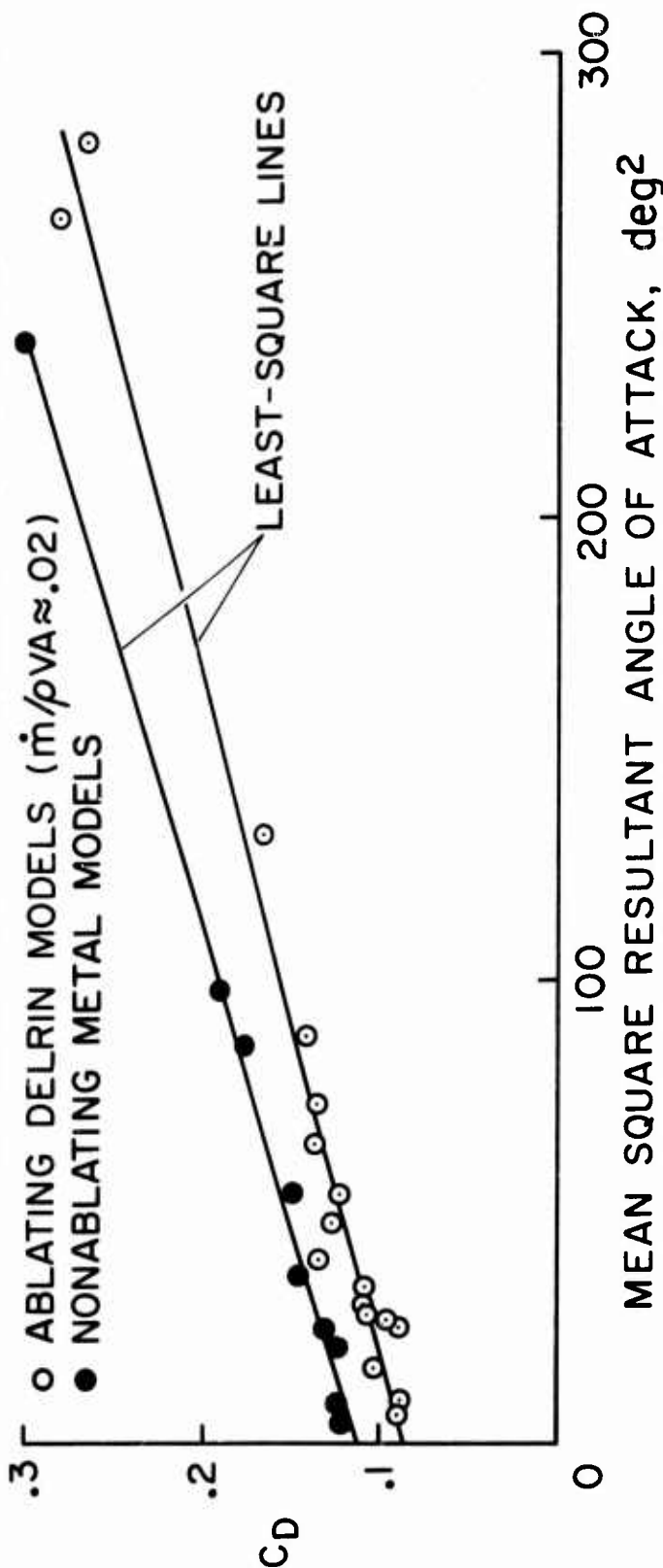


Fig. 4. Quadratic dependence of drag coefficient on angle of attack.

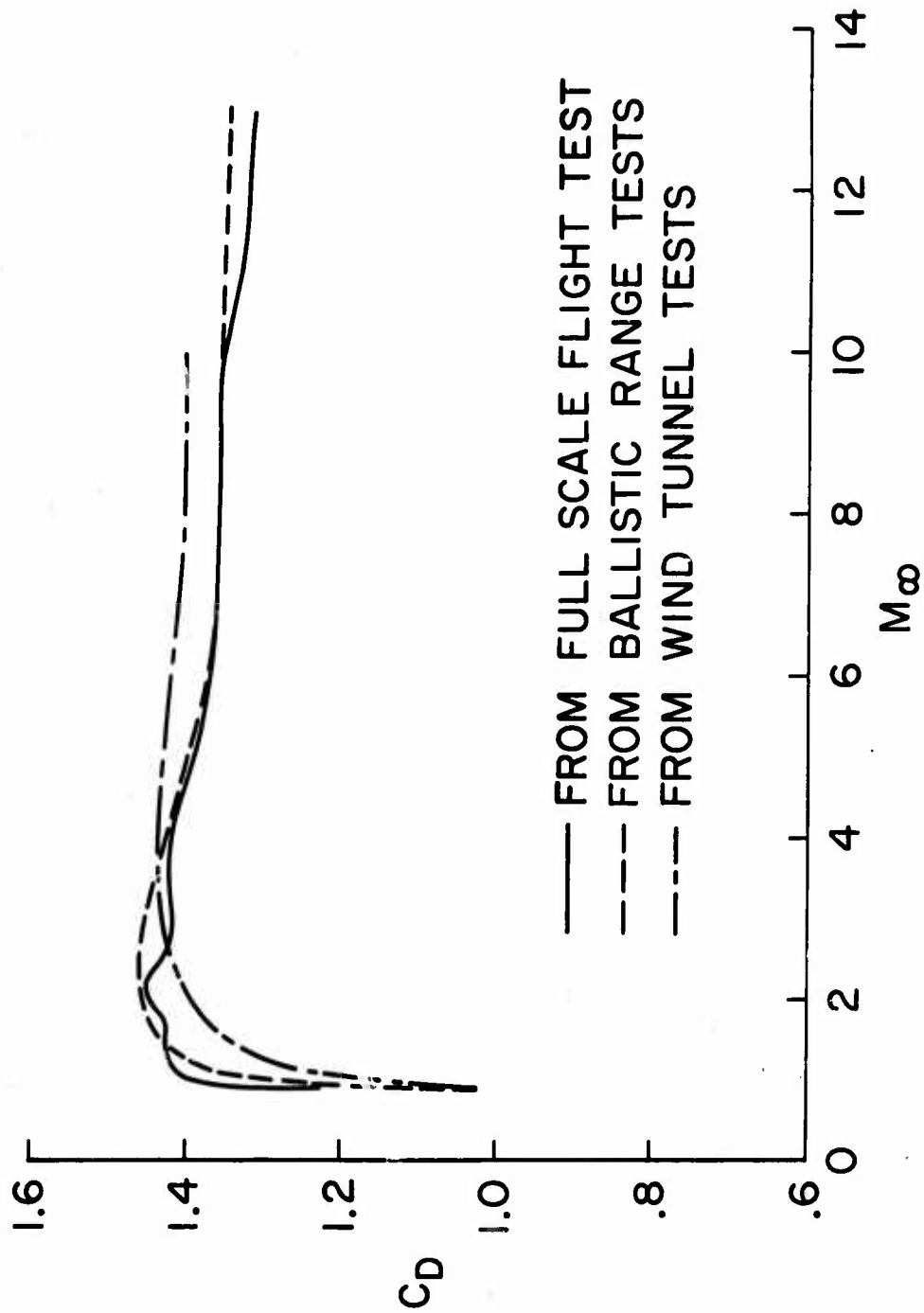


Fig. 5. Drag coefficient comparison.

tolerance. The starting solution for this case is obtained using the Prony method.<sup>11</sup> It will be discussed briefly in a later section involving starting solutions. Some examples of fits to experimental  $\alpha, \beta$  data using Equation (13). Some values of the aerodynamic pitching moment curve slope,  $c_{m\alpha}$ , obtained from similar fits, are shown in Figure 7.<sup>12</sup>

Note  $C_{m\alpha}$  is related to  $\omega_1$  and  $\omega_2$  as

$$C_{m\alpha} = - \frac{2\omega_1\omega_2 I_y}{\rho AL} . \quad (17)$$

We see in Figure 7 that  $C_{m\alpha}$  appears to be nearly constant (i.e.,  $C_m$  is linear) for small angles of attack and hence the use of Equation (13) is justified but, as the angles become large,  $C_{m\alpha}$  starts to decrease and our assumption of linear aerodynamics is no longer valid. This can be handled in a quasilinear manner, as developed by Murphy<sup>13</sup> and Rasmussen and Kirk.<sup>14</sup> These methods relate the values of  $C_{m\alpha}$  determined from a linear analysis using Equation (13) for several flights at different amplitudes to the best nonlinear polynomial representations of pitching moment  $C_m$ . When the procedure of Reference 14 is applied to the data for Mach number 11.5 of Figure (7), we get the results shown in Figure (8).

The method of quasilinear analysis does not handle nonlinear damping very well and entails some other approximations that prevent it from having complete generality. Hence, we are led to the final and most general case of parameter identification.

#### Differential Equation Case - Nonlinear Stability

The most general case of parameter identification occurs when the mathematical model is of sufficient complexity to prohibit the possibility of finding a closed-form solution. To illustrate how a solution for this problem proceeds we will consider the planar motion of a vehicle governed by nonlinear static and dynamic aerodynamic moments. The equation for this case is

$$\ddot{\alpha} + (C_1 + C_2\alpha^2) \dot{\alpha} + (C_3 + C_4\alpha^2)\alpha = 0 \quad (18)$$

with the initial conditions

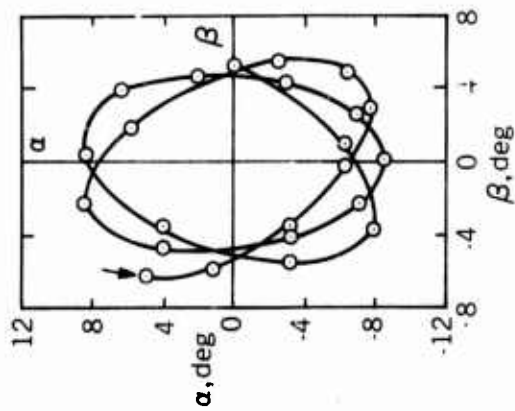
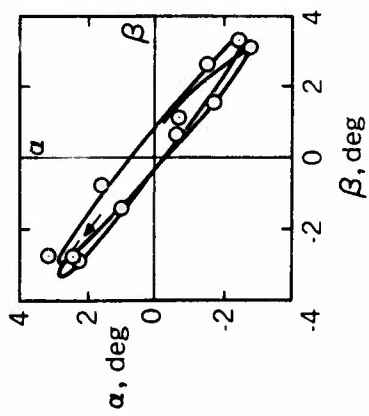
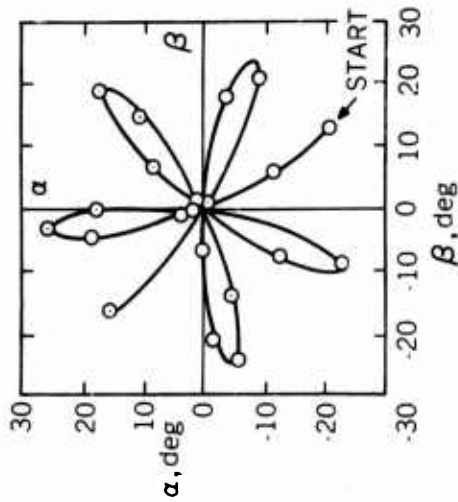
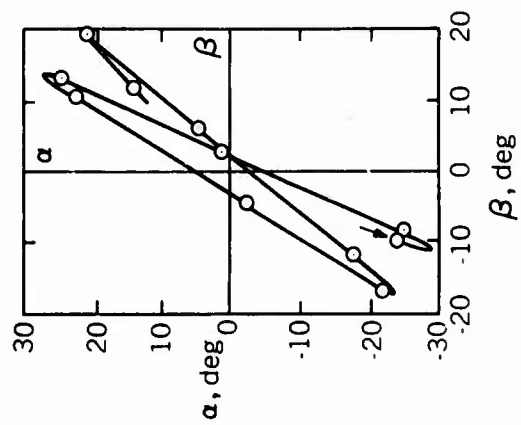


Fig. 6. Typical curve fits to angular motion.



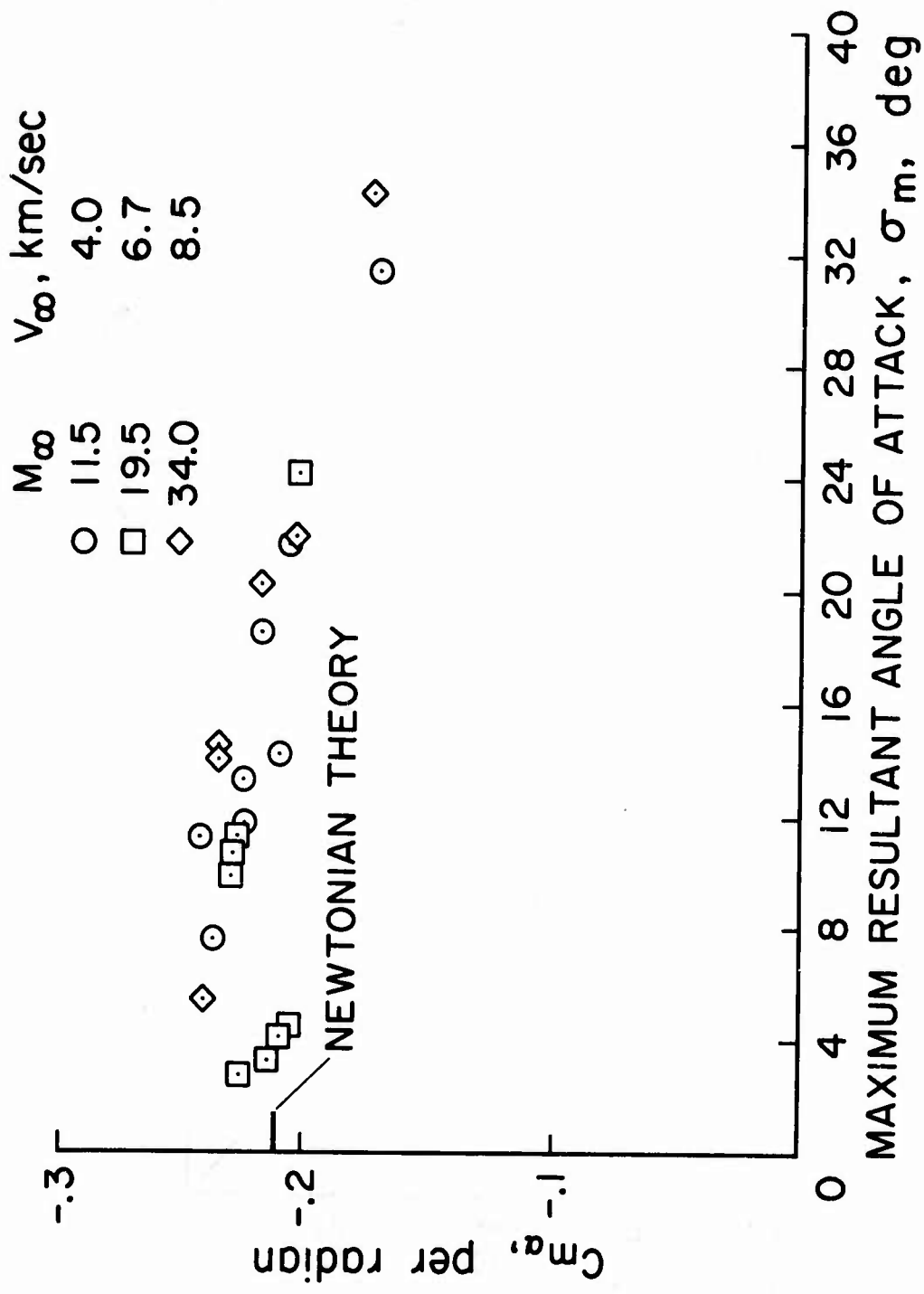


Fig. 7. Typical pitching moment curve slope  $C_{m\alpha}$  for a 30° blunted cone.

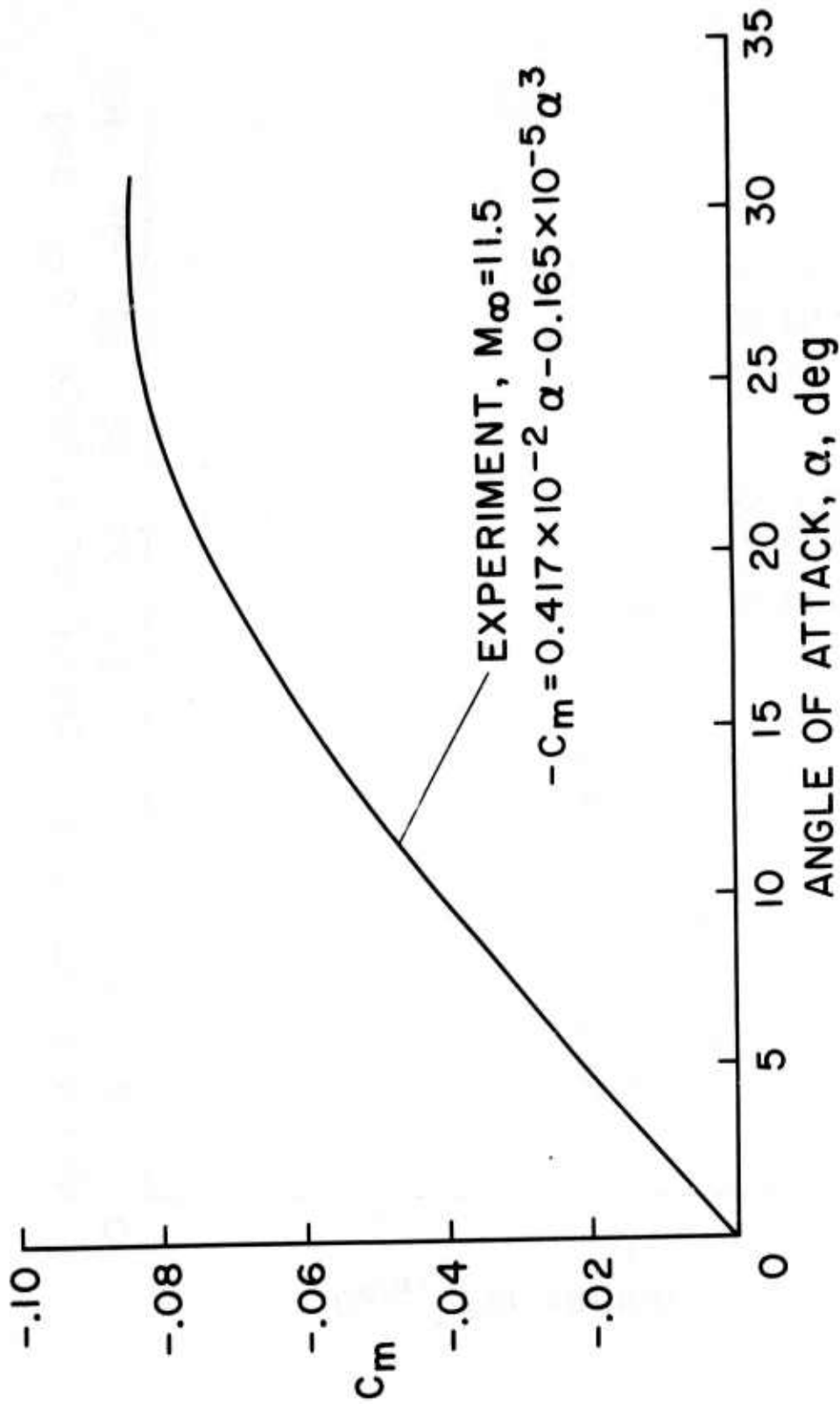


Fig. 8. Nonlinear pitching moment coefficient  $C_m$  obtained using the quasilinear analysis.

$$\dot{\alpha}(0) = C_5$$

$$\alpha(0) = C_6.$$

To apply least squares to this we first use the method of differential corrections (Equation (16)). Note  $f_c$  is replaced by  $\alpha$  and there are only six unknowns to consider. What is needed is an initial solution  $\alpha_{c_0}$  if six approximate values of the C's are known by straightforward integration (a Runge-Kutta integration procedure works well). To obtain the derivatives of  $\alpha_{c_0}$  with respect to the C's we apply parametric differentiation.<sup>15</sup>

PARAMETRIC DIFFERENTIATION. Parametric differentiation starts by differentiation of the equation of interest (Equation (18)) with respect to each of the parameters to be determined.<sup>16</sup> For example, for  $C_3$ , defining  $G_3 = \partial\alpha/\partial C_3$  we obtain

$$\ddot{G}_3 + (C_1 + C_2\alpha^2)\dot{G}_3 + 2C_2\alpha\dot{\alpha}G_3 + (C_3 + 3C_4\alpha^2)G_3 = -\alpha \quad (19)$$

with initial conditions

$$\dot{G}_3(0) = 0$$

$$G_3(0) = 0.$$

There are six of these equations. Using appropriate values for the C's we must integrate simultaneously the six equations for the  $G_i$ 's and the one for  $\alpha$  with appropriate initial conditions. We now have all of the information to proceed with out least square by differential correction procedure.

Before proceeding with some results we will discuss an important factor that affects how the above procedure is applied.

INFORMATION CONTENT. In any single test in a ballistic range where the aerodynamics may be nonlinear there is normally not enough data (information) to determine, with any degree of confidence, the individual nonlinear terms. This will be illustrated with the wave form of a planar oscillation. In Figure 9 we have plotted the wave form of the pitching motion for two cases having the same wave length and with the amplitude normalized out. The upper curve is for linear aerodynamic moment and is a sine wave. The

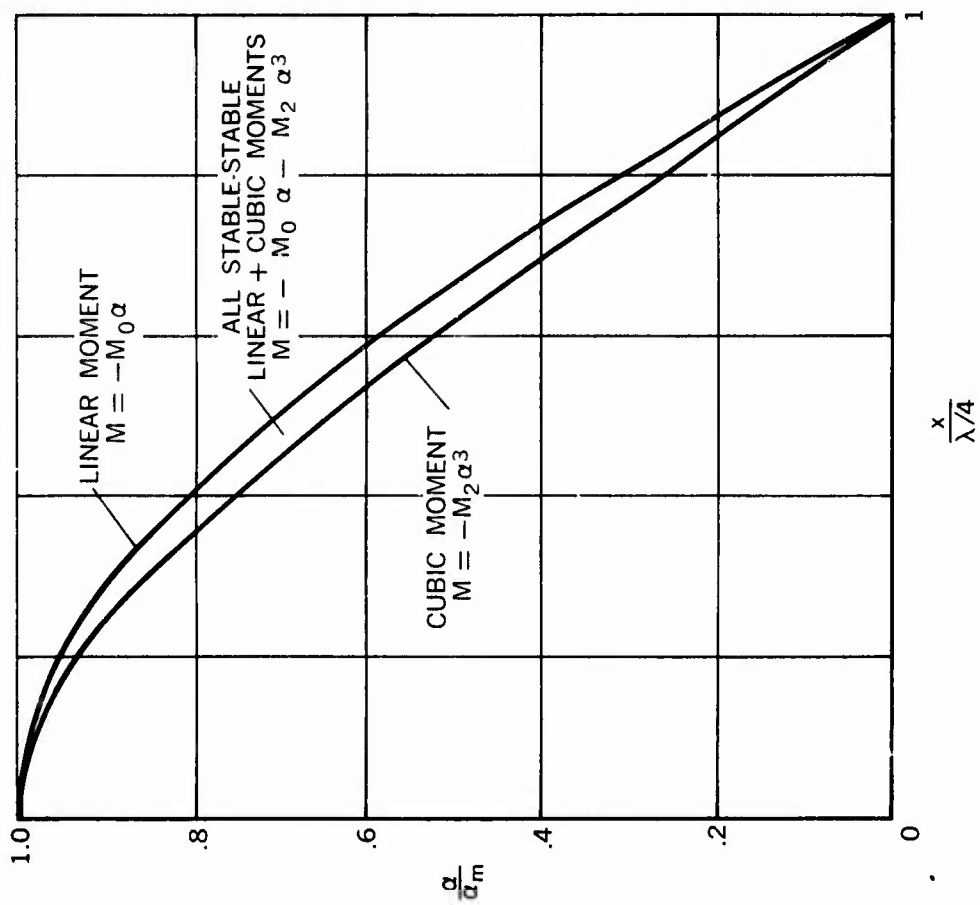


Fig. 9. Effect of nonlinear pitching moment on waveform of oscillation.

other curve is for a pure cubic pitching moment. Note that there is not much difference between the two, particularly when only a few discrete data points on the curve are considered and these contain experimental error. All two-term linear-cubic pitching moments where both terms are stabilizing produce wave forms that fall between these two. Hence we see that the only important way of detecting nonlinearity when the pitching moment is nonlinear is that the wave length is a function of amplitude. In many ballistic range tests however a model may be only lightly damped (small change in amplitude) and/or only a few wave lengths of motion are observed. Therefore the only way to discover the nonlinearity is to reduce simultaneously data from several flights with different amplitudes. In doing this, the aerodynamics (coefficients  $C_1 - C_4$ ) are assumed to be constant from flight to flight but the initial conditions ( $C_5$  and  $C_6$ ) are different. Hence, for example, in fitting four tests simultaneously, we must solve for twelve unknowns - four aerodynamic parameters and eight initial conditions. The fit of Equation (18) to four ballistic range tests of a Gemini capsule is shown in Figure 10. The static pitching moment curves deduced from these flights are shown in Figure 11. Shown for comparison is the result deduced using the quasilinear approach.<sup>14</sup> The slight difference can possibly be attributed to the inclusion of a nonlinear damping. Also included to show the sensitivity is a curve generated with only three runs; one of the small angle runs was deleted.

#### Comparison of $C_{m\alpha}$

To complete the discussion of stability we will again make a comparison of wind tunnel, ballistic range, and full-scale flight. The data are from the same sources as the drag comparison (Figure 5). The comparison shown in Figure 12 is seen to be very good for most of the Mach number range. Again we see that the data from the small-scale tests can be used with confidence in a full-scale test.

#### Additional Cases

Two additional cases that have employed parametric differentiation will be considered briefly to illustrate the versatility of the method. The first involves the data reduction for a ballistic range test of the X-15 airplane. In this case only linear aerodynamics have been used but because the body is not axially symmetric, there are many more unknown aerodynamic parameters. A curve fit of the  $\alpha$  and  $\beta$  data using Equations 7.97 and 7.98 of Reference 1 is shown in Figure 13. Note the odd behavior that is exhibited by the X-15 model and yet the curve fit is very good. Aerodynamic data obtained from the test agree well with flight test results. When this particular test was made in the late 1950's, it could be reduced with more approximate techniques.

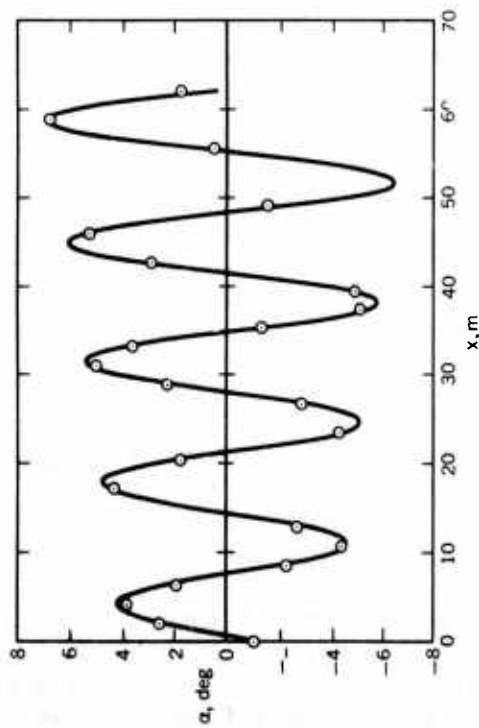
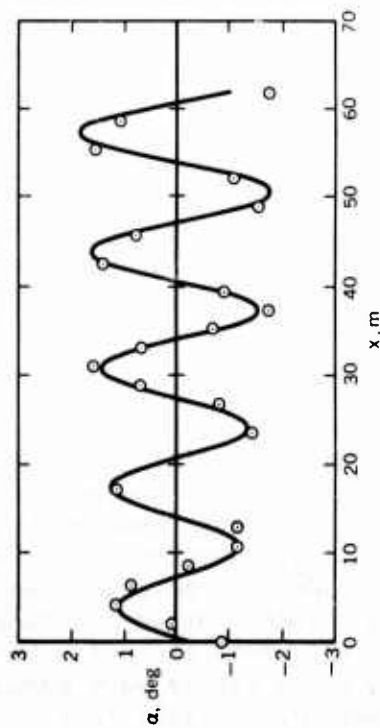
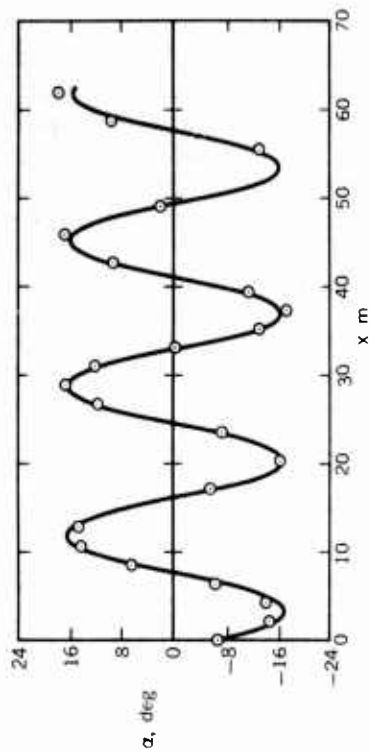
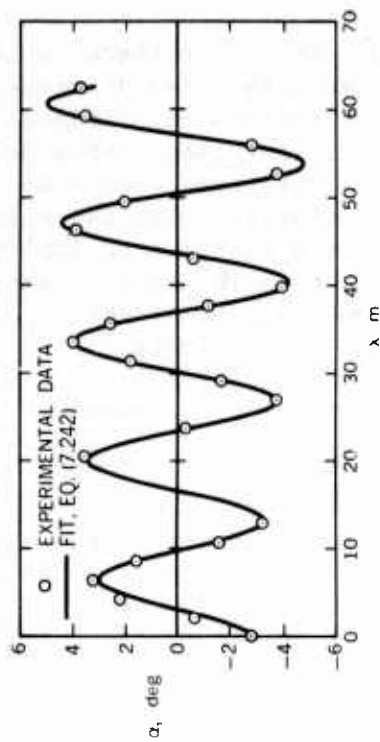


Fig. 10. Curve fits of four Gemini ballistic range tests.

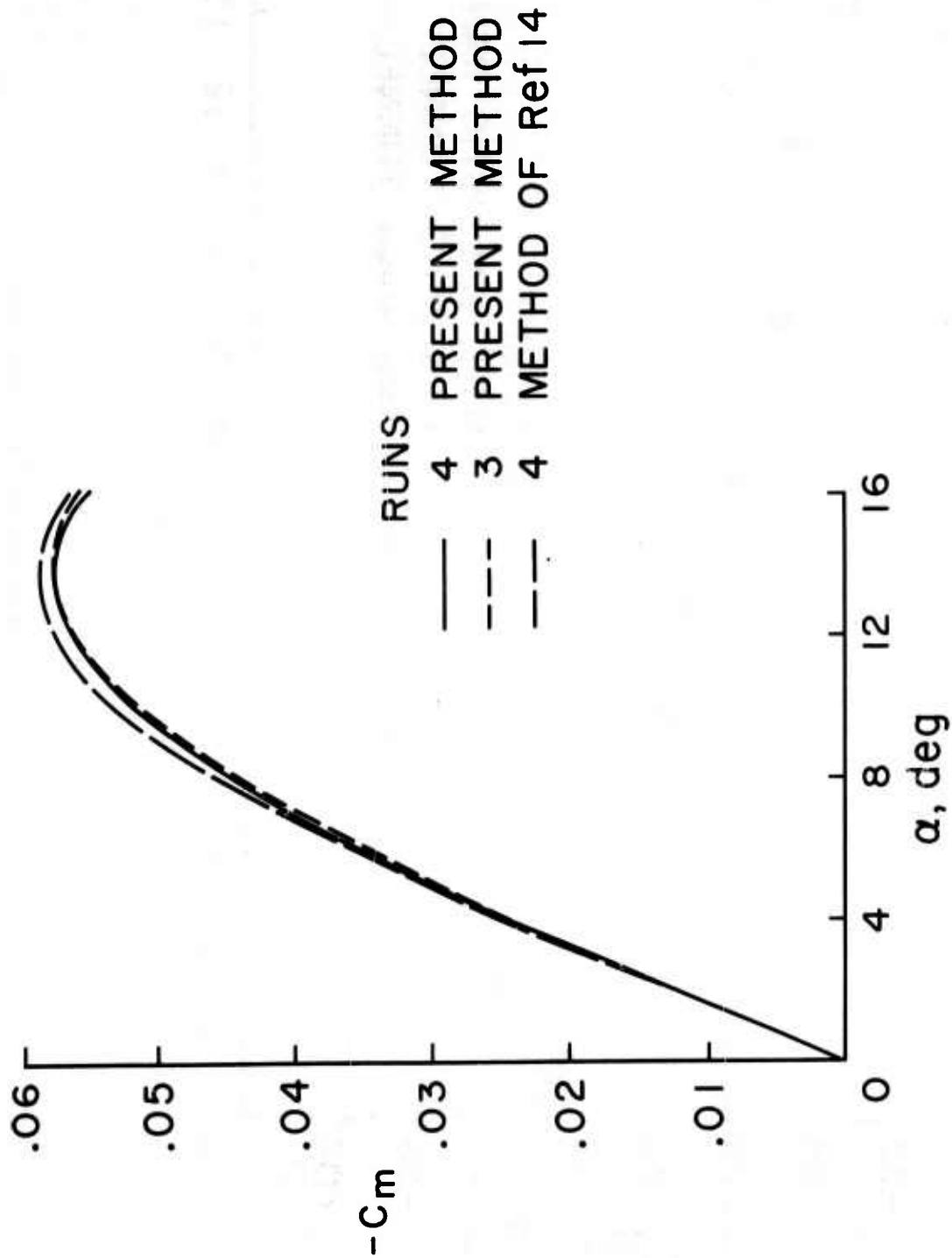


Fig. 11. Pitching moment for Gemini deduced from ballistic range tests.



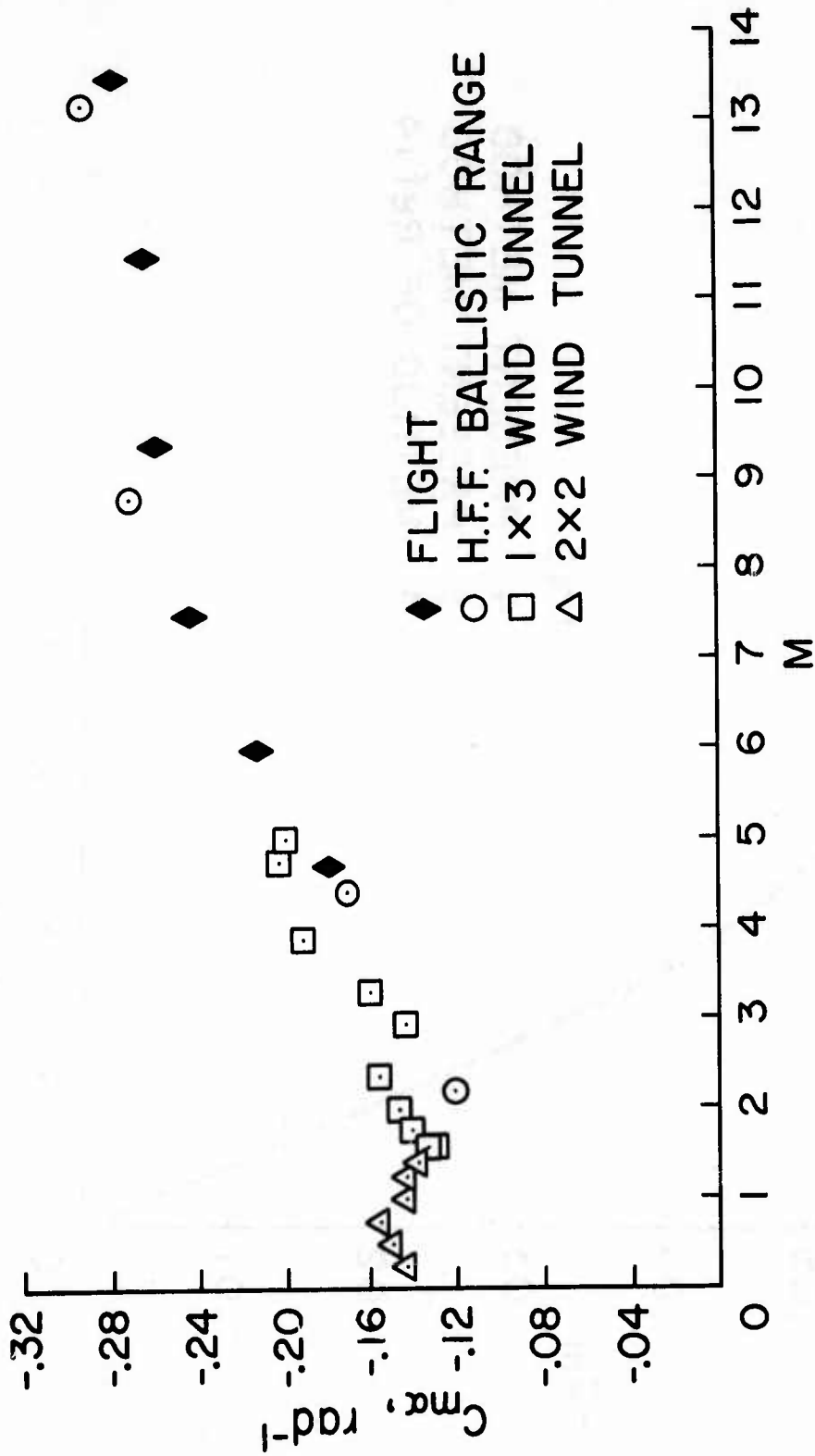


Fig. 12. Static stability comparison.

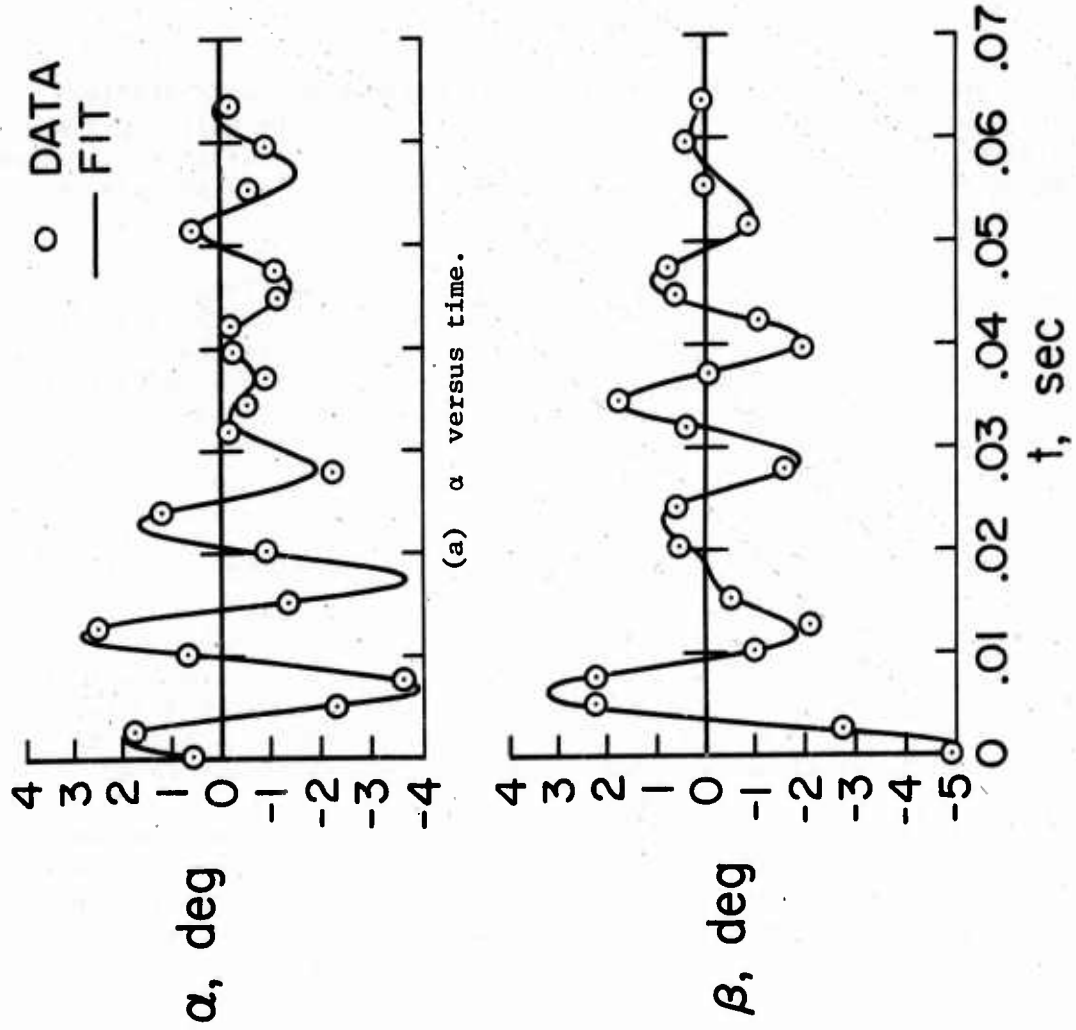


Fig. 13. Curve fit to X-15 airplane motion obtained in a ballistic range.

