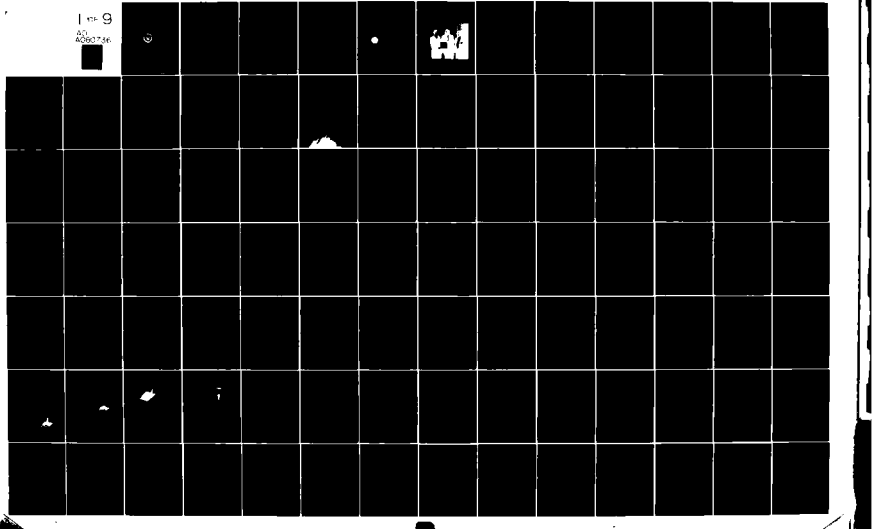
AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL



ARO Report 80-1

LEVEL II

TRANSACTIONS OF THE TWENTY-FIFTH

CONFERENCE OF ARMY MATHEMATICIANS

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.

Sponsored by

The Army Mathematics Steering Committee

on behalf of

THE CHIEF OF RESEARCH, DEVELOPMENT

AND ACQUISITION

(14) ARO-80-1

U. S. ARMY RESEARCH OFFICE

Report No. 80-1

January 1980

(6) TRANSACTIONS OF THE [redacted] CONFERENCE OF ARMY MATHEMATICIANS (25th).

Sponsored by the Army Mathematics Steering Committee

(9) Interim technical rept.

Hosts

U. S. Army Ballistic Research Laboratory

with the

Johns Hopkins University
Baltimore, Maryland

(12) 793

6-8 June 1979

(11) Jan 80

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, North Carolina

040 900

FOREWORD

The Silver Anniversary of the Conferences of Army Mathematicians was conducted on the dates 6-8 June 1979. The U. S. Army Ballistic Research Laboratory and the Johns Hopkins University served as its hosts. It was held on the Homewood Campus of the Johns Hopkins University in Baltimore, Maryland. The first meetings in this series were entitled "Conferences of Arsenal Mathematicians". The initial one was conducted at Watertown Arsenal on 29 October 1954. The host for the meeting this year served in the same capacity for the second conference which was only a one-day meeting with one invited speaker. The eighth meeting in the series was the first one to come under the auspices of the Army Mathematics Steering Committee (AMSC). This Committee requested that these conferences be held on an Army-wide basis and suggested that their title be changed to "Conferences of Army Mathematicians".

"Continuum Mechanics" was the theme selected for the Silver Jubilee. To celebrate this occasion the number of guest speakers were increased, and those individuals who were invited to talk were selected because they are effective researchers who are in the frontiers of their fields. Another important reason for their appearance on the program is that they are interested in current and envisioned U. S. Army materiel research and development problems. As in previous conferences there were a large number of papers contributed by Army scientists. These, on the whole, addressed problems of immediate interest to scientists in the various Army laboratories.

The keynote speaker was Professor R. S. Rivlin, Director of the Center for the Application of Mathematics at Lehigh University. The title of his address was "The Mechanics of Non-Newtonian Fluids". Professor Werner Goldsmith, Department of Applied Mechanics, University of California-Berkeley, gave a featured presentation on "Mathematical Modeling of Some Aspects of the Penetration of Plates by Projectiles". On the morning of the second day of the conference, Professor Daniel D. Joseph, Department of Aerospace Engineering and Mechanics, University of Minnesota, spoke about "Motions which Perturb States of Rest of Viscoelastic Solids". On the afternoon of this same day, Professor S. Nemat-Nasser, Department of Civil Engineering, Northwestern University, gave an address entitled "Finite Deformation Plasticity and Plastic Instability".

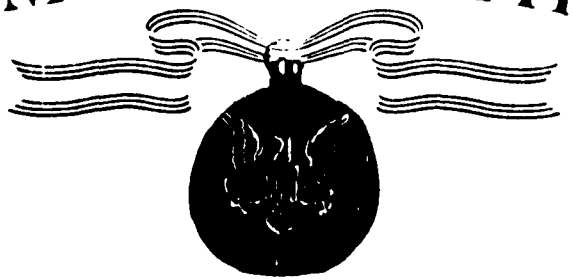
Other invited speakers were Professor George Papanicolaou, Courant Institute of Mathematical Sciences, New York University, Professor Harry F. Tiersten, Department of Mechanical Engineering, Aeronautical Engineering and Mechanics, Rensselaer Polytechnic Institute and Professor Morton Gurtin, Department of Mathematics, Carnegie-Mellon University. These gentlemen addressed the conference on the third and last morning of this meeting.

Dr. Papanicolaou's topic was "Effective Parameters and Fluctuation Phenomena in Continuum Mechanics". Dr. Tiersten reported on the "Theory of Interpenetrating Solid Continua and Some Applications", while Dr. Gurtin addressed the group on the topic "Recent Results in Finite Elasticity".

Another outstanding feature of this Silver Anniversary conference was the awarding of the *Decoration for Distinguished Civilian Service* to Professor Benjamin Noble upon his retirement as Director, Mathematics Research Center (MRC), University of Wisconsin-Madison. This presentation was made at the banquet by Dr. Percy Pierre, Assistant Secretary of the Army for Research and Development. Dr. Pierre pointed to the many outstanding scientific contributions made by Professor Noble to the field of mathematics, and he also stressed the fact that Noble had striven with unprecedented vigor, enthusiasm and innovation to render MRC more responsive to Army needs. [In the photograph on one of the following pages, Professor Noble is shown receiving the award from Dr. Pierre. The gentlemen on the right is Professor John A. Nohel, the incoming Director of MRC.]

Members of the AMSC would like to thank the Ballistic Research Laboratory (BRL) and Johns Hopkins University for serving as hosts of the Twenty-Fifth Conference of Army Mathematicians. Dr. Stephen S. Wolff, Chairman of Local Arrangements, did an outstanding job of arranging for all the needs of the speakers, and for handling the many requests for help by those in attendance. He was ably assisted with these tasks by various staff members of BRL and also employees of Johns Hopkins University. All those that spent time to prepare and present papers are also due their thanks. Their scientific ideas and methods for treating difficult problems were enlighting to members of the audience. The latter group added to the value of this conference by raising interesting questions and making valuable suggestions on possible ways to cope with troublesome problems.

DEPARTMENT OF THE ARMY



Dr. Benjamin Noble

IS AWARDED THE

DECORATION FOR DISTINGUISHED CIVILIAN SERVICE

CITATION:

For distinguished service and exemplary leadership while serving as Director of the Mathematics Research Center, University of Wisconsin, from 1975 to 1979. During this period, Dr. Benjamin Noble exhibited particularly meritorious management ability and scientific knowledge in guiding the Center. He introduced new organizational ideas and made the Center more sensitive to the scientific needs of the Army. His personality and scientific initiative created an atmosphere conducive to high quality research, which led to joint work among the core of researchers at the Center. The prestige and international prominence attained by the Mathematics Research Center reflect favorably on Dr. Noble and the United States Army.

Clifford C. Alexander Jr.

O.D.

Dr. Benjamin Nobel, Dr. Percey Pierre and Dr. John Nohel

TABLE OF CONTENTS*

*This table of contents contains only the papers that are published in this technical manual. For a list of all papers presented at the Twenty-Fifth Conference of Army Mathematicians, see the Program of the meeting.

1979
THE SILVER JUBILEE YEAR OF THE CONFERENCES OF ARMY MATHEMATICIANS

THIS PAMPHLET CONTAINS THE NAMES OF INVITED SPEAKERS

SELECTED FROM THE PROGRAMS OF

THE FIRST TWENTY-FOUR MEETINGS IN THIS SERIES

These Conferences Are Sponsored By
The Army Mathematics Steering Committee

U. S. Army Research Office
Research Triangle Park, North Carolina

CONFERENCES OF ARMY MATHEMATICIANS*

<u>List of Invited Speakers at</u>
<u>the First Twenty-Four Meetings</u>

<u>First Conference</u>: 29 October 1954, Watertown Arsenal

Professor F. J. Murray, Columbia University

COMPUTING METHODS FOR THE SOLUTION OF SYSTEMS OF DIFFERENTIAL EQUATIONS

<u>Second Conference</u>: 24 February 1956, Ballistic Research Laboratory

Professor C. B. Tompkins, University of California

LINEAR PROGRAMMING AND HIGH SPEED COMPUTER APPLICATIONS

[illegible], [illegible] Arsenal

Professor John W. Cell, North Carolina State College

MATHEMATICAL THEORY OF MOTION OF SPINNER ROCKETS DURING THE BURNING PERIOD WHEN THE RATIO OF ANGULAR TO LINEAR VELOCITY MAY BE DIFFERENT FROM THE RATIO OF ANGULAR TO LINEAR ACCELERATION DURING BURNING OUTSIDE THE LAUNCHER

<u>Fourth Conference</u>: 5-6 June 1958, Picatinny Arsenal

Professor Mark Kac, Cornell University

PROBABILITY AND RELATED TOPICS IN PHYSICAL SCIENCE

Dr. D. B. McElwain, Reaction Motors, Inc.

RECENT DEVELOPMENTS IN BOTH LIQUID AND SOLID FUEL MOTORS (*Talk presented at Banquet.*

<u>Fifth Conference</u>: 3-4 June 1959, Rock Island Arsenal

Professor R. G. Langebartel, University of Illinois

EARTH SATELLITE ORBIT DETERMINATION THEORY

---

*The first seven meetings in this series were entitled, "Conferences of Arsenal Mathematicians".

Sixth Conference: 1-2 June 1960, U.S. Army Research Office (Durham)

Professor J. B. Diaz, University of Maryland

ON EXISTENCE, UNIQUENESS AND NUMERICAL EVALUATION OF SOLUTIONS OF ORDINARY AND HYPERBOLIC DIFFERENTIAL EQUATIONS

Seventh Conference: 7-8 June 1961, Mathematics Research Center, University of Wisconsin

Dr. H. F. Bueckner, Mathematics Research Center (MRC)

ON A CLASS OF SINGULAR INTEGRAL EQUATIONS

Eighth Conference: 6-7 June 1962, Mathematics Research Center

Professor B. R. Seth, Indian Institute of Technology and MRC

SIMPLE CASES OF TRANSITION PHENOMENON

Ninth Conference: 5-6 June 1963, Watervliet Arsenal

Professor R. C. DiPrima, Rensselaer Polytechnic Institute

STABILITY OF FLOW BETWEEN ROTATING CYLINDERS AND RELATED TOPICS

Tenth Conference: 10-11 June 1964, U.S. Army Materiels Research Agency

Professor Garrett Birkhoff, Harvard University

WELL-SET PROBLEMS, FUNCTION SPACES AND COMPUTING

Eleventh Conference: 9-10 June 1965, U.S. Army Frankford Arsenal

Professor Frank Harary, University of Michigan

MUTUAL APPLICATIONS OF GRAPHS AND MATRICES

Twelfth Conference: 22-23 June 1966, U.S. Army Cold Regions Research and Engineering Laboratory

Professor J. G. Kemeny, Dartmouth College

THE DARTMOUTH TIME SHARING COMPUTING SYSTEM

Dean Myron Tribus, Dartmouth College

A REEXAMINATION OF STATISTICAL INFERENCE FROM THE POINT OF VIEW OF INFORMATION THEORY

Thirteenth Conference: 7-8 June 1967, U.S. Army Electronics Command

Professor R. J. Duffin, Carnegie Institute of Technology

OPTIMIZATION OF ENGINEERING DESIGN BY GEOMETRIC PROGRAMMING

Fourteenth Conference: 12-13 June 1968, Rock Island Arsenal

Colonel L. M. Orman, U. S. Army Weapons Command

NEW CONCEPTS IN WEAPON DEVELOPMENT (Talk given at banquet)

Professor B. R. Seth, Dibrugarh University, India, and State University, Corvallis, Oregon

NEW CONCEPTS IN CONTINUUM MECHANICS

Fifteenth Conference: 11-12 June 1969, U.S. Army Aviation Systems Command, St. Louis, Missouri

Professor R. E. Meyer, University of Wisconsin

SURF

Sixteenth Conference: 28-29 May 1970, U.S. Army Strategy and Tactics Analysis Group, Bethesda, Maryland

Dr. Guillermo Owen, Rice University

THE INFLUENCE OF GAME THEORY ON GAMING

Professor R. E. Machol, Northwestern University

THE ROLE OF MATHEMATICIANS IN SYSTEMS ENGINEERING

Seventeenth Conference: 26-28 May 1971, U.S. Army Missile Command, Redstone Arsenal, Alabama

Professor J. D. C. Little, Massachusetts Institute of Technology
MANAGERS AND MODELS - A CONCEPT OF DECISION CALCULUS

Professor H. A. Antosiewicz, University of Southern California
STABILITY THEORY: AN OVERVIEW

Eighteenth Conference: 24-26 May 1972, Picatinny Arsenal

Dr. Ira Cochin, Newark College of Engineering
CONNECTIVITY THEORY AND APPLICATIONS

Professor E. L. Reiss, New York University, Courant Institute of Mathematical Sciences
NONLINEAR DYNAMIC STABILITY

Nineteenth Conference: 23-25 May 1973, U.S. Army Training Device Agency, Naval Training Equipment Center, Orlando, Florida

Professor L. A. Segel, Rensselaer Polytechnic Institute
ON COLLECTIVE MOTIONS OF CHEMOTACTIC CELLS

Dr. T. C. Hu, Mathematics Research Center
A NEW PROOF OF THE T-C ALGORITHM

Dr. S. M. Robinson, Mathematics Research Center
COMPUTABLE ERROR BOUNDS FOR NONLINEAR PROGRAMMING

Twentieth Conference: 14-16 May 1975, U.S. Army Natick Laboratories

Professor S. W. Golomb, University of Southern California
THEORY AND APPLICATION OF FINITE FIELDS

Professor J. P. LaSalle, Brown University
RECENT ADVANCES IN THE STUDY OF DYNAMICAL SYSTEMS

Professor Fritz John, Courant Institute of Mathematical Sciences
NONLINEAR WAVE PROPAGATION

Professor J. B. Rosser, Mathematics Research Center
FOURIER SERIES IN THE COMPUTER AGE

Professor Carl de Boor, Mathematics Research Center
HOW TO DIFFERENTIATE NUMERICALLY IF YOU MUST

Professor R. A. Askey, Mathematics Research Center
SUMS OF BINOMIAL COEFFICIENTS

Professor D. A. Sanchez, Mathematics Research Center
FUNCTIONAL ANALYSIS AND THE METHOD OF HARMONIC BALANCE

Twenty-First Conference: 14-16 May 1975, U.S. Army White Sands Missile Range

Professor A. J. Hoffman, IBM-Thomas J. Watson Research Center
ON SPECTRALLY BOUNDED SIGNED GRAPHS

Professor Donald S. Cohen, California Institute of Technology
NONLINEAR PROBLEMS IN CHEMICALLY REACTING DIFFUSIVE SYSTEMS

Dr. A. J. Viterbi, LINKABIT Corporation, San Diego, California
A MAXIMUM LIKELIHOOD DECISION ALGORITHM FOR MARKOV SEQUENCES WITH MULTIPLE APPLICATIONS TO DIGITAL COMMUNICATIONS

Twenty-Second Conference: 12-14 May 1976, Watervliet Arsenal

Professor A. C. Eringen, Princeton University
STATE OF STRESS IN THE NEIGHBORHOOD OF A SHARP CRACK TIP

Professor John Buckmaster, University of Illinois
ACTIVATION ENERGY ASYMPTOTICS AND UNSTEADY FLAMES

Professor Thomas Kailath, Stanford University

SOME NEW METHODS FOR SOLVING LINEAR EQUATIONS

Dr. H. S. Bueckner, General Electric Company

THE WEIGHT FUNCTIONS OF MODE I OF THE PENNY-SHAPED AND OF THE ELLIPTIC CRACK

Professor James Rice, Brown University

RECENT DEVELOPMENTS IN THE THEORY OF ELASTICITY AND RUPTURE OF FLUID INFILTRATED SOLIDS

Twenty-Third Conference: 11-13 May 1977, U.S. Army Mobility R&D Laboratory, Langley Research Center, Hampton, Virginia

Professor M. D. Kruskal, Princeton University

WHAT'S ALL THIS ABOUT SOLITONS

Professor D. H. Sattinger, University of Minnesota

GROUP THEORETIC METHODS IN BIFURCATION THEORY

Professor Michael Crandall, Mathematics Research Center

EVOLUTION GOVERNED BY ACCRETIVE OPERATORS

Professor H. O. Kreiss, Uppsala University, Sweden; Visiting NYU

NUMERICAL SOLUTION OF PROBLEMS WITH DIFFERENT TIME SCALES

Professor Edward Kamen, Georgia Institute of Technology

USE OF ALGEBRAIC METHODS IN THE DESIGN OF CONTROLLERS AND OBSERVERS FOR SYSTEMS WITH TIME DELAYS

Twenty-Fourth Conference: 31 May - 2 June 1976, U.S. Army Forest Science and Technology Center. Held on the campus of the University of Virginia, Charlottesville, Virginia

Professor E. J. McShane, University of Virginia

CHOOSING A MATHEMATICAL MODEL FOR A SYSTEM AFFECTED BY NOISE

Professor R. E. Kalman, University of Florida

NONLINEAR REALIZATION THEORY

Professor Y. K. Lin, University of Illinois

STOCHASTIC THEORY OF ROTOR BLADE DYNAMICS

Professor Roger Brockett, Harvard University

MODELING AND ESTIMATION WITH BILINEAR STOCHASTIC SYSTEMS

Professor Ronald DiPerna, Mathematics Research Center

HYPERBOLIC CONSERVATION LAWS

Professor Eugene Wong, University of California-Berkeley

A MARTINGALE THEORY OF RANDOM FIELDS

PROGRAM

of the

TWENTY-FIFTH

CONFERENCE OF ARMY MATHEMATICIANS

sponsored by

THE JOHNS HOPKINS UNIVERSITY
and
U.S. ARMY ARRADCOM BALLISTIC RESEARCH LABORATORY

June 6-8, 1979

Wednesday, 6 June 1979

0800-0830 Registration

0830-0900 Opening remarks

0900-1000 GENERAL SESSION I
Dr. Ben Noble, Director, Mathematics Research Center, University of Wisconsin

Speaker: Prof. Ronald S. Rivlin, Director, Center for the Application of Mathematics, Lehigh University
THE MECHANICS OF NON NEWTONIAN FLUIDS

1000-1030 Break

1030-1215 TECHNICAL SESSION I
Dr. Aivars Celmins
ARRADCOM Ballistic Research Laboratory

Philip Hall, Rensselaer Polytechnic Institute
Taylor Vortices in Finite Cylinders

John T. Harrison, ARRADCOM Ballistic Research Laboratory
The History of the Utilization of Eulerian Hydrodynamic Computer Codes at the Ballistic Research Laboratory

Abdul R. Kiwan, ARRADCOM Ballistic Research Laboratory
Study of Convergent Flows in Certain Shaped Charge Systems

Janet E. Lacsters and William P. Walters, ARRADCOM Ballistic Research Laboratory
Theoretical and Experimental Studies of Hemispherical Shaped Charge Liners

John Mescall, Army Materials and Mechanics Research Center
Computer Simulation of Adiabatic Shear Processes in Ballistic Impact and Fragmentation

TECHNICAL SESSION II
Mr. Roger Willis
TRADOC Systems Analysis Agency

P. R. Kumar, University of Maryland, Baltimore Campus
Optimal Mixed Strategies in Dynamic Games

M. A. Hirschberg, ARRADCOM Ballistic Research Laboratory and B. F. Cummings, Army Materiel Systems Analysis Activity
The Use of Symbolic Methods to Reduce the Size and Cost of Game Computations

T. N. E. Greville, Mathematics Research Center
Bounds for Eigenvalues of Hermitian Trench Matrices

R. W. Soanes, Jr., ARRADCOM Large Caliber Weapon Systems Laboratory
Efficient Algorithms for Continuous Piecewise Linear Approximations with Variable Knots

Charles R. Leake, US Army Armor and Engineer Board
An Extension C_{-1} of C_1 that Has an Application in Learning Theory

1215-1345 Lunch

1345-1515 TECHNICAL SESSION III
Dr. Thomas E. Simkins
ARRADCOM Large Caliber Weapon Systems Laboratory

John F. Polk, ARRADCOM Ballistic Research Laboratory
An Algorithm for Heat Transfer in Gun Barrels

Ashwani Kapila, Rensselaer Polytechnic Institute
Reactive Diffusive System with Arrhenius Kinetics: Dynamics of Ignition

Geoffrey S. S. Ludford, Cornell University
Premixed Cylindrical Flames

Julian L. Davis, ARRADCOM and Yu Chen, Rutgers University
Thermoelastic Stresses in Gun Barrels

TECHNICAL SESSION IV
Dr. Edward Ross
NARADCOM

John A. Nohel, Mathematics Research Center
A Nonlinear Hyperbolic Volterra Equation for One-Dimensional Viscoelastic Motion

Francis E. Council, MERADCOM
A New Technique for the Solution of Navier's Equations

J. Barkley Rosser, Mathematics Research Center
An Iterative Algorithm for Calculating Potentials near Two Parallel Plates of Equal Length. Part II

C. N. Shen, ARRADCOM Large Caliber Weapon Systems Laboratory
Sensitivity Coefficient of External Ballistics with Velocity Square Damping

1515-1530 Break

1530-1630 GENERAL SESSION II
Dr. Thomas W. Wright, ARRADCOM Ballistic Research Laboratory

Speaker: Prof. Werner Goldsmith, Department of Applied Mechanics, University of California-Berkeley
MATHEMATICAL MODELING OF SOME ASPECTS OF THE PENETRATION OF PLATES BY PROJECTILES

1800-2200 CONFERENCE BANQUET (cash bar 1800-1900)
Caesar's Forum, Holiday Inn of Downtown Baltimore

Thursday, 7 June 1979

0830-1030 — TECHNICAL SESSION V
Dr. Siegfried Lehnigk
US Military Academy

Yoshisuke Nakano, US Army Cold Regions Research and Engineering Laboratory
Propagating Velocity of Singularities Occurring in Certain Degenerate Parabolic Equations

Bernard Fleishman, Rensselaer Polytechnic Institute and Thomas J. Mahar, Northwestern University
A Minimum Principle for Superharmonic Functions Subject to Interface Conditions

James A. Schmitt, ARRADCOM Ballistic Research Laboratory
An Analytic Model for Shock Wave Propagation into Concave Corners

Joseph E. Oliger, Mathematics Research Center
On the Initial Boundary Value Problem for the Equations of Gas Dynamics

Richard E. Ewing, Mathematics Research Center and Ohio State University
Efficient Time Stepping Procedures for Nonlinear Partial Differential Equations with Nonlinear Neumann Boundary Conditions

TECHNICAL SESSION VI
Dr. John Mescall
Army Materials and Mechanics Research Center

Joseph J. Roseman, Tel-Aviv University
Integral Bounds for the Strain Energy in Terms of Surface Tractions or Displacements and Body Forces in Finite Elastostatics

P. C. T. Chen, ARRADCOM Large Caliber Weapon Systems Laboratory
A Finite Difference Approach to Axisymmetric Plane Strain Problems beyond the Elastic Limit

Ram P. Srivastav and A. Gerasoulis, State University of New York - Stony Brook
A Method for the Numerical Solution of Singular Integral Equations with a Principal Value Integral

Dennis M. Tracey and Colin E. Freese, Army Materials and Mechanics Research Center
Developments in Elastic Plastic Finite Element Analysis

M. A. Hussain and S. L. Pu, ARRADCOM Large Caliber Weapon Systems Laboratory and Ben Noble, Mathematics Research Center
Stress Singularity at the Vertex of a Thin Angular Sector Crack by Variational Method

1030-1100 — Break

1100-1200 — GENERAL SESSION III
Dr. Charles H. Murphy, Jr., ARRADCOM Ballistic Research Laboratory

Speaker: Prof. Daniel D. Joseph, Department of Aerospace Engineering and Mechanics, University of Minnesota
MOTIONS WHICH PERTURB STATES OF REST OF VISCOELASTIC SOLIDS

1200-1330 — Lunch

1330-1515 — TECHNICAL SESSION VII
Dr. J. Barkley Rosser
Mathematics Research Center

R. Breuwkes, Jr. and Donald M. Neal, Army Materials and Mechanics Research Center
A Simple Density Function with Finite Distribution Limits

M. M. Al-Hussaini and G. Y. Baladi, Waterways Experiment Station
Distribution of Displacements and Stresses beneath a Rigid Wheel

H. Baussus von Luetzow, Engineering Topographic Laboratory
On the Limitation and Improvement of Present Numerical Weather Prediction

Siegfried Lehnigk, US Military Academy
On Laguerre Type Functions

Shunsuke Takagi, US Army Cold Regions Research and Engineering Laboratory
Some Bessel Function Identities Arising in Ice Mechanics Problems

TECHNICAL SESSION VIII
Dr. F. G. Sharkoff
ARRADCOM Large Caliber Weapon Systems Laboratory

James T. Wong and Richard M. Carlson, AVRADCOM
Some Intrinsic Properties of Exact Solutions for the State Bending of Uniform Rotating Beams

Julian J. Wu and John D. Vasilakis, ARRADCOM Large Caliber Weapon Systems Laboratory
Dynamic Stability of Columns Subjected to Nonconservative Forces

J. D. Vasilakis and P. C. T. Chen, ARRADCOM Large Caliber Weapon Systems Laboratory
Thermo Elastic Plastic Stresses in Hollow Cylinders Due to Quenching

Julian J. Wu and Thomas E. Simkins, ARRADCOM Large Caliber Weapon Systems Laboratory
A Numerical Comparison between Two Unconstrained Variational Formulations

Thomas E. Simkins, ARRADCOM Large Caliber Weapon Systems Laboratory
Unconstrained Variational Statements for Initial and Boundary Value Problems

1515-1530 — Break

1530-1630 — GENERAL SESSION IV
Dr. John A. Nohel, Mathematics Research Center

Speaker: Prof. S. Nemat Nasser, Department of Civil Engineering, Northwestern University
FINITE DEFORMATION PLASTICITY AND PLASTIC INSTABILITY

Friday, 8 June 1979

0830-1030 GENERAL SESSION V
Dr. Stephen Wolff, ARRADCOM Ballistic Research Laboratory

Speaker: Prof. George Papanicolaou, Courant Institute of Mathematical Sciences, New York University
EFFECTIVE PARAMETERS AND FLUCTUATION PHENOMENA IN CONTINUUM MECHANICS

Speaker: Prof. Harry F. Tiersten, Department of Mechanical Engineering, Aeronautical Engineering and Mechanics, Rensselaer Polytechnic Institute
THEORY OF INTERPENETRATING SOLID CONTINUA AND SOME APPLICATIONS

1030-1100 Break

1100-1200 GENERAL SESSION VI
Dr. Moyed Hussain, ARRADCOM Large Caliber Weapon Systems Laboratory

Speaker: Prof. Morton Gurtin, Department of Mathematics, Carnegie-Mellon University
RECENT RESULTS IN FINITE ELASTICITY

1215 ADJOURN

THE MECHANICS OF NON-NEWTONIAN FLUIDS

Ronald S. Rivlin
Center for the Application of Mathematics
Lehigh University
Bethlehem, Pennsylvania

ABSTRACT. The continuum-mechanical theory of non-Newtonian fluids predicts many interesting flow effects which are qualitatively different from those which are observed in Newtonian fluids. A number of these effects are discussed. For the most part, the Rivlin-Ericksen constitutive equation is used as a basis for this discussion.

1. INTRODUCTION. One of the most striking advances in continuum mechanics in recent years has been the development of a rational theory for the mechanics of non-Newtonian fluids. One might think that the predictions of such a theory would consist of no more than minor modifications of those provided by the classical mechanics of Newtonian fluids in which effects, which in Newtonian fluids obey linear laws, for non-Newtonian fluids obey non-linear laws. That this is not the case was made evident from the rod-climbing experiment of Garner and Nissan [1] which showed that, at any rate in certain non-Newtonian fluids, rotation of a cylindrical rod with constant angular velocity in a bath of non-Newtonian fluid would result in a rise of the fluid up the stirrer, rather than a depression of the fluid due to the centrifugal effect. It was quickly realized that such effects may arise from the tensorial character of the non-linear relation between the stress and the kinematic variables describing the flow field. Theories based on such considerations have undergone a considerable development since that time and have lead to many unexpected flow phenomena, some of them of a spectacular character. It is not possible in this paper to introduce the many different formulations of the mechanics of non-Newtonian fluids which have been published. We concentrate on one particular formulation, due to Rivlin and Ericksen [2], and give some examples of the more interesting effects which it predicts.

2. RECTILINEAR SHEAR FLOW. For an incompressible Newtonian fluid in a state of rectilinear shearing flow with velocity gradient κ, direction of shear parallel to the x_1-axis of a rectangular cartesian coordinate system x, and the x_1x_2-plane as the plane of shear, the stress $\underset{\sim}{\sigma} = ||\sigma_{ij}||$ is given by

$$\sigma_{12} = \sigma_{21} = \eta\kappa \ , \quad \sigma_{11} = \sigma_{22} = \sigma_{33} = - \ p \ , \tag{2.1}$$

where η is the viscosity of the fluid, p is an arbitrary hydrostatic pressure and the remaining components of $\underset{\sim}{\sigma}$ are zero. The expressions (2.1) are valid whether or not κ is time-dependent.

More generally, for an incompressible non-Newtonian fluid, we can argue from simple symmetry considerations that, provided κ is time-independent, $\sigma_{12}(=\sigma_{21})$ is an odd function of κ and, apart from an arbitrary hydrostatic pressure, σ_{11}, σ_{22}, σ_{33} are even functions of κ. Again, the remaining components of $\underset{\sim}{\sigma}$ are zero. Thus,

$$\begin{aligned} &\sigma_{12} = \sigma_{21} = \kappa f(\kappa^2) \ , \\ &\sigma_{11} = f_1(\kappa^2) - p \ , \quad \sigma_{22} = f_2(\kappa^2) - p \ , \quad \sigma_{33} = - p \ . \end{aligned} \tag{2.2}$$

(The even function of κ in the expression for σ_{33} has been absorbed into the arbitrary hydrostatic pressure p.)

We note immediately one striking feature which distinguishes the mechanics of a non-Newtonian fluid from that of a Newtonian fluid. While for rectilinear shearing flow of a Newtonian fluid the three normal components of the stress are equal, this is not in general true for a non-Newtonian fluid. Effects which arise from this fact are called normal stress effects. We also note the non-linear dependence of the shear components of the stress on κ.

Provided that the functions f, f_1, f_2 are sufficiently smooth, we may express them as Taylor series and neglect terms of higher degree than the second in κ. We then have

$$\begin{aligned} &\sigma_{12} = \sigma_{21} = \tfrac{1}{2}\alpha_1\kappa \ , \quad \sigma_{11} = \tfrac{1}{4}\alpha_3\kappa^2 - p \ , \\ &\sigma_{22} = (\alpha_2+\tfrac{1}{4}\alpha_3)\kappa^2 - p \ , \quad \sigma_{33} = - p \ , \end{aligned} \tag{2.3}$$

where the α's are constants. (The manner in which the constant coefficients in (2.3) are written is chosen in order to conform with notation used later.) If in (2.3) terms of higher degree than the first in κ are neglected we arrive at the expressions (2.1) for the stress in a Newtonian fluid. This can accordingly be regarded as the first-order approximation to the stress associated with slow time-independent rectilinear shearing flow in a non-Newtonian fluid. Equations (2.3) can be regarded as the second-order approximation. Third and

higher order approximations can be obtained by including in the approximations to the functions f, f_1, f_2 terms of degree three or higher in κ. For example, in the third-order theory we must add to the expression for σ_{12} in (2.3) a term constant $\times \kappa^3$.

If κ is not time-independent, then the functions f, f_1 and f_2 in (2.2) may depend not only on κ, but also on time derivatives of κ of various orders, or alternatively on the history $\kappa(\tau)$ of the velocity gradient up to and including the instant at which $\underset{\sim}{\sigma}$ is measured. In either case the parity of these functions must be such that f changes sign and f_1, f_2 remain unchanged when κ and its time derivatives change sign simultaneously, or when $\kappa(\tau)$ changes sign. The dependence of stress on the time derivatives of κ, or on its history, implies many interesting effects which are not present in Newtonian fluids.

3. VISCOMETRIC FLOWS. There are a number of simple flows which can be generated in non-Newtonian fluids, without the application of body forces, in which the flow, referred to a local rectangular cartesian coordinate system rotating with an appropriate angular velocity, is a rectilinear shearing flow. The stress referred to this coordinate system, or to a local inertial coordinate system instantaneously coinciding with it, is then given by (2.2) if appropriate substitution is made for the velocity gradient κ.

Examples of such flows are:

(i) Poiseuille flow in a straight pipe of circular cross-section, or in the annular region between two coaxial circular pipes, under a constant pressure gradient.

(ii) Couette flow in the annular region between two infinite coaxial cylinders, due to the rotation of the outer cylinder, the inner cylinder, or both, with constant angular velocities.

(iii) Rectilinear shearing flow between infinite coaxial cylinders, in which the flow is produced by the relative uniform motion, with constant longitudinal velocity, of one cylinder relative to the other.

(iv) Superpositions of Couette flow, rectilinear shearing flow, and Poiseuille flow between infinite coaxial cylinders are also viscometric flows.

(v) Torsional flow in which a circular cylindrical mass of fluid is contained between rigid discs, one of which is held stationary, while the other is rotated with constant angular velocity. (This is

strictly a viscometric flow only if the effect of inertial forces is neglected.)

(vi) Biconical torsional flow in which the fluid is contained between two infinite cones with common apexes and axes, one or both of which are rotated with constant angular velocities. As a special case one of the cones may be a flat plate. (This is not strictly a viscometric flow, but is very nearly so if inertial forces are neglected and the semi-vertical angles of the cones are nearly equal.)

For the viscometric flows, the forces and velocity fields resulting from given boundary conditions can be calculated by using equations (2.2) for the stress.

For example, in the case of Poiseuille flow in a straight circular pipe, we find that the fluid particles flow down the pipe in rectilinear paths, with velocity $w(r)$ at radial distance r from the axis of the pipe, where w is given by the differential equation

$$w'f(w'^2) = \frac{1}{2} Pr \ , \tag{3.1}$$

with $w = 0$ at the wall of the pipe. The prime denotes differentiation with respect to r, and P denotes the pressure gradient along the pipe. The longitudinal normal stress component is then given by

$$- f_2 - \int^r f_2/r \, dr + Pz \ , \tag{3.2}$$

where $f_2 = f_2(w'^2)$, and z is distance measured along the pipe. Equation (3.1) leads, in general, to a non-parabolic distribution of velocity over the cross-section of the pipe and a non-proportionality between the rate of discharge and the pressure gradient. Equation (3.2) implies that, in general, the normal force per unit area exerted by the fluid over a cross-section of the pipe varies with r. For a Newtonian fluid it is, of course, constant.

For Couette flow the angular velocity ω varies with radial distance r from the axis of the cylinders in accordance with the formula

$$2\pi r^3 \omega' f(r^2\omega'^2) = M \ , \tag{3.3}$$

where M is the couple per unit length exerted on the cylinders, and ω takes the specified values on the cylindrical boundaries. One of the more interesting

new features which arises is that the normal forces which are exerted by the fluid on the bounding cylinders are no longer equal, even if inertial forces are neglected, as they are in the case of a Newtonian fluid. The difference in the radial normal stress components at the inner and outer cylinders is given by

$$\sigma_{22}\Big|_{R_2}^{R_1} = -\int_{R_2}^{R_1}\left(\frac{f_2-f_1}{r} + \rho r\omega^2\right)dr \ , \tag{3.4}$$

where R_1 and R_2 are the radii of the outer and inner cylinders respectively and ρ is the density of the fluid. Also, if we calculate the longitudinal normal stress component σ_{33}, we obtain

$$\sigma_{33} = -f_2 - \int^r \left[\frac{1}{r}(f_2-f_1) + \rho r\omega^2\right]dr \ . \tag{3.5}$$

The term $\rho r\omega^2$ in the integral is an inertial term and is present whether or not the fluid is Newtonian. The remaining terms in (3.5) vanish if the fluid is Newtonian. An important effect of the non-constancy of σ_{33} even when inertial forces are neglected arises if we consider the axis of the system to be vertical and the fluid to have initially a force-free horizontal surface. Then, the rotation of the inner or outer cylinder may be expected to result in a distortion of the free-surface, apart from that due to inertia. In practice it is generally found that the fluid rises at the inner cylinder and falls at the outer cylinders. This effect is called the rod-climbing effect and was discovered by Garner and Nissan [1] during the Second World War. It was this discovery which was largely responsible for attracting attention to the whole problem of the mechanics of viscoelastic fluids. Results equivalent to those described above were first given by Rivlin [3]. Their derivation was subsequently modified and simplified by a number of people.

Recently the rod-climbing experiment has been analyzed by Joseph and his collaborators [4,5,6] and investigated experimentally in the case when a cylinder of radius R rotates in a half-space of the non-Newtonian fluid ($R_2=R$, $R_1=\infty$). For sufficiently slow flows the profile of the free-surface is given approximately by

$$h = -\sigma_{33}/\rho g + \text{constant} \ , \tag{3.6}$$

where σ_{33} is given by (3.5) and the constant is adjusted so that the constancy of volume of the fluid is preserved. g denotes the gravitational acceleration

and h denotes the height of rise of the fluid above its level when the cylinder is not rotating. f, f_1 and f_2 in (3.4) and (3.5) are replaced by their expressions in the second-order equations (2.3). More accurate calculations were also carried out by Joseph and his collaborators by taking into account the effect of surface tension. They found excellent agreement between the measured and calculated profiles.

In (3.6), ρ denotes strictly the difference between the density of the non-Newtonian fluid and air. A very striking effect arises if a Newtonian fluid of density $\bar{\rho}$ say, very slightly less than that of the non-Newtonian fluid and immiscible with it, is floated on the non-Newtonian fluid. Then in (3.6) we must replace ρ by $\rho - \bar{\rho}$ and it is seen [7] that h can be amplified virtually without limit by choosing the densities of the Newtonian and non-Newtonian fluids to be sufficiently close. Of course, in practice, a limit will be set by the instability of the profile when the magnitude of the gradient dh/dr becomes too large.

4. GENERAL THEORY. The constitutive equations in §2 are limited to flows which are, at any rate locally, rectilinear shearing flows. More general constitutive equations can be formulated in a variety of ways. We shall consider here the constitutive equation for the flow of a non-Newtonian fluid formulated by Rivlin and Ericksen in 1955 [2]. The constitutive assumption which provides the starting point is that the stress $\underset{\sim}{\sigma}$ depends not only on the velocity gradient $\underset{\sim\sim}{\nabla v}$, as in the case of a Newtonian fluid, but also on the gradients of time derivatives of the velocity of various orders, thus:

$$\underset{\sim}{\sigma} = \underset{\sim}{\sigma}(\underset{\sim\sim}{\nabla v}, \underset{\sim\sim}{\nabla \dot{v}}, \ldots, \underset{\sim\sim}{\nabla v}^{(\mu)}) - p\underset{\sim}{\delta} \ . \tag{4.1}$$

The term $-p\underset{\sim}{\delta}$ is introduced to express the fact that for an incompressible fluid the flow is unchanged by the addition of a hydrostatic pressure.

If we superpose on the assumed deformation an arbitrary time-dependent rigid rotation, the stress is rotated by a corresponding amount. This fact leads to a restriction on the manner in which the stress can depend on the kinematic gradients $\underset{\sim\sim}{\nabla v}$, $\underset{\sim\sim}{\nabla \dot{v}}$, etc. It is found that they must depend on them through the so-called Rivlin-Ericksen tensors $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_\mu$ defined by

$$\begin{aligned} \underset{\sim}{A}_1 &= \frac{1}{2}\left[\underset{\sim\sim}{\nabla v} + (\underset{\sim\sim}{\nabla v})^{\dagger}\right] , \\ \underset{\sim}{A}_{\alpha+1} &= \frac{\partial \underset{\sim}{A}_\alpha}{\partial t} + \underset{\sim}{v}\cdot\underset{\sim\sim}{\nabla A}_\alpha + (\underset{\sim\sim}{\nabla v})\underset{\sim}{A}_\alpha + [(\underset{\sim\sim}{\nabla v})\underset{\sim}{A}_\alpha]^{\dagger} , \end{aligned} \tag{4.2}$$

where the dagger denotes the transpose. Thus,

$$\underset{\sim}{\sigma} = \underset{\sim}{\sigma}(\underset{\sim}{A}_1, \underset{\sim}{A}_2, \ldots, \underset{\sim}{A}_\mu) - p\underset{\sim}{\delta} \ . \tag{4.3}$$

Isotropy of the fluid leads to restrictions on the manner in which the stress can depend on the Rivlin-Ericksen tensors. Canonical forms expressing these restrictions have been obtained by Spencer and Rivlin (see, for example, the review article by Spencer [8]). They are rather complicated and will not be given here. For rectilinear shearing flow the constitutive equation (4.3), together with the restrictions imposed by isotropy, reduce to equations (2.2).

Generally boundary value problems based on the constitutive equation (4.3) are non-linear and their solution presents formidable difficulties. Nevertheless, many problems of considerable interest can be solved where linearization of the governing partial differential equation is possible. We mention three classes of problems of this type:

(i) If the material is only slightly non-Newtonian, we can replace equation (4.3) by

$$\underset{\sim}{\sigma} = \alpha_1 \underset{\sim}{A}_1 + \varepsilon \underset{\sim}{\bar{\sigma}}(\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_\mu) - p\underset{\sim}{\delta} \ , \tag{4.4}$$

where ε is a small parameter and $\underset{\sim}{\bar{\sigma}}$ is an isotropic matrix function of the argument matrices. By taking $\varepsilon = 0$, we arrive at the constitutive equation for an incompressible Newtonian fluid. We can solve boundary value problems based on the constitutive equation (4.4) by first solving the corresponding problem for a Newtonian fluid and then obtaining a correction to this solution by a perturbation procedure based on linearization of the governing equations with respect to ε.

(ii) We have seen that it is possible to solve the viscometric flow problems for a non-Newtonian fluid with some generality. Accordingly, problems in which the flows are only slightly different from viscometric flows can be solved, using a perturbation procedure involving linearization in the difference between the actual velocity field and that corresponding to the neighboring viscometric flow.

(iii) It was seen in §2 that, for rectilinear shearing flows, a hierarchy of slow flow approximations to the expressions (2.2) for the stress can be constructed. A corresponding hierarchy based on (4.3) can be set up for more general flows. We assume that the dependence of the stress on the Rivlin-Ericksen tensors is sufficiently smooth so that it can be expressed as an isotropic

matrix polynomial in the latter. We can then write

$$\underset{\sim}{\sigma} = \underset{\sim}{F}_1 + \underset{\sim}{F}_2 + \ldots + \underset{\sim}{F}_\mu - p\underset{\sim}{\delta} \,, \tag{4.5}$$

where

$$\begin{aligned}
&\underset{\sim}{F}_1 = \alpha_1 \underset{\sim}{A}_1 \,, \quad \underset{\sim}{F}_2 = \alpha_2 \underset{\sim}{A}_2 + \alpha_3 \underset{\sim}{A}_1^2 \,, \\
&\underset{\sim}{F}_3 = \beta_1 (\mathrm{tr}\, \underset{\sim}{A}_2)\underset{\sim}{A}_1 + \beta_2 \underset{\sim}{A}_2\underset{\sim}{A}_3 + \beta_3 (\underset{\sim}{A}_1\underset{\sim}{A}_2 + \underset{\sim}{A}_2\underset{\sim}{A}_1) \,, \\
&\underset{\sim}{F}_4 \quad (\gamma_1 \mathrm{tr}\, \underset{\sim}{A}_1^3 + \gamma_2 \mathrm{tr}\, \underset{\sim}{A}_1\underset{\sim}{A}_2 + \gamma_3 \mathrm{tr}\, \underset{\sim}{A}_3)\underset{\sim}{A}_1 \\
&\qquad + (\mathrm{tr}\, \underset{\sim}{A}_2)(\gamma_4 \underset{\sim}{A}_1^2 + \gamma_5 \underset{\sim}{A}_2) + \gamma_6 \underset{\sim}{A}_2^2 + \gamma_7 (\underset{\sim}{A}_1^2\underset{\sim}{A}_2 + \underset{\sim}{A}_2\underset{\sim}{A}_1^2) \\
&\qquad + \gamma_8 (\underset{\sim}{A}_1\underset{\sim}{A}_3 + \underset{\sim}{A}_3\underset{\sim}{A}_1) + \gamma_9 \underset{\sim}{A}_4 \,,
\end{aligned} \tag{4.6}$$

where the α's, β's and γ's are constants and the α's have the same values as those occurring in (2.3). Expressions for $\underset{\sim}{F}_\alpha$ with $\alpha > 4$ can also be readily obtained from the canonical representations of Rivlin and Spencer [8].

We observe, as did Coleman and Noll [9], that $\underset{\sim}{A}_\alpha$, defined in (4.2), is of dimensionality $-\alpha$ in time. Accordingly, $\underset{\sim}{A}_1$ is of dimensionality -1 in time, $\underset{\sim}{A}_2$ and $\underset{\sim}{A}_1^2$ are of dimensionality -2, and so on; i.e. $\underset{\sim}{F}_\alpha$ is formed by taking all the terms in the canonical expression for an isotropic matrix polynomial in the $\underset{\sim}{A}$'s which are of dimensionality $-\alpha$ in time. By considering a sufficiently severe retardation of the given flow, we can approximate the stress by retaining only the term $\underset{\sim}{F}_1$ in (4.5). For a less severe retardation, we retain $\underset{\sim}{F}_1$ and $\underset{\sim}{F}_2$, and so on.

Alternatively [10], in the case when the flow is steady-state, so that $\partial \underset{\sim}{A}_\alpha / \partial t = 0$ in (4.2), we note that $\underset{\sim}{A}_\alpha$ is homogeneous of degree α in $\underset{\sim}{v}$ and its spatial derivatives. Accordingly, $\underset{\sim}{F}_\alpha$ is homogeneous of degree α in $\underset{\sim}{v}$ and its spatial derivatives and the hierarchy of slow flow approximations follows as before. For slow flows we write the velocity $\underset{\sim}{v}$ in the form

$$\underset{\sim}{v} = \varepsilon \underset{\sim}{u}_1 + \varepsilon^2 \underset{\sim}{u}_2 + \varepsilon^3 \underset{\sim}{u}_3 + \ldots \,, \tag{4.7}$$

where ε is a small parameter. Boundary value problems can then be solved for $\underset{\sim}{u}_1$ by linearizing the governing equations in ε (which is equivalent to solving the corresponding problem for a Newtonian fluid). $\underset{\sim}{u}_2$, $\underset{\sim}{u}_3$,... are then obtained as successive regular perturbations. Some examples of the results obtained by this procedure are given in the next section.

5. FLOW IN STRAIGHT PIPES AND RELATED PROBLEMS. If a Newtonian fluid with viscosity $\frac{1}{2}\alpha_1$ is caused to flow in a straight pipe of uniform non-circular cross-section under a uniform pressure gradient P, then, provided the flow is not so fast that it becomes turbulent, each particle of the fluid moves in a rectilinear longitudinal path with velocity w determined by the equation

$$\nabla^2 w = 2P/\alpha_1 \tag{5.1}$$

and the no-slip boundary condition $w = 0$ on the pipe. It was found by Ericksen [11] that if the fluid is non-Newtonian and we assume that, as in the case of a Newtonian fluid, the particles of the fluid travel in rectilinear longitudinal paths, then the differential equations for the determination of the manner in which this longitudinal velocity and the hydrostatic pressure p in the constitutive equation vary over the cross-section of the pipe, have, in general, no solution. The problem of determining the flow field was taken up by Green and Rivlin [12] and by Langlois and Rivlin [13,10]. In [12] and [13] the problem was discussed on the assumption that the fluid is only slightly non-Newtonian and in [10] on the basis of the assumption that the flow is slow, so that the slow flow approximations to the Rivlin-Ericksen constitutive equations discussed in §4 can be used. It was found in [10] that if the flow is sufficiently slow so that the material properties are adequately described by the second-order theory then the flow field is unaltered from that which obtains if the first-order (i.e. Newtonian) theory is used. The third-order theory leads to a change in the detailed distribution of the longitudinal velocity over the cross-section of the pipe. However, the analysis based on the fourth-order theory leads to the conclusion that purely longitudinal flow is no longer possible. Superimposed on the longitudinal flow given by the third-order theory, there is a steady flow in transverse planes, for which the stream-function ψ is given by an equation of the form

$$\nabla^4 \psi = \left(\frac{P}{\alpha_1}\right)^4 \Gamma \phi(x_1, x_2) \ , \tag{5.2}$$

where

$$\Gamma = -\frac{2(\beta_1+\beta_3)}{\alpha_1}\left(\alpha_2 + \frac{1}{4}\alpha_3\right) - \frac{1}{2}(\gamma_4+4\gamma_5+4\gamma_6+2\gamma_7) \tag{5.3}$$

and the function ϕ depends on the shape of the cross-section and can be

calculated explicitly from the third-order solution. In the case when the pipe has an elliptical cross-section

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \ , \tag{5.4}$$

ψ is given by

$$\psi = - \frac{H(b^2x_1^2+a^2x_2^2-a^2b^2)}{3\alpha_1(5a^4+6a^2b^2+5b^4)} \ , \tag{5.5}$$

where

$$H = -12\left(\frac{P}{\alpha_1}\right)^4 \Gamma \frac{a^2b^2(a^2-b^2)}{(a^2+b^2)^3} \ . \tag{5.6}$$

The stream-function (5.5) represents an eddy in each of the four quadrants of the elliptical cross-section.

We note that the transverse flow field involves only a single combination of the constants occurring in the fourth-order Rivlin-Ericksen equations. It has recently been shown [14] that this is generally true if we calculate, according to the fourth-order theory, the transverse secondary flow associated with any steady antiplane primary flow.

Another interesting effect which arises in the flow of a non-Newtonian fluid through a pipe of non-circular cross-section was analyzed by Pipkin and Rivlin [15].

For a Newtonian fluid the normal force exerted, per unit area, by the fluid on the pipe does not change as we move round the periphery of a cross-section of the pipe. This is no longer the case if the fluid is non-Newtonian. Indeed, it was shown in [15] that it is given by

$$- \frac{2(\alpha_2+\frac{1}{4}\alpha_3)}{\alpha_1^2} P^2 \underset{\sim}{\nabla w} \cdot \underset{\sim\sim}{\nabla w} + Px_3 + \text{constant} \ , \tag{5.7}$$

where x_3 is distance measured along the tube.

Now suppose that the elliptical pipe is tilted at a small angle β to the horizontal, the axes of the elliptical cross-sections of length $2a$ also being horizontal. Suppose also that the fluid flows down the pipe slowly under the action of gravity. The thrust τ per unit area exerted on the tilted

mid-plane in the pipe by the fluid above it can be calculated on the basis of the second-order theory as

$$\tau = 2(\alpha_2 + \frac{1}{4}\alpha_3)\left[\frac{\rho g b \sin\beta}{\alpha_1(a^2+b^2)}\right]^2 \left[(2b^2+a^2)x_1^2 - a^2(a^2+b^2)\right] + \rho g x_3 \sin\beta , \qquad (5.8)$$

where x_1 denotes the distance measured horizontally from the center of the cross-section and x_3 is the length measured along the tube from its highest point to the cross-section considered. It follows, from reasoning similar to that employed in discussing the free-surface profile in the rod-climbing experiment, that for slow flow of a non-Newtonian fluid in a tilted channel with semi-elliptical cross-section, the free surface of the fluid filling the channel will, in general, have a curved profile, the height h of this profile above some ambient level being given approximately by

$$h = \frac{2(\alpha_2 + \frac{1}{4}\alpha_3)}{\rho g}\left[\frac{\rho g b \sin\beta}{\alpha_1(a^2+b^2)}\right]^2 \left[(2b^2+a^2)x_1^2 - a^2(a^2+b^2)\right] + \text{constant} . \qquad (5.9)$$

The constant in (5.9) is, of course, determined by the condition that the cross-sectional area of the fluid remains that of the elliptical cross-section. If the fluid is Newtonian, $\alpha_2 + \frac{1}{4}\alpha_3 = 0$ and the profile is no longer curved. For non-Newtonian fluids which have been studied experimentally, it is found that $\alpha_2 + \frac{1}{4}\alpha_3 < 0$, so that, from (5.9), the fluid rises above the ambient level at the center of the free-surface and falls at its edges. Of course, analogous results follow for channels with other cross-sectional shapes. The calculation of the free-surface profiles was first carried out by Wineman and Pipkin [16] and later modified by Sturges and Joseph [17] to include the effects of surface tension. The effect was studied experimentally by Tanner [18].

One may expect that if the calculations were carried out on the basis of the fourth-order Rivlin-Ericksen theory, one would find that a steady secondary flow in the cross-sectional planes will be imposed on the longitudinal flow. That this is indeed the case was shown by Sturges and Joseph [17] except in the case when the channel has a semi-circular cross-section. In that case secondary transverse flows arise only when the sixth-order constitutive equations are adopted. The particular status of the semi-circular cross-section is less surprising when we consider that no transverse secondary flows arise in Poiseuille flow of a non-Newtonian fluid through a pipe of circular cross-section.

We have seen that in Poiseuille flow of a non-Newtonian fluid through a

straight pipe of non-circular cross-section, the velocity distribution, calculated according to the second-order theory, is the same as that given by the first-order - i.e. Newtonian - theory. However, the force per unit area exerted normally by the fluid on the pipe varies as we move round the periphery of a cross-section. The resultant force exerted by the fluid in a transverse direction is however necessarily zero.

We now consider the fluid to flow in the annular region between two infinite circular cylinders with parallel axes, either as the result of the uniform relative longitudinal motion of the cylinders, or as a result of the application of a uniform longitudinal pressure gradient. In either case, it is found that the fluid exerts a resultant transverse force on the inner cylinder and, of course, an equal and opposite transverse force on the outer cylinder.

If the flow is due to a uniform relative motion of the cylinders with velocity V, then the force on the inner cylinder, measured per unit length and calculated on the basis of the second-order constitutive equation, is given by [19]

$$T(\alpha_2 + \tfrac{1}{4}\alpha_3)V^2/R_1 , \tag{5.10}$$

where T is a non-dimensional function of $\bar{R} = R_2/R_1$, the ratio between the radii of the inner and outer cylinders respectively, and of $\varepsilon/(R_1 - R_2)$, where ε is the distance between the cylinder axes. The force is positive if it tends to drive the inner cylinder towards coaxiality with the outer cylinder and is negative if it tends to drive it towards contact. The manner in which T depends on $\bar{R}$ and $\bar{\varepsilon}$ is shown in Fig.1(a). We note that T is negative and, since we may expect that in practical situations $\alpha_2 + \frac{1}{4}\alpha_3$ will be negative, we see that the force tends to drive the cylinders towards coaxiality. This contrasts with the situation which arises when the flow results from a uniform pressure gradient P. Then, the transverse force per unit length is given by [20]

$$T(\alpha_2 + \tfrac{1}{4}\alpha_3)P^2R_1^3 , \tag{5.11}$$

where T is again a non-dimensional function of $\bar{R}$ and $\bar{\varepsilon}$, its dependence on which is shown in Fig.1(b). We see that if $\alpha_2 + \frac{1}{4}\alpha_3 < 0$, this force tends to drive the cylinders towards contact.

6. FLOW BETWEEN ROTATING ECCENTRIC CYLINDERS AND RELATED PROBLEMS. We now suppose that in the eccentric cylinder arrangement of §5, plane flow is

produced in the annular region between the cylinders by rotating one or both of the cylinders about their respective axes with constant angular velocities. According to a theorem due to Tanner [21], the flow field in a steady plane problem, in which the velocities are specified on the boundaries, is the same whether calculated on the basis of the first-order or second-order constitutive equations*. In particular, this applies to the flow field in the annular region between the cylinders. However, the stress fields and consequently the forces exerted on the cylinders by the fluid are different. It is found that, on the basis of the second-order theory, and with the neglect of inertial forces, a resultant transverse force is exerted by the fluid on the inner cylinder. This is given, per unit length of the cylinder, by [23]

$$\alpha_2 R_1 (\bar{T}\Omega_1^2 + \tilde{T}\Omega_1\Omega_2 + \hat{T}\Omega_2^2) \ , \tag{6.1}$$

where $\bar{T}$, $\tilde{T}$ and $\hat{T}$ are the non-dimensional functions of $\bar{R}$ and $\bar{\epsilon}$ plotted in Fig.2, and Ω_1 and Ω_2 are the constant angular velocities of the outer and inner cylinders. If this force is positive it tends to drive the cylinders towards coaxiality. In practice α_2 is negative and, from Fig.2, it is seen that $\bar{T}$, $\tilde{T}$ and $\hat{T}$ are all negative. So, provided that Ω_1 and Ω_2 are of the same sign (i.e. the cylinders are rotating in the same sense) the transverse, or lift, force tends to drive the inner cylinder towards coaxiality with the outer. In fact it can be shown numerically that this is also the case when Ω_1 and Ω_2 are of opposite signs.

We may contrast the manner in which the lift force depends on $\bar{R}$ and $\bar{\epsilon}$ in the case discussed with that which obtains when the fluid is Newtonian but inertial forces are not neglected. The lift force, per unit length of the cylinder, is then given, in the linearized inertial approximation, by [24,25]

$$\rho R_1^3 (\bar{T}\Omega_1^2 + \tilde{T}\Omega_1\Omega_2 + \hat{T}\Omega_2^2) \ , \tag{6.2}$$

where $\bar{T}$, $\tilde{T}$ and $\hat{T}$ are the non-dimensional functions of $\bar{R}$ and $\bar{\epsilon}$ plotted in Fig.3, ρ is the density of the fluid and μ is its viscosity. It is seen that $\tilde{T}$ is negative, while both $\bar{T}$ and $\hat{T}$ are negative for the lower

---

* Strictly, Tanner proved this theorem only in the case when inertial effects are neglected. Recently, however, a slight extension of the theorem was made by Kazakia and Rivlin [22] which renders it also applicable to linearized inertial approximations to the flow fields.

eccentricities and positive for the higher eccentricities. The lift force accordingly tends to drive the inner cylinder towards some intermediate position between coaxiality and contact with the outer cylinder.

The lift force (6.1) has been calculated for a non-Newtonian fluid on the basis of the second-order theory and with the total neglect of inertia. The lift force (6.2) has been calculated for a Newtonian fluid with the linearized inertial approximation. It has been shown [22] that in order to obtain the lift force for a non-Newtonian fluid on the basis of the second-order theory, but including the effect of inertia in the linearized inertial approximation, we have merely to add the forces given in (6.1) and (6.2), with $\mu = \frac{1}{2}\alpha_1$.

From the above results we can derive [25] corresponding results for some different related systems. Suppose that we consider our eccentric cylinder system to sit on a turntable rotating about the axis of the outer cylinder with constant angular velocity $-\bar{\Omega}$. In order that the motion of the fluid, referred to a coordinate system fixed to the turntable, shall be same as that which obtains in the previous problem, body forces must be applied throughout the fluid to balance the inertial and Coriolis forces which are called into play by the motion of the turntable. It emerges that these are derivable from a potential function and can, accordingly, be absorbed into the hydrostatic pressure term in the expression for the stress. The effect of the rotation of the turntable is thus to superpose on the motion previously obtained a rigid rotation with angular velocity $-\bar{\Omega}$ about the axis of the outer cylinder and to modify the stress field and consequently the forces exerted by the fluid on the cylinders. Again from the point of view of an inertial reference system, the outer cylinder is rotating with angular velocity $\Omega_1-\bar{\Omega}$ about its axis, while the inner cylinder is rotating about its axis with angular velocity Ω_2 and is simultaneously executing a planetary motion about the axis of the outer cylinder with angular velocity $-\bar{\Omega}$. This may be regarded as an idealized stirrer arrangement.

Another interesting result can be obtained by considering the limiting situation in which $R_1 \to \infty$, $\varepsilon \to \infty$ and $\bar{\Omega} = \Omega_1 \to 0$ in such a way that $R_1\Omega_1$ has a constant value, V say, and the value of $R_1 - \varepsilon$ remains constant. The configuration which is thus obtained is an infinite cylinder of radius R_2 moving, in an infinite half-space of the fluid bounded by a rigid wall, with velocity V parallel to the wall and perpendicular to its own length, while simultaneously rotating about its axis with angular velocity Ω_2 . While the calculated forces and flow field obtained in this way are correct if inertia is

neglected, they do not properly take account of inertial effects. This is, no doubt, due to the fact that the expression of the solution in powers of the Reynolds number, on which the linearized inertial approximation is based, breaks down when the fluid is of infinite extent.

7. EFFECT OF VIBRATION ON POISEUILLE FLOW. In previous sections we have discussed problems in which the flows are either steady or derivable from steady flows by the superposition of a rigid motion. In the present section we discuss a class of unsteady flows. The analysis of these flows was motivated by an experiment of Manero and Mena [26]. In their experiment a polymer solution flows through a straight pipe of circular cross-section under a constant pressure gradient. Simultaneously, the tube is subjected to a longitudinal vibration and the effect of this vibration on the time-averaged rate of discharge of the fluid is measured. It is found that it may be increased many times by the vibration. In contrast, the time-averaged rate of discharge is unchanged if the fluid is Newtonian.

If the fluid is Newtonian and has viscosity η, then the velocity of the fluid is longitudinal and, at a distance r from the axis of the pipe, is given by

$$v = \frac{P}{4\eta}(a^2-r^2) + L(r)\sin \omega t + M(r)\cos \omega t , \tag{7.1}$$

where

$$\begin{aligned} L(r) &= V \frac{\mathrm{ber}\, \nu a\, \mathrm{ber}\, \nu r + \mathrm{bei}\, \nu a\, \mathrm{bei}\, \nu r}{\mathrm{ber}^2 \nu a + \mathrm{bei}^2 \nu a} , \\ M(r) &= V \frac{\mathrm{ber}\, \nu a\, \mathrm{bei}\, \nu r - \mathrm{bei}\, \nu a\, \mathrm{ber}\, \nu r}{\mathrm{ber}^2 \nu a + \mathrm{bei}^2 \nu a} , \end{aligned} \tag{7.2}$$

with

$$\nu^2 = \rho\omega/\eta . \tag{7.3}$$

In these equations a denotes the radius of the pipe and the vibration velocity of the pipe is $V \sin \omega t$.

It is easily shown that for a non-Newtonian fluid the velocity of the fluid particles is longitudinal, whatever the constitutive equation adopted. Moreover, it emerges that we need only consider the constitutive equation for the shear component, σ say, of the stress in order to calculate the mean rate of discharge of the fluid. We may write this constitutive equation in the form

$$\sigma = \eta\kappa + \varepsilon F[\kappa(\tau)] \ , \tag{7.4}$$

where $\kappa(\tau)$ is the velocity gradient at time τ and F denotes a functional of the history of the velocity gradient from time $\tau = -\infty$ to the time t at which the stress is measured. η and ε are constants. Alternatively, we may write the constitutive equation for σ in the form

$$\sigma = \eta\kappa + \varepsilon f(\kappa,\dot{\kappa},\ddot{\kappa},\ldots) \ , \tag{7.5}$$

where $\kappa,\dot{\kappa},\ddot{\kappa},\ldots$ are the velocity gradient and its time derivatives of various orders measured at time t .

The mean rate of discharge Q is then given by [27]

$$Q = \frac{\pi P a^4}{8\eta} + \frac{\varepsilon\omega}{2\eta} \int_0^a \int_0^{2\pi/\omega} r^2 F[\kappa(\tau)]\,dt\,dr \ , \tag{7.6}$$

or

$$Q = \frac{\pi P a^4}{8\eta} + \frac{\varepsilon\omega}{2\eta} \int_0^a \int_0^{2\pi/\omega} r^2 f(\kappa,\dot{\kappa},\ddot{\kappa},\ldots)\,dt\,dr \ . \tag{7.7}$$

If ε is sufficiently small we can calculate Q by replacing the actual velocity gradient field by that appropriate to the Newtonian case, i.e. by the velocity gradient field calculated from (7.1). In this way the effect on the time-averaged rate of discharge of various types of term in the constitutive equation for a slightly non-Newtonian fluid can be examined. For example, it is seen that if the constitutive equation

$$\sigma = \eta\kappa + \varepsilon(\tilde{\eta}\dot{\kappa} + \hat{\eta}\dot{\kappa} + \bar{\eta}\kappa) \ , \tag{7.8}$$

where the η's are constants, is adopted, then Q is given by [27]

$$Q = \frac{\pi P a^4}{8\eta} - \varepsilon \frac{\pi P^3 a^6 \eta}{48\eta^4} \left[1 + 36\left(\frac{\eta V}{Pa^2}\right)^2 \Delta\right] \ , \tag{7.9}$$

where Δ is defined by

$$\begin{aligned} &a\nu(\mathrm{ber}^2 a\nu + \mathrm{bei}^2 a\nu)(\Delta+1) \\ &\quad = a^2\nu^2(\mathrm{ber}\ a\nu\ \mathrm{ber}'a\nu + \mathrm{bei}\ a\nu\ \mathrm{bei}'a\nu) \\ &\qquad + 2(\mathrm{ber}\ a\nu\ \mathrm{bei}'a\nu - \mathrm{bei}\ a\nu\ \mathrm{ber}'a\nu) \end{aligned} \tag{7.10}$$

and ν is given by

$$\nu^2 = \rho\omega/\eta \ . \tag{7.11}$$

We note that, of the non-Newtonian terms in (7.8), only the term in κ^3 contributes to the change of the time-averaged rate of discharge as a result of the superposed vibration. However, this term contributes to Q through two terms, one of which is independent of the presence of a superposed vibration and the other is not. We note that both of these terms give increases in Q if $\bar{n}$ is negative, as we may expect it to be for a polymer solution.

The analysis of the above problem has led to a number of further predictions regarding the effect, on the time-averaged rate of discharge of a polymer solution, of superposed vibrations [27,28]. For example, we may also expect that rotational vibration of the pipe will increase the time-averaged rate of discharge. Also, non-zero time-averaged rates of discharge may be obtained, even when no pressure gradient is applied, if two, or more, sinusoidal vibrations, with appropriately related frequencies, are applied to the pipe.

Indeed, the calculations in [27,28] draw attention to the whole range of rectification and modulation phenomena in non-Newtonian fluids. These will differ from the usual discussions of rectification and modulation phenomena in many other areas of physics, as a result of their possible three-dimensional character.

For example, Kazakia and Rivlin [29] have studied the effect of a superposed longitudinal vibration on the flow of a non-Newtonian fluid, under a uniform pressure gradient, through a straight pipe of non-circular cross-section. They found that if the second-order Rivlin-Ericksen constitutive equation is adopted, it is predicted that a secondary flow in transverse planes will be superposed on the rectilinear flow. This is, of course, an unsteady flow, but has a non-zero time-averaged velocity field. The stream-lines for this field are shown in Fig.4 in the case when the cross-section of the pipe is rectangular with various aspect ratios. It is seen that in the case when the cross-section is square there are two eddies in each quadrant. As the aspect ratio changes from unity one of these eddies grows at the expense of the other and eventually seems to disappear. It is not, however, certain that it does disappear, since the stream-line patterns shown in Fig.4 were obtained by a finite element calculation and we cannot be certain that there is no eddy of dimensions smaller than those of the finite elements.

ACKNOWLEDGEMENT. The preparation of this paper was supported by a grant DAAG 29-77-G-0204 from the Army Research Office to Lehigh University.

REFERENCES.

[1] F.H. Garner and A.H. Nissan, Nature 158, 634 (1946).

[2] R.S. Rivlin and J.L. Ericksen, J. Rational Mech. Anal. 4, 323 (1955).

[3] R.S. Rivlin, J. Rational Mech. Anal. 5, 179 (1956).

[4] D.D. Joseph and R.L. Fosdick, Arch. Rational Mech. Anal. 49, 321 (1973).

[5] D.D. Joseph, G.S. Beavers and R.L. Fosdick, Arch. Rational Mech. Anal. 49, 381 (1973).

[6] G.S. Beavers and D.D. Joseph, J. Fluid Mech. 69, 475 (1975).

[7] G.S. Beavers and D.D. Joseph, J. Fluid Mech. 8, 265 (1977).

[8] A.J.M. Spencer, Theory of Invariants in "Continuum Mechanics" ed. A.C. Eringen, Academic Press, New York 1, 239 (1971).

[9] B.D. Coleman and W. Noll, Arch. Rational Mech. Anal. 6, 355 (1960).

[10] W.E. Langlois and R.S. Rivlin, Rend. Mat. 22, 169 (1963).

[11] J.L. Ericksen, Q. Applied Math. 14, 318 (1956).

[12] A.E. Green and R.S. Rivlin, Q. Applied Math. 14, 299 (1956).

[13] W.E. Langlois and R.S. Rivlin, Brown University Technical Report No.3, December 1959.

[14] R.S. Rivlin, J. Non-Newtonian Fluid Mech. 1, 391 (1976).

[15] A.C. Pipkin and R.S. Rivlin, Z. Angew. Math. Phys. 14, 738 (1963).

[16] A.S. Wineman and A.C. Pipkin, Acta Mechanica 2, 104 (1966).

[17] L. Sturges and D.D. Joseph, Arch. Rational Mech. Anal. 59, 359 (1975).

[18] R.I. Tanner, Trans. Soc. Rheol. 14, 483 (1970).

[19] B.Y. Ballal and R.S. Rivlin, Rheologica Acta 14, 484 (1975).

[20] B.Y. Ballal and R.S. Rivlin, Rheologica Acta (in the press).

[21] R.I. Tanner, Phys. Fluids 9, 1246 (1966).

[22] J.Y. Kazakia and R.S. Rivlin, J. Non-Newtonian Fluid Mech. 2, 151 (1977).

[23] B.Y. Ballal and R.S. Rivlin, Trans. Soc. Rheology 20, 65 (1976).

[24] B.Y. Ballal and R.S. Rivlin, Arch. Rational Mech. Anal. 62, 237 (1976).

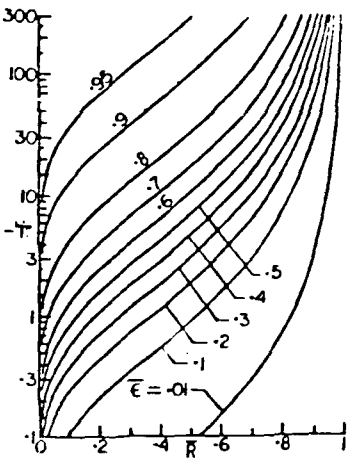
[25] J.Y. Kazakia and R.S. Rivlin, Studies in Applied Math. 58, 209 (1978).

[26] O. Manero and B. Mena, Rheologica Acta 16, 573 (1977).

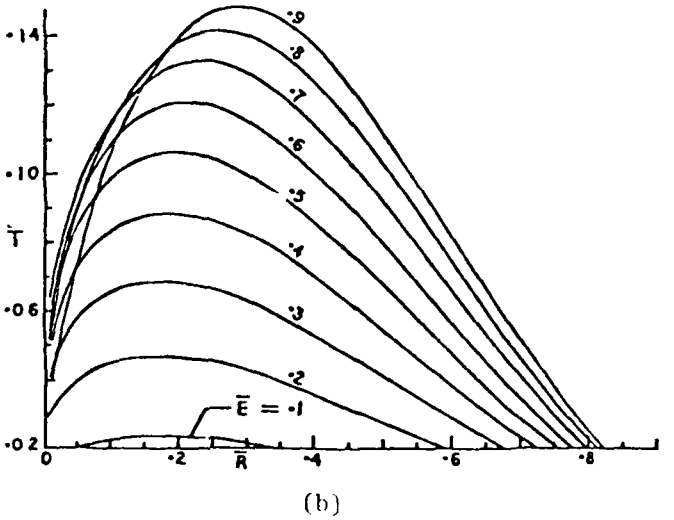
[27] J.Y. Kazakia and R.S. Rivlin, Rheologica Acta 17, 210 (1978).

[28] J.Y. Kazakia and R.S. Rivlin, Rheologica Acta 18, 244 (1979).

[29] J.Y. Kazakia and R.S. Rivlin, J. Non-Newtonian Fluid Mech. 6, 145 (1979).



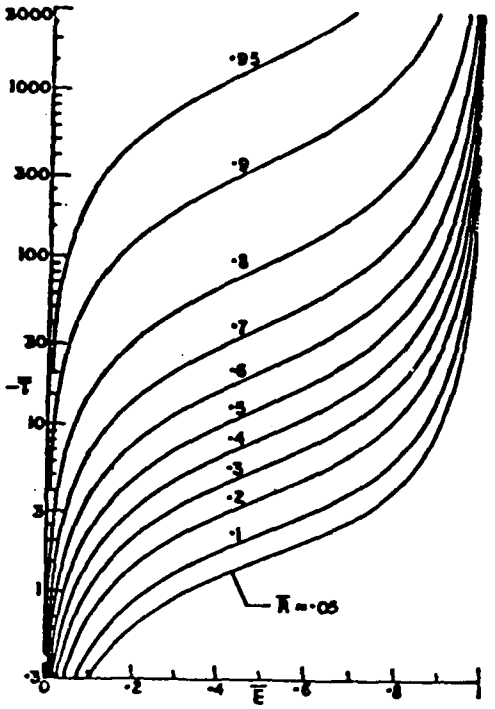
(a)

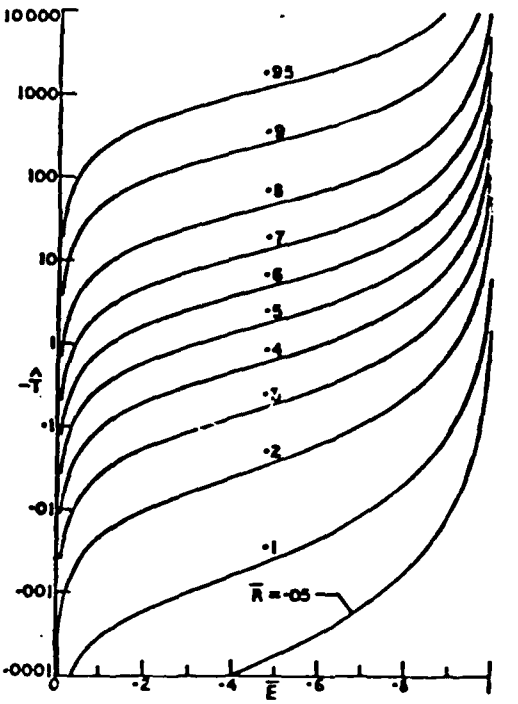


(b)

Fig.1 Plot of $\bar{T}$ vs. $\bar{R}$ for various values of $\bar{\epsilon}$
(a) rectangular shearing flow
(b) Poiseuille flow

Reprinted from Rheologica Acta





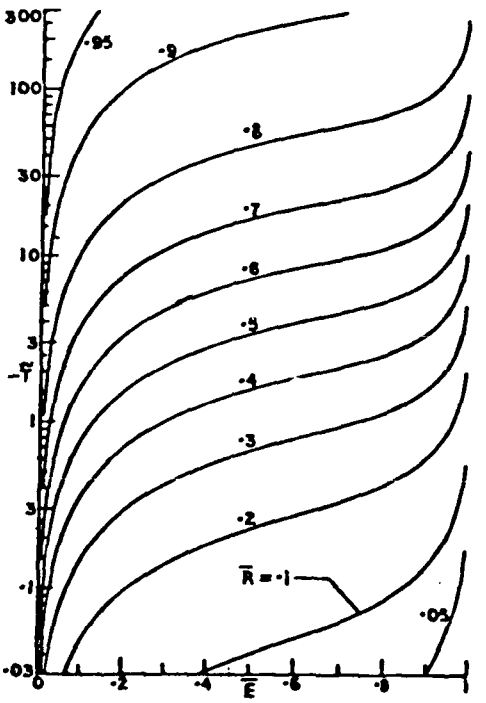


Fig.2 Plot of $\bar{T}$, $\tilde{T}$ and $\hat{T}$ vs $\bar{E}$ for various values of $\bar{R}$, for non-Newtonian fluid in annular region between rotating eccentric cylinders.

Reprinted from Trans. Soc. Rheology

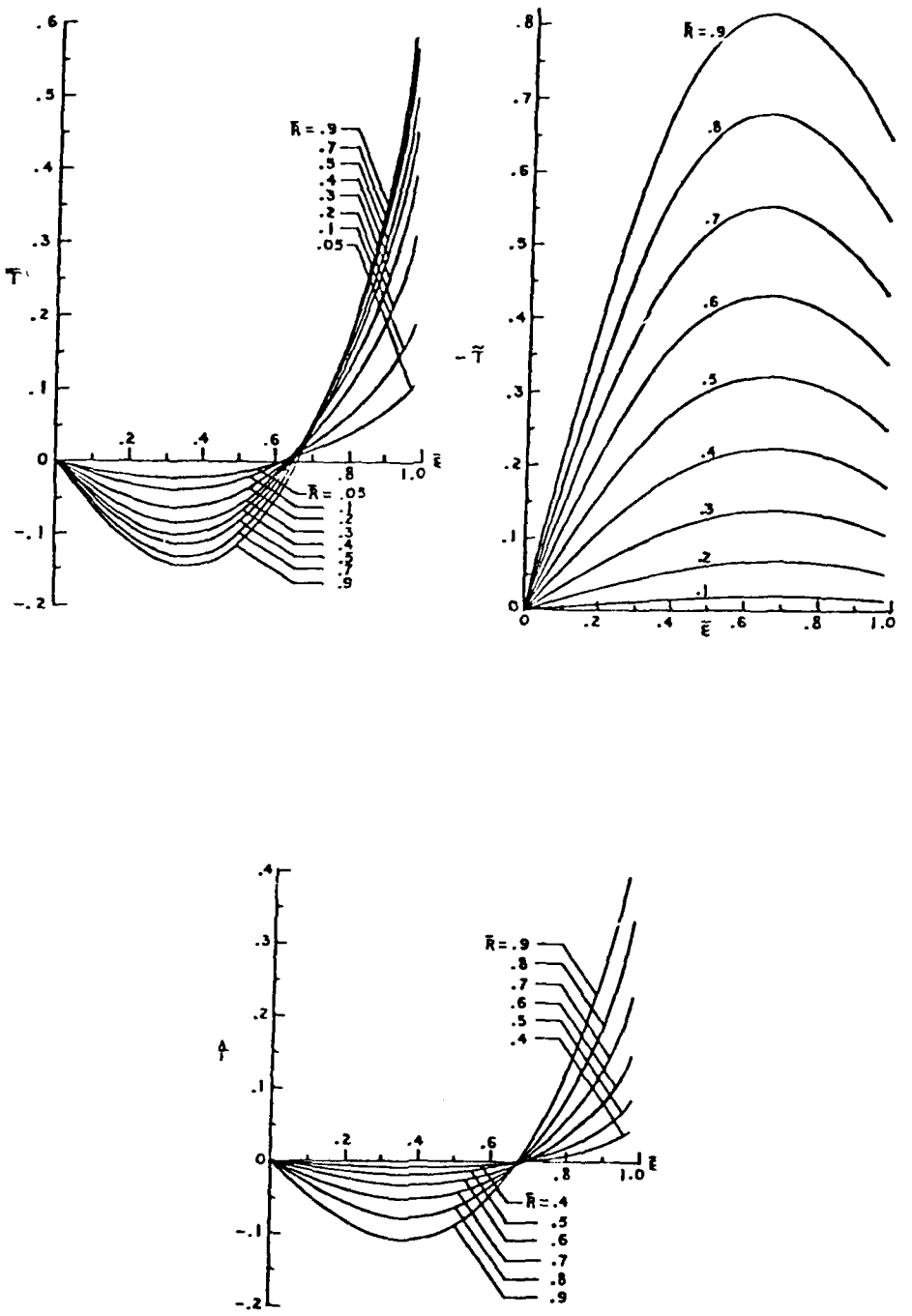


Fig. 5 Plot of $\bar{T}$, $\tilde{T}$ and $\hat{T}$ vs $\bar{\varepsilon}$ for various values of $\bar{R}$, for Newtonian fluid in annular region between rotating eccentric cylinders (linear inertial approximation).

Reprinted from Archive for Rational Mechanics and Analysis

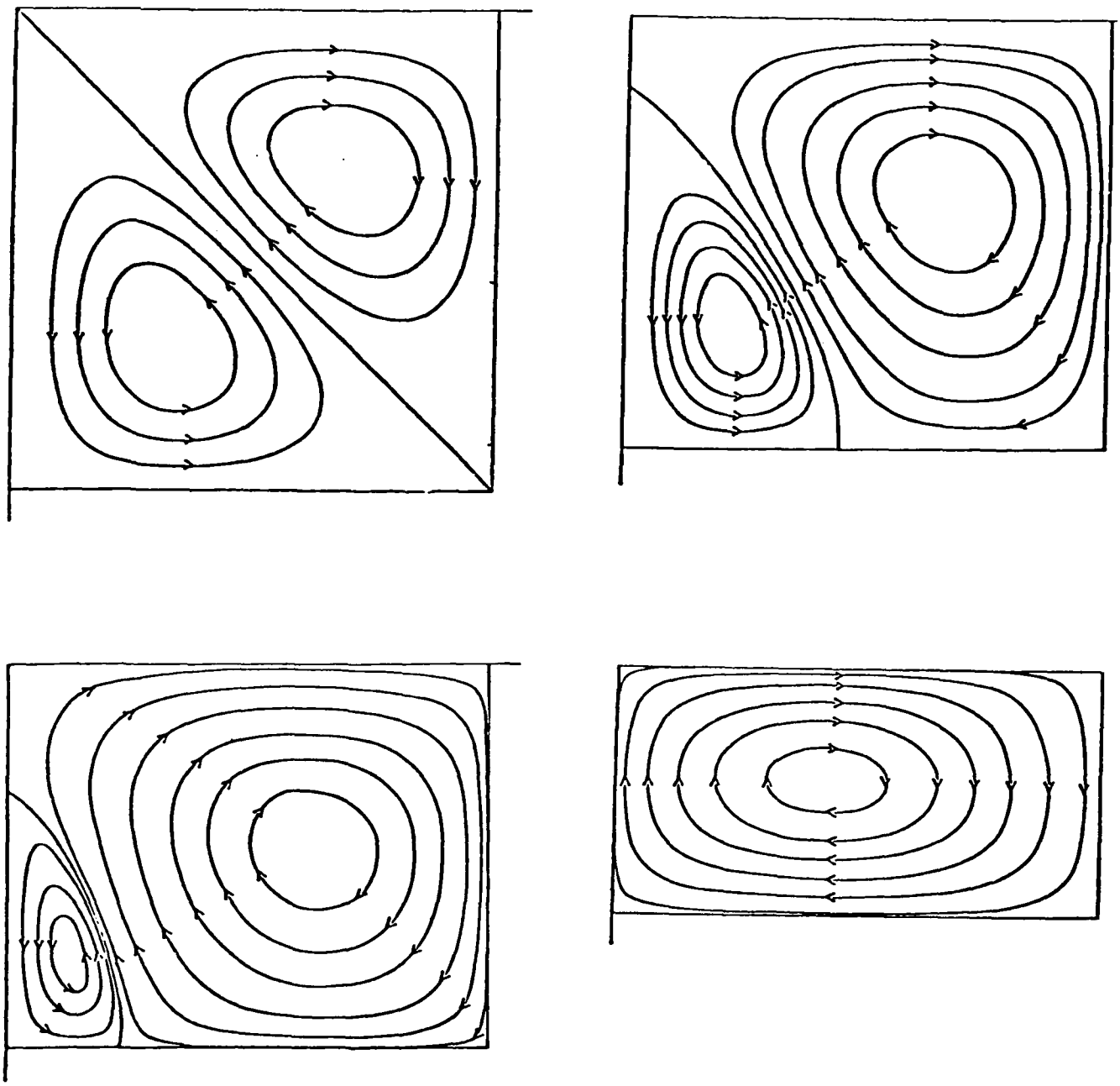


Fig.4 Stream-lines for time-averaged transverse flow in rectangular pipe for aspect ratios 1, 0.9, 0.8, 0.5 and $a(\omega\rho/\eta)^{\frac{1}{2}} = 1$ (top, left quadrant). [a = larger dimension of quadrant, η = viscosity, ω = angular frequency of vibration, ρ = density of fluid.]

Reprinted from J. Non-Newtonian Fluid Mechanics.

CENTRIFUGAL INSTABILITIES IN FINITE CONTAINERS

Philip Hall, Mathematical Sciences Department,
Rensselaer Polytechnic Institute, Troy, New York 12181

Abstract

The effect of endwalls in a Taylor vortex apparatus is investigated using nonlinear stability theory. It is shown that one important consequence of the finiteness of any such apparatus is that the initial vortex motion develops smoothly with increasing Taylor number and not as a bifurcation from circumferential flow. Moreover, in short cylinders, the dominant nonlinearty motions is quadratic and not cubic as is well known for the infinite problem.

1. Introduction

In this paper we discuss the response of circumferential flows in finite cylinders to centrifugal instabilities. In very long containers it is observed experimentally that the disturbed flow is periodic along the axes of the cylinders and this flow is usually called a Taylor-vortex flow. Until recently the large number of theoretical papers on this problem have assumed that the cylinders are infinitely long. If this assumption is made, and the inner cylinder rotates, then linear stability theory predicts that this flow becomes unstable when the angular velocity of the inner cylinder reaches a certain critical value. If the angular velocity is increased slightly above this value, then nonlinear stability theory shows that a stable supercritical equilibrium flow exists. However, if the angular velocity is further increased, the axisymmetric Taylor vortex flow itself becomes unstable to wavy vortex modes which are periodic along the axes of the cylinders and around the cylinders.

Suppose then that we now restrict the cylinders to be of finite length and that the ends of the cylinders are at rest. The basic flow set up when the inner cylinder rotates is now three-dimensional and can only be calculated numerically unless the speed of rotation of the inner cylinder is small. Thus even before the onset of any instabilities, a circulatory flow exists in the cylinders. Related studies in Bénard convection theory by Daniels (1977) and Hall and

Walton (1977) suggest that this flow develops smoothly into a Taylor flow, without any bifurcations taking place. In addition to the three-dimensional nature of the basic flow, any disturbances to the flow will be influenced by the end walls which constrain the possible fluid motions between the cylinders. We shall concentrate on the latter effect here and describe the bifurcation problem in cylinders having end walls rotating such that the basic flow is purely circumferential. For such flows we show that if the length of the cylinders is the same order as their separation, then the amplitude equations determining finite amplitude disturbances have quadratic nonlinearities.

The linear stability theory for the flow which we consider here has recently been discussed by Blennerhassett and Hall (1979), hereafter referred to as I. In that paper it was found that the class of unstable disturbances to the flow could be divided into those having axial velocity even about the mid-plane of the cylinders and those odd about that plane. Depending on L, the nondimensional length of the cylinders, the most dangerous disturbance can be either odd or even. In fact at a given value of L, there exists an infinite sequence of values for Ω, the angular velocity of the inner cylinder, at which disturbances become unstable. For large values of L the results given in I show that the critical value of Ω differs from its value for the infinite problem by an amount of order L^{-2}.

2. Formulation of the problem

We consider the flow of a viscous incompressible fluid of kinematic viscosity ν between concentric cylinders of radii R_0 and $R_0 + d$ and of common length 2Ld. The inner cylinder rotates with angular velocity Ω whilst the outer cylinder is at rest. We restrict our attention to the small gap limit $\frac{d}{R_0} \to 0$ and define dimensionless coordinates x and ϕ by

$$x = \frac{r-R_0}{d} \quad , \qquad \phi = zd^{-1} \quad , \qquad (2.1a,b)$$

where r is the distance of any point from the common axis of the cylinders whilst z measures distance along that axis. The boundaries of the cylinders are defined by $x = 0,1$ and $\phi = \pm L$.

If $\frac{d}{R_0} < 1$, and the cylinders are infinitely long, the fluid moves with azimuthal velocity $\Omega R_0(0,1-x,0)$. We assume that the ends of the cylinders in the problem to be considered here rotate with this velocity so that the basic flow between the cylinders is purely circumferential.

We now suppose that this flow is perturbed in an axisymmetric manner such that the new velocity field is $(\frac{-\nu}{2d}u,\ \Omega R_0(1-x) + \frac{\Omega R_0 v}{2},\ -\frac{\nu}{2d}w)$. We note that the azimuthal component of the disturbance velocity is scaled on ΩR_0 whilst the radial and axial components are scaled on $\frac{\nu}{2d}$, the factor 2 and the minus sign are introduced for convenience. We further note that the perturbation pressure p associated with the above disturbance is independent of the polar angle θ. The Taylor number T associated with the basic flow is defined by

$$T = \frac{2\Omega^2 R_0 d^3}{\nu^2} \cdot \tag{2.2}$$

We define an operator ℓ and a time variable τ^* by

$$\ell \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial \phi^2}, \tag{2.3}$$

and

$$\tau^* = \frac{\nu}{d^2} t . \tag{2.4}$$

The equations which determine u, v, w are obtained by eliminating the pressure from the first and third momentum equations and by using the continuity equation. We obtain

$$(\ell - \frac{\partial}{\partial \tau^*})\, u - T(1-x)v_{\phi\phi} = -\frac{1}{2}Q_{1\ \phi\phi} + \frac{1}{2}Q_{2x\phi},$$

$$(\ell - \frac{\partial}{\partial \tau^*})v - u = -\frac{1}{2}Q_3, \qquad u_x + w_\phi = 0, \tag{2.5a,b,c}$$

where

$$Q_1 = uu_x + wu_\phi - \tfrac{1}{2}Tv^2 ,$$

$$Q_2 = uw_x + ww_4 , \qquad (2.6a,b,c)$$

$$Q_3 = uv_x + wv_\phi .$$

We must solve the above partial differential system subject to the no-slip conditions at the ends and curved surfaces of the cylinders:

$$u = v = w = 0, \quad x = 0,1$$
$$u = v = w = 0, \quad \phi = \pm L . \qquad (2.7a,b)$$

3. Finite amplitude solutions for $L \sim 0(1)$

In I it was shown that depending on the length of the cylinders the most dangerous disturbance can have axial velocity component even or odd about the plane $\phi = 0$. The main results of I are shown here in Figure 1 reproduced from I. We see that in addition to the first odd and even eigencurves, there are further higher modes. In fact, there is an infinite sequence of pairs of odd and even eigencurves which wrap around each other when L varies. We denote the first even and odd critical Taylor numbers by $T_E(L)$ and $T_0(L)$ and near any point of intersection (L^*,T^*) of these curves we can write

$$T_E(L) = T^* + (L-L^*)T_E^* + \ldots ,$$
$$T_0(L) = T^* + (L-L^*)T_0^* + \ldots . \qquad (3.1a,b)$$

We note that all the intersection points of the curves shown in Figure 1 are such that T_E^*, T_0^* are both negative. Thus if for example $T_E^* > T_0^*$ the even mode is the most dangerous for $L < L^*$ whilst the odd mode becomes the most dangerous for $L > L^*$. We now assume that the point (L,T) is close to the intersection point (L^*,T^*). More precisely, we write

$$T = T^* + \varepsilon T_1 + \varepsilon^2 T_2 + \ldots ,$$
$$L = L^* + \varepsilon L_1 , \qquad (3.2a,b)$$

where ε is small and positive whilst L_1, T_1, T_2, etc., are prescribed constants of order ε^0. A straightforward analysis of the linearized form of (2.5) shows that the growth rates of the first odd and even modes are of order ε when $T - T^*$ are of order ε. Thus we define a slow time variable τ by

$$\tau = \varepsilon\tau^* \ . \tag{3.3}$$

We now expand the disturbance velocity $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$ in the form

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \varepsilon \begin{pmatrix} u_e \\ v_e \\ w_e \end{pmatrix} + \varepsilon \begin{pmatrix} u_o \\ v_o \\ w_o \end{pmatrix} + \varepsilon^2 \begin{pmatrix} u_{2o} \\ v_{2o} \\ w_{2o} \end{pmatrix} + \varepsilon^2 \begin{pmatrix} u_{2e} \\ v_{2e} \\ w_{2e} \end{pmatrix} + \ldots \ , \tag{3.4}$$

where subscripts o and e denote velocity fields having axial velocities odd and even about the plane $z = 0$ respectively.

The functions $\underline{u}_e$, $\underline{u}_o$ are just the linear eigenfunctions given in I. In fact, these modes have amplitudes $A(\tau)$ and $B(\tau)$ where these amplitude functions are determined at order ε^2. We find that the "amplitude equations" determining A and B are

$$\begin{aligned} e_1 \frac{dA}{d\tau} &= \{e_2 T_1 + e_3 L_1\}A + e_4 AB \\ f_1 \frac{dB}{d\tau} &= \{f_2 T_1 + f_3 L_1\}B + f_4 A^2 + f_5 B^2 \end{aligned} \tag{3.5a,b}$$

where e_1, f_1, etc., must be calculated numerically at each point of intersection of the odd and even eigencurves. Some values of these coefficients are given by Hall (1979).

4. The solutions of the amplitude equations for $L \sim O(1)$ and discussion.

In this section we shall discuss the equilibrium solutions of (3.5a,b) for the particular values of the coefficients e_1, f_1 obtained by Hall (1979). The three equilibrium solutions (A_E, B_E) of these equations are obtained by setting $\frac{dA}{d\tau} = \frac{dB}{d\tau} = 0$ and solving the resulting equations to give

$$\text{I.} \quad A_E = B_E = 0 \ ,$$

$$\text{II.} \quad A_E = 0 \ , \quad B_E = -\{f_2T_1 + f_3L_1\}f_5^{-1} \ ,$$

(4.1a,b,c)

$$\text{III.} \quad B_E = -\{e_2T_1 + e_3L_1\}e_4^{-1} \ ,$$

$$A_E = \pm\left\{\frac{(e_2T_1 + e_3L_1)}{f_4}\left[\frac{f_2T_1 + f_3L_1}{e_4} - \frac{f_5[e_2T_1 + e_3L_1]}{e_4^2}\right]\right\}^{1/2} .$$

The first solution is the trivial one which corresponds to the unperturbed state. The second solution represents a finite amplitude odd disturbance which exists for all values of L_1 and T_1. (We note that odd and even modes correspond to disturbance velocity fields which are respectively symmetric and antisymmetric about the mid plane $\phi = 0$.) Finally, the third type of solution corresponds to a disturbance which is neither odd nor even. Since A_E must be real, this disturbance does not exist for all values of L_1 and T_1. However, it is easy to show that, for the values calculated by Hall (1979), this solution exists for T_1 in the range between $T_1 = -e_3L_1e_2^{-1}$, and $T_1 = T_{1B} = L_1[e_3f_5 - e_4f_3][e_4f_2 - e_2f_5]^{-1}$. The latter value corresponds to the T_1 coordinate of the point of intersection of the lines $B_E = -\{f_2T_1 + f_3L_1\}f_5^{-1}$ and $B_E = -\{e_2T_1 + e_3L_1\}e_4^{-1}$.

The stability of the equilibrium solutions can be investigated in the usual way. We perturb the equilibrium flow by writing

$$A = A_E + ae^{\sigma\tau} \ , \qquad B = B_E + be^{\sigma\tau}$$

where a,b are small and independent of τ. Substituting for A and B from above into (3.20) and retaining only linear terms in a,b, we obtain

$$e_1\sigma a = a\{e_2T_1 + e_3L_1 + e_4B_E\} + b\{e_4A_E\}$$

$$f_1\sigma b = a\{2f_4A_E\} + b\{f_2T_1 + f_3L_1 + 2f_5B_E\} \ .$$

The growth rates σ_1, σ_2 are determined from the roots of the quadratic equation which ensures the existence of a nontrivial solution to the above homogeneous equations. If the real parts of σ_1 and σ_2 are both negative, the equilibrium flow is stable. If either growth rate has positive real part, the equilibrium flow is unstable. The three equilibrium solutions (4.1) corresponding to the different values of (L^*,T^*) given by Hall (1979) were investigated in this way. The stability properties of the equilibrium solution depend on the coefficients e_1, f_1, etc., at any value of (L^*,T^*). However, surprisingly enough, at any value of (L^*,T^*) there are only two distinct cases to consider.

Case a. The most dangerous mode an odd disturbance

Suppose that L_1 is such that the most dangerous mode is an odd mode. In this case the trivial solution is stable until the bifurcation point $T_1 = - f_3 L_1 f_2^{-1}$ and for T_1 greater than this value it is unstable. The equilibrium solution II is stable for $T_1 > - f_3 L_1 f_2^{-1}$ and unstable for $T_1 < - f_3 L_1 f_2^{-1}$. Thus the zero solution and the type II equilibrium flow exchange stability characteristics at the bifurcation point $T_1 = - f_3 L_1 f_2^{-1}$. The third type of solution, III, is unstable wherever it exists. These results are summarized graphically in Figure 2a where continuous lines represent stable solutions and dotted lines represent unstable solutions. We conclude that when L_1 is held fixed and T_1 increased, then, if the odd mode is the most dangerous, the unperturbed flow is stable for $T_1 < - f_3 L_1 f_2^{-1}$ and for T_1 greater than this value an odd disturbance with magnitude proportional to T_1 exists. This latter flow corresponds to an even number of cells in the range $-L \leq \phi \leq L$.

Case b. The most dangerous mode an even disturbance

The stability characteristics of the equilibrium solutions in this case are summarized in Figure 2b. We note that for $T_1 < -e_3 L_1 e_2^{-1}$ the only stable configuration is the unperturbed state. If T_1 lies in the range $(-e_3 L_1 e_2^{-1}, T_{1B})$, the only stable

equilibrium flow is a type III solution corresponding to a disturbance neither odd nor even in ϕ . However, this solution exists only for T_1 in this range and for $T_1 > T_{1B}$ the only stable solution is an odd disturbance corresponding to II in (4.1).

Thus, we conclude that, in general, the most likely disturbance to be observed is an odd mode; the mixed mode exists for a finite range of values for T_1 and is only stable when the most dangerous mode of linear theory has axial velocity even in ϕ .

Finally, we briefly describe the effect of applying more realistic conditions at the ends of the cylinders. It has been shown by Hall (1979) that if we perturb the end conditions towards the no-slip conditions appropriate to the case of fixed rigid endwalls, then (3.5a,b) are modified to give

$$e_1\frac{dA}{d\tau} = \{ e_2T_1 + e_3L_1 \}A + e_4AB \quad ,$$

$$f_1\frac{dB}{d\tau} = \{ f_2T_1 + f_3L_1 \}B + f_4A^2 + f_5B^2 + f_6 \ .$$

The introduction of the constant term in the second equation is a direct consequence of the more realistic end conditions. The effect of this term on the equilibrium solutions of the amplitude equations is dramatic. We now find that the solution A = B = 0 of (3.5) is no longer a solution of the more realistic problem. An investigation of the above equations shows that there exists a smoothly developing odd mode (A = 0) for all values of T_1. In some cases, this smoothly developing solution becomes unstable to mixed mode solutions with both A and B nonzero.

This work was partially supported by the Army Research Office.

References

Blennerhassett, P. and Hall, P., 1979 PRS(A), 365, 191.
Daniels, P., 1977 PRS(A), 358, 173.
Hall, P. and Walton, I. C., 1977 PRS(A), 353, 199.
Hall, P., 1979; submitted to PRS(A).

FIGURE 1

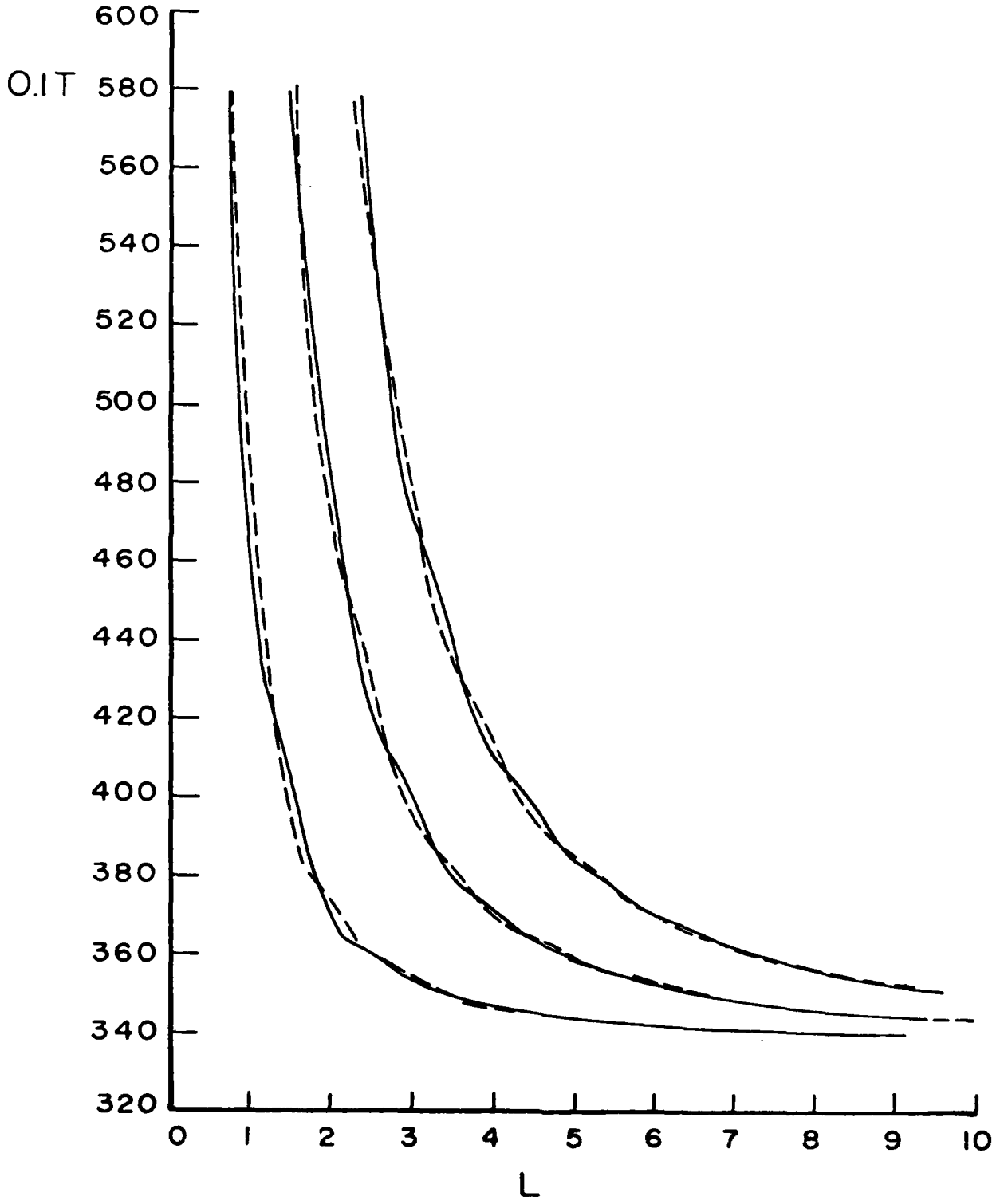

O.I T
600
580
560
540
520
500
480
460
440
420
400
380
360
340
320
0
1
2
3
4
5
6
7
8
9
10
L

FIGURE 2a

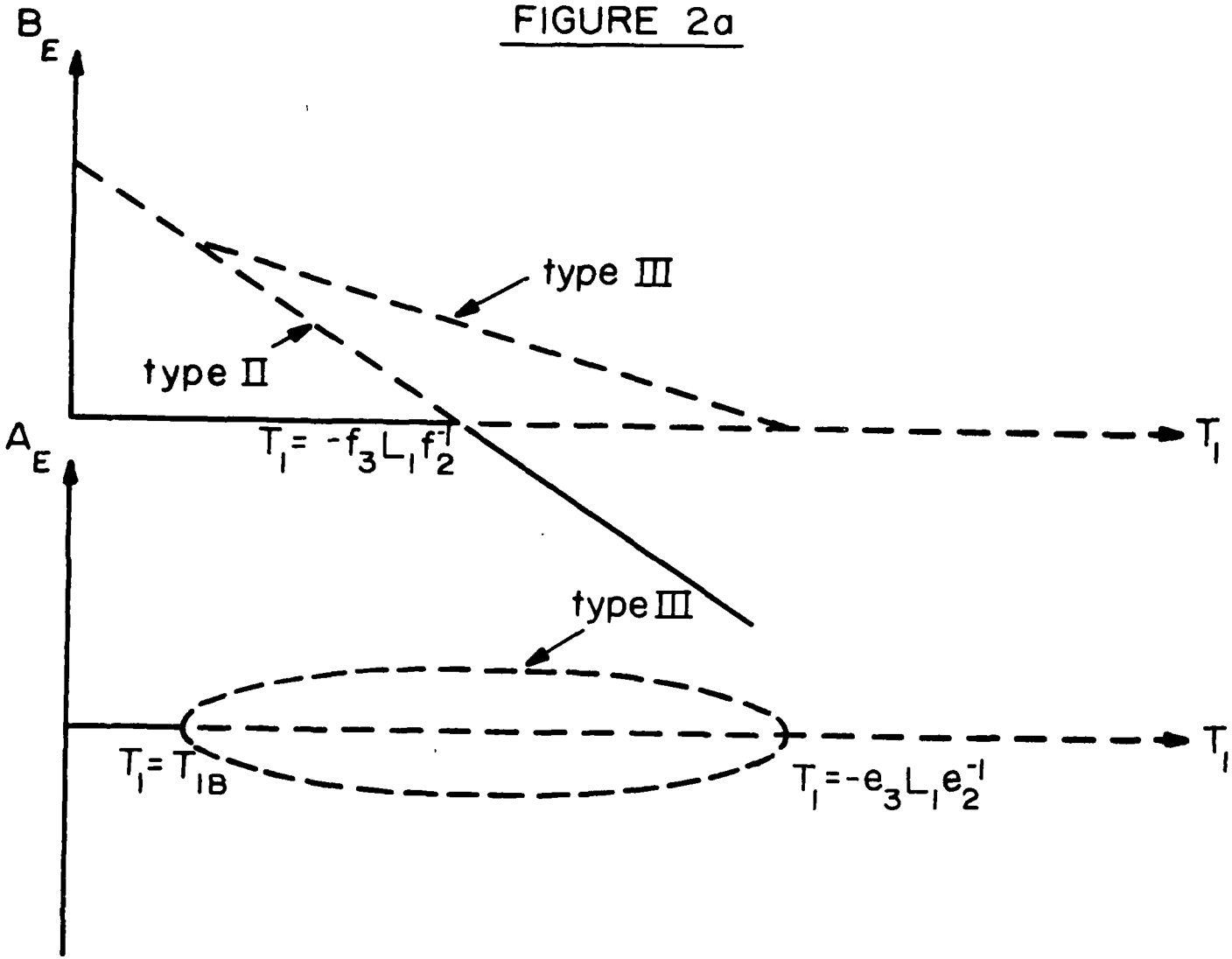
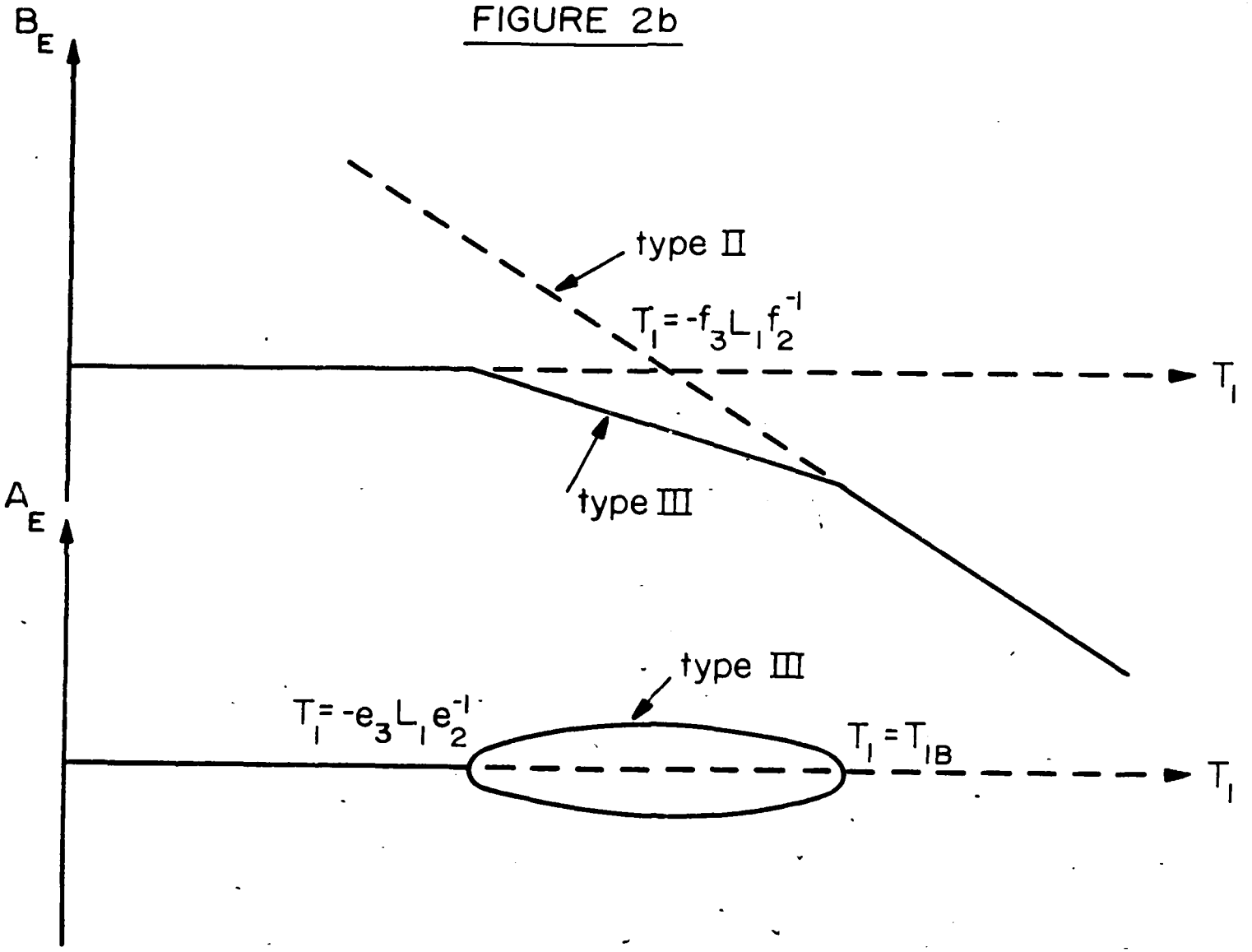

B_E
type III
type II
T_1= -f_3 L_1 f_2^-1
T_1
A_E
type III
T_1= T_1B
T_1=-e_3 L_1 e_2^-1
T_1

FIGURE 2b


B_E
type II
$T_1 = -f_3 L_1 f_2^{-1}$
T_1
type III
A_E
type III
$T_1 = -e_3 L_1 e_2^{-1}$
$T_1 = T_{1B}$
T_1

THE HISTORY OF THE UTILIZATION OF EULERIAN HYDRODYNAMIC COMPUTER CODES AT THE BALLISTIC RESEARCH LABORATORY

John T. Harrison
Shaped Charge Branch
Terminal Ballistics Division
Ballistic Research Laboratory
US Army Research and Development Command
Aberdeen Proving Ground, MD 21005

ABSTRACT. The Ballistic Research Laboratory has been involved with the development and utilization of hydrodynamic computer codes for over twenty years. The final product of one of these developmental efforts is the two-dimensional, Eulerian, finite-difference, hydrodynamic code called 'HELP'. This paper will trace its history and point out some of its applications.

The HELP code has evolved from four major hydrodynamic programs which were developed over a twenty-year period. Its genesis is the original Particle-In-Cell (PIC) code written by M.W. Evans and F. Harlow at the Los Alamos Scientific Laboratory. The next step was the OIL code which replaced the discrete particle mass transport scheme by a continuum. The RPM code, which repesented the material as having rigid perfectly-plastic properties, was third in the series. Finally, the HELP code was developed with a multi-material capability and representing the material properties as elastic-plastic.

These codes have been used in a wide range of problems in continuum mechanics. Their evolution has had a immense infuence on research and development in both government and industry.

I. INTRODUCTION. The Ballistic Research Laboratory (BRL) has a rich history of advancements in computer technology. This has enabled the accomplishment of some remarkable studies in a wide scope of fields, and continuum mechanics is one of these. It is now possible to examine in detail such problems as:

1. The penetration produced by a hypervelocity impact.
2. The cratering produced by a 'chunky' fragment upon impact.
3. The detonation of a stick of explosive and corresponding deformation of the adjacent materials, and many more of similar complexity.

Two things were required to accomplish these investigations. One was the development of the the computers and computer languages themselves. The other was the development of the mathematical techniques that transform the basic differential equations into forms suitable for numerical analysis. This paper will concentrate on the historical development and utilization of the latter, but it is useful to mention briefly the computers which were available to the numerical analyst at the BRL from a historical standpoint [1].

First, the world's premier high-speed,electronic digital computer, the ENIAC, was operational at the BRL from 1947 to the early 50's. This was a decimal computer and programmable only in machine language.

Second, EDVAC was the first binary and stored program computer. It was operational from 1951 to the early 60's. This too was programmable

only in machine language.

Third, the ORDVAC was operational from the early 50's to the mid-1960's. This computer was programmable in both a pusedo-language termed assembly language and a machine language.

Fourth, BRLESC I was operational from the early 60's to the mid 70's. Computer systems analysts at the BRL developed a high level language called Formula and Assembly Translator, FORAST. This high level language was designed for the scientific programmer and was operational until 1968 when FORTRAN IV was adopted.

Fifth, BRLESC II was operational from the mid 60's to the mid 70's when both it and BRLESC I were replaced by a Control Data Corporation, Cyber 170 computing system.

Beyond this, we will not go into computers in any great detail. Instead, we examine the initial comprehensive research project which was instituted to obtain solutions to U.S. Army related problems. This initial project examined, in detail, the numerical scheme used to obtain these solutions. This paper will cover the following:

1. The personnel responsible for and involved with the project.
2. The problems in continuum mechanics for which solutions were needed.
3. The governing equations and underlying approximations considered to obtain suitable solutions to the problems and the numerical schemes used.
4. A brief statement about the results of this inital study and observations concerning future code development work.

As a result of the initial investigation, a series of improvements to this numerical scheme where supported by the BRL over a seventeen year period. The chronological order of development for those Eulerian umerical codes used at the BRL will be presented. Finally, a 'Family Tree' of Eulerian hydrodynamic codes will place each in their respective branch.

Many of these codes are still being used today. It is the intent of this paper to not only familiarize those using these codes with their origin, but also to acquaint them to the part that the BRL played in their evolution.

II. BACKGROUND. The initial theoretical analysis, code development, and experimental work was a combined comprehensive research program in the field of hypervelocity impact. The work was initiated in May 1962 under an Advanced Research Projects Agency (ARPA) Order No. 71-62. This resulted from a series of recommendations by Dr. R.E. Duff, who, as a member of the Institute for Defense Analysis, prepared a planning document to assist personnel at ARPA.

The experimental work and overall responsibility for the project was delegated to the Army's Ballistic Research Laboratories. The theoretical analysis and code development work was completed by personnel at the General Atomic Division (GAMD), General Dynamics under Contract No. DA-04-495-AMC-116(X). In addition to the the ARPA contract, Army funds were used to support a modest theoretical effort at Drexel Institute of Technology under Contract No. DA-36-034-ORD-3672-RD.

This research program was a cooperative venture between eight scientists and the overall coordinator for the project, Dr. F.E. Allison of the

BRL. Others at the BRL included J.H. Kineke, J.T. Frazier, and V.M. Boyle. Work at the General Atomics Division of General Dynamics was under the technical supervision of J.M. Walsh. Others there included J.E. Johnson, J.K. Dienes, and J.H. Tillotson. Work at the Drexel Institute of Technology was under the technical supervision of Pei Chi Chou.

Central to the program was the development and use of the hydrodynamic computer code called 'OIL'[2,3]. OIL is a two-space dimensions and time dependent code with only one-material and based upon the Eulerian or fixed grid numerical formulation. This code is closely related to the Particle-In-Cell (PIC) code [4,5]. It was developed by modifying the General Atomic PIC code named SHELL [6] by the introduction of a continuous mass transport scheme.

The OIL code was used to obtain numerical solutions for the initial interaction between a projectile and target in an hypervelocity situation. This simulated the penetration of a continuous shaped charge jet. A second problem was also studied during this initial investigation. This was the study of craters produced by 'chunky' fragments. It analyzed the cratering effects produced by a shaped charge jet after it has broken into many discrete particles or by a fragmenting munition. Experimental evidence showed that the initial interaction of the projectile and target, in both cases, is a hydrodynamic process, which can be treated as a problem in compressible fluid dynamics. This study of the hydrodynamic phase of the interaction lead to significient conclusions concerning the terminal effects of the impact.

III. GOVERNING EQUATIONS. The penetration of a target by a projectile can be described in terms of classical hydrodynamics. The cratering problem also has a hydodynamic phase during the initial interaction of the target and projectile which can be treated as a problem in compressible fluid dynamics. The leading edge of the interaction is defined by a shock front across which the flow satisfies the Rankine-Hugoniot equations [7]:

$$\rho_0 U = \rho_1 (U - u_1) , \tag{1}$$

$$P_1 - P_0 = \rho_0 U u_1 , \tag{2}$$

$$E_1 - E_0 = \tfrac{1}{2}(P_1 + P_0)(V_0 - V_1) . \tag{3}$$

Elsewhere, the equations for continuous compressible flow [7] are assumed to apply:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 , \tag{4}$$

$$\frac{\partial \vec{u}}{\partial t} + V \nabla P = 0 , \tag{5}$$

$$\frac{dE}{dt} = -P \frac{dV}{dt} . \tag{6}$$

Here the derivates d/dt denote the usual convective derivatives; i.e., $df/dt = \partial f/\partial t + \vec{u} \cdot (\nabla f)$. This set of equations is completed by specifying an equation of state which is taken to be of the general form:

$$P = f(\rho, E). \tag{7}$$

To provide a realistic thermodynamic description of the material, the equation of state is constructed by fitting available experimntal data. The preceeding equations are solved numerically by the OIL code.

An extremely useful property of the preceding equations is that the solutions can be scaled; i.e., An impact, for which the characteristic length of the projectile is L_0, has the same hydrodynamic solution as a geometrically similar impact (same velocity) with the characteristic length FL_0, except that all times and distances are changed by the factor F. Intensive variables such as pressure, density, and velocity remain unchanged under the transformation.

IV. UNDERLYING APPROXIMATIONS. The principal approximations needed to derive Equations 4-6 are: (1) diffusion effects (such as those due to viscosity, thermal conduction, and radiation) are negligible within the continuous flow; and (2) the tensor stresses due to material strength can be negelected. These approximations and the previously stated simple scaling law are closely related.

The former approximations, i.e. the diffusion effects, would introduce higher-order derivatives in the continuous-flow equations. This would cause a departure from the simple scaling laws. Therefore, since the main body of experimental evidence at that time indicated that simple scaling was a valid approximation, neglecting the effects of diffusion was a good approximation.

The latter approximation is valid by definition for the so-called hydrodynamic phase of the impact process, which is taken to be that early part of the interaction during which the equation-of-state pressures are large compared to the material yield strengths. Although strength effects are negligible in the hydrodynamic phase, it is necessary to note that the ordinary stress-strain effects do scale; whereas the time-dependence within the stress-strain relations causes a breakdown in simple scaling.

V. CONCLUSIONS AND RECOMMENDATIONS. Many aspects of experimental and theoretical analysis were taken into consideration during this study. These included the use of one and two-dimensional impact models with an ideal gas equation-of-state, similarity solutions, and experimental-theoretical correlations. As a result the crucial link between the hydrodynamic phase and the terminal effects was first observed in impact calculations from a hydodynamic code based upon the PIC formulation of compressible fluid flow. This result was improved upon by the use of the OIL code. This was the the first significant result attributed to hydro-code calculations and was verified in 1963 by experimental observations [3].

The crucial link was termed 'Late-stage equivalence.' It, simply stated, meant that two impact calculations become indistinguishable when the characteristic dimensions of the projectiles are related to their velocities through the relation $(L_0'/L_0)=(U_0/U_0')^{\alpha}$, then the two flows will become equivalent in their later stages. Figure 1 is a simple one-dimensional curve showing the concept of late-stage equivalence.

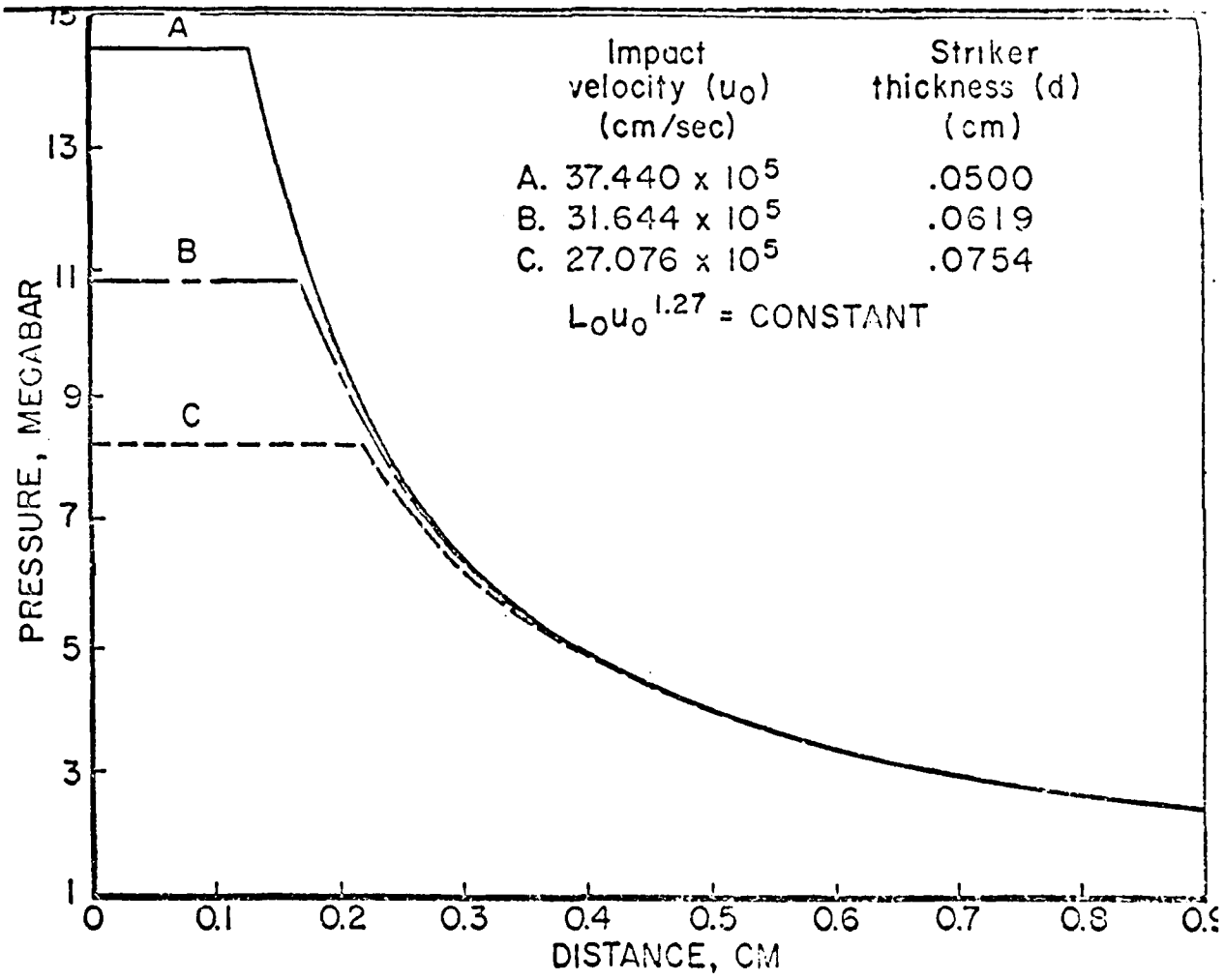


FIG. 1. PEAK SHOCK PRESSURE AS A FUNCTION OF DISTANCE FROM THE POINT OF IMPACT FOR THREE IMPACT PROBLEMS IN WHICH THE PRODUCT $L_0 u_0^{1.27}$ WAS HELD CONSTANT. FOR DISTANCES GREATER THAN 0.2 CM, THE PEAK SHOCK PRESSURE IS INDEPENDENT OF INITIAL CONDITIONS. THIS FIGURE IS FROM REFERENCE 8.

Despite the fact that significant conclusions were obtained by examining the hydrodynamic phase of the impact interaction, it was highly desirable to obtain a detailed description of the actual mechanics for the strength-dependent portion of the interaction. The most promising approach to the problem consisted of generalizing the hydrodynamic codes so that problems in strength-dependent deformation could be computed. An additional area which required attention was that of an impact which was not axi-symmetric. A theoretical understanding of the so-called obliquity effects must be based upon calculations of flows with three space variables and time.

As a result of these recommendations and observations, the BRL sponsored the scientists who are responsible for the development of two of the hyrodynamic computer codes used today. These are the HELP [9] and TRIDORF [10] codes. The HELP code is a two-dimensional, multi-material, Eulerian computer model with an elastic-plastic consituative relationship. The code was developed by L.J. Hageman under the technical supervision of J.M. Walsh at the System, Science, and Software (SSS). The TRIDORF code is a three-dimensional, rigid perfectly-plastic material strength, Eulerian code having the computational capability of handling two materials in a zone. The latter code was a product of the so-called evolutionary computer code development tree. The developer was W.E. Johnson of the Computer Code Consultants.

VI. CHRONOLOGICAL ORDER OF DEVLOPMENT. Table 1 lists the chronological order of development of two and three dimensional, Eulerian, hydrodynamic computer codes. The table presents the date of first publication of the users manuals, the authors name, and a short discription of the code. The genesis of these codes is the particle-in-cell code developed by M.U. Evans and F. Harlow in 1957 for all of the codes listed except the SMITE code [11]. The SMITE code is a high-order accurate differencing scheme which is based upon the Eulerian method. It was developed by S.Z. Burstein and H.S. Schecter from Mathematical Applications Group, Incorporated (MAGI). This code was developed under a BRL contract in 1972.

These codes have been used by researchers at the BRL from their very beginning. A good reason for this is the fact that their development has been very well documented. The bibliography of this paper will contain the references for these codes based upon the first reported date.

VII. THE EULERIAN HYDRO CODE TREE. The Eulerian hydro code tree shows the evolution of the Eulerian, hydrodynamic computer codes. The tree is a direct outgrowth of research sponsored for the most part by the BRL. The development of these numerical schemes has had a great influence on both government and industry.

Table 2 is a representation of the Eulerian hydro code tree. At its trunk rests the PIC code, developed at the Los Alamos Scientific Laboratory. Next, the SHELL code which has the same formulation as PIC, i.e. with discrete mass particles dispersed in fixed, Eulerian cells. It was developed at the General Atomic Division of General Dynamics by W.E. Johnson. The next in the chain came a code that had a large impact on the code development industry, the OIL code. OIL was the catalyst for the outgrowth of many other codes because it replaced the discrete particle mass transport scheme by a continuum. The predominate reason for its impact is that it reduced the amount of computer internal storage requirements to do even a simple calculation.

Thereafter, many computers were then able the support the use of the code.

From this point, the hydro code tree branched in five different directions. First, a material strength branch, the RPM code [12], was added to the OIL code. Second, a multimaterial branch, the TOIL code [13], was added to the OIL code. Third, a branch of the OIL code was added with three space dimensions and time as the independent variable. This code was appropriately called the TRIOIL code [14]. The fourth branch of the OIL code is a new calculational technique called 'the splitting technique'. The new code, SOIL [15] was developed by W.E. Johnson.

A fifth branch is a limb to itself. This limb contains the HULL codes [16]. The HULL code was initially developed at the Air Force Weapons Laboratory (AFWL) by a modifation of the SHELL code. The HULL code has been used by personnel at the BRL since its inception.

The multi-purpose code in much use today is the HELP code [17]. It's a two-dimensional, multimaterial, hydrodynamic code with elastic-plastic formulation for material strength. The code has evolved from four major hydrodynamic programs developed over a long period of time. In Table 2, we trace it from its genesis, the PIC code, to the OIL code, to the RPM code, and then to the HELP code. The HELP code has undergone several programming changes over the years. These changes varied from minor programming changes to major code modifications such as the BRLSC code [18]. The most recent was a change in the numerical differencing proceedure for the calculation of specific internal energies. This version has been modified by personnel at the BRL and thus receives the acronym, BRLHELP code [19].

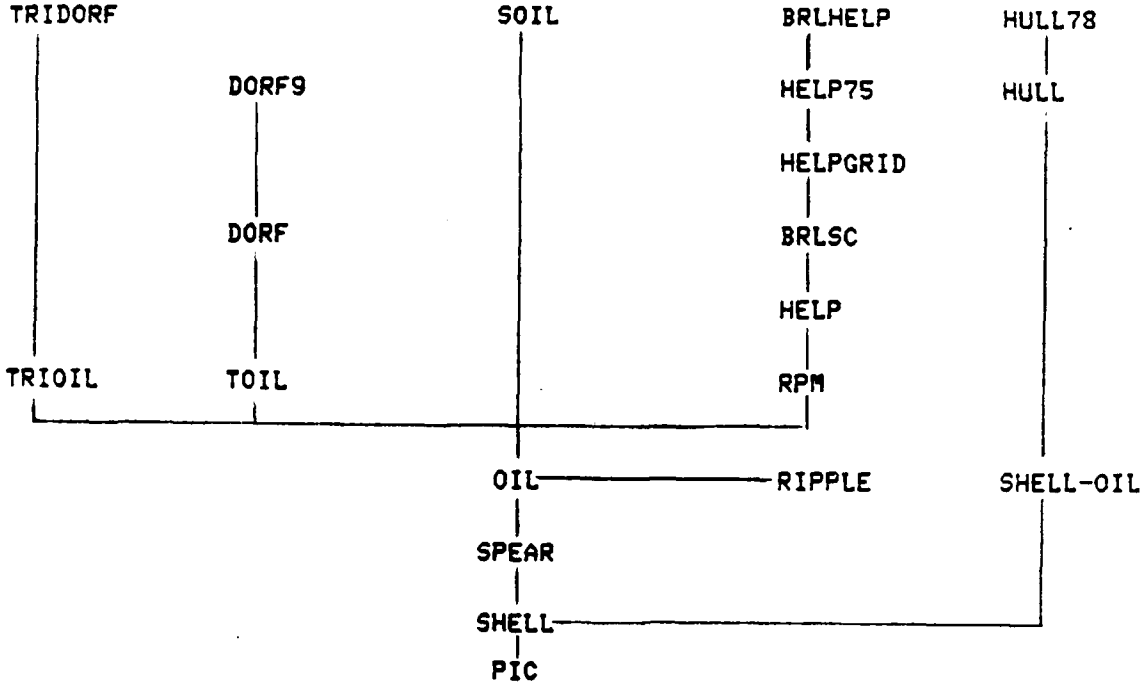
The hydrodynamic codes listed in this paper are a small sample of those in existance. They are those mostly used by scientists at the BRL. Nevertheless they have had a tremendous impact in our national defense. We are on the threshold of a new era in both the use of the hydrodynamic codes and the computing machines that run them. Solutions to a wide variety of theoretical research problems are now forthcoming.

Table 1. The Chronological Order of Development of Two and Three Dimensional Eulerian, Hydrodynamic Computer Codes.

Date	Name	Authors	Place	Description
1957[4]*	PIC	M.W.Evans,F.Harlow	Los Alamos	Particle-in-cell,2-D, Hydro.
1959[6]	SHELL	W.E.Johnson	GAMD	Version of the PIC code
1961[6]	SPEAR	W.E.Johnson	GAMD	Improved SHELL code
1963[2]	OIL	W.E.Johnson	GAMD	Continuous Mass Transport, 2-D,1-Mat.,Hydro
1965[3]	OIL	W.E.Johnson	GAMD	Improved OIL code.
1967[13]	TOIL	W.E.Johnson	GAMD	OIL with 2-materials.
1967[14]	TRIOIL	W.E.Johnson	GAMD	3-D OIL,1-Mat.,Hydro.
1967[20]	RIPPLE	M.W.Evans,L.Hageman &J.A.Williamson	GAMD	The OIL code used for solving detonation problems.
1968[12]	RPM	Johnson,Hageman Evans,Dienes,Walsh	GAMD	OIL Method with rigid perfectly-plastic material strength.
1970[9]	HELP	L.J.Hageman & J.M.Walsh.	SSS	OIL Method,Elastic-plastic strength,Mutimaterial,2-D,Hydro.
1971[21]	DORF	W.E.Johnson	SSS	OIL code,2-Mat.,Rigid perfectly-plastic,2-D.
1971[18]	BRLSC	M.Gettings	SSS	HELP modified to solve the shaped charge problem.
1973[11]	SMITE	S.Z.Burstein, H.S.Schecter & E.L.Turkel.	MAGI	High order accurate differencing scheme, 2-D,2-Mat.,Hydro
1973[22]	HELPGRID	R.T.Segwick, L.J.Walsh,D.Wilkins	SSS	Improved BRLSC code, with grid packaging.
1975[17]	HELP	L.J.Hageman,et.al.	SSS	Improved HELP for the plugging and shaped charge problems.
1976[16]	HULL	M.A.Fry,et.al.	AFWL	Improved SHELL.
1976[10]	TRIDORF	W.E.Johnson	CCC	3-D DORF,Rigid perfectly-plastic,2-Mat.
1977[15]	SOIL	W.E.Johnson	CCC	OIL using the splitting technique.
1978[19]	BRLHELP	J.Lacetera, J.Schmid, & J.Lacetera.	BRL	Improved HELP75 (correcting the internal energy problem).

*[] - The numbers inside of the brackets represent the reference number of the first publication. (see bibliography).

TABLE 2. A hydrodynamic, Eulerian computer code "family" tree.



BIBLIOGRAPHY

1. K. Kempf, "Historical Monograph Electronic Computers within the Ordnance Corps," Aberdeen Proving Ground, Md., November 1961.

2. W.E. Johnson, "Computer Development to Improve SHELL Code," General Atomic Report GA-4673, Contract No. AF29(601)-6028, October 1963.

3. W.E. Johnson, "OIL, A Continuous Two-Dimensional Eulerian Hydrodynamic Code," General Atomic Division, General Dynamics Corp., Report GAMD-5580, Ballistic Research Laboratory Army Contract No. DA-04-495-AMC-116(X), January 1965.

4. M.W. Evans and F.H. Harlow, "The Particle-In-Cell Method for Hydrodynamic Calculations," Los Alamos Scientific Laboratory, Report LA-2139, November 1957.

5. F.H. Harlow, "Two-Dimensional Hydrodynamic Calculations," Los Alamos Scientific Laboratory, Report LA-2301, September 1959.

6. J.M. Walsh, W.E. Johnson, J.K. Dienes, J.H. Tillotson, and D.R. Yates "Summary Report on the Theory of Hypervelocity Impact," General Atomic Division, General Dynamics Corp., Report GA-5119, Ballistic Research Laboratory Army Contract DA-04-495-AMC-116(X), March 1964.

7. R. Courant and K.O. Friedrichs, Supersonic Flow and Shock Waves, Volume I, Interscience Publishers, Inc., New York and London, 1948.

8. F.E. Allison, et.al., Unpublished, Ballistic Research Laboratory Report, Ballistic Research Laboratory, Aberdeen Proving Ground,Md.

9. L.J. Hageman and J.M. Walsh, "HELP, A Multi-Material Eulerian Program for Compressible Fluid and Elastic-Plastic Flows in Two Space Dimensions and Time," Systems, Science and Software Report 3SR-350, Topical Report Report under Contract DAAG 07-68-C-0931, August 1970. (published as Ballistic Research Laboratories Contract Report No. 39, AD 726459, May 1971).

10. W.E. Johnson, "TRIDORF - A Two-Material Version Of the TROIL Code with Strength," Computer Code Consultants, CCC-976, September 1976.

11. S.Z. Burstein, H.S. Schecter, and E.L. Turkel, "The Application of SMITE Code to the BRL-105 Shaped Charge Problem," Mathematical Applications Group, Inc., Ballistic Research Laboratories Contract DAAA05-73-C-0016, 1973.

12. J.K. Dienes, M.W. Evans, L.J. Hageman, W.E. Johnson, and J.M. Walsh, "An Eulerian Method for Calculating Strength Dependent Deformation," General Atomic Report GAMD-8497, Parts I, II, III and Addendum, (AD 678565, 678566, 678567, 678568, February 1968).

13. W.E. Johnson, "TOIL (A Two - Material Version of the OIL Code)," General Atomic Division Report GAMD-8073, Ballistic Research Laboratory Contract No. DA-04-495-AMC-1481(X), July 1967.

14. W.E. Johnson, "TRIOIL, A Three - Dimensional Version of the OIL Code," General Atomic Division Report GAMD-7310, Ballistic Research Laboratory Contract No. DA-04-495-AMC-1481(X), June 1967.

15. W.E. Johnson, Privite Communication, 1977.

16. M.A. Fry, et. al., "The HULL Hydrodynamic Computer Code," AFWL-TR-76-183, Air Force Weapons Laboratory, Kirtland AFB, New Mexico, September 1976.

17. L.J. Hageman, et. al., "HELP- A Multi-Material Eulerian Program for Compressible Fluid and Elastic-Plastic Flows in Two Space Dimensions and Time - Revised Edition," System, Science and Software Report SSS-R-75-2654, July 1975.

18. M.C. Gittings, "BRLSC: An Advanced Eulerian Code For Predicting Shaped Charge Performance," System, Science and Software Report 3SR-642, February 1971. (published as a Ballistic Research Laboratories Contract Report No. 279, AD 023962, December 1975).

19. J.E. Lacetera, J.A. Schmitt, and J.F. Lacetera, "The BRL, CDC 7600 Version of the HELP Code," Report in Publication, Ballistic Research Laboratory, 1979.

20. M.W. Evans, L.J. Hageman and J.A. Williamson, "RIPPLE, A Two-Dimensional Eulerian Computer Program for Calculating Compressible Flow and Detonation Problems," General Atomic Division, General Dynamics, Report No. GAMD-8165, Ballistic Research Laboratories Contract No. DA-04-495-AMC-1481(X), September 1967.

21. W.E. Johnson, "Development and Applications of Computer Programs to Hypervelocity Impact," Systems, Science and Software report 3SR-749, July 1971.

22. R.T. Segwick, L.J. Walsh and D. Wilkins, "Research Study and Analysis for Improvement of the Shaped Charge Code," Ballistic Research Laboratories Contract No. DAAD 05-72-C-0291, Prepared by System, Science and Software, August 1973.

STUDY OF CONVERGENT FLOWS IN CERTAIN SHAPED CHARGE SYSTEMS

By Abdul R. Kiwan
Vulnerability Methodology Team
USA ARRADCOM, BALLISTIC RESEARCH LABORATORY

ABSTRACT

Convergent flows arise in a certain class of shaped charge systems with hemispherical liners which are set in motion by a convergent detonation wave. The study of such flows explains the processes of liner collapse, jet formation, flight and elongation, the resultant jets and their properties in such systems. The mathematical and computational part of this study utilized the HELP code which is a two dimensional Eulerian, multi-material hydrodynamic code. The study showed that the jet forms as a result of the liner compression and differs from the process of jet formation in conical shaped charge systems. Furthermore the resulting jet did not possess an inverse velocity gradient. Jets from such systems possess different distributions of mass, momentum, and energy which make them suitable for certain applications. The study of such flows explained the cause of some of the experimentally observed problems with jets from such systems. The theoretical predictions from this study have been corroborated qualitatively and quantitatively by experimental measurements.

I. INTRODUCTION

Considerable interest was shown recently in the application of convergent flows to the design of a new generation of shaped charge warheads of the future. It was believed that if a metallic hemispherical liner is set in motion by a convergent detonation wave striking it, then the resulting shaped charge jet might possess high velocities and other desirable properties. This consideration has been the primary reason behind the present study. Intuitively it seems that the collapse of the liner might be understood by considering the motion of a typical liner element E relative to an observer at the pole P of the hemisphere as shown in Figure 1. The observer at the pole P which is moving towards the center C with velocity V will see the flow from element E coming at him with velocity V_r. In a stationary coordinate system E is collapsing towards C with velocity V. The liner will continually get thicker during the collapse until eventually it starts jetting to relieve the high compression in the liner material. The dotted line shows the actually computed liner configuration. The numerical simulation of the collapse and jet formation process was achieved with the HELP code which we describe briefly below.

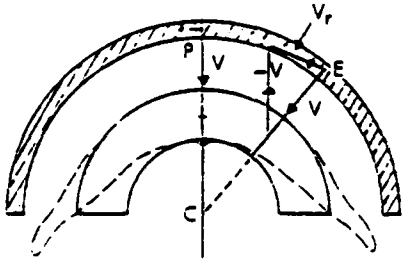


Figure 1: Idealized Schematic of Imploded Hemisphere Collapse

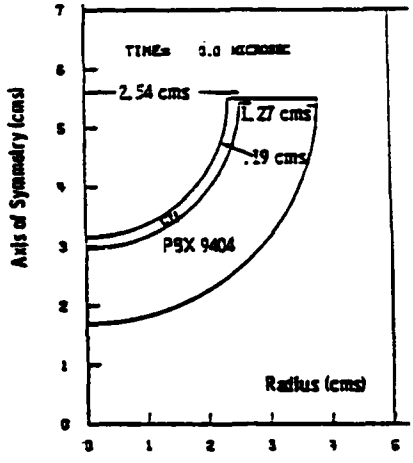


Figure 2: A Schematic of the computational setup

II. DESCRIPTION OF THE COMPUTATIONAL PROCESS

HELP[1] is a two dimensional, Eulerian, multi-material, finite difference hydrodynamic code. The material model in HELP also includes strength effects. The equations that are solved numerically in HELP on the computational grid are the conservation equations of a continuous medium in motion written in the conservation form[2]:

$$\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial x_i} (\rho u_i), \tag{1}$$

$$\frac{\partial}{\partial t} (\rho u_j) = \frac{\partial}{\partial x_i} S_{ij} - \frac{\partial p}{\partial x_j} - \frac{\partial}{\partial x_i} (\rho u_i u_j), \tag{2}$$

$$\frac{\partial}{\partial t} (\rho E_T) = \frac{\partial}{\partial x_i} (S_{ij} u_j) - \frac{\partial}{\partial x_i} (p u_i) - \frac{\partial}{\partial x_i} (\rho u_i E_T). \tag{3}$$

x_i, u_i denote the i^{th} coordinate of position and velocity component, t the time, ρ is the material density, E_T is the total energy, S_{ij} is the deviator stress tensor, and p is the hydrostatic pressure. Equations (1) to (3) together with an equation of state are integrated in three phases. In the first phase (SPHASE) only the terms due to strength effects are considered, while the other terms due to the hydrodynamic pressure forces or transport are temporarily neglected. In the second phase (HPHASE), only the pressure terms are considered while only the transport terms are considered in phase 3 (TPHASE). Thus during phase 2 the considered equations are of the form:

$$\frac{\partial \rho}{\partial t} = 0 \tag{4}$$

$$\frac{\partial}{\partial t}(\rho u_j) = -\frac{\partial p}{\partial x_j} \tag{5}$$

$$\frac{\partial}{\partial t}(\rho E_T) = -\frac{\partial}{\partial x_i}(p u_i). \tag{6}$$

HELP employs the Tillotson[3] equation of state to represent inert material properties. This equation can be represented symbolically by

$$p = f(E_I, \rho), \tag{7}$$

where E_I is the internal energy,

$$E_I = E_T - \frac{1}{2} u_i u_i. \tag{8}$$

Equations (1) to (3) are integrated in a given computational cycle over the volume V of a typical computational cell. The volume integrals arising on the right hand side of the above equations are converted to surface integrals. Executing the three phases of the calculations described above updates cell mass, momenta and energy from cycle n to cycle (n+1). The pressure p at the $(n+1)^{th}$ cycle is then updated from equation (7). The stress deviator S_{ij}, the pressure p, the velocity components u_i, and the energy E_T are calculated at the center of each computational cell. Boundary surface values of those variables that are needed to evaluate the surface integrals are obtained by averaging the adjacent cell centered values of those variables. HELP contains a transmittive and reflective boundary conditions which are optional at some of the grid boundaries. Material interfaces are defined in Lagrangian manner by means of massless tracer particles. Passive tracer particles provide the option to follow the motion of individual material particles in a cell. Tracer particles are moved with the local flow velocity. Sliplines can be introduced along material interfaces. HELP employs the Von Mises yield condition and contains a tensile failure criterion, and has an explosive burn routine based on the JWL equation of state. The time step is determined from a courant stability condition. Additional information about the form of the resulting difference equations, the various options available, and the treatment of boundary and interface cells can be found in the HELP manual[1].

The calculations below were set up in axisymmetric geometry which implies that our representative cell is a torus. Our calculations neglected the strength effects but contained artificial viscosity. J. M. Walsh[4] reported in similar calculations that his results were not substantially affected when the strength effects were incorporated in the calculations.

III. THE COMPUTATIONAL RESULTS AND EXPERIMENTAL COMPARISONS

The problem considered in this study is a shaped charge with a hemispherical copper liner of an outer diameter 2R = 50.8mm, thickness = 1.9mm, and has a hemispherical charge layer 12.7mm thick of PBX 9404 as shown in Figure 2. The computation was set in an axisymmetric grid containing 50 x 90 cells. Each cell was 0.8mm x 0.8mm in the region containing the liner. The cells were gradually enlarged radially and axially to the end of the grid. Ten equally spaced angularly latitude circles, 9° apart, were selected on the outer hemispherical surface of the charge. The simultaneous initiation of these rings was assumed to simulate a convergent detonation wave and the experimental situation.

Figure 3 shows the liner configuration at t = 1.92 μs and the velocity field in the liner and explosive. The liner is seen starting to collapse after being hit by the almost convergent detonation wave about 0.46 μs earlier. The lack of confinement on the equatorial plane of the charge allows the detonation products to expand rapidly from that surface causing a departure from the idealized collapse depicted in Figure 1. The equatorial section of the liner starts to elongate at the explosive-metal-air interface. As the equatorial rarefaction wave travels toward the pole the pressures are relieved. Figure 4 shows the pressure in the flow field at t = 1.92 μs which is seen to be large in the collapsing liner and has a maximum value of 0.328 Mbar. About this

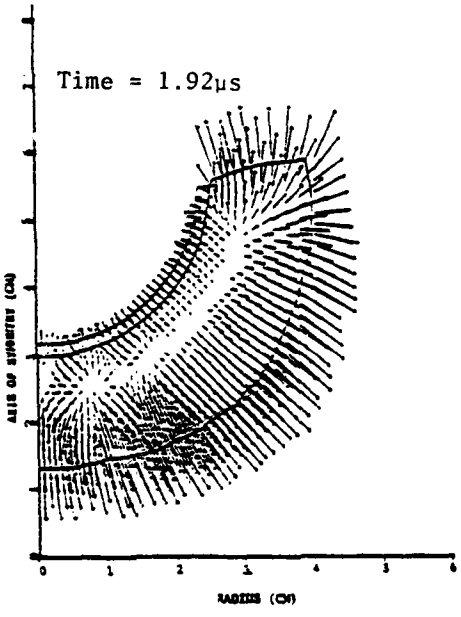


Figure 3: Early stage of liner collapse

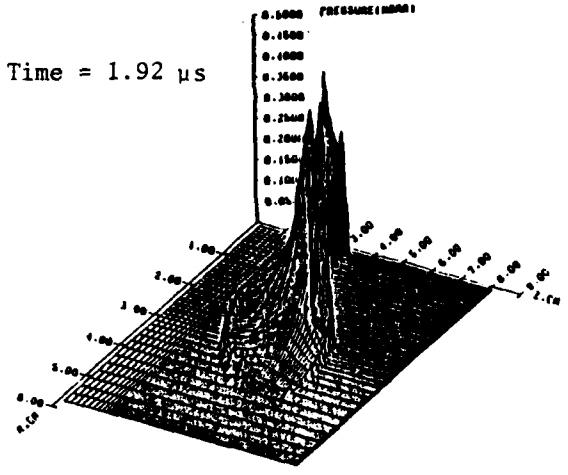


Figure 4: Pressure field at an early stage

time the effect of the rarefaction from the nearest unconfined spherical surface of the charge starts to affect the liner surface and the pressure begins to decrease. Figure 5 shows the flow velocity field at t = 5.8 μs. The liner is observed to be getting thicker at the pole and elongating further in the equatorial region. The pressure field at this time is seen in Figure 6 to have decreased significantly by that time. The maximum pressure in the liner material is found to be 0.15 Mbar. As the liner collapse advances, the pressure in the liner starts to increase again

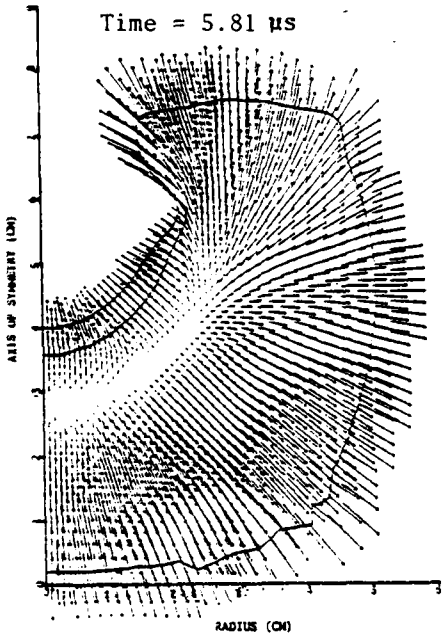


Figure 5: Flow field prior to jet formation

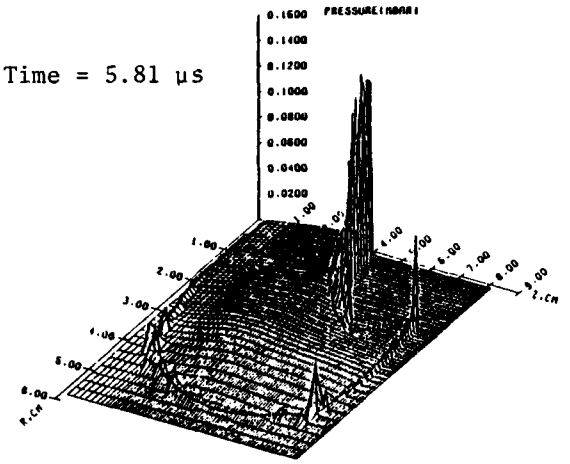


Figure 6: Reduced pressure field due to nonconfinement

due to liner compression and reaches a maximum value of 0.603 Mbar, at t = 9.2 μs as can be seen in Figure 7. The pressure decreases thereafter due to the influence of the equatorial rarefaction wave and the expansion of liner material arising from jet formation and elongation. The equatorial rarefaction wave reaches the polar region about t = 7.2 μs and the jet becomes discernable after that time. Figures 8, 9, and 10 show the process of jet formation and the beginning of its flight. The short arrows show the direction of flow velocity in the different layers of liner material. It is clear from those figures that the jet forms from the innermost liner layer. Figure 11 shows the equatorial section of the liner in the process of impacting the jet. A radiograph of the jet from such a system, at late time, showed part of the jet to be missing. The cause of the missing part was not known then but our calculations

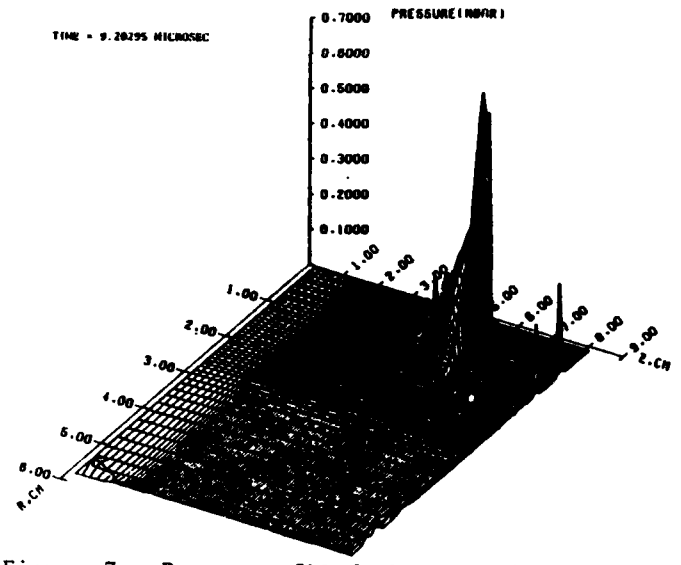


Figure 7: Pressure filed shortly after jet formation

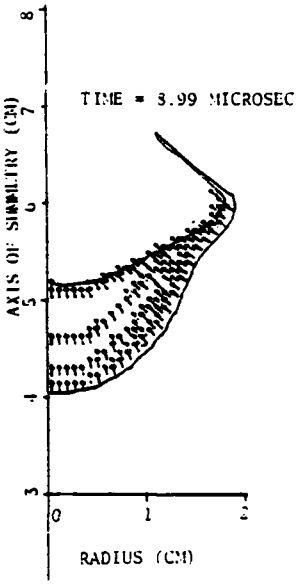


Figure 8: Early jet formation

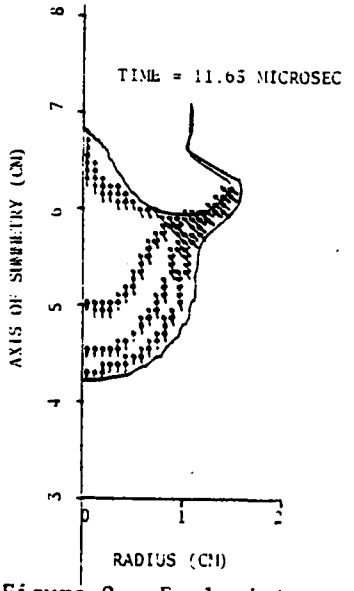


Figure 9: Early jet

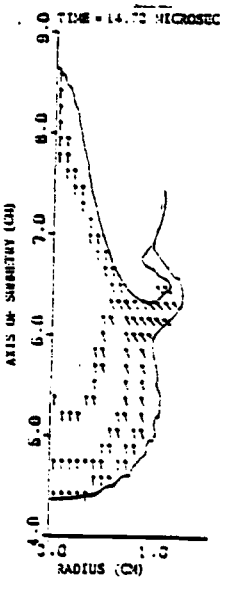


Figure 10: Jet elongation

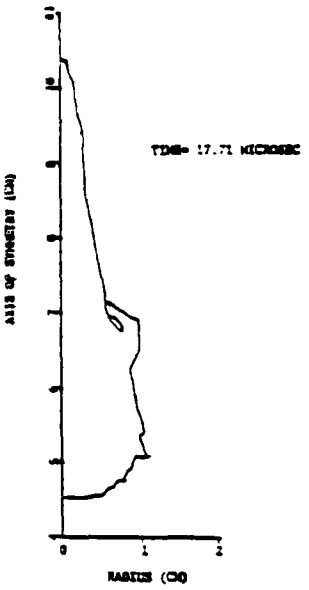


Figure 11: Jet pinchoff

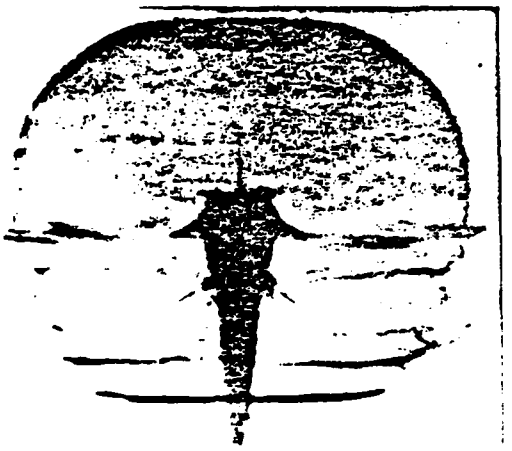


Figure 12: Radiograph of jet pinchoff 17μs

later shown in Figure 11 explained it. Figure 12 shows an experimental radiograph taken subsequently, of the jet being struck by the equator at about t = 17 μs. Figure 13 shows a plot of velocity versus time for ten passive tracer particles placed on the inside surface of the hemispherical liner. As remarked earlier all these particles attain jet velocities. Similar plots of velocities of passive tracers placed across liner thickness reveal that only those placed on the innermost liner surface attain jet velocities (i.e. $v \geq 2\text{mm}/\mu\text{s}$). Figure 14 shows a plot of the average velocity components of the metallic liner as functions of time, while Figure 15 shows the different liner energies as functions of time. The initial rarefaction wave arriving at the liner surface due to the nonconfinement of the spherical charge surface reduces the liner acceleration, while the equatorial rarefaction wave causes the jet to become distinguishable. The continued liner compression converts the liner radial momentum to axial momentum. The total liner energy increases rapidly at first and approaches an asymptotic value later on. Figure 16 shows a plot of jet velocity as a function of cumulative mass at various times. The jet mass continually increases as more metal is accelerated to jet velocities.

It was found computationally that about 18.7% of the hemispherical liner forms the jet (i.e., has velocity $\geq$ 2mm/μs). Experimental measurements estimate the jet mass to be about 18% of the liner mass. Our calculations indicate that about 19.9% of the explosive energy is delivered to the copper liner of which 11.8% is in kinetic energy form. The kinetic energy of the jet is 41% of the total energy of the liner.