A second example is that of an axisymmetric body trimmed near 90° angle of attack. For this case the equation for the resultant angle assuming no swerve (Equation 8.46 of Reference 1) was used with a nonlinear pitching moment and a constant damping parameter. A curve fit of the data from this test is shown in Figure 14. Here again we see a good fit to the data. An important point to be made here is that the choice of equation for the fitting procedure can simplify the problem greatly.

#### Associated Topics

There are several topics associated with parameter identification that have been passed over or only touched on briefly. We will now consider five of these; starting solutions, convergence of iteration procedure, modeling of forces and moments, experimental errors, and sensitivity of results and experiment design.

STARTING SOLUTIONS. The problem of obtaining a good starting solution for the differential correction procedure can be very important to obtaining the correct converged solution (i.e., there may be multiple minimums in a nonlinear multiparameter problem). There are probably as many ways to get starting solutions as there are problems. We will briefly mention a few that have been found to be useful in ballistic work. The first approach would be to look for a linear solution. This was done in the case of drag for small values of  $KE_D$ . The next case arises when the solution is written as a sum of exponentials as in the linear stability case. Here the Prony procedure<sup>11</sup> can be applied. The basic approach here is that for equally spaced data points a recursion formula can be written at a prior data point. This procedure leads to a linear set of equations that is readily solved. If the data are not equally spaced, as is the normal case, a set of equally spaced data can be constructed from the original set either by hand or machine-fairing of the data.

For cases where parametric differentiation is employed one can use quasilinear analysis of Reference 13 or Reference 14 to determine starting values of the static stability parameters, or one can formally integrate the equations and obtain implicit integral equations for  $\alpha$  and  $\beta$ . If the integrals in these equations are determined by fairing the data and performing the indicated integration graphically, a set of linear equations can be constructed to evaluate the unknown parameters. This method was originally used in Reference 17 and is described in Reference 16.

Other sources of starting solutions are existing data, theoretical determination of aerodynamic parameters, and finally, probably the most important, the experience of the analyst.

CONVERGENCE AND STABILITY. In the iterative procedures discussed one normally has to prescribe some realistic criterion by which to judge con-

RUN 799

○ DATA  
— FIT

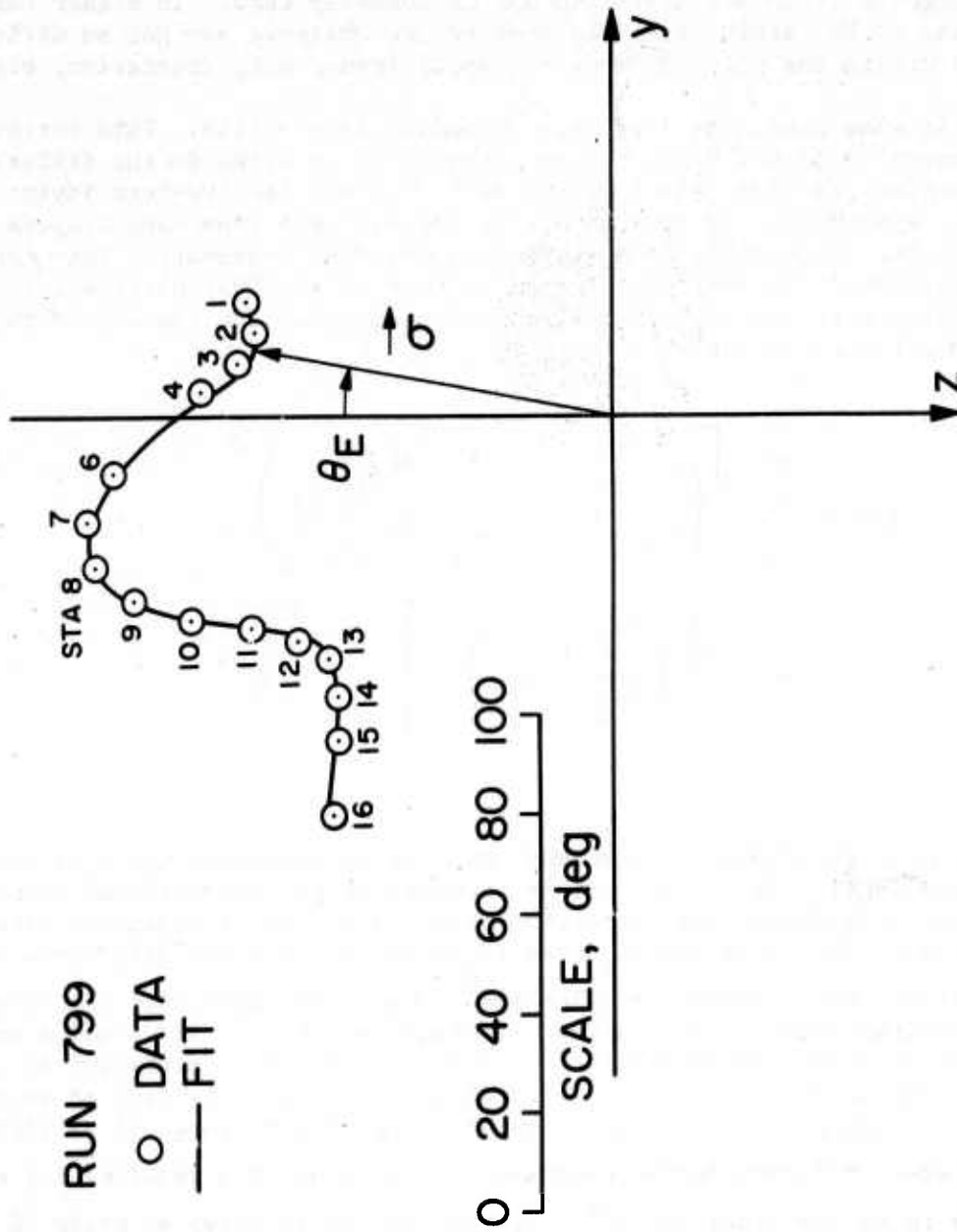


Fig. 14. Large amplitude motion about a near 90° trim point.

vergence. One can either require that the change in the sum of the squares of the residuals is less than some amount or that the changes in the parameters being sought are smaller than some prescribed values. The latter is probably a better criterion but the prior requires fewer convergence tests per iteration and is normally used. In either case one has to be careful that the convergence criteria are not so strict as to be within the calculation error bound (round off, truncation, etc.).

In some cases the iterative procedure is unstable. This instability normally arises because some corrections obtained in the differential correction procedure are so large as to violate the two-term Taylor Series expansion. An obvious way to prevent this from happening is to retard the corrections in a systematic way when necessary. This can be accomplished with what is referred to here as the "Marquardt Algorithm." It starts with the following alternative equation for the sum of the square of the residuals.

$$SSR = \sum_{i=1}^N \left[ \left( f_{e_i} - f_{c_i}_o - \sum_{j=1}^R \left. \frac{\partial f_{c_i}}{\partial a_j} \right|_o \Delta a_j \right)^2 + \lambda \sum_{j=1}^R \left( \left. \frac{\partial f_{c_i}}{\partial a_j} \right|_o \Delta a_j \right)^2 \right] \quad (20)$$

where R is the number of unknowns and  $\lambda$  is an arbitrary positive constant, usually small. The first term in brackets is the conventional least squares by differential correction term. When SSR is minimized with just this term, the  $\Delta a_j$ 's are adjusted to obtain the minimum difference between experiment and calculation. The additional term represents the change from the initial solution to the new calculated solution. If we add a small amount of this term (determined by adjusting  $\lambda$ ) to be minimized at the same time as the first term, we in effect slow down the rate at which the  $\Delta a_j$ 's change and hence keep our solution within the range of validity of the two-term Taylor Series expansion. The value of  $\lambda$  required for most cases is on the order of  $10^{-3}$ . However values as large as order 10 have sometimes been required. It is important to note that the choice of  $\lambda$  will have no effect on the converged solution since it affects only intermediate steps (i.e., the path and rate of convergence are altered but not the end point). This latter is not strictly true if the new path would by chance take the solution to the neighborhood of some other nearby minimum.

The resulting matrix equation for the least squares solution of Equation (20) can be written as

$$\underline{A} \underline{\Delta a} = \underline{r} \quad (21)$$

or

$$\underline{\Delta a} = \underline{A}^{-1} \underline{r}$$

where  $A_{jj} = (1 + \lambda)A'_{jj}$  and  $A_{jk} = A'_{jk}$ ,  $j \neq k$ . Hence the  $A'_{jj}$  and the  $A'_{jk}$  are the matrix elements when only the first term of Equation (20) is retained (conventional least squares with differential correction). Note the simple way the "Marquardt Algorithm" is added to the problem; the diagonal elements are multiplied by  $(1 + \lambda)$ . Additional coding is required to determine when and how large  $\lambda$  should be. One procedure is to use it only when SSR increases and then start with  $\lambda = 1 \times 10^{-3}$ ; if this fails, square  $(1 + \lambda)$ ; if this fails again, cube  $(1 + \lambda)$  and so on. When the convergence is reestablished, return  $\lambda$  to 0 and proceed until it is required again.

**MODELING FORCES AND MOMENTS.** It was noted earlier that modeling of forces and moments was an important area of work to be done and that normally polynomial representations were used. The parametric differentiation approach to the problem is much more general; for example, the moment could be composed of a series of piecewise linear segments. The possibilities are many and I will not pursue the point further. What I will illustrate briefly is the sensitivity to the form of the expression. This will be illustrated by a set of results that was obtained using the quasilinear method.<sup>13,14</sup> (Similar results would be expected to apply to the parametric differentiation approach.) In that study Kirk and Miller<sup>18</sup> used a four-term polynomial; the first term was always the linear term in angle of attack, the remaining three terms were various combinations of powers from 2 to 7. This resulted in 20 possible polynomials. The best five polynomials are shown in Figure 15. All other polynomials produced fits that had considerable larger residuals. It is not really necessary to choose a best representation from these five polynomials as they give nearly identical results within the amplitude range of the experiment.

**EXPERIMENTAL ERROR.** Errors always exist in experimental data. Most of them are random, but they often may be correlated because of some calibration errors or facility problem. The latter can and should be eliminated and the magnitudes of random ones known at least insofar as standard deviations are concerned. A detailed evaluation of the errors is possible on a continued basis when data are being reduced in large quantities.



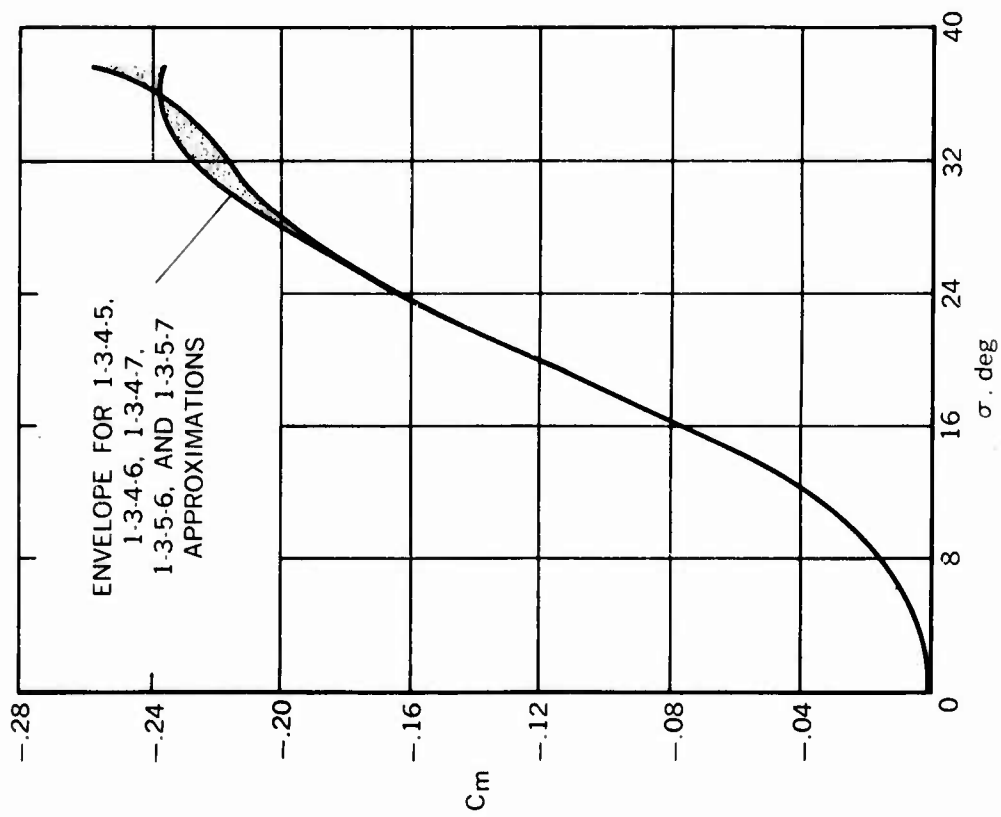


Fig. 15. Sensitivity of pitching moment modeling.



This is accomplished by saving the residuals from each curve fit and periodically examining them. The residuals obtained from a series of tests in the Ames hypervelocity Free Flight Aerodynamic Facility are shown plotted in Figure 16.<sup>1</sup> The number of times an error fell between  $\pm 0.00013$  cm about a value of  $\Delta z$  is indicated by the number of dots plotted at that level. This is done for each station. Curves are faired over each of these distributions. We can see that all of the stations appear to have similar distributions but few are centered at zero. This bias is thought to be associated with the facility. The test section is made in sections each composed of four or five windows. If each section were twisted relative to the other, a slightly different optical distortion would occur in each window. These nonzero means can be taken out of the experimental data by a calibration constant. A periodic facility calibration can be obtained if a heavy sphere is launched at very low pressure in the facility; its trajectory, which is a straight line except for gravity, can be used to calibrate the facility. Continued collection of residual information from ongoing tests can be used to check for reading errors, calibration changes, and modeling errors.

SENSITIVITY AND EXPERIMENT DESIGN. The last point to be considered is the sensitivity of deduced results to experimental setup and error. This could be done using a Monte Carlo approach, by perturbing all the data points in a random manner, analyzing the results, and repeating this process many times to see what the statistical effect on the results is. This process is time-consuming and not really required. When the differential correction procedure is used, the inverse of the A matrix represents the variance and co-variance of the parameter being determined, hence all of the information that is required to determine the sensitivity of the parameters is present when the solution is obtained.

The above procedure can also be looked at in a different manner; namely, what combination of test procedures will minimize the error in a particular parameter. This was done for a simple damped sine wave with some further simplification to allow a closed-form expression to be obtained for the expected error in parameter of interest.<sup>19</sup> The results of this are plotted in Figure 17 ofr the damping parameter  $\xi$ . The expected error in the damping parameter is plotted versus number of cycles of motion observed, N, for various number of data points per cycle, n. Also shown are some data obtained by a Monte Carlo approach. From this we see that, if the number of stations of data are fixed, we are better off with many cycles of motion and only a few points per cycle. Thus we see that we can also use our data reduction procedure to help design optimum experiments.

CONCLUSION. In the foregoing material we have covered in a rather comprehensive manner most aspects of parameter identification as it

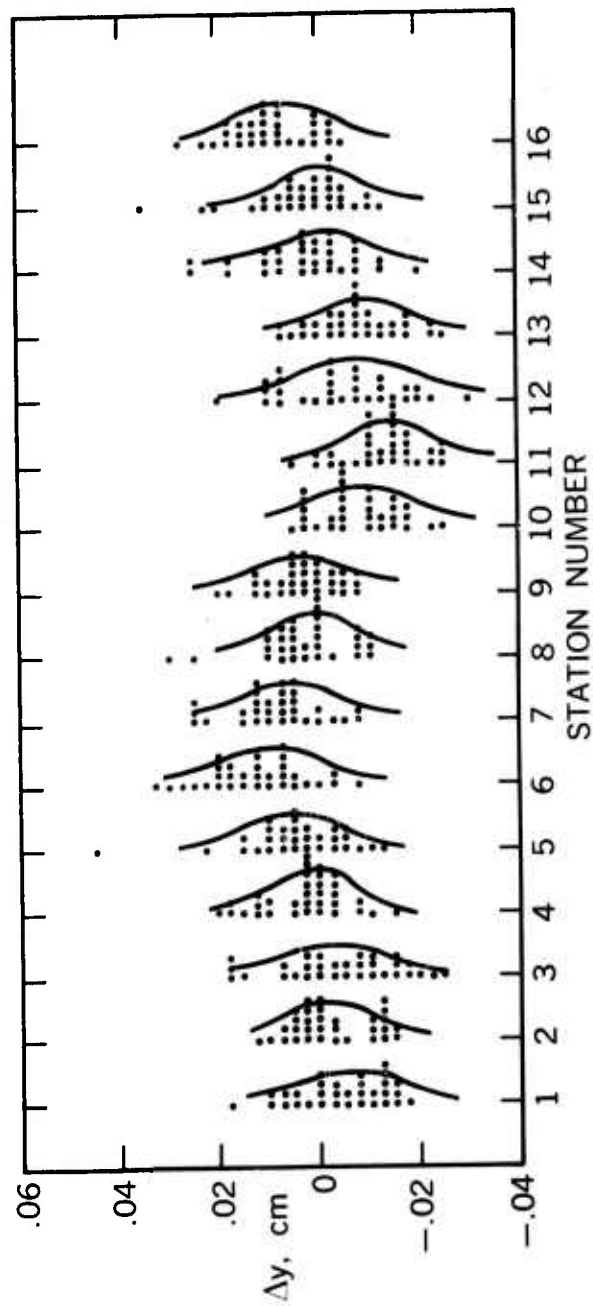


Fig. 16. Y-residuals.

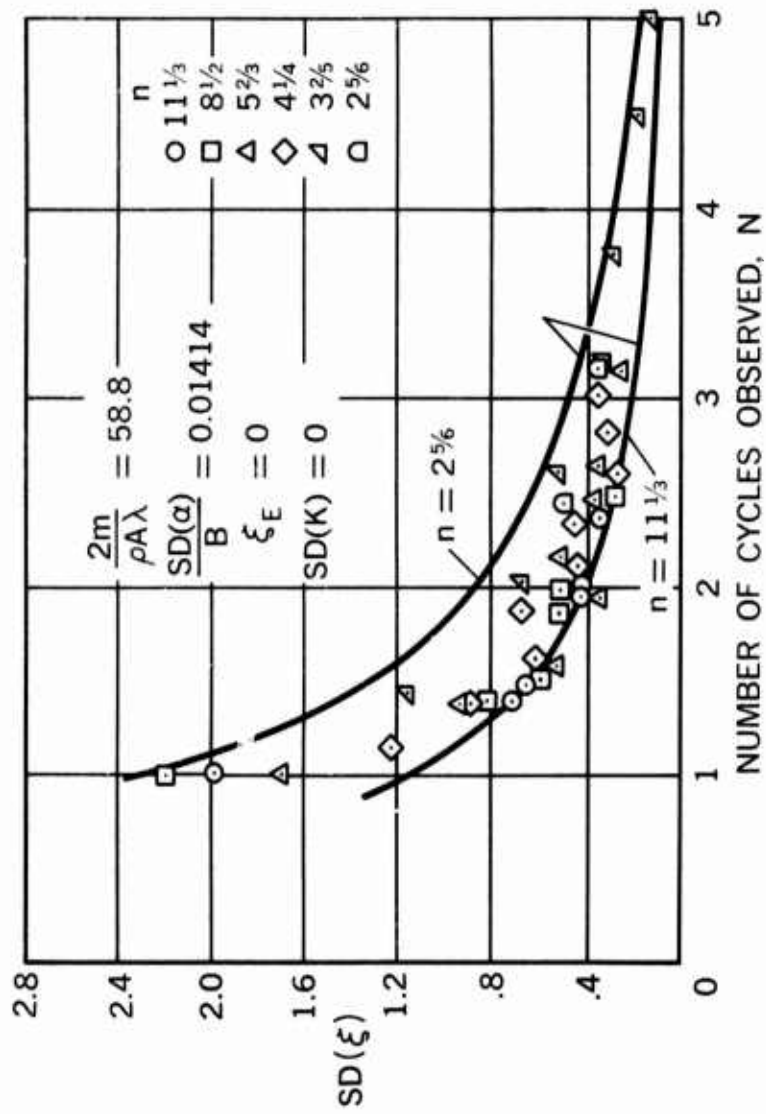


Fig. 17. Sensitivity of damping parameter.

applies to ballistic range data reduction. There are many areas where much work remains to be done but the procedures are sufficiently well-defined to allow aerodynamic results to be obtained with a high degree of confidence in their validity.

ACKNOWLEDGEMENTS. The present paper contains the experience and work of many people accumulated over several years at Ames Research Center. In particular, I would like to thank Donn Kirk, who helped in the development of much of the material discussed.

#### REFERENCES

1. Ballistic-Range Technology. Edited by T. M. Canning, A. Sieff and C. S. James. AGARDograph No. 138 1970.
2. Bridgman, P. W.: Dimensional Analysis. Paperbound ed., Yale University Press, New Haven and London, 1963.
3. Seiff, Alvin and Wilkins, Max E.: Experimental Investigation of a Hypersonic Glider Configuration at Mach Number of 6 and Full-Scale Reynolds numbers. NASA TN D-341, 1961.
4. Maple, C. G. and Synge, J. L.: Aerodynamic Symmetry of Projectiles Quarterly of Appl. Math., Vol. VI, No. 4 Jan. 1949.
5. Tobak, Murray, Schiff, Lewis B., and Peterson, Victor L.: Aerodynamics of Bodies of Revolution in Coning motion. AIAA Journal Vol. 7, No. 1, Jan. 1969.
6. Intrieri, Peter F., Kirk, Donn B., Chapman, Gary T. and Terry, James E.: Ballistic Range Test of Ablating and Nonabating Slender Cones. AIAA Journal Vol. 8, No. 3, March, 1970.
7. Unpublished data from the Planetary Atmosphere Experiment Test (PAET). Data was analyzed by Donn Kirk of Ames Research Center.
8. Unpublished data from Ames Research Center 1' x 3' and 2' x 2' Wind Tunnels, together with Several other Sources too numerous to mention, were used. The compilation was done by Robert Sammonds and Robert Kruse both of Ames Research Center.
9. Nicholaides, John D.: On The Free-Flight Motion of Missiles Having Slight Configurational Asymmetries. BRL Rep. 858, June 1953.
10. Bevington, P. R.; Data Reduction and Error Analysis for the Physical Sciences. McGraw Hill Book Company 1969.

11. Shinbrot, Marvin: A Least Square Curve Fitting Method with Application to the Calculation of Stability Coefficients From Transient Response Data. NASA TN-2341, 1951.
12. Intrieri, Peter F.: Experimental Stability and Drag of a Pointed and a Blunted 30° Half Angle Cone at Mach Numbers from 11.5 to 34 in Air. NASA TN D-3193, Jan. 1966.
13. Murphy, Charles H.: The Effect of Strongly Nonlinear Static Moment on the Combined Pitching and Yawing Motion of a Symmetric Missile. BRL Rep. 114, Aberdeen Proving Ground, Md., Aug. 1960.
14. Rasmussen, Maurice L. and Kirk, Donn B.: On the Pitching and Yawing Motion of a Spinning Symmetric Missile Governed by an Arbitrary Non-linear Restoring Moment. NASA TN D-2135, 1964.
15. Rubert, Peter E. and Landahl, Martin T.: Solution of Transonic Airfoil Problem through Parametric Differentiation. AIAA Journal Vol. 5, No. 3, March 1967.
16. Chapman, Gary T. and Kirk, Donn B.: A Method for Extraction at Aerodynamic Coefficients from Free-Flight Data. AIAA Journal Vol. 8, No. 4, April 1970.
17. Boissevain, A. G. and Intrieri, Peter F.: Determination of Stability Derivatives From Ballistic Range Test of Rolling Aircraft Models. NASA TM X-399, 1961.
18. Kirk, Donn B. and Miller, Robert J.: Aerodynamic Characteristics of a Truncated-Cone Lifting Reentry Body at Mach numbers from 10 to 21. NASA TM X-786, 1963.
19. Chapman, Gary T. and Kirk, Donn B.: Obtaining Accurate Aerodynamic Force and Moment Results from Ballistics Tests. AGARD Conference Proceedings No. 10. The Fluid Dynamic Aspects of Ballistics. Sept. 1966.

## THE WSMR BEST ESTIMATE OF TRAJECTORY - AN OVERVIEW

William S. Agee & Robert H. Turner  
Analysis & Computation Division  
National Range Operations Directorate  
White Sands Missile Range, New Mexico

### INTRODUCTION

The Special Projects Section of the Analysis and Computation Division at WSMR has developed a Best Estimate of Trajectory (BET) program for use in post-flight data reduction. The first question that might be asked is why did we develop a BET program? One reason is because of requests from Range Users (we are currently using the BET for LANCE, SAM-D, and the forthcoming 621B Navigational Satellite Tests). However, besides these requests, BET has several advantages over the conventional single instrumentation system data reduction programs currently in use. In order to see these advantages it is easiest to first review the conventional post-flight reduction procedure and note its deficiencies.

The primary instrumentation systems at WSMR are radar, cinetheodolite or fixed camera, and dovap. In the traditional method of post-flight reduction independent estimates of trajectory parameters (Cartesian components of position, velocity, and acceleration) are obtained for each of the primary systems observing the trajectory. In some cases all three of the primary systems will be tracking so that there would be three independent sets of position, velocity, and acceleration. These independent estimates are bound to disagree. This disagreement presents a difficult problem for the trajectory analyst. Another difficulty is that each of the primary systems provides measurements only for a portion of the trajectory. For example, a radar often will not provide valid tracking data during the boost phase of a missile, the optical measuring systems run out of film, dovap provides unreliable data at low altitudes. Thus, none of the independent trajectory estimates cover the entire trajectory. In addition, requirements on accuracy and precision sometimes cannot be met by reducing data from a single instrumentation system. Instead of developing and procuring new instrumentation to meet these requirements it may be possible and certainly more economical to satisfy the requirements by extending the capability of existing instrumentation by combining data from the various instrumentation systems. Telemetered or in-flight recorded data can also be included in the BET solution to

---

The remainder of this paper has been reproduced photographically from the author's manuscript.

extend the capability of existing instrumentation. In summary, we have pointed out three serious deficiencies of the conventional single instrumentation system type of data reduction procedure.

1. No single set of trajectory parameter estimates is obtained.
2. None of independent sets of trajectory parameter estimates cover the entire trajectory.
3. The inefficient use of measuring resources.

The removal of these deficiencies are the basic advantages we hope to gain by use of a BET program. In addition to the above advantages the BET will provide feedback to the instrumentation people in the form of a plotted time history of instrumentation system performance. Time histories of measurement bias, variance, and residual are a natural output of the BET program. These are plotted for each measurement participating in the trajectory solution.

#### TRAJECTORY MODELLING

As stated previously one of our primary reasons for developing a BET at WSMR is to provide a single estimate of position, velocity, and acceleration through the combination of all available range measurements. Any technique developed for this application must apply to most of the flight test programs at WSMR for which there are data reduction requirements. Obviously, this requires the use of a rather general dynamic model of the flight test trajectory.

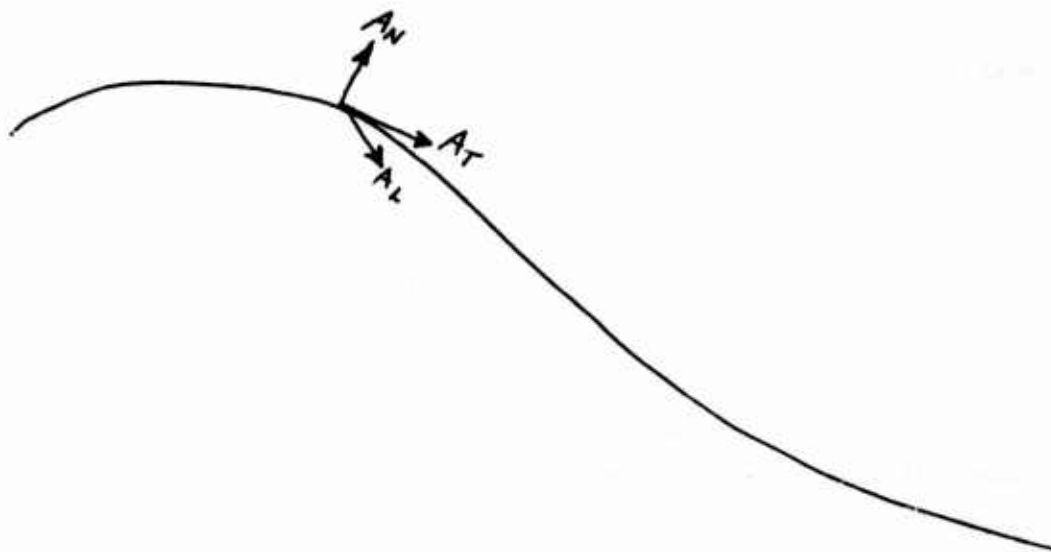
One model which meets this requirement is the second order polynomial model presently used in the data reduction process. In state variable form this model is

$$\begin{array}{rcl}
 x_1 & = & x \\
 x_2 & = & y \\
 x_3 & = & z \\
 x_4 & = & \dot{x} \\
 x_5 & = & \dot{y} \\
 x_6 & = & \dot{z} \\
 x_7 & = & \ddot{x} \\
 x_8 & = & \ddot{y} \\
 x_9 & = & \ddot{z}
 \end{array}
 \quad \dot{x} = f(x) =
 \begin{bmatrix}
 x_4 \\
 x_5 \\
 x_6 \\
 x_7 \\
 x_8 \\
 x_9 \\
 0 \\
 0 \\
 0
 \end{bmatrix}$$



This very simple model has been used with some success in situations where dynamics of the process are not too severe such as aircraft tracking and free flight missile trajectories.

We have had much more success with a dynamic model which, rather than  $\ddot{x}$ ,  $\ddot{y}$ , and  $\ddot{z}$  as states, uses acceleration components close to "where the action is" namely tangent and normal to the trajectory. Let  $A_T$  (tangential acceleration) be tangent to the trajectory,  $A_N$  (normal acceleration) be normal to  $A_T$  and lie in the vertical plane, and  $A_L$  be normal to the  $A_T$  and  $A_N$  directions and complete right-handed system.



We assume that these acceleration components do not contain the effect of gravity. Using these accelerations we define the following dynamic model.

$$\begin{array}{l}
 x_1 = x \\
 x_2 = y \\
 x_3 = z \\
 x_4 = \dot{x} \\
 x_5 = \dot{y} \\
 x_6 = \dot{z} \\
 x_7 = A_T \\
 x_8 = A_N \\
 x_9 = A_L
 \end{array}
 \quad
 \dot{x} = f(x) =
 \begin{bmatrix}
 x_4 \\
 x_5 \\
 x_6 \\
 \frac{x_4 x_7}{V} - \frac{x_4 x_6 x_8}{V_G V} + \frac{x_5 x_9}{V} \\
 \frac{x_5 x_7}{V} - \frac{x_5 x_6 x_8}{V_G V} + \frac{x_5 x_9}{V} \\
 \frac{x_6 x_7}{V} + \frac{x_8 V_G}{V} + g \\
 0 \\
 0 \\
 0
 \end{bmatrix}$$

where

$$V_G = (x_4^2 + x_5^2)^{1/2}$$

$$V = (x_4^2 + x_5^2 + x_6^2)^{1/2}$$

This is the model which we presently use in the BET program. Note that we are still using a constant acceleration assumption as in the quadratic model but that the present model is nonlinear. One reason that this model is considerably more useful is that acceleration measurements, which are made aboard the test vehicle, are usually much easier to model in terms of tangential, normal, and lateral accelerations.

## MEASUREMENT MODELLING

Besides modelling the trajectory, the measurement must also be modelled in terms of the trajectory state variables. Thus, for each measurement we must specify a nonlinear measurement function  $h(x)$  which relates the ideal measurement to the trajectory state. We assume that the position and velocity state variables ( $x_1, x_2, x_3, x_4, x_5, x_6$ ) of the trajectory are with respect to a coordinate system which we will call the launch system.

## RADAR MEASUREMENTS

Radar observations are usually from the FPS-16 instrumentation radars. These radars measure range, azimuth, and elevation of a target in a local radar Cartesian coordinate system. Some of the radars also measure the range rate of the target. The observed range, azimuth, and elevation (RAE) are first corrected for calibration and refraction. Direction cosines are computed from these corrected observations and then related to the launch coordinate system where azimuth and elevation angles are recomputed. In terms of the trajectory state variables which are in the launch coordinate system the radar measurement functions are:

### RANGE

$$h_1(x) = [(x_1 - x_I)^2 + (x_2 - y_I)^2 + (x_3 - z_I)^2]^{1/2}$$

### AZIMUTH

$$h_2(x) = \tan^{-1} \frac{x_1 - x_I}{x_2 - y_I}$$

### ELEVATION

$$h_3(x) = \tan^{-1} \frac{x_3 - z_I}{[(x_1 - x_I)^2 + (x_2 - y_I)^2]^{1/2}}$$

where  $(x_I, y_I, z_I)$  are the coordinates of radar in the launch coordinate system.

### RANGE RATE

$$h_4(x) = \frac{x_4(x_1 - x_I) + x_5(x_2 - y_I) + x_6(x_3 - z_I)}{h_1(x)}$$

### OPTICAL MEASUREMENTS

The fixed cameras and tracking cameras measure azimuth and elevation of the line-of-sight to the target. The observed angles are corrected for calibrations and refractions. The measurement functions for the cameras are the same as for the radar azimuth and elevation measurements.

$$h_2(x) = \tan^{-1} \frac{x_1 - x_I}{x_2 - y_I}$$

$$h_3(x) = \tan^{-1} \frac{x_3 - z_I}{[(x_1 - x_I)^2 + (x_2 - y_I)^2]^{1/2}}$$

where  $(x_I, y_I, z_I)$  are coordinates of camera station in the launch coordinate system.

### DOVAP MEASUREMENTS

The dovap measuring system is a two-way doppler system. The basic digitized measurement is the doppler cycle count over the sampling interval  $(t_1, t_2)$ , which when properly scaled yields the change in loop range from transmitter to target to receiver. If in addition the measurement is divided by  $(t_2 - t_1)$ , the result is the average loop range rate over the interval which approximates the instantaneous loop range rate at

$$\frac{t_1 + t_2}{2}$$

Following this procedure we represent the dovap observation by the measurement function

$$h_5(x) = x_4 \left( \frac{x_1 - x_T}{R_T} + \frac{x_1 - x_R}{R_R} \right) + x_5 \left( \frac{x_2 - y_T}{R_T} + \frac{x_2 - y_R}{R_R} \right) + x_6 \left( \frac{x_3 - z_T}{R_T} + \frac{x_3 - z_R}{R_R} \right)$$

where  $(x_T, y_T, z_T)$  and  $(x_R, y_R, z_R)$  are the coordinates of the dovap transmitter and receiver. The quantities  $R_T$  and  $R_R$  are

$$R_T = [(x_1 - x_T)^2 + (x_2 - y_T)^2 + (x_3 - z_T)^2]^{1/2}$$

$$R_R = [(x_1 - x_R)^2 + (x_2 - y_R)^2 + (x_3 - z_R)^2]^{1/2}$$

#### VELOCIMETER MEASUREMENTS

The velocimeter is much like the doppler except that it is a one-way doppler system. Again the cycle count is scaled and divided by the sampling interval,  $(t_2 - t_1)$ , and the result interpreted as the instantaneous range rate at  $(t_1 + t_2)/2$ . The resulting measurement function is

$$h_6(x) = \frac{x_4(x_1 - x_I) + x_5(x_2 - y_I) + x_6(x_3 - z_I)}{[(x_1 - x_I)^2 + (x_2 - y_I)^2 + (x_3 - z_I)^2]^{1/2}}$$

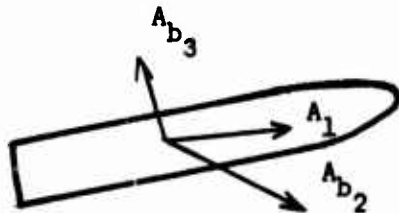
#### ACCELERATION MEASUREMENTS

Several types of acceleration measurements are possible. Often only the longitudinal body acceleration of a missile is measured and telemetered. If the missile is assumed to have zero angle of attack, which is often a good assumption, the measurement function is equal to the tangential acceleration  $A_T$  which is the seventh component of the state vector.

Thus

$$h_7(x) = x_7 \text{ (zero angle of attack)}$$

Sometimes three orthogonal components of missile body accelerations are measured and telemetered as below



If the missile is assumed to have zero angle of attack and zero roll angle, an assumption which must be given careful consideration for each individual case, the measurement functions are

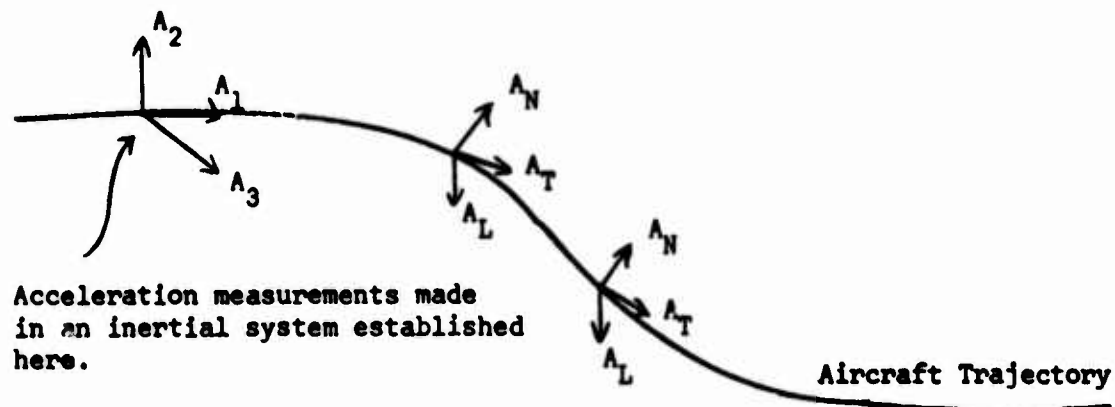
$$h_7(x) = x_7 = A_T$$

$$h_8(x) = x_8 = A_N$$

$$h_9(x) = x_9 = A_L$$

Another important class of acceleration measurements comes from inertial measurement units (IMU). This is the type of measurements on the 621B Navigational Satellite Tests. There are many configurations for IMU measurements. Some IMU's make acceleration measurements in a coordinate system slaved to the local vertical, some in an inertial coordinate system setup at a launch point, etc. In addition to acceleration measurements attitude measurements of the body with respect to the reference coordinate system are also available.

One simple type of IMU measurements which we have processed with the BET program came from a purely inertial system. In this case the inertial system was aligned with the aircraft at a specified time. The future accelerations were then measured in this coordinate system. This measurement configuration is shown below



Although inertial platform misalignments and drifts, accelerometer scale factor errors, and accelerometer zero set errors must be modeled and estimated, we assume these to be zero for the present discussion. In absence of these errors the acceleration measurements may be modelled in terms of the trajectory accelerations as

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{matrix} M_{Ie} & M_{et} \\ 3 \times 3 & 3 \times 3 \end{matrix} \begin{bmatrix} A_T \\ A_N \\ A_L \end{bmatrix}$$

Where  $M_{Ie}$  is the rotation matrix from the earth fixed launch coordinate system to the inertial system and  $M_{et}$  is the velocity dependent rotation matrix from the trajectory coordinate system to the launch coordinate system. In any case the scalar acceleration measurements are linear functions of the trajectory accelerations

$$h_{10}(x) = \begin{matrix} M^T \\ 1 \times 3 \end{matrix} \begin{bmatrix} A_T \\ A_N \\ A_L \end{bmatrix}$$

where the vector  $M = M(x)$  is state dependent.

#### EXTENDED KALMAN FILTER

Our BET program is basically an extended Kalman filter. For the case of systems whose dynamics are linear, the measurement functions are linear and all uncertainties have Gaussian statistics, the Kalman filter is known to provide the optimal recursive estimate of the state. For nonlinear systems the extended Kalman filter obtained by linearizing the nonlinear functions about the current estimate of the state has become a popular and highly useful estimation procedure for nonlinear systems.

For our extended Kalman filter we assume the dynamic trajectory model

$$\dot{x} = f(x) + w$$

where  $f(x)$  was previously given and  $w$  is a white noise term with zero mean and covariance  $Q$ . The presence of the state noise  $w$  or rather its covariance is used to compensate the filter gains for the errors made in modelling, in particular for the errors in the trajectory model due to the constant acceleration assumption.



Observations  $z(K)$  are available, at discrete instants of time  $t_K$ . We assume that the  $z(K)$ 's are statistically independent scalar observations. The processing of scalar observations provides a numerically efficient as well as intuitively appealing method of processing the observations. The assumption of statistical independence of the observations can be removed if necessary. The scalar observations are represented as

$$Z(K) = h(x(K)) + v(K)$$

where  $v(K)$  is a measurement noise term assumed to have zero mean and variance  $r^2(K)$ .

Let  $x^*(K|K-1)$  denote the filtered estimate at time  $t_K$  after processing all observations through  $t_{K-1}$ , and  $x^*(K)$  the filtered estimate at  $t_K$  after processing all observations through  $t_K$ . Assuming that the state estimate  $x^*(K-1)$  has been computed, the predicted state estimate for the next measurement time  $t_K$  is computed by numerically integrating the trajectory model  $\dot{x} = f(x)$  using a second order Taylor series integration procedure

$$x^*(K|K-1) = x^*(K-1) + f(x^*(K-1))\Delta t_K + \frac{1}{2} J(x^*(K-1)) f(x^*(K-1)) (\Delta t_K)^2$$

where

$$\Delta t_K = t_K - t_{K-1}$$

and

$$J(x^*(K-1)) = \begin{bmatrix} \frac{\partial f}{\partial x} \end{bmatrix}_{9 \times 9} x^*(K-1)$$

The covariance matrix of the predicted state estimate, which satisfies a matrix Ricatti differential equation between  $t_{K-1}$  and  $t_K$ , is computed by using a trapezoidal integration procedure. Let  $P_{K-1}^*$  denote the covariance of  $x^*(t_{K-1})$  and  $P_{K|K-1}^*$  the covariance matrix of  $x^*(K|K-1)$ . Then we compute  $P_{K|K-1}^*$  by

$$P_{K|K-1}^* = \Phi_K (P_{K-1}^* + .5Q\Delta t_K) \Phi_K^T + .5Q\Delta t_K$$

where

$$\Phi_K = I + J(x^*(K-1))\Delta t_K + \frac{1}{2} J^2(x^*(K-1)) \frac{(\Delta t_K)^2}{2}$$

and Q is the covariance of the additive state noise w.

Our extended Kalman filter employs a matrix square root formulation of the covariance equations, see Ref 1. We have found that the square root formulation not only provides a numerically stable estimation procedure but it is computationally efficient as well. For the predicted covariance matrix  $P_{K|K-1}^*$  computed above, the matrix square root  $L_{K|K-1}$  such that

$$P_{K|K-1}^* = L_{K|K-1} L_{K|K-1}^T$$

is computed by means of a Choleski decomposition, see Reference 4.

For each scalar observation occurring at the new time  $t_K$  an updated state estimate and an updated square root of the covariance matrix are computed. Let  $x^{*(i)}(K)$  denote the state estimate after processing the  $i^{\text{th}}$  scalar observation at  $t_K$  and let  $L_K^{(i)}$  denote the square root of the covariance of  $x^{*(i)}(K)$ . These quantities are computed from

$$x^{*(i)}(K) = x^{*(i-1)}(K) + \frac{L_K^{(i-1)} L_K^{T(i-1)} H_i^T}{r_i^2(K) + H_i L_K^{(i-1)} L_K^{T(i-1)} H_i^T} \left( Z_i(K) - h_i(x^{*(i-1)}(K)) \right)$$

$$x^{*(0)}(K) = x^*(K|K-1)$$

$i = 1, m = \#$  of observations at  $t_K$

$$H_i = \left[ \frac{\partial h_i(x)}{\partial x} \right]_{1 \times 9} x^{*(i-1)}(K)$$

$$L_K^{(i)} = L_K^{(i-1)} \left( I - \theta_i L_K^{(i-1)} H_i^T H_i L_K^{(i-1)T} \right)$$

$$L_K^{(0)} = L_{K|K-1}$$

$$\theta_i = \left[ 1 + \frac{r_i(K)}{(r_i^2(K) + H_i L_K^{(i-1)} L_K^{T(i-1)} H_i^T)^{1/2}} \right] / H_i L_K^{(i-1)} L_K^{T(i-1)} H_i^T$$

## MEASUREMENT BIAS ESTIMATION

So far we have neglected to discuss one of the most important considerations in the development of a BET program, namely to account for the inconsistencies produced by bias errors in the measurements. There is a natural way of including bias terms in the extended Kalman filter; one merely adds an additional state variable for each bias term to be considered and forms the optimal estimate of the biases in the same way as for the trajectory state variables. This technique is fine for cases where there are only a few bias terms to be estimated. However, a typical application of our BET program has a large number of measurements involved. For example, a LANCE flight test might have two radars, 28 doppler receivers, eight fixed cameras, and eight cinetheodolites. Considering only one bias term per measurement this results in 66 additional state variables to be estimated. With a trajectory state dimension of nine we then would have to compute estimates for 75 state variables. An ordinary Kalman filtering program using 75 dimensional state vector is computationally prohibitive at the present time. Fortunately, Friedland, see Ref 2, has developed a decomposition technique for Kalman filters which we were able to adapt and extend to the measurement bias estimation problem at WSMR. The application of this decomposition procedure has resulted in a computationally feasible BET program which includes estimation of measurement biases.

We will call the filter described previously the zero bias filter and the estimates  $x^*(k)$  obtained from this filter the zero bias estimates. Let  $b$  denote a  $p$ -vector of bias terms. We revise our previous measurement model to include these terms.

$$z_i(k) = h_i(x(k)) + g_i^T(x(k))b + v_i(k)$$

Thus, we allow the bias of each measurement to be a linear function of several bias variables. For example, a model for the bias of a radar azimuth measurement might be

$$\Delta A = g^T b = b_1 + b_2 \tan E_0 \sin A_0 + b_3 \tan E_0 \cos A_0 + b_4 \tan E_0 + b_5 \sec E_0 + b_6 \dot{A}_0$$

We assume the constant dynamic model for the bias

$$b(K+1) = b(K)$$

Note that we have not included a state noise term in the bias dynamics to account for the possible misassumption that the biases are constant. The reason for not including a state noise term will become evident later.

Now let the bias state vector  $b$  be adjoined to the trajectory state  $x$  to form the augmented state vector  $y$

$$y = \begin{bmatrix} x \\ b \end{bmatrix}_{(p+9) \times 1}$$

We could proceed directly and obtain a new extended Kalman filter giving the state estimate

$$\hat{y} = \begin{bmatrix} \hat{x} \\ \hat{b} \end{bmatrix}$$

However, as previously mentioned this is computationally prohibitive for large  $p$ . Instead we employ the filter decomposition procedure developed by Friedland. This procedure attempts to write the optimal estimate  $\hat{x}$ , which includes the effect of biases, as

$$\hat{x}(K) = x^*(K) + T(K)b(K)$$

where  $x^*(K)$  is the zero-bias estimate already obtained and  $T(K)$  is a  $9 \times p$  matrix to be determined. Upon examination we find that the decomposition holds if the filter satisfies certain restrictive conditions. The details of the derivation are tedious and will not be presented. The restriction imposed by the decomposition procedure is merely an assumption we have already made: The bias dynamics must not include a state noise term. This may not seem like much of a restriction since it was assumed to begin with, but this was hindsight. Indeed, this is a very severe restriction since the state noise covariance is used as an adjustable filter parameter to account for mismodelling errors. Fortunately, there

is another way of accounting for mismodelling errors in the bias dynamics for which the decomposition does hold. Specifically, we have been able to extend the decomposition procedure to the case of a fading memory Kalman filter in which deweighting of past observations is accomplished by exponentially weighting past residuals. We use small fading factors to account for bias mismodelling errors and use the state noise covariance of the zero-bias filter to account for trajectory mismodelling errors.

The form of the bias filter is almost identical to the zero bias filter with the residuals from the zero-bias filter forming the observations. Again we employ the square root formulation for the filter. At a new observation time we have the prediction equations

$$\begin{aligned}\hat{b}(K|K-1) &= \hat{b}(K-1) \\ C_b(K|K-1) &= C_b(K-1) \\ T(K|K-1) &= \Phi_K T(K-1)\end{aligned}$$

where  $C_b(K)$  is the square root of the covariance of  $\hat{b}(K)$ ,  $\Phi_K$  is the transition matrix from the zero-bias filter and  $T(K)$  is the combining matrix of the decomposition. For each observation  $Z_i(K)$  at the new observation time new bias estimates  $\hat{b}^i(K)$  and the square root of its covariance  $C_b^{(i)}(K)$  are computed from

$$\hat{b}^{(i)}(K) = \hat{b}^{(i-1)}(K) + w_b^{(i)}(K) \left( r_i^*(K) - \mathbf{s}_i^T(K) \hat{b}^{(i-1)}(K) \right)$$

where

$$r_i^*(K) = z_i(K) - h \left( x^{*(i-1)}(K) \right) = \text{residual from zero bias filter}$$

and  $w_b^{(i)}(K)$  is the vector Kalman gain given by

$$w_b^{(i)}(K) = \frac{C_b^{(i-1)}(K) C_b^{T(i-1)}(K) \mathbf{s}_i(K)}{a_i^2(K) + \mathbf{s}_i^T C_b^{(i-1)}(K) C_b^{T(i-1)}(K) \mathbf{s}_i}$$

$$a_i^2(K) = r_i^2(K) + H_i^T L_K^{(i-1)} L_K^{T(i-1)} H_i = \text{variance of residual}$$

$$s_i^T = g_i^T(x^{*(i-1)}(K)) + H_i^T T_{i-1}(K)$$

$i = 1, m = \# \text{ observations at } T_K$

The square root of the covariance matrix is updated at an observation by

$$C_b^{(i)}(K) = C_b^{(i-1)}(K) \left( I - \theta_i C_b^{(i-1)}(K) s_i^T(K) s_i(K) C_b^{(i-1)}(K) \right)$$

$$\theta_i = \left[ 1 + \frac{a_i}{\left( a_i^2 + s_i^T C_b^{(i-1)}(K) C_b^{(i-1)}(K) s_i(K) \right)^{1/2}} \right] / \left( s_i^T(K) C_b^{(i-1)}(K) C_b^{(i-1)}(K) s_i(K) \right)$$

For each measurement the combining matrix is updated according to

$$T_i(K) = T_{i-1}(K) - w_{*}^{(i)} s_i^T(K)$$

$$T(K) = T_m(K)$$

$$T_0(K) = T(K|K-1)$$

Where  $w_{*}^{(i)}(K)$  is the vector gain from the zero-bias filter.

The optimal state estimate  $\hat{x}^{(i)}(K)$  is computed as

$$\hat{x}^{(i)}(K) = x^{*(i)}(K) + T_i(K) \hat{b}^{(i)}(K)$$

To modify the above equations for the fading memory filter, first choose a fading factor  $\alpha_K \geq 1$ . The above equations are then replaced by

$$C_b(K|K-1) = \alpha_K C_b(K-1)$$

$$P^*(K|K-1) = \phi_K (\alpha_K^2 P^*(K-1) + .5Q\Delta T_K) \phi_K^T + .5Q\Delta T_K$$

#### COMPUTATION OF OBSERVATION VARIANCES

For each of the scalar measurements  $Z_i(K)$  a measurement noise variance  $r_i^2(K)$  must be available for use by the Kalman filter. Several possibilities exist for supplying the variance. An immediately obvious method is use the variance values given in specifications of the instrument or to use variance values computed from past performance history of the instrument. These methods are most useful when the measurement variance is stable from day to day and mission to mission. Another method which is often used is to

compute variances from measurement residuals

$$p_i(K) = r_i^*(K) - s_i^T(K) \hat{b}_i(K-1)$$

produced in the BET program. A method for computing variances from the residuals which is economical in both computing time and storage is the fading memory variance estimate defined by the following equations

$$\bar{p}_i(n) = \bar{p}_i(n-1) + \frac{(p_i(n) - \bar{p}_i(n-1))}{P_n}$$

$$P_n = 1 + wP_{n-1}, P_1 = 1$$

$$S_i(n) = wS_i(n-1) + \left(1 - \frac{1}{P_n}\right) (p_i(n) - \bar{p}_i(n-1))^2$$

$$F_n = P_n - H_n / P_n$$

$$H_n = 1 + w^2 H_{n-1}, H_1 = 1$$

$$\hat{\sigma}_i^2(n) = \frac{S_i(n)}{F_n}$$

In the above  $\bar{p}_i(n)$  is the estimate of the residual mean,  $w$ ,  $0 < w < 1$  is a fading factor, and  $\hat{\sigma}_i^2(n)$  is the variance estimate for the  $i^{\text{th}}$  measurement.

An optional approach to the estimation of the measurement variance is the fading memory variate difference technique, see Ref 3. Let  $y_i(n)$  be the  $k^{\text{th}}$  backward difference of the observation  $Z_i$ . If we assume the mean of the  $k^{\text{th}}$  differences are zero the following equations define the fading memory variate difference method

$$S_i(n) = wS_i(n-1) + y_i^2(n)$$

$$F_n = wF_{n-1} + \frac{(2K)}{K}$$

$$\hat{\sigma}_i^2(n) = \frac{S_i(n)}{F_n}$$



## REFERENCES

1. The Office of the Chief of Research and Development. Matrix Square Root Formulation of the Kalman Filter Covariance Equations, by Wm. S. Agee. Durham, NC, ARO-D. (ARO-D Report 70-1)
2. Fridland, Bernard. "Treatment of Bias in Recursive Estimation," IEEE Transactions on Control. Oct 68.
3. White Sands Missile Range. Exponentially Weighted Variate Differences (A Technique for Real Time Variance Estimation), by Wm. S. Agee. White Sands Missile Range, NM, STEWS-NR-AM, 27 Aug 70. (Technical Report No. 16)
4. R. S. Martinez, G. Peters, and J. H. Wilkinson, "Symmetric Decomposition of a Positive Definite Matrix," Numerische Mathematik, Bund 7, p362-383 (1965).

A COMPUTER PROGRAM TO INVESTIGATE  
EXO-ATMOSPHERIC ENGAGEMENTS OF INTERCEPTORS  
AND RE-ENTRY VEHICLES

LTC M.L. Roberson, CPT C. Van Nostrand and W.A. Barbieri  
Development Branch, ACS/Studies and Analysis  
HQ, US Air Force, Washington, D.C.

Computer modeling of military problems is accomplished by a variety of techniques. The most common is monte carlo. There are disadvantages in the use of all techniques and three frequently cited in the use of monte carlo are; the failure to calculate the number of replications required, disregarding the effects of events that occur only at extreme ends of distributions, and the large amount of computer time that may be required. The technique of building a variance-covariance matrix to calculate total distribution effects from one iteration minimizes these disadvantages while maintaining the flexibility of application that is characteristic of monte carlo models. This presentation describes the application of this technique to the problem of investigating the interactions involved in a "one on one" engagement of an anti-ballistic missile (ABM) versus a re-entry vehicle (RV) in the exo-atmosphere. First, the physical aspects of the problem will be described. Then the application of the solution method will be discussed. Finally, the computer program will be outlined.

Assume that the probability of kill of an RV by an ABM is a function of the miss distance resulting from the geometry associated with a given impact point. Figure one presents the environment as defined by the problem. From launch to impact the RV trajectory is Keplerian on a round non-rotating earth. The launch and initial impact point, the Missile Site Radar (MSR) location, the Parimeter Acquisition Radar (PAR) location, and the ABM launch location are specified in latitude and longitude.

When the RV is within the specified maximum range, maximum off-boresight angle, and minimum elevation angle, the PAR begins tracking. The quality of each tracking point is a function of RV radar cross-section, off-boresight angle, range, and the radar quality parameters such as bandwidth, antenna gain, and wave-length.

Similarly when the RV is within maximum MSR range, maximum off-boresight angle, and minimum elevation angle, the MSR begins tracking. Again the quality of each tracking point is a function of RV radar cross-section, off-boresight angle, range, and the radar quality parameters. Tracking data is taken from detection until the last midcourse correction of the ABM.

# THE EXO - 1 CONCEPT OF THE EXOATMOSPHERIC INTERCEPT

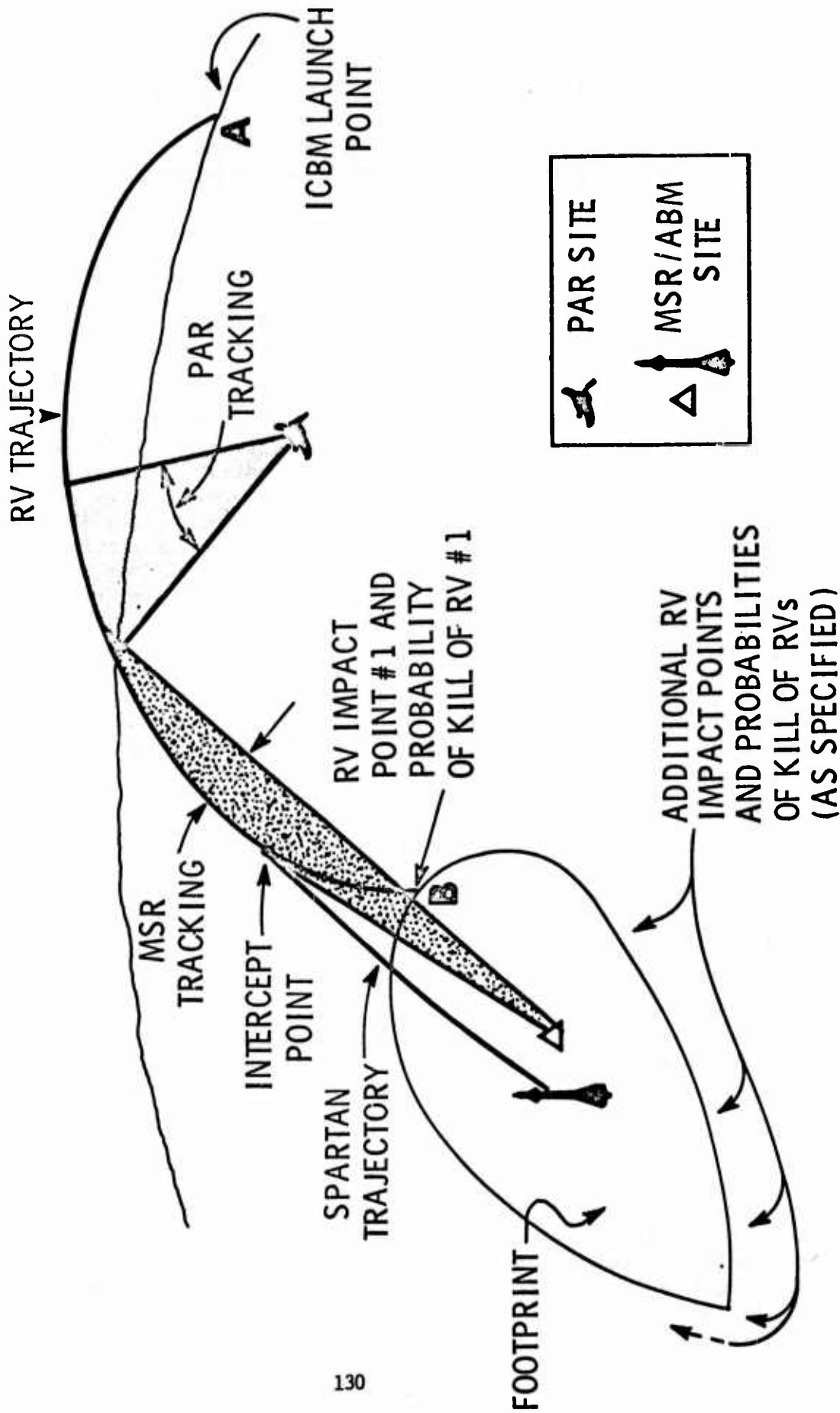


FIGURE 1

Constraints may be placed on the ABM launch such as, minimum track time, minimum intercept altitude, intercept within MSR coverage, etc. When these constraints are fulfilled the ABM is launched. From burnout to intercept the trajectory is assumed to be nearly Keplerian. Tracking data is taken from burnout to the last midcourse correction.

The problem does not include details such as blackout or allocation of radar power although the approach taken to the model makes it readily accept modifications and additional equations. The four major areas of concern in the model are:

1. The geometrical considerations in general such as locations, fly-out curves and times, and performance constraints.
2. The RV tracking points and position prediction error.
3. The ABM tracking points and position prediction error.
4. Determining the  $P_k$  based on prediction errors.

#### KEND AND EPERAN

In this presentation I will describe the mathematical basis for two of the subroutines KEND and EPERAN used in EXO-1, rather than describe them in detail or show results of running them. [See Chart 1]

#### KENDALL

The basic work was published in a Rand Report for ARPA by Dr. William Kendall in February 1963 in a paper entitled "The Probability Distribution of Anti-Missile Missile Miss Distance Due to Observations and Guidance Noise." In the preface he mentioned the problem stemmed from setting accuracy specifications for the observing instruments of a mid-course intercept system intended to operate against submarine-launched ballistic missiles and that results could be applied to a wide variety of intercept situations.

In the problem, it is assumed that the position of the target at some future time is estimated from a set of measurements (possible correlated) of any type, and that the error in the future position estimate is related to the errors in the measurements by the usual linear relation. This implies that if measurement errors are Gaussian, future position errors will also be jointly Gaussian. It is also assumed that the noise in the interceptor guidance system leads to uncertainty in the interceptor's position at intercept characterized by a three-dimensional Gaussian distribution. Under these assumptions, the probability that the miss distance between interceptor and target will be less than any given amount is determined analytically. An important feature of the solution is that the effects of interceptor guidance error, and of the tactical geometry and measurement accuracies can be separated.

**PROBABILITY DISTRIBUTION  
FUNCTION OF  
MISS DISTANCE  
(KEND)**



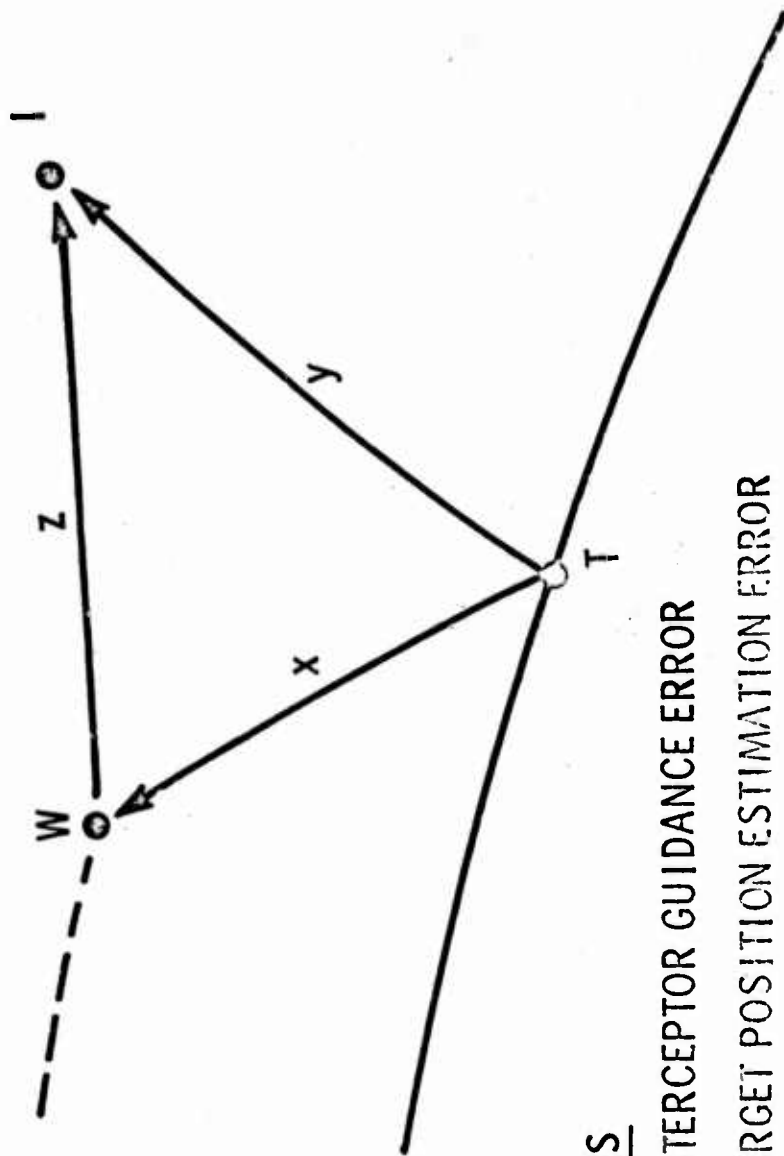
**TRACKING AND POSITION  
ERROR PREDICTION  
(EPERAN)**

CHART 1

# INTERCEPT

## POSITIONS

- W = ESTIMATED POSITION OF TARGET  
= INTERCEPTOR AIM POINT
- I = ACTUAL INTERCEPTOR LOCATION
- T = ACTUAL TARGET LOCATION



## ERRORS

- z = INTERCEPTOR GUIDANCE ERROR
- x = TARGET POSITION ESTIMATION ERROR
- y = MISS DISTANCE = x + z

CHART 2

Here is the situation at intercept. [See Chart 2] The target actually moves along the solid line and at the time of intercept is at position T. The measurement system takes its measurements at an earlier time, and estimates that the target will be at position W at the time of intercept. The interceptor is aimed at the targets estimated position at time t, but due to guidance errors actually arrives at position I. Z is the interceptor guidance error, the difference between where it was aimed and where it arrived. X is the target position error at intercept. The difference between where it is and where it was predicted to be. Y is the distance between actual locations, the miss distance and the vector sum of the two errors.

In chart 3 we state the mathematical problem. If measurement errors have jointly Gaussian distribution, then to a usually good approximation, errors in W, i.e., X, are jointly Gaussian with covariance matrix cap X. The guidance noise induced errors in the rectangular coordinates of the AMM are assumed to be jointly Gaussian random variables with zero means and covariance matrix Z. Thus the rectangular coordinates of the difference in actual positions is  $x + z = y$ , which is Gaussian, zero mean, and assuming statistical independence of the target position estimate and interceptor errors, has covariance matrix  $X + Z = Y$ .

The miss distance squared is shown in terms of its components. [Charts 4 and 5]

re (8) (1) Moments related to cumulants

$$v_1 = k_1$$

$$v_2 = k_2 + k_1^2$$

$$v_3 = k_3 + 3k_1k_2 + k_1^3$$

Cumulants defined by

$$\ln \phi(v) = \sum_{i=1}^{\infty} \frac{(jv)^i k_i}{i!} \quad \text{inf series in } (jv)^i$$

(2) Take log of  $\phi(v)$

Identify relating  $\det[I - 2jvY]$  to

$$\exp \left\{ -\sum_{i=1}^{\infty} \frac{(2jv)^i}{i} \text{tr } Y^i \right\}$$

$$\ln \phi(v) = \frac{1}{2} \sum_{i=1}^{\infty} \frac{(2jv)^i}{i} \text{tr } Y^i$$



# MISS DISTANCE DISTRIBUTION FUNCTION

$$y = x + z$$

$$[X] + [Z] = [Y]$$

$$r^2 = \sum_i (x_i + z_i)^2 = y^T y$$

CALCULATE: Prob.  $y^T y \leq k^2$  FOR ANY  $k$  GIVEN pdf  $y$

$$f(y) = M_q \exp \left\{ -\frac{1}{2} y^T [Y]^{-1} y \right\}$$

$$\text{WHERE } M_q = (2\pi)^{-\frac{q}{2}} (\det Y)^{-\frac{1}{2}}$$

CHART 3

## OUTLINE OF SOLUTION

(1) CHARACTERISTIC FUNCTION OF  $r^2 = \phi(v)$

$$\phi(v) = [\det(I - 2jvY)]^{-\frac{1}{2}}$$

(2) pdf OF  $r^2$  FROM INVERSE FOURIER TRANSFORM

(3)  $q > 2$ , NO CLOSED FORM SOLUTION OF INTEGRAL

(4)  $f(r^2)$  CAN BE EXPANDED IN LAGUERRE SERIES

$$f(r^2) = \sum_{i=0}^{\infty} c_i e^{-t} t^a L_i^a(t) \quad t > 0$$

WHERE

$$t = r^2/b$$

(5) COEFFICIENTS OF SERIES EXPRESSED IN TERMS OF MOMENTS OF  $f(r^2)$  AND  $a, b$ .

(6) COEFFICIENTS OF LAGUERRE POLYNOMIALS

$$c_0 = \frac{1}{b\Gamma(a+1)} = \frac{v_1}{\sigma^2 \Gamma(v_1^2/\sigma^2)}$$

$$c_1 = c_2 = 0$$

$$c_i \quad i \geq 3 \text{ } \} \text{ NEGLIGIBLE}$$

$$\sigma^2 = v_2 - v_1^2$$

CHART 4

## OUTLINE OF SOLUTION

(7) DETERMINE MOMENTS FROM  $\ln \phi(v)$ .

$$v_1 = k_1$$

$$v_2 = k_2 + k_1^2, \text{ ETC.}$$

$$k_i = 2^{i-1} (i-1)! \operatorname{tr} Y^i$$

(8) MOMENTS OF  $f(r^2)$  ARE

$$v_1 = \operatorname{tr} Y$$

$$v_2 = 2 \operatorname{tr} Y^2 + (\operatorname{tr} Y)^2$$

ETC.

$$(9) f(r^2) = \frac{v_1}{\sigma^2} \Gamma\left(\frac{v_1}{\sigma^2}\right) \left[\frac{v_1 r^2}{\sigma^2}\right]^{(v_1/\sigma^2)-1} \exp\left\{-\frac{v_1 r^2}{\sigma^2}\right\} \quad r^2 > 0$$

(10) CUMULATIVE DISTRIBUTION FUNCTION OF  $r^2$

$$F_{r^2}(k^2) = \operatorname{Prob.} (r \leq k) = \operatorname{Prob.} (r^2 \leq k^2) = \int_{-\infty}^{k^2} f(r^2) d(r^2) \\ = I\left[\frac{k^2}{\sigma}, \left(\frac{v_1}{\sigma}\right)^2 - 1\right]$$

INCOMPLETE GAMMA FUNCTION

re (9) This is of the form of the generalized chi-squared distribution with

$$2 (v_1/\sigma)^2 \text{ degrees of freedom.}$$
$$1 \leq 2 (v_1/\sigma)^2 \leq q$$

Parametric solution curves are displayed in chart 6.

This is one of several convenient graphs from Kendall's report. On the ordinate is the probability that the miss distance is less than some value  $k$  and as the abscissa the parameter  $k/\sqrt{\sigma}$ . Curves are parametric in  $v$  and  $\sigma$  and these in turn are simple functions of the covariance matrix of  $Y$ . The effects of target position prediction and interceptor position are separable.

#### Application to our problem

In order to determine  $v_1$  and  $\sigma$  we must know  $\text{tr}Y$  and  $\text{tr}Y^2$  where [See Chart 7]

$$Y = X + Z$$

Kendall's paper also deals with the sensitivities of the errors in measurements and predicted target position. In our case EPERAN provides the covariance matrix of the target,  $X$ , directly. Assume that noise in Interceptor is such that the position uncertainty at time  $t$  has spherical symmetry (e.g., when only the rms value of the interceptor miss distance is known) then

$$Z = (a/q)I$$

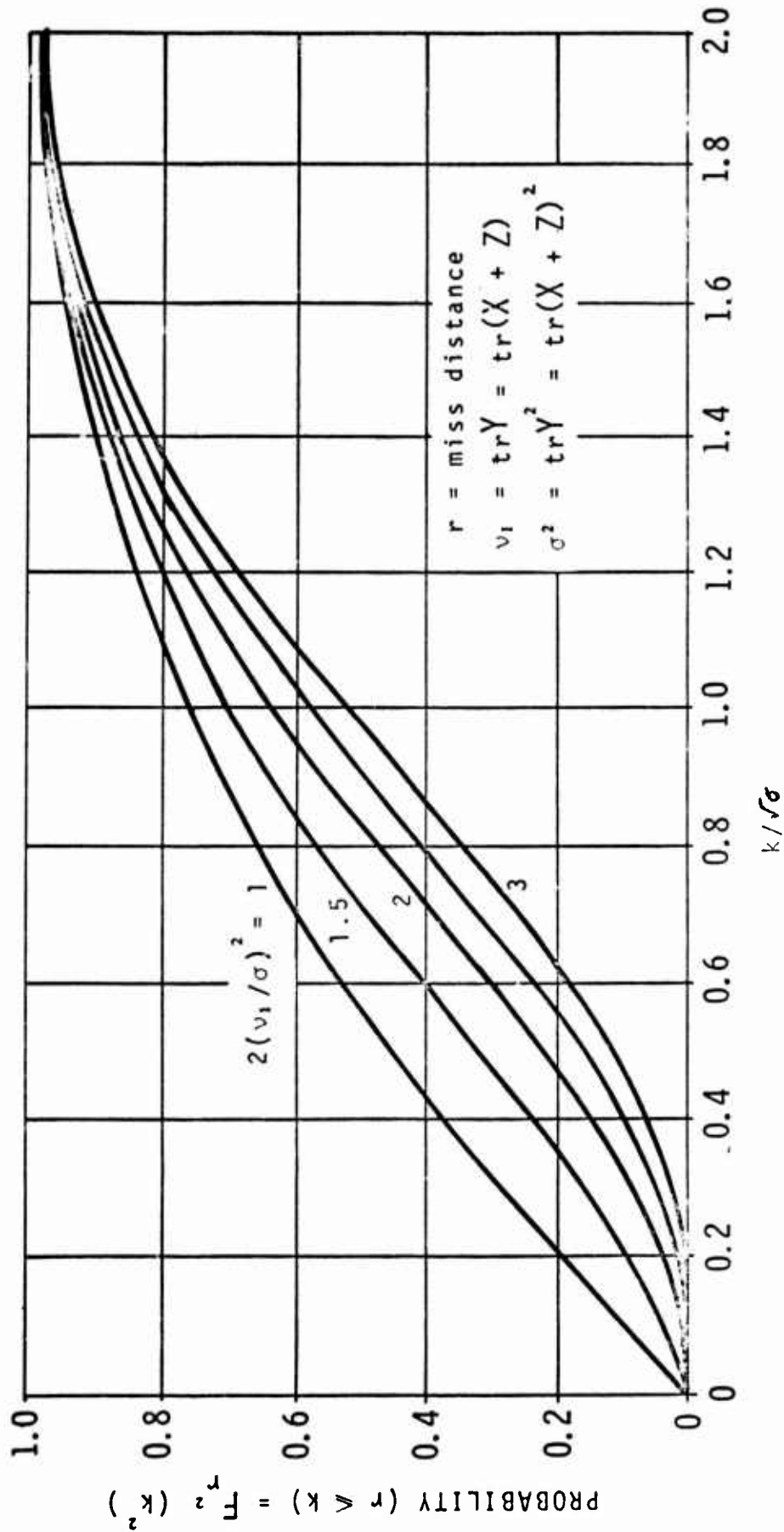
where  $a$  is mean squared miss distance due to interceptor guidance noise. This is a useful approximation when the interceptor guidance errors are much smaller than target position error, and the non-spherical nature of the interceptor error is unimportant.

Trace of  $Y^2$  is the sum of three terms. The first two are immediately available. The last requires a matrix multiplication of known  $3 \times 3$ .

That completes the description of the calculation of the probability distribution miss distance. I will now describe the subroutine EPERAN which provides the covariance matrices. [Chart 8]

EPERAN is a computer program for determining the accuracy of instrumentation used in estimating the position and velocity of a vehicle in near-Keplerian motion.

# CUMULATIVE PROBABILITY DISTRIBUTION OF $r$ ; LINEAR SCALE



## APPLICATION

$$\begin{aligned}\text{tr}Y &= \text{tr}(X + Z) \\ &= \text{tr}X + \text{tr}Z\end{aligned}$$

ASSUME  $X$  GIVEN  
INTERCEPTOR ERRORS SPHERICAL, i.e.

$$Z = \frac{a}{q}I \quad a = \text{ms miss}$$

$$\text{tr}Y = \text{tr}X + a$$

$$\begin{aligned}\text{tr}Y^2 &= \text{tr}(X + Z)^2 \\ &= \text{tr}X^2 + \text{tr}2XZ + \text{tr}Z^2\end{aligned}$$

$$\text{tr}Z^2 = \frac{2a}{3}$$

$$\text{tr}2XZ = \frac{2a}{3} \text{tr}X$$

$$\text{tr}X^2 = \text{MATRIX MULTIPLICATION OF KNOWN } 3 \times 3.$$

**PROBABILITY DISTRIBUTION  
FUNCTION OF  
MISS DISTANCE  
(KEND)**

**TRACKING AND POSITION  
ERROR PREDICTION  
(EPERAN)**



CHART 8

Reproduced from  
best available copy.



An early version of EPERAN was described in a Project Rand Report by Gabler, Belcher, and Johnson in April 1963, entitled "A Computing Program for Determining Certain Statistical Parameters Associated With Position and Velocity Errors for Orbiting and Re-entering Space Vehicles". An improved version appeared two years later in RM-4740-PR by Gabler and Belcher, entitled "A Computer Program for Tracking Error Analysis of Keplerian Trajectories". With some minor modification, this is the version that was used in this application:

Any computer program that relates errors in a series of measurements to the covariance matrix of position errors at the time of intercept can, of course, be used. EPERAN was chosen as the simplest program that provided the needed calculation. It was readily available, well documented, and had been checked against other models because of its use in range instrumentation analysis, and in studies of space tracking.

In EPERAN it is assumed that estimates for orbital parameters have been made by a general least squares (differential correction) routine which weights observations inversely with their assigned standard deviations. All partial derivatives used in error propagation are obtained from analytic formulas. Provision is made for multiple tracking stations that move along great circle paths. The results are given in terms of a coordinate system associated with the trajectory plane of the tracked vehicle.

#### CALCULATION TECHNIQUE

In the process of statistical estimation of parameters by maximum likelihood or by general least squares, a covariance matrix is usually obtained as a representation of the error in the parameters. The inverse of this variance-covariance matrix is called the information matrix. [Chart 9] It was found to be more convenient to use the information matrix than the variance-covariance matrix. When it is non-singular, the information matrix may be inverted to give the variance-covariance matrix.

The information matrix associated with a column vector X is propagated to a column vector y by the transformation

$$S_y = A^T S_x A$$

where  $S_x$  is the information matrix for vector x,  $S_y$  for vector y and A is the matrix of partial derivatives.

Each piece of tracking data contributes to the final information matrix associated with the error at intercept time. The contribution is determined by the variance of the measurement and the functional relationship between the measurement and the estimated parameters.

# ERROR PROPAGATION

EPERAN: TRACKING ERROR ANALYSIS

$S_x$  = INFORMATION MATRIX

= INVERSE OF VARIANCE-COVARIANCE MATRIX

$$S_y = A^T S_x A$$

WHERE

$$A = \begin{bmatrix} \frac{\partial x_1}{\partial y_1}, & \frac{\partial x_1}{\partial y_2}, & \dots, & \frac{\partial x_1}{\partial y_n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial x_m}{\partial y_1}, & \dots, & \dots, & \frac{\partial x_m}{\partial y_n} \end{bmatrix}$$

For example, each range measurement has an assigned variance of the form

$$\sigma_p^2 = d_1 \rho^{d_2} + d_3 \cos \theta_e + d_4$$

where the  $d_i$  are functions of the (radar) or other measurement system.  $\theta_e$  is the elevation angle. [Chart 10] The forms allow for range dependent or independent errors, and can be gotten either from experimental data or from theoretical expressions involving signal to noise ratio and pulse shape factors. The information is measured by the reciprocal of the variance

$$\phi_p = \frac{1}{\sigma_p^2}$$

The information is propagated through a succession of column vectors arranged according to the functional relations in accordance with equations of Chart 11.