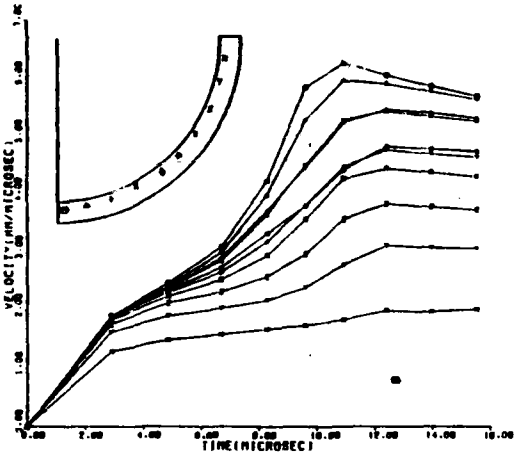


Figure 13: History of velocities of tracer particles on inner liner surface

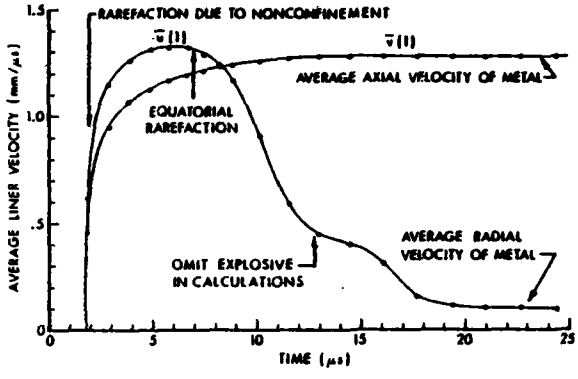


Figure 14: Histories of velocity components of liner

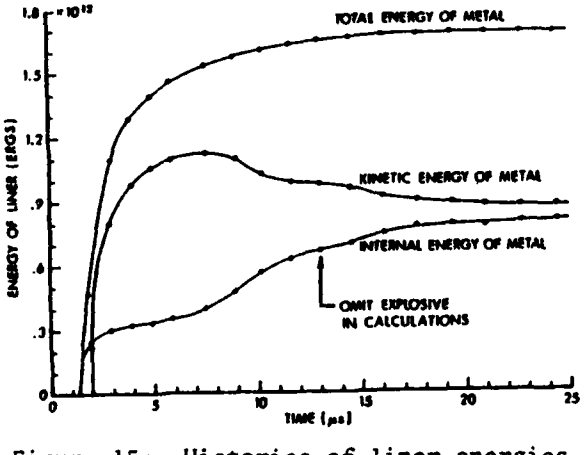


Figure 15: Histories of liner energies

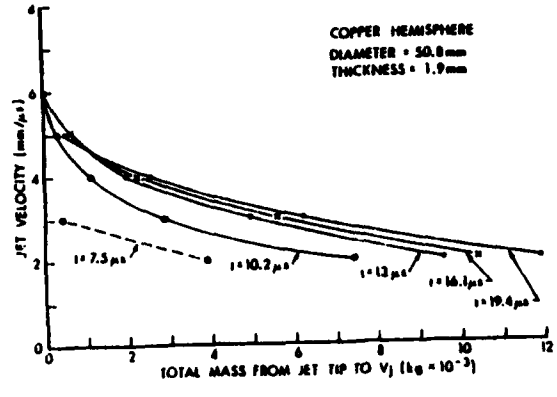


Figure 16: Jet velocities vs cumulative mass of liner

This is in agreement with experimental observations. The jet tip velocity was found to be 6.42 mm/μs. Two experimental measurements of jet tip velocity for such a charge were made at BRL of 7.07 mm/μs and 7.57 mm/μs. In the experimental tests the hemispherical charge used was 0.151 kg of PBX9404 and the initiation package contained 0.2685 kg of composition B-3. In the computations only the energy from the hemispherical charge was incorporated in the calculations. If one adds the kinetic energy of the flier plate in the initiation package to the explosive energy of the charge and accounts also for the effect of confinement provided by the flier plate then the resultant jet tip velocity[5] will be found to be 7.35 mm/μs which is within the range of experimental measurements. Parametric studies which included variations of liner and charge thicknesses, confinement and geometry were conducted and their effects on the resultant jets studied. The details of those studies can be found in reference 5.

IV. CONCLUSIONS

It is clear from the above study that the processes of collapse and jet formation for a hemispherical liner which is set in motion by an almost convergent detonation wave are fundamentally different from those of a cone[6,7]. The various hemispherical liner elements converge in this case towards a single point as they collapse. The jet forms due to the compression of the liner material which squeezes out the jet as in an extrusion process. The jet forms from the innermost layer of liner material[8]. No inverse velocity gradient was observed in jets from such systems and consequently the leading jet particle has a small mass. The total amounts of jet mass and kinetic energy from such a system are similar to those obtained from conical systems (same order of magnitude), although their distributions along the jet length appear to be different. Jets from the present hemispherical systems have greater length. Experimental observations indicate that jets from such systems have longer breakup time. According to jet penetration theory[6] the depth of penetration P by a jet of length ℓ is given by

$$P = \ell\left(\frac{\lambda\rho_j}{\rho}\right)^{1/2}, \qquad (9)$$

where ρ_j is the jet density, ρ is the target density, and λ is a parameter which is equal to one for continuous jet and equals two for dispersed particle jet. Equation (9) indicates that deeper penetrations would be obtained from the longer jets obtained from the present hemispherical systems. This has been observed experimentally to be the case, although the penetration holes have smaller cross sectional area since these jets have smaller cross sections as a consequence of the conservation of mass.

REFERENCES

1. L. J. Hageman, et al, "HELP, A Multi-Material Eulerian Program for Compressible Fluid and Elastic-Plastic Flows in Two Space Dimensions and Time," Systems, Science and Software Report 75-2654 (1975).

2. P. D. Lax, "Weak Solutions of Nonlinear Hyperbolic Equations and Their Numerical Computation," Comm. on Pure and Appl. Math, Vol VII, p159-193 (1954).

3. J. H. Tillotson, "Metallic Equations of State for Hypervelocity Impact," General Atomic Report GA-3216 (1962).

4. J. M. Walsh, Private communication.

5. A. R. Kiwan and A. L. Arbuckle, "Study of Liner Collapse, Jet Formation and Characteristics From Implosive Shaped Charge Systems," BRL R 2028 (1977).

6. G. Birkhoff, D. P. MacDougall, E. M. Pugh, and Sir G. I. Taylor, "Explosives with Lined Cavities," J. Appl. Phys. 19 (1948) p. 563.

7. E. M. Pugh, R. J. Eichelberger, and N. Rostoker, "Theory of Jet Formation by Charges with Lined Conical Cavities," J. Appl. Phys. 23 (1951) P. 532.

8. A. R. Kiwan and H. Wisniewski, "Theory and Computations of Collapse and Jet Velocities of Metallic Shaped Charge Liners," BRL R 1620, (1972).

THEORETICAL AND EXPERIMENTAL STUDIES OF HEMISPHERICAL SHAPED-CHARGE LINERS

Janet E. Lacetera and William P. Walters
Ballistic Research Laboratory
US Army Armament Research and Development Command
Aberdeen Proving Ground, Maryland 21005

The collapse, jet formation and performance of shaped-charge liners are influenced by many factors. These factors relate basically to the geometry and characteristics of the high energy explosive used to collapse the liner as well as to the liner geometry and liner material properties.

Hemispherical liners are especially sensitive to the variation in wall thickness from pole to equator, i.e., wall taper. In fact, the wall taper can dramatically alter the collapse and jet formation process and the performance characteristics of hemispherical liners. In this paper, the collapse and jet formation behavior of Electrolytic Tough Pitch (ETP) copper hemispherical liners with severe wall taper ratios of up to two to one will be examined using the HELP[1,2] code.

A uniform-wall-thickness liner was used as the reference case to assess the effects of severe wall taper ratios. This reference liner had a constant wall thickness of 3.79 mm and an outside diameter of 127 mm. The liner was driven by 75/25 Octol high explosive with 127-mm head height (distance from base of charge to pole of the liner) and sub-calibration (overlap of explosive on each side of liner) of 3 mm. The explosive was point detonated on the axis of symmetry at the base of the charge. In the experimental test assembly, an aluminum casing 190 mm long and 3.2 mm thick surrounded the charge. This casing was omitted from the analytical calculations since the thin aluminum body has a minimal effect on the liner collapse and formation process. Figure 1 illustrates the analytical liner configuration as well as tapered wall configurations of the thick and thin pole designs investigated in this study.

The thick pole designs utilized the same equatorial thickness as the uniform wall hemisphere and achieved the thickened pole effect by offsetting the center of the inner wall of the liner to increase the size of the pole region. Similarly, the thin pole effect was achieved by increasing the equatorial thickness over that of the uniform wall hemisphere and holding the pole region constant by offsetting the center of the semicircle forming the inner liner. No attempt was made in this study to optimize the liner mass. Both thick and thin pole design were more massive than the uniform liner.

It is also feasible to form tapered liners with less mass than the uniform wall liner by causing the maximum thickness of the tapered liner to be the thickness of the uniform wall liner and removing material to obtain the prescribed tapers. This approach was taken in a 1978 study[3] performed at the BRL/ARRADCOM on hemispherical wall variations using two to one taper ratios. In this case the thick pole configuration produced a jet with a higher tip velocity than the jet produced by the uniform wall liner. The thin pole liner did not, however, form a coherent jet due to the extreme two to one taper ratio at the pole. This behavior was quite different from that observed by increasing the liner mass as we presently show.

Other open literature studies regarding tapered hemispherical liners were reported by the Los Alamos Scientific Laboratory (LASL). The LASL design called the TLC (Tapered Liner Charge)[4] investigates tapers of less severe ratios.

In the present study, the reference copper hemisphere was varied from pole to equator using taper ratios of 1.5 to 1, 2.0 to 1 (thick poles) and 1 to 1.5 and 1 to 2.0 (thin poles) as shown in Figure 2. This study was designed as a reference or base case from which further design improvements could be made. Besides wall tapering, other improvements are possible depending on the application of the warhead, such as heavy steel confinement of the explosive,[3,5] a liner taper resulting in less mass,[3] an alternate liner material,[4,5] an optimized mass distribution of the liner material[4] and an optimization of the explosive, explosive geometry, and mode of initiation.[6]

The short-time collapse and jet formation properties were examined with the two-dimensional finite difference Eulerian computer code HELP. HELP is capable of describing unsteady multimaterial interactions and treating compressible fluids or solids in the hydrodynamic or elastic-plastic regime. The compressible, two-dimensional mass, momentum, and energy conservation equations are solved together with an equation of state (in this case Tillotson) that governs the thermodynamic behavior of the liner and body material. For the calculations presented here the code makes use of an explosive burn routine based on JWL equation of state.[1,2]

Figure 3 illustrates the collapse and jet formation sequence at early times for the uniform wall liner as calculated with HELP. Noticeable deformation has occurred by 10 μs after denotation with the penetrator developing a jet tip velocity of 4.0 mm/μs (Table 1) and becoming increasingly well defined as time passes. This collapse and formation sequence as well as the predicted jet tip velocities show excellent agreement with experimentally obtained data.

TABLE I

Liner	Jet Tip Velocity (mm/μs)
Uniform Wall	4.0
Thin Pole	
(1 to 1.5)	3.8
(1 to 2.0)	3.6
Thick Pole	
(1.5 to 1)	3.6
(2.0 to 1)	3.3

After successfully modeling the collapse and jet formation behavior for a uniform wall hemisphere, the tapered designs were analyzed. In the case of the thick pole liners, the intent of these designs was to create a slower but more massive jet. Such a jet should have a longer breakup time, that is, remain continuous longer than uniform wall liners, and should be less susceptible to radial breakup while spinning when gun launched. Figures 4 and 5 present the collapse and jet formation process for thick pole tapers of 1.5 to 1 and 2.0 to 1, respectively. In these cases, the equatorial thickness was conserved at 3.79 mm and the pole thickness set at 1.5 or 2.0 times the equatorial thickness. Both the 1.5 and 2.0 tapers produce substantially more massive jet than in the uniform case, and show a jet tip velocity decrease of 10% and 17.5%, respectively. For the early time computer results it was apparent that the thick pole penetrators would form massive coherent jets and this was verified experimentally.

The thin pole configuration involved the same severe taper ratios, but in these cases the equator was 1.5 or 2.0 times the nominal pole thickness (3.79 mm). Figures 6 and 7 show the collapse and jet formation pattern to be expected. These early time sequences were viewed with some concern because of the appearance of an indentation at the base of the jet which could presage a later instability on the penetrator. However, late time experimental radiographs show the formation of a coherent penetrator unlike the low mass, thin pole configuration reported by Aseltine et al.[3] The improved behavior is undoubtedly due to the thicker wall design. The effect of the thick equator is to slow down the rear of the jet and increase the velocity gradient. The thin pole penetrators are predicted to yield higher tip velocities (3.8 for 1 to 1.5 and 3.6 for 1 to 2.0) than the thick pole designs, but they reach lower velocities than the less massive uniform wall hemisphere. We note also that the thin-pole, thick-equator designs will have a lower jet tail or rear velocity due to the more massive equatorial portion of the liner. This thick equator effect will slow down the collapse of the base of the liner and, for the appropriate explosive geometry, may prevent the jet pinch-off effect observed by Kiwan.[6]

In general, we see that liner taper, both in degree and direction, alters the collapse and jet formation process of the hemispherical liner. Exemplifying this, Figure 8 compares the uniform wall, 2 to 1 taper thick pole and the 1 to 2 taper thin pole designs at 10 and 19 μs. The formation of the penetrator is clearly different for these three cases.

We can conclude from this study and evidence presented in the works cited, that the HELP code can be an effective design tool in creating unique warhead configurations. Unconventional liner wall thickness tapers, various liner geometries, various liner materials, heavy confinement effects, head height effects, explosive geometry variations, types of explosives and modes of explosive initiation can all be simulated with HELP. Here we found that the collapse and jet formation process, as well as the jet tip velocity for hemispherical liners was accurately predicted by HELP. Such information is extremely useful when fabricating new designs and an a priori knowledge of the jet tip velocity is necessary to set the appropriate delay times for the experimental flash radiographs.

Reviewing the results obtained by the severe wall tapers, we note that the thick pole designs result in a slower, more massive jet than the uniform hemisphere. The thin pole designs created by adding mass form a coherent jet with a large breakup time. Further studies involving the properties of tapered hemispherical liners are underway based on the encouraging results of these preliminary investigations.

REFERENCES

(1) Hageman, L. J., et al., "HELP-A Multimaterial Eulerian Program for Compressible Fluid and Elastic-Plastic Flows in Two Space Dimensions and Time," Systems, Science, and Software Report, 1975.

(2) Lacetera, Joseph, Lacetera, J. E., and Schmitt, J, "The BRL CDC 7600 Version of the HELP Code," BRL report in publication, 1979.

(3) Aseltine, C., Walters, W. P., Arbuckle, A., and Lacetera, J. E., "Hemispherical Shaped Charges Utilizing Tapered Liners," Proceedings of the Fourth International Ballistics Symposium, Monterey, CA, Nov 78.

(4) Carter, W., "New Developments in Shaped Charge Technology," Petroleum Engineer International, Apr 78.

(5) DiPersio, R., Jones, W., Merendino, A., and Simon, J., "Characteristics of Jets from Small Caliber Shaped Charge with Copper and Aluminum Liners," BRL Memorandum Report 1866, Sep 67.

(6) Kiwan, A. and Arbuckle, A., "Study of Liner Collapse, Jet Formation and Characteristics from Implosive Shaped Charge Systems," BRL Report 2028, Nov 77.

LINER GEOMETRY

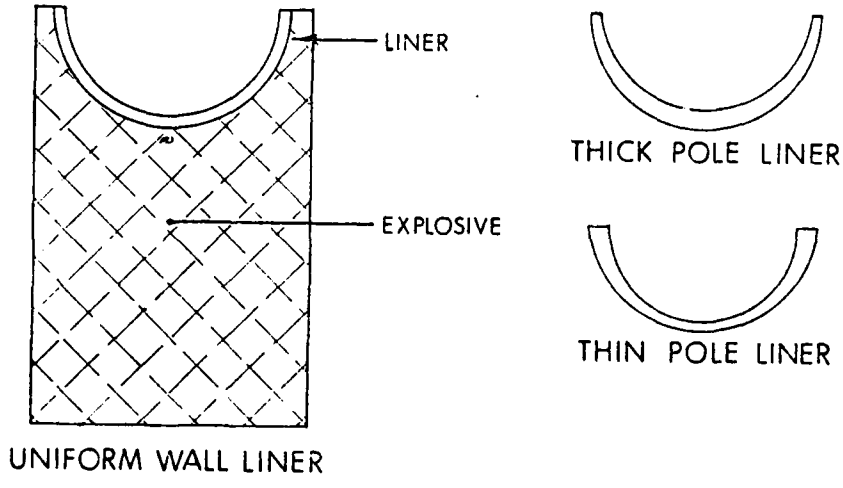


Figure 1

SEVERE TAPER DESIGNS

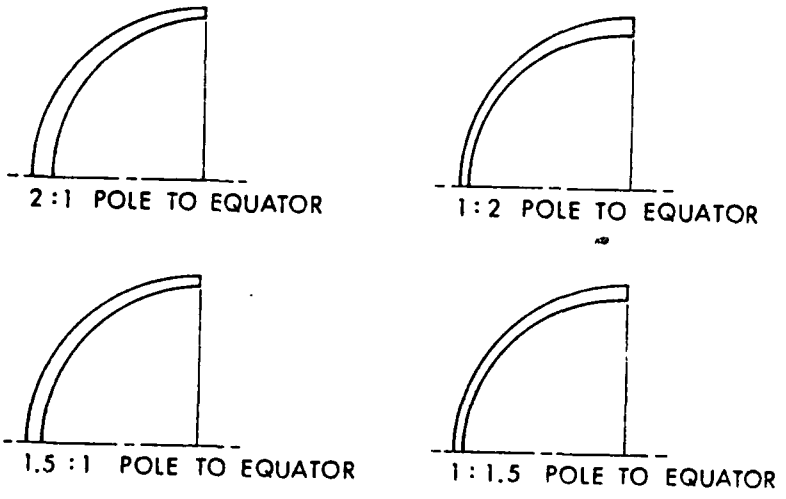


Figure 2

COLLAPSE & JET FORMATION PROCESS FOR UNIFORM HEMISPHERICAL LINER

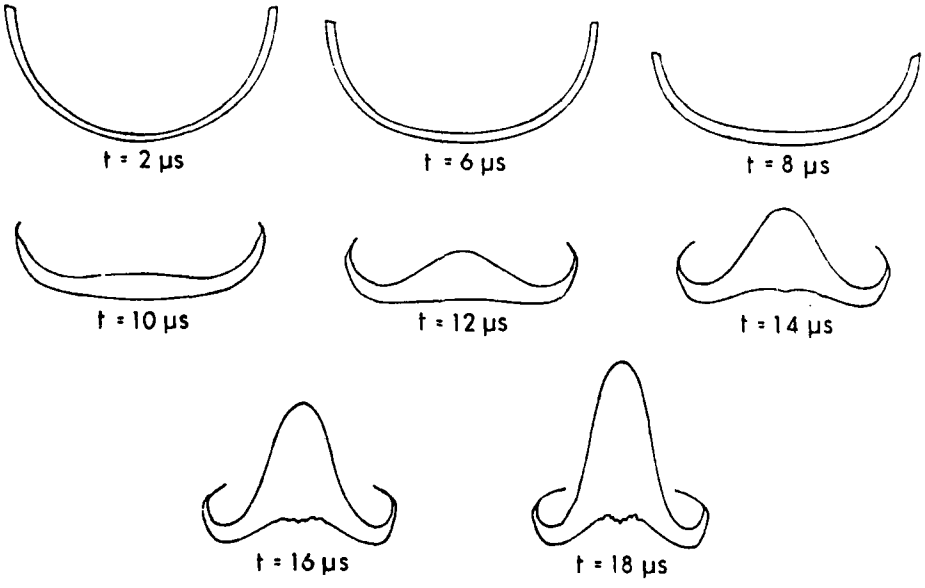


Figure 3

COLLAPSE & JET FORMATION PROCESS FOR A THICK POLE HEMISPHERICAL LINER. (1.5 TO 1)

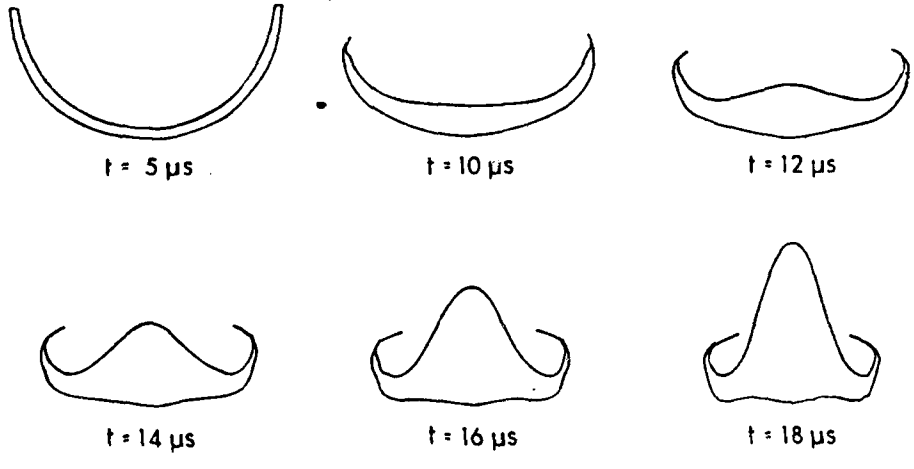


Figure 4

COLLAPSE & JET FORMATION PROCESS FOR THICK POLE HEMISPHERICAL LINER (2:1)

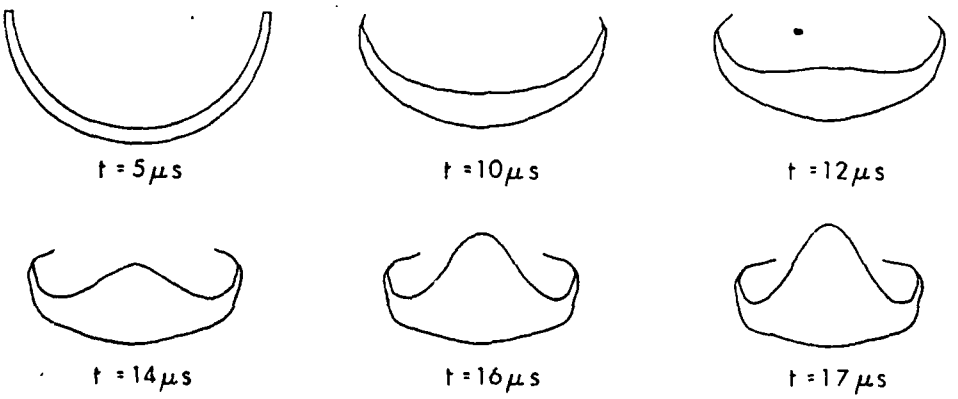


Figure 5

COLLAPSE & JET FORMATION PROCESS OF A THIN POLE HEMISPHERICAL LINER (1 TO 1.5)

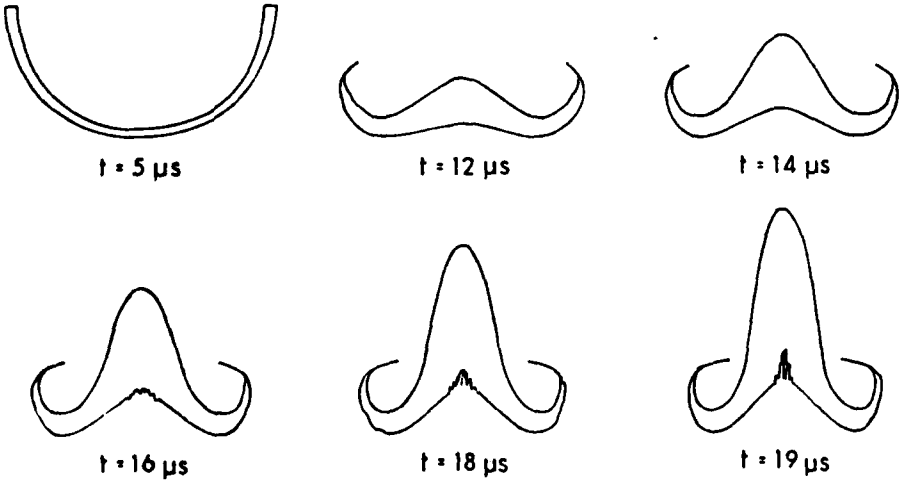


Figure 6

COLLAPSE & JET FORMATION PROCESS FOR THIN POLE HEMISPHERICAL LINER (1:2)

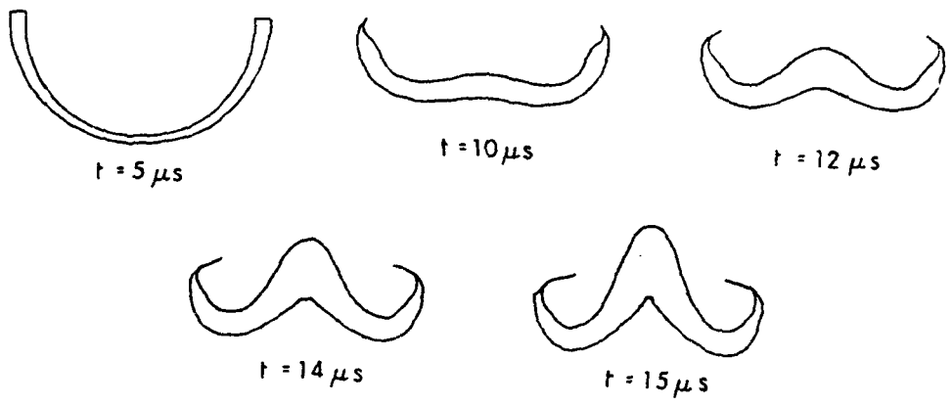


Figure 7

COMPARISON OF UNIFORM, THIN & THICK POLE AT 10 & 14 μs (2:1) (1:2)

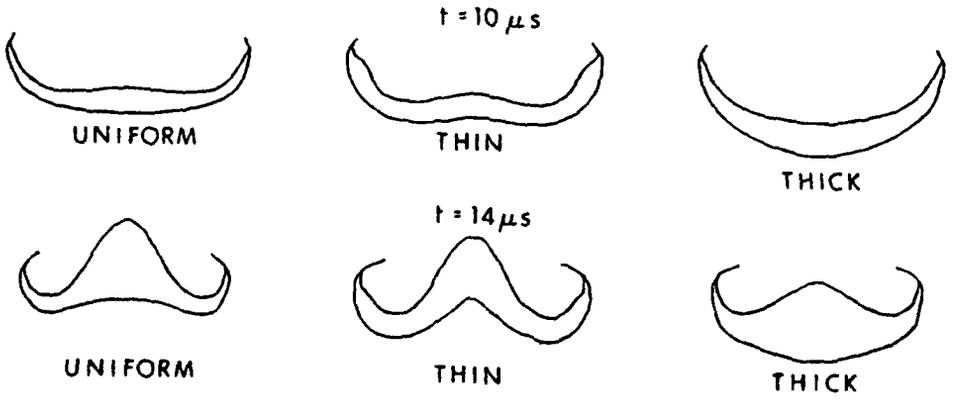


Figure 8

1 JUN 79

OPTIMAL MIXED STRATEGIES IN DYNAMIC GAMES*

P. R. KUMAR
Department of Mathematics
University of Maryland Baltimore County
Baltimore, Maryland 21228

ABSTRACT. We treat a class of two person, zero sum, dynamic games of the type $x_{K+1} = f(x_K, u_K, w_K)$. The two players each wish to steer the system to a different subset of the state-space. The optimal solutions for this class of games have to be sought in the class of mixed strategies. A theory of optimality is developed.

The specific class of games considered is one of the most classic of the problems in game theory. A gun is firing at a moving object. How best should the object move in order to reach a certain destination? Conversely, where should the gun fire in order to prevent the object from reaching its destination? This problem occurs in different guises in a variety of situations. The moving object could, for example, be a ship or a tank. The optimal strategies of the two players have, of necessity, to be mixed.

---

* A more detailed treatment of this presentation can be found in Reference [1]. The research reported in this paper was supported in part by Scientific Services Agreement with Battelle Columbus Laboratories under Contract No. DAAG-29-76-D100 and in part by the U. S. Army Research Office under Grant No. DAAG-29-79-C-0064.

I. INTRODUCTION. Consider the two-person, zero-sum, dynamic game:

$$\begin{bmatrix} x_{K+1} \\ \\ y_{K+1} \end{bmatrix} = \begin{bmatrix} x_K - v(t - s(x_K)) \\ \\ 0 \end{bmatrix} + \begin{bmatrix} u_K \\ \\ w_K \end{bmatrix}$$

$x_0 \geq 0$ specified .

$u_K \in [- vs(x_K), + v_b\, s(x_K)]$

$w_K \in (- \infty, \infty)$

$J(\{u_K\}, \{w_K\}) = 1$ if for some n, $x_n < 0$ and

$(x_i + u_i) \in [y_{i+1} - r, y_{i+1} + r)$ for

$i = 0,1,2,\ldots,n - 1$

$= 0$ otherwise .

Players I and II both observe $\begin{bmatrix} x_K \\ y_K \end{bmatrix}$, the state of the system, and based on these observations choose u_K and w_K respectively to maximize and respectively minimize the cost criterion.

This system models the following dynamic game.

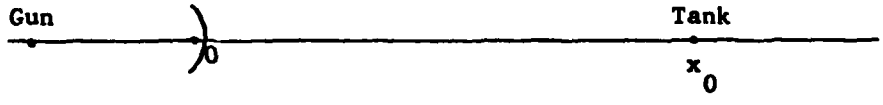


Figure 1

A gun located as shown wishes to prevent a tank located initially at x_0 from entering $(-\infty,0)$. The game proceeds as follows. The gun aims at $y_1 = w_0$ and fires a projectile at the tank. The projectile takes a time period $s(x_0)$ to reach the vicinity of the tank. During this interval, the tank whose forward and

backward velocities are bounded by v and v_b respectively, moves a distance $u_0 \in [-vs(x_0), +v_b\, s(x_0)]$. If the new position of tank $x_0 + u_0$ belongs to a certain neighborhood of the point at which the projectile lands — more precisely if $(x_0 + u_0) \in [y_{i+1} - r, y_{i+1} + r)$ — the tank is destoyed, and the game ends. If not, then assuming that the gun fires every t seconds, there is a certain time period $t - s(x_0)$ during which the gun does not fire. In this time period, the tank moves at full speed and takes up a position $x_1 = x_0 + u_0 - (t - s(x_0))v$ at the next instant of firing of the gun. If $x_1 \in (-\infty, 0)$, then the tank has accomplished its objective and the game ends. If not, the above sequence of events repeats itself. If the tank safely reaches $(-\infty, 0)$, then it wins and the payoff $J = 1$, otherwise $J = 0$.

II. MIXED STRATEGIES.

For any deterministic $\{u_K\}$ chosen by the tank (with each u_K chosen as a function of past history), the gun can choose $\{v_K\}$, to ensure $J = 0$. Similarly, for any deterministic policy $\{v_K\}$ chosen by the gun, the tank can choose $\{u_K\}$ to guarantee $J = 1$. Hence $\min_{\{v_K\}} \max_{\{v_K\}} J = 1 > 0 = \max_{\{u_K\}} \min_{\{v_K\}} J$. Therefore a saddle-point does not exist in the class of pure strategies. Turning therefore to mixed strategies, we define a mixed strategy for the tank as a collection $\{F_x : x \geq 0\}$ where each F_x is a probability distribution on $[x - vs(x), x + v_b\, s(x)]$. F_x is the probability distribution of $x + u$, given that the current position of the tank is x. Similarly, we define a mixed strategy for the gun as a collection $\{G_x : x \geq 0\}$ where each G_x is a prob. dist. on $(-\infty, \infty)$. The gun is assumed to choose v, the position at which it aims its fire according to G_x, when the tank is located at x at a firing instant.

Implicit in these definitions is a restriction of the mixed strategies to be both Markovian and stationary. However, the optimal mixed strategies belong to this class, and therefore we avoid notational complexities. The reader is referred to [1] for greater detail.

III. OPTIMAL SOLUTIONS. $t \equiv s(x)$

We consider the restrictive case where the flight time of the projectile is a constant precisely equal to the interfiring time of the gun. For more general situations the reader is referred to [1]. We also assume $vt = 2nr$ for some integer n.

The following definitions are necessary. Let $F = \{F_x : x \geq 0\}$ and $G = \{G_x : x \geq 0\}$ represent mixed strategies for the tank and gun respectively. Let $1 - K(x_0; F,G) := E[J|F,G]$ represent the probability that the tank achieves its goal, given that its initial position is x_0 and the mixed strategy pair (F,G) is adopted. If $\inf_G 1 - K(x_0; F^0, G) = \sup_F 1 - K(x_0;F,G^0) = 1 - K^0 x_0)$ then (F^0,G^0) will be a saddle-point in mixed-strategies and $1 - K^0(x_0)$ will be the value of the game, for the initial position x_0. Let I(a) denote the largest integer less than or equal to a and $H(\{x\})$ represents the prob. measure under H of the singleton set $\{x\}$.

Proposition 2

The game has a mixed strategy pair (F^0, G^0) which is a saddle-point for every x_0. The value of the game is given by

$$1 - K^0(x_0) = \frac{1}{1+f(2jr)} \quad \text{for} \quad \begin{matrix} 2jr \leq x_0 \\ < 2(j+1)r \end{matrix}$$

where f is given by the linear recurrence,

$$f(2jr) = 0 \qquad \text{for} \quad j = -1, -2, -3, \ldots.$$

$$= \frac{1}{n-1} \left[1 + \sum_{i=j-n}^{j-1} f(2ir) \right] \qquad \text{for} \quad j = 0,1,2,\ldots.$$

Also

$$F_x^0(\{x - 2jr\}) = \left[[1 - K^0(x - 2jr)] \sum_{i=1}^{n} \frac{1}{1 - K^0(x - 2ir)} \right]^{-1}$$

$$\text{for} \quad j = 1,2,\ldots,n$$

$$G_x^0(\{2I(\frac{x}{2r}) - 2jr - r\}) = 1 - \frac{1 - K^0(x)}{1 - K^0(2I(\frac{x}{2r})r - 2jr - r)}$$

$$\text{for} \quad j = 0,1,\ldots,n-1 .$$

Proof See [1] .

IV. SOME INTERPRETATIONS.

We provide below some interpretations of the above result.

Firstly, the tank does not utilize its backward motion capability. Such a capability is therefore unnecessary in the situation $t \equiv s(x)$. However for $t > s(x)$, see [1], such a capability is useful.

Secondly, the optimal prob. dist F_x^0 is almost a discrete unitorm prob. dist. on $[x - vs(x), x]$. Since a discrete uniform distribution is the distribution which renders the tank hardest to "hit", such a distribution is precisely what the tank would adopt if its only goal was survival, and if it did not have a destination. Therefore, the optimal mixed strategy F^0 is close to a pure survival strategy.

However, a similar situation does not hold for the gun.

V. CONCLUSIONS. We have considered a dynamic game of the type $x_{K+1} = f(x_K, u_K, w_K)$ where each player attempts to steer the system to a subset of the state-space. The particular system considered is a model of the encounter between a tank and a gun. Pure strategies are of no use in such a game, and a value does not exist in such a class. However, a value and saddle-point do exist in the class of mixed strategies. Such optimal mixed strategies have been presented.

VI. REFERENCES.

1) P. R. Kumar, "Optimal Mixed Strategies in Dynamic Games," Mathematics Research Report No. 79-3, UMBC, May 1979.

THE USE OF SIMILITUDE METHODS TO REDUCE THE SIZE AND COST OF GAME COMPUTATIONS

Morton A. Hirschberg, USA BRL/ARRADCOM
and
Benjamin E. Cummings, USAMSAA
Aberdeen Proving Ground, MD 21005

ABSTRACT. The Buckingham Pi Theorem was applied to Lanchester Linear Models and conventional war games (e.g. DIVLEV) to reduce their size and/or computer costs. The Lanchester models could not be reduced; however, the approach is a serious alternative for Lanchester modeling. Conventional games are candidates for size and cost reduction when the number of computations performed is large in relation to accounting or bookkeeping in the code. In addition, the use of similitude can play an important role in the formulation of new models and games.

I. INTRODUCTION. The scientific computer codes in use in the Army span the complete range from simple models of physical phenomena through large, complex simulations and games depicting division and Army level combat scenarios. The size and cost of game computations have been of concern to the authors and others for some time. Sophisticated mathematical techniques such as matrix decomposition or manifold reduction are just beginning to find their way into the analysis and structure of new models and games but are rare even in recent (i.e. 3-5 year old) codes.

This paper is an examination of existing models and games and the possible reduction of their size and/or running costs. In addition, we wish to present a methodology for minimizing size and cost of future models and games. The basis of the analysis is the Buckingham Pi Theorem or Pi Theorem, the major tool of dimensional or similitude analysis.

A practical working definition of similitude is the investigation of complex phenomena using experiments or models of similar phenomena which are easier to describe and analyze Using similitude one forms a likeness to the complex phenomena incorporating as many features of that phenomena which adequately represent it in the areas of interest. Representation of some aspects of the phenomena may be distorted or completely lacking. It is not our intention in this paper to deal with the complete theory of similitude, its strengths and weaknesses, or to teach one exactly how to use similitude (especially what constitutes an adequate representation of a complex phenomena). These topics are outside the scope of the paper; however, we suggest the following references for further information on similitude:

L. I. Sedov, Similarity and Dimensional Methods in Mechanics, Academic Press, New York, 1959.

H. L. Langhaar, Dimensional Analysis and Theory of Models, Wiley, New York, 1951.

W. E. Baker, P. S. Westine, and F. T. Dodge, Similarity Methods in Engineering Dynamics, Hayden, New Jersey, 1973.

The Journal of the Franklin Institute, Special Issue, Modern Dimensional Analysis, 292, 6, December, 1971.

II. Buckingham Pi Theorem. The principal tool of similitude and dimensional analysis is the Buckingham Pi Theorem. Its simplest statement, shown below, while not strictly correct, serves well for most instances.

Buckingham Pi Theorem

> Given n variables involving N reference units, these may be combined to form n-N dimensionless parameters each having N+1 variables.

The theorem tells us that if we have a set of n dimensional formulas expressed in terms of N reference units we may reduce this set of formulas to an equivalent set of n-N formulas (called pi terms) each of which contains N+1 of the original n variables. A complete algebraic treatment of the Pi Theorem appears in Langhaar.

As an example from physics, consider the following five dimensional equations expressed in terms of three reference units.

$$F = F$$

$$\mu = FL^{-2}T$$

$$A = L^{2}$$

$$\rho = FL^{-4}T^{2}$$

$$v = LT^{-1}$$

where F = force, μ = viscosity, A = area, ρ = density, v = velocity, L = length, and T = time. These may be combined to form two dimensionless pi terms:

$$\pi_1 = \frac{F}{A\rho v^2} \quad \text{and} \quad \pi_2 = \frac{\mu}{vA^{\frac{1}{2}}\rho}$$

(note π_2 is the Reynold's number).

An important aspect of the Pi Theorem is the nature of the functional relationship of the quantities characterizing the phenomena under investigation. The numerical values of dimensional quantities depend upon a choice of units which has no connection with the substance of the phenomenon. That is, the functional relations are independent of a choice of units. This fact is most germain to our application of the Buckingham Pi Theorem to phenomena which are not necessarily physical. It allows us to introduce nonphysical reference units. In addition, it allows us to introduce artificial reference units into physical or nonphysical phenomenon as long as the functional relationships of the phenomena are not altered. With this knowledge, we may now extend the use of the Buckingham Pi Theorem to arbitrary phenomena such as economics, or in our case military models and games.

III. The Lanchester Model. The authors chose two courses; first, an examination of the Lanchester linear model, second, an examination of DIVLEV, a tradition large scale war game model. Both will be discussed in detail below.

The Lanchester linear model is given by:

$$\frac{dM}{dt} = -a_{11}M - a_{12}N + r_1(t)$$

$$\frac{dN}{dt} = -a_{21}M - a_{22}N + r_2(t)$$

These two equations comprise a mathematical model describing the interaction of two military forces -- Red and Blue -- which inflict an attrition upon each other while each is also undergoing replacement. We shall identify Red forces by M and Blue forces by N.

The first equation, which describes the rate of change, (dM/dt), of a military force indicates that the Red force is modified by: (1) a loss component, represented by $(-a_{11}M)$, which is proportional to the magnitude of the Red force, M, (2) a loss component, represented by $(-a_{12}N)$, which is proportional to the magnitude of the opposing Blue force, N, and (3) a replacement rate, $r_1(t)$, describing the forces added to the Red side. The second equation has a similar interpretation for Blue forces.

As a mathematical model, we note that these equations are characterized by four basic aspects:

(a) Economy of Expression
(b) Availability of Solutions
(c) Determination of Values of Parameter
(d) Applications and Predictions.

To simplify the analysis, the authors selected for study the combat situation involved in the capture of Iwo Jima by US forces[1]. The following equations result:

Japanese Forces without replacement

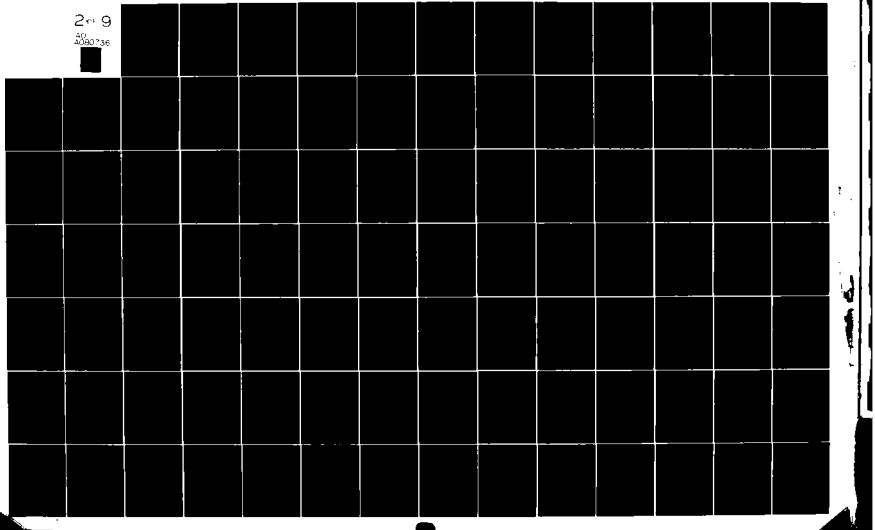
$$\frac{dM}{dt} = -BM$$

US Forces with replacement

$$\frac{dN}{dt} = -AN - DM + R(t)$$

[1] T. E. Oberbeck, Military Operations Research Lecture 1, Lanchester's Equations, IDA, 1964.

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH).(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL



Formulating a difference equation solution for the dimensional version of this model we obtain:

$$M_{k+1} = -BN_k\Delta t + M_k$$

$$N_{k+1} = (1-A\Delta t)N_k - DM_k\Delta t + R_k(t)\Delta t$$

There are six parameters in this model; M, N, R, A, B, and D.

The difference equation solution for the nondimensional version of this model is:

$$M_{k+1} = -\pi_B N_k\Delta\tau + M_k$$

$$N_{k+1} = (1 - \pi_A\Delta\tau)N_k - \pi_D M_k\Delta\tau + \pi_R(\tau)\Delta\tau$$

Here we have four dimensionless parameters π_A, π_B, π_D and π_R; however, we have the following expressions for the π's.

$$\pi_A = \frac{\gamma}{\sqrt{\alpha\beta}}$$

$$\pi_B = \frac{\beta N_o}{\sqrt{\alpha\beta}\ M_o}$$

$$\pi_D = \frac{\alpha M_o}{\sqrt{\alpha\beta}\ N_o}$$

$$\pi_R = \frac{R(\tau)}{\sqrt{\alpha\beta}\ N_o}$$

where: M_o is the initial Japanese force.
N_o is the initial US force.
α is the attrition of the US due to Japanese.
β is the attrition of the Japanese due to US.
γ is the loss factor based on the size of the US forces.
$R(\tau)$ is the replacement of the US forces.

and the dimensions of:

$$\alpha = \frac{US}{Japanese\ T}$$

$$\beta = \frac{Japanese}{US\ T}$$

$$\gamma = \frac{1}{T}$$

$$R = \frac{US}{T}$$

where: T is time.

In comparing the dimensional and nondimensional versions of the Iwo Jima linear Lanchester models we see that the nondimensional version has two less parameters. The computational size of both versions are comparible and both are small. The computational speeds of both are comparible. Furthermore, unless the replacement rate is constant, the bulk of the computational time in both versions may be governed by calculating $R(t)$ or $\pi_R(\tau)$.

The conclusion is that in general either version of the Lanchester linear model should run extremely fast. Cost reductions using the Lanchester model will depend upon the skill of the analyst in formulating as few computer runs as possible. The use of similitude does however present an alternative method of viewing the Lanchester equations.

In passing it should also be noted that when closed form solutions of the Lanchester equations exist they are already in nondimensional form.

IV. DIVLEV DIVLEV[2] is a combined arms war game model that takes player determined organizations and tactical decisions for both forces in the game and determines the movement and attrition that occur based on this information. Its primary purpose is the evaluation of material systems. DIVLEV consists of a main program, 66 subroutines and 10 functions.

The game may be played in an open or closed fashion. It is usually run with a 5 minute clock cycle with status reports generated at every 15 minutes game time and plots of force deployment generated at every 30 minutes of game time. The game is usually stopped after each 30 minutes of game time for analysis. A typical game may require anywhere from one to several months for analysis and involve many hours of computer time (see Figure 1).

Data for DIVLEV exists as in-line code (DATA statements) of a fixed nature and input parameters in the form of cards (as many as 38 different card types involving hundreds of parameters may be needed for one DIVLEV game).

[2]DIVLEV War Game Model Computer Program, USAMSAA, January 1977.

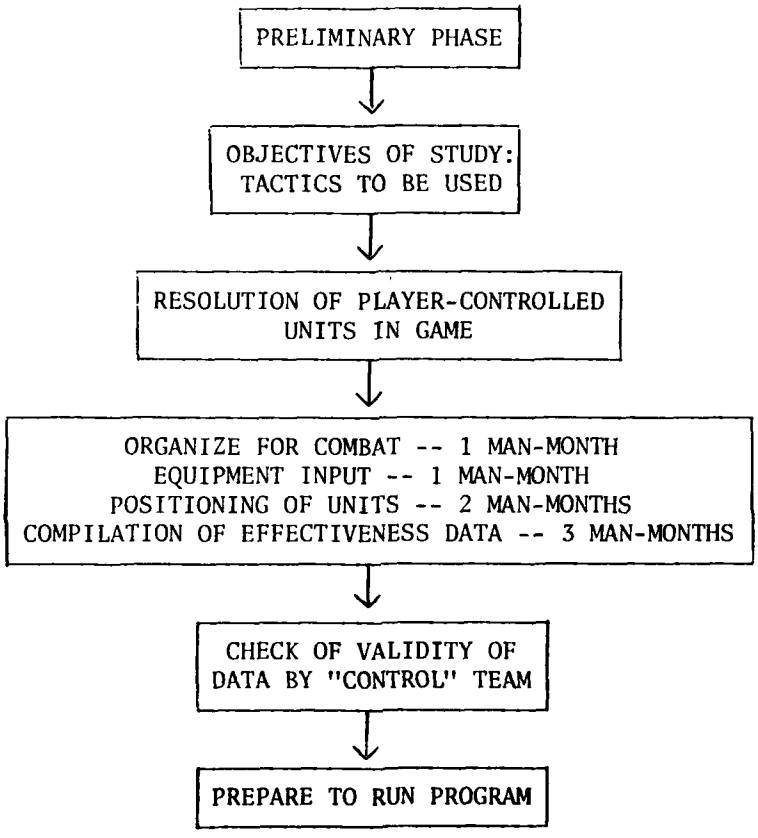


Figure 1. Sequence of Events in Preliminary Phase.

Using parameters from the data statements, 19 dimensional formulas were developed in terms of ten reference units. These are shown in Figure 2. The meanings of the variables and reference units are given in Table I.

The equations listed in Figure 2 are read, e.g., as follows:

IDEPLOYT = TIME
NPBP = NUMBER x WEAPON x KILLPERS

Noting the large number of variables and reference units one sees it is not practical to solve large dimensional problems by hand. A computer program developed by Sloan and Happ[3] to generate integer solutions was embedded into a procedure for solving, permuting, and identifying nondimensional parameters[4]. The importance of permuting dimensional equations is in the fact that an application of the Buckingham Pi Theorem produces one set of π terms for a given ordering of the dimensional equations. A particular ordering of the equations may yield trivally simple nondimensional solutions (e.g., TIME/TIME), while more fruitful solutions go unnoticed. By permuting the dimensional equations one increases the chances of finding interesting nondimensional solutions (of course many more trivial solutions are generated, too). The procedure developed in Reference 4 also attempts to identify solutions by comparing generated solutions with well-known nondimensional numbers (e.g., the Reynolds number (DENSITY x VELOCITY x LENGTH/VISCOSITY) or $\rho v A^{\frac{1}{2}}/\mu$ (see sample problem on page 2). The importance of identifying well-known nondimensional numbers is obvious when considering physical problems; in problems where artificial reference units are introduced the importance is that one may discover meaningful analogs to physical problems which explain the phenomena under investigation.

Returning to DIVLEV, and applying the Buckingham Pi Theorem one sees that there are nine Pi terms each involving 11 variables for each ordering of the dimensional equations. One such Pi term generated is:

$$\frac{\text{STDLP}}{\text{MTAH x BARTER}}$$

The explanation for this Pi term is that it relates the length of time a helicopter is with a unit, to the coverage the helicopter can give to that unit (a very important piece of information).

While this is all very interesting, did it help reduce the size or cost of DIVLEV? The answer is that, when the authors examined the DIVLEV code in detail, they found there were few modeling computations performed in relation to the accounting or bookkeeping in the code. That is, there were few models in comparison to the

---

[3]A. D. Sloan and W. W. Happ, Computer Program for Dimensional Analysis, NASA TN D-5165, April 1969.

[4]M. A. Hirschberg, The Evaluation, Manipulation, and Identification of Nondimsional Numbers, ARBRL-TR-02076, June 1978.

DIVLEV

DIMENSIONAL EQUATIONS

VARIABLES	REFERENCE UNITS									
	TIME	DISTANCE	TONNAGE	NUMBER	WEAPON	ARMORED	KILLPERS	KILLARMO	ARTILLERY	ROUNDS
IDEPLOYT	1									
BARTER	-1	1								
POST		2								
FUEL 1			1							
FUEL 2	-1		1							
NAMPRE				1					1	
NAMBRT				1					1	
STDLP		1								
NPBP				1	1		1			
NPBA				1	1	1	1			
NABP				1	1			1		
NABA				1	1	1		1		
AEL		2								
PCCAS1	-1						1			
PCCAS2	-1							1		
MTAH	1									
ITATAH	1									
RPM	-1									1
RPK								-1		1

Figure 2.

Table I. Meaning of Variables and Reference Units Used in Dimensional Formulas

Variables

IDEPLOYT	-	Time for unit to deploy.
BARTER	-	Barrier.
POST	-	Posture description of unit.
FUEL1	-	Fuel capacity.
FUEL2	-	Fuel usage rate.
NAMPRE	-	Number of artillery battalions.
NAMBRT	-	Number of artillery battalions.
STDLP	-	Distance between concentration points.
NPBP	-	Number of unarmored weapons that kill personnel.
NPBA	-	Number of armored weapons that kill personnel.
NABP	-	Number of unarmored weapons that kill armor.
NABA	-	Number of armored weapons that kill armor.
AEL	-	Area effects for artillery and missiles.
PCCAS1	-	Fraction of acceptable loss rate for personnel.
PCCAS2	-	Fraction of acceptable loss rate for armor.
MTAH	-	Length of time helicopter is with unit.
ITATAH	-	Length of time helicopter is in laager area.
RPM	-	Rounds per minute
RPK	-	Rounds per armor kill

Reference Units

TIME	-	Time
DISTANCE	-	Distance
TONNAGE	-	Tonnage
NUMBER	-	Number
WEAPON	-	Weapon
ARMORED	-	Armored
KILLPERS	-	Kill of personnel.
KILLARMO	-	Kill of armor.
ARTILLERY	-	Artillery
ROUNDS	-	Rounds

number of tests to determine which model to use. The bulk of the code and computer time is used in testing and not model computation. This finding was somewhat shocking but confirmed by examining another conventional game model.

V. DISCUSSION. The typical scaling problem of similitude analysis is one of saving money through testing at reduced size or reduced number of parameters. That is not fully the situation which can be exploited in the similitude analysis of existing games. A most valuable associated result of similitude analyses is reduced cost through improved understanding. One may hope for this result, provided an analyst has the initiative to apply these methods.

The typical scaling that we have referred to starts with a tabulation of relevant parameters; proceeds to the reduction of a set of pi-parameters; obtains the constraints that must be applied for realistic testing; and, finally verifies that the response functions exhibit null variation under changes of scale. The nature of games is to exploit the human variation of the players and so to develop the knowledge of how to win. Thus, only some of the parameters which may be developed can be held constant from run to run. The response function will thus be a function of scaled pi-parameters and uncontrolled or quasi-controlled pi-parameters, or what we might call upsilon-parameters (T_r). The reality of the game is that the scaling constraints are in fact relaxed by the nature of the model (man-in-the-loop) and so the analyst must recognize the loss of these constraints in his methods.

The conventional similitude analysis recognizes that the response functions take the form:

$$F_1(R_1, \pi_1, \ldots, \pi_N) = 0$$
$$\vdots$$
$$F_\ell(R_\ell, \pi_1, \ldots, \pi_N) = 0$$

and the associated response functions are:

$$R_1 = R_1(\pi_i)$$
$$\vdots$$
$$R_\ell = R_\ell(\pi_i).$$

The ordinary scaling of the response functions is achieved if scaling of variables leads to fixed pi-parameters and a resultant null variation in the response functions. Because of the presence of player in the game, it may not be possible to achieve this class of scaling in a similitude analysis of a war game. In that situation an inherently less restrictive analysis is appropriate. Instead of the restriction:

$$\pi_i = \text{constant for } 0 \leq i \leq N;$$

we introduce:

$$\pi_i = \text{constant for } 0 \leq i \leq n$$

$$T_i = \pi_i \text{ for } 0 < i \leq n$$

and

$$T_i \neq \text{constant for } n < i \leq N.$$

Then the variation of the response functions will change from

$$\delta R_j = 0$$

under scaling to

$$\delta R_j = \varepsilon_j \neq 0$$

and

$$\varepsilon_j = \varepsilon_j(T_i).$$

One method of obtaining useful information from these residuals is by making a Taylor's series expansion

$$R(\pi,T) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(\pi-\hat{\pi})^m (T-\hat{T})^n \partial^{m+n} R}{m!n!\partial\pi^m \partial T^n}$$

but of course $\pi = \hat{\pi} = \text{constant}$ so that

$$R(\pi,T) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n R}{\partial T^n} \cdot (T-\hat{T})^n.$$

The finite size of the computed set of games will, of course, truncate the series which can be fitted to the data set.

In many cases it will be desirable to use the residuals to do regression analysis rather than Taylor's series fits to the game results. Other methods should also be considered for analysis of these solutions. For instance solutions of the Lanchester model give the form for the time history of a simple engagement. This form could be used to find a trial solution to the response equations and the Galerkin method applied to optimize the parameters of the solution through driving the weighted error to zero. Most certainly there will be other methods which could be applied. The starting point for all, however, is in the application of the concepts of similitude modeling. Toward that end we offer the following conclusions.

VI. CONCLUSIONS. The conclusions reached during the study of conventional war games are:

a. Use of similitude showed definite results (e.g., air cover through a barrier).

b. Pi terms are useful for structuring future models. Dimensional equations can be written for input and output variables to show relationships for model and game construction and analysis of the outputs.

c. Pi terms are useful for studying doctrine in existing games.

d. Existing conventional games which contain too few physical computations in relation to the amount of testing should be examined in other ways for size and computation reduction (e.g., decision tables).

VII. SUMMARY. We have employed the Buckingham Pi Theorem, the major tool of similitude analysis to various models and war games and find it to be a useful tool, but one which warrants further use and further analysis.

The authors propose to apply dimensional analysis to other more heavily compute bound problems, such as the Hemp[5] code, with the expectation of finding relationships between variables more amenable to size and computation reductions.

In addition, the authors propose to catalog the results of many runs of the same games using similitude analysis applied to the outputs. The notion of saving results from old runs and applying new techniques to those results is not new; however, it deserves more attention than currently given to it.

Finally, the authors strongly feel the techniques presented in this paper offer new alternatives for looking at old problems and a means of looking at new problems and new problem areas in an old but promising way.

[5] E. D. Giroux, Hemp User's Manual, Lawrence Livermore Laboratory, UCRL-51079, Rev. 1, December 1973.

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

BOUNDS FOR EIGENVALUES OF HERMITIAN TRENCH MATRICES

T. N. E. Greville

Technical Summary Report

ABSTRACT

A banded matrix $H = (h_{ij})_{i,j=0}^{N}$ is one such that $h_{ij} = 0$ for $j - i > r$ and for $i - j > s$, where r and s are nonnegative integers. In [5] W. F. Trench and I called it strictly banded if, in addition, $r + s \leq N$. We also showed that a necessary condition for a strictly banded matrix to have a Toeplitz inverse is that it have a certain special structure fully characterized by two polynomials, $A(x)$ of degree r and $B(x)$ of degree s. I call a matrix having this special structure a Trench matrix. It was also shown in [5] that a Trench matrix is nonsingular if and only if $A(x)$ and $B(x)$ have no common zero, and that a strictly banded matrix has a Toeplitz inverse if and only if it is a nonsingular Trench matrix. In this paper there are established bounds for eigenvalues of Hermitian Trench matrices that depend only on the polynomials $A(x)$ and $B(x)$ and not on the order of the matrix.

AMS(MOS) Subject Classification: 15A09, 15A57.

Key Words: Toeplitz matrix, Band matrix, Hermitian matrix, Matrix eigenvalue, Eigenvalue bound.

Work Unit Number 2 - Other Mathematical Methods

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

BOUNDS FOR EIGENVALUES OF HERMITIAN TRENCH MATRICES

T. N. E. Greville

Mathematics Research Center
University of Wisconsin - Madison
Madison, Wisconsin 53706

1. Introduction.

In [5] W. F. Trench and I studied the conditions under which a band matrix has a Toeplitz inverse. More specifically, let $H = (h_{ij})_{i,j=0}^{N}$ be a real or complex matrix, where

$$h_{ij} = 0 \quad \text{for} \quad j - i > r \quad \text{or} \quad i - j > s \ ,$$

with

$$r \geq 0 \ , \qquad s \geq 0 \ .$$

Such a matrix we called a band matrix. We called it strictly banded if, in addition,

$$r + s \leq N \ .$$

Let

$$H_i(x) = \sum_{j=0}^{N} h_{ij} \, x^j$$

be the generating function of the elements of the ith row of H. In this paper I define a Trench matrix as a strictly banded matrix such that

$$H_i(x) = \begin{cases} x^i \, A(x) \sum_{\mu=0}^{i} b_\mu \, x^{-\mu} & (0 \leq i < s) \\ x^i \, A(x) \, B(1/x) & (s \leq i \leq N - r) \\ x^i \, B(1/x) \sum_{\nu=0}^{N-i} a_\nu \, x^\nu & (N - r < i \leq N) \ . \end{cases} \tag{1.1}$$

where

$$A(x) = \sum_{\nu=0}^{r} a_\nu \, x^\nu \ , \qquad B(x) = \sum_{\mu=0}^{s} b_\mu \, x^\mu$$

are polynomials with real or complex coefficients (according as H is real or complex) and $a_0 \, b_0 \neq 0$.

Though the form (1.1) in its full generality previously appeared in a joint paper [5], it was first suggested by Trench, and, for a particular case,

---

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

had been published by him in 1967 [8]. It is therefore appropriately called by his name.

By Lemma 3 of [5] a Trench matrix is nonsingular if and only if $A(x)$ and $x^s B(1/x)$ have no common zero. (Both real and complex zeros must be taken into account even if H is real.) In fact, it is shown in [5] that a strictly banded matrix has a Toeplitz inverse if and only if it is a nonsingular Trench matrix.

It is also shown in [5] that a Trench matrix is persymmetric: that is, symmetric about its secondary diagonal, and is also quasi-Toeplitz. The latter term implies that it has the Toeplitz property

$$h_{i+1,j+1} = h_{ij}$$

so long as neither of these elements is in the s by r submatrix in the upper left corner or the r by s submatrix in the lower right corner.

It is the purpose of this paper to establish certain bounds for the eigenvalues of Hermitian Trench matrices. More specifically, let the polynimials $A(x)$ and $B(x)$ be given, and consider the corresponding family of Trench matrices H_N given by (1.1) of all orders from $r + s + 1$ to ∞. We wish to establish bounds depending on $A(x)$ and $B(x)$, but independent of N, for the eigenvalues of H_N. As there is an extensive literature on bounds for eigenvalues of Toeplitz matrices (see, e.g., [2], [9]), it is tempting to think that in the nonsingular case one could deduce bounds for the Trench matrices from what is known about their Toeplitz inverses. However, it turns out that this would impose severe restrictions on the choice of the polynomials $A(x)$ and $B(x)$.

Consider the family of Toeplitz matrices T_N characterized by the doubly infinite sequence $\{t_\nu\}_{\nu=-\infty}^{\infty}$ so that $T_N = (t_{ij})_{i,j=0}^{N}$, where $t_{ij} = t_{j-i}$, and note that, while the Trench matrices are banded, their Toeplitz inverses are not, so that the entire sequence $\{t_\nu\}$ is involved. The available theorems regarding bounds for eigenvalues of such families of Toeplitz matrices require that the Laurent series

$$\sum_{\nu=-\infty}^{\infty} t_\nu x^\nu \tag{1.2}$$

converge in some fashion in an appropriate region of the complex plane. The convergence may be weak (see, e.g., [9]), but we wish to extend our consideration to cases in which the series (1.2) does not exist or its convergence fails entirely.

Now, for the family of Toeplitz matrices whose inverses belong to the given family of Trench matrices, it was shown in [4] that (1.2) converges in some part of the plane if and only if all zeros of $x^s B(1/x)$ are smaller in absolute value than all zeros of $A(x)$. The case in which this condition is fulfilled is an important one, as we shall see in Theorem 1 and its proof, but by no means do we wish to limit our consideration to that case. Moreover, we shall find it expedient to take full advantage of the very special structure of Trench matrices by working directly with them rather than with their inverses.

We do, however, confine our attention to Hermitian Trench matrices. While the case of greatest practical interest is that of a real symmetric matrix, our results have been extended to Hermitian complex Trench matrices, as this was easily accomplished. It is hoped that someone will pursue a similar investigation for the more difficult case of non-Hermitian Trench matrices. In this connection, some fragmentary results are available. Consider, for example, a real tridiagonal Trench matrix H; this implies that both $A(x)$ and $B(x)$ are linear. If the single zero of $A(x)$ and that of $B(x)$ have the same sign, it is easily shown that H is similar to a real symmetric (tridiagonal) matrix. Thus the eigenvalues are real, and the results of this paper apply to the transformed matrix.

2. The Main Results.

It is easily seen that the Trench matrix H given by (1.1) is Hermitian if and only if $r = s$ and

$$a_\nu b_\mu = \bar{a}_\mu \bar{b}_\nu \qquad (0 \le \mu,\nu \le r) \quad . \tag{2.1}$$

If we define

$$A^*(x) = \sum_{\nu=0}^{r} \bar{a}_\nu x^\nu = \overline{A(\bar{x})} \quad , \tag{2.2}$$

then (2.1) is easily seen to be equivalent to the condition

$$B(x) = cA^*(x) \quad ,$$

where c is a nonzero real constant. If c is negative, we can consider the matrix $-H$, whose eigenvalues are, of course, the negatives of those of

H. If c is positive, $A(x)$ and $B(x)$ can be normalized so that $B(x) = A^*(x)$. Thus there is no loss of generality if we limit consideration to Hermitian Trench matrices with $B(x) = A^*(x)$.

The function

$$h(x) = A(x)\, B(1/x) = A(x)\, A^*(1/x) = \sum_{\nu=-r}^{r} h_\nu\, x^\nu \tag{2.3}$$

will play an important role in this paper, as it did in [4]. Consider the values of this function on the unit circle. If $x = e^{i\theta}$, it follows from (2.2) and from the fact that for this x, $x^{-1} = \overline{x}$, that

$$h(x) = A(x)\, \overline{A(x)} = |A(x)|^2 \qquad (x = e^{i\theta}) \tag{2.4}$$

Therefore $h(x)$ is real and nonnegative on the unit circle, and moreover $h(x) = \phi(\theta)$ is a continuous periodic function of the real variable θ with period 2π. Hence it has a maximum and a minimum value, which we denote by M and m, respectively.

The following two theorems are the main results of this paper. Theorem 1 deals with the "regular" case in which (1.2) converges and Szegö's theorem applies; Theorem 2 asserts a weaker conclusion in a more general context.

Theorem 1. Let H_N be the Hermitian Trench matrix of order $N + 1 > 2r + 1$ characterized by the polynomials $A(x)$ and $B(x) = A^*(x)$ of degree $r > 0$. Then H_N is positive definite if and only if all the zeros of $A(x)$ are outside the unit circle. It is positive semidefinite if and only if all the zeros of $A(x)$ that are not also zeros of $A^*(1/x)$ are outside the unit circle. If it is positive definite, all its eigenvalues are greater than m and less than M, and if ρ_1 and ρ_{N+1} denote the smallest and largest eigenvalue, respectively,

$$\lim_{N\to\infty} \rho_1 = m\,, \qquad \lim_{N\to\infty} \rho_{N+1} = M\,.$$

Theorem 2. Let H_N denote the Hermitian Trench matrix of order $N + 1$ described in Theorem 1, and let σ_N denote its spectral radius. Then $\sigma_N < M$ for all N, and

$$\lim_{N\to\infty} \sigma_N = M\,.$$

3. Some Implications of the Theorems.

Before proceeding to the proofs of the theorems, we shall briefly

discuss a few of their implications. In some applications (see, e.g., [3]) we are interested in matrices of the form

$$G = I - kH \quad , \tag{3.1}$$

where H is a Hermitian Trench matrix and k is a positive constant. In particular we would like to know if the limit

$$G^{\infty} = \lim_{n \to \infty} G^{n} \tag{3.2}$$

exists. We note that Oldenburger [6] and Dresden [1] have shown that, for any square matrix G, G^{∞} exists if and only if either all the eigenvalues of G are inside the unit circle, or else $+1$ is a simple zero of the minimum polynomial of G and all other zeros are inside the unit circle. The following corollary (first conjectured by Trench) is a consequence of Theorems 1 and 2.

Corollary 1. Let G be given by (3.1), where H is the Hermitian Trench matrix described in Theorem 1. Then the limit (3.2) exists for all N if and only if

$$k \leq 2/M \tag{3.3}$$

and no zero of $A(x)$ is inside the unit circle unless it is also a zero of $A^*(1/x)$.

Proof. Let (3.3) and the condition on the zeros of $A(x)$ be satisfied. Then, H is positive semidefinite by Theorem 1, since any zero of $A(x)$ on the unit circle is a zero of $A^*(1/x)$, and therefore its eigenvalues are nonnegative. By Theorem 2 the eigenvalues of H are less than M. Since the eigenvalues of G are obtained by subtracting from unity k times those of H, the former are greater than $1 - kM$ and not greater than 1. In fact, if H is singular, 1 is an eigenvalue of G. Since H (and therefore G) is Hermitian, all zeros of the minimum polynomial are simple, and 1 is at most a simple zero. Since $k \leq 2/M$, $1 - kM \geq -1$ and so the eigenvalues of G are greater than -1. Thus, the condition of Oldenburger and Dresden is satisfied and G^{∞} exists.

On the other hand, if a zero of $A(x)$ that is not a zero of $A^*(1/x)$ is inside the unit circle, by Theorem 1, H has a negative eigenvalue. Since k is positive, this implies that G has an eigenvalue greater than 1, and so G^{∞} does not exist. Alternatively, if $A(x)$ has no zero inside the unit circle, but $k > 2/M$, then, for sufficiently large N, G has a

negative eigenvalue arbitrarily close to $1 - kM < -1$. Thus G^{∞} fails to exist for some N.

4. Proofs of the Theorems.

In these proofs we shall employ a certain special matrix notation. Let

$$P(x) = \sum_{\nu=0}^{d} p_{\nu} x^{\nu}$$

be a given polynomial. Then we define the matrix

$$P_{m,n} = (p_{ij})_{i=1}^{m}{}_{j=1}^{n} ,$$

where

$$p_{ij} = p_{j-i} ,$$

and it is understood that $p_{\nu} = 0$ for $\nu < 0$ and for $\nu > d$.

We shall also need to use the special matrix J_N, which is defined as the square matrix of order N having 1's on its secondary diagonal and 0's elsewhere. Note that multiplying an m by n matrix on the left by J_m reverses the order of the rows, and multiplying it on the right by J_n reverses the order of the columns. Of course, $J_N^2 = I_N$. For convenience we shall often omit the subscript of J when the context makes this clear. A persymmetric matrix Q is characterized by the fact that

$$JQJ = Q^T .$$

In the proof of Theorem 1 we shall find the case of a singular Trench matrix to be more difficult than the nonsingular case, and we shall need a lemma that expresses a singular Trench matrix as the product of a nonsingular Trench matrix and two rectangular matrices. Because singular Trench matrices appear to be interesting in their own right, the lemma is stated with more generality (i.e., without the restriction to Hermitian matrices) than is required for the purposes of this paper.

Lemma 1. Suppose the polynomial

$$E(x) = \sum_{\nu=0}^{q} e_{\nu} x^{\nu}$$

divides both $A(x)$ and $x^s B(1/x)$ and define

$$E^{\#}(x) = x^q E(1/x) ,$$

$$\hat{A}(x) = A(x)/E(x) = \sum_{\nu=0}^{r-q} \hat{a}_\nu x^\nu \tag{4.1}$$

and

$$\hat{B}(x) = B(x)/E^{\#}(x) = \sum_{\mu=0}^{s-q} \hat{b}_\mu x^\mu \quad . \tag{4.2}$$

Let H be the Trench matrix of order $N + 1$ characterized by $A(x)$ and $B(x)$ as in (1.1), and let D be the Trench matrix of order $N - q + 1$ characterized by $\hat{A}(x)$ and $\hat{B}(x)$; thus, the generating function of the elements of the ith row of D is

$$D_i(x) = \begin{cases} x^i \hat{A}(x) \sum_{\mu=0}^{i} \hat{b}_\mu x^{-\mu} & (0 \le i < s - q) \\ x^i \hat{A}(x) \hat{B}(1/x) & (s - q \le i \le N - r) \\ x^i \hat{B}(1/x) \sum_{\nu=0}^{N-i} \hat{a}_\nu x^\nu & (N - r < i \le N - q). \end{cases} \tag{4.3}$$

Then,

$$H = E^{\#T}_{N-q+1,N+1} \, DE_{N-q+1,N+1} \quad . \tag{4.4}$$

Proof. For convenience let us drop the subscripts of the rectangular matrices in (4.4). It follows from (4.3) and from the structure of E that the generating function of the elements of the ith row of DE is $D_i(x) E(x)$. (Note that D is of order $N - q + 1$, and that the "special rows" at the bottom of D are $r - q$ in number, and $(N - q + 1) - (r - q) = N - r + 1$.)

With the understanding that $b_\mu = 0$ for $\mu > s$, the first two parts of (1.1) can both be written in the form

$$H_i(x) = A(x) \sum_{\mu=0}^{i} b_\mu x^{i-\mu} = A(x) \sum_{\mu=0}^{i} b_{i-\mu} x^\mu \quad . \tag{4.5}$$

Thus, by (4.1) and (4.2) we have

$$D_i(x) = \hat{A}(x) \sum_{\mu=0}^{i} \hat{b}_{i-\mu} x^\mu \quad . \tag{4.6}$$

Therefore, the generating function of the elements of the ith row of $E^{\#T} DE$ is

$$\sum_{k=0}^{i} e_{q-i+k}\, D_k(x)\, E(x) = A(x) \sum_{k=0}^{i} e_{q-i+k} \sum_{\mu=0}^{k} \hat{b}_{k-\mu}\, x^{\mu}$$

by (4.1) and (4.6). Reversing the order of summation gives

$$A(x) \sum_{\mu=0}^{i} x^{\mu} \sum_{k=\mu}^{i} e_{q-i+k}\, \hat{b}_{k-\mu} \quad , \tag{4.7}$$

and the summation with respect to k can be rewritten as

$$\sum_{\nu=0}^{i-\mu} e_{q-\nu}\, \hat{b}_{i-\mu-\nu} = b_{i-\mu}$$

by (4.2). Thus (4.7) reduces finally to

$$A(x) \sum_{\mu=0}^{i} b_{i-\mu}\, x^{\mu} = H_i(x)$$

by (4.5). This proves (4.4) for rows 0 to N - r, inclusive, of H.

Let us now consider the matrix JHJ, in which the order of both rows and columns of H is reversed. By means of (1.1) it is not difficult to see that this is a Trench matrix in which, as compared with H, the roles of A(x) and B(x) are interchanged. Therefore by the first part of this proof, the equation

$$JHJ = E^{T}(JDJ)E^{\#} \tag{4.8}$$

holds for rows 0 to N - s, inclusive, of the matrices on both sides. Now, it is easily verified that $JE^{T}J = E^{\#T}$ and $JE^{\#}J = E$. Thus, multiplying (4.8) by J both on the left and on the right gives (4.4). As rows 0 to N - s of JHJ become rows s to N of H (with the order of the elements reversed), this completes the proof of the lemma.

Proof of Theorem 1. This proof consists of three parts. First, we shall use Szegö's theorem to show that if all the zeros of A(x) are outside the unit circle, then H is positive definite, and the inequalities and limiting relations for the eigenvalues follow. Second, we shall prove that if A(x) has one or more zeros on or inside the unit circle that are also zeros of $x^{r} A^{*}(1/x)$ (but all other zeros are outside the unit circle), then H is positive semidefinite. Finally, we shall show that if A(x) has a zero inside the unit circle that is not a zero of $x^{r} A^{*}(1/x)$ as well, H is not positive definite or semidefinite.

Let all the zeros of $A(x)$ be outside the unit circle. Then the zeros of $x^s B(1/x) = x^r A^*(1/x)$ are all inside the unit circle, and it was shown in [4] that $[h(x)]^{-1} = [A(x)\, A^*(1/x)]^{-1}$ has a Laurent expansion (1.2) that converges in an annular region containing the unit circle. It follows from the discussion preceding Theorem 1 that $[h(x)]^{-1}$ is real and positive on the unit circle, its maximum and minimum values there being $1/m$ and $1/M$, respectively. Therefore, by Szegö's theorem (see Chapter 5 of [2]) the eigenvalues of $T_N = H_N^{-1}$ are greater than $1/M$ and less than $1/m$ for all N, and these bounds are the limits of the smallest and the largest eigenvalues as N goes to infinity. As the eigenvalues of H_N are the reciprocals of those of T_N, the statements in Theorem 1 concerning the positive definite case follow at once.

In order to deal with the case in which H is singular, we specialize the formula (4.4) established in Lemma 1. We recall that the zeros of $B(1/x) = A^*(1/x)$ are the conjugates of the reciprocals of those of $A(x)$. Let all the zeros of $A(x)$ that are not also zeros of $A^*(1/x)$ be outside the unit circle. In fact, since the conjugate of a point on the unit circle is also its reciprocal, any zero of $A(x)$ that is on the unit circle is also a zero of $A^*(1/x)$. Therefore, let $A(x) = \hat{A}(x)\, E(x)$, where the zeros of $\hat{A}(x)$ are those of $A(x)$ that are outside the unit circle, and the zeros of $E(x)$ are those of $A(x)$ that are also zeros of $A^*(1/x)$. It follows that $E^*(x)$ and $E^{\#}(x)$ have the same zeros, and are therefore identical. Since $E^{\#}(x)$ is obtained from $E(x)$ by reversing the order of the coefficients, and $E^*(x)$ by taking the conjugates of the coefficients, we must have

$$e_{q-\nu} = \bar{e}_\nu \qquad (\nu = 0, 1, \ldots, q) \ ;$$

It follows that $E^{\#T} = E^{CT}$, and (4.4) becomes

$$H_N = E^{CT}_{N-q+1,N+1}\, D E_{N-q+1,N+1} \ . \tag{4.9}$$

If u is an arbitrary nonzero vector, and $v = Eu$, then by (4.9), $u^{CT} H_N u = v^{CT} D v$, which is nonnegative, since D is positive definite. Therefore, H is positive semidefinite.

We come finally to the third part of the proof. Let $A(x)$ have a zero, $x = \xi$, inside the unit circle such that $A^*(\xi^{-1}) \neq 0$. Since H is a Trench matrix, $a_0 \neq 0$, and so $\xi \neq 0$. It follows that $\bar{\xi}^{-1}$ (which is of course outside the unit circle) is a zero of $A^*(1/x)$. Now let v be the vector whose ith component (starting the numbering with 0) is $\bar{\xi}^{-i}$. It follows from the definition of the generating function that the ith component of Hv is $H_i(\bar{\xi}^{-1})$. For all but the first r components (i.e., those numbered from 0 to $r-1$), $A^*(\bar{\xi}) = 0$ is a factor of $H_i(\bar{\xi}^{-1})$, and so these components vanish. For $0 \leq i < r$,

$$H_i(\bar{\xi}^{-1}) = A(\bar{\xi}^{-1}) \sum_{\mu=0}^{i} \bar{a}_\mu \bar{\xi}^{\mu-i} \quad . \tag{4.10}$$

Now, let the polynomial

$$F(x) = \sum_{\nu=0}^{r-1} f_\nu x^\nu$$

be defined by

$$A(x) = (x - \xi)F(x) = -\xi(1 - x\xi^{-1})F(x) \quad . \tag{4.11}$$

Then,

$$F(x) = -\xi^{-1}(1 - x\xi^{-1})^{-1} A(x) \quad ,$$

and consequently,

$$f_j = -\sum_{\nu=0}^{j} a_\nu \xi^{\nu-j-1} = -\xi^{-j-1} \sum_{\nu=0}^{j} a_\nu \xi^\nu \qquad (0 \leq j < r) \quad ,$$

or

$$\sum_{\nu=0}^{j} a_\nu \xi^{\nu-j} = -\xi f_j \quad . \tag{4.12}$$

Substitution of (4.11) and the conjugate of (4.12) in (4.10) gives

$$H_i(\bar{\xi}^{-1}) = -\bar{\xi}(\bar{\xi}^{-1} - \xi)F(\bar{\xi}^{-1})\bar{f}_i = (\xi\bar{\xi} - 1)F(\bar{\xi}^{-1})\bar{f}_i \quad ,$$

and so

$$v^{CT} Hv = \sum_{i=0}^{r-1} \xi^{-i} H_i(\bar{\xi}^{-1}) = (\xi\bar{\xi} - 1)F(\bar{\xi}^{-1}) \overline{F(\bar{\xi}^{-1})} \quad ,$$

an expression which is clearly negative, since $|\xi| < 1$ and

$$0 \neq \overline{A^*(\bar{\xi}^{-1})} = A(\bar{\xi}^{-1}) = (\bar{\xi}^{-1} - \xi)F(\bar{\xi}^{-1}) \quad ,$$

so that $F(\bar{\xi}^{-1}) \neq 0$. It follows that H is not positive definite or semi-definite. This completes the proof of Theorem 1.

In the proof of Theorem 2 we shall need a lemma that expresses the Hermitian Trench matrix H in terms of simpler matrices. Let us define $\tilde{A}(x)$ by

$$\tilde{A}(x) = x^r A^*(1/x) \quad ,$$

and let us define $\tilde{A}$ as the square matrix of order $N + 1$

$$\tilde{A} = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{A}_{r,r} \end{bmatrix} \quad . \tag{4.13}$$

Then we have

Lemma 2. If H is the Hermitian Trench matrix defined in Theorem 1,

$$H = A^{CT}_{N+1,N+1} A_{N+1,N+1} - \tilde{A}^{CT} \tilde{A} \quad . \tag{4.14}$$

Proof. First we note that H and the first term of the right member of (4.14) agree in all their elements except the square submatrices of order r in the lower right corner. For all but the last r rows, this follows easily from (1.1) taking $B(x) = A^*(x)$. For the last r rows, excluding the square submatrix in the right corner, it follows from the Hermitian symmetry of both matrices. Moreover, the second term of the right member of (4.14) has zeros everywhere except in the corner submatrix mentioned. These observations make (4.14) at least plausible, and permit us to limit our attention to the r by r submatrices in the lower right corner.

In the case of the first term of the right member of (4.14), this corner submatrix is obtained by multiplying the last r rows of the first factor by the last r columns of the second factor. Taking into account that some of the elements of these rows and columns are zeros, this product can be written as

$$\tilde{A}_{r,2r} \tilde{A}^{CT}_{r,2r} \quad .$$

Moreover, by partitioning the first factor of this latter product into the first r columns and the last r columns, and the second factor similarly by rows, we obtain

$$\tilde{A}_{r,2r} \tilde{A}^{CT}_{r,2r} = \tilde{A}_{r,r} \tilde{A}^{CT}_{r,r} + A^{CT}_{r,r} A_{r,r} \quad ,$$

or

$$A_{r,r}^{CT} A_{r,r} = \tilde{A}_{r,2r} \tilde{A}_{r,2r}^{CT} - \tilde{A}_{r,r} \tilde{A}_{r,r}^{CT} \quad . \tag{4.15}$$

Now, the left member of (4.15) is precisely the square submatrix of order r in the upper left corner of H. Since H is Hermitian and persymmetric, the one in the lower right corner is obtained from it by reversing the order of both rows and columns and then taking the conjugate. Accordingly, let us perform these operations on the right member. As the first term is Hermitian and Toeplitz, the effect of the operations is to leave that term unchanged. Coming now to the second term, since

$$J_r \tilde{A}_{r,r} \tilde{A}_{r,r}^{CT} J_r = J_r \tilde{A}_{r,r} J_r J_r \tilde{A}_{r,r}^{CT} J_r \quad ,$$

we can perform the operations on each factor separately. We note also that the effect of the three operations on a matrix $P_{r,r}$ is to take the conjugate transpose. Thus the result is $\tilde{A}_{r,r}^{CT} \tilde{A}_{r,r}$. In view of (4.13), this proves (4.14).

Proof of Theorem 2. Let us denote by K and L the respective products in the right member of (4.14), so that

$$H = K - L \quad .$$

Clearly K is Hermitian positive definite and L is Hermitian positive semidefinite. Let v be an arbitrary nonzero vector of complex elements. Then the Rayleigh quotients satisfy

$$\frac{v^{CT} Hv}{v^{CT} v} = \frac{v^{CT} Kv}{v^{CT} v} - \frac{v^{CT} Lv}{v^{CT} v} \leq \frac{v^{CT} Kv}{v^{CT} v} \quad . \tag{4.16}$$

Let

$$V(t) = \sum_{\nu=0}^{N} v_\nu e^{i\nu t}$$

be the characteristic function of v. Then,

$$A(e^{-it}) V(t) = \sum_{\nu=-r}^{N} w_\nu e^{i\nu t} \quad ,$$

where, for $0 \leq \nu \leq N$, w_ν is the νth component of $A_{N+1,N+1} v$. (It may be helpful to the reader to think of the vector v as being extended by annexing a number of zeros at the bottom.) By Parseval's formula (see [7], p. 699)

$$v^{CT}\, v = |v|^2 = \frac{1}{2\pi}\int_0^{2\pi} |V(t)|^2\, dt \quad ,$$

while

$$v^{CT}\, Kv = |A_{N+1,N+1}\, v|^2 \leq \frac{1}{2\pi}\int_0^{2\pi} |A(e^{-it})\, V(t)|^2\, dt \quad .$$

But by (2.4), $|A(e^{-it})|^2 = h(e^{-it})$, and therefore

$$v^{CT}\, Kv \leq \frac{1}{2\pi}\int_0^{2\pi} h(e^{-it})\,|V(t)|^2\, dt < M\; v^{CT}\, v \quad . \tag{4.17}$$

Note that since the zeros of $V(t)$ are a set of measure zero, the second inequality of (4.17) could be replaced by equality only if $h(x)$ is a constant function, which would imply that H is diagonal, and therefore a scalar matrix. This is tantamount to saying that $r = 0$, contrary to hypothesis (see description of H in Theorem 1).

It follows from (4.16) and (4.17) that

$$\frac{v^{CT}\, Hv}{v^{CT}\, v} < M \quad , \tag{4.18}$$

and, since the greatest eigenvalue of H is the maximum value of the Rayleigh quotient in the left member of (4.18), we have shown that the greatest eigenvalue of H is less than M.

However, it will be noted that in Theorem 2 we have not imposed the condition that would make H positive definite or semidefinite. Thus H may have negative eigenvalues, and it is conceivable that such a negative eigenvalue might exceed M in absolute value. We must prove that this is not the case. The algebraically smallest eigenvalue is the minimum value of the Rayleigh quotient in the left member of (4.18). Since K and L are both Hermitian positive semidefinite, this minimum value is greater than, or at least equal to the negative of the maximum value of the Rayleigh quotient with respect to L.

It follows from (4.13) and (4.14) that the elements of L are all zero with the exception of the square submatrix of order r in the lower right corner. Because of this structure, the eigenvalues of L, other than zero, are those of $\hat{L} = \tilde{A}_{r,r}^{CT}\, \tilde{A}_{r,r}$. Therefore the maximum Rayleigh quotient with respect to $\hat{L}$ is the same as that with respect to L. Therefore, a lower

bound to the eigenvalues of H is $-\hat{\rho}$, where $\hat{\rho}$ is the largest eigenvalue of $\hat{L}$. In order to complete the proof that the spectral radius of H is less than M, we must show that $\hat{\rho} < M$. In fact, if v is an arbitrary vector of r complex components and its characteristic function is

$$V(t) = \sum_{\nu=0}^{r-1} v_\nu e^{i\nu t} ,$$

then, by reasoning closely parallel to that used in the first part of this proof, we conclude that

$$v^{CT} \hat{L} v \leq \frac{1}{2\pi} \int_0^{2\pi} \left| A^*(e^{-it}) V(t) \right|^2 dt < M v^{CT} v .$$

Since $\hat{L}$ is Hermitian, $\hat{\rho}$ is the maximum value of the Rayleigh quotient.

To prove the second part we let M' be an arbitrary positive constant less than M and show that, for a suitable vector v and for sufficiently large N, the Rayleigh quotient $v^{CT} H_N v / v^{CT} v$ can be made larger than M'. Since $h(e^{it})$ is a continuous function of t and M is its maximum value in $[0, 2\pi]$, there is some value $t = \tau$, such that

$$h(e^{i\tau}) > M' .$$

Let us choose $v = [v_0, v_1, \ldots, v_N]^T$ so that $v_\nu = e^{-i\nu\tau}$ for $0 \leq \nu \leq N$. Then, except for the first r and the last r components, the νth component of Hv is $h(e^{i\tau}) v_\nu$. Therefore

$$v^{CT} Hv = (N - 2r + 1) h(e^{i\tau}) + C , \tag{4.19}$$

where C is the contribution of the first r and the last r components. Since every component of v has absolute value 1, an upper bound to the absolute value of C is the sum of the absolute values of the elements in the first r and the last r rows of H. Call this C', and note that C' does not depend on N.

Now choose N sufficiently large so that

$$N + 1 > \frac{C' + 2r\, h(e^{i\tau})}{h(e^{i\tau}) - M'} .$$

Then

$$(N + 1)[h(e^{i\tau}) - M'] > C' + 2r\, h(e^{i\tau}) ,$$

or

$$(N - 2r + 1)h(e^{i\tau}) > (N + 1)M' + |C| \quad ,$$

and consequently

$$(N - 2r + 1)h(e^{i\tau}) + C > (N + 1)M' \quad . \tag{4.20}$$

Since $|v_\nu|^2 = 1$ for every ν,

$$v^{CT} v = N + 1$$

and therefore by (4.19) and (4.20)

$$\frac{v^{CT} Hv}{v^{CT} v} > M'$$

as required. This completes the proof of Theorem 2.

Acknowledgements. My indebtedness to W. F. Trench is obvious. He has also kindly read a draft of this paper and made some suggestions that resulted in great improvements. My thanks are also due to M. V. Subbarao, who graciously arranged for me to visit the University of Alberta under his research grant from Natural Sciences and Engineering Research Council Canada. During that visit an important part of the research described herein was completed, and this paper has benefited significantly from discussions with him.

REFERENCES

[1] A. Dresden, On the iteration of linear homogeneous transformations, Bull. Amer. Math. Soc., 48 (1942), 577-579.

[2] U. Grenander and G. Szegö, Toeplitz Forms and Their Applications, U. of Calif. Press, Berkeley, 1958.

[3] T. N. E. Greville, Moving-weighted-average smoothing extended to the extremities of the data, MRC Technical Summary Report #1786, Mathematics Research Center, University of Wisconsin-Madison, August 1977.

[4] T. N. E. Greville, On a problem concerning band matrices with Toeplitz inverses, to appear in Proc. 8th Manitoba Conf. Numer. Math. Comput., Utilitas Mathematica Publishing Inc., Winnipeg, 1979.

[5] T. N. E. Greville and W. F. Trench, Band matrices with Toeplitz inverses, to appear in Linear Algebra and Appl.

[6] R. Oldenburger, Infinite powers of matrices and characteristic roots, Duke Math. J., 6 (1940), 357-361.

[7] K. Rektorys (ed.), Survey of Applicable Mathematics, MIT Press, Cambridge, Mass., 1969.

[8] W. F. Trench, Weighting coefficients for the prediction of stationary time series from the finite past, SIAM J. Appl. Math., 15 (1967), 1502-1510.

[9] H. Widom, Toeplitz matrices, pp. 179-209 of MAA Studies in Mathematics, Vol. 3, Studies in Real and Complex Analysis, Prentice-Hall, Englewood Cliffs, 1965.

EFFICIENT ALGORITHMS FOR CONTINUOUS PIECEWISE LINEAR APPROXIMANTS WITH VARIABLE KNOTS

Royce W. Soanes Jr.
US Army Armament Research and Development Command
Large Caliber Weapon Systems Laboratory
Benet Weapons Laboratory
Watervliet, NY 12189

ABSTRACT. Algorithms are derived for selecting and moving the knots of a continuous piecewise linear approximant. Knots are selected and moved as a subset of n equally spaced mesh points. As knots are selected, neighboring knots are moved for the purpose of more closely approximating their optimal values. A least squares fit criterion is used and the movement of each knot improves the global error sum of squares. If all the interior knots are moved during an iteration, only O(n) arithmetic operations of any kind are consumed. A simple example of the reduction in computational complexity obtainable with the methods herein discussed is the least squares fitting of a straight line to an arbitrarily large set of data using only five multiplications or divisions.

I. INTRODUCTION. The purpose of this article is to derive efficient and reliable algorithms for attacking the problem of continuous piecewise linear approximation of relatively large amounts of data, consisting of perhaps thousands of points. Approximation is done in the least squares sense, where we strive to make the global error sum of squares (SSE) small. We are not, however, concerned with obtaining a precisely optimal solution, as this would be prohibitively expensive for the large amounts of data involved [1]. The continuous piecewise linear approximants (linear splines) used will be defined over a variable knot mesh with respect to which we will nearly minimize global SSE.

Authors of available algorithms [1,3,5,6] seem reluctant to deal with the large amounts of data which are common as a result of analog to digital conversion. Indeed, they generally consider no more than a few dozen data points and only a handful of knots, neither do they consider data with an appreciable noise level. This reluctance is evidently due to the large computation times that would result in the large sample case or to the unreliability of analytic methods of seeking global optima in the presence of noise. On the other hand, elementary linear smoothing methods are efficient, but they either suffer from Gibb's phenomenon in the case of oscillatory kernel smoothing or are prone to cut off corners in the case of positive kernel smoothing [2].

The algorithms presented here are conceptually simple and therefore not difficult for nonspecialists in this area to implement. The data mesh is uniform as is invariably the case with data converted from analog to digital form. The nonuniform knot mesh is a subset of the uniform data mesh. We do not select knots from a continuum because the unlimited number of analog digital points that can be generated makes continuous variation of the knots unnecessary. Also, the algorithms derived here make exhaustive local searches for global improvements of SSE quite efficient.

In all that follows, let slow operation denote multiplication or division and let fast operation denote addition or subtraction. Slow operation and fast operation will be further abbreviated to SO and FO respectively. When we subsequently refer to $n \pm c$ operations of one kind or another, where c is a small integer constant, we will write n for simplicity. In general, we will refer only to the high order term in computation time expressions i.e. $5n^2 + 2n + 3$ will be "about $5n^2$".

II. A SIMPLE EXAMPLE OF COMPLEXITY REDUCTION. Shamos [4] describes the ordinary algorithm for fitting a straight line as requiring O(n) time. He makes no distinction between SO's and FO's. Since the usual formulas are O(n) for both addition and multiplication, and the O(n) additions cannot be avoided, he remains correct. The O(n) SO's involved in fitting a straight line may be reduced to five, however, as the algorithm following shows; multiplications by small integer constants are not counted here since these may be done faster through addition.

$$S_0 = T_0 = 0$$

$$S_i = S_{i-1} + Y_i \qquad 1 \leq i \leq n$$

$$T_i = T_{i-1} + S_i$$

$$c = 2[S_n + (S_n - 2T_n)/n]/(n + 1)$$

$$v_1 = [2(T_n - S_n)/n - c(n - 2)]/(n - 1)$$

$v_2 = v_1 + 3c$ where v_1 and v_2 are the values of the approximant at the left and right extremes of the equally spaced data. This algorithm is derived using the well known formulas for sums of first and second powers and summation by parts. Summation by parts, for instance, enables

one to reduce the sum

$$\sum_{i=1}^{n} iY_i$$

involving n SO's and n FO's to the expression

$$(n+1)\sum_{i=1}^{n} Y_i - \sum_{i=1}^{n}\sum_{j=1}^{i} Y_j$$

involving only one SO and 2n FO's.

For the sake of brevity and to emphasize the central role played by summation by parts, the algorithms derived here are called SBP algorithms. It is perhaps not particularly important to derive more efficient formulas for fitting a straight line, but it is important to reduce O(n) SO's to a constant number of SO's because the same technique will make it possible to reduce $O(n^2)$ SO's + $O(n^2)$ FO's to O(n) SO's + O(n) FO's in the next algorithm.

III. DESCRIPTION OF FUNDAMENTAL ALGORITHMS SBP (1) and SBP (2). Consider a linear spline consisting of two segments i.e. two fixed end knots and a single variable internal knot between them. There is only one way to find the globally optimal internal knot: for every internal data mesh point, set up the three normal equations, solve them for the three ordinates, and compute the SSE. Programming this algorithm naively (not using summation by parts) may be done in about $7n^2$ SO's and $6n^2$ FO's. On the other hand, the SBP algorithm for this case (SBP(1)), accomplishes exactly the same computation using only 15n SO's and about 35n FO's. The ratio of naive effort to SBP effort for SO's therefore goes approximately as n/2. The effort ratio can obviously be as large as 1000 to 1 for a data set consisting of only a couple of thousand points. The SBP(1) algorithm is used to pick out points of abrupt behavior in the data for the purpose of splitting subintervals and inserting new knots.

The other fundamental algorithm presented here is quite similar to the preceding one except that the end ordinates are held fixed instead of being free. There is therefore only one normal equation instead of three. The optimal internal knot for this case is obtained in 12n SO's. This SBP(2) algorithm assures improvement of global SSE when it is used to move the knots.

IV. THE GENERAL NORMAL EQUATIONS. Let (x_i,y_i) $1 \le i \le n$ be a large set of data with equally spaced x values and let (u_i,v_i) $1 \le i \le N$ be a small set of data whose linear spline approximates the (x,y) data in the least squares sense. The u's are a subset of the x's.

The linear spline approximant on the ith knot subinterval is defined as:

$$L_i(x) = (1 - r_i(x))v_i + r_i(x)v_{i+1}$$

where

$$r_i(x) = (x - u_i)/(u_{i+1} - u_i) .$$

Let M_i be the mesh index of the ith knot. The error sum of squares is given by:

$$SSE = \sum_{i=1}^{N-1} \sum_{m=M_i+1}^{M_{i+1}-1} (L_i(x_m) - y_m)^2 + \sum_{i=1}^{N} (v_i - y_{M_i})^2.$$

The following symbolic abbreviations will hold throughout the article: abbreviate $r_i(x_m)$ to r_i and $1 - r_i(x_m)$ to s_i.

The error sum of squares therefore becomes:

$$SSE = \sum_{i=1}^{N-1} \sum_{m=M_i+1}^{M_{i+1}-1} (s_i v_i + r_i v_{i+1} - y_m)^2 + \sum_{i=1}^{N} (v_i - y_{M_i})^2.$$

Setting the partial derivative of SSE with respect to v_j equal to zero, we have:

$$v_{j-1} \sum_{m=M_{j-1}+1}^{M_j-1} r_{j-1}s_{j-1} + v_j\left(1 + \sum_{m=M_{j-1}+1}^{M_j-1} r_{j-1}^2 + \sum_{m=M_j+1}^{M_{j+1}-1} s_j^2\right)$$

$$+ v_{j+1} \sum_{m=M_j+1}^{M_{j+1}-1} r_j s_j = y_{M_j} + \sum_{m=M_{j-1}+1}^{M_j-1} y_m r_{j-1} + \sum_{m=M_j+1}^{M_{j+1}-1} y_m s_j \qquad (4.1)$$

if v_{j-1}, v_j and v_{j+1} are defined.

If only v_j and v_{j+1} are defined, the result is:

$$v_j\left(1 + \sum_{m=M_j+1}^{M_{j+1}-1} s_j^2\right) + v_{j+1} \sum_{m=M_j+1}^{M_{j+1}-1} r_j s_j = y_{M_j} + \sum_{m=M_j+1}^{M_{j+1}-1} y_m s_j \qquad (4.2)$$

If only v_{j-1} and v_j are defined,

$$v_{j-1} \sum_{m=M_{j-1}+1}^{M_j-1} r_{j-1} s_{j-1} + v_j\left(1 + \sum_{m=M_{j-1}+1}^{M_j-1} r_{j-1}^2\right) =$$

$$y_{M_j} + \sum_{m=M_{j-1}+1}^{M_j-1} y_m r_{j-1} \qquad (4.3)$$

The following abbreviations regarding summation notation will be observed throughout the article.

Abbreviate $\sum_{m=M_{i-1}}^{M_i}$ to $\sum_{-1}$

$\sum_{m=M_i}^{M_{i+1}}$ to $\sum_{0}$

and $\sum_{m=M_{i-1}}^{M_{i+1}}$ to $\sum_{-10}$.

The normal equations therefore become:

$$v_{i-1}\sum_{-1} r_{i-1} s_{i-1} + v_i\left(-1 + \sum_{-1} r_{i-1}^2 + \sum_{0} s^2\right) + v_{i+1}\sum_{0} r_i s_i =$$

$$-y_{M_i} + \sum_{-1} y_m r_{i-1} + \sum_{0} y_m s_i \qquad (v_{i-1},\ v_i \text{ and } v_{i+1} \text{ defined}) \quad (4.4)$$

$$v_i \sum_{0} s_i^2 + v_{i+1} \sum_{0} r_i s_i = \sum_{0} y_m s_i \qquad (v_i \text{ and } v_{i+1} \text{ defined}) \quad (4.5)$$

$$v_{i-1} \sum_{-1} r_{i-1} s_{i-1} + v_i \sum_{-1} r_{i-1}^2 = \sum_{-1} y_m r_{i-1} \quad (v_{i-1} \text{ and } v_i \text{ defined}) \quad (4.6)$$

V. GENERATION AND SOLUTION OF SBP(1) NORMAL EQUATIONS. Considering only three local knots u_{i-1}, u_i and u_{i+1}, we may write down the normal equations for v_{i-1}, v_i and v_{i+1} using (4.6), (4.4), and (4.5) respectively.

$$v_{i-1} \sum_{-1} s_{i-1}^2 + v_i \sum_{-1} r_{i-1} s_{i-1} = \sum_{-1} y_m s_{i-1} \quad (5.1)$$

$$v_{i-1} \sum_{-1} r_{i-1} s_{i-1} + v_i \left(-1 + \sum_{-1} r_{i-1}^2 + \sum_{0} s_i^2\right) + v_{i+1} \sum_{0} r_i s_i =$$

$$-y_{M_i} + \sum_{-1} y_m r_{i-1} + \sum_{0} y_m s_i \quad (5.2)$$

$$v_i \sum_{0} r_i s_i + v_{i+1} \sum_{0} r_i^2 = \sum_{0} y_m r_i \quad (5.3)$$

Adding 5.1 and 5.3 to 5.2 we obtain

$$v_{i-1} \sum_{-1} s_{i-1} + v_i \left(-1 + \sum_{-1} r_{i-1} + \sum_{0} s_i\right) + v_{i+1} \sum_{0} r_i = \sum_{-10} y_m$$

since r+s =1.

The contents of the appendix should be reviewed at this point.

After substituting sums from the appendix into these normal equations and performing a couple of row operations, we may obtain the following simplified normal equations:

$$v_{i-1}(2n_{i-1}+1) + v_i(n_{i-1}-1) = 6(T_{-1}-S_{-1})/(n_{i-1}+1) \quad (5.4)$$

$$v_{i-1} + v_i(n_{i-1}+n_i+2) + v_{i+1} = -2S_{-10} - 6[(T_{-1}-S_{-1})/(n_{i-1}+1)$$

$$+ (T_{-1}-S_{-1}-T_{-10})/(n_i+1)] \quad (5.5)$$

$$v_i(n_i-1) + v_{i+1}(2n_i+1) = 6[S_{-10} + (T_{-1}-S_{-1}-T_{-10})/(n_i+1)] \quad (5.6)$$

Excluding multiplications by small integer constants, the augmented matrix for these equations may be computed in only two SO's.

Using Gaussian elimination on:

$$\begin{matrix} a_{11} & a_{12} & 0 & c_1 \\ 1 & a_{22} & 1 & c_2 \\ 0 & a_{32} & a_{33} & c_3 \end{matrix}$$

we have:

$$a_{22} \leftarrow a_{22} - a_{12}/a_{11}$$

$$c_2 \leftarrow c_2 - c_1/a_{11}$$

$$q \leftarrow a_{32}/a_{22}$$

$$a_{33} \leftarrow a_{33} - q$$

$$c_3 \leftarrow c_3 - c_2 q$$

$$v_{i+1} \leftarrow c_3/a_{33}$$

$$v_i \leftarrow (c_2 - v_{i+1})/a_{22}$$

$$v_{i-1} \leftarrow (c_1 - a_{12}v_i)/a_{11}$$

Hence, v_{i-1}, v_i and v_{i+1} may be computed in eight SO's.

VI. ERROR SUM OF SQUARES FOR SBP(1). The SSE for SBP(1) is given by:

$$SSE = \sum_{m=M_{i-1}}^{M_i-1} (s_{i-1}v_{i-1}+r_{i-1}v_i-y_m)^2 + \sum_{m=M_i+1}^{M_{i+1}} (s_iv_i+r_iv_{i+1}-y_m)^2 + (v_i-y_{M_i})^2$$

$$= \sum_{-1} (s_{i-1}v_{i-1}+r_{i-1}v_i-y_m)^2 + \sum_{0} (s_iv_i+r_iv_{i+1}-y_m)^2 - (v_i-y_{M_i})^2 \quad (6.1)$$

Expansion of these sums, making substitutions from the normal equations, and considerable fortuitous cancellation yields:

$$SSE = \sum_{-10} y_m^2 - v_{i-1} \sum_{-1} s_{i-1} y_m - v_i \left(\sum_{-1} r_{i-1} y_m + \sum_{0} s_i y_m - y_{M_i} \right)$$

$$- v_{i+1} \sum_{0} r_i y_m \qquad (6.2)$$

If we now substitute the required sums from the appendix into 6.2, we get the final expression for SSE:

$$SSE = \sum_{-10} y_m^2 + (v_i - v_{i-1})(T_{-1} - S_{-1})/n_{i-1} + (v_i - v_{i+1})(T_{-1} - S_{-1} + S_{-10} - T_{-10})$$

$$/n_i - v_{i+1} S_{-10} \qquad (6.3)$$

It is obvious that the sum of squares of the data values in 6.3 is a constant component of SSE for any u_i between u_{i-1} and u_{i+1}. We need not therefore actually compute this sum of squares as we exhaustively search for the u_i with the smallest SSE. The variable component of SSE may therefore be computed in only five SO's.

Given a u_i, we may therefore set up the SBP(1) normal equations, solve them and ultimately compute the variable component of SSE in only fifteen SO's. If n is the number of internal mesh points, we may find the best u_i in 15n SO's. It is also important to notice that S_{-10} and T_{-10} do not depend on the position of u_i and that although S_{-1} and T_{-1} do depend on the position of u_i, these sums need only be updated using two FO's per mesh point as we calculate SSE for each internal mesh point. The significance of this is that although the computation of SSE for one and only one u_i is O(n) for FO's, the computation of SSE for all the u_i's is still only O(n) for FO's. Hence the entire SBP(1) algorithm is O(n) for both SO's and FO's.

This is in sharp contrast to the $O(n^2)$ complexity for SO's and FO's which would have been the case had summation by parts not been exploited. This overall reduction of complexity from $O(n^2)$ to O(n) makes the SBP(1) algorithm a viable technique, especially in the context of large sets of data and in spite of the fact that it is an exhaustive, brute force search method.

VII. NORMAL EQUATION AND SSE FOR SBP(2). If the endpoints are fixed instead of free in the 3 knot case, there is only one normal equation (5.2):

$$v_{i-1} \sum_{-1} r_{i-1} s_{i-1} + v_i \left(-1 + \sum_{-1} r_{i-1}^2 + \sum_{0} s_i^2\right) + v_{i+1} \sum_{0} r_i s_i$$

$$= -y_{M_i} + \sum_{-1} y_m r_{i-1} + \sum_{0} y_m s_i$$

Substituting the expressions for the various sums from the appendix, we have:

$$v_{i-1}(n_{i-1}-1/n_{i-1}) + v_i(1/n_{i-1}+1/n_i+2(n_{i-1}+n_i)) + v_{i+1}(n_i-1/n_i)$$

$$= 6[(S_{-1}-T_{-1})/n_{i-1} - (S_{-10}-S_{-1}+T_{-1}-T_{-10})/n_i] \qquad (7.1)$$

The general expression for SSE is given by 6.1.

Expanding 6.1 and substituting 5.2 into it yields:

$$SSE = \sum_{-10} y_m^2 + v_{i-1}(v_{i-1}\sum_{-1} s_{i-1}^2 - 2\sum_{-1} s_{i-1}y_m)$$

$$+ v_{i+1}(v_{i+1}\sum_{0} r_i^2 - 2\sum_{0} r_i y_m) - v_i^2(-1+\sum_{-1} r_{i-1}^2 + \sum_{0} s_i^2) \qquad (7.2)$$

Using the sums from the appendix in 7.2 gives:

$$6SSE = 6\sum_{-10} y_m^2 + v_{i-1}[v_{i-1}(1/n_{i-1}+3+2n_{i-1}) - 12(T_{-1}-S_{-1})/n_{i-1}]$$

$$+ v_{i+1}[v_{i+1}(1/n_i+3+2n_i) - 12(S_{-10}+(S_{-10}+T_{-1}-S_{-1}-T_{-10})/n_i)]$$

$$- v_i^2(1/n_{i-1}+1/n_i+2(n_{i-1}+n_i)) \qquad (7.3)$$

Excluding multiplications by small integer constants, the variable component of SSE in the SBP(2) algorithm may be obtained in 12 SO's.

VIII. A STRATEGY FOR USING SBP(1) AND SBP(2). There are many ways one could employ SBP(1) and SBP(2) for knot selection and movement respectively. Based on experimental results, the following technique seems to be quite reliable - especially when there is considerable variation in noise level or ringing amplitude.

(1) Initialize the knot set to consist of three knots and corresponding ordinates using SBP(1).

(2) Find the knot subinterval which shows the most "promise" for knot insertion.

(3) Insert a knot in this "promising" interval using SBP(2).

(4) Use SBP(2) to move the knots to the left of the newly inserted knot and stop when the position of the moved knot doesn't change significantly (relative to the sum of its left and right knot subinterval lengths).

(5) Use SBP(2) to move the knots to the right of the new knot in the same manner.

(6) Move the new knot once.

(7) Quit or go back to (2).

The amount of "promise" that a subinterval exhibits for knot insertion is determined by first computing the SSE (variable component) for the knot subinterval in question and then tentatively inserting a knot using SBP(1) and noting the SSE associated with this tentative insertion. The difference between these two SSE's gives some measure of how much the global SSE is likely to be ultimately reduced by the knot insertion. The knot subinterval having the largest such reduction is therefore picked for knot insertion.

It should be recalled that when we use SBP(2) for moving a knot, we are also redefining the ordinate of the approximant corresponding to the moved knot. What this amounts to is simply Gauss-Seidel iteration for the ordinates simultaneously mixed with knot movement. This doubly iterative process makes it unnecessary to compute the global normal equations and solve the resulting tridiagonal system.

IX. SSE FOR APPROXIMANT OVER ONE KNOT SUBINTERVAL. The SSE for knot subinterval i-1 is derived here.

$$SSE = \sum_{-1}(s_{i-1}v_{i-1}+r_{i-1}v_i-y_m)^2 = \sum_{-1}y_m^2 + (n_{i-1}+1)[2n_{i-1}((v_{i-1}+v_i)^2 - v_iv_{i-1}) + (v_i-v_{i-1})^2]/(6n_{i-1}) + (2/n_{i-1})(T_{-1}-S_{-1})(v_i-v_{i-1}) - 2v_iS_{-1}$$

The variable component of this SSE can be computed in 9 SO's.

X. FIRST AND LAST ORDINATE ESTIMATION. The normal equation for the first ordinate of the approximant (i=1) is:

$$v_i\sum_0 s_i^2 + v_{i+1}\sum_0 r_is_i = \sum_0 y_ms_i$$

Substituting sums from the appendix gives:

$$v_1(1/n_1+3+2n_1) + v_2(n_1-1/n_1) = 6(T_0-S_0)/n_1 \qquad (10.1)$$

With 10.1 we compute v_1 when v_2 is held fixed. This is essentially a specia' case of the SBP(2) algorithm; we cannot move the first knot but we must estimate its ordinate.

The normal equation for the last ordinate of the approximant (i=N) is:

$$v_{i-1}\sum_{-1} r_{i-1}s_{i-1} + v_i\sum_{-1} r_{i-1}^2 = \sum_{-1} y_m r_{i-1}$$

Substituting sums from the appendix in this gives us:

$$v_{N-1}(n_{N-1}-1/n_{N-1}) + v_N(1/n_{N-1}+3+2n_{N-1}) = 6[S_{-1}+(S_{-1}-T_{-1})/n_{N-1}] \quad (10.2)$$

With 10.2 we compute v_N when v_{N-1} is held fixed.

APPENDIX

Let n_i = the number of data mesh subintervals in the ith knot subinterval

$\therefore n_i h = u_{i+1} - u_i = l_i$ where h is the mesh size

now, $x_m = x_1 + (m-1)h$

and $u_i = x_1 + (M_i - 1)h$

$\therefore r_i(x_m) = (x_m - u_i)/l_i = (m - M_i)/n_i$

similarly, $r_{i-1}(x_m) = (m - M_{i-1})/n_{i-1}$.

These last two identities are used throughout the derivation of the various sums.

All the sums given here in the appendix are calculated using the formulas for sums of first and second powers and summation by parts:

$$\sum_{i=1}^{n} i = n(n+1)/2$$

$$\sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6$$

$$\sum_{i=m}^{n} a_i b_i = a_{n+1} \sum_{i=m}^{n} b_i - \sum_{i=m}^{n} \Delta a_i \sum_{j=m}^{i} b_j$$

The summation by parts formula (due to Abel) may be derived in the following manner:

$$\Delta a_i b_i = a_{i+1} b_{i+1} - a_i b_i$$

$$= (a_{i+1} - a_i + a_i) b_{i+1} - a_i b_i$$

$$= b_{i+1} \Delta a_i + a_i \Delta b_i$$

$\therefore$

$$a_i \Delta b_i = \Delta a_i b_i - b_{i+1} \Delta a_i$$

$$\sum_{i=m}^{n} a_i \Delta b_i = a_{n+1} b_{n+1} - a_m b_m - \sum_{i=m}^{n} b_{i+1} \Delta a_i$$

let $c_i = \Delta b_i$

$$\therefore \quad \sum_{i=m}^{n} c_i = b_{n+1} - b_m$$

$$\therefore \quad \sum_{i=m}^{n} a_i c_i = a_{n+1}\left(b_m + \sum_{i=m}^{n} c_i\right) - a_m b_m - \sum_{i=m}^{n} \Delta a_i \left(b_m + \sum_{j=m}^{i} c_j\right)$$

$$= a_{n+1} b_m + a_{n+1} \sum_{i=m}^{n} c_i - a_m b_m - b_m(a_{n+1} - a_m) - \sum_{i=m}^{n} \Delta a_i \sum_{j=m}^{i} c_j$$

$$= a_{n+1} \sum_{i=m}^{n} c_i - \sum_{i=m}^{n} \Delta a_i \sum_{j=m}^{i} c_j$$

The power sum formulas may also be derived using summation by parts.

Sums:

$$S_{-1} = \sum_{m=M_{i-1}}^{M_i} y_m$$

$$T_{-1} = \sum_{k=M_{i-1}}^{M_i} \sum_{m=M_{i-1}}^{k} y_m$$

$$S_{-10} = \sum_{m=M_{i-1}}^{M_{i+1}} y_m$$

$$T_{-10} = \sum_{k=M_{i-1}}^{M_{i+1}} \sum_{m=M_{i-1}}^{k} y_m$$

$$\sum_{-1} r_{i-1} y_m = S_{-1} + (S_{-1} - T_{-1})/n_{i-1}$$

$$\sum_{-1} s_{i-1} y_m = (T_{-1} - S_{-1})/n_{i-1}$$

$$\sum_{0} r_i y_m = S_{-10} + (S_{-10} - S_{-1} + T_{-1} - T_{-10})/n_i$$

$$\sum_{0} s_i y_m = y_{M_i} - S_{-1} - (S_{-10} - S_{-1} + T_{-1} - T_{-10})/n_i$$

$$\sum_{-1} s_{i-1}^2 = \sum_{-1} r_{i-1}^2 = (n_{i-1}+1)(2n_{i-1}+1)/(6n_{i-1})$$

$$\sum_{0} s_i^2 = \sum_{0} r_i^2 = (n_i+1)(2n_i+1)/(6n_i)$$

$$\sum_{-1} s_{i-1} = \sum_{-1} r_{i-1} = (n_{i-1}+1)/2$$

$$\sum_{0} s_i = \sum_{0} r_i = (n_i+1)/2$$

$$\sum_{-1} r_{i-1}s_{i-1} = (n_{i-1}+1)(n_{i-1}-1)/(6n_{i-1})$$

$$\sum_{0} r_i s_i = (n_i+1)(n_i-1)/(6n_i)$$

$$S_0 = \sum_{m=M_i}^{M_{i+1}} y_m$$

$$T_0 = \sum_{k=M_i}^{M_{i+1}} \sum_{m=M_i}^{k} y_m$$

$$\sum_{0} r_i y_m = [(n_i+1)S_0 - T_0]/n_i$$

$$\sum_{0} s_i y_m = (T_0 - S_0)/n_i$$

REFERENCES

1. Ertel, J.E. and E.B. Fowlkes, "Some algorithms for linear spline and piecewise multiple regression", JASA 71(355),1976,p. 640-648.

2. Hamming, R.W., Digital Filters, Prentice Hall, 1977.

3. Jupp, D.L.B., "Approximation to data by splines with free knots", Siam J. Numer. Anal. 15(2), 1978.

4. Shamos, N.I., "Geometry and statistics: problems at the interface" in Algorithms and Complexity: New Directions and Recent Results, J.F. Traub ed., Academic Press, 1976, p. 251-280.

5. Smith, P.W. and S. Hrncir, "Nonlinear spline regression on minicomputers", Proc. of the 1976 Army Numer. Anal. and Computers Conf., ARO Rept. 76-3, p. 53-90.

6. Wilson, D.G., "Algorithm 510, piecewise linear approximation to tabulated data", ACM Trans. Math. Software, 2(4), 1976.

An Extension $C\alpha$ of C_J That has an Application in Learning Theory

Charles R. Leake
US Army Armor and Engineer Board
Fort Knox, Kentucky 40121

ABSTRACT

$C\alpha$ which is an extension of C_J is discussed. $C\alpha$ is a class of algebras that are commutative, generally nonassociative with inverses for all non zero elements when the ground field is the reals. In general $C\alpha$ contains divisors of zero. Each $C\alpha$ has an involutorial automorphism and is a quadratic extension of the ground field. Moreover, for each element x in any $C\alpha$ there are unique elements T_x and N_x. An example is given showing how previous, experimental and concurrent learning can be resolved and how the magnitude of the learning or training effectiveness can be measured.

1. The concept of C_J and C_N. In a recent paper [9] has shown that C_J and C_N have applications in thermodynamics. C_J is generally a commutative, nonassociative algebra of order J with identity $1\neq 0$ which under the condition that the ground field is the real numbers contains an inverse for each nonzero element as well as zero divisors. In a special case for $J = 2$, $C_J \simeq C$, the complex numbers and for $J = 1$, $C_J \simeq R$, the real numbers. C_N is generally a noncommutative, nonassociative algebra of order N with identity $1\neq 0$ which under the condition that the ground field is R contains an inverse for each nonzero element as well as zero divisors. Under appropriate conditions for $N = 1,2$ & 4; $C_N \simeq R$, $C_N \simeq C$ and $C \simeq Q$, the quaternions. C_N & C_J were also shown in [9] to belong to a broad class of algebras that are known as quadratic extensions Γ of a field K.

Considerable work has been done on sets Γ. In [3] Γ is characterized for the real, complex, quaternions and cayley number systems. In [4] Γ is generalized to a field K where there is an element i such that $i^2-\beta i-\alpha=0$. The concept is then extended to the case when K is a commutative ring with unit that admits an involutorial automorphism in [5] and in [6] the geometry of the place of a quadratic extension Γ of a field K is discussed. In [7] and [8] there are examples of when Γ is a nonassociative algebra.

Quadratic extensions belong to a class of algebras commonly known as Clifford numbers. Vander Waerden in [11] discusses a class of these numbers which he calls hypercomplex numbers.

This paper will be limited to a discussion of $C\alpha$ which is an extension of C_J.

2. The concept of the commutative algebra $C\alpha$ over a field K of characteristic $\neq 2$. $C\alpha$ is an algebra of order $M \geq 1$ where $1, e_2, e_3, \ldots, e_m$ is a basis for $C\alpha$. In the case of $M = 1$, $C\alpha = K$. In addition there exist $\alpha_1, \alpha_2, \ldots, \alpha_M \in K$. Using the operations defined on K, $C\alpha$ has the following sum and products defined on it for all $a, b \in C\alpha$ and $\theta, a_i, b_i \in K$

$$(1)\quad a + b = \sum_{i=1}^{M} (a_i + b_i)\, e_i$$

$$(2)\quad \theta a = \theta \sum_{i=1}^{M} a_i\, e_i$$

$$(3)\quad ab = (\alpha_i^2\, a_1 b_1 - \sum_{i=2}^{M} \alpha_i^2\, a_i b_i)\, 1 + \sum_{i=2}^{M} (\alpha_1\, \alpha_i\, a_1\, b_i + \alpha_i \alpha_1 a_i b_1)\, e_i$$

The automorphism $a \longrightarrow \bar{a}$ is

$$(4)\quad \bar{a} = a.1 - \sum_{i=2}^{M} a_i\, e_i$$

The unit or 1 in $C\alpha$ is

$$(5)\quad 1 = 1 - \sum_{i=2}^{M} 0 e_i$$

The trace T_a and norm N_a are

$$(6)\quad T_a = \bar{a} + a \text{ and } (7)\quad N_a = a\bar{a}$$

T_a and $N_a \in K$ and $ab = ba$ for all $a, b \in C\alpha$. When the characteristic of K is 0, K = T, the set $C\alpha$ has inverses

$$(7) \quad a^{-1} = \frac{\bar{a}}{N_a}$$

Moreover, in general

$$(8) \quad N_{ab} \neq N_a N_b$$

$$(9) \quad N_a{}^2 \neq (N_a)^2$$

For $\alpha_i = 1$, $i = 1, 2, \ldots, M$, $C\alpha \simeq C_J$.

Each element a of $C\alpha$ satisfies the equation,

$$(10) \quad a^2 - T_a a + N_a = 0$$

Thus $C\alpha$ also belongs to the class of algebras known as quadratic extensions Γ of a field K. See [1], [2], [3], [4], [5] and [7] for a more generalized discussion of quadratic extensions.

3. An application of $C\alpha$ to learning theory. One of the problems with learning theory is its lack of a geometric base. Early discoveries in science were related to Euclidean geometry and many of its premises were directly related to observations made in terms of Euclidean geometry. We are no longer so naive as to believe that the universe is Euclidean, but in moving away from this geometry, we have been led into a position where geometries are now abritrary. In some sciences the geometry is related to the law of least squares with a statistical intepretation of the results, but this makes it difficult to combine results from other sources. The least squares approach has also been tried in educational circles, but with rather disappointing results. What is needed is a geometry that incorporates recent learning and affords the researcher the opportunity of combining data from other sources with his theories such as those in [10] and [1] & [2]. In order to do this, it is required to move beyond one and two dimensional space into N-dimensional space. Until now thought in this area has been limited due to the striking results provided by the Cartan-Hurowitz theorem. However, Jordan has provided us with some insights into the problem of multi dimensional algebras, but these have been mainly limited to the physical sciences and biology. $C\alpha$ offers us a geometry which is relevant to experimentally based educational data.

Of prime concern in education theory is coordinating previous learning experiences with those being examined experimentally. Presently this is being done by using a statistical technique known as the analysis of covariance where previous educational or intelligence factors are covariated out, usually in a linear manner. Other techniques such as rotational techniques are designed to remove cross terms. These again are based on a least squares interpretation of the data which attempts to unscramble the interlocking relationships between the variables into a manageable array. The method proposed in this paper is different. Previous experiences as well as concurrent experiences are resolved in the traditional manner by vector addition.

For example, suppose as a start in analyzing subject A with regards to a particular educational goal, we used the classification scheme indicated in [1]. This would require that we measure the subject in six dimensions. We are not limited by any means to a six dimensional analysis and as many as we chose could be subtracted or added to the original array. [1] has six classifications, namely knowledge, comprehension, application, analysis, synthesis and evaluation. We could very easily add to this array I.Q. or any other dimension which we wished to include. For the purposes of this example, we will not. Suppose subject A was doing some work which was connected with a proposed experiment. We could call that concurrent education and measure the effect on the educational goal in six dimensions. Next, we give subject A some treatment. Again we measure him in six dimensions in terms of our educational goal. Call these measurements u_1, u_2 & u_3 respectively. The resolution of subject A's learning is

$$X = u_1 + u_2 + u_3 \tag{12}$$

The magnitude of his learning would be $N\alpha$, where

$$N\alpha = \sum_{i=1}^{6} \alpha_i^2 x_i^2 \tag{13}$$

The α_i's could at first be based upon expert opinion. Ultimately they could be experimentally derived.

Another feature of this method is that it enables us to resolve not only the cognitive part of learning, but the affective as well. In [2] there is a relation between the cognitive dimensions in [1]. If we change the ground field from the reals to the complex, the affective part of learning could become the imaginary component of learning. The definition for the resolution and magnitude of learning

remain the same as given above. However, the possibility for showing negative learning can be developed when using complex numbers as the ground field instead of the real numbers.

The effect of scalar multiplication can be used to establish standards for goals to attain in training programs or refresher courses. Erasing bad training can be established by examining the inverses to learning. Vector multiplication can be used to resolve different educational goals where zero divisors represent conflicts that cancel learning.

The method described so far need not be limited to the applications suggested by [1] and [2], but can be used for multi-tiered learning such as that suggested in [8]. Moreover, a skill such as learning how to fire a main gun of a tank can be broken up into its skill components each of which can be assigned a dimension. Resolution of learning can be obtained by [12], and [13] can be used to predict main gun performance or correlated with it. Or, the skills that are desired can be related to those given in [1] and [2] with resolution of learning and its magnitude as previously defined.

The purpose of this paper was to show that $C\alpha$ had an application in learning theory. Clearly, there are a multitude of applications of $C\alpha$ not only to learning theory but the physical and behavioral sciences as well.

Bibliography

1. Bloom, B. S., et. al., Taxonomy of Educational Objectives, Handbook I: Cognitive Domain, New York, N.Y.: David McKay Co., Inc., 1966.

2. _____, Taxonomy of Educational Objectives, Handbook II: Affective Domain, New York, N.Y.: David McKay Co., Inc., 1965.

3. Curtis, C. W., "The four and eight square problem and division algebras," Studies in Modern Algebra, Vol. 2, Mathematical Association of America, 1963.

4. De Cicco, J. "Introduction to the theory of a quadratic extension of a field K," Universita e Politecnico di Torino Rendiconti del Seminario Matematico, vol. 17, 1957/58, pp. 223-251.

5. De Cicco, J. "Some theorems concerning commutative rings with unit which admit involutorial automorphism," Redle Accademia della Scienze di Torino. Atti Classe di Scienze, Fisiche, Mathematiche e Naturali, vol 92, 1957/58, pp. 225-242.

6. DeCicco, J. "The geometry of the z-plane based on a quadratic extension of a field K," Universita e Politecnico di Torino Rendiconti del Seminario Matematico, vol. 18, 1958/59, pp. 91-119.

7. Jacobson, N., "Structure and representations of Jordan algebras," American Mathematical Society Colloquium Publications, vol. 39, American Mathematical Society, 1968.

8. Kleinfeld, E., "A characterization of the Cayley numbers, "Studies In Modern Algebra, vol 2, Mathematical Association of America, vol 2, 1963.

9. Leake, C. R., Some Notes on an Application of C_J and C_N to the Physical Science (unpublished paper, 1978).

10. Piaget, J., The Psychology of Intelligence, Paterson, N.J.: Littlefield, Adams & Co., 1963.

11. van der Waerden, B. L., Modern Algebra, vol. I, New York, N.Y.: Frederick Ungar Publishing Co., 1953.

AN ALGORITHM FOR HEAT TRANSFER IN GUN BARRELS

John F. Polk
Fragmentation Branch
Terminal Ballistics Division
US Army Ballistic Research Laboratory
Armament Research and Development Command
Aberdeen Proving Ground, MD 21005

ABSTRACT. Experimental measurements indicate that steep temperature gradients exist near the bore surface of gun barrels during operation. These pose severe difficulties for obtaining the true surface temperatures and for developing effective mathematical models of the heat transfer process. Singular perturbation methods provide a natural means for attacking the mathematics underlying such problems and can be used to obtain asymptotic expansions for the bore surface temperature, valid for small times. These expansions have been incorporated into a basic algorithm which can be repeated as frequently as necessary to predict temperatures over longer time intervals. In conjunction with a simple physical model of the interior ballistics this procedure has resulted in temperature predictions showing excellent overall agreement with measured data.

I. INTRODUCTION. On the most elementary level the transient heat transfer occurring in a gun barrel can be described by the following model. The temperature θ at a given axial station will be assumed to depend only on the radial coordinate, R, and on time, T, which vary over the ranges

$$R_0 \leq R \leq R_1 \text{ and } 0 \leq T \leq T_1$$

where

R_0 = inner bore radius

R_1 = exterior barrel radius

T = 0 when the bullet passes the axial station

$T = T_1$ is the maximum time of interest.

In order for the short time asymptotic methods which we shall use to be valid it will be necessary to assume that

$$T_1 << R_0^2/k \tag{1.1}$$

where $k = K/\rho c$ is the coefficient of thermal diffusivity,

K = thermal conductivity

ρ = density

c = specific heat.

More comment will be made on this point shortly, however we should note that longer term solutions can be constructed by repeating the basic approximations over successive time intervals $m\ T_1 \leq T \leq (m+1)\ T_1$.

For convenience we shall henceforth use the non-dimensional form

$$U\ (R,T) = (\theta(R,T) - \theta_o)/(\theta_M - \theta_o)$$

to describe the temperature where

θ_o = ambient temperature

θ_M = melting point for gun barrel steel.

We suppose that this function is known at time $T = 0$ in the form

$$U\ (R,0) = F\ (R) \tag{1.2}$$

and that its subsequent rise is governed by the linear heat equation, in cylindrical form

$$U_T = k\ [U_{RR} + \frac{1}{R}\ U_R]. \tag{1.3}$$

At the bore surface the heat transfer obeys the convective law

$$[U - \frac{K}{H}\ U_R]\ (R_o,T) = G\ (T) = \frac{\theta_{gas}\ (T) - \theta_o}{\theta_M - \theta_o} \tag{1.4}$$

where H is the heat transfer coefficient and $\theta_{gas}\ (T)$ is the instantaneous temperature of the propellant gas in the barrel. At the outer surface the short term heat losses will be considered negligible so that the zero flux condition

$$U_R\ (R_1,T) = 0 \qquad\qquad 0 \leq T \leq T_1 \tag{1.5}$$

applies. The zero flux condition at $R = R_o$ can be obtained as a special case of (1.4) by letting $H \to 0$.

Clearly this model is a simple one since it is linear and implicitly assumes that the gas dynamics and heat transfer to the barrel are separable problems. It also assumes that H and G (T) are known whereas the available experimental data concerning these quantities is very sketchy. Nevertheless, an analysis of the problem, as formulated, is relevant for several reasons: first, it provides a qualitative insight into the relative importance of the various parameters; second, an understanding of the non-linear problem is only possible after a thorough rendering of the linear case; third, more complicated problems can be handled by quasi-linearization in which the equations are treated locally as linear in restricted sub-regions; finally, even with this simple model we are able to obtain reasonable agreement with experimental data, in some cases.

To reduce the problem even further and to obtain explicit solutions we limit our consideration to low order polynomial forms for F (R) and G (T). In this regard we shall let V_n (R,T), n = 0,1 and W_n (R,T) n = 0,1,2 denote the five solutions of problem (1.2) - (1.5) having the particular supplementary data indicated in Table 1.

Table 1

Problem	Solution	Initial Values F (R)	Gas Temp. G (T)
1	V_0	0	1
2	V_1	0	T/T_1
3	W_0	1	0
4	W_1	$(R-R_o)/R_o$	0
5	W_2	$((R-R_o)/R_o)^2/2$	0

Other functions of physical interest could be considered but this table represents a minimal list of functions we should be able to treat. It also appears to be adequate for application to the gun barrel problem.

II. ANALYSIS. As a first step let us non-dimensionalize the foregoing model by introducing the independent variables

$$r = R/R_o$$

$$t = kT/R_o^2$$

which vary over the ranges

$$1 \leq r \leq r_1 \qquad \text{and} \qquad 0 \leq t \leq t_1$$

where

$$r_1 = R_1/R_o \text{ and } t_1 = k\, T_1/R_o^2 .$$

The function U can now be written as

$$u\,(r,t) = U\,(r\, R_o,\ t\, R_o^2/k) = U\,(R,T)$$

and problem (1.2) - (1.5) transforms into

$$u_t = u_{rr} + \frac{1}{r}\, u_r \tag{2.1}$$

$$u\,(r,o) = f\,(r) = F\,(r\, R_o) \tag{2.2}$$

$$\left[u - \frac{K}{HR_o}\, u_r\right](1,t) = g\,(t) = G\,(t\, R_o^2/k) \tag{2.3}$$

$$u_r\,(r_1,t) = 0. \tag{2.4}$$

Note that the coefficients in (2.1) are all of order unity while (1.1) implies that

$$0 \leqslant t << 1 . \tag{2.5}$$

Thus the problem in this form is truly a short-time problem.

In this non-dimensional form problem (2.1) - (2.4) is suitable for analysis using the DESS (Diffusion Equation Solution Sequence) method which was introduced in a separate discussion[1,2]. This is a technique, based on the assumption of small times, in which asymptotic expansions are developed for the solution in those regions where singularities such as steep gradients (boundary layers) are encountered. In the present case such a phenomenon is observed near the bore surface and is caused by the sudden rush of hot gasses over an initially cool surface. The flux of heat to the barrel is so sudden that it cannot be diffused uniformly outward but results in a thin, high-temperature region near the bore surface $R = R_o$ $(r = 1)$. In the mathematical context this condition arises when (1.1) is satisfied. For gun barrels we have the typical values

$$R_o \approx 1 \text{ cm}$$

$$k \approx .1 \text{ cm}^2/\text{sec}$$

and thus our analysis will apply when

$$T_1 << 10 \text{ sec.}$$

For common gun systems this is many times greater than the time during which convective heating of the barrel occurs.

In our previous discussion of the DESS method asymptotic expansions for solutions of the equation

$$u_t = a(x)\, u_{xx} + b(x)\, u_x + c(x)\, u$$

were obtained in explicit form. However, only the Dirichlet type (function value specified) of boundary condition was considered so that the expansions obtained previously do not directly apply to the present case. On the other hand the formal procedures used to derive those expansions can be repeated for problem (2.1) - (2.4) to obtain a different, but still explicit, expansion for its solution. Let us now see how this is done.

The basic procedure is to emphasize the local, short-term behavior of u near r = 1 by introducing the stretched variables

$$\sigma = (r-1)/\varepsilon$$

$$\tau = t/\varepsilon^2$$

and a transformed or "inner" solution

$$\tilde{u}(\sigma,\tau) = u(1 + \varepsilon\,\sigma,\ \varepsilon^2\,\tau) = u(r,t)$$

where, for convenience we have introduced the notation

$$\varepsilon = \sqrt{t_1} = \sqrt{k\ T_1}/R_o\ .$$

From (2.1) - (2.3) we can obtain the new equations

$$\tilde{u}_\tau = \tilde{u}_{\sigma\sigma} + \frac{\varepsilon}{1 + \varepsilon\sigma}\ \tilde{u}_\sigma \qquad \sigma > o,\ \tau > o \qquad (2.6)$$

$$\tilde{u}(\sigma,o) = \tilde{f}(\sigma) = f\ (1 + \varepsilon\ \sigma) \qquad \sigma > o \qquad (2.7)$$

$$\tilde{L}_h\ \tilde{u}(o,\tau) = \tilde{g}(\tau) = g(\varepsilon^2\ \tau) \qquad \tau > o \qquad (2.8)$$

where $\tilde{L}_h$ here denotes the operator

$$\tilde{L}_h\ \tilde{u} = \tilde{u} - \frac{1}{h}\,\tilde{u}_\sigma$$

with

$$h = H\ \sqrt{T_1/K\rho c}\ .$$

The boundary condition at $R = R_1$ is ignored in this new system since it now occurs at the coordinate

$$\sigma_1 = (r_1 - 1)/\varepsilon$$

which, for small values of ε, is very remote. In its place we should require that $\tilde{u}(\sigma,\tau)$ satisfy a growth condition as $\sigma \to \infty$. In the present discussion, however, we are not concerned with the questions of uniqueness and continuous dependence on the data but only wish to explain the mechanics by which a formal expansion is obtained. The only justification of our methods is accomplished *a postoriori* by comparing our computations with experiment.

In connection with problems 1-5 we shall use the notation $\tilde{V}_n(\sigma,\tau)$ and $\tilde{W}_n(\sigma,\tau)$ in place of the general solution $\tilde{u}(\sigma,\tau)$. The supplementary data for these problems, obtained by transforming Table 1, is given in Table 2.

Table 2. Supplementary data for Problems 1-5 in Stretched Variables

Problem	Solution $\tilde{u}(\sigma,\tau)$	Initial Temp. $\tilde{f}(\sigma)$	Gas Temp. $\tilde{g}(\tau)$
1	$\tilde{V}_0(\sigma,\tau)$	0	1
2	$\tilde{V}_1(\sigma,\tau)$	0	τ
3	$\tilde{W}_0(\sigma,\tau)$	1	0
4	$\tilde{W}_1(\sigma,\tau)$	$\varepsilon\sigma$	0
5	$\tilde{W}_2(\sigma,\tau)$	$\varepsilon^2\sigma^2/2$	0

Let us now suppose that $\tilde{u}$ can be [illegible] for small $\varepsilon > 0$ in the asymptotic form

$$\tilde{u} \sim \varepsilon^p [\tilde{u}_0 + [illegible] + \cdots\cdots] . \tag{2.9}$$

The value of the exponent p follows from the particular choice of the supplementary data; from Table 2 we see

$$p = \begin{cases} 0 & \text{for problems 1, 2, 3} \\ 1 & \text{for problem 4} \\ 2 & \text{for problem 5.} \end{cases} \tag{2.10}$$

Substituting (2.9) into (2.6), recalling the expansion

$$\frac{1}{1+\varepsilon\sigma} \sim 1 - \varepsilon\sigma + (\varepsilon\sigma)^2 - (\varepsilon\sigma)^3 + \ldots$$

and collecting by powers of ε results in the following system

$$(\tilde{u}_0)_\tau - (\tilde{u}_0)_{\sigma\sigma} = 0$$

$$(\tilde{u}_1)_\tau - (\tilde{u}_1)_{\sigma\sigma} = (\tilde{u}_0)_\sigma$$

$$(\tilde{u}_2)_\tau - (\tilde{u}_2)_{\sigma\sigma} = (\tilde{u}_1)_\sigma - \sigma\,(\tilde{u}_0)_{\sigma\sigma}$$

and, in general,

$$(\tilde{u}_k)_\tau - (\tilde{u}_k)_{\sigma\sigma} = \sum_{j=1}^{k} \tilde{L}_j\, \tilde{u}_{k-j} \tag{2.11}$$

where $\tilde{L}_j$ is defined by

$$\tilde{L}_j\, \tilde{u} = (-\sigma)^{j-1}\, \tilde{u}_\sigma \ .$$

The sequence $\{\tilde{u}_k : k = 0,1,2, \ldots.\}$ therefore forms a DESS according to the definition in References 1 and 2.

Specializing to the solutions of problems 1-5 let us use the notation $\tilde{V}_{n,k}$ and $\tilde{W}_{n,k}$ in place of the individual terms $\tilde{u}_k$ of (2.9); that is, we seek the expansions

$$\tilde{V}_n\ (\sigma,\tau) \sim [\tilde{V}_{n0} + \varepsilon\, \tilde{V}_{n1} + \varepsilon^2\, \tilde{V}_{n2} + \ldots]\ (\sigma,\tau) \tag{2.12}$$

$$\tilde{W}_n\ (\sigma,\tau) \sim \varepsilon^n\ [\tilde{W}_{n0} + \varepsilon\, \tilde{W}_{n1} + \varepsilon^2\, \tilde{W}_{n2} + \ldots]\ (\sigma,\tau). \tag{2.13}$$

These satisfy the equations

$$(\tilde{V}_{nk})_\tau - (\tilde{V}_{nk})_{\sigma\sigma} = \begin{cases} 0 & k=0 \\ \sum_{j=1}^{k} \tilde{L}_j\, \tilde{V}_{n,k-j} & k=1,2\ldots \end{cases} \tag{2.14}$$

$$(\tilde{W}_{nk})_\tau - (\tilde{W}_{nk})_{\sigma\sigma} = \begin{cases} 0 & k=0 \\ \sum_{j=1}^{k} \tilde{L}_j \tilde{W}_{n,k-j} & k=1,2,\ldots \end{cases} \qquad (2.15)$$

in the domain $\sigma > o$, $\tau > o$. Supplementary data for these terms is obtained by substituting from Table 2 into (2.12) and (2.13) and collecting by powers of ε. This yields

$$\tilde{V}_{n,k}(\sigma,o) = 0 \qquad k=0,1,2,\ldots \qquad (2.16)$$

$$\tilde{L}_h \tilde{V}_{n,k}(o,\tau) = \begin{cases} t^n/n! & k=0 \\ 0 & k=1,2,3,\ldots \end{cases} \qquad (2.17)$$

and

$$\tilde{W}_{n,k}(\sigma,o) = \begin{cases} \sigma^n/n! & k=0 \\ 0 & k=1,2,3,\ldots \end{cases} \qquad (2.18)$$

$$\tilde{L}_h \tilde{W}_{n,k}(o,\tau) = 0 \qquad k=0,1,2,\ldots \qquad (2.19)$$

The systems (2.14), (2.16) and (2.17) for $\tilde{V}_{n,k}$ and (2.15), (2.18) and (2.19) for $\tilde{W}_{n,k}$ can be explicitly solved in terms of certain special functions which have been investigated by the author in separate work. In the next section we shall briefly review these functions and apply them to the present problem.

III. SPECIAL FUNCTIONS. The special functions H_γ, H_γ^*, Z_γ and $Z_\gamma^\#$ (γ is any real number) were defined and investigated in References 3 and 4 which can be consulted for more detail. Their basic significance is that they are solutions of the heat equation which satisfy special initial and boundary conditions. The most basic of these are the functions H_γ which are defined for $t > 0$ by

$$H_\gamma(x,t) = (4\pi t)^{-1/2} \int_o^\infty \frac{s^\gamma}{\gamma!} \exp\left[-(x-s)^2/4t\right] ds$$

when $\gamma > -1$ and recursively by

$$H_\gamma(x,t) = \frac{\partial}{\partial x} H_{\gamma+1}(x,t)$$

when $\gamma \leq -1$. For $t = 0$ and all γ we define

$$H_\gamma(x,0) = h_\gamma(x) \equiv \begin{cases} 0 & x \leq 0 \\ x^\gamma/\gamma! & x > 0 \end{cases} \tag{3.1}$$

The functions H^*_γ are next defined by

$$H^*_\gamma(x,t) \equiv H_\gamma(-x,t)$$

and have the initial values

$$H^*_\gamma(x,0) = h^*_\gamma(x) \equiv h_\gamma(-x). \tag{3.1}^*$$

Along $x = 0$ these functions take on the values

$$H_\gamma(o,t) = H^*_\gamma(o,t) = \sqrt{t}^\gamma/2(\gamma/2)! \tag{3.2}$$

For integer values of γ, $\gamma = n$, the functions H_γ and H^*_γ can be obtained in explicit functional forms. For example

$$\begin{aligned} H_{-1}(x,t) &= (4\pi t)^{-1/2} \exp[-x^2/4t] \\ H_0(x,t) &= (1/2)\,\mathrm{erfc}(-x/\sqrt{4t}) \\ H_1(x,t) &= x H_0 + 2t H_{-1} \\ H_2(x,t) &= [(x^2 + 2t) H_0 + 2xt H_{-1}]/2 \\ H_3(x,t) &= [x^3 + 6xt) H_0 + 2(x^2t + 4t^2) H_{-1}]/3! \\ H_4(x,t) &= [(x^4 + 12x^2t + 12t^2) H_0 + 2(x^3t + 10xt^2) H_{-1}]/4! \end{aligned} \tag{3.3}$$

All of the functions H_γ and H^*_γ are infinitely differentiable with respect to both variables and satisfy the following relations

$$\frac{\partial}{\partial x} H_\gamma = H_{\gamma-1} \qquad \frac{\partial}{\partial x} H^*_\gamma = - H^*_{\gamma-1}$$

$$\frac{\partial}{\partial t} H_\gamma = H_{\gamma-2} \qquad \frac{\partial}{\partial t} H^*_\gamma = H^*_{\gamma-2} .$$

The functions Z_γ and Z^*_γ can now be defined as the transformations

$$Z_\gamma (x,t) = T_h H_\gamma (x,t)$$

$$Z^*_\gamma (x,t) = T_h H^*_\gamma (x,t)$$

where

$$T_h [f(x,t)] \equiv h \int_x^\infty f(s,t)\, e^{h(x-s)}\, ds.$$

This transformation is actually the inverse of the differential operation $L_h u = u - \frac{1}{h} u_x$ so that

$$L_h Z_\gamma = H_\gamma$$

$$L_h Z^\#_\gamma = H^*_\gamma$$

and consequently along $x = 0$ we have from (3.2)

$$L_h Z_\gamma (o,t) = L_h Z^\#_\gamma (0,t) = \sqrt{t}^\gamma / 2(\gamma/2)! \tag{3.4}$$

The differentialiability of the functions H_γ and H^*_γ carries over to their transforms; we have

$$\frac{\partial}{\partial x} Z_\gamma = Z_{\gamma-1} \qquad \frac{\partial}{\partial x} Z^\#_\gamma = - Z^\#_{\gamma-1} \tag{3.5}$$

$$\frac{\partial}{\partial t} Z_\gamma = Z_{\gamma-2} \qquad \frac{\partial}{\partial t} Z^\#_\gamma = Z^\#_{\gamma-2} \tag{3.6}$$

For our applications only the functions $Z^\#_\gamma$, not Z_γ, will be needed. These can be shown to have the series representation

$$Z_\gamma^{\#}(x,t) = -\sum_{k=1}^{\infty} (-h)^k H_{\gamma+k}(x,t) \tag{3.7}$$

and thus from (3.1) and (3.2) we have

$$Z_\gamma^{\#}(x,0) = 0 \qquad x \geq 0 \tag{3.8}$$

$$Z_\gamma^{\#}(o,t) = -\frac{\sqrt{t}^\gamma}{2} \sum_{k=1}^{\infty} \frac{(-h\sqrt{t})^k}{((\gamma+k)/2)!} \qquad t > 0\ ; \tag{3.9}$$

the latter series can be truncated for small values of $h\sqrt{t}$. When γ is an integer we have the following explicit functional forms for $Z_\gamma^{\#}$: for $\gamma = -1$

$$Z_{-1}^{\#}(x,t) = (h/2)\ \mathrm{erfc}\ ((x + 2ht)/\sqrt{4t})\ \exp(hx + h^2t) \tag{3.10a}$$

otherwise

$$Z_n^{\#}(x,t) = (-h)^{-n-1} Z_{-1}^{\#}(x,t) + \begin{cases} \sum_{k=o}^{n} (-h)^{k-n} H_k^*(x,t), & n \geq 0 \\ -\sum_{k=1}^{-n-1} (-h)^{-n-k} H_k^*(x,t), & n \leq -2 \end{cases} \tag{3.10b}$$

Along $x = 0$ we then have

$$Z_n^{\#}(o,t) = \frac{h}{2}\ \mathrm{erfc}\ (h\sqrt{t})\ \exp(h^2t) \tag{3.11a}$$

$$Z_n^{\#}(o,t) = (-h)^{-n-1} Z_{-1}^{\#}(o,t) + \frac{1}{2} \begin{cases} \sum_{k=o}^{n} (-h)^{k-n} \frac{\sqrt{t}^k}{(k/2)!}, & n \geq 0 \\ -\sum_{k=1}^{-n-1} (-h)^{-n-k} \frac{\sqrt{t}^k}{(k/2)!}, & n \leq -2 \end{cases} \tag{3.11b}$$

For large values of $h\sqrt{t}$ we can combine (3.11) with the standard asymptotic formula

$$\mathrm{erfc}\ (z)\ \exp(z^2) \sim \frac{1}{\sqrt{\pi}\ z}\left[1 - \frac{1}{2z^2} + \frac{1\cdot 3}{(2z^2)^2} - \frac{1\cdot 3\cdot 5}{(3z^2)^3} + \ldots\right]$$

as $z \to \infty$ to obtain

$$Z_n^{\#}(o,t) \sim \frac{1}{2}(-h)^{-n} \sum_{k=-\infty}^{n} \frac{(-h\sqrt{t})^k}{(k/2)!} \tag{3.12}$$

as $h\sqrt{t} \to \infty$.

We can now return to the problems formulated at the end of Section II and write explicit solutions for the terms $\tilde{V}_{n,k}$ and $\tilde{W}_{n,k}$ of expansions (2.12) and (2.13). We have

$$\begin{aligned}
\tilde{V}_{0,0}(\sigma,\tau) &= 2 Z_o^{\#} \\
\tilde{V}_{0,1}(\sigma,\tau) &= Z_1^{\#} - 2\tau Z_{-1}^{\#} \\
\tilde{V}_{0,2}(\sigma,\tau) &= \tau^2 Z_{-2}^{\#} + (\sigma^2/2) Z_o^{\#}
\end{aligned} \tag{3.13}$$

$$\begin{aligned}
\tilde{V}_{1,0}(\sigma,\tau) &= 2 Z_2^{\#} \\
\tilde{V}_{1,1}(\sigma,\tau) &= 3 Z_3^{\#} - 2\tau Z_1^{\#} \\
\tilde{V}_{1,2}(\sigma,\tau) &= -2 Z_4^{\#} + (\sigma^2/2 + 4\tau) Z_2^{\#} - \tau^2 Z_o^{\#}
\end{aligned} \tag{3.14}$$

$$\begin{aligned}
\tilde{W}_{0,0}(\sigma,\tau) &= 1 - 2 Z_o^{\#} \\
\tilde{W}_{0,1}(\sigma,\tau) &= 2\tau Z_{-1}^{\#} - Z_1^{\#} \\
\tilde{W}_{0,2}(\sigma,\tau) &= -(\sigma^2/2) Z_o^{\#} - \tau^2 Z_{-2}^{\#}
\end{aligned} \tag{3.15}$$

$$\begin{aligned}
\tilde{W}_{1,0}(\sigma,\tau) &= \sigma + (2/h) Z_o^{\#} \\
\tilde{W}_{1,1}(\sigma,\tau) &= \tau - (2\tau/h) Z_{-1}^{\#} + (1/h) Z_1^{\#} - 2 Z_2^{\#} \\
\tilde{W}_{1,2}(\sigma,\tau) &= 2\tau Z_1^{\#} - 3 Z_3^{\#} + (\tau^2/h) Z_{-2}^{\#} - \sigma\tau + (\sigma^2/2h) Z_o^{\#} - (2/h) Z_2^{\#}
\end{aligned} \tag{3.16}$$

$$\tilde{W}_{2,0}\ (\sigma,\tau) = \sigma^2/2 + \tau - 2\ Z_2^{\#}$$

$$\tilde{W}_{2,1}\ (\sigma,\tau) = \sigma\tau + (2/h)\ Z_2^{\#} + 2\ \tau\ Z_1^{\#} - 3\ Z_3^{\#} \qquad (3.17)$$

$$\tilde{W}_{2,2}\ (\sigma,\tau) = -\ \sigma^2\tau - \tau^2/2 + (3/h)\ Z_3^{\#} + (2\ \tau - \sigma^2/2)\ Z_2^{\#}$$
$$- (2\ \tau/h)\ Z_1^{\#} + \tau^2\ Z_o^{\#}$$

where the functions $Z_n^{\#}$ in the right hand expressions are to be evaluated at σ,τ. Verification of these solutions can be accomplished by direct substitution into the appropriate equations, using (3.4), (3.5), (3.6) and (3.8).

By reversing the derivation in Section II we can express these functions in terms of the original variables R and T. This yields

$$V_n\ (R,T) \sim \tilde{V}_n\ (\sigma,\tau)$$
$$\sim [\tilde{V}_{n0} + \varepsilon\ \tilde{V}_{n1} + \varepsilon^2\ \tilde{V}_{n2}]\ (\sigma,\tau)$$

and

$$W_n\ (R,T) \sim \varepsilon^n\ [\tilde{W}_{n0} + \varepsilon\ \tilde{W}_{n1} + \varepsilon^2\ \tilde{W}_{n2}]\ (\sigma,\tau)$$

where

$$\varepsilon = \sqrt{k\ T_1}/R_o$$

$$\sigma = (R - R_o)/\sqrt{k\ T_1}$$

$$\tau = T/T_1\ .$$

The terms $\tilde{V}_{nk}$ and $\tilde{W}_{nk}$ are evaluated using the above list of functions with the parameter h given by

$$h = H\ \sqrt{T_1/K\rho c}\ .$$

But the parameter T_1 may be considered as a dummy variable since it can be replaced by any value satisfying (1.1). Thus we can replace T_1 by T in the above formulas to obtain the somewhat simpler forms

$$V_n\ (R,T) \sim [\tilde{V}_{n0} + \varepsilon\ \tilde{V}_{n1} + \varepsilon^2\ \tilde{V}_{n2}]\ (\sigma,1) \qquad (3.18)$$

$$W_n\ (R,T) \sim \varepsilon^2\ [\tilde{W}_{n0} + \varepsilon\ \tilde{W}_{n1} + \varepsilon^2\ \tilde{W}_{n2}]\ (\sigma,1) \tag{3.19}$$

where now

$$\varepsilon = \sqrt{kT}/R_o$$

$$\sigma = (R-R_o)/\sqrt{kT} \tag{3.20}$$

and $$h = H\ \sqrt{T/K\rho c}\quad .$$

In particular, along the bore surface $R = R_o$ we can approximate V_n and W_n by

$$V_n\ (R_o,T) \sim \tilde{V}_n\ (0,1)$$

$$\approx \tilde{V}_{no}\ (0,1) + \varepsilon\ \tilde{V}_{n1}\ (0,1) + \varepsilon^2\ \tilde{V}_{n2}\ (0,1) \tag{3.21}$$

$$W_n\ (R_o,T) \sim \tilde{W}_n\ (0,1)$$

$$\approx \varepsilon^n\ [\tilde{W}_{n0}\ (0,1) + \varepsilon\ \tilde{W}_{n1}\ (0,1) + \varepsilon\ \tilde{W}_{n2}\ (0,1)] \tag{3.22}$$

Formulas (3.13) - (3.17) can also be used to obtain approximations for U (R,T) in the special case of zero flux at $R = R_o$ by letting $h \to 0$. Note from (3.7) that

$$Z^{\#}_{\gamma}\ (\sigma,\tau) \to 0$$

and

$$\frac{1}{h}\ Z^{\#}_{\gamma}\ (\sigma,\tau) \to H^{*}_{\gamma+1}\ (\sigma,\tau)$$

as $h \to 0$. Thus formulas (3.13) - (3.17) take on the simplified forms

$$\tilde{V}_{nk}\ (\sigma,\tau) = 0 \qquad \text{for all n and k}$$

$$\tilde{W}_{00}\ (\sigma,\tau) = 1$$

$$\tilde{W}_{01}\ (\sigma,\tau) = 0 \tag{3.15}'$$

$$\tilde{W}_{02}\ (\sigma,\tau) = 0$$

$$\tilde{W}_{10}(\sigma,\tau) = \sigma + 2H_1^*$$
$$\tilde{W}_{11}(\sigma,\tau) = \tau - 2\tau H_0^* + H_2^* \qquad (3.16)'$$
$$\tilde{W}_{12}(\sigma,\tau) = \tau^2 H_{-1}^* - \sigma\tau + (\sigma^2/2) H_1^* - 2H_3^*$$

$$\tilde{W}_{20}(\sigma,\tau) = \sigma^2/2 + \tau$$
$$\tilde{W}_{21}(\sigma,\tau) = \sigma\tau + 2H_3^* \qquad (3.17)'$$
$$\tilde{W}_{22}(\sigma,\tau) = -\sigma^2\tau - \tau^2/2 + 3H_4^* - 2\tau H_2^*$$

when the zero flux condition holds at $R = R_o$.

IV. COMPUTATIONAL ALGORITHM. The preceeding expansions can be incorporated into a numerical algorithm which effectively computes the heat transfer over longer durations. To describe this in more detail let us consider the original formulation, equations (1.2) - (1.5). We suppose that a numerical mesh is constructed as indicated in Figure 1, that values of U are known at the nodal points at time $T = T_o$, denoted U_0, U_1, U_2, U_N, and that $G_0 = G(T_o)$ and $G_1 = G(T_o + \Delta T)$ are known. The values of U at the new time $T = T_o + \Delta T$ will be denoted U_0', U_1', ... U_N'.

We first consider how to obtain U_0'. At time T_o the variation in U near $R = R_o$ can be approximated by a power series

$$U(R,T_o) \approx a_0 + a_1\left(\frac{R-R_o}{R_o}\right) + \frac{a_2}{2}\left(\frac{R-R_o}{R_o}\right)^2$$

where
$$a_0 = U(R_o,T_o)$$
$$a_1 = R_o U_R(R_o,T_o)$$
$$a_2 = R_o^2 U_{RR}(R_o,T_o) .$$

These coefficients can be determined from the knowledge of U_0, U_1 and G_0 in conjunction with boundary condition (1.4); we have

$$a_0 = U_0$$
$$a_1 = (H R_o/K)[U_0 - G_0]$$
$$a_2 = 2R_o^2[U_1 - U_0 - a_1 \Delta R]/\Delta R$$

where ΔR is the distance between the first two nodal points. We can similarly approximate G(T) by

$$G(T) \approx G_0 + (G_1 - G_0)(T - T_o)/\Delta T \ .$$

If we were now to introduce a shifted time variable

$$T' = T - T_o$$

with $0 \leq T' \leq \Delta T$ then the analysis of the last two sections would carry over identically with T' in place of T and ΔT in place of T_1. Using the approximate formulas developed in Section III we would thus obtain

$$U_0' = U\ (R_o,\ \Delta T)$$

$$\approx [a_0 W_0 + a_1 W_1 + a_2 W_2 + G_0 V_0 + (G_1 - G_0) V_1]\ (R_o,\ \Delta T)$$

or

$$U_0' \approx [a_0 \tilde{W}_0 + a_1 \tilde{W}_1 + a_2 \tilde{W}_2 + G_0 \tilde{V}_0 + (G_1 - G_0) \tilde{V}_1]\ (0,1) \qquad (4.1)$$

where $\tilde{W}_n\ (0,1)$ and $\tilde{V}_n\ (0,1)$ are evaluated from (3.21) and (3.22) using

$$\varepsilon = \sqrt{k\ \Delta T}/R_o \text{ and } h = H\ \sqrt{\Delta T/K\rho c}\ .$$

A similar discussion using the DESS method can be used to develop an approximation for U at the exterior boundary $R = R_1$ where the zero flux condition (1.5) applies. (This is not really necessary for the gun barrel problem since no thermal boundary layers are observed in this region. However the derivations are still valid). This results in the following approximate solution

$$U_N' = U\ (R_1,\ T_o + \Delta T)$$

$$\approx U_N + a\ \varepsilon^2\ \left(1 + \frac{4}{3\sqrt{\pi}}\ \varepsilon - \frac{1}{4}\ \varepsilon^2 + \ldots\right) \qquad (4.2)$$

where $a = 2\ R_1^2\ (U_{N-1} - U_N)/\Delta R^2$

$$\varepsilon = \sqrt{k\Delta T}/R_1$$

and ΔR is the distance between the last two nodal points.

Once values of U_0' and U_N' have been determined the problem can be regarded as one with Dirichlet boundary conditions and the solution advanced at interior nodes using any of the standard explicit or implicit

algorithms for parabolic differential equations. We have incorporated this general approach into a computer code written in BASIC for the Hewlett Packard 9845A desk top computer. Some of the resulting calculations will be given in the next section after the introduction of a simple model of the propellant thermodynamics. However, as mentioned earlier, the algorithm which we have just described can also be coupled with a sophisticated interior ballistics code which provides update, values of H and θ_{gas} (T) at each time step of the calculation.

V. A SIMPLE OVER-ALL MODEL. The coefficient of heat transfer H and the propellant gas temperature θ_{gas} (T) appearing in (1.4) are poorly understood physical parameters which may vary considerably during a single firing cycle. Nevertheless we can consider the following simplistic model.

$$H = \begin{cases} H_o & 0 \leq T \leq T_e \\ 0 & T > T_e \end{cases} \tag{5.1}$$

$$F(R) = 0 \qquad R_o \leq R \leq R_1 \tag{5.2}$$

$$G(T) = G_0 \qquad T \geq 0 \tag{5.3}$$

where H_o is a constant and T_e denotes the exposure time after which the heat transfer becomes negligible and (1.4) can be replaced by the zero flux condition. The constant G_0 may be taken as

$$G_0 = (\theta_{flame} - \theta_o)/(\theta_M - \theta_o)$$

where θ_{flame} is the adiabatic flame temperature for the propellant. The most appropriate values for H_o and T_e are not at all clear from currently available physical theories. However we can treat these as adjustable parameters which can be chosen for best agreement with experimental data. If this is done for the 37mm gun studied in Reference 5, for example, we obtain

$$H_o \approx .28 \text{ cal/sec (cm)}^2 \text{ °C}$$

$$T_e \approx .018 \text{ sec.}$$

(These values should be considered preliminary since the data was not available in tabular form and best agreement was obtained by visual comparison with the figures in Reference 5. There is clearly some distortion in the reproduction process and it is suspected that H_o should be somewhat larger and T_e smaller.) The resulting claculations using our numerical

scheme are plotted in Figure 2; run time was approximately one minute. The actual measured data is shown in Figure 2 and shows excellent agreement. Quite similar agreement has been obtained for the 5.56mm and 20mm guns also studied in Reference 5. This would seem to indicate both that our simple model forms a reasonable first approximation to the actual heat transfer process and that our numerical scheme is working properly. It is the author's intention to pursue this matter in more detail in future work so that greater confidence can be gained. For the moment these comparisons are only qualitative.

To simplify things even more one can dispense with the numerical scheme altogether when the exposure time satisfies

$$T_e << R_o^2/k$$

and approximate the temperature rise using only the single term V_0, that is

$$\begin{aligned} U(R,T) &= G_0 V_0 (R,T) \\ &\approx G_0 [\tilde{V}_{00} + \varepsilon \tilde{V}_{01} + \varepsilon^2 \tilde{V}_{02}] (\sigma,1) \end{aligned} \tag{5.4}$$

where formulas (3.20) are used for ε, σ and h. In particular, along the bore surface

$$U(R_o,T) \approx G_0 [\tilde{V}_{00}(0,1) + \varepsilon \tilde{V}_{01}(0,1) + \varepsilon^2 \tilde{V}_{02}(0,1)]$$

for $0 \leq T \leq T_e$. To approximate U for $T > T_e$ note that the variation of U in the R direction at time $T = T_e$ is approximated by

$$U(R,T_e) \approx a_0 + a_1 ((R-R_o)/R_o) + \frac{a_2}{2}((R-R_o)/R_o)^2 \tag{5.5}$$

where a, b and c are obtained by differentiation of (5.4) with respect to R. Explicity we can show that

$$a_0 = G_0 [2 Z_o^{\#} - 2 \varepsilon Z_{-1}^{\#} + \varepsilon Z_1^{\#} + \varepsilon^2 Z_{-2}^{\#}] (0,1)$$

$$a_1 = G_0 [-(2/\varepsilon) Z_{-1}^{\#} + 2 Z_{-2}^{\#} - Z_0^{\#} - \varepsilon Z_{-3}^{\#}] (0,1)$$

$$a_2 = G_0 [(2/\varepsilon^2) Z_{-2}^{\#} - (2/\varepsilon) Z_{-3}^{\#} + (1/\varepsilon) Z_{-1}^{\#} + Z_{-4}^{\#} + Z_0^{\#}] (0,1)$$

where $\varepsilon = \sqrt{k T_e}/R_o$.

Using (5.5) for initial values beginning at time $T = T_e$, assuming a zero flux condition we then have, for $T > T_e$,

$$U(R_o, T) \approx a_0 + a_1 \bar{\varepsilon} [\tilde{W}_{10} + \bar{\varepsilon} \tilde{W}_{11} + \bar{\varepsilon}^2 \tilde{W}_{12}] (0,1)$$
$$+ a_2 \bar{\varepsilon}^2 [\tilde{W}_{20} + \bar{\varepsilon} \tilde{W}_{21} + \bar{\varepsilon}^2 \tilde{W}_{22}] (0,1) \qquad (5.6)$$

where equations (3.16)' and (3.17)' are used to evaluate $\tilde{W}_{nk}$ and

$$\bar{\varepsilon} = \sqrt{k\,(T-T_e)}/R_o \ .$$

Unfortunately (5.6) breaks down quickly because of the inaccuracy in (5.5).

VI. SUMMARY AND CONCLUSIONS. We have used a small parameter analysis to derive an algorithm which can be used to generate numerical solutions for problems (1.2) - (1.5). This algorithm was then used in conjunction with the simplified physical assumptions (5.1) and (5.3) to produce temperature profiles quite similar to experimentally measured temperatures.

An overall predictive model for the likely temperature profiles in a gun barrel is given by formulas (5.4) and (5.6). It involves two adjustable parameters H_o (average heat transfer coefficient) and T_e (effective duration of exposure to hot gasses). It is hoped that this model will provide a useful tool for weapons designers who only need a general qualitative understanding of the gun barrel temperature response. By building up a data base of typical values of H_o and T_e for existing guns and propellants it might be feasible to extrapolate to the expected thermal behavior of proposed weapons systems.

REFERENCES

1. J. F. Polk, "Asymptotic Expansions for the Solutions of Parabolic Differential Equation with a Small Parameter," Ph. D. Dissertation, Department of Mathematics, University of Delaware, Newark, Delaware, 1979.

2. J. F. Polk, "Diffusion Equation Solution Sequences," USAARRADCOM, Ballistic Research Laboratory Technical Report, in preparation.

3. J. F. Polk, "Special Function Solutions of the Diffusion Equation," USAARRADCOM, Ballistic Research Laboratory Technical Report 02182, July 1979.

4. J. F. Polk, "Exact Solutions for Convective Heat Transfer," USAARRADCOM, Ballistic Research Laboratory Technical Report 02186, August 1979.

5. T. L. Brosseau, "An Experimental Method for Accurately Determining the Temperature Distribution and the Heat Transferred in Gun Barrels," USA Ballistic Research Laboratories Report 1740, September 1974.

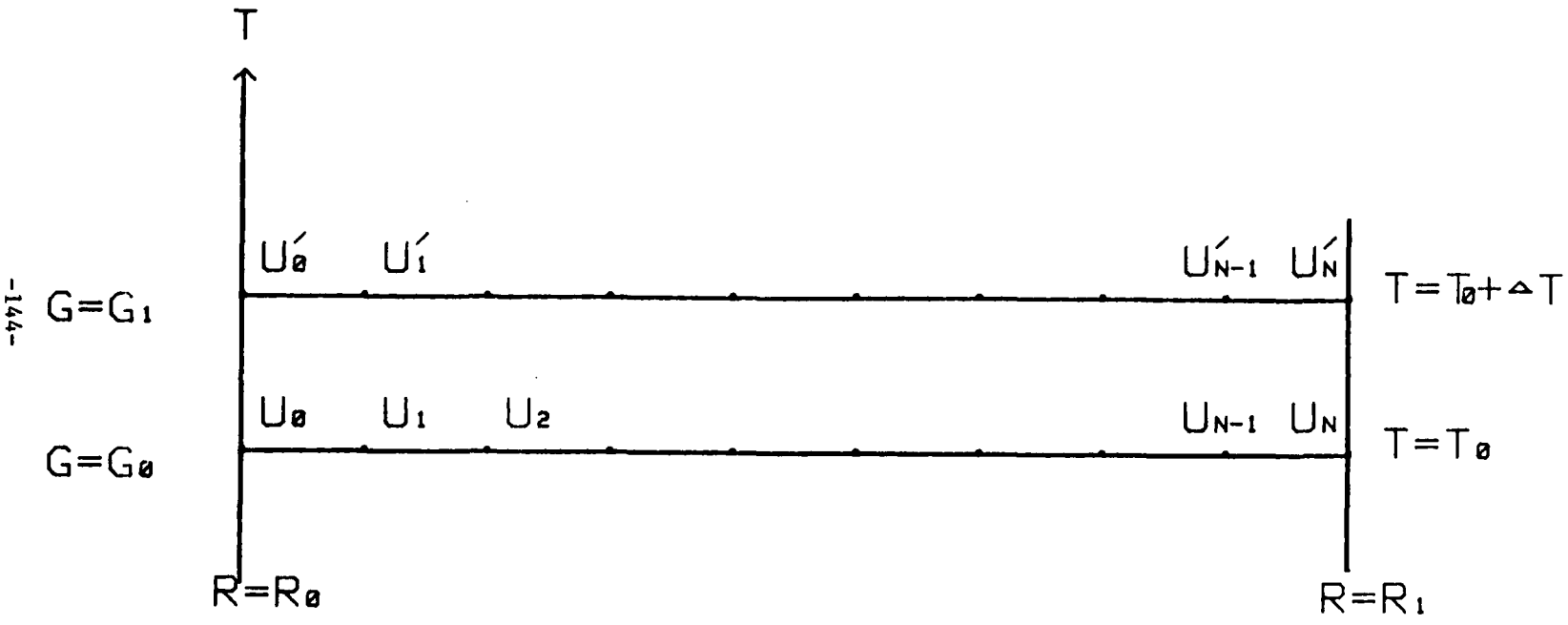


Figure 1. Numerical mesh for solving problem (1.2) - (1.5).

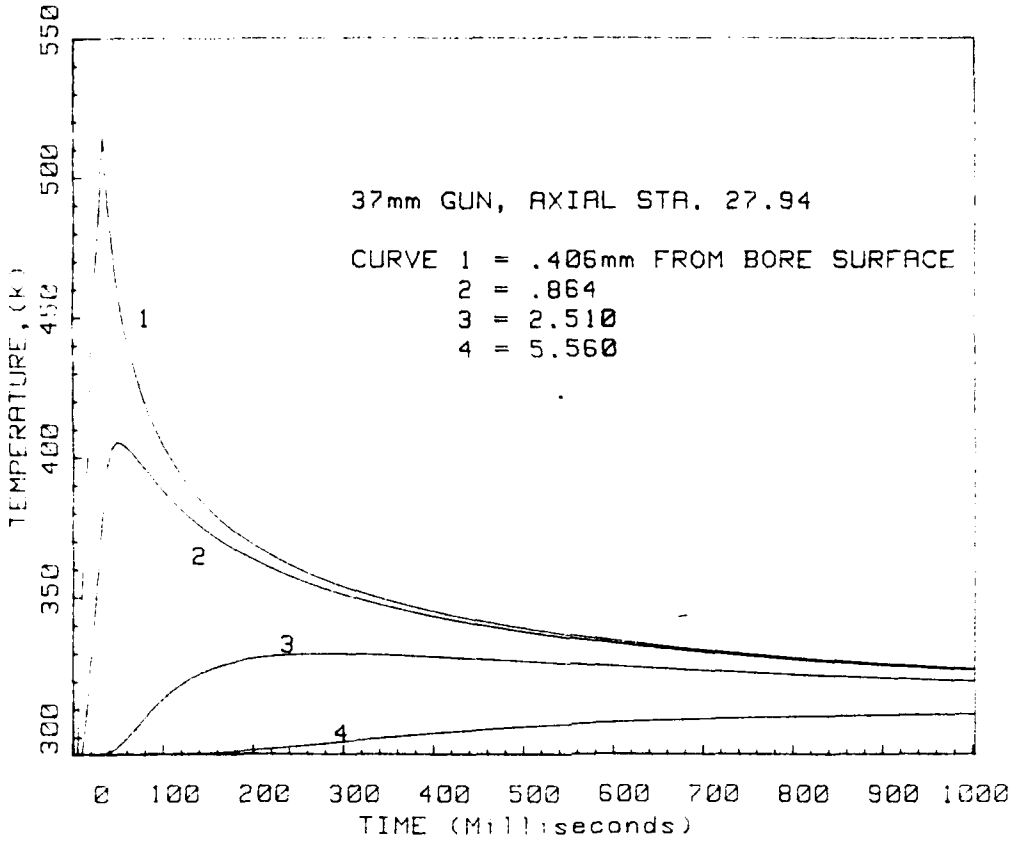


Figure 2. Calculated temperature/time profiles at several distances from bore surface of 37mm gun barrel.

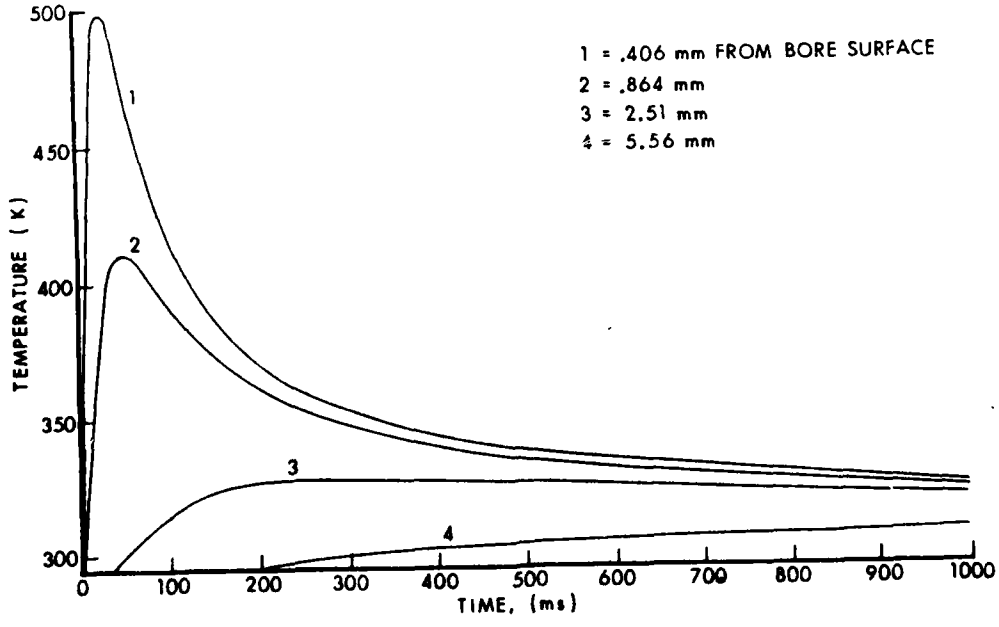


Figure 3. Experimentally measured temperature/time curves from BRL Report No. 1740.

DYNAMICS OF IGNITION

A. K. Kapila
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, New York 12181

ABSTRACT. Activation-energy asymptotics is employed to determine the complete burning history, from ignition to deflagration, of a premixed combustible in a spatially inhomogeneous configuration. The sequence of events consists of a benign induction period, followed by the rapid development and growth of a hot spot. When the entire reactant within the hot spot is consumed, the latter transforms into a practically steady deflagrating wave travelling across the vessel.

I. Introduction. Ignition and subsequent burning of a premixed combustible in a confined space is a complex process. Even for the simple case of a homogeneous, constant-property gas mixture, the mathematical problem is a difficult one, primarily due to the strong coupling between chemistry and gas dynamics. Matters are compounded still further in any realistic situation, such as combustion in a gun barrel.

In an attempt to develop appropriate mathematical techniques, this paper takes a first step by treating an extremely idealized model, where the combustible is assumed to have negligible thermal expansion. This assumption removes gas dynamics from the scene, and reduces the problem to a purely reactive-diffusive one. Large activation energy asymptotics is used to trace the complete burning history of the system. The analysis is the spatially varying counterpart of Kassoy's treatment [1] of the "lumped" version of the problem. It is envisaged that the notions developed here can be extended to include gas-dynamic effects.

A more detailed treatment of this presentation can be found in [2].

II. Formulation. Let a cold combustible mixture, at initially uniform temperature and reactant concentration, be confined to the region between the planes $x = \pm 1$. Let the boundaries of the region be maintained at the initial levels of temperature and concentration for $t>0$. (Thus, heat, fresh mixture and products of combustion are allowed to cross the boundaries.) Taking the Lewis number to be unity and invoking symmetry about $x = 0$, the mathematical problem to be considered is

$$z = (1+\beta-y)/\beta, \tag{1}$$

$$y_t = y_{xx} + \{D/(\beta\gamma)\}(1+\beta-y)\exp(\gamma-\gamma/y), \quad 0<x<1, \quad t>0, \tag{2}$$

$$y_x(0,t) = 0, \quad y(1,t) = 1, \tag{3}$$

$$y(x,0) = 1. \tag{4}$$

This dimensionless system describes a single, one-step Arrhenius reaction (Fuel + Oxidant → Product). Here, y is the temperature and z the concentration of a reactant (say, fuel), while β is the chemical heat release, γ the activation energy and D the Damkohler number. It is assumed that

$$D > 0.878,$$

which assures that the system is potentially explosive [3]. Henceforth, we shall treat eqns. (2-4) for y; z is then given by (1). The object is to determine how the solution evolves in time. The analysis will be based on the asymptotic limit $\gamma\to\infty$.

III. Induction Stage. Equation (2) and the initial condition (4) suggest that in the beginning, $y - 1 = O(\gamma^{-1})$. Therefore we employ the expansion

$$y = 1 + \gamma^{-1}y_1 + \ldots \tag{5}$$

which, to leading order, yields the reduced problem

$$\left.\begin{aligned} y_{1_t} &= y_{1_{xx}} + De^{y_1}, \quad 0<x<1, \quad t>0, \\ y_{1_x}(0,t) &= y_1(1,t) = y_1(x,0) = 0. \end{aligned}\right\} \tag{6}$$

This problem was solved numerically and a typical solution is displayed in Fig. 1. Initially the solution develops gradually, but then the temperature near x=0 begins to rise rapidly, while changes are more leisurely elsewhere. Eventually, at a definite time $t^\infty(D)$, $y_1(0,t)$ becomes unbounded. It is found that t^∞ falls off with increasing D, i.e. higher Damkohler numbers cause the system to explode sooner.

The singularity at x=0, as $t\to t^\infty$, can be examined analyti-

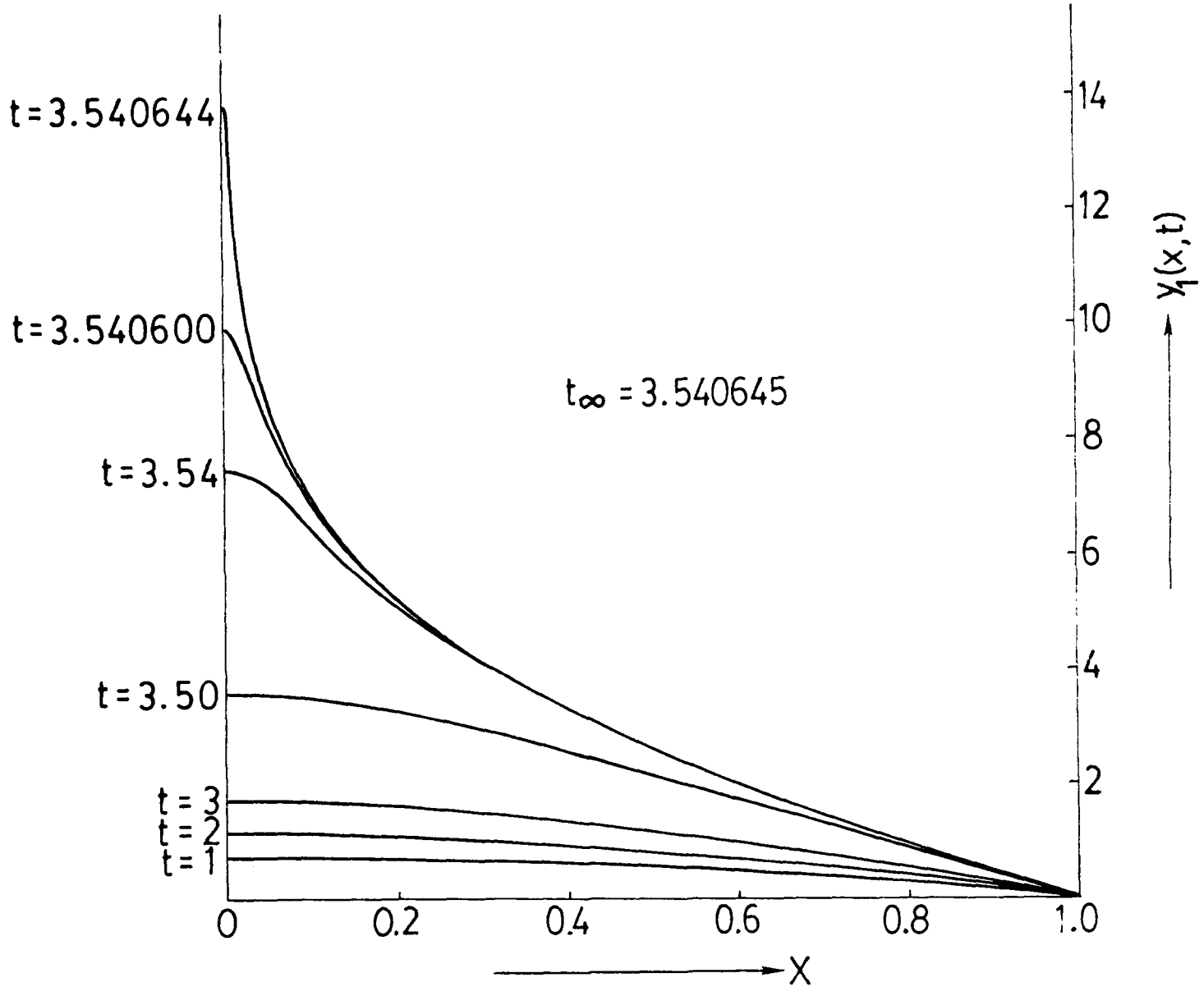


Fig. 1. Induction solution for D = 1.

cally by developing a boundary-layer expansion in variables τ and η, which are defined by

$$\tau = t_\infty - t, \quad \eta = x/\sqrt{\tau}.$$

These reduce the differential equation in (6) to

$$y_{1\eta\eta} - (1/2)\eta y_{1\eta} + y_{1\tau} + \exp(y_1) = 0$$

where y is now a function of η and τ. In the limit $\tau \to 0$, η fixed, this equation describes a boundary layer $O(\sqrt{\tau})$ thick. In this layer y_1 is seen to have the expansion

$$y_1(\eta,\tau) = -\ell n(D\tau) + f_0(\eta) + O(\tau), \quad \tau \to 0, \tag{7}$$

where f_0 is found to satisfy

$$f_0'' - (1/2)\, f_0' + \exp(f_0) - 1 = 0, \quad 0<\eta<\infty,$$

$$f_0'(0) = 0, \quad f_0 = -2\, \ell n\eta + A_0(D) + o(1) \text{ as } \eta \to \infty.$$

The left boundary condition on f_0 is due to symmetry, and the right boundary condition comes from matching with the numerical solution outside the boundary layer. The problem for f_0 can be solved numerically, and yields a monotonically decreasing function.

IV. Explosion Stage. As $\tau \to 0$ the boundary layer solution (7) grows logarithmically, eventually causing the induction-period solution (5) to break down. Further development occurs on an exponentially rapid time scale σ, defined by

$$D\tau = \exp(-\gamma\sigma), \quad \sigma>0 \text{ and } O(1).$$

In terms of the boundary-layer variables η and σ, the full differential equation (2) reduces to

$$y_\sigma = \gamma y_{\eta\eta} - \gamma(\eta/2)y_\eta + \beta^{-1}(1+\beta-y)\exp[\gamma(1-\sigma-y^{-1})].$$

Its solution can be shown to have the expansion

$$y = (1-\sigma)^{-1} + \gamma^{-1}(1-\sigma)^{-2}[f_0(\eta) - \ell n\{(1-\sigma)^2(1+\beta)-(1-\sigma)\}$$

$$+\ell n\beta] + \ldots . \qquad (8)$$

The above expansion justifies the term explosion for this stage, because it shows that the solution undergoes an O(1) change during a period of utmost brevity in the original time variable t. However, this change is confined to an ultra-thin boundary layer at x=0, outside which the system remains essentially stationary at $t=t^{\infty}$. In other words the outer region, governed by the diffusion time t, is incapable of responding to the fast time σ.

The growth of the hot spot (boundary layer) lasts until $\sigma=\beta/(1+\beta)$. Then, the second term of (8) becomes singular and in the boundary layer, y approaches its maximum value $1+\beta$, indicating that the reactant is completely exhausted (see (1)). The hot spot now transforms into a thin zone of reaction which begins to move into the interior of the region.

V. Propagation Stage. The moving reaction zone (i.e. flame or deflagration wave) is surrounded by a slightly thicker "envelope" which, in turn, separates a burnt region behind the flame from a cold region ahead of the flame (see Fig. 2). The portion of the envelope ahead of the reaction zone is a preheat region, where inert heating brings the cold mixture up to the flame temperature $1+\beta$.

It is convenient to shift to a coordinate system in which the flame is stationary, i.e. we let

$$x = x_0(t) + (\delta/\varepsilon)\zeta$$

where x_0 is the flame location (considered to be an O(1) quantity), ζ is the spatial coordinate in the envelope and ε,δ are small parameters defined by

$$\delta = (\beta\gamma/D)^{1/2}\exp[-\beta\gamma/(2+2\beta)], \qquad (9)$$

$$\varepsilon = (1+\beta)^2/\gamma;\ \delta/\varepsilon << 1. \qquad (10)$$

With t scaled via

$$t = t^{\infty} + (\delta/\varepsilon)s,$$

the flame speed is found to be

$$dx_0/dt = (\varepsilon/\delta)dx_0/ds = (\varepsilon/\delta)U(s), \text{ say,}$$

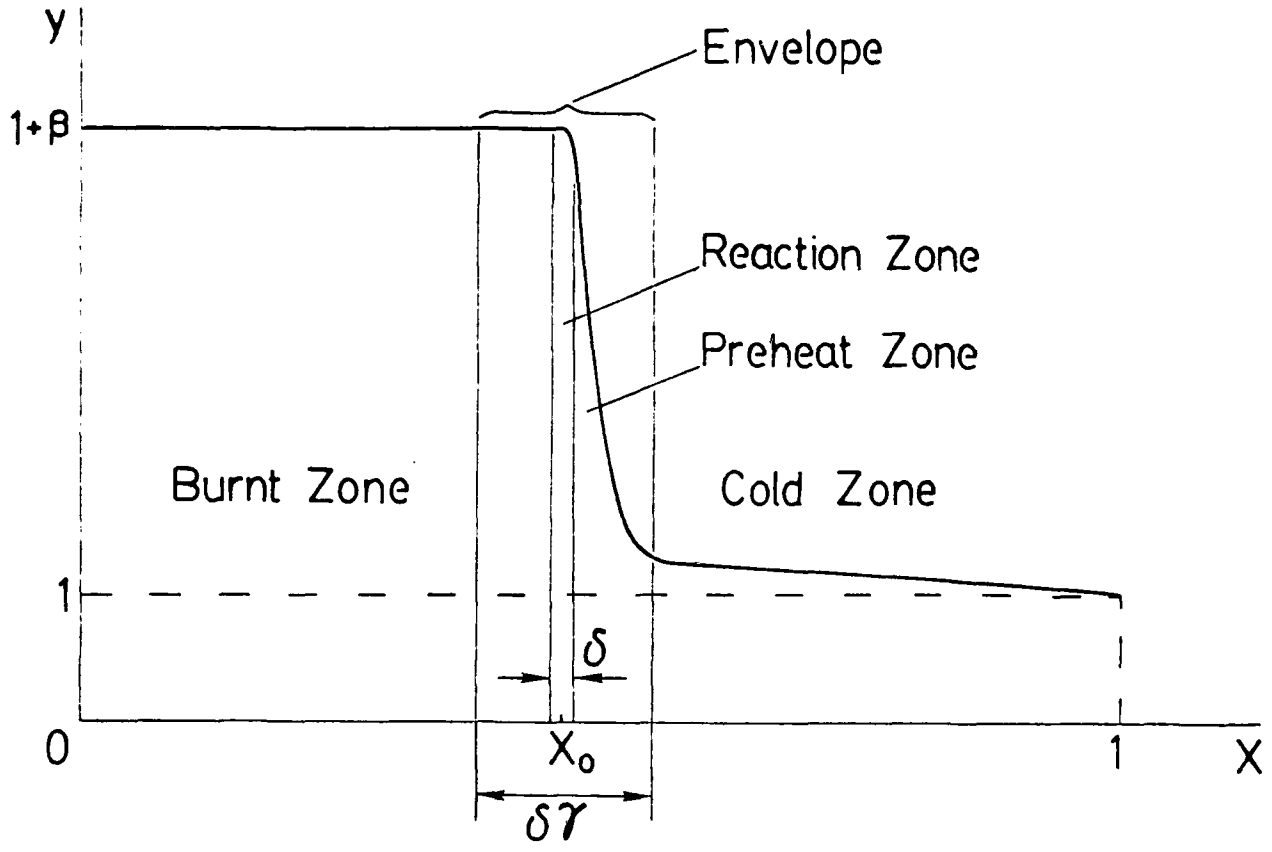


Fig. 2. Temperature profile during the propagation stage.

where U is taken to be O(1). Thus the flame travels across the region at an exponentially rapid rate. In the ζ,s variables, the full equation (2) transforms into

$$(\delta/\varepsilon)y_s = y_{\zeta\zeta} + Uy_\zeta + \varepsilon^{-2}(1+\beta-y)\exp[\gamma/(1+\beta)-\gamma/y],$$

$$-\infty<\zeta<\infty,\ s>0.$$

To leading order, this equation is steady. It can be shown that the solutions in the various regions of Fig. 2 are given by the following expressions:

$$\text{Burnt region } (0<x<x_0):\quad y = 1+\beta; \tag{11}$$

$$\text{Envelope:}\quad y = 1+\beta \text{ for } \zeta<0,\quad y = A(s)+[1+\beta-A(s)]\exp[-U\zeta] \text{ for } \zeta>0; \tag{12}$$

$$\text{Cold region } (x_0<x<1):\quad y = 1+\gamma^{-1}y_1(x,t^\infty)+\ldots . \tag{13}$$

Here, A(s) and U(s) are still to be determined. We note from (13) that in the cold region ahead of the flame, the solution is still frozen at its value at the end of the induction period. Also, (12) shows that the two solutions in the envelope have unequal slopes at $\zeta=0$. This discontinuity is smoothed out by the classical Bush-Fendell flame sheet [4], of thickness δ, located at $\zeta=0$.

Matching between the various spatial regimes ultimately leads to

$$A(s) = 1+\gamma^{-1}y_1(x_0,t^\infty)+\ldots$$

and the expression for the scaled flame speed

$$U(s)= \sqrt{2}/[1+\beta-A(s)] = \sqrt{2}/\beta+O(\gamma^{-1}).$$

Thus the flame speed is essentially constant, and has precisely the same value that it would have in an unconfined plane geometry

V . Transition and Relaxation Stages. The propagation-stage analysis assumes that the flame is away from both x=0 and x=1; it breaks down when the flame approaches either of these locations. When x_0 is within an $O(\delta)$ distance of x=0, the boundary condition at x=0 intrudes into the flame zone. The flame structure then is no longer of the steady Bush-

Fendell variety; rather, the balance in the flame is unsteady-diffusive-reactive. The interested reader can find the details of this transition stage in [2]. Suffice it to say that the corresponding problem can only be solved numerically, and that the solution describes the manner in which the hot spot, generated during the explosion stage, transforms into a moving reaction front.

The propagation stage also becomes invalid when the flame has moved too far to the right, because the boundary x=1 then enters the preheat zone. Once again we omit the details except to state that the preheat zone now becomes time-dependent (while the flame is still steady), and causes the reaction zone to come to rest. This relaxation occurs at a time scale that is small compared to the propagation time s. The ultimate location of the flame is given by

$$x_0 = 1 - (\delta/\varepsilon)\beta/\sqrt{2}$$

where δ and ε are the same as before. The reason for the existence of the steady flame is the boundary condition

$$y(1,t) = z(1,t) = 1, \quad t>0$$

which provides for a continuous supply of cold, fresh reactant at the walls of the vessel. No-flux boundary conditions would lead to an ultimately fully-burnt state.

VII. Final Remarks. Through large activation-energy asymptotics, the analysis presented above succeeds in describing the complete combustion history of a confined mixture, from an initial cold state to a final deflagrating state. In partcular, the time scale of each stage of evolution is identified and its structure determined. It is envisaged that investigations of fluid-dynamical effects in the fully-coupled problem of confined gaseous combustion should be feasible in this framework.

REFERENCES

1. D. R. Kassoy, Extremely rapid transient phenomena in combustion, ignition and explosion, SIAM-AMS Proceedings 10 (1976), pp. 61-72.
2. A. K. Kapila, Reactive-diffusive system with Arrhenius kinetics: dynamics of ignition, to appear.
3. A. K. Kapila and B. J. Matkowsky, Reactive-diffusive system with Arrhenius kinetics: multiple solutions, ignition and extinction, SIAM J. Appl. Math. 36 (1979), pp. 373-389.
4. W.B. Bush and F.E. Fendell, Asymptotic analysis of laminar flame propagation for general Lewis numbers, Comb. Sci. Technol. 1 (1970), pp. 421-429.

PREMIXED CYLINDRICAL FLAMES*

G.S.S. Ludford
Center for Applied Mathematics
Cornell University, Ithaca NY 14853

ABSTRACT. Although it has not been studied to the same extent as the plane premixed flame [cf. Ludford, J. Mécanique 16 (1977), 553], the cylindrical flame is almost as easy to produce. The reacting mixture is supplied through the surface of a circular cylinder and is induced to flow radially by means of sufficiently close end plates. The flame then forms a coaxial cylinder and can be observed through the end plates, which should be transparent and good thermal insulators.

Analytically the cylindrical flame stands between the plane and spherical flames. The structure of its reaction zone is the same as that of a plane flame but, like the spherical flame, it does not exhibit the cold-boundary difficulty (loc. cit.) since the mixture must be introduced at a finite radius.

The present paper will show, on the basis of activation-energy asymptotics, how cylindrical geometry modifies a premixed flame. For simplicity we shall consider a single reactant (monopropellant), which decomposes in one step irreversibly.

I. INTRODUCTION.

Although it has not been studied to the same extent as the plane premixed flame, the cylindrical flame is in principle almost as easy to produce. The reacting mixture is supplied through the surface of a circular cylinder and is induced to flow radially by means of sufficiently close end plates. The flame then forms a coaxial cylinder and can be observed through the end plates, which should be transparent and good thermal insulators.

Analytically the cylindrical flame stands between the plane and spherical flames. The structure of its reaction zone is the same as that of the plane flame, with temperature constant beyond; so that there is no curvature effect as for the spherical flame (cf. Ludford, Yannitell & Buckmaster 1976). On the other hand, like the spherical flame it does not exhibit the cold-boundary difficulty: the mixture must be introduced at a finite radius, which can however be so small that a line source is effectively formed. Ironically enough, in their attempt to treat curved flames Spalding & Jain (1959) use

*This work was supported by the U.S. Army Research Office under Contract No. DAAG29-79-C-0121

plane-flame analysis on the spherical flame, where it is never valid, and neglect the cylindrical flame, where it is always valid.

The object of the present paper is to show how cylindrical geometry modifies a premixed flame. For simplicity we shall consider a simple reactant (monopropellant), which decomposes in one step irreversibly.

II. THE BASIC CYLINDRICAL FLAME

The notation will be that used by Ludford (1977a), who gives a derivation of the equations with which we have to deal. Non-dimensionalization is based on the radius, a, of the supply cylinder. If v is the radial velocity, the equation of continuity admits $r\rho v$ being constant. When M is the mass flux at the supply cylinder we may therefore write

$$r\rho v = M \quad \text{or} \quad \rho v = M/r, \tag{1}$$

the latter giving the mass flux at every other radius r. Once the temperature T has been determined, the density $\rho = 1/T$ and $v = MT/r$ follow immediately, while variations in pressure about its constant level can be calculated from the momentum equation of the mixture. There remain then the energy and reactant species equations

$$\mathcal{L}(Y) = -\mathcal{L}(T) = DY \exp(-\theta/T), \tag{2}$$

where Y is the mass fraction of reactant and D is the Damköhler number (which may be taken constant). Here

$$\mathcal{L} \equiv \frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right) - \frac{M}{r}\frac{d}{dr} \tag{3}$$

in the cylindrical geometry; for simplicity we have assumed unit Lewis number and a first-order decomposition. As is now common, the equations will be solved in the limit where the activation energy θ tends to infinity.

First note that the Shvab-Zeldovich variable $\tilde{Y} = Y + T$ satisfies

$$\mathcal{L}(\tilde{Y}) = 0 \quad \text{for} \quad 1 < r < \infty. \tag{4}$$

Since the only solutions which remain bounded at infinity are constants, we have

$$Y + T = T_\infty = Y_s + T_s \tag{5}$$

for a reaction that goes to completion, where s denotes conditions at the supply $r = 1$. The fact that the Y, T-relation is identical to that for plane

flames (Ludford 1977b) is responsible for the structures of the reaction zones being the same.

The asymptotic analysis of the equations (2) proceeds as for a plane flame. The temperature beyond the flame sheet must be constant, i.e.

(6) $\quad Y = 0,\ T = T_\infty \quad \text{for} \quad r > r_*,$

while up to the flame sheet the reaction is frozen, i.e. $\mathcal{L}(Y) = \mathcal{L}(T) = 0$ so that

(7) $\quad Y = Y_s + L(1 - r^M),\ T = T_s + L(r^M - 1) \quad \text{for} \quad 1 < r < r_*.$

Here $L = M^{-1}T_s'$, with T_s' the temperature gradient at the supply, is the heat conducted back into the supply per unit mass of mixture. These two pairs of formulas give the same values at

(8) $\quad r_* = [1 + (T_\infty - T_s)/L]^{1/M};$

as expected, the stand-off distance for a plane flame is recovered as the radius a of the supply cylinder tends to infinity when due attention is paid to a mass-flux unit proportional to 1/a. Consistency requires $r_* > 1$, i.e.

(9) $\quad L > 0,$

which means the supply must be a conductive heat sink.

As usual the interior of the flame is investigated with the expansions

(10) $\quad Y = \delta y(\xi) + o(\delta),\quad T = T_\infty + \delta t(\xi) + o(\delta)$

where

(11) $\quad \xi = (r - r_*)/\delta \quad \text{and} \quad \delta = T_\infty^2/\theta.$

The structure is thereby found to satisfy the equation

(12) $\quad d^2t/d\xi^2 = -\tilde{D}\, y\, e^t, \quad \text{where} \quad y + t = 0,$

and the boundary conditions

(13) $\quad t = MJ_s\xi/r_* + o(1) \quad \text{as} \quad \xi \to -\infty\ , \quad t = o(1) \quad \text{as} \quad \xi \to +\infty,$

which come from matching with the expressions (6) and (7) outside. Here

(14) $\quad \tilde{D} = \delta^2 e^{-\theta/T_\infty} D \quad \text{while} \quad J_s = Y_s - Y_s'/M$

is the reactant flux fraction $Y - rY'/M$ at the supply (usually 1). Exactly the same problem is obtained for a plane flame except that

the coefficient MJ_s/r_* in the condition (13a) is replaced by MJ_* (which equals MJ_s there). Noting that here $J_* = J_s/r_*$ (because the total reactant flux $2\pi rMJ$ is conserved for frozen chemistry) shows that the cylindrical reaction zone is locally plane. For spherical flames the y,t-relation (12b), which derives from the Y,T-relation (5), is changed so that the reaction zone is quite different from its plane counterpart.

From the solution of the corresponding plane-flame problem we deduce

$$\tilde{D} = (J_s/r_*)^2 M^2/2 \tag{15}$$

so that

$$D = (J_s^2\theta^2 e^{\theta/T_\infty}/2T_\infty^4 r_*^2)M^2, \tag{16}$$

which is the required M,D-relation. For fixed ML, i.e. heat conducted back to the supply, it has the general shape of the parabola obtained for the plane flame, because the factor

$$1/r_*^2 = [1 + (T_\infty - T_s)/L]^{-2/M} \tag{17}$$

varies only between $\exp[-2(T_\infty - T_s)/ML]$ and 1 as M increases from 0 to ∞. The parabola is useful for determining the speed (i.e.M) with which a plane flame propagates into fresh mixture at given pressure (i.e. D), but there is no equivalent use here. On the other hand, for the set-up envisaged in the opening paragraph both M and D are prescribed (along with T_s and J_s), and the formula determines the final temperature T_∞ (note $L = J_s + T_s - T_\infty$).

III. NEAR-SURFACE AND SURFACE FLAMES. REMOTE FLAMES.

For the above solution to be valid the parameter values must be such that r_*, as given by the formula (8), lies between 1 and ∞. As for the plane flame (Ludford 1977b) three limiting cases arise, two of which are essentially the same as there. The third leads to an interesting new phenomenon.

When M and D become large, with all other parameters held fixed, the flame approaches the surface. An intense convective-diffusive zone of thickness $O(M^{-1})$ forms near the surface, bounded by a reaction zone whose thickness is $O(\theta^{-1}M^{-1})$.

By contrast a true surface flame can be produced for any M by adjusting the pressure so as to make $T_\infty \to T_s$. Similarly remote flames can be produced

by making $T_\infty \to J_s + T_s$ (i.e. $L \to 0$). At both extremes the preceding analysis becomes invalid: either the boundary intrudes into the reaction zone and there is no frozen region between them or the isothermal limit of (7) is not uniformly valid in the unbounded frozen region. In either case the asymptotics must be reworked.

The analysis of the surface flame is identical to that for the plane case (Ludford 1977b) provided x is changed to $r-1$. We conclude that $\tilde{D}$ will change from $J_s^2 M^2/2$, the value (15) when $r_* = 1$, to ∞ as the temperature difference $(T_\infty - T_s)/\delta$, measured on the δ-scale, decreases from ∞ to 0.

By contrast, the remote flame cannot be treated as in the plane case since the asymptotic analysis breaks down earlier, in fact as soon as L becomes $O(\delta)$. The difference lies in the reactant flux $M J_s/r_*$ at the flame, which now tends to zero like $\delta^{1/M}$ as $r_* \to \infty$; the condition (13a) loses its effectiveness unless a different scale is used for the structure. Setting

$$r = r_* + \delta^{1-1/M}\xi \tag{18}$$

gives the new condition

$$t = MJ_s(\ell/J_s)^{1/M}\xi + o(1) \quad \text{as } \xi \to -\infty. \tag{19}$$

where $L = \delta\ell$; the corresponding change

$$\tilde{D} = \delta^{2(1-1/M)}\, e^{-\theta/T_\infty} D \tag{20}$$

must also be made to keep the structure equation balanced.

We therefore end with the same problem, except that r_* is replaced by $(J_s/\ell)^{1/M}$; so that $\tilde{D}$ is given by the formula (15) with the same replacement and

$$D = J_s^{2(1-1/M)}\, \theta^{2(1-1/M)}\, e^{\theta/T_\infty} \ell^{2/M}/2T_\infty^{4(1-1/M)}. \tag{21}$$

Thus D varies from 0 to ∞ on the scale $\theta^{2(1-1/M)}\, e^{\theta/T_\infty}$ as ℓ increases from 0 to ∞, the upper end of the range corresponding to 0 on the previous scale $\theta^2 e^{\theta/T_\infty}$.

The most interesting feature is the spreading of the zone in which there is chemical activity, as M decreases. The transformation (18) implies that its

thickness is $O(\theta^{1/M-1})$, so that for $M \leq 1$ it is no longer a sheet; indeed for $M < 1$ it has infinite extent [remember $r_* = O(\theta^{1/M})$ is larger still]. Such a phenomenon should be easily observable.

IV. CONCLUDING REMARK.

We have seen that, except when remote, the cylindrical flame is locally plane, unlike the spherical flame. These results stem from the diffusion-convection operator

$$(22) \qquad \mathcal{L} \equiv \frac{1}{r^\alpha}\left[\frac{\partial}{\partial r}\left(r^\alpha \frac{\partial}{\partial r}\right) - M\frac{\partial}{\partial r}\right]$$

governing the reactionless field behind the flame. Here $\alpha = 0$ (plane, when $r = x$), 1 (cylindrical) or 2 (spherical) and

$$(23) \qquad \mathcal{L}(T) = 0$$

has the general solution

$$(24) \qquad T = \begin{cases} A + B\, e^{Mr} & (\alpha = 0), \\ A + B\, r^M & (\alpha = 1), \\ A + B\, e^{-M/r} & (\alpha = 2) \end{cases}$$

in the three cases. Boundedness of T makes $B = 0$ in the first two cases but not in the third, where T_∞ becomes an assignable parameter in addition to any others. It is this difference between the convection-diffusion process in plane and cylindrical geometries on the one hand and spherical geometry on the other which accounts for the similarities and differences of the corresponding flames.

REFERENCES.

Ludford, G.S.S. 1977a Combustion: basic equations and peculiar asymptotics. J. Mécanique 16, 531-551.

Ludford, G.S.S. 1977b The premixed plane flame. J. Mécanique 16, 553-573.

Ludford, G.S.S., Yannitell, D.W. & Buckmaster, J.D. 1976 The decomposition of a cold monopropellant in an inert atmosphere. Combust. Sci. Tech. 14, 133-145.

Spalding, D.B. & Jain, V.K. 1959 Theory of the burning of monopropellant droplets, Aero. Res. Council CP447, Tech. Rept. 20,176.

THERMOELASTIC STRESSES IN GUN BARRELS

Julian Davis
US Army Armament Research and Development Command
Dover, N. J. 07801

Yu Chen
Dept. of Mechanics and Materials Science
Rutgers, The State University of New Jersey
New Brunswick, N. J. 08903

ABSTRACT. This problem arose from an attempt to get a better understanding of the thermoelastic stress distribution in gun barrels during rapid fire due to the impact of hot propellant gases on the interior of the barrel. It is formulated as a coupled dynamic thermoelastic problem. The coupling between the thermal and elastic effects cannot be ignored due to the high rate of pressure and thermal inputs at the boundary. The mathematical model consists of a pair of partial differential equations for the stress and temperature distribution which is solved by a perturbation method. Solutions are discussed for particular cases.

I. INTRODUCTION. When a gun is under rapid fire, the hot propellant gases that build up in the barrel supply the boundary conditions of unsteady cyclic pressure and heat flux at the inner wall. The "radiation" boundary condition of heat flux (in the sense of Carslaw and Jaeger[1]p.18) along with the gas pressure produces a complex thermal stress field in the barrel. This field depends on the boundary and initial conditions (assumed homogeneous), the equations of motion, the coupled energy equations (which couples thermal and mechanical energy), and the Duhamel-Neumann constitutive equations. The coupling effect in the energy equation cannot be ignored due to the highly energetic and cyclic nature of the boundary conditions. In the usual thermal stress problems the rate of deformation is slow such that the thermal energy predominates the energy balance equation. As a consequence, the energy balance equation is just the thermal conduction equation from which the temperature distribution can be determined. This temperature distribution can then be introduced in the conservation equation of momentum as a body force derivable from a potential. Thus, the problem of solving the stress distribution can be treated independently without having to interact with the energy balance equation. This is the uncoupled case.

In this paper a general formulation of the coupled thermoelastic problem is given in terms of scalar and vector potentials which satisfy wave equations. The wave equation for the scalar potential is non-homogeneous-the non-homogeneous term

depending on the temperature. The problem is specialized to a cross section of a gun barrel with appropriate boundary conditions. It is then further specialized, for simplicity in illustrating the method, to a finite one-dimensional slab. Two equivalent formulations are given: (1) a fourth order p.d.e. for the stress which has both wave-like and diffusion-like properties; (2) a pair of second order p.d.e.'s for the temperature and stress. This is put in dimensionless form with appropriate boundary conditions and solved by a perturbation method wherein the temperature and stress are expanded in power series with respect to a dimensionless perturbation parameter. The uncoupled temperature distribution is solved for by constructing the appropriate Green's function using the method of weak solutions. From this we obtain the uncoupled temperature distribution under repeated heating. Next, coupled thermal stress fields are obtained both for the case of a unit step function in stress and unit gas temperature. The first case illustrates multiple reflecting waves.

II. MATHEMATICAL MODEL. We present a mathematical formulation of the structure of a coupled linearized unsteady thermoelastic stress field. Consider a three dimensional region R bounded by a surface S on which are prescribed surface tractions and a linear combination of temperature and temperature gradient (called the radiation boundary condition). The surface tractions are prescribed functions of time t. The radiation boundary condition is given in the form $-K \text{grad} T = -h\,(T-T_g)$, where T is the temperature in R in the neighborhood of S (to be solved for as a function of the particle coordinates given by the vector $\underline{x}(x_1,x_2,x_3)$ and t, K is the thermal conductivity, h-the heat transfer coefficient with respect to the particular surface environment, T_g-the temperature of the gas external to S-is a prescribed function of t; K and h are assumed constant. For a gun tube S consists of the outer and inner walls of the tube. On the outer wall T_g is assumed to be the ambient temperature; on the inner wall T_g (the propellant gas temperature) is a prescribed function of t given from interior ballistic considerations. Also for the gun tube, the surface tractions reduce to the t varying prescribed pressure on the inner wall (given from the solution of the interior ballistic problem for the pressure of the propellant gas). The outer wall is considered to be stress free. ... a differential element of material in the interior of R. This is ... le of the medium. Its position at time t is given by $\underline{x}$. The con... ... the equations of motion, the energy balance equation, and the ... which hold for a material particle.

... ation in vector form is

$$\mu\nabla^2\underline{u} + (\lambda+\mu)\ \text{grad div}\ \underline{u} + \underline{F} - \gamma\ \text{grad}\ T = \rho\underline{u}_{tt} \quad (1)$$

where $\underline{u}$ (u_1, u_2, u_3) is the particle displacement vector, T-the temperature, $\underline{F}$-the body force vector, (μ,λ) the Lame constants, ∇^2-the Laplacian in three dimensions, and γ is a constant given by

$$\gamma = (3\lambda + 2\mu)\alpha$$

where α is the coefficient of thermal expansion. The last term on the left hand side of Eq. (1) represents the force per unit volume due to the temperature gradient in the medium. In an uncoupled problem the temperature field is prescribed in the medium and the energy equation given below does not directly involve mechanical energy. In a coupled problem, the energy equation involves mechanical energy represented by the time rate of change of dilatation and the temperature field cannot be prescribed but must be solved for along with the stress and strain fields.

The coupled energy equation is

$$\nabla^2 T - \kappa^{-1}T_t - \eta\ \text{div}\ \underline{u}_t = -\kappa^{-1}Q \quad (2)$$

where the constant η is given by

$$\eta = \frac{\gamma T}{\rho c \kappa}$$

κ is the thermal diffusivity, c-the heat capacity, Q-the heat source. The last term on the left hand side is proportional to the time rate of change of dilation, since div $\underline{u}_t = \theta$, where the dilation θ is given by $\theta = u_{i,i}$ (the tensor summation convention is used). This term represents the coupling of mechanical energy. If $\alpha = 0$ then $\eta = 0$ and Eq. (2) reduces to the unsteady Fourier heat transfer equation with a heat source.

The Duhamel-Neumann constitutive equations for an isotropic material are

$$\sigma_{ij} = \delta_{ij}\lambda\theta + 2\mu u_{i,j} - \gamma\delta_{ij}T \qquad i,j = 1,2,3 \quad (3)$$

where σ_{ij} is the ijth component of the stress tensor, δ_{ij} the Kronecker delta ($\delta = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}$), and $u_{i,j} \equiv \frac{\partial u_i}{\partial x_j}$. The last term of the right hand side of Eq. (3) represents the linearized contribution of the temperature to the stress field. Note that the temperature only affects the principle stresses. The shear stresses are not affected by temperature. This derives from the assumption that the temperature field only changes the volume of the material particle.

*Throughout this paper subscripts indicate partial differentiation with respect to the subscript.

We may now formulate the general thermoelastic problem: Given the thermal and elastic properties of the material (the constants κ, K, ρ, c, λ, μ), the appropriate boundary conditions on S, the initial conditions on T, $\underline{u}$, $\underline{u}_t$, and the definition of the linearized strain $\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,2})$, solve the above conservation equations for the stress and temperature in R.

SCALAR AND VECTOR POTENTIALS

The general problem formulated above is difficult to solve. Under certain conditions a simplification is obtained if a scalar potential ϕ and a vector potential $\underline{\psi}$ are introduced according to the following definition: We decompose $\underline{u}$ into the sum of the gradient of a scalar and the curl of a vector.
Thus,

$$\underline{u} = \text{grad } \phi + \nabla x \underline{\psi} \tag{4}$$

Inserting Eq. (4) into Eqs. (1) and (2) (for the case $\underline{F} = Q = 0$) gives the following:

$$\Box_1 \phi = mT \tag{5}$$

$$\Box_2 \underline{\psi} = 0 \tag{6}$$

where

$$m = \frac{\gamma}{\lambda + 2\mu}, \ \Box_i = \nabla^2 - c_i^{-2} \frac{\partial^2}{\partial t^2}, \ i = 1,2, \ c_1^2 = \frac{\lambda + 2\mu}{\rho}, \ c_2^2 = \frac{\mu}{\rho} \tag{7}$$

$\Box_1$ is the longitudinal wave operatior, $\Box_2$-the shear wave operator, c_1-the velocity of a longitudinal wave, c_2-the velocity of a shear wave. Eq. (5) is the wave equation for ϕ coupled with T and represents longitudinal waves in ϕ. Eq. (6) represents shear wave in $\underline{\psi}$ independent of T. These are rotational waves. This is consistent with the Duhamel-Neumann relationships which assume T only affects volume change and does not produce shear. These wave equations along with Eq.s (4) and (2) allow us to obtain the ϕ, $\underline{\psi}$ and T fields. The stress tensor can then be calculated from Eq. (3). The appropriate initial and boundary conditions must be used.

GUN BARREL

We now apply the above theory to a typical gun barrel. Consider a cross section of the barrel. The region R (the barrel wall) is represented by the annulus bounded by the inner radius r_1, and outer radius r_2. The boundary conditions are $r = r_1$, $-KT_r = -h_1(T-T_g(t))$, $\sigma_{rr} = p(t)$, $r = r_2$, $-KT_r = h_2(T - T_0)$, $\sigma_{rr} = 0$

σ_{rr} is the principal stress in the radial direction, T_g-the temperature of the propellant gas in the interior of the barrel and p is the pressure on the inner wall due to the propellant gas. T_g and p are prescribed functions of time, to be determined from interior ballistic considerations. This requires a study of the heat transfer of the hot propellant gas by convection through the turbulent boundary layer at the inner wall.

ONE-DIMENSIONAL CASE

For simplicity we neglect the curvature of the barrel and replace the annulus by a one-dimensional finite slab. The wall of the barrel is thus represented by the region R: $0 \leq x \leq \ell$. 0 is the position of the inner wall, ℓ-the outer wall. ℓ is the wall thickness. The equations of motion reduce to a single equation:

$$\sigma_x = \rho u_{tt} \tag{8}$$

where σ is the one-dimensional stress and u-the displacement. All dependent variables are functions of (x, t). Let e be the linearized one-dimensional strain. Then $e = u_x$ so that Eq. (8), after differentiation with respect to x, becomes

$$\sigma_{xx} = \rho e_{tt} \tag{9}$$

The coupled energy equation becomes

$$T_{xx} - \kappa^{-1} T_t = \eta e_t \tag{10}$$

The Duhamel-Neumann constitutive equation is

$$\sigma = Ee - \gamma T \tag{11}$$

where E is the Young's modulus of the material. $E = \rho c_1^2$

Eqs. (9), (10), (11) are equivalent to the following coupled system of second order partial differential equations for T and σ:

$$L_1 T = \bar{\eta} \sigma_t \tag{12}$$

$$\Box_1 \sigma = \bar{\gamma} T_{tt} \tag{13}$$

where

$$L_1 = \frac{\partial^2}{\partial x^2} - \kappa^{-1}(1 + a)\frac{\partial}{\partial t}, \quad \bar{\eta} = \frac{\eta}{\rho c_1^2}, \quad \bar{\gamma} = \frac{\gamma}{c_1^2}, \quad a = \gamma\bar{\eta}\kappa$$

L_1 is the modified linear unsteady heat transfer operator. Operating on Eq. (13) by L_1 and using Eq. (12) to eliminate T gives the following fourth order p.d.e. for σ:

$$L_1 \Box_1 \sigma = \bar{\gamma}\eta\sigma_{ttt}$$

Written in extended form Eq. (14) becomes

$$\kappa(c_1^2\sigma_{xx} - \sigma_{tt})_{xx} - (c_2^2\sigma_{xx} - \sigma_{tt})_t = 0$$

where

$$c_2^2 = c_1^2 \left[1 + \left(\frac{3\lambda + 2\mu}{\rho c_1}\right)^2 \frac{T_o\alpha^2}{c}\right] \tag{14}$$

The boundary conditions are as follows:

$$x = 0 \quad \sigma(0,t) = f(t), \; T_x(0,t) - \beta_1 T(0,t) = g(t) \tag{15a}$$

$$x = 1 \quad \sigma(1,t) = 0, \; T_x(1,t) + \beta_2 T(1,t) = 0 \tag{15b}$$

where

$$\beta_1 = \frac{h_1}{\kappa}, \; \beta_2 = \frac{h_2}{\kappa}, \; f(t) = -p(t), \; g(t) = \beta_1 T_g(t)$$

One formulation of the problem of getting the thermoelastic stress distribution in the finite slab is to solve Eqs. (12) and (13) for T and σ subject to the boundary conditions given by Eqs. (15-a,b) and homogeneous initial conditions on T, σ, σ_t. Another formulation is to solve the fourth order p.d.e. Eq. (14) for σ and manipulate Eq. (11) and the boundary conditions to get the appropriate boundary conditions on σ. However, instead of using these approaches we take a different tack. Making use of the weak coupling of the mechanical term $\eta\theta_t$ in the energy equation, we develop a perturbation method which involves expansions of σ and T in powers of a small dimensionless perturbation parameter ε which describes the effect of this coupling. The reason for this approach is to attempt to determine the conditions for which we have a coupled thermoelastic stress field. This method will be described below. In order to set the scene we recast the problem in dimensionless variables.

DIMENSIONLESS VARIABLES

We define dimensionless x, t, T, σ, α by the corresponding barred quantities as follows:

$$\bar{x} = \frac{x}{\ell}, \; \bar{t} = \frac{c_1^2 t}{\kappa}, \; \bar{T} = \frac{T}{T_o}, \; \bar{\sigma} = \frac{\sigma}{E}, \; \bar{\alpha} = \alpha T_o \tag{16}$$

where T_o is the initial or reference temperature. We also define the dimensionless parameters ε, r and the time t_M, t_T by

$$\varepsilon = r\bar{\alpha}, \; r = \frac{\kappa}{\ell c_1} = \frac{t_M}{t_T}, \; t_M = \frac{\ell}{c_1}, \; t_T = \frac{\ell^2}{\kappa} \tag{17}$$

t_M, the "mechanical time," is the time for an elastic stress wave to travel a

distance ℓ with a velocity c_1. t_T, the "thermal time," is the time for a temperature pulse to decay to e^{-1} of its original value.

Inserting Eqs. (16) and (17) into Eqs. (12) and (13) the coupled second order p.d.e.'s for T and σ become in dimensionless form (omitting the bars)

$$r^2 T_{xx} - (1 + a_1\varepsilon^2)T_t = a_2\varepsilon\sigma_t \tag{18}$$

$$r^2\sigma_{xx} - \sigma_{tt} = a_3\varepsilon T_{tt} \tag{19}$$

$a_1 = (3\lambda+2\mu)^2/T_0\rho c(\lambda+2\mu)r^2$, $a_2 = 3\lambda+2\mu/\rho cT_0 r$, $a_3 = 3\lambda+2\mu/(\lambda+2\mu)r$

The initial conditions are

$$t = 0, \quad \sigma(x,0) = \sigma_t(x,0) = T(x,0) = T_t(x,0) = 0 \tag{20}$$

The boundary conditions given by Eqs. (15-2,b) are put in dimensionless form by using Eq. (16). They have the same form except β_i is replaced by $\ell\beta_i$, i = 1,2, f(t) by $\frac{p(t)}{E}$ and g(t) by $\frac{\ell\beta_1 T_g(t)}{T_0}$.

Concerning the smallness of the parameters r and ε: For a given ℓ and $c_1 r_1 \to 0$ as $\kappa \to 0$ so that when r=0 with a loss in boundary conditions we have a singular perturbation. However, if r>0 and ε=0, we do not have a singular perturbation, since this means α=0 and the system becomes uncoupled with no loss in boundary conditions.

PERTURBATION METHOD

We apply the following perturbation technique to the system given by Eqs. (18), (19) and the boundary and initial conditions: We expand σ and T in power series in the perturbation parameter ε. Thus

$$\sigma(x,t) = \sum_{n=0}^{\infty} \sigma^n(x,t)\varepsilon^n, \quad T(x,t) = \sum_{n=0}^{\infty} T^n(x,t)\varepsilon^n \tag{21}$$

$\sigma^n(x,t)$ and $T^n(x,t)$ are the nth order expansion parameters for $\sigma(x,t)$ and $T(x,t)$ and are to be solved for as functions of dimensionless (x,t). The series expansions given by Eq. (21) are assumed to be convergent for small ε. For ε small enough only the first few expansion coefficients are sufficient to give a good approximation to the solution. For ε=0 we have $\sigma(x,t) = \sigma^0(x,t)$ and $T(x,t) = T^0(x,t)$ which represents the uncoupled problem. Inserting Eq. (21) into Eqs. (18), (19) and the dimensionless form of (15-a,b) we get the following iterative coupled system of p.d.e.'s for the expansion coefficients $T^n(x,t)$, $\sigma^n(x,t)$ with the appropriate boundary conditions:

$$r^2 T^n_{xx} - T^n_t = a_1 T^{n-2}_t + a_2 \sigma^{n-1}_t \quad , \quad 0 \le x \le 1 \tag{22}$$

$$r^2 \sigma^n_{xx} - \sigma^n_{tt} = a_3 T^{n-1}_{tt}$$

$$T^n = \sigma^n = 0 \quad n = -1, -2, \ldots\ldots$$

The boundary conditions are

$$x = 0 \qquad T^0_x(0,t) - \beta_1 T^0(0,t) = g(t), \qquad \sigma^0(0,t) = f(t)$$

$$T^n_x(0,t) - \beta_1 T^n(0,t) = 0, \qquad \sigma^n(0,t) = 0, \quad n = 1,2,\ldots \tag{22a}$$

$$x = 1 \qquad T^n_x(1,t) + \beta_2 T^n(1,t) = 0, \qquad \sigma^n(1,t) = 0, \quad n = 0,1,2,\ldots \tag{22b}$$

The initial conditions are homogeneous. Eqs. (22) and (22-a,b) represent a coupled set of second order p.d.e.'s for the set $T^n(x.t)$, $\sigma^n(x,t)$ that is iterative in the sense that the solution for T^n depends on T^{n-} and σ^{n-1}, and σ^n depends on T^{n-1}.

UNCOUPLED SYSTEM

Setting n=0 (the zeroth order perturbation) in Eqs. (22) and (22-a,b) gives the following boundary value problems for T^0 and σ^0:

$$r^2 T^0_{xx} - T^0_t = 0, \; 0 < x < 1 \tag{23}$$

$$x = 0, \qquad T^0_x(0,t) - \beta_1 T^0(0,t) = g(t) \tag{23a}$$

$$x = 1, \qquad T^0_x(1,t) + \beta_2 T^0(1,t) = 0 \tag{23b}$$

$$t = 0, \qquad T(x,0) = 0 \qquad 0 \le x \le 1 \tag{23c}$$

$$r^2 \sigma^0_{xx} - \sigma^0_{tt} = 0 \qquad o \le x \le 1 \qquad 0 \le t \tag{24}$$

$$x=0, \qquad \sigma^0(0,t) = f(t); \qquad x=1, \qquad \sigma^0(1,t) = 0 \tag{24a,b}$$

$$t=0, \qquad \sigma(x,0) = 0, \qquad \sigma_t(x,0) = 0 \tag{24c}$$

The solution of Eqs. (23) and (24) yields the uncoupled temperature and stress distribution in the slab with the appropriate boundary and initial conditions.

UNCOUPLED TEMPERATURE DISTRIBUTION - USE OF GREEN'S FUNCTION

We are interested in solving Eqs. (23), ... (23-c) for a prescribed g(t). The solution can be obtained by using standard operational techniques such as Laplace transforms. However, this approach although apparently straightforward, is too unwieldy because of the computational difficulty in obtaining the inverse transform. A more elegant and efficient approach is to make use of the method of weak solutions. This involves constructing the appropriate Green's function for the adjoint system, and then using this Green's function in an integral manner to calculate $T^0(x,t)$.

The first step in solving for the uncoupled temperature field is to define this Green's function. Let (x,t) be the coordinates of a field point imbedded in the region R: $(0 \leq \xi \leq 1,\ 0 \leq \tau \leq t_1$ (t_1 is an upper bound in time). We define $G(\xi,x;\ \tau,t)$ as the solution to the following adjoint boundary value problem:

$$r^2 G_{\xi\xi} + G_\tau = -\delta(\xi-x)\delta(\tau-t), \qquad 0 \leq \xi,\ x \leq 1,\ 0 \leq \tau \leq t < t_1 \tag{25}$$

$$\tau = 0,\ G = 0;\ \xi = 0,\ G_\xi - \beta_1 G = 0;\ \xi = 1,\ G_\xi + \beta_2 G = 0 \tag{25a}$$

The operator $L^* = r^2\frac{\partial^2}{\partial\xi^2} + \frac{\partial}{\partial\tau}$ operating on G is adjoint to the heat conduction operator $L = r^2\frac{\partial^2}{\partial\xi^2} - \frac{\partial}{\partial\tau}$ operating on $T^0(\xi,\tau)$ in Eq. (23). (Note that (ξ,τ) are variables and (x,t) is a field point in R). L is not self-adjoint. The non-homogeneous term in Eq. (25) consists of the Dirac delta functions $\delta(\xi-x)\delta(\tau-t)$. As seen below, this form for the non-homogeneous term has the property of picking out the value of T^0 at the field point (x,t) upon integration over the region R. Green's function depends on the region involved, i.e. whether $\xi<x$ or $\xi>x$ so that for $\xi<x$, $G = G_\ell$ (the left hand Green's function) and for $\xi>x$, $G = G_r$. G is continuous across $\xi=x$ but there is a finite jump discontinuity of G_ξ at $\xi=x$. Also G is symmetric in the sense that $G_\ell(\xi,x;\ \tau,t) = G_r(x,\xi;\ \tau,t)$. The interpretation of G is that it is a solution of the "backward" heat equation ($G = 0$ for $\tau = 0$, $\tau > t$) with an impulsive source at $\xi=x$ at time $\tau=t$, as shown by the backward heat operator L^* in Eq. (25) and the delta functions for the non-homogeneous or source term. Eq. (25-a) shows us that G is moreover the solution for homogeneous boundary conditions of the "radiation type." Thus, the problem of solving for $T^0(x,t)$ from Eqs. (23), ... (23-c) is first reduced to solving the simpler problem for G.

Having obtained G, we calculate $T^0(x,t)$ from the following integral identity (making use of Green's identity):

$$\int_0^{t_1}\int_0^1 G(r^2T_{\xi\xi}-T_\tau)d\xi d\tau = \int_0^{t_1}[GT_\xi - G_\xi T]_0^1 d\tau - \int_0^1 [GT]_0^{t_1} d\xi$$

$$+ \int_0^{t_1}\int_0^1 T(r^2G_{\xi\xi}+G_\tau)d\xi d\tau \qquad (26)$$

where the region of integration is R: ($0\le\xi\le1$, $0\le\tau\le t_1$). The left hand side of Eq. (26) vanishes because $T(\xi,\tau)$ satisfies the heat equation. The last integral on the right becomes $-T(x,t)$, by virtue of Eq. (25) using the properties of the delta functions. The first two integrals on the right involve boundary conditions. The second integral vanishes because $G=0$ at $\tau=0$ and $\tau=t_1$ ($t_1>t$). Hence Eq. (26) becomes

$$T^o(x,t) = \int_0^{t_1} -T^o(1,\tau)[G_\xi(1,x;\tau,t)+\beta_2 G(1,x;\tau,t)] + T^o(0,\tau)\begin{bmatrix} G_\xi(0,x;\tau,t) \\ -\beta_1 G(0,x;\tau,t)\end{bmatrix} d\tau$$

$$+\beta_1 G(0,x;\tau,t)T_g(\tau)$$

Using the homogeneous boundary conditions on G given by Eq. (25-a) we get the final expression for T^o:

$$T^o(x,t) = \beta_1 \int_0^t G(0,x;\tau,t)T_g(\tau)d\tau \qquad (27)$$

since $G(0,x;\tau,t) = 0$ for $\tau>t$. (Note $t_1>t$).

To obtain $G(0,x;\tau,t)$ we must solve the boundary value problem for G given by Eqs. (25, 25-a). We do this by taking the Laplace transform of this system. We obtain the following boundary value problem:

$$\bar{G}_{\xi\xi} + q^2\bar{G} = \delta(\xi-x)e^{-st}, \quad 0\le\xi, \quad x\le1, \quad t>0 \qquad (28)$$

$$\xi = 0, \quad \bar{G}_\xi-\beta_1\bar{G} = 0; \qquad \xi = 1, \quad \bar{G}_\xi + \beta_2\bar{G} = 0 \qquad (28a)$$

where $\bar{G}$ is the Laplace transform of G with respect to τ. The solution is

$$\bar{G}(\xi,x,s,t) = \begin{cases} A\cos q\xi + B\sin q\xi\ , & \xi<x \\ A\cos q\xi + B\sin q\xi + \dfrac{e^{-st}}{q}\sin(\xi-x), & \xi>x \end{cases} \qquad (29)$$

where

$$A = \frac{e^{-st}[\cos q(1-x) + \frac{\beta_2}{q}\sin q(1-x)]}{(\frac{\beta_1\beta_2}{q} - q)\sin q + (\beta_1+\beta_2)\cos q}, \quad B = \frac{B_1 A}{q}, \quad q^2 = \frac{s}{r^2} \tag{30}$$

The values for s are the roots of the following transcendental equation:

$$\tan q = \frac{(\frac{\beta_1\beta_2}{q})q}{q^2-\beta_1\beta_2} \tag{31}$$

These roots are denoted by the set $\{q_n\}$, n=1,2, ... The inverse Laplace transform is obtained by use of the inversion integral. The result is

$$G(\xi,x;\ \tau,t) = \begin{cases} G_\ell & \xi<x \\ G_r & \xi>x \end{cases}$$

$$G_\ell(\xi,x;\ \tau,t) = \sum_{n=1}^{\infty} \frac{e^{-s_n(t-\tau)}}{Q'(s_n)} \cdot (q_n\cos q_n\xi + \beta_1\sin q_n\xi)[q_n\cos q_n(1-x) + \beta_2\sin q_n(1-x)]$$

$$Q'(s_n) = \frac{1}{2qr^2}\{[\beta_1\beta_2-q^2(\beta_1+\beta_2+3)]\sin q + q\,[2(\beta_1+\beta_2)+\beta_1\beta_2-q^2]\cos q\} \tag{32}$$

and the eigenvalues s_n are obtained from the corresponding roots of Eq. (31). In using Eq. (27) to solve for T^o we need only set ξ=0 in the above expression for G_ℓ and calculate the integral in the right hand side, knowing the gas temperature as a function of time.

TEMPERATURE DISTRIBUTION UNDER REPEATED HEATING

In the case of a slab subjected to repeated heating, $T_g(t)$ may be described by a series of Dirac delta functions displaced in time:

$$T_g(t) = T_g \sum_{m=1}^{N} \delta[t-(m-1)t_R] \tag{33}$$

where t_R is the time elapsed between cycles, N is the number of cycles of heating, and T_g is a constant gas temperature. Substituting Eq. (33) into (27) gives

$$T^0(x,t) = \beta_1 T_g \sum_{n=1}^{\infty} \sum_{m=1}^{N} e^{-s_n[t-(m-1)t_R]} \frac{q_n[q_n \cos q_n(1-x) + \beta_2 \sin q_n(1-x)}{Q'(s_n)} \tag{34}$$

Eq. (34) is the uncoupled unsteady temperature distribution in a finite slab under repeated heating of the inner surface for the "radiation" boundary condition given by Eq.s (23-b,c).

UNIT STEP FUNCTION IN STRESS, f(t)=1

We now turn to the coupled expansion coefficients T^n, σ^n, $n>0$, in the system given by Eqs. (22, 22-a,b). As an example of a coupled system we consider a unit step function in stress applied at x=0, the other boundary conditions being homogeneous. We wish to solve the boundary value problem given by Eqs. (22,22-a,b) for the case g(t) = 0, f(t) = 1. It is easily seen that the series expansions given by Eq. (21) become

$$\sigma(x,t) = \sum_{n=0}^{\infty} \sigma^{2n}(x,t)\varepsilon^{2n}, \quad T(x,t) = \sum_{n=0}^{\infty} T^{2n+1}(x,t)\varepsilon^{2n+1} \tag{35}$$

The uncoupled field is given by $T^0(x,t) = 0$ and $\sigma^0(x,t)$ the solution of the boundary value problem given by Eqs. (24-a,b,c) for f(t) = 1. This is easily obtained by taking the Laplace transform with respect to t which is designated as $\tilde{\sigma}(x,s)$. The boundary value problem for $\tilde{\sigma}(x,s)$ involves the following ordinary differential equation:

$$\tilde{\sigma}^{0''} - \lambda^2 \tilde{\sigma}^0 = 0, \quad 0<x<1, \quad \lambda = \frac{s}{r}$$
$$x = 0, \quad \tilde{\sigma}^0(0,s) = \frac{1}{s}, \qquad x = 1, \quad \tilde{\sigma}(1,s) = 0 \tag{36}$$

whose solution is

$$\tilde{\sigma}^0(x,s) = \frac{1}{s} \frac{\sinh\lambda(1-x)}{\sinh\lambda}$$

This can be expanded into the following series of exponentials:

$$\tilde{\sigma}^0(x,s) = \frac{e^{\lambda(1-x)} - e^{-\lambda(1-x)}}{se^{\lambda}(1-e^{-2\lambda})} = \frac{1}{s} \sum_{n=0}^{\infty} [e^{-\lambda(2n+x)} - e^{-\lambda[2(n+1)-x]}] \tag{37}$$

The inverse Laplace transform of Eq. (37) gives

$$\sigma^o(x,t) = \sum_{n=0}^{\infty} [s_{2n+x}(t) - s_{2(n+1)-x}(t)], \quad s_k(t) = \begin{cases} 0 & 0<t<k \\ 1 & t>k \end{cases} \tag{38}$$

(Note that t is replaced by rt). The interpretation of this solution is that multiple reflecting stress waves propagate in the strip $0 \leq t$, $0 \leq x \leq 1$ which satisfy the boundary and initial conditions. This is shown in Fig. 1 "Characteristic Diagram for Uncoupled Stress Distribution for Unit Stress Input." The solution for $\sigma^o(x,t)$, as seen from the figure, is best interpreted in terms of characteristic theory.

The region n is the triangle whose boundaries are the characteristics $x+t = 2n$, $x-t = -2n$ and the line $x = 1$, $2n-1 \leq t \leq 2n+1$. In this region $\sigma^o(x,t) = 0$. The neighboring region n+1 is the triangle bounded by the characteristics $x-t = -2n$, $x+t = 2(n+1)$ and the line $x = 0$, $2n \leq t \leq 2n+2$. In this region $\sigma^o(x,t) = 1$. We see that σ^o alternately jumps from zero to one to zero, etc. The jumps are across the characteristics. This means that $\sigma^o_t(x,t) = \pm 1$ across each characteristic depending on whether the characteristic has a positive or negative slope, and $\sigma^o_t(x,t) = 0$ elsewhere.

As an example of a typical firing condition, we take a stress input at the inner boundary as a series of uniform pulses of unit strength, pulse width = 2 millisec, time between pulses = 0.1 sec., thickness of gun barrel = 1.6 cm, $c_1 = 3 \times 10^5$ cm/sec. The time for a stress wave to travel from the interior to the exterior boundary is 5 μsec. This means that a stress wave initiated at the inner boundary will travel back and forth 200 times during one pulse, so that multiple reflections (due to the finiteness of the region) are important.

The first order coupled temperature distribution $T'(x,t)$ is obtained from Eqs. (22-a,b) for n = 1.

Using the jump condition of $\sigma^o_t(x,t)$ at the characteristics, the b.v. problem for T' becomes

$$r^2 T'_{xx} - T'_t = a_2 \delta(x-t) + a_2 \sum_{n=1}^{\infty} [\delta(2n+x-t) - \delta(2n-x-t)] \tag{39}$$

$$x = 0, \quad T'_x(o,t) - \beta_1 T'(0,t) = 0; \quad x = 1, \quad T'_x(1,t) + \beta_2 T'(1,t) = 0$$

where $\delta(\ )$ is the Dirac delta function. The solution for this non-homogeneous p.d.e. for T' with homogeneous boundary and initial conditions is obtained by using the Green's function technique (described previously for the uncoupled temperature distribution). We make use of the Green's identity given by Eq. (26) where the

Green's function $G(\xi, x; \tau, t)$ is given by Eq. (32). Inserting Eq. (39) in the left hand side of Eq. (26) (replacing x by ξ and t by τ in Eq. (39)) and inserting the homogeneous b.c. $g(\tau) = 0$ and Eq. (25) in the right hand side, we get the solution for T':

$$T'(x,t) = \int_0^t \int_0^1 G(\xi, x; \tau, t) f(\xi \pm \tau) d\xi d\tau$$

$$f(\xi \pm \tau) = a_2 \delta(\xi - \tau) + a_2 \sum_{n=1}^{\infty} [\delta(2n + \xi - \tau) - \delta(2n - \xi - \tau)] \qquad (40)$$

The higher order perturbation expansion coefficients $\sigma^{2n}(x,t)$ and $T^{2n+1}(x,t)$ are obtained by successively solving the iterated system given by Eqs. (22, 23-a,b). For example, to calculate σ^2, insert the solution for T' in the right hand side of the second equation of Eq. (22). σ^2 is then the solution of a non-homogeneous p.d.e. with homogeneous b.c.'s. Then use the solution for T' and σ^2 to calculate the right hand side of the first of Eq. (22). This gives a non-homogeneous p.d.e. with homogeneous bc's for T^3. Working in this manner we calculate as many of the expansion coefficients T^{2n+1} and σ^{2n} as we need.

UNIT GAS TEMPERATURE

We consider the case of zero stress at the boundaries and unit gas temperature at the inner boundary. We want to solve the iterated system Eqs. (22, 23-a,b) for the expansion coefficients T^n, σ^n using the boundary conditions $f(t) = 0$, $g(f) = 1$ in Eq. 23-a). Inserting the perturbation expansions for $\sigma(x,t)$ and $T(x,t)$ given by Eq. (21) into Eqs. (22, 23-a,b) for the appropriate b.c.'s gives the following expansion:

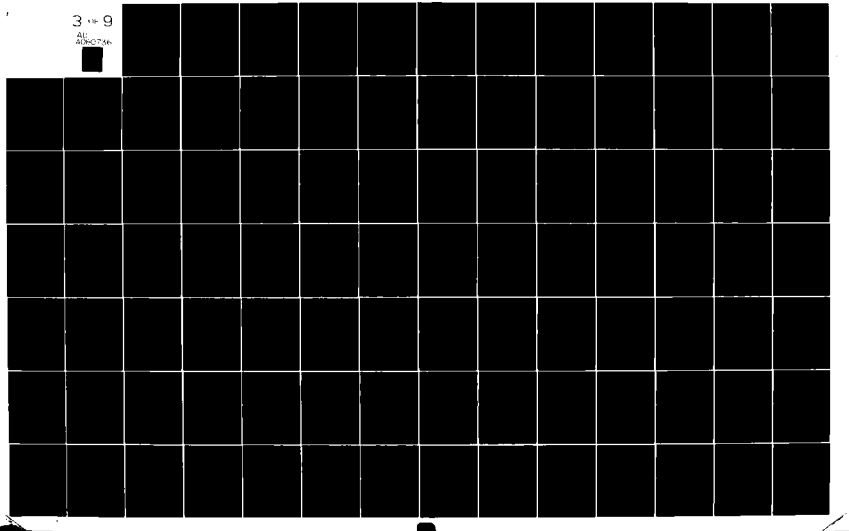
$$\sigma(x,t) = \sum_{n=0}^{\infty} \sigma^{2n+1}(x,t)\varepsilon^{2n+1}, \quad T(x,t) = \sum_{n=0}^{\infty} T^{2n}(x,t)\varepsilon^{2n} \qquad (41)$$

Note that for this case only the odd expansion coefficient for σ and the even expansion coefficients for T are non-zero which is the reverse of the first case for $f(t) = 1$, $g(t) = 0$. The same procedure is used to solve this system iteratively for the σ^{2n+1} and T^{2n}.

REFERENCES

1. Carslaw, H. S. and Jaeger, J. C. Conduction of Heat in Solids, Oxford at the Clarendon Press, Second Edition, 1959.
2. Nowacki, W. Thermoelasticity, Addison-Wesley Publishing Co., 1962, p. 296.
3. Hildebrand, F. B. Advanced Calculus for Applications, Second Edition, Prentice Hall, 1976, p. 669.
4. Stakgold, I., Boundary Value Problems of Mathematical Physics, Vol. 2, MacMillan, 1967, p. 199.
5. Churchill, R. V., Modern Operational Mathematics in Engineering, First Edition, 1944, pp. 91.

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL



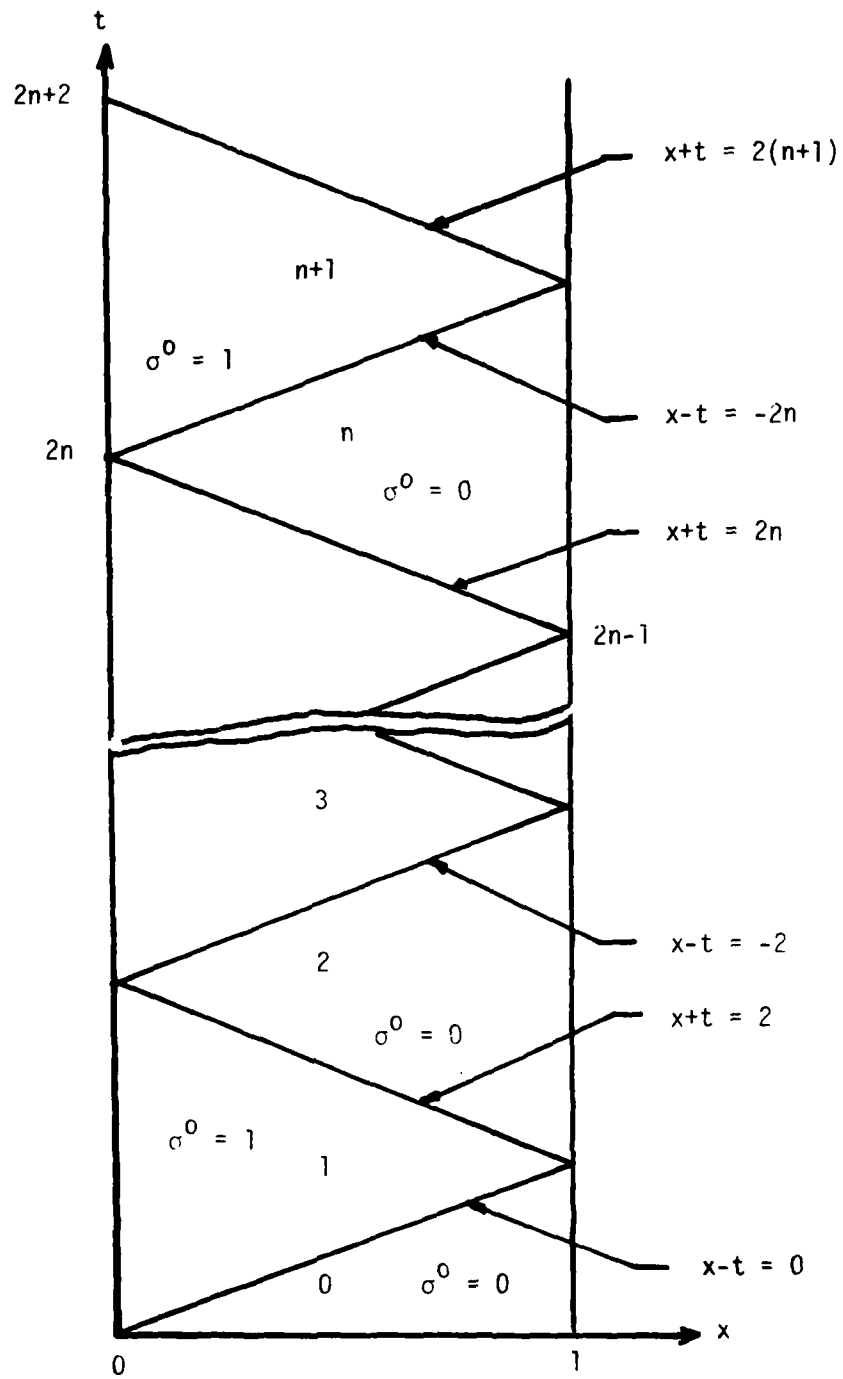


Figure 1. CHARACTERISTIC DIAGRAM FOR UNCOUPLED STRESS DISTRIBUTION FOR UNIT STRESS INPUT

A NONLINEAR HYPERBOLIC VOLTERRA EQUATION OCCURRING IN VISCOELASTIC MOTION.

John A. Nohel
Mathematics Research Center
University of Wisconsin
Madison, Wisconsin 53706

ABSTRACT. A mathematical model for the motion of a nonlinear one dimensional viscoelastic rod is formulated and analysed by an energy method developed by C. M. Dafermos and the author. Global existence, uniqueness, boundedness, and the decay of smooth solutions as $t \to \infty$ are established for sufficiently smooth and "small" data.

1. INTRODUCTION. In this paper we motivate and summarize results on the global existence, uniqueness, boundedness, and decay as $t \to \infty$ of smooth solutions of the nonlinear Cauchy problem:

$$\text{(VE)} \qquad \begin{cases} u_{tt}(t,x) = \sigma(u_x(t,x))_x + \int_0^t a'(t-\tau)\sigma(u_x(\tau,x))_x d\tau + g(t,x) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (0 < t < \infty,\ x \in \mathbb{R}) \\ u(0,x) = u_0(x)\ ,\quad u_t(0,x) = u_1(x) \qquad (x \in \mathbb{R}), \end{cases}$$

for appropriately small, smooth data u_0, u_1, g; $a : [0,\infty) \to \mathbb{R}^+$, $\sigma : \mathbb{R} \to \mathbb{R}$ $(\sigma(0) = 0)$, $g : [0,\infty) \times \mathbb{R} \to \mathbb{R}$, u_0, $u_1 : \mathbb{R} \to \mathbb{R}$ are given functions satisfying assumptions motivated by physical considerations sketched below and partly by the method of analysis. In (VE) subscripts denote partial derivatives and u is the unknown function. In addition to the Cauchy problem (VE), we will comment on several closely related initial-boundary value problems.

The results stated in section 2 were established recently by a general energy method for the study of nonlinear hyperbolic Volterra equations developed jointly with C. M. Dafermos [4].

Problem (VE) arises in the following physical context. Consider one dimensional motion of an unbounded viscoelastic rod of unity density. According to the theory of materials of "fading memory" type (see Coleman and Gurtin [1]) the stress $S(t,x)$ at time t and position x is given by a functional of the history of the strain, $u_x(t-\tau,x)$ $(\tau \geq 0)$, where $x+u(t,x)$ denotes the position at time t of a section of the rod which is at position x in the unstretched configuration. In the nonlinear case the theory suggests assuming that the stress functional S has the form

$$\text{(1.1)} \qquad S(t,x) = \sigma(u_x(t,x)) - \int_0^\infty b(\tau)\varphi(u_x(t-\tau,x))d\tau \qquad (t > 0),$$

with the history of the displacement $u(t,x)$ prescribed for $t < 0$ and $x \in \mathbb{R}$. Relaxation experiments of materials indicate that $\sigma, \varphi : \mathbb{R} \to \mathbb{R}$ are smooth functions which satisfy the assumptions $\sigma(0) = \varphi(0) = 0$, $\sigma'(\xi) > 0$, $\varphi'(\xi) > 0$

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and Grant No. DAAG29-77-G-0004.

($\xi \in \mathbb{R}$), and that the "influence" (or memory) function $b:[0,\infty) \to \mathbb{R}^+$ satisfies $b(t) > 0$, $b'(t) < 0$ for $t \in \mathbb{R}^+$ and that $b \in L^1(0,\infty)$ (e.g. b is a linear combination of decaying exponentials with positive coefficients). We remark that a standard assumption of linear theory is that $\sigma(\xi) = c_1\xi$, $\varphi(\xi) = c_2\xi$ where $c_1, c_2 > 0$ are constants [2].

If the rod is also subjected to an external force $F(t,x)$, then the equation of motion for the rod is

$$u_{tt}(t,x) = S_x(t,x) + F(t,x), \qquad (0 < t < \infty,\ x \in \mathbb{R}), \tag{1.2}$$

together with prescribed initial values $u(0,x)$, $u_t(0,x)$, where S is the stress functional defined by (1.1). Recalling that the history of displacement is prescribed for $t < 0$ and defining

$$g(t,x) = F(t,x) - \int_t^\infty b(\tau)\varphi(u_x(t-\tau,x))_x d\tau \tag{1.3}$$

for $t > 0$, $x \in \mathbb{R}$ shows that the motion of the unbounded viscoelastic rod is described by the Cauchy problem

$$\begin{cases} u_{tt} = \sigma(u_x)_x - b*\varphi(u_x)_x + g & (0 < t < \infty,\ x \in \mathbb{R}) \\ u(0,x) = u_0(x), \quad u_t(0,x) = u_1(x) & (x \in \mathbb{R}), \end{cases} \tag{1.4}$$

where $*$ denotes the convolution, defined by

$$(b * \varphi(u_x)_x)(t,x) = \int_0^t b(t-\tau)\varphi(u_x(\tau,x))_x d\tau.$$

Our method of analysis requires us to make the further assumption

$$\varphi(\xi) = c\sigma(\xi) \qquad (\xi \in \mathbb{R}), \tag{1.5}$$

where $c > 0$ is a constant. Assumption (1.5) is satisfied in the linear case and is reasonable for certain nonlinear problems. We shall be primarily interested in the "genuinely nonlinear case" $\sigma''(\xi) \not\equiv 0$ $(\xi \in \mathbb{R})$.

Consider next the Cauchy problem (1.4) with $g \equiv 0$ (or $\lim g(t,x) = 0$, uniformly in x), under assumption (1.5). The corresponding steady state problem is meaningful if it is assumed that

$$1 - c\int_0^\infty b(\tau)d\tau > 0\ ; \tag{1.6}$$

assumption (1.6) has the interpretation that the static modulus elasticity is positive (see Dafermos [2], [3] where the same assumption is made in the linear case). With the assumptions (1.5), (1.6), and those concerning b made above, we can reduce the Cauchy problem (1.4) to the equivalent form (VE) as follows: define $a : [0,\infty) \to \mathbb{R}^+$ by

$$(1.7)\qquad \begin{cases} a(t) = a_\infty + A(t)\ ; \quad a_\infty = 1 - c\int_0^\infty b(\tau)d\tau > 0\ ; \\ A(t) = c\int_t^\infty b(\tau)d\tau;\quad a(0) = 1\ ;\quad A(t) > 0\ ,\quad A'(t) < 0 \\ \qquad\qquad\qquad\qquad \text{for}\quad 0 \le t < \infty;\quad A(\infty) = 0\ ; \end{cases}$$

define g by (1.3); then (1.4)-(1.7) is equivalent to (VE). The analysis which follows will be concerned with (VE), where a satisfies the physically reasonable assumptions implied by (1.7); for technical reasons based on our analysis we shall require that a satisfy somewhat stronger assumptions.

To motivate our result for (VE) we begin with some general remarks. If $a(t) \equiv 1$, (VE) reduces to the equation of nonlinear elasticity

$$(E)\qquad u_{tt} = \sigma(u_x)_x + g\ ,\quad u(0,x) = u_0(x)\ ,\quad u_t(0,x) = u_1(x)\ .$$

If $g \equiv 0$ in (E) and if σ is "genuinely nonlinear" Lax [6], has shown that (E) fails to have global smooth solutions in time, no matter how smooth one takes the initial data u_0, u_1, due to the development of "shocks" (the first derivatives of solutions generally develop singularities when characteristics cross).

Nishida [13] has shown that for the wave equation with "frictional" damping

$$u_{tt} + u_t = \sigma(u_x)_x,\quad u(0,x) = u_0(x),\quad u_t(0,x) = u_1(x)$$

the dissipation precludes the development of shocks if the initial data are sufficiently smooth and "small", resulting in global smooth solutions. The proof which rests on the concept of Riemann invariant is restricted to one space dimension.

MacCamy [10] has recently studied (VE) on $(0,\infty) \times (0,1)$ and homogeneous Dirichlet boundary conditions at $x = 0$, $x = 1$, by combining Nishida's method with certain a priori estimates under suitable assumptions on the kernel $a(t)$ and the forcing term g. His object is to show that the memory term in (VE) induces a dissipative mechanism which guarantees global existence for "small" initial data and forcing term. The problem of obtaining the existence and uniqueness of a suitable local solution of (VE) to be continues with the aid of the derived a priori estimates is not discussed in [10], but this gap can be filled by the method outlined in Nohel [14] where the result of [13] is extended.

The object of this work is to study (VE) by a different approach based entirely on energy estimates and not on Riemann invariants. While the exposition is restricted to the one-dimensional problem (VE) for clarity, the method can be applied to problems in any number of space dimensions, provided estimates on derivatives of sufficiently high order are computed. This approach to (VE) may be regarded as a generalization of recent work of Matsumura [11], [12] who studies multi-space-dimensional nonlinear wave equations with frictional damping for small data by a similar energy method. We are grateful to Professor Nishida for explaining this approach to us during a recent visit to Madison.

We remark that the special case of (VE) resulting from the kernel $a(t) = \frac{1}{2}(1 + \exp(-t))$ can easily be shown to be equivalent to the Cauchy problem

$$u_{ttt} + u_{tt} = \sigma(u_x)_{xt} + \tfrac{1}{2}\sigma(u_x)_x + g_t + g \ ,$$

$$u(0,t) = u_0(x) \ , \quad u_t(0,x) = u_1(x) \ , \quad u_{tt}(0,x) = \sigma(u_{0x})_x \ ,$$

which was studied by Greenberg [5]; his result is a special case of ours.

Finally, we note that (VE) is of the abstract form

$$\text{(A)} \qquad \begin{cases} u''(t) + Au(t) + \int_0^t a'(t-\tau)Au(\tau)d\tau = F(t), & 0 < t < \infty, \\ u(0) = u_0 \ , \quad u'(0) = u_1 \end{cases}$$

where Au is a nonlinear operator ($Au = -\frac{\partial}{\partial x}\sigma(u_x)$ together with conditions at $\pm\infty$, or boundary conditions at say $x = 0,1$). Abstract problems of the form (A) have been studies by Londen [7], [8] for a class of kernels $a(\cdot)$ which are positive, decreasing, convex on $[0,\infty)$ and satisfy the crucial assumption $a'(0^+) = -\infty$; the latter assumption is not satisfied by most "memory" functions in viscoelasticity.

2. STATEMENT OF RESULTS. We make the following assumptions. Concerning σ let

$$(\sigma) \qquad \sigma \in C^3(\mathbb{R}) \ , \quad \sigma(0) = 0 \ , \quad \sigma'(0) > 0 \ ,$$

the first for technical reasons and the remaining on physical grounds. Concerning the kernel a assume

$$\text{(a)} \qquad \begin{cases} \text{(i)} \quad a \in B^{(3)}[0,\infty) \ , \\ \text{(ii)} \quad a(t) = a_\infty + A(t) \ , \quad a_\infty > 0 \ , \quad a(0) = 1 \ , \quad a'(0) < 0 \ , \\ \text{(iii)} \quad (-1)^j A^{(j)}(t) \geq 0 \qquad (0 \leq t < \infty \ ; \ j = 0,1,2), \\ \text{(iv)} \quad t^j A^{(m)}(t) \in L^1(0,\infty) \qquad (m,j = 0,1,2,3), \end{cases}$$

where $B^{(m)}[0,\infty)$ is the set of functions with bounded, continuous derivatives on $[0,\infty)$ up to and including order m. The meaning of assumptions (a) is that the kernel a in (VE) is positive, smooth, decreasing and convex on $[0,\infty)$, and that the part $A(t)$ of a and three of its derivatives have moments up to order three integrable on $[0,\infty)$. The forcing term g is assumed to satisfy

$$\text{(g)} \qquad g, g_t \in L^1([0,\infty); L^2(\mathbb{R})), \ g_x, g_{tt}, g_{tx} \in L^2([0,\infty); L^2(\mathbb{R})),$$

meaning that g and some of its distributional derivatives decay sufficiently reapidly at infinity. The initial data u_0, u_1 satisfy

$$(u_0) \qquad u_0 \in H^3(\mathbb{R}), \qquad (u_1) \qquad u_1 \in H^2(\mathbb{R}) \ .$$

Our result concerning (VE) is (see [4; Theorem 5.1]):

Theorem 2.1. Let the assumptions (σ), (a), (g), (u_0), (u_1) hold. If the $H^2(\mathbb{R})$ norms of u_{0x}, u_1, the $L^1([0,\infty);\ L^2(\mathbb{R}))$ norms of g, g_t, and the $L^2([0,\infty);\ L^2(\mathbb{R}))$ norms of g_x, g_{tt}, g_{tx} are sufficiently small then (VE) has a unique solution $u \in C^2([0,\infty) \times \mathbb{R})$ having the following properties:

(2.1) $\quad u_t, u_x, u_{tt}, u_{tx}, u_{xx}, u_{ttt}, u_{ttx}, u_{txx}, u_{xxx} \in L^\infty([0,\infty);\ L^2(\mathbb{R}))$,

(2.2) $\quad u_{tt}, u_{tx}, u_{xx}, u_{ttt}, u_{ttx}, u_{txx}, u_{xxx} \in L^2([0,\infty);\ L^2(\mathbb{R}))$,

(2.3) $\quad u_{tt}(t,\cdot), u_{tx}(t,\cdot), u_{xx}(t,\cdot) \to 0$ in $L^2(\mathbb{R})$ as $t \to \infty$.

(2.4) $\quad u_t(t,x), u_x(t,x), u_{tt}(t,x), u_{tx}(t,x), u_{xx}(t,x) \to 0$ uniformly in $\mathbb{R}$ as $t \to \infty$.

We remark that conclusions (2.3), (2.4) are an easy consequence of (2.1), (2.2). It also follows from the proof of the theorem that the solution u has a finite speed of propagation. In addition, we note that the same result hold (and with the same proof) for the following two problems of a viscoelastic rod of unit length:

(i) (VE) on $(0,\infty) \times (0,1)$ with homogeneous Neumann boundary conditions at $x = 0$ and $x = 1$, and with initial data prescribed on $[0,1]$.

(ii) (VE) on $(0,\infty) \times (0,1)$ with homogeneous Dirichlet boundary conditions $u(t,0) = u(t,1) \equiv 0$, and initial data prescribed on $[0,1]$, provided one also assumes that the forcing term g also satisfies $g(t,0) = g(t,1) \equiv 0$. Finally, we observe that a comparison of Theorem 2.1 and its proof in [4] with Mac Camy's results in [10] shows that our approach, in addition to being simpler, more direct, and not restricted to one space dimension, yields a more general result.

3. COMMENTS ON THE PROOF OF THEOREM 2.1. Our procedure can be outlined as follows. We first reduce the problem (VE) to the equivalent form

$$(3.1)\quad \begin{cases} u_{tt}(t,x) + \dfrac{\partial}{\partial t}\,(k * u_t)(t,x) = \sigma(u_x(t,x))_x + \Phi(t,x) \\ \qquad\qquad (0 < t < \infty\ ,\ x \in \mathbb{R}) \\ u(0,x) = u_0(x)\ ,\quad u_t(0,x) = u_1(x) \qquad (x \in \mathbb{R})\ , \end{cases}$$

where k is the resolvent kernel of a' defined as the unique solution of the linear equation

$$(k)\qquad k(t) + (a' * k)(t) = -a'(t) \qquad (0 \le t < \infty)\ .$$

By standard harmonic analysis methods, and by a frequency domain argument to obtain the last conclusion (see Nohel and Shea [15, Theorem 1]), the resolvent kernel has the following properties:

Lemma 3.1. Let assumptions (a) be satisfied. Then

(i) $k \in B^{(2)}[0,\infty)$, $k(0) > 0$;

(ii) $k^{(m)} \in L^1(0,\infty)$ $(m = 0,1,2)$;

(iii) For every $T > 0$ and for every $v \in L^2(0,T)$ one has

$$\int_0^T v(t) \frac{d}{dt} (k * v)(t)dt \geq 0 .$$

The function Φ in (3.1) is determined in terms of g and k ; assumptions (g) and the properties of k in Lemma 3.1 imply that Φ satisfies

$$(\Phi) \qquad \Phi, \Phi_t \in L^1([0,\infty;L^2(\mathbb{R})), \ \Phi_x, \Phi_{tt}, \Phi_{tx} \in L^2([0,\infty); L^2(\mathbb{R})] .$$

The (non-physical) assumption (1.5) which is crucial for the reduction of (VE) to (3.1), is not used anywhere else in the analysis.

The next step is to prove the existence and uniqueness of a sufficiently regular, smooth local solution u of (3.1) on $[0,T] \times \mathbb{R}$ for some $T > 0$; this is done with the aid of the Banach fixed point theorem and a fairly standard energy argument (See [4; Theorems 3.1, 3.2].

The essential and rather tedious part of the proof is to establish a series of energy estimates for derivatives of u by elementary methods which allow the extension of the local solution into a global one. Unfortunately, this requires yet another transformation of the Cauchy problem (3.1) (equivalent to (VE)), because property (iii) of Lemma 3.1 only allows us to obtain uniform bounds on $\int_{-\infty}^{\infty} u_t^2(s,x)dx$, $\int_{-\infty}^{\infty} u_x^2(s,x)dx$, but not on $\int_0^s \int_{-\infty}^{\infty} u_{tt}^2(t,x)dx\,dt$, on any interval $0 \leq s < T$ on which the local solution exists (see [4, estimate (5.7) and remarks following). The additional transformation is elementary but technical (see [4, Section 2, part II, especially Lemma 2.3]). The long series of energy estimates for the newly transformed problem (see [4, estimates (5.8)-(5.25)] which allow continuation of the local solution and from which one obtains at the same time conclusions (2.1), (2.2) of Theorem 2.1, each have the form

$$(3.2) \qquad E(t) - E(0) \leq -\int_0^t \int_{\infty} Q[u,u]]dxd\tau + \int_0^t \int_{\infty} P[u,u]dxd\tau + \int_0^t \int_{\infty} \Pi[u,\Phi]dxd\tau$$

where $E(t)$ is an "energy" that controls the growth of the solution; $Q[u,u]$, the dissipation term induced by the memory term, is a positive definite quadratic form in a set of derivatives of $u(t,x)$; $P[u,u]$, the remainder term due to the nonlinearity of the problem, is a quadratic form in the same derivatives as $Q[u,u]$ and with coefficients that are small whenever the "energy" E is small; finally, $\Pi[u,\Phi]$ is a bilinear form in the set of derivatives of $u(t,x)$ involved in $Q[u,u]$ and in $\Phi(t,x)$ and some of its derivatives. The idea now is that for as long as $E(t)$ is small, $P[u,u]$ is dominated by $-Q[u,u]$. Moreover, the Cauchy-Schwarz inequality allows us to dominate the u-part in $\Pi[u,\Phi]$ by $-Q[u,u]$. Then, if $E(0)$ and Φ are "small", (3.2) shows that $E(t)$ remains small and the cycle closes.

4. CONCLUDING REMARKS. We have in Section 3 indicated the crucial role of assumption (1.5) in our analysis. Since this assumption is not really physical, considerable effort is being spent in current research to remove it by attempting to apply energy methods directly to the physical equations (1.4). If these efforts are successful, there is hope of being able to apply such (possibly modified) energy methods to treat the considerably more complicated system of nonlinear hyperbolic Volterra equations which describe nonlinear viscoelastic motion in two and three space dimensions.

REFERENCES

[1] B. D. Coleman and M. E. Gurtin, Waves in materials with memory. II On the growth and decay of one-dimensional acceleration waves. Arch. Rat. Mech. and Analysis 19 (1965), 239-265.

[2] C. M. Dafermos, An abstract Volterra equation with applications to linear viscoelasticity, J. Differential Equations 7 (1970), 554-569.

[3] C. M. Dafermos, Asymptotic stability in viscoelasticity, Arch. Rat. Mech. and Analysis 37 (1970), 297-308.

[4] C. M. Dafermos and J. A. Nohel, Energy methods for nonlinear hyperbolic Volterra integrodifferential equations, Communications in P.D.E. 4 (1979) 219-278.

[5] J. M. Greenberg, A priori estimates for flows in dissipative materials, J. Math. Anal. and Appl. 60 (1977), 617-630.

[6] P. D. Lax, Development of singularities of solutions of nonlinear hyperbolic partial differential equations, J. Math. Phys. 5 (1964), 611-613.

[7] S. O. Londen, An existence result for a Volterra equation in Banach space, Trans. Amer. Math. Soc. 235 (1978), 285-305.

[8] S. O. Londen, An integrodifferential Volterra equation with a maximal monotone mapping, J. Differential Equations (to appear).

[9] R. C. Mac Camy, An integro-differential equation with applications in heat flow, Q. Appl. Math. 35 (1977), 1-19.

[10] R. C. Mac Camy, A model for one-dimensional, nonlinear viscoelasticity, Ibid 35 (1977), 21-33.

[11] A Matsumura, Global existence and asymptotics of the solutions of the second order quasilinear hyperbolic equations with first order dissipation (to appear).

[12] A. Matsumura, Energy decay of solutions of dissipative wave equations (to appear).

[13] T. Nishida, Global smooth solutions of the second-order quasilinear wave equations with the first-order dissipation (unpublished).

[14] J. A. Nohel, A forced quasilinear wave equations with dissipation, Proceedings of EQUADIFF 4, Lecture Notes Vol 703 (1979), 318-327, Springer-Verlag.

[15] J. A. Nohel and D. F. Shea, Frequency domain methods for Volterra equations, Advances in Math. 22 (1976), 278-304.

A NEW TECHNIQUE FOR THE SOLUTION OF NAVIER'S EQUATIONS

Francis E. Council, Jr.
Management Information Systems Directorate
U. S. Army Mobility Equipment Research and Development Command
Fort Belvoir, VA 22060

ABSTRACT. A solution for the partial differential equations otherwise known as Navier's equations is obtained by means of Fourier transforms and Parseval's relation which are used to form a Green's tensor. The displacement functions that are obtained by this technique are used with a forcing function with a randomly occurring phase and amplitude to synthesize the accelerations and frequency spectrums of earthquakes. The techniques that have been developed in this paper have an applicability to other types of partial differential equations.

NOTATIONS. The following notations are used in this unless otherwise specified.

F_1, F_2, F_3	=	components of force associated with the forcing function
F	=	symbol for the Fourier transform
F^{-1}	=	symbol for the inverse Fourier transform
$[G]$	=	Green's tensor
G_k	=	base vector in curvalinear frame
g_K	=	base vector in curvalinear frame
$\vec{U}$	=	displacement vector
U_1, U_2, U_3	=	components of displacement vector
X_1, X_2, X_3	=	coordinates of a point in undeformed coordinates
X^I	=	one of the coordinates of a point in the undeformed state
x_1, x_2, x_3	=	coordinates of a point in deformed coordinates
x^i	=	one of the coordinates of a point in the deformed state
T	=	time variable
Z_I	=	a coordinate of the Cartesian coordinate system, undeformed
z_i	=	a coordinate of the Cartesian coordinate system, deformed
α	=	prestrained strain amplitude
λ, μ	=	Lame's constants
ρ	=	density
W	=	angular frequency
$[\ ]$	=	matrix

I. INTRODUCTION. This paper addresses the problem of solving some partial differential equations, namely Navier's equations with a specific application being synthesis of accelerations and frequency spectra associated with earthquakes. Solutions for these partial differential equations were obtained by means of Fourier transforms and Parseval's relation being used with a Green's tensor. The techniques that are developed in this paper are of sufficiently general nature such that other partial differential equations may be solved by means of these techniques. The justification for the equations that were developed as well as the results that were obtained help to give some insight as to the causes and effects of earthquakes.

Before discussing the solution of the Navier's equations, some background should be given as to the system of coordinates and the forcing function that were used. The model that has been chosen in this paper as representative of an earthquake, namely a dislocation, implies that there is a release of energy associated with an earthquake. This release of enercy is not instantaneous but occurs over a finite period of time. The random nature of seismic disturbances is an indication that the periods of this oscillatory process occur with randomly occuring lengths. This then suggests that some function such as a sine function with a different period for each oscillation be used such that a time history of the accelerations in the surrounding medium can be obtained. The previous discussion would seem to indicate that a frequency or a phase modulation could be used with this sine function. Since actual seismograms display peaks, a factor for amplitude modulation can be included. It then follows that if the times for an acceleration to go from zero to a maximum and then to zero are represented as a random sequency of numbers, i.e., T_1, T_2, T_3,...,T_i, then the total time for a disturbance can be represented as $T = \sum_{i}^{n-1} T_i + t$ where $t \leq T_i$.

Continuing in this manner, then a function that can be used as a forcing function for synthesizing the main portion of an earthquake by means of the equations developed in this paper for describing the displacement in the surrounding medium can be expressed as

$$F(t) = \frac{aT_i}{2(.64)}(\sin\pi(n-1) + \frac{t}{T_i}) \qquad (1.1)$$

where a is the constant for gravitational acceleration.

The specific choice of coordinates can no longer be deferred. In order to facilitate computations, a Cartesian coordinate system is chosen such that two of the axes are in the plane of the discontinuity associated with a dislocation. For instance, if a slip strike is being considered, then the X_1 of the undeformed state and x_1 of the deformed state are parallel and perpendicular to the plane involving the dislocation. The X_3 and x_3 coordinates, referring to the undeformed and deformed states, respectively, are coincident with the plane of the dislocation and perpendicular to the surface of the Earth. The third members of the triads for describing undeformed and deformed states, X_2 and x_2, respectively, are parallel to the plane of the dislocation and are colinear. If a dip strike is being considered, then displacement is with respect to the X_3 and x_3 coordinates. In this paper, there is only one force, F_3, for a dip strike which is a function of the magnitude and which is directed along the X_3 coordinate. A slip strike has a force, F_2, directed along the X_2 coordinate.

II. MODIFICATION OF STANDARD DEFORMATION THEORY FOR A PRESTRAINED MEDIA. The residual strain field present after an earthquake is a contributing factor to the magnitude of the stress waves associated with an earthquake, the volume adjacent to a fault being no longer isotropic if it was previously isotropic. For the propagation of certain stress waves, a prestrained media is required. Although it is more accurate to consider the fall-off of residual strain energy as a function of distance by means of an expoential function, in order to facilitate computations, one relationship between the undeformed and the deformed coordinates can be given as

$$x^i = (1-\alpha) X^I \delta_I^i \tag{2.1}$$

where axial changes are considered. In this last equation, α is the relative amount of change of one of the coordinates with respect to the other. The simplest model that can be considered for the residual stress field remaining after a slip strike or dip strike has occurred is that of uniaxial compression. The next model that can be considered, and a possibly more realistic model, is that of both a compression and a dilation.

For a dip strike, $i = 3$, and for a slip strike $i = 2$. Since a solid is in general incompressible, this implies that the other two coordinates, involving dilation, of the deformed state are related to the undeformed state as

$$x^i = \frac{1}{(1-\alpha)^{\frac{1}{2}}} X^I \delta_I^i \tag{2.2}$$

such that for a dip strike, i = 1,2. This type of model offers a way of describing the force that is associated with the release of energy of an earthquake. A fixed orthogonal Carte coordinate system is introduced by means of the real single v and reversible transformations where

$$X^K = X^K(Z_J)$$
$$Z_K = Z_K(X^J) \quad , \quad K,J = 1,2,3 \tag{2.3}$$

and

$$x^k = x^k(Z_J)$$
$$z_j = z_j(x^k) \quad , \quad k,j = 1,2,3 \tag{2.4}$$

with $$\left|\frac{\partial z_k}{\partial Z_K}\right| \equiv J \neq 0 \qquad \left|\frac{\partial Z_K}{\partial z_k}\right| \equiv J^{-1} \neq 0 \tag{2.5}$$

A coordinate point is described in the undeformed reference system, is described as $R = Z^K{}_K$ *and in the deformed reference* system as $\vec{R} = z^k \epsilon_k$. Base vectors in the curvilinear frames X^K AND x^k are defines as

$$\vec{G}_J = \frac{\partial \vec{R}}{\partial X^J} = \frac{\partial Z_K}{\partial X_J} \vec{\epsilon}_K \quad , \quad \vec{g}_j = \frac{\partial \vec{r}}{\partial x_j} = \frac{\partial Z_k}{\partial x_j} \vec{\epsilon}_k \tag{2.6}$$

with reciprocal base vectors being defined as

$$\vec{G}^K = \frac{\partial X^K}{\partial Z_J} \vec{\epsilon}_J \quad , \quad \vec{g}^k = \frac{\partial x^k}{\partial z_j} \vec{\epsilon}_j \tag{2.7}$$

$$G^K G_J = \delta^K_J \quad , \quad g^k g_j = \delta^k_j \tag{2.8}$$

A displacement vector, $\vec{U}$, is expressed as

$$\vec{U} = \vec{r} - \vec{R} \tag{2.9}$$

These relationships are shown in Figure 2.1.

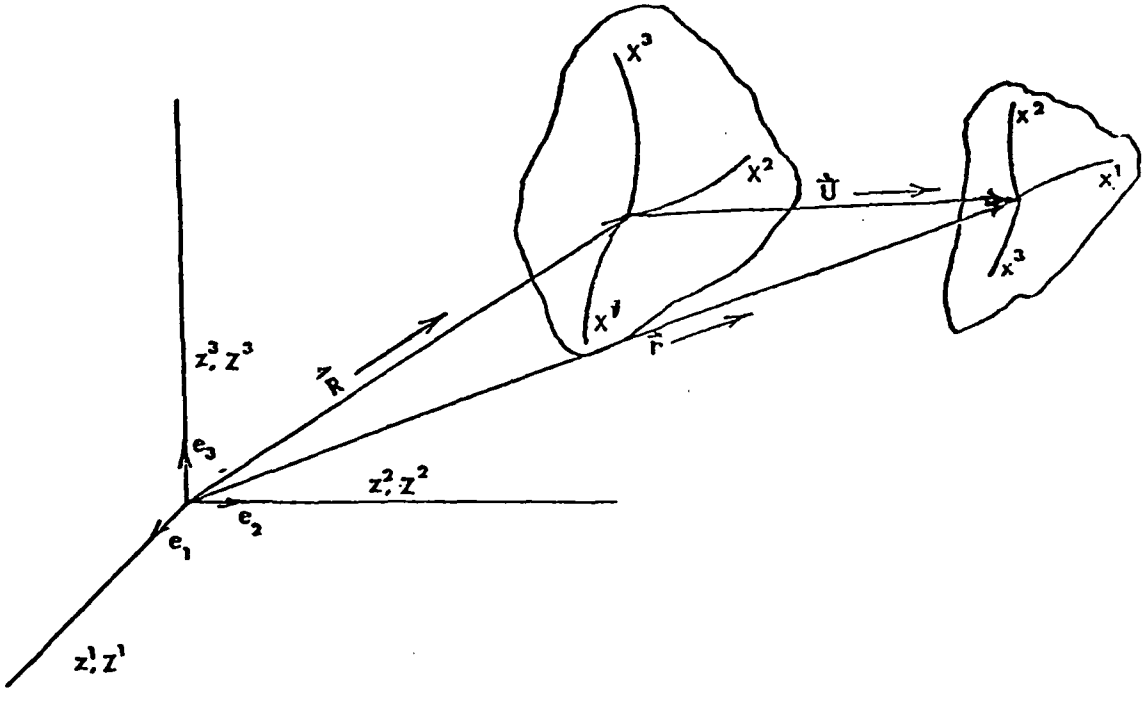


Figure 2.1

Relationship of Undeformed and Deformed Coordinates

III. MODIFICATION OF NAVIER'S EQUATIONS FOR A PRESTRAINED MEDIA. One way of relating the stress tensor in deformed state coordinates to the coordinates of the undeformed state is by means of the Piola-Kirchoff stress tensor, defined by the relationship

$$T^{KL} = JX^{K}{}_{,k}X^{K}{}_{,\ell}T^{k\ell} = T^{LK} \tag{3.1}$$

such that

$$(x^{k}{}_{,K}T^{KL})_{;L} - F^{k} - f^{k} = 0 \tag{3.2}$$

with the semi-colon indicating total covariant differentiation. In this equation, F^{k} refers to externally applied forces and f^{k} refers to inertial forces. The mixed stress tensor (Piola) is defined by the relationship

$$T^{Kk} = JX^{K}{}_{,\ell}T^{k\ell} \tag{3.3}$$

with

$$T^{Kk}{}_{;K} - F^{k} - f^{k} = 0 \tag{3.4}$$

The Piola-Kirchoff stress tensor and the Piola mixed tensor are discussed further by Eringin (3) and Truesdell (5). Another way of relating a tensor of a deformed state to that of an undeformed state is by using finite deformation theory as developed by Green and Zerna (4).

Since a prestrained medium is being considered, the relationships between the undeformed and deformed states as expressed by equations (2.1) and (2.2) are used in equation (3.2). As previously discussed a single forcing function is considered, one that is consistent with equation (3.4) and that is acting along the X_3 coordinate axes. In view of these equations, then three equations are obtained from equation (3.4) such that

$$\frac{(\lambda+2\mu)}{(1-\alpha)}U_{1,11} + \frac{2\mu}{(1-\alpha)}U_{2,12} + (1-\alpha)^{\frac{1}{2}}2\mu U_{3,13} - \rho\frac{\partial^{2}U_{1}}{\partial t^{2}} = 0 \tag{3.5}$$

$$\frac{2\mu}{(1-\alpha)}U_{1,21} + \frac{(\lambda+2\mu)}{(1-\alpha)}U_{2,22} + (1-\alpha)^{\frac{1}{2}}2\mu U_{3,23} - \rho\frac{\partial^{2}U_{2}}{\partial t^{2}} = 0 \tag{3.6}$$

$$(1-\alpha)^{\frac{1}{2}}\ 2\mu U_{1,31} + (1-\alpha)^{\frac{1}{2}}\ 2\mu U_{2,23} + (1-\alpha)^{2}(\lambda+2\mu)U_{3,33} \tag{3.7}$$

$$-\rho\frac{\partial^2 U_3}{\partial t^2} = F_3$$

IV. OPERATORS USED WITH THE FOURIER TRANSFORMS. The set of partial differential equations that have been previously developed, equations (3.5), (3.6), (3.7), are most easily solved by means of Fourier transforms although the equations can be solved by means of iterative processes used with a computer.

Prior to demonstrating the solution by means of Fourier transforms, a brief description of the transforms used in this dissertation is given here. If

$$Ff(x_i) = \int_{-\infty}^{\infty} e^{ip_i x_i} f(x_i)dx_i = g(p_i) \tag{4.1}$$

and

$$F^{-1}g(p_i) = \int_{-\infty}^{\infty} e^{-ip_i x_i} g(p_i)dp_i \tag{4.2}$$

then

$$F\frac{\partial}{\partial x_i} f(x_i) = ip_i Ff(x_i) \tag{4.3}$$

since

$$F^{-1}F\frac{\partial}{\partial x_i} f(x_i) = \int ie^{-ip_i x_i} p_i\, g(p_i)dp_i = \tag{4.4}$$

$$\frac{\partial}{\partial x_i} F^{-1}Ff(x_i) = \frac{\partial}{\partial x_i} f(x_i)$$

The factor $\frac{1}{\sqrt{2\pi}}$ has been changed to unity because of the contour of integration that is used to evaluate these equations. Consider the contour of integration to be over the upper half plane. If

$$f(x_i) = \frac{1}{x_i}$$

such that $$F\,\frac{1}{x_i} = \int \frac{e^{ip_i x_i}}{x_i}\,dx_i = \pi i \tag{4.5}$$

then $$F^{-1}F\,\frac{1}{x_i} = \frac{1}{x_i} \tag{4.6}$$

and

$$F^{-1}\pi i = \pi i \int e^{ip_i x_i} dp_i = \pi i \delta(x_i) = \frac{1}{x_i} \tag{4.7}$$

The slight modifications to the definitions of the Fourier transform, inverse Fourier transform, and the Dirac delta function, $\delta(x-x')$, are made in order to maintain a consistency of development. Continuing in this manner, then

$$F\,\frac{1}{(x_i)^2} = \frac{e^{ip_i x_i} dxi}{(x_i)^2} = -\,p_i \tag{4.8}$$

Also

$$F^{-1}\,\frac{1}{p_i} = \pi i \tag{4.9}$$

$$F^{-1}\,\frac{1}{(p_i)^2} = \pi x_i \tag{4.10}$$

Another important relationship that is used is Parseval's relation for Fourier transforms ( ) where

$$\int_{-\infty}^{\infty} F(\tau)G(\tau)\,d\tau = \int_{-\infty}^{\infty} f(-n)g(n)\,dn \tag{4.11}$$

At this time some operations must be defines. Since $F\,\frac{\partial}{\partial x_1}\,f(X_1) = ipf(X_1)$ and $F\,\frac{\partial^2}{\partial x_1^2}\,f(X_1) = -\,p^2\,Ff(X_1)$ then this implies that $\frac{1}{ip}\,Ff(X_1) = F\int f(X_1)dX_1$ and that $-\,\frac{1}{(p_1)^2}\,Ff(X_1) = F\iint f(X_1)dx_1dx_1$. Continuing in this manner, then the expression $(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)})$ can be expressed as an operator. For example, the wave equation in one dimensional space where $\psi(X,T)$ is the wave function is expressed as

$$\left(\frac{\partial^2}{\partial X^2} - \frac{\omega^2}{C^2}\right)\psi(X,T) = C\delta(X)\delta(T) \tag{4.12}$$

Taking the Fourier transforms of this equation,

$$\left(p^2 - \frac{\omega^2}{C^2}\right)FF_T\psi(X,T) = CFF_T\delta(X)\delta(T) \tag{4.13}$$

and rearranging,

$$FF_T\psi(X,T) = \frac{CFF_T\delta}{\left(p^2 - \frac{\omega^2}{C^2}\right)}(X)\delta(T) \tag{4.14}$$

If now the inverse Fourier transforms are taken,

$$(X,T) = CF^{-1}F_T^{-1}\frac{FF_T\delta(X)\delta(T)}{\left(p^2 - \frac{\omega^2}{C^2}\right)} \tag{4.15}$$

with the usual way of representing the wave function as

$$(X,T) = Ce^{i(X-CT)} \tag{4.16}$$

while for a wave traveling in the other direction,

$$(X,T) = Ce^{i(X+CT)} \tag{4.17}$$

such that

$$F^{-1}F_T^{-1}\frac{CFF_T\delta(X)\ \delta(T)}{\left(\frac{P_1P_1}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}\right)} = Ce^{\,i\left(X_1(1-\alpha)^{\frac{1}{2}} - \left(\frac{\lambda+2\mu}{\rho}\right)^{\frac{1}{2}}T\right)} \tag{4.18}$$

which when extended to three dimensions is

$$F^{-1}F_T^{-1}\frac{CFF_T\delta(X)\delta(T)}{\left(\frac{P_1P_1}{(1-\alpha)} - \left(\frac{\rho}{\lambda+2\mu}\right)\omega^2\right)\left(\frac{P_2P_2}{(1-\alpha)} - \left(\frac{\rho}{\lambda+2\mu}\right)\omega^2\right)\left(P_3P_3(1-\alpha)^2 - \left(\frac{\rho}{\lambda+2\mu}\right)\omega^2\right)} =$$

$$Ce^{\,i\left(X_1(1-\alpha)^{\frac{1}{2}} + X_2(1-\alpha)^{\frac{1}{2}} + X_3(1-\alpha) - \left(\frac{\lambda+2\mu}{\rho}\right)^{\frac{1}{2}}T\right)} \tag{4.19}$$

The constant C is set equal to 1 since dimensionally consistent results are obtained with this value.

V. THE GREEN'S TENSOR. One way of solving equations (3.5), (3.6) and (3.7) is by means of a Green's tensor (2). If operator notation is used

$$[L]\vec{U} = \vec{F} \tag{5.1}$$

and

$$[L]\ [G] = -\ [\delta] \tag{5.2}$$

then if the operator $[L]$ is a matrix of order three and the delta function is of order three, then equation (5.2) is an indication that the Green's tensor, $[G]$, is a matrix of order three. If Fourier transforms with respect to the spatial coordinates are taken, then equation (5.2) is written as a system of equations as

$$\left(\frac{(\lambda+2\mu)}{(1-\alpha)}\ p_1p_1 + \rho\ \frac{\partial^2}{\partial T^2}\right)FG_{11} + \frac{2\mu}{(1-\alpha)}\ p_1p_1FG_{21} + (1-\alpha)^{\frac{1}{2}}\ 2\mu p_1p_3FG_{31} =$$

$$F\delta(\vec{X}-\vec{X})\ \delta\ (T-t) \tag{5.3}$$

$$\left(\frac{(\lambda+2\mu)}{(1-\alpha)}\ p_1p_1 + \rho\ \frac{\partial^2}{\partial T^2}\right)FG_{12} + \frac{2\mu}{(1-\alpha)}\ p_1p_2FG_{22}$$

$$+ (1-\alpha)^{\frac{1}{2}}\ 2\mu p_1p_3FG_{32} = 0 \tag{5.4}$$

$$\left(\frac{(\lambda+2\mu)}{(1-\alpha)}\ p_1p_1 + \rho\ \frac{\partial^2}{\partial T^2}\right)FG_{13} + \frac{2\mu}{(1-\alpha)}\ p_1p_2FG_{23}$$

$$+ (1-\alpha)^{\frac{1}{2}}\ 2\mu p_1p_3FG_{33} = 0 \tag{5.5}$$

$$\frac{2\mu}{(1-\alpha)}\ p_2p_1FG_{11} + \left(\frac{(\lambda+2\mu)}{(1-\alpha)}\ p_2p_2 + \rho\ \frac{\partial^2}{\partial T^2}\right)FG_{21}$$

$$+ (1-\alpha)^{\frac{1}{2}}\ 2\mu p_2p_3FG_{31} = 0 \tag{5.6}$$

$$\frac{2\mu}{(1-\alpha)} p_2 p_1 FG_{12} + \left(\frac{(\lambda+2\mu)}{(1-\alpha)} p_2 p_2 + \frac{\partial^2}{\partial T^2}\right) FG_{22}$$

$$+ (1-\alpha)^{\frac{1}{2}} 2\mu p_2 p_3 FG_{32} = F\delta(\vec{X}-\vec{x})\delta(T-t) \qquad (5.7)$$

$$\frac{2\mu}{(1-\alpha)} p_2 p_1 FG_{13} + \left(\frac{(\lambda+2\mu)}{(1-\alpha)} p_2 p_2 + \rho \frac{\partial^2}{\partial T^2}\right) FG_{23}$$

$$+ (1-\alpha)^{\frac{1}{2}} 2\mu p_2 p_3 FG_{33} = 0 \qquad (5.8)$$

$$(1-\alpha)^{\frac{1}{2}} 2\mu p_3 p_1 FG_{11} + (1-\alpha)^{1^2} 2\mu p_3 p_2 FG_{21}$$

$$+ \left((1-\alpha)^2 (\lambda+2\mu) p_3 p_3 + \rho \frac{\partial}{\partial T^2}\right) FG_{31} = 0 \qquad (5.9)$$

$$(1-\alpha)^{\frac{1}{2}} 2\mu p_3 p_1 FG_{12} + (1-\alpha)^{\frac{1}{2}} 2\mu p_3 p_2 FG_{22}$$

$$+ \left((1-\alpha)^2 (\lambda+2\mu) p_3 p_3 + \rho \frac{\partial}{\partial T^2}\right) FG_{32} = 0 \qquad (5.10)$$

$$(1-\alpha)^{\frac{1}{2}} 2\mu p_3 p_1 FG_{13} + (1-\alpha)^{\frac{1}{2}} 2\mu p_3 p_2 FG_{23}$$

$$+ \left((1-\alpha)^2 (\lambda+2\mu) p_3 p_3 + \rho \frac{\partial}{\partial T^2}\right) FG_{33} = F\delta(\vec{X}-\vec{x})\delta(T-t) \qquad (5.11)$$

Fourier transforms with respect to time are now taken of equations (5.3) through (5.11). As an example, a Fourier transform with respect to time of equation (5.3) is shown as an example as follows:

$$\left(\frac{\lambda+2\mu}{(1-\alpha)}\,p_1p_1 - \rho\omega^2\right)FF_TG_{11} + \frac{2\mu}{(1-\alpha)}\,p_1p_2FF_TG_{12}$$

$$+ (1-\alpha)^{\frac{1}{2}}\,2\mu p_1p_3FF_TG_{13} = FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t) \qquad (5.12)$$

At first glance, equation (5.2) and some related equations would seem to be sufficient for determining the individual elements of the Green's tensor if Parseval's relation and equation (4.19) are used. In actuality, an additional technique must be used. The group of equations, (5.3) through (5.11). in view of equation (5.2) can be expressed as

$$[P]\,[FF_TG] = FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t)] \qquad (5.13)$$

where the matrices $[P]$, $[FF_TG]$, and $[FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t)]$ are expressed as

$$[P] = \begin{bmatrix} \left(\frac{(\lambda+2\mu)p_1p_1}{(1-\alpha)} - \rho\omega^2\right) & \frac{2\mu p_1p_2}{(-\alpha)} & (1-\alpha)^{\frac{1}{2}}\,2\mu p_2p_3 \\ \frac{2\mu}{(1-\alpha)}\,p_1p_2 & \left(\frac{(\lambda+2\mu)p_2p_2-\rho\omega^2}{(1-\alpha)}\right) & (1-\alpha)^{\frac{1}{2}}\,2\mu p_1p_3 \\ (1-\alpha)^{\frac{1}{2}}\,2\mu p_1p_3 & (1-\alpha)^{\frac{1}{2}}\,2\mu p_2p_3 & ((\lambda+2\mu)(1-\alpha)^2p_3p_3-\rho\omega^2) \end{bmatrix} \qquad (5.14)$$

$$[F_tF\,G] = \begin{bmatrix} F_TFG_{11} & F_TFG_{12} & F_TFG_{13} \\ F_TFG_{21} & F_TFG_{22} & F_TFG_{23} \\ F_TFG_{31} & F_TFG_{32} & F_TFG_{33} \end{bmatrix} \qquad (5.15)$$

and

$$[F_TF\delta(\vec{X}-\vec{x})\,\delta(T-t)] = \begin{bmatrix} F_TF\delta(\vec{X}-\vec{x})\,\delta(T-t) & 0 & 0 \\ 0 & F_TF\delta(\vec{X}-\vec{x})\,\delta(T-t) & 0 \\ 0 & 0 & F_TF\delta(X-x)\,\delta(T-t) \end{bmatrix} \qquad (5.16)$$

If now the inverse of the matrix $[P]$ is given as $[P]^{-1}$, then

$$F_T F[G] = [P]^{-1}\left[F_T F \delta(\vec{X}-\vec{x})\,\delta(T-t)\right] \tag{5.17}$$

where

$$F_T F G_{11} = \frac{1}{D}\{(\frac{(\lambda+2\mu)}{(1-\alpha)}p_2p_2-\rho\omega^2)\ ((1-\alpha)^2(\lambda+2\mu)p_3p_3-\rho\omega^2) \tag{5.18}$$
$$-\ (1-\alpha)\ (2\mu)^2(p_2)^2(p_3)^2\}\ FF_T\delta[\vec{X}-\vec{x}]\,\delta(T-t)$$

$$F_T F G_{12} = \frac{1}{D}\ (2\mu)^2(1-\alpha)p_1p_2(p_3)^2$$
$$-\ ((\lambda+2\mu)(1-\alpha)^2p_3p_3-\rho\omega^2)(1-\alpha)^{\frac{1}{2}}2\mu p_1p_3FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t) \tag{5.19}$$

$$F_T F G_{13} = \frac{1}{D}\{\frac{(2\mu)}{(1-\alpha)^{\frac{1}{2}}}p_1(p_2)^2p_3 - (1-\alpha)^{\frac{1}{2}}2\mu p_1p_3(\frac{(\lambda+2\mu)}{(1-\alpha)}\cdot$$
$$\cdot\ p_2p_2-\rho\omega^2)\}FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t) \tag{5.20}$$

$$F_T F G_{21} = \frac{1}{D}\{(1-\alpha)\ (2\mu)^2p_1p_2(p_3)^2 - \frac{2\mu}{(1-\alpha)}\ p_1p_2\ \cdot$$
$$\cdot\ ((1-\alpha)^2(\lambda+2\mu)\ (p_3p_3) - \rho\omega^2)\}FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t) \tag{5.21}$$

$$F_T F G_{22} = \frac{1}{D}\{(\frac{(\lambda+2\mu)}{(1-\alpha)}\ p_1p_1-\rho\omega^2)\ ((1-\alpha)\ (\lambda+2\mu)p_3p_3-\rho\omega^2)$$
$$-\ (1-\alpha)\ (2\mu)^2(p_1)^2(p_3)^2\}FF_T\delta(X-x)\ \delta(T-t) \tag{5.22}$$

$$F_T F G_{23} = \frac{1}{D}\{\frac{(2\mu)^2}{(1-\alpha)^{\frac{1}{2}}}\ (p_1)^2p_2p_3$$
$$-\ (1-\alpha)^{\frac{1}{2}}\ 2\mu p_2p_3(\frac{(\lambda+2\mu)}{(1-\alpha)}\ p_1p_1-\rho\omega^2)\}\ FF_T\delta(\vec{X}-\vec{x})\,\delta(T-t) \tag{5.23}$$

$$F_T FG_{31} = \frac{1}{D}\{\frac{(2\mu)^2}{(1-x)^{\frac{1}{2}}} p_1 (p_2)^2 p_3 - (1-\alpha)^{\frac{1}{2}} 2\mu p_1 p_3 \cdot$$

$$\cdot (\frac{(\lambda+2\mu)}{(1-\alpha)} p_2 p_2 - \rho\omega^2)\} FF_T \delta(\vec{X}-\vec{x})\delta(T-t) \qquad (5.24)$$

$$F_F FG_{32} = \frac{1}{D} \frac{(2\mu)^2}{(1-\alpha)^{\frac{1}{2}}} (p_1)^2 p_2 p_3 - (1-\alpha)^{\frac{1}{2}} 2\mu p_2 p_3 \cdot$$

$$\cdot (\frac{(\lambda+2\mu)}{(1-\alpha)} p_1 p_1 - \rho\omega^2)\} FF_T \delta(\vec{X}-\vec{x})\delta(T-t) \qquad (5.25)$$

$$F_T FG_{33} = \frac{1}{D}\{(\frac{(\lambda+2\mu)}{(1-\alpha)} p_1 p_1 - \rho\omega^2)(\frac{(\lambda+2\mu)}{(1-\alpha)^2} (p_2 p_2 - \rho\omega^2)$$

$$- \frac{(2\mu)_2}{(1-\alpha)^2} (p_1)\ (p_2)^2\} FF_T \delta(\vec{X}-\vec{x})\delta(T-t) \qquad (5.26)$$

$$D = (\frac{\lambda+2\mu}{(1-\alpha)} p_1 p_1 - \rho\omega^2)(\frac{(\lambda+2\mu)}{(1-\alpha)} p_2 p_2 - \rho\omega^2)(\lambda+2\mu)(1-\alpha)^2 p_3 p_3 - \rho\omega^2)$$

$$+ (2\mu)^3 (p_1 p_2 p_3)^2$$

$$- (2\mu)^2 \{(1-\alpha)\ (p_1)^2 (p_3)^2 (\frac{(\lambda+2\mu)}{(1-\alpha)} p_2 p_2 - \rho\omega^2)$$

$$+ (1-\alpha)(p_2)^2 (p_3)^2 (\frac{(\lambda+2\mu)}{(1-\alpha)} p_1 p_1 - \rho\omega^2)$$

$$+ \frac{(p_1)^2 (p_2)^2}{(1-\alpha)^2} ((1-\alpha)^2 (\lambda+2\mu) p_3 p_3 - \rho\omega^2) \qquad (5.27)$$

As an example, the elements of the Green's tensor $[G]$ are obtained from equation (5.18) with equations (5.19) through (5.27) being developed in a similar manner. First rewrite equation (5.18) as

$$\{(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)})\ (\frac{p_2p_2}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)})\ (\frac{p_3p_3}{(1-\alpha)}(1-\alpha)^2$$

$$- \frac{\rho\omega^2}{(\lambda+2\mu)})\ (\lambda+2\mu)^3 + (2\mu)^3(p_1p_2p_3)^2$$

$$- (2\mu)^2(\lambda+2\mu)\ \{(1-\alpha)\ (p_1)^2(p_3)^2(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}\}$$

$$+ (1-\alpha)(p_2)^2(p_3)^2(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}$$

$$+ \frac{(p_1)^2(p_2)^2}{(1-\alpha)^2}\ ((1-\alpha)^2p_3p_3 - \frac{\rho\omega^2}{(\lambda+2\mu)})\ \}FF_TG_{11}$$

$$= \{(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)})\ (p_3p_3(1-\alpha)^2-\rho\omega^2) - (1-\alpha)\ (2\mu)^2(p_2)^2\ \cdot$$

$$\cdot\ (p_3)^2\}FF_T\delta(\vec{X}-\vec{x})\delta(T-t) \qquad (5.28)$$

equation (5.28) is now developed as

$$\{(\alpha+2\mu)^3 + (2\mu)^3\left(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right)\left(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right)(p_3p_3(1-\alpha)^2$$

$$- \frac{\rho}{(\lambda+2\mu)}\omega^2)(p_1p_2p_3)^2 - (2\mu)^2(\lambda+2\mu) \cdot$$

$$\cdot \{(1-\alpha)\left(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right)(p_3p_3(1-\alpha)^2$$

$$- \frac{\rho}{(\lambda+2\mu)}\omega^2)(p_1)^2(p_3)^2 + (1-\alpha)\left(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right) \cdot$$

$$\cdot (p_3p_3(1-\alpha)^2 - \frac{\rho}{(\lambda+2\mu)}\omega^2)(p_2)^2(p_3)^2\} \cdot$$

$$\left\{\frac{1}{\left(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right)\left(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right)(p_3p_3(1-\alpha)^2 - \frac{\rho}{(\lambda+2\mu)}\omega^2)}\right\} FF_TG_{11}$$

$$= \{\left(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}\right) - (1-\alpha)(2\mu)^2\left(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho}{(\lambda+2\mu)}\omega^2\right) \cdot$$

$$\cdot \left(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}\right)(p_3p_3(1-\alpha)^2 - \frac{\omega^2}{(\lambda+2\mu)})(p_2)^2(p_3)^2\}$$

$$\left\{\frac{1}{\left(\frac{p_1p_1}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}\right)\left(\frac{p_2p_2}{(1-\alpha)} - \frac{\rho\omega^2}{(\lambda+2\mu)}\right)(p_3p_3(1-\alpha)^2 - \frac{\rho\omega^2}{(\lambda+2\mu)}}\right\} FF_T\delta(\vec{X}-\vec{x})\delta(T-t)$$

(5.29)

which with the aid of Parseval's relation and with the operator's previously defined is evaluated as

$$G_{11}(X_1,X_2,X_3,T;\ x_1,x_2,x_3,t) = \frac{1}{DP}\{(\exp(i(X_2(1-\alpha)^{\frac{1}{2}} + \frac{X_3}{(1-\alpha)}$$

$$- (\frac{\lambda+2\mu}{\rho})\ T))^{\frac{1}{2}} - (2\mu)^2\}\ [\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.30)$$

where

$$DP = (\lambda+2\mu)^3\{\exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}} + \frac{(X_3)}{(1-\alpha)} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}$$

$$+ (2\mu)^3 - (\lambda+2\mu)^2(2\mu)^2\{\exp(i((X_1(1-\alpha)^{\frac{1}{2}}$$

$$- (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) + \exp(i((X_2(1-\alpha)^{\frac{1}{2}} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))$$

$$+ \exp((i(\frac{X_3}{(1-\alpha)} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\} \qquad (5.31)$$

Continuing in the same manner, then the other elements of the Green's tensor are given as

$$G_{12}(X_1,X_2,X_3,T;\ x_1,x_2,x_2,t) = \frac{1}{DP}\{(2\mu)^2\exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}}$$

$$+ \frac{X_3}{1-\alpha} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - (\lambda+2\mu)^2\mu(\exp(i((X_3(1-\alpha)^{\frac{1}{2}}$$

$$- (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\{\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.32)$$

$$G_{13}(X_1,X_2,X_3,T;\ x_1,x_2,x_2,t) = \frac{1}{DP}\{(2\mu)^2\exp(i(((X_1+X_2)(1-\alpha)^{\frac{1}{2}}$$

$$+ \frac{X_3}{1-\alpha} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - 2\mu(\lambda+2\mu)\exp((i(X_2(1-\alpha)^{\frac{1}{2}} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\cdot$$

$$\cdot\{\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.33)$$

$$G_{21}(X_1,X_2,X_3,T;\ x_1x_2x_3t) = \frac{1}{DP}\{(2\mu)^2(\exp(i((X_1+X_2)^{\frac{1}{2}})(1-\alpha)^{\frac{1}{2}}$$
$$+ X_3(1-\alpha) - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - (\lambda+2\mu)^2\ \mu\ \exp(i((X_3(1-\alpha)^{-1}$$
$$- (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\{\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.34)$$

$$G_{22}(X_1,X_2,X_3,T;\ x_1,x_2,x_3,t) = \frac{1}{DP}\{(\lambda+2\mu)^2\ \exp(i((X_1)(1-\alpha)^{\frac{1}{2}}$$
$$+ X_3(1-\alpha)^{-1} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - (2\mu)^2\ \exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}}$$
$$+ X_3(1-\alpha)^{-1} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\{\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.35)$$

$$G_{23}(X_1,X_2,X_3,T;\ x_1,x_2,x_3,t) = \frac{1}{DP}\{(2\mu)^2\ \exp(i((X_1+X_2)\ (1-\alpha)^{\frac{1}{2}}$$
$$+ X_3(1-\alpha)^{-1} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - (2\mu)\ (\lambda+2\mu)\ \exp(i((X_1(1-\alpha)^{\frac{1}{2}}$$
$$- (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\{\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.36)$$

$$G_{31}(X_1,X_2,X_3,T;\ x_1,x_2,x_2,t) = \frac{1}{DP}\{(2\mu)^2\ \exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}}$$
$$+ X_3(1-\alpha)^{-1} - (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - 2\mu(\lambda+2\mu)(\exp(i(X_2(1-\alpha)^{\frac{1}{2}}$$
$$- (\frac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\{\delta(\vec{X}-\vec{x})\delta(T-t)\} \qquad (5.37)$$

$$G_{32}(X_1,X_2,X_3,T;\ x_1,x_2,x_3,t) = \frac{1}{DP}\{(2\mu)^2(\exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}} + X_3(1-\alpha)^{-1} - (\tfrac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))] - (\lambda+2\mu)\exp(i(X_1(1-\alpha)^{\frac{1}{2}} - (\tfrac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))][\delta(\vec{X}-\vec{x})\delta(T-t)] \tag{5.38}$$

$$G_{33}(X_1,X_2,X_3,T;\ x_1x_2x_3t) = \frac{1}{DP}\{(\lambda+2\mu)^2\exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}} - (\tfrac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T)) - (2\mu)^2\exp(i((X_1+X_2)(1-\alpha)^{\frac{1}{2}} + X_3(1-\alpha)^{-1} - (\tfrac{\lambda+2\mu}{\rho})^{\frac{1}{2}}T))\}\{(\vec{X}-\vec{x})\delta(T-t)\} \tag{5.39}$$

VI. THE DISPLACEMENT FUNCTIONS.

With a Green's tensor having been developed, then the displacement functions can be obtained. Consider the force that is causing the displacement to be expressed as a vector, i.e.

$$\vec{F} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} \tag{6.1}$$

with the displacements being obtained as

$$\vec{U} = [G]\ \vec{F}\ d\vec{X}\ dT \tag{6.2}$$

Expressed in component form, then

$$U_1 = \int (G_{11}F_1 + G_{12}F_2 + G_{13}F_3)\,dX_1\,dX_2\,dX_3\,dT \tag{6.3}$$

$$U_2 = \int (G_{21}F_2 + G_{22}F_2 + G_{23}F_3)dX_1dX_2dX_3dT \quad (6.4)$$

$$U_3 = \int (G_{31}F_3 + G_{32}F_2 + G_{33}F_3)dX_1dX_2dX_3dT \quad (6.5)$$

The displacement functions as derived here satisfy initial conditions since they are equal to zero when t = 0. If a dip strike is considered F_1 and F_2 are equal to zero and F_3 is obtained from Equation (1.1) such that

$$U_1 = \int G_{13}F_3(t)dX_1dX_2dX_3dT \quad (6.6)$$

$$U_2 = \int G_{23}F_3(t)dX_1dX_2dX_3dT \quad (6.7)$$

$$U_3 = \int G_{33}F_3(t)dX_1dX_2dX_3dT \quad (6.8)$$

VII. DISSCUSSION OF RESULTS. The rather crude model for describing an earthquake as developed in this dissertation does contribute to an understanding of observed seismograms. The computed results which were obtained by a program by Council (1) were for a ten second period of time which is consistent with the period of time that the maximum effects of an earthquake are observed in which the maximum accelerations associated with a slip strike would increase from zero to .5g with reversals of the direction of the acceleration occurring at random time intervals. Figure 7.1 represents a time history of the accelerations associated with a dip strike as obtained by using Equation (1.1) for the displacement function equations, Equations (6.6), (6.7) and (6.8). Time histories of the accelerations obtained from Equations (6.6), (6.7) and (6.8) for a ten second period of time are displayed in Figures 7.2, 7.3 and 7.4. A comparison of these last three figures with the one previous show the effect of the medium through which the disturbance is propagated. A one second time history of the accelerations as obtained from Equations (6.6), (6.7) and (6.8) are shown superimposed in Figure 7.5. The tendency for a lack of coherence of time histories of the accelerations along each of the coordinate axis can be considered as a reflection of the surrounding medium being prestressed and is consistent with observed seismograms. If instead, these time histroies for the accelerations along each of the coordinate axis were considered to be accelerations resulting from multiple foci, then this would help to explain some of the complexities of observed seismograms. The amplitudes of the frequency spectrums of the accelerations along each of the coordinate axis are shown in Figures 7.6, 7.7 and 7.8. The differences between the results presented here and actual frequency spectrums associated with free fields is an indication that a frequency dependent attenuation factor should have been included in the equations that were derived.