Finally, the contribution of each piece of tracking data to the final information matrix is accumulated in  $S$  and at the end of the time interval  $R = S^{-1}$  inverted to give the variance-covariance matrix of position and velocity errors.

$J_1$  is the required covariance matrix of position errors.

Eigen values and 50 and 95 percent confidence regions associated with position and velocity errors are also determined.

Chart 12 shows a typical input setup. Chart 13 shows an intermediate tracking point printout and the  $R$  matrix.

In summary, then, the probability of kill distribution is found in terms of covariance matrices of target and interceptor errors. The covariance matrices are calculated from the set of measurements.

Colonel Roberson will now describe other aspects of the computer model.

#### PART IV

The scenario for EXO-1 has been described and the method for calculating the effects of error in position has been presented. I will describe the model organization and show characteristic results.

The EXO-1 Model, simulating (ABM) interceptors controlled by radar engaging a ballistic missile re-entry vehicle (RV), was configured from several existing programs. These programs, outlined on Chart 14, are:

EACH MEASUREMENT HAS ASSOCIATED VARIANCE.

EXAMPLE, RANGE:

$$\sigma_{\rho}^2 = d_1^2 + d_2^2 + d_3^2 \cos^2 \theta + d_4^2$$

$\{d_i\}$  ARE SENSOR SYSTEM PARAMETERS.

INFORMATION:

$$\varphi_{\rho} = \frac{1}{\sigma_{\rho}^2}$$

CHART 10

RANGE

$$S_1 = E^T D^T C^T B^T A^T \varphi_\rho \text{ A B C D E}$$

RANGE RATE

$$S_2 = E_1^T D_1^T C_1^T B_1^T A_1^T \varphi_\rho \text{ A}_1 \text{ B}_1 \text{ C}_1 \text{ D}_1 \text{ E}_1$$

ELEVATION ANGLE

$$S_3 = E^T D^T C^T B^T A_2^T A_1^T \varphi_{\theta.e} \text{ A}_1 \text{ A}_2 \text{ B C D E}$$

AZIMUTH ANGLE

$$S_4 = E^T D^T C^T B^T A_4^T A_3^T \varphi_{\theta.a} \text{ A}_3 \text{ A}_4 \text{ B C D E}$$

$$R = S^{-1}$$

CHART 11

SAMPLE OF COMPUTER PRINTOUT  
CHART 12

INITIAL CONDITIONS

XO C.127040CRF-CC -0.90144505E 00 ZC 0.52173198E 00 XC DOT 0.7179R393E-01 YO DOT 0.1179296E-01 ZO DOT 0.40264504E-02 TA 0.  
 A 0.105871E 01 0.9955575E-02 F 0.30700000E 02 G 0.50746283E 02 H 0.19497049E 03 I 0.33924629E 02 J 0.17492600E 01 K 0.4754981E-13  
 SO DOT  
 G.7250717CF-C1

ECCENTRIC ANCPALY AT ZERO = 0.3361C145F 02

TH = C.100000CF 01 VA = 0.375C7874E 02 ECCENTRIC ANOMALY = 0.37557231E 02 MEAN ANOMALY 0 = 0.33292995E 02

MEAN ANOMALY = 0.37207539E 02

TRACKER NUMBER 1

LATITUDE	LONGITUDE	HEIGHT	S	R	T
C.31523467E 02	-0.81392970E 02	0.10000000E 01	0.	0.	0.
01 THROUGH 016	-0.	-0.	0.91347870E-12		
-0.	-0.	-0.	0.32885233E-12		
-0.	-0.	-0.	-0.		
-0.	-0.	-0.	-0.		

TRACKER NUMBER 2

LATITUDE	LONGITUDE	HEIGHT	S	R	T
C.2252827E 02	-0.81236221E 02	0.10000000E 01	0.	0.	0.
01 THROUGH 016	-0.	-0.	0.91347870E-12		
-0.	-0.	-0.	0.32885233E-12		
-0.	-0.	-0.	-0.		
-0.	-0.	-0.	-0.		

TRACKER NUMBER 3

LATITUDE	LONGITUDE	HEIGHT	S	B	T
0.29973467E 02	-0.77482333E 02	0.10000000E 01	0.	0.	0.
01 THROUGH 016	-0.	-0.	0.91347870E-12		
-0.	-0.	-0.	0.32885233E-12		
-0.	-0.	-0.	-0.		
-0.	-0.	-0.	-0.		

X 0.12427313F-C0 -0.90022417E 00 Y 0.90022417E 00 Z 0.52212283E 00 X DOT 0.71735839E-01 Y DOT 0.12417338E-01 Z DOT 0.37762955E-02

XG 0.13382925E-CC -0.80022430E 00 YG 0.80022430E 00 ZG 0.52212283E 00 XG DOT 0.87802326E-01 YG DOT 0.1173381E-01 ZG DOT 0.37762955E-02

R 0.10493006E 01 S DOT 0.7230333E-01 LAT 0.29970771E 02 LONG 0.91544758E 02

# CHART 13 OUTPUT

RHO = 0.811174C3F-01 RMC DOT = 0.47771621E-01  
 THETA E = 0.35955211E 02 THETA A = 0.60364657E 02  
 RHO = 0.45565960E-01 RMC DOT = 0.76063966E-02  
 THETA F = 0.83717034E 02 THETA A = 0.52665554E 02  
TIME = 0.99999999F 00

X                    Y                    Z                    X DOT                    Y DOT                    Z DOT  
 0.20564383E-C0 -0.88565446E 00 0.52464946E 00 0.70922267E-01 0.16686759E-01 0.12664049E-02  
 XG                    YG                    ZG                    XG DOT                    YG DOT                    ZG DOT  
 0.20117879E-0C -0.88663391E CC 0.52464946E 00 0.67122490E-01 0.15464145E-01 0.12664049E-02  
 R                    R DOT                    S DOT                    LAT                    LUNG  
 0.10497287E 01 0.44815654E-03 0.72869891E-01 0.29986460E 02 -0.77203642E 02  
 V = 0.38335617E 02 W = 0.50054873E 02 OMEGA = 0.19496508E 03 THETA G = 0.27575292E-00

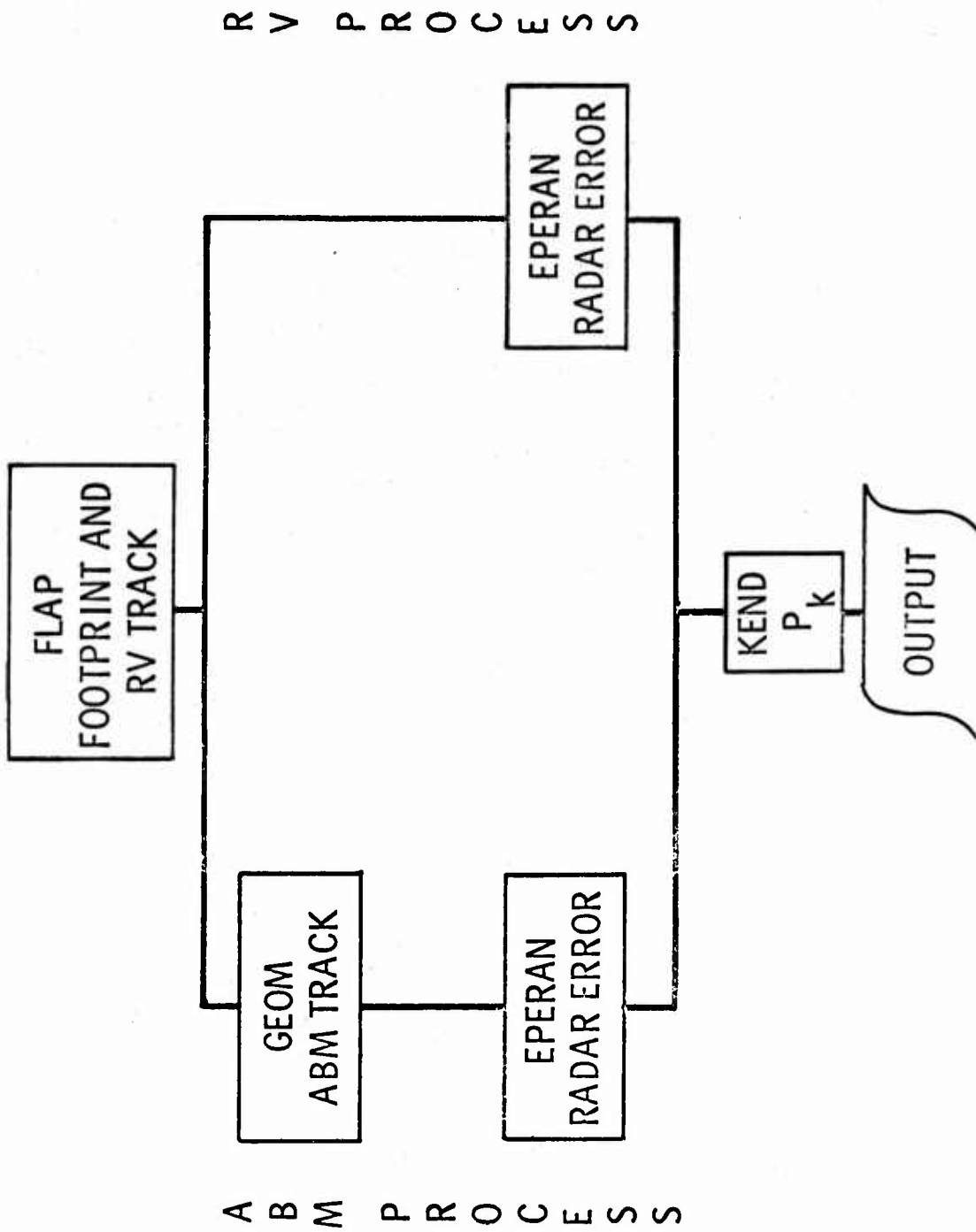
ECCENTRIC ANOMALY = 0.37951851E 02 MEAN ANOMALY = 0.37599487E 02  
 RHO = 0.86192462F-C1 RMC DOT = 0.49603913E-01  
 THETA E = 0.33248547E 02 THETA A = 0.11284576E 03  
 RHO = 0.86026015E-01 RMC DOT = 0.50285959E-01  
 THETA E = 0.3332635E 02 THETA A = 0.6283052E 02  
 RHO = 0.5117C33E-C1 RMC DOT = 0.16276391E-01  
 THETA E = 0.76C25873F 02 THETA A = 0.90395642E 02

ST MATRIX  
 C.81629305F C7 0.12526485E 05 0.77357558E 05 -0.30230C72E 05 -0.86669378E 03 0.60615495E 05  
 0.12526479E 05 0.57189552E 06 -0.27697057E 05 0.70048100E 02 -0.59050627E 04 0.19209278E 03  
 0.7735756CE 05 -0.27897057E 05 0.30053692E 07 -0.14150591E 05 -0.62470804E 01 -0.15006651E 05  
 -0.30230770E 05 0.70046059E 02 -0.14150591E 05 0.45757961E 03 0.32198735E 01 -0.14824587E 03  
 -0.86669307F 03 -0.59050628E 04 -0.62471852E 01 0.32198734E 01 0.14170725E 03 -0.56494631E 01  
 0.50815676E 05 0.19209274E C3 -0.15006651E C5 -0.14824588E C3 -0.56494631E C1 0.58670194E C3

R INVERSE MATRIX  
 0.17359742F-C5 0.81281744E-09 -0.10107936E-05 0.16071796E-04 0.19409443E-05 -0.20335974E-03  
 0.81281744E-09 0.30722581E-05 0.34793475E-07 0.11466975E-05 0.12807136E-03 0.10395410E-05  
 -0.10107936E-05 0.34793455E-07 0.11349625E-05 0.14276689E-04 0.64929593E-06 0.14192208E-03  
 0.16071796E-04 0.11466975E-05 0.14276689E-04 0.3555217E-02 0.65629213E-05 -0.4108190E-03  
 0.19409443E-05 0.12807136E-03 0.64929593E-06 0.65629213E-05 0.12401084E-01 -0.10915732E-03



# EXO - 1



MODEL ORGANIZATION

CHAPTER 17

Redstone Arsenal's "Fast Look Analysis Program" (acronym FLAP); Rand Corporation's programs for tracking error analysis (EPERAN) and computation of ABM probability of kill (KEND). A routine was written to produce the geometry of the ABM flyout curves (GEOM).

Flap constructs the scenario to be studied and calculates the RV Keplerian track up to the earliest possible ABM interception point. Flap also calculates initial detection time. This procedure is repeated for each point desired around the maximum intercept foot print. A system parameter may then be modified and the new set of detection and intercept points and times will be calculated.

Eperan determines the accuracy of estimating the position and velocity of a vehicle in Keplerian motion by the method just discussed. To compute that portion of the total error that is attributed to errors in the prediction of the position of the RV at expected time of intercept, Eperan gets the initial tracking conditions (detection position, velocity and time) and ending time (intercept time) from FLAP. The initial tracking conditions for the ABM are not produced by FLAP. Therefore an intermediate program (GEOM) is used to calculate the burnout velocity vector of an ABM boost profile to the RV interception point. The assumptions made in GEOM are:

1. All of the ABM's fuel is consumed.
2. Flight of the interceptor after burnout is in a vacuum.
3. The earth is spherical and not rotating.
4. Linear interpolation between values of parameters defining trajectories to two adjacent points will define an additional realizable trajectory to a point in-between.

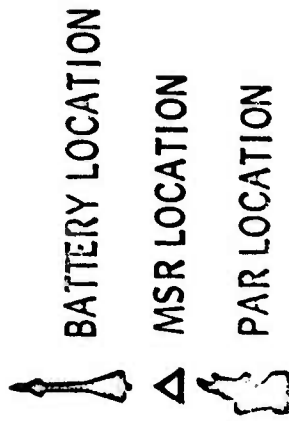
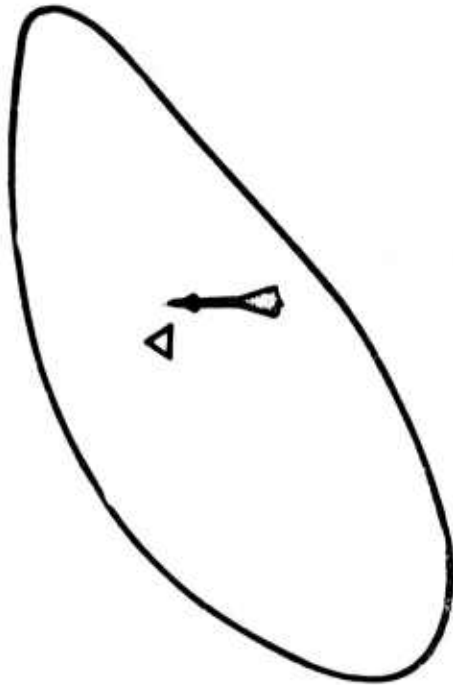
From Geom and Flap, Eperan gets sufficient conditions to compute the error in the prediction of the position of the ABM at expected time of intercept. Eperan also requires RV and ABM cross sections and radar description parameters. The output from Eperan in each case is a variance co-variance matrix of position errors. This information is used by the KEND program to calculate the mean and standard deviation of the miss distance. Finally, the probability of kill for the given weapon radius is computed and printed.

Chart 15 shows a maximum coverage foot print that is a result of FLAP. The additional information that can be obtained from EXO-1 is shown on Chart 16 with isoquants super-imposed over the maximum coverage contour.

What we have presented is a problem that was solved by simulation. However, from the start it appeared to us, that the use of the Monte Carlo technique would require such extensive computer time that use of the model would be greatly restricted. The analytical solution via the

# FLAP OUTPUT

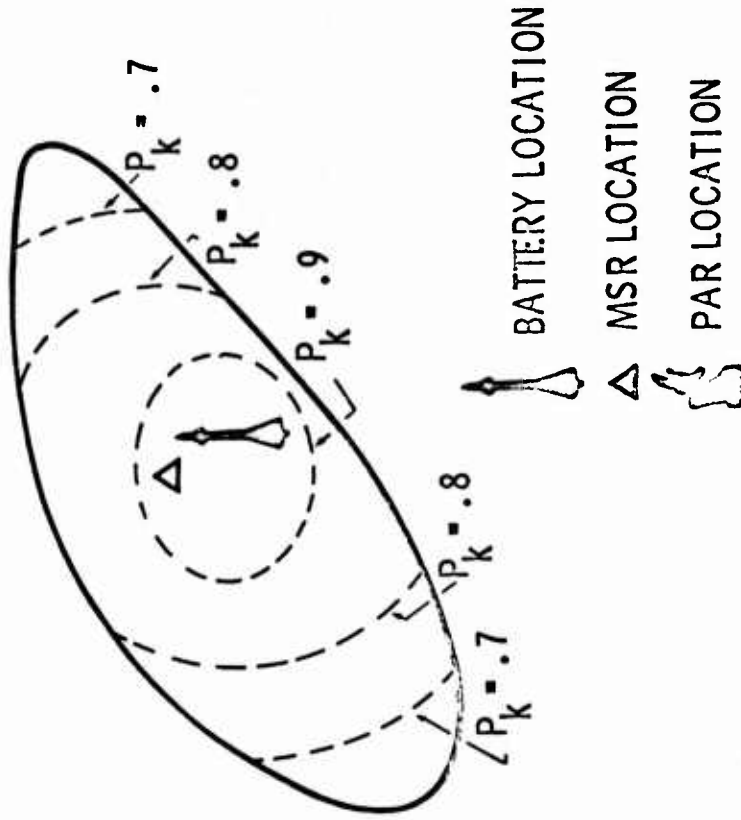
CHART 15



MAXIMUM COVERAGE CONTOUR

# EXO - 1 OUTPUT

CHART 16



3

PROBABILITY OF KILL ISOQUANTS SUPERIMPOSED OVER THE  
MAXIMUM COVERAGE CONTOUR

variance co-variance technique eliminated this problem. EXO-1 on the GE-635 computer requires about 1 minute of central processor time to produce the probability of kill at one point. A similar program for endo-atmospheric interceptions generating 30 Monte Carlo iterations requires 10 minutes per point. We estimate that the effort required to build a computer model for a new application of this technique could be as much as two man years.

We are interested in any solution method that may be used as an alternate to Monte Carlo which will produce reliable information in less computer time.

#### EXO-1 REFERENCES

A Computer Program to Investigate Exoatmospheric Engagements of Interceptors and Re-entry Vehicles, Washington, D. C.: Air Force Assistant Chief of Staff, Studies and Analysis, Strategic Unmanned, 1971.

Gabler, R. T. and S. J. Belcher, A Computing Program for Tracking Error Analysis of Keplerian Trajectories, The RAND Corporation, RM-4740-PR, October 1965.

Kendall, W. B., The Probability Distribution of Anti-Missile Missile Miss Distance Due to Observation and Guidance Noise, The RAND Corporation, RM-3505-ARPA, February 1963.

The EXO-1 briefing was given by:

Barbieri, W. A., The RAND Corporation, 2100 M Street, N. W. Washington, D. C. 20037

Roberson, M. L. Lt. Col., Asst. Chief of Staff, Studies and Analysis, Hq USAF, Washington, D. C. 20330

Van Nostrand, C., Capt., Asst. Chief of Staff, Studies and Analysis, Hq USAF, Washington, D. C. 20330

THE WIND TUNNEL FREE FLIGHT TESTING TECHNIQUE

A. S. Platou  
 Research Aerospace Engineer  
 U. S. Army Ballistic Research Laboratories  
 Aberdeen Proving Ground, Maryland

Symbols

a Speed of sound, ft/sec  
 d Model reference diameter, ft  
 $I_x$  Axial moment of inertia of model  
 $I_y$  Transverse moment of inertia of model  
 $k_x$  Axial radius of gyration of model =  $\sqrt{I_x/md^2}$   
 $k_y$  Transverse radius of gyration of model =  $\sqrt{I_y/md^2}$   
 L Total length of flight, ft  
 m Mass of model, slugs or grams  
 Ma Mach No.  $U_\infty/a$   
 H, M, P, } See section entitled, The Yaw Reduction  
 T, G }  
 p Spin rate of model radians per second or revolutions per minute  
 $P_0$  Tunnel stagnation pressure, psia  
 q Free stream dynamic pressure =  $\frac{1}{2} \rho U_\infty^2$  or missile angular rotation  
 S Dimensionless distance along flight path =  $\frac{1}{L} \int_0^t U_\infty dt$   
 s Area =  $\pi d^2/4$  feet<sup>2</sup> or inches<sup>2</sup>  
 t Time, seconds or milliseconds  
 $T_0$  Tunnel stagnation temperature, °R  
 T Tunnel static temperature, °R  
 $U_\infty$  Free stream tunnel air velocity  
 V Model velocity in X direction  
 X } Coordinates axes, X is along tunnel axis, positive upstream; Y is horizontal and perpendicular to X; Z is vertical and perpendicular to X, positive down; (right hand rule applies to directions and rotations)  
 Y }  
 Z }  
 $\rho$  Air density, slugs per ft<sup>3</sup>  
 $\alpha$  Angle of attack in X Z plane  
 $\beta$  Angle of attack in X Y plane  
 $\delta$  Total angle of attack or yaw =  $\beta + i\alpha$   
 $\xi$  sin  $\delta$   
 $\bar{\delta}^2$  Mean square yaw =  $\frac{1}{L} \int_{-L/2}^{+L/2} \delta^2 ds$   
 $v$   $pd/2U_\infty$   
 $\phi$  Roll angle  
 $\xi$   $e^{i\phi}$   
 $\mu$  Viscosity  
 $C_D$  Drag coefficient =  $D/q s$   
 $C_{D0}$  Drag at zero lift  
 $C_{L\alpha}$  Lift curve slope =  $L/q s \sin \delta$   
 $C_{N\alpha}$  Normal force coefficient =  $N/q s \sin \delta$   
 $C_{N_{P\alpha}}$  Magnus force coefficient =  $M.F./q s v \sin \delta$   
 $C_{M\alpha}$  Pitching moment coefficient =  $M/q s d \sin \delta$

$C_{M_{P\alpha}}$  Magnus moment coefficient =  $M_p/q s d v \sin \delta$   
 $C_{L_P} v + C_{L_\delta} \delta$  Rolling moment coefficient =  $l/q s d$   
 $C_{M_q} (qd/2V) + C_{M_{\dot{\alpha}}} (\dot{\alpha}d/2V)$  = Damping moment coefficient =  $M/q s d$   
 $Re = \rho U_\infty d/\mu$   
 C.G. Location of C.G. from model base, % of length  
 $C_{(i)}^* = \frac{\rho s d}{2m} C_{(i)}$

Abstract

The free flight wind tunnel technique has been used successfully to obtain aerodynamic coefficients on a variety of configurations with and without spin. High mass to moment of inertia ratio models are electroformed as .001 inch thick nickel shells with tungsten cores at the center of gravity. This permits wide C.G. variations and provides models which can withstand the high speed, high temperature flows. The model launcher is based on the principle of an inverted pea shooter. A  $\frac{1}{4}$ -inch diameter tube is inserted into the aft portion of the model through the model base, and compressed air acting through the tube on the tungsten core propels the model forward. Spin can be imparted to the model with an air turbine just prior to launch. Cones have been launched with success up to 10,000 rpm, while high fineness ratio models have been launched at 40,000 rpm with moderate success.

Introduction

In recent years a free flight wind tunnel testing technique has been developed which is advantageous for obtaining aerodynamic data on certain configurations<sup>1</sup>. The largest and most desirable advantage over other wind tunnel testing techniques is the sting interference-free data, and relative ease of testing long bodies. Advantages over range testing techniques are the large number of cycles per flight, the large number of photographic stations (up to 500) per flight, and the low model accelerations necessary to launch the model. The effective range  $(U_\infty + V)t$  for this type of testing is 200 to 400 feet, depending on the tunnel Mach number and the time of flight. The most popular use of this free flight technique has been for dynamic stability data; however, it also appears attractive for Magnus and roll data on spinning configurations.

Magnus data have been obtained in wind tunnels on some ballistic shapes using sting supported models and strain gage balances<sup>2</sup>; however, in some cases sting interference is indicated. In hypersonic tunnels the complication of cooling both bearings and balances is a further deterrent to standard type Magnus testing. Roll damping measurements in a wind tunnel have depended on low friction air bearings, and here again free flight

This paper appeared in the Proceedings of the AIAA Third Aerodynamic Testing Conference held in San Francisco, California on 8-10 April 1968. We would like to thank the American Institute of Aeronautics and Astronautics for permission to photographically reproduce this article.

ceeding page blank



testing avoids this complication. It is the object of this paper to present the BRL (Ballistic Research Laboratories) plans for free flight testing of spinning and nonspinning ballistic shapes in our hypersonic wind tunnel.

It is possible to obtain all of these measurements (dynamic stability, Magnus force and moment, and roll damping) from one flight, for each of these measurements requires a model having low moments of inertia and small mass. Dynamic stability measurements require low transverse moments of inertia; Magnus measurements require low mass, and low transverse moments of inertia and roll damping require low axial moments of inertia. Models of this type are built as thin, lightweight shells containing a heavy metal core close to the center of gravity of the model. The only change between a nonspinning dynamic stability model and a spinning one to obtain all three pieces of data is the addition of spin rate counting pins in the base of the model or darkening of one side of the model.

#### The Optical System

The optical system is used to obtain a series of consecutive pictures of the model in flight so that the trajectory including the angle of attack history of the model can be examined. It is a dual path orthogonal system so that three dimensional coordinates of the model are recorded during the flight. The orthogonal system makes it possible to obtain data on all types of ballistic shapes including spinning and nonspinning bodies having both planar and nonplanar motions.

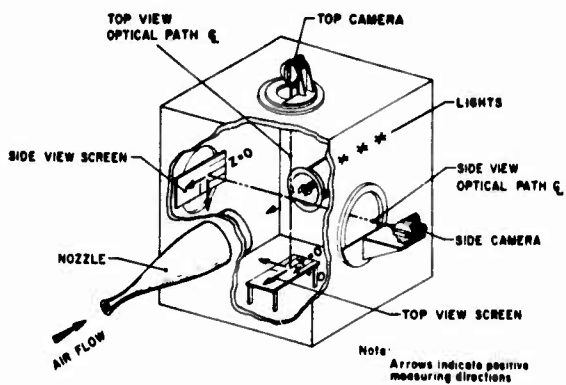


Figure 1. Instrumentation for Free Flight Testing in the BRL Hypersonic Tunnel

The optical system requires two 35mm Fastax cameras mounted (Figure 1) with the side view camera photographing the vertical motion of the model and the top view camera photographing the horizontal motion. Both views see a silhouette of the model against lighted, white screens containing the measuring reference lines (Figure 1). Optical alignment of the cameras is made before each test period and is such that the model coordinates in each individual frame can be read to .03 inch. Each of the screens is illuminated by three 1000-W photographic Sun Gun lights located approximately 4 feet from the screens. For correct exposure of the film (Tri-X-Negative), the camera lens is set at  $f/8$ , using a 2 inch lens. The cameras are

approximately 60 inches from the tunnel centerline. Identical one-millisecond timing marks, which are coded from a prescribed zero time, are placed on each film during the flight. In this manner the time-motion histories from each film can be linked together so that the three dimensional time-motion history of the model can be obtained.

#### The Models

The construction of the models is based on the fact that they must be lightweight, have low axial and transverse moments of inertia, and have a high mass to moment of inertia ratio. These characteristics are obtained by forming the model as a lightweight, thin-skin shell, with a heavy metal core placed around the model center of gravity (Figure 2). These models, however, must withstand the rigors of exposure to hypersonic flow for a short time, so that certain precautions must be taken to insure this. First, it is estimated that portions of the model may reach  $800^{\circ}\text{F}$  during the flight; and second, the pressure distributions over the model may be quite severe as the model emerges from the launch chamber into the main tunnel flow.



Figure 2. Free Flight Model Design

A survey of the field shows that there are a number of materials which will withstand the expected temperatures; however, there are only a few of these which are sufficiently lightweight to keep the moments of inertia low. The most numerous high temperature lightweight materials found are a series of porous materials made by Emerson and Cuming Company. These materials have specific gravities from .24 to .75 and will withstand temperatures of at least  $800^{\circ}\text{F}$ , and in some cases  $3000^{\circ}\text{F}$ . They are quite fragile though, having flexural strengths between 500 and 1000 psi, and therefore, must be reinforced in order to withstand the pressure distributions.

The first models used in our experiments were made of thin walled steel or aluminum tubing machined down to 3 mil walls with the nose and tail being made of the high temperature porous material. Later models have been made by electroforming nickel on an aluminum mandrel to a thickness of 1 or 2 mils and then eroding the mandrel leaving the nickel shell. It is believed that thin skinned electroformed models can be made for any configuration for which an aluminum mandrel can be machined. To date, electroformed models of  $10^{\circ}$  half angle cones, 20 and 30 caliber cone cylinder flares and a finned configuration (Figure 3) have been made.

The first models used lead for the metal core, while later models have used tungsten. Lead is easier to machine but has a lower mass to moment of inertia ratio than tungsten for the same geometric configuration. Lead also lowers the maximum allowable model temperature, thereby increasing the risk of losing a model during the prelaunch period in the hypersonic tunnel. To secure the model to the



launcher, one end of a 5 mil copper wire is imbedded into the metal core.

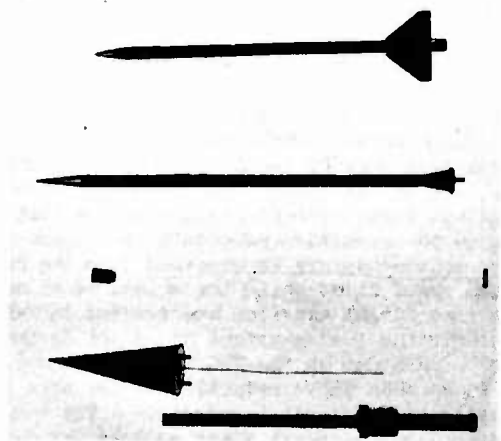


Figure 3. An Example of the Models Used with the Free Flight Launcher

#### The Launcher

After several attempts at launching models into the hypersonic test section were made, it became apparent that a number of precautions must be taken in order that successful launches occur a large percentage of the time. The launcher and model must be protected from the hot high speed tunnel air by inclosing it within an insulated air cooled chamber. Doors on the front of the chamber must open and close automatically during the launch cycle. Symmetric flow conditions in the door region is required to hold down velocity attenuation and lateral jump. Some type of guidance to maintain model attitude during the acceleration phase is necessary. For spinning models a turbine drive must be incorporated.

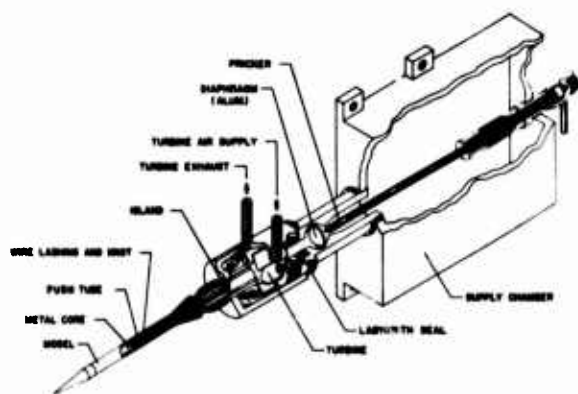


Figure 4. Free Flight Launcher

A launcher design (Figure 4), which incorporates the above features, has been developed and used by the Ballistic Research Laboratories for several months. The launcher is housed in an insulated, air-cooled chamber which keeps the ambient temperature below 200°F, while the tunnel stagnation

temperature is 1300°F. Doors on the front of the chamber automatically open before the launch and close after the launch so that the chamber temperatures remain low. The launcher is mechanically simple and contains very few moving parts. It is unique in that no part other than the model is accelerated forward during the launch, and the model is guided during the full period of acceleration by an internal push tube. The launcher will accept any shape model from  $\frac{1}{4}$ -inch diameter up to at least 1 inch diameter, and up to  $7\frac{1}{2}$  inches long; the only requirement on these models being that they will accept a  $\frac{3}{16}$ -inch minimum diameter cylindrical push tube mounted through the model base, and resting against the metal sleeve which is common to all of these models. Each model requires a pretest assembly and wire lashing to the cylindrical tube, but can be readily installed on the launcher just prior to launch. The launcher is a single-shot launcher, but means of reloading it while the tunnel is running are being considered now.

The principle of operation of the launcher is that of a shock tube. A normal shock is generated by breaking, with a pricker, an aluminum foil diaphragm which is mounted at one end of a small diameter, hollow cylinder (Figure 4) called the push tube. The strength of the normal shock and, in turn, the model launch velocity is governed by the initial pressure in the supply chamber behind the diaphragm. The normal shock travels the length of the push tube and reflects from the model metal core which is lashed to the other end of the push tube. The lashing is accomplished by imbedding a 5-mil diameter knotted copper wire in the metal core and soldering the other end of the wire to an island mounted in the center of the push tube. When the aluminum diaphragm is broken, the increased pressure on the metal core base breaks the wire at the knot, and the model is launched. The exact aerodynamic action which takes place during the several milliseconds required for the model to leave the push tube is quite complicated; however, it is known that the shock is traveling at much higher velocities than the model, so that several shock reflections must take place in the push tube during the launch time. The strength and number of shock reflections are a function of the cylinder and supply chamber configuration, and therefore, each launcher configuration must be calibrated to determine the correct pressures for launching models at the correct velocity. The launcher will launch models with speeds from 20 feet per second to 75 feet per second.

To adapt the nonspinning launcher to a spinning launcher requires that a custom-built air turbine be inserted between the aluminum diaphragm and the push tube. The air turbine is designed the same as a dental air drill, except that it is approximately four times larger (Figure 4). The enlargement is necessary so that the main shaft can be hollow and form the first portion of the push tube. Also, the enlarged turbine blades increase the available torque and lower the maximum allowable spin rate. Since the dental drills have been developed for spin rates of 400,000 rpm, the enlarged version will still easily spin up to 50,000 rpm. It is necessary to place a labyrinth seal between the spinning turbine shaft and the stationary diaphragm in order to prevent large attenuation of the launch pressure after the diaphragm is broken. The remainder of the push tube, including the model, rotates with the turbine, so that once the supply air passes through the diaphragm it is subject to a rotating wall

boundary condition. Monitoring of the prelaunch spin rate will be accomplished using a photoelectric pick-up, while the actual spin rate during flight will be read from the high speed motion pictures.

The launching procedure for both the nonspinning and the spinning launcher is essentially the same. In both cases the model is placed on the launcher, the tunnel is started and brought up to the desired temperature and pressure. Just prior to launch the predicted supply pressure is set, and in the case of a spinning launch the model is brought up to the desired spin rate. At this point the actual launch is actuated by a six-channel timer which controls the photographic lights, the launch pricker, the cameras, the timing mark coder and the launcher doors. The launch is triggered by the pricker rupturing the diaphragm; however, all of these events must be triggered accurately so that the flight is photographed on the high speed portion of the film.

In launching models into the tunnel airstream we wish to obtain flights in which the model will oscillate through the maximum number of cycles. The number of cycles is given by

$$N = \frac{1}{\pi} \left( \frac{C_{m\alpha}}{C_D} \times \frac{d}{2} \times \frac{m}{I} \times S \right)^{\frac{1}{2}}$$

$C_{m\alpha}$  and  $C_D$  are fixed by the configuration being tested,  $d$  and  $m/I$  are maximized by making the model as large as possible and by concentrating as much mass as possible close to the C.G.,  $S$  is maximized by finding the longest path or the longest flight time. The longest flight path is obtained by adjusting the launch velocity so that the model reaches the upper portion of the test rhombus and returns downstream before gravity has time to pull the model out of the test rhombus. The launch velocity for constant deceleration is  $V_0 = \sqrt{2S D/m}$  and the time of flight is  $t = \sqrt{\frac{2S}{D/m}}$ .

For the  $10^\circ$  half angle cone models the  $D/m$  values range from 150 to 300 ft/sec<sup>2</sup> for the cases tested. The launch velocities are 40 to 50 ft/sec and the flight times are such that the models would not quite fall through the test rhombus due to gravity. For the high fineness ratio models the  $D/m$  values range from 50 to 100 ft/sec<sup>2</sup>, and the launch velocities are in the range 15 to 25 ft/sec. This plus the relatively long push tube, especially for the 30 caliber models, decreased the launch pressure such that breaking of the lash wire is marginal.

To overcome this, the launching system is being modified so that the launch occurs in two steps. First, the present launcher is given a vertical and a downstream horizontal velocity. Second, the present launcher is activated so that the resulting model launch motion is forward and upward with the model at low angle of attack.

The model will be launched from the lower downstream portion of the test rhombus such that the motion will carry the model to the top forward portion of the rhombus before gravity and the drag pull the model down towards the launch point.

The reason for the initial horizontal downstream

velocity imparted by the motion of the launcher is to increase the launch velocity,  $V_0$ , thus increasing the required launch pressures insuring breakage of the lashing wire.

#### Data Reduction

Data reduction of each model launched in the wind tunnel is accomplished by fitting the time motion history of the flight to the equations of motion for a ballistic missile. The reduction is basically the same reduction used by the BRL aerodynamic ranges<sup>3,4</sup>. However, due to differences in methods of recording the data and other characteristics of the test, some changes in the reduction procedure are necessary. The time motion history is obtained from the orthogonal high speed films while the equations of motion for the free flight missiles are derived in reference 4. Fitting the equations of motion to the data has been programmed on the BRL high speed computer, BRLESQUE, so that rapid reduction of the data to aerodynamic coefficients is possible. The reduction is separated into several steps so that various aspects of the flight can be considered separately.

#### Tunnel Aerodynamic Conditions

This portion of the program concerns the reduction of the tunnel operating conditions to parameters which can be used for reduction of the model geometric data to aerodynamic coefficients. The tunnel aerodynamic reduction uses the compressible fluid flow equations for a de Laval nozzle which are outlined and tabularized in reference 5. Quantities which are either set during the tunnel test or computed from the tunnel conditions are: Mach number, stagnation pressure, temperature, dynamic pressure, test section air velocity and Reynolds number.

#### Model Geometric Reduction

The geometric data reduction consists of the evaluation of the motion of the model to define the space coordinates of the model center of gravity and the variation of angle of attack ( $\beta + i\alpha$ ) as a function of time. The  $x$ ,  $y$  and  $z$  coordinates of the model nose and base are read from each of the Fastex films (Figure 5) using an optical film comparator. These data are used to determine the lateral motion of the model center of gravity and the angle of attack motion of the model. The film speed or the frame time is obtained from the 1000 cycle timing marks which have been placed on the side of each film during the operation of the Fastex cameras. The time histories obtained in the geometric data reduction define the complex yaw ( $\beta + i\alpha$ ) and complex swerve motion ( $y + iz$ ) of the model in flight and, also, define the velocity and spin variation with time. Examples of this information are presented in Figures 7, 8, 9 and 10. The data accuracy is governed by the resolution of the Fastex cameras, which in this case, correspond to coordinate accuracies of 0.03 inch. The results of the model geometric data reduction are used as input information for the remainder of the data reduction.

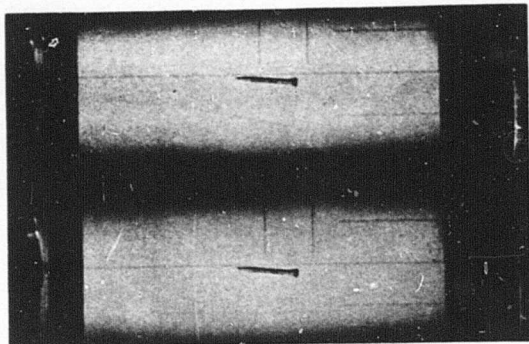


Figure 5. High Speed Photographs Taken for Measuring the Model Coordinates

### Drag and Roll Reduction

The drag and roll reduction are two separate reductions; however, since the procedures are the same and their reductions are independent of the remaining reductions, they can be discussed as a group.

The average drag force acting on the model during the flight can be obtained by assuming constant deceleration during the flight. Constant deceleration is not strictly true for it will vary with angle of attack; however, this variation is within the accuracy of the data. Computing the deceleration (least square fit) from the velocity time curve (Figure 6) and using Newtons equation

$$D = -m \ddot{X}$$

$$C_D = \frac{D}{qS}$$

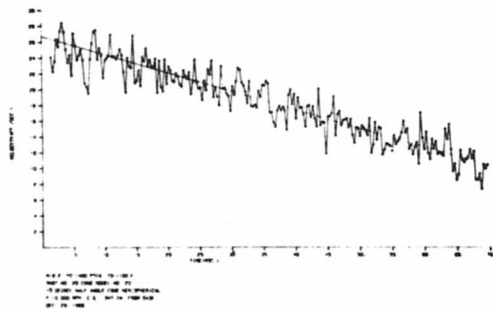


Figure 6. Velocity-Time Curve

Along with the average drag computation it is necessary to compute the mean square yaw so that the drag curve for the configuration can be obtained. The mean square yaw is obtained from:

$$\bar{\delta}^2 = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} \delta^2 ds$$

and since  $C_D = C_{D_0} + C_{D_\delta} \bar{\delta}^2$

we can obtain  $C_{D_0}$  once, two or more flights of the same configuration have been made at different mean square yaw values (Figures 10 and 11).

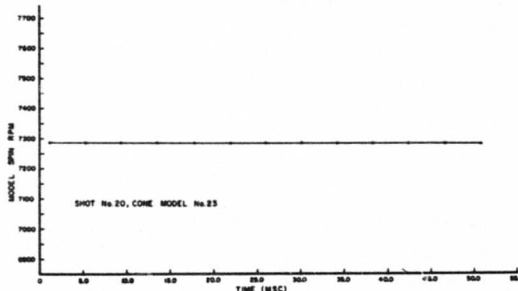


Figure 7. Spin-Deceleration Curve

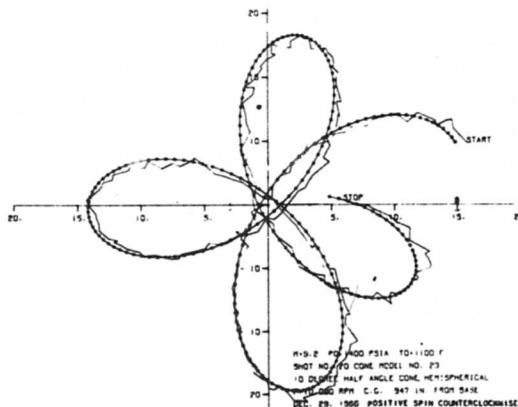


Figure 8. Complex Yaw Plot

The roll reduction depends on measuring the roll deceleration of the model during the flight. This is done by measuring the rate of rotation of a white strip on the model or the counting pins (Figure 7). Again, assuming constant deceleration, the rolling moment can be obtained from

$$l = I_x \dot{p}$$

$$C_{l_p} v + C_{l_\delta} \delta = \frac{l}{q S b}$$

$C_{l_\delta}$  is the rolling moment due to fin cant and needs

only to be considered for finned configurations.

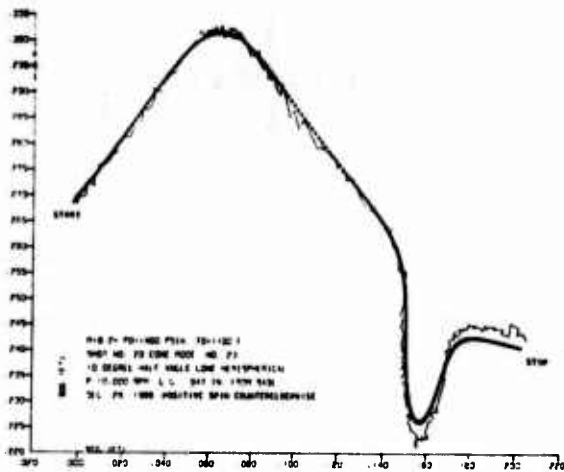


Figure 9. Swerve Plot

#### The Yaw Reduction

The yaw reduction is designed to obtain from the model motion the pitching moment, the damping moment, and the Magnus moment coefficients. This is accomplished by fitting the complex yawing motion of the model (Figure 8) to the equations of motion of the model.

$$\ddot{\xi}'' + (H - i P) \dot{\xi}' - (M + i P T) \xi = G$$

$$\text{where } H = \left[ C_{L\alpha}^* - C_D^* - k_y^{-2} (C_{Mq}^* + C_{M\dot{\alpha}}^*) \right]$$

$$M = \left[ k_y^{-2} C_{M\alpha}^* \right]$$

$$T = \left[ C_{L\alpha}^* + k_x^{-2} C_{M_{P\alpha}}^* \right]$$

$$P = \frac{I_x}{I_y} v$$

$$G = \frac{P g d}{U_\infty^2}$$

The above is equation 6.8 of reference 4 where its' derivation is shown. It is too lengthy to present here, and it will have to suffice that it is similar to the more familiar planar equation of motion

$$I_y \ddot{\alpha} + \mu \dot{\alpha} + M = 0$$

Equation 6.8 takes into account the nonplanar motion of the model and for ease of reduction the derivatives have been taken with respect to distance rather than with respect to time.

The fit of the equation of motion to the complex yaw data is illustrated in Figure 8.

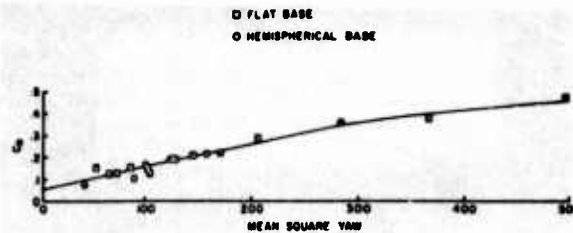


Figure 10. The Drag of the 10° Half Angle Cone  
d = .9"

#### The Swerve Reduction

The swerve reduction analysis studies the lateral motion of the model center of gravity and determines the lift, Magnus and damping forces acting on the model. The reduction involves fitting the equation of motion of the center of gravity lateral coordinates to the actual lateral motion of center of gravity. This equation of motion is equation 9.8 of reference 4.

$$\frac{y_e + iz_e}{l} e = B_0 + B_1 s - \left[ C_{L\alpha}^* + i \frac{Pl}{V} C_{N_{r\alpha}} \right] I_1$$

$$- (C_{N_q}^* + C_{N_{\dot{\alpha}}}^*) I_3 + (C_{Y_0}^* + i C_{Z_0}^*) I_4$$

$$+ i \left( \frac{gd}{U_\infty^2} \right) \frac{s^2}{2}$$

$$I_1 = \int_0^s \int_0^s \xi ds_1 ds_2$$

$$I_3 = \int_0^s \xi ds_1$$

$$I_4 = \int_0^s \int_0^s e^{i\phi} ds_1 ds_2$$

$B_0, B_1$  are complex constants.

The Coriolis force term has been omitted because in the wind tunnel free flight technique the motion of the model with respect to the earth is negligible. An example of the swerve fit to the actual motion is shown in Figure 9. Again, for further explanation of this reduction procedure, the reader is referred to reference 4.

#### Single Plane Reduction

Recently we have started modifying one of our supersonic tunnels  $M = 1.25$  to  $5.0$  so as to increase our free flight capabilities to the complete supersonic speed range. In order to obtain orthogonal views, it would entail considerable effort to locate a viewing window in the ceiling of the test section. In an attempt to eliminate this complica-



tion we are investigating the possibility of using a single view reduction system in which only the projected coordinates of the model in the X-Z plane would be recorded. By solving only the imaginary part of the epicyclic equations of motion

$$\xi'' + (H - i P) \xi' - (M + i PT) \xi = G$$

the aerodynamic coefficients will be obtained. The ease and accuracy of this solution is now being compared for us with the full viewing system method by the University of Notre Dame.

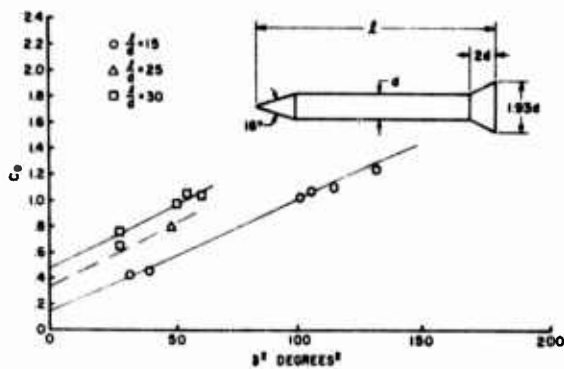


Figure 11. Drag of Cone Cylinder Flare Configuration

Model Flights and Aerodynamic Data

A number of flights of  $10^\circ$  half angle cone models and high fineness ratio cone cylinder flare models have been reduced and analyzed. Most of the flights have been made at  $M = 9.2$  and have included variable Reynolds number, variable center of gravity and variable spin. The cone models included both flat base and hemispherical base configurations in order to determine the variation in the  $C_{p6}$  pitch damping moment due to the base configuration, and the cone cylinder flare models included length to diameter ratios of 15 to 30. Launching the cones proved to be quite successful. However, some difficulties were experienced in launching the high fineness ratio bodies. The longer launching distance and in the case of spinning models, difficulty of controlling the dynamic balance proved to be critical. Long models having shorter launch distances and better balance are now being fabricated.

The aerodynamic data obtained on these models is shown in Figures 10, 11, 12 and 13. The drag coefficients are linear with the mean square yaw for the cone cylinder flare configuration and up to  $15^\circ$  for the  $10^\circ$  cone. The pitching moment coefficients for all of the models are constant over the Reynolds number range tested (Figures 12 and 13), while the pitch damping coefficients for all models vary with Reynolds number. The flat base cone pitch damping increases with Reynolds number, while the hemispherical base cone data indicate decreased damping with Reynolds number, but a large increase in damping over the flat base cone. The data also indicate that the cone base configuration may not influence the pitch damping at higher Reynolds numbers. These data also agree with data obtained at AEDC<sup>7,8</sup>. The cone cylinder flare pitch damping

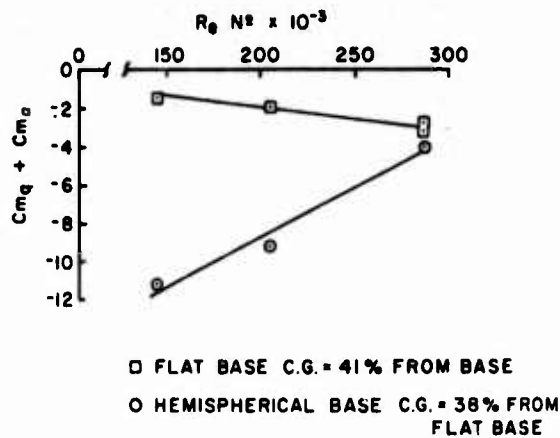
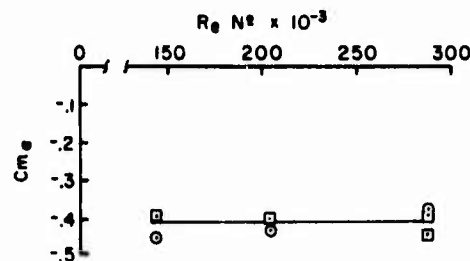


Figure 12. Reynolds Number Variation of Pitching and Damping Moments on the  $10^\circ$  Half Angle Cone  $M = 9.2$



□ C.G.=7.2 CAL FROM BASE  
○ C.G.=6.6 CAL FROM BASE

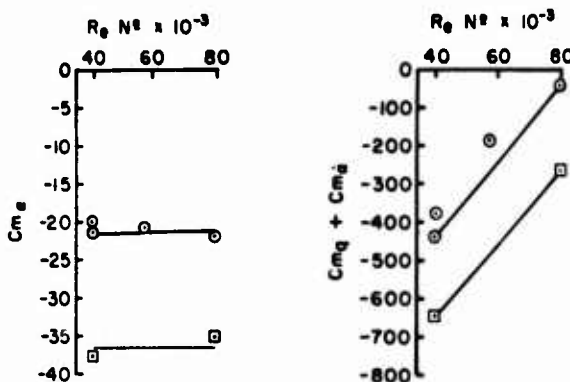


Figure 13. Reynolds Number Variation of Pitching and Damping Moments on the 15 Caliber Cone Cylinder Flare Configuration  $M = 9.2$

coefficients decrease with Reynolds number and the indication is that the pitch damping will become positive at slightly higher Reynolds numbers. Magnus and roll derivatives have also been obtained but are so far too few to be presented here.

While analyzing the epicyclic motion of these models, it became apparent that the epicyclic arm rates of rotation ranged from 5 to 60 revolutions per second. These rates are in the same range as a considerable amount of the turbulence of most wind tunnels which indicates the possibility of turbulence influence on the epicyclic motion and the computed aerodynamic coefficients. An examination of the pitching moments and pitch damping moments of the flights obtained to date indicate that the tunnel turbulence can influence these coefficients provided that the arm rates are below 20 cycles. Further tests where the tunnel turbulence level is lower are planned for the future.

#### Conclusions

The analysis of the flights made to date are very encouraging. The free flight technique permits the evaluation of damping, Magnus, and roll derivatives on all types of configurations without the presence of a sting. Flights in the BRL wind tunnel have included spinning models with spin rates up to 40,000 rpm. So far, configurations have been limited to  $10^\circ$  half angle cones, 1 inch base diameter, and 15 to 30 caliber cone cylinder flares,  $\frac{1}{4}$  inch body diameter. Low model weights and low moments of inertia have been achieved by electroforming the model shells from .001 inch thick nickel. The model fabrication technique assures us of being able to electroform any model shape for which an electroforming mandrel can be made. The reduction procedure permits computation of all of the aerodynamic coefficients including the damping, Magnus and roll derivatives.

#### References

1. Dayman, Bain, Jr., "Free Flight Testing in High Speed Wind Tunnels," AGARD 113 (May 1955).
2. Platou, A. S., "Magnus Characteristics of Finned and Nonfinned Projectiles," AIAA J. 1, 3 (January 1965).
3. Murphy, C. H., "Data Reduction for the Free Flight Spark Ranges," BRL Report No. 900 (February 1954).
4. Murphy, C. H., "Free Flight Motion of Symmetric Missiles," BRL Report No. 1216 (July 1963).
5. Ames Research Staff, "Equations, Tables, and Charts for Compressible Flow," National Advisory Committee for Aeronautics Report 1135 (1953).
6. Deitrick, R. E., "Effect of a Hemispherical Base on the Aerodynamic Characteristics of Shell," BRL Memorandum Report No. 947 (November 1955).
7. Uselton, B. L. and Gregson, J. H., "Dynamic Stability Characteristics of a  $10^\circ$  Cone at Large Amplitudes of Oscillation at Mach Number 10," Arnold Engineering Development Center TR-66-49 (March 1966).
8. Hodapp, A. E., Jr., and Burt, G. E., "Dynamic Stability Characteristics of a  $10^\circ$  Cone at Mach Number 10," Arnold Engineering Development Center TDR-64-98 (May 1964).

SOME ASPECTS OF REGULARIZATION AND APPROXIMATION OF  
SOLUTIONS OF ILL-POSED OPERATOR EQUATIONS<sup>1</sup>

M. Z. Nashed<sup>2</sup>  
Mathematics Research Center  
The University of Wisconsin, Madison

**ABSTRACT.** In this paper we attempt a classification of the various approaches that have been proposed for the investigation (solutions and approximations) of ill-posed problems. The classification is only descriptive; we do not go into technical details. The author hopes to prepare a more extensive survey of regularization methods for operator equations where the technical aspects will be considered. The bibliography, although extensive, is not intended to be complete.

**1. ILL-POSED PROBLEMS: REMARKS AND EXAMPLES.** The notion of a well-posed (correct properly-posed) problem introduced by Hadamard at the beginning of this century plays an important role in the theory and numerical approximation of various operator equations arising from problems of mathematical physics, engineering, and analysis.

Let  $X$  and  $Y$  be two metric spaces, and let  $A$  be an operator on  $X$  into  $Y$ . The operator equations  $Ax = y$  is said to be well-posed by Hadamard's definition if the following conditions hold:

- (i) there exists a solution of the equation for all  $y \in Y$ ;
- (ii) the solution is unique in the space  $X$ ;
- (iii) the solution depends continuously on the right-hand side  $y$ .

If any of these conditions are not satisfied, then the operator equation is said to be ill-posed (incorrect, improperly posed).

The first requirement of well-posedness means that the problem is not overdetermined, and superfluous conditions are not imposed. The third requirement of well-posedness is important in problems of mathematical physics and natural phenomena since the data  $y$  is obtained from measurements made with instruments and is therefore known only approximately. The requirement guarantees that a small error in  $y$  cannot produce a big change in the solution  $x$ .

---

<sup>1</sup>Sponsored by the United States Army under Contract No.: DA-31-124-ARO-D-462, while the author was a visiting member of the Mathematics Research Center.

<sup>2</sup>Present address: School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332.



It should be emphasized that the notion of a well-posed problem for a given operator depends on the spaces considered. Thus the operator equation  $Ax=y$  may be well-posed relative to the spaces  $(X,Y)$ , while being ill-posed relative to the spaces  $(X',Y')$ , i.e. when  $A$  is considered as a map on  $X'$  into  $Y'$ , so that the data is drawn from  $Y'$ , and the notion of continuity is that induced by the metrics on  $X'$  and  $Y'$ . Mathematically this is clear since a mapping may be continuous relative to one topology while failing to be so in another. Physically, this gives a clue as to the type of measurements that are meaningful, and provides a framework where small errors in measurements are tolerable.

Hadamard gave an example of a problem (the Cauchy problem for Laplace's equation), which has now become classical, which is ill-posed in any of the usual function spaces (the space of continuously differentiable functions,  $L_p$  spaces, Sobolev spaces  $W_p^b$  spaces  $H_p^q$  of analytic functions on the unit disc., etc.). On the basis of this example, Hadamard concluded that this problem, and other problems exhibiting a similar dependence of the solution on the data, do not correspond to any real formulations, i.e. they are not problems of mathematical physics. In other words, there is something wrong with the mathematical model and not with the physical problems which it portrays. It was discovered later, however, that Hadamard's conclusion was erroneous and that many situations in physics and in the study of natural phenomena, as well as in some areas of analysis, lead to problems which are ill-posed. We mention next some examples of ill-posed problems which appear frequently in the literature. Some of these were cited by the Soviet Academician A.N. Tikhonov in his invited address at the International Congress of Mathematicians in Moscow in 1966.

#### Some Classes of Ill-Posed Problems in Analysis

1. Determination of a uniform approximation to the derivative  $u'$  under approximate data in the metric of the space of continuous functions.
2. Determination of the sum of a Fourier series at a given point in terms of arbitrary values in the space  $\ell_2$  for the Fourier coefficients.
3. Uniform approximations of solutions of integral equations of the first kind and other problems leading to them (analytic continuation, conformal mapping, operational calculus in the real domain) under perturbation of the data in the metric of  $L_2$  (the space of square-integrable functions).
4. Numerical inversion of Laplace transforms.
5. The problem of determining the input to a system when we know the impulse response and the output.
6. A wide range of unstable problems of optimization (unstable problems of optimal control and filtering, linear and dynamic programming).

7. Linear problems on the spectrum under the usual supplementary conditions determining a unique solution.

8. Ill-conditioned algebraic systems, Fredholm alternative problems with nonuniqueness, etc.

9. Some classes of "inverse problems". A typical situation here is of finding  $Z$ , which is not accessible to direct measurement, from a physically determined manifestation  $u$  of  $Z$ , where  $u = AZ$ , and  $A$  is a completely continuous operator. Then the inverse of  $A$ , is not continuous.

Note that there is a considerable overlap among the classes mentioned above, and that from the operator-theoretic point of view many of these classes share the same characteristic.

#### Some Areas of Real Phenomena Which Lead to Ill-Posed Problems

(1) Some classical problems of mathematical physics: for example the Cauchy problem for nonnegative time, the nonhyperbolic Cauchy problem for the wave equations, etc. (see John [93], Lavrentiev [111], and Mikhlin [133] for an elucidation of some of these examples). These include problems from potential theory, hydrodynamics, magnetohydrodynamics, etc.

(2) Problems arising from geophysics, atmospheric studies, meteorology, reservoir engineering, seismology, etc. The bibliography contains numerous references on such applications.

(3) Problems of control of a system governed by partial differential equations where the control appears on the boundary, and boundary-value problems with overabundant data on one part of the boundary and insufficient data on the rest of the boundary.

2. SOME APPROACHES TO THE INVESTIGATION OF ILL-POSED PROBLEMS. The lack of a continuous dependence of the solution in an ill-posed problem, or the lack of uniqueness (or even of the existence of a solution in the classical sense) make direct investigation (and particularly approximation) of ill-posed problems difficult. The "regularization" of such problems has been cited by Bellman as one of the important concepts of modern analysis. Intuitively speaking, what this means is to analyze an ill-posed problem via an analysis of a well-posed, or a sequence of well-posed problems, provided this analysis gives a suitable approximation to the given problem. This suggests several approaches which are, generally speaking, based on

- (a) a change of the concept of a solution;
- (b) a change of the spaces in question;
- (c) a change of the operator itself;
- (d) the concept of a regularizer or "regularization operators";
- (e) probability or well-posed stochastic extensions of ill-posed problems.

These intuitive ideas manifest themselves explicitly in several approaches which are available at the present for the investigation and approximation of ill-posed problems;

(1) The earliest approach, due to Tikhonov [188], is based on the assumption that there exists a prior information restricting the class of solutions to a compact set  $U$ . In this case the operator equation is well-posed if the operator  $A$  is considered as a map from  $U$  onto  $A(U)$ . An investigation of the various problems which are amenable to this approach was carried out by Lavrentiev [111], John [92] and others (see the bibliography in [111]).

(2) Another approach is based on changing the notion of a "solution" of a problem (for instance to "quasisolution", Ivanov [81]; a brief exposition of this approach is given in the monographs of Holmes [77] and Lavrentiev [111]; or to the notion of least squares solution of minimal norm in the case the underlying spaces are Hilbert spaces; see for instance Nashed [144], Kammerer and Nashed [95], the latter approach lends itself readily to mathematical programming.

(3) The use of regularizing parametric operators introduced by Tikhonov [188], and further explored by Bakušinskii [11], [12], Nashed and Wahba [200], and others.

(4) Another approach closely related to (3) is that of replacing the operator equation by a stable minimization problem depending on a parameter (see Bellman, Glicksberg and Gress [22], [23], Phillips [162], Ribiere [166] and others).

(5) Stochastic and probabilistic approaches involve questions of measurement of error and disturbances, and provide well-posed extensions of ill-posed problems (see for instance Franklin [58], Lavrentiev [111] and Bakušinskii [13]).

(6) The method of quasirevisibility of Lattes and Lions [107] is based on the idea of modifying the differential or integrodifferential operators arising in boundary-value problems and unstable control and optimization problems, in order to impart regularity. This approach is closely connected with some of the preceding approaches.

(7) A new approach to regularization based on the notion of pseudosolution (least squares solution of minimal norm) in reproducing kernel Hilbert spaces has been proposed and investigated recently by Nashed and Wahba [200], [147]. This approach coincides in philosophy with some of the approaches mentioned above (in the sense that the notion of a solution is changed and the problem is considered in new spaces), even though it differs sharply in technical details. In this approach the geometry of reproducing kernel Hilbert spaces is exploited (in an optimal way), and the results obtained are the best possible in this context.

Some of the above approaches are easily carried out on computing machines thereby providing effective methods for the numerical solution of a wide class of problems. This remark applies in particular to the approaches (2), (3), (4), (6), and (7).

## Bibliography on Regularization Methods and Related Approximations

1. M. S. Agranovich and I. M. Visik, Elliptic Problems with Parameter and General Parabolic Problems, *Uspekhi Mat. Nauk*, XIX (1965), pp. 53-161 (in Russian).
2. A. Albert and R. Sittler, A Method for Computing Least Squares Estimators that keep up with the Data, *SIAM J. Control* 3 (1965), pp. 384-417.
3. L. Amerio, Sui Problemi Di Cauchy e di Dirichlet per l'equazione di Laplace in due Variabili, *Atti della Reale Accad. d'Italia XXI* (1943), pp. 393-425.
4. Ju. T. Antohin, Analytic approach to the problem of equations of the first kind, *Dokl. Akad. Nauk SSSR*, 167(1966), pp. 727-730; *Soviet Math. Dokl.*, 7(1966), pp. 445-448.
5. N. Aronszajn, Theory of reproducing kernels, *Trans. Amer. Math. Soc.*, 68(1950), pp. 337-404.
6. V. Ja. Arsenin and V. V. Ivanov, On optimal regularization, *Dokl. Akad. Nauk SSSR* 182 (1968), pp 9-12, *Soviet Math. Dokl.* 9 (1968), pp. 1067-1070.
7. G. Backus, Inference from inadequate and inaccurate data, *Proc. Nat. Acad. Sci.*, 65 (1970), pp. 1-7.
8. C. Baiocchi, Regularita e unicità della soluzione di una equazione differenziale astratta, *Rend. Sem. Padova*, 35 (1965), pp. 380-417.
9. C. T. H. Baker, L. Fox, D. F. Mayers and K. Wright, Numerical Solution of Fredholm integral equations of the first kind, *Comput. J.*, 7 (1964), pp. 141-148.
10. A. B. Bakušinskiĭ, On a method of solving Fredholm integral equations of the first kind. *Zh. vĭ chisl. Mat. mat. Fiz.* 5,4 (1965), pp. 744-748.
11. A. B. Bakušinskiĭ, Regularization algorithms for linear equations with unbounded operators, *Dokl. Akad. Nauk SSSR*, 133 (1968), pp. 12-14; *Soviet Math. Dokl.*, 9 (1968), pp. 1298-1300.
12. A. B. Baušinskiĭ, A general method of constructing regularizing algorithms for a linear incorrect equation in Hilbert space, *Ž. Vyčisl. Mat. i. Mat. Fiz.*, 7 (1967), pp. 672-677; *U. S. S. R. Comp. Math. Phys.*, 7 (1967), pp. 279-284.
13. A. B. Bakušinskiĭ, The construction of regularizing algorithms in the case of random noise, *Dokl. Akad. Nauk SSSR*, 189 (1969), pp. 231-233; *Soviet Math.*, 10 (1969), pp. 1376-1378.

14. A. V. Balakrishnan, The epsilon technique - a constructive approach to optimal control, in Control Theory and the Calculus of Variations, edited by A. V. Balakrishnan, Academic Press, New York, 1969, pp. 343-383.
15. M. S. Baouendi, These, Paris (1966).
16. R. Bellman, An introduction to matrix analysis, 2nd edition, McGraw-Hill, New York, 1970.
17. R. Bellman, An introduction to the Mathematical Theory of Control Processes, Vol. 1, Academic Press, New York, 1967.
18. R. Bellman, Introduction to the Mathematical Theory of Control Processes, vol. II, Nonlinear Processes, Academic Press, New York, 1971.
19. R. E. Bellman, R. E. Kalaba, and Jo Ann Lockett, Numerical Inversion of the Laplace Transform: Applications to Biology, Economics, Engineering, and Physics, American Elsevier, New York, 1966.
20. R. E. Bellman, R. E. Kalaba, and J. Lockett, Dynamic programming and ill-conditioned linear systems, J. Math. Anal. Appl., 10 (1965), pp. 206-215.
21. R. E. Bellman, R. E. Kalaba, and J. Lockett, Dynamic programming and ill-conditioned linear systems, II, J. Math. Anal. Appl., 15 (1965), pp. 393-400.
22. R. Bellman, I. Glicksberg, and O. Gross, On some variational problems occurring in the theory of dynamic programming, Rend. Circ. Mat. Palermo, 3 (1954), pp. 1-35.
23. R. Bellman, I. Glicksberg, and O. Gross, Some Aspects of the Mathematical Theory of Control Processes, the RAND Corporation, R-313, 1958.
24. R. Bellman, B. Bluss, and R. Roth, Sequential differential approximation and the "Black Box" problem, J. Math. Anal. Appl., 12 (1965), pp. 91-104.
25. A. Bensoussan and P. Kenneth, Sur l'analogie entre les méthodes de régularisation et de pénalisation, to appear.
26. S. Bergman, The Kernel Function and Conformal Mapping, Mathematical Surveys, Vol. 5, Amer. Math. Soc., Providence, R. I. 1950.
27. H. Bialy, Iterative Behandlung linearer Funktionalgleichungen, Arch. Rational Mech. Anal., 4 (1959), pp. 166-176.
28. A. V. Brcadze, Normally solvable elliptic boundary value problems, Dokl. Akad. Nauk SSSR, 164 (1966), pp. 1218-1200; Soviet Math. Dokl., 6 (1966), pp. 1347-1350.

29. A. S. Blagovesčenskii, Some correct problems for the ultra-hyperbolic and wave equations with data given on the characteristic cone, Dokl. Akad. Nauk SSSR, 140 (1961), pp. 990-993; Soviet Math. Dokl. 2 (1961), pp. 1284-1287.
30. Ju. E. Bojarincev and V. G. Vasiliev, Convergence of approximate solutions of an operator equation, Dokl. Math. Nauk SSSR 180 (1968), pp. 1027-1028; Soviet Math. Dokl., 9 (1968), pp. 707-708.
31. C. de Boor and R. E. Lynch, On splines and their minimum properties, J. Math. Mech., 15 (1966), pp. 953-969.
32. V. M. Borok and Ja. I. Zitomirskii, stability of the correctness classes for the cauchy problem for an equation correct in the sense of Petrovskii, Dokl. Math. Nauk SSSR, 167(1966), pp. 495-498; Soviet Math. Dokl. 7 (1966), pp. 399-402.
33. B. M. Burdak, A. Vignoli, and Ju. L. Gaponeko, On a method of regularizing the extremal problem for a continuous convex functional, Dokl. Math. Nauk SSSR, 184 (1969), pp. 12-15; Soviet Math. Dokl., 10 (1969), pp. 4-7.
34. B. M. Burdak, A. Vignoli, and Ju. L. Gaponeko, A Certain method of regularization for a continuous convex functional, Z. Vyčisl. Mat. Fiz., 9 (1969), pp. 1046-1056.
35. T. Butler and A. V. Martin, On a method of Courant for minimizing functionals, J. Math. and Phys., 41 (1962), pp. 291-299.
36. H. S. Cabayan and G. G. Belford, On computing a stable least squares solution to the inverse problem for a planar Newtonian potential, SIAM J. Appl. Math., 20 (1971), pp. 51-61.
37. A. P. Calderon, Uniqueness in the Cauchy problem for partial differential equations, Amer. J. Math., 80 (1958), pp. 16-35.
38. Z. B. Caljuk, On the stability of system of linear integral equations of Volterra type, Dokl. Akad. Nauk SSSR, 150 (1963), pp. 268-270; Soviet Math. Dokl., 4 (1963), pp. 662-644.
39. J. Cea. Approximation Variationnella des problemes aux limites, Ann. Inst. Fourier, 14, 2 (1964), pp. 345-344.
40. B. L. Chalmers, Subspace kernels and minimum problems in Hilbert spaces with kernel function, Pacif. J. Math., 31 (1969), pp. 619-628.
41. P. G. Ciarlet and R. S. Varga, Discrete variational Green's functions II., Numer. Math., 16 (1970), pp. 115-128.



42. R. Courant, Variational Methods for the solution of problems of equilibrium and vibrations, Bull. Amer. Math. Soc., 49 (1943), pp. 1-23.
43. J. Cullum, Numerical differentiation and regularization, SIAM J. Numer. Anal., 8 (1971), pp. 254-265.
44. Ju. M. Danilin, On an approach to minimization problems, Dokl. Akad. Nauk SSSR, 188 (1969), pp. 1221-1222; Soviet Math. Dokl. 10 (1969), pp. 662-663.
45. J. V. Dave, Determination of size distribution of spherical polydispersions using scattered radiation data, Applied Optics, 10 (1971), pp. 2035-2044.
46. P. J. Davis, Interpolation and Approximation, Blaisdell, Waltham, Mass., 1963.
47. R. Denchev, The stability of linear equations on a compact set, Zh. vŷ chisl. Mat. Mat. Fiz. 7,6 (1967), pp. 1367-1370; USSR Comp. Math. Math. Phy. 7 (1967), pp. 201-205.
48. J. B. Diaz and F. T. Metcalf, On iteration procedures for equations of the first kind,  $Ax=y$ , and Picardi's criteria for the existence of a solution, Math. Comp., 24 (1970), pp. 923-935.
49. I. N. Dombrovskaja, The approximate solution of Fredholm integral equations of the first kind, Ural, Gos. Univ. Mat. Zap., 4 (1964), pp. 30-35; Translated as TSR # 652, Mathematics Research Center, The University of Wisconsin.
50. J. Douglas, Jr., Approximate continuation of harmonic and parabolic functions, in "Numerical Solution of Partial Differential Equations," J. H. Bramble, editor, pp. 353-364, Academic Press, New York, 1966.
51. J. Douglas, Jr., Mathematical programming and integral equations, in "Symposium on the Numerical Treatment of Ordinary Differential Equations, Integral and Integro-Differential Equations," pp. 269-274, Birkh user, Basel, 1960.
52. J. Douglas, Jr., A numerical method for analytic continuation, in "Boundary Problems in Differential Equations, pp. 179-189, University of Wisconsin Press, Madison, Wisconsin, 1960.
53. J. Douglas, Jr. and S. Jones, The determination of a coefficient in a parabolic differential equations, J. Math. Mech. 11 (1962), pp. 919-926.
54. J. Douglas and T. M. Gallie, An approximate solution of an improper boundary value problem, Duke Math. J., 26 (1959), pp. 339-347.
55. N. Dunford and J. Schwartz, Linear Operators I: General Theory, Interscience, New York, 1958.



56. S. Elubaev, On the inverse problem for the telegraph equation, Dokl. Akad. Nauk SSSR, 189 (1969), pp. 461-463; Soviet Math. Dokl. 10 (1969), pp. 1421-1423.
57. H. E. Fleming and W. L. Smith, Inversion technique for remote sensing of atmospheric temperature profiles, to appear.
58. J. N. Franklin, Well-posed stochastic extensions of ill-posed linear problems, *J. Math. Anal. Appl.*, 31 (1970) pp. 682-716.
59. V. M. Fridman, Method of successive approximations for Fredholm integral equations of the first kind, *Uspehi Mat. Nauk*, 11 (1956), pp. 233-234.
60. K. O. Friedrichs, Asymptotic phenomena in mathematical physics, *Bull. Amer. Math. Soc.*, 61 (1955), pp. 485-504.
61. H. Fujita, On the blowing up of solutions of the Cauchy problem for  $u_t = \Delta u + u^{1+d}$ , *Journal of the Faculty of Science, University of Tokyo*, Vol. 13, 1967.
62. M. Furi, M. Martelli, and A. Vignoli, On minimum problems for families of functionals, *Annali di Matematica Pura ed Applicata*, 86 (1970), pp. 181-187.
63. M. Furi and A. Vignoli, A characterization of well-posed minimum problems in a complete metric space. *J. Optimization Theory and Applications* 5 (1970), pp. 452-461.
64. M. Furi and A. Vignoli, About well-posed optimization problems for functionals in metric spaces, *J. Optimization Theory and Applications* 5 (1970), pp. 225-229.
65. M. Furi and A. Vignoli, On the regularization of a nonlinear ill-posed problem in Banach spaces, *J. Optimization Theory and Applications* 4 (1969), pp. 206-209.
66. P. Garabedian and M. Lieberstein, *J. Aeron. Sciences*, 25 (1958), pp. 109-118.
67. V. B. Glasko and Yu. M. Timofeyev, Possibilities of the regularization method in solving the problem of atmospheric thermal sounding, *Izv. atmospheric and Oceanic Physics*, 4 (1968), pp. 1243-1253.
68. M. Golomb and H. F. Weinberger, Optimal approximation and error bounds, in "On Numerical Approximation", *Proc. Sympos. (Madison, Wisconsin, 1958)*, R. E. Langer, editor, Univ. of Wisconsin Press, Madison, Wisconsin, 1959, pp. 117-190.
69. A. V. Gončarskiĭ and A. G. Jagola, Uniform approximation of monotonic solutions of incorrect problems, *Dokl. Akad. Nauk SSSR* 184 (1969), pp. 771-773; *Soviet Math. Dokl.* 10 (1969), pp. 155-157.

70. A. V. Gončarskiĭ and A. G. Jagola, On the solution of integral equations of the form  $\int_a^b k(x,s) dg(s) = u(x)$ , Dokl. Akad. Nauk SSSR 193 (1970), No. 2; Soviet Math. Dokl. 11 (1970), pp. 901-903.
71. T. N. E. Greville, The pseudoinverse of a rectangular matrix and its application to the solution of systems of linear equations, SIAM Rev., 1 (1959), pp. 38-43.
72. T. N. E. Greville, Some applications of the pseudoinverse of a matrix, SIAM Review, 2 (1960), pp. 15-22.
73. J. Hadamard, Lectures on Cauchy's Problem in Linear Partial Differential Equations, Dover, New York, 1952.
74. R. Hanson, A numerical method for solving Fredholm integral equations of the first kind using singular values, SIAM J. Numer. Anal., 9 (1970), pp. 616-622.
75. P. Henrici, An algorithm for analytic continuation, SIAM J. Numer. Anal., 3 (1966), pp. 67-78.
76. B. M. Herman, S. R. Browning, and J. A. Reagan, Determination of aerosol size distributors for Lidar measurements, J. Atmosph. Sci., 28 (1971), pp. 763-771.
77. R. B. Holmes, A Course on Optimization and Best Approximation, Lecture Notes in Mathematics 257, Springer-Verlag, Berlin-Heidelberg-New York, 1972.
78. L. Hörmander. Linear Partial Differential Operators, Springer-Verlag, Berlin, 1963.
79. B. R. Hunt, Extension of a method of Phillips and Twomey to solving convolution-type integral equations by the use of fast Fourier transform, Los Alamos Scientific Laboratory, Los Alamos, New Mexico, 1971.
80. B. V. Hvedildze, The regularization problem in the theory of integral equations with Cauchy kernel, Dokl. Akad. Nauk SSSR, 140 (1961), pp. 66-68; Soviet Math. Dokl., 2 (1961), pp.1169-1171.
81. V. K. Ivanov, On linear problems which are not well posed, Dokl. Akad. Nauk SSSR, 153 (1963), pp. 49-52; Soviet Math. Dokl., 3 (1962), pp. 981-983.
82. V. K. Ivanov, The approximate solution of operator equations of the first kind, Zh. Vychisl. Mat. Mat. Fiz. 6 (1966), pp. 1089-1094; USSR Comp. Math. Math. Phy., 6 (1966), pp. 197-205.
83. V. K. Ivanov, Integral equations of the first kind and the approximate solution of the inverse potential problem, Dokl. Akad. Nauk SSSR 142 (1962) pp. 998-1000; Soviet Math. Dokl., 3 (1962), pp. 210-212.

84. V. K. Ivanov, Ill-posed problems in topological spaces, *Sibirsk. Mat. Zh.*, 10 (1969), pp. 1065-1074.
85. V. K. Ivanov, On uniform regularization of unstable problems, *Sibirskii Matem. Zh.*, 7 (1966), pp. 547-558.
86. V. K. Ivanov and T. I. Korolyuk, Error estimates for solutions of incorrectly posed linear problems, *Zh. Vychisl. Mat. Mat. Fiz.*, 9 (1969), pp. 30-41; *USSR Comp. Math. Math. Phys.*, 9 (1969), 35-49.
87. V. V. Ivanov and V. Yu. Kudrinskii, Approximate solution of linear operator equations in Hilbert space by the method of least squares. I, *Zh. Vychisl. Mat. Mat. Fiz.*, 6 (1966), pp. 831-944; *USSR Comp. Math. Math. Phys.*, 6 (1966), pp. 60-75.
88. V. V. Ivanov and V. Yu. Kudrinskii, Approximate solution of linear operator equations in Hilbert spaces by the method of least squares. II, *Zh. Vychisl. Mat. Mat. Fiz.* 7,3 (1967), pp. 475-496; *USSR Comp Math Math. Phys.* 7 (1967), pp. 1-29.
89. V. V. Ivanov, On the convergence of some computational algorithms of the method of least squares, *Zh. vychisl. Mat. Mat. Fiz.* 1,6 (1961), pp. 963-971.
90. N. N. Janeko and Ju. I. Sokin, Correctness of first differential approximations of difference schemes, *Dokl. Akad. Nauk SSSR*, 182 (1968), pp. 776-778; *Soviet Math. Dokl.*, 9 (1968), pp. 1215-1217.
91. F. John, Non-admissible data for differential equations with constant coefficients, *Comm. Pure Appl. Math.* 10, (1957), pp. 391-398.
92. F. John, Continuous dependence on data for solutions of partial differential equations with a prescribed bound. *Comm. Pure Appl. Math.*, 13 (1960), pp. 551-585.
93. F. John, Partial differential equations, Chapter V in *Mathematics Applied to Physics*, E. Roubine, editor, Springer-Verlag, New York, 1970.
94. R. Kalman, Contributions to the theory of optimal control, *Bol. Soc. Mat. Mexicano*, 1960, pp. 102-119.
95. W. J. Kammerer and M. Z. Nashed, Iterative methods for best approximate solutions of linear integral equations of the first and second kinds, *MRC Technical Summary Report #1117*, Mathematics Research Center, The University of Wisconsin; to appear in *J. Math. Anal. Appl.*, 1972.
96. W. J. Kammerer and M. Z. Nashed, Steepest descent for singular linear operators with nonclosed range, *Applicable Analysis*, 1 (1971), pp. 143-159.