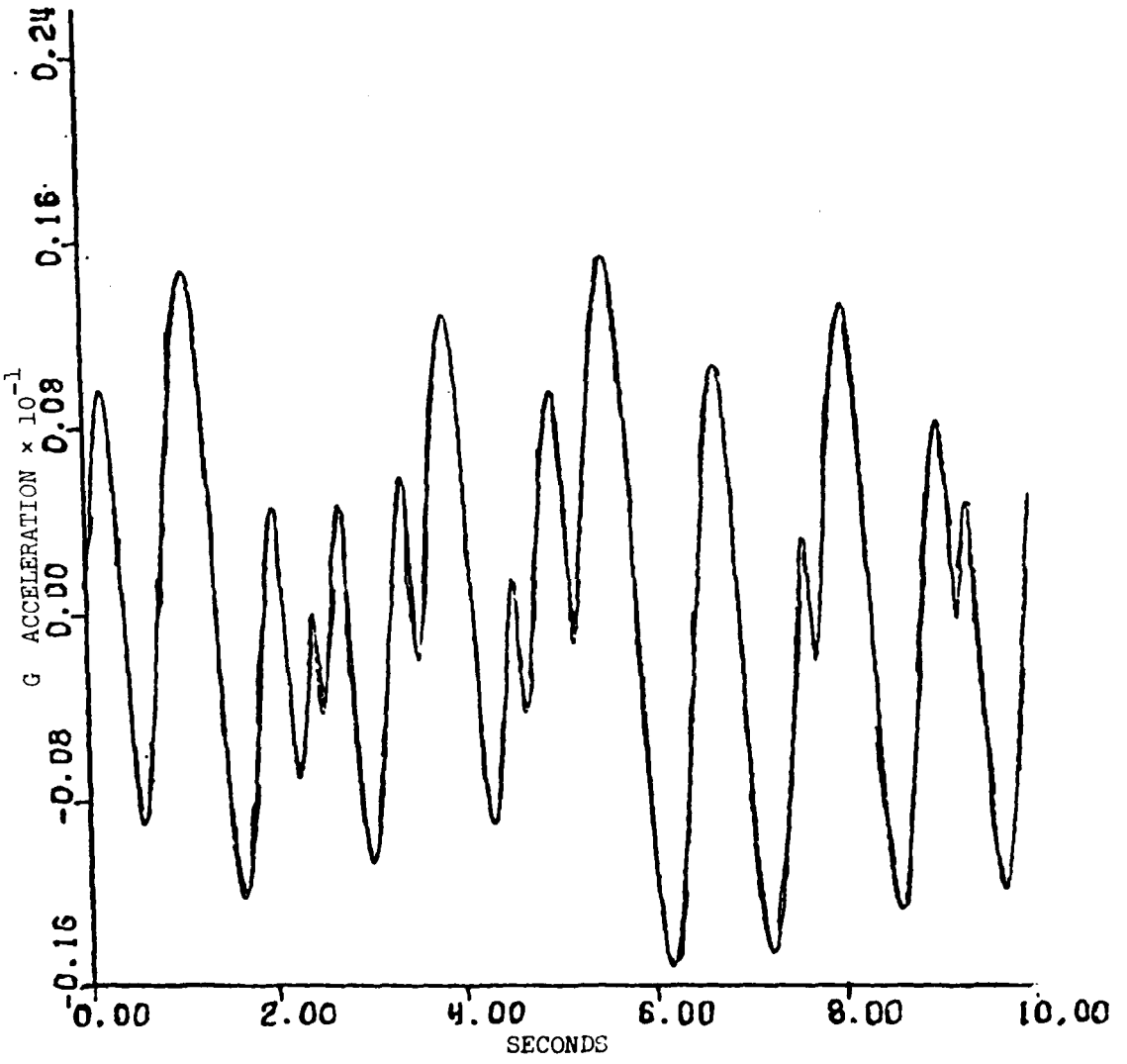


FIGURE 7.1 - TIME HISTORY OF THE FORCING FUNCTION REPRESENTING ACCELERATIONS ASSOCIATED WITH A DIP STRIKE

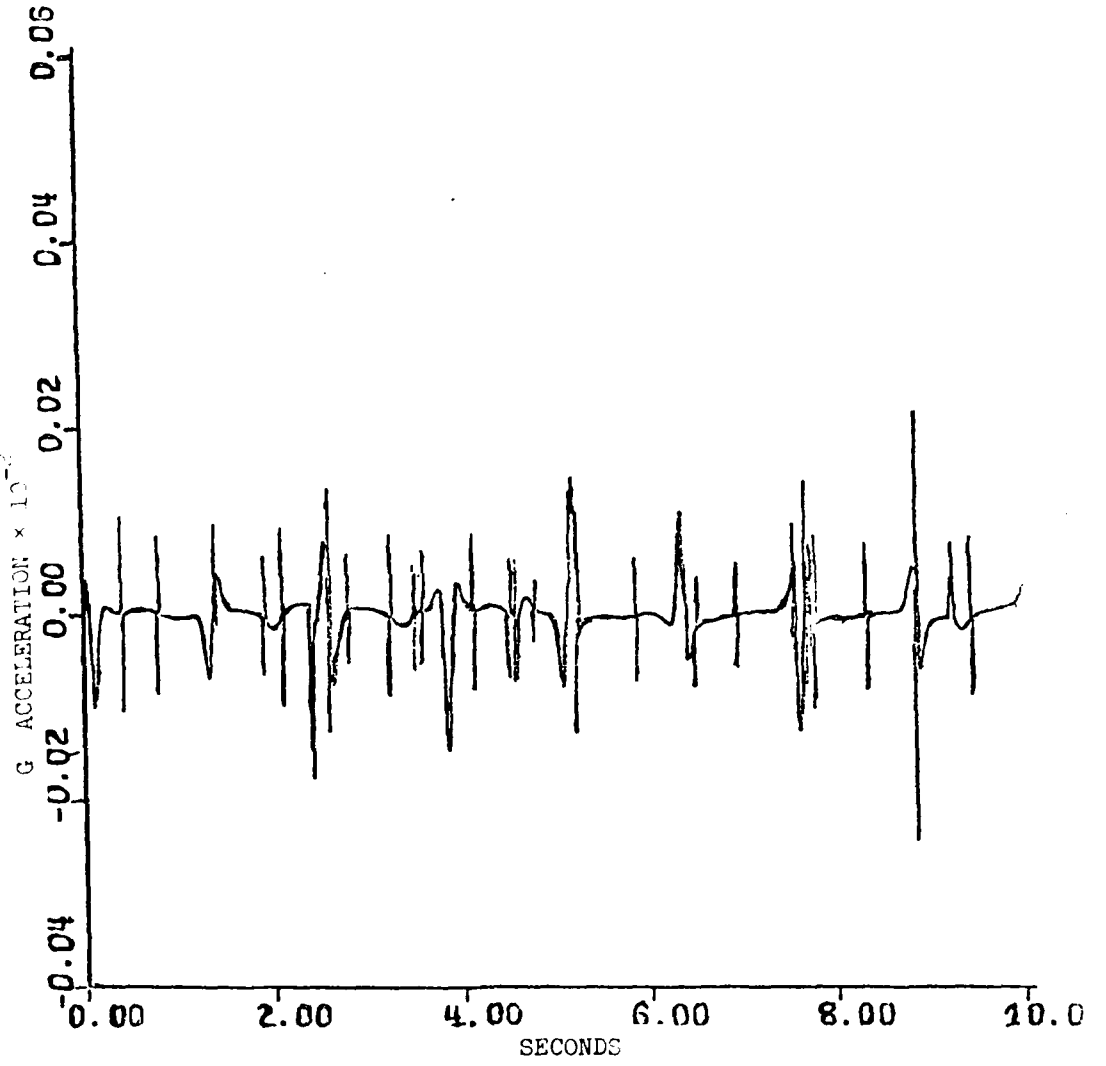


FIGURE 7.2 - TIME HISTORY OF THE ACC. IN SURROUNDING MEDIUM ALONG THE X_1 COORDINATE AXIS

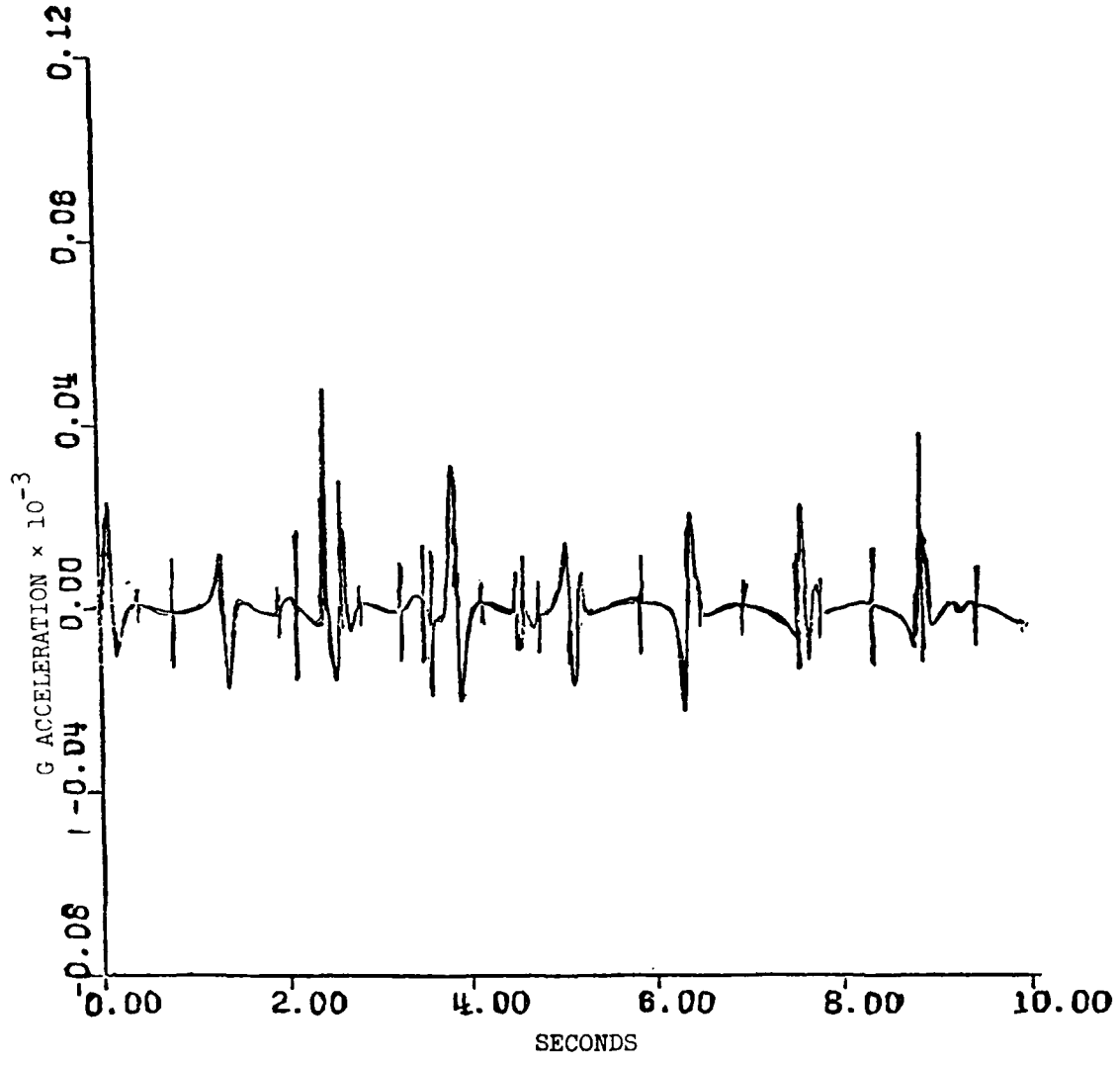


FIGURE 7.3 - TIME HISTORY OF THE ACC. IN SURROUNDING MEDIUM ALONG THE X_2 COORDINATE AXIS

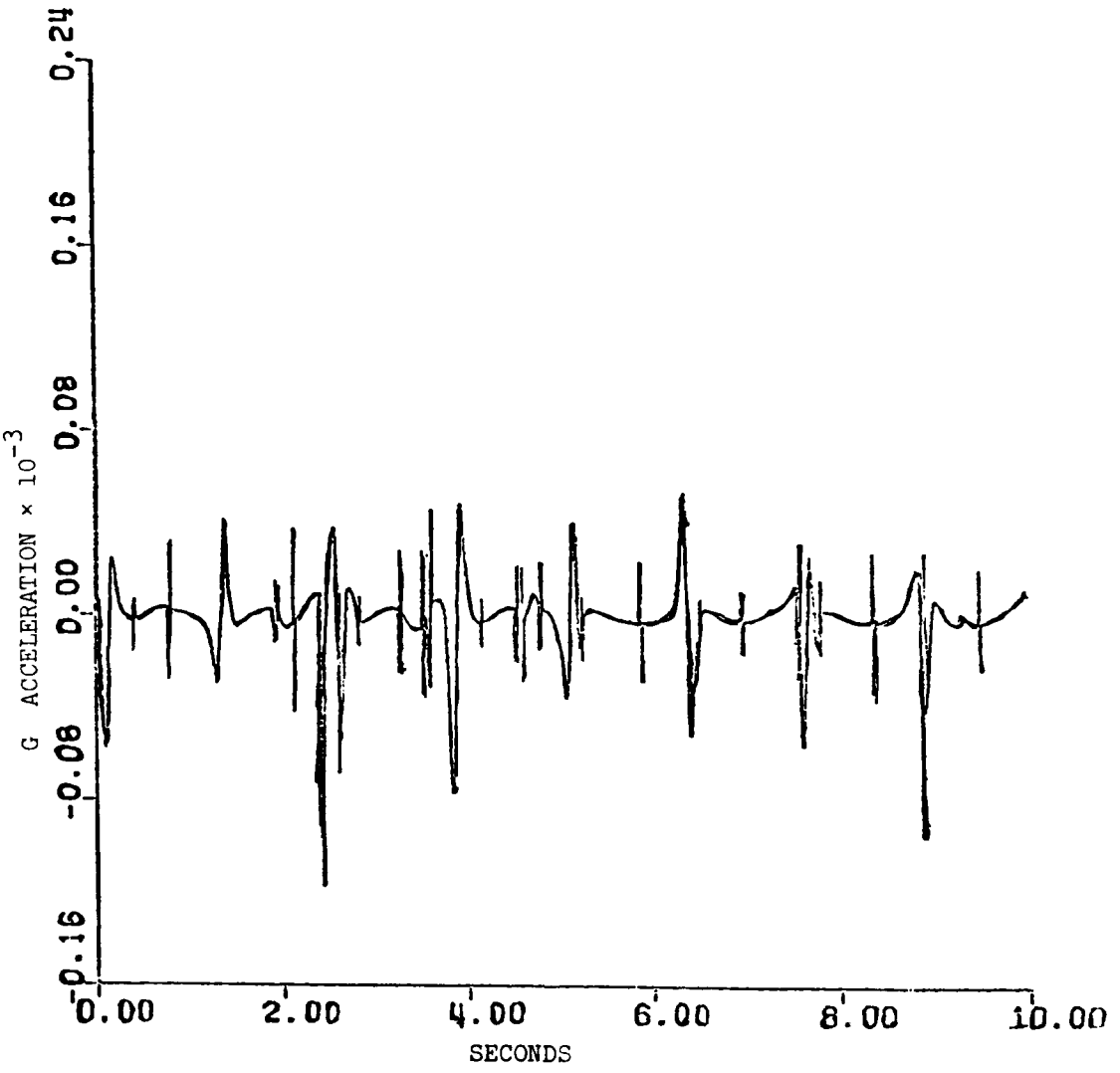


FIGURE 7.4 - TIME HISTORY OF THE ACC. IN SURROUNDING MEDIUM ALONG THE X_3 COORDINATE AXIS

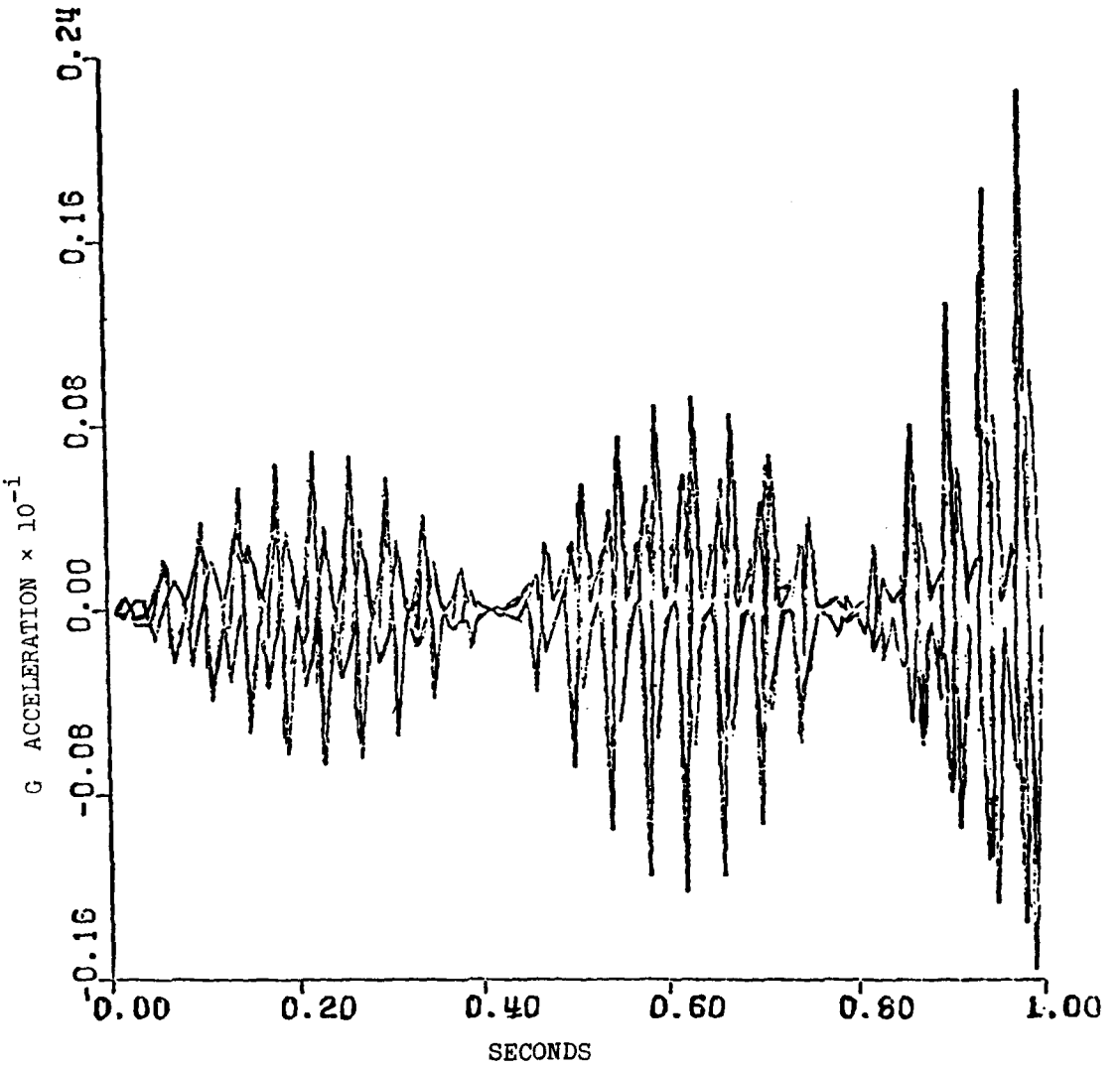


FIGURE 7.5 - ACCELERATION IN THE SURROUNDING MEDIUM, SUPERIMPOSED

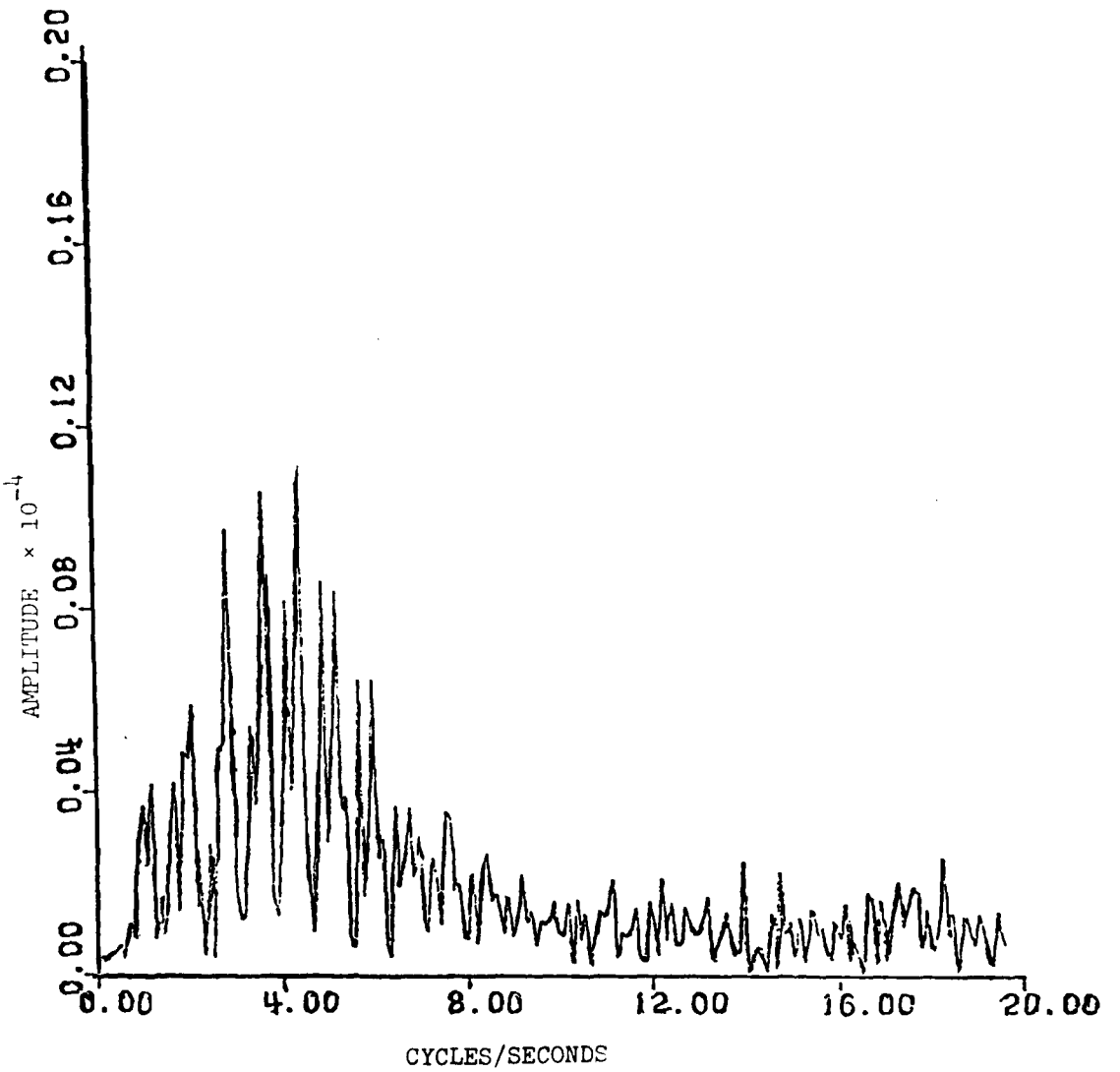


FIGURE 7.6 - FREQ. SPECTRUM OF ACC. IN SUR. MED.
ALONG THE X_1 COORDINATE AXIS

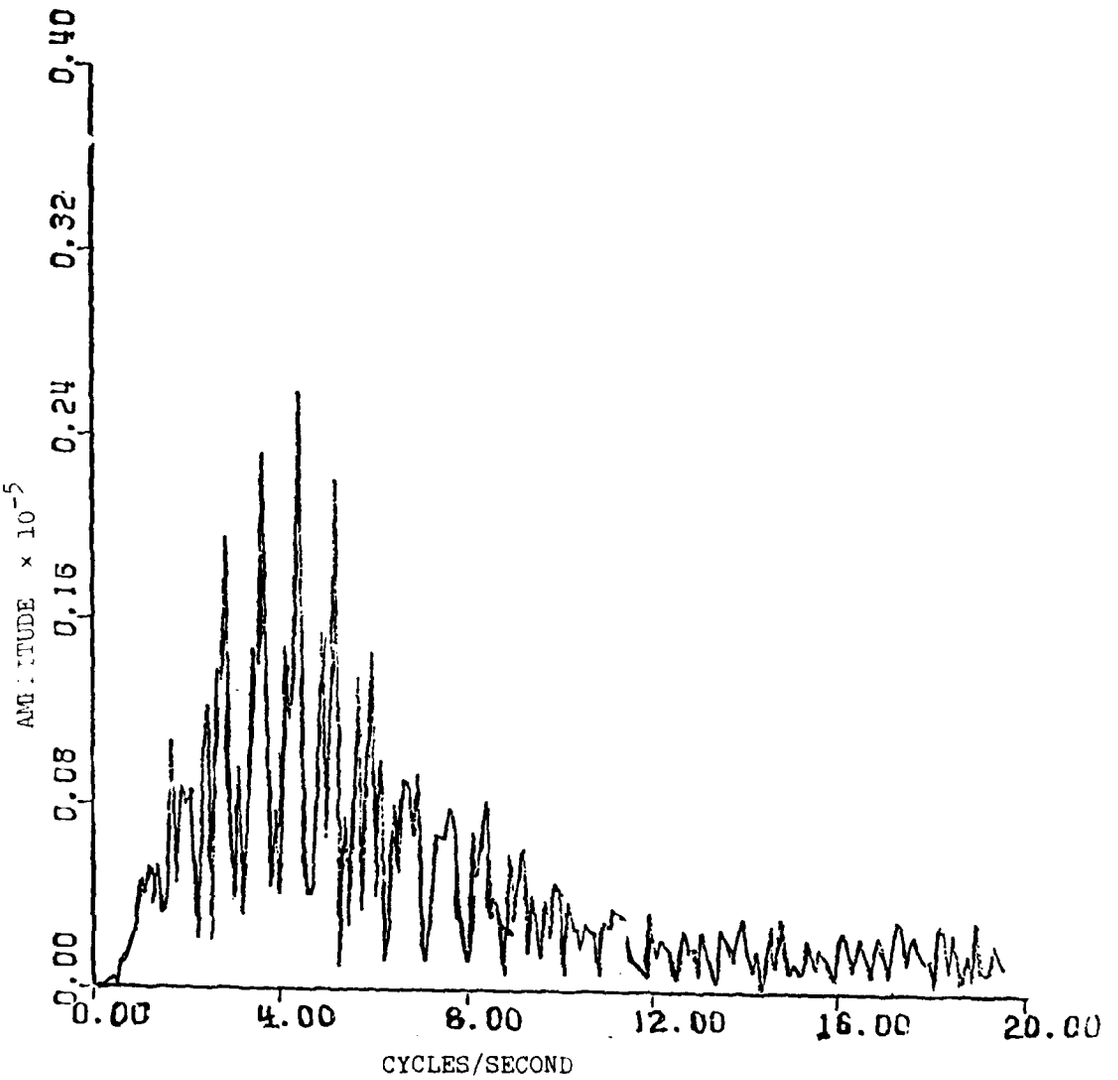


FIGURE 7.7 - FREQ. SPECTRUM OF ACC. IN SUR. MED.
ALONG THE X_2 COORDINATE AXIS

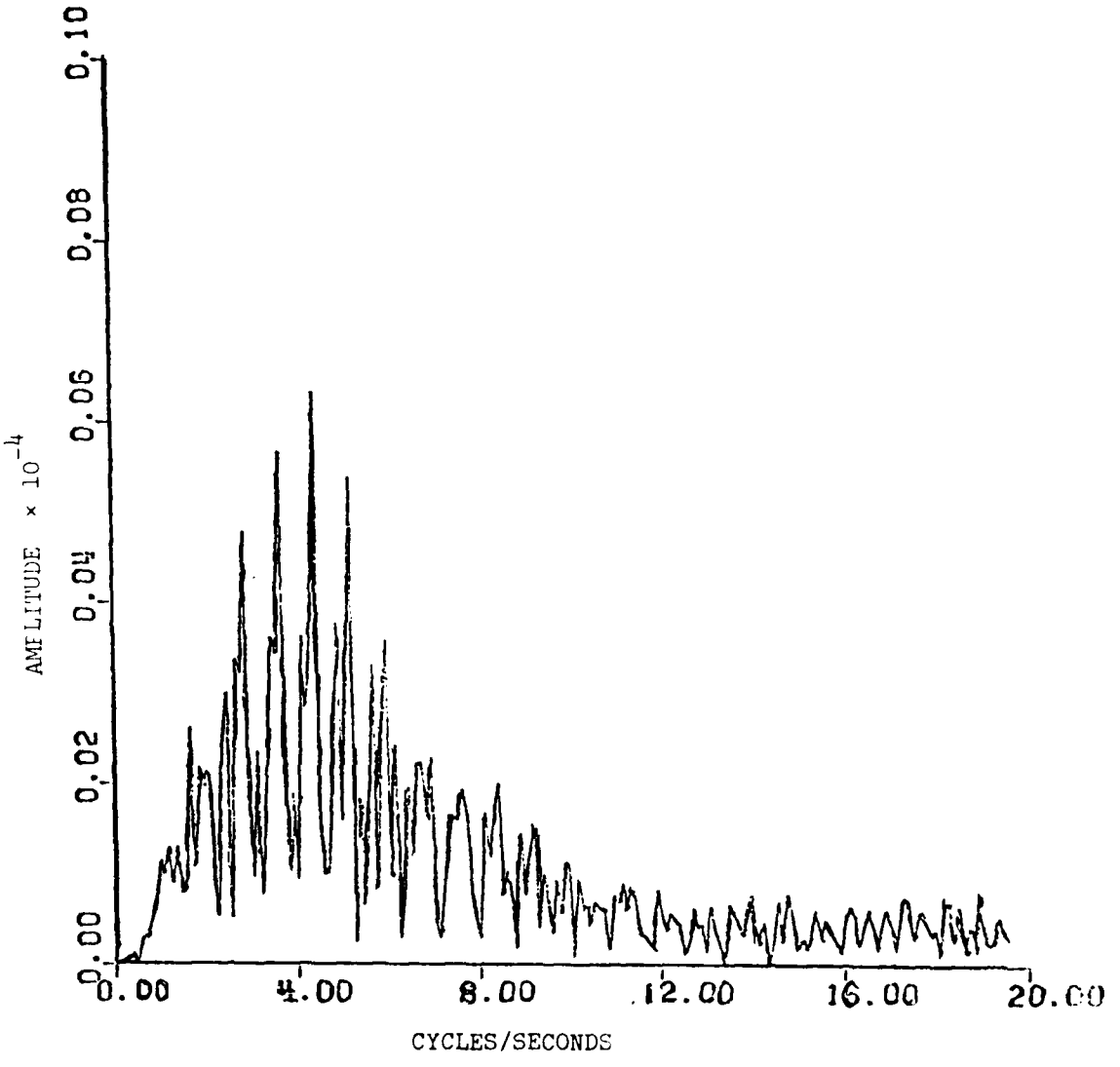


FIGURE 7.8 - FREQ. SPECTRUM OF ACC. IN SUR. MED.
ALONG THE X_3 COORDINATE AXIS

Although a dip strike was considered in this discussion, the equations can be modified such that a slip strike or a combination of a slip strike and a dip strike are considered. Also, a multi foci situation can be considered which is more consistant with an actual earthquake.

LIST OF REFERENCES

(1) Council, F. E., Dynamics of A Flat Plate Foundation During An Earthquake, Dissertation, Geroge Washington University, Washington, D. C. (1978).

(2) Courant, R. and Hilbert, D., Methods of Mathematical Physics, Interscience, New York, NY, (1953), Chapter 5.

(3) Eringin, A. Cemal, Nonlinear Theory of Continuous Media, McGraw Hill, New York, NY, (1962), Chapter 3.

(4) Green, A. E. and Zerna, W., Theoretical Elasticity, 2nd ed., Oxford, London, (1968), Chapters II, III.

(5) Truesdell, C., "General and Exact Theory of Waves in Finite Elastic Strain," Archive for Rational Mechanics and Analysis, Vol. 8, (1951).

(6) Sneddon, Ian S., Fourier Transforms, McGraw Hill, New York, NY, (1951), p 25.

AN ITERATIVE ALGORITHM FOR CALCULATING POTENTIALS NEAR TWO PARALLEL PLATES OF EQUAL WIDTH, PART II

J. Barkley Rosser
Mathematics Research Center
University of Wisconsin, Madison, Wisconsin

ABSTRACT. This is an extension of the report "An iterative algorithm for calculating potentials near small groups of finite charged plates" by Acton and Rosser, that was presented at the Seventeenth Conference of Army Mathematicians in 1971. Since the first report, it has been possible to determine the rate of convergence of the algorithm. It is reported how to accelerate the convergence by the ε-algorithm. It is also shown how to use the Fast Fourier Transform to reduce the labor of calculation.

1. SUMMARY. Good approximations for the potentials around condenser plate arrangements found in many pieces of equipment in electron optics can be got by passing a plane through the middle of the condenser and solving for a 2-dimensional potential in this plane. In [1], an iterative procedure was explained which can be used to get the required 2-dimensional potential in the neighborhood of a finite number of finite charged plates, however they are arranged.

In the electrostatic lenses of cathode-ray tubes, one has the particularly simple case of two parallel plates of equal width, directly opposite each other. One can find the potential by elliptic integrals; see [2]. One can also get a numerical approximation by calculations with a singular integral; see [3]. The iterative procedure of [1] was tried for such a condenser in which the separation of the plates was 1.25 times the plate width. The convergence was very fast, requiring about 1/10 the computational labor of either of the methods given in [2] or [3].

In this part, this situation of parallel plates is studied for general separation ratios; the separation ratio is the ratio of the distance between the plates to the width of the plates. If one has a solution for a condenser of one size with a given separation ratio, one can get a solution for a larger or smaller condenser with the same separation ratio by using the obvious scale factor.

It will be shown that the iterative procedure of [1] converges for each separation ratio. A formula is determined which gives the rate of convergence in terms of the separation ratio. As the separation ratio decreases, the rate of convergence also decreases; also, the calculations become more extensive for each step of the iteration. For small separation ratios, the calculations can be considerably abridged by using the Fast Fourier Transform. Details will be given. Also, for small separation ratios, the rate of convergence can be much accelerated by using the ε-algorithm. Details will be given. Because of these improvements, the iterative procedure of [1] appears to be relatively efficient, whatever the separation ratio.

I wish to acknowledge the help of H.-S. Hung and T.-J. Huang in the preparation of this part, and the assistance of Dianne Hollenbeck in the programming and calculations.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

Calculations were made, and are summarized, for each of the separatic ratios 1.25, 0.5, 0.2, 0.1, and 0.01. As a check, for each of these, the key results were verified by the methods of [2].

2. BASIC IDEAS. It will be good first to review some of the ideas of [1]. In space, we have the two parallel plates. We pass a plane perpendicular to both, and seek a 2-dimensional potential in the plane. In the plane, the plates appear as two parallel straight line segments, as in Figure 2.1. Here we have two plates, each of width ℓ, separated by a space of 2 units. The separation ratio is $2/\ell$. If a different size is required, one scales up or down. We wish to find a 2-dimensional potential in the plane of which Figure 2.1 is a part. The actual

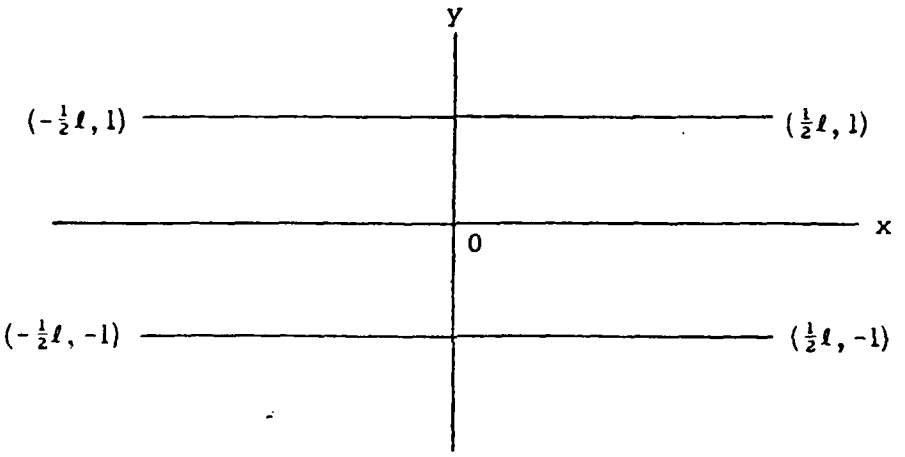


Figure 2.1.

plates are considered to extend indefinitely in both directions in the direction perpendicular to the plane of Figure 2.1.

Since the plates are conductors, the potential must be constant along each plate. We also require a zero potential at infinity, from which we conclude that the potentials along the two plates must be $+C$ and $-C$ respectively. If one had preferred potentials of $+K$ and $-K$, one could multiply the potential for the present case by K/C at all points.

So, we wish to find a real function X which will be continuous in the entire plane, and harmonic except on the two "plates"; we further require that X approach zero as $x^2 + y^2$ goes to infinity, that X be $+C$ along the upper "plate", and that X be $-C$ along the lower "plate", with $C \neq 0$.

We introduce the complex variable

$$z = x + iy ;$$

we will define X as the real part of a certain function of z. In the z-plane the upper "plate" is the line segment connecting the two points $(\pm\frac{1}{2}\ell + i)$ and the lower "plate" is the line segment connecting the two points $(\pm\frac{1}{2}\ell - i)$.

The discussion in [1] leads to the following approach. Define $w(s)$ for complex s by

$$s = w + \frac{1}{w} , \tag{2.1}$$

$$|w| < 1 . \tag{2.2}$$

Except for s on a cut consisting of the line segment from -2 to $+2$, (2.1) and (2.2) define w as a single valued analytic function of s. The function $w(s)$ maps all of the s-plane except the cut conformally in a one-to-one manner into the interior of the unit circle in the w-plane. We shall shortly consider the details of this mapping. In particular, if s approaches the point $2\cos\theta$ on the cut from above, then w approaches $e^{-i\theta}$, with $0 \leq \theta \leq \pi$, from the interior of the unit circle. If s approaches the point $2\cos\theta$ on the cut from below, then w approaches $e^{i\theta}$ from the interior of the unit circle.

Let $S(s)$ be a function with the properties:

$$\mathcal{R}\{S(s)\} \text{ is harmonic except on the cut;} \tag{2.3}$$

$$\mathcal{R}\{S(s)\} \text{ is continuous everywhere;} \tag{2.4}$$

$$\lim_{s\to\infty} \mathcal{R}\{S(s) - S(s+\alpha)\} = 0 \quad \text{for each } \alpha . \tag{2.5}$$

Let us suggest

$$\mathcal{R}\{S(4(z-i)/\ell) - S(4(z+i)/\ell)\} \tag{2.6}$$

as the first approximation for the function X that we seek. By (2.4), (2.6) is continuous in the entire plane. By (2.3), (2.6) is harmonic except on the two "plates". By (2.5), (2.6) goes to zero as $x^2 + y^2$ goes to infinity. What we lack is that (2.6) should be constant along each "plate".

Can we find a correction for (2.6) that makes it more nearly constant along each "plate"? A way to do this is as follows. We transform the z-plane, except for the upper "plate", conformally into the interior of the unit circle in the w-plane. This is done in two steps. First transform the z-plane, minus the upper "plate", into the s-plane, minus the cut, by putting

$$z = \frac{\ell s}{4} + i . \tag{2.7}$$

Then transform the s-plane, minus the cut, into the interior of the unit circle in the w-plane by (2.1) and (2.2).

Under this transformation, the function (2.6) goes into a real function $T(r,\theta)$, for $0 < r < 1$, in the w-plane, where we are taking

$$w = r\, e^{i\theta} .$$

We have that $T(r,\theta)$ is continuous for $0 \leq r < 1$. Since a harmonic function goes into a harmonic function under a conformal map, $T(r,\theta)$ for $0 < r < 1$ is harmonic except on the image in the w-plane of the lower "plate".

By continuity, we can extend $T(r,\theta)$ to $r = |w| = 1$. At $w = e^{i\theta}$, for $-\pi < \theta \leq \pi$, we assign $T(1,\theta)$ the value

$$\mathcal{R}\{S(2\cos\theta) - S(2\cos\theta + \tfrac{8}{\ell}i)\} .$$

This makes $T(r,\theta)$ continuous for $0 \leq r \leq 1$. Also, $T(1,\theta)$ is an even function of θ, continuous, and with period 2π.

Expand $-T(1,\theta)$ in a Fourier series

$$-T(1,\theta) = \beta_0 + \sum_{n=1}^{\infty} (\beta_n \cos n\theta + \gamma_n \sin n\theta). \tag{2.8}$$

As $T(1,\theta)$ is an even function of θ, the γ_n will all be 0. We have of course

$$\beta_n = -\frac{1}{\pi}\int_{\pi}^{\pi} T(1,\theta)\cos n\theta \; d\theta$$

for $n \geq 1$. As $T(1,\theta)$ is an even function of θ, we have for $n \geq 1$

$$\beta_n = -\frac{2}{\pi}\int_{0}^{\pi} T(1,\theta)\cos n\theta \; d\theta . \tag{2.9}$$

Let, temporarily,

$$u(w) = \sum_{n=1}^{\infty} \beta_n w^n . \tag{2.10}$$

Then obviously $u(w)$ is analytic inside the unit circle and continuous inside and including the unit circle. Also $u(0) = 0$. We have of course

$$\mathcal{R}\{u(e^{i\theta})\} = \sum_{n=1}^{\infty} \beta_n \cos n\theta . \tag{2.11}$$

So, since the γ_n are 0, we have by (2.8)

$$T(1,\theta) + \mathcal{R}\{u(e^{i\theta})\} = -\beta_0 \tag{2.12}$$

for all θ. That is, $u(w)$ has completely smoothed out the fluctuations of $T(r,\theta)$ around the unit circle. If we transform $T(r,\theta) + \mathcal{R}\{u(w)\}$ back to the z-plane, we have something which is constant along the upper "plate". Of course, we have produced additional fluctuations along the lower "plate". These reflect the fluctuations of $u(w)$ along the image in the w-plane of the lower "plate". But here $|w| < 1$, so that from the definition (2.10) one would expect smaller fluctuations than those we smoothed out for $|w| = 1$.

Transforming $u(w)$ back to the z-plane gives

$$u(w(4(z-i)/\ell)) = \sum_{n=1}^{\infty} \beta_n (w(4(z-i)/\ell))^n \tag{2.13}$$

of course, by (2.10).

This suggests using

$$\bar{S}(s) = S(s) + u(w(s)) \tag{2.14}$$

as an improvement for $S(s)$. We verify that $\bar{S}(s)$ satisfies (2.3) and (2.5) by appealing to the properties of $u(w)$ cited below (2.10). Specifically, we verify (2.3) since for s not on the cut we have $|w(s)| < 1$, so that $u(w(s))$ is analytic. Hence $\mathcal{R}\{u(w(s))\}$ is harmonic. We verify (2.5) since $w(s)$ goes to zero as s goes to infinity. To verify (2.4) for $\bar{S}(s)$, the only difficulty is along the cut. But at $s = 2\cos\theta$ along the cut, we have

$$w(s) = e^{\pm i\theta}$$

so that

$$\mathcal{R}\{w(s)\} = \sum_{n=1}^{\infty} \beta_n \cos n\theta = -T(1,\theta) - \beta_0 .$$

So things are O.K.

Now the new approximation for the function X that we seek is

$$\mathcal{R}\{S(4(z-i)/\ell) - S(4(z+i)/\ell)\} + \mathcal{R}\{u(w(4(z-i)/\ell))\} - \mathcal{R}\{u(w(4(z+i)/\ell))\}. \qquad (2.15)$$

If we should leave off the final term of (2.15), the result would be constant along the top "plate". For reasons given above, we expect the fluctuations caused along the top "plate" by the final term of (2.15) to be less than we had for (2.6). For analogous reasons, we expect that the final term of (2.15) will considerably reduce the fluctuations along the lower "plate".

If this is really so (and we will prove that it is), we have improved the situation. Then, of course, we should repeat the operation, to try for further improvement. So, again, we transform the z-plane, except for the upper plate, conformally into the interior of the unit circle in the w-plane. As before, the first term of (2.15) goes into $T(r,\theta)$. The second term of (2.15) goes into $\mathcal{R}\{u(w)\}$. The third term of (2.15) goes into

$$-\mathcal{R}\{u(w(w + \frac{1}{w} + \frac{8}{\ell} i))\}. \qquad (2.16)$$

We now extend to $|w| = 1$. For $w = e^{i\theta}$, the first two terms of (2.15) together go to $-\beta_0$, because that is how we chose $u(w)$. So, for $w = e^{i\theta}$, the entire formula (2.15) goes to

$$-\beta_0 - \mathcal{R}\{u(w(2\cos\theta + \frac{8}{\ell} i))\} . \qquad (2.17)$$

We expand the negative of this in a Fourier series

$$\bar{\beta}_0 + \sum_{n=1}^{\infty} \bar{\beta}_n \cos n\theta$$

(as before, the sine terms drop out) where for $n \geq 1$

$$\bar{\beta}_n = \frac{2}{\pi} \int_0^{\pi} \mathcal{R}\{u(w(2\cos\theta + \frac{8}{\ell} i))\} \cos n\theta \, d\theta.$$

In (2.10), the β_n's are bounded (by (2.9)). Also

$$|w(2\cos\theta + \frac{8}{\ell} i)| < 1$$

by (2.2). So the terms in the series for

$$\mathcal{R}\{u(w(2\cos\theta + \frac{8}{\ell}\, i))\}$$

are bounded by a geometric series. Hence we can interchange the order of summation and integration, which gives us for $n \geq 1$

$$\bar{\beta}_n = \frac{2}{\pi} \sum_{m=1}^{\infty} \beta_m \int_0^{\pi} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell}\, i))^m\} \cos n\theta \, d\theta. \tag{2.18}$$

We could now define a $\bar{u}(w)$ in terms of the $\bar{\beta}_n$ analogously to (2.10). Then, analogously to (2.14), we could get a still better approximation by using $\bar{S}(s) + \bar{u}(w(s))$.

We proceed in this way, successively, until we get to an approximation that is as near constant as we wish along each of the two "plates". The important thing to observe is that each new set of β's is defined in terms of the previous set by (2.18). Hence, we can write a computer program to calculate successive sets of β's. That is, we can do that as soon as we learn enough about the function $w(s)$. We turn now to that.

3. PROPERTIES OF $w(s)$. Somewhat more generally than (2.1), we define s as a function of w by

$$s = w + \frac{1}{w} \tag{3.1}$$

for the entire w-plane. Clearly this defines w as a double valued function of s by the equation

$$w = \tfrac{1}{2}\{s - (s^2 - 4)^{\frac{1}{2}}\}. \tag{3.2}$$

To make this single valued, we make the determination

$$(s^2 - 4)^{\frac{1}{2}} = i\sqrt{5} \quad \text{when} \quad s = i\,; \tag{3.3}$$

it suffices then to make suitable cuts in the s-plane. Two choices are useful: a singly connected cut (SCC), and a doubly connected cut (DCC); see Figure 3.1.

With SCC, $(s^2 - 4)^{\frac{1}{2}}$ is an odd function of s;

$$((-s)^2 - 4)^{\frac{1}{2}} = -(s^2 - 4)^{\frac{1}{2}} \quad \text{for SCC.} \tag{3.4}$$

With DCC, $(s^2 - 4)^{\frac{1}{2}}$ is an even function of s;

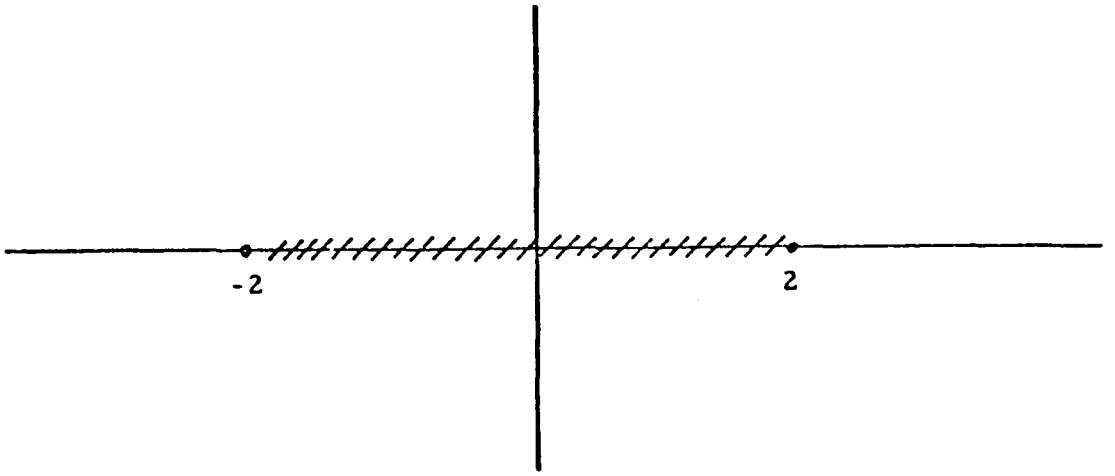
$$((-s)^2 - 4)^{\frac{1}{2}} = (s^2 - 4)^{\frac{1}{2}} \quad \text{for DCC.} \tag{3.5}$$

Since $(s^2 - 4)^{\frac{1}{2}}$ is single valued if we make the determination (3.3), whether we use SCC or DCC, we see by (3.2) that $w(s)$ is likewise single valued. However, the mapping from the s-plane into the w-plane induced by $w(s)$ will be quite different, according as we use SCC or DCC.

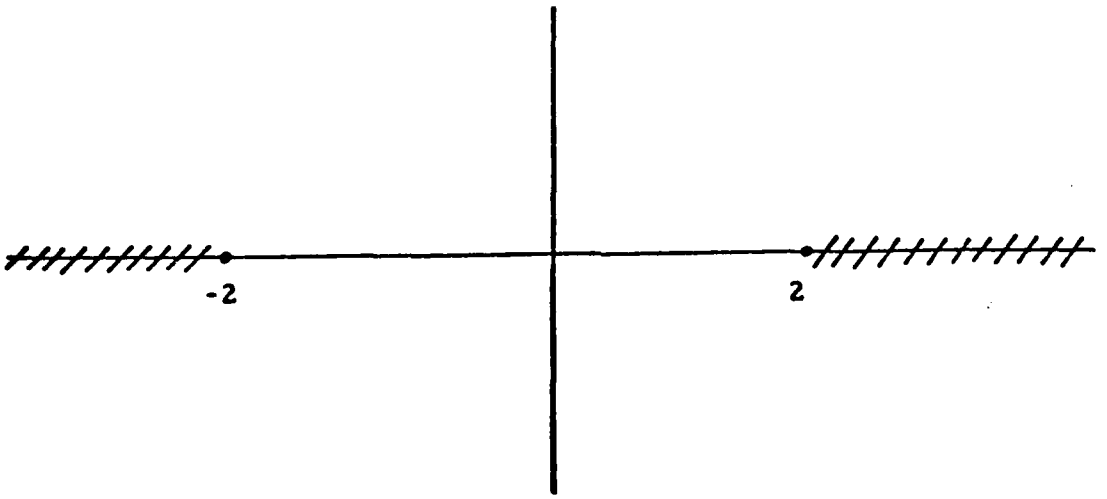
First, let us see what we get for SCC. By (3.4) and (3.2), $w(s)$ is an odd function of s. That is

$$w(-s) = -w(s). \tag{3.6}$$

From (3.3), we conclude



Singly connected cut



Doubly connected cut

Figure 3.1.

$$w(it) = \frac{i}{2}(t - \sqrt{t^2 + 4}) \quad \text{for} \quad t > 0 . \tag{3.7}$$

Then by (3.6), we have

$$w(it) = \frac{i}{2}(t + \sqrt{t^2 + 4}) \quad \text{for} \quad t < 0 . \tag{3.8}$$

We easily conclude

$$w(s) = \tfrac{1}{2}(s - \sqrt{s^2 - 4}) \quad \text{for} \quad s > 2 , \tag{3.9}$$

whence by (3.6) we get

$$w(s) = \tfrac{1}{2}(s + \sqrt{s^2 - 4}) \quad \text{for} \quad s < -2. \tag{3.10}$$

If we put

$$w = re^{i\theta} \tag{3.11}$$

into (3.1), taking r fixed with $0 < r < 1$, and letting θ vary, we see that s lies on the ellipse

$$s = (r + \frac{1}{r})\cos\theta - i(\frac{1}{r} - r)\sin\theta . \tag{3.12}$$

This has foci at -2 and $+2$, and center at the origin. The major axis lies along the real axis and is of length $2(r + (1/r))$. The minor axis lies along the imaginary axis and is of length $2((1/r) - r)$.

Thus we see that the interior of the unit circle in the w-plane is mapped into the s-plane minus SCC. In other words, with SCC, (3.2) defines a function which carries the s-plane minus the cut into the interior of the unit circle in the w-plane. That is, with SCC, the function defined by (3.2) is the same as the one defined by (2.1) and (2.2).

By (3.2), we see that $\frac{d}{ds}w(s) \neq 0$. So by the usual theory of conformal mapping, the map of the s-plane, minus SCC, into the interior of the unit circle in the w-plane is one-to-one and comformal.

To see what happens on the cut SCC, note by (3.12) that as w goes around the circle of radius r counterclockwise, s goes around the ellipse clockwise. Let r be very close to 1. Then the ellipse becomes very thin, practically indistinguishable from the cut SCC. Taking the limit as $r \to 1$, we see that if s approaches the point $2\cos\theta$ on the cut SCC from above, with $0 \leq \theta \leq \pi$, then w approaches $e^{-i\theta}$ from the interior of the unit circle. If s approaches the point $2\cos\theta$ on the cut SCC from below, then w approaches $e^{i\theta}$ from the interior of the unit circle.

We will usually be operating with the cut SCC, so that we will be using $w(s)$ as defined by (2.1) and (2.2). However, occasionally we will wish to extend $w(s)$ across the cut SCC from above. To do this, we operate with the cut DCC. Let us investigate what happens in this case.

When we use DCC, it turns out that the s-plane minus the cut DCC is mapped into the lower half of the w-plane, one-to-one and conformally. For example, we have

$$w(it) = \frac{i}{2}(t - \sqrt{t^2 + 4}) \quad \text{all real} \quad t . \tag{3.13}$$

Thus, the entire imaginary axis in the s-plane goes into the negative half of the imaginary axis in the w-plane.

More generally, let (3.11) hold with $r > 0$ and θ fixed with $-\pi/2 < \theta < 0$ Put $s = x + iy$ in (3.12) and conclude

$$\frac{x}{\cos \theta} = r + \frac{1}{r} , \tag{3.14}$$

$$\frac{y}{\sin \theta} = r - \frac{1}{r} . \tag{3.15}$$

So s is on the hyperbola

$$\frac{x^2}{4 \cos^2\theta} - \frac{y^2}{4 \sin^2\theta} = 1 . \tag{3.16}$$

By (3.14), x is positive, so that s must be on the right branch of the hyperbol If s is on the upper part of the branch (that is, $y > 0$), then $r < 1$ by (3.15 since $-\pi/2 < \theta < 0$, so that w is inside the unit circle. If s is on the lower part of the branch, then $r > 1$, so that w is outside the unit circle. We have w hitting the unit circle, $w = e^{i\theta}$, when $s = 2 \cos \theta$. If we now hold θ fixed with $-\pi < \theta < -\pi/2$, then similarly s will be on the left branch of the hyperbola (3.16).

The foci of the hyperbola (3.16) are $s = \pm 2$. So we verify that the s-plane minus the cut DCC is mapped into the lower half of the w-plane.

Let us recall the key formula (2.18). There, of course, we were using the cut SCC. We wish a formula for

$$\mathcal{R}\{(w(2 \cos \theta + \frac{8}{\ell} i))^m\} \tag{3.17}$$

for $0 \leq \theta \leq \pi$. I wish to express my appreciation to Dr. H. -S. Hung for the following derivation.

We take $s = x + iy$, and undertake generally to find $w(s) = r e^{i\theta}$ with $0 < r < 1$. This satisfies (2.1) and (2.2), so that we get

$$\frac{x}{\cos \theta} = r + \frac{1}{r} > 0 , \tag{3.18}$$

$$\frac{y}{\sin \theta} = r - \frac{1}{r} < 0 . \tag{3.19}$$

Subtracting these gives

$$r = \frac{2}{\frac{x}{\cos \theta} - \frac{y}{\sin \theta}} . \tag{3.20}$$

If we square (3.18) and (3.19) and subtract, we get

$$\frac{x^2}{4 \cos^2\theta} - \frac{y^2}{4 \sin^2\theta} = 1 . \tag{3.21}$$

This gives

$$4\cos^4\theta - (4 + x^2 + y^2)\cos^2\theta + x^2 = 0 .$$

So

$$\cos^2\theta = \frac{4 + x^2 + y^2 \pm \sqrt{(4 + x^2 + y^2)^2 - 16x^2}}{8} .$$

If we use the + sign above, then for fixed x and very large y, $\cos^2\theta$ would be greater than unity. So for very large y, the minus sign is required. However, $\cos\theta$ varies continuously with y for fixed x. As

$$(4 + x^2 + y^2)^2 - 16x^2 = (x^2 + y^2 - 4)^2 + 16y^2 ,$$

we see that continuity requires that we use a minus sign for all y. So

$$\cos^2\theta = \frac{2x^2}{x^2 + y^2 + 4 + \sqrt{(x^2 + y^2 - 4)^2 + 16y^2}} .$$

By (3.18), $x/\cos\theta > 0$. So

$$\frac{x}{\cos\theta} = + \sqrt{\tfrac{1}{2}\left\{ x^2 + y^2 + 4 + \sqrt{(x^2 + y^2 - 4)^2 + 16y^2}\right\}}. \qquad (3.22)$$

Remembering that $y/\sin\theta < 0$ by (3.19), we get by (3.21)

$$\frac{y}{\sin\theta} = - \sqrt{\tfrac{1}{2}\left\{ x^2 + y^2 - 4 + \sqrt{(x^2 + y^2 - 4)^2 + 16y^2}\right\}} . \qquad (3.23)$$

If $x^2 + y^2$ is less than 4 and $16y^2$ is very small, there can be serious cancellation of significant figures in using (3.23), and one would do better to use

$$\frac{y}{\sin\theta} = \frac{-\sqrt{8y^2}}{+\sqrt{4 - x^2 - y^2 + \sqrt{(4 - x^2 - y^2)^2 + 16y^2}}} . \qquad (3.24)$$

We now undertake to evaluate $w(2\cos\theta + \frac{8}{\ell}\, i)$. Here the θ is a parameter in the s-plane, and should not be confused with an angle in the w-plane. So we try to find $\bar{r}$ and $\bar{\theta}$ so that

$$w(2\cos\theta + \frac{8}{\ell}\, i) = \bar{r}\, e^{i\bar{\theta}} .$$

In the analysis above, we take

$$x = 2\cos\theta , \qquad y = \frac{8}{\ell} .$$

So (3.20) becomes

$$\bar{r} = \frac{1}{\frac{\cos\theta}{\cos\bar{\theta}} - \frac{4}{\ell \sin\bar{\theta}}}\ ; \tag{3.25}$$

for the evaluation of this, we use (3.22) to get

$$\frac{\cos\theta}{\cos\bar{\theta}} = +\sqrt{1 + \frac{8}{\ell^2} - \frac{\sin^2\theta}{2} + \sqrt{\left(\frac{8}{\ell^2} - \frac{\sin^2\theta}{2}\right)^2 + \frac{16}{\ell^2}}} \tag{3.26}$$

and we use (3.23) and (3.24) to get

$$\frac{4}{\ell \sin\bar{\theta}} = -\sqrt{\frac{8}{\ell^2} - \frac{\sin^2\theta}{2} + \sqrt{\left(\frac{8}{\ell^2} - \frac{\sin^2\theta}{2}\right)^2 + \frac{16}{\ell^2}}} \tag{3.27}$$

if $16 \geq \ell^2 \sin^2\theta$, and

$$\frac{4}{\ell \sin\bar{\theta}} = \frac{-4}{\ell\sqrt{\frac{\sin^2\theta}{2} - \frac{8}{\ell^2} + \sqrt{\left(\frac{\sin^2\theta}{2} - \frac{8}{\ell^2}\right)^2 + \frac{16}{\ell^2}}}} \tag{3.28}$$

if $16 \leq \ell^2 \sin^2\theta$. Then (3.26) gives

$$\cos\bar{\theta} = \frac{\cos\theta}{+\sqrt{1 + \frac{8}{\ell^2} - \frac{\sin^2\theta}{2} + \sqrt{\left(\frac{8}{\ell^2} - \frac{\sin^2\theta}{2}\right)^2 + \frac{16}{\ell^2}}}}\ . \tag{3.29}$$

This gives us finally

$$\mathcal{R}\{\ell n(w(2\cos\theta + \tfrac{8}{\ell}\, i))\} = \ell n\, \bar{r} \tag{3.30}$$

$$\mathcal{R}\{(w(2\cos\theta + \tfrac{8}{\ell}\, i))^m\} = \bar{r}^m \cos m\,\bar{\theta}\ . \tag{3.31}$$

4. A CHOICE FOR S(s). For the S(s), that we discussed in Section 2, let us propose

$$S(s) = \ell n(w(s))\ . \tag{4.1}$$

To make this single valued, we choose the branch of the ℓn such that $\ell n(w(3)) = \ell n(\frac{1}{2}(3 - \sqrt{5}))$ (see (3.9)) and make a cut leftward along the real s-axis from $s = 2$.

Obviously, $\mathcal{R}\{S(s)\}$ is harmonic (and hence continuous) everywhere except on the cut left from $s = 2$. As s approaches the part of the cut from $s = -2$ to $s = 2$, $w(s)$ approaches the unit circle, and hence $\mathcal{R}\{S(s)\}$ approaches 0 continuously. If $s_0 < -2$, then by (3.10) $\mathcal{R}\{\ell n(w(s))\}$ approaches $\ell n(-\frac{1}{2}(s_0 + \sqrt{s_0^2 - 4}))$ continuously as s approaches s_0 . So $\mathcal{R}\{S(s)\}$ is continuous everywhere, verifying (2.4).

One can continue $\ell n(w(s))$ analytically downward across the cut left from $s = -2$. Of course, the continuation will fail to agree with the value we assigned

to $\ell n(w(s))$ below the cut. Then $\mathcal{R}\{\ell n(w(s))\}$ will be harmonic for the continuation. It will take the value $\ell n\ r$, with r given by (3.20), using the values given in (3.22) and (3.23), because we continue $\ell n(w(s))$ by first continuing $w(s)$, which encounters no difficulties at the cut, and then taking ℓn of it. But the continuation of r below the cut is the same as was assigned. So $\mathcal{R}\{\ell n(w(s))\}$ proceeds harmonically across the cut. Thus we have verified (2.3).

For s very large, $w(s)$ is very close to $1/s$, by (2.1). So, though $\ell n(w(s))$ and $\ell n(w(s+\alpha))$ could differ by nearly $2\pi i$ for large s (if s is below the cut and $s + \alpha$ is above it), $\mathcal{R}\{\ell n(w(s))\}$ and $\mathcal{R}\{\ell n(w(s+\alpha))\}$ must be nearly equal. So we verify (2.5).

So our first approximation for $X(x,y)$ is $\mathcal{R}\{V(x + iy)\}$, where

$$V(z) = \ell n(w(\frac{4(z-i)}{\ell})) - \ell n(w(\frac{4(z+i)}{\ell})) . \tag{4.2}$$

Now, as in Section 2, we transform the z-plane into the w-plane by first using (2.7) and then (2.1). We then continue out to the unit circle in the w-plane. On this, we get

$$\mathcal{R}\{\ell n(w(\frac{4(z-i)}{\ell}))\} = \mathcal{R}\{\ell n(w)\} ,$$

which is 0 because w is on the unit circle. So $T(1,\theta)$ is $-\mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\}$, which we can calculate by (3.30). Indeed, by (3.30), $T(1,\pi-\theta) = T(1,\theta)$. So, by (2.9), the β_n are zero for odd n. So we dispense with them altogether, and define $\beta_n^{(0)}$ to be β_{2n}. That is, by (2.9)

$$\beta_n^{(0)} = \frac{2}{\pi} \int_0^{\pi} \mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\}\cos 2n\theta d\theta \tag{4.3}$$

for $n \geq 1$. Then

$$u(w) = \sum_{n=1}^{\infty} \beta_n^{(0)} w^{2n} .$$

If we follow through our previous discussion to (2.14), we see that we take

$$\bar{S}(s) = \ell n(w(s)) + \sum_{n=1}^{\infty} \beta_n^{(0)} (w(s))^{2n} \tag{4.4}$$

as the improved $S(s)$.

Recall that we proceeded to successive improvements. We have dispensed with the β_n for odd n, and taken $\beta_n^{(0)}$ to be β_{2n}. So (2.18) takes the form

$$\bar{\beta}_n = \frac{2}{\pi} \sum_{m=1}^{\infty} \beta_m^{(0)} \int_0^{\pi} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\cos n\theta d\theta .$$

This reminds us of (3.31). Recall that $\cos 2m\bar{\theta}$ is a polynomial in $(\cos\bar{\theta})^2$. So, by (3.29), $\cos 2m\bar{\theta}$ is symmetric about $\theta = \pi/2$ as a function of θ. So, by (3.31), we conclude that $\mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}$ is symmetric about $\theta = \pi/2$. Hence $\bar{\beta}_n$ is 0 for odd n. We dispense with them altogether, and define $\beta_n^{(1)}$ to be $\bar{\beta}_{2n}$. So we get finally for $n \geq 1$

$$\beta_n^{(1)} = \frac{2}{\pi} \sum_{m=1}^{\infty} \beta_m^{(0)} \int_0^{\pi} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\cos 2n\theta d\theta. \tag{4.5}$$

This gives

$$\tilde{\tilde{S}}(s) = \ell n(w(s)) + \sum_{n=1}^{\infty} (\beta_n^{(0)} + \beta_n^{(1)})(w(s))^{2n} \tag{4.6}$$

as a still further improvement of $S(s)$.

By now a pattern has emerged. Define

$$\gamma_{n,m} = \frac{2}{\pi} \int_0^{\pi} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\cos 2n\theta d\theta. \tag{4.7}$$

Define $\beta_n^{(\lambda)}$ by the iteration that $\beta_n^{(0)}$ is given by (4.3) and for $n \geq 1$

$$\beta_n^{(\lambda+1)} = \sum_{m=1}^{\infty} \beta_m^{(\lambda)} \gamma_{n,m} \tag{4.8}$$

(see (4.5)). Define for $n \geq 1$

$$\alpha_n^{(\lambda)} = \sum_{\mu=0}^{\lambda} \beta_n^{(\mu)} . \tag{4.9}$$

Then, for larger and larger λ

$$\ell n(w(s)) + \sum_{n=1}^{\infty} \alpha_n^{(\lambda)} (w(s))^{2n}$$

will be a better and better improvement for $S(s)$ (see 4.6)).

We define for $n \geq 1$

$$\alpha_n = \lim_{\lambda \to \infty} \alpha_n^{(\lambda)} , \tag{4.10}$$

provided that the limit exists (and we will show that it does). Then we take

$$U(s) = \ell n(w(s)) + \sum_{n=1}^{\infty} \alpha_n (w(s))^{2n} \tag{4.11}$$

as the ultimate improvement for $S(s)$, provided the series on the right converges (we will prove that it does, and indeed uniformly in s). So we will take $X(x,y)$ as $\mathcal{R}\{W(x+iy)\}$, where

$$W(z) = U(4(\qquad - U(4(z+i)/\ell) . \tag{4.12}$$

We note that $\mathcal{R}\{U(s)$ [illegible] function of s, and $\mathcal{R}\{W(z)\}$ is an odd function of z. We can [illegible] ts already set forth to conclude that $U(s)$ has the properti[illegible].4), and (2.5) that were postulated for $S(s)$. Hence $X(x,y)$ is continuo[illegible] x and y, is harmonic except on the two "plates" and goes to 0 [illegible] goes to infinity.

The crucial question [illegible]r $X(x,y)$ is a non-zero constant C along the upper "plate". If it is, then it will be $-C$ along the lower plate, since $\mathcal{R}\{W(z)\}$ is an odd function of z.

By (4.3) and (4.9), we have for $n \geq 1$

$$\alpha_n^{(0)} = \frac{2}{\pi}\int_0^{\pi} \mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\}\cos 2n\theta d\theta. \tag{4.13}$$

By (4.9), $\alpha_n^{(-1)} = 0$. So by (4.9)

$$\alpha_n^{(\lambda+1)} = \alpha_n^{(\lambda)} + \beta_n^{(\lambda+1)} \tag{4.14}$$

for $\lambda \geq -1$ and $n \geq 1$. Hence, by induction on λ , we can prove by (4.8) that

$$\alpha_n^{(\lambda+1)} = \alpha_n^{(0)} + \sum_{m=1}^{\infty} \alpha_m^{(\lambda)} \gamma_{n,m} \tag{4.15}$$

for $\lambda \geq -1$ and $n \geq 1$.

We will show that $\alpha_n^{(\lambda)}$ goes to α_n with sufficient uniformity in n that we can let $\lambda \to \infty$ in (4.15) and conclude for $n \geq 1$

$$\alpha_n = \alpha_n^{(0)} + \sum_{m=1}^{\infty} \alpha_m \gamma_{n,m} . \tag{4.16}$$

Along the upper "plate" we have

$$z = i + \tfrac{1}{2}\ell \cos\theta. \tag{4.17}$$

So

$$w(4(z-i)/\ell) = e^{\pm i\theta} \tag{4.18}$$

$$w(4(z+i)/\ell) = w(2\cos\theta + \frac{8}{\ell} i) . \tag{4.19}$$

So, along the upper plate X is a function of θ given by

$$\begin{aligned} X = C(\theta) &= \mathcal{R}\{W(i + \tfrac{1}{2}\ell\cos\theta)\} \\ &= \sum_{m=1}^{\infty} \alpha_m \cos 2m\theta - \mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\} \\ &\quad - \sum_{m=1}^{\infty} \alpha_m \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\} . \end{aligned} \tag{4.20}$$

Clearly $C(\theta)$ is an even function of θ. So,

$$\int_{-\pi}^{\pi} C(\theta)\sin n\theta d\theta = 0 .$$

$$\int_{-\pi}^{\pi} C(\theta)\cos n\theta d\theta = 2\int_0^{\pi} C(\theta)\cos n\theta d\theta .$$

By an analysis we carried out earlier, $C(\theta)$ is symmetric with respect to θ about the point $\theta = \pi/2$. So

$$2\int_0^{\pi} C(\theta)\cos n\theta d\theta = 0$$

for odd n. By (4.7), (4.13) and (4.20),

$$2\int_0^\pi C(\theta)\cos 2n\theta d\theta = \pi[\alpha_n - \alpha_n^{(0)} - \sum_{m=1}^{\infty} \alpha_m \gamma_{n,m}]$$

for $n \geq 1$. So, by (4.16),

$$\int_{-\pi}^{\pi} C(\theta)\cos n\theta d\theta = 0$$

for $n \geq 1$. Hence $C(\theta)$ has a Fourier series expansion whose only non-vanishing coefficient is that of the constant term.

Thus we see that (4.20) holds with a constant C for z on the upper "plate". To show that $C \neq 0$, we argue as follows. Suppose $C = 0$. Then $\mathcal{R}\{W(z)\} = 0$ on the upper "plate", whence by (4.11) and (4.12) it must also be 0 on the lower "plate". As $\mathcal{R}\{W(z)\}$ is a harmonic function except on the "plates", and approaches zero as z goes to ∞, $\mathcal{R}\{W(z)\}$ must be identically zero. So by the Cauchy-Riemann differential equations, $W(z)$ is a constant, W.

But $U(4(z+i)/\ell)$ is analytic and single valued except on the lower "plate" and its extension to the left. By (4.12),

$$U(4(z-i)/\ell) = W + U(4(z+i)/\ell) .$$

So $U(4(z-i)/\ell)$ is analytic and single valued in the neighborhood of the upper "plate". But in the neighborhood of the upper "plate" $w(4(z-i)/\ell)$ is single valued as long as one stays off the "plate", whereas $\ell n(w(4(z-i)/\ell))$ will go from one branch to another as one encircles the upper "plate". So by (4.11), $U(4(z-i)/\ell)$ cannot be single valued in the neighborhood of the upper plate.

Thus we have our contradiction, and can conclude that $C \neq 0$. Integrating both sides of (4.20) from 0 to π gives

$$C = -\frac{1}{\pi}\int_0^\pi \mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\}d\theta$$

$$-\frac{1}{\pi}\sum_{m=1}^{\infty} \alpha_m \int_0^\pi \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}d\theta. \qquad (4.21)$$

The "C-test", to see if the calculation of the α_n has proceeded without numerical mistakes, is to substitute a number of different values of θ into the right side of (4.20) and see if it takes the value C as given by (4.21), or by one of the other values of θ.

We might remark that if one gets the same value of C by (4.20) for all values of θ, one has found the desired values of α_m, whatever might have been the method of calculation; the desired potential will certainly be given by $\mathcal{R}\{W(x+iy)\}$ (see (4.12)). If one gets nearly the same value of C for all values of θ, then by the maximum principle $\mathcal{R}\{W(x+iy)\}$ will give a good approximation to the potential. If one gets nearly the same value of C for a considerable number of values of θ, it is not likely that C would be far off for any of the other values of θ. Thus, again by the maximum principle, it is likely that $\mathcal{R}\{W(x+iy)\}$ will give a good approximation to the potential.

5. THE FAST FOURIER TRANSFORM. In (4.13), we need to calculate an approximation for

$$\phi_n = \frac{2}{\pi} \int_0^{\pi} J(\theta) \cos 2n\theta d\theta. \tag{5.1}$$

Combining (4.7) with (4.8), we find again that we wish to calculate an approximation for (5.1).

Expand $J(\theta)$ in a Fourier series,

$$J(\theta) = \beta_0 + \sum_{n=1}^{\infty} (\beta_n \cos n\theta + \gamma_n \sin n\theta)$$

for $-\pi \leq \theta \leq \pi$. In all the cases where we wish to evaluate (5.1), $J(\theta)$ is an even function. So the γ_n are all zero. Also, we have $J(\pi-\theta) = J(\theta)$, so that $\beta_n = 0$ for odd n and $\beta_{2m} = \phi_m$ for positive m. In effect, the ϕ_n are Fourier coefficients, and we have

$$J(\theta) = \tfrac{1}{2}\phi_0 + \sum_{n=1}^{\infty} \phi_n \cos 2n\theta. \tag{5.2}$$

We will find, for the $J(\theta)$'s that we are interested in, that the ϕ_n's decrease rapidly in absolute value. Choose Ω large enough so that

$$\sum_{n=\Omega}^{\infty} |\phi_n|$$

is negligible for purposes of computation. Then, to a high degree of accuracy, $J(\theta) \cong J_\Omega(\theta)$, where

$$J_\Omega(\theta) = \tfrac{1}{2}\phi_0 + \sum_{n=1}^{\Omega-1} \phi_n \cos 2n\theta. \tag{5.3}$$

Define

$$\sum_{\Omega} I(\theta) = \sum_{k=0}^{2\Omega-1} I\left(\frac{k\pi}{2\Omega}\right). \tag{5.4}$$

One can verify that

$$\sum_{\Omega} \cos 2m\theta \cos 2n\theta = \Omega\delta_{nm} \tag{5.5}$$

for $0 \leq m < \Omega$ and $1 \leq n < \Omega$. So multiply both sides of (5.3) by $\cos 2m\theta$ and apply the $\sum_{\Omega}$ operator. We get

$$\phi_m = \frac{1}{\Omega} \sum_{\Omega} J_\Omega(\theta) \cos 2m\theta$$

for $0 \leq m < \Omega$. As $J(\theta) \cong J_\Omega(\theta)$, we get

$$\phi_m \cong \frac{1}{\Omega} \sum_{\Omega} J(\theta) \cos 2m\theta \tag{5.6}$$

for $0 \leq m < \Omega$. As the ϕ_m are real, we have $\phi_m \cong \mathcal{R}\{\bar{\phi}_m\}$, where

$$\bar{\phi}_m = \frac{1}{\Omega} \sum_{k=0}^{2\Omega-1} J\left(\frac{k\pi}{2\Omega}\right) e^{2mk\pi i/2\Omega} \tag{5.7}$$

for $0 \leq m < \Omega$.

The Fast Fourier Transform (FFT) provides an efficient way to calculate such sums as occur in (5.7). A collection of articles on the FFT is given on pages 312-382 of [4]. Amongst them, our reference [5] contains a useful discussion.

In one of our examples, we had to calculate ϕ_m for m going up somewhat in excess of 200. We took $\Omega = 256$, and the FFT resulted in a saving of computational labor by a factor of 10 to 20. Maximum efficiency results when one can take Ω to be a power of 2, as we did. A FORTRAN program for calculation in this case is given in [6].

By Theorem 8.9, we will be able to show that

$$\sum_{n=\Omega}^{\infty} |\alpha_n^{(\lambda)}| \tag{5.8}$$

can be made as small as desired by taking Ω sufficiently large, uniformly in λ. Take Ω large enough that (5.8) is negligible for purposes of computation. As (4.13) has the form (5.1), we conclude that

$$\alpha_n^{(0)} \cong \bar{\alpha}_n^{(0)} = \frac{1}{\Omega} \sum_{\Omega} \mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\}\cos 2n\theta \tag{5.9}$$

for $0 \le n < \Omega$, to a high degree of approximation. Thus we may use the FFT to calculate approximations for the $\alpha_n^{(0)}$.

Using (4.13) and (4.7) in (4.15) gives

$$\alpha_n^{(\lambda+1)} = \frac{2}{\pi}\int_0^{\pi}\left[\mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\} + \sum_{m=1}^{\infty} \alpha_m^{(\lambda)} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\right]\cos 2n\theta d\theta .$$

The interchange of order of summation and integration is justified because the $\alpha_n^{(\lambda)}$ are bounded (see (5.8)) and $|w(2\cos\theta + \frac{8}{i}\ell)|$ is less than unity uniformly by (2.2). As (5.8) is negligible for $\lambda + 1$, we can argue as before to conclude that

$$\alpha_n^{(\lambda+1)} \cong \frac{1}{\Omega}\sum_{\Omega}\left[\mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\} + \sum_{m=1}^{\infty} \alpha_m^{(\lambda)} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\right]\cos 2n\theta.$$

to a high degree of approximation. Using the definition in (5.9), we write this as

$$\alpha_n^{(\lambda+1)} \cong \bar{\alpha}_n^{(0)} + \frac{1}{\Omega}\sum_{\Omega}\left[\sum_{m=1}^{\infty} \alpha_m^{(\lambda)} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\right]\cos 2n\theta.$$

But (5.8) is negligible, and $|w(2\cos\theta + \frac{8}{\ell} i)| < 1$, so that we get

$$\alpha_n^{(\lambda+1)} \cong \bar{\alpha}_n^{(0)} + \frac{1}{\Omega}\sum_{\Omega}\sum_{m=1}^{\Omega-1} \alpha_m^{(\lambda)} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m} \cos 2n\theta.$$

Let us define

$$\bar{\gamma}_{n,m} = \frac{1}{\Omega}\sum_{\Omega} \mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}\cos 2n\theta . \tag{5.10}$$

Then we have finally

$$\alpha_n^{(\lambda+1)} \cong \bar{\alpha}_n^{(0)} + \sum_{m=1}^{\Omega-1} \alpha_m^{(\lambda)}\, \bar{\gamma}_{n,m} . \tag{5.11}$$

In this, replace λ by $\lambda-1$, and subtract from (5.11). By (4.14), we get

$$\beta_n^{(\lambda+1)} \cong \sum_{m=1}^{\Omega-1} \beta_m^{(\lambda)} \bar{\gamma}_{n,m} . \tag{5.12}$$

This is quite similar to (4.8). Note that we are not claiming that $\bar{\gamma}_{n,m} \cong \gamma_{n,m}$. We merely point out that the use in (5.12) gives a good approximation. And, of course, the $\bar{\gamma}_{n,m}$ can be calculated very efficiently by the FFT.

Our method of operation is first to compute approximations for the $\beta_n^{(0)}$. We do this by (5.9), since $\beta_n^{(0)} = \alpha_n^{(0)}$ by (4.9). Then we use (5.12) to calculate $\beta_n^{(\lambda)}$ for larger and larger λ. Then we get the $\alpha_n^{(\lambda)}$ for larger and larger λ by (4.14).

Of course, this is done only for $n < \Omega$. However, because (5.8) is negligible and $|w(s)| < 1$, this gives good approximations for $U(s)$ by (4.11). Then we get $X(x,y)$ by taking $\Re\{W(x,iy)\}$, with $W(z)$ given by (4.12).

Although we have shown that (5.12) gives a good approximation for $\beta_n^{(\lambda+1)}$, the buildup of the $\alpha_n^{(\lambda+1)}$ by use of (4.14) allows the possibility of accumulation of errors, possibly to a harmful degree. So there remains a question of how accurate are our final approximations for the α_n. One can use the methods of Section 8 to show that we can come as close as we wish to the values of the α_n, for $n < \Omega$, by taking a large enough Ω. However, the analysis is quite complex. One would wish to apply the C-test (see above) at the end, to make sure one had not committed numerical mistakes. But if we get good results from the C-test, we are assured that we have good approximations for the α_n. For the five computations that we tried, we had good success with the C-test. For further confirmation, we checked the results by the method of [2].

6. COMPUTATIONAL RESULTS. We postpone the proofs of convergence, and that sort of thing, to Section 8. Here we will summarize some of the numerical results which we obtained.

For the early computations, we had not yet obtained the information about convergence, etc., which is in Section 8. In order for (5.7) to give adequate accuracy, we have to choose Ω large enough so that

$$\sum_{n=\Omega}^{\infty} |\phi_n|$$

is negligible for purposes of computation. Before we had learned our rates of convergence, we had no basis for choosing Ω. So at first we just choose Ω quite large, and hoped for the best. Subsequently, after we learned the rates of convergence, it turned out that we had taken Ω larger than need be in all but one case; there it seemed about right. This did not invalidate any of our results, but just meant that we had done more calculation than was necessary. Even without knowing the rates of convergence, the success of the C-test would have assured us that we had not taken too small a value for Ω.

Our first calculation was for $\ell = 1.6$. Table 6.1 contains values of various α's. We will derive (8.34) to give bounds for the $|\alpha_m^{(0)}|$. A list of these bounds, rounded to two significant decimals, is given in Table 6.1. We note that by (8.34)

Table 6.1

Values and bounds for $\ell = 1.6$.

m	$\alpha_m^{(0)}$	(8.34)	α_m	(8.36)
1	-3.0608×10^{-2}	1.3×10^{-1}	-3.0798×10^{-2}	1.5×10^{-1}
2	3.4597×10^{-4}	4.3×10^{-3}	3.5537×10^{-4}	4.8×10^{-3}
3	-2.7782×10^{-6}	1.4×10^{-4}	-3.1322×10^{-6}	1.6×10^{-4}
4	-4.1446×10^{-8}	4.7×10^{-6}	-3.0523×10^{-8}	5.2×10^{-6}
5	2.8131×10^{-9}	1.5×10^{-7}	2.5373×10^{-9}	1.7×10^{-7}
6	-8.0717×10^{-11}	5.0×10^{-9}	-7.5444×10^{-11}	5.6×10^{-9}
7	1.3296×10^{-12}	1.6×10^{-10}	1.2800×10^{-12}	1.8×10^{-10}
8	2.6301×10^{-15}	5.4×10^{-12}	1.2321×10^{-15}	6.0×10^{-12}
9	-1.0895×10^{-15}	1.8×10^{-13}	-9.9615×10^{-16}	2.0×10^{-13}
10	4.9699×10^{-17}	5.8×10^{-15}	4.6757×10^{-17}	6.4×10^{-15}

$$|\alpha_{10}^{(0)}| \leq (5.8 \pm 0.05) \times 10^{-15} .$$

As the individual terms of the sum (5.9) are not much less than unity, this means that in calculating $\alpha_{10}^{(0)}$ at least twelve or thirteen significant decimal digits, perhaps more, must be lost off the front due to cancellation. At best, this would leave very few correct significant digits, even in a double precision calculation (none at all in a single precision calculation).

In fact, $|\alpha_{10}^{(0)}|$ is likely a good bit smaller than 5.8×10^{-15}. Although double precision was used in the computation, it is likely that the combination of cancellations and round off errors is so great that the approximation calculated for $\alpha_{10}^{(0)}$, and shown in Table 6.1, has not a single significant digit correct.

To get some notion of the size of round off error, the computation was repeated with single precision. By comparison with the double precision values, the amount of cancellation and round off error combined for the single precision calculation could be determined. One would suppose that the double precision calculation is affected with similar errors.

The single precision approximation for $\alpha_1^{(0)}$ had five significant decimal digits correct, for $\alpha_2^{(0)}$ had four, for $\alpha_3^{(0)}$ had one, and for all other $\alpha_m^{(0)}$ had none. The approximation for $\alpha_4^{(0)}$ had the right sign and the right order of magnitude, but the approximation for $\alpha_5^{(0)}$ had neither the right sign nor the right order of magnitude. From this, one is tempted to conjecture that the approximations calculated by double precision and listed for $\alpha_8^{(0)}$ and $\alpha_9^{(0)}$ in Table 6.1 may each have one, or possibly two, significant decimal digits correct. This order of accuracy is corroborated by the C-test and the comparison with the values computed by elliptic integrals (as described in [2]).

At the beginning, we had no notion how fast $\alpha_m^{(\lambda)}$ converges to α_m. So the $\beta_m^{(\lambda)}$ were computed for $\lambda \leq 11$. As it turned out, $\lambda \leq 7$ would have been entirely adequate. Some values of $\beta_m^{(\lambda)}$ are shown in Table 6.2.

It will be noted that while the bounds given by (8.34) and (8.36) in Table 6.1 are appreciably too large, they are not ridiculously so. On the other hand, the bounds given by (8.47) in Table 6.2 are preposterously too large. However, they suffice to assure that $\alpha_n^{(\lambda)}$ converges to α_n, for $1 \leq n < \Omega$.

The values of α_n were computed from

$$\alpha_n = \sum_{\lambda=0}^{\infty} \beta_n^{(\lambda)} , \tag{6.1}$$

which follows from (4.9) and (4.10). The $\beta_n^{(\lambda)}$ for this purpose had been computed by (5.12), starting from (5.9); this latter gets us started since $\beta_n^{(0)} = \alpha_n^{(0)}$.

The $\bar{\gamma}_{n,m}$ had been computed by (5.10). This raises the question of what sorts of errors can arise from the use of (5.10). Of course, we verified at the end, by the C-test, that we had arrived at good values for the α_n. However, during the course of the calculation, it is well to be assured that we have taken Ω large enough so that we will not be disappointed by the C-test at the end of

Table 6.2

Values for $\ell = 1.6$.

λ	$\beta_1^{(\lambda)}$	$\beta_2^{(\lambda)}$	$\beta_3^{(\lambda)}$	$\beta_4^{(\lambda)}$	$\beta_5^{(\lambda)}$	(e.47)
1	-1.89×10^{-4}	9.35×10^{-6}	-3.518×10^{-7}	1.0854×10^{-8}	-2.741×10^{-10}	3.0×10^{-2}
2	-1.17×10^{-6}	5.82×10^{-8}	-2.21×10^{-9}	6.89×10^{-11}	-1.78×10^{-12}	7.2×10^{-3}
3	-7.2×10^{-9}	3.6×10^{-10}	-1.4×10^{-11}	4.3×10^{-13}	-1.1×10^{-14}	1.7×10^{-3}
4	-4.5×10^{-11}	2.2×10^{-12}	-8.5×10^{-14}	2.6×10^{-15}	-6.8×10^{-17}	4.2×10^{-4}
5	-2.8×10^{-13}	1.4×10^{-14}	-5.2×10^{-16}	1.6×10^{-17}	-4.2×10^{-19}	1.0×10^{-4}
6	-1.7×10^{-15}	8.6×10^{-17}	-3.2×10^{-18}	1.0×10^{-19}	-2.6×10^{-21}	2.5×10^{-5}
7	-1.1×10^{-17}	5.3×10^{-19}	-2.0×10^{-20}	6.3×10^{-22}	-1.6×10^{-23}	6.0×10^{-6}
8	-6.6×10^{-20}	3.3×10^{-21}	-1.2×10^{-22}	3.9×10^{-24}	-1.0×10^{-25}	1.5×10^{-6}
9	-4.1×10^{-22}	2.0×10^{-23}	-7.7×10^{-25}	2.4×10^{-26}	-6.2×10^{-28}	3.5×10^{-7}
10	-2.5×10^{-24}	1.3×10^{-25}	-4.8×10^{-27}	1.5×10^{-28}	-3.9×10^{-30}	8.5×10^{-8}
11	-1.6×10^{-26}	7.8×10^{-28}	-3.0×10^{-29}	9.2×10^{-31}	-2.4×10^{-32}	2.1×10^{-8}

the calculation.

If we calculate the $\bar{\gamma}_{n,m}$ by (5.10), then using (5.12) is exactly the same as using (5.11). So it does not matter if (5.10) gives good approximations for the $\gamma_{n,m}$ individually. Whether or not (5.10) makes the $\bar{\gamma}_{n,m}$ good approximations for the $\gamma_{n,m}$ individually, if one uses (5.10) and then uses (5.12) one is in effect using (5.11); if (5.8) is small uniformly in λ, then (5.11) gives a good approximation.

The analysis above presupposes that there are no cancellations or round off errors made in the use of (5.10). Any such errors made in the use of (5.10) tend to invalidate the equivalence of (5.12) with (5.11).

As the $\beta_m^{(\lambda)}$ decrease rapidly with λ, one can allow the percent of error of the $\beta_m^{(\lambda)}$ to increase as λ increases without producing much error in the α_m by the use of (6.1). Thus, one is less concerned with the accumulation of cancellations or round off errors in using (5.10) than in using (5.9). The accumulation of cancellations and round off errors ought to be about the same for both, and hence of less consequence for (5.10). Interestingly enough, the cancellation appears to be less with (5.10) than with (5.9), especially for large n. The reason is that the term $\mathcal{R}\{\ell n(w(2\cos\theta + \frac{8}{\ell} i))\}$ in (5.9) does not vary greatly, so that one is adding nearly equal terms, which is conducive to cancellation. However, the term $\mathcal{R}\{(w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}$ in (5.10) varies much more, especially for large m. Thus a few terms predominate, and the cancellation from these few terms is not too severe.

This was brought out most strikingly in the single precision calculations which we performed. While cancellations and round off errors left no correct significant digits in $\alpha_m^{(0)}$ for $m \geq 4$, the $\gamma_{4,m}$ each had at least three correct significant decimal digits. All the $\gamma_{6,m}$ had at least one correct significant decimal digit except for $m = 0$. Though $\alpha_{10}^{(0)}$ probably has no correct significant digits even in a double precision calculation, the single precision calculation of the $\gamma_{10,m}$ gave at least one correct significant decimal digit for both $m = 9$ and $m = 10$.

In any case, as we noted above, the values of α_m which were computed checked out very well in both the C-test and comparison with the values calculated by use of elliptic integrals (as described in [2]). Incidentally, the value of C is approximately

$$1.6772\quad 45213\quad 00067\quad 57.$$

The next case considered was $\ell = 4$. A first calculation was done with assorted large values of Ω, some running into the hundreds. It was realized about then that this was not advisable, and the computation was repeated, this time with $\Omega = 20$ uniformly. The results are summarized in Table 6.3. From the values of (8.34) and (8.36), it is clear that indeed $\Omega = 20$ is adequately large, even for a double precision calculation, such as we were doing. Although we have listed approximations from the calculated values of $\alpha_m^{(0)}$ and α_m for $m = 16$ and $m = 17$, it is not too likely that they have any correct significant digits. The values calculated for $m = 18$ and 19 can almost be guaranteed not to have any correct significant digits, and it seemed pointless to list them.

Table 6.3

Values and bounds for $l = 4.0$.

m	$\alpha_m^{(0)}$	(8.34)	α_m	(8.36)
1	-9.0179×10^{-2}	4.4×10^{-1}	-9.7090×10^{-2}	9.0×10^{-1}
2	9.4071×10^{-5}	5.2×10^{-2}	1.0469×10^{-3}	1.1×10^{-1}
3	2.2969×10^{-4}	6.2×10^{-3}	1.6123×10^{-4}	1.3×10^{-2}
4	-8.3332×10^{-6}	7.5×10^{-4}	-7.5227×10^{-6}	1.5×10^{-3}
5	-1.1821×10^{-6}	8.9×10^{-5}	-8.5594×10^{-7}	1.8×10^{-4}
6	1.1400×10^{-7}	1.1×10^{-5}	9.2086×10^{-8}	2.2×10^{-5}
7	5.1314×10^{-9}	1.3×10^{-6}	3.6342×10^{-9}	2.6×10^{-6}
8	-1.2905×10^{-9}	1.5×10^{-7}	-1.0263×10^{-9}	3.2×10^{-7}
9	1.0298×10^{-11}	1.8×10^{-8}	1.1448×10^{-11}	3.8×10^{-8}
10	1.2703×10^{-11}	2.2×10^{-9}	1.0015×10^{-11}	4.5×10^{-9}
11	-7.2763×10^{-13}	2.6×10^{-10}	-6.0434×10^{-13}	5.4×10^{-10}
12	-1.0233×10^{-13}	3.2×10^{-11}	-7.9780×10^{-14}	6.5×10^{-11}
13	1.3206×10^{-14}	3.8×10^{-12}	1.0696×10^{-14}	7.8×10^{-12}
14	4.9899×10^{-16}	4.5×10^{-13}	3.7473×10^{-16}	9.3×10^{-13}
15	-1.8145×10^{-16}	5.4×10^{-14}	-1.4620×10^{-16}	1.1×10^{-13}
16	4.8572×10^{-18}	6.5×10^{-15}	4.4561×10^{-18}	1.3×10^{-14}
17	6.5919×10^{-18}	7.8×10^{-16}	6.1783×10^{-18}	1.6×10^{-15}

The $\beta_m^{(\lambda)}$ were computed for $\lambda \leq 21$. It turned out that $\lambda \leq 14$ would have sufficed. Some values are shown in Table 6.4. As before, the bounds given by (8.47) are worthless, except to prove convergence in (6.1), which is equivalent to (4.10).

Table 6.4

Values for $\ell = 4.0$.

λ	$\beta_1^{(\lambda)}$	$\beta_5^{(\lambda)}$	$\beta_9^{(\lambda)}$	(8.47)
1	-6.4×10^{-3}	3.2×10^{-7}	1.2×10^{-12}	1.8×10^{-1}
5	-1.9×10^{-7}	2.5×10^{-12}	-3.9×10^{-17}	1.1×10^{-2}
9	-5.7×10^{-12}	7.4×10^{-17}	-1.1×10^{-21}	7.0×10^{-4}
12	-2.3×10^{-15}	3.0×10^{-20}	-4.6×10^{-25}	8.8×10^{-5}
15	-9.2×10^{-19}	1.2×10^{-23}	-1.9×10^{-28}	1.1×10^{-5}
18	-3.7×10^{-22}	4.9×10^{-27}	-7.5×10^{-32}	1.4×10^{-6}
21	-1.5×10^{-25}	2.0×10^{-30}	-3.0×10^{-35}	1.7×10^{-7}

The C-test and comparison with values calculated by elliptic integrals corroborate the accuracy of the α_m . The value of C is approximately

$$0.96265 \quad 43980 \quad 34667 \quad 78 \ .$$

In Tables 6.5, 6.6, and 6.7 are given selected values of the $\alpha_m^{(0)}$ and α_m for $\ell = 10$, 20, and 200 respectively, with bounds calculated by (8.34), Theorem 8.10 and Theorem 8.11. For these tables, use of (8.36) to get bounds for $|\alpha_m|$ is not possible. It will be noted that Theorems 8.10 and 8.11 give much poorer bounds. However, except for the case $\ell = 200$, Theorem 8.10 is not too bad. For example, for both $\ell = 10$ and $\ell = 20$, one can show by Theorem 8.10 that a larger value was used for Ω than necessary.

For $\ell = 10$, we took $\Omega = 60$. However, $\Omega = 30$ would have been more than adequate for a double precision calculation, such as we made. For a single precision calculation, $\Omega = 12$ would have sufficed. The $\beta_m^{(\lambda)}$ were computed for $\lambda \leq 61$. For double precision, $\lambda \leq 32$ would have been quite adequate, and for single precision a good bit less would have sufficed. The value of C was approximately

$$0.48401 \quad 62679 \quad 54256 \quad 77 \ .$$

Table 6.5

Values and bounds for $\ell = 10.0$.

m	$\alpha_m^{(0)}$	(8.34)	α_m	Thm. 8.10	Thm. 8.11
1	-1.3101×10^{-1}	9.2×10^{-1}	-1.8046×10^{-1}	14	11
3			5.4098×10^{-4}	1.0	2.3
4	3.6910×10^{-4}	1.8×10^{-2}			
6	-8.7083×10^{-6}	1.4×10^{-3}	-5.4906×10^{-6}	2.0×10^{-2}	2.2×10^{-1}
9	1.4733×10^{-7}	2.7×10^{-5}	8.9578×10^{-8}	4.1×10^{-4}	2.1×10^{-2}
11	-3.6197×10^{-9}	2.0×10^{-6}	-2.3683×10^{-9}	3.0×10^{-5}	4.5×10^{-3}
14	9.9592×10^{-11}	4.0×10^{-8}	6.2094×10^{-11}	6.0×10^{-7}	4.3×10^{-4}
16			-1.1583×10^{-12}	4.4×10^{-8}	9.0×10^{-5}
17	-1.8414×10^{-12}	8.1×10^{-10}			
19	7.8477×10^{-14}	6.0×10^{-11}	4.9388×10^{-14}	8.9×10^{-10}	8.7×10^{-6}
22	-1.7750×10^{-15}	1.2×10^{-12}	-1.1009×10^{-15}	1.8×10^{-11}	8.4×10^{-7}
24	6.2566×10^{-17}	8.9×10^{-14}	3.9518×10^{-17}	1.3×10^{-12}	1.8×10^{-7}

Table 6.6

Values and bounds for $\ell = 20.0$.

m	$\alpha_m^{(0)}$	(8.34)	α_m	Thm. 8.10	Thm. 8.11
1	-1.2355×10^{-1}	1.3	-2.3237×10^{-1}	377	70
4			2.3773×10^{-4}	25	21
5	3.4185×10^{-4}	3.5×10^{-2}			
8	-4.9366×10^{-6}	2.3×10^{-3}	-2.5934×10^{-6}	6.5×10^{-1}	4.3
12	1.3359×10^{-7}	6.0×10^{-5}	6.4113×10^{-8}	1.7×10^{-2}	8.8×10^{-1}
16	-3.0641×10^{-9}	1.6×10^{-6}	-1.4636×10^{-9}	4.5×10^{-4}	1.8×10^{-1}
19	9.6190×10^{-11}	1.0×10^{-7}	4.7528×10^{-11}	3.0×10^{-5}	5.5×10^{-2}
23	-2.8192×10^{-12}	2.7×10^{-9}	-1.3618×10^{-12}	7.9×10^{-7}	1.1×10^{-2}
26	8.1741×10^{-14}	1.8×10^{-10}	4.0730×10^{-14}	5.1×10^{-8}	3.4×10^{-3}
30	-2.9296×10^{-15}	4.7×10^{-12}	-1.4243×10^{-15}	1.4×10^{-9}	6.9×10^{-4}
34	8.0793×10^{-17}	1.3×10^{-13}	3.8349×10^{-17}	3.6×10^{-11}	1.4×10^{-4}

Table 6.7

Values and bounds for $\ell = 200.0$.

m	$\alpha_m^{(0)}$	(8.34)	n	α_n	Thm. 8.10	Thm. 8.11
1	-8.9018×10^{-2}	2.4	1	-3.1424×10^{-1}	5.0×10^{14}	8.5×10^{3}
16	4.3573×10^{-5}	3.4×10^{-2}	13	1.0087×10^{-5}	1.7×10^{13}	5.2×10^{3}
27	-7.8296×10^{-7}	1.5×10^{-3}	26	-1.2793×10^{-7}	4.2×10^{11}	3.1×10^{3}
38	1.9221×10^{-8}	6.7×10^{-5}	37	3.2736×10^{-9}	1.7×10^{10}	2.0×10^{3}
49	-5.4367×10^{-10}	3.0×10^{-6}	48	-9.5140×10^{-11}	8.3×10^{8}	1.3×10^{3}
61	1.7091×10^{-11}	9.9×10^{-8}	60	2.9451×10^{-12}	2.8×10^{7}	8.0×10^{2}
72	-5.6778×10^{-13}	4.4×10^{-9}	71	-9.8066×10^{-14}	1.2×10^{6}	5.1×10^{2}
83	1.9529×10^{-14}	1.9×10^{-10}	82	3.3667×10^{-15}	5.4×10^{4}	3.3×10^{2}
94	-6.8877×10^{-16}	8.6×10^{-12}	93	-1.2087×10^{-16}	2.4×10^{3}	2.1×10^{2}
105	2.4749×10^{-17}	3.8×10^{-13}				
116	-9.0187×10^{-19}	1.7×10^{-14}				
127	3.3215×10^{-20}	7.5×10^{-16}				
139	-1.2605×10^{-21}	2.5×10^{-17}				
150	4.8256×10^{-23}	1.1×10^{-18}				
161	-1.8584×10^{-24}	4.9×10^{-20}				
172	7.1908×10^{-26}	2.2×10^{-21}				
183	-2.7964×10^{-27}	9.6×10^{-23}				
194	1.0930×10^{-28}	4.3×10^{-24}				

For $\ell = 20$, we took $\Omega = 100$. However, $\Omega = 40$ would have been more than adequate for the double precision calculation we made, and $\Omega = 16$ for single precision. The $\beta_m^{(\lambda)}$ were computed for $\lambda \leq 101$. For double precision, $\lambda \leq 64$ would have been quite adequate. The value of C was approximately

$$0.26894 \quad 13477 \quad 20025 \quad 97.$$

For $\ell = 200$, we took $\Omega = 130$, which was hardly any too large for double precision. A special triple precision computation was made of $\alpha_m^{(0)}$, using $\Omega = 256$ and the FFT. The $\beta_m^{(\lambda)}$ were computed for $\lambda \leq 300$. This turned out not to be enough, and it was extended to $\lambda = 550$, which apparently sufficed. We say "apparently", because the C-test did not work out quite as well as for the smaller values of ℓ. However, the C-test worked well enough for us to feel that the value of C is approximately

$$0.03068 \quad 67555 \quad 10411 \ .$$

It appears that one cannot have much confidence in the values of the α_m beyond the fifteenth decimal. Whether this means we should have gone to a still larger value of λ, or whether it merely reflects a large accumulation of round off error from an extended calculation is not clear. As usual, the bound (8.47) is of little value. According to it, one would have had to carry λ past 2500 to begin to get the α_m correct to more than fifteen decimals.

One can also wonder if we should have used a value of Ω greater than 130. The approximations shown for $\alpha_m^{(0)}$ in Table 6.7 were rounded from a triple precision calculation using $\Omega = 256$. According to (8.34), this was a large enough value of Ω even for triple precision. So we may have confidence in the values listed for $\alpha_m^{(0)}$, except for the latter digits of the last two or three entries. From the listed values of $\alpha_m^{(0)}$, it appears that $\Omega = 120$ would be more than adequate for a double precision calculation of the $\alpha_m^{(0)}$.

What of the other $\alpha_m^{(\lambda)}$, and of the α_m? By Theorem 8.10, we might have needed to take Ω as large as 300. However, the bounds given by Theorem 8.10 appear to be quite excessively large. Certainly, if our value of 130 for Ω had been seriously too small (like less than half enough), we could not have done as well as we did on the C-test. Actually, our C-test, though a bit disappointing, was really fairly good. We shall present some other evidence to suggest that $\Omega = 130$ was adequately large. Indeed, one could likely have taken $\Omega = 128 = 2^7$ in safety, and hence been able to make good use of the FFT.

This evidence is based on certain properties which apparently all the sequences $\{\alpha_m^{(0)}\}$ and $\{\alpha_m\}$ have, whatever the value of ℓ. However, the properties become much more prominent for the larger values of ℓ.

To describe these properties, we make certain definitions. Let $a_1, a_2, a_3, \ldots$ be a sequence. We will say that a_m is a local minimum if $a_m < a_{m-1}$ and $a_m < a_{m+1}$. We will say that a_m is a local maximum if $a_m > a_{m-1}$ and $a_m > a_{m+1}$. For each of the sequences $\{\alpha_m^{(0)}\}$ and $\{\alpha_m\}$ shown in Tables 6.1, 6.3, 6.5, 6.6, and 6.7, the first term (for $m = 1$) is less than all the rest. We will count it also as a local minimum; it is likewise a global minimum.

In Tables 6.5, 6.6, and 6.7, only values of $\alpha_m^{(0)}$ or α_m are listed which are either local minima or local maxima. We refer to these as local extrema. Every local extremum is listed as far as the tables extend.

From the definitions just given, it is a triviality that in going from a local minimum to the next local maximum, the terms increase monotonically, and in going from a local maximum to the next local minimum, the terms decrease monotonically. (It happens that in all our sequences, there are no cases where consecutive terms are equal.) However, in Table 6.7 it is uniformly around eleven steps from each local minimum to the next local maximum and around eleven more steps to the next local minimum. In such a case, the triviality becomes a very striking phenomenon, which is seen for both the sequences $\{\alpha_m^{(0)}\}$ and $\{\alpha_m\}$ when $\ell = 200$ The same thing, with a smaller number of steps from one local extremum to the next local extremum, occurs for smaller values of ℓ .

We refer to this relative uniformity in the number of steps from one local extremum to the next local extremum as property A.

As far as our calculations go, each of the sequences $\{\alpha_m^{(0)}\}$ and $\{\alpha_m\}$ for $\ell = 1.6$, 4, 10, 20, and 200 has the following further properties.

B. Each local minimum is negative, and each local maximum is positive.

C. If a_m is a local extremum and a_n is the next local extremum, then $|a_m| > |a_n|$.

As it happens, property B is a logical consequence of property C .

It follows from properties B and C (and hence from property C alone) that each local minimum is less than all subsequent terms, and that each local maximum is greater than all subsequent terms.

The rate of decrease from one local maximum to the next local maximum is not markedly different for different values of ℓ . To illustrate this we list in Table 6.8 the fourth local maxima of the various sequences. A similar thing is true of local minima, as illustrated in Table 6.9, where we have listed the fourth local minima of the various sequences.

In Table 6.1, we listed all terms of the sequences out to the fourth local maximum. In Table 6.3, we listed all terms of the sequences out to the fifth local maximum. However, it is questionable if the listings for $m = 16$ and $m = 17$ in Table 6.3 are significant; most likely, Table 6.3 goes out only to the fifth local minimum, but has a couple of extraneous extra entries. In Tables 6.5, 6.6, and 6.7, we listed local minima and local maxima, and no other values, since this sufficed to show the general behavior of the sequences; recall that the sequences progress monotonely from one term shown to the next term shown. Indeed, they do so with surprising regularity.

In the case of $\ell = 200$, the $\alpha_m^{(0)}$ were computed with triple precision, which is why this sequence has such a long listing in Table 6.7.

The remarkably parallel behavior of $\alpha_m^{(0)}$ and α_m for a given value of ℓ can certainly not be fortuitous, though we have not a clue as to an explanation

Table 6.8

The fourth local maxima.

ℓ	$\{\alpha_m^{(0)}\}$	$\{\alpha_m\}$
1.6	5.0×10^{-17}	4.7×10^{-17}
4	1.3×10^{-14}	1.1×10^{-14}
10	7.8×10^{-14}	4.9×10^{-14}
20	8.2×10^{-14}	4.1×10^{-14}
200	2.0×10^{-14}	3.4×10^{-15}

Table 6.9

The fourth local minima.

ℓ	$\{\alpha_m^{(0)}\}$	$\{\alpha_m\}$
1.6	-1.1×10^{-15}	-1.0×10^{-15}
4	-7.3×10^{-13}	-6.0×10^{-13}
10	-1.8×10^{-12}	-1.2×10^{-12}
20	-2.8×10^{-12}	-1.4×10^{-12}
200	-5.7×10^{-13}	-9.8×10^{-14}

Table 6.10

Ratios of local extrema.

$\ell = 1.6$	$\ell = 4$	$\ell = 10$	$\ell = 20$	$\ell = 200$
1.006	1.077	1.377	1.881	3.530
1.027	4.558	1.466	0.695	0.231
1.127	0.903	0.631	0.525	0.163
0.902	0.808	0.608	0.480	0.170
0.935	0.795	0.654	0.478	0.175
0.963	0.901	0.623	0.494	0.172
0.914	0.831	0.629	0.483	0.173
0.941	0.810	0.629	0.498	0.172
	0.806	0.620	0.486	0.175
	0.937	0.632	0.475	

of it. Not only do local extrema of $\alpha_m^{(0)}$ and α_m occur at nearly equal values of m, but their relative size behaves very uniformly. To illustrate this, we have listed in Table 6.10 for each value of ℓ the ratio of the n-th local extremum of α_m divided by the n-th local extremum of $\alpha_m^{(0)}$. The relative uniformity of the final entry under $\ell = 4$ suggests that the entries for $m = 16$ and 17 in Table 6.3 are perhaps not entirely spurious, though they cannot possibly have more than their first significant digits correct.

The behavior of the ratios in Table 6.10 suggests very strongly that for a given value of ℓ, each of the sequences $\alpha_m^{(0)}$ and α_m are values at integer arguments from two damped oscillatory functions, and that the ratio of the n-th extrema of these functions approaches a limit as the argument tends to infinity. As the extrema of the functions will usually not occur at integer values of the arguments, the ratios of the n-th extrema of the sequences will move somewhat randomly about this limit.

If this, or something a bit like it, is the case, it would be of much value to learn how to prove it. It would certainly give us much more information than we now have as to bounds for the $\alpha_m^{(\lambda)}$ and α_m, and hence as to the value we should take for Ω. If sufficiently precise information were available, we might even know how to compute the α_m directly, without proceeding through the limiting procedure which we now use.

In the absence of any sort of proofs, the extremely uniform behavior of the ratios for $\ell = 200$ in Table 6.10 suggests most strongly that the sizes of the local minima and local maxima in Table 6.7 are approximately right. As the $\alpha_m^{(\lambda)}$ behaved quite similarly to $\alpha_m^{(0)}$ and α_m, it would appear confirmed (though of course not proved) that $\Omega = 130$ was a large enough value for double precision for $\ell = 200$.

7. THE ε-ALGORITHM. The ε-algorithm is a transformation which has been used with much success to accelerate the convergence of sequences. The reader should be warned that the ε-algorithm is a nonlinear transformation. The definition and a basic property are given in [7], p. 30. The entire Chapter III of [8], pp. 37-95, is taken up with properties of the ε-algorithm. The ε-algorithm was discovered by Peter Wynn; see [9], which contains the definition and a number of basic properties.

Let $\{A_\lambda\}$ be a sequence whose convergence is to be accelerated. Specifically, for a fixed n, let $A_\lambda = \alpha_n^{(\lambda)}$; for large ℓ this converges very slowly to α_n, and we wish to speed up the process. We define quantities $\varepsilon_s^{(\lambda)}$ for $\lambda \geq 0$ and $s \geq -1$; we use recursion on s according to the scheme

$$(\varepsilon_{s+1}^{(\lambda)} - \varepsilon_{s-1}^{(\lambda+1)})(\varepsilon_s^{(\lambda+1)} - \varepsilon_s^{(\lambda)}) = 1 \ . \tag{7.1}$$

It is usual to picture the $\varepsilon_s^{(\lambda)}$ in a triangular array, as in Figure 7.1. Then the relation (7.1) involves the four corners of a rhombus in the array. It is called the "rhombus rule".

	$\varepsilon_0^{(0)}$					
		$\varepsilon_1^{(0)}$				
$\varepsilon_{-1}^{(1)}$						
	$\varepsilon_0^{(1)}$		$\varepsilon_2^{(0)}$			
		$\varepsilon_1^{(1)}$		$\varepsilon_3^{(0)}$		
$\varepsilon_{-1}^{(2)}$						
	$\varepsilon_0^{(2)}$		$\varepsilon_2^{(1)}$		$\varepsilon_4^{(0)}$	
		$\varepsilon_1^{(2)}$		$\varepsilon_3^{(1)}$		$\varepsilon_5^{(0)}$
$\varepsilon_{-1}^{(3)}$						
	$\varepsilon_0^{(3)}$		$\varepsilon_2^{(2)}$		$\varepsilon_4^{(1)}$	
		$\varepsilon_1^{(3)}$		$\varepsilon_3^{(2)}$		
$\varepsilon_{-1}^{(4)}$						
	$\varepsilon_0^{(4)}$		$\varepsilon_2^{(3)}$			
		$\varepsilon_1^{(4)}$				
$\varepsilon_{-1}^{(5)}$						
	$\varepsilon_0^{(5)}$					

Figure 7.1.

To get the recursive definition of the $\varepsilon_s^{(\lambda)}$ started, we define

$$\varepsilon_{-1}^{(\lambda)} = 0 \qquad (\lambda \geq 1) \tag{7.2}$$

$$\varepsilon_0^{(\lambda)} = A_\lambda \qquad (\lambda \geq 0). \tag{7.3}$$

With (7.2) and (7.3), we can calculate the $\varepsilon_1^{(\lambda)}$ by (7.1); then we can calculate the $\varepsilon_2^{(\lambda)}$, then the $\varepsilon_3^{(\lambda)}$, and so on.

The point of this is that if

$$\lim_{\lambda \to \infty} A_\lambda = A , \tag{7.4}$$

then very often

$$\lim_{s \to \infty} \varepsilon_{2s}^{(\lambda)} = A . \tag{7.5}$$

Various sets of sufficient conditions for this to take place are given in [8]. One of these is discussed on pp. 43-50 of [15]. Surprisingly often, the convergence in (7.5) is much faster than in (7.4). We found this to be the case when we took $A_\lambda = \alpha_n^{(\lambda)}$ for a fixed n.

We used a computer program supplied by Peter Wynn. If one has an ALGOL compiler, one can use the ALGOL program given in [10]. Or one can use the programs given in [11] and [12]. They are attributed to Peter Wynn, but he disclaims knowledge of their existence.

The ε- algorithm was invoked first for $\ell = 20$. Not knowing what to expect, we had the computer calculate and print out the $\varepsilon_{2s}^{(\lambda)}$ up to $2s = 60$ for the case

$$\varepsilon_0^{(\lambda)} = \alpha_1^{(\lambda)} .$$

This was far more than needed. Already at $2s = 14$, all the $\varepsilon_{2s}^{(\lambda)}$ were essentially equal to α_1. They differ from it and each other by only a few units in the final decimal place. With the inevitable round off errors, one could not hope to do better than this. To calculate $\varepsilon_{14}^{(0)}$ requires knowledge of $\varepsilon_0^{(\lambda)}$ only for $\lambda = 0,1,\ldots,15$. That is, from knowing $\alpha_1^{(0)},\alpha_1^{(1)},\alpha_1^{(2)},\ldots,\alpha_1^{(15)}$, one could get a fine estimate for α_1, despite the fact that one must take λ as large as 60 to 70 before $\alpha_1^{(\lambda)}$ is equally close to α_1.

Based on this, we tried the cases $m = 2,3,\ldots,15$, taking

$$\varepsilon_0^{(\lambda)} = \alpha_m^{(\lambda)} ,$$

in accordance with (7.3). In each case we took $\lambda = 0,1,\ldots,15$. This allows us to compute $\varepsilon_{14}^{(0)}$ for each value of m. For each m, we got an approximation for α_m good to at least 18 decimal digits to the right of the decimal place. To get $\alpha_m^{(\lambda)}$ this close to α_m, we generally needed λ greater than 60.

Thus we got quite satisfactory acceleration of convergence in these cases.

For $\ell = 200$, we had to take λ up to about 550 to get $\alpha_m^{(\lambda)}$ approximately equal to α_m, to the accuracy of the calculation (which was double precision). In general, for $1 \le m \le 20$, $\alpha_m^{(61)}$ would not agree with α_m to more than two significant decimal digits, sometimes only one. Then for $1 \le m \le 20$, we took for each m

$$\varepsilon_0^{(\lambda)} = \alpha_m^{(\lambda)} ,$$

in accordance with (7.3) for $0 \le \lambda \quad 61$. The $\varepsilon_{60}^{(\lambda)}$ agree fairly well with the corresponding α_m; for $m = 3$, there was agreement in the first 11 decimal digits after the decimal point, and for the other values of m, there was agreement to 12 or more decimal digits after the decimal point.

This certainly was a very, very major improvement for a fairly minor computing investment. It looked as if we could have done still better by carrying λ on beyond 61. However, instead we tried an iteration of the ε-algorithm. This is discussed in [13], and in some cases has had sensational success. In our case, it worked out quite poorly. This whole area needs to be explored more carefully.

Possibly more critical is the question how sensitive the ε-algorithm is to the build up of round off error. Our experience seems to suggest that this is not serious. However, the question deserves more careful study. We might remark that in our calculations for $\ell = 200$, the $\varepsilon_{2s}^{(\lambda)}$ had erratic irregularities in their latter digits, especially for larger values of s. We did not explore the matter enough to discover the cause of these. The irregularities did not seem to be such as one would expect from round off errors. Another possible contributing factor was the fact that for $0 \leq \lambda \leq 10$ we computed the $\alpha_m^{(\lambda)}$ by (5.11) by the FFT, using triple precision, while for larger λ we used

$$\alpha_m^{(\lambda)} = \alpha_m^{(\lambda-1)} + \beta_m^{(\lambda)}$$

where the $\beta_m^{(\lambda)}$ were computed by (5.12) using double precision. While this could certainly cause irregularities, they should be much smaller than those observed.

We can only repeat that, while we had good success with the ε-algorithm, there is need for more careful study. We should point out that it is safe to experiment, if one takes the precaution of applying the C-test after values for the α_n have been calculated. As we noted before, if the C-test works, one certainly has accurate enough values for the α_n, no matter how they were obtained.

Given what seem to be suitable approximations for the α_m, one must still sum two series to calculate $W(z)$; see (4.11) and (4.12). Indeed, if we wish to invoke the C-test as a test of our approximations for the α_m, we must sum two series; see (4.20). One of these is a Fourier series. A method to apply the ε-algorithm to accelerate the convergence of Fourier series is given in [14].

For theoretical reasons, which are explained in [14], and on pp. 55-56 of [15], the thing to do is to write the Fourier series as the real part of an exponential series. In our case, (4.20) comes already in this form. In fact (4.20) can be rewritten as

$$C(\theta) = \Re\{\sum_{m=1}^{\infty} \alpha_m e^{2m\theta i} - \ell n(w(2\cos\theta + \frac{8}{\ell} i)) - \sum_{m=1}^{\infty} \alpha_m (w(2\cos\theta + \frac{8}{\ell} i))^{2m}\}. \quad (7.6)$$

So we seek the values of

$$\sum_{m=1}^{\infty} \alpha_m e^{2m\theta i} \quad (7.7)$$

and

$$\sum_{m=1}^{\infty} \alpha_m (w(2\cos\theta + \frac{8}{\ell} i))^{2m} . \quad (7.8)$$

Clearly (7.7) is

$$\lim_{\lambda \to \infty} A_\lambda \quad (7.9)$$

if we define

$$A_\lambda = \sum_{m=1}^{\lambda} \alpha_m e^{2m\theta i} . \quad (7.10)$$

So we apply the ε-algorithm just as before to accelerate the convergence of the A_λ. The only difference is that since the A_λ are complex numbers, the $\varepsilon_{2s}^{(\lambda)}$ will be also. However, the computer can be set to do complex arithmetic. After the convergence has been accelerated, one just takes the real part.

From Table 6.7, we see that to get an approximation for (7.7) correct to 16 decimal digits to the right of the decimal point we can take

$$\Re\{A_\lambda\}$$

for $\lambda \geq 100$. By contrast

$$\Re\{\varepsilon_{40}^{(0)}\}$$

gave 16 correct decimal digits for $\theta = \frac{1}{2}\pi$, and 12 for $\theta = 0$. Probably

$$\Re\{\varepsilon_{50}^{(0)}\}$$

would have given 16 correct decimal digits for both values of θ .

As the calculation of the $\varepsilon_s^{(\lambda)}$ involved complex arithmetic, it would have used less calculation to calculate (7.7) directly, if the first 100 α_m's are known. On the other hand, to calculate $\varepsilon_{50}^{(0)}$ requires knowledge of only the first 50 α_m's.

As it is fairly laborious to calculate each α_m , it may be faster to calculate 50 α_m's and then do an ε-algorithm with complex arithmetic than to calculate 100 α_m's and then use (7.7) directly.

What about (7.8)? The theory in [14] and [15] depends basically on the fact that we are dealing with powers. So it is perfectly applicable to (7.8). We proceed with this quite analogously to the way we proceeded with (7.7). In this case

$$\Re\{\varepsilon_{40}^{(0)}\}$$

gave 17 decimals both when $\theta = \frac{1}{2}\pi$ and when $\theta = 0$.

In the above, we have used two ε-algorithms. One is for a sequence $\{A_\lambda\}$, defined by (7.10). The other is for a sequence $\{B_\lambda\}$ defined analogously from (7.8). We wish the limit of the sequence

$$\{A_\lambda - B_\lambda\} \quad .$$

Why not apply a single ε-algorithm directly to this sequence? This was tried, but gave much poorer acceleration of convergence. The reason for this can be explained by some of the theory in [8]. It is also explained on pp. 58-59 of [15]. It is pretty complex, so we skip it. However, it is the case that it is advisable to use the ε-algorithm on each of (7.7) and (7.8) separately, rather than combining them so as to use a single ε-algorithm The same will apply to the two series appearing in (4.12). Actually, unless one is very close to one of the "plates", each series will converge fairly rapidly, and it is likely not worth bothering with the ε-algorithm

8. RATE OF CONVERGENCE. By (4.13), (4.7), and (4.15), we have

$$\alpha_m^{(\lambda+1)} = \Re\{\frac{1}{\pi}\int_0^\pi \ell n(w(2\cos\theta + \frac{8}{\ell}i))[e^{2mi\theta} + e^{-2mi\theta}]d\theta\}$$

$$+ \Re\{\sum_{n=1}^{\infty}\alpha_n^{(\lambda)}\frac{1}{\pi}\int_0^\pi (w(2\cos\theta + \frac{8}{\ell}i))^{2n}[e^{2mi\theta} + e^{-2mi\theta}]d\theta\}. \qquad (8.1)$$

Define

$$v(u) = u + \frac{1}{u} + \frac{8}{\ell}i. \qquad (8.2)$$

Then, by setting $u = e^{i\theta}$, we may rewrite (8.1) as

$$\alpha_m^{(\lambda+1)} = \Re\{\frac{1}{\pi i}\int \ell n(w(v(u)))[u^{2m-1} + u^{-2m-1}]du\}$$

$$+ \Re\{\frac{1}{\pi i}\sum_{n=1}^{\infty}\alpha_n^{(\lambda)}\int (w(v(u)))^{2n}[u^{2m-1} + u^{-2m-1}]du\} \qquad (8.3)$$

where the integration is counter clockwise along the top half of the unit circle. In the u^{-2m-1} part, set $t = u^{-1}$. As

$$v(u) = v(u^{-1}),$$

we may rewrite (8.3) as

$$\alpha_m^{(\lambda+1)} = \Re\{\frac{1}{\pi i}\oint \ell n(w(v(u)))u^{2m-1}du\} + \Re\{\frac{1}{\pi i}\sum_{n=1}^{\infty}\alpha_n^{(\lambda)}\oint (w(v(v)))^{2n}u^{2m-1}du\} \qquad (8.4)$$

where the integration is around the unit circle.

Suppose we take $0 < a < 1$, and deform the contour of integration in (8.4) to go around a circle of radius a, with center at the origin (if we can). If the summation in (8.4) behaves decently, then the u^{2m-1} term in (8.4) will let us conclude that

$$|\alpha_m^{(\lambda+1)}| \leq K a^{2m}, \qquad (8.5)$$

for some K. This will be useful in establishing convergence.

If we can deform the contour as indicated, so that $|u| = a$ in (8.4), then by (8.2) v lies on the ellipse

$$\frac{x^2}{(a + a^{-1})^2} + \frac{(y - 8/\ell)^2}{(a^{-1} - a)^2} = 1. \qquad (8.6)$$

More generally, if $a < |u| < 1/a$, then v lies inside the ellipse (8.6). Put

$$\alpha = (a + a^{-1})^2. \qquad (8.7)$$

Then the equation of the ellipse is

$$\frac{x^2}{\alpha} + \frac{(y - 8/\ell)^2}{\alpha - 4} = 1. \qquad (8.8)$$

If this ellipse contains no points of the DCC in its interior (see Figure 3.1), then one easily sees that $w(v(u))$ is analytic for $a < |u| < 1/a$. Also, since w carries the s-plane minus DCC into the lower half of the w-plane, $\ell n(w(v(u)))$ is also analytic for $a < |u| < 1/a$. Thus, in (8.4) one can deform the contour of integration into any circle of radius between a and $1/a$.

If one decreases a from 1 towards 0, the ellipse (8.6) increases in size, but always with the foci at $(\pm 2, 8/\ell)$. So there will be a unique a, $a = A$, for which the ellipse passes through $(\pm 2, 0)$, on the DCC. For $a < A$, there will be points of the DCC inside the ellipse, but for $A \leq a \leq 1/A$ there will be no points of DCC inside the ellipse. We undertake to determine A.

Define

$$k = k(\ell) = \sqrt{2 + \sqrt{4 + \ell^2}} \,. \tag{8.9}$$

We have

$$k^4 = 4k^2 + \ell^2 \,. \tag{8.10}$$

We will show that

$$A = A(\ell) = \frac{\ell}{k(2 + k)} \,. \tag{8.11}$$

Note that $k > \sqrt{\ell}$, so that

$$0 < A < 1 \,. \tag{8.12}$$

As we pointed out above, it suffices to show that if $a = A$, then the ellipse (8.6) passes through $(\pm 2, 0)$. We have

$$A + \frac{1}{A} = \frac{\ell^2 + k^2(2 + k)^2}{\ell k(2 + k)} = \frac{\ell^2 + 4k^2 + 4k^3 + k^4}{\ell k(2+k)} \,.$$

If we substitute the right side of (8.10) into this, we get

$$A + \frac{1}{A} = \frac{2k^2}{\ell} \,. \tag{8.13}$$

Taking $a = A$ in (8.7) gives

$$\alpha = \frac{4k^4}{\ell^2} \,. \tag{8.14}$$

By (8.10), this gives

$$\alpha - 4 = \frac{16k^2}{\ell^2} \,. \tag{8.15}$$

If $(\pm 2, 0)$ is to be on the ellipse (8.6), which is the same as the ellipse (8.8), we must have

$$\frac{4}{\alpha} + \frac{64}{\ell^2(\alpha - 4)} = 1 \,.$$

If we substitute α from (8.14) and $\alpha-4$ from (8.15) into this, we find by (8.10) that it is satisfied.

So we have proved the following theorem.

Theorem 8.1. If A is defined by (8.11) then $w(v(u))$ and $\ell n(w(v(u)))$ are analytic for $A < |u| < 1/A$.

We will have

$$|w(v(u))| \leq 1 \quad \text{for} \quad |u| = a$$

if we choose a so that the ellipse (8.6) does not contain any points of SCC. For this, it suffices that the ellipse not enter the lower half plane. This can be assured by choosing a so that the ellipse passes through the origin. Then, since $0 < a < 1$, $8/\ell = (1/a) - a$.

Theorem 8.2. If we define

$$\bar{A} = \bar{A}(\ell) = \frac{\ell}{4 + \sqrt{16 + \ell^2}} ,$$

then $|w(v(u))| \leq 1$ for

$$\bar{A} \leq |u| \leq \frac{1}{\bar{A}} .$$

The argument above shows that $A(\ell) < \bar{A}(\ell)$. So

$$0 < A(\ell) < \bar{A}(\ell) < \frac{\ell}{4+\ell} < 1 . \tag{8.16}$$

We wish to find the bounds on $|w(v(u))|$ for $|u| = a < 1$. By (3.1),

$$v(u) = w(v(u)) + \frac{1}{w(v(u))} .$$

So, by (3.1) and (3.12),

$$|w(v(u))| = b \tag{8.17}$$

if and only if v lies on the ellipse

$$\frac{x^2}{\beta} + \frac{y^2}{\beta-4} = 1 , \tag{8.18}$$

where

$$\beta = (b + b^{-1})^2 . \tag{8.19}$$

Thus, if we choose β so that the ellipse (8.18) intersects the ellipse (8.8), then for some u with $|u| = a$ we will have (8.17) satisfied. Conversely, if the ellipses (8.18) and (8.8) have no points in common, then there is no u with $|u| = a$ for which (8.17) holds.

We shall be interested in the extreme case in which $a = A(\ell)$. Since we may take a as close to $A(\ell)$ as we wish, we can, by taking limits in the results obtained, in effect take $a = A(\ell)$.

Let us first find the minimum b for a given a . One might think that the minimum b for which the two ellipses would intersect would be the one for which the top point of the ellipse (8.18) coincides with the top point of the ellipse (8.8). For this, one would have $b^{-1} - b = (8/\ell) + A^{-1} - A$. By (8.11) and (8.10), $b^{-1} - b = 4(2+k)/\ell$, so that

$$b = \frac{\ell}{4 + 2k + \sqrt{(4 + 2k)^2 + \ell^2}} \quad . \tag{8.20}$$

However, there are two possible configurations when the top points of the two ellipses coincide. In one configuration, the ellipse (8.8) lies entirely inside the ellipse (8.18) except for the single point of tangency at the top. In this case, a decrease in b would cause an increase in β , which would enlarge the ellipse (8.18), so that then (8.8) would lie entirely inside it. In this case, the b which makes the top points coincide would be the least b for which the two ellipses would intersect, and so would be the least value of $|w(v(u))|$ for $|u| = A$.

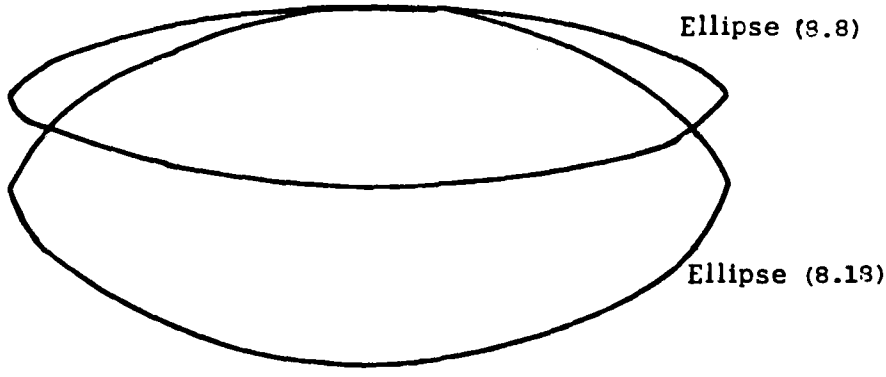


Figure 8.1.

In the other possible configuration, illustrated in Figure 8.1, the ellipse (8.8) lies partly outside the ellipse (8.18). By decreasing b and increasing β , we can enlarge the ellipse (8.18) to a size such that the ellipse (8.8) lies entirely inside except for TWO points of tangency. The value of b that produces this condition is the minimum of $|w(v(u))|$ for $|u| = A$.

To find the y-coordinates of the points of intersection of the two ellipses, we eliminate x between (8.8) and (8.18). This gives

$$4(\beta - \alpha)y^2 - \frac{16\alpha(\beta - 4)}{\ell} y + \frac{64\alpha(\beta - 4)}{2} + (\beta - \alpha)(\alpha - 4)(\beta - 4) = 0.$$

But when $a = A$, we have by (8.14) and (8.15)

$$\alpha = \frac{4k^4}{\ell^2} = 4 + \frac{16k^2}{\ell^2} \quad .$$

So the equation for y reduces to

$$4(\beta - \alpha)y^2 - \frac{16\alpha(\beta - 4)}{\ell}y + \frac{16k^2(\beta - 4)^2}{\ell^2} = 0. \tag{8.21}$$

At $b = 1$, we have $\beta = 4$. For $A < b < 1$, we have $4 < \beta < \alpha$. In this range, we see that (8.21) has both a positive and a negative solution. As β approaches α, the negative solution approaches $-\infty$ and the positive solution approaches $4/\ell$. For $\beta > \alpha$, there are two positive solutions, until we reach $\beta = \alpha + (\alpha^2/k^2)$. By (8.14) and (8.15), this equals $\alpha^2/4$. At this value of β, (8.21) has a double root, and for larger values of β there are only non-real solutions of (8.21). So, the largest value of β for which (8.21) has real solutions is

$$\beta = \frac{4k^8}{\ell^4} = 4 + \frac{32k^2}{\ell^2} + \frac{64k^4}{\ell^4} \quad . \tag{8.22}$$

By our earlier analysis, the value of β at which the top point of the ellipse (8.18) coincides with the top point of the ellipse (8.8) is

$$\beta = 4 + \frac{16}{\ell^2}(2 + k)^2 \quad . \tag{8.23}$$

If we give β the value (8.23), then there is certainly one common point i the two ellipses, namely their common top point. So (8.21) has at least one real root. So (8.23) cannot be greater than (8.22), else when β has the value (8.23 there would be no real roots of (8.21).

The question is whether we have the situation of Figure 8.1 when we give β the value (8.23). For this to happen, the two solutions of (8.21) would have to be the y-coordinate of the common top point of the two ellipses and a SMALLER number. The y-coordinate of the common top point is $\sqrt{\beta-4} = 4(2+k)/\ell$, by the result just before (8.20). The product of the two solutions of (8.21) is

$$\frac{16k^2(\beta - 4)^2}{\ell^2} \div 4(\beta - \alpha) \ ,$$

which is

$$\frac{16k^2}{\ell^2}\left\{\frac{16}{\ell^2}(2 + k)^2\right\}^2 \div 4\left\{\frac{64}{\ell^2}(1 + k)\right\},$$

which is

$$\frac{16k^2(2 + k)^4}{\ell^4(1 + k)} \quad .$$

Then the other solution of (8.21) would be

$$\frac{4k^2(2 + k)^3}{\ell^3(1 + k)} \quad .$$

Thus we get the situation of Figure 8.1 if and only if

$$\frac{4}{\ell}(2 + k) > \frac{4k^2(2 + k)^3}{\ell^3(1 + k)} ,$$

which is equivalent to $\ell^2(1 + k) > k^2(2 + k)^2$. As $\ell^2 = k^4 - 4k^2$ by (8.10), this is equivalent to $(k - 2)(1 + k) > 2 + k$. As $k > 2$, this is equivalent to $k > 1 + \sqrt{5}$. As we said, when this holds, we have the situation of Figure 8.1. In this case we should define β by (8.22). Otherwise, we should use (8.23).

Theorem 8.3. Let us define $B = B(\ell)$ to be

$$B = \frac{\ell}{4 + 2k + \sqrt{(4 + 2k)^2 + \ell^2}} \tag{8.24}$$

if $\ell^2 \leq 16(2 + \sqrt{5})$ and take it to be

$$B = \frac{\ell^2}{k^4 + \sqrt{8k^2\ell^2 + 16k^4}} \tag{8.25}$$

$$= \frac{\ell^2}{k^2(k^2 + \sqrt{8\sqrt{4 + \ell^2}})}$$

$$= \frac{\ell^2}{k^2(\sqrt{2} + (4 + \ell^2)^{\frac{1}{4}})^2}$$

otherwise. Then for $A < |u| < 1/A$, we have $|w(v(u))| > B$.

The proof of the theorem follows from the analysis above. This same analysis shows that the maximum b would be the one for which the bottom point of the ellipse (8.18) coincides with the bottom point of the ellipse (8.8). Thus we can conclude:

Theorem 8.4. Let us define

$$\bar{B} = \bar{B}(\ell) = \frac{2k - 4 + \sqrt{(2k-4)^2 + \ell^2}}{\ell} . \tag{8.26}$$

Then for $A < |u| < 1/A$, we have $|w(v(u))| < \bar{B}$.

To avoid cancellation in calculating $2k-4$, we can use (8.10) to get

$$2k-4 = \frac{2(k^2-4)}{k-2} = \frac{2\ell^2}{k^2(k+2)} .$$

Theorem 8.5. We have $A\bar{B} < 1$.

Proof. If $\bar{B}$ were as large as $1/A$, then the ellipse (8.18) would be at least as large as the ellipse (8.8). But then their bottom points could not coincide.

Put

$$w(v(u)) = \bar{r}e^{i\bar{\theta}} , \tag{8.27}$$

with $0 < \bar{r}$.

Recall that we are using DCC in the v-plane. Therefore, by the formula above (8.17), the entire v-plane minus DCC is mapped into the lower half of the w-plane; see our earlier discussion of (3.1). So in (8.27), we must have

$$-\pi < \bar{\theta} < 0 . \tag{8.28}$$

Recall that $\alpha_m^{(-1)} = 0$ (see (4.9)). So by (8.4)

$$\alpha_m^{(0)} = \mathcal{R}\{\frac{1}{\pi i}\oint \ell n(w(v(u)))u^{2m-1}du\} ,$$

where the path of integration is the unit circle. Take d a constant. Then for $1 \leq m$

$$\alpha_m^{(0)} = \mathcal{R}\{\frac{1}{\pi i}\oint \{\ell n(w(v(u))) + d\}u^{2m-1}du\} . \tag{8.29}$$

Take ε positive and very small. Then, by Theorem 8.1, we can deform the path of integration to the circle $|u| = A + \varepsilon$.

Let us take

$$d = \tfrac{1}{2}\pi i - \tfrac{1}{2}\ell n(B\bar{B}),$$

$$D = D(\ell) = \tfrac{1}{2}\left|\pi i + \ell n(\bar{B}/B)\right| . \tag{8.30}$$

Then by (8.28) and Theorems 8.3 and 8.4 $\left|\ell n(w(v(u))) + d\right| < D$ for $u = A + \varepsilon$. So by (8.29) $|\alpha_m^{(0)}| < 2D(A + \varepsilon)^{2m}$ for $1 \leq m$. Letting ε tend to zero gives

$$|\alpha_m^{(0)}| \leq 2DA^{2m} \tag{8.31}$$

for $1 \leq m$. Then for $1 \leq N$

$$\sum_{m=N}^{\infty} \bar{B}^{2m}\,|\alpha_m^{(0)}| \leq \frac{2D(A\bar{B})^{2N}}{1 - (A\bar{B})^2} . \tag{8.32}$$

Put

$$\gamma = \frac{2(A\bar{B})^2}{1 - (A\bar{B})^2} . \tag{8.33}$$

Theorem 8.6. For $0 \le \lambda$ and $1 \le m$,

$$|\alpha_m^{(\lambda)}| \le 2DA^{2m} \frac{\gamma^{\lambda+1} - 1}{\gamma - 1}, \tag{8.34}$$

$$\sum_{m=1}^{\infty} \bar{B}^{2m} |\alpha_m^{(\lambda)}| \le D \frac{\gamma^{\lambda+2} - \gamma}{\gamma - 1}. \tag{8.35}$$

Proof by induction on λ. By (8.31), (8.32), and (8.33), the theorem holds for $\lambda = 0$. So assume the theorem for λ. Then, by Theorems 8.4 and 8.5, the right side of (8.4) converges uniformly for $A + \varepsilon \le |u| \le 1$, for sufficiently small positive ε. So we deform the contours of integration to $|u| = A + \varepsilon$. On this circle, $|\ell n(w(v(u))) + d| < D$, as we had before, and $|w(v(u))| < \bar{B}$ by Theorem 8.4. So by (8.4) and (8.35)

$$|\alpha_m^{(\lambda+1)}| < 2D(A + \varepsilon)^{2m} + 2D \frac{\gamma^{\lambda+2} - \gamma}{\gamma - 1} (A + \varepsilon)^{2m}.$$

Letting ε tend to zero gives (8.34) for $\lambda + 1$. By (8.34) for $\lambda + 1$, using Theorem 8.5, we get

$$\sum_{m=1}^{\infty} \bar{B}^{2m} |\alpha_m^{(\lambda+1)}| \le D \frac{\gamma^{\lambda+2} - 1}{\gamma - 1} \frac{2(A\bar{B})^2}{1-(A\bar{B}^2)}.$$

By (8.33), this is (8.35) for $\lambda+1$.

If $\gamma = 1$, the fractions on the right sides of (8.34) and (8.35) should be replaced by $\lambda+1$, of course.

For the smaller values of ℓ, we will have $(A\bar{B})^2 < 1/3$. In such cases, $0 < \gamma < 1$. If (4.10) holds, and we shall show that it does, we can let $\lambda \to \infty$ in (8.34) and (8.35). We conclude

$$|\alpha_m| \le \frac{2DA^{2m}}{1-\gamma} \quad \text{if} \quad 0 < \gamma < 1, \tag{8.36}$$

$$\sum_{m=1}^{\infty} \bar{B}^{2m} |\alpha_m| \le \frac{D\gamma}{1-\gamma} \quad \text{if} \quad 0 < \gamma < 1. \tag{8.37}$$

We proceed to show that (4.10) holds.

Theorem 8.7. Let $f(z)$ be analytic inside and on the unit circle. Let 1 be the maximum variation of $\Re\{f(z)\}$ for z on the unit circle. That is,

$$|\Re\{f(e^{i\theta})\} - \Re\{f(e^{i\phi})\}| \le 1 \tag{8.38}$$

for all real θ and ϕ. Let $0 < r \le 1$. Then the maximum variation of the [illegible] $f(z)$ for $|z| \le r$ is bounded by H, where

$$H = H(r) = \frac{4}{\pi} \arctan r. \tag{8.39}$$

$$|\Re\{f(z_1)\} - \Re\{f(z_2)\}| \le H \tag{8.40}$$

for each z_1 and z_2 with $|z_1| \leq r$ and $|z_2| \leq r$.

Proof. As $\mathcal{R}\{f(z)\}$ is a harmonic function, we can conclude by the princip of the maximum that its greatest and least values are assumed on the unit circle So the maximum variation of $\mathcal{R}\{f(z)\}$ for z inside or on the unit circle is 1 So our theorem follows from Problem 289 on page 140 of Part 3 of [16].

Define

$$q^{(\lambda)}(t) = \sum_{n=1}^{\infty} \beta_n^{(\lambda)} t^{2n} . \tag{8.41}$$

By (4.14) and Theorem 8.6, $q^{(\lambda)}(t)$ is analytic for $|t| < A^{-2}$; hence $q^{(\lambda)}(t)$ is analytic inside and on the unit circle.

By (4.7) and (4.8), we see that the $\beta_m^{(\lambda+1)}$ are in effect Fourier coeffi-cients, so that for a suitably chosen (real) $\beta_0^{(\lambda+1)}$

$$\sum_{n=0}^{\infty} \beta_n^{(\lambda+1)} \cos 2n\theta = \sum_{n=1}^{\infty} \beta_n^{(\lambda)} \mathcal{R}\{(w(2 \cos \theta + \frac{8}{\ell} i))^{2n}\} .$$

So by (8.2)

$$\beta_0^{(\lambda+1)} + \mathcal{R}\{q^{(\lambda+1)}(u)\} = \mathcal{R}\{q^{(\lambda)}(w(v(u)))\} \tag{8.42}$$

for $|u| = 1$.

Similarly, by (4.3)

$$\beta_0^{(0)} + \mathcal{R}\{q^{(0)}(u)\} = \ell n|w(v(u))| \tag{8.43}$$

for $|u| = 1$.

Theorem 8.8. For $|u| = 1$, $A \leq |w(v(u))| \leq \bar{A}$.

Proof. By (8.2), for $u = 1$, we have $v(u)$ on the line segment connectin

$$-2 + \frac{8i}{\ell} \quad \text{with} \quad 2 + \frac{8i}{\ell} .$$

By (8.17), if $|w(v(u))| = b$, then v lies on the ellipse (8.18). If $b = \bar{A}$, then the ellipse passes through the midpoint of the line segment, and if $b = A$, then the ellipse passes through the end points.

Define

$$v^{(\lambda)} = \text{max. var. of } \mathcal{R}\{q^{(\lambda)}(t)\} \text{ for } |t| = 1 . \tag{8.44}$$

By (8.43) and Theorem 8.8 ,

$$v^{(0)} = \ell n \frac{\bar{A}}{A} . \tag{8.45}$$

In Theorem 8.7, we take $f(z) = \mathcal{R}\{q^{(\lambda)}(z)\} / v^{(\lambda)}$. By (8.44), the hypothesis of Theorem 8.7 is satisfied. So, by Theorem 8.8, the maximum variation of

$\mathcal{R}\{q^{(\lambda)}(w(v(u)))\}$ for $|u| = 1$ is less than or equal to $H(\bar{A})v^{(\lambda)}$. By (8.42) and (8.44) the maximum variation of $\mathcal{R}\{q^{(\lambda)}(w(v(u)))\}$ for $|u| = 1$ is $v^{(\lambda+1)}$. So $v^{(\lambda+1)} \leq H(\bar{A})v^{(\lambda)}$. Hence, by (8.45),

$$v^{(\lambda)} \leq (H(\bar{A}))^{\lambda} \ell n \frac{\bar{A}}{A} \quad . \tag{8.46}$$

As $\mathcal{R}\{q^{(\lambda)}(t)\}$ is harmonic for $|t| \leq 1$, its value at the origin (namely zero) is the average of its values on the unit circle. So by (8.44)

$$|\mathcal{R}\{q^{(\lambda)}(t)\}| \leq v^{(\lambda)}$$

for $|t| = 1$. By (8.41)

$$\mathcal{R}\{q^{(\lambda)}(e^{i\theta})\} = \sum_{n=1}^{\infty} \beta_n^{(\lambda)} \cos 2n\theta.$$

So, by the usual formula for determining Fourier coefficients,

$$\beta_n^{(\lambda)} = \frac{1}{\pi} \int_0^{2\pi} \mathcal{R}\{q^{(\lambda)}(e^{i\theta})\}\cos 2n\,\theta\, d\theta$$

for $n \geq 1$. So

$$|\beta_n^{(\lambda)}| \leq 2v^{(\lambda)} \tag{8.47}$$

for $n \geq 1$. Then by (6.1) and (8.46), α_n exists, and

$$|\alpha_n| \leq \frac{2\ell n(\bar{A}/A)}{1-H(\bar{A})} \quad . \tag{8.48}$$

Thus we have established that (4.10) is valid. We still need to prove that (5.8) can be made as small as desired, uniformly in λ.

Write K for the right side of (8.48). The same argument which established (8.48) gives

$$|\alpha_n^{(\lambda)}| \leq K \ . \tag{8.49}$$

Theorem 8.9. Suppose that A, C, and D have the properties that $A(\ell) \leq A < 1$, $AC < 1$, and for some d and for $A < |u| < 1/A$ we have

$$|\ell n(w(v(u))) + d| \leq D \ , \tag{8.50}$$

$$|w(v(u))| \leq C \ . \tag{8.51}$$

Choose M an integer such that

$$4(AC)^{2M} \leq 1 - (AC)^2 \ . \tag{8.52}$$

Define

$$L = L(M) = 4\left(D + \frac{C^{2M} - C^2}{C^2 - 1} K\right)/(1 - (AC)^2) \ . \tag{8.53}$$

Then for $1 \leq m$, $1 \leq \Omega$, and $0 \leq \lambda$

$$|\alpha_m^{(\lambda)}| \leq (1 - (AC)^2)LA^{2m} , \tag{8.54}$$

$$\sum_{m=\Omega}^{\infty} C^{2m}|\alpha_m^{(\lambda)}| \leq (AC)^{2\Omega}L , \tag{8.55}$$

$$|\alpha_m| \leq (1 - (AC)^2LA^{2m} , \tag{8.56}$$

$$\sum_{n=\Omega}^{\infty} C^{2m}|\alpha_m| \leq (AC)^{2\Omega}L . \tag{8.57}$$

Proof. We first prove (8.54) and (8.55) simultaneously by induction on λ. Using (8.50), we can derive $|\alpha_m^{(0)}| \leq 2DA^{2m}$ from (8.29) in the same way that (8.31) was derived. By (8.52), we have $M \geq 1$. So then (8.54) for $\lambda = 0$ follows by (8.53). From (8.54) we infer (8.55) for $\lambda = 0$. So assume that (8.54) and (8.55) hold for λ. In (8.4), take the contour of integration to be $|u| = A + \varepsilon$, with ε positive and very small. Then

$$|\alpha_m^{(\lambda+1)}| \leq 2(D + \sum_{n=1}^{\infty} C^{2n}|\alpha_n^{(\lambda)}|)(A + \varepsilon)^{2m} .$$

Letting ε tend to zero gives

$$|\alpha_m^{(\lambda+1)}| \leq 2(D + \sum_{n=1}^{\infty} C^{2n}|\alpha_n^{(\lambda)}|)A^{2m} .$$

By (8.49) and (8.53)

$$2(D + \sum_{n=1}^{M-1} C^{2n}|\alpha_n^{(\lambda)}|) \leq \tfrac{1}{2}(1 - (AC)^2)L .$$

By (8.55), with $\Omega = M$,

$$2\sum_{n=M}^{\infty} C^{2n}|\alpha_n^{(\lambda)}| \leq 2(AC)^{2M}L .$$

So by (8.52)

$$2\sum_{n=M}^{\infty} C^{2n}|\alpha_n^{(\lambda)}| \leq \tfrac{1}{2}(1 - (AC)^2)L .$$

Thus we are able to conclude that (8.54) holds for $\lambda + 1$. From it, we can deduce (8.55) for $\lambda+1$. Having shown that (8.54) holds for all λ, we can let $\lambda \to \infty$ and conclude that (8.56) holds. From it, we can deduce (8.57).

Corollary. The series

$$\sum_{m=1}^{\infty} C^{2m}|\alpha_m^{(\lambda)}|$$

converges uniformly in λ.

Proof. Use (8.53) and (8.55).

Theorem 8.10. In Theorem 8.9, we can take $A = A(\ell)$, $C = \bar{B}(\ell)$, and $D = D(\ell)$.

Proof. The condition (8.50) follows from the result just below (8.30). We get (8.51) and $AC < 1$ by Theorems 8.4 and 8.5.

Theorem 8.11. In Theorem 8.9, we can take $A = \bar{A}(\ell)$, $C = 1$, and $D = \bar{D}(\ell)$, where

$$\bar{D} = \bar{D}(\ell) = \tfrac{1}{2}\left|\pi i - \ell n\, b(\ell)\right|, \tag{8.58}$$

in which

$$b = b(\ell) = \frac{\ell}{8 + \sqrt{64 + \ell^2}} \tag{8.59}$$

if $\ell^2 \leq 32$, and

$$b = b(\ell) = \frac{\ell^2}{k^2\sqrt{16 + \ell^2} + \sqrt{k^4(16 + \ell^2) - \ell^4}} \tag{8.60}$$

$$= \frac{\ell^2}{k^2(2k + \sqrt{16 + \ell^2})}$$

otherwise. Also $L(M)$ has to be written as

$$L = L(M) = 4(\bar{D} + (M-1)K)/(1 - \bar{A}^2) .$$

Proof. By Theorem 8.2 we conclude that (8.51) is satisfied. To find the minimum of $|w(v(u))|$ for $|u| = \bar{A}$, we parallel the proof of Theorem 8.3. With $a = \bar{A}$, we have $a = 4 + (64/\ell^2)$. So, if we wish the ellipses (8.6) and (8.18) to have the same top point, we will take

$$\beta = 4 + \frac{256}{\ell^2} \tag{8.61}$$

For α as given just above, the equation for the y-coordinates of the points of intersection of the two ellipses reduces to

$$4(\beta-\alpha)y^2 - \frac{16\alpha(\beta-4)}{\ell}\, y + \frac{64\beta(\beta-4)}{\ell^2} = 0 .$$

The largest value of β for which this has real solutions is

$$\beta = \frac{4(16 + \ell^2)k^4}{\ell^4} . \tag{8.62}$$

The y-coordinate of the top point of (8.8) is $16/\ell$. The other y-coordinate of intersection will be

$$\frac{16(64 + \ell^2)}{3\ell^3} .$$

We get the configuration of Figure 8.1 if and only if

$$\frac{16}{\ell} > \frac{16(64 + \ell^2)}{3\ell^3} \quad .$$

This holds if and only if $\ell^2 > 32$. So $|w(v(u))| \geq b(\ell)$ for

$$\bar{A} \leq |u| \leq \frac{1}{\bar{A}} \quad .$$

So, as $b(\ell) \leq |w(v(u))| \leq 1$, we can take $d = \frac{1}{2}(\pi i - \ell n\, b(\ell))$ and conclude that (8.50) holds with $D = \bar{D}(\ell)$.

Theorem 8.12. In Theorem 8.6 and the results (8.36) and (8.37) we can replace A, $\bar{B}$, D, and γ by $\bar{A}$, 1, $\bar{D}$, and $\bar{\gamma}$ respectively, where

$$\bar{\gamma} = \frac{2\bar{A}^2}{1 - \bar{A}^2} \quad .$$

Proof. As we have just seen, (8.50) holds if we replace A and D by $\bar{A}$ and $\bar{D}$. So we can parallel the proof of Theorem 8.6, using Theorem 8.2 instead of Theorem 8.4 and 8.5.

9. AFTERTHOUGHT. Let us write $\tilde{\beta}^{(\lambda)}$ for the (infinite) vector with components $\beta_m^{(\lambda)}$, and $\tilde{\gamma}$ for the (infinite) matrix with components $\gamma_{n,m}$. Then (4.8) can be written as

$$\tilde{\beta}^{(\lambda+1)} = \tilde{\gamma}\tilde{\beta}^{(\lambda)} \quad .$$

So (6.1) can be written as

$$\tilde{\alpha} = \sum_{\lambda=0}^{\infty} \{\tilde{\gamma}^{\lambda}\tilde{\beta}^{(0)}\} \quad , \tag{9.1}$$

where we write $\tilde{\alpha}$ for the (infinite) vector with components α_m. The proof that we gave that the right side of (9.1) converges does not depend on what the starting vector $\tilde{\beta}^{(0)}$ is. Hence

$$(1 - \tilde{\gamma})^{-1} = \sum_{\lambda=0}^{\infty} \tilde{\gamma}^{\lambda}$$

exists, and we may write (9.1) as

$$\tilde{\alpha} = (1 - \tilde{\gamma})^{-1}\tilde{\alpha}^{(0)} \quad , \tag{9.2}$$

since $\tilde{\alpha}^{(0)} = \tilde{\beta}^{(0)}$.

If we write $\tilde{\alpha}^{(\lambda)}$ for the (infinite) vector with components $\alpha_m^{(\lambda)}$, then by (4.13) and (4.7), we may write (4.15) as

$$\tilde{\alpha}^{(\lambda+1)} = \tilde{\alpha}^{(0)} + \tilde{\gamma}\tilde{\alpha}^{(\lambda)} \quad .$$

Letting $\lambda \to \infty$ gives

$$\tilde{\alpha} = \tilde{\alpha}^{(0)} + \tilde{\gamma}\tilde{\alpha} \quad .$$

This agrees with (9.2).

If it were not for the fact that we can accelerate the convergence of the right side of (9.1) by means of the ε-algorithm, it would probably be quicker to compute an approximation for $\tilde{\alpha}$ by means of (9.2).

REFERENCES

[1] F. S. Acton and J. Barkley Rosser, An iterative algorithm for calculating potentials near small groups of finite charged plates, Proceedings of the Seventeenth Conference of Army Mathematicians, 1971, pp. 111-125.

[2] A. E. H. Love, Some electrostatic distributions in two dimensions, Proc. London Math. Soc., Ser. 2, XXII (1923), pp. 337-369.

[3] B. Noble, The numerical solution of singular integral equations, MRC Technical Summary Report #730, January 1966, University of Wisconsin.

[4] Bede Liu, Digital filters and the Fast Fourier Transform, Dowden, Hutchinson & Ross, Inc., 1975, distributed by Halsted Press, a division of John Wiley & Sons, Inc.

[5] W. M. Gentleman and G. Sande, Fast Fourier Transforms - for fun and profit, 1966 Fall Joint Computer Conference, AFIPS Proceedings, Vol. 29, pp. 563-578, Spartan, Washington, D. C., 1966.

[6] G. D. Bergland, A radix-eight fast Fourier transform subroutine for real-valued series, IEEE Transactions on Audio and Electro-acoustics, Vol. AU-17, June 1969, pp. 138-144.

[7] W. B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, SIAM Review, vol. 14 (1972), pp. 1-62.

[8] C. Brezinski, Accélération de la convergence en analyse numérique, Lecture notes in mathematics, 584, Springer-Verlag, 1977.

[9] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, Math. Tables and Aids to Computation, vol. 10 (1956), pp. 91-96.

[10] P. Wynn, An arsenal of ALGOL procedures for the evaluation of continued fractions and for effecting the epsilon algorithm, MRC Technical Summary Report #537, January 1965, University of Wisconsin. Also published in the Revue Francaise de Traitement de l'Information (Chiffres), vol. 9 (1966), pp. 327-362.

[11] IBM Application Program, GH20-0205-4, Program Number 360A-CM-03X, Programmer's Manual.

[12] IBM Application Program, H20-0586-0, Program Number 360A-CM-07X, Program Description and Operations Manual.

[13] P. Wynn, A note on programming repeated application of the ε-algorithm, MRC Technical Summary Report #527, November 1964, University of Wisconsin. Also published in the Revue Francaise de Traitement de l'Information (Chiffres), vol. 8 (1965), pp. 23-62.

[14] P. Wynn, Transformations to accelerate the convergence of Fourier series, MRC Technical Summary Report #673, July 1966, University of Wisconsin. Also published in the Gertrude Blanch Anniversary Volume, Wright-Patterson Air Force Base Publication, dated 1966.

[15] J. Barkley Rosser, Potentials of charged plates, Part III, MRC Technical Summary Report #1352, December 1973, University of Wisconsin.

[16] G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, Springer, Berlin, 1925.

SENSITIVITY COEFFICIENT OF EXTERIOR BALLISTICS WITH VELOCITY SQUARE DAMPING

C. N. Shen
U. S. Army Armament Research and Development Command
Benet Weapons Laboratory, LCWSL
Watervliet Arsenal, Watervliet, NY 12189

ABSTRACT. The principal equation of exterior ballistics has a drag term which, in this case, is proportional to the square of the velocity in the tangential direction of the trajectile. The sensitivity coefficient is expressed as the ratio of the initial elevation angle deviation to the initial percentage velocity deviation. The work in this paper is to find analytically the sensitivity coefficient of the exterior ballistics with velocity square damping which comes from the nonlinear air resistance for a projectile. This principal equation is integrated analytically in obtaining the solution for tangential velocity in terms of the elevation angle, together with all the necessary initial conditions. The horizontal range and the vertical range are also expressed as integrals of certain function of the elevation angles. In order to obtain the sensitivity coefficient it is necessary to find the perturbations of the horizontal and vertical ranges. This procedure is similar to that of evaluating differentiation under the integral sign. The perturbation of the ranges is the sum of the perturbations due to the initial velocity, the initial elevation angle and the impact elevation angle. By setting to zeroes the range perturbations we can group the coefficients of the perturbations into two separate equations. The ratio of the perturbations for initial elevation angle to that for initial velocity is the sensitivity coefficient for exterior ballistics that we are seeking.

I. INTRODUCTION. The design of a gun involves numerous parameters. These parameters should be in such a combination that it gives the best first round accuracy. The shell while it leaves the gun has perturbations for the muzzle elevation angle and the muzzle velocity. The ratio of the two is the sensitivity coefficient of the interior and the exterior ballistics. It is desired to compensate the errors due to uncertain changes of muzzle velocity, by the automatic response of the muzzle elevation angle within the gun system. With a correct design this can be made by matching the exterior ballistics to the interior ballistics through the analysis of gun dynamics. This is what is called passive control, since there is no external measurement involved, nor instrumentation needed for control. This general problem can be formulated by first investigating the sensitivity coefficients for exterior ballistics with velocity square damping.

II. DYNAMICAL EQUATIONS FOR TRAJECTORIES. For a constant mass travelling in a vertical plane with no lift and applied thrust, but having drag and velocity vectors contained in the plane of symmetry as shown in Figure 1, the dynamical equations of motion are [1]:

$$\frac{dx}{dt} - v\cos\theta = 0 \tag{1}$$

$$\frac{dy}{dt} - v\sin\theta = 0 \tag{2}$$

$$m(g\cos\theta + v\frac{d\theta}{dt}) = 0 \tag{3}$$

$$\frac{d^2x}{dt^2} = -\frac{D\cos\theta}{m} \tag{4}$$

Where m = the mass of the projectile
g = the acceleration due to gravity
D = the drag of the projectile
v = the velocity of the projectile
θ = the path inclination (elevation angle)
x = the horizontal distance of the projectile
y = the altitude or vertical distance of the projectile

It is noticed that deviations due to anomalies in the azimuth direction is not considered here.

By differentiating Equation (1) with respect to t one obtains

$$\frac{d^2x}{dt^2} = \frac{d}{dt}(v\cos\theta), \tag{5}$$

Substituting into Eq. (4) we have

$$\frac{d}{dt}(v\cos\theta) = -\frac{D\cos\theta}{m}. \tag{6}$$

Solving or $d\theta/dt$ in Eq. (3) one obtains

$$\frac{d\theta}{dt} = \frac{-g\cos\theta}{v} \tag{7}$$

Equation (7) indicates that the differential equations can be transformed from the time domain in t to the angle domain in θ. Equations (6), (1) and (2) are divided by Equation (7) in achieving this transformation as

$$\frac{d(v\cos\theta)}{d\theta} = \frac{Dv}{mg} \tag{8}$$

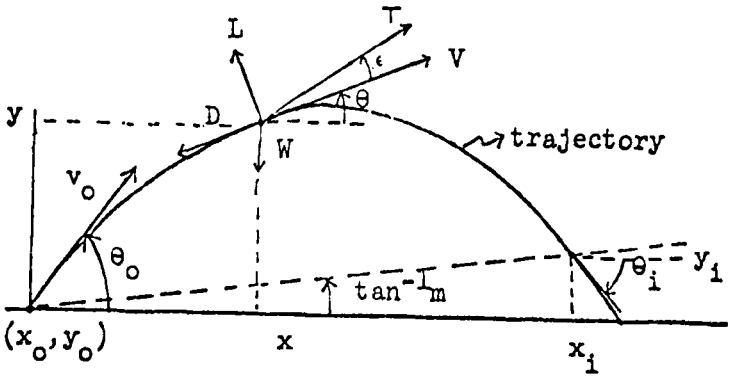


Fig. 1; Forces, Slopes, and Initial and Final Parameters for a Trajectory

$$\frac{dx}{d\theta} = -\frac{v^2}{g} \tag{9}$$

$$\frac{dy}{d\theta} = -\frac{v^2}{g}\tan\theta \tag{10}$$

Equation (8) is called the principal equation of exterior ballistics [2]. It can be integrated if the drag D is a known function of velocity v.

III. TRANSFORMATION OF VARIABLES UNDER HEAD WIND DRAG. The head wing drag D is a velocity square damping term given as

$$D = mcv^2 \tag{11}$$

where

$$c = c_w\left(\frac{\pi}{4}d^2\right)(\rho/2) \tag{12}$$

c_w = the dimensionless resistant coefficient
d = the diameter of the projectile
and ρ = the air density.

Thus the principal equation of exterior ballistics (Equation (8)) becomes:

$$\frac{d}{d\theta}(v\cos\theta) = \frac{cv^3}{g} \tag{13}$$

A further transformation of the dependent variable is necessary by letting

$$u = v\cos\theta \tag{14}$$

where u is the horizontal component of the projectile velocity. Then the dynamical Equations (13), (9) and (10) become

$$\frac{du}{d\theta} = \frac{c}{g}u^3\sec^3\theta \tag{15}$$

$$\frac{dx}{d\theta} = -\frac{u^2}{g}\sec^2\theta \tag{16}$$

$$\frac{dy}{d\theta} = -\frac{u^2}{g}\sec^2\theta\tan\theta \tag{17}$$

To simplify further the form of the dynamical equations another transformation of the independent variable is made by letting

$$q = \tan\theta \tag{18}$$

where q is the projectile slope.

Thus

$$\frac{dq}{d\theta} = \sec^2\theta = 1+q^2 \tag{19}$$

Equations (15), (16) and (17) are divided by Eq. (19) to give:

$$du = \frac{c}{g} u^3 (1+q^2)^{1/2} dq \tag{20}$$

$$dx = -\frac{u^2}{g} dq \tag{21}$$

$$dy = -\frac{u^2}{g} qdq \tag{22}$$

Solution for u in Equation (20) can be readily integrated in closed form. Solutions for x and y can be expressed in the form of integrals once u^2 is obtained.

IV. SOLUTION FOR HORIZONTAL COMPONENT OF PROJECTILE VELOCITY AS FUNCTION OF TRAJECTORY SLOPE. The solution for horizontal component of velocity u can be obtained by integrating Equation (20)

$$-\frac{1}{2}(u^{-2}-u_o^{-2}) = \frac{1}{2}\frac{c}{g}\{q(1+q^2)^{1/2} + \ell n[q+(1+q^2)^{1/2}]\}_{q_o}^{q} \tag{23}$$

where q_o equals projectile slope initially at launch and

$$u_o^2 = v_o^2 \sec^{-2}\theta_o = v_o^2(1+q_o^2)^{-1} \tag{24}$$

by virtue of Equations (14) and (19).

Equation (23) can be written as

$$\frac{1}{u^2} = \frac{1}{u_o^2}\{1-u_o^2\frac{c}{g}[p(q) - p_o(q_o)\} \tag{25}$$

where $$p(q) = q(1+q^2)^{1/2} + \ell n[q + (1+q^2)^{1/2}] \tag{26a}$$

and $$p_o(q_o) = q_o(1+q_o^2)^{1/2} + \ell n[q_o + (1+q_o^2)^{1/2}] \tag{26b}$$

Finally, Equation (25) becomes in the form

$$u^2 = \frac{v_o^2}{1+q_o^2}\{1 + \frac{H(q,q_o,v_o^2,c/g)}{1 - H(q,q_o,v_o^2,c/g)}\} \tag{27}$$

where

$$H(q,q_o,v_o^2,c/g) = \frac{c}{g}\frac{v_o^2}{1+q_o^2}[p(q) - p_o(q_o)] \tag{28}$$

It is noted from the above equation that as

$$q \to q_o \qquad H_{q=q_o} = 0 \tag{29}$$

For the case with no air resistance we have as

$$c \to 0 \qquad H_{c=0} = 0 \tag{30}$$

which implies that the horizontal component of projectile velocity u at any time is a constant.

V. SOLUTION FOR NONDIMENSIONAL RANGE. In determining the range x for the trajectory the closed form solution of u^2 in Equation (27) can be substituted into Equation (21) to obtain the solution in integral form as

$$x_i - x_o = -\frac{v_o^2}{g(1+q_o^2)}\left[(q_i-q_o) + \int_{q_o}^{q_i}\frac{H(q,q_o,v_o^2,c/g)}{1-H}\,dq\right] \tag{31}$$

where x_i = range at impact point
x_o = range at initial point
and q_i = projectile slope at impact point.

To non-dimensionize the range, Equation (31) is divided by the factor v_o^2/g as

$$X(x_i,x_o,v_o)/\Lambda(q_i,q_o) = G_x(q_i,q_o,v_o^2,c/g) \tag{32}$$

where the nondimensional range is

$$X(x_i,x_o,v_o) = (x_i-x_o)g/v_o^2\ , \tag{33}$$

the slope function is

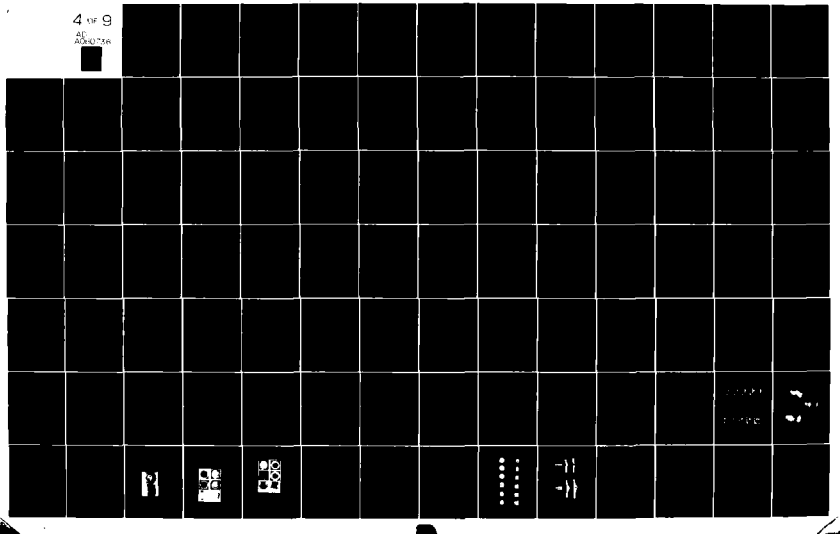
$$\Lambda(q_o,q_i) = \frac{q_o-q_i}{1+q_o^2} \tag{34}$$

and the range drag function due to air resistance is

$$G_x(q_i,q_o,v_o^2,c/g) = 1 - \frac{1}{q_o-q_i}\int_{q_o}^{q_i}\frac{H(q,q_o,v_o^2,c/g)}{1-H}\,dq \tag{35}$$

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL

4 OF 9
AD A080736



It is noted that the left side of Eq. (32) contains no drag coefficient c. It is the term G_x that is a function of H, which in turn is a function of the drag coefficient c. The numerator of the integral in Eq. (35) is H. Since $H_{c=0} = 0$ in Eq. (30), the range drag function at this condition is

$$G_{x(c=0)} = 1 \tag{36}$$

A separate form of Eq. (32) can be written as

$$\frac{g(x_i-x_o)}{v_o^2}\frac{(1+q_o^2)}{(q_o-q_i)} = 1 - \frac{1}{(q_o-q_i)}\int_{q_o}^{q_i}\frac{H}{1-H}\,dq \tag{37}$$

VI. VARIATION OF THE NONDIMENSIONAL RANGE AND THE SLOPE FUNCTION. In order to obtain a first round hit of the target one of the conditions is that the variation of the range should be zero, i.e., from Eq. (33)

$$\delta(x_i-x_o) = 0 \tag{38}$$

We take the perturbation for the nondimensional range from Eq. (33) as

$$\frac{\delta X}{X} = \frac{\delta(x_i-x_o)}{x_i-x_o} - \frac{2\delta v_o}{v_o}$$

$$\frac{\delta X}{X} = 0 - \frac{2\delta v_o}{v_o} \tag{39}$$

The variation of the slope function Λ in Eq. (34) becomes

$$\frac{\delta\Lambda}{\Lambda} = -\frac{\delta q_i}{q_o-q_i} + \frac{\delta q_o}{q_o-q_i} - \frac{2q_o\delta q_o}{1+q_o^2} \tag{40}$$

Next, taking the variation of Eq. (32) and using the expressions given in Eq. (39) and (40) we have

$$\frac{\delta X}{X} - \frac{\delta\Lambda}{\Lambda} = \frac{\delta G_x}{G_x} \tag{41}$$

or

$$-\frac{2\delta v_o}{v_o} + \frac{\delta q_i}{q_o-q_i} - \frac{\delta q_o}{q_o-q_i} + \frac{2q_o\delta q_o}{1+q_o^2} = \frac{\delta G_x}{G_x} \tag{42}$$

It is noticed that in the absence of air damping the range drag function is unity from Eq. (36) and the variation $\delta G_x = 0$. Under this condition the solutions for Eq. (42) was given by the author in the paper entitled, "On the Sensitivity Coefficient of Exterior Ballistics and Its Potential Matching to Interior Ballistics Sensitivity". This paper was presented at the Second U.S. Army Symposium on Gun Dynamics, September 1978. With the velocity square damping the variation of G_x is not zero, i.e.,

$$\delta G_x \neq 0 \tag{43}$$

VII. VARIATION OF THE RANGE DRAG INTEGRAND. The range drag function in Eq. (35) can be written as

$$G_x(q_i,q_o,v_o^2,c/g) = 1 - \frac{1}{q_o-q_i}\int_{q_o}^{q_i} F(q,q_o,v_o^2,c/g)dq \tag{44}$$

where the range integrand is

$$F(q,q_o,v_o^2,c/g) = \frac{H(q,q_o,v_o^2,c/g)}{1 - H} \tag{45}$$

The variation of G_x involves the initial velocity v_o, the initial slope q_o, and the impact slope q_i.