97. W. J. Kammerer and M. Z. Nashed, On the convergence of the conjugate gradient method for singular linear operator equations, *SIAM J. Numer. Anal.*, 9 (1972), pp. 165-181.
98. W. J. Kammerer and M. Z. Nashed, A generalization of a matrix iterative method of G. Cimmino to best approximate solution of linear integral equations of the first kind, *Accad. Nuz. Die Lincei, Serie VIII*, 51 (1971), pp. 20-25.
99. S. Kaplan, On growth of solutions, *Comm. Pure Appl. Math.*, 16 (1963), pp. 305-330.
100. Yu. I. Khudak, On the regularization of solutions of integral equations of the first kinds, *Zh. Vychisl. Mat. Mat. Fiz.*, 6 (1966), pp. 766-769; *USSR Comp. Math. Math. Phys.*, 6(1966), pp.217-221.
101. G. Kimeldorf and G. Wahba, Some results on Tchebycheffian spline functions, *J. Math. Anal. Appl.*, 33 (1971), pp. 82-95.
102. A Klinger, Approximate pseudoinverse solutions to ill-conditioned linear systems, *J. Optimization theory and Applications* 2 (1968), pp. 117-123.
103. S. G. Krein, On correctness classes for some boundary problems, *Dokl. Akad. Nauk, SSSR*, 114 (1957), pp. 1162-1165.
104. S. G. Krein and O. I. Prozorovskaja, Approximate Methods of Solving ill-posed problems, *Ž. Vycisl. Mat. i Mat. Fiz.* 3 (1963), pp. 120-130.
105. L. Landweber, An iteration formula for Fredholm integral equations of the first kind, *Amer. J. Math.*, 73 (1951), pp. 615-624.
106. F. M. Larkin, Optimal approximation in Hilbert spaces with reproducing kernel functions, *Math. Comp* 24 (1970), pp. 911-921.
107. R. Lattes and J. L. Lions, *Theory and Applications of the Method of Quasi-Reversibility*, American Elsevier Publishing Co., New York, 1969.
108. M. M. Lavrentiev, On integral equations of the first kind, *Dokl, Akad. Nauk SSSR*, v. 127, 1959.
109. M. M. Larrentiev, On certain ill-posed problems of mathematical physics, pp. 258-276 of *Problems of Numerical Mathematics and Applications*, G. I. Marchouk, editor, Novosibirsk, 1966 (in Russian).
110. M. M. Lavrentiev, On the Cauchy problem for the Laplace equation, *Izv, Akad. Nauk.* 20 (1956), pp. 819-842 (in Russian).
111. M. M. Lavrentiev, *Some Improperly Posed Problems of Mathematical Physics*, Izdat. Sibirsk. Otdel. Akad. Nauk SSSR, Novosikbirsk, 1962; English Transl. *Springer Tracts in Natural Philosophy*, Vol. 11, Springer-Verlag, Berlin, 1967.

112. M. M. Lavrentiev and V. G. Vasiliev, On the formulation of some improperly posed problems of mathematical physics, *Sibirsk. Mat. Z.*, 7 (1966), pp. 559-576.
113. M. M. Lavrentiev, V. G. Romanov, and A. M. Vasiliev, *Multidimensional Inverse Problems in Differential Equations, Lecture Notes in Mathematics* 167, Springer-Verlag, Berlin-Heidelberg-New York, 1970.
114. P. D. Lax and B. Wendroff, On the stability of difference schemes with variable coefficients, *Comm. Pure Appl. Math.*, 14 (1961), pp. 497-520.
115. V. I. Lebedev, On the solution on compact sets of some restoration problems, *Zh. vĥisl. Mat. mat. Fiz.* 6 (1968), pp. 1002-1010; *USSR Comp. Math. and Math. Phys* 6 (1968), pp. 79-101.
116. M. K. Likht, Dependence in the minimum problem of quadratic functionals, *Zh. vĥisl. Mat. Mat. Fiz.* 3,6 (1963), pp. 979-987.
117. M. K. Likht, Conditionality in the minimum problem for a quadratic functional, *Zh. vĥisl. Mat. mat. Fiz.* 3,6(1963), pp. 939-987.
118. M. K. Likht, On the calculation of functionals in the solution of linear equations of the first kind, *Zh. vĥisl. Mat. Mat. Fiz.* 7,3(1967), pp. 667-672; *USSR Comp. Math. Math Phys.* 7 (1967), pp. 211-218.
119. M. K. Likht, Resolvent of the problem of the minimum of a quadratic functional in a special case, *Zh. vĥisl. Mat. Mat. Fiz.* 9 (1969), pp. 201-203; *USSR. Comp. Math. and Math. Phys.* 9, (1969), pp. 269-272.
120. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, 1970.
121. J. L. Lions, Approximation numérrique de la solution des problèmes d'équations aux dérivées partielles, ch. VII in "Mathematics Applied to Physics," E. Roubine, editor, Springer-Verlag, New York, 1970.
122. W. S. Loud, Generalized inverses and generalized Green's functions, *SIAM J. Appl. Math.*, 14 (1966), pp. 342-369.
123. W. S. Loud, Some examples of generalized Green's functions and generalized Green's matrices, *SIAM Review*, 12(1970), pp. 194-210.
124. G. I. Marcuk, Formulation of some converse problems, *Soviet Math.*, 5 (1964), pp. 675-678.
125. G. I. Marcuk and S. A. Atanbaev, Some problems of "global" regularization, *Dokl. Akad. Nauk SSSR* 190 (1970), *Soviet Math. Dokl.* 11 (1970), pp. 148-152.

126. G. I. Marčuk and Ju. A. Kuznecov, On the question of optimal iteration processes, Dokl. Akad. Nauk SSSR 181 (1968), pp. 1331-1334; Soviet Math. Dokl. 9 (1968), pp. 1041-1044.
127. G. I. Marčuk and V. G. Vasil'ev, On an approximate solution for operator equations of the first kind, Dokl. Akad. Nauk SSSR 195 (1970); Soviet Math. Dokl. 11 (1970), pp. 1562-1565.
128. K. Marton and L. Varga, Regularization of certain operator equations by filters, Studies Sci. Math. Haugarica, 6 (1971), pp. 457-465.
129. V. P. Maslov, Regularization of incorrect problems for singular integral equations, Dokl. Akad. Nauk SSSR, 176 (1967), pp. 1012-1014; Soviet Math. Dokl., 8 (1967), pp. 1251-1253.
130. V. P. Maslov, Regularization of the Cauchy problem for pseudodifferential equations, Dokl. Akad. Nauk SSSR, 177 (1967), pp. 1277-1280, Soviet Math Dokl., 8 (1967), pp. 1588-1591.
131. H. Meschkowski, Hilbertsche Räume mit Kernfunktion die Grundlehren der Math. wissenschaften, Band 113, Springer-Verlag, Berlin, 1962.
132. S. G. Mikhlin, The Problem of the Minimum of a Quadratic Functional, Holden-Day, San Francisco, 1965.
133. S. G. Mikhlin, Mathematical Physics, An Advanced Course, North Holland Publishing Co., Amsterdam-London, 1970.
134. K. Miller, Three circle theorems in partial differential equations and applications to improperly posed problems, Arch. Rat. Mech. Anal., 16 (1964), pp. 126-154.
135. K. Miller, An eigenfunction expansion method for problems with over-specified data, Annali della Scuola Normale Superiore di Pisa, Serie III, 19, No. 3, 1965.
136. S. Mizohata, some remarks on the Cauchy problem, Jour. of Math., of Kyoto University, 1 (1961), pp. 110-112.
137. S. Mizohata, Lectures on the Cauchy Problems, Tata Institute of Fundamental Research, Bombay, 1965.
138. E. H. Moore, General Analysis, Memoirs of the American Philosophical Society, Part I, 1935, Part II, 1939.
139. V. A. Morozov, On the solution of functional equations by the method of regularization, Dokl. Akad. Nauk SSSR, 167 (1966), pp. 510-512., Soviet Math. Dokle. 7 (1966), pp. 414-416.



140. V. A. Morozov, Choice of parameter for the solution of functional equations by the regularization method, Dokl. Akad. Nauk SSSR, 175 (1967); Soviet Math. Dokl. 8 (1967), pp. 1000-1003.
141. V. A. Morozov, Regularization of incorrectly posed problems and the choice of regularization parameters, Zh. Vychisl. Mat. Mat. Fiz., 6 (1966), pp. 170-175; USSR Comp. Math. Math. Phys., 6 (1966), pp. 242-251.
142. V. A. Morozov, The error principle in the solution of operator equations by the regularization method, Zh. Vychisl. Mat. Mat. Fiz., 8 (1968), pp. 295-309.
143. V. A. Morozov, A Stable Method for computation of values of unbounded operators, Dokl. Akad. Nauk SSSR, 185 (1969), pp. 267-270; Soviet Math. Dokl. 10 (1969), pp. 339-341.
144. M. Z. Nashed, Generalized inverses, normal solvability and iteration for singular operator equations, in Nonlinear Functional Analysis and Applications, L. B. Rall, editor, Academic Press, New York, 1971, pp. 311-359.
145. M. Z. Nashed, Steepest descent for singular linear operator equations, SIAM J. Numer. Anal., 7 (1970), pp. 358-362.
146. M. Z. Nashed, Differentiability and related properties of nonlinear operators: some aspects of the role of differentials in nonlinear functional analysis, in Nonlinear Functional Analysis and Applications, L. B. Rall, editor, Academic Press, New York, 1971, pp. 103-309.
147. M. Z. Nashed, and G. Wahba, Approximate regularized pseudosolutions of linear operator equations in reproducing kernel spaces, MRC Technical Summary Report in preparation; Mathematics Research Center, The University of Wisconsin-Madison.
148. M. Z. Nashed and G. Wahba, Rates of convergence of approximate least squares solutions of linear integral and operator equations, MRC Technical Summary Report #1260 Mathematics Research Center, The University of Wisconsin-Madison.
149. I. P. Nedyalkov, An approach to the theory of incorrect problems, C. R. Acad. Bulgare Sci., 23 (1970), pp 5-8.
150. J. von Neumann and R. Richtmyer, A method for the numerical calculations of hydrodynamical shocks, J. Appl. Phys., 21 (1950), pp. 232-247.
151. B. Noble, A bibliography on : "Methods for solving integral equations" Author listing, MRC Tech. Summary Rept. #1176, Mathematics Research Center, The University of Wisconsin-Madison, September 1971.



152. B. Noble, A bibliography on: "Methods for solving integral equations" Subject listing, MRC Tech. Summary Rept. #1177, Mathematics Research Center, The University of Wisconsin-Madison, September 1971.
153. O. A. Oleink, Discontinuous solutions of nonlinear partial differential equations, *Uspekhi Mat. Nauk*, 12 (1957), pp. 3-73.
154. O. A. Oleink, A Problem of Fichera, *Dokl. Akad. Nauk*, 157 (1963), pp. 1297-1301; *Soviet Math.* 5 (1964), 1129-1133.
155. V. P. Palamodov, Regularization and the problem of division, *Dokl. Akad. Nauk SSSR* 132 (1960), pp. 295-298; *Soviet Math. Dokl.*, 1 (1960), pp. 560-563.
156. O. I. Panić, Equivalent regularization of boundary value problems by means of potentials, *Dokl. Akad. Nauk, SSSR*, 184 (1969), pp. 554-557; *Soviet Math. Dokl.*, 10 (1969), pp. 142-145.
157. E. Parzen, A new approach to the synthesis of optimal smoothing and prediction systems, Chapter 5 of *Mathematical Optimization Techniques*, R. Bellman, editor, University of California Press, 1963.
158. E. Parzen, Statistical inference on time series by RKHS methods, in *Proceedings of the 12th Biennial Canadian Mathematical Seminar*, R. Pyke, editor, 1971, pp. 1-37.
159. G. B. Pedrick, Theory of reproducing kernel spaces of vector valued functions, University of Kansas, Tech. Report 19, 1957.
160. A. P. Petrov and A. V. Khovanskii, Lower bound of the error in the solution of operator equations in compact spaces, *Zh. vychisl. Mat. mat. Fiz.* 9 (1969), pp. 194-201; *USSR Comp. Math and Math Phys.*, 9 (1966), pp. 257-268.
161. I. Petrovsky, *Lectures on Partial Differential Equations*, Interscience Publ., New York, 1954.
162. D. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, *J. Assoc. Comp. Mach.*, 9 (1962), pp. 84-97.
163. C. Pucci, Sui problemi di Cauchy non "ben posti," *Rend. Accad. Naz. Lincei* (8), 18 (1955), pp. 473-477.
164. C. Pucci, Discussione del problema di Cauchy per le equazioni di tipo ellitico, *annali Mat. Pura Appl.*, 46 (1958), pp. 131-153.
165. W. T. Reid, Generalized inverses of differential and integral operators, in *Theory and Applications of Generalized Inverse of Matrices*, T. L. Boullion and P. L. Odell, editors, Texas Technological College, Mathematics Series, No. 4, Lubbock, Texas, 1968, pp. 1-25.

166. G. Ribiere, Regularisation d'operateurs, Rev. Franfarse Informat. Recherche operationnelle, 1 (1967), pp. 57-79.
167. V. S. Ryabenkii, Necessary and sufficient conditions for boundary value problems to be well posed, Zh. vychisl. Mat. mat. Fiz., 4 (1964), pp. 242-255.
168. W. W. Schmaedeke, A new approach to unstable problems using variational techniques, J. Math. Anal, 25 (1969), pp. 272-275.
169. L. Schwartz, Sous-espaces Hilbertiens d'espaces vectoriel topologiques et noyaux associes (noyaux reproduisants), J. Analyse Math., 13 (1964), pp. 115-256.
170. H. L. Shapior, Topics in Approximation Theory, Springer-Verlag Lecture Notes in Mathematics 187, Berlin, 1970.
171. D. W. Showalter and A. Beu-Israel, Representation and computation of the generalized inverse of a bounded linear operator between two Hilbert spaces, Accad. Naz. Die Lincei, 48 (1970), pp. 184-194.
172. D. Slepian and H. O. Polak, Prolate spheroidal wave functions, and uncertainty-I, Bell System Technical Journal, 40 (1961), pp. 43-64.
173. F. Smithies, Integral Equations, Cambridge University Press, Cambridge, 1958.
174. S. L. Sobolev, Applications of Functional Analysis in Mathematical Physics, Leningrad, 1950; English transl., Transl. Math. Monographs No. 7, Amer. Math. Soc., Providence R. I., 1963.
175. F. W. Stallmann, Numerical Solution of Integral Equations, Numer. Math., 15 (1970), pp. 297-305.
176. V. N. Strahov, Methods of the approximate solution of linear conditionally well posed problems, Dokl. Akad. Nauk. SSSR, 196 (1971), pp. 786-788; Soviet Math. 12 (1971), pp. 271-274.
177. O. N. Strand, Theory and Methods for Operator Equations of the First Kind, Ph. D. Thesis, Colorado State University, Fort Collins, Colorado, 1972.
178. O. N. Strand and E. R. Westwater, Statistical estimation of the numerical solution of a Fredholm integral equation of the first kind, Jour. Assoc. Comp. Mach., 15 (1968), pp. 100-114.
179. O. N. Strand and E. R. Westwater, Minimum-RMS estimation of the numerical solution of a Fredholm integral equation of the first kind, SIAM J. Numer. Anal., 5 (1968), pp. 287-295.

180. V. P. Tanana, Incorrectly posed problems and the geometry of Banach spaces, Dokl. Akad. Nauk SSSR 193 (1970); Soviet Math. Dokl. 11 (1970), pp 864-867.
181. A. N. Tikhonov, On the stability of the functional optimization problem, USSR Computational Mathematics and Mathematical Physics, Vol. 6, No. 4, 1966.
182. A. N. Tikhonov, Methods for regularization of optimal control problems, Dokl. Akad. Nauk SSSR 162 (1965), pp. 763-765; Soviet Math. Dokl. 6 (1965), pp. 761-763.
183. A. N. Tikhonov, On incorrectly formulated problems, Computational Methods and Programming, No. 8 Moscow, 1967.
184. A. N. Tikhonov, On the solution of incorrectly formulated problems and the regularization method, Dokl. Akad. Nauk SSSR, 151 (1963), pp. 501-504; Soviet Math., 4 (1963), pp. 1035-1038.
185. A. N. Tikhonov, On nonlinear equations of the first kind, Dokl. Akad. Nauk SSSR, 161 (1965), pp. 1023-1027; Soviet Math. Dokl. 6 (1965), pp. 559-562.
186. A. N. Tikhonov and V. B. Glasko, Use of the regularization method in nonlinear problems, USSR Comput. Math and Math. Phys. 5 (1965), no. 3, pp. 93-107.
187. A. N. Tikhonov, Regularization of incorrectly posed problems, Dokl. Akad. Nauk. SSSR, 153 (1963), pp. 49-52; Soviet Math., 4 (1963), pp. 1624-1627.
188. A. N. Tikhonov, On the stability of inverse problems, Dokl. Akad. Nauk, SSSR, 39 (1944), pp. 195-198.
189. A. N. Tikhonov, On methods of solving incorrect problems, Amer. Math. Soc. Translations, Series 2 70 (1968), pp. 222-224.
190. A. N. Tikhonov and V. B. Glasko, An approximate solution of Fredholm integral equations of the first kind, Z. Vycisl. Mat. i Mat. Fiz., 4 (1964), pp. 564-571.
191. S. Twomey, On the numerical solution of Fredholm integral equations of the first kind by inversion of the linear system produced by quadrature, J. Assoc. Comp. Mach., 10 (1963), pp. 97-101.
192. V. V. Vasin, Regularization of numerical differentiation problem, Ural. Gos. Univ. Mat. Zap., 8 (1969-70), pp. 29-33.
193. V. A. Vinokurov, On a necessary condition for regularizability in the sense of Tihonov, Dokl. Adad. Nauk SSSR, 195 (1970); Soviet Math. Dokl., 11 (1970), pp. 1495-1496.

194. V. V. Voevodin, The method of regularization, Zh. Vychisl. Mat. Mat. Fiz., 9 (1969), pp. 673-685; USSR Comp. Math. Math. Phys., 9 (1969), pp. 228-238.
195. G. Wahba, A class of approximate solutions to linear operator equations, J. Approx. Theory, to appear.
196. G. Wahba, Convergence rates for certain approximate solutions to Fredholm integral equations of the first kind, J. Approx. Theory, to appear.
197. G. Wahba, On the regression design problem of Sacks and Ylvisaker, Ann. Math. Statist., 42 (1971), pp. 1035-1053.
198. J. H. Wilkinson, Rounding Error in Algebraic Processes, Prentice-Hall, Englewood Cliff, N. J., 1963.
199. K. Yosida, Functional Analysis, 2nd ed., Springer-Verlag, Berlin-New York, 1968.
200. M. Z. Nashed and G. Wahba, Generalized inverses in reproducing kernel spaces; an approach to the regularization of linear operator equations, MRC Tech. Summary Rept. #1200, Mathematics Research Center, The University of Wisconsin-Madison, 1971.

# Kinetics Analysis by Digital Parameter Estimation

---

Y. BARD  
*IBM Corporation  
New York Scientific Center  
New York, New York*

and

L. LAPIDUS  
*Department of Chemical Engineering  
Princeton University  
Princeton, New Jersey*

I. KINETIC AND MODEL DEFINITIONS . . . . .	69
II. PARAMETER ESTIMATION . . . . .	73
A. Least-Squares Minimization . . . . .	74
B. Maximum Likelihood . . . . .	77
C. Bayesian Estimation . . . . .	80
III. METHODS OF NUMERICAL SOLUTION . . . . .	81
A. Direct-Search Methods . . . . .	82
B. Gradient Methods . . . . .	82
C. Constraints . . . . .	87
D. Interpretation of Parameter Estimates . . . . .	88
IV. IMPROVING PARAMETER ESTIMATES BY PROPER EXPERIMENTAL DESIGN . . . . .	89
V. COMPUTATIONAL RESULTS IN PARAMETER ESTI- MATION . . . . .	92
VI. DESIGN OF EXPERIMENTS FOR MODEL DISCRIMINA- TION . . . . .	103
A. Termination of a Sequence of Experiments . . . . .	105
B. Numerical Example . . . . .	105
Appendix . . . . .	108
References . . . . .	109

---

This paper appeared on pages 67-112 of the book entitled *Catalysis Reviews, Volume 2*, published in 1969 by Marcel Dekker, Inc., New York. We would like to thank this company for permission to photographically reproduce this article.

One of the more important problems that a chemist or chemical engineer may encounter is the determination of the mechanism and/or model for a chemical reaction system. This problem may be encompassed within the broader area of identification [30-32] which ranges from the one extreme where no a priori information regarding the system representation is known, to the other extreme where much is known about the system. The development of a suitable chemical kinetic representation for the experimental reaction system lies somewhere between these two extremes. This results from the fact that there usually exist certain postulated models to represent experimental kinetic data; however, the determination of the best constants in these mechanistic models and perhaps the discrimination among a number of alternative models usually remain as vexing questions.

We shall adopt the view here that for any given reaction system there exists a large but finite number of permissible mechanisms, i.e., mathematical models which, on an a priori basis, could describe the reaction. While this number may be formidable, the investigator can usually focus his attention on only those selected models which he considers appropriate to analyze in detail. The results of this analysis produce one or more plausible models which are adequate descriptions of the experimental data. If there is only one such model, the investigator might, after further consideration, accept this as the correct model. More often than not, there are several models which are adequate. These must be subjected to further study, perhaps along with other models not previously studied. Through this additional study, including more extensive experimentation and other independent information, the investigator hopes to be able to discriminate among alternative models to find the correct model, assuming one exists.

The difficulties of this analysis are numerous. They rest, for example, on the proper design of experiments which is used to generate the data. Box and co-workers [8-14] and Blakemore and Hoerl [5] have emphasized the need for appropriately designed experiments since the damage of poor design is irreparable and may negate the subsequent analysis no matter how ingenious this analysis. Assuming several models satisfy all the features of the analysis, the selection of an adequate model can be made only tentatively. Further experimentation must be conducted or independent information brought to bear to attempt to discriminate among competitive models and thus to determine that model which best fits the data.

In this context, one can recognize the process of kinetic investigation as an iterative one in which experimentation and a proposed model lead to data analysis, which in turn leads to further experi-

mentation. This combined approach continues in a presumably converging cycle between analysis and experimentation toward a choice of the most adequate model. Since the investigation cannot be exhaustive in examining all possibilities, no proof exists that the correct model has been found. The investigator must be satisfied that he has determined only the most adequate representation of the experimental kinetic data.

It is apparent that the advent of the modern digital computer has greatly enlarged the scope of methods for such data analysis and model construction and discrimination. In the present article we shall point out how the high-speed capabilities of the computer can be used as an integral part of the overall cyclic procedure mentioned above. In other words we shall show how the computer may first be used to construct the kinetic models (nonlinear parameter estimation) which represent in a statistical sense the experimental kinetic data. Application to both homogeneous and heterogeneous kinetic systems will be discussed as well as a variety of alternative statistical formulations. It will then be shown how the computer can be used to specify the experimental conditions for obtaining further kinetic data so as to aid the discrimination among the alternative models or to improve the accuracy of a single model.

### I. KINETIC AND MODEL DEFINITIONS

To characterize the models which we shall consider in this article, we shall first specify the types of variables and then the forms of equations relating the variables, and we shall finally show how the kinetic equations are encompassed within this formulation. Thus we define the following types of variables:

1. Parameters. These are constants within a model whose numerical values are unknown. The (column) vector of  $p$  parameters is denoted by  $\theta = \{\theta_1, \theta_2, \dots, \theta_p\}$ .

2. Independent Variables. These are variables which are either fixed arbitrarily for each experiment or which are known precisely for each experimental observation. Typical independent parameters in a kinetics experiment might be the reaction time or the space velocity. The (column) vector of  $k$  independent variables for the  $\mu$ th experiment is denoted by  $x_\mu = \{x_{\mu 1}, x_{\mu 2}, \dots, x_{\mu k}\}$ .

3. Dependent Variables. These are the variables which the model will predict on the basis of known values of the parameters and independent variables. Thus, given a specified set of parameter values and reaction time, the model will predict a concentration of a component in the kinetic model. The (column) vector of  $n$  dependent



variables is denoted by  $\eta = \{\eta_1, \eta_2, \dots, \eta_n\}$ .

4. Observed Variables. Those dependent variables which are actually measured in an experiment are called the observed variables. The (column) vector of  $r$  observed variables for the  $\mu$ th experiment is denoted by  $y_\mu = \{y_{\mu 1}, y_{\mu 2}, \dots, y_{\mu r}\}$ .

An experiment consists of the measurement of all observed variables for a given set of values of the independent variables.

The models which we shall consider relate the dependent variables to the independent variables and the parameters. In concise form we may write the model as

$$q(\eta_\mu, x_\mu, \theta) = 0 \quad \mu = 1, 2, \dots, m \quad (1)$$

where  $q$  is a vector of functions of dimension equal to that of the dependent variables and  $m$  the number of experiments. An alternative formulation is to write the reduced form

$$\eta_\mu = f(x_\mu, \theta) \quad \mu = 1, 2, \dots, m \quad (2)$$

While Eqs. (1) and (2) are equivalent mathematically, care must be used in terms of computer algorithms since parameter estimates may not be invariant under the transformation from one form to the other.

Let us now show how kinetic rate equations fit into the formulation of Eqs. (1) or (2). Suppose we have a unidirectional, unimolecular homogeneous reaction of the type



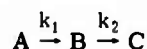
For isothermal conditions the rate equation for this reaction is

$$\frac{d[N]}{dt} = -k[N] \quad (3)$$

where  $[N]$  is the concentration or partial pressure,  $t$  the reaction time, and  $k$  the reaction velocity constant. Integrating this equation with a given initial  $[N]$  at time  $t = 0$ ,  $[N_0]$ , we have

$$[N] = [N_0] \exp(-kt) \quad (4)$$

If we now relate  $\eta = \{[N]\}$ ,  $x = \{t, [N_0]\}$ , and  $\theta = \{k\}$  (the subscript  $\mu$  has been dropped for convenience), we see that Eq. (4) is equivalent to Eq. (2). Now consider the isothermal rate equations corresponding to the reaction



They are

$$\frac{d[A]}{dt} = -k_1[A]$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B]$$

$$\frac{d[C]}{dt} = k_2[B]$$

When initial conditions are established such as

$$[A_0] = 1, [B_0] = [C_0] = 0 \quad t = 0$$

the corresponding integrated forms are given by

$$[A] = \exp\{-k_1 t\}$$

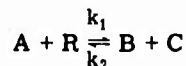
$$[B] = [k_1/(k_2 - k_1)](\exp\{-k_1 t\} - \exp\{-k_2 t\})$$

$$[C] = 1 - [1/(k_2 - k_1)](k_2 \exp\{-k_1 t\} - k_1 \exp\{-k_2 t\})$$

Now defining the vectors  $\eta = \{[A], [B], [C]\}$ ,  $x = \{t\}$ , and  $\theta = \{k_1, k_2\}$ , we see that we have once again the form of Eq. (2).

A feature of this comparison which deserves some further consideration is that in both cases the rate equations can be integrated analytically. Under this requirement we may ascertain the resulting form without any difficulty. But what if we introduce a temperature-dependent velocity constant or nonunimolecular reactions. The probability of an analytical integration decreases to zero rapidly. However, since we shall always consider that we have a digital computer available, this need not distress us greatly. We can always integrate the set of rate equations numerically to yield values of  $\eta$  at any selected  $x$ . Stated in another way, once we have the rate equations it is in a sense immaterial whether we integrate these rate equations analytically or numerically. The end result of either path is a functional equation of the form of Eq. (2).

Let us now briefly turn to a heterogeneous-type kinetic system. As an example, the following bimolecular reversible reaction in the gas-solid phase is considered



If the controlling mechanism is that of surface reaction on dual sites without dissociation, the Langmuir-Hinshelwood expression becomes

$$d[A]/d(W/F_A) = r_A = \frac{k_1[A][R] - k_2[B][C]}{(1 + K_A[A] + K_R[R] + K_B[B] + K_C[C])^2}$$

where the dependent variables denoted by concentrations are normally expressed as partial pressures, the independent variable is reciprocal space velocity,  $W/F_A$ , and the parameters, in addition to those previously encountered, include the adsorption equilibrium coefficients,  $K_A$ ,  $K_R$ ,  $K_B$ , and  $K_C$ .

As compared with homogeneous noncatalytic reactions, it is apparent that  $W/F_A$  assumes the same role as  $t$ , that the numerator in the rate expression is identical, and that the only substantial difference lies in the introduction of the denominator term. Thus, in functional notation, if we define  $\eta = \{[A], [R], [B], [C]\}$ ,  $x = \{W/F_A, [A_0], [R_0], [B_0], [C_0]\}$ , and  $\theta = \{k_1, k_2, K_A, K_R, K_B, K_C\}$ , we have the equivalent of Eq. (2).

Let us close out this section with one further example to illustrate the models which can be encompassed within the form of Eq. (2). Consider that we have three components in a homogeneous reaction and that it is postulated that the rate equations are

$$\begin{aligned} \frac{d\eta_1}{dt} &= -k_1\eta_1\eta_2 + k_2\eta_3 - k_3\eta_1 \\ \frac{d\eta_2}{dt} &= -k_1\eta_1\eta_2 + k_2\eta_3 + k_3\eta_1 \\ \frac{d\eta_3}{dt} &= k_1\eta_1\eta_2 - k_2\eta_3 \end{aligned} \quad (5)$$

where  $\eta_1$ ,  $\eta_2$ , and  $\eta_3$  correspond to, e.g.,  $[A]$ ,  $[B]$ , and  $[C]$  used previously. In addition we have a set of unknown initial concentrations:

$$\begin{aligned} \eta_1 &= a \\ \eta_2 &= b \\ \eta_3 &= 1 - a - b \end{aligned} \quad t = 0 \quad (6)$$

At time  $t$  (representing the  $\mu$ th experiment) we withdraw three samples from the reaction mixture and perform the following tests:

- We determine  $\eta_{\mu 1} = \eta_1(t_\mu)$  directly by titration.
- A different person independently determines  $\eta_{\mu 1}$  by titration.
- We determine the light absorptivity of the solution, assuming this to be a linear function with unknown coefficients of the concentrations.

Denoting the results of these three measurements as  $y_{\mu_1}$ ,  $y_{\mu_2}$ , and  $y_{\mu_3}$ , we have

$$\begin{aligned} y_{\mu_1} &= \eta_{\mu_1} \\ y_{\mu_2} &= \eta_{\mu_1} \\ y_{\mu_3} &= \beta_0 + \beta_1 \eta_{\mu_1} + \beta_2 \eta_{\mu_2} + \beta_3 \eta_{\mu_3} \end{aligned} \quad (7)$$

Equations (5)-(7) together constitute the model; functionally these are related by the equation

$$y_{\mu} = h(\eta_{\mu}, \theta) = f(x_{\mu}, \theta) \quad (8)$$

where  $\theta = \{a, b, k_1, k_2, k_3, \beta_0, \beta_1, \beta_2, \beta_3\}$  and  $x_{\mu} = \{t_{\mu}\}$ . Equation (8) has a slightly different form than Eq. (2) because we have written the left-hand side in terms of the observed variables and not the model dependent variables. This is of no importance, since dependent variables which are not observed play no role in the parameter estimation procedure.

Thus we see that the kinetic models under consideration in this work may be written in the functional form of Eq. (1) or (2) [or (8)] and that all kinetic rate equations can, in one form or another, be encompassed within this formulation.

## II. PARAMETER ESTIMATION

Having established the form of our predictive model and assuming that experimental data are available, the next area of discussion here is to outline efficient methods for determining the parameters  $\theta$  such that the model fits the data. Due to errors in the measurements and inaccuracies in the model, it is impossible to hope for an exact fit. Instead we shall try to find values of  $\theta$  which minimize some appropriate measure of the errors involved. This is parameter estimation. Such a minimization may be carried out in a number of ways, but here we wish to point out the prodigious increase in our ability to do this estimation via computer analysis.

To illustrate the available procedures, let us rewrite our model equation in the form

$$y_{\mu} = f(x_{\mu}, \theta) + \epsilon_{\mu} \quad (9)$$

where  $\epsilon_{\mu}$  is the vector of errors, or residuals, between the observations and the predicted dependent variables. Parameter estimation

then tries to find a set of parameters  $\theta$  such that some scalar function of the errors is minimized. In general we shall write this function as  $S(\theta)$  to indicate the dependence on the chosen parameters.

#### A. Least-Squares Minimization

In the least-squares approach to this minimization the function  $S(\theta)$  is defined by the sum of squares of the errors or

$$S_{LS}(\theta) = \sum_{\mu=1}^m \epsilon_{\mu}^2 = \sum_{\mu=1}^m [y_{\mu} - f(x_{\mu}, \theta)]^2 \quad (10)$$

One seeks to find those  $\theta$  which minimize  $S_{LS}(\theta)$ ; the resulting "best" parameters are indicated by  $\hat{\theta}$ .

It is possible to recognize two cases associated with Eq. (10). In the first case the parameters  $\theta$  enter in  $f(x_{\mu}, \theta)$  in a linear fashion (the result of the minimization is then referred to as linear least squares); in the second case no such linearity occurs and we have nonlinear least squares.

##### 1. Linear Least Squares

Linear least squares is also known as multiple linear regression. The approach generally treated in most kinetic books deals with the very special case where the parameters enter the model equations in a linear fashion. In such a case the model need only be substituted into Eq. (10), the derivative of  $S(\theta)$  taken with respect to  $\theta$  and the result set equal to zero. This yields  $p$  linear algebraic equations corresponding to the  $p$  parameters (called the normal equations). By means of a single matrix inversion it is then possible to solve for the best  $\hat{\theta}$ .

If the  $\epsilon_{\mu}$  are independently and identically distributed errors with zero means, then by this procedure the  $\hat{\theta}$  are efficient linear unbiased estimates of  $\theta$ . In other words the parameter estimates which minimize the sum of squares of the residuals will, on the average, equal the true values of the parameters (unbiased) and will be estimated with maximum precision (minimum variance).

Unfortunately this method has a number of decided defects. Foremost among these is the fact that few kinetic models of any complexity occur in the desired linear form. As a result the model equations must be linearized or rearranged to be handled by this method. Thus for the familiar expression [see Eq. (4)]

$$\eta = \exp\{-kt\}$$

a conventional logarithmic transformation leads to

$$\eta' = \log \eta = -kt$$

which is linear in the parameter  $k$  and is therefore amenable to treatment by linear least squares. However, if the corresponding error equation for the original expression is given by

$$y_{\mu} = \exp\{-kt_{\mu}\} + \varepsilon_{\mu}$$

it is apparent that the logarithmic transformation does not preserve the original distribution of errors which might have been judged appropriate.

This situation is further aggravated when a temperature-dependency condition is introduced through the Arrhenius relation. Then a two-step estimation process with logarithmic transformation for each step is required.

From  $\log \eta = -kt$ , estimate  $k$

From  $\log k = \ln k_0 - (E/R)(1/T)$ , estimate  $k_0$  and  $E$

A similar difficulty ensues when heterogeneous catalytic reactions represented by the Langmuir-Hinshelwood relations are studied by linear least squares. Thus suppose the model of a heterogeneous reaction is taken as

$$r_A = \frac{k'(p_A - p_R p_S / K_{\text{equil}})}{(1 + K_A p_A + K_R p_R + K_S p_S)^2} \quad (11)$$

where the  $p_A, \dots$ , are partial pressures and  $K_{\text{equil}}$  is a known constant. Obviously the parameters  $k', K_A, K_R$ , and  $K_S$  do not appear linearly in the equation. If, however, the equation is rearranged to the form

$$[(p_A - p_R p_S / K_{\text{equil}}) / r_A]^{1/2} = \frac{1}{k'} + \frac{K_A p_A}{k'} + \frac{K_R p_R}{k'} + \frac{K_S p_S}{k'} \quad (12)$$

or

$$\eta_A = \theta_1 + \theta_2 p_A + \theta_3 p_R + \theta_4 p_S$$

then combinations of parameters occur linearly. However, now the partial pressures appear on both sides of the equation and therefore

play the role of both dependent and independent variables. Furthermore, the errors of Eq. (12) are minimized and not the errors of Eq. (11).

As a result of such transformations the method of linear least squares will give parameter estimates which guarantee neither a best prediction of reaction rates [the left side of Eq. (11)] nor reasonably satisfactory extrapolations. Furthermore, in models involving several competing reactions, linearization is usually quite impossible.

In summary, we see that linear least squares is applicable for parameter estimation in only very special cases and that it cannot be prescribed as a viable method except under special circumstances.

## 2. Nonlinear Least Squares

Since the linear least-squares approach has these obvious defects, it seems natural to turn to a method which does not require that the model equations be linear in the unknown parameters. We refer to this method as nonlinear least squares since the parameters may occur in any fashion, linear or nonlinear.

In a broad-gauged description of this method, one attempts to minimize  $S(\theta)$  in an iterative fashion rather than in a single step. As such, and because of the extensive calculations which are then required, the method almost certainly requires implementation via a computer. Methods for performing this minimization are described in a subsequent section.

It is apparent that nonlinear least squares, in which the dependency of the dependent variables on the parameters is essentially unrestricted, obviates many of the difficulties inherent in linear least squares. The kinetic rate equations are used in the form originally proposed without imposing arbitrary transformations; the minimization of the sum of squares of errors is appropriate in terms of untransformed variables. Furthermore, the functional rate models may be formulated explicitly or implicitly, either as integrated or differential rate equations.

Nonlinear least squares, however, does not overcome all problems inherent in parameter estimation. It is not valid when several variables are observed at each experiment; it does not make sense to add together sums of squares in Eq. (10) of, e.g., pressures and temperatures. This problem may be overcome by assigning a weight factor to each variable and minimizing the weighted sum of squares, i.e., instead of Eq. (10) use

$$S_{wLS}(\theta) = \sum_{\mu=1}^m \sum_{i=1}^r W_i [y_{\mu i} - f_i(x_{\mu}, \theta)]^2 \quad (13)$$



Unfortunately one usually does not know what weights to assign, i.e., what numerical values of  $W_i$  to use. On theoretical grounds these weights should be inverses of the variances of the measurement errors, assuming these are known. This problem can be overcome by the maximum likelihood approach to be described below which enables one to estimate not only the parameters but the weights as well.

### B. Maximum Likelihood

The maximum likelihood principle offers a powerful and versatile tool to the parameter estimator. A price must be paid, however, in that explicit assumptions concerning the form of the probability distribution of the errors must be made. Once the form has been assumed, any parameters appearing in the distribution function may be estimated along with the unknown parameters in the model. This is really as it should be; the errors in the observations are as much a part of physical reality as are the reactions being observed. Meaningful parameter estimation can proceed only if we provide a mathematical model not only of the reactions, but also of the errors, and a probability distribution is an appropriate mathematical model for the errors.