Eq. (44) has the parameters q_o and q_i in the denominator as well as in the integral of F. By chain rule we have,

$$\delta G_x = -\frac{1}{q_o-q_i}\,\delta\left[\int_{q_o}^{q_i} F(q,q_o,v_o^2,c/g)dq\right]$$

$$-(-1)(q_o-q_i)^{-2}\delta(q_o-q_i)\int_{q_o}^{q_i} F(q,q_o,v_o^2,c/g)dq \tag{46}$$

The variation of the integral of F is given in the next section.

VIII. VARIATION OF THE RANGE DRAG INTEGRAL. The parameters in the integral are q_o, q_i and v_o. The variation of the range drag integral follows the rules of differentiation under the integral sign.

$$\delta\left[\int_{q_o}^{q_i} F(q,q_o,v_o^2,c/g)dq\right]$$

$$= \left[\int_{q_o}^{q_i} \frac{\partial F}{\partial v_o} dq\right]\delta v_o + F(q=q_i,q_o,v_o^2,c/g)\delta q_i$$

$$+ \left[\int_{q_o}^{q_i} \frac{\partial F}{\partial q_o} dq\right]\delta q_o - F(q=q_o,q_o,v_o^2,c/g)\delta q_o \qquad (47)$$

The last term of this equation is zero by virtue of Equations (29) and (45). Substituting Eq. (47) into Eq. (46) gives

$$\delta G_x = \left[-\frac{1}{q_o-q_i}\int_{q_o}^{q_i} \frac{\partial F}{\partial v_o} dq\right]\delta v_o$$

$$+ \left[-\frac{1}{q_o-q_i} F(q=q_i) - \frac{1}{(q_o-q_i)^2}\int_{q_o}^{q_i} F\, dq\right]\delta q_i$$

$$+ \left[-\frac{1}{q_o-q_i}\int_{q_o}^{q_i} \frac{\partial F}{\partial q_o} dq + \frac{1}{(q_o-q_i)^2}\int_{q_o}^{q_i} F\, dq\right]\delta q_o \qquad (48)$$

IX. EVALUATION OF THE PARTIAL DERIVATIVES OF THE RANGE DRAG INTEGRAND. The partial derivatives of F with respect to v_o can be found by using Eqs. (45) and (28).

$$\frac{\partial F}{\partial v_o} = \frac{(1-H)\frac{\partial H}{\partial v_o} - H\left(-\frac{\partial H}{\partial v_o}\right)}{(1-H)^2} = \frac{\frac{\partial H}{\partial v_o}}{(1-H)^2} \qquad (49)$$

where

$$\frac{\partial H_o}{\partial v_o} = \frac{c}{g}\frac{2v_o}{1+q_o^2}\left[p(q)-p(q_o)\right] = \frac{2}{v_o} H \qquad (50a)$$

Combining the above we have

$$\frac{\partial F}{\partial v_o} = \frac{2}{v_o}\frac{H}{(1-H)^2} \qquad (50b)$$

Similarly the partial derivatives of F with respect to q_o is

$$\frac{\partial F}{\partial q_o} = \frac{\frac{\partial H}{\partial q_o}}{(1-H)^2} \tag{51}$$

where

$$\frac{\partial H}{\partial q_o} = \frac{c}{g} v_o^2 (-1)(1+q_o^2)^{-2} 2q_o [p(q)-p_o(q_o)] + \frac{c}{g} \frac{v_o^2}{1+q_o^2} (-1) \frac{dp_o}{dq_o}$$

$$= -\frac{2q_o}{1+q_o^2} H - \frac{c}{g} \frac{v_o^2}{1+q_o^2} \frac{dp_o}{dq_o} \tag{52}$$

Substituting Eq. (52) into Eq. (51) one obtains

$$\frac{\partial F}{\partial q_o} = -\frac{2q_o}{1+q_o^2} \frac{H}{(1-H)^2} - \frac{c}{g} \frac{v_o^2}{(1+q_o^2)} \frac{dp_o}{dq_o} \frac{1}{(1-H)^2} \tag{53}$$

X. VARIATION OF THE RANGE DRAG FUNCTION. By substituting Eqs. (45), (50) and (53) into Eq. (48), we have

$$\delta G_x = \left[-\frac{1}{q_o-q_i} \int_{q_o}^{q_i} \frac{H}{(1-H)^2} dq\right] \frac{2\delta v_o}{v_o}$$

$$+ \left[-\frac{H_{q=q_i}}{1-H_{q=q_i}} - \frac{1}{q_o-q_i} \int_{q_o}^{q_i} \frac{H}{1-H} dq\right] \frac{\delta q_i}{q_o-q_i}$$

$$+ \left[\frac{2q_o}{1+q_o^2} \int_{q_o}^{q_i} \frac{H}{(1-H)^2} dq + \frac{1}{q_o-q_i} \int_{q_o}^{q_i} \frac{H}{1-H} dq\right.$$

$$\left. + \frac{c}{g} \frac{v_o^2}{1+q_o^2} \frac{dp_o}{dq_o} \int_{q_o}^{q_i} \frac{dq}{(1-H)^2}\right] \frac{\delta q_o}{q_o-q_i} \tag{54}$$

It is noted that the difference of the end slope is not zero, i.e., $q_o-q_i \neq 0$. Therefore, the problem does not become singular. We have expressed the variation δG_x in terms of the variational parameter δv_o,

δq_i, and δq_o. However, the variational parameter δq_i at the impact point is not known explicitly and must be eliminated by using another variation in the direction of the elevation y.

XI. VARIATIONAL EQUATION FOR THE RANGE. By substituting Eq. (54) into Eq. (42) and grouping the coefficients for the variational terms, we have

$$I_v \frac{2\delta v_o}{v_o} + I_{q_i} \frac{\delta q_i}{q_o - q_i} + I_{q_o} \frac{\delta q_o}{q_o - q_i} = 0 \tag{55}$$

where

$$I_v = 1 - \frac{1}{G_x} \frac{1}{q_o - q_i} I_{12}(q_o, q_i) \tag{56}$$

$$I_{q_i} = -1 - \frac{1}{G_x} \frac{1}{q_o - q_i} I_{11}(q_o, q_i) - \frac{1}{G_x} \frac{H_i}{1 - H_i} \tag{57}$$

and

$$I_{q_o} = 1 + \frac{1}{G_x} \frac{1}{q_o - q_i} I_{11}(q_o, q_i) + \frac{1}{G_x} \left(\frac{2q_o}{1+q_o^2}\right) I_{12}(q_o, q_i)$$

$$+ \frac{1}{G_x} \frac{c}{g} \frac{v_o^2}{1+q^2} \frac{dp_o}{dq_o} I_{02}(q_o, q_i) - \frac{2q_o}{1+q_o^2} (q_o - q_i) \tag{58}$$

In turn, the integral I_{11}, I_{12} and I_{02}, and other terms are given as follows.

$$I_{11}(q_o, q_i) = \int_{q_o}^{q_i} \frac{H(q, q_o, v_o^2, c/g)}{1-H} \, dq \, , \tag{59}$$

$$I_{12}(q_o, q_i) = \int_{q_o}^{q_i} \frac{H}{(1-H)^2} \, dq \, , \tag{60}$$

$$I_{02}(q_o, q_i) = \int_{q_o}^{q_i} \frac{1}{(1-H)^2} \, dq \, , \tag{61}$$

$$H_i = H(q = q_i) \tag{62}$$

and

$$G_x = 1 - \frac{1}{q_o - q_i} I_{11} \tag{63}$$

It is noted that δq_i in Eq. (55) has to be eliminated in solving the sensitivity problem. Similar variation equation may be obtained by considering the variation of the elevation.

XII. THE SOLUTION FOR ELEVATION. The differential equation for elevation was given in Eq. (22) and the solution for u is in Eq. (27).

Substituting Eq. (27) into Eq. (22) gives

$$dy = - \frac{v_o^2}{g(1+q_o^2)} \left[1 + \frac{H(q,q_o,v_o^2,c/g)}{1 - H}\right] q dq \tag{64}$$

Integrating the above one obtains

$$y_i - y_o = - \frac{v_o^2}{g(1+q_o^2)} \left[\frac{q_i^2 - q_o^2}{2} + \int_{q_o}^{q_i} \frac{qH}{1-H} dq\right] \tag{65}$$

Rearranging yields the relationship between the range Y, the end slope function Λ, and the elevation drag function G_y.

$$Y(y_i,y_o,v_o)/\Lambda(q_i,q_o) = \frac{1}{2}(q_o+q_i) + G_y(q_i,q_o,v_o^2,c/g) \tag{66}$$

where the nondimensional elevation is

$$Y(y_i,y_o,v_o) = \frac{g(y_i-y_o)}{v_o^2} \tag{67}$$

Λ is given in Eq. (20) and the elevation drag function is

$$G_y(q_i,q_o,v_o^2,c/g) = - \frac{1}{q_o-q_i} \int_{q_o}^{q_i} qF(q,q_o,v_o^2,c/g) dq \tag{68}$$

A separate form of Eq. (66) may be written as

$$\left[\frac{g(y_i-y_o)}{v_o^2}\right] / \left[\frac{q_o-q_i}{1+q_o^2}\right] = \frac{1}{2}(q_o+q_i) + \frac{1}{q_o-q_i} \int_{q_o}^{q_i} qF \, dq \tag{69}$$

It is noted that left side of Eq. (66) contains no drag coefficient c. It is the term G_y which is a function of drag coefficient c.

For $$c = 0 \qquad G_y = 0 \tag{70}$$

XIII. TERRAIN SLOPE FROM LAUNCH POINT TO TARGET POINT. If Eq. (69) is divided by Eq. (37) with the aid of Eqs. (35) and (68), one obtains

$$\frac{y_i-y_o}{x_i-x_o} \triangleq m = \frac{(1/2)(q_o+q_i)+G_y}{G_x} \tag{71}$$

where m is the terrain slope from launch point to target point, a constant parameter. Therefore,

$$(1/2)(q_o+q_i) + G_y = mG_x \tag{72}$$

It is noted that for m = 0,

$$q_i = -q_o - 2G_y \tag{73}$$

From Equations (36), (70) and (72) we have for c = 0,

$$q_i + q_o = m \ . \tag{74}$$

We use Equation (71) to find the variational equation for the elevation. Taking the variation of Eq. (72) for any given m, we have

$$(1/2)(\delta q_o+\delta q_i) + \delta G_y - mG_x\frac{\delta G_x}{G_x} = 0 \tag{75}$$

where $\delta G_x/G_x$ is given in Equation (42) and

$$\delta G_y = -\frac{1}{q_o-q_i}\,\delta\left[\int_{q_o}^{q_i} qF\,dq\right]$$

$$-(-1)(q_o-q_i)^{-2}\delta(q_o-q_i)\int_{q_o}^{q_i} qF\,dq \tag{76}$$

obtained from Eq. (68).

It is noted that for c = 0, both δG_y and δG_x are zero in Eq. (75).

It can also be proved that the result for δG_y is

$$\delta G_y = \left[- \frac{1}{q_o - q_i} \int_{q_o}^{q_i} \frac{qH}{(1-H)^2}\, dq \right] \frac{2\delta v_o}{v_o}$$

$$+ \left[- \frac{q_i H_{q=q_i}}{1 - H_{q=q_i}} - \frac{1}{q_o - q_i} \int_{q_o}^{q_i} \frac{qH}{1-H}\, dq \right] \frac{\delta q_i}{q_o - q_i}$$

$$+ \left[\frac{2q_o}{1+q_o^2} \int_{q_o}^{q_i} \frac{qH}{(1-H)^2}\, dq + \frac{1}{q_o - q_i} \int_{q_o}^{q_i} \frac{H}{1-H}\, dq \right.$$

$$\left. + \frac{c}{g} \frac{v_o^2}{1+q_o^2} \frac{dp_o}{dq_o} \int_{q_o}^{q_i} \frac{q\,dq}{(1-H)^2} \right] \frac{\delta q_o}{q_o - q_i} \qquad (77)$$

XIV. VARIATIONAL EQUATION FOR THE ELEVATION. By substituting Equations (77) and (42) into Equation (75) and grouping the coefficients for the variational terms, we have

$$J_v \frac{2\delta v_o}{v_o} + Jq_i \frac{\delta q_i}{q_o - q_i} + Jq_o \frac{\delta q_o}{q_o - q_i} = 0 \qquad (78)$$

where

$$J_v = - \frac{1}{q_o - q_i} J_{12} + mG_x \qquad (79)$$

$$Jq_i = - \frac{1}{2} (q_o - q_i) - \frac{1}{q_o - q_i} J_{11} - \frac{q_i H_i}{1 - H_i} - mG_x \qquad (80)$$

and

$$Jq_o = \frac{1}{2} (q_o - q_i) + \frac{1}{q_o - q_i} J_{11} + \frac{2q_o}{1+q_o^2} J_{12}$$

$$+ \frac{c}{g} \frac{v_o^2}{1+q_o^2} \frac{dp_o}{dq_o} J_{02} - mG_x \left[-1 + \frac{2q_o(q_o - q_i)}{1+q_o^2} \right] \qquad (81)$$

In turn, the integral J_{11}, J_{12} and J_{02}, and other terms are given as follows.

$$J_{11} = \int_{q_o}^{q_i} qH(q, q_o, v_o^2, c/g)/(1-H)\, dq\ , \qquad (82)$$

$$J_{12} = \int_{q_o}^{q_i} qH/(1-H)^2 \, dq \, , \tag{83}$$

$$J_{02} = \int_{q_o}^{q_i} q/(1-H)^2 \, dq \, , \tag{84}$$

and

$$G_x = 1 - \frac{1}{q_o - q_i} I_{11} . \tag{85}$$

XV. THE SENSITIVITY COEFFICIENT FOR EXTERIOR BALLISTICS. Eliminating $\delta q_i/(q_i - q_o)$ from Equations (55) and (78) we have

$$[(I_v/Iq_i) - (J_v/Jq_i)] \frac{2\delta v_o}{\delta v_o} + [(Iq_o/Iq_i) - (Jq_o/Jq_i)] \frac{\delta q_o}{q_o - q_i} = 0 \tag{86}$$

From which one obtains the sensitivity coefficient through the aid of Equation (19)

$$S = \frac{\delta\theta_o}{\delta v_o/v_o} = \frac{(I_v/Iq_i) - (J_v/Jq_i)}{(Iq_o/Iq_i) - (Jq_o/Jq_i)} \left[\frac{-2(q_o - q_i)}{1+q_o^2}\right] \tag{87}$$

It is noted that Equation (87) requires the evaluation of the integrals I and J, which are given in Equations (56) through (63) and Equations (78) through (85).

XVI. SUMMARY. The following results are concluded in this paper:

1. The principal equation of Exterior Ballistics is derived with the Trajectory Slope as independent variables.

2. The closed form solution for the horizontal component of Trajectory Velocity is determined for the case of Exterior Ballistics with velocity square damping.

3. The nondimensional range is obtained in terms of an end slope function and a range drag function.

4. Variations of the nondimensional range are expressed as variations of launch velocity.

5. Variations of the range drag function are in terms of the variations of the range drag integral.

6. The range drag integral has parameters in the integrand as well as the upper and lower limits. The variations of this integral are found.

7. The partial derivatives of the range drag integrand are evaluated.

8. The variational equation for the range are in terms of elements involving three integrals as coefficients of three variational parameters.

9. The variational parameters are that of launch velocity, the launch elevation angle, and the impact elevation angle.

10. The average of the end slopes is equal to the terrain slope times the range drag function minus the elevation drag function.

11. Variations of the nondimensional elevation are expressed as variations of the end slopes and the variations of the drag function.

12. The variational equations for the elevation are determined similar to that for the range.

13. Eliminating the variations of impact slope, δq_i, from the set of two variational equations gives the ratio of the coefficients of $\delta v_o/v_o$ and $\delta q_o/(q_o-q_i)$.

14. The sensitivity $\delta\theta_o/(\delta v_o/v_o)$ may be obtained by dividing this ratio $\delta q_o/(\delta v_o/v_o)$ by the quantity $(1+q_o^2)$.

Numerical calculations of this problem will be carried out in the future.

REFERENCES

1. Shen, C. N., "On the Sensitivity Coefficient of Exterior Ballistics and Its Potential Matching to Interior Ballistics Sensitivity," Proceedings of the Second US Army Symposium on Gun Dynamics at the Institute on Man and Science, Rensselaerville, NY, September 1978, sponsored by USA ARRADCOM.

2. Rheinmetall Staff, "Rheinmetall Weapons Engineering Handbook," September 1975. Translation from German. Original Second Edition by Rheinmetall GmbH. Dusseldorf FRG.

MATHEMATICAL MODELLING OF SOME ASPECTS OF PLATE PERFORATION

Werner Goldsmith
Department of Mechanical Engineering
University of California, Berkeley

INTRODUCTION

An examination of the effect of projectiles on targets is one of the most important problems in military strategy, and the subject has also recently become technologically significant in such areas as impact riveting, impulsive anchoring of bolts in rigid foundations, protection of industrial equipment from fragments generated by accidents, integrity of space structures in view of possible collisions with meteorites, and a host of other applications. While the topic is receiving a continual review and input as needed to analyze new phenomena introduced by recent scientific development, there has currently evolved a spurt of interest in this field resulting in special meetings and publications devoted to the area. A comprehensive survey article has covered the entire spectrum of the penetration of projectiles into all types of targets,[1]* two sessions of a recent meeting of the Society of Engineering Science brought together experts in this area,[2] and a special issue of the International Journal of Engineering Science was dedicated entirely to penetration mechanics.[3] Ref. (1) presents a balanced quantitative analytical and experimental treatment of the subject, divided into topics encompassing methodology, characteristics of projectiles and targets, and applications to semi-infinite, thick, intermediate and thin targets. Ref. (2) contains brief discourses on a variety of penetration subjects, while Ref. (3) provided a substantially

* Numbers in superiors refer to the references

more detailed description of numerical methods of penetration performance (including the governing parameters), projectile deformation and mass loss determination[4], soil penetration, effects of yawing, hydrodynamic approaches, long-rod striker investigations, and a compendium of published information on penetration described briefly in a qualitative manner[5].

The present contribution will focus on certain phenomenological aspects of the normal penetration and perforation of thin plates and those of intermediate thickness by kinetic energy projectiles that will facilitate improvement of relatively simple models of the process which will predict the history and terminal state of the event with sufficient accuracy to be acceptable without the need of invoking complicated and expensive numerical schemes. In accordance with the definitons cited in Ref. (1), thin plates are those where stress and deformation gradients throughout the thickness can be completely neglected, while intermediate thicknesses are characterized by a noticeable influence of the rear surface on target deformation.

Projectile impact on such plates will produce the deformation patterns of bulging and dishing, resulting from the effects of bending and shear and, at velocities above the ballistic limit, a variety of perforation modes, such as shown in Fig. 1, in addition to the gross target deflection. The particular perforation mechanism found in any particular situation depends upon the material properties of target and striker, including hardness, the nose shape of the projectile, and the impact velocity.

Fracture on the distal side of the target due to compressive stress waves with amplitudes exceeding the ultimate compressive strength of the plate could conceivably be initiated in weak, low-density targets, while radial failure can only occur in materials with pronounced lower tensile than compressive strengths. Spalling represents tensile failure of the target due to

reflection of the initial compressive transient and is frequently found in loading resulting from the impact of projectiles or from contact explosions. Scabbing fractures have a similar appearance, but result from deformation rather than excess stress and are due to local inhomogeneities or anisotropies. Plugging occurs due to shearing failure produced by normally-striking blunt penetrators and is most frequently found in thin or intermediate plates of substantial hardness. Petalling is produced by high axi-symmetric tensile stresses after passage of the initial pulse occurring near the lip of the penetrator. This is produced by bending moments in thin plates most frequently generated by sharp-nosed projectiles traveling at relatively low velocities, and is generally accompanied by bulging or dishing. Fragmentation occurs only in extremely brittle targets struck by projectiles at normal incidence. Ductile hole enlargement is both an analytical concept following initial penetration as well as an observed mechanism occurring alone only under special circumstances; however, a combination of ductile hole enlargement and cratering or plugging appears to be characteristic for the perforation of thick plates of medium or low hardness.

At obliquity, the projectile may embed itself in the target, and either ricochet or perforate while remaining intact or else fracturing into one or more components. This process is illustrated by the phase diagram shown in Fig. 2 derived experimentally for a typical impact situation.[1]

Correlation of experimental results for normal impact on thin and moderately thick plates has occurred by means of empirical relations, by analytical models based either on rigid-body mechanics or hydrodynamic representations, the latter for initial ultra-high speeds, by completely numerical methods,

and by combination of these techniques. Treatment by means of simple models has the longest history and potentially the highest cost/benefit ratio; this topic can be subdivided into approaches either neglecting or involving projectile deformation with single- or multiple-effect forcing functions (the latter acting either simultaneously or consecutively) applied to elastic/brittle, elastic/plastic or elastic/viscoplastic targets. In the case of non-deforming projectiles, a blunt nose shape has generally been observed to produce plugging, whereas a sharp nose generates petals in the target. Contributions to the ballistic limit velocity at normal incidence, v_{50}, due to bulging, dishing, plugging and penetrator deformation obtained experimentally for several projectile-target configurations are presented in Fig. 3.[6]

The sequel will detail some experimental results which provide guidance for the construction of suitable models for several different impact configurations. It will then concentrate on a discussion of the analysis of the normal perforation of very thin plates by spherical- and conical-nosed projectiles at speeds just above the ballistic limit and on theoretical representations of such situations by blunt-nosed strikers at speeds within the usual ordnance range, but substantially above this limit, where plugging is expected to occur. Critiques of current phenomenological descriptions will be included with suggestions for improvements by combination of effects without exertion of excessive computational effort.

EXPERIMENTAL OBSERVATIONS

A number of experimental investigations have been conducted by the author and his associates to ascertain the perforation characteristics of various types of targets. In one such test series, thin, fully-annealed 2024-0 aluminum sheets were subjected to normal impact and perforation by spherical

and cylindro-conical projectiles at velocities attainable with pneumatic laboratory guns[7][8]. The quasi-static tensile yield and ultimate strength of the target material are 12,800 and 33,000 psi, respectively, with an ultimate shear strength of about 19,000 psi. The substance exhibits significant work-hardening, but is relatively strain-rate insensitive. The 0.05 in. thick, 14.5 in. diameter plates were clamped in a rigid frame at the 14 in. diameter. For the majority of the tests, the projectile diameter was 0.5 in.; either a ball bearing with a hardness of R_C67 or a hard-steel (drill rod) cylindrical striker with a cone angle of 60^o at the tip and an overall length of 0.75 in. were utilized, with masses of 4.78×10^{-5} and 6.66×10^{-5} lb-s^2/in, respectively. The ballistic limits for the two configurations were found to be 400 and 150 ft/s, respectively.

Three tests were executed with the two 0.25 in. diameter projectiles; here, the sphere had a mass of 0.60×10^{-5} lb-s^2/in, while that of the 60^o conically-tipped cylinder, with an overall length of 0.625 in. was 1.744×10^{-5} lb-s^2/in. The plate deformation history and projectile position were observed by means of a high-speed framing camera, with initial striker velocities determined independently by means of the signal recorded from the interruption by bullet passage of two sets of lights transmitted through slots in the barrel near the muzzle end and focused onto photosensors. Final velocities were obtained in many cases from the signal of two coils wound around a tube through which the projectile passed after impact, in addition to the photographic data. Strain and displacement gages were also employed to monitor the process, and quasi-static tests were conducted to compare the resultant deformation with that obtained under dynamic conditions.

A summary of the plate perforation runs involving all projectiles is presented in Table 1. Selected photographs of the target behavior during perforation of the two types of 0.5-in. diameter projectiles are presented in Figs. 4 and 5, and Fig. 6 portrays the post-mortem appearance of projectiles and craters. No plastic deformation of the strikers was found in any of the tests. The deformation history of the plate, including petal formation, has also been plotted from camera data in Figs. 7 and 8 for representative initial conditions.

Initially, the force history acting on the rigid projectile was determined by double differentiation of camera data, smoothed by a least mean-square process and constrained by the independently-measured initial velocity. This procedure was subsequently discarded in view of the large inherent errors in such a process; instead, an empirical relation for the force history F(t) was assumed in the form

$$F(t) = m_b \ddot{w}_b = m_b(B_1 e^{-B_2 nt} \sin^2 nt) \qquad (1)$$

where m_b and $\ddot{w}_b$ are the mass and deceleration of the projectile, B_1 and B_2 are empirical constants, and n is a scale factor so that the force vanishes when perforation is complete at time $t = T$. This is defined as the instant beyond which no further increase in the size of the crater occurs which can be ascertained from the photographic data. Evaluation of the projectile trajectory requires a double integration of Eq. (1) which introduces two additional constants; however, two of the four empirical parameters are fixed by the matching conditions for the initial and terminal projectile velocity -- which automatically insures the identity of the impulse and

momentum change for the striker. Thus, only two constants need to be determined from the position data which was accomplished by a computer program employing a least mean square fit requirement [9]. Typical results for the two striker geometries are exhibited in Figs. 9 and 10.

Perforation by both sphere sizes occurs by tearing and separation of a cap whose shape conforms closely to the configuration of the ball. The plate initially deforms in the same manner as for non-perforating impact at lower velocities, but as the plate cannot absorb the larger amounts of energy transferred with sufficient rapidity, the ultimate stress is reached and fracture of this cap, apparently by shear at 45^{o} to the deformed plate surface, takes place. Both the cap diameter and mass of the plug increase with increasing impact velocity, while the change of momentum and reduction of cap thickness decrease, all indicative of less severe overall loading of the target at higher speeds. The tangential strain at the cap tip varies from 8.7 to 7.3 percent over the velocity range tested, corresponding to that at the ultimate tensile or ultimate shear strength under quasistatic conditions. This suggests the existence of a strength criterion for the perforation phenomenon in this case.

Strain gage results show a propagation velocity of about 730 ft/s for the peak of the pulse whether or not the plate is perforated; this is also the value of the plastic hinge velocity for the unperforated samples in the central region, obtained from camera data, while displacement gages farther out yield a value of 390 ft/s at a radius of 5 in. In contrast, the hinge ring generated under conditions of perforation propagates outward with a speed decreasing from a value of 13,000 in/s at the origin (compared to 12,000 in/s for the 1/4-in. diameter sphere fired at a velocity of 659 ft/s) to 11,850 in/s at a distance of 1.73 in. from the center, indicating the

dependence of this speed on loading rate. In Fig. 7, perforation occurs at approximately 95 μs, substantially beyond the value of the maximum force of 1500 lbs. at about 47 μs based on Fig. 9. This peak force compares to a value of 1165 lbs. required to statically perforate such a plate with a 1/2-in. diameter sphere. Impact at velocities just below the ballistic limit furnished identical deformation patterns for both 14-in. diameter clamped and 4 ft. x 4 ft. freely suspended plates, indicating that plastic flow was confined here to a region smaller than a 7-in. radius so that the boundary had no influence on the phenomenon.

The plate behavior under attack by the conical-nosed projectiles is significantly different, exhibiting piercing followed by radial fractures with the formation of petals, ranging in number from four to six. These fractures also occur at 45^{0} to the plate surface, suggesting a shearing type of failure. As shown in Fig. 8, at $t = 75$ μs, piercing commences for this striker geometry when the slope of the plate at the tip of the projectile has attained the magnitude of its half-cone angle. Thus, in contrast to the strength criterion apparently controlling the event for the case of blunt-nosed strikers, a geometrical requirement appears to govern initiation of piercing for sharp-tipped projectiles, with subsequent hole enlargement resulting from the outward push of the diverging portion of the striker.

The momentum drop of the cylindro-conical penetrator is substantially smaller at comparable initial velocities than for a sphere of the same diameter. This results in less severe loading of the target as manifested by a smaller outward spread of the plastic zone and lower value of the peak force, about 630 lb. for the case shown in Fig. 10 where, moreover, the peak force occurs after commencement of perforation, in contrast to the spherical

projectile situation. The propagation speed of the plastic hinge was found to be nearly constant at 463 ft/s, also significantly lower than corresponding values for spherical impact.

A second group of tests was executed with an evacuated 50 caliber powder gun capable of firing 0.375-in. and 0.25-in. diameter steel spheres with the aid of a sabot at velocities up to 8500 ft/s for the smaller projectile[10]. Initial and final velocities of the projectile were measured with two sets of velocity coils and the event was observed photographically by means of a six-frame Kerr cell camera using a focusing shadowgraph back-lighting scheme. In general, an adequate photographic record of the perforation process for a particular set of projectile and target parameters required interpolation of data from several rounds. Radial strain gages were mounted on both impact and distal faces of the targets outside the plastic zone, and representative targets were sectioned to provide a contour of the deformation and crater produced; crater dimensions were ascertained for most of the tests.

Targets consisted of 12-in. square plates of SAE 1020 steel, both large-grained and small-grained, SAE 4130 quenched and tempered steel (armor plate), and 2024-0, 2024-T3 and 2024-T4 aluminum. The thickness and mechanical properties of these substances are presented in Table 2. The plates were frequently tested several times with impact positions sufficiently far apart so as to avoid interference from the effects of a previous run as well as from the plate edges. The samples were clamped at two points of a single edge onto a rigid stand and placed centrally in the path of the projectile, at least for the initial impact on the specimen.

Fig. 11 presents photographs of target sections struck by 1/4-in. diameter steel spheres at velocities of about 2900 ft/s which bulged, but did not

perforate the steel targets. Fig. 12 exhibits the impact and exit sides and sections of steel and aluminum targets perforated by this striker at an initial velocity of about 2800 ft/s. The first figure shows the highly strained cap removed from the thin aluminum target, essentially a continuation of the process occurring at lower impact velocities, with a thinning of the plug near the edges and ring-like petal formation occurring on both faces, a curved cross-section at the impact side and a straight lip on the distal side. Caps were obtained for all thin targets struck at this velocity, but the fragment sphericity decreased with plate ductility. The plug punched out in the thicker plate is both fractured and severely flattened; the target exhibits thickening in both directions with minor ring-like petalling on the impact side and major effects on the exit face. Bending of the plate, if it occurs at all, is confined to the region immediately exterior to the crater; plug formation involves compression and shear, the latter also evident in causing a portion of the cap fracture for the aluminum plate.

The high-velocity data shown in Fig. 13 exhibit some ring-like fragmentation patterns for both targets, distinctly evident in the case of the alloy steel plate. These are indicative of shear failure, but the large rings punched out in 0.25-in. thick coarse-grained SAE steel plates are the result of tensile rather than shear failure. Sections of the mild and alloy steel targets with embedded projectiles fired at speeds of about 2900 ft/s, Fig. 11, show significant petalling, flattening of the projectile, and a bulge on the distal side. However, the alloy steel severely deformed and cracked the striker and also exhibited a smaller bulge on the distal side, by virtue of its greater resistance to perforation, but with more extensive plate bending than found in the case of mild steel. These patterns portray the effect of

the dominant stresses active in each of the situations depicted.

Strain waves propagating outwards from the crater show the presence of an initial symmetric component followed by an antisymmetric pulse. It was found that the maximum symmetric strain at a given position decreases slightly with impact velocity until the ballistic limit is reached, beyond which it increases, whereas the maximum bending strain acts precisely in opposite fashion. Close to the ballistic limit, the rise time of the symmetric strain component was about equal to the perforation time (based on an average projectile velocity), whereas at substantially higher speeds, this rise occurred in about half this interval. The time of occurrence of the peak antisymmetric pulse was found to be independent of initial projectile velocity. Both peaks decreased approximately exponentially with distance from the impact point. The maximum symmetric strain in both thin and thick targets decreased by an average of about 65 percent from the peak to the minimum perforation velocity for the present tests, whereas the corresponding maximum antisymmetric strain doubles over this range. Thus, the radial motion of the target due to the increasing size of the hole becomes more dominant at higher velocity, while the effects of bending diminish. Still, at these speeds, both of these manifestations can be shown to be very small in any energy balance of the process. Although the perforation process occurs primarily by shear, the characteristics cited do indicate that at least some initial portion of the process takes place as a hole enlargement.

Fig. 14 shows the results of tests designed to recover the plugs separated from the target plates. The central plug thickness was found to vary inversely with impact velocity in the regime just above the ballistic limit, levelling at some asymptotic value beyond a certain threshhold. Due to severe

fragmentation, the cap thickness was found to vary across its width with the maximum dimension at the center for the more ductile materials, and either a uniform thichness or larger at the edges for the more brittle materials. A plot of the velocity drop Δv as a function of initial velocity v_o is presented in Fig. 15. All targets exhibit a drop in this velocity difference just beyond the ballistic limit up to a critical value of initial velocity beyond which this quantity increases again.

All physical quantities measured, the terminal projectile velocity, the thickness of the separated cap, strain gage data, the pictorial history of the process and post mortem examination of sections support the observation of the change of the deformation pattern of the plate from dishing to punching corresponding to a change of the dominant mechanism from bending of the plate to compression and shearing. At higher velocities, fragmentation of the plug occurs, but the terminal velocity of the plug and the projectile are nearly the same over the entire range of test velocities. Both tensile and shear failure of the plug were observed in steels subjected to high impact velocity, the dominant pattern apparently depending on heat treatment (or ductility) of the target.

An experimental investigation of the deformation and mass loss of cylindrical projectiles has been executed in conjunction with the modelling of the process utilizing a procedure [4] that employs elastic and perfectly plastic wave propagation concepts developed earlier[11][12]. With increased impact speed, during plate perforation, such cylinders acquire progressively more rounded fronts as well as both shorter undeformed and overall lengths; mass loss was not found below a certain velocity, but increased beyond this threshhold. The terminal shapes for a typical experiment are presented in

Fig. 16; corresponding results involving only slight penetration without perforation against a harder and thicker target exhibited greater flattening of the projectile nose, substantially shorter terminal lengths and rearward curving petals extending outward from the approximate region of the original projectile diameter[4].

CRITICAL PARAMETERS OF THE PROJECTILE/TARGET CONFIGURATION AND DEFORMATION MECHANISM

For the projectiles considered here which do not contain warheads, ballistic performance depends upon initial velocity, orientation relative to the target, shape and appropriate material characteristics. Ideally, the more sharp-nosed the projectile, the higher the initial velocity, the longer the rod, and the more normal the orientation, the greater will be the efficiency of penetration. Impact at sufficient obliquity may not only result in failure to adequately penetrate the target, but may even result in ricochet, in accordance with the data shown in Fig. 2; in addition, the projectile may both pitch and yaw, severely reducing the penetrability of the striker. Efficiency increases with both length and density of the projectile, since this will concentrate the maximum energy on a given target area. However, increased length introduces both instability in spin-stabilized projectiles and the additional possibility of bending mode failure. Reduced risk from such fracturing or by shattering in the contact zone, amounting to defeat by the target, demands both high strength and substantial ductility of the projectile material, representing contradictory requirements that require optimization.

The targets considered here are flat plates that constitute or simulate elements of larger structures; curvature or irregular profiles are not

expected to have a significant influence on the phenomenon except very close to or below the ballistic limit. The thickness of a target is classified in terms of the number of traversals of elastic waves therein normal to the faces relative to one transit in the projectile. A thin plate is defined, somewhat arbitrarily, as one where this ratio is greater than 5; this value is chosen to insure the maintenance of a nearly constant stress level in the plate during contact. Intermediate targets are those with values of this ratio between 1 and 5, where the process is influenced by the presence of the rear surface, but without achieving dynamic equilibrium. Thick targets, on the other hand, exhibit ratios of such traverses less than unity, so that distal surface reflections return no faster than those in the penetrator.

The resistance to penetration by targets increases with increasing density, thickness, acoustic impedance and strength. Frictional effects for sharp-pointed projectile penetration of thin plates have been found to represent less than 3 percent of the total energy [13], although this proportion may loom larger for thicker plates struck at speeds just above the ballistic limit.

The experimental results presented indicate the deformation patterns that must be considered both in the striker and the target under conditions of normal impact at velocities at and above the ballistic limit of thin and moderately thick plates when initial speeds are restricted well below those of the hypervelocity regime. As a first approximation, the striker might be considered as rigid, particularly when impinging upon a much softer material, but the evidence is overwhelming that significant plastic deformation of the projectile occurs in most practical circumstances, and ablation or extrusion

generate increasingly larger mass loss at successively higher velocities.

Target deformation is more complex: For thin, relatively soft targets struck at velocities just above the ballistic limit, there is substantial plastic deformation outside of the central crater region resulting from the separation of a cap with a shape closely conforming to that of a gently curved nose of a hard projectile. This slug is produced by shear with minimal petal formation, whereas petalling is dominant in the case of sharp-nosed strikers. Initial fracture occurs in the case of cap formation when some critical strength value is exceeded as the result of projectile motion, but is manifested as the result of a geometric criterion in the second instance, i.e., when the slope of the plate deformation at the impact point begins to match that of the projectile tip.

The perforation of intermediate targets is substantially more complex, involving initial compression of both striker and target, subsequent shearing of the plate, dishing of the target outside the contact region, and deformation and possible fracturing of the striker. The plastic distortion of the projectile produces a shape that bears some similarity to the form of a high-speed jet entering a thick target[1]. The patterns of deformation described above are represented in Fig. 16, while Fig. 17 presents photographs of the successive deformation of a 22 caliber, 11 grain steel cylinder fired against the same target plate at increasingly higher velocity.* Furthermore, the crucial importance of material characteristics is demonstrated in Fig. 18 where the deformation pattern of both striker and target is drastically damaged by merely increasing the hardness of the former, all other parameters remaining the same.*

*R.F. Recht, personal communication.

Clearly, a suitable analytical model must not only consider appropriate material properties but also must encompass a variety of mechanisms, some of which might be neglected for certain regimes of initial velocity and/or projectile target configurations.

PLATE DEFORMATION DUE TO BULGING AND DISHING

This pattern which assumes maximum importance in the velocity regime just below and just above the ballistic limit, may well represent an even more important target failure process than the formation of a crater, particularly when it spreads substantially outward from the impact point. At a given velocity, this effect is the larger the smaller the thickness of the target. The analysis of the plate deformation pattern such as that shown in Fig. 16b is generally carried out by the application of the theory of plasticity, frequently with the neglect of elastic effects, a constitutive assumption called rigid-plastic. Moreover, in many instances, work-hardening is ignored so that the material is described as rigid/perfectly plastic.

A substantial number of analyses have been carried out for the case of impulsive or blast loading on a uniform plate of thickness h_o and mass density ρ. As shown in Fig. 19 for the case of axisymmetric loading under uniform distributed pressure, p, per unit area of undeflected plate of initial thickness h_o, the equilibrium equations in polar coordinates r and θ with rotational inertia neglected are [14]-[18]

$$(\alpha_\theta N_r)' - \alpha_\theta' N_\theta - \alpha_r \alpha_\theta Q/R_r - \alpha_r \alpha_\theta p \sin\phi + \rho h_o \alpha_\theta \alpha_r \ddot{w} \sin\phi - \rho h_o \alpha_\theta \alpha_r \ddot{u} \cos\phi = 0 \quad (2)$$

$$(\alpha_\theta Q)' + \alpha_r \alpha_\theta \left[\frac{N_r}{R_r} + \frac{N_\theta}{R_\theta}\right] - \alpha_r \alpha_\theta p \cos\phi + \rho h_o \alpha_\theta \alpha_r \ddot{w} \cos\phi + \rho h_o \alpha_\theta \alpha_r \ddot{u} \sin\phi = 0 \quad (3)$$

$$(\alpha_\theta M_r)' - \alpha_\theta' M_\theta - \alpha_r \alpha_\theta Q = 0 \quad (4)$$

where a prime denotes a derivative with respect to r and a dot a derivative with respect to time. Here, M_r, M_θ and N_r, N_θ are the radial and circumferential bending moments and radial and circumferential membrane forces, all per unit length, Q is the shear force per unit length, and u and w are the deflection along r and normal to the plate, both in the undeformed state. Further, ϕ is the slope of the plate midplane in the plane passing through $r = 0$ and normal to the plate surface, and R_r and R_θ are the principal radii of curvature. The latter and quantities α_r and α_θ are defined in terms of radial and circumferential strains ε_r and ε_θ as

$$\alpha_r = 1 + \varepsilon_r\ , \quad \alpha_\theta = r + u = r\,(1 + \varepsilon_\theta);\ \frac{1}{R_r} \equiv \phi'\,(1 + \varepsilon_r);\ \frac{1}{R_\theta} = (\sin\phi)/r \tag{5}$$

Rotational inertia may be included by use of the equations of motion given in Ref. (19). Existence of large strains and deflections dictates application of Eqs. (2) - (5).

For small strains and moderate deflections, Eqs. (5) become

$$\alpha_r \simeq 1; \quad \alpha_\theta = r; \quad \frac{1}{R_r} = \phi'; \quad \frac{1}{R_\theta} = \frac{\sin\phi}{r}; \quad \alpha' = \cos\phi \tag{6}$$

and strains ε_r and ε_θ and curvatures κ_r and κ_θ are given by

$$\varepsilon_r = u' + \tfrac{1}{2} w'^2 \quad \text{or} \quad \dot{\varepsilon}_r = \dot{u}' + w'\,\dot{w}';\ \varepsilon_\theta = \frac{u}{r} \quad \text{or} \quad \dot{\varepsilon}_\theta = \frac{\dot{u}}{r};$$

$$\kappa_r = (1 + u')\,w'' - u''\,w' \quad \text{or} \quad \dot{\kappa}_r = (1 + u')\,\dot{w}'' + \dot{u}'w'' - u''w' - u''\dot{w}';$$

$$\kappa_\theta = \frac{w'}{r} \quad \text{or} \quad \dot{\kappa}_\theta = \frac{\dot{w}'}{r} \tag{7}$$

With the further assumption that $\cos \phi \approx 1$ and $\sin \phi \approx - w'$, and neglecting small quantities in Eqs. (2) - (4), these relations may be written as

$$rN_r' + N_r - N_\theta = rpw' + \rho h_o r\ddot{w}w' + \rho h_o r\ddot{u} \tag{8}$$

$$rM_r'' + 2M_r' - M_\theta' - 4N_\theta w'/h_o = - rp - \rho h_o r\ddot{w} + h_o \rho r\ddot{u}w' \tag{9}$$

upon elimination of shear force

$$Q = \frac{1}{r}\left[r\, M_r\right]' - M_\theta \tag{10}$$

For a solution of the problem, it is necessary to employ a four-dimensional yield surface characterizing the relation between M_r, M_θ, N_r and N_θ. While interaction exists between all four variables, it has been found expedient to assume a separate Tresca yield condition - which stipulates initiation of yielding when the maximum shear stress attains the yield value - for the moments and for the in-plane forces[20], as shown in Fig. 20. This approximation provides an upper bound to the solution for the case of a uniform shell, while a similar set of yield curves reduced in size to 61.8 percent constitutes a lower bound [21]. The size of the hexagon, i.e., its intercepts along the axis is characterized by the fully plastic moment $M_Y = \frac{1}{4}\sigma_Y h_o^2$ and force $N_Y = \sigma_Y h_o$, where σ_Y is the yield stress in simple tension If $\sigma_Y = \sigma_o$ (or $M_Y = M_o$) is a constant, the material is perfectly plastic; if σ_Y increases with the amount of plastic work performed, the material work-hardens.

Solutions for rigid-perfectly plastic materials have been obtained for a number of cases involving both the concept of travelling hinges and expressions for the deflection separable in time and the radial coordinate. In the first case, radial motion was neglected; it was further assumed that bending effects predominate during the hinge motion, while membrance action governed beyond this phase until the plate motion had completely ceased. The analysis of the response of a rigid-viscoplastic strain-hardening annular plate loaded impulsively by a linear initial velocity profile indicated that strain hardening is important, rate effects play an even larger role, and the influence of membrane forces is dominant in reducing permanent deflections over a wide range of loading parameters and up to deflections of twice the plate thickness.

Several approaches have also been developed for the delineation of plastic deformation of plates under impact loading by cylindrical projectiles, of radius R_b, although none thus far have incorporated the basic equations of motion for the target involving large deflections, as given by Eqs. (2) - (4). A simplification similar to that employed for Eqs. (8) and (9), but involving retention of the N_r term and assuming no motion in the radial direction yields

$$\frac{1}{r}\frac{d}{dr}\left[rQ + rN_r\frac{\partial w}{\partial r}\right] = \frac{1}{r}\frac{d}{dr}\left[rM_r' + M_r - M_\theta + rN_r\frac{\partial w}{\partial r}\right] = \rho h_o\frac{\partial^2 w}{\partial t^2} \qquad (11)$$

that incorporates both bending and membrane forces, with the maximum values of M_r and M_θ given by $\frac{1}{4}\sigma_Y h_o^2$ and that for N_r given by $h_o\sigma_Y$. In the general case when both effects must be considered, the admissible velocity (or deformation) field for each stress component must be determined for the various segments of the Tresca regime and combined to obtain the overall deflection.

The waves generated in the target by the impact of a non-perforating rigid projectile of radius R_b will divide the plate into five zones when analyzed in terms of Eq. (11), as indicated in Fig. 21: (1) An outermost zone $r>r_C = R_b + c_D t$ is beyond the front of the elastic compressional waves travelling with dilatational velocity $c_D = \sqrt{(\lambda + 2G)/\rho}$ (with λ and G as the Lamé constants) which is completely stress free, (2) the range $r_S = R_b + c_S t < r < r_C$ consisting of elastic compression without transverse deflection, with shear wave velocity $c_S = (G/\rho)^{\frac{1}{2}}$, (3) the annulus $r_B = R_b + c_w t < r < r_S$, where only elastic bending occurs, with c_w as a plastic wave velocity $(\sigma_Y/\rho)^{\frac{1}{2}}$, (4) the domain $R_b < r < r_B$ where plastic deformation takes place, and (5) the region $r<R_b$ which travels with bullet velocity. Elastic deformations are neglected relative to permanent deflections; thus, only zone (4) requires further analysis for the determination of plate response.

Experiments performed on rigid 1/2-in. diameter projectiles with masses from 15 - 100 g fired against several types of aluminum alloy plates up to 3/16 in. thick at initial velocities v_o from 83 - 335 ft/s indicated values of $\frac{\partial w}{\partial r}$ between 0.1 and 0.2. On this basis, it is considered that the effects of membrane action due to N_r overshadow those of bending, which is neglected in Eq. (11). Thus, the motion of the plastic zone (representing yielding throughout the entire plate thickness h_o is given by the wave equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\,\frac{\partial w}{\partial r}\right) = \frac{1}{c_w^2}\,\frac{\partial^2 w}{\partial t^2}\;; \qquad c_w = \sqrt{\sigma_Y/\rho} \tag{12}$$

The initial and boundary conditions, including the joint motion of the striker and target in the area of contact, $r \leq R_b$ are given by

$$w(r, t) = 0; \quad \frac{\partial w}{\partial r}(r, t) = 0 \qquad \text{for } r > r_B \tag{13}$$

$$2\pi R_b \sigma_Y h_o \frac{\partial w}{\partial r} = (m_b + \pi R_b^2 \rho h_o) \frac{\partial^2 w}{\partial t^2} \qquad \text{for } r = R_b \tag{14}$$

$$w(r, 0) = 0, \; r > R_b; \quad \dot{w}(r, 0) = \begin{matrix} 0 & r>R_b \\ v_o, & r \leq R_b \end{matrix} \tag{15}$$

A finite difference solution of the equations presented was in very satisfactory correspondence with data obtained for a suitable preselected regime of physical parameters, but for thicker plates or low-strength materials, divergence of predicted and measured deflections suggested the need for inclusion of both elastic and plastic bending effects for better correlation.

In contrast, deformation due to projectile impact of thin plates in other velocity regimes or under conditions of perforation frequently has been treated by neglect of membrane effects and inclusion only of bending. A model based on point loading will lead to a singularity at this position both for a perfectly-plastic or a work-hardening solid; in consequence, representations of this type hypothesize or specify a central portion of the plate, in contact with the projectile, to be perfectly rigid. It is generally assumed that the region of the target touching a flat striker impinging at normal incidence is considered to be rigid. An analysis of the motion under these conditions conceived by Prof. J. L. Kelly* [23] is based on the concept of both stationary and moving yield hinges in the plate; the motion of the plate is derived from the kinematically admissible fields deduced from several regions of the Tresca yield criterion that is assumed

* Department of Civil Engineering, University of California, Berkeley

to define the plastic behavior of the target material.

For the solution, it is assumed that three yield circles are formed as the result of the impact as shown in Fig. 22: One at the edge of the rigid central region $r = R_b$ in contact with the striker, one at an intermediate position $r = r_1(t)$ in the deformed zone, and one at its edge $r = R(t)$ beyond which the plate acts in a rigid manner. Certain continuity conditions obtain for w, its time and spatial derivatives, and for M_r and M_θ[24] [25]. The quantities w, $\dot{w}$, and M_r are continuous in r and t; w', w'', $\dot{w}'$, $\ddot{w}$ and M_θ are continuous in t and piecewise continuous in r with discontinuities occurring at the hinge circle of radius $r_\gamma(t)$; jumps in these quantities, symbolized by $\Delta\{\ \}$ must satisfy relations

$$\Delta\{\dot{w}'\} + \dot{r}_\gamma(t)\,\Delta\{w''\}=0; \quad \Delta\{\ddot{w}\} + \dot{r}_\gamma(t)\,\Delta\{\dot{w}'\}= 0; \quad \Delta\{\dot{M}_r\}+\dot{r}_\gamma(t)\,\Delta\{M_r'\}= 0 \qquad (16)$$

For a discontinuous loading function, all quantities except w and $\dot{w}$ must admit of time discontinuities.

The first hinge circle at $r = R_b$ is located at A in the Tresca hexagon, Fig. 20a; within, the plate is rigid and moves with projectile velocity $v_b(t)$, with $v_b(0) = v_o$. The second, at $r = r_1$ is located at point B of this diagram, and the outermost hinge circle at $r = R_t(t)$ occurs at point C; in between, plate motion is governed by the segments AB and BC, respectively. Here, the curvature rates $\dot{\kappa}_r$ and $\dot{\kappa}_\theta$, defined by

$$\dot{\kappa}_r \equiv - \dot{w}''(r, t); \qquad \dot{\kappa}_\theta \equiv - \frac{1}{r}\,\dot{w}'(r, t) \qquad (17)$$

are governed by the condition of normality for these quantities relative to the yield surface. Thus, as indicated in Fig. 22, the conditions for the various regimes are

$$r \leq R_b \text{ (rigid):} \quad w(r,t) = \int_0^t v_b(t)dt; \quad \dot{w}(r,t) = v_b(t); \quad \ddot{w}(r,t) = \dot{v}_b(t) \tag{18}$$

$$R_b \leq r < r_1(t), \text{ (Region AB of Fig. 20a)}; \quad \dot{\kappa}_r = -\dot{w}''(r,t) = 0 \tag{19}$$

$$r_1(t) \leq r \leq R(t), \text{(Region BC of Fig. 20a)}; \quad \dot{\kappa}_r + \dot{\kappa}_\theta = -(\dot{w}'' + \frac{1}{r}\dot{w}') = 0 \tag{20}$$

$$r > R(t), \text{ (rigid):} \quad w(r,t) = \dot{w}(r,t) = \ddot{w}(r,t) = 0 \tag{21}$$

During the application of a positive pressure against the plate by a projectile, the hinge circles remain stationary, while $r_1(t)$ and $R(t)$ move outward upon load removal until the plate has attained its terminal deformation.

The displacements in regions AB and BC are determined by integration of Eqs. (19) and (20) and use of the appropriate boundary conditions. This yields for the velocities in

$$\text{Region AB:} \quad \dot{w}(r,t) = v_b \left[\ell n \frac{r_1(t)}{R_t(t)} + \frac{r(t)}{r_1(t)} - 1\right] / \left[\ell n \frac{r_1(t)}{R_t(t)} + \frac{R_b}{r_1(t)} - 1\right] \tag{22}$$

$$\text{Region BC:} \quad \dot{w}(r,t) = v_b \left[\ell n\, r(t) - \ell n\, R_b(t)\right] / \left[\ell n \frac{r_1(t)}{R_t(t)} + \frac{R_b}{r_1(t)} - 1\right] \tag{23}$$

The acceleration fields are obtained by differentiation of Eqs. (22) and (23) with respect to time; for stationary hinges, $\dot{r}_1(t) = \dot{R}_t(t) = 0$. Plate velocities and accelerations are functions of projectile speed and hinge circle location; for moving hinges, the accelerations are also functions of the hinge circle velocities.

The problem is solved by substituting a prescribed force history $F(t)$, or its equivalent pressure, $p = F(t)/\pi R_b^2$, into the equation of transverse motion which yields the shear as

$$Q = -\frac{1}{r}\int_0^r [p - \rho h_o \ddot{w}(r, t)]\, r dr = (-p + \rho h_o \ddot{w})\frac{r}{2} \tag{24}$$

and a second integration involving the moment equation (9) with stretching forces and in-plane motion neglected over both regimes AB and BC provides the expressions determining the hinge locations r_1 and R. Upon load removal, a similar procedure yields the velocity histories for the intermediate and outer hinge when the terminal conditions at the end of the stationary hinge stage are employed as the initial conditions for the moving hinge phenomenon. If the plate loading condition is given in terms of the mass, radius and initial velocity of a projectile, an iterative calculation process is required at each stage of the motion to insure that the pressure exerted by the projectile on the plate produces a velocity in the target equal to that of the rigid projectile whose deceleration is determined by Newton's law.

A somewhat different interpretation of the plastic deformation of a thin plate [4], due to the nonperforating normal impact of a flat projectile is based on rigid-workhardening plate behavior as deduced from a linear stress-plastic strain curve. The deformation pattern shown in Fig. 23 is assumed where a central yield hinge encloses a rigid region of the plate of mass $\rho\pi R_b^2 h_o$ from which a plastic shear wave moves outward with constant velocity $c_s^p = (G^p/\rho)^{\frac{1}{2}}$. From conservation of momentum, the initial velocity of the projectile mass and plate plug $\pi\rho R_b^2 h_o$ is given as

$$v_i = v_o/\left(1 + \pi\rho \frac{R_b^2 h_o}{m_b}\right) \tag{25}$$

The ratio of distances traversed by the wave and plug in unit time is v_i/c_S^P which equals initial shear strain γ_i resulting in material hardening; this zone will not strain further unless the applied stress τ exceeds the level corresponding to $\gamma_i(\tau_{\gamma_i})$. Here, τ_γ is the initially attained yield stress in simple shear, taken conventionally as $\tau_\gamma = \sigma_\gamma/\sqrt{3}$ (which is strictly true only for the von Mises yield criterion where yielding starts when the energy of distortion attains a critical value). The outward motion of the hinge results in successively lower stress levels in the material; but at any time t, the velocity of the entire plastically-deformed annulus $r = R$ behaving as a rigid mass is a constant, $v(r) = \text{constant} = C(t) = v_i$. The decrease in the velocity of this in time dt is due to (a) the communication of an impulse $F\,dt = 2\pi r h_o \tau_{\gamma_i}\,dt$ to the rigid exterior portion of the plate and (b) the growth of plate mass dm acquiring motion. The momentum balance for this system is

$$m_b v_o = (m_b + \pi\rho R_b^2 h_o + m)v + \int_o^t F\,dt \quad \text{with } dt = dr/(c_S^P) \tag{26}$$

where m is the mass of the plate between R_b and R.

Use of Eq. (26) and integration over the region from R_b to r yields for plate velocity v

$$v/v_o = \frac{1 + (\pi h_o \tau_{\gamma_i} R_b^2)\,(1 - \frac{r^2}{R_b^2})/c_S^P v_o m_b}{1 + (\pi r^2 \rho h_o)/m_b} = \frac{1 + (\xi/\beta)(1 - \frac{r^2}{R_b^2})}{1 + \xi(r^2/R_b^2)} \tag{27}$$ where

$$\xi = \pi\rho h_o R_b^2/m_b;\ \beta = \frac{\rho c_S^P v_o}{\tau_{\gamma_i}} \quad \text{and } r = R_b + (c_S^P)t \tag{28}$$

The displacement of the plate at time t when the plastic zone has extended $r = R$ is

$$w = \int_0^t v\,dt = \frac{v_o}{c_S^P} \int_R^r \frac{1 + \frac{\xi}{\beta}\left(1 - \frac{r^2}{R_b^2}\right)}{1 + \xi\,(r^2/R_b^2)}\,dr \quad \text{or} \quad \frac{w}{R_b} = \left[\frac{v_o}{c_S^P} \Big/ \xi\beta\right]\left[\xi\left(\frac{R}{R_b} - \frac{r}{R_b}\right)\right.$$

$$\left. - \sqrt{\xi}\,(1 + \beta + \xi)\,\left(\tan^{-1}\sqrt{\xi}\,\frac{R}{R_b} - \tan^{-1}\sqrt{\xi}\,\frac{r}{R_b}\right)\right] \tag{29}$$

These solutions apply unless the stress at any position r exceeds that experienced during initial work-hardening upon passage of the plastic wave front. The stress is

$$\tau(r) = \frac{(m_b + \rho\pi R_b^2\, h_o + m)}{2\pi h r}\,\frac{dv}{dt} \tag{30}$$

from which

$$\frac{\tau(r)}{\tau_{Y_i}} = \frac{R/R_b}{r/R_b}\left[\frac{1 + \xi\,(r/R_b)^2}{1 + \xi(R/R_b)^2}\right] \tag{31}$$

where τ_{Y_i} is the shear yield stress attained initially. Inspection of Eq. (31) indicates the validity of the assumption of the propagation of a single hinge circle provided the ratio $\pi\rho h R_b^2/m$ exceeds 1/3; otherwise the ratio of Eq. (31) exceeds unity, first attained at the edge of the projectile, and a second yield hinge will propagate outward at that time. Experimental results support the conclusion of the analysis which is considered to apply when membrane stresses can be neglected and surface cratering is absent.

The maximum central deflection of elastic-perfectly plastic or linearly work-hardening plates has also been calculated on the basis of an energy balance, but required an assumed deformation pattern of the target that was based on experimental results.[8] Some of the mathematical difficulties

encountered in an analysis of projectile loading of elasto-plastic plates have been surmounted by inclusion of rate dependence in the constitutive relation,[26]-[28] exemplified for the uniaxial case by

$$\dot{\varepsilon} = \frac{2}{T^*\sqrt{3}} (\sigma - \sigma_o)/\sigma_o \tag{32}$$

where T^* is the viscoplastic relaxation time. The velocity v of the midsurface of a viscoplastic plate of radius R_t, restrained at the edges and subjected to central impact by a projectile under the assumption of small deflections and neglect of membrane forces is governed by [28][29]

$$\nabla^4 v + \frac{3\sqrt{3}}{2h_o M_o T^*} \left(\rho h_o \frac{\partial v}{\partial t} + \frac{m_b}{\pi R_b^2} \left.\frac{\partial v}{\partial t}\right|_{r \le R_b} + p^S\right) = 0, \qquad 0 < r \le R_b \tag{33}$$

$$\nabla^4 v + \frac{3\sqrt{3}}{2h_o M_o T^*} \rho \frac{\partial v}{\partial t} = 0, \qquad R_b \le r \le R_t \tag{34}$$

where $M_o = \frac{h_o^2}{4}\sigma_o$ is the unit perfectly plastic yield moment and p^S is the static pressure distribution corresponding to dynamic loading p, or collapse load, amounting to $4\pi M_o/\sqrt{3}$ for a concentrated load at $r = 0$. Then Eq. (33) may be replaced by the relation

$$\lim_{r \to o} 2\pi r \left(\frac{2h_o M_o T^*}{3\sqrt{3}}\right)\left(\frac{\partial}{\partial r}[\nabla^2 v]\right) = -\frac{4\pi M_o}{\sqrt{3}} - m \left.\frac{\partial v}{\partial t}\right|_{r=0} \tag{35}$$

Eqs. (34) and (35) together with initial conditions due to projectile impact and boundary conditions for a clamped plate given by

$$v(0,\,0) = v_0;\quad v(r,\,0) = 0,\quad r \neq 0;\quad v(R_t,\,t) = 0;\ \frac{\partial v}{\partial r} = (R_t,\,t) = 0 \qquad (36)$$

can be solved approximately in closed form [29]; however, such solutions are not applicable for very thin plates. Here, a solution for the velocity in the Laplace transform space has been obtained as [28]

$$\bar{v}(r,s) = \frac{4}{\pi}\left[\frac{4\pi M_o/m_b\sqrt{3}}{s^{1.5}(b_1+s^{\frac{1}{2}})} - \frac{v_o}{s^{\frac{1}{2}}(b_1+s^{\frac{1}{2}})}\right] \mathrm{kei}\ (\alpha_1 s^{\frac{1}{4}} r) \qquad (37)$$

where $\mathrm{kei}\ x = \mathrm{Re}\left|\frac{1}{2}\pi\, H_1^{(1)}(xe^{\frac{3\pi i}{4}})\right|$ is the Kelvin function, $H_1^{(1)}$ is the Hankel function of the first kind and order one, s is the transform parameter, a bar over a symbol denotes its transform, and

$$\alpha_1^4 = \frac{3\sqrt{3}\rho}{2T^2 M_o}, \quad b_1 = \frac{16\rho h_o}{\alpha^2 m_b} \qquad (38)$$

This relation can be numerically inverted to obtain $v(r, t)$ and the velocity field is integrated to obtain the deflection history.

The solution of Eq. (37) at the origin $r = 0$ is

$$\bar{v}\ (0,\,s) = \frac{v_o}{s^{\frac{1}{2}}(b_1 + s^{\frac{1}{2}})} - \frac{4\pi M_o/m_b\ \sqrt{3}}{s^{3/2}\ (b_1 + s^{\frac{1}{2}})} \qquad (39)$$

whose inversion yields

$$v(0,t) = \left[v_o - \frac{4\pi M_o}{\sqrt{3}m_b\ b_1^{\,2}}\right] e^{b_1^{\,2}t}\ \mathrm{erfc}\ (b_1 t^{\frac{1}{2}}) + \frac{4\pi M_o}{\sqrt{3}m_b b_1^{\,2}}\left[1 - 2b_1\left(\frac{t}{\pi}\right)^{\frac{1}{2}}\right] \qquad (40)$$

Integration of Eq. (40) yields

$$w(0,t) = \left(v_0 - \frac{4\pi M_0}{\sqrt{3}m_b b_1^2}\right)\left[\frac{1}{b_1^2} e^{b_1^2 t} \operatorname{erfc}\left(b_1 t^{\frac{1}{2}}\right) + \frac{2}{b_1}\left(\frac{t}{\pi}\right)^{\frac{1}{2}} - \frac{1}{b_1^2}\right]$$

$$+ \frac{4\pi M_0 t}{\sqrt{3} b_1^2 m_b} - \frac{4\pi M_0 t^{3/2}}{3\sqrt{3\pi} m_b t_1} \tag{41}$$

while differentiation of the same equation provides the central plate acceleration as

$$\ddot{w}(0,t) = \left(v_0 - \frac{4\pi M_0}{\sqrt{3}m_b b_1^2}\right)\left[b_1^2 e^{b_1^2 t} \operatorname{erfc}\left(b_1 t^{\frac{1}{2}}\right) - \frac{b_1}{\sqrt{\pi t}}\right] - \frac{4\pi M_0}{m b_1 \sqrt{3\pi t}} \tag{42}$$

Eq. (42) exhibits an infinite deceleration at the instant of impact, a result inherently due to the choice of the material behavior and the requirement that the velocity at the center change instantaneously from $v = 0$ to v_0. This deficiency can be circumvented by treating the central region of the plate as rigid under projectile impact, employing conditions such as Eq. 25, and considering viscoplastic deformation for the plate outside the contact region. It should also be mentioned that the viscoplastic theory is singular in that it does not reduce precisely to the rigid-perfectly plastic case as $T^* \to 0$, corresponding to the absence of strain-rate effects. Furthermore, in view of a constant collapse load in Eq. (40), the velocity will decrease monotonously with time; however, the result has physical significance only up to the instant t_f when the plate center reaches zero velocity, determined implicity from this equation by setting v equal to zero:

$$\left(v_0 - \frac{4\pi M_0}{\sqrt{3}m_b b_1^2}\right) e^{b_1^2 t_f} \operatorname{erfc}\left(b_1 t_f^{\frac{1}{2}}\right) + \frac{4\pi M_0}{\sqrt{3}m_b b_1^2}\left[1 - 2b_1\left(\frac{t_f}{\pi}\right)^{\frac{1}{2}}\right] = 0 \tag{43}$$

PENETRATION MODELS INVOLVING AXIAL AND RADIAL PLATE MOTION

Pointed projectiles displace a significant amount of target material radially after penetration is initiated at the tip; the further process of separation consists either of (a) the enlargement of a hole in the target during projectile passage, or (b) the development of radial fractures travelling outward from the piercing point to form petals. Both events primarily occurring in relatively thin plates have been modelled in terms of displacements of a rigid/perfectly-plastic target material of tensile yield strength σ_o. A quasistatic analysis of symmetrical hole enlargement provided the work performed to expand a pin hole to one of radius r as [30][31]

$$W = 1.33\pi r^2 h_o \sigma_o \tag{44}$$

However, most penetration phenomena for thin and moderately thick sheets under both quasistatic [32] and ballistic penetration speeds occur in the antisymmetric mode, where the material of the crater is displaced axially and only on the exit side. For this case, the corresponding work performed is much smaller, i.e.

$$W = \frac{1}{2}\pi r^2 h_o \sigma_o \tag{45}$$

Another analysis that included dynamic effects, but specifies that the enlargement process occurs initially at constant velocity v_o and subsequently at constant acceleration, utilized a similarity solution to express hole radius r and the local thickness h as a function of radial particle velocity $\dot{u}$ by[33]

$$\frac{r}{t} = u - \sqrt{2}c_w \tanh\left[\sqrt{2}\,(\dot{u}-v_o)/c_w\right] \;;\quad \frac{h}{h_o} = \frac{\sqrt{2}c_w\, t\, \sinh\left(\sqrt{2}\, v_o/c_w\right)}{\left(v_o/c_w\right)\cosh\left(\sqrt{2}(v_o-\dot{u})/c_w\right)} \tag{46}$$

where c_w is once more $\sqrt{\sigma_Y/\rho}$.

The antisymmetric pattern depicted in Fig. 24 is not found in practice, even for thin plates struck just above the ballistic limit, but has been analyzed on the basis of an energy balance for a rigid/perfectly-plastic material and its predictions have frequently been cited in the literature[34]. If H is designated as the crater height and z* the distance from the crater tip in question, and if the radial stress is neglected, so that an approximate uniaxial state of stress exists, then incompressibility dictates that

$$h/h_o = (r/R_b)^{\frac{1}{2}} = (z^*/H)^{\frac{1}{2}} \quad \text{and} \quad H = 0.75R_b,\;\; h = 1.15h_o\,(z^*/R_b)^{\frac{1}{2}} \tag{47}$$

with the static work of plastic deformation given by Eq. (45) for $r = R_b$. The force acting on the mass $m^* = \pi\rho h_o r_b^{\,2}$ displaced at time t is

$$F = m^* \frac{d^2 r_b}{dt^2} + \frac{dm^*}{dt}\frac{dr_b}{dt} \tag{48}$$

with t_f as the time of final hole size attainment; the dynamic work of perforation is

$$W^D = \pi\rho h_o \int_o^{R_b(t_f)} r_b^{\,2}\, \frac{d^2 r_b}{dt^2}\, dr_b + 2\pi\rho h_o \int_o^{R_b(t_f)} r_b\left[\frac{dr_b}{dt}\right]^2 dr_b \tag{49}$$

Typical sharp-tipped projectiles are the conical and ogival nose, of length L_N, preceding a uniform cylindrical base of radius R_b; these shapes are defined by

$$r_b = (R_b/L_N)z = z \tan \beta \quad \text{for the cone and} \quad r_b = R_b \sin \frac{\pi}{2}\left(\frac{z}{L_N}\right) \text{ for the ogive} \qquad (50)$$

where β is the half-cone angle. The projectile velocity during penetration may be obtained by equating (49) to the change of kinetic energy of the projectile $\frac{1}{2} m \Delta v_b^2 = \frac{1}{2} m \Delta\left(\frac{dz}{dt}\right)^2$ using either of the shape functions given by Eq. (50). Ref. (34) calculated the total work performed for the two cases for the unrealistic assumption of a constant v as

$$W = \pi h_o R_b^2 \left[\rho\left(\frac{v_o R_b}{L_N}\right)^2 + \tfrac{1}{2}\sigma_o\right] \qquad \text{(cone)} \quad \text{and} \qquad (51)$$

$$W = \pi h_o R_b^2 \left[\frac{\rho\pi}{T\sigma}\left(\frac{v_o R_b}{L_N}\right)^2 + \tfrac{1}{2}\sigma_o\right] \qquad \text{(ogive)} \qquad (52)$$

If, on the other hand, the conical projectile is assumed to have a constant deceleration calculated from the work-energy principle as

$$-\ddot{w} = \pi R_b h_o \tan \beta \,(\tfrac{1}{2}\sigma_o + \rho v_o^2 \tan^2\beta)/(m_b + \pi h_o R_b^2 \rho \tan^2\beta) \qquad (53)$$

then the total energy loss is given by $m_b \ddot{w} R_b/\tan\beta$, and the deceleration of the ogival projectile could be similarly expressed, albeit in much more complicated fashion.

A description of the petaling process in the absence of plugging has been provided that neglects static strength effects and invokes a momentum

balance, but requires an a priori stipulation of both the deformation pattern of the plate as well as a radius R_t beyond which the target is not affected by the action of a striker. A schematic of the model is presented in Fig. 25; the principal mechanisms of energy absorption are fracture and plastic deformation in the zone $r \leq R$. Clearly, the credibility of the results depends upon the accuracy of the assumed plate deflection. Fracture directions at the tip have been found to be related to the direction of planar isotropy of the material.[7][8][32] The momentum balance in the direction of motion z of a normally-impinging sharp-tipped projectile is given by[35]

$$m_b \Delta v = m_b(v_o - v) = m_e(z)v \qquad \text{with} \quad m_e = 2\pi\rho h_o \int_o^R r \frac{\partial w}{\partial z}\, dr \tag{54}$$

where m_e is an effective target mass and $w(r_o, z)$ is the axial displacement of a deformed plate element initially at r_o. This relation can be solved for any assumed target deformation. As an example, if it is assumed that the plate conforms to the shape of the nose of a penetrating conically-tipped projectile, as shown in Fig. 26, involving no stretching of the petals and thus ignoring material strength effects, the plate deformation w is given by $w = (z \tan \beta - r) \cos \beta$ and the velocity drop is

$$\Delta v = \frac{v_o \pi \rho h_o}{m_b} [z \tan \beta]^2 \sin \beta \quad , \text{ or}$$

$$(\Delta v)_f = v_o - v_f = \frac{\pi \rho R_b^2 v_o}{m_b} \sin \beta \qquad \text{for} \quad \frac{\Delta v}{v_o} \ll 1 \tag{55}$$

Even this quasi-empirical approach does not provide good correlation with experimental data at impact velocities in the range just above the ballistic limit, indicating the need of incorporating strength parameters into the

momentum balance. This can be effected by addition to the right-hand side of the first of Eqs. (54) of the impulse communicated to the rigid plate section at R_t, namely $2\pi\rho h_o \int_o^R \frac{rdr}{c^P}$ where $c^P = c_w$ for a perfectly-plastic material. The total line load per unit length F^* at the base of the petals is approximated by

$$F^* = (\rho h v_o^2 \tan^2\beta)(2[1-\sin\beta])^{\frac{1}{2}} \tag{56}$$

inclined to the projectile axis at angle $\frac{\pi}{4} - \frac{1}{2}\beta$; this is evaluated on the hypothesis that the contact pressure between projectile and petal vanishes.

Several at least partly empirical equations have been suggested to delineate the force acting on the projectile when target failure initiates at the tip of sharp projectiles, based on the simultaneous action of a variety of physical mechanisms that, however, uniformly neglect projectile deformation.[36][37] One of these includes the contributions of compression (C), distortion (X), friction (F) and inertia (I) and specifies the differential total force dF in terms of the notation exhibited in Fig. 25 as

$$dF = dF_G + dF_X + dF_F + dF_I + [b_1 \pi r_b \sigma_o \tan\phi\sqrt{1+\sec^2\phi} + (2\pi E^* \tan\phi \sec^2\phi \cdot$$

$$\cdot(\cos\phi - r_b\kappa) + 4b_2\rho r_b \tan\phi\,[r_b\tan\phi\,\dot{v} + \tan^2\phi - \frac{r_b\kappa}{\cos^3\phi})]) \cdot$$

$$\cdot\{1 + \frac{\sin\zeta}{\sin\phi\cos(\phi+\zeta)}\}]dz \tag{57}$$

where z denotes projectile position $v = \dot{z}$, $\Phi = \tan^{-1} \frac{dr_b}{dz}$ is the local slope of the projectile nose at station $r_b(w) = r$, with $\kappa = -\cos\Phi \frac{d\Phi}{dz}$ as the corresponding curvature, and E^* is the specific target surface energy. The constants utilized were approximated as $b_1 \approx 1$ and $b_2 \approx 2.5$. Friction angle ζ diminishes rapidly with penetration as the result of temperature rise in the plate; the effect of plate distortion was also considered to be negligible. An even simpler representation has been proposed as

$$F = \pi r^2 \langle z \rangle \left[6\overline{BH}\left(1 - \frac{z}{h_o}\right)^2 - 4\overline{BH}\left(1 - \frac{z}{h_o}\right)^3 + \rho v^2\right] \qquad (58)$$

that accounts only for motion acquired by the target and plastic indentation characterized by Brinell hardness $\overline{BH}$. Friction, tearing and wave effects are neglected, and separation of the target and the surface of the projectile tip could occur.

Another simple prescription for such a force law, involving a conical bullet, that accounts solely for static strength and virtual mass effects, frequently cited in the literature, is given by [39]

$$\frac{dF}{dA} = (P_\pi^S + \rho v^2 \sin^2\beta) \sin\beta \qquad (59)$$

Here P_π^S is the average contact pressure required to perforate the target quasistatically, empirically determined as $5.3h_o$ MPa for soft aluminum, with h_o in mm. This parameter might also be derived from a slip-line solution of the event for a perfectly-plastic or work-hardening solid. The form of Eq. (59) is also valid for blunt-nosed projectiles.

It is easy to criticize this approach as not fundamental, relying on a number of empirical quantities, limiting attention to some effects while neglecting others, and frequently requiring step-wise numerical solutions in any case. However, the representation is usually based on some form of physical model, the empirical parameters can frequently (or at least have the potential to) be related to fundamental physical properties of the system, and predicted results are often in good agreement over limited ranges of the governing variables when the empirical constants are properly chosen. However, it must be emphasized extrapolation outside such limits usually is not acceptable.

PLUGGING OF TARGETS

This perforation phenomenon is defined as the condition when a section of the target involving the zone of contact and possibly also its immediate vicinity are separated from the remainder of the plate due to, at least in part, the shear produced by the penetration of the striker. The most elemental concept of this process involving only conservation of momentum for the identical terminal velocity v_f of a blunt striker of initial length L_o and density ρ_o is given by [40]

$$\frac{v_f}{v_o} = \rho_b L/(\rho_b L + \rho h_o) \tag{60}$$

while a fluid model of the element resisting with force $F = \pi R_b^2 \rho v^2$ yields a value of v_f given by

$$v_f = \exp\left[-\pi R_b^2 \rho h_o/m_b\right] \tag{61}$$

An energy balance including that of separation of the plug of mass $m_q \approx \pi\rho R_b^2 h_o$ from the remaining target, assumed constant for any particular target-striker combination, and that expended in plastic deformation of the components to permit the target and striker attaining a common velocity yields the final velocity for normal impact as [41]

$$v_f = \frac{m_b}{m_b + m_q} (v_o^2 - v_{50}^2)^{\frac{1}{2}} \tag{62}$$

For impact of such blunt cylinders at an angle of obliquity θ, the terminal velocity is expressed as

$$v_f = \frac{(v_o^2 - v_{50}^2)^{\frac{1}{2}} \cos\theta^*}{1 + m_b/m_q} \tag{63}$$

where the change in the direction of travel, θ^* is approximated by the expression

$$\sin 2\theta^* = \frac{\sin 2\theta}{\left(\frac{v}{v_{50}}\right) + \frac{v}{v_{50}} \left(\frac{v^2}{v_{50}^2} - 1\right)^{\frac{1}{2}}} \tag{64}$$

Consideration of only shearing in the plugging process leads to the equation of motion for normal impact of a blunt projectile and associated initial and boundary conditions

$$\rho \ddot{w} = \tau_{,r} + \frac{\tau}{r} \tag{65}$$

$$\dot{w}(r,0) = \begin{cases} 0\ , & r > R_b\ ; \quad v(\infty, t) = 0 \quad \text{for an infinite plate} \\ v_i, & r \leq R_b\ ; \quad v(R_t, t) = 0 \quad \text{for a plate clamped at radius } R_t \end{cases} \tag{66}$$

$$\text{and } (m_b + m_q)\,\dot{v} = 2\pi R_b\, h_o \tau_\gamma \tag{67}$$

where the initial velocity of the combined projectile-target system, v_i, is given by v_f in Eq. (60). Eq. (65) is a modification of Eq. (11) where the plate motion was considered to be changed by membrane forces. The constitutive equations employed have included that for a perfectly-plastic material, so that $\tau_\gamma = \tau_o$ [42], a linear work-hardening solid[43], a Bingham-type material featuring a dynamic viscosity term ν, given by [44] - [46]

$$\tau = \left[\operatorname{sgn} \frac{\partial v}{\partial r}\right] \tau_\gamma + \nu \dot{\gamma} = -\tau_\gamma + \nu \frac{\partial^2 w}{\partial r \partial t} \tag{68}$$

elasto-viscoplastic solids[47], and empirical relations based on experimental results. In general, closed-form solutions have been obtained only when severely restrictive assumptions were invoked, including such hypotheses as constant projectile velocity, neglect of plug mass relative to projectile mass, or neglect of target strength. Results not limited in this fashion have been obtained by numerical techniques.

A relatively simple, yet reasonably comprehensive model for the perforation of both thin sheets and those of intermediate thickness by a deformable projectile divides the process into three consecutive stages [48] - [50], as shown in Fig. 27. The first phase involves compression and indentation of the target, terminating at an empirically determined depth $h_o - h_1$ when plug shear is initiated. The second phase continues the compression process in addition to plug shear and ends when projectile

and plug attain the same velocity. The third stage involves only shear and ceases upon complete plug ejection; all other effects are ignored, including friction, target flexure, and wave propagation. Furthermore, the deformation of the projectile, while permitted, is accounted for only in an empirical manner.

During the first stage, the target material ahead of the projectile is compressed to its ultimate strength σ_{UC} and a section of the target acquires simultaneously some of the momentum transferred by the bullet, leading to the equation of motion

$$F_1(t) = -\frac{1}{2}\rho C_1 A_u v^2 - \sigma_{UC} A_u = \rho A_u v^2 + (m_{bo} + \rho A_u z) v\frac{dv}{dz} \quad \text{for} \quad 0 \leq z \leq h_o - h_1 \qquad (69)$$

Here m_{bo} and A_u are the initial mass and projected area of the striker, h_1 the thickness of the target at the end of the first phase (substantially equal to the plug thickness so that $h_1 = h_q$), C_1 is a constant depending on striker geometry accounting for virtual target mass which takes on the values of $\frac{1}{2}$, 1, and $\cos^2\beta$ for a spherical, cylindrical, and conical-tipped projectile, respectively, and may be ascertained for other nose shapes. However, the flattening of non-armor-piercing projectiles is so severe that a value of unity is probably an appropriate choice for almost any unjacketed shape. D_1 is the projectile diameter for each stage 1, 2, 3.

Integration of Eq. (69) yields the projectile velocity and a second provides the displacement in quadratures as

$$v = \left[\left\{v_o^2 + \frac{\sigma_{UC}}{\rho(1+\frac{1}{2}C_1)}\right\}\left\{\frac{m_{bo}/\rho A_u}{z + m_{bo}/\rho A_u}\right\}^2 + C_1 - \frac{\sigma_{UC}}{\rho(1+\frac{1}{2}C_1)}\right]^{1/2}, \quad 0 \leq z \leq h_o - h_1 \qquad (70)$$

$$t(z) = \int_0^x \frac{dz}{v} \qquad \text{or } t_1 = \int_0^{h_0 - h_1} \frac{dz}{v} \tag{71}$$

Substitution of Eq. (70) in Eq. (69) yields the force-displacement relation, and use of Eq. (71) provides an implicit equation for the force history that can be determined numerically.

During stage 2, the total force consists of compressive, shear and inertial components; the last term acts on plug area A_q, whose diameter may be approximated as that of the original projectile base, or else as varying linearly with z from that value to the measured final plug base diameter, if significantly different. A choice of $C_1 = \frac{1}{2}$, corresponding to a spherical striker shape in this domain appears appropriate, but results in a small discontinuity in force at the end of the first stage. The compressive force is reduced from $\sigma_{UC}A_q$ to zero in this interval, with an assumed parabolic variation. The shear force at the plug periphery is considered to be of the viscoplastic type represented by Eq. (63) with the shear strain rate $\dot{\gamma} = v/\Delta r_s$, with Δr_s as the width of the shear zone, also labelled the radial clearance. This term can be ascertained either from experiments or the analyses of Refs. (47) and (48), since the final results are quite insensitive with respect to its numerical selection; in fact, its value was such that the shear force could be neglected during stage 2. The complete equation of motion is given by

$$F_2(t) = -\tfrac{1}{2}C_1 \rho A_q v^2 - (\tau_Y + \nu \frac{v}{\Delta r_s})\pi D_2 [z - (h_o - h_1)] - \sigma_{UC}A_q \cdot$$

$$\cdot\{1 - \frac{[z - (h_o - h_1)]^2}{\Delta r_s} \qquad \text{for } h_o - h_1 \leq z \leq h \tag{72}$$

and the velocity during this interval can be evaluated from

$$\frac{dv}{dz}(z) = [- (1 + \tfrac{1}{2}C_1) A_q v^2 - \tau_\gamma \pi D_2 z - \frac{\nu \pi D_2 v z}{\Delta r_s} + \frac{\nu \pi D_2 (h_o - h_1) v}{\Delta r_s}$$

$$+ \tau_\gamma \pi D_2 (h_o - h_o) - \sigma_{UC} A_q \{1 - \frac{[z - (h_o - h_1)]^2}{h_1} - 1\}] / (m_o + \rho A_q z) v] \qquad (73)$$

The effective mass at the end of the first stage is $m_{e1} = m_o + \rho A_u (h_o - h_1)$, D_2 is the cavity diameter at the end of phase two, and the duration of this interval can be determined from the integral $t_2 = \int_{h_o - h_1}^{h_o} (1/v)\, dz$.

During the third period, the projectile and plug move together due to the action of shear stress τ acting on the area $\bar{A}_q^* = \pi \bar{D}_2 h_1$ in the shear zone of depth Δr_s around the plug of the average cavity diameter $\bar{D}_2$, taken as the final plug diameter, governed by the relation

$$m_{e2} \frac{d^2 z^*}{dt^2} = F_3 = -\tau \bar{A}_q^* = -(\tau_\gamma + \nu \dot{\gamma}) A_q^* \quad \text{with } z^* = z - h_o = \gamma \Delta r_s \qquad (74)$$

Further, the displacement for material failure, z_f^* is reached at the ultimate shear strain of the material $\gamma_U = z_f^*/\Delta r_s$ beyond which no further resisting force acts on the system. The solution of Eq. (74), with v_{2f} as the velocity at the end of stage 2, is given by

$$z^* = (v_{2f} + \frac{\tau_\gamma \Delta r_s}{\nu})(\frac{m_{e2} \Delta r_s}{\nu A_q^*})[1 - \exp(\frac{-\bar{A}_q^* \nu t}{m_{e2} \Delta r_s})] - \frac{\tau_\gamma \Delta r_s}{\nu} t; \quad 0 \leq z^* \leq z_f^* \qquad (75)$$

and the force in this phase acting during the interval t_3 is

$$F_3 = - \bar{A}_q^* (\tau_Y + \nu \frac{v_{2f}}{\Delta r_s}) \exp [- \frac{\bar{A}_q^* \nu t}{m_{e2} \Delta r_s}] ; \quad 0 \leq t \leq t_3 \qquad (\ )$$

The period required for the plug to leave the target is $(h_1 - z_f^*)/v_f$ with v_f as the final plug velocity at t_3, and the total time for ending the plug ejection from the instant of contact is

$$t_f = t_1 + t_2 + t_3 + (h_1 - z_f^*)/v_f \qquad (\ )$$

This will be followed by the ejection of fragments representing the effective mass added to the projectile in phase 1, and then by the projectile itself.

Representative results using this procedure are presented in Fig. 28, with values taken from both ballistic tests and direct measurement, with Δr_s and D_i, i = 1, 2, 3 measured.[50] The value of $\bar{D}_2$ corresponded closely to the average of D_1 and D_q. Tests have also indicated that the ratio $\bar{D}_2/h_o$ and h_1/h_o for a given striker-target combination appear to be nearly constant over the range of velocities examined.

A different approach of the plugging process during the normal impact of blunt projectiles also considers the simultaneous action of compression and shear under conditions of lateral constraint and divides the phenomenon into two phases:[52] (a) Concurrent compression of the plug to its terminal thickness, shear that moves an equivalent symmetrically deformed plug section to the distal edge of the plate, and an acceleration of the

plug so that its velocity is equal to that of the blunt-nosed projectile, all occurring at a constant force level, and (b) Plug shear and compression until complete ejection occurs with a common projectile and slug velocity, at constant compressive force level, but with a linearly diminishing shear force to an equivalent level dictated by the assumed equivalent symmetrical model during this stage. Elastic behavior is governed by Hooke's law with the plastic domain characterized by a parabolic stress-strain relation augmented by a constraint factor to compensate for the confinement that prevents side flow. The ejection velocity can be evaluated by an energy balance involving the work of compression, shear, and friction during these two stages. The model appears to be somewhat artificial and also depends on both empirically determined plug dimensions and confinement parameters as well as an assumed material behavior pattern.

A still different representation portrays the failure of targets struck normally by sharp-nosed projectiles as an adiabatic shear plugging mode.[53] This may occur when the work-hardening rate of a substance is less than the rate of thermal softening due to heat generated as the result of plastic flow; if it is confined to a narrow annulus, it will result in severe strength reduction of the target. Two modes of failure are examined: a ductile hole enlargement as previously described [30][31][33][34] or a plug shear, the actual process being specified as that requiring minimum energy. Thus, two plate thicknesses are specified in terms of the basic projectile radius; a "thick" target in which a combination of the two mechanisms takes place, and a "thin" plate in which only shearing eventuates. This concept is very appealing; unfortunately, the work of hole enlargement, correctly written in the form of Eq. (45) was overstated by a factor of four that

renders suspect the model and, in particular, the correlation of the predictions of the analysis based on a power-law of stress and plastic strain with experimental data.

PROJECTILE DEFORMATION AND FRACTURE

As demonstrated in Figs. 17 and 18, another important phenomenon in the perforation process requiring modelling is the deformation (and/or fracture) of the projectile. Two general approaches have been developed to provide some predictive capability with respect to changes in projectile shape: (a) Use of plastic wave propagation analysis, occasionally in conjunction with elastic transients, and (b) Hydrodynamic description of the behavior of the striker, based on extensions of the Munroe jet effect. The first technique can be used to predict failure of penetrator rods that occurs when the local stress exceeds the ultimate strength, particularly in tension. While some obvious successes have been scored with this type of investigation, this aspect of the complete phenomenon is the least well-known and understood, particularly with respect to initiation and propagation of fracture in the projectile and its decomposition into a few sizeable or very large number of small fragments. As also recommended elsewhere[5], there is an urgent need to delineate these failure phenomena under conditions of rapid loading, based upon fundamental information in the domains of constitutive behavior, material science and fracture mechanics.

The initial analysis of this deformation considered only normal impact against ideally rigid targets and was developed primarily to deduce the dynamic yield strength of the striker from its observed distortion[11][12], based on plastic-rigid or plastic-work-hardening behavior. The approach has been extended to incorporate the normal impact of blunt-nosed cylinders against non-rigid targets of finite thickness at velocity v_o[4]. Here, the

striker is assumed to exhibit a bilinear stress-strain curve corresponding to a plastic wave velocity $c_p = \sqrt{(1/\rho)(d\sigma/d\varepsilon)}$ in that regime. A fixed reference frame is used to measure all absolute speeds and the target is at rest prior to impact.

The essential mechanism of deformation is illustrated in Fig. 29, together with the nomenclature. Impact of the striker on targets of semi-infinite or finite depth produces a local interface velocity, v_I, together with two plane waves travelling in the cylindrical striker towards the free surface: an elastic compression wave with speed c_o and a slower plastic wave propagating with speed c_p that produces sidewise permanent deformation of the frontal region of the projectile. The stress just forward of the plastic wave is the yield stress σ_Y, the amplitude of the elastic wave, whose initial transit reduces the particle velocity in the rearward portion of the striker from v_o to $v_o - \Delta v$, where $\Delta v = \sigma_Y/\rho_b c_o$. After the first reflection from the free distal surface of the cylinder, the wave cancels stress σ_Y while again reducing the particle velocity by another increment Δv. Upon interaction with the advancing plastic wave front, the process is repeated, resulting in the pattern indicated in the diagram and an instantaneous velocity v at time t for that section of the elastic remainder located between the elastic and plastic wave fronts.

Two specific target models are examined here[4]: (a) A deformable half-space as shown in Fig. 29 in which wave dispersion is ignored and a constant value of v_I as determined from momentum considerations is stipulated, and (b) A rigid plug removed from a plate of thickness h_o, as shown in Fig. 30, where v_I varies according to Newton's law for the plug and the latter is considered to be that sheared by a cylinder with a radius 25 percent larger than bullet radius R_b, in accordance with experimental

evidence[54]. When model (b) predicts a greater deformation than (a), it is replaced by (a).

In considering the case when $v_L \equiv v - v_I \leq c_p$, where Fig. 29 applies, the striker mass m_b remains constant, plastic deformation occurs only within the increment

$$dz = (c_p - [v - v_I])\ dt = (c_p - v_L)\ dt \tag{78}$$

during time dt and conservation of mass for the incompressible cylinder requires that

$$A/A_o = \left[1 - \frac{v_L}{c_p}\right]^{-1} \tag{79}$$

Here, A and A_o denote the areas on the plastic and elastic sides of the plastic wave front, respectively. If L^E represents the length of the elastic portion of the cylinder, of mass $m^E = \rho_b A_o L^E$, then the force acting on m^E is

$$F = -\sigma_Y A_o = m^E(dv/dt) \tag{80}$$

Eq. (80) permits the evaluation of the instantaneous relative velocity v_L as

$$v_L = v_o + (\sigma_Y/\rho_b c_p)\ \ln\left(1 - \frac{c_p t}{L_o}\right) - v_I \tag{81}$$

For the half-space model (a), a momentum balance at the interface yields the value of v_I as

$$v_I = \frac{v_o}{1 + (\rho c_H / \rho_b c_p)} \qquad \text{where} \quad c_H = (K/\rho)^{\frac{1}{2}} \tag{82}$$

here c_H is the bulk velocity, ρ the density, and K the bulk modulus of the target. For model (b) with negligible impulse transferred to the target, found to be an excellent approximation when the initial velocity was at least 25 percent greater than the ballistic limit, a momentum balance yields

$$m_{b_o} v_o = m^E v + (m_{b_o} - m^E + m_q) v_I \quad \text{or} \quad v_I = (m_{b_o} v_o - m^E v)/(m_{b_o} - m^E - m_q) \tag{83}$$

with plug mass $m_q = \pi\rho(1.25\ R_b)^2\ h_o$. The length of the terminal undeformed portion of the cylinder for the two cases is given by

$$\text{(a)} \quad L_f^E = L_o \exp\ (-v_o \rho_b c_p / 7\sigma_Y) \quad . \qquad \text{and} \tag{84}$$

$$\text{(b)} \quad L_f^E = L_o \exp\ [(-v_o \rho_b c_p / \sigma_Y)\{m_q/(m_q + m_o)\}] \tag{85}$$

where the impedance factor $Z = 1 + (\rho_b c_p / \rho c_H)$. Integration of Eq. (78) yields $z = f(t)$, and Eq. (79) can then be combined with this result to provide an implicit relation for $t = t(A/A_o)$ which, in turn permits relation of A/A_o to z. This yields for the two cases

$$\text{(a)} \quad \frac{z}{L_o} = \eta^* + \left(1 - \frac{v_o}{c_p Z}\right) - \left(\eta^* + \frac{A_o}{A}\right) \exp\left\{\frac{1}{\eta^*}\left[1 - \frac{A_o}{A} - \frac{v_o}{c_p Z}\right]\right\}$$

$$\text{where} \quad \eta^* = \sigma_Y / \rho_b\ c_p^{\,2} \tag{86}$$

$$(b)\quad \frac{z}{L} = \frac{c_p t}{L_o} + \frac{\eta^*}{1.54}\left[\left(1 - \frac{c_p t}{L_o}\right) \ell n \left(\frac{1 - c_p t/L_o}{\xi}\right) + \frac{c_p t}{L_o} + \ell n \xi\right]$$

$$+ 0.96 \frac{\eta^* c_p t}{L_o} - \left[\eta^* \xi \ell n \xi + \frac{m_q}{m_{b_o}} \frac{v_o}{c_p}\right] \ell n \left(1 + \frac{c_p t}{L_o} \Big/ \frac{m_q}{m_{b_o}}\right) \qquad (87)$$

where $\xi \equiv 1 + (m_q/m_{b_o})$ and where an approximation in the integration for case (b) has been employed, $[\ln\zeta/(1 - \zeta)] \approx \{(\ln \zeta)/1.54\} - 0.96$ which is satisfatory for $Z > 0.1$; otherwise, the cylinder is nearly disintegrated. Furthermore,

$$A/A_o = \left[\frac{m_q}{m_{b_o}} + \frac{c_p t}{L_o}\right] \Bigg/ \left[\frac{m_q}{m_{b_o}} + \frac{c_p t}{L_o} - \eta^* \xi \ell n \left(1 - \frac{c_p t}{L_o}\right) - \frac{m_q}{m_{b_o}} \frac{v_o}{c_p}\right] \qquad (88)$$

and a combination of Eqs. (87) and (88) permits expression of $z = z(A/A_o)$ at specified values of time.

When the value of $v_L \equiv v - v_I > c_p$, the plastic wave cannot travel away from the interface, and a standing shock wave is produced at some distance from the interface which erodes the material passing through this section by disintegration, ablation, melting, or flashing.[4][24] When the velocity v_L has dropped sufficiently so that $v_L \leq c_p$, the previous analysis applies. Thus, the problem can be solved in two steps: An erosion model is utilized for the case when $v_L > c_p$ that determines the residual length L_1, mass $m_{b_1} = (L_1/L_o)m_{b_o}$ and velocity $v_1 = Zc_p$ when v_L has been reduced to c_p. These quantities serve as the initial conditions for the analysis described by Eqs. (78) - (88) for the evaluation of the terminal projectile shape, as shown in Fig. 30. During

subsequent plate perforation, the laterally expanded segment of the striker may be sheared off, and this will be also evaluated.

For the half-space, neglecting the distance between the shock wave and the impact surface and utilizing the equation of motion for the remaining portion of the cylinder $m = \rho A_o L^E$ with $dL^E = - v_L\, dt$ and integrating between the initial value $v_L = v_o/Z$ and the final value c_p, for corresponding cylinder lengths L_o and L_{f_1} yields

$$L_{f_1} = L_o \exp \{[-(\rho/2\sigma_Y)]\,[(v_o/Z)^2 - c_p^2]\} \quad \text{with } m_{b_f} = (L_f/L_o) m_{b_o} \tag{89}$$

The plug sheared from the plate is accelerated during erosion; the impulse transmitted to the plate is ignored, and it is further assumed that all eroded material has a velocity component in direction z equal to the instantaneous plug velocity v_q. Thus a momentum balance gives

$$\sigma_Y A_o\, dt + (v - v_q)(- dm) = m_q dv_q \tag{90}$$

while

$$v - v_q = - (dm/dt)/\rho A_o \tag{91}$$

If in the resulting equation for $m(t)$, the ratio m_q/m is neglected, corresponding to the case of a thin plate, the solution where $m_b = m_{b_1}$, i. e., when $v - v_q = c_p$ is

$$m_{b_1} = m_{b_o} + \tfrac{1}{2} m_q \ln ([1 + (\eta^*)^{-1}]/\{1 + (v_o/c_p)^2/\eta^*\}) \tag{92}$$

where the new length after erosion L_{f_1} is

$$L_{f_1} = (m_q/m_{b_o})L_o \tag{93}$$

In thin plate perforation, a ring of material is frequently separated from the deformed cylinder as indicated in Fig. 30. Test results show that the final deformed projectile diameter is 25 percent greater than the initial value, corresponding to an area ratio $A/A_o = 1.56$. The time t_f corresponding to this deformation is computed, and the corresponding length $z_f(t_f)$ can then be determined. The residual mass of the cylinder after losses due to erosion and/or shear is given by

$$m_g = \rho A_o (L_g - c_p t_f + 1.56 z_f) \tag{94}$$

where L_g is the remaining cylinder length when erosion has just stopped. The predictions of this analysis were in good agreement with test results on 0.22 caliber steel projectiles fired at steel plates at speeds up to 3850 ft/s.

The other popular treatment of projectile deformation consists of a hydrodynamic description of the penetration of long rods into thick targets, representing the essential elements of shaped charge action. The model, shown in Fig. 31, involves the motion with velocity $v(z)$ of a jet of density ρ_j and length L_j which penetrates a target of density ρ to a distance $P(z)$ at the rate $U(t)$. If ξ denotes the initial z coordinate of the element arriving at the stagnation point P_Z at time t, and if

$v_j(\xi)$ is assumed to be constant prior to impact, the element displacement at time t is

$$tv(\xi) = \xi(t) + P(t) \tag{95}$$

Under steady-state conditions, the event is governed by a modified form of the Bernoulli equation that accounts for the dynamic yield strength of striker σ_{Y_j} and target σ_Y

$$\sigma_{Y_j} + \tfrac{1}{2}\rho_j (v_j - U)^2 = \tfrac{1}{2}\rho U^2 + \sigma_Y \tag{96}$$

Neglect of these strengths permits an integral expression for $P(t)$ in the form [55]

$$P = L_j\left[\left(\frac{\rho_j}{\rho} + 1\right)/(v_R)^{\rho_j/\rho} \int_0^1 v^{\rho_j/\rho}(\zeta)\, d\zeta - 1\right] + S\left[\left(\frac{v_F}{v_R}\right)^{\rho_j/\rho} - 1\right] \tag{97}$$

where v_F and v_R are the velocities of the front and rear of the jet, respectively, and S is the standoff distance. If the jet velocity is

$$P = L_j \sqrt{\frac{\rho_j}{\rho}}\left[1 - \sqrt{\frac{\rho_j}{\rho}\left(1 - \frac{2\sigma_Y}{\rho_j v_j^2}\{1 - \frac{\rho_j}{\rho}\}\right)}\right]\left[\sqrt{1 - \frac{2\sigma_Y}{\rho_j v_j^2}(1 - \frac{\rho_j}{\rho})}\right.$$

$$\left. - \sqrt{\frac{\rho}{\rho_j}}\right]^{-1} \tag{98}$$

The change in behavior of a projectile from an undeformable rod to the jet inversion shown in Fig. 31 occurs in some critical velocity range v_0

experimentally found to lie between 3200 and 6500 ft/s[56]. Contrary to the assumption for shaped-charge effect that leads to Eq. (96), for an initial velocity v_o just above v_Q, the crater bottom does not move supersonically either with respect to striker or target, and hence there exists a second critical speed v_Q' above which such supersonic comportment occurs. This velocity is derivable from Eq. (96) by replacing $v_j - U$ with the sound velocity of the striker, and is equal to twice this velocity for identical projectile and target materials. Between the two limits, the rod inverts as shown, but an additional force is transferred to the rod that affects penetration velocity U.

In addition to the penetration given by Eqs. (97) or (98) (the latter frequently being used with the strength neglected), that occurs in the relatively short interval

$$t_1 = (L_j/v)(1 + \sqrt{\frac{\rho_j}{\rho}}) \tag{99}$$

there exist two other stages in such a hypervelocity impact: (a) a secondary penetration or cavitation when the projectile has been totally deformed, but the shock wave and cavity continue to expand, and (b) a recovery stage when contraction of the cavity due to elastic-plastic or brittle restitution occurs, sometimes ending in brittle failure. The secondary penetration has been approximated by half the crater diameter, in reasonable agreement with experimental results on steels and aluminum in the velocity range from 6500 to 22,000 ft/s.

The behavior of a rod travelling with initial speed v_o and characterized by incompressible hydrodynamic material behavior with initial

and instantaneous lengths of L_o and $L(t)$, respectively, may be derived by use of Eq. (96) and an equation of motion for the striking projectile retarded by a braking force at the base of the crater, given by[57]

$$\rho_b L(t) \frac{dv_b}{dt} = - \sigma_{Y_b} \quad \text{where} \quad dL(t)/dt = - (v_b - U) = (dz/dt) - v_b \tag{100}$$

The value of U is obtained here by assuming the strength difference $\sigma_Y - \sigma_{Y_b}$ to be small compared to the dynamic loads, yielding

$$U = (dz/dt) = B_1 v_b - B_2 (v_o^2/v_b) \quad \text{where} \quad B_1 = (1 + \sqrt{\rho/\rho_b})^{-1}$$

$$B_2 \quad (\frac{1}{v_o^2}) \frac{\sigma_Y - \sigma_{Y_b}}{\sqrt{\rho}\ \sqrt{\rho_b}} \tag{101}$$

With the definitions

$$B_3 \equiv \frac{\sigma_{Y_b}}{\rho_b v_o^2} \; ; \quad \chi \equiv v_b/v_o \leq 1 \; ; \quad \text{and} \quad f(\chi) = \exp\left[\frac{B_1 - 1}{2B_3} (1 - \chi^2)\right] \tag{102}$$

it follows that

$$L(t) = L_o (\chi)^{B_2/B_3} f(\chi) \qquad \text{and} \qquad t = (L_o/v_o B_3) \int_\chi^1 (\chi)^{B_2/B_3} f(\chi)\, d\chi \tag{103}$$

$$\text{and} \qquad z = (L_o/B_3) \int_\chi^1 (\chi)^{B_2/B_3} (B_1\chi - B_2/\chi) f(\chi)\, d\chi \tag{104}$$

The crater radius R_c is determined from a balance of the deformation energy of the target and the energy lost by the projectile and is given by

$$R_c = R_b \sqrt{\frac{(\chi^2 - 2B_3)[(1 - B_1)\chi^2 + B_2]}{B_4(B_1\chi^2 - B_2)}} \qquad \text{where} \quad B_4 \equiv \tilde{2E}/\rho_b v_o^2 \qquad (105)$$

and $\tilde{E}$ is the work required to displace unit volume of the material. The validity of the relations diminishes as $v_o \to v_Q'$ since, under these circumstances, the shock moves slowly toward the rear of the cylinder, and the braking force cannot be treated as continuous, as indicated by Eq. (100). It is interesting to observe that the perforation of a target by a projectile initially travelling at a velocity $v_o < v_Q$ occurs with a significant drop in velocity, but no decrease in mass. On the other hand, for a shaped charge exhibiting a velocity $v_j = v_o > v_Q'$, penetration takes place without a velocity drop, but with a substantial decrease in mass; in between these extremes, a slender rod with a velocity $v_C \leq v_o \leq v_Q'$ loses both mass and speed. Similar penetration analyses have also been reported in Refs. (58) and (59).

Investigations of initiation of failure in the projectile have thus far been limited to the same phenomenological basis as for the target, involving fracture into components when the local stress in tension, shear or compression exceeds the respective ultimate strength. The stress state in the striker is generally considered to be uniaxial which may represent a degree of realism for long rod kinetic energy penetrators, but is probably not applicable for projectiles with small aspect ratios where stress states are distinctly multi-dimensional. Furthermore, these simplistic failure criteria are known to be inapplicable to short-duration, high-amplitude loading where the fracture phenomena depend very definitely on time and probably also significantly on temperature[5]. Spalling of the

material results from the tension generated by rarefaction waves created by the impact phenomenon, whereas shear or bending failures occur due to antisymmetric loading.

The fracture process has been analyzed from three distinct viewpoints: (a) Methods of continuum mechanics based on an instantaneous or cumulative damage criterion, (b) Crack nucleation at the microscopic level, and (c) Crack propagation descriptions based on both continuum and fracture mechanics precepts. The author is also currently engaged in the construction of a model for the combined tensile, shear and crushing failure of polycrystalline masses bonded by cementitious substances, such as represented by most natural rocks, under projectile impact. However, knowledge in all these areas, as well in suitable constitutive descriptions of both projectiles and targets under high and rapidly applied loads below the failure level need still to be substantially expanded to permit better predictions of the phenomena of impact and perforation of targets by projectiles.

RECOMMENDATIONS AND CLOSURE

Suggestions for needed investigative activities in the area of projectile penetration into targets have recently been given in both Refs. (1) and (5); these include the development of more rational bases of specifying modeling parameters, improved description of interface phenomena and analytical modeling of force contributions, more emphasis on the effects of obliquity and flight orientation of the projectile, better constitutive representations, including thermo-mechanical coupling and heat dissipation as well as delineation of failure criteria. In addition, improvements in computer codes and better utilization of their capability have been outlined in a

variety of areas. In closing the present paper, however, the writer would like to specify his concept of an immediate area of improvement in currently available phenomenological models based on relatively simple mathematical representations that will combine a number of the developments surveyed here. These problems will be attacked by the author and his associates in the near future.

The first and most obvious step is a combination of the three stage compression-shear-plugging model with the flexure of the target. This can be accomplished by using an axisymmetric shear loading of the plate at the periphery of the penetrating cylinder based on a perfectly-plastic or work-hardening material behavior concept, and referring the motion of the cylinder to a moving coordinate system embedded in the plate at the crater edge. A second effort would be an attempt to eliminate the current empirical (or assumed) entrance and exit diameters of the crater by analyzing the cylinder deformation on the basis of the solid continuum, described previously, or some combination of this with hydrodynamic concepts that will permit prediction both of projectile deformation and hole geometry in conjunction with the other phenomenological modelling steps. There appears to be a distinct need to characterize the limiting states of material strength as a function of the amount and rapidity of prior plastic work performed so that the adiabatic shear phenomenon can be better characterized.

A parallel investigation should be conducted into the geometrical and field variables that control hole formation in thin plates due to sharp-pointed strikers and the formation and propagation of cracks generating petals. Here, also, thermo-mechanical relations are undoubtedly required

to specify material behavior. This is but one aspect of crack propagation in targets that needs substantial further study.

The subject of penetration and perforation of targets by projectiles is one of the most complicated phenomena in the field of mechanics and there is little danger that a sufficiently high level of understanding in all its aspects will be attained in the next few decades to relegate the field to investigative oblivion. However, the past history of the subject covering more than two centuries has not been marked by major breakthroughs (with the probable exception of the development of numerical codes), as has happened in certain areas of physics, but rather has been characterized by the slow, steady accumulation of knowledge provided by a host of investigators, and this is also the prospect for the future. Furthermore, the vast plurality of the contributions have been initiated by government-sponsored activities in the weapons area. The topic has numerous industrial uses and much potential for further development in this domain. It would be highly desirable for further progress in this area to seek new and profitable applications of perforation processes in non-military technological areas which would attract a much larger fraction of the scientific community to work on the challenging unsolved problems in this field.

ACKNOWLEDGMENT

The author would like to acknowledge the assistance of Mr. M. E. Backman of the Naval Weapons Center, China Lake, California, whose collaboration in an earlier paper contributed substantially to the present work, and of Mr. R. Recht of the University of Denver Research Institute who furnished original photographs. A portion of this presentation was synthesized under the auspices of the Naval Weapons Center, China Lake.

LIST OF SYMBOLS

A	Area
B,C	Constants
$\overline{BH}$	Brinell Hardness Number
E*	Specific Energy
E	Work to displace Unit Volume
F	Force
G	Shear Modulus
H	Crater Height
K	Bulk Modulus
L	Length
M	Bending Moment per Unit Length
N	Membrane Force per Unit Length
P	Penetration
Q	Shear Force per Unit Length
R	Radius, Radius of Curvature
S	Standoff Distance
T	Time
T*	Viscoplastic Relaxation Time
U	Penetration Rate
W	Work
Z	Impedance
b	Plug Thickness
c	Wave Speed
c_o	Rod Wave Velocity
h	Plate Thickness
m	Mass
n	Scale Factor

p	Pressure
r	Radius, Radial Coordinate
s	Transform Parameter
t	Time
u	Displacement in Plane of Plate
v	Velocity
w	Transverse Plate Displacement
z	Coordinate normal to Plate, Position
z^*	Distance from Crater Tip
α	Stretch Parameter
β	Half-Cone Angle
γ	Shear Strain
ε	Normal Strain
ζ	Variable
θ	Polar Coordinate, Angle of Obliquity
θ^*	Angular Change in Travel Direction
κ	Curvature
λ	Lamé Constant
ν	Dynamic Viscosity
ρ	Density
σ	Normal Stress
τ	Shear Stress
Δ	Difference
Φ	Slope of Projectile Nose

Subscripts

C	Compression
D	Dilatation
F	Front

H	Hydrodynamic
I	Interface
L	Relative
N	Nose
Q	Critical
R	Rear
S	Shear
U	Ultimate
Y	Yield
Z	Stagnation
b	Projectile
c	Crater
e	Effective
f	Final
g	Remaining after Erosion
i	Initial Combined
j	Jet
o	Initial
q	Plug
r	In r Direction
u	Projected
w	Perfectly Plastic
50	Ballistic Limit
θ	In θ Direction
π	Contact

Superscripts

D Dynamic

E Elastic

P Plastic

S Static

A dot over a symbol denotes its derivative with respect to time

A prime after a symbol denotes its derivative with respect to the argument

REFERENCES

1. Backman, M.E., and Goldsmith, W., "The Mechanics of Penetration of Projectiles into Targets", Int. J. Engng. Sci., v. 16, pp. 1-99, 1978.

2. Sih, G.C., ed. Proceedings of the 14th Annual Meeting of the Society of Engineering Science. Bethlehem, Pa., Lehigh University Press, pp. 3-109, 1977.

3. Eringen, A.C., ed. "Penetration Mechanics", Special Issue, Int. J. Engng. Sci., v. 16, n. 11, pp. 793-920, 1978.

4. Recht, R.F., "Taylor Ballistic Impact Modelling Applied to Deformation and Mass Loss Determinations", Ref. 3, pp. 809-828.

5. Jonas, G.H., and Zukas, J.A., "Mechanics of Penetration: Analysis and Experiment", Ref. 3, pp. 879-904.

6. Recht, R.F., and Ipson, T.W., Ballistic Penetration Resistance and its Measurement. Report NWC TP 5648. Naval Weapons Center, China Lake, CA,1974. See also Proc. 1st Int. Symp. on Ballistics., IV-315-330, 1974.

7. Goldsmith, W., Liu, T.W., and Chulay, S., "Plate Impact and Perforation by Projectiles", Exp. Mech., v. 5, pp. 385-404, 1965.

8. Calder, C.A., and Goldsmith, W., "Plastic Deformation and Perforation of Thin Plates Resulting from Projectile Impact", Int. J. Solids Structures, v. 7, pp. 863-881, 1971.

9. Calder, C.A., Plastic Deformation and Perforation of Thin Plates Resulting from Projectile Impact. Dissertation (Ph.D.) University of California, Berkeley, 1969.

10. Goldsmith, W., and Finnegan, S.A., "Penetration and Perforation Processes in Metal Targets at and Above Ballistic Velocities", Int. J. Mech. Sci., v. 13, pp. 843-866, 1971.

11. Taylor, G.I., "The Use of Flat-ended Projectiles for Determining Dynamic Yield Stress. I. Theoretical Considerations", Proc. Roy. Soc. Lond., A., v. 194, pp. 289-299, 1948.

12. Lee, E.H., and Tupper, S.J., "Analysis of Plastic Deformation in a Steel Cylinder Striking a Rigid Target", J. Appl. Mech., v. 19, pp. 63-70, 1954.

13. Krafft, J.M., "Surface Friction in Ballistic Penetration", J. Appl. Phys., v. 26, pp. 1248-1253, 1955.

14. Reissner, E., "On Finite Deflections of Circular Plates", Proc. Symp. Appl. Math. v. 1, Amer. Mathematical Soc., N.Y., pp. 213-219, 1949.

15. Jones, N., "Impulsive Loading of a Simply Supported Circular Rigid Plastic Plate", J. Appl. Mech., v. 35, pp. 59-65, 1968.

16. Jones, N., "Finite Deflections of a Simply Supported Rigid-Plastic Annular Plate Loaded Dynamically", Int. J. Solids Structures, v. 4, pp. 593-603, 1968.

17. Jones, N., "Finite Deflections of a Rigid-Viscoplastic Strain-Hardening Annular Plate Loaded Impulsively", J. Appl. Mech., v. 35, pp. 349-356, 196

18. Jones, N., "A Theoretical Study of the Dynamic Plastic Behavior of Beams and Plates with Finite Deflections", Int. J. Solids and Structures, v. 7, pp. 1007-1029, 1971.

19. Jahsman, W.E., "Propagation of Abrupt Circular Wave Fronts in Elastic Sheets and Plates", Proc. 3rd U.S. Natl. Congr. Applied Mechanics, New York, ASME, pp. 195-202, 1958.

20. Hodge, P.G., "Yield Conditions for Rotationally Symmetric Shells Under Axisymmetric Loading", J. Appl. Mech., v. 27, pp. 323-331, 1960.

21. Onat, E.T., and Prager, W., "Limit Analysis of Shells of Revolution", Proc. Roy. Netherlands Academy of Science, Parts I and II, Ser. B, v. 57, pp. 534-541, 542-548, 1954.

22. Beynet, P., and Plunkett, R., "Plate Impact and Plastic Deformation by Projectiles", Exp. Mech., v. 11, pp. 64-70, 1971.

23. Shawa, O., The Application of Dynamic Plastic Analysis to Problems of Structural Impact., Dissertation (Ph.D.), University of California, Berkeley, 1977.

24. Goldsmith, W., Impact, London, E. Arnold, 1960.

25. Wang, A.J., and Hopkins, R.G., "On the Plastic Deformation of Built-in Circular Plates under Impulsive Loads", J. Mech. Phys. Solids, v. 3, pp. 22-37, 1954.

26. Kelly, J.M., and Wierzbicki, T., "Motion of a Circular Viscoplastic Plate Subject to Projectile Impact", Z. ang. Math. Phys., v. 18, pp. 236-246, 1967.

27. Wierzbicki, T., and Florence, A.L., "A Theoretical and Experimental Investigation Of Impulsively Loaded Clamped Circular Viscoplastic Plates", Int. J. Solids Structures, v. 6, pp. 553-568, 1970.

28. Calder, C.A., Kelly, J.M., and Goldsmith, W., "Projectile Impact on an Infinite, Viscoplastic Plate", Int. J. Solids Structures, v. 7, pp. 1143-1152, 1971.

29. Kelly, J.M., and Wilshaw, T.R., "A Theoretical and Experimental Study of Projectile Impact on Clamped Circular Plates", Proc. Roy. Soc., London, A., v. 306, pp. 435-447, 1968.

30. Bethe, H., "An Attempt at a Theory of Armor Penetration", Rept. No. UN-41-4-23, Frankford Arsenal, Ordnance Laboratory, 1941.

31. Taylor, G.I., "The Formation and Enlargement of a Circular Hole in a Thin, Plastic Sheet", Quart. J. Mech. Appl. Math., v. 1, pp. 103-124, 1948.

32. Johnson, W., Chitkara, N.R., and Bex, P.A., "Characteristic Features in the Hole Flanging and Piercing of Thin and Thick Circular Plates Using Conical and Ogival Punches", Proc. 15th Int. Machine Tool Design and Research Conference, ed. by S. A. Tobias and F. Koenigsberger., pp. 695-701, 1974.

33. Freiberger, W., "A Problem in Dynamic Plasticity: The Enlargement of a Circular Hole in a Flat Sheet", Proc. Camb. Phil. Soc., v. 48, pp. 135-148, 1952.

34. Thomson, W.T., "An Approximate Theory of Armor Penetration", J. Appl. Phys., v. 26, pp. 80-82, 1955.

35. Zaid, M., and Paul, B., "Mechanics of High-speed Projectile Perforation", J. Franklin Inst., v. 264, pp. 117-126, 1957.

36. Richter, H., La théorie de la perforatian des blindages, Rept. LRSL 3/46 (ISL). Franco-German Armament Research Establishment, St. Louis, France, Ballistics Institute, 1946.

37. Richter, H., Recherches sur la théorie de la perforation des blindages, Rept. LRSL 20/50 (ISL). Franco-German Armament Research Establishment, St. Louis, France, Ballistics Institute, 1950.

38. Gabeaud, M.L., Mémorial de l'Artillerie française, v. 54, 1935, cited in: Sutterlin, R., Les Projectiles, ibid, v. 40, pp. 569-650, 850-922, 1966 and v. 41, pp. 12-110, 1967.

39. Nishiwaki, J., "Resistance to the Penetration of a Bullet Through an Aluminum Plate", J. Phys. Soc. Japan, v. 6, pp. 374-378, 1951.

40. Spells, K.E., "Velocities of Steel Fragments After Penetration of Steel Plates", Proc. Phys. Soc. Lond., B, v. 64, pp. 212-218, 1951.

41. Recht, R.F., and Ipson, T.W., "Ballistic Perforation Dynamics", J. Appl. Mech., v. 30, pp. 384-390, 1963.

42. Kochetkov, A.M., On the Propagation of Elastic-Viscoplastic Shear Waves in Plates (in Russian), Prikladn. Matem. y. Mekhan., v. 14, pp. 203-208, 1950.

43. Cristescu, N., Dynamic Plasticity. Amsterdam, North-Holland, 1967.

44. Bakhshiyan, F.A., "Visco-plastic Flow by Impact of a Cylinder upon a Plate" (in Russian), Prikladn. Matem. y. Mekhan., v. 12, pp. 47-52, 1948.

45. Chou, P.C., "Viscoplastic Flow Theory in Hypervelocity Perforation of Plates", Proc. 5th Symposium on Hypervelocity Impact", v. 1, pt. 1, pp. 307-328, 1962.

46. Pytel, A., and Davids, N., "A Viscous Model for Plug Formation in Plates", J. Frankl. Inst., v. 276, pp. 394-406, 1963.

47. Thomson, R.G., Analysis of Hypervelocity Perforation of a Viscoplastic Solid Including the Effects of Target Material Strength, NASA Tech. Rept. R-221, 1965.

48. Awerbuch, J., "A Mechanics Approach to Projectile Penetration", Israel J. Technol., v. 8, pp. 375-383, 1970.

49. Awerbuch, J., and Bodner, S.R., "Analysis of the Mechanics of Perforation of Projectiles in Metallic Plates", Int. J. Solids Structures, v. 10, pp. 671-684, 1974.

50. Awerbuch, J., and Bodner, S.R., "Experiments on the Normal Perforation of Projectiles in Metallic Plates", Int. J. Solids and Structures, v. 10, pp. 685-699, 1974.

51. Chou, P.C., "Perforation of Plates by High-Speed Projectiles", Developments in Mechanics, ed. by J.E. Lay and L.E. Malvern, v. 1, Amsterdam, North-Holland, pp. 286-295, 1961.

52. Woodward, R.L., and De Morton, M.E., "Penetration of Targets by Flat-ended Projectiles", Int. J. Mech. Sciences, v. 18, pp. 119-128, 1976.

53. Woodward, R.L., "The Penetration of Metal Targets which fail by Adiabatic Shear Plugging", Int. J. Mech. Sciences, v. 30, pp. 599-607, 1978.

54. Ipson, T.W., Deformation and Reduction in Weight of Compact Steel Fragments Perforating Thin, Mild Steel Plates. Naval Weapons Center, China Lake, CA, NWC-TP-4533, 1968.

55. Abramson, G.K., and Goodier, J.N., "Penetration by Shaped Charge Jets of Non-uniform Velocity", J. Appl. Phys., v. 34, pp. 195-199, 1963.

56. Defourneaux, M., Pénétration d'un projectile dans un matériau plastique, Sciences et techniques de l'armament, Mémorial de l'Artillerie française v. 45, pp. 645-671, 1971.

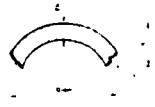
57. Alexeievski, V.P., "Penetration of a Rod into a Target at High Velocity", Combustion, Explosion and Shock Waves, v. 2, n. 2, pp. 63-66, 1966.

58. Tate, A., "A Theory for the Deceleration of Long Rods After Impact", J. Mech. Phys. Solids, v. 15, pp. 387-399, 1967.

59. Tate, A., "Further Results in the Theory of Long Rod Penetration", J. Mech. Phys. Solids, v. 17, pp. 141-150, 1969.

TABLE 1. Perforation Data for Aluminum Plates

RUN NO.	PLATE SIZE	PROJECTILE VELOCITY ft/s INITIAL	PROJECTILE VELOCITY ft/s FINAL	PROJECTILE TYPE	MOMENTUM CHANGE, Δmv, lb-s	NO. OF PETALS	PLUG (CAP) MEASUREMENTS MASS g	DIAMETER in.	2h in.	d in.
1	a	371	163	c	0.0435					
2	a	381	200	c	0.0379					
3	b	259	113	d	0.1176					
4	b	261	106	d	0.125					
5	b	379	163	c	0.045					
6	b	497	315	e	0.1044	-	0.135	0.326	0.0413	0.093
7	b	682	587	e	0.0545	-	0.187	0.378	0.0425	0.110
8	b	933	864	e	0.0396	-	0.269	0.425	0.0427	0.145
9	b	1289	1016	f	0.0197					
10	b	301	195	d	0.0847	6				
11	b	391	302	d	0.0711	4				
12	b	497	445	d	0.0416	4				
13	b	570	521	d	0.0392	4				
14	b	840	803	d	0.0296	4				
15	b	861	806	c	0.0115	4				



Plates: (a) 4 ft x 4 ft x 0.050 in. freely suspended
(b) 14.5 in. diameter x 0.050 in. thick clamped on a 14 in. diameter

Projectiles: (c) 60° cylindro-conical, 1/4 in. diameter, 5/8 in. long, $m = 1.74 \times 10^{-5}$ lb-s^2/in
Hard (d) 60° cylindro-conical, 1/2 in. diameter, 3/4 in. long, $m = 6.66 \times 10^{-5}$ lb-s^2/in
Steel (e) 1/2 in. diameter sphere, $m = 4.78 \times 10^{-5}$ lb-s^2/in
(f) 1/4 in. diameter sphere, $m = 0.60 \times 10^{-5}$ lb-s^2/in

TABLE 2. Mechanical Properties of the Target Plates

MATERIAL	h_o THICKNESS (in.)	E YOUNG'S MODULUS (10^6 lb/in^2)	ν POISSON'S RATIO	σ_p PROPORTIONAL LIMIT (10^3 lb/in^2)	σ_U ULTIMATE STRENGTH* (10^3 lb/in^2)	σ_S ULTIMATE STATIC SHEAR STRENGTH (10^3 lb/in^2)	ROCKWELL HARDNESS
2024-0 Aluminum	0·05	10·6	0·33	12·8	34·0	19·0	H88
2024-T3 Aluminum	0·05, 0·125	10·6	0·33	53·0	70·0	41·0	B76
2024-T4 Aluminum	0·25	10·6	0·33	53·0	69·2	41·0	B78
SAE 1020 Steel, large-grained	0·25	29·6	0·29	64·0	71·0	45·2	B90
SAE 1020 Steel, small-grained	0·062, 0·25	29·6	0·29	48·0	59·5	37·9	B77
SAE 4130 Steel, quenched and tempered	0·25	29·6	0·29	170·0	189·0	113·5	C40

*Quasistatic tension.

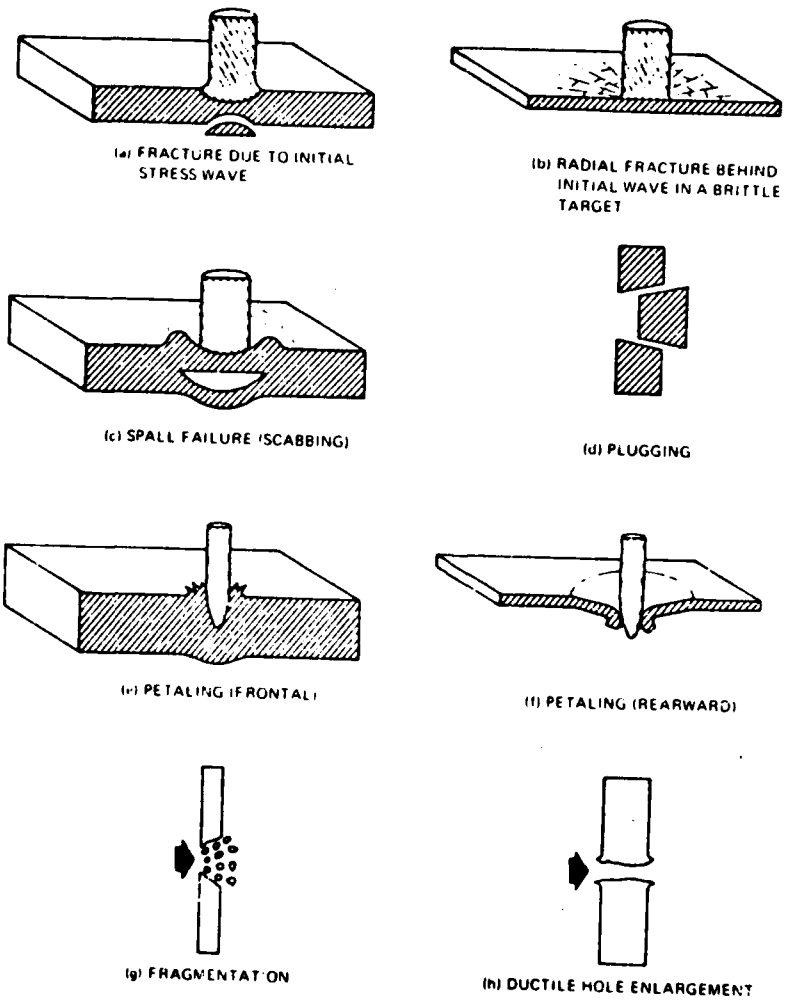


Fig. 1 Perforation Modes

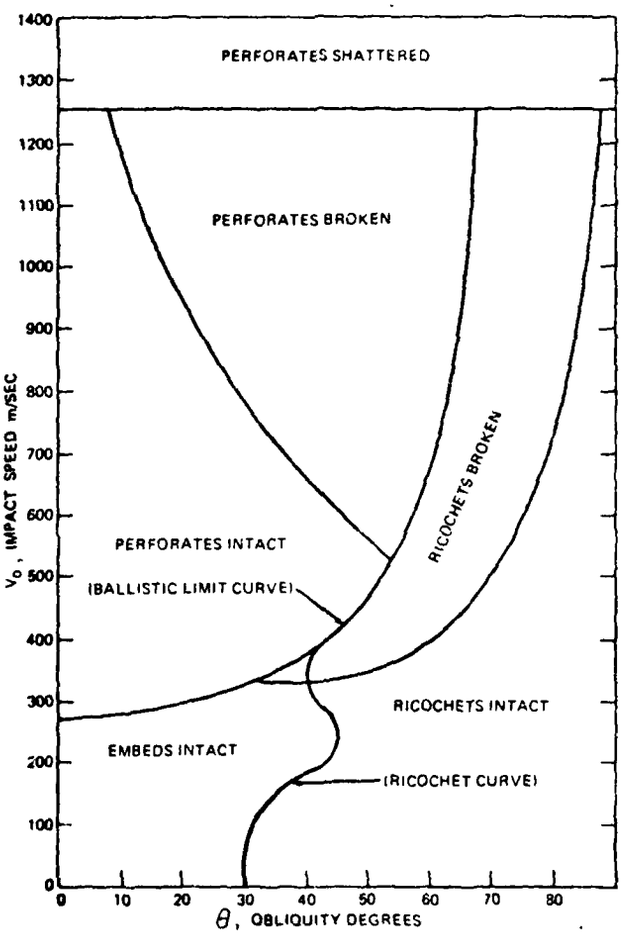


Fig. 2 Phase Diagram for a 6.35 mm Dia. Ogival-Nosed Projectile and a 6.35 mm thick Aluminum Alloy Target

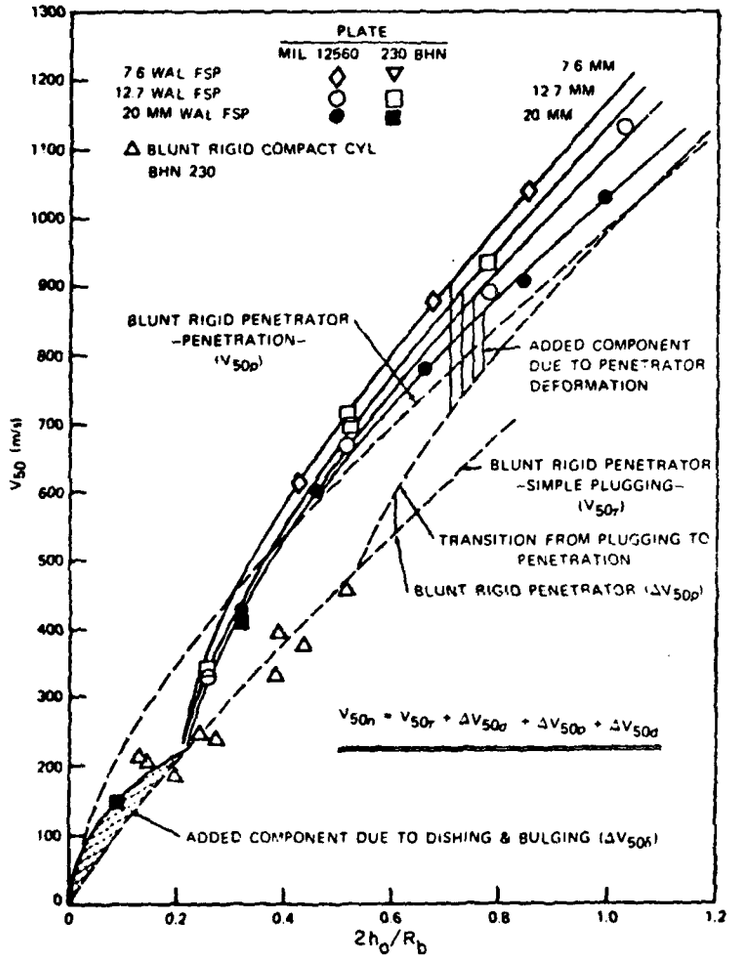


Fig. 3 Contributions to the Ballistic Limit for Rigid and Deforming Blunt, Compact Fragments Striking Steel Plates (Ref. 6)

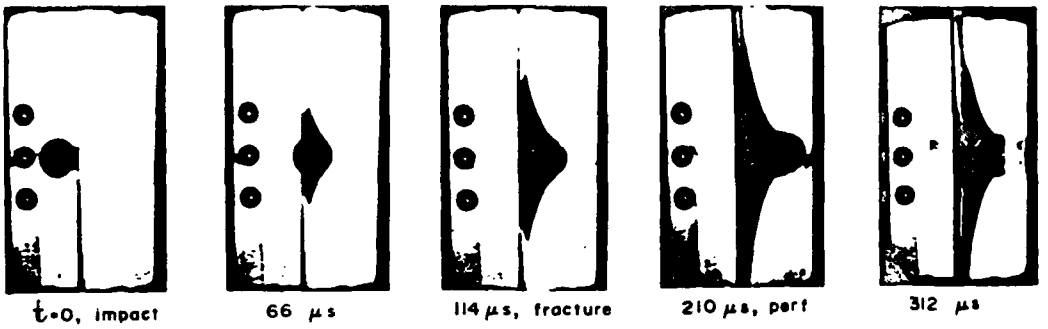


Fig. 4 Perforation of a 0.050-in. thick Aluminum Plate Clamped at a Radius of 7 in. by a ½-in. Dia. Steel Sphere. Initial velocity: 494 ft/s. Final velocity: 315 ft/s. Framing rate: 168,000/s

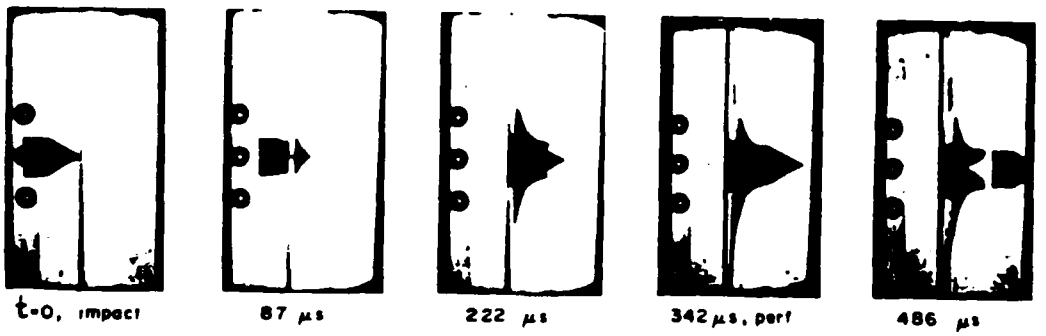


Fig. 5 Perforation of a 0.050-in. thick Aluminum Plate Clamped at a Radius of 7 in. by a ½-in. Dia. Cylindro-conical Steel Projectile with a Half-cone Angle of 30°. Initial velocity: 300 ft/s. Final velocity: 195 ft/s. Framing rate: 110,000/s

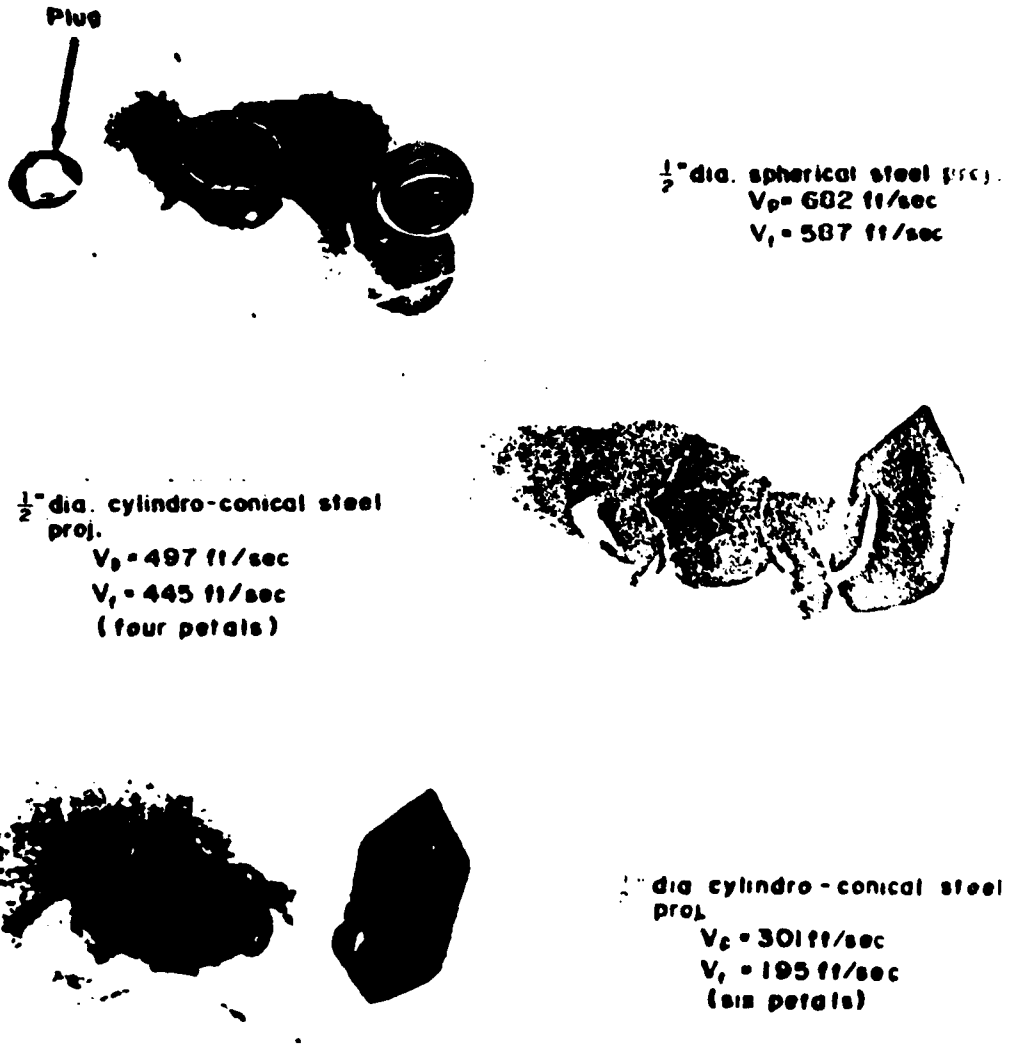


Fig. 6 Post-mortem Perforation Patterns for a Clamped 2024-0 Aluminum Plate, 0.050 in. thick

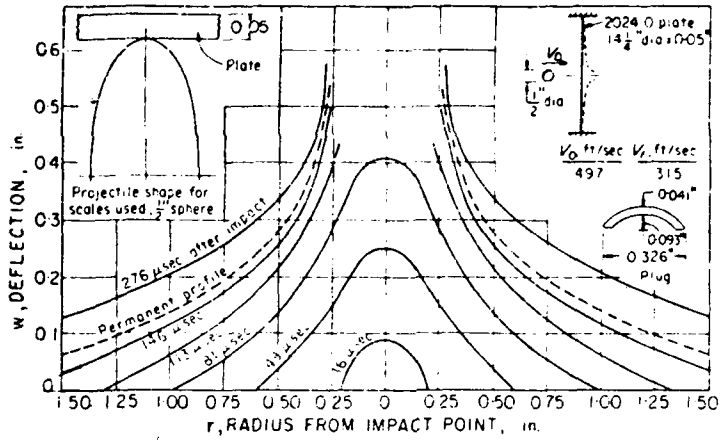
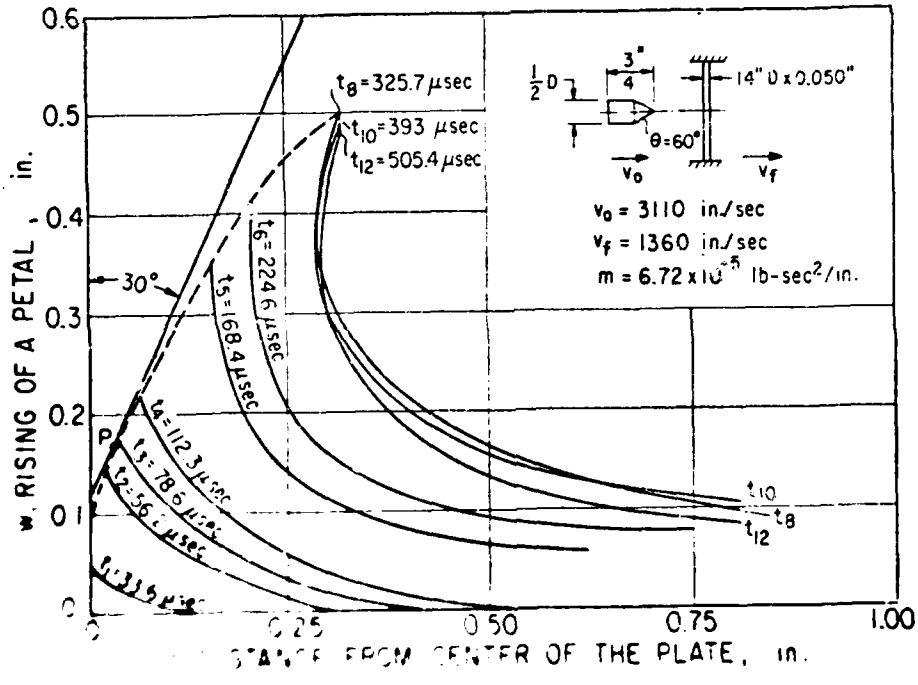


Fig. 7 Contour History of a 0.050-in. thick 2024-0 Aluminum Plate Clamped on a 7 in. Radius during Perforation by a ½-in. Dia. Steel Sphere



...-0 Aluminum Plate Clamped

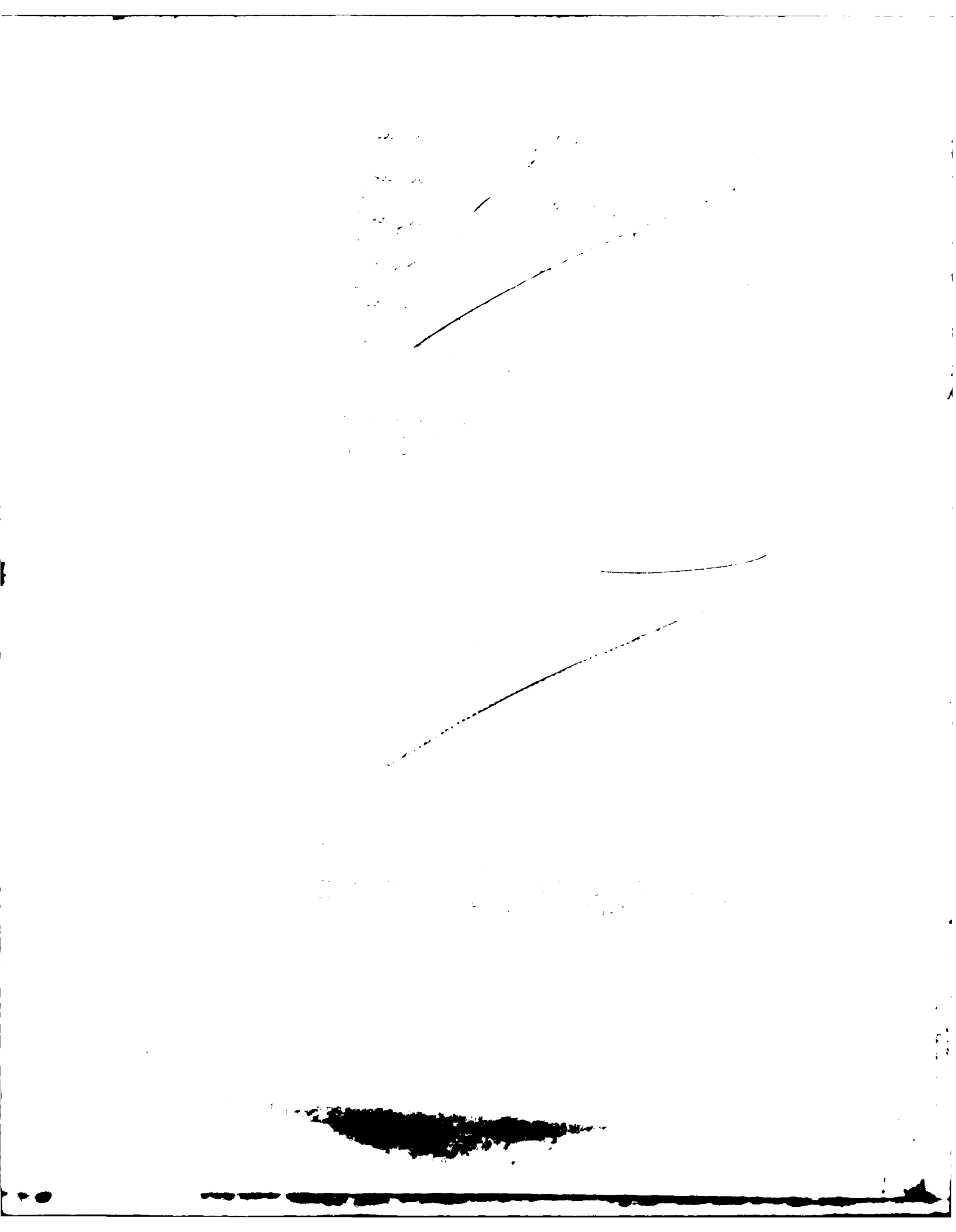


Fig. 11 Sections of ¼-in. thick Steel Plates Struck by a ½-in. Dia. Hard-Steel Sphere at a Velocity of about 2900 ft/s. (a) SAE 1020 steel plate (b) SAE 4130 steel plate

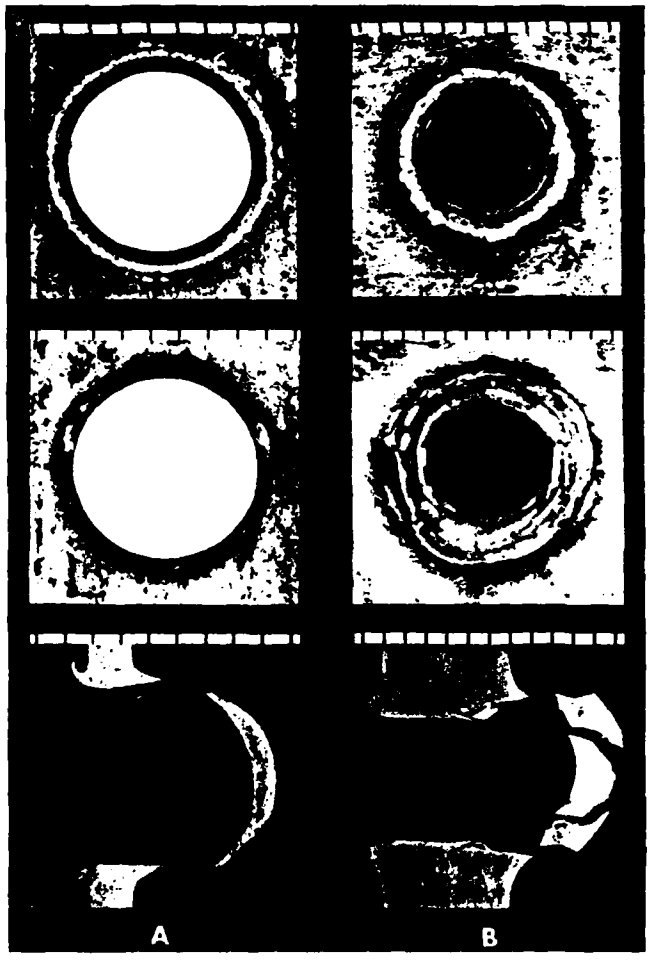


Fig. 12 Impact and Exit Side Photographs and Sections of Targets Perforated by a ¼-in. Dia. Hard-Steel Sphere at a Velocity of about 2800 ft/s. (a) 0.062-in. thick SAE steel plate (b) ¼-in. thick 2024-T4 aluminum plate

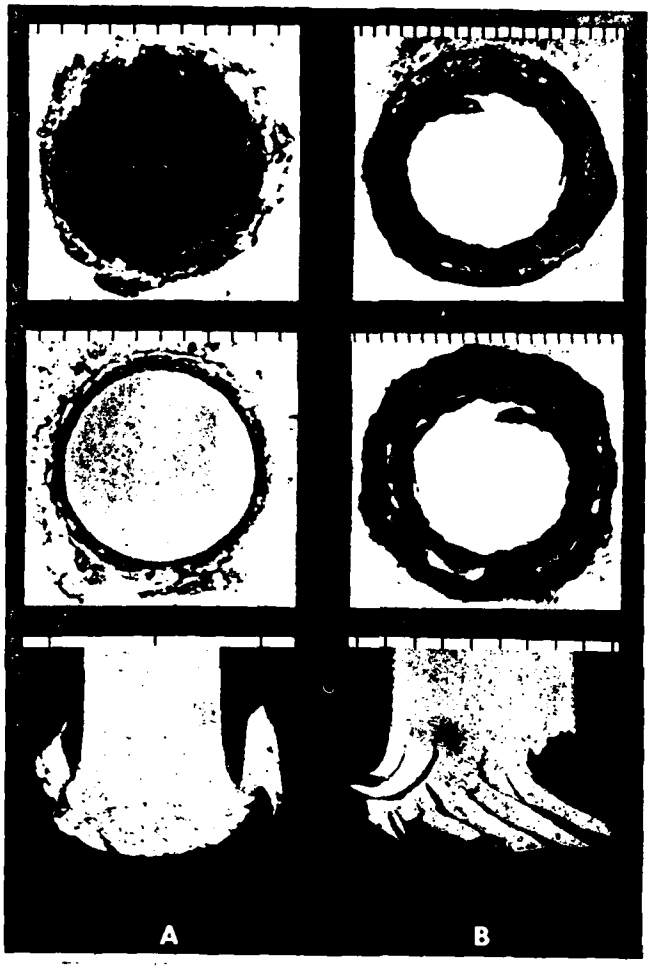
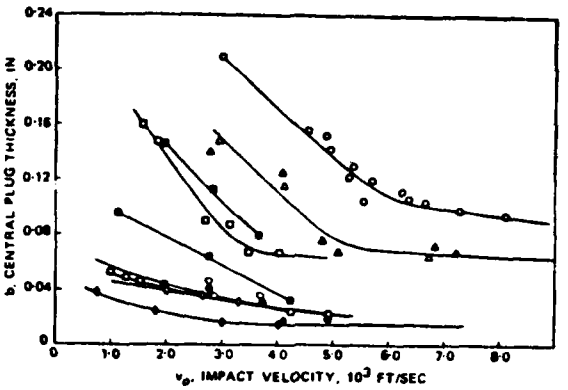
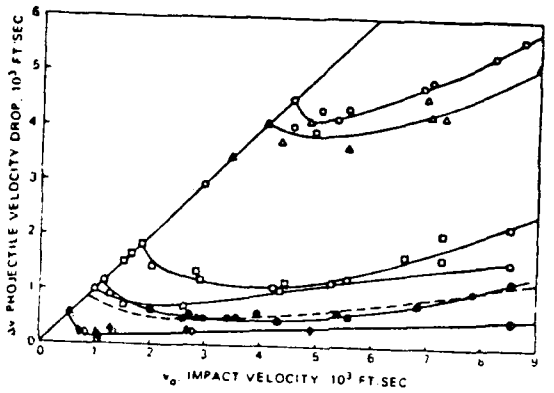


Fig. 13 Impact and Exit Side Photographs and Sections of 0.25-in thick Targets Perforated by a ½-in. Dia. Hard-Steel Sph... striking at a Velocity of about 8600 ft/s. (a) 2024-T4 aluminum plate (b) SAE 4130 steel plate



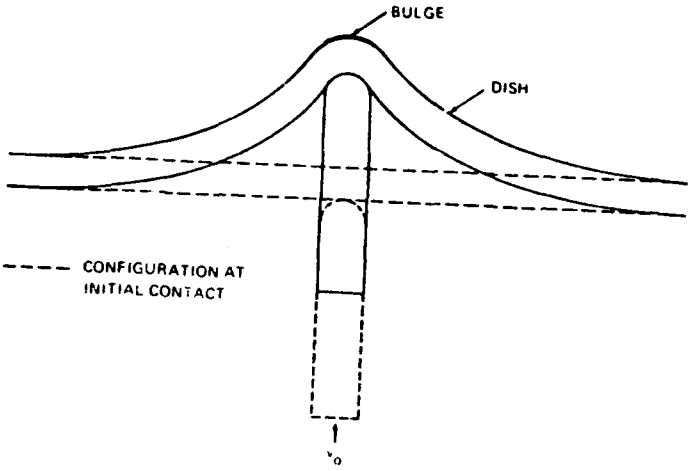
SYMBOL	TARGET THICKNESS, IN	TARGET MATERIAL	STEEL PROJECTILE DIAMETER, IN
○	0·250	4130 STEEL ARMOR	0·250
△	0·250	1020 STEEL (LARGE GRAINED)	0·250
▲	0·250	1020 STEEL (SMALL GRAINED)	0·250
□	0·250	2024-T4 ALUMINUM	0·250
■	0·250	2024-T4 ALUMINUM	0·375
▮	0·125	2024-T3 ALUMINUM	0·250
⌂	0·062	1020 STEEL (SMALL GRAINED)	0·250
●	0·062	1020 STEEL (SMALL GRAINED)	0·375
◊	0·050	2024-T3 ALUMINUM	0·250
◆	0·050	2024-0 ALUMINUM	0·250
○	0·050	2024-T3 ALUMINUM	0·375

Fig. 14 Measured Central Thickness of Plugs Separated from Target Plates by Projectile Perforation as a Function of Impact Velocity

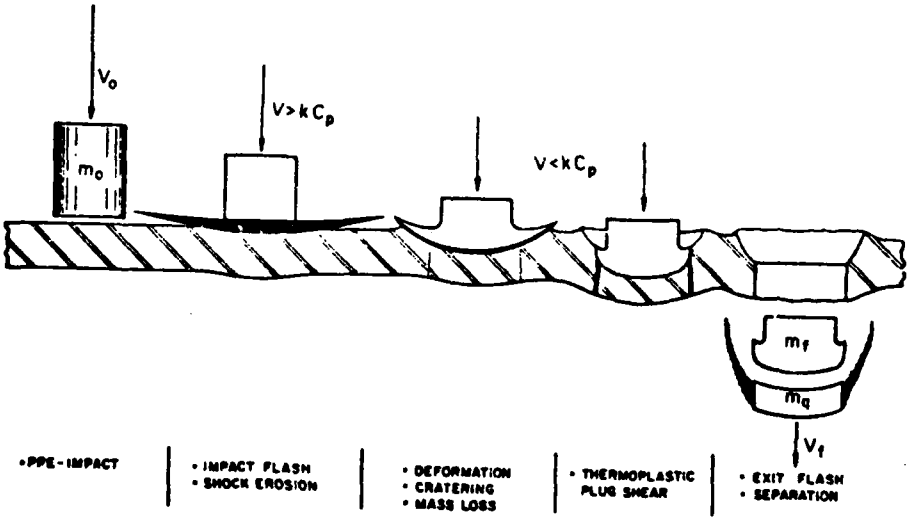


SYMBOL	TARGET THICKNESS IN	MATERIAL	STEEL PROJECTILE DIAMETER IN
○	0.250	4130 STEEL ARMOR	0.250
△	0.250	1020 STEEL (LARGE GRAINED)	0.250
□	0.250	2024 T4 ALUMINUM	0.250
▢	0.125	2024 T3 ALUMINUM	0.250
■	0.062	1020 STEEL (SMALL GRAINED)	0.250
●	0.062	1020 STEEL (SMALL GRAINED)	0.375
◇	0.050	2024 T3 ALUMINUM	0.250
◆	0.050	2024 0 ALUMINUM	0.250
◆₁	FROM REF 1		

Fig. 15 Projectile Velocity Drop during Perforation as a Function of Initial Velocity for various Targets



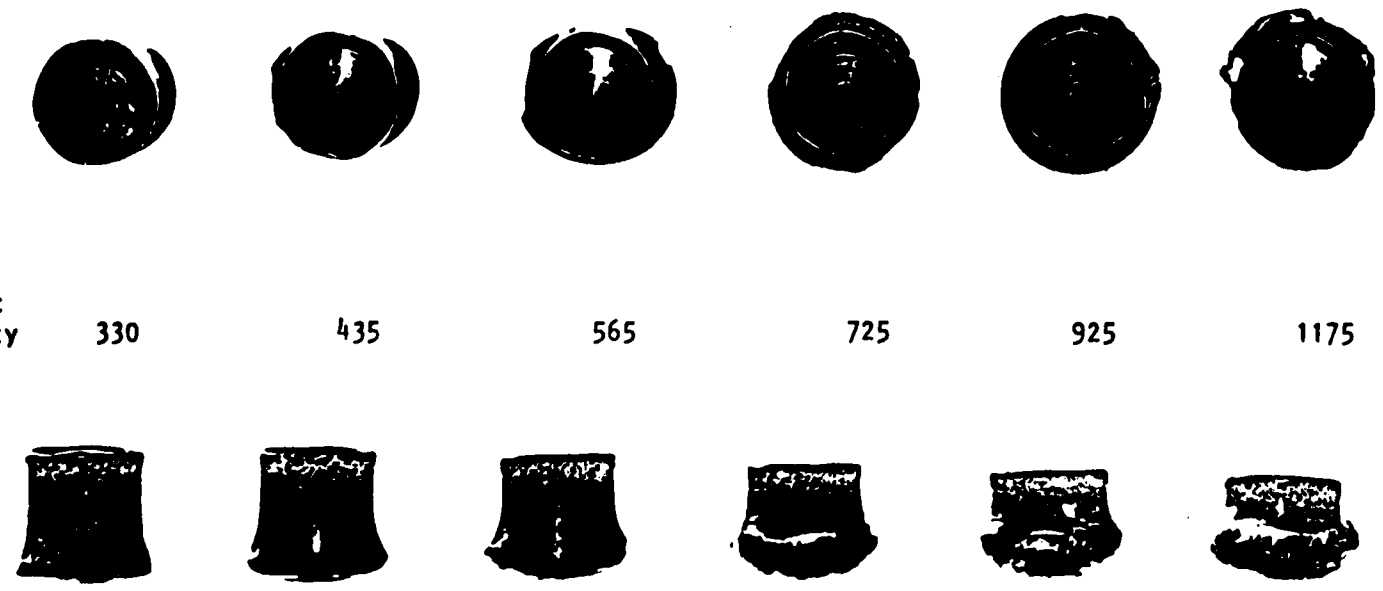
(a)



(b)

Fig. 16 Target Deformation due to Normal Impact of a Blunt-nosed Projectile (a) Bulging and dishing (b) Plugging and striker distortion (Ref. 4)

Fig. 17 Deformation of 17 grain, 0.22 Caliber Blunt Cylindrical Projectiles due to Perforation of ¼-in. thick Mild Steel Plates as a Function of Initial Velocity (Courtesy, R. F. Recht)



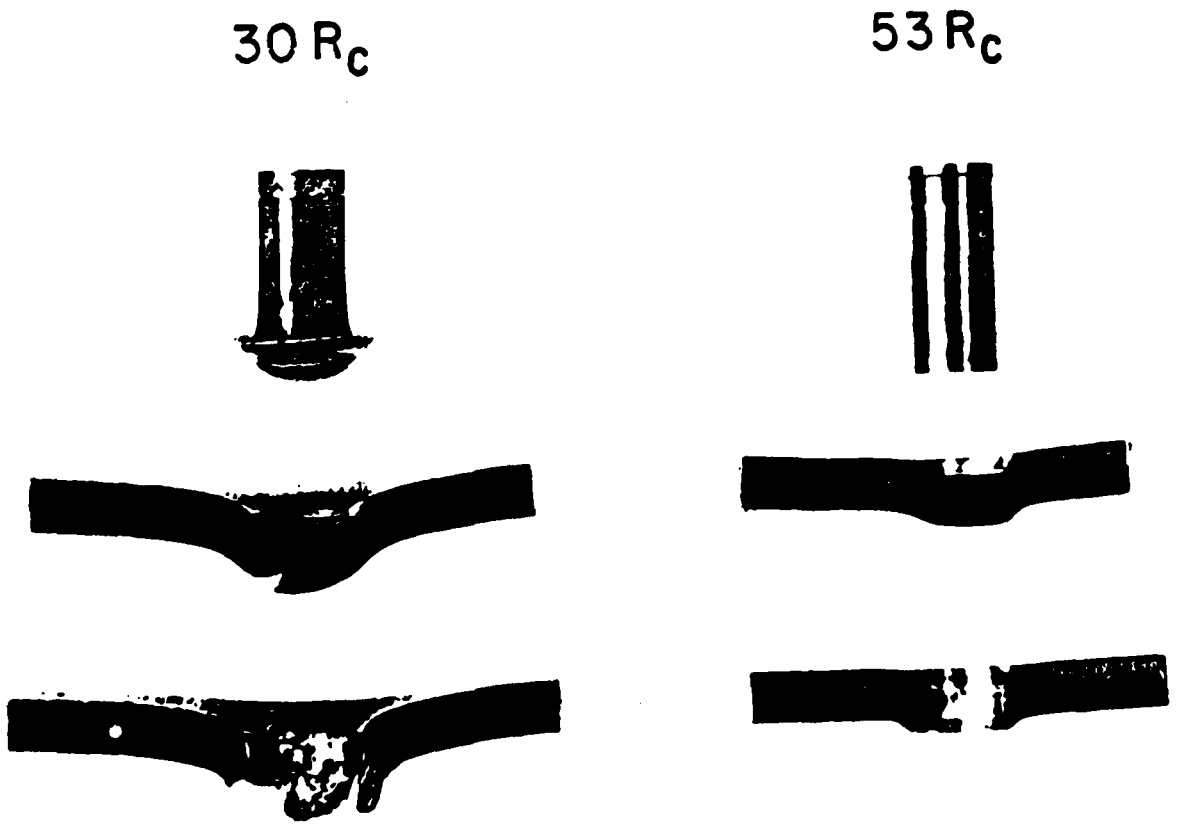


Fig. 18 Effects of Penetrator Hardness on Deformation Patterns in Normal Impact of Blunt-nosed Projectiles (Courtesy, R. F. Recht)

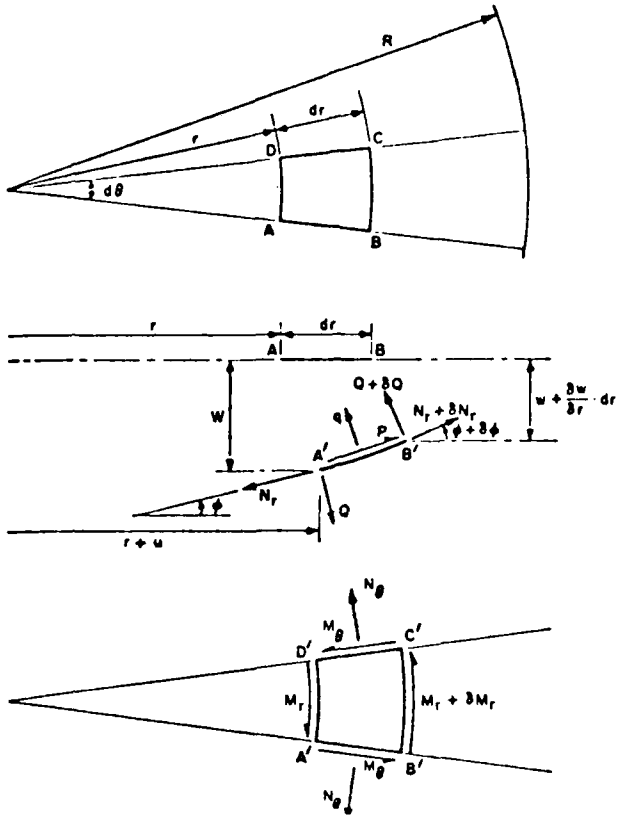


Fig. 19 Elements of Plastic Plate Deformation and Corresponding Nomenclature

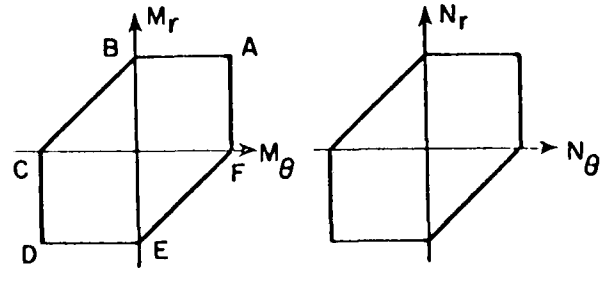


Fig. 20 Tresca Yield Criterion

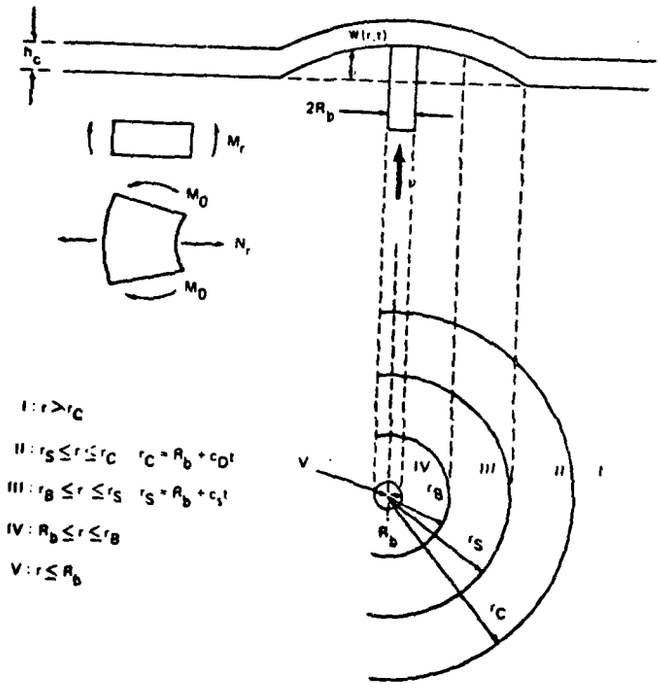


Fig. 21 Regions of Differing Response for an Elastic-perfectly Plastic Plate due to Waves Generated by the Normal Impact of a Flat-nosed Non-perforating Projectile

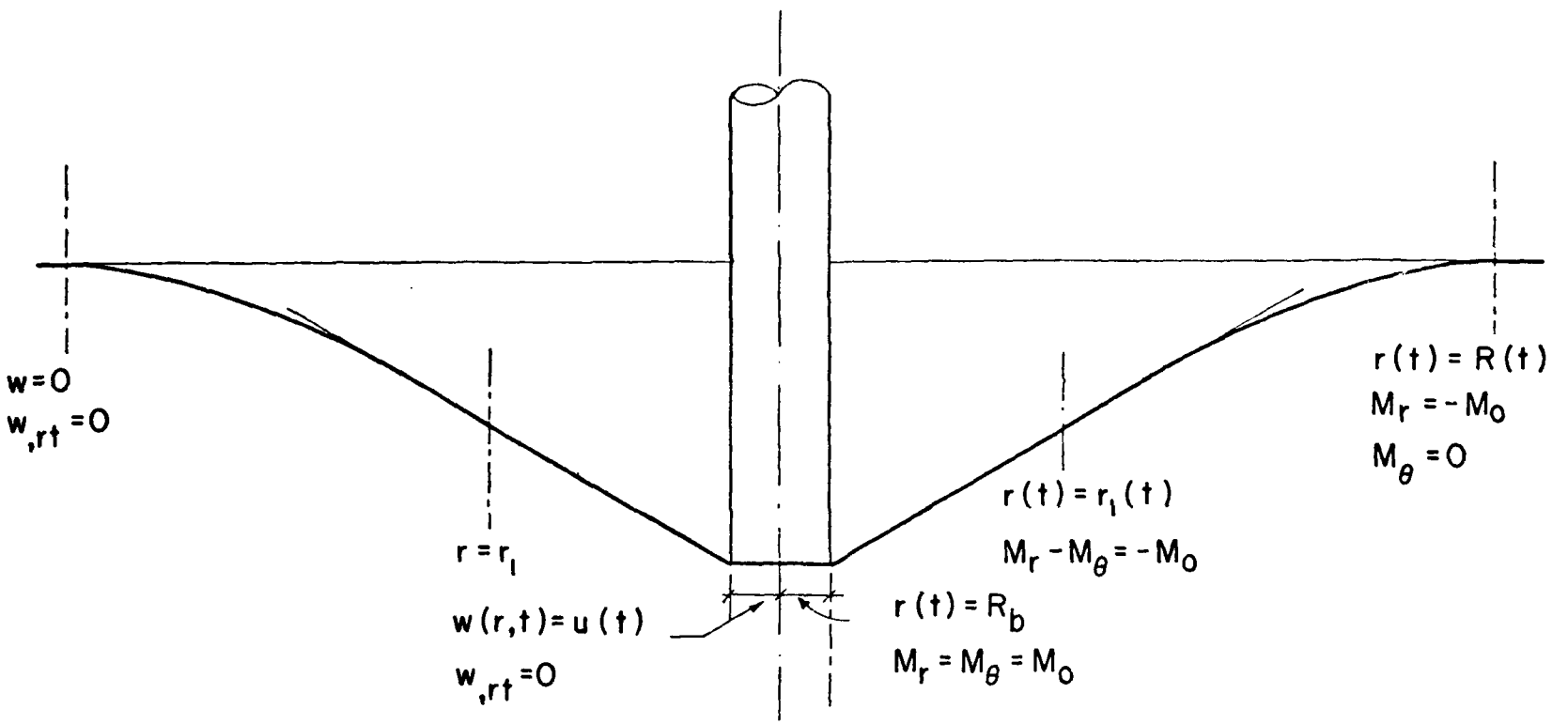


Fig. 22 Bending of a Perfectly-plastic Plate under Normal Impact of a Hard Flat-ended Projectile

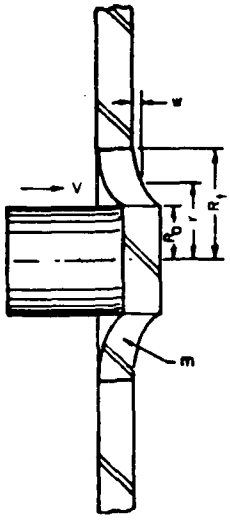


Fig. 23 Plastic Plate Deformation Model due to Non-perforating Normal Impact of a Flat-nosed Projectile

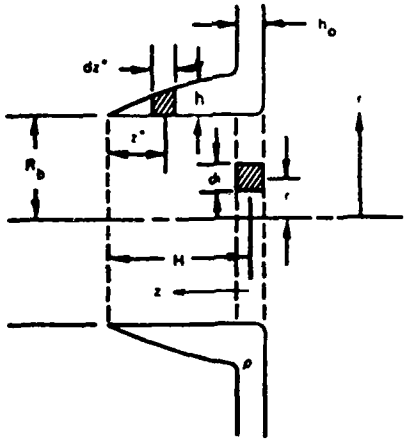
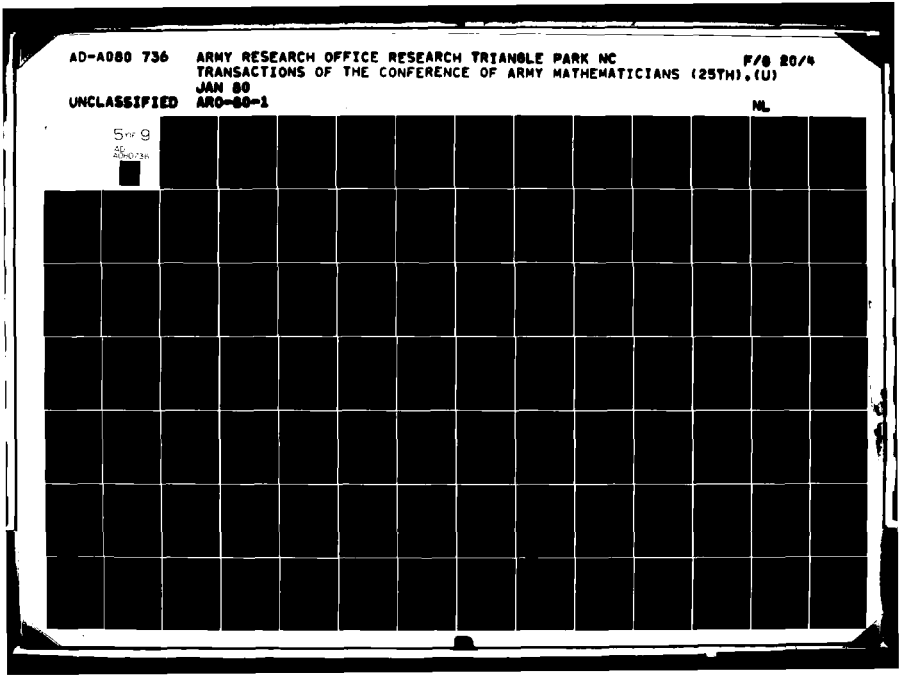


Fig. 24 Assumed Antisymmetric Deformation Pattern for the Perforation of a Thin Target by a Projectile Striking at Normal Incidence Modelled as a Ductile Hole Enlargement


AD-A080 736
ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC
F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED
ARO-80-1
NL
5 OF 9
AD-A080736

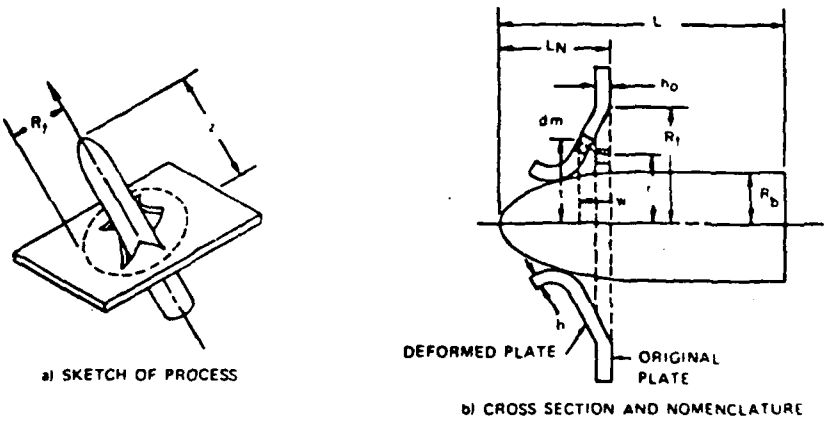


Fig. 25 Petalling Model for Thin Plates due to Normal Impact of Sharp-nosed Projectiles

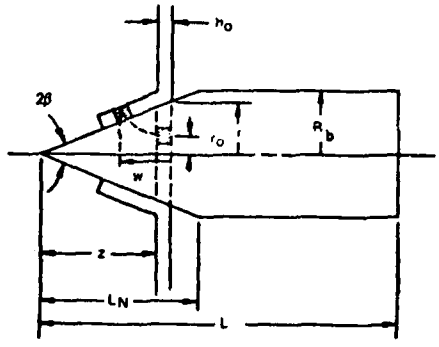


Fig. 26 Assumed Petalling Deformation Pattern for a Conical-nosed Striker

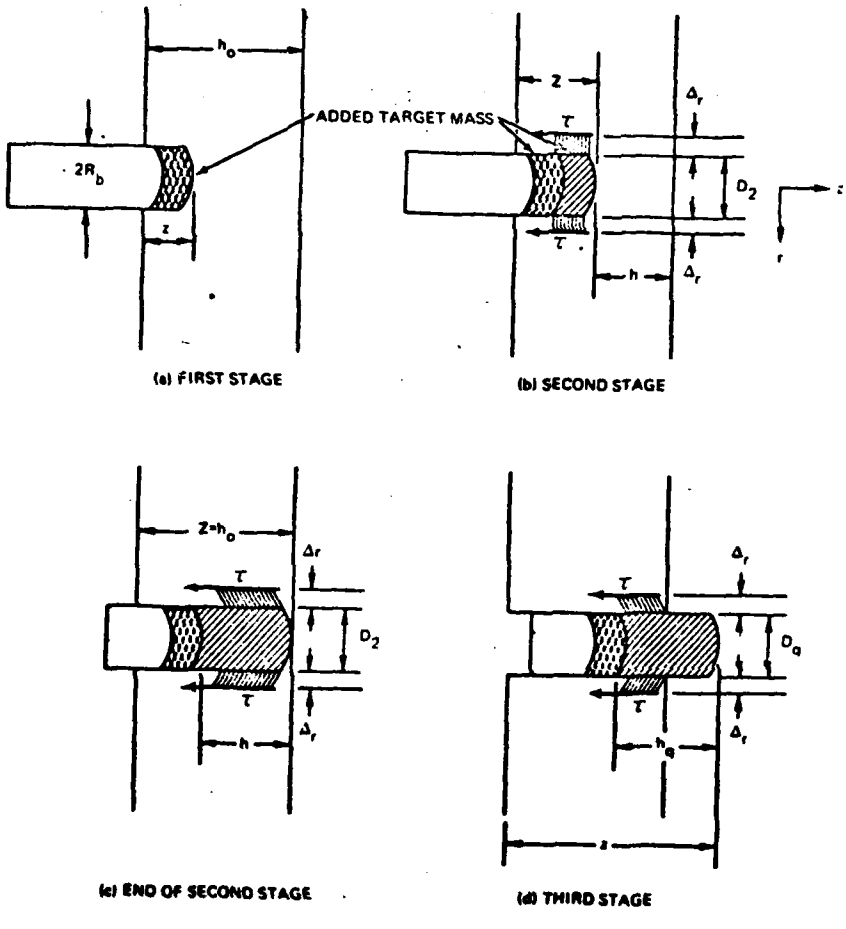


Fig. 27 Three Consecutive Stages of a Model of Plugging Deformation
(a) Compression (b) Compression and shear (c) Shear only

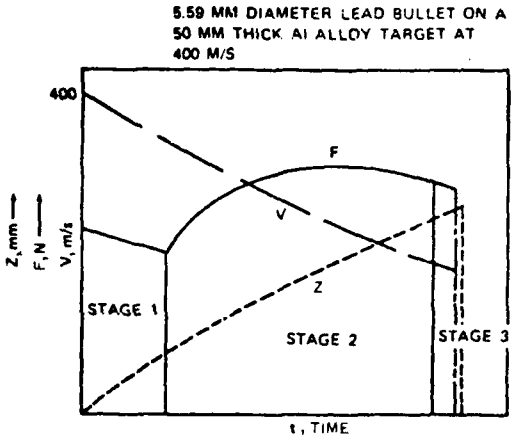


Fig. 28 Typical Kinematic and Kinetic Histories for the 3-Stage Plugging Model

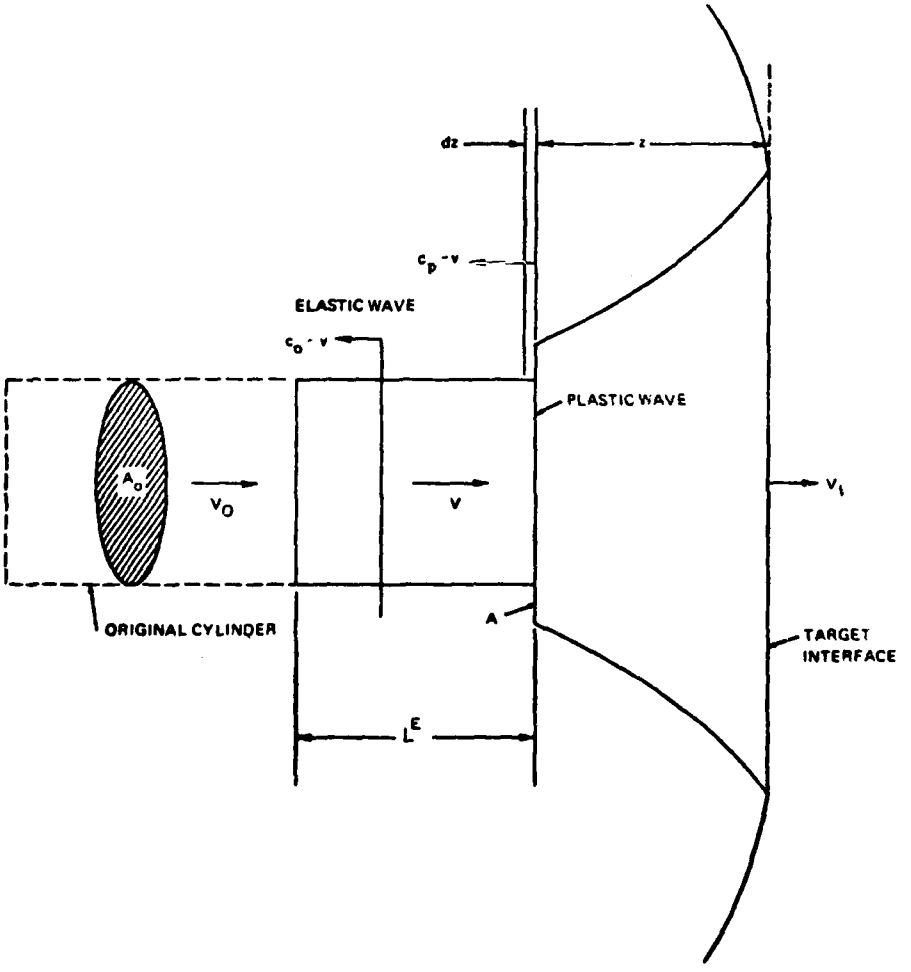


Fig. 29 Projectile Deformation Model when Striking a Deformable Half-space

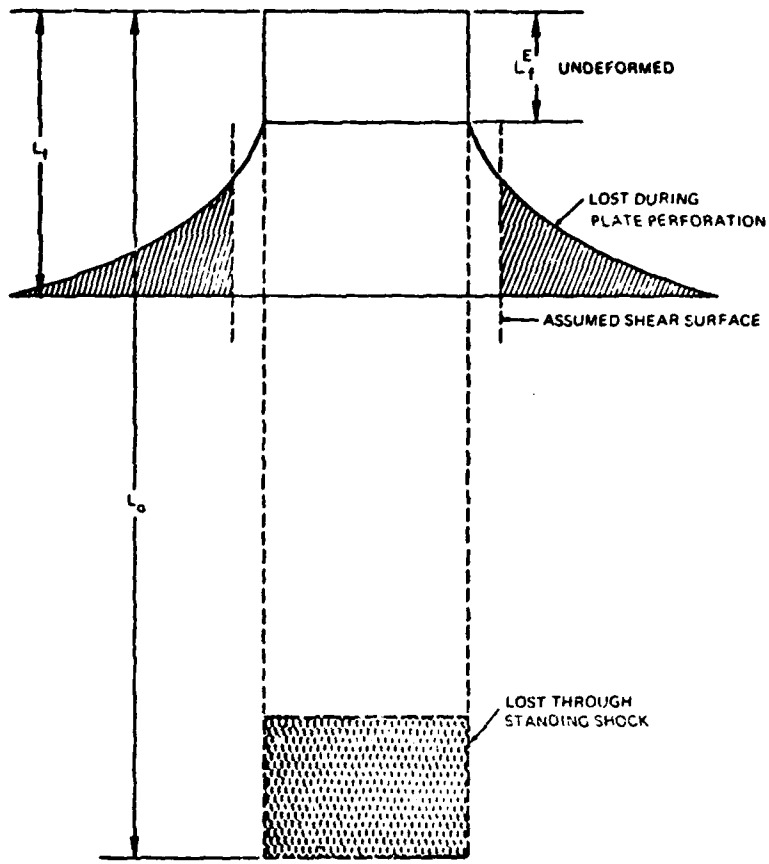


Fig. 30 Projectile Deformation and Mass Loss Model for Target Plugging Failure Mechanism

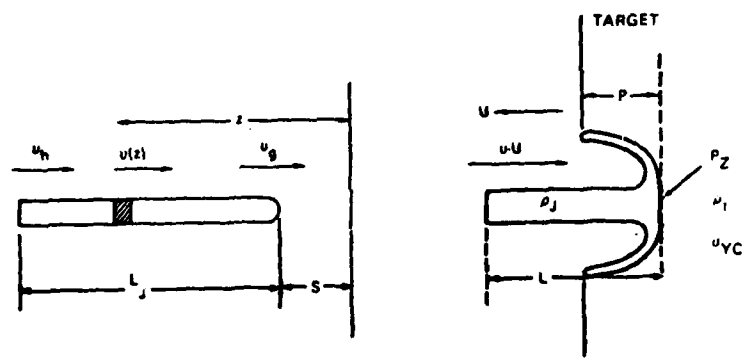


Fig. 31 Penetration of a Target by a High-speed Jet

PROPAGATING VELOCITY OF SINGULARITY OCCURRING IN CERTAIN DEGENERATE PARABOLIC EQUATIONS

Yoshisuke Nakano

U.S. Army Cold Regions Research and
Engineering Laboratory, Hanover, N.H. 03755

ABSTRACT. It has been shown that the Cauchy problems for certain degenerate parabolic equations describing flow through porous media may not have classical solutions and that the singularities occurring in the solutions may be interpreted as propagating acceleration waves. The propagating velocity of such singularities is studied base upon the theory of acceleration waves and the validity of such a viewpoint is examined for explicit solutions to the Cauchy problems.

I. INTRODUCTION. Consider the Cauchy problem

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x}\left(m\theta^{m-1}\frac{\partial \theta}{\partial x}\right) \qquad \text{in } S = R^1 \times (0, \infty) \tag{1}$$

with the initial condition

$$\theta(x, 0) = \theta_o(x) \qquad \text{in } R^1 \tag{1a}$$

where R^1 denotes the one dimensional Euclidean space and $m \geq 1$ is a constant. θ_o is a given bounded continuous non-negative function that satisfies the condition

$$\theta_o > 0 \text{ on } I, \text{ otherwise } \theta_{io} = 0 \tag{1b}$$

where $I = (a_1, a_2)$ is a bounded open interval in R^1.

Eq. 1 describes infiltration of water into a dry porous medium where θ is the volumetric water content and mass flux of water q is given as

$$q = \rho\theta v = -\rho\,\partial\theta^m/\partial x \tag{2}$$

where v is the velocity of water relative to the solid phase and ρ is the density of water. An interface between a wet part and a dry part of the medium is called a wetting front.

Eq. 1 is a nonlinear equation which is parabolic for $\theta > 0$, but which degenerates when $\theta = 0$. Oleinik et al. [1958] have defined a class of weak solutions to problem (1), and have proved the existence and uniqueness of solutions in that class. They have shown that if θ_o has compact support then θ also has compact support in R^1 for each $t > 0$. This implies that the wetting fronts propagate with a finite speed. Moreover, in the neighborhood of any point of S where θ is positive, θ satisfies Eq. 1 in the classical sense. In general, the transition from a region where $\theta > 0$ to one where $\theta = 0$ is not smooth and it is therefore necessary to use the term "weak solution".

The purpose of the present paper is to present the conditions that determine the velocity of wetting fronts and serve also as the moving boundary conditions at the wetting fronts based upon the theory of singular surfaces and acceleration waves [Truesdell and Toupin, 1963].

II. ACCELERATION WAVES. We assume that the flow is in one direction only. We examine the condition in the neighborhood of a wetting front that divides the whole region V into a part V^+ where water is present and a part V^- where water is absent. Let $x = \xi(t)$ be the location of the wetting front. Then the condition of mass balance at the wetting front is given as

$$[\![\rho\theta(v - u)]\!] = 0 \qquad \text{on } x = \xi(t) \tag{3}$$

where $u \neq 0$ is the propagating velocity of the wetting front and a double bracket is defined as

$$[\![A]\!] = A^+ - A^- \tag{4a}$$

$$A^+ = \lim_{x \to \xi} A, \qquad x \text{ in } V^+ \tag{4b}$$

$$A^- = \lim_{x \to \xi} A, \qquad x \text{ in } V^- \tag{4c}$$

Since $v^- = 0$ Eq. 3 reduces to

$$\rho^+ \theta^+ (v^+ - u) = 0 \qquad \text{on } x = \xi(t) \tag{5}$$

It follows from Eq. 5 that we have either

$$\rho^+ \theta^+ = 0 \tag{6a}$$

or

$$v^+ = u \tag{6b}$$

In terms of flux q, Eq. 5 may be written as

$$q^+ = \rho^+ \theta^+ u \tag{7}$$

From Eq. 7 it is clear that Eq. 6a is a necessary and sufficient condition for $q^+ = 0$ as long as u remains finite, which is the case in physical problems. On the other hand, Eq. 6b does not necessarily imply that $q^+ \neq 0$ unless $\rho^+ \theta^+ \neq 0$. Thus we have two possible cases: Case 1, $\rho^+ \theta^+ = 0$, then $q^+ = 0$. Case 2 $\rho^+ \theta^+ \neq 0$, then $u = v^+$, and $q^+ \neq 0$.

In Case 2 ρ, θ, v and q are all discontinuous at $x = \xi(t)$, and the wetting front can be interpreted as a shock wave. Since reported experimental data have not confirmed the actual occurrence of such shock waves, we exclude this case from our discussion and concentrate on Case 1.

There are several ways to define acceleration. However we prefer to define it in terms of q for the sake of convenience. In the case where q is continuous, but q_x is not at the wetting front, then u is given as

$$u = - \lim_{x \to \xi} \left(\frac{q_t}{q_x}\right), \qquad x \text{ in } V^+ \tag{8}$$

Eq. 8 also serves a moving boundary condition at the wetting front in this case and was obtained by Nakano [1978, 1979]. q_t stands for acceleration in terms of flux q and the wetting front can be interpreted as a propagating acceleration wave of the first order.

As it will be shown later, Eq. 8 does not apply to all the cases studied here since singularities of yet higher order occur. In general, the propagating velocity of accleration waves of the nth order is given as [Truesdell and Toupin, 1963]

$$u = - \lim_{x \to \xi} \frac{Q^n{}_t}{Q^n{}_x}, \qquad x \text{ in } V^+ \tag{9}$$

where Q^n is defined in V^+ as

$$Q^n = \underbrace{q_{x,\ x,\ x,\ \ldots,\ x}}_{n-1} \tag{9a}$$

or

$$Q^1 = q$$

$$Q^2 = q_x$$

$$Q^3 = q_{x,x} \quad \text{etc.}$$

It should be mentioned that Eq. 9 is not the only way to determine the propagating velocity. In the situation which we are considering here, u might be equal to v^+. Also since $\rho\theta$ is continuous, if $(\rho\theta)_x$ is not continuous, then u can be determined by

$$u = \lim_{x \to \xi} - \frac{(\rho\theta)_t}{(\rho\theta)_x}, \qquad x \text{ in } V^+ \tag{10}$$

Although Eq. 9 is not a unique way to determine u, Eq. 9 should be able to determine u for all possible cases. We will show that this is the case in the following section.

III. VELOCITY OF WETTING FRONTS IN PARTICULAR SOLUTIONS. We examine Eq. 9 by applying it to two kinds of explicit solutions of Eq. 1 that were originally derived by Barenblatt [1952, 1953].

Solution 1

Eq. (1) has a particular solution:

$$\theta(x, t) = \begin{cases} \{\lambda \frac{m-1}{m} (\lambda t - x + B)\}^{\frac{1}{m-1}}, & 0 \leq x \leq \lambda t + B \\ 0 & x > \lambda t + B \end{cases} \tag{11}$$

where $t \geq t_o > 0$, λ is a positive constant and B a real number. The location of the wetting front is given as

$$x = \xi(t) = \lambda t + B \tag{11a}$$

It follows from Eq. 11a that the wetting front travels with a constant velocity λ.

We examine the regularity properties of Eq. 11. From Eq. 11 we obtain in V^+

$$q = \rho\lambda\theta \tag{12a}$$

$$q_x = -\frac{\rho}{m}\lambda^2\theta^{2-m} \tag{12b}$$

$$q_t = \frac{\rho}{m}\lambda^3\theta^{2-m} \tag{12c}$$

$$q_{x,x} = \frac{(2-m)}{m^2}\rho\lambda^3\theta^{3-2m} \tag{12d}$$

$$q_{x,t} = -\frac{(2-m)}{m^2}\rho\lambda^4\theta^{3-2m} \tag{12e}$$

$$q_{x,x,x} = -\frac{(2-m)(3-2n)}{m^3}\rho\lambda^4\theta^{4-3m} \tag{12f}$$

etc.

If $m \geq 2$, although q is continuous, but q_x is discontinous at the wetting front. Thus the wetting front is an acceleration wave of the first order. The velocity u is determined by Eq. 9 as

$$u = -\lim_{x\to\xi-}\frac{Q_t^1}{Q_x^1} = -\lim_{x\to\xi-}\frac{q_t}{q_x} = \lambda \tag{13}$$

If $2 > m \geq \frac{3}{2}$, although q and q_x are continuous, but $q_{x,x}$ is discontinuous at the wetting front. Therefore the wetting front is an acceleration wave of the second order. The velocity u is determined by Eq. 9 as

$$u = -\lim_{x\to\xi-}\frac{Q_t^2}{Q_x^2} = -\lim_{x\to\xi-}\frac{q_{xt}}{q_{xx}} = \lambda \tag{14}$$

Generally if $\frac{n}{n-1} > m \geq \frac{n+1}{n}$ where n is an integer, the wetting front is an acceleration wave of the nth order and the velocity u is determined by Eq. 9.

Solution 2

Eq. 1 has a particular solution:

$$\theta(x, t) = \begin{cases} \beta\, t^{\frac{-1}{m+1}} (\eta_0^2 - \eta^2)^{\frac{1}{m-1}} & 0 \leq \eta \leq \eta_0 \\ 0 & \eta > \eta_0 \end{cases} \tag{15}$$

whre $t \geq t_o$,

$$\eta = x/t^{1/(m+1)} \tag{15a}$$

$$\beta = \left(\frac{m - 1}{2m(m + 1)}\right)^{\frac{1}{m-1}} \tag{15b}$$

and

$$\eta_0 = \left\{2\left(\frac{2m(m + 1)}{m - 1}\right)^{\frac{1}{m-1}} \Gamma\left(\frac{1}{2} + \frac{m}{m-1}\right) \{\Gamma\left(\frac{m}{m-1}\right) \Gamma\left(\frac{1}{2}\right)\}^{-1}\right\}^{\frac{m-1}{m+1}} \tag{15c}$$

where Γ is the gamma function defined as

$$\Gamma(r) = \int_o^\infty t^{r-1} e^{-t}\, dt \qquad r > 0 \tag{15d}$$

We examine the regularity properties of Eq. 15. From Eq. 15 we obtain in V^+

$$q = \alpha\, m \rho\, t^{-1}\, x\, \theta \tag{16a}$$

$$q_x = \alpha\, m \rho\, t^{-1}\, (\theta - \alpha t^{-1}\, x^2\, \theta^{2-m}) \tag{16b}$$

$$q_t = \frac{\alpha\, m \rho}{m+1}\, t^{-2} x\, [\alpha t^{-1}\, x^2\, \theta^{2-m} - (m+2)\theta] \tag{16c}$$

etc.

where $\alpha = 2\beta^{m-1}/(m-1)$

If $m \geq 2$, although q is continuous, but q_x is discontinuous at the wetting front. Therefore the wetting front is an acceleration wave of the first order. The velocity u is determined by Eq. 9 as

$$u = - \lim_{\eta \to \eta_0^-} \frac{Q_t^1}{Q_x^1} = \frac{\eta_o}{m+1}\, t^{-m/m+1)} \tag{17}$$

generally if $\frac{n}{n-1} > m \geq \frac{n+1}{n}$, where n is an integer, the wetting front is an acceleration wave of the nth order and the velocity u is determined by Eq. 9.

IV. CONCLUDING REMARKS. We have presented the conditions that govern the propagating velocity of a wetting front and that serve also as the moving boundary conditions at the wetting front. We have shown that such conditions hold true for reported particular solutions in the literature.

We have found that the wetting front is generally an acceleration wave of the nth order, where n is an interger. Since unfortunately there exists no general and easy method to determine the order of these acceleration waves without knowing the exact form of solutions, the definite determination of the order has to rely upon strictly mathematical analysis case by case. It is possible that a physical law which might determine the order will be found in the future. Further research is needed to understand the true nature of wetting fronts.

V. REFERENCES

1. Barenblatt, G.I. 'On certain non-steady movement of liquid and gas in porous media'. Prikl. Mat. Mekh., Vol. 16, 67-78, 1952.

2. Barenblatt, G.I. 'On a class of exact solutions to one-dimensional problems of non-steady filtration of gas in porous media'. Prikl. Mat. Mekh., Vol. 17, 739-742, 1953.

3. Oleinik, O.A., A.S. Kalashinikov, and Chzhou Yui-Lin, The Cauchy problem and boundary problems for equations of the type of non-stationary infiltration. Izv. Akad. Nauk.SSSR Ser. Mat. 22, pp. 667-704, 1958.

4. Nakano, Y. Theory and numerical analysis of moving boundary problems in the hydrodynamics of porous media. Water Resour. Res., 14 (1) 125-134, 1978.

5. Nakano, Y. Some recent results from functional analysis contrary to the traditional viewpoint on wetting fronts. A manuscript accepted for publication in Water Resour. Res., 1979.

6. Truesdell, C. and R.A. Toupin. The classical field theories, Handbuch der Physik III/1, pp. 226-793, particularly p. 523, Section 190. Springer-Verlag, Berlin. 1963.

A MINIMUM PRINCIPLE FOR SUPERHARMONIC FUNCTIONS SUBJECT TO INTERFACE CONDITIONS*

Bernard A. Fleishman
Rensselaer Polytechnic Institute
Troy, New York 12181

Thomas J. Mahar
Northwestern University
Evanston, Illinois 60201

ABSTRACT. Let D be a bounded domain in R^2 with smooth boundary. Let $B_1,\ldots,B_m$ be non-intersecting smooth Jordan curves contained in D, and let D' denote the complement of $\bigcup_{i=i}^{m} B_i$ with respect to D. Suppose that $u \in C^2(D') \cap C(\bar{D})$ and $\Delta u \leq 0$ in D' (where Δ is the Laplacian), while across each "interface" B_i, i = 1,...,m, there is "continuity of flux" (as suggested by the theory of heat conduction). It is proved here that the presence of the interfaces does not alter the conclusions of the classical minimum principle (for $\Delta u \leq 0$ in D). The result is extended in several regards. Also it is applied to an elliptic free boundary problem and to the proof of uniqueness for steady-state heat conduction in a composite medium. Finally this minimum principle (which assumes "continuity of flux") is compared with one due to Collatz and Werner which employs an alternative interface condition.

1. INTRODUCTION. To prove a minimum principle in a domain with interfaces (or internal boundaries) we shall make repeated use of the classical result. Let us therefore state the classical minimum principle for functions satisfying $\Delta u \leq 0$ (so-called superharmonic functions), where Δ is the Laplacian operator.

*Sponsored by U.S. Army Research Office under Contract No. DAAG29-79-C-0012.

CLASSICAL MINIMUM PRINCIPLE (CMP). Suppose D is a bounded domain in R^n, with smooth (for example, C^2) boundary B. If $u \in C^2(D) \cap C(\bar{D})$ and $\Delta u \leq 0$ in D, then $\min_{\bar{D}} u$ (which we denote by μ) is assumed on the boundary B; it is assumed in D only if $u \equiv \mu$ in $\bar{D}$. Furthermore, when $u \not\equiv \mu$, at a point of B where $u = \mu$ the exterior normal derivative of u, u_ν, is negative, where $\vec{\nu}$ denotes the outward-directed unit normal.

Of course u has a minimum value, μ, in $\bar{D}$ because u is continuous in $\bar{D}$.

Recently, in the course of investigating some free boundary problems for nonlinear elliptic equations, we found that we needed a minimum principle when D contains internal boundaries on which Δu is not defined, but across which certain interface conditions hold. We prove such a minimum principle here.

Results of this type have appeared in the literature. Oleinik [6] discusses a maximum principle for elliptic problems with interfaces, but requires the equation to contain a non-homogeneous term. Our result below does not have such a requirement. Littman [3] develops a generalized maximum principle for smooth equations which have adjoints. Problems with interfaces do not appear to be covered by this result. Rubinstein [8] studies existence and uniqueness of solutions to free boundary problems for the Laplace equation; the interface conditions, however, differ from those we use below in that he specifies the values of the dependent variable on the interfaces.

Our minimum principle is formulated and proved in Section 2. Some extensions are described in Section 3, and in Sections 4 and 5

we present two applications. The first deals with the uniqueness of a steady-state temperature distribution in a composite medium, the second with a simple diffusion-reaction equation containing a discontinuous reaction term. Finally, in Section 6 we compare our interface condition ("continuity of flux") on normal derivatives with an alternative condition used by Collatz [1] and Meyn and Werner [4].

2. A MINIMUM PRINCIPLE. For ease of exposition the result is formulated and proved for $n = 2$; the minor modification required for $n > 2$ is described in the next section.

Let $B_1,\ldots,B_{m+1}$ be non-intersecting smooth Jordan curves in R^2 such that for $i = 1,\ldots,m$,

$$B_i \subset \text{int } B_{m+1} = D.$$

$B_1,\ldots,B_m$ are the interfaces. Note that now D is a simply-connected domain.

The complement of $\bigcup_{i=1}^{m} B_i$ with respect to D, $D' = D/\bigcup_{i=1}^{m} B_i$, may also be written

$$D' = \bigcup_{i=1}^{m+1} A_i ,$$

where $A_1,\ldots,A_{m+1}$ are the disjoint subdomains into which $B_1,\ldots,B_m$ divide D, and A_i is the one immediately interior to B_i (see Figure 1 for illustration).

In order to introduce a "continuity of flux" condition (suggested by heat conduction) to hold across the interfaces $B_1,\ldots,B_m$, we define in $\bar{D}$ a positive-valued, piecewise-continuous function k such that $k = k_i(x,y)$ $(i = 1,\ldots, m+1)$ is continuous

in A_i and may be extended continuously to $\bar{A}_i$. As in Section 1, at a point P of any B_i (i=1,...,m+1), let $\vec{\nu}$ denote the unit normal directed out of A_i, and u_ν the corresponding normal derivative of u at P. We now formulate the interface condition, after which the minimum principle is stated and proved.

CONTINUITY-OF-FLUX (COF) CONDITION. At every point of B_i (i=1,...,m) ku_ν is continuous across B_i.

MINIMUM PRINCIPLE. Suppose that $u \in C^2(D') \cap C(\bar{D})$, that $\Delta u \leq 0$ in D', that in A_i (i=1,...,m+1) u_x and u_y may be extended continuously up to the boundary, and that COF holds. Then (i) $\mu = \min_{\bar{D}} u$ is assumed on B_{m+1}; (ii) μ is assumed in D only if $u \equiv \mu$ in $\bar{D}$; and (iii) in case $u \not\equiv \mu$, at a point of B_{m+1} where $u = \mu$ we have $u_\nu < 0$.

In other words, when the interface conditions are continuity of u and continuity of flux the presence of interfaces leaves unchanged the conclusions of the classical minimum principle

PROOF. Either $u \equiv \mu$ in $\bar{D}$ or not. In the first case, (i) holds trivially. To show that (i) holds also in case $u \not\equiv \mu$, suppose (ii) is true; then $u \not\equiv \mu$ implies that $u \neq \mu$ in D, therefore $u = \mu$ at a point of B_{m+1}. Also to prove (iii) it is enough to know that (ii) holds; for then, as indicated in the argument below, $u \not\equiv \mu$ in $\bar{D}$ implies $u \not\equiv \mu$ in A_{m+1}, while $u = \mu$ somewhere on B_{m+1}. Then the classical minimum principle (CMP) applied to A_{m+1} yields result (iii).

Thus, to complete the proof it suffices to show that (ii) is true. Suppose $u = \mu$ somewhere in D; this may occur at interface

points, non-interface points, or both. If $u = \mu$ in some A_i, then from $\Delta u \leq 0$ in A_i and $u \in C(\bar{A}_i)$, CMP implies that $u \equiv \mu$ in $\bar{A}_i$ and therefore on the boundary of A_i, which includes at least one $B_j \neq B_{m+1}$. Thus, if $u = \mu$ at a non-interface point in D then $u = \mu$ at some interface point.

Suppose then that $u = \mu$ at a point Q of some B_i $(i=1,\ldots,m)$. We show that in this case $u \equiv \mu$ in both subdomains of D' bordering B_i (A_i immediately interior to B_i and, say, A_j immediately exterior to B_i).

Let $(u_\nu)_i$ and $(u_\nu)_j$ denote the limiting values of u_ν (at Q on B_i) from the interiors of A_i and A_j, respectively. Since by definition $\vec{\nu}$ is directed exterior to A_i (therefore interior to A_j) $(u_\nu)_j$ represents a normal derivative interior to A_j. Now COF may be expressed in the form

$$k_i(u_\nu)_i = k_j(u_\nu)_j \ . \tag{1}$$

Also, applying CMP to u in A_i and A_j yields

$$(u_\nu)_i \leq 0 \text{ and } -(u_\nu)_j \leq 0 \tag{2}$$

respectively.

From the positivity of k_i and k_j, it follows from (1) that $(u_\nu)_i$ and $(u_\nu)_j$ have the same sign. This is consistent with (2), however, only if

$$(u_\nu)_i = (u_\nu)_j = 0. \tag{3}$$

If now $u \not\equiv \mu$ in A_i, CMP implies $(u_\nu)_i < 0$ at Q, contradicting (3).

Thus, $u = \mu$ at a point Q of B_i $(i=1,\ldots,m)$ implies $u \equiv \mu$ in A_i and, similarly, $u \equiv \mu$ in A_j, the domain immediately exterior to

B_i. Now $u = \mu$ also on the other boundaries of A_i and A_j, the argument may be repeated (a finite number of times), and we conclude that $u \equiv \mu$ in $\bar{D}$. This completes the proof of part (ii) and therefore the entire theorem.

3. EXTENSIONS. We give several extensions of the minimum principle just proved.

1) As in the classical case, when the sense of the inequality is changed from $\Delta u \leq 0$ to $\Delta u \geq 0$, the minimum principle is replaced by a maximum principle.

2) To obtain a minimum principle in R^n, $n > 2$, the interface curves $B_1,\ldots,B_{m+1}$ must be replaced by appropriate surfaces. Specifically we want each B_i to be a closed surface which separates R^n into two disjoint domains, an (unbounded) exterior and a (bounded) interior. Thus, if the $B_1,\ldots,B_{m+1}$ are non-intersecting "Jordan manifolds" which are C^2, and therefore possess the interior-ball property (see [5], p. 7), we shall have the minimum principle in R^n.

3) As in the classical case minimum (or maximum) principles for more general elliptic operators are obtainable in the interface case.

EXAMPLE. Let the hypotheses of the minimum principle in Section 2 be unaltered except that instead of $\Delta u \leq 0$, we assume

$$Lu + fu \leq 0 \text{ in } D' ,$$

where L is a uniformly elliptic operator of the form

$$Lu \equiv au_{xx} + bu_{xy} + cu_{yy} + du_x + eu_y ,$$

a,b,c,d,e,f are functions continuous in each $\bar{A}_i$ (i=1,...,m+1), and $f < 0$ in D.

The classical results for a function u satisfying $Lu + fu \leq 0$ in D under these conditions (e.g., see [7]) are that u can not assume a negative minimum in D, and that $u_\nu < 0$ at a point of the boundary where the minimum occurs, unless $u \equiv$ constant. (Since $f < 0$, $u \equiv$ constant < 0 is no longer a possibility.) Because this is the situation in each A_i, an argument like that used in the proof yields a similar result for D in the interface case.

4) Minimum and maximum principles are also obtainable for parabolic inequalities in the presence of interfaces, for example, by arguments like those used above for elliptic inequalities. In fact much of the work on parabolic free boundary problems makes use of maximum principles in one form or another. We shall not pursue this here; for references to the extensive literature on the subject the interested reader is referred to [8] and [9].

4. APPLICATION: UNIQUENESS RESULT FOR STEADY-STATE HEAT CONDUCTION IN A COMPOSITE MEDIUM. Let D be a two-dimensional region divided into sub-domains A_i by curves B_i, as in Section 2. Let $f_i(x,y)$ represent heat sources (or sinks) in A_i, and k_i the constant conductivity of region A_i. If u(x,y) denotes the temperature at the point (x,y), then $k_i \Delta u = f_i(x,y)$ in D', and u is specified on B_{m+1}, the outer boundary of D. The continuity of flux interface condition will hold if there are no heat sources or sinks distributed along the interior curves B_i.

To derive a uniqueness result for such a linear Poisson interface problem, we show that $u \equiv 0$ on B_{m+1} and $\Delta u = 0$ in D' imply $u \equiv 0$ in D. Since $\Delta u = 0$ in D', both $\Delta u \leq 0$ and $\Delta u \geq 0$ hold in D'. Combining the maximum and minimum principles, we conclude that u attains its extreme values on B_{m+1}. As $u \equiv 0$ on B_{m+1}, we have $u \equiv 0$ in D, thus proving the uniqueness theorem.

Note that a similar proof shows that solutions to problems of this type depend continuously on the boundary data specified along B_{m+1}.

5. APPLICATION: SOLUTION BY ITERATION OF A SIMPLE DIFFUSION-REACTION PROBLEM WITH DISCONTINUOUS REACTION TERM. Consider the following boundary value problem in $D = \{0 \leq r < 1\}$.

$$P(\varepsilon): \begin{cases} \Delta u + H(u-\mu) = 0 & \text{in } D/\Gamma \\ u(1,\theta) = \varepsilon h(\theta) \ , & 0 \leq \theta \leq 2\pi. \end{cases}$$

Here H denotes the Heaviside step function ($= 0$ for $u < \mu$, $= 1$ for $u \geq \mu$), $\mu > 0$ a given constant; Γ is the set of points in D (not known a priori) where $u = \mu$; $\varepsilon \geq 0$ is a parameter; and h is a given function, continuous, periodic with period 2π, and satisfying $0 < h(\theta) < 1$.

$P(\varepsilon)$ may be regarded as governing the steady states of a simple reaction-diffusion system in which the reaction rate changes abruptly when the state variable u reaches the triggering value μ.

Suppose $\varepsilon < \mu$. Then if $u > \mu$ somewhere in D, there will be one or more interfaces in D across which $H(u-\mu)$ changes discontinuously; in this case $P(\varepsilon)$ is a (nonlinear) free boundary problem (FBP), the solution of which requires also the determination of

the interface(s) (defined by $u = \mu$). If $u < \mu$ throughout D, $H(u-\mu) \equiv 0$ and $P(\varepsilon)$ reduces to a linear Dirichlet problem for Laplace's equation.

In [2] we have used an iterative method to establish the existence of a solution to the FBP $P(\varepsilon)$, and we wish to indicate here how the minimum principle proved in Section 2 may be applied to show that the iterates form a monotone sequence.

Consider the "reduced problem" $P(0)$, with boundary condition $u(1,\theta) \equiv 0$. If we restrict ourselves to symmetric solutions $u = u(r)$, $P(0)$ takes the form

$$(ru')' + rH(u-\mu) = 0 , \qquad 0 < r < 1 , \tag{4}$$

$$u'(0) = u(1) = 0 . \tag{5}$$

Suppose a C^1 solution exists for which $u(0) > \mu$. From (4), $(ru')' \leq 0$. But because $(ru')' \not\equiv 0$, u is strictly decreasing on the interval (0,1), so that $u(r_o) = \mu$ has exactly one root r_o in (0,1). By solving $(ru')' = -r$ on $(0,r_o)$ and $(ru')' = 0$ on $(r_o,1)$ subject to (5), then requiring u and u' to be continuous at $r = r_o$ (and also $u(r_o) = \mu$), we find the following (see [2]).

<u>If</u> $\mu < 1/4e$, <u>the</u> BVP (4-5) <u>has two</u> C^1 <u>solutions of the form</u>

$$u_o(r) = \begin{cases} -\frac{r_o^2}{2} \ln r_o + \frac{r_o^2}{4} - \frac{r^2}{4} , & 0 \leq r \leq r_o \qquad (6) \\ -\frac{r_o^2}{2} \ln r , & r_o \leq r \leq 1 \qquad (7) \end{cases}$$

<u>each corresponding to a root</u> $r_o^2 = \xi$ <u>of</u>

$$-\xi \ln \xi = 4\mu . \tag{8}$$

For fixed $\mu \in (0, 1/4e)$, equation (8) has two distinct roots ξ

in (0,1); to each of these roots $\xi = r_o^2$ corresponds a solution (6-7).

Let $u_o = u_o(r)$ denote the function (6-7) corresponding to the larger root $\xi = r_o^2$ of (8). u_o is chosen as the first term of an iterative sequence $u_o, u_1, u_2, \ldots$, in which $u_n(r,\theta)$, $n = 1,2,\ldots$ is defined as the (unique) solution of the linear Poisson interface problem

$$P_n: \begin{cases} \Delta u + H(u_{n-1}-\mu) = 0 & \text{in } D/\Gamma_{n-1}\ , \\ u(1,\theta) = \varepsilon h(\theta)\ , & 0 \le \theta \le 2\pi\ . \end{cases}$$

A solution $u(r,\theta)$ of the FBP $P(\varepsilon)$ is then sought as the limit of the sequence $\{u_n\}$.

By Γ_n is meant the set of points in D at which $u_n = \mu$; thus Γ_o is the circle $r = r_o$. It is not obvious that $\Gamma_1, \Gamma_2, \ldots$ are simple closed curves; the proof of this fact is part of the analysis in [2].

The minimum principle will be applied to the differences $(u_{n+1}-u_n)$, $n = 0,1,\ldots$. We note first that there is a unique C^1 solution $u_1(r,\theta)$ of the BVP P_1: $\{\Delta u+H(u_o-\mu) = 0$ in D/Γ_o, $u(1,\theta) = \varepsilon h(\theta)\}$. On the other hand, $u_o(r)$ is the unique solution of the BVP $\{\Delta u+H(u_o-\mu) = 0$ in D/Γ_o, $u(1,\theta) = 0\}$. Therefore

$$\Delta(u_1-u_o) = 0 \qquad \text{in } D/\Gamma_o\ ,$$

$$u_1(1,\theta) - u_o(1,\theta) = \varepsilon h(\theta) \ge \varepsilon\alpha > 0,\ 0 \le \theta \le 2\pi\ ,$$

where $\alpha = \min h(\theta) > 0$. It follows from our minimum principle that

$$u_1 - u_o \ge \varepsilon\alpha > 0 \text{ in } D.$$

Since $u_1 < \mu$ on $r = 1$ and $u_1(r,\theta) > u_o(r) > \mu$ for $0 \le r < r_o$, only in the annulus $r_o < r < 1$ are there points where $u_1 = \mu$. In fact, these points may be shown to lie in a thinner annulus, S_ε: $r_o < r < r_\varepsilon < 1$, where Γ_ε: $r = r_\varepsilon$ is the free boundary in the one-dimensional BVP

P_ε: $\{(ru')' + rH(u-\mu) = 0,\ 0 < r < 1;\ u(1,\theta) \equiv \varepsilon\}$.

In order to apply the minimum principle to the next difference, $u_2 - u_1$, we must know that the set $\Gamma_1 = \{(r,\theta): u_1(r,\theta) = \mu\}$ forms a smooth simple closed curve $r = r_1(\theta)$ (which clearly encircles Γ_o: $r = r_o$). This is established (see [2]) by a) showing that in S_ε $\partial u_1/\partial r < 0$, thus, that along any ray $\theta = \theta_c$ (constant), there is exactly one value of r at which $u_1(r,\theta_c) = \mu$; and then b) utilizing bounds on $\partial u_1/\partial r$ and $\partial u_1/\partial\theta$ to allow application of the implicit function theorem to prove the smooth connectedness of the points of Γ_1. (An integral representation of u_1 is used to obtain bounds on these derivatives of u_1.)

Again, there is a unique C^1 solution $u_2(r,\theta)$ of the linear Poisson interface problem P_2: $\{\Delta u + H(u_1-\mu) = 0$ in D/Γ_1; $u(1,\theta) = \varepsilon h(\theta)\}$. It follows that

$$\Delta(u_2-u_1) = -[H(u_1-\mu) - H(u_o-\mu)]$$
$$\le 0\ (\not\equiv 0) \text{ in } D/(\Gamma_o \cup \Gamma_1)\ ,$$
$$u_2(1,\theta) - u_1(1,\theta) = 0\ , \qquad 0 \le \theta \le 2\pi\ .$$

Note that $\Delta(u_2-u_1) \not\equiv 0$ because in the annular domain bounded by Γ_o and Γ_1, $S_1 = \{(r,\theta): r_o < r < r_1(\theta)\}$, $u_o < \mu < u_1$, so that

$$H(u_1-\mu) - H(u_o-\mu) = 1.$$

Application of the minimum principle then gives

$u_2 - u_1 > 0$ in D.

By arguments like the preceding, it may be shown that the points where $u_2 = \mu$ form a smooth simple closed curve Γ_2: $r = r_2(\theta)$, with $r_1(\theta) < r_2(\theta) < r_\varepsilon$.

Proceeding iteratively, we establish that there is a sequence of C^1 functions $\{u_n\}$ which is monotone and bounded:

$u_o(r) < u_1(r,\theta) < u_2(r,\theta) <\ldots< u_\varepsilon(r)$ in D, where $u_\varepsilon(r)$ is the larger solution of P_ε. Similarly, the interfaces Γ_n: $r = r_n(\theta)$, $n = 1,2,\ldots$, form a monotone and bounded sequence of smooth simple closed curves:

$$r_o < r_1(\theta) < r_2(\theta) <\ldots< r_\varepsilon < 1 , \quad 0 \leq \theta \leq 2\pi .$$

Finally, the respective limits of these sequences, $u(r,\theta)$ and Γ: $r = \bar{r}(\theta)$, may be seen to have appropriate regularity and to form a solution of the FBP $P(\varepsilon)$, as follows.

THEORFM. Suppose $\mu\varepsilon(0,1/4e)$. For $\varepsilon > 0$ small enough, the sequence $\{u_n(r,\theta)\}$ converges monotonically and uniformly to a limit $u(r,\theta)$, where $u_o(r) \leq u(r,\theta) \leq u_\varepsilon(r)$, and the sequence $\{r_n(\theta)\}$ to a (closed) limit curve Γ: $r = \bar{r}(\theta) \in C^1$, where $r_o \leq \bar{r}(\theta) \leq r_\varepsilon$. Then $u(r,\theta)$ is a solution of the BVP $P(\varepsilon)$, with free boundary Γ.

6. COMPARISON OF THE CONTINUITY OF FLUX WITH AN ALTERNATIVE INTERFACE CONDITION. In order to prove the minimum principle in Section 2 we have invoked two requirements to link a solution u across an interface: the continuity of u itself (which has not been stressed but should not be taken for granted) and the continuity of flux (COF). It is possible to assume alternative

interface conditions, depending on the applications one has in mind.

In particular, we wish to compare COF with a condition used by Collatz [1] and Meyn and Werner [4] to obtain maximum/minimum principles and monotonicity results for functions satisfying elliptic differential inequalities in regions with interfaces.

In terms of our notation ($\vec{\nu}$ representing the unit normal directed form A_i to A_j) the function satisfying $\Delta u \leq 0$ in D' is shown to take its minimum on the boundary of D when across the interfaces u is continuous and

$$(u_\nu)_j \leq (u_\nu)_i \quad . \tag{9}$$

A maximum principle holds when the sense of the inequality is reversed in both the differential inequality and the interface inequality (9). (Note that by contrast the COF interface condition is an equation, which may be used for both minimum and maximum principles.)

A simple geometric interpretation may be given for (9). With respect to the graph of u as a function of the normal variable ν, (9) says that when the interface is crossed the slope u_ν can not increase (see Figure 2). Indeed, if the slope decreases (discontinuously) the concave-down corner (see the figure) disallows a minimum value for u at the interface.

The COF condition and the Collatz-Werner condition (9) are alternative interface conditions (on the normal derivatives); either one, together with the continuity of u, is sufficient to yield a minimum principle (for $\Delta u \leq 0$ and similar elliptic

inequalities). COF neither implies nor is implied by (9). A dramatic illustration of this fact is that the Collatz-Werner condition is inadequate for treating the application in Section 4 (uniqueness for steady-state heat conduction in a composite medium), as we now show.

Recall that $\Delta u = 0$ (therefore both $\Delta u \leq 0$ and $\Delta u \geq 0$ hold) in each A_i. Now the Collatz-Werner interface condition for a minimum (maximum) principle is $(u_\nu)_j \leq (u_\nu)_i$ $((u_\nu)_j \geq (u_\nu)_i)$. Thus, to have both a minimum and maximum principle one would have to require

$$(u_\nu)_j = (u_\nu)_i \quad ,$$

continuity of the normal derivative, which is simply not the case when (if u represents temperature) heat transfer takes place between adjacent media with different conductivities.

In closing we remark that the COF condition is motivated by some important physical processes while the Collatz-Werner condition has a strong geometric motivation.

REFERENCES

1. Collatz, L., "Monotonicity with discontinuities in partial differential equations," Springer Lecture Notes 415 (A. Dold and B. Eckmann, eds.), 85-102 (1974).

2. Fleishman, B.A., and Mahar, T.J., "On the existence of classical solutions to an elliptic free boundary problem," Differential Equations and Applications (W. Eckhaus and E.M. deJager, eds.), 39-57 (1978).

3. Littman, W., "A strong maximum principle for weakly L-subharmonic functions," J. Math. & Mech. 8, 761-770 (1959).

4. Meyn, K.-H., and Werner, B., "Maximum and monotonicity principles for elliptic boundary value problems in partitioned domains," Institut für Angewandte Mathematik, Universität Hamburg (1978).

5. Miranda, C., Partial Differential Equations of Elliptic Type (2nd Revised Edition), Springer-Verlag, New York (1970).

6. Oleinik, O. A., "Boundary-value problems for linear elliptic and parabolic equations with discontinuous coefficients," A.M.S. Translations (Series 2) 42, 175-194 (1964).

7. Protter, M.H., and Weinberger, H.F., Maximum Principles in Differential Equations, Prentice-Hall, Englewood Cliffs (1967).

8. Rubinstein, L.I., The Stefan Problem (Translations of Math. Monographs 27), American Math. Society, Providence (1971).

9. Wilson, D.G., Solomon, A.D., and Boggs, P.T. (eds.), Moving Boundary Problems, Academic Press, New York (1978).

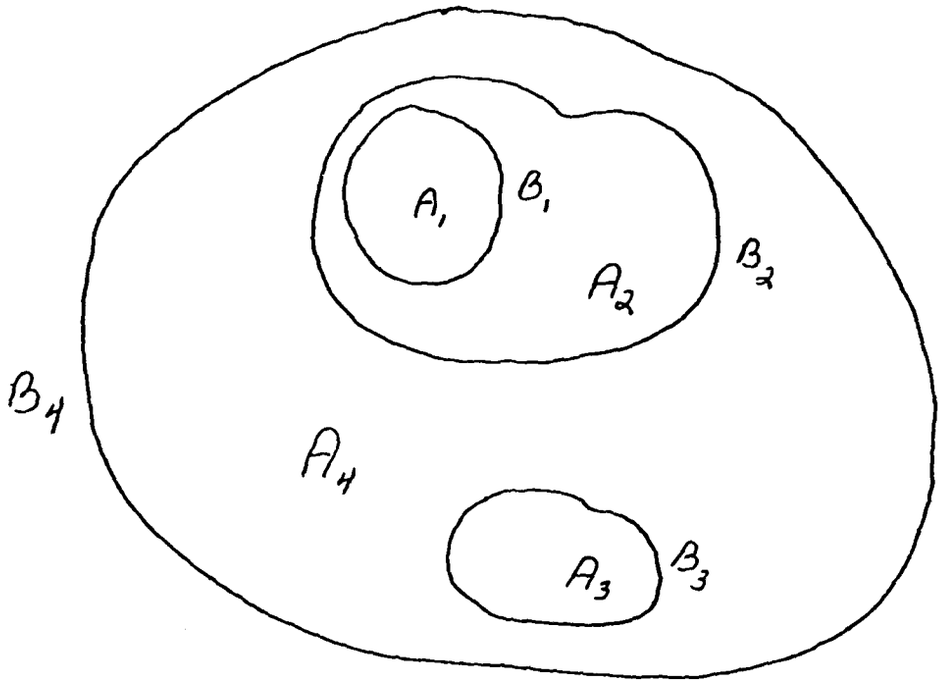


FIGURE 1 m = 3

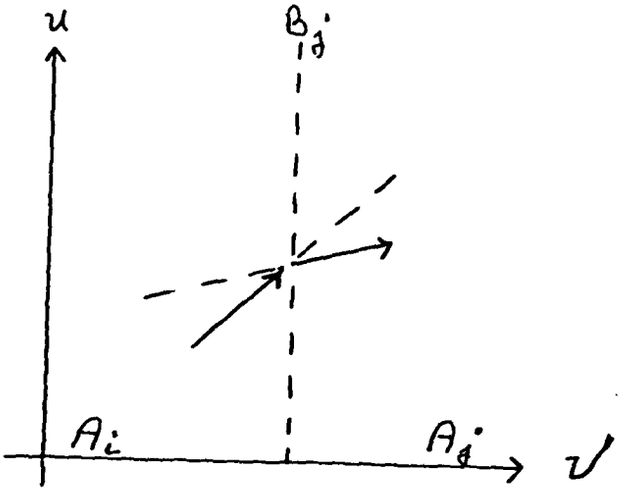


FIGURE 2

ANALYTIC MODEL FOR SHOCK WAVE PROPAGATION INTO CONCAVE CORNERS

James A. Schmitt
Ballistic Modeling Division
US Army ARRADCOM Ballistic Research Laboratory
Aberdeen Proving Ground, MD 21005

ABSTRACT. When a shock wave is incident upon a concave corner, multiple reflections occur and the pressure distribution on its walls are increased dramatically. A two-dimensional analytic model is described which, when no diffracted shocks occur, determines the exact flow field values within the corner, and otherwise, approximates the peak corner pressure. The solution is achieved by the repeated use of the oblique shock relations. An advantage of the model is that solutions involving both regular and Mach reflections can be obtained by algebraic means alone. Comparisons of the model with experimental shock tube data are given.

I. INTRODUCTION. When a shock wave propagates into a concave corner, it is reflected one or more times from the walls forming the corner. Upon reaching the corner, the direction of the shock propagation is reversed, one or more additional reflections may occur, and, in general, the last reflected shock is diffracted. These multiple reflections can cause tremendous increases in the pressure along the walls. Therefore, such corners are very susceptible to damage from blast waves.

When the propagation direction of the incident shock lies in the cross-sectional plane of a re-entrant corner of infinite width, a two-dimensional model of this phenomenon is appropriate. See Figure 1. The mathematical problem corresponding to this model with the additional assumption of infinitely long walls has been solved analytically in several special cases. Lighthill[1] considered an arbitrary strength shock propagating into a corner with an apex angle which deviated only slightly from 180°. Keller and Blank[2] considered weak shock waves (acoustic waves) propagating into any corner. Later, Keller[3] considered the special cases where no diffractions of the regular reflected shock waves occur and determined the exact solution by algebraic means. Schniffman et al[4] considered a series of re-entrant corner problems most of which involved corners formed at a right angle (some of these corners had one finite length wall). For corners formed at non-right angles, they considered only regular reflection within an infinitely long corner and used approximations to the oblique shock relations in order to obtain estimates of the resulting pressure field.

The purpose of this paper is to determine the peak pressure at the apex of a corner formed at a general angle and for an arbitrary strength shock. The only restriction on the corner angle and shock strength is that complex and double Mach reflections do not occur within the corner. Although, in concept, the present model may be extended to these cases, the model currently includes only regular and simple Mach reflections. Under the assumptions of the analytic model, the flow field within the corner can be analyzed as a cascading series of straight line shock reflections, except for possibly the final reflected shock. The model enables one to trace the propagation of all the shocks within the corner, to determine the type of reflection occuring at each reflection point within the corner and to calculate the gas and shock wave properties associated with each reflection. The flow field

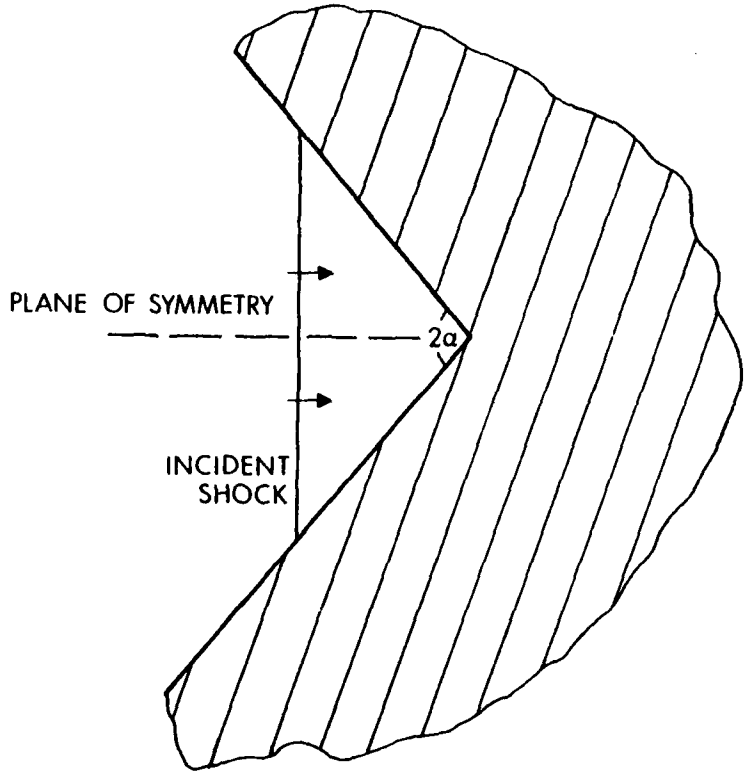


Figure 1. Schematic of Incident Shock in a Concave Corner of Infinite Width.

resulting from a shock wave propagating into an infinite two-dimensional corner can be solved algebraically provided that the final reflected shock is not diffracted as shown by Keller. However, if any shock is diffracted by either the leading edge of a finite corner or the final reflection process, no exact analytic treatment is possible. In these cases, an approximate technique (numerical or analytic) must be used. The present analytic analysis will provide an exact solution of the flow field if no diffraction occurs and an approximation of the apex pressure which is comparable with experimental results if diffraction occurs.

II. ASSUMPTIONS AND INITIAL CONDITIONS. The assumptions made in this analytic model are four:

a. The incident shock propagates with constant velocity and is symmetrically placed within the corner. See Figure 1. This hypothesis permits the analysis of a shock propagating perpendicularly to the plane of symmetry (a rigid wall) into a corner which has an acute angle equal to the bisected angle of the physical corner. This assumption, of course, can be ignored, if the incident shock is already propagating perpendiculary to a wall. Since the incident shock velocity is constant, its propagation can be considered pseudo-stationary if a frame of reference is attached to the shock.

b. We presume the medium in which the shock wave propagates is a perfect gas with neglible viscosity. The latter part of this assumption excludes the formation of boundary layers along the wall, and enables us to idealize shock waves as discontinuous surfaces. Following the derivation in Thompson[5] or Courant and Friedricks[6], the jump conditions across a planar discontinuity are:

$$\rho_b(\vec{v}_b - \vec{w}) \cdot \vec{n} - \rho_a(\vec{v}_a - \vec{w}) \cdot \vec{n} = 0, \tag{1}$$

$$\rho_b \left[(\vec{v}_b - \vec{w}) \cdot \vec{n}\right]^2 - \rho_a\left[(\vec{v}_a - \vec{w}) \cdot \vec{n}\right]^2 = P_a - P_b, \tag{2}$$

$$\vec{v}_b \cdot \vec{t} - \vec{v}_a \cdot \vec{t} = 0, \tag{3}$$

$$h_b + \tfrac{1}{2}(\vec{v}_b \cdot \vec{n})^2 - h_a - \tfrac{1}{2}(\vec{v}_a \cdot \vec{n})^2 = \vec{w} \cdot (\vec{v}_b - \vec{v}_a), \tag{4}$$

where ρ, $\vec{v}$, P, h, $\vec{w}$, $\vec{n}$, $\vec{t}$ are the gas density, gas velocity in laboratory coordinates, gas pressure, gas specific enthalpy, the shock wave velocity, the unit outer normal vector to the shock wave and the unit tangential vector to the shock wave, respectively. The gas properties immediately ahead of the shock wave are denoted by the subscript a and those immediately behind by the subscript b. The perfect gas postulate requires an equation of state of the form $h = \gamma P/\left[(\gamma-1)\rho\right]$ where γ is the ratio of two constants (the specific heat at constant pressure c_p and specific heat at constant volume c_v). The sound speed in a perfect gas is given by $a = (\gamma P/\rho)^{\frac{1}{2}}$. Equations (1) - (4) are commonly known as the oblique shock relations and are valid at any point Q on the shock. If the shock is a straight line shock in the immediate vicinity of Q, then the flow is uniform in this neighborhood and the flow properties computed by the oblique shock relations are also valid in this neighborhood.

c. We assume the walls forming the corner are rigid and infinite. The rigidity is a physical reality in most cases. The infinite extent of the walls eliminates the rarefaction wave which is generated at the leading edge of the corner and causes the curvature of some reflected shocks.

d. The fourth assumption is that only regular and simple Mach reflections occur within the corner. The restriction is necessary since, at present, only these types of reflections are modeled. The local conditions for the initiation and termination of regular and simple Mach reflections are stated. Thus the procedure of the analytic model will assign the type of reflection and the method to analyze it.

The initial conditions are the absolute pressure P_0 and temperature T_0 in the undisturbed medium, the incident shock strength and the angle of the apex 2α. From the initial pressure value and shock strength, the pressure behind the incident shock can easily be computed. With the assumption of a perfect gas, the initial density ρ_0 is given by the relation $\rho_0 = P_0/(T_0 R^*)$, where R^* is the gas constant.

III. REGULAR REFLECTION, MACH REFLECTION AND DISTINGUISHING CRITERIA. The theory of regular reflection from a solid boundary is well known[7,8]. Consider a plane shock wave I which is propagating with a constant velocity, is incident at point Q upon an infinite plane rigid wedge making an angle θ_w with the horizontal, and causes a regular reflected shock R to arise from the wedge. If we attach a frame of reference to the point Q, the incident shock velocity is zero and the flow in region 0 approaches I parallel to the wedge surface. See Figure 2. We define the region upstream of I as region 0, downstream of I and upstream of R as region 1 and downstream of R as region 2. We wish to relate the properties in regions 0, 1, and 2 in a neighborhood of the reflection point Q. While passing through the incident shock at an angle of $\phi_0 = 90° - \theta_w$, the flow is deflected towards I by an angle θ_1 from its original direction and its dynamic and thermodynamic properties are changed. These properties are related by the oblique shock relations (1) - (4) in the neighborhood of point Q. In these circumstances, the oblique shock relations can be simplified since $\vec{u}_a = \vec{v}_a - \vec{w}$, $\vec{u}_b = \vec{v}_b - \vec{w}$, $\vec{u}_a \cdot \vec{n} = u_0 \sin\phi_0$, $\vec{u}_b \cdot \vec{n} = u_1 \sin(\phi_0 - \theta_1)$, $\vec{u}_a \cdot \vec{t} = u_0 \cos\phi_0$ and $\vec{u}_b \cdot \vec{t} = u_1 \cos(\phi_0 - \theta_1)$, and can be rewritten as:

$$\rho_1 u_1 \sin(\phi_0 - \theta_1) = \rho_0 u_0 \sin\phi_0, \tag{5}$$

$$P_1 + \rho_1\left[u_1 \sin(\phi_0 - \theta_1)\right]^2 = P_0 + \rho_1\left[u_0 \sin\phi_0\right]^2, \tag{6}$$

$$u_1 \cos(\phi_0 - \theta_1) = u_0 \cos\phi_0, \tag{7}$$

$$h_1 + 0.5\left[u_1 \sin(\phi_0 - \theta_1)\right]^2 = h_0 + 0.5\left[u_0 \sin\phi_0\right]^2, \tag{8}$$

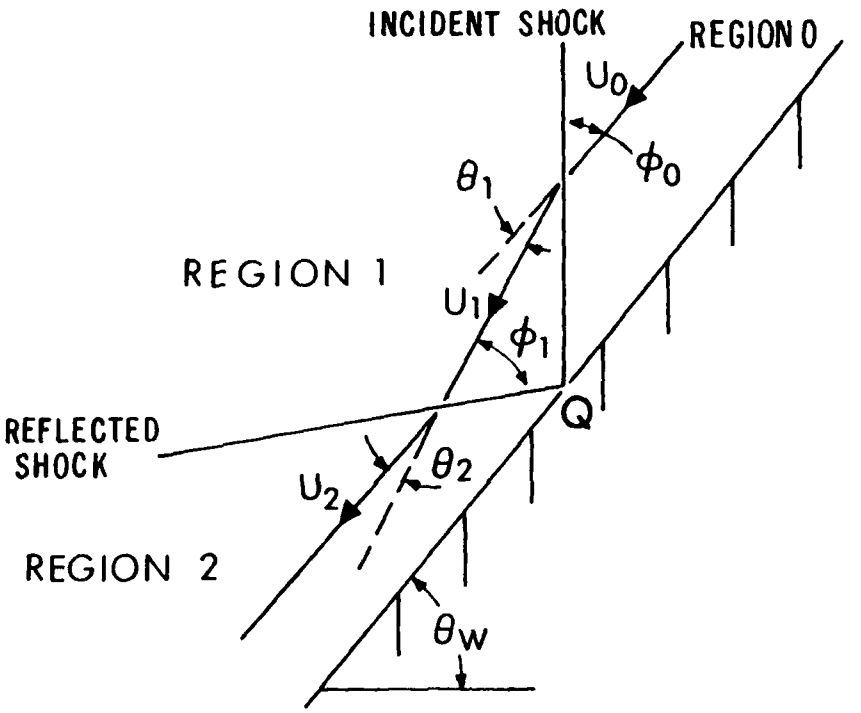


Figure 2. Schematic of Regular Reflection in a Neighborhood of the Reflection Point Q.

where we define $u_i = |\vec{u}_i|$. When the initial conditions of the analytic model (P_0, P_1, T_0, and θ_w) are given, the following quantities are also known $\rho_0 = P_0/(R^*T_0)$ and $h_0 = \gamma P_0/[(\gamma-1)\rho_0]$. Thus, equations (5) - (8) represent a system of four nonlinear algebraic equations in four unknowns, u_0, ρ_1, u_1 and θ_1, since $h_1 = \gamma P_1/[(\gamma-1)\rho_1]$. The solution of this system is obtained easily. The explicit formulas for the unknowns are derived in the Appendix. We note that the formulas in the Appendix are independent of the type of reflection occurring at the point Q.

The flow deflection across the incident shock causes the flow in region 1 in the neighborhood of Q to approach the reflected shock obliquely at an angle ϕ_1. While passing through the reflected shock, the flow is deflected towards R by an angle θ_2 from its region 1 trajectory and its dynamic and thermodynamic properties are altered. These properties are related by the oblique shock relations (1) - (4) in the neighborhood of Q. In this case, the velocities are $\vec{u}_a \cdot \vec{n} = u_1 \sin \phi_1$, $\vec{u}_b \cdot \vec{n} = u_2 \sin (\phi_1 - \theta_2)$, $\vec{u}_a \cdot \vec{t} = u_1 \cos \phi_1$ and $\vec{u}_b \cdot \vec{t} = u_2 \cos (\phi_1 - \theta_2)$. In order that the resulting flow in region 2 adjacent to the wall is parallel to the wall in the neighborhood of Q, the deflection angles must be equal, that is, $\theta_1 = \theta_2$. In this framework both θ_1 and θ_2 are positive angles and the difference in deflection direction is incorporated in the formulism. The oblique shock relations can be written in the form:

$$\rho_2 u_2 \sin (\phi_1 - \theta_1) = \rho_1 u_1 \sin \phi_1, \tag{9}$$

$$u_2 \cos (\phi_1 - \theta_1) = u_1 \cos \phi_1, \tag{10}$$

$$P_2 + \rho_2[u_2 \sin (\phi_1 - \theta_1)]^2 = P_1 + \rho_1[u_1 \sin \phi_1]^2, \tag{11}$$

$$h_2 + 0.5[u_2 \sin (\phi_1 - \theta_1)]^2 = h_1 + 0.5[u_1 \sin \phi_1]^2. \tag{12}$$

With the solution of system (5) - (8), ρ_1, u_1, P_1, and θ_1, are known. Thus, equations (9) - (12) represent four nonlinear equations in four unknowns P_2, ρ_2, u_2, and ϕ_1. The solution to this system is more difficult to obtain, since P_2 is not known (previously P_0 and P_1 were known). The solution was obtain numerically by utilization of the ISML subroutine ZSYSTM[9]. ZSYSTM solves a system of N simultaneous nonlinear equations in N unknowns by using Brown's[10] technique. Thus, the entire flow field in the neighborhood of Q can be determined uniquely in this shock fixed coordinate system from the given initial conditions. Furthermore, this flow configuration can be verified experimentally for a class of incident shock strengths and wall angles (or incident angles).

The theory of single Mach reflections from a solid boundary is discussed in References 8, 11 and quite completely by 12. Throughout this paper we will refer to single Mach reflection as Mach reflections and state more complex Mach reflections explicitly. Consider a plane shock wave I which is propagating with a constant velocity, is incident upon a plane rigid wall making an angle θ_w with the horizontal, and causes a Mach reflection to arise from the wall. The frame of reference is attached to the triple point T. See Figure 3. The incident shock I, reflected shock R and the Mach stem M emanate from T as well as the slipline. The trajectory of T is along a constant angle χ from the leading edge of the wall surface. The region upstream of the I and M is denoted by region 0, upstream of R and downstream of I by region 1, downstream of R by region 2, and downstream of M by region 3. The slipline divides regions 2 and 3 which have equal pressures and flow directions but different velocities. We wish to correlate the properties in these four regions in the immediate vicinity of the triple point. In this shock fixed coordinate system, the incident shock velocity is zero and the gas velocity in region 0 relative to the wall's velocity is parallel to the wall surface. The portion of the flow in region 0 which passes through I makes an angle

$$\phi_0 = 90^\circ - (\theta_w + \chi) \tag{13}$$

with I. The resulting flow is then very similar to that described in the regular reflection case except that the flow in region 2 need not be parallel to the wall surface itself, but only parallel relative to the wall's motion. With the assumption that the incident and reflected shocks are straight line shocks at least in the neighborhood of T, the oblique shock relations which now relate uniform flow properties in regions 0, 1 and 2 are:

$$\rho_1 u_1 \sin(\phi_0 - \theta_1) = \rho_0 u_0 \sin\phi_0, \tag{14}$$

$$P_1 + \rho_1\left[u_1 \sin(\phi_0 - \theta_1)\right]^2 = P_0 + \rho_1\left[u_0 \sin\phi_0\right]^2, \tag{15}$$

$$u_1 \cos(\phi_0 - \theta_1) = u_0 \cos\phi_0, \tag{16}$$

$$h_1 + 0.5\left[u_1 \sin(\phi_0 - \theta_1)\right]^2 = h_0 + 0.5\left[u_0 \sin\phi_0\right]^2, \tag{17}$$

$$\rho_2 u_2 \sin(\phi_1 - \theta_2) = \rho_1 u_1 \sin\phi_1, \tag{18}$$

$$P_2 + \rho_2\left[u_2 \sin(\phi_1 - \theta_2)\right]^2 = P_1 + \rho_1\left[u_1 \sin\phi_1\right]^2, \tag{19}$$

$$u_2 \cos(\phi_1 - \theta_2) = u_1 \cos\phi_1, \tag{20}$$

$$h_2 + 0.5\left[u_2 \sin(\phi_1 - \theta_2)\right]^2 = h_1 + 0.5\left[u_1 \sin\phi_1\right]^2. \tag{21}$$

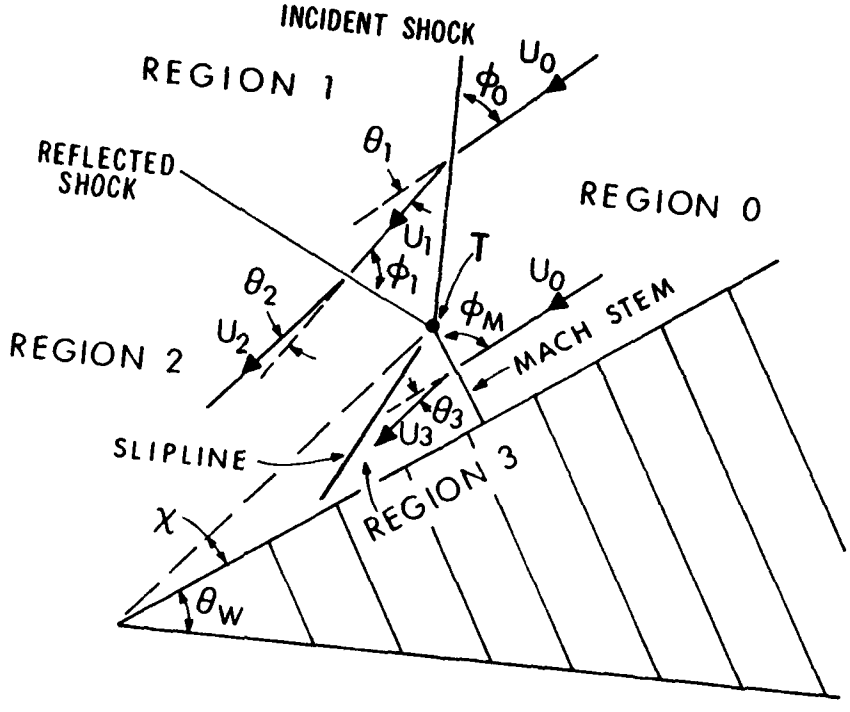


Figure 3. Schematic of Mach Reflection in a Neighborhood of the Triple Point T.

The portion of the flow in region 0 which passes through the Mach stem makes an angle ϕ_M with M. In general, the Mach stem will be a curved shock so that ϕ_M will vary with position along the Mach stem. While passing through the Mach stem, the flow is deflected by an angle θ_3 from its original direction and its dynamic and thermodynamic properties are changed. If the Mach stem is straight at least in the neighborhood of T, the flow is uniform in the vicinity. These properties are related by the oblique shock relations (1) - (4). The oblique shock relations can be simplified with $\vec{u}_a \cdot \vec{n} = u_0 \sin \phi_M$, $\vec{u}_b \cdot \vec{n} = u_3 \sin (\phi_M - \theta_3)$, $\vec{u}_a \cdot \vec{t} = u_0 \cos \phi_M$ and $\vec{u}_b \cdot \vec{t} = u_3 \cos (\phi_M - \theta_3)$ and rewritten as:

$$\rho_3 u_3 \sin (\phi_M - \theta_3) = \rho_0 u_0 \sin \phi_M, \tag{22}$$

$$u_3 \cos (\phi_M - \theta_3) = u_0 \cos \phi_M, \tag{23}$$

$$P_3 + \rho_3 \left[u_3 \sin (\phi_M - \theta_3)\right]^2 = P_0 + \rho_0 \left[u_0 \sin \phi_M\right]^2, \tag{24}$$

$$h_3 + 0.5 \left[u_3 \sin (\phi_M - \theta_3)\right]^2 = h_0 + 0.5 \left[u_0 \sin \phi_M\right]^2. \tag{25}$$

Furthermore, the flow fields in regions 2 and 3 are related across the slipline, namely, equal pressures and the same flow direction occur:

$$P_3 = P_2, \tag{26}$$

$$\theta_3 = \theta_1 - \theta_2. \tag{27}$$

For the special case where $\chi = 0$, the triple point T attaches to the wall, the slipline and region 3 are nonexistent, and if one allows $\theta_3 = 0$ equations (13) - (21) and equation (27) reduce to the regular reflection case. For $\chi \neq 0$, equations (13) - (27) represent 15 equations in 16 unknowns χ, ϕ_0, u_0, ρ_1, u_1, θ_1, ϕ_1, θ_2, ρ_2, P_2, u_2, ϕ_M, θ_3, ρ_3, P_3, and u_3, when the initial conditions (P_0, P_1, T_0 and θ_w) of the analytic model are given. The perfect gas relations $\rho = P/(R^*T)$ and $h = \gamma P/[(\gamma-1)\rho]$ are assumed. Thus, only nonunique solutions exist for this system. In order to obtain an unique solution, a simplification can be made: the Mach stem is assumed to be a straight line shock. Except for strong diffractions, the Mach stem is only slightly curved.[11] Thus, this assumption will not introduce gross errors and will allow a uniform flow field about the Mach stem. Since the flow adjacent to the wall must remain relatively parallel to the wall's surface in the laboratory coordinate system after passing through the Mach stem, the Mach stem must intersect the wall at 90°. Consequently, we have the following geometric relation

$$\phi_M = 90 - \chi \qquad (28)$$

along the entire Mach stem. Equations (13) - (28) form a system of 16 nonlinear equations in 16 unknowns which can be solved uniquely in the neighborhood of T when the initial conditions of the analytic model are given.

For nonstationary flow, the criteria for distinguising the sundry types of reflections are contained in References 8 and 12. Reflection occurs if the flow behind the incident shock is nonsubsonic in the shock fixed coordinate system. Figure 4 delineates the regions of regular reflection (bottom section) from Mach reflections (top section) in the angle of incidence - inverse shock strength plane. The curve labeled ϕ_e is the limiting curve above which regular reflection is theoretically impossible. The curve labeled ϕ_c is the boundary below which the past history cannot affect the reflection process. The experimental points indicate the smallest incident angle at which Mach reflection has been observed. The termination of Mach reflection occurs when the Mach number in the shock fixed coordinate system of region 2 is nonsubsonic.

The implementation of the regular and Mach reflection theories to form the analytic model for shock wave propagation into a re-entrant corner is best illustrated by an example. In the next section, the model is used to simulate two shock tube experiments.

IV. EXAMPLES AND COMPARISON WITH EXPERIMENTS. A series of shock tube experiments which had a straight shock propagating in air perpendicularily along a shock tube wall into a corner with apex angle of 50° were performed at the ARRADCOM Ballistic Research Laboratory. In reference to Figure 1, the experiments simulated the corner with $2\alpha = 100°$ and with the shock tube wall substituting for the plane of symmetry. The length of the rigid material forming the wall of the corner was 0.166 m. A pressure gage was inserted at the apex and the pressure-time history of the apex was obtained.[13] The extent of the corner's width was long enough as to consider it infinite. In an experimental series, the weakest incident shock had a pressure ratio of $P_1/P_0 = 1.1231$ and the strongest incident shock had a pressure ratio of $P_1/P_0 = 2.3699$. In the experiments, assumption a of the analytic model is satisfied and assumption d will be shown to be satisfied. Consequently, in the simulation of these experiments by the model, we note that air does not strictly satisfy the perfect inviscid gas assumption and the walls forming the corners do not have infinite extent. The latter restricts the model's prediction to the calculation of the peak apex pressure value, since the amplified pressure value of the apex will be decreased by the arrival of the rarefaction wave from the leading edge of the experimental corner. To obtain the time-dependent decrease of the apex pressure, a hydrodynamic computer code simulation of the entire experiment must be performed.

Let us consider the straight line shock of strength $P_1/P_0 = 2.3699$. See Figure 5. From the geometry, the incident angle is $\phi_0 = 40°$. The medium is assumed to

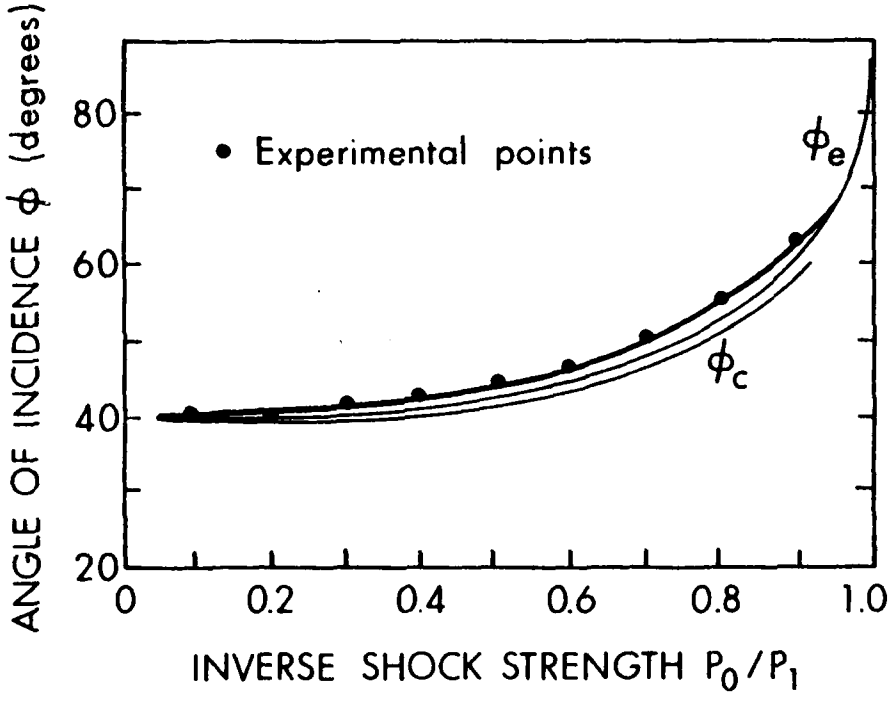


Figure 4. Criterion for Determining the Presence of Regular and Mach Reflection (Adapted from Reference 8).

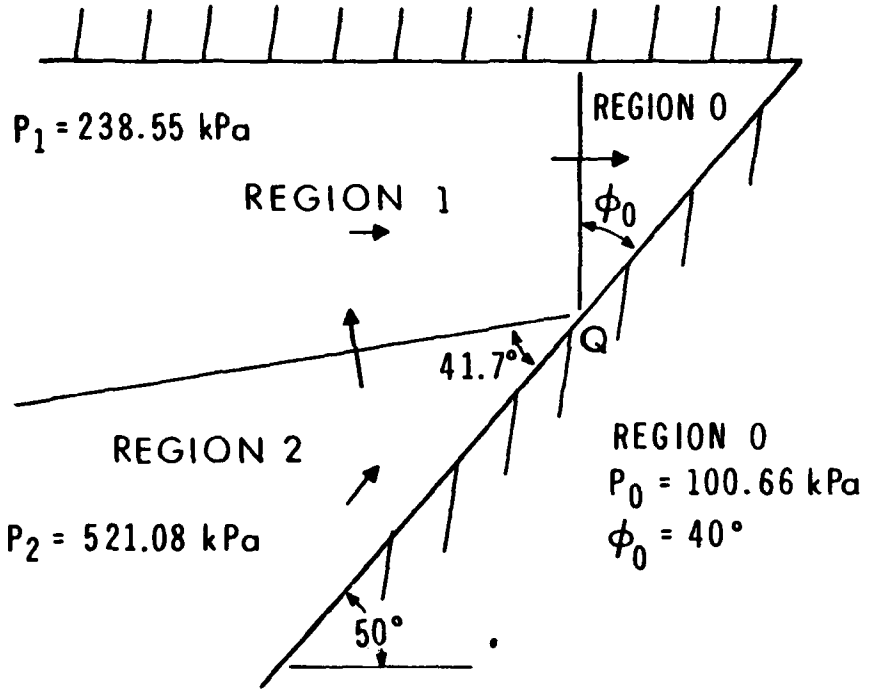


Figure 5. Schematic of the Flow Field Before Incident Shock (P_1/P_0 = 2.3699) Reaches the Apex.

have constant specific heats c_v = 714.0 J/(kg • K) and c_p = 1001. J/(kg • K), and a gas constant R* = 287.03 J/(kg • K). The initial state of region 0 is characterized by P_0 = 100.66 kPa, T_0 = 295.48K, and zero gas velocity in the laboratory coordinate system with its origin at the apex. Consider a point Q on the corner wall at which the incident shock impinges. If we make a Galilean transformation at Q, we can apply the formulas in the Appendix. In the shock fixed coordinate system, we compute ρ_1 = 2.1561 kg/m^3, u_1 = 667.35 m/s, T_1 = 385.45K, θ_1 = 15.208° and a_1 = 393.84 m/s. Since the flow is supersonic in region 1, reflection occurs at Q. The point ϕ_0 = 40° and P_1/P_0 = 0.4168 falls below the ϕ_e curve and thus, regular reflection occurs at Q. Solving the four equations governing regular reflection, equations (9) - (12), in the neighborhood of Q with ZSYSTM termination parameters EPS = 10^{-10}, NSIG = 13 and ITMAX = 100, we obtain ρ_2 = 3.7132 kg/m^3, u_2 = 488.09 m/s, P_2 = 521.08 kPa and ϕ_1 = 56.893°. From geometric considerations, the angle of reflection is 41.685°. With the assumption of an infinite corner, rarefaction waves do not presently exist within the corner and the incident and reflected shock remain straight. Thus, the gas properties behind these shocks are uniform and the values of the flow properties calculated in the neighborhood of Q are those behind the entire extent of the shocks. Upon transforming back to the laboratory coordinates, we obtain the configuration depicted in Figure 5. The gas properties in regions 0, 1 and 2 are summarized in Table 1. (The velocities in the shock fixed coordinate system are denoted by u but in the laboratory coordinate system by v.) The shock wave speeds of the incident and first reflected shocks are denoted by v_I and v_{R1}, respectively. We note: that the velocities in the laboratory coordinates in regions 1 and 2 are parallel to the plane of symmetry and the corner wall, respectively; that the angle of incidence is not equal to the angle of reflection; and that this one reflection process has already increased the pressure near the wall by a factor of 5.18.

The pseudo-steady flow of Figure 5 remains unchanged until the incident shock reaches the apex. At that instant only the first reflected shock remains (only regions 1 and 2). This shock continues to propagate along the plane of symmetry with an angle of 8.315° and a speed of 3637.0 m/s. With an inverse pressure strength of P_2/P_1 = .4578, regular reflection occurs at any reflection point Q′ according to Figure 4. See Figure 6. Since the flow properties are already calculated in regions 1 and 2, only the flow in region 3 must be calculated. We make a Galilean transformation at Q′. In this shock fixed coordinate system, the velocity magnitudes are u_1 = 3865.52 m/s in region 1 and u_2 = 3838.63 m/s in region 2 and the flow deflection angle across the shock is 3.404°. Solving the four equations governing regular reflection, equations (9) - (12), in the neighborhood of point Q′ with identical ZSYSTM termination parameters as before, we obtain ρ_3 = 6.0394 kg/m^3, u_3 = 3808.8 m/s, P_3 = 1.045 MPa and ϕ_2 = 9.069°. From geometric considerations, the angle of reflection is 5.605°. With the infinite corner assumption, the second reflected shock remains straight and the gas properties behind the shock are uniform. Thus, the values of the flow properties calculated in the neighborhood of Q′ are those behind the entire extent of the second reflected shock. Upon transforming back to the

Table 1. Regional Flow Properties in Laboratory Coordinates.

Region 0	Region 1	Region 2
P_0 = 100.66 kPa	P_1 = 238.55 kPa	P_2 = 521.08 kPa
ρ_0 = 1.1869 kg/m^3	ρ_1 = 2.1561 kg/m^3	ρ_2 = 3.7132 kg/m^3
T_0 = 295.48 K	T_1 = 385.45 K	T_2 = 488.91 K
a_0 = 344.82 m/s	a_1 = 393.84 m/s	a_2 = 443.56 m/s
v_{0x} = 0 m/s	v_{1x} = 228.52 m/s	v_{2x} = 194.62 m/s
v_{0y} = 0 m/s	v_{1y} = 0 m/s	v_{2y} = 231.94 m/s
v_I = 508.36 m/s	v_{R1} = 525.96 m/s	v_{R2} = 355.24 m/s

Region 3	Region 4*	Region 5*
P_3 = 1.0447 MPa	P_4 = 1.8913 MPa	P_5 = 1.8913 MPa
ρ_3 = 6.0394 kg/m^3	ρ_4 = 9.1668 kg/m^3	ρ_5 = 8.7707 kg/m^3
T_3 = 602.65 K	T_4 = 718.80 K	T_5 = 751.26 K
a_3 = 492.46 m/s	a_4 = 537.83 m/s	a_5 = 549.84 m/s
v_{3x} = 171.8 m/s	v_{4x} = -33.492 m/s	v_{5x} = -101.90 m/s
v_{3y} = 0 m/s	v_{4y} = -75.358 m/s	v_{5y} = -121.42 m/s
v_{R3} = 479.713 m/s		v_{R4} = 497.18 m/s

*These values are valid only in the neighborhood of the triple point.

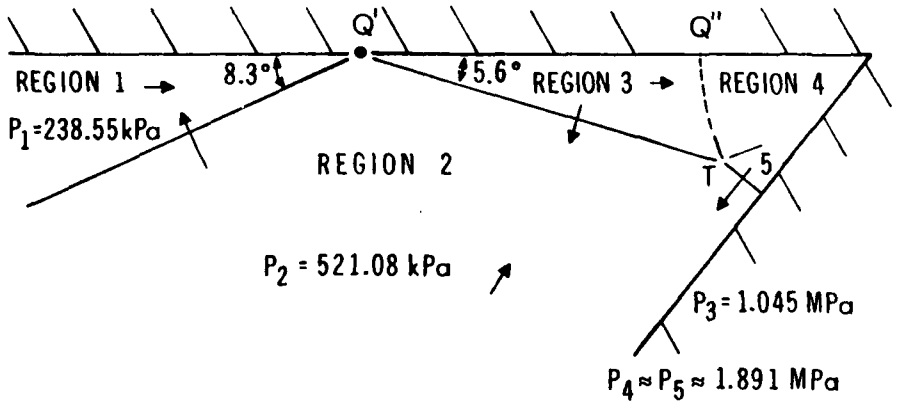


Figure 6. Schematic of Flow Field Shortly After Incident Shock

laboratory coordinates, we obtain the flow field in Figure 6 with respect to regions 1, 2 and 3. The gas properties of region 3 are given in Table 1. We note that the gas velocity in region 3 is parallel to the plane of symmetry. The speed of the second reflected shock is denoted by v_{R2}.

The second reflected shock will subsequently impinge on the corner wall at an angle of 55.605°. With an inverse shock strength of $P_2/P_3 = 0.4988$, Mach reflection occurs at the wall according to Figure 4. For Mach reflections, a Galilean transformation is made at the triple point T. The velocity of the triple point depends on an unknown χ of the configuration (see Figure 3). Thus, in the shock fixed coordinate system for Mach reflections, the values corresponding to ϕ_0, u_0, u_1, θ_1 are not known, even though in the laboratory coordinate system they are known. Since the thermodynamic properties are independent of the coordinate system, the values corresponding to P_0, P_1, ρ_0, ρ_1 are known along with the wall angle $\theta_w = 90° - 55.605°$. In reference to Section III, we have now 15 equations and 15 unknowns (ρ_1 is now known). It was found by the author that the solution of these equations is more simply obtained by the following procedure: (a) guess an initial value of χ, χ_G, and compute the corresponding ϕ_{M_G} from equation (28); (b) solve an appropriate subset of the equations corresponding to equations (13) - (27) for ϕ_1, θ_2, ρ_2, P_2, u_2, ϕ_{M_c}, θ_3, ρ_3, P_3 and u_3 with the subroutine ZSYSTM; (c) iterate on χ until $|\phi_{M_G} - \phi_{M_c}|$ is zero within a given tolerance. Following this technique (tolerance $= 10^{-4}$), the solution of the flow field in the neighborhood of T in the shock fixed coordinate system is $\chi = 11.083°$, $\phi_0 = 48.664°$, $\phi_M = 83.059°$, $\theta_1 = 13.710°$, $\phi_1 = 80.807°$, $\theta_2 = 4.608°$, $\theta_3 = 9.101°$, $u_0 = 805.86$ m/s, $u_1 = 649.33$ m/s, $u_2 = 434.86$ m/s, $u_3 = 352.39$ m/s, $\rho_2 = 9.1668$ kg/m^3, $\rho_3 = 8.7707$ kg/m^3 and $P_2 = P_3 = 1.8913$ MPa.

Table 1 lists the values of the flow variables in the laboratory coordinate system. From geometric considerations the angle between the second reflected shock and the reflected shock in the Mach configuration is 64.239°. If we extend this Mach reflected shock in a straight line to the plane of symmetry, the incident angle at the point of intersection Q'' is 110.156° - an obtuse angle. Thus, no more reflection can occur. This reflected shock must then impinge at Q'' at a right angle to the plane of symmetry in order that the gas flow in the neighborhood of Q'' remains parallel to the plane of symmetry. Consequently, the reflected shock must be curved to satisfy the required angles at points T and Q'', and the flow properties in region 4 are not uniform. See Figure 6. The exact values of the flow properties in region 4 must be obtained by a hydrodynamic computer code simulation. However, an approximation of the pressure values in region 4, and, thus, at the apex can be obtained by taking the pressure value calculated in the intersection of region 4 and the neighborhood of T as the apex value. Although obviously incorrect, the resulting pressure value gives a comparable peak pressure value to those of experiments while retaining the simplicity of the model. We note

that if the incident angle at Q'' was acute, Q'' would be another reflection point and the method of analysis would have continued. If the incident angle at Q'' was 90°, no further reflection would have occurred and the final shock would not have been diffracted. In such a case, the reflected shock would remain straight, the gas properties in region 4 would be uniform, and the method would give EXACT values of the flow field within the entire infinite re-entrant corner. The Mach stem is straight and intersects the wall at 90° according to the discussion in Section III. Its speed is denoted by v_{R4}. The velocity in region 5 is parallel to the wall. The pressure near the wall behind the Mach stem has increased by a factor of 18.8. The curved slipline separates regions 4 and 5 which has identical pressure values across it. The velocities in regions 4 and 5 relative to the unsteady motion of the slipline are parallel to the slipline. Because region 4 is nonuniform, the computed values are valid only in a neighborhood of the triple point and care must be used in any extrapolation. For example, the velocity in region 4 given in Table 1 is obviously wrong near the plane of symmetry. The speed of the Mach reflected shock v_{R3} is also correct only in the neighborhood of T.

The second experiment with an incident shock strength of $P_1/P_0 = 1.1231$ was analyzed in a similar fashion. Conceptually, the only difference in the analyses is that at the last reflection point, regular reflection occurred instead of Mach reflection. As before, the last shock wave was diffracted and the pressure at the final reflection point was taken as the apex pressure value.

Initial Strength	Peak Apex Pressure Values	
P_1/P_0	Analytic Model	Experiment
1.1231	165.44 kPa	(154.58 ± 2.69) kPa
2.3699	1.8913 MPa	(1.8195 ± 0.0859) MPa

Table 2. Analytical and Experimental Peak Apex Pressure Values for Two Incident Shocks.

Table 2 lists both the experimental peak apex pressure values with error bars and the approximate apex pressure value of the analytic model. For the weaker shock, the difference between the pressure values is 7.0 percent and for the stronger shock, the difference is 3.9 percent. The model predicted value for the stronger shock is within the experimental error. The comparison of the stronger shock results indicates that the closeness of the point ($\phi = 40°$, $P_0/P_1 = 0.4168$) associated with the first reflection to the ϕ_c curve of Figure 4 does not invalidate the model. In both comparisons, the analytic pressure values are greater than the experimental pressure values because the analytic value is the pressure value at the shock front (the last reflection point) of the expansion wave emanating from the apex which is larger than the pressure values behind the shock front.

V. SUMMARY. A simple, widely applicable analytic model for shock wave propagation into re-entrant corners is discussed. The mathematical techniques used in the model are simple; namely Galilean transformations and a method to solve a nonlinear system of algebraic equations. Its wide applicability stems from the fact that both regular and Mach reflections are modeled. When the final reflected shock is not diffracted, the model calculates the exact solution of the entire flow field within an infinite corner, or the exact solution of the flow field near the apex in a finite corner until the rarefaction wave(s) reaches the apex. When the final reflected shock is diffracted, the model provides an estimate of the peak apex pressure value which is shown to be comparable to experimental values. The model's delineation of the corner into distinct regions is verifiable by a hydrodynamic computer code simulation. See Reference 14. Because of the model's predictive capabilities, the model could be used as an aid to experimental design and as a benchmark problem for hydrodynamic computer codes.

REFERENCES

1. Lighthill, M. J., "The Diffraction of Blast II", Proc. Roy. Soc., Series A, Vol. 198, pp. 554-65, 1950.

2. Keller, J. B. and Blank, A., "Diffraction and Reflection of Pulses by Wedges and Corners", Communs. Pure and Appl. Math., Vol. IV, No. 1, pp. 75-94, 1951.

3. Keller, J. B. "Multiple Shock Reflection in Corners", Journal of Applied Physics, Vol. 25, No. 5, pp. 558-590, 1954.

4. Schniffman, T., Heyman, R. J., Sherman, A., and Weimer, D., "Pressure Multiplication in Re-Entrant Corners", in Proceedings of the First Shock Tube Symposium, Air Force Weapons Center Report No. SWR-TM-57-2 (AD467-201), 1957.

5. Thompson, P. A., Compressible Fluid Dynamics, McGraw-Hill Book Company, New York, 1972.

6. Courant, R. and Friedricks, K.O., Supersonic Flow and Shock Waves, Vol. 1, Interscience Publishers, Inc., New York, 1948.

7. Bleakney, W. and Taub, A. H., "Interaction of Shock Waves", Review of Modern Physics, 21, pp. 584-605, 1949.

8. Polachek, H. and Seeger, R., "Shock Wave Interactions" in Fundamentals of Gas Dynamics, H. W. Emmons, Ed., Princeton University Press, pp. 494-504, 1958.

9. International Mathematical and Statistical Libraries, Inc., ISML Library 3, Edition 6, ISML, Houston, Texas, 1977.

10. Brown, K. M., "A Quadratially Convergent Newton-Like Method Based Upon Gaussian Elimination", SIAM Journal on Numerical Analysis, Vol. 6, No. 4, pp. 560-569, 1969.

11. Law, C. K., "Diffraction of Strong Shock Waves by a Sharp Compressive Corner", University of Toronto Institute for Aerospace Studies Technical Note No. 150, 1970.

12. Ben Dor, G., "Regions and Transitions of Nonstationary Oblique Shock Wave Diffractions in Perfect and Imperfect Gases", University of Toronto Institute for Aerospace Studies Technical Report No. 232, 1978.

13. Taylor, W., ARRADCOM, BRL, Private Communication.

14. Schmitt, J. A., Goodman, H., and Lottero, R., "Analytic Model and Numerical Simulation of Shock Wave Propagation into a Re-Entrant Corner", Draft ARRADCOM BRL Technical Report.

APPENDIX

We multiply equation (6) by the quantity $\frac{1}{2}\left[\frac{1}{\rho_1} + \frac{1}{\rho_0}\right]$, use equation (5) and obtain:

$$\tfrac{1}{2}(P_1 - P_0)\left[\frac{1}{\rho_1} + \frac{1}{\rho_0}\right] = 0.5\left[u_0^2 \sin^2 \phi_0 - u_1^2 \sin^2(\phi_0 - \theta_1)\right]. \tag{A1}$$

Rewriting equation (8) in terms of pressure and density instead of enthalpy, we have

$$\frac{\gamma}{\gamma-1}\left[\frac{P_1}{\rho_1} - \frac{P_0}{\rho_0}\right] = 0.5\left[u_0^2 \sin^2 \phi_0 - u_1^2 \sin^2(\phi_0 - \theta_1)\right]. \tag{A2}$$

We equate the left hand sides of (A1) and (A2). Multiplying the resulting equation by the ratio ρ_1/P_0, we obtain the density in region 1 in terms of γ, P_1/P_0 and ρ_0:

$$\rho_1 = \rho_0 \left\{\left[\frac{\gamma+1}{\gamma-1}\frac{P_1}{P_0} + 1\right] \Big/ \left[\frac{P_1}{P_0} + \frac{\gamma+1}{\gamma-1}\right]\right\}. \tag{A3}$$

The ratio of equations (5) and (6) is

$$\rho_1 \tan(\phi_0 - \theta_1) = \rho_0 \tan \phi_0. \tag{A4}$$

From equation (A4), the deflection angle θ_1 can be expressed as:

$$\theta_1 = \phi_0 - \tan^{-1}\left[(\rho_0 \tan \phi_0)/\rho_1\right]. \tag{A5}$$

The velocity magnitude in region 0 can be expressed in terms of the quantities u_1, ϕ_0 and θ_1 from equation (7):

$$u_0 = u_1 \cos(\phi_0 - \theta_1)/(\cos \phi_0). \tag{A6}$$

Using equation (A6), we can rewrite equation (6) as:

$$(P_1 - P_0)/\cos^2(\phi_0 - \theta_1) = u_1^2 \left[-\rho_1 \tan^2(\phi_0 - \theta_1) + \rho_0 \tan^2 \phi_0\right]. \tag{A7}$$

Using equation (A5), we rewrite equation (A7) as:

$$u_1 = + \left[\frac{P_1 - P_0}{\rho_1 - \rho_0} \frac{\rho_0}{\rho_1}\right]^{1/2} \Big/ \sin(\phi_0 - \theta_1). \tag{A8}$$

By solving equations (A3), (A5), (A8), and (A6) in order and by applying the perfect gas relations $h_1 = \gamma P_1/[(\gamma-1)\rho_1]$, $T_1 = P_1/(R^*\rho_1)$ and $a_1 = (\gamma P_1/\rho_1)^{1/2}$, we obtain all the values of the flow variables in the intersection of region 1 and a neighborhood of the reflection point Q.

ON THE INITIAL BOUNDARY VALUE PROBLEM FOR THE EQUATIONS OF GAS DYNAMICS

Joseph Oliger*
Mathematics Research Center
University of Wisconsin-Madison
and
Department of Computer Science**
Stanford University
Stanford, CA 94305.

ABSTRACT. A priori estimates for the initial boundary value problem for the linearized equations for gas dynamics in three space dimensions are discussed. The strengths of the different estimates which are obtainable with different specifications of boundary data will be emphasized. The estimates presented are obtained by the energy method techniques of Friedrichs and/or by extensions of normal mode analysis techniques of Kreiss. Limitations of the two techniques are discussed. Boundary conditions for subsonic and supersonic flows with rigid wall and open boundaries are included.

I. THE PROBLEM. The Eulerian equations for gas dynamics in three space dimensions can be written in the form

$$\begin{aligned} &\frac{d}{dt}\,\underline{u} + \alpha \operatorname{grad} p + \underline{F} = 0 \\ &\frac{d}{dt}\,\alpha - \alpha \operatorname{div} \underline{u} = 0 \\ &\frac{d}{dt}\,p + p\,\gamma \operatorname{div} \underline{u} = 0 \end{aligned} \tag{1}$$

where $\underline{u} = (u_1, u_2, u_3)$ is the velocity vector, $\alpha = 1/\rho$ is the specific volume, ρ is the density, and p is the pressure. $\gamma = c_p/c_v$ is the ratio of specific heats at constant pressure, c_p, and at constant volume, c_v. For the material derivative d/dt we have

$$\frac{d}{dt} = \partial_t + \underline{u} \cdot \operatorname{grad} .$$

The term $\underline{F}$ contains zero order or undifferentiated terms which might represent

* The author has been sponsored in the course of this work by the United States Army under Contract No. DAAG29-75-C-0024 and by the Office of Naval Research under Contract No. N00014-75-C-1132.

** Current address.

coriolis forces, etc.

We will consider the initial boundary value problem for the linearized equations corresponding to system (1). If we let $\underline{q} = (u_1,u_2,u_3,\alpha,p)'$ be a perturbation of the state $\bar{\underline{q}} = (\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{\alpha},\bar{p})'$ the linearized equations corresponding to (1) can be written

$$\partial_t \underline{q} + \sum_{j=1}^{3} A_j(\bar{\underline{q}})\, \partial_{x_j} \underline{q} + C\underline{q} = \underline{f} \qquad (2)$$

where

$$A_1(\bar{\underline{q}}) = \begin{bmatrix} \bar{u}_1 & 0 & 0 & 0 & \bar{\alpha} \\ 0 & \bar{u}_1 & 0 & 0 & 0 \\ 0 & 0 & \bar{u}_1 & 0 & 0 \\ -\bar{\alpha} & 0 & 0 & \bar{u}_1 & 0 \\ \bar{p}\gamma & 0 & 0 & 0 & \bar{u}_1 \end{bmatrix}, \qquad A_2(\bar{\underline{q}}) = \begin{bmatrix} \bar{u}_2 & 0 & 0 & 0 & 0 \\ 0 & \bar{u}_2 & 0 & 0 & \bar{\alpha} \\ 0 & 0 & \bar{u}_2 & 0 & 0 \\ 0 & -\bar{\alpha} & 0 & \bar{u}_2 & 0 \\ 0 & \bar{p}\gamma & 0 & 0 & \bar{u}_2 \end{bmatrix},$$

and

$$A_3(\bar{\underline{q}}) = \begin{bmatrix} \bar{u}_3 & 0 & 0 & 0 & 0 \\ 0 & \bar{u}_3 & 0 & 0 & 0 \\ 0 & 0 & \bar{u}_3 & 0 & \bar{\alpha} \\ 0 & 0 & -\bar{\alpha} & \bar{u}_3 & 0 \\ 0 & 0 & \bar{p}\gamma & 0 & \bar{u}_3 \end{bmatrix}.$$

The matrix C arises from the linearization process and the zero order term $\underline{F}$. This is discussed in detail in [6]. The matrices A_j, $j = 1,2,3$, are not symmetric but they are simultaneously symmetrizable.

If we let R be the positive definite matrix

$$R = \bar{\alpha}^{-1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & a^2\bar{p}\gamma/\bar{\alpha} & a^2 \\ 0 & 0 & 0 & a^2 & (1+a^2)\dfrac{\bar{\alpha}}{\bar{p}\gamma} \end{bmatrix}$$

where a^2 is an arbitrary positive real parameter and define T by $T^{*-1}T^{-1} = R$, then $T^{-1}A_jT$ is symmetric for $j = 1,2,3$. The matrices A_j have the real eigenvalues $\bar{u}_j,\bar{u}_j,\bar{u}_j,\bar{u}_j+\bar{c}$ and $\bar{u}_j - \bar{c}$ where $\bar{c} = (\bar{p}\bar{\alpha}\gamma)^{\frac{1}{2}}$ is the sound speed. It follows that (2) is a symmetrizable hyperbolic system but it is not strictly hyperbolic.

We are generally interested in estimates of the solution of equation (2) on domains $[0,\infty] \times \Omega$ where $\Omega \subset R^3$ is open with a compact C^∞ boundary, $\partial\Omega$, in terms of initial data $\underline{q}_0(x)$ defined on Ω; suitable boundary data $\underline{g}$ defined on $(0,\infty) \times \partial\Omega$; and the forcing term $\underline{f}$. In order to simplify our notation and discussion here we will consider the quarter-space problem for Ω, i.e., we will consider solutions of (2) for $(t,\underline{x}) \in (0,\infty) \times (0,\infty) \times R^2$ which satisfy initial conditions

$$\underline{q}(0,x) = \underline{q}_0(x) \quad 0 \leq x_1 < \infty, \; -\infty < x_2, x_3 < \infty \qquad (3)$$

and boundary conditions

$$M(\underline{q})\underline{q} = \underline{g}(t,x) \quad t \in (0,\infty), \quad x_1 = 0, -\infty < x_2, x_3 < \infty, \qquad (4)$$

where M is a matrix. Our main interest here is the strength of the estimates we can get for different M. We assume compatibility of (3) and (4) in the space-time corner $x_1 = 0$, $t = 0$.

Extensions to more general domains are immediate for those results obtained via the classical energy method of Friedrichs, see Oliger and Sundström [6]. The extension to more general domains of those results obtained via the normal mode analysis technique of Kreiss is discussed by Majda and Osher [4]. The form of the estimates is not affected.

Define

$$\|\underline{q}\|_{j,G,\eta} = \|e^{-\eta t/2}\,\underline{q}\|_{H^j(G)}$$

where $\|\cdot\|_{H^j(G)}$ is the usual Sobolev norm,

$$\|\underline{q}\|^2_{H^j(G)} \; \overset{=}{|\alpha| \leq j} \sum \|\partial^\alpha \underline{q}\|^2_{L_2(G)},$$

for integer j with α a multi-index with n=dim G components. We define $\|\underline{q}\|_{H^j(G)}$ in the usual way for fractional j. Let $\underline{v}$ denote the projection of $\underline{q}$ on Ker A_1. We seek estimates of the solution of (2) satisfying (3) and (4) of the form

$$\|\underline{q}(t)\|_{0,\Omega,\eta} + \sqrt{\eta}\,\|\underline{q}\|_{0,[0,t]\times\Omega_c,\eta} + \|\underline{v}\|_{0,[0,t]\times\partial\Omega,\eta} \qquad (6)$$

$$\leq C\left(\|\underline{q}_0\|_{0,\Omega,\eta} + \frac{1}{\sqrt{\eta}}\|\underline{f}\|_{0,[0,t]\times\Omega,\eta} + \|\underline{g}\|_{j,[0,t]\times\partial\Omega,\eta}\right)$$

for a constant C and η sufficiently large where
$\Omega_c = \{\underline{x} \in R^3 : 0 \leq c < x_1 < \infty, \; -\infty < x_2, x_3 < \infty\}$ and $\Omega = \Omega_0$.

This is much like the form of estimates used in Majda and Osher [4] and we will use their notation and development extensively here.

[illegible] unless the boundary is characteristic and we will only comment on one form of characteristic boundary problem here.

We need to make several assumptions about equation (1) and its boundary conditions (4). We assume that the coefficient matrices [illegible] are uniformly bounded [illegible] constant for $|x| + |t|$ large. We assume that the matrix [illegible] is smoothly varying.

It is also necessary that we choose our boundary conditions, or M, such that [illegible] annihilates all vectors in the null space of A_1 and such that M has as many rows as the dimension of the positive eigenspace of A_1.

III. [illegible] THE ENERGY METHOD. The results mentioned in this section are from [illegible] and Sundström [illegible] and many of these estimates were established earlier by [illegible].

Estimates of the form (6) with $\epsilon = 0$ and $j = 0$ can be obtained easily and also [illegible] in certain cases from a growth equation

$$\partial_t \|\underline{q}\| \leq K \|\underline{q}\| + \|\underline{f}\| \tag{7}$$

where $\|\cdot\|$ is an inner product norm

$$\|\underline{q}\|^2 = \int_\Omega \underline{q}^* R\, \underline{q}$$

which is equivalent to the L_2 norm. R is the matrix defined in Section I. The estimate (6) will follow from (7) using integration by parts if

$$B(\underline{q}) \equiv \int_{\partial\Omega} \underline{q}^* R A_1 \underline{q} \leq 0 \quad \text{when} \quad \underline{g} \equiv 0 . \tag{8}$$

If this is not the case this technique provides us with no information.

There are several important problems for which (8) can be easily verified.

Case 1. A solid wall boundary. The physical boundary condition is that the normal velocity should vanish at the boundary, in our situation $u_1 = 0$ at $x_1 = 0$. In this case Ker A_1 has dimension 3 and $B(\underline{q}) \equiv 0$. We obtain (6) with $\epsilon = 0$ and $j = 0$. This is the only case we will consider here with $v \not\equiv 0$.

Case 2. Supersonic inflow or outflow. If $u_1 > c > 0$ we have supersonic inflow. If we take $M = I$, i.e., we specify all quantities, then $B(\underline{q}) \equiv 0$ and

we obtain (6) with $\epsilon = 0$ and $j = 0$. If $u_1 < c < 0$ we have supersonic outflow. If we take $M = 0$, i.e., we specify nothing, then $B(q) \leq 0$ and we obtain (6) with $\epsilon = 0$ and $j = 0$.

Case 3. Subsonic outflow. If $c < u_1 < 0$ we have subsonic outflow. In this case the positive eigenspace of A_1 has rank one and an examination of (8) shows that we can obtain (6) via the energy method only if (4) yields a relationship of the form

$$u_1 - p(\bar{\alpha}/(\bar{\rho}\gamma))^{\frac{1}{2}} = b_1 u_2 + b_2 u_3 + b_3[u_1 + p(\bar{\alpha}/\bar{\rho}\gamma))^{\frac{1}{2}}]$$

$$+ b_4[\alpha(\bar{\rho}\gamma/\bar{\alpha})^{\frac{1}{2}} + p(\bar{\alpha}/(\bar{\rho}\gamma))^{\frac{1}{2}}] + g(t,x_2,x_3)$$

where the b_i must be chosen to satisfy (8). It is easy to see that (8) is satisfied if we choose $b_1 = b_2 = b_4 = 0$ and $b_3 = -1$. This is simply giving the normal velocity u_1 as data. In this case we again obtain (6) with $\epsilon = 0$ and $j = 0$ (8) is also satisfied if we give the pressure as data, or set $b_1 = b_2 = b_4 = 0$ and $b_3 = 1$, and we obtain the same estimate.

Case 4. Subsonic inflow. If $0 < u_1 < c$ we have subsonic inflow. In this case the dimension of the positive eigenspace of A_1 is 4 . If we examine (8) we find that M must be chosen to yield relations of the form

$$\begin{aligned} u_2 &= a_1 w + g_1(t,x_2,x_3) \\ u_3 &= a_2 w + g_2(t,x_2,x_3) \\ u_1 - p(\bar{\alpha}/(\bar{\rho}\gamma))^{\frac{1}{2}} &= a_3 w + g_3(t,x_2,x_3) \\ \alpha(\bar{\rho}\gamma/\bar{\alpha})^{\frac{1}{2}} + p(\bar{\alpha}/(\bar{\rho}\gamma))^{\frac{1}{2}} &= a_4 w + g_4(t,x_2,x_3) \end{aligned} \tag{9}$$

where $w = u_1 + p(\bar{\alpha}/(\bar{\rho}\gamma))^{\frac{1}{2}}$. Furthermore, we will be able to satisfy (8) iff

$$u_1(a_1^2 + a_2^2 + a^2 a_4^2) + \tfrac{1}{2}(c-u_1)a_3^2 \leq \tfrac{1}{2}(c + u_1) \tag{10}$$

where a is the positive, real number in the definition of R .

We can certainly satisfy (9) and (10) and thus (8) to obtain (6) with $\epsilon = 0$ and $j = 0$ if we choose $a_1 = a_2 = a_3 = a_4 = 0$. There are clearly many other choices of the a_j which will yield the same estimate via (8); however, there is no possibility of obtaining an estimate via (8) if we want to specify data for any four of the components of q directly. This is a major point of this paper and provides the motivation for the next section.

III. RESULTS VIA NORMAL MODE ANALYSIS. We have noted in the last section that the energy method fails if we want to analyze boundary conditions for subsonic inflow such as giving u_1, u_2, u_3 and α ; or u_2, u_3, α and p . The corresponding M matrices satisfy our earlier hypotheses and their convenience in terms of utilizing commonly measured quantities is obvious.

To analyze these boundary conditions we turn to the normal mode analysis techniques developed by Kreiss [2,3], and extended by Majda and Osher [4] and Miyatake [5] . Kreiss' work only covers the strictly hyperbolic case, and as we have remarked long ago, the system (1) is not strictly hyperbolic. However, it does satisfy conditions developed by Majda and Osher, their "block structure" hypothesis. These problems still don't quite fit in the Majda and Osher framework either, because they don't satisfy the "uniform Kreiss condition". However, they are like systems examined by Kreiss and Miyatake and the Majda and Osher framework can be extended to cover the boundary conditions we are now considering. The technical details of this development will be contained in Oliger and Sundström [7] .

In this section we will only discuss the subsonic inflow problem but we remark that the cases previously discussed could also be treated using the techniques of this section with the exception of the rigid wall boundary case. At this time there is no theoretical justification for this application. The boundary is characteristic, but it does not usually belong to the class of uniformly characteristic boundary problems which have been studied by Majda and Osher.

The estimates (6) are obtained in the normal mode analysis technique by applying freezing arguments utilizing pseudo differential operators, PsDOp's. The construction of the appropriate PsDOp's needed here follows from that of Majda and Osher but a modification like that of Miyatake is needed because of the existence of generalized eigenvalues [3]. Finally, the proper form of the estimates for these two cases follows using the techniques of Kreiss [3] for the generalized eigenvalue case.

The estimates are obtained via the machinery mentioned above by estimating solutions of the frozen ordinary differential equations

$$A_1(t_0,x_0)\frac{d}{dx_1}\hat{q} + (i\,\xi_2 A_2(t_0,x_0) + i\,\xi_3 A_3(t_0,x_0) + s\,I)\hat{q} = 0 \qquad (11)$$

$$M(x_0,t_0)\hat{q} = \hat{g}$$

which is formally obtained for each (t_0,x_0) on $(0,\infty) \times \partial\Omega$ by Laplace transform

in t with dual variable s and Fourier transform in x_2, x_3 with real dual variables ξ_2, ξ_3 after dropping all undifferentiated terms. $A_1(x_0, t_0)$ can be transformed to diagonal form in the cases we consider via a transformation $P(t_0, x_0)$ so we need only estimate solutions of the ordinary differential equations

$$\frac{d}{dx_1} \underline{\psi} = L \underline{\psi} \tag{12}$$

where $\underline{\psi}(s, \xi_1, \xi_2) = P^{-1} \hat{\underline{q}}$. This is accomplished using the standard techniques for linear systems of equations with constant coefficients. It is algebraically messy since a number of different eigenvalue-eigenvector representations must be used for different values of s, ξ_1 and ξ_2.

We now state our results for the two cases we are considering here.

Theorem 1. If u_1, u_2, u_3 and α are given as data for subsonic inflow the estimate (6) holds with either $\epsilon = 0$ and $j = \frac{1}{2}$ or for arbitrary $\epsilon > 0$ and $j = 0$. Furthermore, these estimates cannot be improved in the sense that (6) does not hold for $\epsilon = 0$ and $j = 0$.

In this case there is, in Kreiss' terminology [3], a generalized eigenvalue of the first kind.

Theorem 2. If u_2, u_3, α and p are given as data for subsonic inflow the estimate (6) holds with $\epsilon = 0$ and $j = 1$. Furthermore, this estimate cannot be improved in the sense that (6) does not hold for $j < 1$ for any $\epsilon \geq 0$.

In this case there is, in Kreiss' terminology, a generalized eigenvalue of the second kind. As pointed out by Kreiss the situation is even worse in this last case because the introduction of another boundary and boundary condition on it can lead to further loss of derivatives as the waves reflect from one boundary to the other.

IV. DISCUSSION AND SUMMARY. We begin with a few remarks about the two methods we have used to obtain estimates in the last two sections. The classical energy method is definitely easier to push through when it works but doesn't give as much insight into just what choice of boundary conditions, M, might work as the normal mode analysis method. The energy method, since it is based on integration by parts, is easier to extend to boundaries which are only piecewise smooth. Only the energy method is justified at present for characteristic boundaries which are not uniformly characteristic, see [4]. However, we are more limited with regard to the form of M which we can treat with the energy method and we can only obtain sufficient results.

Finally, a few words about the boundary conditions treated in the two theorems of Section III. In both cases, giving u_1, u_2, u_3, and α, or u_2, u_3, α, p; we obtain weaker estimates than we got for the other boundary conditions. There is a "loss of smoothness". In the first case, Theorem 1, this is not a serious problem since we maintain internal regularity and the introduction of a second boundary does not cause further problems. If we use the boundary conditions of Theorem 2 we can have a continued loss of smoothness globally if we introduce a second boundary, i.e., treat a bounded region. There seems to be a big difference in these two boundary conditions if difference approximations are used to approximate the solution. Elvius and Sundström [1] were able to successfully implement approximations of boundary conditions analogous to those of Theorem 1 for the shallow water equations where the same type of estimates hold but were not able to successfully compute with boundary conditions analogous to those of Theorem 2 when, once again, an estimate like that of Theorem 2 held. John Strikwerda, private communication, has had similar experiences with the equations for gas dynamics, i.e., the boundary conditions of Theorem 1 can be successfully approximated but it seems that those of Theorem 2 cannot be.

The stronger form of (6), ($\epsilon = 0$, $j = 0$) cannot be obtained for the subsonic inflow problem if boundary data is given directly in terms of any four of the five quantities u_1, u_2, u_3, α, p. The same situation exists if we were to use potential temperature in lieu of α, say. We must give data in terms of linear combinations of these variables if we want the stronger estimates. While the estimate given in Theorem 1 may often be satisfactory, it seems worthwhile to look for other prognostic variables in which the equations can be written in order to obtain stronger estimates with boundary data that has physical significance and/or is readily measured in a straightforward manner.

REFERENCES.

[1] T. Elvius and A. Sundström, "Computationally efficient schemes and boundary conditions for a fine-mesh baratropic model based on the shallow-water equations", Tellus, 25 (1973), pp. 132-156.

[2] H.-O. Kreiss, "Initial boundary value problems for hyperbolic equations", Comm. Pure Appl. Math., 23 (1970), pp. 277-298.

[3] H.-O. Kreiss, "Initial boundary value problems for hyperbolic equations", Conference on the Numerical Solution of Differential Equations, A. Dold and B. Eckman, eds., Lecture Notes in Mathematics, No. 363, Springer-Verlag, Berlin, 1974, pp. 64-74.

[4] A. Majda and S. Osher, "Initial boundary value problems for hyperbolic equations with uniformly characteristic boundaries", Comm. Pure and Appl. Math., 28 (1975), pp. 607-675.

[5] S. Miyatake, "Mixed problem for hyperbolic equation of second order", Jour. Math. Kyoto Univ., 13 (1973), pp. 435-487.

[6] J. Oliger and A. Sundström, "Theoretical and practical aspects of some initial boundary value problems in fluid dynamics", SIAM J. Appl. Math., 35 (1978), pp. 419-446.

[7] J. Oliger and A. Sundström, "The initial-boundary value problem for the inviscid Eulerian equations for fluid dynamics", to appear.

[8] J. Serrin, "On the uniqueness of compressible fluid motions", Arch. Rational Mech. Anal., 3 (1959), pp. 271-288.

EFFICIENT MULTISTEP PROCEDURES FOR NONLINEAR PARABOLIC PROBLEMS WITH NONLINEAR NEUMANN BOUNDARY CONDITIONS

Richard E. Ewing
Mathematics Research Center
University of Wisconsin-Madison
Madison, Wisconsin 53706
and
Department of Mathematics
The Ohio State University
Columbus, Ohio 43210

ABSTRACT. Efficient multistep procedures for time-stepping Galerkin methods for nonlinear parabolic partial differential equations with nonlinear Neumann boundary conditions are presented and analyzed. The procedures involve using a preconditioned iterative method for approximately solving the different linear equations arising at each time step in a discrete time Galerkin method. Optimal order convergence rates are obtained for the iterative methods. Work estimates of almost optimal order are obtained.

I. Introduction. We shall consider the numerical solution of nonlinear parabolic partial differential equations with nonlinear Neumann boundary conditions of the form

$$\begin{aligned} &\text{a)} \quad c(x,u)\frac{\partial u}{\partial t} - \nabla \cdot [a(x,u)\nabla u + b(x,u)] = f(x,t,u), \quad x \in \Omega, \ t \in J \ , \\ &\text{b)} \quad a(x,u)\frac{\partial u}{\partial \nu} + b(x,u) \cdot \nu = g(x,t,u), \quad x \in \partial\Omega, \ t \in J \ , \\ &\text{c)} \quad u(x,0) = u_0(x), \quad x \in \Omega \ , \end{aligned} \tag{1.1}$$

where Ω is a bounded domain in $\mathbb{R}^d$, $d \leq 3$, with boundary $\partial\Omega$, ν is the outward unit normal to $\partial\Omega$, $J \equiv (0,T]$, and c, a, b, f, g, and u_0 are prescribed. We shall use a Galerkin approximation in the space variable and high-order, efficient, multistep time-stepping procedures. We first present basic multistep time-stepping procedures which produce a different linear system of equations to be solved at each time step. We then modify the basic procedures by using a preconditioned iterative method to approximate the solution of the linear equations. The use of a time-independent preconditioning matrix eliminates the need to refactor a new matrix at each time step, while the iterative procedure stabilizes the resulting algorithm. Using this modification, we obtain the same order error estimates as for the base scheme with greatly reduced computational requirements. We obtain very nearly optimal possible work estimates for our procedure.

Galerkin procedures for parabolic problems with nonlinear Neumann boundary conditions were first considered by Douglas and Dupont in [8]. Then, in [17], Luskin extended the work of [8] to quasilinear equations similar to those considered here. Luskin used Crank-Nicolson time-stepping methods which are second

Sponsored by the United States Army under Contract Nos. DAAG29-75-C-0024 and DAAG29-78-G-0161. This material is based on work supported by the National Science Foundation under Grant No. MCS78-09525.

order correct in the time discretization. In [12], the author used the iterative stabilization techniques developed in [9, 10] to present computationally efficient variants of the methods of Luskin and extended these methods to treat coupled systems of nonlinear partial differential equations with nonlinear boundary conditions. In this paper, we present time-stepping procedures which are higher-order in time than those analyzed in [8-13, 17]. These time-stepping schemes are based on the backward differentiation multistep schemes [cf. 15, 14, 19]. They have been presented and analyzed for quasilinear parabolic equations by Bramble and Sammon in [2, 7]. Very efficient alternating direction variants for use on rectangular domains will appear in [4, 5].

The efficient time-stepping techniques presented here can also be used to analyze approximation procedures for initial boundary value problems for many other types of nonlinear partial differential equations. The author has applied iterative stabilization techniques to equations of Sobolev type (in [10]) which have applications in thermodynamics, fluid flow in fissured rock, and shearing of second order fluids. In [11, 12], the methods are applied to coupled systems of equations which model miscible displacement in porous media. Also the author has used iterative methods successfully for second order in time equations (in [13]) which have applications in vibrational problems and nonlinear viscoelasticity.

In Section 2 we introduce certain notational preliminaries and present the base time-stepping Galerkin schemes. In Section 3 we present our iterative modifications of the base methods and analyze the effect of the iterative approximation on a single time step. In Section 4 we obtain global error estimates for a particular multistep method. Section 5 contains a brief discussion of the computational complexity of the methods presented.

II. Preliminaries and Description of Galerkin Methods. Let $(\varphi,\psi) = \int_\Omega \varphi\psi dx$, $\|\psi\|^2 = (\psi,\psi)$, $\langle\varphi,\psi\rangle = \int_{\partial\Omega}\varphi\psi ds$, and $|\psi|^2 = \langle\psi,\psi\rangle$. Let $W_s^k(\Omega)$ be the Sobolev space on Ω with norm

$$\|\psi\|_{W_s^k} = \left(\sum_{|\alpha|\le k}\left\|\frac{\partial^\alpha\psi}{\partial x^\alpha}\right\|_{L^s(\Omega)}^s\right)^{1/s} \tag{2.1}$$

with the usual modification for $s = \infty$. When $s = 2$, let $\|\psi\|_{W_2^k} = \|\psi\|_{H^k} = \|\psi\|_k$ and $|\psi|_{W_2^k} = |\psi|_{H^k} = |\psi|_k$. If $\nabla F = (F_1,F_2)$, write $\|\nabla F\|_{W_s^k}$ in place of $\left(\|F_1\|_{W_s^k}^s + \|F_2\|_{W_s^k}^s\right)^{1/s}$. For definitions of corresponding fractional order spaces, see [16].

Let $\{M_h\}$ be a family of finite-dimensional subspaces of $H^1(\Omega)$ with the following property:

For $p = 2$ or $p = \infty$, there exist an integer $r \ge 2$ and a constant K_0 such that, for $1 \le q \le r$ and $\psi \in W_p^q(\Omega)$,

$$\inf_{\chi\in M_h} \{\|\psi-\chi\|_{W_p^0} + h\|\psi-\chi\|_{W_p^1}\} \le K_0\|\psi\|_{W_p^q} h^q \quad . \tag{2.2}$$

We also assume that $\{M_h\}$ satisfies the following so-called "inverse assumptions":

if $\psi \in M_h$,

$$\begin{aligned} &\text{a)} \quad \|\psi\|_1 \le h^{-1}K_0\|\psi\| \quad , \\ &\text{b)} \quad |\psi| \le h^{-1/2}K_0\|\psi\| \quad , \\ &\text{c)} \quad \|\psi\|_{L^\infty(\Omega)} + h\|\nabla\psi\|_{L^\infty(\Omega)} \le K_0 h^{-\frac{d}{2}}\|\psi\| \quad . \end{aligned} \tag{2.3}$$

Restrict Ω as follows (with (S) denoting the collection of restrictions):

(S) : 1) Ω is H^2-regular.

2) $\partial\Omega$ is Lipschitz.

3) There exists a constant K_0 such that

$$|\varphi|^2 \le K_0\|\varphi\|\,\|\varphi\|_1 \quad . \tag{2.4}$$

If X is a normed space on Ω with norm $\|\cdot\|_X$ and $\varphi : [0,T] \to X$, then we define

$$\begin{aligned} &\text{a)} \quad \|\varphi\|_{L^s(J;X)} = \left[\int_0^T \|\varphi(t)\|_X^s dt\right]^{1/s}, \quad 1 \le s < \infty \quad , \\ &\text{b)} \quad \|\varphi\|_{L^\infty(J;X)} = \sup_{t\in[0,T]} \|\varphi(t)\|_X \quad . \end{aligned} \tag{2.5}$$

Throughout the paper we shall assume that a and c are bounded above and below by positive constants and that a, b, c, and g are smooth functions of their arguments. We shall also assume that the solution u is sufficiently smooth for our arguments to hold. For typical explicit smoothness assumptions on u and the coefficients, see [8-12, 17].

As in [18], we shall introduce an auxiliary elliptic problem to aid in our analysis. Let $\lambda > 0$ be chosen sufficiently large that the bilinear form

$$N(\psi;\varphi,\chi) \equiv (a(\psi)\nabla\varphi,\nabla\chi) + \lambda(\varphi,\chi) - (g(t,\varphi),\chi)$$

satisfies

$$N(\psi;\varphi,\varphi) \ge K_0\|\varphi\|_1^2, \quad \varphi,\psi \in M_h \quad .$$

Let $W \in M_h$ be the projection of u into M_h, defined, for each $t \in J$, by

$$N(u(\cdot,t); W(\cdot,t),\chi) = N(u(\cdot,t); u(\cdot,t),\chi)$$
$$= -(c(u)\frac{\partial u}{\partial t},\chi) + (b(u),\nabla\chi) + (f(u),\chi) + \lambda(u,\chi), \quad \chi \in M_h . \tag{2.6}$$

Then, as in [8, 9, 12, 18], we can obtain the following lemma.

Lemma 2.1. There exists a constant $K_1 = K_1(u)$ such that if $\eta = u - W$, $s = 0$ or $s = 1$, and $2 \le q \le r$,

$$\text{a)} \quad \|\eta\|_{L^\infty(J;H^s)} \le K_1 h^{q-s} \|u\|_{L^\infty(J;H^q)}$$
$$\text{b)} \quad \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(J;H^s)} \le K_1 h^{q-s}\{\|u\|_{L^2(J;H^q)} + \left\|\frac{\partial u}{\partial t}\right\|_{L^2(J;H^q)}\} . \tag{2.7}$$

In order to require weak smoothness assumptions on $\frac{\partial u}{\partial t}$, we shall need to use some duality theory and obtain some approximation theory results in negative-indexed norms. For these results, assume that Ω, a, b, c, and g are sufficiently smooth [16] that for each $t \in J$, if

$$\text{a)} \quad -\nabla \cdot [a(x,u)\nabla u] + \lambda_1 u = \psi_1, \quad x \in \Omega ,$$
$$\text{b)} \quad a(x,u)\frac{\partial u}{\partial \nu} = \psi_2, \quad x \in \partial\Omega , \tag{2.8}$$

then

$$\|u\|_{k+2} \le K(u)\{\|\psi_1\|_k + |\psi_2|_{k+\frac{1}{2}}\} . \tag{2.9}$$

If (2.8)-(2.9) holds, we shall say that Ω is H^{k+2}-regular. Next, define for $k \ge 0$,

$$\text{a)} \quad \|\psi\|_{-k} \equiv \sup\{(\psi,\varphi) : \|\varphi\|_k = 1\} ,$$
$$\text{b)} \quad |\psi|_{-k} \equiv \sup\{(\psi,\varphi) : |\varphi|_k = 1\} . \tag{2.10}$$

Lemma 2.2. If Ω is H^{k+2}-regular for $k \le 1$, there exists a constant $K(u)$ such that for $1 \le q \le r$ and $t \in J$,

$$\|\eta\|_{-k} + |\eta|_{-(k+\frac{1}{2})} + \left\|\frac{\partial \eta}{\partial t}\right\|_{-k} \le K(u)h^{q+k}\{\|u\|_q + \left\|\frac{\partial u}{\partial t}\right\|_q\} . \tag{2.11}$$

Proof: See [12].

We also make the assumption on $\{M_h\}$ and u that there exists a constant K_2 such that

$$\|w\|_{L^\infty(J;L^\infty)} + \|\nabla w\|_{L^\infty(J;L^\infty)} + \left\|\frac{\partial w}{\partial t}\right\|_{L^\infty(J;L^\infty)} + \left\|\nabla\frac{\partial w}{\partial t}\right\|_{L^2(J;L^\infty)}$$

$$+ \left\|\frac{\partial^2 w}{\partial t^2}\right\|_{L^\infty(J;H^1)} + \left\|\frac{\partial^3 w}{\partial t^3}\right\|_{L^\infty(J;H^1(\partial\Omega))} + \left\|\frac{\partial^4 w}{\partial t^3}\right\|_{L^\infty(J;H^1)} \le K_2 \quad .$$

Sufficient conditions for the above to hold can be found in [9, 10, 18].

We next consider discrete-time Galerkin approximations. Let $\Delta t > 0$, $N = T/\Delta t \in \mathbf{Z}$ and $t^\sigma = \sigma\Delta t$, $\sigma \in \mathbb{R}$. Also let $\psi^n \equiv \psi^n(x) \equiv \psi(x,t^n)$, and

$$\text{a)}\quad d_t\psi^{n+1} = \frac{\psi^{n+1}-\psi^n}{\Delta t}$$

$$\text{b)}\quad \delta\psi^{n+1} = \psi^{n+1} - \psi^n$$

$$\text{c)}\quad \delta^2\psi^{n+1} = \psi^{n+1} - 2\psi^n + \psi^{n-1} \qquad (2.12)$$

$$\text{d)}\quad \delta^3\psi^{n+1} = \psi^{n+1} - 3\psi^n + 3\psi^{n-1} - \psi^{n-2}$$

$$\text{e)}\quad \delta^4\psi^{n+1} = \psi^{n+1} - 4\psi^n + 6\psi^{n-1} - 4\psi^{n-2} + \psi^{n-3} \quad .$$

We next define a family of extrapolated coefficient backwards differentiation multi-step discrete time methods.

Let $U : \{t_0,\ldots,t_N\} \to M_h$ be an approximate solution of (1.1). Assume that U^k are known for $k \le n$. Then, given certain choices of parameters β, α_1, α_2, α_3, and α_4 and an extrapolation $\hat{U}^{n+1}$, we determine U^{n+1} to satisfy

$$\begin{aligned}(c(\hat{U}^{n+1})\,\frac{U^{n+1}-U^n}{\Delta t},\ \chi) + \beta(a(\hat{U}^{n+1})\nabla U^{n+1},\ \nabla\chi)\\ = \beta(g(t^{n+1},\ \hat{U}^{n+1}),\ \chi) + (c(\hat{U}^{n+1})\,\frac{1}{\Delta t}\,[\alpha_1 U^n+\alpha_2 U^{n-1}+\alpha_3 U^{n-2}+\alpha_4 U^{n-2}],\ \chi) \qquad (2.13)\\ - \beta(b(\hat{U}^{n+1}),\ \nabla\chi) + \beta(f(t^{n+1},\ \hat{U}^{n+1}),\ \chi),\quad \chi \in M_h \quad .\end{aligned}$$

A parcicular example from this family of methods is the choice $\hat{U}^{n+1} = U^n$, $\beta = 1$ and $\alpha_i = 0$, $i = 1,2,3,4$. This choice is the the well-known backward Euler method with lagged coefficients which is known to have time-discretization error of order Δt. Other choices of the parameters and extrapolation in the coefficients yield temporal errors of order $(\Delta t)^2$, $(\Delta t)^3$, and $(\Delta t)^4$.

We present these special choices in the following table.

Table 1: Selected Multistep Methods

Extrapolation $\hat{U}^{n+1}$	β	α_1	α_2	α_3	α_4	Time-discretization Error $(\Delta t)^{\mu}$
$U^{n+1}-\delta U^{n+1}$	1	0	0	0	0	Δt
$U^{n+1}-\delta^2 U^{n+1}$	2/3	1/3	-1/3	0	0	$(\Delta t)^2$
$U^{n+1}-\delta^3 U^{n+1}$	6/11	7/11	-9/11	2/11	0	$(\Delta t)^3$
$U^{n+1}-\delta^4 U^{n+1}$	12/25	23/25	-36/25	16/25	-3/25	$(\Delta t)^4$

We note that by extrapolating the coefficients in (2.13), we have reduced each of the above problems to the solution of a different set of linear equations at each time step.

III. Iterative Stabilization Procedures. In this section we consider efficient methods for solving the linear equations arising from (2.13). We note that the coefficient matrices from (2.13) change with each time step. In order to avoid the factorization of different matrices at each time step to solve the different systems of linear equations, we shall present an iterative method for approximating their solution to sufficient accuracy.

Let $\{\varphi_i\}_{i=1}^{M}$ be a basis for M_h and let U^m from (2.13) be written as

$$U^m = \sum_{i=1}^{M} \xi_i^m \varphi_i \quad . \tag{3.1}$$

Using (3.1), (2.13) can be written as

$$L^n(\xi)(\xi^{n+1}-\xi^n) \equiv [C^n(\xi)\{\sum_{i=1}^{4} \alpha_i \xi^{n+1-i}\} + \Delta t F_1^n(\xi)] \tag{3.2}$$

$$\equiv F^n(\xi)$$

where the matrices and vectors are of the form

$$
\begin{aligned}
&\text{a)}\quad L^n(\xi) = C^n + \Delta t \beta A^n \quad , \\
&\text{b)}\quad C^n(\xi) = \widehat{((c(\sum_{k=1}^{M} \xi_k^n \varphi_k)\varphi_j, \varphi_i))} \quad , \\
&\text{c)}\quad A^n(\xi) = \widehat{((a(\sum_{k=1}^{M} \xi_k^n \varphi_k)\nabla\varphi_j, \nabla\varphi_i))} \quad , \\
&\text{d)}\quad F_1^n(\xi) = \beta(\widehat{(g(t^{n+1}, \sum_{k=1}^{M} \xi_k^n \varphi_k), \varphi_i)} - \widehat{(b(\sum_{k=1}^{M} \xi_k^n \varphi_k), \nabla\varphi_i)} \\
&\qquad\qquad + \widehat{(f(t^{n+1}, \sum_{k=1}^{M} \xi_k^n \varphi_k), \varphi_i)}) \quad ,
\end{aligned}
\tag{3.3}
$$

for $i, j = 1,\ldots,M$.

Instead of solving (3.2) exactly, we shall approximate its solution by using an iterative procedure which has been preconditioned by L^0, the associated matrix with coefficients evaluated at $t = 0$, for each time step. The preconditioning process eliminates the need for factoring new matrices at each time step, while the iterative procedure stabilizes the resulting problem. The stabilization process requires iteration only until a predetermined norm reduction is achieved.

Denote by

$$
V^m = \sum_{k=1}^{M} \gamma_k^m \varphi_k \tag{3.4}
$$

the approximation to U^m produced by only approximately solving (3.2). An iterative procedure for obtaining the necessary V^k starting values using the iterative procedure described here will appear in [3]. We assume such a starting procedure has been used to obtain sufficiently accurate (see (4.7)) starting values. Thus assume $V^0,\ldots,V^n$ have been determined. We shall determine the M-dimensional vector γ^{n+1} (and thus V^{n+1}) using a preconditioned iterative method to approximate ξ^{n+1} from (3.2). As an initial guess for $\xi^{n+1} - \xi^n$ we shall extrapolate from previously determined values. Specifically, for a particular method having time-truncation error $(\Delta t)^\mu$, we shall use as the initialization for our iterative procedure

$$
X_0 = (\gamma^{n+1} - \gamma^n) - \delta^{\mu+1}\gamma^{n+1} \quad , \tag{3.5}
$$

where the m^{th} backward difference operator δ^m is defined in (2.12) for $m = 1,\ldots,4$. Since we are using previously determined γ^i in the coefficient matrices to determine γ^{n+1}, our errors accumulate.

In order to estimate the cumulative error, we first consider the single step error. We define $\bar{\gamma}^{n+1}$ to satisfy

$$L^n(\gamma)(\bar{\gamma}^{n+1} - \gamma^n) = F^n(\gamma), \quad n \geq \mu \quad . \tag{3.6}$$

We can use any preconditioned iterative method which yields norm reductions of the form

$$\|L^n(\gamma)^{1/2}(\bar{\gamma}^{n+1} - \gamma^{n+1})\|_e \leq \rho_n \|L^n(\gamma)^{1/2}(\bar{\gamma}^{n+1} - \gamma^{n+1} + \delta^{\mu+1}\gamma^{n+1})\|_e \tag{3.7}$$

where $0 < \rho_n < 1$ and the subscript e denotes the Euclidean norm of the vector. A specific iterative procedure for obtaining (3.7) is the preconditioned conjugate gradient method analyzed in [1, 9, 10].

Let

$$\text{a)} \quad \|\varphi\|^2_{c^n} \equiv (c(\hat{V}^{n+1})\varphi,\varphi) \quad ,$$

$$\text{b)} \quad \|\varphi\|^2_{a^n} = (a(\hat{V}^{n+1})\nabla\varphi,\nabla\varphi) \quad , \tag{3.8}$$

$$\text{c)} \quad |||\varphi|||_n = \|\varphi\|_{c^n} + (\Delta t)^{1/2}\|\varphi\|_{a^n}$$

be special norms and seminorms. Note that $\|\cdot\|_{c^n}$ and $\|\cdot\|_{a^n}$ are uniformly equivalent to $\|\cdot\|$ and $\|\nabla\cdot\|$, respectively. Then letting

$$\bar{V}^m = \sum_{i=1}^{M} \bar{\gamma}_i^m \varphi_i \quad , \tag{3.9}$$

with $\bar{\gamma}^m$ defined in (3.6), we see that $\bar{V}^{n+1}$ satisfies

$$\begin{aligned}(c(\hat{V}^{n+1})\,\frac{\bar{V}^{n+1}-V^n}{\Delta t},X) + \beta(a(\hat{V}^{n+1})\nabla\bar{V}^{n+1},\nabla X) + \beta(b(\hat{V}^{n+1}),\nabla X) \\ = \beta(g(t^{n+1},\hat{V}^{n+1}),X) + \beta(f(t^{n+1},\hat{V}^{n+1}),X) + (c(\hat{V}^{n+1})\,\frac{1}{\Delta t}\sum_{i=1}^{4}\alpha_i V^{n+1-i},X), \quad X \in M_h \quad .\end{aligned} \tag{3.10}$$

Also using (3.8), our single-step error (3.7) becomes

$$|||\bar{V}^{n+1}-V^{n+1}|||_n \leq \frac{\rho_n}{1-\rho_n}\,|||\delta^{\mu+1}V^{n+1}|||_n, \quad n \geq \mu + 1 \quad . \tag{3.11}$$

We note that as in [6, 12], there is a Q depending upon bounds for the coefficients, such that

a) $\rho_n \leq 2Q^\kappa$, with $0 < Q < 1$, and

b) $\dfrac{\rho_n}{1+\rho_n} \equiv \rho_n' \leq n\Delta t, \quad n \geq 1$. (3.12)

IV. A Priori Error Estimates. In this section we develop a priori bounds for the errors $V^n - u^n$ for the procedures defined in (3.10) using the base schemes defince in (2.13). The techniques for treating the nonlinearities in the coefficients of a, b, and f are tedious and appear in [7, 9, 12]. Therefore, for simplicity of exposition, we shall consider the simplified problem

a) $c(x,t)\dfrac{\partial u}{\partial t} - \nabla \cdot [a(x,t)\nabla u] = 0 \quad , \quad x \in \Omega,\ t \in J \ ,$

b) $a(x,t)\dfrac{\partial u}{\partial \nu} = g(x,t,u) \quad , \quad x \in \partial\Omega,\ t \in J \ ,$ (4.1)

c) $u(x,0) = u_0(x) \quad , \quad x \in \Omega \ .$

We can thus examine the higher-order efficient time-stepping procedures without the added complexity of nonlinearities, except in the Neumann boundary condition.

Also, for simplicity, we shall present the details for the particular method whose choice of parameters yields time-discretization error of order $(\Delta t)^\mu$ where $\mu = 3$. Proofs of stability and convergence for the other methods follow similarly and can be derived from the proofs of similar problems which appear in [7].