Suppose we assume  $p(\varepsilon, \psi)$  to be the joint probability density function of all the errors  $\varepsilon$  in the observed variables. Here  $\psi$  represents a set of unknown parameters (e.g., means, variances) which appear in the formulas for the distribution. Given any values for the model parameters  $\theta$ , we substitute the residuals  $y - f(x, \theta)$  for the errors  $\varepsilon$  in the expression for the probability density. This yields a function depending on  $\theta$  and  $\psi$  (the  $x$  and  $y$  being given by the observations):

$$L(\theta, \psi) = p[y - f(x, \theta), \psi]$$

We refer to  $L(\theta, \psi)$  as the likelihood function. The maximum likelihood method simply consists of finding those values of  $\theta$  and  $\psi$  which maximize  $L$ . For reasons of convenience one usually maximizes  $S_{ML} = \log L$ . This is clearly equivalent to maximizing  $L$ . Thus, we have

$$S_{ML}(\theta, \psi) = \log p[y - f(x, \theta), \psi] \quad (14)$$

If, as is frequently assumed, the observation errors in different experiments are uncorrelated, the joint probability density  $p$  is the product of the individual experiment probability densities  $p_\mu$ , and

$$S_{ML}(\theta, \psi) = \sum_{\mu=1}^m \log p_\mu[y_\mu - f(x_\mu, \theta), \psi] \quad (15)$$

The most frequently used distribution is the normal distribution. If the errors of each experiment are normally distributed with zero means and covariance matrix  $V_\mu$ , then

$$p_\mu(\epsilon_\mu, V_\mu) = (2\pi)^{-r/2} (\det V_\mu)^{-1/2} \exp(-\frac{1}{2} \epsilon_\mu^T V_\mu^{-1} \epsilon_\mu) \quad (16)$$

where  $\det V$  means the determinant of the matrix  $V$  and the superscript  $T$  means the transpose of  $\epsilon$ . Substituting Eq. (16) in Eq. (15), we obtain

$$S_{ML}(\theta, V_\mu) = -\frac{rm}{2} \log 2\pi - \frac{1}{2} \sum_{\mu=1}^m \log(\det V_\mu) - \frac{1}{2} \sum_{\mu=1}^m \epsilon_\mu^T V_\mu^{-1} \epsilon_\mu \quad (17)$$

where we have written  $\epsilon_\mu$  for  $y_\mu - f_\mu = y_\mu - f(x_\mu, \theta_\mu)$ ;  $r$  is the number of observed variables and  $m$  the number of experiments.

We shall now apply the maximum likelihood principle to some special cases of Eq. (17):

a. Suppose all the matrices  $V_\mu$  are known. The only nonconstant term in Eq. (17) is

$$-\frac{1}{2} \sum_{\mu=1}^m \epsilon_\mu^T V_\mu^{-1} \epsilon_\mu$$

and therefore maximizing Eq. (17) is equivalent to minimizing:

$$S(\theta) = \sum_{\mu=1}^m \epsilon_\mu^T V_\mu^{-1} \epsilon_\mu = \sum_{\mu=1}^m \sum_{i=1}^r \sum_{j=1}^r [y_{\mu i} - f_i(x_\mu, \theta)] [y_{\mu j} - f_j(x_\mu, \theta)] V_{\mu ij}^{-1} \quad (18)$$

where  $V_{\mu ij}^{-1}$  is the  $i, j$  element of the matrix  $V_\mu^{-1}$ . This is the most general form of weighted least squares.

b. Suppose the covariance matrices for all experiments are identical, i.e.,  $V_1 = V_2 = \dots = V$ . Then Eq. (17) becomes

$$S_{ML}(\theta, V) = -\frac{rm}{2} \log 2\pi - \frac{nm}{2} \log(\det V) - \frac{1}{2} \sum_{\mu=1}^m \epsilon_\mu^T V^{-1} \epsilon_\mu \quad (19)$$

Consider first the case where all errors are independent. This is

equivalent to saying that the matrix  $V$  is diagonal. Let  $v_i = V_{ii}$  be the  $i$ th diagonal element of  $V$ . Then Eq. (19) becomes

$$S_{ML}(\theta, V) = -\frac{rm}{2} \log 2\pi - \frac{m}{2} \sum_{i=1}^r \log v_i - \frac{1}{2} \sum_{\mu=1}^m \sum_{i=1}^r \epsilon_{\mu i}^2 v_i^{-1} \quad (20)$$

If the  $v_i$  are known, we minimize

$$\sum_{\mu=1}^m \sum_{i=1}^r \epsilon_{\mu i}^2 v_i^{-1}$$

the usual form of weighted least squares. If the  $v_i$  are unknown, we maximize Eq. (20) first with respect to the  $v_i$  by setting  $(\partial S / \partial v_i) = 0$ . Thus

$$\frac{\partial S_{ML}}{\partial v_i} = -\frac{m}{2v_i} + \frac{1}{2} \sum_{\mu=1}^m \frac{\epsilon_{\mu i}^2}{v_i^2} = 0$$

Solving for  $v_i$  we find

$$v_i = \frac{1}{m} \sum_{\mu=1}^m \epsilon_{\mu i}^2 \quad (21)$$

and substituting Eq. (21) into Eq. (20),

$$\begin{aligned} S_{ML}(\theta) &= -\frac{rm}{2} \log 2\pi - \frac{m}{2} \sum_{i=1}^r \log \frac{1}{m} \sum_{\mu=1}^m \epsilon_{\mu i}^2 \\ &\quad - \frac{1}{2} \sum_{i=1}^r \left( m / \sum_{\mu=1}^m \epsilon_{\mu i}^2 \right) \sum_{\mu=1}^m \epsilon_{\mu i}^2 \\ &= -\frac{rm}{2} \left( 1 + \log \frac{2\pi}{m} \right) - \frac{m}{2} \sum_{i=1}^r \log \sum_{\mu=1}^m \epsilon_{\mu i}^2 \end{aligned} \quad (22)$$

Maximizing Eq. (22) is equivalent to minimizing

$$S(\theta) = \sum_{i=1}^r \log \sum_{\mu=1}^m \epsilon_{\mu i}^2 \quad (23)$$

This may be regarded as solving a weighted least-squares problem with unknown weights. Once the maximizing values of  $\theta$  have been found, the unknown weights can be estimated from Eq. (21).

c. Analogously, it can be shown [24] that if  $V$  is an unknown, non-diagonal matrix, maximizing Eq. (19) is equivalent to minimizing

$$S(\theta) = \log(\det M) \quad (24)$$

where  $M$  is the moment matrix of the residuals

$$M_{ij} = \sum_{\mu=1}^m \varepsilon_{\mu i} \varepsilon_{\mu j}$$

and  $V$  is estimated by

$$V = \frac{1}{m} M \quad (25)$$

If the observed variables are linearly dependent (as concentrations are related in material balances), the matrix  $M$  will be nearly singular and minimization of Eq. (24) will be difficult, if not impossible. It is thus recommended that Eq. (23) be used instead of Eq. (24) in such cases. Alternatively a linearly independent subset of the observed variables may be chosen, and "pseudo-observations" for these may be calculated by linear regression from the totality of observed variables for each experiment. These pseudo-observations may then be used safely in Eq. (24).

### C. Bayesian Estimation

It occurs frequently that even before we start estimating the parameters from current data, some information concerning the values of the parameters is available from previous experiments, or from general physical considerations. It is frequently possible to summarize this a priori information in the form of a relative probability density function  $p_0(\theta)$ . If, e.g.,  $p_0(\theta_1)/p_0(\theta_2) = 10$ , this is interpreted to mean that we think  $\theta = \theta_1$  to be 10 times more likely than  $\theta = \theta_2$ . We refer to  $p_0(\theta)$  as the prior distribution. The following are typical examples:

a. A reaction rate constant must be positive. Hence  $p_0(k) = 0$  for  $k \leq 0$ . If we have no further information on  $k$ , we set  $p_0(k) = c$  (a positive constant whose value is immaterial) for  $k > 0$ . Computational convenience often requires that an upper bound  $\alpha$  be prescribed for  $k$ . In this case  $p_0(k) = 0$  also for  $k > \alpha$ .

b. The ratio  $K$  of forward and reverse reaction constants may have been estimated as being  $K_0 \pm \sigma$  from measurements of equilibrium concentrations. It is then reasonable to use the normal prior distribution

$$p_0(K) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(K - K_0)^2}{2\sigma^2}\right]$$

When we come to estimate the parameters on the basis of the new data, we take the prior information into account by multiplying the likelihood function by the prior distribution. In terms of the logarithms, we maximize

$$S_B(\theta, \psi) = \log p[y - f(x, \theta), \psi] + \log p_0(\theta) \quad (26)$$

In case (a) above,  $\log p_0$  is negatively infinite outside the prescribed bounds. Maximizing Eq. (26) then consists of finding the maximum of  $\log p$  subject to the constraints, e.g.,  $0 \leq k \leq \alpha$ .

Sequential estimation, i.e., the reestimation of the parameters after the results of each in a series of experiments become available, is an important application of the Bayesian technique. Suppose after  $N$  experiments the parameters are estimated to be  $\theta = \theta_N$ , with covariance matrix  $V_N$  (the matrix is contained in the output of most parameter estimation programs). The relevant information contained in the data from these experiments may be summarized in the posterior distribution  $p_N(\theta) \approx c \exp[-\frac{1}{2}(\theta - \theta_N)^T V_N^{-1}(\theta - \theta_N)]$ , where  $c$  is an irrelevant constant. When results of subsequent experiments become available, the likelihood function is constructed from these alone, and  $p_N(\theta)$  is used as the prior distribution.

One cannot overstate the importance of using all available prior information, in the form of either constraints on the parameters or of prior densities. Use of such information frequently spells the difference between convergence and nonconvergence of the estimation procedure.

### III. METHODS OF NUMERICAL SOLUTION

In the preceding section the parameter estimation problem was formulated as that of finding those values of the parameters (possibly subject to constraints) which minimize (or maximize) a certain objective function, whose explicit form we have derived for a number of cases. We now describe a number of numerical methods which are suitable for finding the minimum [or maximum; maximizing  $F(x)$  is equivalent to minimizing  $-F(x)$ ] of the objective function. We restrict our attention mainly to unconstrained minimization, but where constraints exist, they may be incorporated in the objective function by methods discussed below. In the Appendix we detail certain computer programs which incorporate the features to be discussed.

### A. Direct-Search Methods

Methods for finding the minimum of  $S(\theta)$  that do not require the computation of the derivatives  $\partial S/\partial \theta$  are known as direct-search methods, in contrast to gradient methods, which require derivative evaluations. This distinction is not always clean-cut since gradient methods in which derivatives are computed by finite difference approximations may be regarded as direct-search methods. Among proper direct-search methods one may mention those due to Hooke and Jeeves [36], Rosenbrock [64], and Powell [61,62], the last reference being specific to least-squares problems. Box [15] reported particularly favorably on the two Powell methods. The results are based, however, on somewhat unrepresentative sample problems.

At present there exists no conclusive evidence for preferring one method, or set of methods, over another. Direct-search methods have the obvious advantage of not requiring differentiations, and they seem to perform well on well-conditioned problems. In difficult problems, however, precise knowledge of the derivatives can be crucial, and gradient methods tend to be more reliable. For this reason we omit any detailed description of the direct-search methods.

### B. Gradient Methods

The problem under investigation here is to find the minimum of  $S(\theta)$ . It may help the reader to visualize a set of mountains and the need to locate the lowest valley within these mountains. To solve the problem we proceed in an iterative sequence: given an initial value  $\theta_0$  of the parameters, we seek a new value of  $\theta_1$  which is nearer the minimum, in the sense that  $S(\theta_1) < S(\theta_0)$ . Once  $\theta_1$  has been obtained, we proceed to find  $\theta_2, \theta_3, \dots$ : each, in turn, having the property of being closer to the minimum. In the class of methods which have proved successful for parameter estimation, the formula used for finding the new value is

$$\theta_1 = \theta_0 - \lambda Rg \quad (27)$$

where  $\lambda$  is a scalar,  $R$  a matrix, and  $g$  the gradient vector of  $S$ , i.e.,  $g_i = \partial S/\partial \theta_i$ . Gradient methods differ from each other in the choice of  $R$  and of  $\lambda$  and we shall now examine both items. Before doing so, however, we see that  $R$  when it premultiplies the vector  $g$  twists  $g$  in vector space to produce a new vector;  $R$  thus determines the direction to go from  $\theta_0$ .  $\lambda$ , being a scalar, merely defines how far along this direction to go and determines the length of the step.

### 1. Choice of Direction

The choice  $R = I$  (the identity matrix), i.e.,  $\theta_1 = \theta_0 - \lambda g$ , constitutes the method of steepest descent. It converges very slowly in most practical problems.

The choice  $R = Q^{-1}$ , where  $Q$  is the Hessian matrix of  $S$  [i.e.,  $Q_{ij} = (\partial^2 S / \partial \theta_i \partial \theta_j)$ ] constitutes the Newton-Raphson method. It performs beautifully when one is near the minimum and possesses some desirable properties elsewhere (see Greenstadt [29]) but suffers from two major difficulties:

1. Except near the minimum, a step taken along the Newton-Raphson direction is not guaranteed to reduce  $S(\theta)$ , no matter what value is chosen for  $\lambda$ .

2. The method requires computation of the second derivatives of  $S$ , usually a laborious procedure.

These difficulties may be overcome in the following ways:

1. If  $R$  is positive-definite and  $g \neq 0$ , then one can show that  $S(\theta_1) < S(\theta_0)$  for sufficiently small  $\lambda$ . Unfortunately  $Q$  (and therefore  $Q^{-1}$ ) is not necessarily positive-definite away from the minimum. However, let  $\mu_1, \mu_2, \dots, \mu_p$  be the eigenvalues of  $Q$ , and let  $v_1, v_2, \dots, v_p$  be the corresponding eigenvectors. The inverse of  $Q$  may be computed from its spectral representation

$$Q^{-1} = \sum_{k=1}^p \mu_k^{-1} v_{ki} v_{kj}$$

where  $v_{ki}$  is the  $i$ th component of  $v_k$ . Greenstadt [24,29] has recommended that one define  $R$  by

$$R_{ij} = \sum_{k=1}^p |\mu_k^{-1}| v_{ki} v_{kj} \quad (28)$$

If some  $\mu_k = 0$ , we replace it by a small positive number. As defined by Eq. (28),  $R$  is positive-definite and coincides with  $Q^{-1}$  where the latter is also positive-definite. We thus retain all the advantages of the Newton-Raphson method, at the cost of having to compute the eigenvalues and vectors of  $Q$ , instead of simply solving the set of simultaneous linear equations  $Q\Delta\theta = -g$ . This cost is often a small one, since the computation of  $S(\theta)$  itself is the major time-consuming operation.

An alternative solution, suggested by Levenberg [53], Marquardt [54], and Goldfeld et al. [28] is to make  $R = (Q + \nu I)^{-1}$ , where  $I$  is the identity matrix. If  $\nu$  is greater in magnitude than the largest negative eigenvalue of  $Q$ , the matrix  $R$  will be positive-definite. The required



value of  $\nu$  may be determined by trial and error, without actually computing the eigenvalues of  $\mathbf{Q}$ .

2. It is fortunately possible in most parameter estimation problems to obtain a reasonable approximation to the second-derivative matrix  $\mathbf{Q}$  without computing any second derivatives of the model equations. We illustrate this point by means of the least-squares criterion, where

$$\begin{aligned} S(\theta) &= \sum_{\mu=1}^m \varepsilon_{\mu}^2 = \sum_{\mu=1}^m [y_{\mu} - f(x_{\mu}, \theta)]^2 \\ \frac{\partial S}{\partial \theta_i} &= -2 \sum_{\mu=1}^m \varepsilon_{\mu} \frac{\partial f_{\mu}}{\partial \theta_i} \quad [f_{\mu} = f(x_{\mu}, \theta)] \\ Q_{ij} &= \frac{\partial^2 S}{\partial \theta_i \partial \theta_j} = -2 \sum_{\mu=1}^m \varepsilon_{\mu} \frac{\partial^2 f_{\mu}}{\partial \theta_i \partial \theta_j} + 2 \sum_{\mu=1}^m \frac{\partial f_{\mu}}{\partial \theta_i} \frac{\partial f_{\mu}}{\partial \theta_j} \end{aligned} \quad (29)$$

If the fit of the model to the experimental data is at all good, the  $\varepsilon_{\mu}$  will be small, and the first term of Eq. (29) will be negligible compared to the second. Thus, we replace  $\mathbf{Q}$  by the approximation  $\mathbf{Q}^*$  (neglect second-order terms):

$$Q_{ij}^* = 2 \sum_{\mu=1}^m \frac{\partial f_{\mu}}{\partial \theta_i} \frac{\partial f_{\mu}}{\partial \theta_j} \quad (30)$$

Using  $\mathbf{R} = (\mathbf{Q}^*)^{-1}$  constitutes the Gauss-Newton method. An alternative interpretation of the method is to view it as replacing the model equations by their tangents [i.e., neglecting  $(\partial^2 f_{\mu} / \partial \theta_i \partial \theta_j)$ ] and solve the resultant linear regression problem to obtain the starting point for the next iteration.

The same treatment can be applied to weighted least squares and to more general maximum likelihood problems. In all cases discussed in the previous section, it can be shown that in the expression for  $(\partial^2 S / \partial \theta_i \partial \theta_j)$ , the quantities  $(\partial^2 f_{\mu k} / \partial \theta_i \partial \theta_j)$  are always multiplied by some residual  $\varepsilon_{\mu l}$  and may therefore be neglected. Thus the first derivatives of the model functions suffice to determine an approximation  $\mathbf{Q}^*$  for the matrix  $\mathbf{Q}$ .

An alternative approach to approximating  $\mathbf{Q}$  without calculating second derivatives is contained in the Davidon-Fletcher-Powell [20,25] method. We start with an arbitrary symmetric positive-definite matrix  $\mathbf{R}_0$ . Let  $\theta_i$  and  $\mathbf{g}_i$  denote, respectively, the values of the parameters and of the gradient of  $S$  at the  $i$ th iteration. Then  $\theta_{i+1} = \theta_i - \lambda_i \mathbf{R}_i \mathbf{g}_i$ , with  $\lambda_i$  chosen so as to minimize  $S(\theta_{i+1})$  along the chosen direction. Let  $\sigma_i = \theta_{i+1} - \theta_i$  and  $\gamma_i = \mathbf{g}_{i+1} - \mathbf{g}_i$ . Then

$$R_{i+1} = R_i - \frac{R_i \gamma_i \gamma_i^T R_i}{\gamma_i^T R_i \gamma_i} + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \gamma_i} \quad (31)$$

For a well-behaved function  $S(\theta)$ , the sequence of matrices  $R_i$  ( $i = 0, 1, 2, \dots$ ) converges to  $Q^{-1}$ . The precise conditions which  $S(\theta)$  must satisfy for the above statement to be true are not known. However, it can be shown that in principle the  $R_i$  are positive-definite as long as no  $\lambda_i$  vanishes. Numerical difficulties, however, require that the initial matrix  $R_0$  approximate  $Q^{-1}$  at least in the magnitude of its diagonal terms.

In tests on various kinetics models, Bard [2] has found the generalized Gauss-Newton method to be somewhat more efficient than its competitors.

## 2. Choice of Step Length

In Davidon's method,  $\lambda$  is chosen so as to minimize  $S(\theta)$  along the chosen direction. However, experience has shown that the accuracy with which  $\lambda$  needs be determined is to about one part in a thousand.

In the original Newton-Raphson method,  $\lambda = 1$ . This is usually not a satisfactory procedure, except very near the minimum. In Marquardt's method also,  $\lambda = 1$  is used. If it turns out that  $S(\theta_1) \geq S(\theta_0)$ , the value of  $\nu$  is increased and  $\theta_1$  is recomputed, until  $S(\theta_1) < S(\theta_0)$ .

In the Gauss-Newton method we initially set  $\lambda = 1$  (if there are constraints, a smaller initial value may be required to guarantee that  $\theta_1$  is in the feasible region). If  $S(\theta_0 - \lambda Rg) \geq S(\theta_0)$ , we choose a smaller value of  $\lambda$  [e.g.,  $\lambda/2$ , or the value that would minimize a parabolic approximation to  $S(\theta)$  based on its computed values at  $\lambda = 0$  and  $\lambda = 1$ , and on its gradient at  $\lambda = 0$ ]. We repeat until  $S(\theta_0 - \lambda Rg) < S(\theta_0)$ , which will always hold for sufficiently small  $\lambda$ . In the Newton-like methods discussed above (Davidon's method excepted), it seems that an extensive search for the value of  $\lambda$  that minimizes  $S(\theta_0 - \lambda Rg)$  is not justified. It rarely pays to test more than one additional value of  $\lambda$  beyond the first one to give an improvement.

## 3. Initiation

Equation (27), combined with the choices of  $R$  and  $\lambda$ , defines the iterative procedure. To start the procedure, however, we need initial estimates or guesses for  $\theta_0$ . The success of the minimization procedure often depends on these guesses, and it is clearly advan-

tageous to make these guesses as close to the true values (the minimum) as possible. Several methods for arriving at reasonable initial guesses have been summarized by Kittrell, et al. [45]. Among these are:

1. Use all available prior information.
2. Use the results of linear -least-squares estimates based on linearized or rearranged equations.
3. Fix the values of some of the parameters (e.g., those on which most prior information is available) and estimate only the other parameters. Use these results as initial guesses for estimating all parameters simultaneously.
4. Compute the objective function on a sparse grid of parameter values. Use the grid point with optimal objective function value as the initial guess.
5. Use an analog computer to simulate the reaction. Search for the optimum by turning the analog elements which set the parameter values. This method may also be combined with the grid search described above.

If the number of unknown parameters is too large to make a grid search feasible, it is possible to conduct a random search instead.

None of these methods are infallible, and which one will work best in any given situation can be determined only by experience.

#### 4. Termination and Convergence

Once the iterative procedure has been started, it will continue to run on the computer until some termination criterion is satisfied. At the minimum of  $S$ , the gradient  $g$  should vanish. Due to round-off errors, however, the condition  $g = 0$  can never be attained precisely, and cannot be used as a termination criterion. In practice it seems preferable to terminate whenever the iterative procedure ceases to cause significant changes in the  $\theta$ 's, i.e., (as suggested by Marquardt [54]) when

$$|\theta_1 - \theta_0|(\epsilon_1 + |\theta_0|)^{-1} \leq \epsilon_2$$

for all elements of  $\theta$ , where  $\epsilon_1$  and  $\epsilon_2$  are predetermined vectors with small elements. It will be found that frequently even very close to the minimum the gradient still has large elements. A good test of stationarity is then given by the dimensionless quantities  $|g_i/\theta_i Q_{ii}|$ , which are all required to be small compared to unity. If the matrix  $Q$  is available, positivity of all its eigenvalues indicates that the solution (if stationary) is indeed a (local) minimum. There is usually no way of proving that a global minimum has been reached.

Failure to converge to a stationary point is usually due to one of the following reasons:

1. A constraint has been reached. If the constraint is one about which there is no doubt (e.g., the condition that a rate constant must be positive), it is likely that the mechanism chosen is inappropriate.
2. The matrix  $R$  is not, or is insufficiently, positive-definite. We have described above how this problem may be overcome.
3. The termination criterion is not stringent enough. Vectors of  $\epsilon_1$  and  $\epsilon_2$  with elements of  $10^{-3}$  and  $10^{-4}$ , respectively, have worked well in many problems but may fail with others.
4. The gradient is computed with insufficient precision. Whenever possible, gradients should be computed by analytic differentiation, rather than by means of finite difference approximations.

We note that in kinetics analysis the model equations are generally given in the form of solutions of differential equations, e.g.,

$$\frac{d\eta_i}{dt} = g_i(\eta, \theta) \quad i = 1, 2, \dots, n \quad (32)$$

The question then is how to obtain analytic derivatives of these solutions. Differentiating Eq. (32) with respect to  $\theta_j$ , we have, using the chain rule,

$$\frac{\partial}{\partial \theta_j} \frac{d\eta_i}{dt} = \sum_{k=1}^n \frac{\partial g_i}{\partial \eta_k} \frac{\partial \eta_k}{\partial \theta_j} + \frac{\partial g_i}{\partial \theta_j}$$

Interchanging the orders of differentiation yields

$$\frac{d}{dt} \frac{\partial \eta_i}{\partial \theta_j} = \sum_{k=1}^n \frac{\partial g_i}{\partial \eta_k} \left( \frac{\partial \eta_k}{\partial \theta_j} \right) + \frac{\partial g_i}{\partial \theta_j} \quad (33)$$

Equation (33) is a set of linear ordinary differential equations in the unknown functions  $\partial \eta_k / \partial \theta_j$  ( $k = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, p$ ). These may be integrated numerically alongside the original Eqs. (32) to obtain the desired values of  $\partial \eta_k / \partial \theta_j$  at any time  $t$ . Although one can scarcely call these "analytic derivatives," they will be determined to the same accuracy as the  $\eta_i$  themselves, and that is the best one can hope for. The alternative procedure is to integrate Eqs. (32) for slightly perturbed values of the  $\theta$ 's, and estimate the  $\partial \eta_k / \partial \theta_j$  by finite differences. The total number of integrations performed is the same in both methods, with the "analytic" one usually providing greater accuracy.

### C. Constraints

To handle constraints within the framework of unconstrained minimization, it is necessary to adjust the objective function in such a

way that one pays a penalty whenever one comes near to violating the constraint. Carroll's method [16] uses the following device: Let  $S(\theta)$  be the function whose minimum is to be found subject to a set of constraints  $g_i(\theta) \geq 0$ ,  $i = 1, 2, \dots$ . We introduce a new objective function:

$$S^*(\theta) = S(\theta) + \sum_i \frac{\alpha_i}{g_i(\theta)} \quad (34)$$

where the  $\alpha_i$  are suitably chosen small positive constants. As one approaches say the  $j$ th constraint,  $g_j(\theta)$  approaches zero, and  $S^*(\theta)$  increases beyond bounds. At a point far from any constraint, the function  $S^*(\theta)$  differs but little from  $S(\theta)$ , and the minima of the two should nearly coincide. After finding the minimum of  $S^*(\theta)$ , one may reduce the  $\alpha_i$  by, e.g., a factor of 10, find the minimum of the new  $S^*(\theta)$ , and repeat the procedure until the contribution of the "penalty function"

$$\sum_i \frac{\alpha_i}{g_i(\theta)}$$

becomes negligible. If the minimum of  $S(\theta)$  is actually on a constraint, this procedure may approach the true minimum as closely as one wishes.

In many special cases the constraints may be handled by other methods. For example, if it is required that  $\theta_1 \geq 0$ , we may substitute  $\psi^2$  for  $\theta_1$  and minimize with respect to the unconstrained variable  $\psi$ . As shown by Box [15], similar substitutions are possible in many cases but often require considerable ingenuity.

#### D. Interpretation of Parameter Estimates

The best estimated values of the parameters are, in themselves, actually of little use. It is essential to know not only what the estimates are, but also, and more importantly, how reliable they are. The observations are random variables, and hence the estimates which are computed from them are also random variables; thus it is meaningful to try and estimate their probability distribution. Fortunately it turns out that the distribution usually approaches the normal as the number of observations is increased. The means of the estimated distribution constitute the estimated values of the parameters; the covariance matrix of the distribution is a measure of the reliability of the estimates. This matrix expresses the manner in which variations in the observations would affect the parameter estimates.

The alternative approach to evaluating the reliability is by exploring the dependency of the objective function on the parameters. If the objective function is little affected by changes in a certain parameter, one would have doubts concerning its value. Since at the final estimates the objective function has a minimum, its gradient vanishes and the effect of the parameters on the objective function is summarized in the second derivative matrix. It turns out that these two measures of reliability are equivalent; in fact if  $V$  is the covariance matrix of the estimates, then

$$[V^{-1}]_{ij} \approx - \frac{\partial^2 \log L}{\partial \theta_i \partial \theta_j}$$

where  $L$  is the likelihood function.

Most parameter estimation programs print out estimates of the matrix  $V$ . The diagonal elements of  $V$  are the variances, and their square roots are the standard deviations of the parameter estimates. Off-diagonal elements indicate the interdependence of the estimates of the various parameters. It is convenient to eliminate these dependencies by finding those linear combinations of the parameters which are statistically independent. These are known as principal components. Let  $\mu_i$  be an eigenvalue of  $V$  with eigenvector  $v_i$ , whose components are  $v_{ij}$ . Then the estimate of the quantity  $\psi_i = \sum v_{ij} \theta_j$  has variance  $\mu_i$ , and the estimates of the different  $\psi_i$  are independent. The  $\psi_i$  are, then, the desired principal components. Examination of the  $\mu_i$  will reveal which linear combinations of the parameters are well determined (small  $\mu_i$ ) and which are doubtful in value (large  $\mu_i$ ).

Once again we point out that the Appendix contains a discussion of available computer programs which have some or all the features discussed in this section.

#### IV. IMPROVING PARAMETER ESTIMATES BY PROPER EXPERIMENTAL DESIGN

Examination of the posterior distribution (i.e., of the covariance matrix of the parameter estimates) will sometimes reveal that some of the parameters, or some linear combinations of the parameters, are ill determined. We may distinguish three causes for such an occurrence:

1. The model chosen to fit the data is inappropriate. This will be marked by the appearance of large systematic deviations (residuals) of the experimentally measured from the predicted values of the observed variables. The obvious remedy is to modify the model. Those

terms whose parameters are the most ill determined should be prime targets for elimination or modification. The nature of the deviations of the residuals may also hold clues as to how the model should be modified. Hunter and co-workers [38, 39, 44] have provided an excellent example of how residual analysis can lead to a systematic modification of the model.

2. The measurement precision is low. This will be characterized by large random residuals. If no improvement of the measurement techniques is feasible, the only remedy is to make more measurements. Unfortunately a tenfold increase in precision may require a 100-fold increase in the number of experiments. It should be noted, however, that the attainable precision in the estimates for a given number of experiments is maximized when the experimental conditions are chosen properly. This point is discussed below.

3. The experiments were not properly designed. This is the conclusion that must be reached if some parameters have a large variance, even though the overall fit is good, i.e., even though the residuals are small. We cite two examples of how this may happen:

a. Suppose species A is supposed to decompose, in two parallel reactions, to species B and C, with rate constants  $k_1$  and  $k_2$ , respectively. If measurements on the concentrations of A alone are available, it is clearly possible to determine  $k_1 + k_2$ , but not  $k_1$  and  $k_2$  individually. While this example may appear trivial, similar effects may arise and be less obvious in more complicated situations.

b. Suppose the model equation is

$$y_{\mu} = \theta_1 x_{\mu 1} + \theta_2 x_{\mu 2}$$

and in all experiments it happened that  $x_{\mu 1} \approx x_{\mu 2}$ . Then the equation

$$y_{\mu} = (\theta_1 + \theta_2)x_{\mu 1} = \theta_3 x_{\mu 1} \quad (35)$$

would represent the data just as well as the original equation. On this basis it is impossible to estimate  $\theta_1$  and  $\theta_2$  individually, but rather only as their sum.

In both these cases proper planning of the experiments would have eliminated the difficulties: in the first case, by measuring concentrations of B and/or C; in the second case, by varying  $x_{\mu 1}$  and  $x_{\mu 2}$  independently.

When we attempt to systematize the selection of appropriate experiments, we are naturally led into the realm of information theory. This follows from the fact that the purpose of an experiment is to gain information. In the case of parameter estimation the relevant information is contained in the posterior distribution of the parameters. When one tries to formalize intuitive notions concerning the amount of information contained in a given distribution (e.g., a dis-



tribution with a small variance contains more information concerning the value of a parameter than one with a large variance) one is led [1] to the following formula:

$$I = E(\log p) = \int p \log p \, d\theta \quad (36)$$

where  $I$  is the measure of information,  $p$  is the probability density function, and  $E$  denotes the expected value. If  $p$  is a multivariate normal distribution with covariance matrix  $V$ , then the information is given by

$$I = C - \frac{1}{2} \log(\det V) \quad (37)$$

where  $C$  is a constant. To derive the maximum amount of relevant information, we must plan our experiments so as to minimize the value of  $(\det V)$ , where  $V$  is the expected variance of the posterior distribution.

Suppose, now, that we wish to plan a series of, e.g.,  $m$  experiments. Our current information on the values of the parameters is summarized in a prior distribution  $p_0(\theta)$ , e.g., a normal distribution with covariance matrix  $V_0$ . In many cases,  $p_0(\theta)$  will actually be the posterior distribution obtained by estimating the parameters on the basis of experiments conducted to date. Then it follows from Eq. (37) that the expected covariance of the posterior distribution is given approximately by

$$[V^{-1}]_{ij} \approx [V_0^{-1}]_{ij} + \sum_{\mu=1}^m \sum_{k,l} \frac{\partial f_{\mu k}}{\partial \theta_i} \frac{\partial f_{\mu l}}{\partial \theta_j} [U^{-1}]_{kl} \quad (38)$$

where  $U$  is the covariance matrix of the observations for each experiment. The derivatives  $\partial f_{\mu k} / \partial \theta_i$  are to be evaluated at the current estimates of the  $\theta$ 's. The simplest case occurs when there is only one observed variable per experiment. Then the subscripts  $k$  and  $l$  may be omitted, and the matrix  $U$  becomes a single number  $\sigma^2$ , where  $\sigma$  is the standard deviation of the measurements. Then Eq. (38) becomes

$$[V^{-1}]_{ij} = [V_0^{-1}]_{ij} + \frac{1}{\sigma^2} \sum_{\mu=1}^m \frac{\partial f_{\mu}}{\partial \theta_i} \frac{\partial f_{\mu}}{\partial \theta_j} \quad (39)$$

This formula was derived by Draper and Hunter [22].

Equation (38) or (39) reveals that  $(\det V)$  is a very complicated function of the  $x_{\mu i}$ , i.e., of the experimental conditions proposed for

the  $m$  desired experiments. Proper design of these experiments requires that one select that set of feasible experimental conditions  $x_{\mu i}$ ,  $\mu = 1, 2, \dots, m$ , that minimizes  $(\det V)$  [or equivalently maximizes  $(\det V^{-1})$ ].

Most of the work in this field to date is contained in a series of papers by Box, Hunter, Kittrell, and their co-workers at the University of Wisconsin [12,14,21,47]. They derive the relevant formulas from several alternative points of view and apply them to several computer-simulated chemical reaction models. The problem of finding the maximum of the information function is complicated by the fact that it usually possesses multiple local maxima and that its derivatives are usually too complicated to calculate. Considerable additional work in the field is required.

#### V. COMPUTATIONAL RESULTS IN PARAMETER ESTIMATION

In the previous sections we have developed the conceptual ideas associated with nonlinear least squares as applied to kinetic models. The present section will outline some of the important results which have been obtained to date in this area. The depth of coverage is minimal but, it is hoped, sufficient to highlight the various points.

Some of the earlier work in the area of parameter estimation in kinetic systems is due to such authors as Box [8], Box and Coutie [9], Blakemore and Hoerl [5], Cull and Brenner [18], Hartley [33], Marquardt [54], Peterson [57,58] and Rubin [65]. In particular we should comment on the pioneering works by Box and by Peterson. Both these authors pointed out the positive and negative aspects of nonlinear parameter estimation and the important advantages associated with the use of a digital computer for the analysis. While Peterson's work was directed toward the computer aspects, Box derived and discussed in detail all the statistical features and developments. In these two papers are contained many of the basic ideas for work carried out to date in the nonlinear estimation area.

Blakemore and Hoerl analyzed the famous or infamous hydrogenation of codimer originally analyzed by Hougen and Watson [37] via linear least squares (for a further discussion see a later part of this paper). For this system, at least 20 alternative models may be postulated as representing the rate-determining step. These authors showed that it was impossible to select any one model as being the best model of the entire set. The need for further and more extensive experimental data was conclusively shown, thus pointing out that once a fixed set of experimental data is available the most sophisticated analysis of this data may be inadequate to isolate a single

model. Cull and Brenner investigated hexane isomerization kinetics via the nonlinear approach. In this case they were able to conclude that one of their postulated reaction steps was the rate-controlling steps within the level of precision of the data.

In a slightly more recent paper Freeh et al. [27] analyzed the hydrogenation reaction in a stirred reactor. By selecting that model whose standard deviation of the fit to experimental data was lowest, they were able to isolate a single model as representative of the reaction mechanism.

In the last few years much of the nonlinear parameter estimation in kinetics systems has been carried out by two groups of researchers; the first group includes Lapidus and Peterson [52,59] and the second group includes Kittrell, Hunter, Mezaki, and Watson [42-49]. In addition, nonlinear estimation methods have been applied in other allied areas. Thus the work of Heineken et al. [34] and Bellman et al. [4] deal with an analysis of the kinetics of biological reactions via some of the computer methods discussed here (see also Refs. [17] and [63] for the adaptive control of a batch reactor).

The work of Peterson [57,58] is worth discussing in some detail since it represents a concrete case where nonlinear estimation was able to provide new insight into a kinetics mechanism. The integral conversion data of both D'Alessandro and Farkas [19] and Ioffe and Sherman [40] on the vapor-phase homogeneous and noncatalytic oxidation of naphthalene were analyzed. The first study concerned naphthalene depletion in which the reaction kinetics were free of extraneous effects; the second study treated the mechanism of the complete reaction, including reaction products, and revealed certain complicating factors which influence the behavior of the reaction.

D'Alessandro and Farkas obtained data for the vapor-phase oxidation of naphthalene in the presence of a catalyst consisting of vanadium pentoxide in a flow reactor. Measurements were made directly on total anhydride, maleic anhydride, 1,4-naphthoquinone, and off-gases. Phthalic anhydride was calculated by difference between total and maleic anhydrides. Residual naphthalene in the produce stream was determined by overall difference. Thus, concentrations of naphthalene (N), phthalic anhydride (P), maleic anhydride (M), naphthoquinone (Q), and off-gases (G) constituted the dependent variables.

Since no measurement errors were given, equal statistical weights were assigned to the observed values of concentration.\* Data were recorded for a given temperature at various reaction times ranging from 0.020-2.44 sec. Temperatures were also explored from 340-

\*The availability of maximum likelihood computer programs would now make such an assignment unnecessary.

475°C. Six observations were made on each chemical species at 340°C, five at 375°C, six at 410°C, four at 450°C, and three at 475°C. In all, 24 observations were made on each chemical species.

Naphthalene depletion behavior was analyzed by examining each set of data first at constant temperature to determine the specific reaction rate at that temperature; second, the kinetics were examined for temperature dependency and for the order of the reaction. The typical rate equation for the second case is given by

$$\frac{d[N]}{dt} = -k_0[N]^\alpha \exp[-E/RT] \quad (40)$$

which when integrated yields for  $\alpha \neq 1$ :

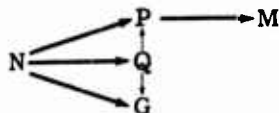
$$\left[ \frac{[N_0]^{1-\alpha}}{1-\alpha} \right] \left[ 1 - \left( \frac{[N]}{[N_0]} \right)^{1-\alpha} \right] = kt \quad (41)$$

Typical results are shown in Table 1, indicating satisfactory agreement between the nonlinear estimation analysis and that carried out by D'Alessandro and Farkas; this agreement holds for each temperature as well as over the entire temperature range.

TABLE 1  
Constant-Temperature Kinetics

Temp., °C	$\hat{k}$ , sec <sup>-1</sup>	
	Nonlinear estimation	D'Alessandro and Farkas
340	0.52	0.56
375	1.84	1.70
410	7.40	7.36
450	21.3	20.4
475	34.8	32.9

The second phase of this study dealing with an analysis of the overall reaction mechanism was next initiated. The mechanism considered most plausible by D'Alessandro and Farkas was given by



To analyze this and possibly other mechanisms, nonlinear estimation was applied to the experimental data by proceeding from simple mechanisms to increasingly more complex ones. The purpose of this cautious approach was to treat mechanisms with few parameters at the outset which could then be augmented by additional parameters as the analysis continued. In this way control could be exercised to avoid difficulties in convergence of the nonlinear estimation technique for those parameter estimates which were poorly determined. The complete analysis is shown in Table 2, where the circled species indicate that measurements on these species were taken into account over the whole temperature range. In Table 2  $k_0'$ , defined as  $k_0' = k_0 \exp\{-E/RT\}$ , was evaluated with  $1/\bar{T} = 1.42 \times 10^{-3} \text{ }^\circ\text{K}^{-1}$ . Furthermore, since values of  $[N]$  were derived by differences, these values are not independent of other measured concentrations. As a result "observations" on  $[N]$  were excluded. In all, 96 observations were used resulting from the 24 observations made on each of the four chemical species.

Several conclusions can be drawn from these tabulated results. First and foremost is that a slightly abbreviated form of the proposed mechanism is an adequate representation of the data:



The overall fit as measured by the standard deviation of residuals  $s$  is satisfactory when compared to values of  $s$  for other mechanisms. Also, the least-squares parameter estimates  $\theta$  obtained appear fairly well determined as indicated by the values of the associated standard deviations of the parameter estimates  $s_{\hat{\theta}}$ .