For $\mu = 3$, the base approximation scheme for (4.1) from (2.13) can be written as

$$(c_{n+1}\delta U^{n+1}, X) + \frac{6}{11}\Delta t(a_{n+1}\nabla U^{n+1}, \nabla X) = \frac{6}{11}\Delta t\langle g(t^{n+1}, \hat{U}^{n+1}), X) + (c_{n+1}[\frac{7}{11}\delta U^n - \frac{2}{11}\delta U^{n-1}], X), \ X \in M_h \ , \tag{4.2}$$

where $c_{n+1} \equiv c(x,t^{n+1})$, $a_{n+1} \equiv a(x,t^{n+1})$, and $\hat{U}^{n+1} = 3U^n - 3U^{n-1} + U^{n-2}$. Let $\eta^n = u^n - W^n$ and $\zeta^n = V^n - W^n$. We know from Lemma 2.1 that W is a function in M_h which is sufficiently close to u. We next estimate how close V and W are. From (2.6), (4.1), and (4.2), we obtain the following error equation

$$
\begin{aligned}
&(c_{n+1}\delta\zeta^{n+1},\chi) + \frac{6\Delta t}{11}(a_{n+1}\nabla\zeta^{n+1},\nabla\chi) \\
&\quad = (c_{n+1}[\tfrac{7}{11}\,\delta\zeta^{n} - \tfrac{2}{11}\,\delta\zeta^{n-1}],\chi) + [\tfrac{6}{11}\,\lambda\Delta t(\eta^{n+1},\chi) \\
&\qquad + \Delta t(c_{n+1}[d_t\eta^{n+1} - \tfrac{7}{11}\,d_t\eta^{n} + \tfrac{2}{11}\,d_t\eta^{n-1}],\chi)] \\
&\qquad + (c_{n+1}[\tfrac{6}{11}\,\Delta t\,\frac{\partial u^{n+1}}{\partial t} - \delta u^{n+1} + \tfrac{7}{11}\,\delta u^{n} - \tfrac{2}{11}\,\delta u^{n-1}],\chi) \\
&\qquad + \tfrac{6}{11}\,\Delta t(g(t^{n+1},\hat{V}^{n+1}) - g(t^{n+1},W^{n+1}),\chi) \\
&\qquad + [(c_{n+1}(V^{n+1} - \bar{V}^{n+1}),\chi) + \tfrac{6}{11}\,\Delta t(a_{n+1}\nabla(V^{n+1} - \bar{V}^{n+1}),\nabla\chi)] \\
&\quad \equiv (c_{n+1}[\tfrac{7}{11}\,\delta\zeta^{n} - \tfrac{2}{11}\,\delta\zeta^{n-1}],\chi) + T_1^{n+1}(\chi) + T_2^{n+1}(\chi) \\
&\qquad + T_3^{n+1}(\chi) + T_4^{n+1}(\chi), \quad \chi \in M_h \ .
\end{aligned}
\tag{4.3}
$$

Term T_1 enters because we are comparing V to W instead of directly to u. Term T_2 measures how well the multistep scheme approximates $\frac{\partial u}{\partial t}$ and term T_3 arises from the nonlinearity of g. Finally, the single-step error made by using the iterative procedure to approximately solve the linear equations appears in term T_4.

We shall first present a few lemmas which will help separate the various parts of our analysis. First we note that the parameters $\beta(\mu)$ and $\alpha_i(\mu)$, $i = 1,\ldots,4$, are chosen in (2.13) to insure the following consistency result.

Lemma 4.1. For each $\mu = 1,2,3,4$, the choice of parameters $\beta(\mu)$ and $\alpha_i(\mu)$, $i = 1,2,3,4$ given in Table 1 yields

$$
\|\beta(\mu)\Delta t\,\frac{\partial u^{n+1}}{\partial t} - [\delta u^{n+1} - \sum_{i=1}^{4}\alpha_i(\mu)u^{n+1-i}]\| \le K_3(\Delta t)^{\mu+1} \ . \tag{4.4}
$$

We next consider the following lemma which will provide the estimates for the basic stability of our methods.

Lemma 4.2. Assume that z^n satisfies, for $m \ge 2$,

$$
\begin{aligned}
&\sum_{n=m}^{\ell-1}[(c_{n+1}\delta z^{n+1},\chi) + \tfrac{6}{11}\,\Delta t(a_{n+1}\nabla z^{n+1},\nabla\chi)] \\
&= \sum_{n=m}^{\ell-1}[(c_{n+1}[\tfrac{7}{11}\,\delta z^{n} - \tfrac{2}{11}\,\delta z^{n-1}],\chi) + (F^{n+1},\chi)], \quad \chi \in M_h \ .
\end{aligned}
\tag{4.5}
$$

Then there exist constants K_4, K_5 and K_6 such that setting $X = Z^{n+1}$ yields

$$\|Z^{\ell}\|^2 + \sum_{n=m}^{\ell-1} [\|\delta Z^{n+1}\|^2 + \|Z^{n+1}\|^2_{a^{n+1}\Delta t}] \tag{4.6}$$
$$\leq K_4 [\|Z^m\|^2 + \sum_{n=m-2}^{\ell-1} \|\delta Z^{n+1}\|^2 + \sum_{n=m-2}^{\ell-1} \|Z^{n+1}\|^2_{\Delta t} + |\sum_{n=m}^{\ell-1} (F^{n+1}, Z^{n+1})|] \quad ;$$

setting $X = \delta Z^{n+1}$ yields

$$\sum_{n=m}^{\ell-1} |||\delta Z^{n+1}|||^2_n + \Delta t \|Z^{\ell}\|^2_1 \leq K_5 [\Delta t \|Z^m\|^2_1 \tag{4.7}$$
$$+ \sum_{n=m-2}^{m-1} |||\delta Z^{n+1}|||^2_n + \sum_{n=m-2}^{\ell-1} \|Z^{n+1}\|^2_1 (\Delta t)^2 + |\sum_{n=m}^{\ell-1} (F^{n+1}, \delta Z^{n+1})|] \quad ;$$

also, setting $X = (n+1)\delta Z^{n+1}$ yields

$$\sum_{n=m}^{\ell-1} (n+1) |||\delta Z^{n+1}|||^2_n + \ell \Delta t \|Z^{\ell}\|^2_1 \leq K_6 [m\Delta t \|Z^m\|^2_1 + \sum_{n=m-2}^{\ell-1} \|\delta Z^n\|^2 \tag{4.8}$$
$$+ \sum_{n=m}^{\ell-1} \frac{6\Delta t(1+\Delta t)}{22} \|Z^{n+1}\|^2_{a^{n+1}} + |\sum_{n=m}^{\ell-1} (n+1)(F^{n+1}, \delta Z^{n+1})|] \quad .$$

Proof: See [7].

The following version of the discrete Gronwall lemma is trivial.

Lemma 4.3. Let $f_j \geq 0$, $\beta_j \geq 0$, and $\gamma > 0$. Assume that for $n = 1, \ldots, \ell$,

$$f_n \leq \sum_{j=m}^{n-1} \beta_j f_j \Delta t + \gamma$$

and

$$\sum_{j=m}^{n-1} \beta_j \Delta t \leq M \quad .$$

Then, $f_n \leq \gamma \exp M$, $n = m-1, \ldots, \ell$.

We shall assume that an efficient start-up procedure using the same preconditioned iterative methods as described in Section 3 has been used to determine initial approximations satisfying

$$\sum_{i=0}^{3} |||\zeta^i|||^2_i + \sum_{i=1}^{3} |||\delta\zeta^i|||^2_{i-1} \leq K[h^{2r} + (\Delta t)^6] \quad . \tag{4.9}$$

For the description of such a start-up procedure and proof of the given estimates, see [3].

We next state the major result of the paper.

Theorem 4.1. Let u and U satisfy (4.1) and (4.2), respectively. Let V be the iterative variant of U satisfying (4.9), (3.10), and (3.11) with ρ_n satisfying (4.21) below. Let $u \in L^2(J;H^r) \cap W_\infty^4(J;W_3^1)$ and either

a) $\int_0^T t\left\|\frac{\partial u}{\partial t}(\cdot,t)\right\|_r^2 dt \le K$ when Ω is H^3-regular and $h^2 \le C\Delta t$, or

b) $\frac{\partial u}{\partial t} \in L^2(J;H^r)$.

Then there exist constants $K_8(u)$, depending upon the norms of u, and h_0 and τ_0 such that if $r > d/2$, $\Delta t \le \min\{\tau_0, h^{d/6}\}$, and $h \le h_0$,

$$\sup_n \|u^n - V^n\| \le K_8(u)[h^r + (\Delta t)^3] \quad .$$

Proof: Letting $\chi = \zeta^{n+1}$ in (4.3) with $m = 3$ and using (4.6), we obtain

$$\begin{aligned} &\|\zeta^\ell\|^2 + \sum_{n=3}^{\ell-1}[\|\delta\zeta^{n+1}\|^2 + \Delta t\|\zeta^{n+1}\|_{a^{n+1}}^2] \\ &\le K_4[\|\zeta^3\|^2 + \sum_{n=1}^{\ell-1}\{\|\delta\zeta^{n+1}\|^2 + \|\zeta^{n+1}\|^2\} + |\sum_{n=3}^{\ell-1}\sum_{i=1}^{4} T_i^{n+1}(\zeta^{n+1})|] \quad . \end{aligned} \tag{4.10}$$

Next, we see that from (2.7) and (2.11),

$$\begin{aligned} \sum_{n=3}^{\ell-1}|T_1^{n+1}(\zeta^{n+1})| &\le K\sum_{n=3}^{\ell-1}[\|\eta^{n+1}\|\,\|\zeta^{n+1}\| + \sum_{j=0}^{2}\|d_t\eta^{n+1-j}\|_{-1}\|\zeta^{n+1}\|_1]\Delta t \\ &\le K_9(u)\,h^{2r} + \frac{1}{8}\sum_{n=3}^{\ell-1}\|\zeta^{n+1}\|_{a^{n+1}}^2\Delta t \quad , \end{aligned} \tag{4.11}$$

where $K_9 = K_9(\|u\|_{L^2(J;H^r)} + \left\|\frac{\partial u}{\partial t}\right\|_{L^2(J;H^{r-1})})$. We note that use of (2.7b) instead of Lemma (2.2), would have required the assumption $\frac{\partial u}{\partial t} \in L^2(J;H^r)$, a much stronger smoothness assumption. From Lemma (4.1) we see that

$$\sum_{n=3}^{\ell-1}|T_2^{n+1}(\zeta^{n+1})| \le K(u)(\Delta t)^6 + \frac{1}{8}\sum_{n=3}^{\ell-1}\|\zeta^{n+1}\|_{a^{n+1}}^2\Delta t \quad . \tag{4.12}$$

We next use (2.4) and smoothness of W to obtain the bound

$$\sum_{n=3}^{\ell-1} |T_3^{n+1}(\zeta^{n+1})| \le K \sum_{n=3}^{\ell-1} (|\zeta^{n+1}| + K_2(\Delta t)^3 + \sum_{j=0}^{2} |\delta\zeta^{n+1-j}|)|\zeta^{n+1}|\Delta t \tag{4.13}$$

$$\le K(u)\{(\Delta t)^6 + \sum_{n=1}^{\ell-1} [\|\zeta^{n+1}\|^2 + \|\delta\zeta^{n+1}\|_1^2]\Delta t\} + \frac{1}{8}\sum_{n=3}^{\ell-1} \|\zeta^{n+1}\|_{a^{n+1}}^2 \Delta t \quad .$$

Using (3.8), (3.11) and (3.12) we see that

$$\sum_{n=3}^{\ell-1} |T_4^{n+1}(\zeta^{n+1})| \le \sum_{n=3}^{\ell-1} |||V^{n+1} - \tilde{V}^{n+1}|||_n |||\zeta^{n+1}|||_n$$

$$\le \sum_{n=3}^{\ell-1} \rho'_{n+1} |||\delta^4 V^{n+1}|||_n |||\zeta^{n+1}|||_n \tag{4.14}$$

$$\le \sum_{n=3}^{\ell-1} K(u)\rho'_{n+1}\{\sum_{i=0}^{3} |||\delta\zeta^{n+1-i}|||_{n-i} + (\Delta t)^4\}|||\zeta^{n+1}|||_n$$

$$\le K(u)[(\Delta t)^6 + \sum_{n=3}^{\ell-1} \|\zeta^n\|^2 \Delta t] + \frac{1}{8}\sum_{n=3}^{\ell-1} \|\zeta^n\|_{a^{n+1}}^2 \Delta t$$

$$+ \sum_{n=0}^{\ell-1} \frac{\rho'_{n+1}}{16\Delta t} |||\delta\zeta^{n+1}|||_n^2 \quad .$$

Noting that the multiplier in the last term on the right side of (4.14) is bounded by (n+1)/16 using (3.12), we combine (4.10) - (4.14) and use (4.9) to obtain

$$\|\zeta^\ell\|^2 + \frac{1}{2}\sum_{n=3}^{\ell-1} [\|\delta\zeta^{n+1}\|^2 + \|\zeta^{n+1}\|_{a^{n+1}}^2 \Delta t]$$

$$\le K(u)[h^{2r} + (\Delta t)^6 + \sum_{n=3}^{\ell-1} \|\zeta^n\|^2 \Delta t] \tag{4.15}$$

$$+ K_4 \sum_{n=3}^{\ell-1} |||\delta\zeta^{n+1}|||_n^2 + \sum_{n=3}^{\ell-1} \frac{(n+1)}{16} |||\delta\zeta^{n+1}|||_n^2 \quad .$$

We note that if we can bound the last two terms on the right of (4.15), we can then use the discrete Gronwall Lemma to obtain our result. In order to bound the next to the last term on the right side of (4.15) we let $\chi = \delta\zeta^{n+1}$ in (4.3) and use (4.7) to obtain

$$\sum_{n=3}^{\ell-1} |||\delta\zeta^{n+1}|||_n^2 + \Delta t\|\zeta^\ell\|_1^2 \le K_5[\Delta t\|\zeta^3\|_1^2 + |||\delta\zeta^2|||_1^2 + |||\delta\zeta^3|||_2^2 \tag{4.16}$$

$$+ \sum_{n=1}^{\ell-1} \|\zeta^{n+1}\|_1^2 (\Delta t)^2 + |\sum_{n=3}^{\ell-1} \sum_{i=1}^{4} T_i^{n+1}(\delta\zeta^{n+1})|] \quad .$$

As in (4.11) we use (2.7) and (2.11) to obtain

$$\sum_{n=3}^{\ell-1} |T_1^{n+1}(\delta\zeta^{n+1})| \le K(u)h^{2r} + \frac{1}{8}\sum_{n=3}^{\ell-1} |||\delta\zeta^{n+1}|||_n^2 \quad . \tag{4.17}$$

Similarly we see that

$$\sum_{n=3}^{\ell-1} |T_2^{n+1}(\delta\zeta^{n+1})| \le K(u)(\Delta t)^6 + \frac{1}{8}\sum_{n=3}^{\ell-1} |||\delta\zeta^{n+1}|||_n^2 \quad . \tag{4.18}$$

Using (2.4) we then see that

$$\begin{aligned}\sum_{n=3}^{\ell-1} |T_3^{n+1}(\delta\zeta^{n+1})| &\le K\sum_{n=3}^{\ell-1}(|\zeta^{n+1}| + K_2(\Delta t)^3 + \sum_{j=0}^{2}|\delta\zeta^{n+1-j}|)|\delta\zeta^{n+1}|\Delta t \\ &\le \frac{1}{8}\sum_{n=1}^{\ell-1} |||\delta\zeta^{n+1}|||_n^2 + K(u)\{(\Delta t)^6 + \sum_{n=1}^{\ell-1}[\|\zeta^{n+1}\|^2\Delta t + \|\zeta^{n+1}\|_1^2(\Delta t)^2]\} \quad .\end{aligned} \tag{4.19}$$

Then, as (4.14), we use (3.8), (3.11), and (3.12) to obtain

$$\begin{aligned}\sum_{n=3}^{\ell-1} |T_4^{n+1}(\delta\zeta^{n+1})| &\le \sum_{n=3}^{\ell-1} \rho_n' |||\delta^4 v^{n+1}|||_n |||\delta\zeta^{n+1}|||_n \\ &\le \sum_{n=3}^{\ell-1} 3K_{10,n}\rho_n'\{\sum_{i=0}^{3} |||\delta\zeta^{n+1-i}|||_{n-i} + (\Delta t)^4\}|||\delta\zeta^{n+1}|||_n \\ &\le K(u)(\Delta t)^3 + 12\sum_{n=0}^{\ell-1} \rho_n' K_{10,n} |||\delta\zeta^{n+1}|||_n^2\end{aligned} \tag{4.20}$$

where $K_{10,n}$ depends upon local upper and lower bounds for the coefficients a_n and c_n (see (4.21)). Then iterating on the preconditioned iterative procedure sufficiently often that

$$\rho_n \le (48K_{10,n})^{-1} \equiv \frac{\min\{a(t^j) : j=n+1,n,n-1,n-2\}}{48\ \sup\{a(t^j) : j=n+1,n,n-1,n-2\}} \quad , \tag{4.21}$$

combining (4.16) - (4.20), and using (4.9) we see that

$$\begin{aligned}&\sum_{n=3}^{\ell-1} |||\delta\zeta^{n+1}|||_n^2 + \Delta t\|\zeta^{\ell}\|_1^2 \\ &\qquad \le K(u)[h^{2r} + (\Delta t)^6 + \sum_{n=3}^{\ell-1}\{\|\zeta^{n+1}\|^2\Delta t + \|\zeta^{n+1}\|_1^2(\Delta t)^2\}] \quad .\end{aligned} \tag{4.22}$$

In order to bound the last term on the right side of (4.15), we let $X = (n+1)\delta\zeta^{n+1}$ and use (4.8) to obtain

$$\sum_{n=3}^{\ell-1}(n+1)\left|\!\left|\!\left|\delta\zeta^{n+1}\right|\!\right|\!\right|_n^2 + \ell\Delta t\left\|\zeta^{\ell}\right\|_1^2 \le K_6[3\Delta t\left\|\zeta^3\right\|_1^2 + \sum_{n=1}^{\ell-1}\left\|\delta\zeta^n\right\|^2$$
$$+ \sum_{n=3}^{\ell-1}\frac{7\Delta t}{22}\left\|\zeta^{n+1}\right\|_{a^{n+1}}^2 + \left|\sum_{n=3}^{\ell-1}\sum_{i=1}^{4}T_i^{n+1}((n+1)\delta\zeta^{n+1})\right|] \quad . \tag{4.23}$$

We note that (4.9) and (4.22) can be used to bound the first term on the right side of (4.23). We next obtain

$$\left|\sum_{n=3}^{\ell-1}(n+1)T_1^{n+1}(\delta\zeta^{n+1})\right| \le K\sum_{n=3}^{\ell-1}[\|\eta^{n+1}\|\,\|\delta\zeta^{n+1}\| + \sum_{j=0}^{2}\|d_t\eta^{n+1-j}\|_{-1}\|\delta\zeta^{n+1}\|_1](n+1)\Delta t \tag{4.24}$$
$$\le \frac{1}{16}\sum_{n=3}^{\ell-1}(n+1)\left|\!\left|\!\left|\delta\zeta^{n+1}\right|\!\right|\!\right|_n^2 + K\sum_{n=3}^{\ell-1}\left\|\eta^{n+1}\right\|^2\Delta t + K\sum_{n=1}^{\ell-1}\left\|d_t\eta^{n+1}\right\|_{-1}^2(n+1)\Delta t \quad .$$

If $u \in L^2(J;H^r)$, we have from (2.7) that

$$\sum_{n=1}^{\ell-1}\left\|\eta^{n+1}\right\|^2\Delta t \le K(u)h^{2r} \quad . \tag{4.25}$$

Then, using (2.11) we have, if $h^2 \le \Delta t$,

$$\sum_{n=1}^{\ell-1}\left\|d_t\eta^{n+1}\right\|_{-1}^2(n+1)\Delta t \le K\sum_{n=1}^{\ell-1}[\left\|u^{n+1}\right\|_r^2 + \left\|\frac{\partial u}{\partial t}^{n+1}\right\|_r^2]h^{2r+2}(n+1)\Delta t$$
$$\le K(\int_0^T t[\|u(\cdot,t)\|_r^2 + \left\|\frac{\partial u}{\partial t}(\cdot,t)\right\|_r^2]dt)h^{2r} \quad . \tag{4.26}$$

Note that $h^2 \le \Delta t$ is not a strong restriction for these high order time-stepping methods. The constant on the right of (4.26) determines the smoothness assumptions we need on u and $\frac{\partial u}{\partial t}$ for this argument. We note that for linear, time-dependent problems the assumption

$$\int_0^T t\left\|\frac{\partial u}{\partial t}(\cdot,t)\right\|_r^2 dt \le K \tag{4.27}$$

is roughly equivalent to $\frac{\partial u}{\partial t} \in L^2(J;H^{r-1})$, the assumption needed for (4.11), and much weaker than the assumption $\frac{\partial u}{\partial t} \in L^2(J;H^r)$ which has been made in [7, 11, 17] for similar estimates. Using (4.4) we see that

$$\sum_{n=3}^{\ell-1}|T_2^{n+1}((n+1)\delta\zeta^{n+1})| \le \frac{1}{16}\sum_{n=3}^{\ell-1}(n+1)\left|\!\left|\!\left|\delta\zeta^{n+1}\right|\!\right|\!\right|_n^2 + K(u)(\Delta t)^6 \quad . \tag{4.28}$$

We next consider the T_3 term from (4.23). Note that

$$\begin{aligned}\Big|\sum_{n=3}^{\ell-1} T_3^{n+1}((n+1)\delta\zeta^{n+1})\Big| &\le \Big|\sum_{n=3}^{\ell-1}\langle\frac{\partial g}{\partial u}\,\delta^3 w^{n+1},\delta\zeta^{n+1}\rangle(n+1)\Delta t\Big| \\ &+ \Big|\sum_{n=3}^{\ell-1}\langle\frac{\partial g}{\partial u}\,\delta^3\zeta^{n+1},\delta\zeta^{n+1}\rangle(n+1)\Delta t\Big| + \Big|\sum_{n=3}^{\ell-1}\langle\frac{\partial g}{\partial u}\,\zeta^{n+1},\delta\zeta^{n+1}\rangle(n+1)\Delta t\Big| \\ &\equiv T_5 + T_6 + T_7 \quad .\end{aligned} \tag{4.29}$$

T_6 can be bounded, using (2.4), as follows

$$\begin{aligned}T_6 &\le K\sum_{n=3}^{\ell-1}\sum_{j=0}^{2}|\delta\zeta^{n+1-j}|\;|\delta\zeta^{n+1}|(n+1)\Delta t \\ &\le \frac{1}{16}\sum_{n=1}^{\ell-1}\|\delta\zeta^{n+1}\|^2(n+1) + K\sum_{n=1}^{\ell-1}\|\delta\zeta^{n+1}\|_1^2(n+1)(\Delta t)^2 \quad .\end{aligned} \tag{4.30}$$

Then using a technical summation by parts argument and estimates like those used in (4.30) we can obtain (see [12, p. 27-29] for details)

$$\begin{aligned}T_5 + T_7 &\le \frac{1}{16}\sum_{n=1}^{\ell-1}\{\|\delta\zeta^{n+1}\|^2(n+1) + \|\zeta^{n+1}\|_{a^{n+1}\Delta t}^2\} + \frac{1}{22}\|\zeta^{\ell}\|_1^2\ell\Delta t \\ &+ K\{\|\zeta^3\|_1^2\Delta t + (\Delta t)^6 + \sum_{n=1}^{\ell-1}[\|\delta\zeta^n\|_{L^\infty}^2 n + \Delta t]\,[\|\zeta^{n+1}\|^2 + \|\zeta^{n+1}\|_1^2(n+1)\Delta t]\} \\ &+ K_{12}\|\zeta^{\ell}\|^2\ell\Delta t \quad .\end{aligned} \tag{4.31}$$

As in (4.20), we see that

$$\begin{aligned}\sum_{n=3}^{\ell-1}|T_4^{n+1}((n+1)\delta\zeta^{n+1})| &\le \sum_{n=3}^{\ell-1}K_{10,n}\rho_n'\{\sum_{i=0}^{3}|||\delta\zeta^{n+1-i}|||_{n-i} + (\Delta t)^4\}|||\delta\zeta^{n+1}|||_n(n+1) \\ &\le K(u)(\Delta t)^6 + 4K_{10,n}\sum_{n=0}^{\ell-1}\rho_n'K_{10,n}|||\delta\zeta^{n+1}|||_n^2(n+1) \quad .\end{aligned} \tag{4.32}$$

Next, by iterating sufficiently often to satisfy (4.21), combining (4.23) - (4.32), and using (4.9), we obtain

$$\frac{1}{2}\{\sum_{n=1}^{\ell-1}(n+1)\,|||\delta\zeta^{n+1}|||_n^2 + \ell\Delta t\|\zeta^\ell\|_1^2\} \le \frac{4}{11}\sum_{n=3}^{\ell-1}\|\zeta^{n+1}\|_{a^{n+1}\Delta t}^2$$

$$+ K(u)[h^{2r} + (\Delta t)^6] + K_{12}\|\zeta^\ell\|^2\ell\Delta t \qquad (4.33)$$

$$+ K(u)\sum_{n=1}^{\ell-1}\{\|\zeta^{n+1}\|^2 + \|\zeta^{n+1}\|_1^2(n+1)\Delta t\}\{\|\delta\zeta^n\|_{L^\infty}^2 n + \Delta t\} \quad .$$

Now adding inequalities (4.15) and (4.33) to K_4 times inequality (4.22) and simplifying we obtain

$$\|\zeta^\ell\|^2 + \|\zeta^\ell\|_1^2\ell\Delta t + \sum_{n=3}^{\ell-1}\{|||\delta\zeta^{n+1}|||_n^2(n+1) + \|\zeta^{n+1}\|_{a^{n+1}\Delta t}^2\}$$

$$\le \bar{K}(u)\{h^{2r} + (\Delta t)^6\} + 4K_{12}\|\zeta^\ell\|^2\ell\Delta t \qquad (4.34)$$

$$+ \bar{\bar{K}}(u)\sum_{n=1}^{\ell-1}\{\|\zeta^{n+1}\|^2 + \|\zeta^{n+1}\|_1^2 n\Delta t\}\{\|\delta\zeta^n\|_{L^\infty}^2 n + \Delta t\} \quad .$$

We next indicate how to treat the term multiplied by $4K_{12}$ on the right side of (4.34). Note that for some $\epsilon_1 > 0$,

$$\|\zeta^{n+1}\|^2(n+1)\Delta t - \|\zeta^n\|^2 n\Delta t - \|\zeta^{n+1}\|^2\Delta t \qquad (4.35)$$

$$\le \epsilon_1(n+1)\|\delta\zeta^{n+1}\|^2 + K\|\zeta^n\|^2\Delta t \quad .$$

We sum (4.35) from $n = 3$ to $n = \ell - 1$, multiply the results by $4K_{12} + \frac{1}{2}$, and add the final inequality to (4.34). Then take $\epsilon_1 \le (8K_{12} + 1)^{-1}$. Next, we make the induction hypothesis that

$$\sum_{n=1}^{\ell-1}\|\delta\zeta^n\|_{L^\infty}^2 n < 1 \quad . \qquad (4.36)$$

Then it follows from (4.34) - (4.36) and Lemma 4.3 that

$$\sum_{n=1}^{\ell-1}\|\delta\zeta^n\|^2 n \le 2\exp\{(1+T)\bar{\bar{K}}(u)\}\bar{K}(u)[h^{2r} + (\Delta t)^6] \quad . \qquad (4.37)$$

It then follows from (4.37) and the inverse hypothesis (2.3.c) that

$$\sum_{n=1}^{\ell-1}\|\delta\zeta^n\|_{L^\infty}^2 n \le K_0 h^{-d}\sum_{n=1}^{\ell-1}\|\delta\zeta^n\|^2 n \le Kh^{-d}[h^{2r} + (\Delta t)^6] \quad . \qquad (4.38)$$

We note that the right hand side of (4.38) tends to zero as h tends to zero if

$$r > \frac{d}{2} \quad \text{and} \quad \Delta t < h^{\frac{d}{6}} \tag{4.39}$$

which justifies the induction hypothesis. Since this implies

$$\|\zeta^{\ell}\|^2 + \|\zeta^{\ell}\|_1^2 \ell\Delta t \leq K[h^{2r} + (\Delta t)^6] \quad , \tag{4.40}$$

the result follows from (4.40), Lemma 2.1, and the triangle inequality.

We note that similar theorems hold for the original nonlinear problem and for the other various multistep methods presented. Also, if Ω is a rectangle, rectangular solid, or unions of these regions, alternating direction variants of the multistep methods presented here are even more computationally efficient. See [4, 5] for these results.

V. Computational Considerations. In this section we shall consider some rough operation counts to estimate the computational complexity of the methods presented here. We shall see that the preconditioned iterative methods allow us to obtain very nearly optimal order work estimates and are thus very efficient computationally.

We shall give estimates for $d = 2$. The procedures of setting up and factoring L^n requires $O(M^{3/2})$ operations, where $M = \dim M_h$. The solution of (3.2), given the factorization, requires $O(M \log M)$ operations. Such bounds have been shown to be minimal. If we conjecture the validity of the above estimates for our problem and refactor L^n and solve (3.2) at each time step, the total amount of work done is

$$O(N\{M^{3/2} + M \log M\}) = O(NM^{3/2}) \quad , \tag{5.1}$$

where N is the total number of time steps $(N \approx (\Delta t)^{-1})$. Note that the work of factorization dominates the estimate.

Using the preconditioned iterative procedure presented here, only the preconditioner, L^0, must be factored. Let κ_n be the number of iterations needed to achieve the necessary norm reductions in (3.11) and (3.12). We note that κ_n can be bounded by a fixed constant κ which is independent of h, n, and Δt. Using this method the total work done is

$$O(M^{3/2} + N\kappa M \log M) \quad . \tag{5.2}$$

Since balancing the spatial and temporal errors yields

$$N \approx (\Delta t)^{-1} \approx h^{-\frac{r}{\mu}} = O(M^{r/2\mu})$$

we note that for $r \geq \mu$, the work of solving dominates the estimate, while for $r < \mu$ the amount of work of solving is even less than the work to factor one

matrix, a necessary piece of work. Clearly, in any case, (5.2) is much preferable to (5.1). Also, since the total number of unknowns in the problem is

$$O(NM) \quad ,$$

(5.2) represents a nearly optimal order work estimate when the work is at least as much as factoring one matrix. If alternating direction variants of these methods can be used, the $\log M$ term can be removed from (5.2) and optimal order work estimates are obtained (see [4, 5]).

It is computationally wasteful to iterate exactly κ times at each time step in order to achieve the pessimistic bounds on ρ_n given in (4.21). Instead, one can monitor the norm reduction actually produced at each time step of the iteration and stop iterating when sufficient norm reduction is achieved. Additional stopping criteria can be imposed in this monitoring process. See [9] for a discussion of stopping criteria for related methods.

REFERENCES

1. O. Axelsson, "On preconditioning and convergence acceleration in sparce matrix problems," CERN European Organization for Nuclear Research, Geneva, 1974.

2. J. H. Bramble, "Multistep methods for quasilinear parabolic equations," Proc. Second Int. Conf. on Comp. Meth. in Nonlinear Mechanics, Austin, Texas, March 26-29, 1979. (to appear).

3. J. H. Bramble and R. E. Ewing, "Efficient starting procedures for high order time-stepping methods for differential equations," (to appear).

4. J. H. Bramble and R. E. Ewing, "Alternating direction multistep methods for parabolic problems - iterative stabilization," (to appear).

5. J. H. Bramble and R. E. Ewing, "Direct alternating direction multistep methods for parabolic problems," (to appear).

6. J. H. Bramble and P. H. Sammon, "Efficient higher order single-step methods for parabolic problems: part I," Math. Res. Center Rep. #1958, Madison, Wisconsin, 1979.

7. J. H. Bramble and P. H. Sammon, "Efficient higher order multistep methods for parabolic problems: part I," (to appear).

8. J. Douglas, Jr., and T. Dupont, "Galerkin methods for parabolic equations with nonlinear boundary conditions," Numer. Math. 20 (1973), pp. 213-237.

9. J. Douglas, Jr., T. Dupont and R. E. Ewing, "Incomplete iteration for time-stepping a Galerkin method for a quasilinear parabolic problem," SIAM J. Numer. Anal. 16 (1979), pp. 503-522.

10. R. E. Ewing, "Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations," SIAM J. Numer. Anal. 15 (1978), pp. 1125-1150.

11. R. E. Ewing, "Efficient time-stepping procedures for miscible displacement problems in porous media," Math. Res. Center Rep. #1934, Madison, Wisconsin (1979) and SIAM J. Numer. Anal. (to appear).

12. R. E. Ewing, "Efficient time-stepping methods for miscible displacement problems with nonlinear boundary conditions," Math. Res. Center Rep. #1952 (1979) and Calculo (to appear).

13. R. E. Ewing, "On efficient time-stepping methods for nonlinear partial differential equations," Computers and Math. with Appl. (to appear).

14. C. W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, New Jersey, 1971.

15. P. Henrici, Discrete Variable Methods in Ordinary Differential Equations, John Wiley and Sons, New York, 1962.

16. J. L. Lions and E. Magenes, Non-homogeneous Boundary Value Problems and Applications, Vol. I, Springer-Verlag, New York, 1972.

17. M. Luskin, "A Galerkin method for nonlinear parabolic equations with nonlinear boundary conditions," SIAM J. Numer. Anal. 16 (1979) pp. 284-299.

18. M. F. Wheeler, "A priori L^2-error estimates for Galerkin approximations to parabolic partial differential equations," SIAM J. Numer. Anal. 10 (1973), pp. 723-759.

19. M. Zlámal, "Finite element multistep discretizations of parabolic boundary value problems," Math. Comp. 29 (1975) pp. 350-359.

INTEGRAL BOUNDS FOR THE STRAIN ENERGY IN TERMS OF SURFACE TRACTIONS OR DISPLACEMENTS AND BODY FORCES IN FINITE ELASTOSTATICS

Joseph J. Roseman*
Department of Mathematical Sciences
Tel-Aviv University

ABSTRACT. A number of results are discussed in which a bound for the strain energy is obtained in terms of integral norms of the given surface data and body forces in the context of finite elasticity, where the displacement gradients are assumed to be small, but not infinitesimal.

I. INTRODUCTION. We consider a homogeneous, isotropic elastic body which occupies a domain $\mathcal{D}_R$, with boundary $\partial\mathcal{D}_R$, in an unstressed undeformed state and which is mapped smoothly, one to one, onto a domain $\mathcal{D}$, with boundary $\partial\mathcal{D}$, under the action of surface tractions and body forces. Let $x = (x_1, x_2, x_3)$ be a point in $\mathcal{D}_R$ which is mapped onto a point $y = (y_1, y_2, y_3)$ in $\mathcal{D}$. The displacement vector u is defined as

$$u = y - x, \tag{1}$$

and, since the mapping is assumed to be one to one, u may be regarded as a function of either x or y. The Jacobian matrix of the transformation p_{ij} is given by

$$p_{ij} = \delta_{ij} + u_{i,j}\,^{\dagger} \quad , \tag{2}$$

---

* The author was on a leave of absence at Georgia Institute of Technology, Mathematics Department, at the time of the Conference. This manuscript was prepared while visiting at Georgia Tech and at the University of Delaware, Mathematics Department. The author's participation in the conference was supported by Georgia Tech. and the U.S. Army Math. Research Center in Madison, Wisconsin.

† Tensor notation is used throughout this paper; all indices may take on the values 1, 2, and 3 and a repeated index in any term is summed over all values of the index. Differentiation with respect to x_k is denoted by a comma, as $u_{i,k} = \partial u_i/\partial x_k$.

the metric tensor, g_{ik}, by

$$g_{ik} = p_{ji}\, p_{jk} = g_{ki} \ , \tag{3}$$

and the strain matrix e_{ik} is defined as

$$e_{ik} = 1/2\ (g_{ik} - \delta_{ik}). \tag{4}$$

It is assumed also that the body is hyperelastic, i.e., there exists a positive strain energy density function $W = W(e_{ik})$ with units of energy/volume which describes the energy in any subdomain of $\mathcal{D}$ in terms of its preimage in $\mathcal{D}_R$ (cf. [1], [2]).

The equations of linear elasticity are obtained by formally considering the displacement gradient $u_{i,k}$ to be infinitesimally small so that quadratic and higher order terms are assumed to be negligable when compared with first order terms.

In the first (displacement) boundary value problem of elastostatics, the boundary displacements and internal body forces are considered to be given data. In the second (traction) boundary value problem the boundary stresses and internal body forces are given. The mixed problem has displacement data on part of the boundary and tractions on the complement. The question of whether the boundary stresses and body forces refer to the deformed or undeformed domain is important in nonlinear elasticity; in linear elasticity the question does not arise since the difference is considered to be of negligably small order.

Integral bounds for the strain energy in terms of the given surface data and body forces have been obtained by Bramble and Payne [3], [4], Dou [5], and Gutierrez [6] in the context of linear elasticity. Here we shall discuss some recent work in which similar bounds are obtained in the context of nonlinear finite elasticity where the displacement gradients are assumed to be sufficiently small relative to the size and geometry of the domain, but not infinitesimal. The various results are obtained by a variety of different techniques and approaches.

In Section II, the strain energy density function is assumed to satisfy the hypotheses of Villagio's inequality [7], which states that

$$\int_{\mathcal{D}_R} \frac{\partial W}{\partial u_{i,k}}\, u_{i,k}\, d\Omega \geq k \int_{\mathcal{D}_R} \epsilon_{ik}\, \epsilon_{ik}\, d\Omega \ , \tag{5a}$$

where

$$\varepsilon_{ik} = 1/2\ (u_{i,k} + u_{k,i}) \quad , \tag{5b}$$

and k is a positive constant which depends upon the domain and the material properties. This assumption puts some restrictions on the class of solids (see [7]).

The Villagio assumption is not used in Sections III and IV. There the only assumption on $W(e_{ij})$ is that it be of the form

$$W = \mu e_{ik}\, e_{ik} + 1/2\ \lambda e_{ii}\, e_{kk} + 0(|e|^3) \quad , \tag{6}$$

where $0(|e|^3)$ is a smooth term of order of magnitude $|e|^3$ for sufficiently small $|e|$ and λ and μ are positive material constants.

II. THE FIRST BOUNDARY VALUE PROBLEM WITH ZERO BOUNDARY DISPLACEMENT. Aron and Roseman [8] studied the first B.V.P. for the case where the boundary displacements are everywhere zero. When the mass density in the reference state is constant, the displacement vector u satisfies the Navier equation

$$\frac{\partial}{\partial x_k}\left(\frac{\partial W}{\partial u_{i,k}}\right) = -F_i(u) = \frac{\partial V}{\partial y_i}(y) \quad , \tag{7}$$

where F_i is the body force in $\mathcal{L}_R$ (force/volume) and V is the potential of the body force (In [8], the density is actually allowed to be variable; the units of F_i and W in [8] are force/mass and energy/mass, respectively). If the validity of Villagio's inequality (5) is assumed, then it is easy to show that

$$||u|| \leq \frac{1}{k}\, ||F(u)|| \quad , \tag{8}$$

when u is a classical smooth solution of the problem and the indicated norm is the L_2 integral norm over $\mathcal{D}_R$. However, in [8], (8) is derived for the more general case where u is a weak solution with the aid of a theorem of Ekeland [9]. It is shown that for every $\epsilon>0$, there exists a vector function u_ϵ such that if

$$J(u) = \int_{\mathcal{D}_R} \left[W(e_{ik}) + V(y)\right] d\Omega \quad , \tag{9}$$

for all $u \in M = \{u/u \in C^2(\bar{\mathcal{D}}_R)$ and $u = 0$ on $\partial\mathcal{D}_R\}$, then

a) $$|| \, DJ \, (u_\epsilon^*) \, ||^* \leq \epsilon \quad , \tag{10}$$

where D is the Gateau derivative of the functional J and $|| \quad ||^*$ is the norm defined by

$$|| \, DJ(u_\epsilon) \, ||^* = \sup_{\substack{v \in M \\ v \neq 0}} \frac{\int_{\mathcal{D}_R} \left[\overline{\frac{\partial W}{\partial u_{i,k}}} \Big|_{u = u_\epsilon} v_{i,k} - F_i(u_\epsilon) \, v_i \right] d\Omega}{||v||} \quad , \tag{11}$$

and

b) $$|| \, u_\epsilon || \leq \frac{1}{k} \, || \, F(u_\epsilon) \, || \, . \tag{12}$$

If it is also assumed that the inequality

$$W(e_{ik}) < \frac{\partial W}{\partial u_{i,k}} \dot{u}_{i,k} \qquad u \in M, \; u \neq 0 \quad , \tag{13}$$

is valid (see [10]), then one may obtain

$$\int_{\mathcal{D}_R} W(e_{ij}) \Big|_{u = u_\epsilon} d\Omega \leq \frac{1}{k} \, || \, F(u_\epsilon) \, ||^2 \tag{14}$$

for every $\epsilon > 0$ and, by allowing $\epsilon \to 0$, we obtain the inequality even in the sense of weak solutions.

III. THE FIRST BOUNDARY VALUE PROBLEM WITH ZERO BODY FORCE. Breuer and Roseman [11] considered the first B.V.P. when body forces are absent and boundary displacements are prescribed. The displacement vector u satisfies

$$\frac{\partial}{\partial x_k} \left(\frac{\partial W}{\partial u_{i,k}} \right) = 0 \quad \text{in } \mathcal{D}_R \quad , \tag{15}$$

and

$$u_i = U_i \, (\xi_1, \xi_2) \quad \text{on } \Gamma \subset \partial\mathcal{D}_R \tag{16a}$$

$$u_i = 0 \qquad \text{on } \partial\mathcal{D}_R - \Gamma, \tag{16b}$$

the ξ_1 and ξ_2 being suitable orthogonal surface coordinates chosen such that $d\sigma$, the element of surface area on Γ, is given by $d\sigma = d\xi_1\, d\xi_2$.

The approach used in [11] is based on a minimum principle of Fritz John [12] in finite elastostatics which is described below. Breuer and Roseman [11] prove the following:

Theorem: Consider an isotropic homogeneous elastic body, which in its undeformed state occupies $\mathcal{D}_R \subset E^3$ and which is mapped onto $\mathcal{D}$ in accordance with (15) and (16). The domain $\mathcal{D}_R$ is of bounded eccentricity in the sense of John [12] and has sufficient regularity for the application of the divergence theorem. The patron $\Gamma \subset \partial\mathcal{D}_R$ has two continuous derivatives; there exists an h_o such that h_o is less than half the minimum radius of curvature of $\overline{\Gamma}$ and at every point of Γ it is possible to place a tangent sphere of radius h_o whose interior is contained in $\mathcal{D}_R$.

Then there exists an $\varepsilon > 0$, depending on $\mathcal{D}_R$ and W, such that if the maximum strain in $\mathcal{D}$ is less than ε, and if the boundary displacements U_i satisfy

$$U_i \in C^2(\Gamma), \tag{17a}$$

$$U_i, \partial U_i, \partial^2 U_i \quad \text{all vanish at } \partial\Gamma, \tag{17b}$$

$$|U_i| < \varepsilon h_o, \tag{17c}$$

$$|\partial U_i| < \varepsilon, \tag{17d}$$

where ∂U_i and $\partial^2 U_i$ represent any first and second derivatives with respect to the surface coordinates ξ_1 and ξ_2 of (16), then

$$E \le B_1 h_o \left(\left|\left| \frac{\partial U}{\partial \xi_1} \right|\right|_\Gamma^2 + \left|\left| \frac{\partial U}{\partial \xi_2} \right|\right|_\Gamma^2 \right) + \frac{B_2}{h_o} \left|\left| U \right|\right|_\Gamma^2 , \tag{18}$$

where E is the total strain energy in $\mathcal{D}$, $||\ \ ||_\Gamma$ is the L_2 integral norm over Γ, and B_1 and B_2 depend only upon the specific strain energy density function W.

Proof: In [12], John proved for a wide class of domains and energy density functions that if K is the subset of vector functions in $C^2(\mathcal{D}_R) \cap C(\overline{\mathcal{D}}_R)$ which satisfy the given boundary data and produce sufficiently small strain, then the solution of the Navier equation (15) is the one and only function in K for which the energy

functional is a minimum. It is assumed here that a solution u exists and, with the assumptions made above on U_i, a vector $v \in K$ is constructed. The quantity $E(v)$, the energy associated with the vector v, is computed and, by John's minimum principle, gives an upper bound for $E(u)$.

We note that the constants in (18) depend only upon the geometry and size of that portion of the boundary which contains the non-zero surface data.

IV. THE SECOND BOUNDARY VALUE PROBLEM. The second B.V.P. in finite elastostatics is considered by Breuer and Roseman in [13], who obtain a result similar to that obtained by Bramble and Payne [3] in the context of linear elasticity. Breuer and Roseman assume here that the body is elastic, isotropic, and homogeneous, and that its reference domain $\mathcal{D}_R$ is convex. The displacement gradient is assumed to satisfy an a priori bound of the form

$$\left| \frac{\partial u_i}{\partial x_k} \right| \leq \delta \left(\frac{d}{D} \right)^{3/2} , \tag{19}$$

where d and D are the inner and outer diameters of $\mathcal{D}_R$ respectively and δ is a sufficiently small universal constant. It is proved that

$$E \leq B\, dD \left(\frac{D}{d} \right)^q \left[||F||^2 + \frac{1}{D} ||T||^2 \right] , \tag{20}$$

where

i) E is the strain energy in $\mathcal{D}$,
ii) $||F||$ is the L_2 integral norm of the body forces in $\mathcal{D}$,
iii) $||T||$ is the L_2 integral norm of the surface tractions over $\partial\mathcal{D}$,
iv) q is a sufficiently large positive universal constant, and
v) B is a constant which depends only the physical properties of the body.

The bound (20) was obtained by a combination of the linear techniques of Bramble and Payne [3], the work of F. John ([14],[15]) on the relation between rotation and strain in nonlinear elasticity, and by a priori estimate techniques.

In a forthcoming paper [16], the same authors consider the second B.V.P. when the body forces are zero and the surface tractions are zero on all but a connected subdomain Γ of $\partial\mathcal{D}_R$. By an extension of the arguments used in [13], they prove that

$$E \leq B \left(\frac{H}{h}\right)^{q} h \, ||T||_{\Gamma}^{2} \quad , \tag{21}$$

where h and H are the inner and outer diameters of a closed subregion of $\overline{\mathcal{D}}_R$ whose boundary includes Γ, $||T||_{\Gamma}$ is the L_2 integral norm of the given tractions on Γ and B is a constant which depends only upon the material.

The bound (21) does not depend on the total size of the body and is, therefore, the type of bound which is desired when one wishes to consider unbounded domains as limiting cases of large finite domains.

Finally, we remark that in the context of linear elasticity the arguments in [13] and [16] will go through without the requirement that $\mathcal{D}_R$ be convex.

REFERENCES

[1] Stoker, J. J., Topics in Nonlinear Elasticity, New York University, Courant Institute of Mathematical Sciences Lecture Notes, 1964.

[2] Novozhilov, V.V., Foundations of the Nonlinear Theory of Elasticity, Graylock Press, Rochester, N.Y., 1953.

[3] Bramble, J. H., and L. E. Payne, Some inequalities for vector functions with applications in elasticity. *Archive for Rational Mechanics and Analysis* 11 (1962) 16-26.

[4] Bramble, J. H. and L. E. Payne, A priori bounds in the first boundary value problem in elasticity. *Journal of Research of the National Bureau of Standards, Section B* 65B (1961) 269-276.

[5] Dou, A., Upper estimate of the potential elastic energy of a cylinder. *Communications on Pure and Applied Mathematics* 19 (1966) 83-93.

[6] Gutierrez, A., A bound on the elastic potential energy of a cylinder with square cross section. *Rev. Real Acad. Ci. Exact. Fis. Natur. Madrid* 70 (1976) 549-573.

[7] Villagio, P., Energetic bounds in finite elasticy. *Archive for Rational Mechanics and Analysis* 45 (1972) 282-293.

[8] Aron, M. and J. J. Roseman, Integral estimates for the displacement and strain energy in nonlinear elasticity in terms of the body force. *International Journal of Engineering Science* 15 (1977) 317-322.

[9] Ekeland, I., On the variational principle. *Journal of Mathematical Analysis and Applications* 47 (1974) 324-353.

[10] Aron, M., On the uniqueness of solutions in finite elasticity. *Inst. Lombardo Accad. Sci. Lett. Rend. A* 109 (1975) 424-432.

[11] Breuer, S. and J. J. Roseman, An integral bound for the strain energy in non-linear elasticity in terms of the boundary displacements. *Journal of Elasticity* 9 (1979) 21-27.

[12] John, F., Uniqueness of nonlinear elastic equilibrium for prescribed boundary displacements and sufficiently small strains. *Communications on Pure and Applied Mathematics* 25 (1972) 617-634.

[13] Breuer, S., and J. J. Roseman, Integral bounds on the strain energy for the traction problem in finite elasticity. *Archive for Rational Mechanics and Analysis* 68 (1978) 333-342.

[14] John, F., Rotations and strains. *Communications on Pure and Applied Mathematics* 14 (1961) 391-413.

[15] John, F., Bounds for deformations in terms of average strains. In Inequalities, Vol. III, 129-144. Academic Press, Inc., 1972.

[16] Breuer, S., and J. J. Roseman, A bound on the strain energy for the traction problem in finite elasticity with localized non-zero surface data. *Journal of Elasticity*, to appear.

A FINITE-DIFFERENCE APPROACH TO AXISYMMETRIC PLANE-STRAIN PROBLEMS BEYOND THE ELASTIC LIMIT

P. C. T. Chen
U. S. Army Armament Research and Development Command
Benet Weapons Laboratory, LCWSL
Watervliet, NY 12189

ABSTRACT. A new finite-difference approach has been developed for solving the axisymmetric plane-strain problems subjected to internal or external pressure beyond the elastic limit. The theory used in this paper includes the Prandtl-Reuss flow rules, von Mises' yield criterion and the effective stress-strain data of a material. The stresses and strains in all principal directions can be computed as functions of loading history. The numerical scheme is stable for ideally-plastic as well as strain-hardening materials. The desired accuracy can be achieved by reducing the grid sizes and/or load increments.

I. INTRODUCTION. Based on a detailed study reported in [1], the best material model for gun tubes under high pressure operation is an elastic-plastic material which obeys the Mises yield criterion and the Prandtl-Reuss incremental stress-strain relation. A literature survey indicates that no closed form solution exists even for the axisymmetric plane strain problems. And, in such situations, one has to rely on numerical methods. Both the finite-difference method [2,3] and the finite-element method [4,5] have been used to solve the elastoplastic problem considered here. The finite-element method is more powerful and can be used to solve more general elastoplastic problems. Since the displacement function is assumed and the programming is complicated, the accuracy of the finite element approach has to be verified. This is usually done by comparing with more rigorous solutions to simpler problems. For the problem considered here, rigorous solutions based on the finite-difference method were obtained by Hodge and White [2] for ideally-plastic materials, and by Chu [3] for strain-hardening materials.

In the present paper, a new finite-difference approach is developed for solving the axisymmetric plane strain problems subjected to internal or external pressure beyond the elastic limit. An incremental approach is used and the numerical scheme is stable for ideally-plastic as well as strain-hardening materials. The desired accuracy has been achieved by reducing the grid sizes and load increments.

II. BASIC EQUATIONS. Assuming small strain and no body forces in the axisymmetric state of plane strain, the radial and tangential stresses, σ_r and σ_θ, must satisfy the equilibrium equation,

$$r(\partial\sigma_r/\partial r) = \sigma_\theta - \sigma_r \ ; \tag{1}$$

and the corresponding strains, ε_r and ε_θ, are given in terms of the radial displacement, u, by

$$\varepsilon_r = \partial u/\partial r \ , \quad \varepsilon_\theta = u/r \ . \tag{2}$$

It follows that the strains must satisfy the equation of compatibility

$$r(\partial\varepsilon_\theta/\partial r) = \varepsilon_r - \varepsilon_\theta \ . \tag{3}$$

Whereas the differential equations (1), (2) and (3) hold throughout the tube regardless of the material properties, the constitution equations assume various forms according to the adopted form of yield function, hardening rule, total or incremental theory of plasticity. In the present paper, the material is assumed to be elastic-plastic, obeying Mises' yield criterion, Prandtl-Reuss flow theory and isotropic hardening law. The complete stress-strain relations are [6]:

$$d\varepsilon_i' = d\sigma_i'/2G + (3/2)\sigma_i' d\sigma/(\sigma H') \tag{4}$$

$$d\sigma \geq 0 \quad \text{for} \quad i = r,\theta,z$$

$$d\varepsilon_m = E^{-1}(1-2\nu)d\sigma_m \tag{5}$$

where E, ν Young's modulus, Poisson's ratio, respectively,

$$2G = E/(1+\nu)$$

$$\varepsilon_m = (\varepsilon_r+\varepsilon_\theta+\varepsilon_z)/3 \ , \quad \varepsilon_i' = \varepsilon_i-\varepsilon_m \ ,$$

$$\sigma_m = (\sigma_r+\sigma_\theta+\sigma_z)/3 \ , \quad \sigma_i' = \sigma_i-\sigma_m \ , \tag{6}$$

$$\sigma = (1/\sqrt{2})[(\sigma_r-\sigma_\theta)^2 + (\sigma_\theta-\sigma_z)^2 + (\sigma_z-\sigma_r)^2]^{1/2} \geq \sigma_o \ ,$$

and σ_o is the yield stress in simple tension or compression. For a strain hardening material, H' is the slope of the effective stress/plastic strain curve

$$\sigma = H(\int d\varepsilon^p) \ . \tag{7}$$

For an ideally-plastic material ($H'=0$), the quantity $(3/2)d\sigma/(\sigma H')$ is to be replaced by $d\lambda$, a positive factor of proportionality. When $\sigma < \sigma_o$ or $d\sigma < 0$, the state of stress is elastic and the second term in equation (4) disappears. Following Yamada et al [7], equations (4) and (5) can be rewritten in an incremental form

$$d\sigma_i = d_{ij} d\varepsilon_j \quad \text{for } i,j = r,\theta,z$$

and

$$d_{ij}/2G = \nu/(1-2\nu) + \delta_{ij} - \sigma_i'\sigma_j'/S \ , \tag{8}$$

where

$$S = \frac{2}{3}\left(1 + \frac{1}{3} H'/G\right)\sigma^2 \ , \tag{9}$$

and δ_{ij} is the Kronecker delta.

This form was used in the finite-element formulation for solving elastic-plastic thick-walled tube problems [5]. In the following section, the incremental stress-strain matrix will be used in the finite difference formulation.

III. FINITE-DIFFERENCE FORMULATION. Consider a thick-walled cylinder of inner radius a and external radius b. The tube is subjected to inner pressure p and/or external pressure q. The elastic solution for this problem is well-known and the pressure p* or q* required to cause initial yielding can be determined by using the Mises' yield criterion. For pressure beyond the elastic limit, an incremental approach of the finite-difference formulation is used. The analysis starts with the applied pressure p or q and the loading path is divided into m increments with

$$\Delta p = (p-p^*)/m \ , \quad \Delta q = (q-q^*)/m \ . \tag{10}$$

The cross section of the tube is divided into n rings with

$$r_1=a, r_2, \ldots, r_k=\rho, \ldots, r_{n+1}=b, \tag{11}$$

where ρ is the radius of the elastic-plastic interface. At the beginning of each increment of loading, the distribution of displacements, strains and stresses are assumed to be known and we want to determine Δu, $\Delta\varepsilon_r$, $\Delta\varepsilon_\theta$, $\Delta\sigma_r$, $\Delta\sigma_\theta$, $\Delta\sigma_z$ at all grid points. Since the incremental stresses are related to the incremental strains by the incremental form (8) and $\Delta u = r\Delta\varepsilon_\theta$, there exists only two unknowns at each station that have to

be determined for each increment of loading. The unknown variables in the present formulation are $(\Delta\varepsilon_\theta)_i$, $(\Delta\varepsilon_r)_i$, for $i = 1,2,\ldots n,n+1$.

The equation of equilibrium (1) and the equation of compatibility (3) are valid for both the elastic and the plastic regions of a thick-walled tube. The finite-difference forms of these two equations at $i = 1,\ldots,n$ are given in [3] by

$$(r_{i+1}-2r_i)(\Delta\sigma_r)_i - (r_{i+1}-r_i)(\Delta\sigma_\theta)_i + r_i(\Delta\sigma_r)_{i+1}$$
$$= (r_{i+1}-r_i)(\sigma_\theta-\sigma_r)_i - r_i[(\sigma_r)_{i+1} - (\sigma_r)_i] \tag{12}$$

for the equation of equilibrium, and

$$(r_{i+1}-2r_i)(\Delta\varepsilon_\theta)_i - (r_{i+1}-r_i)(\Delta\varepsilon_r)_i + r_i(\Delta\varepsilon_\theta)_{i+1}$$
$$= (r_{i+1}-r_i)(\varepsilon_r-\varepsilon_\theta)_i - r_i[(\varepsilon_\theta)_{i+1} - (\varepsilon_\theta)_i] \tag{13}$$

for the equation of compatibility.

With the aid of the incremental stress-strain relations (8), equation (12) can be rewritten as

$$[(r_{i+1}-2r_i)(d_{12})_i + (-r_{i+1}+r_i)(d_{22})_i](\Delta\varepsilon_\theta)_i$$
$$+ [(r_{i+1}-2r_i)(d_{11})_i + (-r_{i+1}+r_i)(d_{21})_i](\Delta\varepsilon_r)_i$$
$$+ r_i(d_{12})_{i+1}(\Delta\varepsilon_\theta)_{i+1} + r_i(d_{11})_{i+1}(\Delta\varepsilon_r)_{i+1}$$
$$= (r_{i+1}-r_i)(\sigma_\theta-\sigma_r)_i - r_i[(\sigma_r)_{i+1} - (\sigma_r)_i] \; . \tag{14}$$

The boundary conditions for the problem are

$$\Delta\sigma_r(a,t) = -\Delta p \quad , \quad \Delta\sigma_r(b,t) = -\Delta q \; . \tag{15}$$

Using the incremental relations (8), we rewrite (15) as

$$(d_{12})_1(\Delta\varepsilon_\theta)_1 + (d_{11})_1(\Delta\varepsilon_r)_1 = -\Delta p \; . \tag{16}$$

and

$$(d_{12})_{n+1}(\Delta\varepsilon_\theta)_{n+1} + (d_{11})_{n+1}(\Delta\varepsilon_r)_{n+1} = -\Delta q \; . \tag{17}$$

Now we can form a system of 2(n+1) equations for solving 2(n+1) unknowns, $(\Delta\varepsilon_\theta)_i$, $(\Delta\varepsilon_r)_i$, for i = 1,2,...,n,n+1. Equations (16) and (17) are taken as the first and last equations, respectively, and the other 2n equations are set up at i = 1,2,...,n using (13) and (14). The final system is an unsymmetric band matrix with the nonzero terms clustered about the main diagonal, two below and one above. In the computer program which was developed, the Gaussian elimination method was used to solve these equations. All calculations were carried out on IBM 360/ Model 44 with double precision to reduce round-off errors.

IV. NUMERICAL RESULTS. The axisymmetric plane-strain problems subjected to internal pressure p and external pressure q beyond the elastic limit were solved. The numerical results were based on the following parameters: b/a = 2, ν = 0.3, H' = 0 or E/19. Various values of m and n were used to test the convergence of the numerical solution. The incremental loadings were applied until the fully plastic state was reached. The values for p or q corresponding to this final state were denoted by p** or q**. It was found that the results converge by increasing m and/or n. To achieve 1% accuracy in p** for an ideally-plastic tube with n=50, we shall have m > 200. The results shown in Figures 1 to 5 were based on n=100 and m=200. Figure 1 shows the relations between internal pressure p, external pressure q and elastic-plastic boundary ρ in an elastic-perfectly plastic tube. Figure 2 shows the bore radial and tangential strains as functions of internal pressure p in an ideally-plastic as well as a strain hardening tube. Figure 3 shows the distributions of radial, tangential and axial stress components in a partially-plastic tube subjected to internal pressure. The dotted curves correspond to initial yielding and the solid curves correspond to the case when half of the tube is plastic. For an ideally-plastic tube subjected to external pressure q, the results were presented graphically in Figures 1, 4 and 5.

REFERENCES

1. Davidson, T. E. and Kendall, D. P., "The Design of Pressure Vessels For Very High Pressure Operation," Watervliet Arsenal Report WVT-6917. Also in Mechanical Behavior of Materials Under Pressure (edited by Pugh, H.L.D.), Elsevier Co., 1970, Chapter 2.

2. Hodge, P. G. and White, G. N., "A Quantitative Comparison of Flow and Deformation Theories of Plasticity," J. Appl. Mech., Vol. 17, 1950, pp. 180-184.

3. Chu, S. C., "A More Rational Approach to the Problem of an Elasto-plastic Thick-Walled Cylinder," J. of the Franklin Institute, Vol. 294, 1972, pp. 57-65.

4. Meijus, P., "Elastic-Plastic Deformation of Thick-Walled Cylinders," 1st International Conference on Pressure Vessel Technology, ASME, 1969, Part 1, pp. 19-34.

5. Chen, P.C.T., "The Finite Element Analysis of Elastic-Plastic Thick-Walled Tubes," Proceedings of Army Symposium on Solid Mechanics, 1972, The Role of Mechanics in Design-Ballistic Problems, pp. 243-253.

6. Hill, R., Mathematical Theory of Plasticity, Oxford University Press, 1950.

7. Yamada, Y., Yoshimura, N., and Sakurni, T., "Plastic Stress-Strain Matrix and Its Application for the Solution of Elastic-Plastic Problems by the Finite Element Method," Int. J. Mech. Sci., Vol. 10, 1968, pp. 343-354.

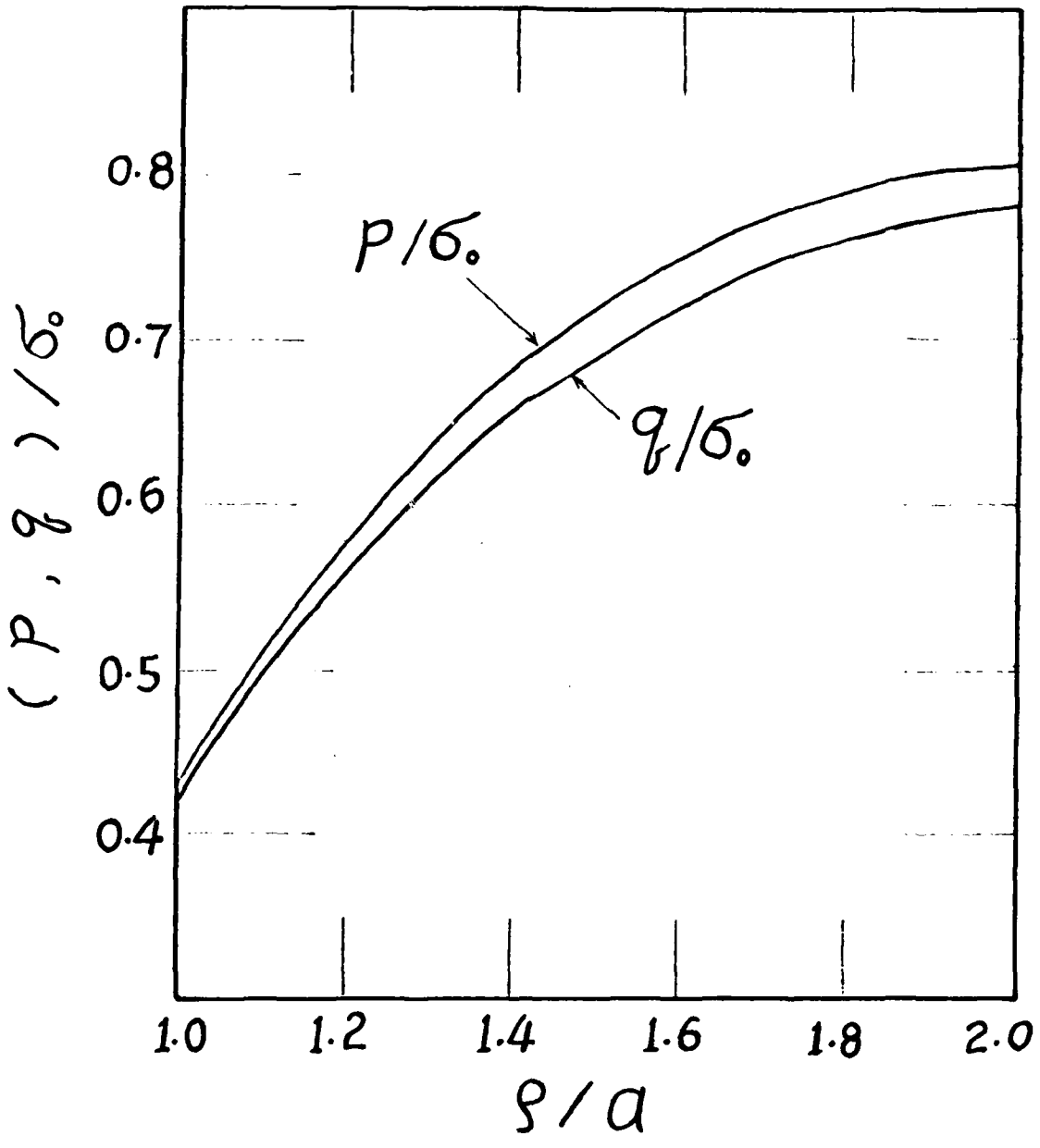


Figure 1. Relations between internal pressure p, external pressure q and elastic-plastic boundary ρ

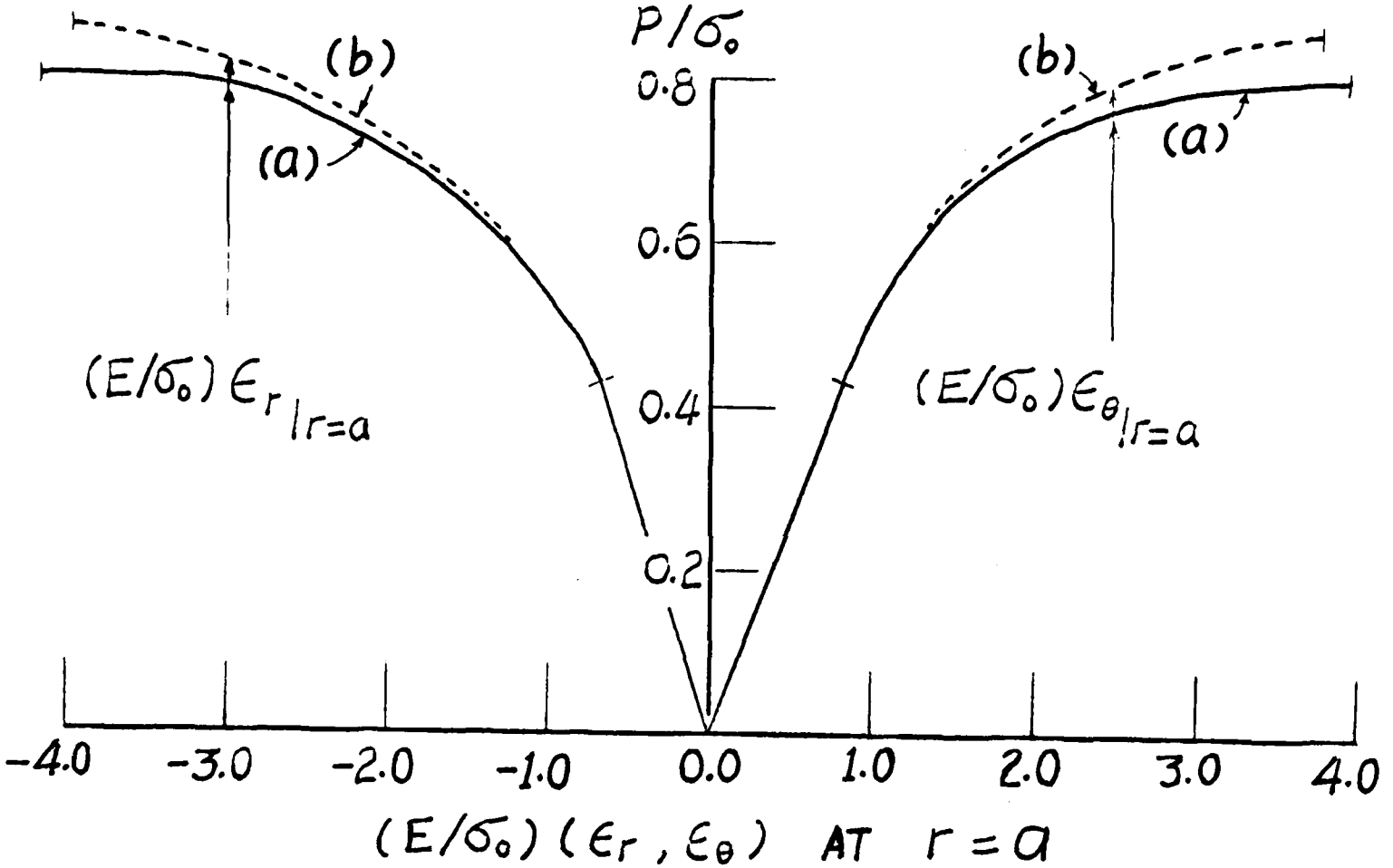


Figure 2. Bore radial and tangential strains as functions of internal pressure p
(a) ideally-plastic tube, (b) strain-hardening tube, H'=E/19

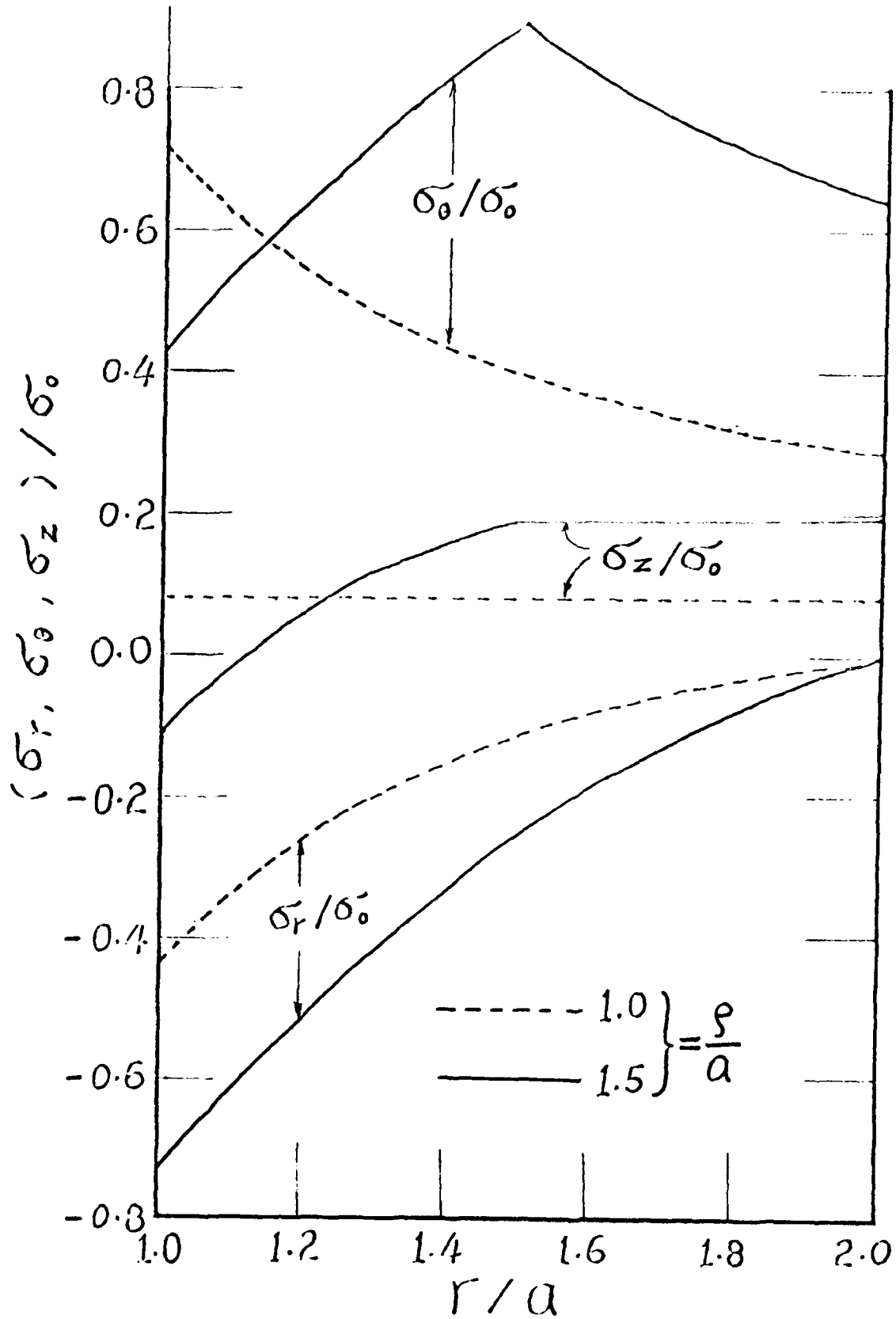


Figure 3. Distributions of radial, tangential and axial stress components in a partially-plastic tube subjected to internal pressure

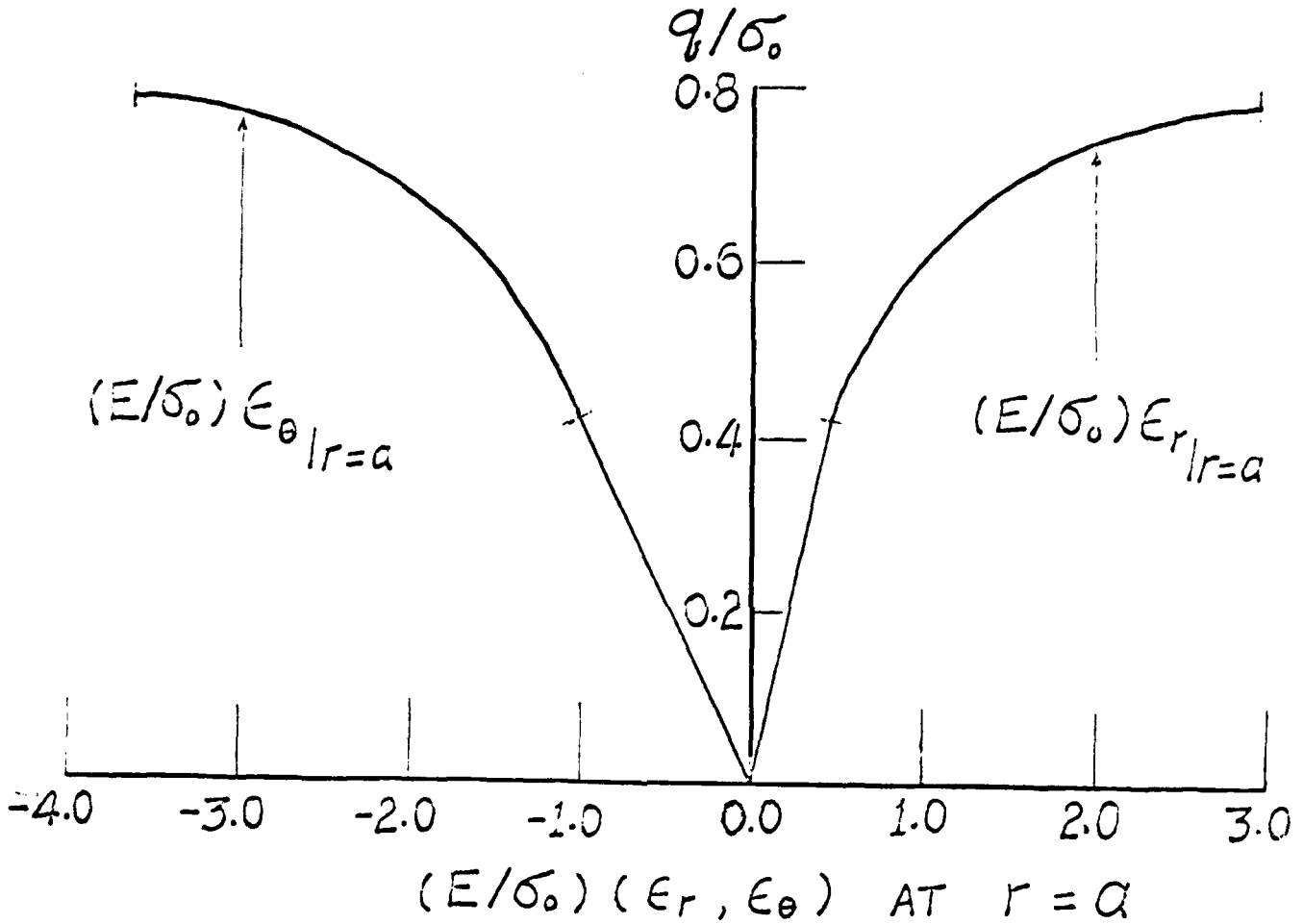


Figure 4. Bore radial and tangential strains as functions of external pressure q in an ideally-plastic tube

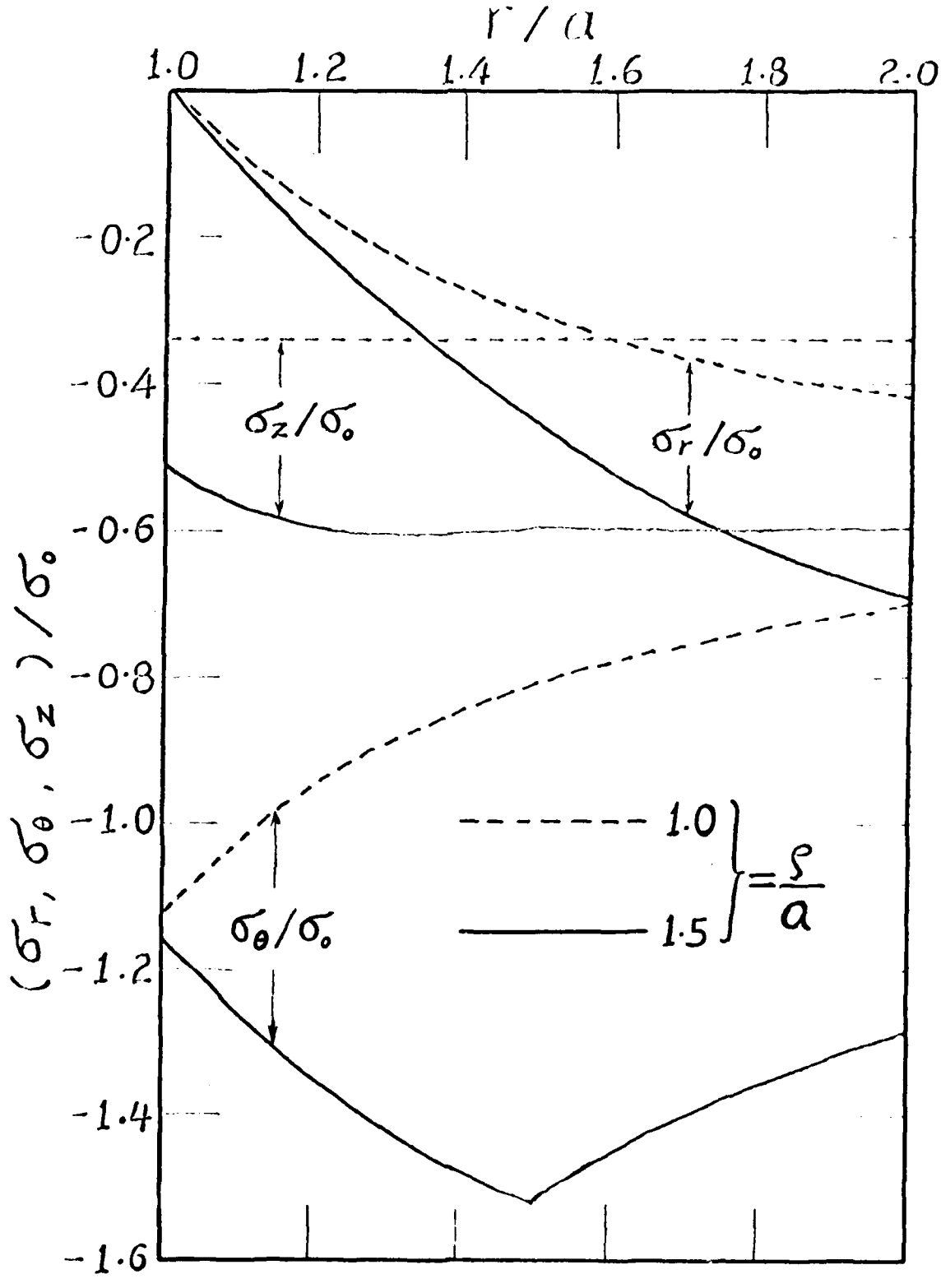


Figure 5. Distributions of radial, tangential and axial stress components in a partially-plastic tube subjected to external pressure q

DEVELOPMENTS IN ELASTIC-PLASTIC FINITE ELEMENT ANALYSIS

Dennis M. Tracey and Colin E. Freese

Mechanics and Engineering Laboratory
Army Materials and Mechanics Research Center
Watertown, Massachusetts 02172

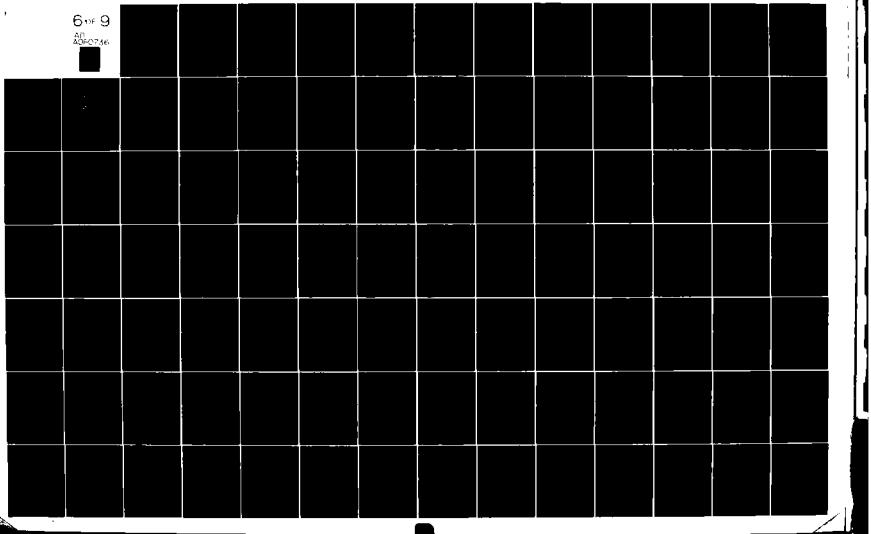
ABSTRACT. The dependence of a solution upon load path discretization is an important consideration in incremental finite element analysis. This issue will be addressed in the context of elastic-plastic finite element analysis using the tangent modulus method. A variable load step approach has been developed which successfully discretizes a given load path during the course of the numerical solution by restricting the structural stiffness approximation at each load step. It produces a set of incremental solutions which are consistently spaced along the load path, as opposed to the arbitrary spacing that often results with approaches requiring a priori discretization. Example solutions demonstrate the accuracy and efficiency benefits derived from the variable load step approach.

INTRODUCTION. We are concerned here with elastic-plastic analysis of material bodies which are governed by the Prandtl-Reuss constitutive relationships. These relationships have an incremental form; increments of stress are related to increments of strain in a manner dependent upon the current stress state and the history of plastic deformation. Hence we concern ourselves with incremental loading finite element formulations which trace the solution history stepwise along the loading path. In particular in this paper, we are concerned with the solution error which is associated with load path discretization, and we restrict attention to this issue for tangent modulus formulations.

Here the view is taken that load path discretization can best be accomplished during the course of the numerical solution by employing a constraint condition which restricts the level of structural stiffness approximation at each load step. In contradistinction, the usual formulation in incremental analysis requires the analyst to specify the step sizes,

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL

6 OF 9
AD A080736



even though in general there is no basis for doing so. In our approach step size is treated as variable at each stage of loading, and is determined along with the usual nodal variables.

We employ a constraint which regulates a field variable known to strongly influence the structural stiffness approximation. Regardless of its form, in the assumed displacement finite element method, the constraint can be represented as

$$g(\underline{\Delta U}_i) = 0 \tag{1}$$

where $\underline{\Delta U}_i$ is the nodal displacement increment vector for the step labeled i. It is convenient to discuss the constraint in the general form; as will become clear it allows discussion of the approach in the broad sense for the general stepwise nonlinear tangent modulus formulation. In application we have concentrated on non-hardening plasticity and have chosen to constrain the yield surface deviatoric stress change in a step. Before discussing our application, however, we will elaborate on the variable load step approach as it would apply to an elementary stepwise linear tangent modulus formulation, one which forms the stiffness matrix $\underline{K}$ on the basis of the initial state alone.

We consider discretization of a proportional loading segment of the load path having a net load change given by the vector $\underline{P}$. The state data sufficient to establish $\underline{K}$ at the beginning of the segment is assumed known. The governing equilibrium equation can be written as

$$\underline{K}\ \underline{\Delta U}_i = \lambda_i\ \underline{P} \tag{2}$$

where λ_i is the undetermined step size scalar. Since $\underline{K}$ is constant, this is a straightforward problem; an arbitrary value λ_i is chosen, equation (2) is solved for $\underline{\Delta U}_i$, and thereupon a scale factor is found so that equation (1) is satisfied, and the scaled $\underline{\Delta U}_i$, λ_i pair constitute the solution for the step. Actually, early elastic-plastic formulations[1,2] took this form, with the constraint limiting the spread of the elastic-plastic boundary.

We will discuss the algorithm developed for formulations which employ stepwise average stiffness matrices defined in terms of the unknown $\underline{\Delta U}_i$, and thus have nonlinear equilibrium equations of the form

$$\underline{K}(\underline{\Delta U}_i)\ \underline{\Delta U}_i = \lambda_i\ \underline{P} \tag{3}$$

The elastic-plastic formulation used in our application has the above stepwise nonlinear form. This results from the averaging of constitutive matrices for points yielding during a step[3], and the use of secant approximations to the Mises yield surface[4] for satisfaction of the yield criterion at the end of a step.

CONSTRAINT CONDITION FOR ELASTIC-PLASTIC PROBLEMS. In the non-hardening elastic-plastic problem, the structural stiffness changes both as a result of elastic-plastic boundary movement and plastic flow direction changes within the plastic zone. We have commented on formulations[1,2] which properly trace boundary motion by a posteriori scaling of load to have yielding occur element by element. Actually, for these formulations there is also corrective reanalysis necessary for consistency of the load/unload decision made for stress states satisfying the yield criterion. Alternately, the boundary movement involving an arbitrary number of grid points can be approximately accommodated by treating the movement as part of the problem and thus solving a nonlinear problem at each step. This is the character of the formulation proposed by Marcal and King[3], a formulation which allows arbitrary step size by employing average constitutive matrices for "transition" points. We have adopted this averaging scheme in the formulation used for our test problems.

Another nonlinear aspect of our numerical formulation follows from use of the Rice and Tracey[4] average flow rule procedure which uses (undetermined) secant approximations to the yield surface for points undergoing plastic flow. Separately or combined, these schemes render the elastic-plastic problem nonlinear with an equilibrium equation of the form (3). Although these averaging methods provide a solution which is consistent with the basic constitutive requirements, the solution nevertheless is approximate - there is load path

discretization dependency. Hence we consider the issues involved in regulating this approximation.

We first consider the changing elastic-plastic boundary and the approximate local stiffnesses of elements involved. The issue is treated relative to the changes of state of points representing the elements, and again we are here allowing for arbitrary sized load steps. In general, a point can experience a number of different stages of deformation during an arbitrary load interval, corresponding to elastic response to yield, followed by various phases of plastic deformation, elastic unloading and reyielding. The simplest interval history would consist of solely elastic behavior with the stress state always below yield. In this case the local stiffness remains constant. A more involved history is depicted in Figure (1). The deviatoric stress space plot illustrates the important states for a point which begins a load interval below yield at stress $\underline{S}_o$, deforms elastically until the incipient yield state $\underline{S}_1$, and then undergoes continuous elastic-plastic deformation reaching a final yield surface stress state $\underline{S}_f$. For this case the local stiffness changes abruptly when $\underline{S}_1$ is reached, and thereafter it changes gradually as the stress state follows the yield surface to $\underline{S}_f$. A complex history entailing more than a single stage of plastic deformation with elastic unloadings would be represented in the plot by distinct stress excursions along and inside of the yield surface from $\underline{S}_1$ to $\underline{S}_f$.

While the above outlines what is possible in an arbitrary step, let us now examine what our averaging techniques can accommodate. The Marcal and King elastic/elastic-plastic partitioning of a step is performed on the basis of the vector $\underline{\Delta U}_i$, the total displacement change over a step. Without any information about variations of displacement rate within the interval, the procedure corresponds to that which applies if the rate were in fact constant. If there is variability, then the partitioning is approximate: e.g., the procedure can accommodate only a single phase of plasticity. Hence there is the need to restrict load step sizes so that they encompass portions of the solution history which involve mildly varying rates in plastically deforming regions. We have not yet considered the

approximate flow rule and how it causes discretization dependency of the solution. It is clear that this approximation is controlled by restricting the stress change $\underline{S}_f - \underline{S}_1$. However, importantly, while $\underline{S}_f - \underline{S}_1$ is being reduced throughout a structure, the displacement rate nonuniformities are likewise being reduced. Hence we have identified a single field variable which directly governs the level of stiffness approximation for our problems, and we are now in a position to define a constraint condition.

If the modulus of $\underline{S}_f - \underline{S}_1$ is denoted by ΔS_{sec}, we regulate load step size so that the maximum value of ΔS_{sec} in the grid results equal to a specified fraction α times the yield stress Y. Hence the constraint equation (1) takes the form

$$g(\underline{\Delta U}_i) = \Delta S_{sec}^{max} - \alpha Y = 0 \tag{4}$$

We have not attempted to establish the relationship between structural stiffness approximation level and the constraint parameter α. For the general problem this does not appear to be possible. We can only state that convergence to the exact solution can be achieved with decreasing α values.

VARIABLE LOAD STEP SOLUTION ALGORITHM. Our elastic-plastic formulation has a nonlinear equilibrium equation which takes the general form (3); however, the explicit form of the equation at each step is itself undetermined. It must be established during an iterative solution process. This problem character follows from the undetermined nature of the plastic zone movement in a step, and the associated stiffness discontinuity that a point experiences at incipient yield.

The customary solution algorithm (for fixed step formulations) involves successively solving for trials $\underline{\Delta U}_i^j$ using a stiffness matrix based upon $\underline{\Delta U}_i^{j-1}$, until convergence. Specifically, at the j-th iteration cycle, the governing equilibrium equation takes the form

$$\underline{K}(\underline{\Delta U}_i^{j-1})\ \underline{\Delta U}_i^j = \lambda_i\ \underline{P} \tag{5}$$

The process starts with a guess $\underline{\Delta U}_i^0$ to establish the first cycle stiffness equation. Strictly speaking, cycling must continue until successive trial solutions are found which are identical, so that $\underline{\Delta U}_i^j$ then satisfies (3).

In our variable load step approach we adjust λ_i during the course of solution to find that $\underline{\Delta U}_i$ which satisfies both the equilibrium equation (3) and the constraint condition (1). This iterative process starts with an estimate of the load step, λ_i^0, as well as with the guess $\underline{\Delta U}_i^0$. For the first cycle the matrix equation for $\underline{\Delta U}_i^1$ takes the form

$$\underline{K}(\underline{\Delta U}_i^0)\ \underline{\Delta U}_i^1 = \lambda_i^0\ \underline{P} \tag{6}$$

In general $\underline{\Delta U}_i^1$ will not satisfy the constraint condition, although there always is a scalar multiple of this vector which will. The operations required to determine the appropriate scale factor depends upon the nature of the constraint. Regardless, the scale factor is found and the correspondingly scaled displacement solution is used as the trial vector for the next cycle of iteration. The next step size trial follows from interpreting the scale factor as being equal to λ_i^1/λ_i^0, as suggested by the linear nature of (6). In general terms the problem after equation (6) is solved to determine λ_i^1 which satisfies

$$g\ (\underline{\Delta U}_i^1 \cdot \lambda_i^1/\lambda_i^0) = 0 \tag{7}$$

The above operations for the first cycle of iteration sets the pattern for subsequent cycles. At cycle j a stiffness matrix is formed according to the estimated displacement $\underline{\Delta U}_i^{j-1} \cdot \lambda_i^{j-1}/\lambda_i^{j-2}$, and a new displacement $\underline{\Delta U}_i^j$ is determined from

$$\underline{K}\ (\underline{\Delta U}_i^{j-1} \cdot \lambda_i^{j-1}/\lambda_i^{j-2})\ \ \underline{\Delta U}_i^j = \lambda_i^{j-1}\ \underline{P} \tag{8}$$

Once $\underline{\Delta U}_i^j$ is obtained, λ_i^j follows from

$$g\ (\underline{\Delta U}_i^j \cdot \lambda_i^j/\lambda_i^{j-1}) = 0 \tag{9}$$

This iterative process is continued until convergence which can be conveniently monitored by the cycle to cycle change in λ_i^j, being that $\underline{\Delta U}_i^j$ and λ_i^j converge concurrently. In our work the convergence test was that iteration terminates when the relative change in λ_i^j in two successive cycles falls below a given tolerance δ. Numerical results have indicated a direct relation between the value of δ and applied load-internal stress imbalance. This is traced to the inherent inconsistency of satisfying the equilibrium equations using a stiffness $\underline{K}(\underline{\Delta U}_i^{j-1})$, while calculating stresses on the basis of $\underline{\Delta U}_i^j$. Hence the value assigned to δ weighs heavily on the ultimate accuracy of a solution, and this must be considered in the accuracy/cost deliberations when undertaking an analysis.

To complete our discussion of the variable load step procedure, we consider some additional restrictions that should be placed on the allowable magnitude of λ_i. When a definite total load vector $\underline{P}$ is specified, there is of course the need to restrict $\lambda_i < 1$. Furthermore, $\Sigma\lambda_i$ over all steps must equal unity. The final step to reach the total load $\underline{P}$ will usually be smaller than that allowed by our constraint condition. For this case the algorithm reverts to the standard fixed load procedure. When the load vector $\underline{P}$ is indefinite in the sense that the final magnitude of its components are not specified, then there is no basis for restricting the values of λ_i. This latter case applies to the test problems considered below. There the vector $\underline{P}$ serves merely to specify load direction and the magnitude increases step by step without restriction until limit load is detected.

NUMERICAL RESULTS. Solutions to two elastic-plastic problems are discussed here to demonstrate the viability of the variable load step approach. First we consider a plate in plane strain which is under imposed uniaxial extension. Hill[7] has given the exact solution to this biaxial stress problem. An interesting aspect of the solution is the load-extension relationship from incipient yield to limit load. Whereas only a slight increase in applied tension is possible after yielding and before uncontrolled plastic deformation occurs, the displacement necessary to reach this limit load state is unbounded. Hence this provides a valuable test case for our solution approach. In Figure

(3), four numerical solutions are given along with the exact solution. The solutions correspond to α values of 0.2, 0.1, 0.05 and 0.025, where α is the freely specified constraint parameter in (4). The solutions were generated for a Poisson's ratio equal to 0.3 and data are presented in plots depicting imposed stress/yield stress versus imposed strain/yield strain, (P/T)/Y vs. (U/L)/(Y/E). These solutions result regardless of whether P or U is taken as the independent loading parameter. The numerical data, labeled with their associated step numbers, are connected to form piecewise linear approximations to the exact solution which is given by the dashed curve on each plot. As would be expected the approximations improve, the step sizes decrease, and the number of steps to final load increases as α is reduced. All solutions were generated by specifying a convergence tolerance value δ equal to 10^{-5}. Three cycles were required to meet this convergence test at each step of each solution.

The plot which dramatically illustrates the worth of the variable load step approach is given in Figure (4). The plot gives the $\alpha = 0.025$ discretization results, in the form of step size versus step number, for both the force P and the displacement U loading conditions. Unless one has a detailed knowledge of the exact solution the discretizations have an unexpected character, suggesting that this problem would entail involved trial and error reanalysis with the standard fixed load approach. The results are displayed relative to P_1 and U_1, the applied force and displacement at incipient yield. The force boundary condition $\Delta P/P_1$ data (marked with triangles) is plotted using the left axis scale. The right axis applies to the displacement boundary condition $\Delta U/U_1$ data (marked with X's). The $\Delta P/P_1$ values vary from 0.0066 for step 2, the first step after incipient yield is reached, to 0.0009 at limit load. The corresponding $\Delta U/U_1$ values range from 0.136 to 0.940.

The second problem involves the plane stress uniaxial tension of a square sheet with a centered circular cutout. This problem is more typical of those encountered in practical analysis in the sense that there are undoubtedly complicated spatial variations in the structure which significantly change character as the yielding progresses. However, a compre-

hensive spatial convergence study was not undertaken, for as throughout this paper the emphasis was placed on load discretization and how it affects the solution for an arbitrary finite element model. A model consisting of four node isoparametric elements was used, having a total of 28 nodes.

Load-extension results are given in Figure (5) for α choices of 0.05 and 0.15. The data are plotted in the normalized form (P/H)/Y vs. (V/H)/(Y/E), where P/H is the uniform tension across the ends of the plate, and V is the displacement of the center of the loaded edge. The step data are numbered for the α = .15 solution and the spread of the plastic zone is illustrated by the shaded elements in the sketches of the model. As can be seen, the significant differences in the two solutions begin at the knee of the curve when yielding loses its localized character.

Loading was continued until step size was reduced to very small fractions of the incipient yield load P_1. The discretization results are plotted in Figure (6). As in the previous problem, it is unlikely that a priori judgement would suggest the form of the results, with $\Delta P/P_1$ starting at values of 0.447 and 0.188 and ending at values in the neighborhood of 0.009 and 0.001 for α equal to 0.15 and 0.05, respectively.

This problem provided useful data concerning the convergence properties of the solution algorithm at a load step. As would be expected, it was found that the typical number of cycles to meet the δ test increases with α value, and furthermore the required number increases at steps near limit load. A δ value of 10^{-3} was used for this problem. Cycle counts averaged close to 5 for α = 0.05, and in the neighborhood of 10 for α = 0.15.

As mentioned earlier, the solutions were obtained using a formulation which employs an average yield surface normal[4] to define the plastic flow rule for yielded points. Special considerations were necessary in adapting this averaging technique to the plane stress problem. In plane stress there is the need to establish the average normal in terms of the out-of-plane direct strain increment, but this strain component depends upon the flow rule for its definition. A method was devised for defining these quantities in a way which insures that both the planar stress condition and the yield condition are satisfied. The details of the method will be described in a forthcoming report[6].

CONCLUSIONS. The numerical results, which now include the load path discretization and corresponding field solution, demonstrate the viability of our variable load step solution algorithm in elastic-plastic analysis. We expect that the algorithm will apply equally well to other nonlinear problems treated by stepwise nonlinear tangent modulus formulations. It is clear that the success of the approach is predicated upon identifying a field variable which controls the level of stiffness approximation, and suitably constraining the variable at each step of the solution. In view of the unpredictable character of the results obtained, the customary fixed load step approach now appears tenuous. The variable step approach not only eliminates the requirement for a priori discretization but also provides a series of consistent incremental solutions according to the desired level of approximation. The analyst need only specify the proportional loading segments of the load path and supply values for the constraint parameter (α in our elastic-plastic problems) and the convergence tolerance parameter δ for the iterative solution at each step. With these parameters the analyst can efficiently examine and control solution accuracy both as regards stiffness approximation level and overall satisfaction of equilibrium.

REFERENCES

1. G. G. Pope, "The Application of the Matrix Displacement Method in Plane Elastic-Plastic Problems," in Proc. Conf. Matrix Methods in Structural Mechanics, Wright-Patterson Air Force Base, 1965, AFFDL-TR-66-80, pp. 635-654.

2. Y. Yamada, N. Yoshimura, and T. Sakurai, "Plastic Stress-Strain Matrix and its Application for the Solution of Elastic-Plastic Problems by the Finite Element Method," Int. J. Mech. Sci., 1968, 10, pp. 343-354.

3. P. V. Marcal and I. P. King, "Elastic-Plastic Analysis of Two-Dimensional Stress Systems by the Finite Element Method," Int. J. Mech. Sci., 1967, 9, pp. 143-155.

4. J. R. Rice and D. M. Tracey, "Computational Fracture Mechanics," in Numerical and Computer Methods in Structural Mechanics, ed. S. J. Fenves et al., Academic Press, 1973, pp. 585-623.

5. R. Hill, The Mathematical Theory of Plasticity, Clarendon Press, Oxford, 1950, pp. 77-79.

6. D. M. Tracey and C. E. Freese, "Implementation of an Elastic-Plastic Finite Element Formulation," in preparation.

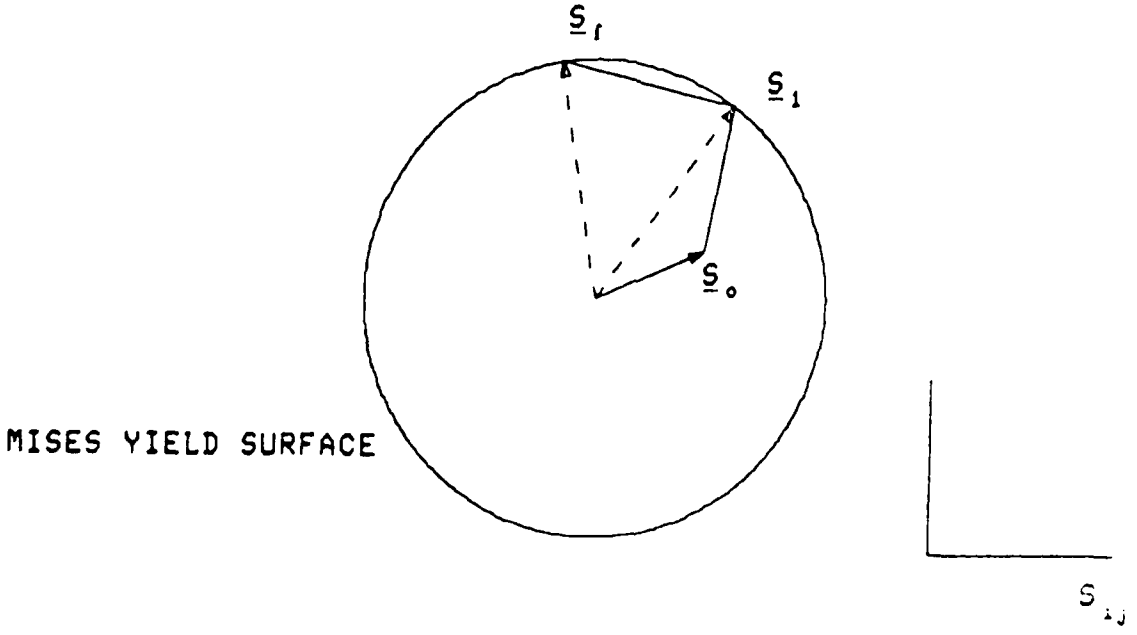
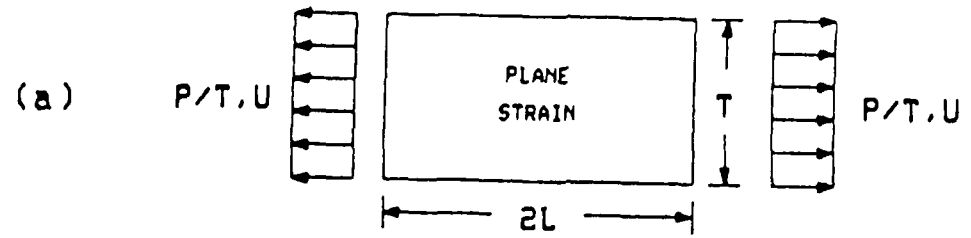
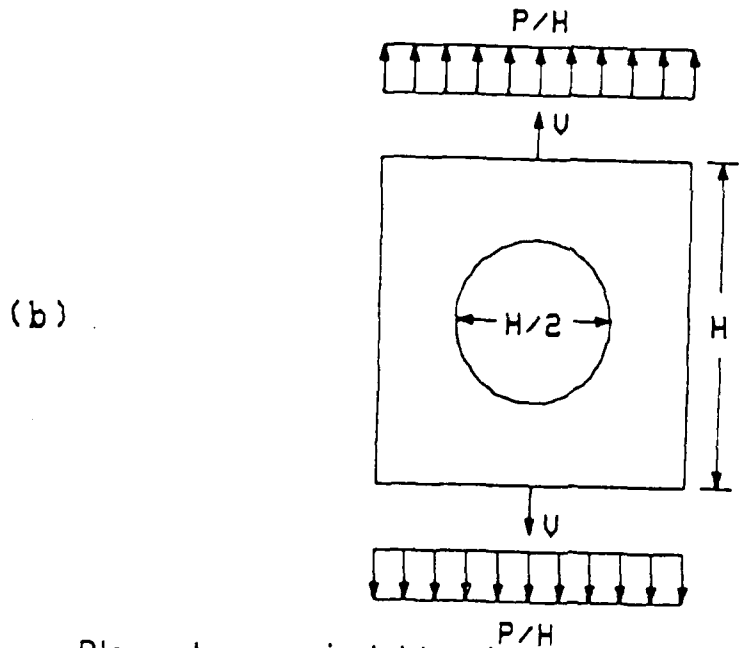


Figure 1. Illustration of deviatoric stress states in example load interval

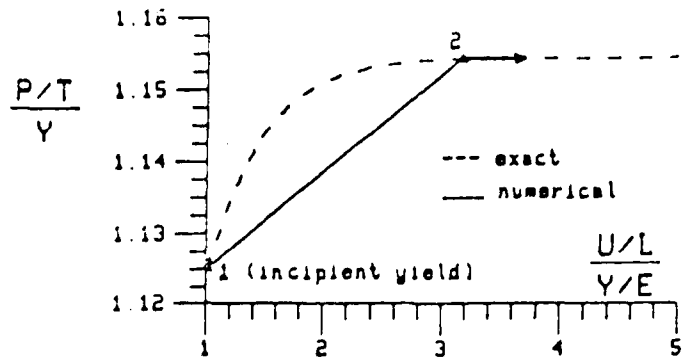


Plane strain uniaxial extension elastic-plastic test problem

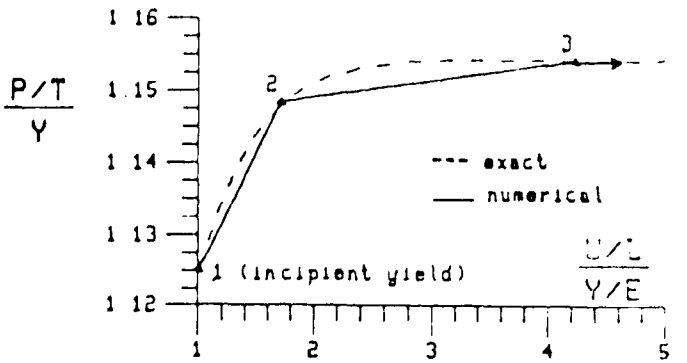


Plane stress uniaxial tension of elastic-plastic weakened sheet

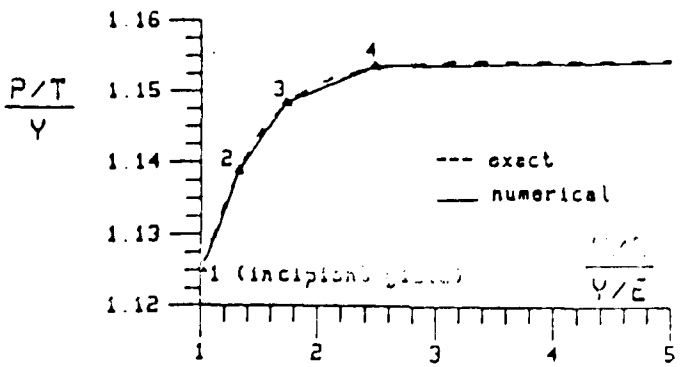
Figure 2



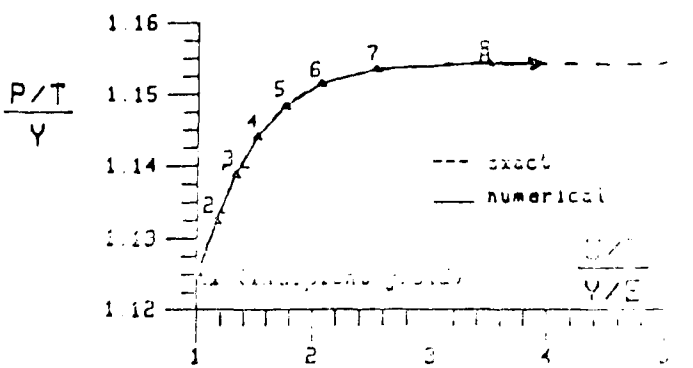
$\alpha = .2$



$\alpha = .1$



$\alpha = .05$



$\alpha = .025$

Figure 3. Load-extension results from incipient yield to limit load for plane strain [illegible]

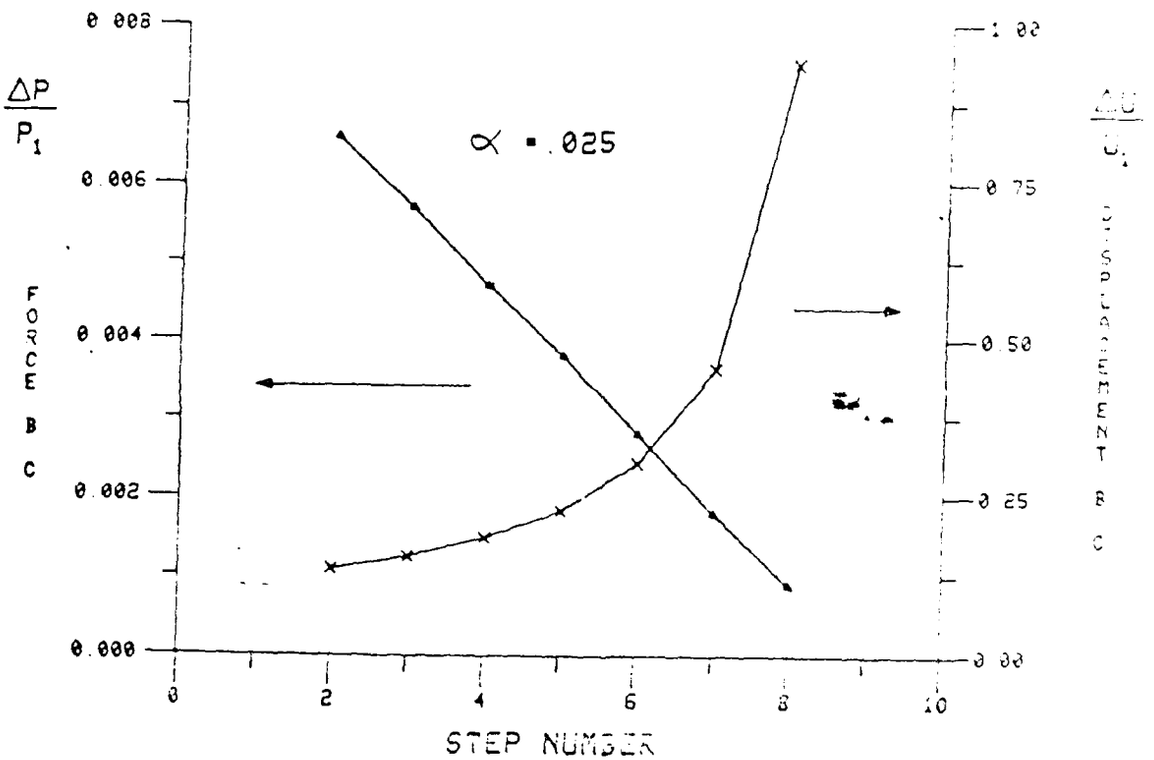


Figure 4. Load discretization results for plane strain problem with constraint parameter equal to 0.025

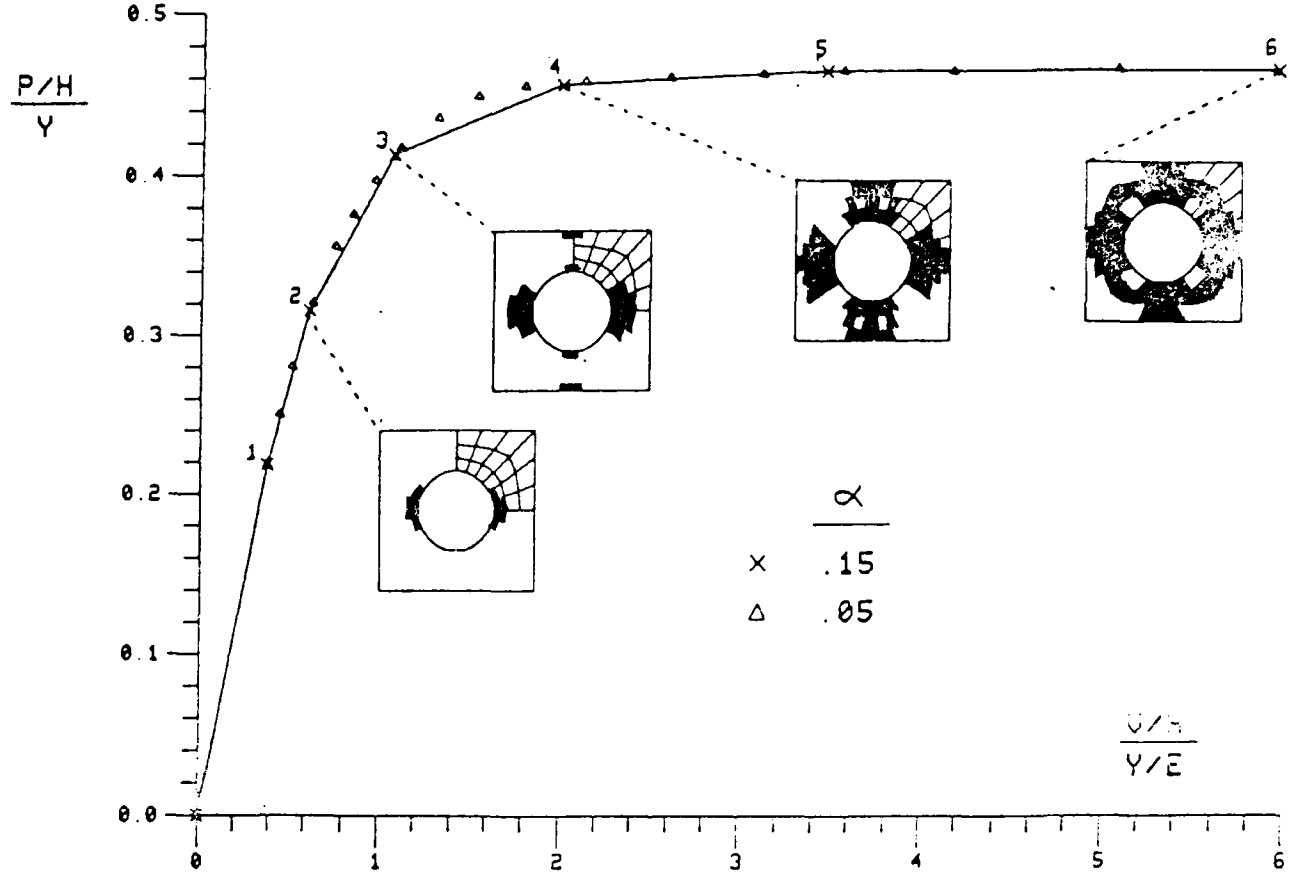


Figure 5. Load-extension results from load free state to limit load for plane stress example corresponding to constraint parameter values of 0.15 and 0.05

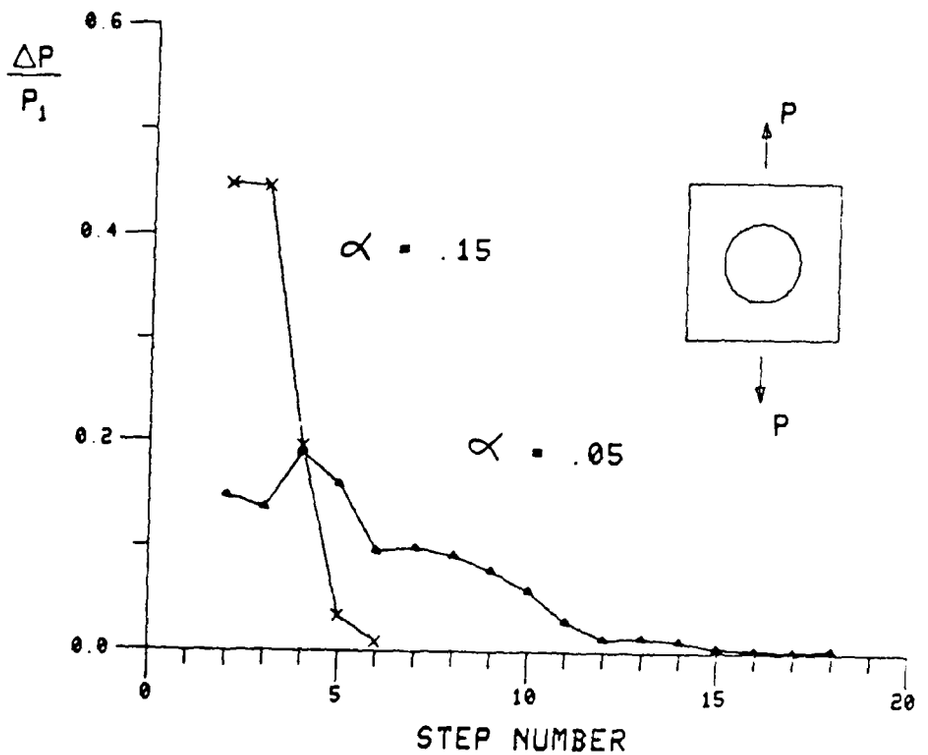


Figure 6. Load discretization results for plane stress problem

STRESS SINGULARITY AT THE VERTEX OF A FLAT WEDGE-SHAPED CRACK BY VARIATIONAL METHOD

M. A. Hussain and S. L. Pu
U.S. Army Armament Research and Development Command
Benet Weapons Laboratory
Watervliet, New York

B. Noble
Mathematics Research Center
Madison, Wisconsin

ABSTRACT. Three dimensional elasticity problems are generally complex. In this paper we present the analysis for the stress singularity at the apex of a three dimensional, flat, wedge-shaped crack under general loadings. The problem is reduced to a set of coupled dual integral equations. Because of the complexity they are not amenable to a closed form solution. A variational method is developed to handle such problems. The physical interpretation of the results are also presented.

I. INTRODUCTION. The theory of fracture mechanics has been a very successful tool in engineering application in recent years. This is mainly due to the use of a single characteristic parameter namely the stress intensity factor, that is the coefficient of the stress singularity at the tip of a crack in the linear theory of elasticity. In most of the two dimensional cracks, in homogeneous media, the singularity is of the order one half. For the three dimensional cracks, however, the singularity depends upon the geometric configurations.

In this paper we study the singularities at the apex of a thin wedge-shaped crack shown in Figure 1 under three loading conditions. Using the near field approach, the problem is reduced to an eigenvalue problem for coupled dual integral equations. The results indicate that cracks tend to straighten out at the apex.

In the following sections we first prove the variational theorm by which the eigenvalue of coupled integral equations is to be obtained. This eigenvalue problem is not the linear one commonly encountered in mathematical physics. Next we present Papkovitch stress function approach to three dimensional theory of elasticity. Then coupled dual series relations are obtained by using mixed boundary conditions. These series are transformed to coupled singular integral equations. Finally the variational method is applied to the coupled integral equations to obtain eigenvalues.

The variational method completely avoids the solution of complex singular integral equations. In this study Macsyma was found to be an indispensable tool at all levels of analysis.

II. VARIATIONAL THEOREM. Consider the following pair of homogeneous coupled integral equations with Fredholm kernels.

$$\int_0^\alpha K_{11}(\phi,\psi;\mu)f(\psi)d\psi + \int_0^\alpha K_{12}(\phi,\psi;\mu)g(\psi)d\psi = 0 \qquad (1)$$

$$\int_0^\alpha K_{21}(\phi,\psi;\mu)f(\psi)d\psi + \int_0^\alpha K_{22}(\phi,\psi;\mu)g(\psi)d\psi = 0 \qquad (2)$$

where eigenvectors f and g and the eigenvalue μ are unknown and $K_{11}(\phi,\psi;\mu)$ etc. involve μ in a linear or nonlinear fashion. Construct the following characteristic equation for the determination of μ^* with appropriate trial functions $f^*(\psi)$ and $g^*(\psi)$.

$$\left(\int_0^\alpha f^*(x)\int_0^\alpha K_{11}(x,y)f^*(y)dydx\right)\left(\int_0^\alpha g^*(\phi)\int_0^\alpha K_{22}(\phi,\psi)g^*(\psi)d\psi d\phi\right)$$
$$-\left(\int_0^\alpha f^*(x)\int_0^\alpha K_{12}(x,y)g^*(y)dydx\right)\left(\int_0^\alpha g^*(\phi)\int_0^\alpha K_{21}(\phi,\psi)f^*(\psi)d\psi d\phi\right) = 0 \qquad (3)$$

If f^* and g^* vary around the exact solutions f and g as

$$f^*(\phi) = f(\phi) + \delta\xi(\phi) \quad , \quad g^*(\phi) = g(\phi) + \delta\eta(\phi) \qquad (4)$$

then $(\mu^*-\mu)$ is stationary around δ as δ approaches zero. That is

$$\mu^* = \mu + O(\delta^2)$$

provided

$$K_{11}(\phi,\psi) = K_{11}(\psi,\phi)$$
$$K_{22}(\phi,\psi) = K_{22}(\psi,\phi) \qquad (5)$$
$$K_{12}(\phi,\psi) = K_{21}(\psi,\phi)$$

Proof: Symbolically we write equation (3) as

$$(f^*K_{11}f^*)(g^*K_{22}g^*) - (f^*K_{12}g^*)(g^*K_{21}f^*) = 0 \qquad (6)$$

where

$$f^* = f + \delta\xi \quad , \quad g^* = g + \delta\eta \tag{7}$$

Expanding the kernels around μ, we have

$$K_{11}(\mu) = K_{11}(\mu^*) + \Delta K'_{11}(\mu^*) + 0(\Delta^2) \quad , \quad \text{etc} \tag{8}$$

where

$$\Delta = \mu - \mu^* \tag{9}$$

Substituting from (7) and (8) into (6) and using (5) we obtain

$$\begin{aligned}
&[(fK_{11}f)(gK_{22}g) - (fK_{12}g)(fK_{12}g)] \\
&+ 2\delta[(\xi K_{11}f)(gK_{22}g)-(\xi K_{12}g)(fK_{12}g)+(\eta K_{22}g)(fK_{11}f)-(fK_{12}\eta)(fK_{12}g)] \\
&+ L_1\Delta + L_2\delta^2 + \text{higher order terms} = 0
\end{aligned} \tag{10}$$

where

$$L_1 = (fK_{11}f)(gK'_{22}g)+(fK'_{11}f)(gK_{22}g)-2(fK_{12}g)(fK'_{12}g) \tag{10a}$$

$$\begin{aligned}
L_2 = {} & (\xi K_{11}\xi)(gK_{22}g)-(\xi K_{12}g)^2+4(fK_{11}\xi)(gK_{22}\eta) \\
& + (fK_{11}f)(\eta K_{22}\eta)-(fK_{12}\eta)^2-4(fK_{12}\eta)(\xi K_{12}g)
\end{aligned} \tag{10b}$$

Using equations (1) and (2) it is seen that the first two terms in (10) vanish. Hence

$$\Delta = \delta^2(L_2/L_1) + 0(\delta^3) \tag{11}$$

The above proves that Δ is of the second order in δ. This completes the proof of the theorem. Because the eigenvalues involved in kernels can be in a nonlinear fashion we cannot prove its bounds as can be done for the linear case in [1].

III. PAPKOVITCH STRESS FUNCTIONS. In the absence of body force the equation of equilibrium of a homogeneous isotropic elastic solid is given by

$$\nabla^2\bar{u} + \frac{1}{1-2\nu}\nabla\nabla\cdot\bar{u} = 0 \tag{12}$$

Using Helmholtz decomposition theorem the solution of (12) can be written as [2,3]

$$2G\bar{u} = \nabla B_o + \nabla(\bar{r}\cdot\bar{B}) - 4(1-\nu)\bar{B} \tag{13}$$

where G and ν are shear modulus and Poisson's ratio, B_o and $\bar{B}$ are known as Papkovitch stress functions satisfying

$$\nabla^2 B_o = 0 \ , \quad \nabla^2 \bar{B} = 0 \ , \quad \bar{B} = \bar{i}\psi + \bar{j}\omega + \bar{k}\lambda \tag{14}$$

For computational purposes it is convenient to write the solution as a superposition of the following basic solutions.

$$\begin{aligned} &\text{1st Basic solution:} && 2G\bar{u} = \nabla B_o \\ &\text{2nd Basic solution:} && 2G\bar{u} = \nabla(x\psi) - 4(1-\nu)\psi\bar{i} \\ &\text{3rd Basic solution:} && 2G\bar{u} = \nabla(y\omega) - 4(1-\nu)\omega\bar{j} \\ &\text{4th Basic solution:} && 2G\bar{u} = \nabla(z\lambda) - 4(1-\nu)\lambda\bar{k} \end{aligned} \tag{15}$$

These basic solutions were transformed to the following spherical coordinate system

$$x = r\sin\theta\cos\phi \ , \quad y = r\sin\theta\sin\phi \ , \quad z = r\cos\theta \tag{16}$$

with the origin at the apex of the crack as shown in Figure 2. This enables us to obtain near field solutions and to study the stress singularity at the apex. The complete results of components of displacement and stress for basic solutions are given in Appendix 1. Since we are interested in the power singularity at the apex, so we choose Papkovitch potentials in the form

$$H(r,\theta,\phi) = r^{\mu} H_1(\theta,\phi) \tag{17}$$

where μ is the eigenvalue to be determined. As can be seen from Appendix 1, stress components will be of the form

$$\sigma = O(r^{\mu-1}) \tag{18}$$

which will be singular when $\mu < 1$. For the displacements to be finite we seek positive eigenvalues between 0 and 1.

The near field geometry surrounding the apex permits us to write (17) as the separation of variables solution

$$H(r,\theta,\phi) = \sum_{m} r^{\mu} P_{\mu}^{m}(\cos\theta) \begin{matrix}\cos m\phi \\ \text{or} \\ \sin m\phi\end{matrix} \qquad (19)$$

where $P_{\mu}^{m}(x)$ is the associated Legendre function of the first kind with degree μ and order m. When (19) is substituted into the four basic solutions given in Appendix 1 we see that the forms of some solutions thus obtained are not convenient to work with. The final solutions used in this analysis are designated as solutions A, B, C and D which are given in Appendix 2. Solution A was obtained by replacing μ by $\mu+1$ after the substitution from (19) into the first basic solution. Solution B is obtained by adding the second and the third basic solutions with proper trigonometric functions in (19), replacing m by m+1 and then using Legendre recursion formulae. A similar process was used to obtain solution C. Solution D simply comes from the fourth basic solution with the use of (19).

Solutions A, B, C, and D must not be linearly independent. This is due to the fact that the condition that the vector point function is solenoidal has not been used explicitly in Papkovitch stress functions approach. We found these solutions are indeed not independent. A relation among them can be written symbolically in the form

$$(\mu+m+1)[C] = [B] + 2(\mu+m+1)[D] - 2(\mu-3+4\nu)[A] \qquad (20)$$

Hereafter solution C is replaced by solutions A, B, D using (20).

IV. THREE MODES OF CRACKS AND COUPLED INTEGRAL EQUATIONS. For a crack shown in Figure 2, the leading edges of the crack are $\phi = \pm \alpha$ and the crack is in the x-y plane ($\theta = \pi/2$). Let D^- and D^+ be the cracked and uncracked region of the plane $\theta = \pi/2$. Within the cracked region, the displacement is discontinuous. If the discontinuity is in the z-direction ($u_{\theta}^{+} - u_{\theta}^{-}$ = finite), the crack is under mode I; if the discontinuity is in the x-direction ($u_{x}^{+} - u_{x}^{-}$ = finite), the crack is defined to be under mode II; and if $u_{y}^{+} - u_{y}^{-}$ = finite, the crack is defined to be under mode III. Boundary conditions for various modes are tabulated below.

BOUNDARY CONDITIONS ON $\theta = \pi/2$

	Non Mixed Conditions (in $D^- + D^+$)	Mixed Conditions in D^-	in D^+	
Mode I	$\tau_{\theta r} = \tau_{\theta\phi} = 0$	$\sigma_{\theta} = 0$	$u_{\theta} = 0$	(21)
Mode II	$\sigma_{\theta} = 0$	$\tau_{\theta r} = \tau_{\theta\phi} = 0$	$u_r = u_{\phi} = 0$	(22)
Mode III	$\sigma_{\theta} = 0$	$\tau_{\theta r} = \tau_{\theta\phi} = 0$	$u_r = u_{\phi} = 0$	(23)

For mode I, u_θ is even in ϕ. This leads to the use of the trigonometric functions at the top of (19). The boundary conditions of (22) and (23) are identical, but the symmetric properties are different for mode II and mode III. In the former case, u_r is even and u_ϕ is odd in ϕ while in the latter, the reverse is true. Hence the proper set of quantities should be selected in (19), for each case.

IV.A. Mode I. Using Eq. (19) and the non-mixed conditions of (21), we have

$$B_m = 0 \ , \quad A_m = (m+\mu+1)^{-1}(1-2\nu)D_m \tag{24}$$

The mixed boundary conditions of (21) and using (24) and (19), yield

$$\begin{aligned} \sum b_m \cos m\phi &= 0 \qquad 0\leq\phi<\alpha \\ \sum q_m b_m \cos m\phi &= 0 \qquad \alpha<\phi\leq\pi \end{aligned} \tag{25}$$

where $\sum$ denotes the summation with respect to m for m=0,1,2...∞, and

$$b_m = (-m+\mu+1)D_m P^m_{\mu+1} \ , \quad q_m = (-m+\mu+1)^{-1}P^m_\mu/P^m_{\mu+1} \ , \quad P^m_\mu = P^m_\mu(0) \tag{26}$$

IV.B. Mode II. For the homogeneous condition of σ_θ in (22), using (19), the coefficients A, B and D are related in the form

$$2(1-\nu)D_m = (m+\mu+1)A_m + [\mu^2+2\nu(\mu+1)-2(1-\nu)m-1]B_m$$

This relation and the mixed conditions of (22) yield the following coupled dual series:

$$\begin{aligned} \sum E_m \cos m\phi &= 0 \qquad 0\leq\phi<\alpha \\ \sum (R_m E_m + S_m F_m)\cos m\phi &= 0 \qquad \alpha<\phi\leq\pi \end{aligned} \tag{27}$$

$$\begin{aligned} \sum{}' F_m \sin m\phi &= 0 \qquad 0\leq\phi<\alpha \\ \sum{}' (U_m F_m + T_m E_m)\sin m\phi &= 0 \qquad \alpha<\phi\leq\pi \end{aligned} \tag{28}$$

where $\sum'$ denotes the summation with respect to m for m = 1, 2, ...∞ and

$$E_m = -(m+\mu+1)\{\mu A_m + [4m(1-\nu)^2+\mu(\mu+4\nu-m-3)]B_m\}P^m_\mu \tag{29}$$

$$F_m = (m+\mu+1)\{mA_m + [4\mu(1-\nu)^2+m(\mu+4\nu-m-3)]B_m\}P^m_\mu \tag{30}$$

$$R_m = [m^2-\mu(\mu+1)(1-\nu)]V_m \ , \quad S_m = m(1-\nu-\nu\mu)V_m$$
$$U_m = [(1-\nu)m^2-\mu(\mu+1)]V_m \ , \quad T_m = m(1+\nu\mu)V_m \tag{31}$$

in (31) V_m stands for $P^m_{\mu+1}/P^m_\mu/[(m+\mu+1)(m^2-\mu^2)]$.

IV.C. Mode III. Similar to the preceding case, we have

$$\sum{}' E_m \sin m\phi = 0 \qquad 0\leq\phi<\alpha$$
$$\sum{}' (R_mE_m+S_mF_m)\sin m\phi = 0 \qquad \alpha<\phi\leq\pi \tag{32}$$

$$\sum F_m \cos m\phi = 0 \qquad 0\leq\phi<\alpha$$
$$\sum (U_mF_m+T_mE_m)\cos m\phi = 0 \qquad \alpha<\phi\leq\pi \tag{33}$$

It was found that the dual series (25) for mode I is identical to that of a potential problem studied in [4], [5], [6] and will not be discussed here.

For convenience we make the following change of variables

$$\phi = \pi-\omega \ , \quad \alpha' = \pi-\alpha \ , \quad (-1)^mE_m = E_m' \ , \quad (-1)^mF_m = F_m'$$
$$R_m = R_m'(1-\nu-\nu\mu) \ , \quad S_m = S_m'(1-\nu-\nu\mu) \tag{34}$$
$$U_m = U_m'(1+\nu\mu)(-1) \ , \quad T_m = T_m'(1+\nu\mu)(-1)$$

The dual series of (27), (28) now become

$$\sum E_m' \cos m\omega = 0 \qquad \alpha'<\omega\leq\pi \tag{35}$$
$$\sum{}' F_m' \sin m\omega = 0 \tag{36}$$
$$\sum (R_m'E_m'+S_m'F_m')\cos m\omega = 0 \qquad 0\leq\omega<\alpha' \tag{37}$$
$$\sum{}' (T_m'E_m'+U_m'F_m')\sin m\,\omega = 0 \tag{38}$$

Let the right hand sides of (35) and (36) for the interval $0<\omega<\alpha'$ be denoted by unknown functions $f(\omega)$ and $g(\omega)$, respectively. The Fourier inversion gives

$$E_o' = \frac{1}{\pi}\int_0^{\alpha'} f(\psi)d\psi \ , \qquad E_m' = \frac{2}{\pi}\int_0^{\alpha'} f(\psi)\cos m\psi d\psi$$
$$F_m' = \frac{2}{\pi}\int_0^{\alpha'} g(\psi)\sin m\psi d\psi \tag{39}$$

Substituting from (39) into (37), (38) and interchanging the order of summation and integration we have the following coupled integral equations for the determination of f and g,

$$\int_0^{\alpha'} K_{11}(\omega,\psi;\mu)f(\psi)d\psi + \int_0^{\alpha'} K_{12}(\omega,\psi;\mu)g(\psi)d\psi = 0 \tag{40}$$

$$\int_0^{\alpha'} K_{21}(\omega,\psi;\mu)f(\psi)d\psi + \int_0^{\alpha'} K_{22}(\omega,\psi;\mu)g(\psi)d\psi = 0 \tag{41}$$

where

$$\begin{aligned} K_{11}(\omega,\psi;\mu) &= \frac{1}{2} R_o' + \sum{}' R_m' \cos m\omega \cos m\psi \\ K_{12}(\omega,\psi;\mu) &= \sum{}' S_m' \cos m\omega \sin m\psi \\ K_{21}(\omega,\psi;\mu) &= \sum{}' T_m' \sin m\omega \cos m\psi \\ K_{22}(\omega,\psi;\mu) &= \sum{}' U_m' \sin m\omega \sin m\psi \end{aligned} \tag{42}$$

It can be shown that [7]

$$\frac{P_{\mu+1}^m}{P_\mu^m} = \frac{-2}{m-\mu-1} \Gamma\left(\frac{m-\mu+1}{2}\right)\Gamma\left(\frac{m+\mu+2}{2}\right)/\Gamma\left(\frac{m-\mu}{2}\right)/\Gamma\left(\frac{m+\mu+1}{2}\right) . \tag{43}$$

Using (43) and (34) the following asymptotic expansions can be established.

$$\begin{aligned} R_m' &= -(1-\nu-\mu\nu)^{-1}(1/m) + O(1/m^2) \quad , \quad S_m' = 1/m^2 + O(1/m^3) \\ U_m' &= [-(1-\nu)/(1+\nu\mu)](1/m) + O(1/m^2) \quad , \quad T_m' = 1/m^2 + O(1/m^3) \end{aligned} \tag{44}$$

Substituting from (44) into (42) and summing the dominant part of the series we have [4], [8]

$$K_{11}(\omega,\psi) = \frac{1}{1-\nu-\nu\mu}\frac{1}{2} \log 2|\cos\omega-\cos\psi| + \text{regular terms} \tag{45}$$

$$K_{22}(\omega,\psi) = \frac{-(1-\nu)}{2(1+\nu\mu)} \log\left|\sin\left(\frac{\omega+\psi}{2}\right)/\sin\left(\frac{\omega-\psi}{2}\right)\right| + \text{regular terms} \tag{46}$$

The previous expressions show the kernels have logarithmic singularities. Similar analysis can be carried out for equations (32) and (33) for mode III (interchanging R_m' and U_m', S_m' and T_m').

V. APPLICATION OF VARIATIONAL PRINCIPLE. Equations (40) and (41) are identical to equations (1) and (2) and all the conditions of the theorem required for the kernels are satisfied. In this section we shall apply the theorem to obtain approximate eigenvalues by assuming approximate trial functions. Without causing ambiguity we shall drop the asterisks and assume the following trial functions,

$$\begin{aligned} f(t) &= (\beta_o + \cos t)\cos(t/2)/(\cos t - \cos\alpha')^{1/2} \\ g(t) &= [(1-\cos\alpha') + 2\cos t]\sin(t/2)/(\cos t - \cos\alpha')^{1/2} \end{aligned} \tag{47}$$

where

$$\beta_o = -\cos^2(\alpha'/2)\,\frac{1+(1-\nu-\nu\mu)R_o' + \log\sin^2(\alpha'/2)}{(1-\nu-\nu\mu)R_o' + \log\sin^2(\alpha'/2)} \tag{48}$$

The above trial functions are the first approximations to the integral equations (40) and (41) with kernels (42) replaced by their dominant parts given by (44). The method of obtaining such solutions by direct computation is illustrated in [9].

Substituting from (47) into the characteristic equation (3), changing the order of summation and integration and using the integral representation of Legendre functions [7], we obtain

$$I_{11}I_{22} - I_{12}^2 = 0 \tag{49}$$

where

$$\begin{aligned} I_{11} &= 2R_o'(1+2\beta_o+P_1)^2 + {\sum}' R_m'[P_{m+1} + (1+2\beta_o)(P_m+P_{m+1}) + P_{m-2}]^2 \\ I_{22} &= 4{\sum}' U_m'[P_{m+1} - P_1(P_m-P_{m-1}) - P_{m-2}]^2 \\ I_{12} &= -2{\sum}' S_m'[P_{m+1} + (1+2\beta_o)(P_m+P_{m-1}) + P_{m-2}][P_{m+1} - P_1(P_m-P_{m-1}) - P_{m-2}] \\ P_m &= P_m(\cos\alpha') \end{aligned} \tag{50}$$

For different values of ν and α' (the complementary angle to half of the vertex angle of the wedge-shaped crack) we get approximate values of μ from equation (49) for both modes II and III.

The form of (49) is very well suited for Macsyma evaluation and using eight terms for summations in (50), the results for μ are marked by x in Figure 3 where the solid lines are results obtained by another method [9]. The results by both methods are in good agreement.

A further refinement of results can be obtained by selecting

$$f(t) = (A + B\cos t)\frac{\cos(t/2)}{(\cos t - \cos\alpha')^{1/2}}$$

$$g(t) = (C + D\cos t)\frac{\sin(t/2)}{(\cos t - \cos\alpha')^{1/2}} \qquad (51)$$

and formally extending the variational technique to a characteristic equation obtained from the vanishing of the determinant of a four by four system (A, B, C, D in (51) must not all vanish). The results, using Macsyma and summing to a maximum of eight terms in (50), are compared in the following table. They also are shown as · in Figure 3.

VALUES OF μ FOR $\nu = 0.25$

Mode	Half Vertex Angle α	Variational Method Using (47)	Variational Method Using (51)	Direct Method [9]
II	0.1π	0.9333	0.9616	0.9582
	0.3π	0.6489	0.6961	0.6953
	0.5π	0.4752	0.5107	0.5017
	0.7π	0.3585	0.3749	0.3654
	0.9π	0.2213	0.2469	0.2137
III	0.1π	0.9739		
	0.3π	0.7642	0.8253	0.8270
	0.5π	0.4881	0.5192	0.5027
	0.6π	0.3829	0.4034	0.3914
	0.7π	0.3021		0.3015
	0.8π	0.2372		0.2335

Even the rigorous proof of the extension of the variational technique from the two by two system using trial functions (47) to the four by four system using trial functions (51) is still to be done, the results thus obtained are in good agreement with results achieved by using trial functions (47) or other methods [9].

VI. CONCLUSIONS. For modes II and III, the results show that the stress singularities are dominated by the vertex angle as well as the elastic constant ν of the material. The results further indicate that when the apex angle is greater than 180°, the stress singularity is stronger than one half enhancing the tendency of crack front to straighten out. Similarly, when the vertex angle is less than 180°, the stress singularity is less severe than one half and, again, this will tend to retard the growth at the vertex until the crack front straightens out.

Macsyma was extensively used throughout the analysis, especially in the generation and use of special functions such as Legendre functions, Gamma functions, Bessel functions from Share directory, in the summation of series, in the solution of linear equations, in seeking roots of determinants of matrices, in the plot routine and in the creation of file for Batch with Teco, etc. This investigation would have been extremely tedious without Macsyma. The methods as well as results in the full entirety, to our knowledge, do not seem to have appeared in literature.

REFERENCES

1. Barlett, C. C. and Noble, B., "A Variational Method for the Solution of Eigenvalue Problems Involving Mixed Boundary Conditions," Applied Science Research, Section B, Vol. 9, 1962.

2. Green, A. E. and Zerna, W., THEORETICAL ELASTICITY, 2nd Edition, Oxford, 1968.

3. Lure', A. T., THREE DIMENSIONAL PROBLEMS OF THE THEORY OF ELASTICITY, Interscience Publishers, 1964.

4. Noble, B., "The Potential and Charge Distribution Near the Tip of a Flat Angular Sector," EM-135, New York University, NY, 1959.

5. Brown, S. N., and Stewartson, K., "Flow Near the Apex of a Plane Delta Wing," Journal of Institute of Mathematics and Its Applications, Vol. 5, p. 206, 1969.

6. Morrison, J. A., and Lewis, J. A., "Charge Singularity at the Corner of a Flat Plate," SIAM, Journal of Applied Mathematics, Vol. 31, p. 233, 1976.

7. Magnus, W. and Oberhettinger, F., SPECIAL FUNCTIONS OF MATHEMATICAL PHYSICS, Chelsea Publishing Co., 1949.

8. Jolly, L. B. W., SUMMATION OF SERIES, Dover Publications, p. 126, 1961.

9. Noble, B., Hussain, M. A., and Pu, S. L., "Apex Singularities for Corner Cracks Under Opening, Sliding and Tearing Modes," to be published in the Proceedings of International Conference on Fracture Mechanics in Engineering Application, Bangalore, India, 1979.

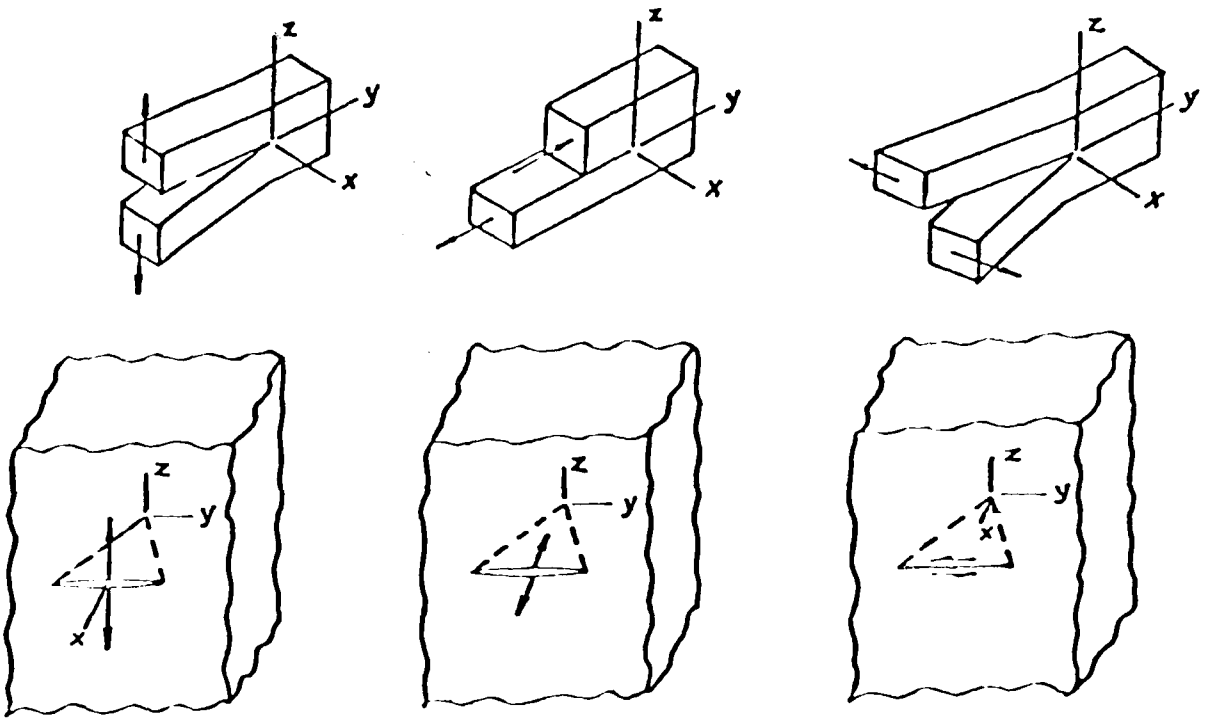


Figure 1. A flat wedge-shaped crack under three different modes and its two-dimensional counterparts.

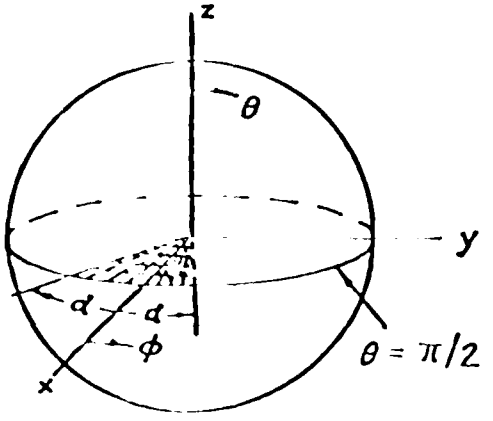


Figure 2. A spherical coordinate surrounding the apex of a thin angular sector crack.

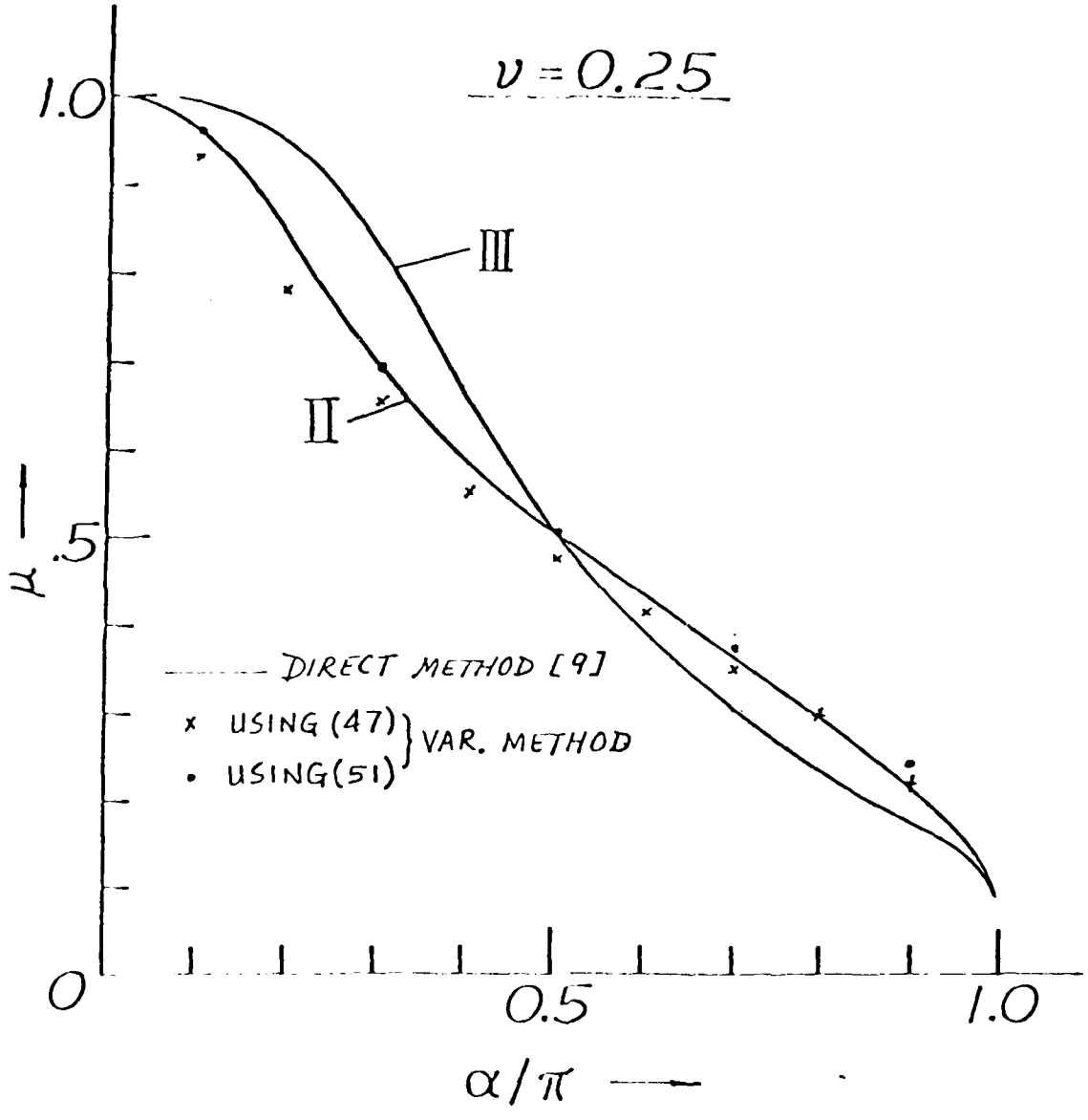


Figure 3. The eigenvalue μ as a function of α/π for $\nu = 0.25$.

(a) Variational results indicated by x and · for mode II.

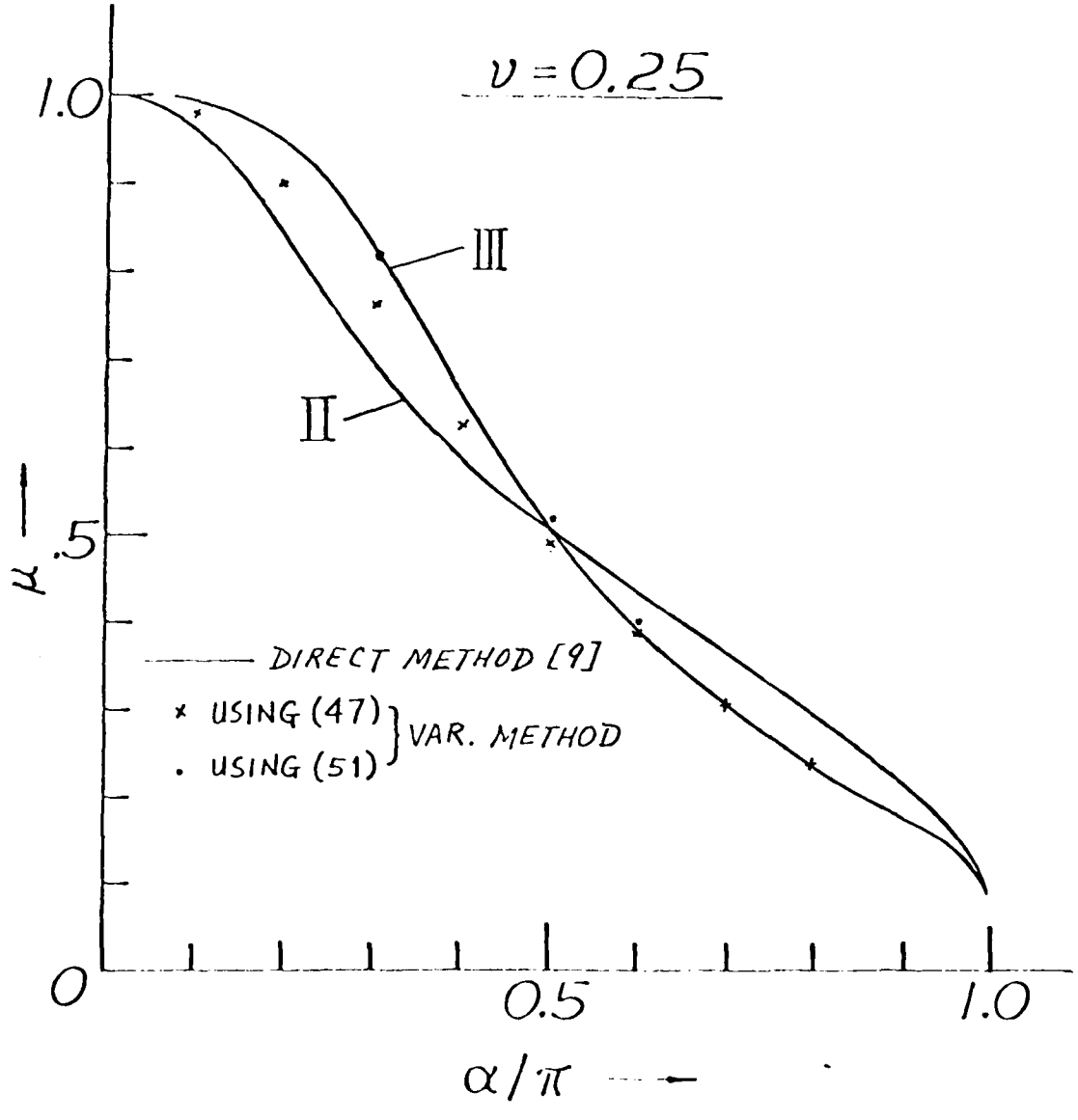


Figure 3. The eigenvalue μ as a function of α/π for ν = 0.25.

(b) Variational results indicated by x and · for mode III.

APPENDIX 1

In this appendix we give components of displacement and stress in terms of Papkovitch stress functions ϕ, ψ, ω and λ. A subscript to a stress function means the partial derivative of the stress function with respect to the variable represented by the subscript, e.g. $\phi_r = \partial\phi/\partial r$, $\phi_{rp} = \partial(\partial\phi/\partial r)/\partial p$. The variable $p = \cos\theta$ is used in the place of θ in both appendix 1 and appendix 2. The notation $\bar{p} = \sin\theta = (1-p^2)^{1/2}$ is also used.

The First Basic Solution:

$$2Gu_r = \phi_r$$

$$2Gu_\theta = -\bar{p}\phi_p/r$$

$$2Gu_\phi = \phi_\phi/(r\bar{p})$$

$$\sigma_r = \phi_{rr}$$

$$\sigma_\theta = r^{-2}(\bar{p}^2\phi_{pp} - p\phi_p + r\phi_r)$$

$$\sigma_\phi = r^{-2}(\phi_{\phi\phi}/\bar{p}^2 + r\phi_r - p\phi_p)$$

$$\tau_{\theta\phi} = -r^{-2}(\phi_{p\phi} + p\phi_\phi/\bar{p}^2)$$

$$\tau_{r\phi} = r^{-2}\bar{p}^{-1}(r\phi_{r\phi} - \phi_\phi)$$

$$\tau_{r\theta} = r^{-2}\bar{p}(-r\phi_{rp} + \phi_p)$$

The Second Basic Solution:

$$2Gu_r = [r\psi_r - (3-4\nu)\psi]\bar{p}\cos\theta$$

$$2Gu_\theta = -[\bar{p}^2\psi_p + (3-4\nu)p\psi]\cos\phi$$

$$2Gu_\phi = \psi_\phi\cos\phi + (3-4\nu)\psi\sin\phi$$

$$\sigma_r = [r\psi_{rr} - 2(1-\nu)\psi_r + 2\nu r^{-1}p\psi_p]\bar{p}\cos\phi + 2\nu(rp)^{-1}\psi_\phi\sin\phi$$

$$\sigma_\theta = r^{-1}[\bar{p}^2\psi_{pp} + (1-2\nu)p\psi_p + (1-2\nu)r\psi_r]\bar{p}\cos\phi + 2\nu(r\bar{p})^{-1}\psi_\phi\sin\phi$$

$$\sigma_\phi = (r\bar{p})^{-1}[\psi_{\phi\phi} + (1-2\nu)r\bar{p}^2\psi_r - (1-2\nu)p\bar{p}^2\psi_p]\cos\phi + 2(1-\nu)(r\bar{p})^{-1}\psi_\phi\sin\phi$$

$$\tau_{\theta\phi} = (-r\bar{p})^{-1}[\bar{p}^2\psi_{p\phi} + 2(1-\nu)p\psi_\phi]\cos\phi - (1-2\nu)r^{-1}\bar{p}\psi_p\sin\phi$$

$$\tau_{r\phi} = r^{-1}[r\psi_{\phi r} - 2(1-\nu)\psi_\phi]\cos\phi + (1-2\nu)\psi_r\sin\phi$$

$$\tau_{r\theta} = [-\bar{p}^2\psi_{rp} - (1-2\nu)p\psi_r + 2(1-\nu)r^{-1}\bar{p}^2\psi_p]\cos\phi$$

The Third Basic Solution:

$$2Gu_r = [r\omega_r - (3-4\nu)\omega]\bar{p}\sin\phi$$

$$2Gu_\theta = -[\bar{p}^2\omega_p + (3-4\nu)p\omega]\sin\phi$$

$$2Gu_\phi = \omega_\phi\sin\phi - (3-4\nu)\omega\cos\phi$$

$$\sigma_r = [r\omega_{rr} - 2(1-\nu)\omega_r + 2\nu r^{-1}p\omega_p]\bar{p}\sin\phi - 2\nu(r\bar{p})^{-1}\omega_\phi\cos\phi$$

$$\sigma_\theta = r^{-1}[\bar{p}^2\omega_{pp} + (1-2\nu)p\omega_p + (1-2\nu)r\omega_r]\bar{p}\sin\phi - 2\nu(r\bar{p})^{-1}\omega_\phi\cos\phi$$

$$\sigma_\phi = r^{-1}[\omega_{\phi\phi}/\bar{p}^2 + (1-2\nu)r\omega_r - (1-2\nu)p\omega_p]\bar{p}\sin\phi - 2(1-\nu)(r\bar{p})^{-1}\omega_\phi\cos\phi$$

$$\tau_{\theta\phi} = (-r\bar{p})^{-1}[\bar{p}^2\omega_{p\phi} + 2(1-\nu)p\omega_\phi]\sin\phi + (1-2\nu)r^{-1}\bar{p}\omega_p\cos\phi$$

$$\tau_{r\phi} = r^{-1}[r\omega_{r\phi} - 2(1-\nu)\omega_\phi]\sin\phi - (1-2\nu)\omega_r\cos\phi$$

$$\tau_{r\theta} = [-\bar{p}^2\omega_{rp} - (1-2\nu)p\omega_r + 2(1-\nu)r^{-1}\bar{p}^2\omega_p]\sin\phi$$

The Fourth Basic Solution:

$$2Gu_r = [r\lambda_r - (3-4\nu)\lambda]p$$

$$2Gu_\theta = [-p\lambda_p + (3-4\nu)\lambda]\bar{p}$$

$$2Gu_\phi = p\bar{p}^{-1}\lambda_\phi$$

$$\sigma_r = rp\lambda_{rr} - 2(1-\nu)p\lambda_r - 2\nu r^{-1}\bar{p}^2\lambda_p$$

$$\sigma_\theta = r^{-1}p\bar{p}^2\lambda_{pp} + (1-2\nu)p\lambda_r + r^{-1}(p^2+2\nu\bar{p}^2-2)\lambda_p$$

$$\sigma_\phi = r^{-1}[p\bar{p}^{-2}\lambda_{\phi\phi} + (1-2\nu)rp\lambda_r - (p^2+2\nu\bar{p}^2)\lambda_p]$$

$$\tau_{\theta\phi} = r^{-1}[-p\lambda_{p\phi} + (2-2\nu-\bar{p}^{-2})\lambda_\phi]$$

$$\tau_{r\phi} = (r\bar{p})^{-1}[rp\lambda_{\phi r} - 2(1-\nu)p\lambda_\phi]$$

$$\tau_{r\theta} = \bar{p}[-p\lambda_{rp} + 2(1-\nu)r^{-1}p\lambda_p + (1-2\nu)\lambda_r]$$

APPENDIX 2

In the following solutions, the selection of cos mφ or sin mφ depends on the geometry of the problem. The sign on top goes with the trigonometric function on the top and vice versa.