These results were confirmed further by considering a number of alternative mechanisms including the originally proposed one and one suggested by Mars and van Krevelen [56] in which the rate-determining step is the chemisorption of oxygen. Without presenting the explicit data here, we merely state that these additional calculations showed that the mechanism proposed by D'Alessandro and Farkas was an adequate representation of the experimental data; furthermore, the reaction rates and activation energies for the primary branching of naphthalene to phthalic anhydride and naphthoquinone were not equal, particularly at the higher temperatures, in contrast to the values determined by D'Alessandro and Farkas which were equal over the temperature range. However, it was also shown that an equally acceptable model is that proposed by Ioffe and Sherman

TABLE 2

Mechanism	$k_0' = k_0 \exp(-E/RT)$					
	1	2	3	4	5	6
	$\hat{\theta}$	8.79	0.842			
	$s_{\hat{\theta}}$	0.26	0.163			
	$\hat{\theta}$	9.23	0.873	0.833		
	$s_{\hat{\theta}}$	0.26	0.146	0.151		
	$\hat{\theta}$	8.10	2.16	0.882	18.8	
	$s_{\hat{\theta}}$	0.61	0.59	0.142	7.8	
	$\hat{\theta}$	8.10	3.00	0.881	28.0	0.831
	$s_{\hat{\theta}}$	0.73	0.74	0.124	8.9	0.221
	$\hat{\theta}$	8.26	3.55	0.926	31.3	0.750
	$s_{\hat{\theta}}$	0.88	1.11	0.142	10.4	0.236
	$\hat{\theta}$	8.09	2.76	1.20	29.2	0.861
	$s_{\hat{\theta}}$	0.72	0.74	0.48	9.0	0.227
						6.10

[40] in which naphthalene undergoes direct and simultaneous oxidation to phthalic anhydride and naphthoquinone, but not to evolved gases, i.e., involving a binary branching



Ioffe and Sherman similarly studied the vapor-phase oxidation of naphthalene in the presence of vanadium oxide catalyst. For fixed

## Mechanism Evaluation

E' = E/R							
1	2	3	4	5	6	$\alpha$	s
14.5	12.4						0.0326
0.4	2.8						
14.6	12.7	9.68					0.0278
0.4	2.4	2.34					
15.0	11.9	9.80	7.41				0.0259
1.0	3.2	2.09	5.60				
15.6	13.0	8.86	9.69	16.8			0.0217
1.1	2.6	3.96	1.80	3.5			
15.9	13.5	9.93	17.6	9.05		1.13	0.0218
1.3	2.8	1.85	3.9	3.8		0.15	
15.6	12.8	11.1	16.6	8.84	10.9		0.0219
1.1	2.7	3.2	3.6	3.90	12.1		

initial concentrations of naphthalene and air of  $2.69 \times 10^{-4}$  and  $8.75 \times 10^{-3}$  moles/liter, respectively, extensive integral conversion data were obtained in a flow reactor under constant volume and temperature conditions over a temperature range of 260-400°C and reaction times of 0.047-0.499 sec.

Measured were conversions to phthalic anhydride (P), 1,4-naphthoquinone (Q), maleic anhydride (M), and off-gases (G). Residual naphthalene (N) was obtained by difference. For each chemical species, four to seven observations were recorded at each of nine tem-



peratures. In all, 55 measurements were obtained for each of the chemical species. However, maleic anhydride did not appear in measurable quantities except at temperatures of 380 and 400°C. Because of omissions in the data of Ioffe and Sherman, some adjustments were necessary in this study at the higher temperatures.

Constant-temperature and then temperature-dependent effects were analyzed by nonlinear estimation with the order of the reaction both prescribed as first order and left unassigned to be estimated from the data. The results obtained were in substantial agreement with the proposals of Ioffe and Sherman, the kinetics being plausibly first order.

Proceeding further, Ioffe and Sherman suggested that the activation energy is not constant with temperature, but decreases from an initial value of 27.4 kcal/mole as the temperature rises and that while first-order kinetics dominated at low temperatures, internal pore diffusion became increasingly important until at about 350°C half-order kinetics prevailed. To test this assertion, data from proximate temperature intervals were examined by nonlinear estimation for both first-order and unassigned-order kinetics. The results are shown in Table 3.

TABLE 3  
Temperature-Dependent Kinetics,  
Proximate Temperature Intervals

Temp. range, °C	First-order kinetics, E, kcal/mole	Unassigned-order kinetics				
		$\hat{E}$	$s_{\hat{E}}$	$\hat{\alpha}$	$s_{\hat{\alpha}}$	s
260-270	28.2	26.6	3.6	0.77	0.24	0.023
270-290	21.6	19.7	1.8	0.59	0.12	0.030
290-310	11.2	10.9	2.8	0.91	0.13	0.046
310-330	17.4	17.4	2.7	1.15	0.12	0.039
330-350	14.0	14.0	3.2	1.06	0.16	0.042
350-360	13.2	12.7	4.9	0.85	0.12	0.029
360-380	12.8	11.4	3.7	0.76	0.15	0.040
380-400	-1.9	-0.8	1.1	0.32	0.16	0.019

Estimates of the activation energy confirm the systematic trend with temperature. The results also show that a significant change occurs in the vicinity of 400°C. Pinchbeck [60], using a similar catalyst in a fluidized bed, observed this effect at about the same temperature and attributed the departure to a modification of the mechanism of reaction. Usbakova et al. [68] demonstrated that the mech-

anism is dependent on the valence state of vanadium and that the state of the oxide is not uniform throughout the catalyst. Tandy [67], in an investigation of catalysts for the oxidation of sulfur dioxide to trioxide, reported an approximate melting point of 400°C for mixtures of vanadium oxide-potassium sulfate. The behavior at this elevated temperature is therefore presumed to be related to a marked alteration in the physical character of the catalyst.

Except at the highest temperature, estimates of the order of reaction do not confirm the half-order kinetics established by Ioffe and Sherman, but rather support largely first-order kinetics at both low and high temperatures. However, the activation energy decreases rapidly with temperature to about one-half the initial value. As noted by Wheeler [70], this behavior is consistent with diffusion occurring in the pores of the catalyst. Because internal pore diffusion is present, systematic deviations are evident at the lowest temperature, but the anomaly at the highest temperature is obscured by the comparatively few observations made.

As before, the next step in the analysis was consideration of the overall reaction mechanism. Nonlinear estimation analysis was carried out as previously, starting with simpler mechanisms and terminating with the most complex mechanism the data could support. Also, as in the previous analysis, "observations" on [N] were excluded, yielding a total number of 220 observations, with 55 measurements made on each of four species.  $k_0'$  was evaluated at  $1/\bar{T} = 1.67 \times 10^{-3} \text{ }^\circ\text{K}^{-1}$ .

A complete tabulation of results is given in Table 4 where  $s\hat{\theta}$  are omitted for all but the terminal mechanisms. The temperature range was initially largely restricted to 260-350°C until maleic anhydride was introduced as a chemical species, at which point the temperature range was extended to 260-400°C. These results indicate that the mechanism proposed by Ioffe and Sherman [Eq. (43)] is plausible. As shown previously there is the possibility that other mechanisms such as Eq. (42) will also fit the data. The results for the two mechanisms, while not shown here, confirm that each fits the data equally well.

Using these results and others obtained by Peterson, the following conclusions can be drawn from the nonlinear estimation analysis: (1) Internal pore diffusion is a plausible explanation of the variation of activation energy with temperature, although contrary to the analysis of Ioffe and Sherman, first-order kinetics largely dominate throughout the temperature range. (2) The mechanism proposed is an adequate representation of the data, but equally acceptable is that proposed by D'Alessandro and Farkas in which naphthalene also undergoes simultaneous oxidation directly to evolved gases. (3) The

TABLE 4

Mechanism	Temp. range, °C	$\hat{\theta}$	$k_f$			
			1	2	3	4
	260-400	$\hat{\theta}$	3.84	0.150		
	260-400	$\hat{\theta}$			2.96	8.65
	260-350	$\hat{\theta}$	5.08		1.47	
	260-350	$\hat{\theta}$	4.40		3.37	3.44
	260-350	$\hat{\theta}$	4.32		3.79	0.73
	260-350	$\hat{\theta}$	4.37		3.86	0.94
	260-400	$\hat{\theta}$ $s\hat{\theta}$	4.03 0.19	0.014 0.013	3.15 0.21	0.98 0.18
	260-400	$\hat{\theta}$ $s\hat{\theta}$	4.01 0.27	0.012 0.012	3.02 0.30	1.02 0.19

fit of the data for 1,4-naphthoquinone is suspect at temperatures in the vicinity of 350°C and suggests the presence of a closely related reaction product, 1,2-naphthoquinone, undetermined by Ioffe and Sherman, but later reported by Shelstad [66].

The above has been concerned with homogeneous kinetic studies; by contrast, Lapidus and Peterson [52, 59] considered the heterogeneous case in which Langmuir-Hinshelwood relations were used to describe the rate-controlling mechanism. Two studies were under-

## Mechanism Evaluation

						E'			
5	1	2	3	4	5	$\alpha$	s		
	7.27	10.7							0.062
			6.26	6.62					0.045
	8.81		4.98						0.067
	9.78		7.72		3.21				0.201
3.52	9.92		7.98	8.54	2.80				0.042
3.45	9.83		7.83	4.18	3.02				0.036
3.51	8.41	22.7	6.34	5.79	4.01				0.042
0.45	0.40	5.3	0.58	1.62	1.36				
3.28	8.14	23.4	6.45	5.41	4.69	0.96			0.042
0.48	0.47	5.6	0.60	1.61	1.40	0.11			

taken, both relating to ethanol. In the first study the experimental data of Kabel [41] on ethanol dehydration were analyzed; in the second study the data of Franckaerts and Froment [26] on the dehydrogenation of ethanol were analyzed. In both cases extensive integral conversion data were available.

In the first study four different models were used to represent the experimental data. Those included three heterogeneous models and one homogeneous model. In the nonlinear estimation analysis,

a preliminary screening of data was made of various mechanisms on the basis of obtaining random residuals and a suitable residual error variance. Several subsets of the original data were identified for further study. A homogeneous-type rate mechanism was found to be as adequate in describing the data as the heterogeneous mechanisms. Using only heterogeneous mechanisms, it proved impossible to conclude that any one mechanism was more suitable than another unless external information was also used. This conclusion, while not unknown to researchers in kinetics and mechanism studies, may seem rather surprising since a large amount of data was available for estimating relatively few parameters.

It was thus concluded that no discrimination among mechanisms was possible, presumably as a result of either or both of these effects: (1) The experimental data examined were constrained to be at only one constant pressure. (2) An accidental arrangement of the physical parameters led to a degeneracy in the rate equations for various heterogeneous mechanisms, which therefore became equivalent to the homogeneous model.

In the second study, nonlinear estimation led to substantially the same results as the initial rate method combined with linear estimation. Since fewer assumptions and transformations were involved in the analysis by nonlinear estimation, the results are possibly more valid.

It has also been shown that data at different pressures as well as different feed compositions measurably improve the ability to discriminate among mechanisms and provide suitable estimates for parameters. Constant-pressure data alone do not appear to be adequate for the more definitive study of kinetics and mechanism.

In a most interesting and important series of papers on kinetics modeling, Kittrell et al. [42-48] first analyzed the isothermal and nonisothermal heterogeneous reduction of nitric oxide via Langmuir-Hinshelwood models. Three different models were postulated and then approximated to the experimental data by linear and by nonlinear least squares. Comparison of the two approaches showed that nonlinear least squares was more useful for a rational selection of an acceptable single model and estimation of its parameters. This was particularly true if the rate equations determined by each procedure were extrapolated beyond the actual regions of the experimental data; here the two approaches yielded rate curves which differed quite widely. Also, these authors indicated explicitly the need for further data to reduce the confidence regions for the estimated parameters and make the analysis even more efficient.

In a summary paper [49] these same authors surveyed model building techniques in general with a primary emphasis on mecha-

nistic models. This work represents a convenient outline of problems involved in model representation of kinetic systems, the methods of solution, and the results obtained.

#### VI. DESIGN OF EXPERIMENTS FOR MODEL DISCRIMINATION

It was stated above that our efforts to estimate parameters for a model may be defeated by data obtained from poorly designed experiments. The same statement applies with even greater force to the problem of selecting the best model from among a set of candidates.

As was the case with parameter estimation, it is the experimenter's task at each stage to seek out that experiment which is likely to yield the greatest amount of relevant information. For the purpose of discriminating between two proposed models, a measure of the relevant information is (see Kullback [50])

$$I = E^{(1)} \log \frac{p^{(1)}}{p^{(2)}} + E^{(2)} \log \frac{p^{(2)}}{p^{(1)}} \quad (44)$$

where  $E^{(i)}$  denotes expectation under the assumption that model  $i$  is true, and  $p^{(i)}$  is the probability density function under the same assumption ( $i = 1, 2$ ).

Suppose we have performed  $N-1$  experiments and have fitted the data to both models 1 and 2, yielding parameter sets  $\theta^{(1)}$  and  $\theta^{(2)}$ . Let  $x_N$  represent the values of the independent variables for a proposed  $N$ th experiment. We can compute the predicted values  $y_N^{(1)}$  and  $y_N^{(2)}$  of the observed variables for the proposed experiments under hypothesis 1 and 2, respectively:

$$y_N^{(1)} = f^{(1)}[x_N, \theta^{(1)}]$$

$$y_N^{(2)} = f^{(2)}[x_N, \theta^{(2)}]$$

Let the covariance matrices of these two predictions be  $V^{(1)}$  and  $V^{(2)}$ , respectively, that is,

$$E^{(i)}[y_N - y_N^{(i)}][y_N - y_N^{(i)}]^T = V^{(i)} \quad (i = 1, 2)$$

Furthermore, let  $U^{(i)} = [V^{(i)}]^{-1}$ . Assuming normal distributions, it is easily shown that Eq. (44) reduces to

$$I = -r + \frac{1}{2} \text{Tr}[U^{(2)}V^{(1)}] + \frac{1}{2} \text{Tr}[U^{(1)}V^{(2)}] + \frac{1}{2} [y_N^{(1)} - y_N^{(2)}]^T [U^{(1)} + U^{(2)}] [y_N^{(1)} - y_N^{(2)}] \quad (45)$$

where

$$\text{Tr}(UV) = \sum_{i,j} U_{ij} V_{ij}$$

Application of Eq. (45) requires estimation of the matrices  $V^{(i)}$  which are measures of the uncertainties in the predicted values  $y_N^{(i)}$ . There are two sources for these uncertainties:

1. The uncertainty in the values of the estimated parameters  $\theta^{(i)}$ , measured by means of the covariance matrix of the estimates  $W^{(i)}$ .
2. The difference between the measured values of the variables  $y_N^{(i)}$  and their true values, due to experimental errors. A measure of these errors is given by the covariance matrix of the residuals  $R^{(i)}$ .

The matrices  $W^{(i)}$  and  $R^{(i)}$  are obtained as by-products of estimating the parameters for the  $i$ th model, based on the experiments already conducted. Now, assuming that these two causes bring about independent errors, we have approximately

$$V^{(i)} = R^{(i)} + \left[ \frac{\partial y_N^{(i)}}{\partial \theta^{(i)}} \right] W^{(i)} \left[ \frac{\partial y_N^{(i)}}{\partial \theta^{(i)}} \right]^T \quad (i = 1, 2) \quad (46)$$

or, written out in full,

$$V_{\alpha\beta}^{(i)} = R_{\alpha\beta}^{(i)} + \sum_{p,q} \frac{\partial y_N^{(i)}}{\partial \theta_p^{(i)}} W_{pq}^{(i)} \frac{\partial y_N^{(i)}}{\partial \theta_q^{(i)}}$$

An interesting case is that of a single observed variable, where  $V^{(i)}$  and  $R^{(i)}$  are single numbers,  $\sigma_1^2$  and  $s_1^2$ , respectively, with

$$\sigma_1^2 = s_1^2 + \sum_{p,q} W_{pq}^{(i)} \frac{\partial y_N^{(i)}}{\partial \theta_p^{(i)}} \frac{\partial y_N^{(i)}}{\partial \theta_q^{(i)}}$$

and Eq. (45) becomes

$$I = \frac{1}{2} \left[ \frac{(\sigma_1^2 - \sigma_2^2)^2}{\sigma_1^2 \sigma_2^2} + (y_N^{(1)} - y_N^{(2)})^2 \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \right] \quad (47)$$

If  $\sigma_1$  and  $\sigma_2$  do not vary much from one set of experimental conditions to another, the information is essentially proportional to  $[y_N^{(1)} - y_N^{(2)}]^2$ , i.e., to the square of the difference between the predicted values under the two hypothesis. The experiment to be performed is the one for which this difference is the greatest—a common sense result. More generally we seek those feasible values of the independent vari-



ables for which the information function [Eqs. (44), (45), or (47)] attains its maximum.

Equation (47) has been derived by Box and Hill [13], who have been the pioneers in the application of sequential design to kinetics experimentation. Hill and Hunter also provide generalizations for discrimination among more than two models and for combining the parameter estimation and model discrimination criteria in a single function [35]. The idea here is to form a linear combination of the two criteria, which is initially weighted to favor model discrimination. As one model becomes increasingly favored, the weight is shifted toward estimating the parameters for that model.

#### A. Termination of a Sequence of Experiments

The maximum information principle permits us to choose the experimental conditions for the Nth experiment after N-1 experiments have been performed. We need a criterion to decide whether the Nth experiment should be performed at all, or whether we may already prefer one of the models with a sufficient degree of certainty. Such a criterion is provided by Wald's likelihood ratio test [69]: Let  $L^{(i)}$  be the maximum value of the likelihood function based on model  $i$  using the data obtained to date. If  $\rho = L^{(1)}/L^{(2)}$ , then we adopt the following procedure:

If  $\rho \geq \frac{\alpha}{1-\alpha}$ , we accept model 1 with confidence  $\alpha$

If  $\rho \leq \frac{1-\alpha}{\alpha}$ , we accept model 2 with confidence  $\alpha$

If  $\frac{1-\alpha}{\alpha} < \rho < \frac{\alpha}{1-\alpha}$ , we continue experimentation

Note: for a confidence level of 99%, we set  $\alpha = 0.99$ .

In the case of a single observed variable with a normal distribution, we have, after N experiments,

$$\rho = \left(\frac{\sigma_2}{\sigma_1}\right)^N \exp \frac{1}{2}(p_1 - p_2)$$

where  $\sigma_i$  is the standard deviation of the residuals and  $p_i$  the number of parameters in model  $i$ .

#### B. Numerical Example

There do not seem to be, as yet, any published results of kinetics experiments actually carried out according to the above-detailed

prescriptions. Tests on computer-simulated experiments, however, show that use of the proper experimental design method can increase spectacularly our ability to select a proper model. In a computer-simulated experiment, a computer subroutine replaces the laboratory equipment. The selected experimental conditions (values of the independent variables) are fed as input to the subroutine, which computes the hypothetical results of the experiment, using one of the alternative models. The subroutine then adds to these results a pseudo-random variable having prescribed statistical properties (e.g., variance). These results are accepted by the parameter estimation and experimental design programs as though they had been produced by an actual experiment.

A system selected for simulation was that of the catalytic hydrogenation of mixed isooctenes, described by Hougen and Watson [37], and further analyzed by Blakemore and Hoerl [5]. The latter authors reduced the number of acceptable rate equations to two:

$$\text{Model 1: } y^{(1)} = \frac{\theta_1^{(1)} x_1 x_2}{(1 + \theta_2^{(1)} \sqrt{x_1} + \theta_3^{(1)} x_2 + \theta_4^{(1)} x_3)^3}$$

$$\text{Model 2: } y^{(2)} = \frac{\theta_1^{(2)} x_1 x_2}{(1 + \theta_2^{(2)} x_1 + \theta_3^{(2)} x_2 + \theta_4^{(2)} x_3)^2}$$

where  $y$  is the rate of the reaction and  $x_1$ ,  $x_2$ , and  $x_3$  are, respectively, the partial pressures of hydrogen, isooctene, and isooctane. The available experimental data were insufficient for preferring one of these models. For the computer-simulated experiments, the partial pressures of the reactants were confined to the same region as in the data used by Blakemore and Hoerl, namely,

$$0.1 \leq x_1 \leq 2.5$$

$$0.1 \leq x_2 \leq 3.0$$

$$0.05 \leq x_3 \leq 2.7$$

For the first six experiments, the fractional design given in Table 5 was used and the results (i.e., the rate of the reaction  $y$ ) were computed assuming model 1 was correct. Subsequent experiments were chosen so as to maximize the information measure given by Eq. (47) [in this case, at least, precisely the same results are obtained by maximizing  $(y_N^{(1)} - y_N^{(2)})^2$ ]. The results of those experiments were computed assuming model 2 was correct; in this way one desired to find out how soon the sequence of experiments would overcome the "misleading" results of the first six experiments and choose model 2

TABLE 5

Experiment	$x_1$	$x_2$	$x_3$
1	0.1	1.55	1.375
2	2.5	1.55	1.375
3	1.3	0.1	1.375
4	1.3	3.0	1.375
5	1.3	1.55	0.05
6	1.3	1.55	2.7

as the correct one. A total of 27 experiments were "performed." Also, 27 experiments constituting a  $3 \times 3 \times 3$  factorial design were carried out. Table 6 compares the levels of confidence in model 2 achieved after 27 experiments at various levels of simulated experimental error.

As expected, the discrimination power of each procedure diminishes with increasing experimental error; the factorial designs fail completely to discriminate at 3% error, whereas the sequential design fails only at 10%. Below this level the performance of the sequential design is spectacularly better than that of the factorial.

It is interesting to list the number of experiments that were needed to reach a 90% confidence level by the sequential design. These were 15, 17, and 30 for experimental errors of 1, 3, and 6%, respectively. In contrast, the factorial design never exceeded a 60% level.

The above results show that sequential experimental design holds great promise. Whether this promise will be fulfilled in practice is not yet known. The main theoretical question that arises when these methods are applied in practice arises from the fact that rarely is any of the proposed mechanisms exactly correct, and therefore one

TABLE 6

Experimental error, %	Confidence in model 2 after 27 experiments, %	
	Factorial design	Sequential design
1	60	99.8
3	50	98
6	50	80
10	50	50

should not exaggerate the confidence placed on any model selected as a result of analyzing the data obtained in a series of experiments. Clearly, if the correct model is not among those proposed, it cannot be the one selected. Frequently, however, the data themselves may lead to systematic modifications of the model, until a suitable model is found. Such an analysis is described in detail by Hunter and co-workers [38,44].

#### APPENDIX

There is currently available a number of computer programs which include some or all of the features discussed in this paper. It is felt worthwhile to point out briefly these programs and their features.

1. The Lapidus-Peterson program suitable for the IBM 7090/94 computers [51]. In brief, this program, which is an extension of an earlier nonlinear estimation program [7], consists of three conceptual parts which are linked together to perform the necessary computations: (a) a kinetics language for input of the kinetics reaction model and experimental data; (b) a differential equation solver for numerical integration of the rate equations; and (c) a nonlinear estimation algorithm for obtaining least-squares estimates of the parameters via the Gauss method. First derivatives are obtained in the program by finite difference approximation.

This program has been found very advantageous since the input is entered in a manner which is close to the kineticist's way of thinking and the emphasis is placed on the physical chemistry aspects rather than the mathematical formulation. Thus the investigator enters card formats which typically read as follows:

Program entry	meaning
$A = R + S$	$A \rightleftharpoons R + S$
Surface reaction control	surface-reaction controlling model
$(1 + ( )0.28A^{0.5})^2$	denominator term in rate expression $(1 + 0.28p_A^{0.5})^2$

Here  $K_A = 0.28$  is a starting value in the nonlinear estimation of  $K_A$ .  $\alpha = 0.5$  and  $\beta = 2.0$  are considered fixed in this particular format statement, but may also be estimated if desired.

From this input and additional data such as values of initial concentrations, settings of the independent variables and the observations, the computer program automatically (a) develops the kinetic rate expression; (b) integrates this expression numerically to yield, in effect, a specific form of the model equation; and (c) provides parameter estimates which fit the observations in the least-squares sense.

As output from the program, various statistics are produced. Among those found useful are the parameter estimates  $\hat{\theta}$ , the standard deviations of the parameter estimates  $s_{\hat{\theta}}$ , the standard deviation of the errors  $s$ , and the residuals between the observed and calculated values of the dependent variable  $\epsilon$ .

2. The Eisenpress-Greenstadt nonlinear maximum likelihood program suitable for the IBM 7090/94 computers [23]. This program was developed originally for large-scale economic problems and has no provisions for integrating differential equations. Therefore it is applicable to kinetics problems only where differential rate measurements are available or when analytic solutions to the rate equations are available. This program maximizes the likelihood function given by Eq. (24), corresponding to the following assumptions: (a) errors are normally distributed; (b) errors in different experiments are independent; and (c) the same unknown covariance matrix applies to the different observed variables in each experiment. This matrix is estimated along with the parameters.

The program uses Greenstadt's modification of the Newton-Raphson method, as described previously. This requires the evaluation of the first and second derivatives. The computer itself, using the FORMAC system of algebraic formula manipulations [6], performs all the required differentiations analytically.

3. The Bard nonlinear parameter estimation program [3], suitable for any computer accepting FORTRAN IV programs. It solves least-squares, weighted least-squares, maximum likelihood, and Bayesian estimation problems of the types discussed above. It incorporates an integration routine and special routines for generating the rate equations and their derivatives in kinetics problems. It uses the generalized Gauss-Newton method, or, optionally, the Davidon-Fletcher-Powell method [20,25].

4. The Marquardt program [55], written in FORTRAN IV. It uses Marquardt's method [54] and may be coupled easily to integration routines.

#### REFERENCES

- [1] R. Ash, *Information Theory*, Wiley (Interscience), New York, 1965.
- [2] Y. Bard, *IPM New York Scientific Center Rept. 320-2902(1967)*.

- [3] Y. Bard, *Nonlinear Parameter Estimation and Programming*, Program 360D 13.6.003, IBM, Hawthorne, N.Y.
- [4] R. Bellman, J. Jacquez, R. Kalaba, and S. Schwimmer, *Math. Biosci.*, **1**, 71 (1967).
- [5] J. W. Blakemore and A. E. Hoerl, *Chem. Eng. Progr. Symp. Ser.*, **59**(42), 14 (1963).
- [6] E. Bond, in *Proc. 19th ACM Natl. Conf.* (August, 1964).
- [7] G. W. Booth and T. I. Peterson, *Share Program No. 687* (1958).
- [8] G. E. P. Box, *Ann. N.Y. Acad. Sci.*, **86**, 792 (1960).
- [9] G. E. P. Box and G. A. Coutie, *Proc. IEE*, **103**, 100 (1956).
- [10] G. E. P. Box and N. R. Draper, *Biometrika*, **52**, 355 (1965).
- [11] G. E. P. Box and W. G. Hunter, *Technometrics*, **4**, 301 (1962).
- [12] G. E. P. Box and W. G. Hunter, in *Proceedings IBM Scientific Computing Symposium on Statistics, 1963*, pp. 113-137.
- [13] G. E. P. Box and W. J. Hill, *Technometrics*, **9**, 57 (1967).
- [14] G. E. P. Box and H. L. Lucas, *Biometrika*, **46**, 77 (1959).
- [15] M. J. Box, *Computer J.*, **9**, 67 (1966).
- [16] C. W. Carroll, *Operations Res.*, **9**, 169 (1961).
- [17] H. Chien and R. Aris, *Automatica*, **2**, 42 (1964); **3**, 59 (1965).
- [18] N. L. Cull and H. H. Brenner, *Ind. Eng. Chem.*, **53**, 833 (1961).
- [19] A. F. D'Alessandro and A. Farkas, *J. Colloid Sci.*, **11**, 653 (1956).
- [20] W. C. Davidon, *A.E.C. Res. Develop. Rept. ANL-5990* (1959).
- [21] N. R. Draper and W. G. Hunter, *Biometrika*, **53**, 525 (1966).
- [22] N. R. Draper and W. G. Hunter, *Biometrika*, **54**, 147 (1967).
- [23] H. Eisenpress, A. Bomberault, and J. Greenstadt, *IBM 7090, Share Program 7090 IBM 0035* (1966).
- [24] H. Eisenpress and J. Greenstadt, *Econometrica*, **34**, 851 (1966).
- [25] R. Fletcher and M. J. D. Powell, *Computer J.*, **6**, 163 (1963).
- [26] J. Franckaerts and G. F. Froment, *Chem. Eng. Sci.*, **19**, 807 (1964).
- [27] E. J. Freeh, H. G. Krane, and A. Syverson, *A.I.Ch.E. J.*, **9**, 400 (1963).
- [28] S. M. Goldfeld, R. E. Quandt, and N. F. Trotter, *Econometrica*, **34**, 541 (1966).
- [29] J. Greenstadt, *Math. Computation*, **21**, 360 (1967).
- [30] P. H. Hammond and M. J. Duckenfield, *Automatica*, **1**, 147 (1963).
- [31] G. Harris and L. Lapidus, *A.I.Ch.E. J.*, **13**, 291 (1967).
- [32] G. Harris and L. Lapidus, *Ind. Eng. Chem.*, **59**, 66 (1967).
- [33] H. O. Hartley, *Technometrics*, **3**, 269 (1961).
- [34] G. Heineken, H. M. Tsuchiya, and R. Aris, *Math. Biosci.*, **1**, 115 (1967).

- [35] W. J. Hill and W. G. Hunter, University of Wisconsin, Dept. of Statistics *Tech. Rept. 69* (1966).
- [36] R. Hooke and T. A. Jeeves, *JACM*, **8**, 212 (1961).
- [37] O. A. Hougen and K. M. Watson, *Chemical Process Principles - Part 3*, Wiley, New York, 1947.
- [38] W. G. Hunter and R. Mezaki, *A.I.Ch.E. J.*, **10**, 315 (1964).
- [39] W. G. Hunter and A. M. Reiner, *Technometrics*, **7**, 307 (1965).
- [40] I. I. Ioffe and Y. Sherman, *Zh. Fiz. Khim.*, **28**, 2095 (1954).
- [41] R. L. Kabel, Ph.D. thesis, University of Washington, Seattle, 1961.
- [42] J. R. Kittrell and R. Mezaki, *Can. J. Chem. Eng.*, **44**, 286 (1966).
- [43] J. R. Kittrell and R. Mezaki, *Ind. Eng. Chem.*, **59**, 28 (1967).
- [44] J. R. Kittrell, W. G. Hunter, and R. Mezaki, *A.I.Ch.E. J.*, **12**, 1014 (1966).
- [45] J. R. Kittrell, R. Mezaki, and C. C. Watson, *Ind. Eng. Chem.*, **57**, 18 (1963).
- [46] J. R. Kittrell, W. G. Hunter, and C. C. Watson, *A.I.Ch.E. J.*, **11**, 1051 (1965).
- [47] J. R. Kittrell, W. G. Hunter, and C. C. Watson, *A.I.Ch.E. J.*, **12**, 5 (1966).
- [48] J. R. Kittrell, R. Mezaki, and C. C. Watson, *Ind. Eng. Chem.*, **58**, 50 (1966).
- [49] J. R. Kittrell, R. Mezaki, and C. C. Watson, *Brit. Chem. Eng.*, **11**, 15 (1966).
- [50] S. Kullback, *Information Theory and Statistics*, Wiley, New York, 1959.
- [51] L. Lapidus and T. I. Peterson, *Share Program 7090 T2 IBM 0014* (1964).
- [52] L. Lapidus and T. I. Peterson, *A.I.Ch.E. J.*, **11**, 891 (1965).
- [53] K. Levenberg, *Quart. Appl. Math.*, **2**, 164 (1944).
- [54] D. W. Marquardt, *J. SIAM*, **11**, 431 (1963).
- [55] D. W. Marquardt, *Share Program No. 3094* (1963).
- [56] P. Mars and D. W. Van Krevelen, *Chem. Eng. Sci.*, **3** (*Spec. Suppl.*), 41 (1954).
- [57] T. I. Peterson, *Chem. Eng. Progr. Symp. Ser.*, **56**, 111 (1960).
- [58] T. I. Peterson, *Chem. Eng. Sci.*, **17**, 203 (1962).
- [59] T. I. Peterson and L. Lapidus, *Chem. Eng. Sci.*, **21**, 655 (1966).
- [60] P. H. Pinchbeck, *Chem. Eng. Sci.*, **6**, 105 (1957).
- [61] M. J. D. Powell, *Computer J.*, **7**, 155 (1964).
- [62] M. J. D. Powell, *Computer J.*, **7**, 303 (1965).
- [63] W. H. Ray and R. Aris, *Automatica*, **3**, 53 (1966); **4**, 139 (1967).
- [64] H. H. Rosenbrock, *Computer J.*, **3**, 125 (1960).
- [65] D. Rubin, *Chem. Eng. Progr. Symp. Ser.*, **59**, 90 (1963).
- [66] K. A. Shelstad, *Can. J. Chem. Eng.*, **38**, 102 (1960).



- [67] C. H. Tandy, *J. Appl. Chem.*, **6**, 68 (1956).  
[68] V. P. Ushakova, *Ukr. Khim. Zh.*, **23**, 191 (1957).  
[69] A. Wald, *Sequential Analysis*, J. Wiley, New York, 1947.  
[70] A. Wheeler, *Advan. Catalysis*, **3** (1951).

*Received by editor November 29, 1967*

*Submitted for publication January 17, 1968*

LIST OF ATTENDEES  
1972 ARMY NUMERICAL ANALYSIS CONFERENCE

AGEE, WILLIAM S.  
Analysis and Computations Division  
National Range Operations Directorate  
White Sands Missile Range,  
New Mexico 88002

BARBIERI, W. A.  
Rand Corporation  
Washington, DC

BENZKOFER, PHILIP D.  
US Army Weapons Command  
Rock Island Arsenal  
Rock Island, Illinois 61201

BRAMLEY, JENNY  
US Army Engineering Topographic  
Laboratories  
Fort Belvoir, VA. 22060

CHANDRA, JAGDISH  
Mathematics Division  
US Army Research Office-Durham  
Box CM, Duke Station  
Durham, North Carolina 27706

CHAPMAN, G. T.  
Ames Research Laboratory  
Mail Stop N213-4, NASA  
Moffett Field, California 94035

CLARK, DAVID  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD. 21005

CLAY, WALLACE  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD. 21005

COLEMAN, NORMAN  
US Army Weapons Command  
Rock Island Arsenal  
Rock Island, Illinois 61201

COLLINGER  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

CUTHILL, ELIZABETH  
Carderock Laboratory  
Code 1805  
Department of the Navy  
Bethesda, Maryland 20034

CLINTON, FRANK  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

DONNELLY, J. D. P.  
Mathematics Research Center  
University of Wisconsin  
610 Walnut Street  
Madison, Wisconsin 53706

DRESSEL, FRANCIS G.  
Mathematics Division  
US Army Research Office-Durham  
Box CM, Duke Station  
Durham, North Carolina 27706

GALBRAITH, ALAN S.  
Mathematics Division  
US Army Research Office-Durham  
Box CM, Duke Station  
Durham, North Carolina 27706

GOLDSTEIN, HENRY  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

GREENE, PATRICIA  
US Army Missile Command  
Redstone Arsenal, Alabama 35809

GREENLAND, C. MAXSON  
Analysis Group, Plans Office  
System Concepts/Studies Division  
Edgewood Arsenal, Maryland 21010

GIESE, J. H.  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

HARRIS, BENJAMIN J.  
Technical Director  
Edgewood Arsenal  
Edgewood Arsenal, Maryland 21010

HEINMETS, F.  
Head, Biophysics Group  
Pioneering Research Laboratory  
US Army Natick Laboratories  
Natick, Massachusetts 01760

HOWES, DAVID R.  
US Army Strategy and Tactics  
Analysis Group  
8120 Woodmont Avenue  
Bethesda, Maryland 20014

JAMESON, JOHN W.  
Biophysics Division  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

KALABA, ROBERT E.  
University of Southern California  
Department of Electrical Engineering  
University Park  
Los Angeles, California 90007

KASS, DONALD  
US Army Natick Laboratories  
Natick, Massachusetts 01762

KITCHENS, CLARENCE  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

KIWAN, A. R.  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

LARKIN, ROBERT  
D/Management Information Systems  
Fort Monmouth, New Jersey 07703

LAPIDUS, LEON  
Chemical Engineering Department  
Princeton University  
Princeton, New Jersey 08540

LILLY, CARLYLE  
Biophysics Division  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

MACKEY, STEVE  
Bel Air High School  
Bel Air, Maryland 21014

McMAMEE, FREDERICK M.  
Office of the Chief of Engineers  
HQ DA (DAEN-DST)  
Washington, DC 20314

McNEILL, JOSEPH W.  
US Army Land Warfare Laboratory  
Aberdeen Proving Ground, MD 21005

MEWILLI, ROBERT  
Bel Air High School  
Bel Air, Maryland 21014

MASAITIS, C.  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

MENDELSON, ERIC C.  
DA, US Army Mobility Equipment  
Research and Development Center  
Fort Belvoir, Virginia 22060

MERMAGEN, WILLIAM H.  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

MESERVE, C. L.  
CRD, Office of the Chief of Naval Operations  
Systems Analysis Division (OP-963)  
Room 4A720, Pentagon  
Washington, DC 20350

MILHOLLAND, ARTHUR  
Applied Mathematics-Statistics Group  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

MITTELMAN, LOU  
DA, US Army Mobility Equipment  
Research and Development Center  
Fort Belvoir, Virginia 22060

MOORE, RICHARD L.  
US Army Weapons Command  
Rock Island Arsenal  
Rock Island, Illinois 61201

MOSHANG, EDWARD  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

MULLIN, A. A.  
Mathematics Branch  
Office of Chief of Research  
and Development  
DA, Washington, DC 20310

METKER, LeROY W.  
Biophysics Division  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

NASHED, M. Z.  
Mathematics Research Center  
University of Wisconsin  
610 Walnut Street  
Madison, Wisconsin 53706

NOSTRAND, C. VAN CPT  
Development Branch  
ACS Studies and Analysis  
HQ, US Air Force  
Washington, DC 20330

NOLAN, JOHN  
Bel Air High School  
Bel Air, Maryland 21014

PLATOU, A. S.  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

PLICHTA, GEORGE E.  
Medical Research Institute  
of Infectious Diseases  
Fort Detrick  
Frederick, Maryland 21701

RALL, LOUIS  
Mathematics Research Center  
University of Wisconsin  
610 Walnut Street  
Madison, Wisconsin 53706

REDUS, KENNETH S.  
Applied Mathematics-Statistics Group  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

ROBERSON, M. L. LTC  
Chief, Development Branch  
ACS Studies and Analysis  
HQ, US AIR FORCE  
Washington, DC 20330

ROWBERG, ALAN H.  
Medical Research Institute  
of Infectious Diseases  
Fort Detrick  
Frederick, Maryland 21701

ROBERTSON, DONALD  
US Army Natick Laboratories  
Natick, Massachusetts 01762

SACCO, WILLIAM J.  
Applied Mathematics-Statistics Group  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

STOVER, RICHARD  
VISICON Inc., Box 1008  
State College, Pennsylvania 16801

SCHIFF, MYRA  
Biophysics Division  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

SHEAR, RALPH  
US Army Aberdeen Research  
and Development Center  
Ballistic Research Laboratories  
Aberdeen Proving Ground, MD 21005

STEGA, MARY  
Bel Air High School  
Bel Air, Maryland 21014

TACKLE, KNUTE  
Code MAT 0335  
Department of the Navy  
Washington, DC 20360

THRALL, ROBERT M.  
Department of Mathematical Sciences  
Rice University  
Houston, Texas 77001

TIEN-YU TSUI  
US Army Materials and Mechanics  
Research Center  
Watertown, Massachusetts 02172

TYLER, JOSEPH S.  
Applied Mathematics-Statistics Group  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

STURDIVAN, LARRY  
Biophysics Division  
Biomedical Laboratory  
Edgewood Arsenal, Maryland 21010

WALGER, JOHN A.  
Office of the Chief of Engineers  
HQ DA (DAEN-DST)  
Washington, DC 20314

WILLIAMS, C. M.  
VISICON Inc., Box 1008  
State College, Pennsylvania 16801

ZUNDEL, BROWN  
Deseret Test Center  
Building 100, Soldiers' Circle  
Fort Douglas, Utah 84113