Solution [A]:

$$2Gu_r = (\mu+1)r^{\mu}P^m_{\mu+1}\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$2Gu_\theta = (r^\mu/\bar{p})[(\mu+1)pP^m_{\mu+1} - (m+\mu+1)P^m_\mu]\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$2Gu_\phi = \bar{+}(r^\mu/\bar{p})mP^m_{\mu+1}\begin{matrix}\sin m\phi\\ \cos m\phi\end{matrix}$$

$$\sigma_r = \mu(\mu+1)r^{\mu-1}P^m_{\mu+1}\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$\sigma_\theta = (r^{\mu-1}/\bar{p}^2)\{[m^2-\mu-1-\bar{p}^2\mu(\mu+1)]P^m_{\mu+1} + p(m+\mu+1)P^m_\mu\}\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$\sigma_\phi = (r^{\mu-1}/\bar{p}^2)[(-m^2+\mu+1)P^m_{\mu+1} - p(m+\mu+1)P^m_\mu]\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$\tau_{\theta\phi} = \bar{+}(r^{\mu-1}/\bar{p}^2)m[\mu pP^m_{\mu+1} - (m+\mu+1)P^m_\mu]\begin{matrix}\sin m\phi\\ \cos m\phi\end{matrix}$$

$$\tau_{r\phi} = \bar{+}(r^{\mu-1}/\bar{p})m\mu P^m_{\mu+1}\begin{matrix}\sin m\phi\\ \cos m\phi\end{matrix}$$

$$\tau_{r\theta} = (r^{\mu-1}/\bar{p})\mu[(\mu+1)pP^m_{\mu+1} - (m+\mu+1)P^m_\mu]\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

Solution [B]:

$$2Gu_r = r^\mu(\mu-3+4\nu)[(-m+\mu+1)P^m_{\mu+1} - (m+\mu+1)pP^m_\mu]\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$2Gu_\theta = (r^\mu/\bar{p})\{(-m-4+4\nu)(-m+\mu+1)pP^m_{\mu+1} + (m+\mu+1)[m+4-4\nu-\bar{p}^2(\mu+4-4\nu)]P^m_\mu\}\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$2Gu_\phi = \bar{+}(r^\mu/\bar{p})(m+4-4\nu)[(-m+\mu+1)P^m_{\mu+1} - (m+\mu+1)pP^m_\mu]\begin{matrix}\sin m\phi\\ \cos m\phi\end{matrix}$$

$$\sigma_r = r^{\mu-1}[(-m+\mu+1)(\mu^2-3\mu+2\mu\nu-2m\nu-2\nu)P^m_{\mu+1} + (m+\mu+1)(-\mu^2+3\mu+2\nu)pP^m_\mu]\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$$

$$\sigma_\theta = (r^{\mu-1}/\bar{p}^2)\{(-m+\mu+1)[(-\mu^2-2\mu\nu-(3-2\nu)(m+1)\bar{p}^2 + (m^2+(5-4\nu)m+4-4\nu)]P^m_{\mu+1}$$
$$+ (m+\mu+1)p[-m^2-5m+4m\nu-4+4\nu+\bar{p}^2(\mu^2+3\mu+3-2\nu)]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\sigma_\phi = (r^{\mu-1}/\bar{p}^2)\{(-m+\mu+1)[-m^2-5m+4m\nu-4+4\nu+\bar{p}^2(1-2\nu)(m+\mu+1)]P^m_{\mu+1}$$
$$+ (m+\mu+1)p[m^2+5m-4m\nu+4-4\nu-(1-2\nu)(2\mu+1)\bar{p}^2]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\tau_{\theta\phi} = \pm (r^{\mu-1}/\bar{p}^2)\{(-m+\mu+1)p[m^2+(5-4\nu)m+4(1-\nu)]P^m_{\mu+1}$$
$$+ (m+\mu+1)[\bar{p}^2(m\mu+(3-2\nu)m+2(1-\nu)\mu+4(1-\nu))-m^2-(5-4\nu)m-4(1-\nu)]P^m_\mu\}^{\sin m\phi}_{\cos m\phi}$$

$$\tau_{r\phi} = \mp (r^{\mu-1}/\bar{p})[m(\mu-2+2\nu)+2\mu(1-\nu)-2(1-\nu)][(-m+\mu+1)P^m_{\mu+1} - (m+\mu+1)pP^m_\mu]^{\sin m\phi}_{\cos m\phi}$$

$$\tau_{r\theta} = (r^{\mu-1}/\bar{p})\{(-m+\mu+1)p[-m\mu+2(1-\nu)(m-\mu+1)]P^m_{\mu+1}$$
$$+ (m+\mu+1)[m\mu+2(1-\nu)(-m+\mu-1)+\bar{p}^2(-\mu^2+2-2\nu)]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

Solution [C]:

$$2Gu_r = r^\mu(\mu-3+4\nu)(-P^m_{\mu+1} + pP^m_\mu)^{\cos m\phi}_{\sin m\phi}$$

$$2Gu_\theta = (r^\mu/\bar{p})\{(-m+4-4\nu)pP^m_{\mu+1} + [(m-4+4\nu)+\bar{p}^2(\mu+4-4\nu)]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$2Gu_\phi = \mp (r^\mu/\bar{p})(m-4+4\nu)(-P^m_{\mu+1} + pP^m_\mu)^{\sin m\phi}_{\cos m\phi}$$

$$\sigma_r = r^{\mu-1}\{[-\mu^2+\mu(3-2\nu)-2m\nu+2\nu]P^m_{\mu+1} + (\mu^2-3\mu-2\nu)pP^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\sigma_\theta = (r^{\mu-1}/\bar{p}^2)\{[\bar{p}^2(\mu^2+2\nu\mu-(3-2\nu)(m-1)) + (-m^2+(5-4\nu)m-4(1-\nu))]P^m_{\mu+1}$$
$$+ [\bar{p}^2(-\mu^2-3\mu-3+2\nu) + m^2-(5-4\nu)m+4(1-\nu)]pP^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\sigma_\phi = (r^{\mu-1}/\bar{p}^2)\{[\bar{p}^2(1-2\nu)(m-\mu-1) + m^2-(5-4\nu)m+4(1-\nu)]P^m_{\mu+1}$$
$$+ [\bar{p}^2(1-2\nu)(2\mu+1) - m^2+(5-4\nu)m-4(1-\nu)]pP^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\tau_{\theta\phi} = \pm (r^{\mu-1}/\bar{p}^2)\{[m^2-(5-4\nu)m+4(1-\nu)]pP^m_{\mu+1}$$

$$+ [\bar{p}^2(-m\mu-(3-2\nu)m+2(1-\nu)\mu+4(1-\nu))-m^2+(5-4\nu)m-4(1-\nu)]P^m_\mu\}^{\sin m\phi}_{\cos m\phi}$$

$$\tau_{r\phi} = \pm (r^{\mu-1}/\bar{p})[-m\mu+2(1-\nu)m+2(1-\nu)(\mu-1)](-P^m_{\mu+1} + pP^m_\mu)^{\sin m\phi}_{\cos m\phi}$$

$$\tau_{r\theta} = (r^{\mu-1}/\bar{p})\{[-m\mu+2(1-\nu)m+2(1-\nu)(\mu-1)]pP^m_{\mu+1}$$

$$+ [m\mu-2(1-\nu)m-2(1-\nu)(\mu-1)-\bar{p}^2(-\mu^2+2-2\nu)]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

Solution [D]:

$$2Gu_r = (\mu-3+4\nu)r^\mu pP^m_\mu{}^{\cos m\phi}_{\sin m\phi}$$

$$2Gu_\theta = (r^\mu/\bar{p})\{(-m+\mu+1)pP^m_{\mu+1} + [\bar{p}^2(\mu+4-4\nu)-(\mu+1)]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$2Gu_\phi = \mp (mp/\bar{p})r^\mu P^m_\mu{}^{\sin m\phi}_{\cos m\phi}$$

$$\sigma_r = r^{\mu-1}\{2\nu(-m+\mu+1)P^m_{\mu+1} + (\mu^2-3\mu-2\nu)pP^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\sigma_\theta = (r^{\mu-1}/\bar{p}^2)\{(m-\mu-1)[1-(3-2\nu)\bar{p}^2]P^m_{\mu+1} + [m^2+\mu+1+\bar{p}^2(-\mu^2-3\mu-3+2\nu)]pP^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\sigma_\phi = (r^{\mu-1}/\bar{p}^2)\{(-m+\mu+1)[1+\bar{p}^2(-1+2\nu)]P^m_{\mu+1} + [-m^2-\mu-1+(1-2\nu)(2\mu+1)\bar{p}^2]pP^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

$$\tau_{\theta\phi} = \pm (r^{\mu-1}/\bar{p}^2)m\{(m-\mu-1)pP^m_{\mu+1} + [\mu+2-\bar{p}^2(\mu+3-2\nu)]P^m_\mu\}^{\sin m\phi}_{\cos m\phi}$$

$$\tau_{r\phi} = \pm (2-2\nu-\mu)mp(r^{\mu-1}/\bar{p})P^m_\mu{}^{\sin m\phi}_{\cos m\phi}$$

$$\tau_{r\theta} = (r^{\mu-1}/\bar{p})\{(-m+\mu+1)(\mu-2+2\nu)pP^m_{\mu+1} + [-\mu^2+\mu(1-2\nu)+2(1-\nu)$$

$$+\bar{p}^2(\mu^2-2+2\nu)]P^m_\mu\}^{\cos m\phi}_{\sin m\phi}$$

MOTIONS PERTURBING STATES OF REST OF VISCOELASTIC SOLIDS

P.M. Dixit and D.D. Joseph
Department of Aerospace Eng. and Mechanics
University of Minnesota, MPLS. 55401

1. Summary

Our goal is to derive the canonical forms of the stress and equations of motion governing the motions which perturb the rest state (of elastostatic deformation) and the natural (unstressed and undeformed) state of viscoelastic solids. In this theory nonlinear elasticity appears as a special case of nonlinear viscoelasticity which arises when the prescribed data is steady. The domain of deformations on which the constitutive equation for viscoelastic solids reduces to the constitutive equation for elasticity is the set of all kinematically admissible states of rest. We find the forms of the stress and discuss some properties of the equations of motion which perturb states of rest (elastostatics). There are too many unknown functions in the theory of perturbations of states of rest of viscoelastic solids to make the theory attractive to material scientists interested in rheometrical measurements. For this purpose, the theory of perturbations of the natural state of viscoelastic solids is more attractive. We develop a detailed theory of perturbations of the natural state and derive equations governing perturbations of zero displacements which may be solved sequentially. At each stage there are three equations for each component of displacement when the material is compressible and, when incompressible, there is an additional unknown, the reaction pressure, and an additional differential equation expressing the incompressibility. We identify the material parameters which must be measured to distinguish one solid from another. In the second order theory for compressible materials there are six elastic constants and twelve viscoelastic material functions. The number of constants is reduced to two and the number of material functions to three

in the incompressible case. For incompressible solids the leading operator which must be inverted at each stage in the perturbation is characterized by one constant and one function. If the constant and function satisfy some mild and physically natural conditions the solutions of the perturbation equations will be stable and unique. We show how to use the perturbation equations for material studies by deriving several problems of possible application in the design of rheometers for viscoelastic solids in motions which perturb the natural state.

In the last part of the paper we derive the linearized equations governing viscoelastic perturbations of elastostatic solutions with an eye to potential applications in the dynamic theory of stability and bifurcation. We give a heuristic and completely physical argument that solids undergoing static deformations cannot bifurcate into time-periodic motions. We argue that critical eigenvalues for the stability of elastostatic solutions cannot be complex-valued and suggest this principle: any choice of the material parameters leading to complex-valued critical values is the wrong choice. Finally, as an example of the nature of the application of the linearized theory, we derive the spectral problem governing the stability of Rivlin's (1949) universal solution of the torsion of an elastic cylinder.

Notation

We have used the same symbol for functions and their values. Sometimes for brevity we have suppressed some or all of the arguments of some of the functions. But they are easily understood from the context.

Relation to previous work

Our work falls in the framework of what elasticians call "small on large". This means small unknown deformations are superposed on large known ones. The unknown deformations are generally treated as elastic but in at least two important works, Coleman and Noll (1961) and Pipkin (1964), the unknown deformations are presumed to be viscoelastic. Some of our best results do not fit in framework of "small on large" or even of "small on small" but, instead, fall into the frame of "small on zero". In the theory of solids, treated here, the "small" is arbitrary and viscoelastic and the "zero" corresponds to the natural state of the body. Many of the results of Green & Rivlin (1957), Coleman and Noll (1961), Pipkin ard Rivlin (1961) and Pipkin (1964) can also be interpreted as falling under the theory of "small on zero".

We can also argue that the results given by the authors mentioned in the last paragraph are contributions to the asymptotic theory of the stress in simple materials. In such studies the norm of some measure of deformations is presumed small and one seeks canonical forms of the stress which perturb the zero norm. Naturally, these canonical forms are ordered in powers of the relevant measure of deformation expressed in terms of multilinear functionals simplified to the degree required by considerations of material symmetry.

The purpose of the asymptotic theory is to specify relatively simple forms of the stress when the motion is one which is small in the appropriate norm. The asymptotic theory is approximate, since the exact conditions allowed by nature, under which the appropriate norm is small, is left unspecified.

In contrast, we have tried to establish the consequences of the fact that the small norm must somehow arise from small prescribed data, like small external forcing, filtered through the equations of motion. So we come up with an ordered asymptotic sequence of boundary value problems in which redundant terms are purged from the canonical forms of the stress, and we derive algorithms in which the ordered computation of motion and deformation is reduced to a recipe. So our theory is a formally exact asymptotic theory in which the prescribed data rather than the unknown motion is presumed small.

The heart of our theory is the perturbation equations of motion. So we are obliged to consider the usual interesting questions of existence, stability and uniqueness. Naturally, these questions cannot come up in studies which are abandoned at the point where the stress is arranged into a suitably invariant series of powers of the deformation.

In solids it is possible to choose several different measures of deformation to express the stress. These different choices lead to different but possibly equivalent expressions for the stress. We have thought the choice of measure of deformation to be important and we followed Coleman and Noll (1961) because in their formulation the constitutive equation

for all static deformations of viscoelastic materials obviously and easily reduces to elasticity. So there is no difference between nonlinear elastostatics and nonlinear viscoelastostatics. There is no doubt that the theory of Green & Rivlin (1957), Pipkin and Rivlin (1961) and Pipkin (1964) also reduces to elastostatics for static deformations but the reduction is less obvious. Suprisingly, this obvious connection between elasticity and viscoelasticity, which is certainly well-known to people working in mechanics, seems not to have been stressed (or even mentioned) by any of the authors cited above.

The parts of the theory of Coleman & Noll (1961) which we use are summarized in the introduction. In this formulation the stress depends on the left Cauchy-Green tensor and the history of the right relative Cauchy-Green tensor. Green and Rivlin (1957), Pipkin and Rivlin (1961) and Pipkin (1964) express the stress in terms of the history of the right Cauchy - Green tensor (not the relative tensor) which is defined is

$$\underline{\underline{C}}(\underline{X},\tau) = \underline{\underline{F}}(\tau)^{T}\underline{\underline{F}}(\tau)$$

where $\underline{\underline{F}}$ is defined in the first equation of the introduction. For some reason that we don't understand Pipkin works with the time derivative $\dot{\underline{\underline{C}}}$. The time derivative of kinematic tensors is good for fluids but not for solids (Joseph & Beavers, 1977).

2. Introduction

The viscoelastic solids which we study are simple isotropic materials of the type discussed in §33 of the treatise by Truesdell and Noll (1965). The material properties of such solids are independent of time and the state of stress is determined by the first spatial gradient of the deformation

$$\underline{\underline{F}} = \nabla \underline{x}, \quad F_{ij} = \partial x_i / \partial X_j$$

where $\underline{X}_i$ are the cartesian coordinates of a particle $\underline{X}$ of the body in the undeformed isotropic state and $x_i(\underline{X},\tau)$ are the cartesian coordinates of the position $\underline{x}$ of the same particle at the time τ. The Cauchy stress at a particle is given by an expression of the form

$$\underline{\underline{T}} = \overset{\infty}{\underset{s=0}{\underline{\underline{\mathcal{F}}}}} [\underline{\underline{B}};\ \underline{\underline{G}}(s)] \tag{2.1}$$

where

$$\underline{\underline{B}} = \underline{\underline{F}}\ \underline{\underline{F}}^T$$

is the left Cauchy-Green strain tensor at the present time t,

$$\underline{\underline{G}}(s) = \underline{\underline{C}}_t(\tau) - \underline{\underline{1}}, \quad s = t - \tau, \quad -\infty < \tau \leq t.$$

$\underline{\underline{C}}_t(\tau)$ is the right, relative Cauchy-Green strain tensor ;

$$\underline{\underline{C}}_t(\tau) = \underline{\underline{F}}^T{}_t(\tau)\ \underline{\underline{F}}_t(\tau), \quad \underline{\underline{C}}_t(t) = \underline{\underline{1}}$$

where

$$\underline{\underline{F}}_t(\tau) = \nabla_x\ \underline{\chi}_t(\underline{x},\tau), \quad \underline{\underline{F}}_t(t) = \underline{\underline{1}}$$

is the relative deformation gradient tensor defined in terms of the relative position vector $\underline{\chi}_t(\underline{x},\tau)$ of a particle which at the time $\tau = t$ is at place $\underline{x} = \underline{\chi}_t(\underline{x},t)$. $\underline{\underline{\mathcal{F}}}$ is a functional of the history $\underline{\underline{G}}(s)$ of a particle depending on the tensor parameter $\underline{\underline{B}}$ and is such that

$$\underline{\underline{T}} = \underline{\underline{\mathfrak{F}}}[\underline{\underline{1}}, \underline{\underline{0}}] = \underline{\underline{0}}. \tag{2.2}$$

Eq. (2.2) says that there is no stress in the undeformed state of the body. The stress-free, undeformed state of the body is called the natural state of the body.

Nonlinear elasticity arises from (2.1) when $\underline{\underline{G}}(s)$ is put to zero:

$$\underline{\underline{T}}(t) = \underline{\underline{\mathfrak{F}}}[\underline{\underline{B}}(t), \underline{\underline{0}}] \ . \tag{2.3}$$

It is possible and, in some asymptotic limits, it is useful to regard (2.3) as defining the dynamical response of a nonlinearly elastic body. But we also note that an elastostatics

$$\underline{\underline{T}} = \underline{\underline{\mathfrak{F}}}[\underline{\underline{B}}, \underline{\underline{0}}] \tag{2.4}$$

of viscoelastic solids arises automatically from (2.1) for every deformation such that $\underline{x} = \underline{x}(\underline{X}, t)$ is independent of t. (If $\underline{x}$ is independent of t, then $\underline{\underline{F}}_t(\tau) = \underline{\underline{1}}$ and $\underline{\underline{G}}(s) = \underline{\underline{0}}$.) We think of the class of t-independent deformations defining the rest state of a viscoelastic solid as coinciding with elasticity; that is, all solids are at least viscoelastic but the constitutive equation for viscoelastic solids reduces to elasticity when the deformations are restricted to t-independent ones. So from the point of view of material science we do not think it useful to admit dynamic elasticity as a viable subject. After a time vibrating solid bodies always come to rest, unless forced, and when they are in motion these bodies satisfy a constitutive equation which is at least as complicated as (2.1). So we think that nonlinear dynamic viscoelasticity and nonlinear elastostatics are not different subjects but just different realizations of the

same governing equations corresponding to unsteady or steady solutions which arise in response to the given data: the initial and boundary conditions and the prescribed forcing. In any case, that is the nature of the theory which we shall now develop.

3. Frechét expansion of the stress

In the state of rest (elastic deformation) $\underline{\underline{G}}(s) = \underline{\underline{0}}$. We assume that the stress perturbing states of rest is expressible as a Frechét expansion of $\underline{\underline{\mathcal{F}}}$ in powers of $\underline{\underline{G}}(s)$. Thus,

$$\underline{\underline{\mathcal{F}}}[\underline{\underline{B}}, \underline{\underline{G}}(s)] = \underline{\underline{\mathcal{F}}}[\underline{\underline{B}},\underline{\underline{0}}] + \underline{\underline{\mathcal{F}}}_1[\underline{\underline{B}},\underline{\underline{0}}|\underline{\underline{G}}(s)] + \frac{1}{2}\underline{\underline{\mathcal{F}}}_2[\underline{\underline{B}},\underline{\underline{0}}|\underline{\underline{G}}(s_1)|\underline{\underline{G}}(s_2)] + O(\|\underline{\underline{G}}(s)\|^3) \tag{3.1}$$

where $\underline{\underline{\mathcal{F}}}_1[\underline{\underline{B}},\underline{\underline{0}}|\cdot]$ is a linear operator and $\underline{\underline{\mathcal{F}}}_2[\underline{\underline{B}},\underline{\underline{0}}|\cdot|\cdot]$ a bilinear operator evaluated on the zero history. Green and Rivlin (1957) assumed an expansion in the form (3.1) with the Frechét derivatives expressed as multiple integrals. They appealed to the Stone-Weirestrass theorem for functionals for mathematical justification. Coleman and Noll (1961) also arrived at an integral expansion. They introduced a Hibert space of histories endowed with a weighted scalar product (fading memory) and appealed to the Riesz representation theorem to justify an integral representation of the first term.

We follow the authors just named and assume that the terms in (3.1) can be expressed as integrals.

$$\underline{\underline{T}} = \underline{\underline{f}}(\underline{\underline{B}}) + \int_0^\infty \underline{\underline{K}}(s, \underline{\underline{B}}(t))\underline{\underline{G}}(s)ds + \int_0^\infty \int_0^\infty \underline{\underline{\Gamma}}(s_1,s_2,\underline{\underline{B}}(t))\underline{\underline{G}}(s_1)\underline{\underline{G}}(s_2)ds_1ds_2 + O(\|\underline{\underline{G}}\|^3) \tag{3.2}$$

where $\underline{\underline{K}}(s,\underline{\underline{B}})$ is an isotropic tensor function of $\underline{\underline{B}}$ of order four whose components $K_{ijk\ell}$ are symmetric in successive pairs of indices, $\underline{\underline{\Gamma}}(s_1,s_2,\underline{\underline{B}}) = \underline{\underline{\Gamma}}(s_2,s_1,\underline{\underline{B}})$ is an isotropic tensor function of $\underline{\underline{B}}$ of order six whose components $\Gamma_{ijk\ell mn}$ are symmetric in successive pairs of indices.

If the integrals in (3.2) are set to zero we are left with

$\underline{T} = \underline{\underline{f}}(\underline{B})$ which is supposed to be the response of bodies which are said to be purely elastic. In mathematical studies various conditions are proposed about $\underline{\underline{f}}(\underline{B})$ to insure appropriate properties of existence and uniqueness of the solutions of the equations governing the dynamic response of nonlinearly elastic bodies. In our study of nonlinear viscoelasticity we are also required to introduce small nonlinear effects of $\underline{\underline{f}}(\underline{B})$. But in our study special assumptions are not required. Instead, the determination of some nonlinear properties of $\underline{\underline{f}}(\underline{B})$ is left as an open question for rheometrical measurements and experiments.

The methods for finding the most general form of isotropic tensor-valued functions of many tensors have been given by Wineman and Pipkin (1964), based on earlier work of Rivlin, Smith and Spencer (see R.S. Rivlin, 1969; Truesdell & Noll, 1965, §13). Dixit (1979) applied these methods to (3.2). The reduction of (3.2) to isotropic form is like the Hamilton-Cayley reduction of a tensor polynomial of degree m > 2 to m = 2. Suppose we have a tensor* $\underline{A}$ which is a function of tensors $\underline{B}$, $\underline{C}$, ...

$$\underline{\underline{A}} = \underline{\underline{g}}(\underline{\underline{B}}, \underline{\underline{C}}, \ldots)$$

The dependence is such that $\underline{\underline{g}}$ satisfies

$$\underline{\underline{Q}}\,\underline{\underline{g}}(\underline{\underline{B}},\underline{\underline{C}},\ldots)\underline{\underline{Q}}^T = \underline{\underline{g}}(\underline{\underline{Q}}\underline{\underline{B}}\underline{\underline{Q}}^T, \underline{\underline{Q}}\underline{\underline{C}}\underline{\underline{Q}}^T,\ldots) , \quad \underline{\underline{Q}} \in \mathcal{O} .$$

$\mathcal{O}$ is the set of all orthogonal tensors.†

The method of finding the most general form of $\underline{\underline{g}}$ is as follows: First introduce an auxiliary second-order tensor $\underline{\underline{\phi}}$. Let $\alpha = \underline{\underline{\phi}} \cdot \underline{\underline{g}}(\underline{\underline{B}},\underline{\underline{C}},\ldots)$. Now α is a scalar invariant of

*Here we consider only the second-order tensors. But the method is applicable even for tensors of other orders.

†More generally, $\mathcal{O}$ is the symmetry-group of the material.

tensors $\underline{\underline{\phi}}$, $\underline{\underline{B}}$, $\underline{\underline{C}}$... . A set of scalar invariants $H_j(\underline{\underline{\phi}}, \underline{\underline{B}}, \underline{\underline{C}},...)$, $j = 1, ...k$ is called a functional basis if every scalar invariant of $\underline{\underline{\phi}}$, $\underline{\underline{B}}$, $\underline{\underline{C}}$.....can be expressed as a function of H_j, $j = 1, ... k$.

If $\underline{\underline{g}}$ is a polynomial in $\underline{\underline{B}}$, $\underline{\underline{C}}$, ..., then α is a polynomial scalar invariant of $\underline{\underline{\phi}}$, $\underline{\underline{B}}$, $\underline{\underline{C}}$, ... A set of polynomial scalar invariants $I_j(\underline{\underline{\phi}}, \underline{\underline{B}}, \underline{\underline{C}}, ...)$, $j = 1 ..., n$ is called an integrity basis if every polynomial scalar invariant of $\underline{\underline{\phi}}$, $\underline{\underline{B}}$, $\underline{\underline{C}}$...can be expressed as a polynomial in I_j, $j = 1,...,n$.

Wineman and Pipkin (1964) have shown that an integrity basis is also a functional basis.

An integrity basis for an arbitrary number of symmetric second-order isotropic tensors was given by Spencer, Smith & Rivlin (see Rivlin, 1969). Integrity bases for an arbitrary number of tensors and vectors and for the case in which the symmetry-group is not the group of all orthogonal tensors was given by Spencer, Smith, Rivlin, Adkins and Weyl (see Wineman and Pipkin, 1964).

Once an integrity basis for the tensors $\underline{\underline{\phi}}$, $\underline{\underline{B}}$, $\underline{\underline{C}}$,... has been found, the elements which are functions of $\underline{\underline{B}}$, $\underline{\underline{C}}$, ... alone are singled out. Call these I_γ, $\gamma = 1,...m$. (These form an integrity basis for the tensors $\underline{\underline{B}}$, $\underline{\underline{C}}$,... .) Then the elements which are linear in $\underline{\underline{\phi}}$ are selected. Each such invariant is of the form $\phi_{ij} f_{ij}^{(\beta)}(\underline{\underline{B}}, \underline{\underline{C}},...)$, $\beta = 1,...\ell$. Then

$$\underline{\underline{g}} = \sum_{\beta=1}^{\ell} F_\beta(I_1, I_2,...I_m)\underline{\underline{f}}^{(\beta)}(\underline{\underline{B}}, \underline{\underline{C}},...) .$$

Applying this method to (3.2) we find that

$$(3.3) \qquad \underline{\underline{f}}(\underline{\underline{B}}) = f_0\underline{\underline{1}} + f_1\underline{\underline{B}} + f_2\underline{\underline{B}}^2 ,$$

and

$$(3.4) \qquad \underline{\underline{K}}(s,\underline{\underline{B}})\underline{\underline{G}}(s) = \mathrm{tr}[(\phi_{00}\underline{\underline{1}} + \phi_{01}\underline{\underline{B}} + \phi_{02}\underline{\underline{B}}^2)\underline{\underline{G}}(s)]\underline{\underline{1}}$$

$$+ \mathrm{tr}[(\phi_{10}\underline{\underline{1}} + \phi_{11}\underline{\underline{B}} + \phi_{12}\underline{\underline{B}}^2)\underline{\underline{G}}(s)]\underline{\underline{B}}$$

$$+ \mathrm{tr}[(\phi_{20}\underline{\underline{1}} + \phi_{21}\underline{\underline{B}} + \phi_{22}\underline{\underline{B}}^2)\underline{\underline{G}}(s)]\underline{\underline{B}}^2$$

$$+ (\phi_{30}\underline{\underline{1}} + \phi_{31}\underline{\underline{B}} + \phi_{32}\underline{\underline{B}}^2)\underline{\underline{G}}(s)$$

$$+ \underline{\underline{G}}(s)(\phi_{30}\underline{\underline{1}} + \phi_{31}\underline{\underline{B}} + \phi_{32}\underline{\underline{B}}^2)$$

where the f_i are functions of the three principal invariants of $\underline{\underline{B}}$, $I_B = \mathrm{tr}\underline{\underline{B}}$, $II_B = \frac{1}{2}[(\mathrm{tr}\underline{\underline{B}})^2 - \mathrm{tr}\underline{\underline{B}}^2]$, $III_B = \det \underline{\underline{B}}$, and the ϕ_{ij} are functions of the same three invariants and the time lag $s = t - \tau$. The isotropic form of $\underline{\underline{\Gamma}}(s_1,s_2;\underline{\underline{B}})\underline{\underline{G}}(s_1)\underline{\underline{G}}(s_2)$ is lengthy and will not be given here (see Dixit, 197[c]).

The forms of the stress which perturb the natural state have a simpler structure than the forms (3.3,4) which perturb states of rest (elastic deformation). The natural state is a state of rest in which the body is undeformed and unstressed so that $\underline{\underline{G}}(s) = \underline{\underline{0}}$, $\underline{\underline{B}}(t) = \underline{\underline{1}}$ and $\underline{\underline{\mathcal{F}}}[\underline{\underline{1}},\underline{\underline{0}}] = \underline{\underline{0}}$. To compute stresses relative to the natural state it is convenient to expand the tensor functions of $\underline{\underline{B}}(t)$ in (3.2) into a series of powers of the perturbation tensor

$$(3.5) \qquad \underline{\underline{b}}(t) = \underline{\underline{B}}(t) - \underline{\underline{1}} .$$

This procedure reduces the problem of finding the most general isotropic forms of (3.2) to a problem of finding isotropic tensor coefficients for multilinear forms. At the end of the

analysis one finds that

$$\underline{\underline{f}} = \beta\underline{\underline{b}} + \beta^{[1]}\underline{\underline{1}}\,\mathrm{tr}\underline{\underline{b}} + \beta^{[2]}\,\underline{\underline{bb}} + \beta^{[3]}\underline{\underline{1}}\,(\mathrm{tr}\underline{\underline{b}})^2 + \beta^{[4]}\underline{\underline{1}}\,\mathrm{tr}(\underline{\underline{bb}}) + \beta^{[5]}\,(\mathrm{tr}\underline{\underline{b}})\underline{\underline{b}} + O(|\underline{\underline{b}}|^3). \tag{3.6}$$

$$\begin{aligned}\underline{\underline{K}}\left(s,\underline{\underline{B}}(t)\right)\underline{\underline{G}}(s) = {} & \zeta(s)\underline{\underline{G}}(s) + \zeta^{[1]}\underline{\underline{1}}\,\mathrm{tr}\underline{\underline{G}}(s) \\ & + \zeta^{[2]}(s)\,\{\underline{\underline{b}}(t)\underline{\underline{G}}(s) + \underline{\underline{G}}(s)\underline{\underline{b}}(t)\} \\ & + \zeta^{[3]}(s)\underline{\underline{G}}(s)\,\mathrm{tr}\underline{\underline{b}}(t) \\ & + \zeta^{[4]}(s)\underline{\underline{b}}(t)\,\mathrm{tr}\underline{\underline{G}}(s) \\ & + \zeta^{[5]}(s)\underline{\underline{1}}[\mathrm{tr}\underline{\underline{b}}(t)][\mathrm{tr}\underline{\underline{G}}(s)] \\ & + \zeta^{[6]}(s)\underline{\underline{1}}\,\mathrm{tr}[\underline{\underline{b}}(t)\,\underline{\underline{G}}(s)] \\ & + O(|\underline{\underline{b}}|^2|\underline{\underline{G}}|),\end{aligned} \tag{3.7}$$

and

$$\begin{aligned}\underline{\underline{\Gamma}}\left(s_1,s_2,\underline{\underline{B}}(t)\right)\underline{\underline{G}}(s_1)\underline{\underline{G}}(s_2) = {} & \alpha(s_1,s_2)\underline{\underline{G}}(s_1)\underline{\underline{G}}(s_2) \\ & + \alpha^{[1]}(s_1,s_2)\underline{\underline{1}}\,[\mathrm{tr}\underline{\underline{G}}(s_1)][\mathrm{tr}\underline{\underline{G}}(s_2)] \\ & + \alpha^{[2]}(s_1,s_2)\underline{\underline{1}}\,\mathrm{tr}[\underline{\underline{G}}(s_1)\underline{\underline{G}}(s_2)] \\ & + \alpha^{[3]}(s_1,s_2)\underline{\underline{G}}(s_1)\mathrm{tr}\underline{\underline{G}}(s_2) \\ & + \alpha^{[4]}(s_1,s_2)\underline{\underline{G}}(s_2)\mathrm{tr}\underline{\underline{G}}(s_1) \\ & + O(|\underline{\underline{b}}|\,|\underline{\underline{G}}|^2).\end{aligned} \tag{3.8}$$

When the solid is incompressible the density is a constant and

(3.9) $\det \underline{\underline{F}} = 1$.

In this case the stress is constitutively determined only up to a scalar field p

(3.10) $\underline{\underline{T}} = - p \underline{\underline{1}} + \underline{\underline{\mathcal{F}}}[\underline{\underline{B}}(t), \overset{\infty}{\underset{s=0}{\underline{\underline{G}}}}(s)]$.

The scalar field p is an additional unknown and (3.9) is the additional equation necessary to determine this field.

The forms of $\underline{\underline{\mathcal{F}}}$ perturbing the rest state and the natural state are the ones already derived for the compressible case with two differences. The first difference is that all the terms in the expansions (3.3) through (3.8) which are proportional to $\underline{\underline{1}}$ may be grouped with p. We may regard the new coefficient of $\underline{\underline{1}}$ in $\underline{\underline{T}}$ as a new "pressure", say π, which is constitutively indeterminate and is to be determined ultimately from the solutions of the equations of motion. So in the incompressible case we take the forms of $\underline{\underline{\mathcal{F}}}$ given by (3.3) through (3.8) modulo terms proportional to $\underline{\underline{1}}$. A second difference between the stress in the compressible and incompressible case arises as a consequence of (3.9). This second difference will be discussed in §7.

4. Equations of motion for the perturbations of the natural state

In solid bodies the natural state is important because the elastic stresses are measured relative to the undeformed, unstressed state of the body. So if $\underline{t}_n$ is the traction vector on the boundary $\partial\mathcal{V}$ of the region of space occupied by the deformed body, then

$$(4.1) \qquad \int_{\partial\mathcal{V}} \underline{t}_n \, da = \int_{\partial\mathcal{V}_0} \underline{\underline{S}}^T \cdot \underline{N} \, dA$$

where $\mathcal{V}_0$ is the region occupied by the undeformed body and $\underline{N}$ is the outward normal on $\partial\mathcal{V}_0$; $\underline{n}$ is the outward normal on $\partial\mathcal{V}$ and

$$(4.2) \qquad \underline{n} \, da = \det \underline{\underline{F}} \, (\underline{\underline{F}}^T)^{-1} \cdot \underline{N} \, dA \, .$$

The Piola-Kirchoff stress $\underline{\underline{S}}^T$ is given in terms of the Cauchy stress by

$$(4.3) \qquad \underline{\underline{S}}^T = \underline{\underline{T}}^T \, (\underline{\underline{F}}^T)^{-1} \det \underline{\underline{F}} = \underline{\underline{T}} (\underline{\underline{F}}^T)^{-1} \det \underline{\underline{F}} \, .$$

The balance of momentum in any small part of $\mathcal{V}$ (also called $\mathcal{V}$) may be written as

$$(4.4) \qquad \int_{\mathcal{V}} \rho \ddot{\underline{u}} \, d\mathcal{V} = \int_{\mathcal{V}} \underline{b} \, d\mathcal{V} + \int_{\partial\mathcal{V}} \underline{t}_n \, da$$

where $\underline{b}$ is the body force per unit mass,

$$(4.5) \qquad \underline{u} = \underline{x} - \underline{X}$$

is the displacement vector of the partical $\underline{X}$,

$$\ddot{\underline{u}} = \partial^2 \underline{u}(\underline{X},t)/\partial t^2 ,$$

the acceteration, is a derivative following the particle (at fixed $\underline{X}$). In a loose notation, we use the symbol $\underline{u}(\underline{x},t)$ and $\underline{u}(\underline{X},t)$ for different functions whose values $\underline{u}$ are identical when $\underline{X}$ is the particle presently in the place $\underline{x}$. The density $\rho(\underline{x},t)$ in (4.4) is related to the density ρ_o of the same particle in the natural state by

$$\rho_o(\underline{X}) = \rho(\underline{x},t) \det \underline{\underline{F}}(t) . \tag{4.6}$$

Eq. (4.6) implies

$$\rho \, d\mathcal{V} = \rho_o d\mathcal{V}_0 . \tag{4.7}$$

Inserting (4.1) and (4.7) into (4.4) we find the Piola-Kirchoff equations of motion

$$\rho_o(\underline{X})\ddot{\underline{u}}(\underline{X},t) = \rho_o \underline{b}(\underline{X},t) + \operatorname{div} \underline{\underline{S}}^T(\underline{X},t) . \tag{4.8}$$

Solutions of (4.8) are driven by the prescribed data: the force field $\underline{b}$, the boundary conditions and the initial history. Our purpose is to develop an algorithm to compute solutions of (4.8) which perturb the zero data giving rise to the natural state. And in the usual way we serve our purpose by requiring that the prescribed data be proportional to a small

parameter ε so that solutions of (4.8) with $\varepsilon \neq 0$ reduce to the natural state in which $\underline{u}$ and S both vanish when $\varepsilon = 0$. For example, we may say that the deformations are driven by $\underline{b}(\cdot, \underline{X}, t) \varepsilon \underline{\beta}(\underline{X}, t)$. Naturally the computation of fields at $\varepsilon = 0$ means that the perturbation problems are all posed on the domain $\mathcal{V}_0$ of the natural state. It is perhaps of interest to remark that our method of solution introduces the natural state automatically through the data and there is no particular advantage gained by starting with the Piola-Kirchoff equations of motion. We arrive at exactly the same equations of motion if we start with Cauchy's equations. In fact, it is more natural to prescribe conditions on $\partial\mathcal{V}$, the boundary of the deformed body, than on $\partial\mathcal{V}_0$, the boundary of the body in the natural state.

Turning now to the aforementioned boundary conditions we declare our interest in boundary value problems of the mixed type. In specifiying "mixed type" boundary conditions we decompose the boundary of the deformed body into two parts.

$$\partial\mathcal{V}(t,\varepsilon) = \partial\mathcal{V}_1(t,\varepsilon) \bigcup \partial\mathcal{V}_2(t,\varepsilon) \tag{4.9}$$

where the deformation is prescribed on $\partial\mathcal{V}_1(t,\varepsilon)$,

$$\underline{x} \in \partial\mathcal{V}_1(t,\varepsilon) \text{ is prescribed} ; \tag{4.10}$$

and the traction vector is prescribed on $\partial\mathcal{V}_2(t,\varepsilon)$,

$$\underline{\underline{T}}\cdot\underline{n}(\underline{x},t,\varepsilon) = \underline{t}_n(\underline{x},t,\varepsilon) \text{ is prescribed for} \tag{4.11}$$

$$\underline{x} \in \partial\mathcal{V}_2(t,\varepsilon) \text{ where } \underline{t}_n(\underline{x},t,0) = \underline{0} .$$

The attentive reader will notice that the prescription of the traction vector in (4.11) is given in terms of the Cauchy stress rather than the Piola-Kirchoff stress. We have already noted that in our local theory the distinction between the Cauchy and Piola-Kirchoff stress is downgraded because both forms lead to exactly the same perturbation equations.

In the same spirit it is convenient to prescribe displacements of the boundary $\partial\mathcal{V}_1(t,\varepsilon)$ of the deformed body where for simplicity we require that

$$\underline{x} - \underline{X} = \varepsilon \underline{U}(\underline{X}, t, \varepsilon), \qquad \underline{X} \in \partial\mathcal{V}_{10}, \ \underline{x} \in \partial\mathcal{V}_1(t,\varepsilon) \tag{4.12}$$

where

$$\partial\mathcal{V}_{10} = \partial\mathcal{V}_1(t,0)$$

is a portion of the boundary

$$\partial\mathcal{V}_0 = \partial\mathcal{V}_{10} \cup \partial\mathcal{V}_{20}$$

of the body in the natural state. Of course $\partial\mathcal{V}_{10}$ and $\partial\mathcal{V}_{20}$ are independent of time. Equation (4.12) says that the set of boundary points for which displacements are prescribed is a material set and no new material points, points on $\partial\mathcal{V}_{20}$, can enter this set as ε is varied.

To complete the prescription of the data for the initial-history problem we prescribe the initial history :

$$\underline{u}_0(\underline{X},t) \text{ is prescribed for } \underline{X} \in \mathcal{V}_0, \ t \leq 0$$

and

$$\underline{u}(\underline{X},t,\varepsilon) = \varepsilon \underline{u}_0(\underline{X},t) \text{ for } \underline{X} \in \mathcal{V}_0, \ t \leq 0 \ . \tag{4.13}$$

5. Kinematics for perturbations of the natural state

In our perturbation we develop a sequence of equations which may be systematically associated with a perturbation of data giving rise to the natural state. The data is all important and when we perturb it we induce a perturbation of the kinematics as well as of the constitutive equation. The perturbation formulas for the kinematic variables are easy to derive. Only the results are listed below.

$$\underline{u}(\underline{X},\tau,\varepsilon) = \varepsilon \underline{u}^{<1>}(\underline{X},\tau) + \varepsilon^2 \underline{u}^{<2>}(\underline{X},\tau) + 0(\varepsilon^3) \ . \tag{5.1}$$

$$\underline{\underline{F}}(\underline{X},\tau,\varepsilon) = \underline{\underline{1}} + \nabla \underline{u}(\underline{X},\tau) = \underline{\underline{1}} + \varepsilon \underline{\underline{F}}^{<1>}(\underline{X},\tau) + \varepsilon^2 \underline{\underline{F}}^{<2>}(X,\tau) + 0(\varepsilon^3) \tag{5.2}$$

where

$$\underline{\underline{F}}^{<n>}(\underline{X},\tau,\varepsilon) = \nabla \underline{u}^{<n>}(\underline{X},\tau), (F_{ij}^{<n>} = \partial u_i^{<n>}/\partial X_j) \ .$$

$$\underline{\underline{F}}^{-1} = \underline{\underline{1}} - \underline{\underline{F}}^{<1>}\varepsilon + (-\underline{\underline{F}}^{<2>} + \underline{\underline{F}}^{<1>}\,\underline{\underline{F}}^{<1>})\varepsilon^2 + 0(\varepsilon^3) \ . \tag{5.3}$$

$$\underline{\underline{G}}(s,\varepsilon) = \underline{\underline{F}}_t^{\,T}(\tau,\varepsilon)\,\underline{\underline{F}}_t(\tau,\varepsilon) - \underline{\underline{1}} = \varepsilon \underline{\underline{G}}^{<1>}(s) + \varepsilon^2 \underline{\underline{G}}^{<2>}(s) + 0(\varepsilon^3) \tag{5.4}$$

where

$$\underline{\underline{G}}^{<1>}(s) = 2\{\underline{\underline{E}}^{<1>}(t-s) - \underline{\underline{E}}^{<1>}(t)\} \ ,$$

$$\underline{\underline{G}}^{<2>}(s) = 2\{\underline{\underline{E}}^{<2>}(t-s) - \underline{\underline{E}}^{<2>}(t)\} + \underline{\underline{\xi}}^{<2>}(t,s) \ ,$$

$$\underline{\underline{E}}^{<n>} = \frac{1}{2}(\underline{\underline{F}}^{<n>} + \underline{\underline{F}}^{T<n>}) \ ,$$

and

$$\underline{\xi}^{<2>}(t,s) = \underline{F}^{T<1>}(t-s)\underline{F}^{<1>}(t-s) + \underline{F}^{T<1>}(t)\underline{F}^{<1>}(t) - 2\underline{F}^{T<1>}(t)\underline{E}^{<1>}(t-s)$$

$$+ \underline{F}^{<1>}(t)\underline{F}^{<1>}(t) + \underline{F}^{T<1>}(t)\underline{F}^{T<1>}(t) - 2\underline{E}^{<1>}(t-s)\underline{F}^{<1>}(t).$$

$$(5.5) \qquad \underline{B}(t,\varepsilon) = \underline{F}(t,\varepsilon)\underline{F}^{T}(t,\varepsilon) = \underline{1} + 2\varepsilon\underline{E}^{<1>}(t) + \varepsilon^{2}\{2\underline{E}^{<2>}(t) + \underline{F}^{<1>}(t)\underline{F}^{T<1>}(t)\} + 0(\varepsilon^{3})$$

$$(5.6) \qquad \det \underline{F} = 1 + \varepsilon \operatorname{tr} \underline{F}^{<1>} + \varepsilon^{2}\{\operatorname{tr} \underline{F}^{<2>} + \frac{1}{2}[\operatorname{tr} \underline{F}^{<1>}]^{2} - \frac{1}{2}\operatorname{tr}[\underline{F}^{<1>}\underline{F}^{<1>}]\} + 0(\varepsilon^{3}) .$$

$$\rho(\underline{x},t,\varepsilon) = \rho_{o}(\underline{X}) + \varepsilon\rho^{<1>}(\underline{X},t) + \varepsilon^{2}\rho^{<2>}(\underline{X},t) + 0(\varepsilon^{3}) .$$

Since ρ_o is independent of ε we may expand $\rho \det \underline{F} = \rho_o$ in powers of ε. Identifying independent powers of ε we find that

$$(5.7) \qquad \rho^{<1>} + \rho_{o} \operatorname{tr} \underline{F}^{<1>} = 0$$

and

$$(5.8) \qquad \rho^{<2>} + \rho^{<1>} \operatorname{tr} \underline{F}^{<1>} + \rho_{o} \operatorname{tr} \underline{F}^{<2>} + \frac{\rho_{o}}{2}[\operatorname{tr} \underline{F}^{<1>}]^{2} - \frac{\rho_{o}}{2}\operatorname{tr}[\underline{F}^{<1>}\underline{F}^{<1>}] = 0 .$$

We shall also need a formula for the pertrubation of the normal

$$(5.9) \qquad \underline{n} = \underline{N} + \varepsilon\underline{n}^{<1>} + 0(\varepsilon^{2}) .$$

To find $\underline{n}^{<1>}$ we write $\tilde{J} = da/dA$ and

$$\underline{n}\tilde{J} = \det \underline{F}(\underline{F}^{T})^{-1}.\underline{N} .$$

Combining (5.3), (5.6), (5.9) and $\tilde{J} = 1 + \varepsilon \tilde{J}^{<1>} + 0(\varepsilon^2)$ we get

$$\underline{n}^{<1>} + \underline{N}\,\tilde{J}^{<1>} = (\det \underline{\underline{F}}^{<1>})\underline{N} - \underline{\underline{F}}^{T<1>}\cdot\underline{N} \ .$$

Since $\underline{n}$ is a unit vector, $\underline{N}\cdot\underline{n}^{<1>} = 0$ and

$$\tilde{J}^{<1>} = \det \underline{\underline{F}}^{<1>} - \underline{N}\cdot\underline{\underline{F}}^{T<1>}\cdot\underline{N} \ .$$

Hence

$$\underline{n}^{<1>} = (\underline{N}\cdot\underline{\underline{F}}^{T<1>}\cdot\underline{N})\underline{N} - \underline{\underline{F}}^{T<1>}\cdot\underline{N} \ . \tag{5.10}$$

Using (5.10), we may write prescribed conditions for the traction vector $\underline{t}_n(\underline{x},t,\varepsilon)$ for $\underline{x} \in \partial\mathcal{V}_2(t,\varepsilon)$ in terms of perturbed Cauchy stresses $\underline{\underline{T}}^{<n>}(\underline{X},t)$ for $X \in \partial\mathcal{V}_{20}$. Thus

$$\begin{aligned}
\underline{t}_n = \underline{\underline{T}}\cdot\underline{n} &= (\varepsilon\underline{\underline{T}}^{<1>} + \varepsilon^2\underline{\underline{T}}^{<2>} + \ldots)\cdot(\underline{N} + \varepsilon\underline{n}^{<1>} + \ldots) \qquad (5.11)\\
&= \varepsilon\,\underline{\underline{T}}^{<1>}\cdot\underline{N} + \varepsilon^2(\underline{\underline{T}}^{<2>}\cdot\underline{N} + \underline{\underline{T}}^{<1>}\cdot\underline{n}^{<1>}) + 0(\varepsilon^3)\\
&= \varepsilon\,\underline{\underline{T}}^{<1>}\cdot\underline{N} + \varepsilon^2\{(\underline{\underline{T}}^{<2>} - \underline{\underline{T}}^{<1>}\,\underline{\underline{F}}^{T<1>})\cdot\underline{N}\\
&\qquad + (\underline{N}\cdot\underline{\underline{F}}^{T<1>}\cdot\underline{N})\underline{\underline{T}}^{<1>}\cdot\underline{N}\} + 0(\varepsilon^3)
\end{aligned}$$

gives the series expansion of $\underline{t}_n$ on $\partial\mathcal{V}_2$ in terms of $\underline{\underline{T}}^{<n>}$ and geometric quantities defined on $\partial\mathcal{V}_{20}$.

6. Canonical forms for the perturbation stresses and equations of motion for compressible solids

The canonical forms of the Cauchy stress for perturbations of the natural state

$$\underline{\underline{T}} = \varepsilon \underline{\underline{T}}^{<1>} + \varepsilon^2 \underline{\underline{T}}^{<2>} + 0(\varepsilon^3) \tag{6.1}$$

can be obtained by identification by combining (3.2, 6, 7, 8) with (5.4) and (5.5). We find that

$$\underline{\underline{T}}^{<1>}(t) = \underline{\underline{\tau}}[\underline{u}^{<1>}(t)] \overset{\text{def}}{=} 2\beta \underline{\underline{E}}^{<1>}(t) + 2\beta^{[1]} \operatorname{div} \underline{u}^{<1>}(t) \underline{\underline{1}} + \int_0^\infty \{\zeta(s) 2[\underline{\underline{E}}^{<1>}(t-s) - \underline{\underline{E}}^{<1>}(t)] + 2\zeta^{[1]}(s) \operatorname{div} [\underline{u}^{<1>}(t-s) - \underline{u}^{<1>}(t)] \underline{\underline{1}}\} ds \, , \tag{6.2}$$

and

$$\underline{\underline{T}}^{<2>} = \underline{\underline{\tau}}[\underline{u}^{<2>}] + \underline{\underline{h}}[\underline{u}^{<1>}] \tag{6.3}$$

where

$$\underline{\underline{h}}[\underline{u}^{<1>}] \overset{\text{def}}{=} \beta \underline{\underline{F}}^{<1>} \underline{\underline{F}}^{T<1>} + \beta^{[1]} \underline{\underline{1}} \operatorname{tr}[\underline{\underline{F}}^{<1>} \underline{\underline{F}}^{T<1>}] + \int_0^\infty \{\zeta(s) \underline{\underline{\xi}}^{<2>}(t,s) + \underline{\underline{1}} \zeta^{[1]}(s) \operatorname{tr} \underline{\underline{\xi}}^{<2>}(t,s)\} ds + \beta^{[2]} \underline{\underline{B}}^{<1>} \underline{\underline{B}}^{<1>} + \underline{\underline{1}} \beta^{[3]} (\operatorname{tr} \underline{\underline{B}}^{<1>})^2 + \beta^{[4]} \underline{\underline{1}} \operatorname{tr} [\underline{\underline{B}}^{<1>} \underline{\underline{B}}^{<1>}] + \beta^{[5]} \underline{\underline{B}}^{<1>} \operatorname{tr} \underline{\underline{B}}^{<1>}$$

$$+ \int_0^\infty \{\zeta^{[2]}(s)\,[\underline{\underline{B}}^{<1>}(t)\underline{\underline{G}}^{<1>}(s) + \underline{\underline{G}}^{<1>}(s)\underline{\underline{B}}^{<1>}(t)]$$

$$+ \zeta^{[3]}(s)\ \underline{\underline{G}}^{<1>}(s)\ \mathrm{tr}\ \underline{\underline{B}}^{<1>}(t) + \zeta^{[4]}(s)\ \underline{\underline{B}}^{<1>}(t)\mathrm{tr}\ \underline{\underline{G}}^{<1>}(s)$$

$$+ \zeta^{[5]}(s)\ \underline{\underline{1}}\ [\mathrm{tr}\ \underline{\underline{B}}^{<1>}(t)]\,[\mathrm{tr}\ \underline{\underline{G}}^{<1>}(s)]$$

$$+ \zeta^{[6]}(s)\underline{\underline{1}}\ \mathrm{tr}\ [\underline{\underline{B}}^{<1>}(t)\underline{\underline{G}}^{<1>}(s)]\}ds$$

$$+ \int_0^\infty\int_0^\infty \{\alpha(s_1,s_2)\ \underline{\underline{G}}^{<1>}(s_1)\underline{\underline{G}}^{<1>}(s_2)$$

$$+ \alpha^{[1]}(s_1,s_2)\ \underline{\underline{1}}\ [\mathrm{tr}\ \underline{\underline{G}}^{<1>}(s_1)]\,[\mathrm{tr}\ \underline{\underline{G}}^{<1>}(s_2)]$$

$$+ \alpha^{[2]}(s_1,s_2)\underline{\underline{1}}\ \mathrm{tr}\ [\underline{\underline{G}}^{<1>}(s_1)\ \underline{\underline{G}}^{<1>}(s_2)]$$

$$+ \alpha^{[3]}(s_1,s_2)\ \underline{\underline{G}}^{<1>}(s_1)\ \mathrm{tr}\ \underline{\underline{G}}^{<1>}(s_2)$$

$$+ \alpha^{[4]}(s_1,s_2)\ \underline{\underline{G}}^{<1>}(s_2)\ \mathrm{tr}\ \underline{\underline{G}}^{<1>}(s_1)\}ds_1\,ds_2 \quad .$$

The Piola-Kirchoff stress tensor is now given by (4.3), (5.3), (5.6) and (6.1,2,3) as

$$(6.4) \qquad \underline{\underline{S}}^{T} = \varepsilon\underline{\underline{T}}^{<1>} + \varepsilon^2\{\underline{\underline{T}}^{<2>} - \underline{\underline{T}}^{<1>}\ \underline{\underline{F}}^{T<1>} + \underline{\underline{T}}^{<1>}\ \mathrm{tr}\ \underline{\underline{F}}^{<1>}\} + 0(\varepsilon^3)$$

To characterize the motion of a particular compressible viscoelastic solid at first order we need values for

2 elastic constants β and $\beta^{[1]}$

and

2 material functions $\zeta(s)$ and $\zeta^{[1]}(s)$.

To characterize the motion of a particular compressible viscoelastic solid at second order we need values for

6 elastic constants β; $\beta^{[n]}$, n = 1,2,3,4,5

and

12 material functions $\zeta(s)$; $\zeta^{[n]}(s)$, n = 1,2,3,4,5,6; $\alpha(s_1,s_2)$ and $\alpha^{[\ell]}(s_1,s_2)$, ℓ = 1,2,3,4 . Obviously, the rheometrical problem of material characterization in the second order theory of motions of viscoelastic solids perturbing the natural state is very hard because there are so many material functions and constants.

We have seen that the expansion of $\underline{u}(\underline{X},\tau,\varepsilon)$ in powers of ε ultimately induces an expansion of the stresses $\underline{\underline{T}}$ and $\underline{\underline{S}}^T$ in powers of ε. The expansion of $\underline{u}(\underline{X},\tau,\varepsilon)$ for $\underline{X} \in \mathcal{V}_0$ was presumed given, but it is not given; it must be determined from solutions of the equations which perturb the natural state. The perturbation of the data driving the motion is given and it induces the expansion of all the other interlocked quantities.

The equations which perturb the natural state arise from (4.8), (4.11), (4.12), (4.13) with $\underline{b} = \varepsilon \underline{b}$ by identification when all of the variables are expanded in powers of ε. The expansion of the stress is given by (6.1-3) and the expansion of the kinematic variables by (5.1-11). At first order

$$\rho_o \ddot{\underline{u}}^{<1>} = \rho_o \underline{b} \qquad + \operatorname{div} \underline{\underline{T}}[\underline{u}^{<1>}] \quad \text{for} \quad \underline{X} \in \mathcal{V}_0, \; t > 0 \; ;$$

$$\underline{u}^{<1>}(\underline{X},t) \qquad \text{is prescribed for } \underline{X} \in \partial\mathcal{V}_{10},\ t > 0\ ;$$

$$\underline{\underline{T}}^{<1>}(\underline{X},t)\cdot\underline{N} = \underline{t}_{n}^{<1>}(\underline{X},t) \text{ is prescribed for } \underline{X} \in \partial\mathcal{V}_{20},\ t > 0\ ;$$

(6.5)

$$\underline{u}^{<1>}(\underline{X},t) \text{ is prescribed for } \underline{X} \in \mathcal{V}_0,\ t \leq 0\ .$$

For orders $n > 1$ we find that in $\mathcal{V}_0$ and for $t > 0$

$$\rho_o\, \ddot{\underline{u}}^{<n>} = \operatorname{div} \underline{\underline{\tau}}[\underline{u}^{<n>}] + \text{terms of lower order} \tag{6.6}$$

where

$$\underline{u}^{<n>}(\underline{X},t) \text{ is prescribed in terms of lower order for}$$

$$X \in \partial\mathcal{V}_{10}, \quad t > 0\ ;$$

$$\underline{\underline{T}}^{<n>}(\underline{X},t)\cdot\underline{N} \text{ is prescribed in terms of lower order}$$

$$\text{for } X \in \partial\mathcal{V}_{20},\ t > 0\ ;$$

and

$$\underline{u}^{<n>}(\underline{X},t) = \underline{0} \quad \text{for } X \in \mathcal{V}_0, \quad t \leq 0\ .$$

For example when $n = 2$, the terms of lower order in $(6.6)_1$ are

$$\operatorname{div}\{\underline{\underline{\mathcal{L}}}[\underline{u}^{<1>}] - \underline{\underline{T}}^{<1>}\,\underline{\underline{F}}^{T<1>} + \underline{\underline{T}}^{<1>}\,\operatorname{tr}\,\underline{\underline{F}}^{<1>}\} \ . \tag{6.7}$$

The perturbation equations can be solved sequentially and at each step of the sequence there are three equations for the three unknown components of $\underline{u}^{<n>}$.

Changes in density due to deformation may be expressed by (5.7) and (5.8). Similar formulas hold at higher orders.

The practical utility of a theory which requires knowing the value of 6 elastic constants and 12 material functions is debatable. Pipkin (1964) noticed that there is a big reduction in the number of unknown material parameters when the material is incompressible.

7. Canonical forms for the perturbation stresses and equations of motion for incompressible solids

Incompressible solids have been discussed in §3. We have already explained that in the incompressible case we may group all terms of

$$\underline{\underline{T}} = - p\underline{\underline{1}} + \underline{\underline{\mathcal{F}}}[\underline{\underline{B}}(t), \underline{\underline{G}}(s)]_{s=0}^{\infty}$$

which are proportional to $\underline{\underline{1}}$ with -p. The $\underline{\underline{\mathcal{F}}}$ part of $\underline{\underline{T}}$ is constitutively determined while the spherically symmetric $p\underline{\underline{1}}$ part of $\underline{\underline{T}}$ is to be determined from the equations of motions. So one simplification in the form of $\underline{\underline{\mathcal{F}}}$ comes from dumping terms of (3.6,7,8) proportional to $\underline{\underline{1}}$ into -p. A second simplification comes from setting det $\underline{\underline{F}} = 1$ in (5.6). Then

$$\text{(7.1)} \qquad \operatorname{tr} \underline{\underline{F}}^{<1>} = \operatorname{div} \underline{u}^{<1>} = 0 \quad ,$$

and

$$\text{(7.2)} \qquad \operatorname{tr} \underline{\underline{F}}^{<2>} = \operatorname{div} \underline{u}^{<2>} = \frac{1}{2} \operatorname{tr}[\underline{\underline{F}}^{<1>} \underline{\underline{F}}^{<1>}] \ .$$

It follows from (5.4), (5.5) and (7.1) that

$$\text{(7.3)} \qquad \operatorname{tr} \underline{\underline{G}}^{<1>}(s) = \operatorname{tr} \underline{\underline{B}}^{<1>} = 0 \quad ,$$

and from (5.5) and (7.2) that

$$\text{(7.4)} \qquad \operatorname{tr} \underline{\underline{B}}^{<2>} = 2 \operatorname{div} \underline{u}^{<2>} + \operatorname{tr}[\underline{\underline{F}}^{<1>} \underline{\underline{F}}^{T<1>}]$$

$$= \operatorname{tr}[\underline{\underline{F}}^{<1>} \underline{\underline{F}}^{<1>} + \underline{\underline{F}}^{<1>} \underline{\underline{F}}^{T<1>}] \ ,$$

and from (5.4) and (7.2) that

$$\text{tr } \underline{\underline{G}}^{<2>}(s) = 2 \text{ div } [\![\underline{u}^{<2>}]\!] + \text{tr } \underline{\underline{\Xi}}^{<2>}(t,s) \tag{7.5}$$

$$= \text{tr}[\![\underline{\underline{F}}^{<1>}\underline{\underline{F}}^{<1>}]\!] + \text{tr } \underline{\underline{\Xi}}^{<2>}(t,s)$$

where $[\![\cdot]\!]$ is a jump operator on whose domain are functions a of t.

$$[\![a]\!] \overset{\text{def}}{=} a(t-s) - a(t) \ .$$

The perturbed stresses are given by

$$\underline{\underline{T}}^{<1>} = -p^{<1>}\underline{\underline{1}} + 2\beta\underline{\underline{E}}^{<1>}(t) + 2\int_0^\infty \zeta(s)\,[\![\underline{\underline{E}}^{<1>}]\!]ds \tag{7.6}$$

$$= -p^{<1>}\underline{\underline{1}} + 2\gamma\underline{\underline{E}}^{<1>}(t) + 2\int_0^\infty \zeta(s)\underline{\underline{E}}^{<1>}(t-s)ds$$

where

$$\gamma = \beta - \int_0^\infty \zeta(s)ds \ , \tag{7.7}$$

and

$$\underline{\underline{T}}^{<2>} = -\pi^{<2>}\underline{\underline{1}} + 2\gamma\underline{\underline{E}}^{<2>}(t) + 2\int_0^\infty \zeta(s)\underline{\underline{E}}^{<2>}(t-s)ds \tag{7.8}$$

$$+ \beta\underline{\underline{F}}^{<1>}\,\underline{\underline{F}}^{T<1>} + \beta^{[2]}\underline{\underline{B}}^{<1>}\,\underline{\underline{B}}^{<1>} + \int_0^\infty \zeta(s)\underline{\underline{\Xi}}^{<2>}(t,s)ds$$

$$+ \int_0^\infty \zeta^{[2]}(s)\,[\underline{\underline{B}}^{<1>}(t)\underline{\underline{G}}^{<1>}(s) + \underline{\underline{G}}^{<1>}(s)\underline{\underline{B}}^{<1>}(t)]ds$$

$$+ \int_0^\infty \int_0^\infty \alpha(s_1,s_2)\underline{\underline{G}}^{<1>}(s_1)\ \underline{\underline{G}}^{<1>}(s_2)ds_1\ ds_2 \tag{7.8}$$

where

$$\pi^{<2>} = p^{<2>} - \beta^{[1]}\ tr(\underline{\underline{F}}^{<1>}\underline{\underline{F}}^{<1>} + \underline{\underline{F}}^{<1>}\underline{\underline{F}}^{T<1>}) - \beta^{[4]}tr[\underline{\underline{B}}^{<1>}\underline{\underline{B}}^{<1>}] \tag{7.9}$$

$$- \int_0^\infty \zeta^{[1]}(s)\{tr\ [\![\underline{\underline{F}}^{<1>}\underline{\underline{F}}^{<1>}]\!] + tr\ \underline{\underline{\mathcal{S}}}^{<2>}(t,s)\}ds$$

$$- \int_0^\infty \zeta^{[6]}(s)\ tr[\underline{\underline{B}}^{<1>}(t)\underline{\underline{G}}^{<1>}(s)]ds$$

$$- \int_0^\infty \int_0^\infty \alpha^{[2]}(s_1,s_2)\ tr\ [\underline{\underline{G}}^{<1>}(s_1)\underline{\underline{G}}^{<1>}(s_2)]\ ds_1\ ds_2 \quad .$$

To characterize the motion of a particular incompressible viscoelastic solid at first order we need values for

1 elastic constant β

and

1 material function $\zeta(s)$.

To characterize the motion of an incompressible solid at second order we need

2 elastic constants β and $\beta^{[2]}$

and

3 material functions $\zeta(s)$, $\zeta^{[2]}(s)$ and $\alpha(s_1,s_2)$.

The material constants appearing in $\pi^{<2>}$ are constitutively undetermined since $\pi^{<n>}$ is to be determined as one of the four unknown fields in the canonical problems governing the perturbation displacements.

Turning next to these canonical problems we find that

$$\text{(7.10)}\qquad \mathbb{J}\underline{u}^{\langle 1\rangle} + \nabla p^{\langle 1\rangle} \overset{*}{=} \underline{0}\ ,\quad \operatorname{div}\underline{u}^{\langle 1\rangle} = 0 \quad \text{in } \mathcal{V}_0\ ,\ t > 0\ ;$$

$$\underline{u}^{\langle 1\rangle}(\underline{X},t) \text{ is prescribed for } \underline{X} \in \partial\mathcal{V}_{10}\ ,\ t > 0\ ;$$

$$\underline{t}_n{}^{\langle 1\rangle} = \underline{\underline{T}}^{\langle 1\rangle}\cdot\underline{N} \text{ is prescribed for } \underline{X} \in \partial\mathcal{V}_{20}\ ,\ t > 0\ ;$$

$$\underline{u}^{\langle 1\rangle}(\underline{X},t) = \underline{u}_0(\underline{X},t) \quad \text{for } \underline{X}\in\mathcal{V}_0\ ,\quad t \le 0\ .$$

In (7.10) and elsewhere

$$\text{(7.11)}\qquad \mathbb{J}(\cdot) \overset{\text{def}}{=} \rho_o\,(\ddot{\cdot}) - \gamma\nabla^2(\cdot) - \nabla^2\int_0^\infty \zeta(s)(\cdot)(t-s)ds\ .$$

At second order

$$\text{(7.12)}\qquad \mathbb{J}\underline{u}^{\langle 2\rangle} + \nabla\pi^{\langle 2\rangle} = \underline{M}_2,\quad \operatorname{div}\underline{u}^{\langle 2\rangle} = \theta_2 \quad \text{in } \mathcal{V}_0\ ,\ t > 0\ ;$$

$$\underline{u}^{\langle 2\rangle}(\underline{X},t) \text{ is prescribed for } \underline{X} \in \partial\mathcal{V}_{10},\ t > 0\ ;$$

$$\underline{t}_n{}^{\langle 2\rangle} = \underline{\underline{T}}^{\langle 2\rangle}\cdot\underline{N} + \underline{\underline{T}}^{\langle 1\rangle}\cdot\underline{n}^{\langle 1\rangle} = (\underline{\underline{T}}^{\langle 2\rangle} - \underline{\underline{T}}^{\langle 1\rangle}\,\underline{\underline{F}}^{T\langle 1\rangle})\cdot\underline{N}$$

$$+ (\underline{N}\cdot\underline{\underline{F}}^{T\langle 1\rangle}\cdot\underline{N})\underline{\underline{T}}^{\langle 1\rangle}\cdot\underline{N} \text{ is prescribed for } \underline{X} \in \partial\mathcal{V}_{20}, t>0\ ;$$

and

$$\underline{u}^{\langle 2\rangle}(\underline{X},t) = \underline{0} \quad \text{for } X \in \mathcal{V}_0\ ,\quad t \le 0\ .$$

In (7.12)

*Henceforth, we neglect the body force.

$$\theta_2 = \frac{1}{2}\,\mathrm{tr}[\underline{\underline{F}}^{<1>}\underline{\underline{F}}^{<1>}] \ ,$$

and

$$\underline{M}_2 = \underline{\underline{F}}^{T<1>}\cdot\nabla p^{<1>} + \mathrm{div}\,\{\beta^{[2]}\ \underline{\underline{B}}^{<1>}\underline{\underline{B}}^{<1>}$$

$$+ \int_0^\infty \zeta(s)\,[-\underline{\underline{G}}^{<1>}(s)\ \underline{\underline{B}}^{<1>}(t) + 2[\![\underline{\underline{F}}^{T<1>}]\!]\ \underline{\underline{E}}^{<1>}(t-s)]\,ds$$

$$+ \int_0^\infty \zeta^{[2]}(s)\,[\underline{\underline{B}}^{<1>}(t)\underline{\underline{G}}^{<1>}(s) + \underline{\underline{G}}^{<1>}(s)\ \underline{\underline{B}}^{<1>}(t)]\,ds$$

$$+ \int_0^\infty\int_0^\infty \alpha(s_1,s_2)\underline{\underline{G}}^{<1>}(s_1)\ \underline{\underline{G}}^{<1>}(s_2)\,ds_1\,ds_2\} \ . \tag{7.13}$$

In deriving $\underline{M}_2$ we made use of the identity

$$\frac{1}{2}\,\nabla\mathrm{tr}\ \underline{\underline{F}}^{<1>}\underline{\underline{F}}^{<1>} = \mathrm{div}\,\{\underline{\underline{F}}^{\,T<1>}\underline{\underline{F}}^{T<1>}\}$$

which holds whenever $\mathrm{div}\ \underline{u}^{<1>} = 0$. Many terms in the expression $\mathrm{div}\,\{\underline{\underline{T}}^{<2>} - \underline{\underline{T}}^{<1>}\ \underline{\underline{F}}^{T<1>}\}$ vanish.

At every order $n \geq 1$ we find that $\underline{u}^{<n>} = \underline{v}$ and $\pi^{<n>} = \pi$ satisfy

$$\left.\begin{aligned} \mathcal{J}\underline{v} + \nabla\pi &= \underline{f}_1(\underline{X},t) \\ \mathrm{div}\ \underline{v} &= f_2(\underline{X},t) \end{aligned}\right\} \quad \text{for } \underline{X}\in\mathcal{V}_0\ ,\ t>0\ ; \tag{7.14}$$

$$\underline{v}(\underline{X},t) = \underline{f}_3(\underline{X},t) \quad \text{for } \underline{X}\in\partial\mathcal{V}_{10}\ ,\ t>0\ ;$$

$$(-\pi\underline{\underline{1}} + 2\gamma\,\underline{\underline{E}}[\underline{v}(t)] + 2\int_0^\infty \zeta(s)\underline{\underline{E}}[\underline{V}(t-s)]\,ds)\cdot\underline{N}$$

$$= \underline{f}_4(\underline{X},t) \quad \text{for } \underline{X}\in\partial\mathcal{V}_{20}\ ,\ t>0\ ;$$

$$\underline{v}(\underline{X},t) = \underline{f}_5(\underline{X},t) \text{ for } \underline{X} \in \mathcal{V}_0, \; t \leq 0$$

where $\underline{\underline{E}}[\underline{v}(t)] = \frac{1}{2}(\nabla \underline{v}(\underline{X},t) + \text{transpose})$ and the right hand sides of (7.14) are known from the prescribed data and lower order solutions. At each order we must solve four equations for the fields $\underline{v}(X,t)$ and $\pi(\underline{X},t)$ in $\mathcal{V}_0$.

8. Stability and uniqueness of solutions of the canonical equations perturbing the natural state

The main aim of this section is to show unique solvability for the sequence of perturbation problems (7.14) for incompressible, viscoelastic solids. We do not consider the problem of existence, however, and restrict ourselves to a discussion of uniqueness and the related problem of stability. It is probable that an existence and uniqueness theory of the type recently given by Slemrod (1977) for Joseph's (1976) theory of motions which perturb the rest state of simple fluids can be adapted to the present problem. But here we follow a different path.

To motivate the analysis we remind the reader that the theory of slow flow of Navier-Stokes fluids is a relatively uncomplicated subject because when the flow is slow (or the Reynolds number is small) there is just one solution in the long run and it is uniquely determined by the boundary conditions and body forces, independent of initial conditions. The unique solution is globally stable in the sense that all disturbances, small or large, of this solution ultimately decay. So we expect to observe in nature what we calculate from the equations when the flow is slow. And we do. But this simplicity is lost when the flow is not slow because there are many solutions for prescribed boundary conditions and body forces and many of these are unstable.

The situation of viscoelastic fluids and solids is not so different. In any event, when the forcing data is small, the rest state or natural state is stable in the linearized approximation provided only that material parameters and their derivatives have

the expected sign . For larger forcing data the problem of stability is probably at least as complicated as in the Navier-Stokes theory. In fact, viscoelastic materials can exhibit shock-ups and loss of existence of smooth solutions without parallel in the Navier-Stokes theory.

Of course, stability can never be asserted on the basis of linearized equations alone because linearized equations do not govern the evolution of large disturbances. So our statements about stability are at best conditional, subject to the restriction that disturbances are sufficiently small. In fact, since conditional stability theorems are not known for viscoelastic materials, it has to be assumed that the analysis of the linearized equations applies to the nonlinear equations when the nonlinear part is small.

The linearized stability problem for the stability of the natural state may be obtained from the linearization of the initial history problem (4.8), (4.11), (4.12), (4.13) and (3.9) for an infinitesimal disturbance $\underline{v}$ of $\underline{u} \equiv \underline{0}$:

$$\left.\begin{aligned} &\mathbb{J}\underline{v} + \nabla\pi = \underline{0} \\ &\operatorname{div} \underline{v} = 0 \end{aligned}\right\} \quad \text{for } \underline{X} \in \mathcal{V}_0, t > 0;$$

(8.1) $$\underline{v}(\underline{X},t) = \underline{0} \text{ for } \underline{X} \in \mathcal{V}_{10},\ t > 0;$$

$$\{-\pi\underline{\underline{1}} + 2\gamma\underline{\underline{E}}[\underline{v}(t)] + 2\int_0^\infty \zeta(s)\,\underline{\underline{E}}[\underline{v}(t-s)]ds\} \cdot \underline{N}$$

$$= \underline{0} \text{ for } \underline{X} \in \mathcal{V}_{20},\ t > 0;$$

$$\underline{v}(\underline{X},t) = \underline{v}_0(\underline{X},t) \text{ is prescribed for } \underline{X} \in \mathcal{V}_0 \text{ and } t \le 0.$$

We are interested in finding the conditions under which $\underline{v} \to 0$ as $t \to \infty$.

The problem (8.1) is identical to the problem which governs

the uniqueness of solutions of the initial-history problem (7.14) if $\underline{v}_0(\underline{X},t)$ is set to zero. To study uniqueness we consider two solutions of (7.14) with same prescribed data and initial histories. The difference between these two solutions satisfies (8.1) with $\underline{v}_0(\underline{X},t) = \underline{0}$.

Uniqueness theorems for linearized initial-history problems for viscoelastic solids under slightly different conditions and constitutive assumptions have been given by Edelstein and Gurtin (1964), Odeh and Tadjbakhsh(1965), Gurtin and Sternberg (1962), Breuer and Onat (1962) and Onat and Breuer (1963). If $\underline{v} = \underline{0}$ is asymptotically stable then the solution $\underline{v} = \underline{0}$ corresponding to a zero initial history is automatically unique. On the other hand uniqueness for the initial history problem does not imply stability since new solutions with $\underline{v} \neq \underline{0}$ may arise from the loss of stability of $\underline{v} = \underline{0}$. Loss of stability is associated with the evolution of disturbances, new initial conditions, which are the inevitable results of fluctuations in the prescribed data.

Unique solvability is intimately connected with stability and has almost no relation to the problem of uniqueness of the initial-history problem. In the present circumstances unique solvability comes down to a verification that $\mathbb{J}$ is uniquely invertible, and new solutions with $\underline{v} \neq \underline{0}$ cannot bifurcate.

Existence, uniqueness and asymptotic stability of <u>generalized</u> solutions of equations like ours have been given by Dafermos (1970) for problems in which displacements are prescribed on the entire boundary $\partial\mathcal{V}_0$ of $\mathcal{V}_0$. In the problem treated by Dafermos the vector field $\underline{v}$ is not necessarily solenoidal and $\pi = 0$. Our problem can be reduced to the one considered by Dafermos by projection techniques used in mathematical studies of the Navier-

Stokes equations (Fujita and Kato, 1964 ; Ladyzhenskaya, 1963). In this method one introduces a Hilbert space H of vectors with scalar product

$$\langle \underline{a}, \underline{b} \rangle = \int_{\mathcal{V}_0} \underline{a}(\underline{X}) \cdot \underline{b}(\underline{X}) d\mathcal{V}_0 \quad ,$$

and norm $||a|| = \langle \underline{a}, \underline{a} \rangle^{1/2}$ by completing the $C^\infty(\mathcal{V}_0)$ vectors with compact support. The compact support is natural for problems like the one which governs $\underline{v}$ when $\partial\mathcal{V}_{10} = \partial\mathcal{V}_0$ so that $\underline{v} = \underline{0}$ for $\underline{X} \in \partial\mathcal{V}_0$, in which the function is prescribed to be zero on the boundary $\partial\mathcal{V}_0$ of $\mathcal{V}_0$. For such problem it is possible to decompose H into orthogonal subspaces of solenoidal vectors(H_1) and gradients(H_2), $H = H_1 \oplus H_2$. There is then the orthogonal projection $\mathbb{P}$ which commutes with $\mathbb{J}$ and annihilates gradients. So we get

$$\begin{aligned} &\mathbb{J}\mathbb{P}\underline{v} = \underline{0} \text{ in } \mathcal{V}_0, \ t > 0; \\ &\mathbb{P}\underline{v} = \underline{0} \text{ for } \underline{X} \in \partial\mathcal{V}_0, \ t > 0; \\ &\mathbb{P}\underline{v} = \mathbb{P}\underline{v}_0 \text{ is prescribed in } \mathcal{V}_0, \ t \leq 0. \end{aligned} \tag{8.2}$$

This problem, (8.2), falls in the frame of the study of Dafermos (1970) who shows that $\langle \dot{\underline{v}}, \dot{\underline{v}} \rangle (t) \to 0$ and $\langle \nabla\underline{v}, \nabla\underline{v} \rangle (t) \to 0$ provided that

$$\begin{aligned} &1. \quad \beta > 0 \ , \\ &2a. \quad \zeta, \dot{\zeta} \in C^0[0,\infty) \cap L^1[0,\infty) \ , \\ &2b. \quad \zeta(s) \leq 0 \text{ for } s \in [0,\infty) \ , \\ &2c. \quad \dot{\zeta}(s) \geq 0 \text{ for } s \in [0,\infty) \ , \\ &2d. \quad \zeta \text{ does not vanish identically} \ . \end{aligned} \tag{8.3}$$

The conditions required by Dafermos for asymptotic stability do not disagree with conditions which rheologists would require on physical grounds using experience and intuition.

The conditions on β and $\zeta(s)$ derived by Dafermos are the desired conditions which guarantee unique solvability of (7.14) when displacements are prescribed over all of $\partial\mathcal{V}_0$ ($\partial\mathcal{V}_{10}=\partial\mathcal{V}_0$, $\partial\mathcal{V}_{20} = 0$). We therefore turn next to the mixed problem using elementary methods in which the source of restrictions on β and $\zeta(s)$ will be easy to interpret.

We start by deriving a spectral problem for solutions of (8.1) using the method of the exponential time factor:

$$\underline{v}(\underline{X},t) = e^{\sigma t}\,\underline{v}(\underline{X}) \ ,$$

$$\pi(\underline{X},t) = e^{\sigma t}\pi(\underline{X}) \tag{8.4}$$

where

$$\sigma = \xi + i\omega \ ,$$

and σ, $v_i(\underline{X})$ and $\pi(\underline{X})$ satisfy equations obtained from (8.1) and (8.4).

$$\rho_o\sigma^2 v_i(\underline{X}) = \kappa(\sigma)\nabla^2 v_i(\underline{X}) - \frac{\partial\pi}{\partial X_i}(\underline{X}) \ ,$$

and

$$\partial v_i/\partial X_i = 0$$

in $\mathcal{V}_0$ and

$$v_i(\underline{X}) = 0 \text{ for } \underline{X} \in \partial\mathcal{V}_{10},$$

and

$$\{-\pi\delta_{ij} + \kappa(\sigma)\left(\frac{\partial v_i}{\partial X_j} + \frac{\partial v_j}{\partial X_i}\right)\} \quad N_j = 0 \text{ for } \underline{X} \in \partial\mathcal{V}_{20}$$

where

$$\kappa(\sigma) = \beta + \int_0^\infty \zeta(s)\{e^{-\sigma s}-1\}ds = \gamma + \int_0^\infty \zeta(s)e^{-\sigma s}ds.$$

After introducing new variables

$$\lambda = -\kappa(\sigma)/\rho_o\sigma^2;\ \tilde{\pi}(\underline{X}) = \pi(\underline{X})/\rho_o\sigma^2$$

we may rewrite the spectral problem as

$$\underline{v} + \lambda\nabla^2\underline{v} + \nabla\tilde{\pi} = \underline{0},\ \text{div}\ \underline{v} = 0\ \text{in}\ \mathcal{V}_0\ ,$$

$$(8.5)\qquad v_i = 0 \qquad \text{for}\ \underline{X} \in \partial^{\mathcal{N}}\mathcal{V}_{10}\ ,$$

$$\{\tilde{\pi}\delta_{ij} + \lambda(\frac{\partial v_i}{\partial X_j} + \frac{\partial v_j}{\partial X_i})\}\ N_j = 0 \ \text{for}\ \underline{X} \in \partial\mathcal{N}_{20}.$$

Eqs. (8.5) define a spectral problem. We seek the values of λ for which (8.5) has solutions $(\underline{v},\tilde{\pi}) \neq (\underline{0},0)$. The following properties characterize the spectrum of (8.5):

(1) The spectrum $\Sigma(\lambda)$ of (8.5) is a pure point spectrum, that is, λ are eigenvalues of (8.5) .

(2) The eigenvalues λ of (8.5) are real-valued .

(3) The number of eigenvalues is countably infinite. They are of finite multiplicity, all semi-simple and may be arranged as a decreasing sequence clustering at zero:

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq 0.$$

If the measure of $\partial^{\mathcal{N}}\mathcal{V}_{10}$ is greater than zero then $\lambda_1 < \infty$.

Proof: To prove the asserted properties we show that the λ_i are the critical points of the functional

$$(8.6)\qquad \lambda[\underline{v}] \stackrel{\text{def}}{=} \frac{\langle|\underline{v}|^2\rangle}{\langle|\underline{\underline{A}}(\underline{v})|^2\rangle}$$

where $\langle \cdot \rangle = \int_{\mathcal{V}_0} (\cdot) d\mathcal{V}_0$,

$$|\underline{v}|^2 = v_i v_i, \quad |\underline{\underline{A}}(\underline{v})|^2 = A_{ij}(\underline{v}) A_{ij}(\underline{v}) ,$$

$$A_{ij}(\underline{v}) = 2E_{ij}(\underline{v}) = \frac{\partial v_i}{\partial X_j} + \frac{\partial v_j}{\partial X_i} ,$$

among functions

$$\underline{v} \in H = [\underline{v}: \operatorname{div} \underline{v} = 0, \ \underline{v}\Big|_{\partial\mathcal{V}_{10}} = \underline{0}, \ \langle |\underline{\underline{A}}(\underline{v})|^2 \rangle < \infty] .$$

Every solution of (8.5) satisfies

$$\langle \underline{v}, \ \underline{v} + \lambda \operatorname{div} \underline{\underline{A}}(\underline{v}) + \nabla\tilde{\pi} \rangle = \langle |\underline{v}|^2 \rangle + \lambda \langle v_i \frac{\partial A_{ij}(\underline{v})}{\partial X_j} \rangle$$

$$+ \langle v_i \frac{\partial \tilde{\pi}}{\partial X_i} \rangle = \langle |\underline{v}|^2 \rangle - \lambda \langle A_{ij}(\underline{v}) A_{ij}(\underline{v}) \rangle$$

$$+ \int_{\partial\mathcal{V}_{20}} v_i N_j [\tilde{\pi}\delta_{ij} + \lambda A_{ij}(\underline{v})] d\Sigma$$

$$= \langle |\underline{v}|^2 \rangle - \lambda \langle |\underline{\underline{A}}(\underline{v})|^2 \rangle = 0$$

where the scalar product $\langle , \rangle$ is as defined before. The eigenvalues $\lambda_n \neq \lambda$ are critical points of the Rayleigh quotient

$$(8.7) \qquad \lambda_n = \max_{H_n} \lambda[\underline{v}]$$

where H_n is the subspace of the space H which is orthogonal to eigensubspaces of the first n-1 eigenvectors. The asserted properties are all known consequences of the variational characterization of eigenvalues (see, for example, Reisz-Nagy, 1955, page 232). The condition measure $\partial\mathcal{V}_{10} > 0$ rules out the solution $\underline{v} = \text{const} \neq \underline{0}$ and guarantees the existence of a finite least upper bound for $\lambda[\underline{v}]$, $\underline{v} \in H$.

Next, we show that solutions of (8.7) are also the eigenvalues of Eqs. (8.5). We reformulate (8.7) as

$$\lambda_1 = \max_{\underline{v} \in \mathcal{H}} \hat{\lambda}[\underline{v}] \qquad \text{where} \tag{8.8}$$

$$\mathcal{H} = [\underline{v}: \underline{v}|_{\partial\mathcal{V}_{10}} = \underline{0},\ \langle|\underline{\underline{A}}(\underline{v})|^2\rangle < \infty] \ , \tag{8.9}$$

$$\hat{\lambda}[\underline{v}] = \{\langle|\underline{v}|^2\rangle - 2\langle q \text{ div } \underline{v}\rangle\}/\langle|\underline{\underline{A}}(\underline{v})|^2\rangle , \tag{8.10}$$

and

$q \in C^1(\mathcal{V}_0)$ is a Lagrange multiplier.

Let $\underline{v}_1$ be the maximizing function. It satisfies the constraint div $\underline{v}_1 = 0$. For any $\underline{\phi} \in \mathcal{H}$ and any real ε

$$\frac{d}{d\varepsilon}\hat{\lambda}[\underline{v}_1 + \varepsilon\underline{\phi}]\Big|_{\varepsilon=0} = \frac{2}{\langle|\underline{\underline{A}}(\underline{v}_1)|^2\rangle}\{\langle\underline{v}_1, \underline{\phi}\rangle - \langle q \text{ div } \underline{\phi}\rangle \tag{8.11}$$

$$- \lambda_1\langle\underline{\underline{A}}(\underline{v}_1) \cdot \underline{\underline{A}}(\underline{\phi})\rangle\} = 0 \ .$$

Hence

$$\langle\underline{v}_1, \underline{\phi}\rangle - \langle q \text{ div } \underline{\phi}\rangle - \lambda_1\langle\underline{\underline{A}}(\underline{v}_1) \cdot \underline{\underline{A}}(\underline{\phi})\rangle = 0 \ . \tag{8.12}$$

After some integration by parts, we get

$$\langle(\underline{v}_1 + \lambda_1\nabla^2\underline{v}_1 + \nabla q), \underline{\phi}\rangle - \int_{\partial\mathcal{V}_{20}} N_j\phi_i[q\delta_{ij} + \lambda_1 A_{ij}(\underline{v}_1)] = 0 \ . \tag{8.13}$$

Eq. (8.13) vanishes for all $\underline{\phi} \in \mathcal{H}$; in particular for $\underline{\phi}$ such that $\underline{\phi} = 0$ on $\partial\mathcal{V}_{20}$. The fundamental lemma of the calculus of variation implies that

$$\underline{v}_1 + \lambda_1\nabla^2\underline{v}_1 + \nabla q = \underline{0} \text{ in } \mathcal{V}_0 \ .$$

Now $\int_{\partial\mathcal{V}_{20}} N_j\phi_i[q\delta_{ij} + \lambda_1 A_{ij}(\underline{v}_1)] = 0$ for arbitrary ϕ.
Hence the boundary condition $(8.5)_3$

$$[q\,\underline{\underline{1}} + \lambda_1\underline{\underline{A}}(\underline{v}_1)] \cdot \underline{N} = \underline{0}$$

arises as a natural boundary condition for the variational problem.

It follows now that (8.5) are Euler equations for $\hat{\lambda}[\underline{v}]$, $\underline{v} \in \mathcal{H}$ subject to the constraint div $\underline{v} = 0$. In a similar fashion it can be shown that solutions of Eq. (8.7) are also the eigenvalues of Eqs. (8.5). Or equivalently all the eigenvalues of Eqs. (8.5) may be characterized variationally by Eq. (8.7).

Having determined the eigenvalues $\lambda_n > 0$ of (8.5) we return to problem of stability of $\underline{u} = \underline{0}$. In the context of the spectral problem $\underline{u} = \underline{0}$ is stable if Re $\sigma = \xi < 0$, neutrally stable if $\xi = 0$ and unstable if $\xi > 0$. The determination of the sign of ξ may be made by analysis of the functional equation

$$-\frac{\rho_o \sigma^2}{\kappa(\sigma)} = 1/\lambda > 0, \quad \sigma = \xi + i\omega,$$

that is,

$$m\{\beta + \int_0^\infty \zeta(s)[e^{-\sigma s} - 1]ds\} = -\sigma^2 \tag{8.19}$$

where $m = 1/(\rho_o \lambda) > 0$.

We need to determine the conditions on β and $\zeta(s)$ which will guarantee $\xi < 0$ for all $m > 0$. If there is only elasticity, and no viscoelasticity, ($\zeta(s) = 0$), then (8.19) reduces to $m\beta = -\sigma^2$. If $\beta < 0$ then $\sigma = \xi = \pm \sqrt{m|\beta|}$ so that $\underline{u} = 0$ is unstable if $\beta < 0$. Hence, for stability $\beta > 0$. This is condition (1) of Dafermos.

The other conditions of Dafermos also follow from the analysis of (8.19). When $\omega = 0$, Eq. (8.19) reduces to

$$m\{\beta + \int_0^\infty \zeta(s)[e^{-\xi s} - 1]ds\} = -\xi^2 \tag{8.20}$$

If $\zeta(s) \leq 0$ for $s \in [0,\infty)$ then $\xi \geq 0$ cannot be a solution of (8.20). $\xi \geq 0$ makes the left hand side of (8.20) positive but the right hand side of (8.20) is always nonpositive.

When $\omega \neq 0$, we decompose Eq. (8.19) into real and imaginary parts:

$$m\{\beta + \int_0^\infty \zeta(s)[e^{-\xi s}\cos\omega s - 1]ds\} = -(\xi^2 - \omega^2) \ ,$$

$$m\{\zeta(s)e^{-\xi s}\sin\omega s\} = 2\xi\omega \ . \tag{8.21}$$

If $\zeta(s) \leq 0$ for $s \in [0,\infty)$ and $|\zeta(s)| \to 0$ as $s \to \infty$ monotonically then $m\int_0^\infty \zeta(s)e^{-\xi s}\frac{\sin\omega s}{\omega}ds$ is negative for $\xi \geq 0$. This is so because $\zeta(s)e^{-\xi s}\frac{\sin\omega s}{\omega}$ is negative when s is small, it changes sign at each zero of $\sin\omega s$ and the contribution to the value of the integral on each interval is of decreasing magnitude. The negative contributions are therefore larger than the positive ones. This shows that Eq. $(8.21)_2$ is not satisfied for $\xi \geq 0$.

It is better for understanding to proceed less generally and to determine the sign of ξ for a relaxation modulus

$$\zeta(s) = -\mu e^{-\nu s}, \ \mu > 0, \ \nu > 0 \tag{8.22}$$

of the Maxwell type. We find that

(i) finite solution of (8.19) and (8.22) have

$$\nu + \xi > 0 \ , \tag{8.23}$$

(ii) finite solutions of (8.19) and (8.22) have at least one and at most three real values of $\sigma = \sigma_n$ for $m = 1/(\rho_o \lambda_n)$ and

$$\xi_n < 0. \tag{8.24}$$

We may restate these results as follows: Given $\beta > 0$, $\zeta(s) < 0 (\mu > 0)$, $\dot{\zeta}(s) > 0 (\nu > 0)$ there are a countably infinite number of finite solutions $\sigma_n = \xi_n + i\omega_n$ of (8.19), and $\xi_n \leq 0$ for all such solutions.

Proof: Substitution of (8.22) into (8.19) leads to

$$(8.25) \qquad m\{\beta - \mu\int_0^\infty [e^{-(\nu+\sigma)s} - e^{-\nu s}]ds\} = -\sigma^2 .$$

If $\nu + \xi \leq 0$ the integral diverges. To prevent this we admit as solutions only those σ for which (8.23) holds. Assuming now that $\nu + \xi > 0$ we evaluate (8.25) as

$$(8.26) \qquad -\sigma^2 = m\{\beta + \mu\sigma/\nu(\nu + \sigma)\}.$$

Eq. (8.26) is a cubic in σ which has to be solved subject to the constraint (8.23). When $\mu = 0$, $\sigma = \pm i\sqrt{m\beta}$. When μ is small, these two roots split into a conjugate pair with $\omega \neq 0$. The real and imaginary parts of (8.26) are

$$(8.27) \qquad \xi^2 - \omega^2 = -m\{\beta + \frac{\mu}{\nu}\frac{\xi^2+\xi\nu+\omega^2}{[(\nu+\xi)^2+\omega^2]}\},$$

and

$$(8.28) \qquad 2\xi\omega = -\mu m\omega/[(\nu+\xi)^2 + \omega^2] .$$

When $\omega \neq 0$, (8.28) shows that $\xi < 0$. When $\omega = 0$, (8.27) reduces to

$$(8.29) \qquad \xi^2 + m\beta = -m\mu\xi/\nu(\nu+\xi) .$$

Since $\nu + \xi > 0$, (8.29) shows that $\xi < 0$. When μ is small, there is only one real root of (8.29).

Now we will give a formal argument, based on Laplace Transforms, to show that the criteria $\xi < 0$ for all eigenvalues σ in the spectrum of (8.5) implies that $\underline{u} = 0$ is asympotically stable. We first rewrite the problem (8.1) as

$$\rho_0 \ddot{\underline{v}} - \gamma \nabla^2 \underline{v} - \nabla^2 \int_0^t \zeta(s)\underline{v}(\underline{X},t-s)ds - \nabla^2 \int_{-\infty}^0 \zeta(t-\tau)\underline{v}(\underline{X},\tau)d\tau$$

$$+ \nabla\pi = \underline{0} \quad ,$$

$$\text{(8.30)} \qquad \operatorname{div} \underline{v} = 0 \quad ,$$

$$\underline{v}(\underline{X},t)\Big|_{\partial\mathcal{V}_{10}} = \underline{0} \quad ,$$

$$\{-\pi\underline{\underline{1}} + 2\gamma\underline{\underline{E}}[\underline{v}(t)] + 2\int_0^t \zeta(s)\ \underline{\underline{E}}[\underline{v}(t-s)]\,ds$$

$$+ 2\int_{-\infty}^0 \zeta(t-\tau)\underline{\underline{E}}[\underline{v}(\tau)]d\tau\} \cdot \underline{N}\Big|_{\partial\mathcal{V}_{20}} = \underline{0} \quad .$$

Since $\underline{v}(\underline{X},\tau) = \underline{v}_0(\underline{X},\tau)$ for $\underline{X} \in \mathcal{V}_0, \tau \leq 0$ is known

$\nabla^2 \int_{-\infty}^0 \zeta(t-\tau)\underline{v}(\underline{X},\tau)d\tau$ and $2\int_{-\infty}^0 \zeta(t-\tau)\underline{\underline{E}}[\underline{v}(\tau)]d\tau$ are known.

Let

$$\text{(8.31)} \qquad \nabla^2 \int_{-\infty}^0 \zeta(t-\tau)\underline{v}(\underline{X},\tau)d\tau = \underline{f}_1(\underline{X},t) \quad ,$$

$$\{2\int_{-\infty}^0 \zeta(t-\tau)\underline{\underline{E}}[\underline{v}(\tau)]d\tau\} \cdot \underline{N}\Big|_{\partial\mathcal{V}_{20}} = -\underline{f}_2(\underline{X},t) \quad .$$

Now we have

$$\rho_0 \ddot{\underline{v}} - \gamma \nabla^2 \underline{v} - \nabla^2 \int_0^t \zeta(s)\underline{v}(\underline{X},t-s)ds + \nabla\pi = \underline{f}_1(\underline{X},t) \quad ,$$

$$\text{(8.32)} \qquad \operatorname{div} \underline{v} = 0 \quad ,$$

$$\underline{v}(\underline{X},t)\Big|_{\partial\mathcal{V}_{10}} = \underline{0} \quad ,$$

$$\{-\pi\underline{\underline{1}} + 2\gamma\underline{\underline{E}}[\underline{v}(t)] + 2\int_0^t \zeta(s)\underline{\underline{E}}[\underline{v}(t-s)]ds\} \cdot \underline{N}\Big|_{\partial\mathcal{V}_{20}}$$

$$= \underline{f}_2(\underline{X},t) \quad .$$

Now we assume that $\underline{v}$, $\ddot{\underline{v}}$, $\operatorname{div} \underline{v}$, $\underline{\underline{E}}[\underline{v}]$, $\nabla^2\underline{v}$, π, ζ, $\underline{f}_1$ and $\underline{f}_2$ possess Laplace transforms.

Let

$$(8.33)\qquad \begin{pmatrix} \underline{V}(\underline{X},\hat{\sigma}) \\ \hat{\pi}(\underline{X},\hat{\sigma}) \\ \hat{\zeta}(\hat{\sigma}) \\ \underline{F}_1(\underline{X},\hat{\sigma}) \\ \underline{F}_2(\underline{X},\hat{\sigma}) \end{pmatrix} = \int_0^\infty e^{-\hat{\sigma}t} \begin{pmatrix} \underline{v}(\underline{X},t) \\ \pi(\underline{X},t) \\ \zeta(t) \\ \underline{f}_1(X,t) \\ \underline{f}_2(\underline{X},t) \end{pmatrix}$$

Then we have

$$\rho_o \hat{\sigma}^2 \underline{V} - \kappa(\hat{\sigma})\nabla^2 \underline{V} + \nabla\hat{\pi} = \underline{F}_1(\underline{X},\hat{\sigma}) + [\hat{\sigma}\underline{v}(\underline{X},0) + \dot{\underline{v}}(\underline{X},0)]\rho_o ,$$

$$\operatorname{div} \underline{V} = 0 ,$$

$$(8.34)\qquad \underline{V}(\underline{X},\hat{\sigma})\Big|_{\partial\mathcal{V}_{10}} = 0 ,$$

$$\{-\hat{\pi}\,\underline{\underline{1}} + 2\kappa(\hat{\sigma})\underline{\underline{E}}[\underline{V}]\} \cdot \underline{N}\Big|_{\partial\mathcal{V}_{20}} = \underline{F}_2(\underline{X},\hat{\sigma}) .$$

In deriving (8.34), we have used the convolution property

$$\int_0^\infty \left[\int_0^t \zeta(s)\underline{v}(\underline{X},t-s)ds\right]e^{-\hat{\sigma}t}\,dt = \hat{\zeta}(\hat{\sigma})\underline{V}(\underline{X},\hat{\sigma}) ,$$

and

$$\int_0^\infty \ddot{\underline{v}}\, e^{-\hat{\sigma}t}\,dt = \hat{\sigma}^2 V(\underline{X},\hat{\sigma}) - \hat{\sigma}\,\underline{v}(\underline{X},0) - \dot{\underline{v}}(\underline{X},0) .$$

Equations (8.34) can be rewritten as

$$\underline{V} + \lambda\nabla^2\underline{V} + \nabla p = F_3(\underline{X},\hat{\sigma}) ,$$

$$(8.35)\qquad \operatorname{div} \underline{V} = 0 ,$$

$$\underline{V}(\underline{X},\hat{\sigma})\Big|_{\partial\mathcal{V}_{10}} = 0 ,$$

$$\{p\underline{\underline{1}} + 2\lambda\underline{\underline{E}}[\underline{V}]\} \cdot \underline{N}\Big|_{\partial\mathcal{V}_{20}} = \underline{F}_4(\underline{X},\hat{\sigma})$$

where

$$\lambda = \kappa(\hat{\sigma})/(-\rho_o\hat{\sigma}^2) ,$$

$$p = \hat{\pi}/\rho_o\hat{\sigma}^2 ,$$

$$\underline{F}_3(\underline{X},\hat{\sigma}) = [\underline{F}_1(\underline{X},\hat{\sigma})/\rho_o + \hat{\sigma}\underline{v}(\underline{X},0) + \dot{\underline{v}}(\underline{X},0)]/\hat{\sigma}^2 ,$$

$$\underline{F}_4(\underline{X},\hat{\sigma}) = \underline{F}_2(\underline{X},\hat{\sigma})/(-\rho_o\hat{\sigma}^2) .$$

The spectrum of the linear operator defined by (8.35) is the collection of complex values $\hat{\sigma} = \sigma_n$ for which (8.35) is not uniquely invertible with inverse depending continuously on $\underline{v}_0(\underline{X},\tau)$, $\tau \leq 0$ through $\underline{F}_3(\underline{X},\hat{\sigma})$ and $\underline{F}_4(\underline{X},\hat{\sigma})$. These are the eigenvalues σ_n which we have already characterized variationally through the functional equation (8.19). We learned that $Re\sigma_n = \xi_n < 0$ for all σ_n when $\beta > 0$, $\zeta(s) \leq 0$ and $|\zeta(s)| \to 0$ as $s \to \infty$ monotonically. It follows that Eqs. (8.33) hold for all $\hat{\xi}$ such that $\hat{\xi} > \xi_1$.

For the other values of $\hat{\sigma}$, not in the spectrum of (8.35), (8.35) is uniquely invertible and

$$\underline{V}(\underline{X},\hat{\sigma}) = \underline{\underline{R}}_{\sigma}^{\wedge}\,\underline{F}_3(\underline{X},\hat{\sigma}) + \underline{\underline{S}}_{\sigma}^{\wedge}\,\underline{F}_4(\underline{X},\hat{\sigma}) \tag{8.36}$$

depends continuously on $\underline{v}_0(\underline{X},\tau)$, $\tau \leq 0$ through, $\underline{F}_3(\underline{X},\hat{\sigma})$, $\underline{F}_4(\underline{X},\hat{\sigma})$ and matrix-valued resolvent operators $\underline{\underline{R}}_{\sigma}^{\wedge}$ and $\underline{\underline{S}}_{\sigma}^{\wedge}$. The values $\hat{\sigma}$ not in the spectrum are said to be in the resolvent set.

Given $\underline{V}(\underline{X},\hat{\sigma})$, we may use Laplace inversion integral to compute

$$\underline{v}(\underline{X},t) = \frac{1}{2\pi i}\int_{\hat{\xi}-i\infty}^{\hat{\xi}+i\infty} e^{\hat{\sigma}t}\,\underline{V}(\underline{X},\hat{\sigma})d\hat{\sigma} \tag{8.37}$$

$$= \frac{1}{2\pi i}\int_{\hat{\xi}-i\infty}^{\hat{\xi}+i\infty} e^{\hat{\sigma}t}[\underline{\underline{R}}_{\sigma}^{\wedge}\underline{F}_3(\underline{X},\hat{\sigma}) + \underline{\underline{S}}_{\sigma}^{\wedge}\underline{F}_4(\underline{X},\hat{\sigma})]d\hat{\sigma} .$$

Eq. (8.37) holds for any value of $\hat{\sigma} = \hat{\xi} + i\hat{\omega}$ for which Eqs. (8.33) hold; that is for $\hat{\xi} > \xi_1$. Since $\xi_1 < 0$, we may choose $\hat{\xi} < 0$. Then (8.37) shows that $\underline{v}(\underline{X},t)$ is asymptotically stable with exponential decay. Eigenvalues associated with (8.35) appear as singularities of the resolvent operators $\underline{\underline{R}}_{\hat{\sigma}}$ and $\underline{\underline{S}}_{\hat{\sigma}}$. If all the eigenvalues are simple then (8.37) may be evaluated by residues.

$$\underline{v}(\underline{X},t) = \sum_n e^{\hat{\sigma}_n t} a_n[\underline{v}_0(\underline{X},\tau)\Big|_{\tau=-\infty}^{0}, \hat{\sigma}_n] \tag{8.38}$$

where the coefficients a_n are functionals of the initial history $\underline{v}_0(\underline{X},\tau)$, $X \in \mathcal{V}_0$, $\tau \leq 0$.

9. Free surface problems perturbing the natural state

Many problems in elastostatics and viscoelastic dynamics can be solved using the equations derived in §7. From these, we have selected two problems in which the second order theory is required for the computation of the change in the shape of a stress-free surface due to nonlinear effects of inertia and stress. Free surface problems are of interest to material scientists because the distortion of the free surface due to deformation can be a sensitive mirror into the state of stress and the measurement of the distorted shape may provide a rheometrical device for measuring material constants and material functions. This hope we have for solids is a fact in fluids (see Joseph and Beavers, 1977).

The problems to be derived here involve distortion due to deformation in viscoelastic solids which are right circular cylinders in the natural state. The effects we compute may be regarded as analogous to Weissenberg effects in fluids and as in the fluids problem the Weissenberg effects appear first at second order. The second order problems require that we solve certain fourth order linear partial differential equations which in the simplest cases reduce to the biharmonic ones. These problems are probably best suited to analysis in biorthogonal series. Solving the problems is a major job which requires considerable analysis unrelated to the physical problem being studied here. So we defer the computation of solutions to a later work and concentrate on deriving the boundary-value problems which need solving under general circumstances.

The problems we treat are axisymmetric. Cylindrical coordinates are natural to such problems and it is necessary to compute the components of displacements

$$\underline{x} = \underline{X} + \underline{u}(\underline{X},t,\varepsilon) = r\underline{e}_r(\theta) + z\underline{e}_z, \tag{9.1}$$

where $\underline{e}_r(\theta)$ and $\underline{e}_z$ are a cylindrical basis in the coordinates (r,θ,z) of the distorted configuration, relative to a cylindrical basis $(\underline{e}_R, \underline{e}_\Theta, \underline{e}_Z)$, in the natural state

$$\underline{X} = \underline{e}_R(\Theta)R + \underline{e}_Z Z. \tag{9.2}$$

The components of $\underline{u}(\underline{X},t,\varepsilon)$ are independent of Θ:

$$\begin{aligned} r &= R + r^{<1>}(R,Z,t)\varepsilon + r^{<2>}(R,Z,t)\varepsilon^2 + O(\varepsilon^3), \\ \theta &= \Theta + \theta^{<1>}(R,Z,t)\varepsilon + \theta^{<2>}(R,Z,t)\varepsilon^2 + O(\varepsilon^3), \\ z &= Z + z^{<1>}(R,Z,t)\varepsilon + Z^{<2>}(R,Z,t)\varepsilon^2 + O(\varepsilon^3). \end{aligned} \tag{9.3}$$

To expand $\underline{u}(\underline{X},t,\varepsilon)$ it is necessary to compute the expansion of

$$\underline{e}_r(\theta) = \underline{e}_R + \varepsilon\,\underline{e}_\Theta\theta^{<1>} + \varepsilon^2\{\underline{e}_\Theta\theta^{<2>} - \underline{e}_R\theta^{<1>2}/2\} + O(\varepsilon^3) \tag{9.4}$$

induced by $(9.3)_2$. The term $-\varepsilon^2 e_R\theta^{<1>2}/2$ can be regarded as an effect of "inertia". Combining (9.1,2,3,4) we find that

$$\underline{u} = \underline{u}^{<1>} + \varepsilon^2\underline{u}^{<2>} + O(\varepsilon^3)$$

where

$$\underline{u}^{<1>} = r^{<1>}\underline{e}_R + R\theta^{<1>}\underline{e}_\Theta + z^{<1>}\underline{e}_Z \ , \tag{9.5}$$

and

$$\underline{u}^{<2>} = (r^{<2>} - R\theta^{<1>2}/2)\underline{e}_R + (R\theta^{<2>} + r^{<1>}\theta^{<1>})\underline{e}_\Theta + z^{<2>}\underline{e}_Z \ . \tag{9.6}$$

The components of $\underline{\underline{F}}^{<1>}$ in the basis $(\underline{e}_R, \underline{e}_\Theta, \underline{e}_Z)$ are

$$[\underline{\underline{F}}^{<1>}] = \begin{pmatrix} r_R^{<1>} & -\theta^{<1>} & r_Z^{<1>} \\ (R\theta^{<1>})_R & r^{<1>}/R & R\theta^{<1>}{}_Z \\ z_R^{<1>} & 0 & z_Z^{<1>} \end{pmatrix} . \tag{9.7}$$

Here and elsewhere $(\cdot)_R$ and $(\cdot)_Z$ denote partial derivatives of $(\cdot)$ with respect to R and Z respectively.

The computation of components in a cartesian basis is slightly less involved because "inertia" terms are absent.

To proceed further it is necessary to be more specific about the two problems under consideration.

(I) Distortion of the cylindrical free surface on a viscoelastic right circular cylinder induced by torsional oscillation of rigidly bonded end plates

A right circular cylinder of viscoelastic solid material of radius a is bonded to rigid parallel plates separated always by distance 2. The plates may rotate around the axis of the cylinder in a more or less arbitrary manner. The natural state is sketched in Fig. 9.1.a and the distorted shape is shown in Fig. 9.1.b.

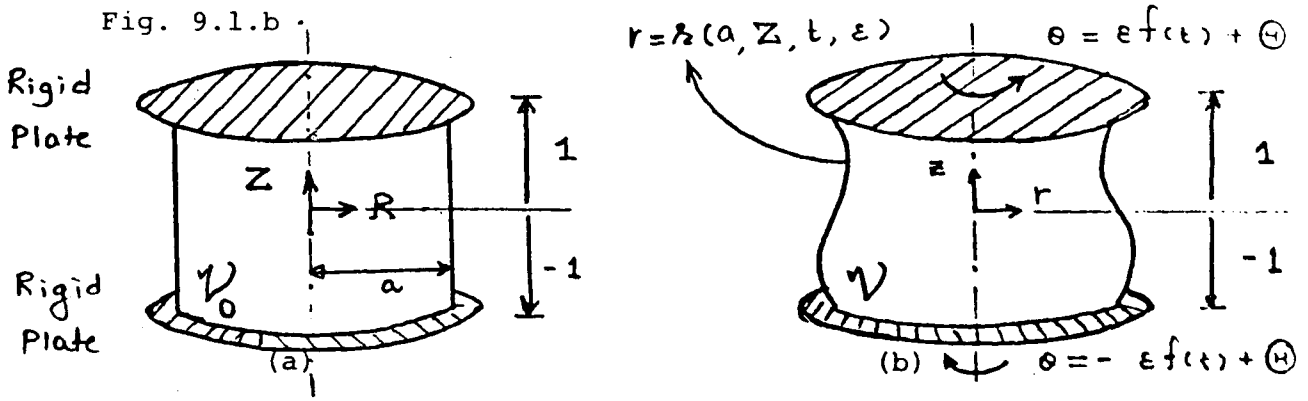


Fig. 9.1 A viscoelastic cylinder is sheared by the rotation of two end plates: (a) is the natural state

$$\mathcal{V}_0 = \{R,Z:\ 0 \le R < a,\ -1 < Z < 1\},$$

and (b) is the deformed state

$$\mathcal{V} = \{r,z:\ 0 \le r < \mathcal{R}(a,Z,t,\varepsilon),\ -1 < z < 1\}$$

where $\mathcal{R}(a,Z,t,\varepsilon) \overset{\text{def}}{=} r(R=a,Z,t,\varepsilon)$ is the radial displacement function given by $(9.3)_1$.

Conservation of volume in the deformed incompressible viscoelastic solid may be expressed through the requirement that

$$(9.8) \qquad \frac{1}{2}\int_{-1}^{1} \mathcal{R}^2(a,Z,t,\varepsilon)dZ = a \quad .$$

The displacement boundary condition are as follows:

$$(9.9) \qquad \theta(R, \pm 1,t,\varepsilon) = \pm\varepsilon f(t) + \Theta \ ,$$

$$r(R, \pm 1,t,\varepsilon) = R \ .$$

It follows then that

$$(9.10) \qquad r^{<n>}(R, \pm 1,t) = 0, \quad n = 1,2,3...,$$

$$\theta^{<1>}(R, \pm 1,t) = \pm f(t) \ ,$$

and

$$\theta^{<n>}(R, \pm 1,t) = 0, \quad n \geq 2.$$

The stress on $r = \mathcal{R}(a,Z,t,\varepsilon)$ must vanish.

If $f(t) = 1$, the cylinder undergoes a steady displacement and the problem falls in the class of universal deformations in nonlinear elastostatics found by Rivlin (1949). These deformations are independent of the constitutive equation provided that the material is undergoing elastic deformation; then, globally, without perturbations we get

$$(9.11) \qquad r = R, \ z = Z, \ \theta = \Theta + \varepsilon Z.$$

So the free surface will not change shape under a static twist. The distortion shown in 9.1.b is entirely dynamic in this (but not other) problems.

Now we shall solve the equations (7.10) governing the first order perturbation

$$p^{<1>} = r^{<1>} = z^{<1>} = 0,$$

(9.12)
$$\theta^{<1>}(R,Z,t) = \phi(Z,t) \text{ is independent of } R.$$

Then $(7.10)_1$ reduces to

$$\rho_o R\ddot{\phi} = \hat{L}\{\nabla^2(R\phi) - \phi/R\} = R\hat{L}(\partial^2\phi/\partial Z^2)$$

where

(9.13)
$$\hat{L}(\cdot) \stackrel{\text{def}}{=} \gamma(\cdot)(t) + \int_0^\infty \zeta(s)(\cdot)(t-s)ds \ ,$$

so that

$$\rho_o\ddot{\phi} = \partial^2\hat{L}(\phi)/\partial Z^2 \ ,$$

(9.14)
$$\phi(\pm 1,t) = +f(t) \ .$$

All other conditions on the first order perturbation are satisfied identically when (9.12) and (9.14) hold; in particular $\underline{\underline{T}}^{<1>} \cdot \underline{N} = \underline{\underline{T}}^{<1>} \cdot \underline{e}_R = \underline{0}$ on $R = a$ where

(9.15)
$$\underline{\underline{T}}^{<1>} = 2\hat{L}(\underline{\underline{E}}^{<1>}) = (\underline{e}_\Theta\underline{e}_Z + \underline{e}_Z\underline{e}_\Theta)\ R\hat{L}(\partial\phi/\partial Z).$$

To solve the second order perturbation problem we must compute the inhomogeneous terms in (7.12). All of these terms may be computed from

(9.16)
$$[\underline{\underline{F}}^{<1>}] = \begin{bmatrix} 0 & -\phi & 0 \\ \phi & 0 & R\partial\phi/\partial Z \\ 0 & 0 & 0 \end{bmatrix} \ ,$$

and

(9.17)
$$\underline{u}^{<2>} = (r^{<2>} - R\phi^2/2)\underline{e}_R + R\theta^{<2>}\underline{e}_\Theta + z^{<2>}\underline{e}_Z$$

using the expressions given in §5.

The equation $(7.12)_2$ expressing the conservation of volume simplifies to

(9.18)
$$\frac{1}{R}\frac{\partial(Rr^{<2>})}{\partial R} + \frac{\partial z^{<2>}}{\partial Z} = 0 \ ,$$

because the term proportional to ϕ^2 in (9.17) cancels θ_2. To compute $\underline{M}_2$ in (7.13) we need only to note that $p^{<1>} = 0$ and the rest follows from simple operations with (9.16). We find that the components (M_{2R}, $M_{2\theta}$ and M_{2Z}) of $\underline{M}_2$ are given by

$$
\begin{aligned}
(9.19)\quad M_{2R} &= -\beta^{[2]} R\phi'^2(t) \\
&\quad + R\int_0^\infty \zeta(s)\,\{[\phi'(t-s)[\![\phi]\!]]' + \phi'(t)[\![\phi']\!]\}ds \\
&\quad - 2R\int_0^\infty \zeta(s)^{[2]}\phi'(t)[\![\phi']\!]ds \\
&\quad - R\int_0^\infty\int_0^\infty \alpha(s_1,s_2)[\phi'(t-s_1) - \phi'(t)] \\
&\qquad\qquad [\phi'(t-s_2) - \phi'(t)]ds_1ds_2 \ , \\
M_{2\theta} &= 0, \\
M_{2Z} &= \beta^2R^2[\phi'^2(t)]' + R^2\int_0^\infty \zeta(s)\ ([\![\phi']\!]^2)'ds \\
&\quad + 2R^2\int_0^\infty \zeta^{[2]}(s)\ [\phi'(t)[\![\phi']\!]]'ds \\
&\quad + R^2\int_0^\infty\int_0^\infty \alpha(s_1,s_2)\{[\phi'(t-s_1) - \phi'(t)] \\
&\qquad\qquad [\phi'(t-s_2) - \phi'(t)]\}'ds_1ds_2
\end{aligned}
$$

where prime denotes partial derivative w. r. t. Z. The R, θ and Z components of $(7.12)_1$ may be written as

$$
(9.20)\quad \rho_o\ddot{r}^{<2>} - \rho_o R\,\overline{\ddot{\phi}^2}/2 = -\partial\pi^{<2>}/\partial R + \hat{L}(\nabla^2 r^{<2>} - r^{<2>}/R^2) - \hat{L}(R[\phi\phi']') + M_{2R} \ ,
$$

$$
(9.21)\quad \rho_o R\ddot{\theta}^{<2>} = \hat{L}[\nabla^2(R\theta^{<2>}) - \theta^{<2>}/R] \ ,
$$

and

$$
(9.22)\quad \rho_o\ddot{z}^{<2>} = -\partial\pi^{<2>}/\partial Z + \hat{L}(\nabla^2 z^{<2>}) + M_{2Z}.
$$

The displacement boundary conditions are

(9.23) $\quad r^{<2>}(R, \pm 1, t) = \theta^{<2>}(R, \pm 1, t) = z^{<2>}(R, \pm 1, t) = 0.$

To form the stress boundary conditions $(7.12)_4$ we note that

(9.24) $\quad \underline{N} = \underline{e}_R,\ \underline{\underline{T}}^{<1>} \cdot \underline{N} = \underline{0},\ \underline{\underline{T}}^{<1>}\underline{\underline{F}}^{T<1>} \cdot \underline{N} = -R\phi\hat{L}(\phi')\underline{e}_Z \ .$

$\underline{\underline{T}}^{<2>}$ is given by (7.8). To compute $\underline{\underline{T}}^{<2>} \cdot \underline{e}_R$ we note that $\underline{\underline{E}}^{<2>} = \frac{1}{2}[\nabla \underline{u}^{<2>} + (\nabla \underline{u}^{<2>})^T]$ and using (9.17) find

$$[2\underline{\underline{E}}^{<2>}] = \begin{pmatrix} 2r_R^{<2>} - \phi^2 & R\theta_R^{<2>} & r_Z^{<2>} + z_R^{<2>} - R\phi' \\ R\theta_R^{<2>} & 2r^{<2>}/R - \phi^2 & R\theta_Z^{<2>} \\ r_Z^{<2>} + z_R^{<2>} - R\phi\phi' & R\theta_Z^{<2>} & 2z_Z^{<2>} \end{pmatrix},$$

so that

(9.25) $\quad 2\underline{\underline{E}}^{<2>} \cdot \underline{e}_R = 2r_R^{<2>}\underline{e}_R + R\theta_R^{<2>}\ \underline{e}_\Theta + (r_Z^{<2>} + z_R^{<2>})\underline{e}_Z$

$$- (\phi^2 \underline{e}_R + R\phi\phi'\ \underline{e}_Z) \ .$$

The remaining terms of $\underline{\underline{T}}^{<2>} \cdot \underline{e}_R$ are formed from simple manipulations using (9.16) .

(9.26) $\quad \underline{\underline{T}}^{<2>} \cdot \underline{e}_R = \underline{e}_R \{-\pi^{<2>} + \beta\phi^2 + \int_0^\infty \zeta(s) [\![\phi^2]\!] ds\}$

$$+ \underline{e}_Z \{R \int_0^\infty \zeta(s)\phi'(t-s)[\![\phi]\!] ds\} + \hat{L}(2\underline{\underline{E}}^{<2>} \cdot \underline{e}_R)$$

$$= \underline{e}_R[-\pi^{<2>} + \hat{L}(2r_R^{<2>})] + \underline{e}_\Theta\ R\hat{L}(\theta_R^{<2>})$$

$$+ \underline{e}_Z\{\hat{L}(r_Z^{<2>} + z_R^{<2>}) - R\phi\hat{L}(\phi')\} \ .$$

Combining (9.24) and (9.26), we get

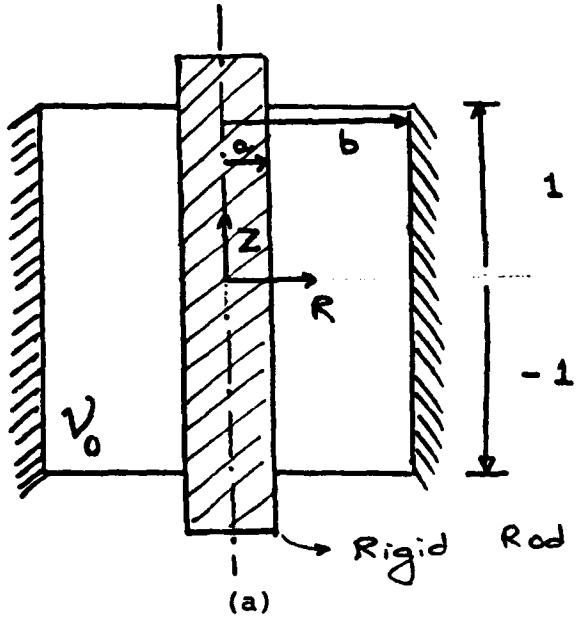
(9.27) $\quad \underline{t}_n^{<2>} = \underline{e}_R[-\pi^{<2>} + \hat{L}(2r_R^{<2>})] + \underline{e}_\Theta\ R\hat{L}(\theta_R^{<2>})$

$$+ \underline{e}_Z\ \hat{L}(r_Z^{<2>} + z_R^{<2>}) = \underline{0} \quad \text{at } R = a.$$

We note that the equation (9.21) governing $\theta^{<2>}$ is homogeneous ($M_{2\theta} = 0$) with homogeneous boundary conditions [equations(9.23) and (9.27)]. Hence $\theta^{<2>} \equiv 0$. To find $r^{<2>}$ and $z^{<2>}$, we must solve (9.18), (9.20), (9.22) subject to the boundary conditions (9.23) and (9.27).

(II) <u>Distortion of the plane surfaces perpendicular to the axis of viscoelastic cylindrical annulus rigidly bonded at the inner and outer radii and undergoing torsional oscillations at the inner radius</u>

A right circular cylindrical annulus of viscoelastic material of inner radius a and outer radius b is rigidly bonded at both the radii. Its initial length is 2. Both the radii remain fixed during the motion. The rigid rod to which the annulus is bonded at the inner radius rotates around the axis of the annulus in a more or less arbitrary manner. The natural state is sketched in Fig. 9.2.a and the distorted shape is shown in Fig. 9.2.b.



(a)

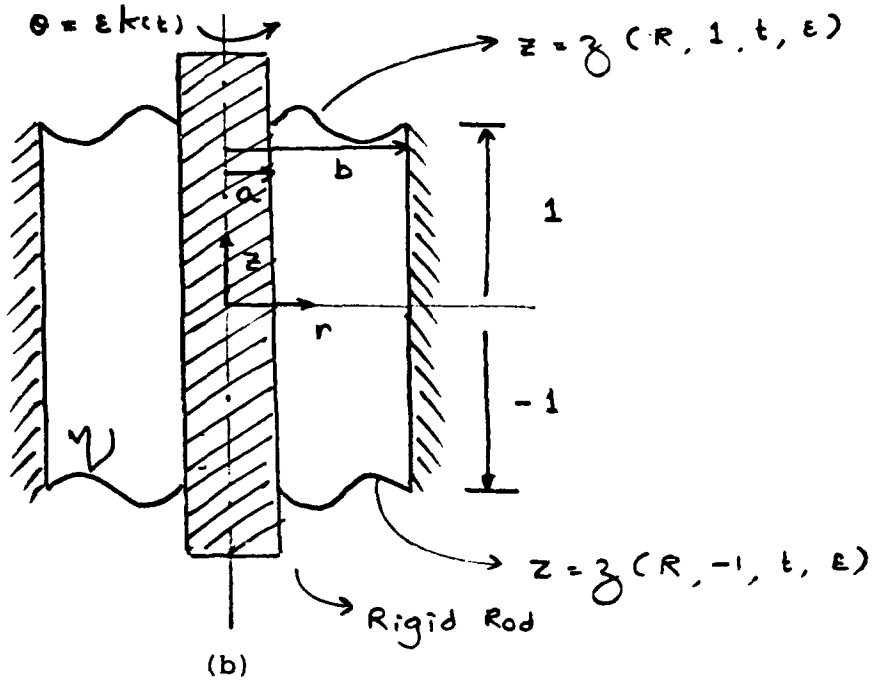


Fig. 9.2 A viscoelastic cylindrical annulus is sheared by the rotation of the rigid rod, bonded to the annulus at the inner radius: (a) is the natural state

$$\mathcal{V}_0 = \{R,Z:\ a < R < b,\ -1 < Z < 1\}\ ,$$

and (b) is the deformed state

$$\mathcal{V} = \{r,z:\ a < r < b,\ \zeta(R,-1,t,\varepsilon) < z < \zeta(R,1,t,\varepsilon)\}$$

where $\zeta(R,\pm1,t,\varepsilon) \overset{\text{def}}{=} z(R,Z = \pm1,\ t,\varepsilon)$ is the axial displacement function given by $(9.3)_3$.

Conservation of volume in the deformed incompressible viscoelastic solid may be expressed through the requirement that

$$(9.28)\qquad \int_a^b [\zeta(R,1,t,\varepsilon) - \zeta(R,-1,t,\varepsilon)]\ R\,dR = (b^2-a^2)\ .$$

The displacement boundary conditions are as follows:

$$(9.29)\qquad \theta(a,Z,t,\varepsilon) = \varepsilon k(t) + \Theta\ ,$$

$$\theta(b,Z,t,\varepsilon) = \Theta\ ,$$

$$z(R,Z,t,\varepsilon) = Z \quad \text{at } R = a \ \&\ b\ .$$

It follows that

(9.30) $\vartheta^{<1>}(a,Z,t) = k(t)$,

$\theta^{<n>}(a,Z,t) = 0$ for $n \geq 2$,

$\theta^{<n>}(b,Z,t) = 0$ for $n = 1,2,3...$,

$z^{<n>}(R,Z,t) = 0$ at $R = a$ & b for $n = 1,2,3...$.

The stress on $z = (R, \pm 1, t, \varepsilon)$ must vanish.

Now we shall solve the equations (7.10) governing the first order perturbation .

(9.31) $p^{<1>} = r^{<1>} = z^{<1>} = 0$,

$\theta^{<1>}(R,Z,t) = \phi(R,t)$ is independent of Z.

Then (7.10) reduces to

(9.32) $\rho_o R\ddot{\phi} = \hat{L}[\nabla^2(R\phi) - \phi/R)] = \hat{L}[\frac{\partial^2(R\phi)}{\partial R^2} + \frac{1}{R}\frac{\partial(R\phi)}{\partial R} - \frac{\phi}{R}]$,

$\phi(a,t) = k(t)$,

$\phi(b,t) = 0$

where $\hat{L}(\cdot)$ is as defined by (9.13). All other conditions on the first order perturbation are satisfied identically when (9.31) and (9.32) hold; in particular $\underline{\underline{T}}^{<1>} \cdot \underline{N} = \underline{\underline{T}}^{<1>} \cdot \underline{e}_Z = \underline{0}$ on $Z = +1$ (and $\underline{\underline{T}}^{<1>} \cdot \underline{N} = -\underline{\underline{T}}^{<1>} \cdot \underline{e}_Z = \underline{0}$ on $z = -1$) where

(9.33) $\underline{\underline{T}}^{<1>} = 2\hat{L}(\underline{\underline{E}}^{<1>}) = (\underline{e}_R\underline{e}_\Theta + \underline{e}_\Theta\underline{e}_R)\,\hat{L}(R\partial\phi/\partial R)$.

To solve the second order perturbation problem we must compute the nonhomogeneous terms in (7.12). All of these may be computed from

(9.34) $[\underline{\underline{F}}^{<1>}] = \begin{bmatrix} 0 & -\phi & 0 \\ \phi + R\partial\phi/\partial R & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

and

$$(9.35)\qquad \underline{u}^{<2>} = (r^{<2>} - R\phi^2/2)\underline{e}_R + R\theta^{<2>}\underline{e}_\Theta + z^{<2>}\,\underline{e}_Z$$

using the expressions given in §5.

The equation $(7.12)_2$ expressing the conservation of volume simplifies to

$$(9.36)\qquad \frac{1}{R}\frac{\partial(Rr^{<2>})}{\partial R} + \frac{\partial z^{<2>}}{\partial Z} = 0$$

because the term proportional to ϕ^2 in (9.35) cancels θ_2. To compute $\underline{M}_2$ in (7.13), we need only to note that $p^{<1>} = 0$ and the rest follows from simple operations with (9.34). We find that the components $(M_{2R}, M_{2\Theta}, M_{2Z})$ of $\underline{M}_2$ are given by

$$\begin{aligned}(9.38)\qquad M_{2R} = \frac{d}{dR}\{&\beta^{[2]}c(t)^2 + \int_0^\infty \zeta(s)[c(t-s)\,[\![c+a]\!] - c(t)\,[\![c]\!]]ds \\ &+ 2\int_0^\infty \zeta^{[2]}(s)[(c(t)\,[\![c]\!]]\,ds \\ &+ \int_0^\infty\int_0^\infty \alpha(s_1,s_2)[c(t-s_1)-c(t)] \\ &[c(t-s_2)-c(t)]ds_1ds_2\} \\ &+ \frac{1}{R}\{\int_0^\infty \zeta(s)c(t-s)([\![c+a]\!] + [\![a]\!])ds\}\ ,\end{aligned}$$

$$M_{2\Theta} = M_{2Z} = 0$$

where

$$(9.39)\qquad c(t) = R\,\frac{\partial\phi(R,t)}{\partial R}\ ,$$

$$a(t) = \phi(R,t)\ .$$

The R, Θ and Z components of $(7.12)_1$ may be written as

$$(9.40)\quad \rho_o \ddot{r}^{<2>} - \rho_o R\,\overline{\ddot{\phi}^2}/2 = -\partial\pi^{<2>}/\partial R + \hat{L}(\nabla^2 r^{<2>} - r^{<2>}/R^2) + \hat{L}[R\phi\partial^2\phi/\partial R^2 + R(\partial\phi/\partial R)^2 + 3\phi\partial\phi/\partial R] + M_{2R} ,$$

$$(9.41)\quad \rho_o R\ddot{\theta}^{<2>} = \hat{L}[\nabla^2(R\theta^{<2>}) - \theta^{<2>}/R] ,$$

and

$$(9.42)\quad \rho_o \ddot{z}^{<2>} = -\partial\pi^{<2>}/\partial Z + \hat{L}(\nabla^2 z^{<2>}) .$$

The displacement boundary conditions are

$$(9.43)\quad r^{<2>}(R,Z,t) = \theta^{<2>}(R,Z,t) = z^{<2>}(R,Z,t) = 0 \text{ at } R = a \text{ \& } b.$$

To form the stress boundary conditions $(7.12)_4$, we note that $\underline{N} = \pm\,\underline{e}_Z$ (at $Z = \pm 1$), $\underline{\underline{T}}^{<1>}\cdot\underline{N} = \underline{0}$, $\underline{\underline{F}}^{<1>T}\cdot\underline{N} = \underline{0}$. Hence

$$(9.44)\quad \underline{\underline{T}}^{<1>}\cdot\underline{n}^{<1>} = \underline{0}.$$

To compute $\underline{\underline{T}}^{<2>}\cdot\underline{e}_Z$, we note that $\underline{\underline{E}}^{<2>} = \frac{1}{2}[\nabla\underline{u}^{<2>} + (\nabla\underline{u}^{<2>})^T]$ and using (9.34) find

$$(9.45)\quad [2\underline{\underline{E}}^{<2>}] = \begin{pmatrix} 2r_R^{<2>} - \phi^2 - 2R\phi\phi_R & R\theta_R^{<2>} & r_Z^{<2>} + z_R^{<2>} \\ R\theta_R^{<2>} & 2r^{<2>}/R - \phi^2 & R\theta_Z^{<2>} \\ r_Z^{<2>} + z_R^{<2>} & R\theta_Z^{<2>} & 2z_Z^{<2>} \end{pmatrix} ,$$

so that

$$2\underline{\underline{E}}^{<2>}\cdot\underline{e}_Z = (r_Z^{<2>} + z_R^{<2>})\underline{e}_R + R\theta_Z^{<2>}\underline{e}_\Theta + 2z_Z^{<2>}\underline{e}_Z .$$

The remaining terms of $\underline{\underline{T}}^{<2>}\cdot\underline{e}_Z$ are formed form simple manipulations using (9.16).

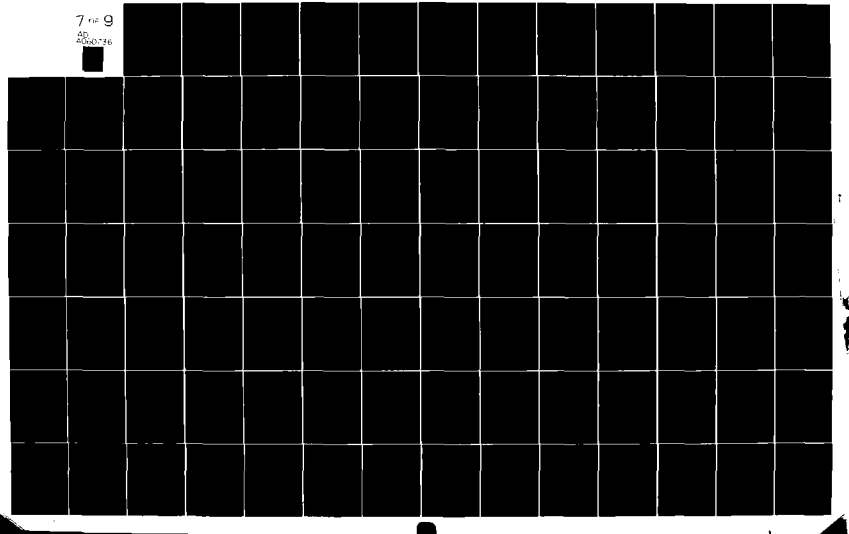
$$(9.46)\quad \underline{\underline{T}}^{<2>}\cdot\underline{e}_Z = -\pi^{<2>}\underline{e}_Z + \hat{L}(2\underline{\underline{E}}^{<2>}\cdot\underline{e}_Z)$$
$$= \underline{e}_R\,\hat{L}(r_Z^{<2>} + z_R^{<2>}) + \underline{e}_\Theta\,R\,\hat{L}(\theta_Z^{<2>})$$
$$+ \underline{e}_Z\,[-\pi^{<2>} + \hat{L}(2z_Z^{<2>})] .$$

combining (9.44) and (9.46), we get

$$(9.47)\quad \underline{t}_n^{<2>} = \underline{e}_R \hat{L}(r_Z^{<2>} + z_R^{<2>}) + \underline{e}_\Theta R\hat{L}(\theta_Z^{<2>}) + \underline{e}_Z [-\pi^{<2>} + \hat{L}(2z_Z^{<2>})] = \underline{0} \quad \text{at } Z = \pm 1.$$

We note that the equation (9.41) governing $\theta^{<2>}$ is homogeneous ($M_{2\Theta} = 0$) with homogeneous boundary conditions [equations (9.43) and (9.47)]. Hence $\theta^{<2>} \equiv 0$. To find $r^{<2>}$ and $z^{<2>}$, we must solve (9.36), (9.40), (9.42) subject to the boundary conditions (9.43) and (9.47).

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL



10. Linearized theory of perturbation of the rest state

In this section we derive the first order equations of motion for incompressible solids for motions perturbing the rest state. Since the rest state contains all static deformations and, in fact, coincides with the set of all elastic deformations, the linearized equations derived here form the basis for the discussion of stability and bifurcation of elastostatic deformations of solids. The trouble we find when carrying out a correct analysis of the linearized theory is that so many (2) material constants and (7) functions are needed to characterize the material.

For incompressible materials $\det \underline{\underline{F}} = 1$. In the rest state

$$\det \underline{\underline{F}}^{<0>} = 1, \tag{10.1}$$

where

$$\underline{\underline{F}}(t) = \nabla \underline{x}(\underline{X},t), \quad \underline{\underline{F}}^{<0>} = \nabla \underline{u}^{\circ}(\underline{X}) + \underline{\underline{1}},$$

$$\underline{x} = \underline{X} + \underline{u}^{\circ}(\underline{X}) + \varepsilon \underline{u}^{<1>}(\underline{X},t) + O(\varepsilon^2). \tag{10.2}$$

In §5, we derived the perturbation formulas for the kinematic variables for perturbations of the natural state. But these are valid only when $\underline{\underline{F}}^{<0>} = \underline{\underline{1}}$. In the case of perturbations of the rest state, the derivation of the perturbation formulas for the kinematic variables is similar to the one in §5. Here we list only the results.

$$\underline{\underline{F}}(\underline{X},\tau,\varepsilon) = \underline{\underline{F}}^{<0>}(\underline{X}) + \varepsilon \underline{\underline{F}}^{<1>}(\underline{X},t) + O(\varepsilon^2). \tag{10.3}$$

$$\underline{\underline{F}}^{-1}(\tau,\varepsilon) = (\underline{\underline{F}}^{<0>})^{-1} - \varepsilon (\underline{\underline{F}}^{<0>})^{-1} (\underline{\underline{F}}^{<1>}(\tau))^{-1} (\underline{\underline{F}}^{<0>})^{-1} + O(\varepsilon^2). \tag{10.4}$$

$$\underline{\underline{G}}(s,\varepsilon) = \varepsilon \underline{\underline{G}}^{<1>}(s) + O(\varepsilon^3), \tag{10.5}$$

where

$$\underline{\underline{G}}^{<1>}(s) = \underline{\underline{F}}^{<1>}(\tau)(\underline{\underline{F}}^{<0>})^{-1} + (\underline{\underline{F}}^{<0>^T})^{-1}\,\underline{\underline{F}}^{<1>^T}(\tau) - \underline{\underline{F}}^{<1>}(t)(\underline{\underline{F}}^{<0>})^{-1} - (\underline{\underline{F}}^{<0>^T})^{-1}\,\underline{\underline{F}}^{<1>^T}(t). \tag{10.6}$$

$$\underline{\underline{B}}(\tau,\varepsilon) = \underline{\underline{F}}^{<0>}\underline{\underline{F}}^{<0>^T} + \varepsilon(\underline{\underline{F}}^{<1>}\underline{\underline{F}}^{<0>^T} + \underline{\underline{F}}^{<0>}\underline{\underline{F}}^{<1>^T}) + O(\varepsilon^2). \tag{10.7}$$

$$\det \underline{\underline{F}}(\tau,\varepsilon) = 1 + \varepsilon\{\mathrm{tr}[(\underline{\underline{F}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>}(\tau)]\} + O(\varepsilon^2) \tag{10.8}$$

where we have used the equation (10.1). We do not need the perturbation formula for ρ. To find the perturbation formula for $\underline{n}$, we use

$$da^2 = (\det \underline{\underline{F}})^2(\underline{N}\cdot\underline{\underline{C}}^{-1}\cdot\underline{N})dA^2. \tag{10.9}$$

where $\underline{\underline{C}} = \underline{\underline{F}}^T\underline{\underline{F}}$ is the right Cauchy-Green strain tensor. This formula reduces to

$$da^2 = (\underline{N}\cdot\underline{\underline{C}}^{-1}\cdot\underline{N})dA^2 \tag{10.10}$$

because of (10.1). Expanding both sides of (10.10) into powers of ε and identifying independent powers of ε, we get

$$da^{<0>} = [\,\underline{N}\cdot(\underline{\underline{C}}^{<0>})^{-1}\cdot\underline{N}]^{1/2}dA, \tag{10.11}$$

$$da^{<1>} = -\underline{N}\cdot[(\underline{\underline{F}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>}(\underline{\underline{C}}^{<0>})^{-1} + (\underline{\underline{C}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>^T}(\underline{\underline{F}}^{<0>^T})^{-1}] \cdot \underline{N}\,dA/2[\underline{N}\cdot(\underline{\underline{C}}^{<0>})^{-1}\cdot\underline{N}]^{1/2}. \tag{10.12}$$

Next we use

$$\underline{n}\,da = (\det \underline{\underline{F}})(\underline{\underline{F}}^{-1})^T\cdot\underline{N}\,dA\ , \tag{10.13}$$

and (10.1) to derive

$$\underline{n}^{<0>} = (\underline{\underline{F}}^{<0>^T})^{-1}\cdot\underline{N}/[\underline{N}\cdot(\underline{\underline{C}}^{<0>})^{-1}\cdot\underline{N}]^{1/2}, \tag{10.14}$$

$$(10.15)\qquad \underline{n}^{<1>} = \{\underline{N}\cdot[(\underline{\underline{F}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>}(\underline{\underline{C}}^{<0>})^{-1} + (\underline{\underline{C}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>T}(\underline{\underline{F}}^{<0>T})^{-1}]\cdot\underline{N}$$

$$(\underline{\underline{F}}^{<0>T})^{-1}\cdot\underline{N}/2[\underline{N}\cdot(\underline{\underline{C}}^{<0>})^{-1}\cdot\underline{N}]$$

$$-(\underline{\underline{F}}^{<0>T})^{-1}\,\underline{\underline{F}}^{<1>T}(\underline{\underline{F}}^{<0>T})^{-1}\cdot\underline{N}\}/[\underline{N}\cdot(\underline{\underline{C}}^{<0>})^{-1}\cdot\underline{N}]^{1/2}.$$

Finally

$$(10.16)\qquad \underline{t}_n = \underline{t}_n^{<0>} + \varepsilon\underline{t}_n^{<1>} + O(\varepsilon^2)$$

where

$$(10.17)\qquad \underline{t}_n^{<0>} = \underline{\underline{T}}^{<0>}\cdot\underline{n}^{<0>}\ ,$$

and

$$(10.18)\qquad \underline{t}_n^{<1>} = \underline{\underline{T}}^{<1>}\cdot\underline{n}^{<0>} + \underline{\underline{T}}^{<0>}\cdot\underline{n}^{<1>}.$$

We note that when $\underline{\underline{F}}^{<0>} = \underline{\underline{1}}$, all of these formulas reduce to those in §5.

To find the canonical forms for the perturbation stresses, we need to expand equations (3.3) and (3.4) into powers of ε. So we need perturbation formulas for f_i, $i = 0, 1, 2$ and ϕ_{ij}, $i = 0, 1, 2, 3$, $j = 0, 1, 2$.

$$(10.19)\qquad f_i = f_i^{<0>} + \varepsilon f_i^{<1>} + O(\varepsilon^2)$$

where

$$(10.20)\qquad f_i^{<0>} = f_i(\mathrm{tr}\,\underline{\underline{B}}^{<0>},\ \mathrm{tr}\,\underline{\underline{B}}^{<0>2})$$

and

$$(10.21)\qquad f_i^{<1>} = (\partial f_i/\partial I_B\big|_{\varepsilon=0})\mathrm{tr}\,\underline{\underline{B}}^{<1>} + 2(\partial f_i/\partial II_B\big|_{\varepsilon=0})\mathrm{tr}(\underline{\underline{B}}^{<0>}\underline{\underline{B}}^{<1>}).$$

$$(10.22)\qquad \phi_{ij} = \phi_{ij}^{<0>} + \varepsilon\phi_{ij}^{<1>} + O(\varepsilon^2)$$

where

$$\phi_{ij}^{<0>} = \phi_{ij}(\text{tr}\,\underline{\underline{B}}^{<0>}, \text{tr}\,\underline{\underline{B}}^{<0>2}, s) \,, \tag{10.23}$$

and

$$\phi_{ij}^{<1>} = (\partial\phi_{ij}/\partial I_B \Big|_{\varepsilon=0}) \text{tr}\underline{\underline{B}}^{<1>} + 2(\partial\phi_{ij}/\partial II_B\Big|_{\varepsilon=0}) \text{tr}(\underline{\underline{B}}^{<0>}\underline{\underline{B}}^{<1>}) \,. \tag{10.24}$$

As stated earlier, we may group all terms of $\underline{\underline{T}}$ proportional to $\underline{\underline{1}}$ with -p. Another simplification comes from combining equations (10.1) and (10.8). Then

$$\text{tr}[(\underline{\underline{F}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>}(\tau)] = 0. \tag{10.25}$$

Hence

$$\text{tr}[\underline{\underline{G}}^{<1>}(s)] = 0. \tag{10.26}$$

But $\text{tr}\,\underline{\underline{B}}^{<1>} \neq 0$.

The perturbed stresses are given by

$$\underline{\underline{T}}^{<0>} = -p^{<0>}\,\underline{\underline{1}} + f_1^{<0>}\underline{\underline{B}}^{<0>} + f_2^{<0>}\underline{\underline{B}}^{<0>2} \,, \tag{10.27}$$

and

$$\begin{aligned}\underline{\underline{T}}^{<1>} = {} & -p^{<1>}\underline{\underline{1}} + (f_1^{<1>}\underline{\underline{B}}^{<0>} + f_1^{<0>}\underline{\underline{B}}^{<1>}) \\ & + (f_2^{<1>}\underline{\underline{B}}^{<0>2} + f_2^{<0>}\underline{\underline{B}}^{<1>}\underline{\underline{B}}^{<0>} + f_2^{<0>}\underline{\underline{B}}^{<0>}\underline{\underline{B}}^{<1>}) \\ & + \int_0^\infty \{\text{tr}[\phi_{11}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{12}^{<0>}\underline{\underline{B}}^{<0>2})\underline{\underline{G}}^{<1>}(s)]\underline{\underline{B}}^{<0>} \\ & + \text{tr}[\phi_{21}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{22}^{<0>}\underline{\underline{B}}^{<0>2})\underline{\underline{G}}^{<1>}(s)]\underline{\underline{B}}^{<0>2} \\ & + (\phi_{30}^{<0>}\underline{\underline{1}} + \phi_{31}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{32}^{<0>}\underline{\underline{B}}^{<0>2})\underline{\underline{G}}^{<1>}(s) \\ & + \underline{\underline{G}}^{<1>}(s)(\phi_{30}^{<0>}\underline{\underline{1}} + \phi_{31}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{32}^{<0>}\underline{\underline{B}}^{<0>2})\}ds^* \,.\end{aligned} \tag{10.28}$$

*We note that when $\underline{\underline{F}}^{<0>} = \underline{\underline{B}}^{<0>} = \underline{\underline{1}}$, equation (10.25) reduces to (7.1). Hence $\text{tr}\underline{\underline{B}}^{<1>} = 0$. It implies $f_i^{<1>} = 0$ for i = 1,2.

(con't)

Equation (10.28) shows that for the linearized theory of perturbation of the rest state of incompressible viscoelastic solids, we need 2 elastic constants: f_1 & f_2 and 7 viscoelastic material functions: $\phi_{ij}(s)$, $i = 1,2,3$, $j = 1,2$ and $\phi_{30}(s)$*.

To find the perturbed form of Piola-Kirchoff stress, we expand $\underline{\underline{S}}^T = \underline{\underline{T}}(\underline{\underline{F}}^T)^{-1}$ in powers of ε.

$$\underline{\underline{S}}^T = \underline{\underline{S}}^{<0>^T} + \varepsilon\underline{\underline{S}}^{<1>^T} + O(\varepsilon^2) \tag{10.29}$$

where

$$\underline{\underline{S}}^{<0>^T} = \underline{\underline{T}}^{<0>}(\underline{\underline{F}}^{<0>^T})^{-1}, \tag{10.30}$$

and

$$\underline{\underline{S}}^{<1>^T} = \underline{\underline{T}}^{<1>}(\underline{\underline{F}}^{<0>^T})^{-1} - \underline{\underline{T}}^{<0>}(\underline{\underline{F}}^{<0>^T})^{-1}\,\underline{\underline{F}}^{<1>}(\underline{\underline{F}}^{<0>^T})^{-1}. \tag{10.31}$$

Substitution of (10.27) and (10.28) into (10.31) yields:

---

footnote con't

It reduces $\underline{\underline{T}}^{<1>}$ to

$$\underline{\underline{T}}^{<1>} = -p^{<1>}\underline{\underline{1}} + (f_1^{<0>} + 2f_2^{<0>})\underline{\underline{B}}^{<1>} + \int_0^\infty 2(\phi_{30}^{<0>} + \phi_{31}^{<0>} + \phi_{32}^{<0>})\underline{\underline{G}}^{<1>}(s)\,ds .$$

Here we have used equation (10.26). This form of $\underline{\underline{T}}^{<1>}$ is same as that given by equation (7.6) where

$$\beta = f_1^{<0>}(3,3) + 2f_2^{<0>}(3,3),$$

and

$$\zeta(s) = 2[\phi_{30}^{<0>}(3,3,s) + \phi_{31}^{<0>}(3,3,s) + \phi_{32}^{<0>}(3,3,s)].$$

Here we have used the fact that when $\underline{\underline{B}}^{<0>} = 1$, $\mathrm{tr}\underline{\underline{B}}^{<0>} = \mathrm{tr}\underline{\underline{B}}^{<0>^2} = 3$.

---

* Exactly the same number of material parameters arise in Pipkin's (1964) viscoelastic perturbation of elastostatic deformations. Elastic constants appear when the stress is evaluated on deformations which are independent of time.

$$
\begin{aligned}
(10.32)\qquad \underline{\underline{S}}^{<1>^T} &= [p^{<0>}(\underline{\underline{F}}^{<0>^T})^{-1}\,\underline{\underline{F}}^{<1>^T}(\underline{\underline{F}}^{<0>^T})^{-1} - p^{<1>}(\underline{\underline{F}}^{<0>^T})^{-1}] \\
&+ f_1^{<1>}\underline{\underline{F}}^{<0>} + f_1^{<0>}\underline{\underline{F}}^{<1>} + f_2^{<1>}\underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<0>} \\
&\qquad + f_2^{<0>}(\underline{\underline{B}}^{<1>}\underline{\underline{F}}^{<0>} + \underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<1>}) \\
&+ \int_0^\infty \{\mathrm{tr}[\phi_{11}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{12}^{<0>}\underline{\underline{B}}^{<0>^2})\underline{\underline{G}}^{<1>}(s)\,]\underline{\underline{F}}^{<0>} \\
&+ \mathrm{tr}[(\phi_{21}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{22}^{<0>}\underline{\underline{B}}^{<0>^2})\underline{\underline{G}}^{<1>}(s)]\,\underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<0>} \\
&+ (\phi_{30}^{<0>}\underline{\underline{1}} + \phi_{31}^{<0>}\underline{\underline{B}}^{<0>} + \phi_{32}^{<0>}\underline{\underline{B}}^{<0>^2})\underline{\underline{G}}^{<1>}(s)(\underline{\underline{F}}^{<0>^T})^{-1} \\
&+ \underline{\underline{G}}^{<1>}(s)(\phi_{30}^{<0>}(\underline{\underline{F}}^{<0>^T})^{-1} + \phi_{31}^{<0>}\underline{\underline{F}}^{<0>} + \phi_{32}^{<0>}\underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<0>})\}ds
\end{aligned}
$$

where we have used the identities

$$
(10.33)\qquad \underline{\underline{B}}^{<0>}(\underline{\underline{F}}^{<0>^T})^{-1} = \underline{\underline{F}}^{<0>},
$$

and

$$
\underline{\underline{B}}^{<1>}(\underline{\underline{F}}^{<0>^T})^{-1} = \underline{\underline{F}}^{<1>} + \underline{\underline{F}}^{<0>}\underline{\underline{F}}^{<1>^T}(\underline{\underline{F}}^{<0>^T})^{-1}.
$$

Now we can write down the linearized problem of perturbation of the rest state of incompressible viscoelastic solid:

$$
(10.34)\qquad \left.\begin{aligned} \rho_o\,\ddot{\underline{u}}^{<1>} &= \mathrm{div}\,\underline{\underline{S}}^{<1>^T} \\ \mathrm{tr}[(\underline{\underline{F}}^{<0>})^{-1}\,\underline{\underline{F}}^{<1>}] &= 0 \end{aligned}\right\} \quad \text{in } \mathcal{V}_0 \text{ for } t > 0\,,
$$

$$
\underline{u}^{<1>} \text{ is specified on } \partial\mathcal{V}_{10} \text{ for } t > 0\,,
$$

$$
\underline{t}_n^{<1>} \text{ is specified on } \partial\mathcal{V}_{20} \text{ for } t > 0,
$$

$$
\underline{u}^{<1>} \text{ is specified in } \mathcal{V}_0 \text{ for } t \in (-\infty, 0]\,,
$$

where $\underline{\underline{S}}_1^{<1>^T}$ is given by (10.32), $\underline{\underline{F}}^{<1>} = \nabla\underline{u}^{<1>}(\underline{X},t)$, and $\underline{t}_n^{<1>}$ is given by (10.18).

11. The linearized theory and elastic stability

The linearized theory of perturbations of the rest state is a good place to start the study of stability and bifurcation of elastostatic solutions of viscoelastic problems. We have maintained that the use of elastic equations for unsteady motions of simple solids has no good justification and is probably unjustified, except as an approximation which is valid in certain asymptotic limits. It is often true that these asymptotic limits contain all the points at issue in certain studies. But as a matter of principle in the study of stability it is necessary at least to test the stability of a solution to small disturbances of arbitrary frequency. Such time dependent disturbances lead to the linearized equations derived in §10 and not to equations of "dynamic elasticity".

Some interesting points about the stability of elastic solutions of viscoelastic problems emerge from general considerations arising in the theory of stability and bifurcation. To develop these points it is necessary to assume that the stability criteria which are associated with the linearized equations are valid for small disturbances governed by the exact nonlinear equations. So if all solutions of the linearized equations are asymptotically stable then the rest state is stable to small disturbances (conditionally stable) but if one of these disturbances grows without bound the rest state is actually unstable. In the exact theory of stability one goes a step further. In this step, the linearized equations are replaced with spectral equations which arise formally from substituting solutions of the form

(11.1) $\underline{u}^{<1>}(\underline{X},t) = e^{\sigma t}\underline{v}(\underline{X})$, $p^{<1>}(\underline{X},t) = e^{\sigma t}\, p^{<1>}(\underline{X})$

into the linearized equations of motions. The values of $\sigma = \xi + i\omega$ for which the resulting problem has solutions are said to be in the spectrum of that problem. We say that the rest state is stable by criteria of the linearized theory if there are no values σ for which $\xi > 0$ and is unstable if there are some such values. In the exact theory one proves that stability and instability by spectral criteria imply actual stability and instability for the correct nonlinear problem when disturbances are small.

In the problems which come up in mechanics the spectral values $\sigma(k)$ depend on a parameter k. The value σ_1 with the largest real part is called the principal spectral value. The loss of stability is associated with a critical value $k = k_o$ at which $\xi_1(k) = \mathrm{re}\ \sigma_1(k)$ passes through zero from negative to positive. In most of the problems studied in mechanics the spectrum which crosses over in this way is of eigenvalues.

In bifurcation theory we usually assume that the principal spectral value σ_1 is isolated and has only one eigenfunction $\underline{v}(\underline{X})$. In this case, if $d\xi_1(k_o)/dk \neq 0$, we get steady bifurcating solutions if $\sigma_1(k_o) = 0$ and time-periodic ones if $\sigma_1(k_o) = i\omega_o$.

Now we are going to assume all is good with the linearized theory of the stability of the rest state (elastostatics), and that the properties relating the spectral problem to true bifurcation are as in the general theory of bifurcation at a simple eigenvalue.

We may derive the spectral problem by substituting (11.1) into (10.34). It is easy to verify that

$$\underline{\underline{S}}^{<1>^T}(\underline{u}^{<1>}) = \underline{\underline{S}}^{<1>^T}(e^{\sigma t}\underline{v}) = e^{\sigma t}\underline{\underline{\mathcal{L}}}(\underline{v})$$

where $\underline{\underline{\mathcal{L}}}(\underline{v})$ is defined by (11.4), and

$$\underline{t}_n^{<1>}(\underline{u}^{<1>}) = \underline{t}_n^{<1>}(e^{\sigma t}\underline{v}) = e^{\sigma t}\underline{\mathcal{B}}(\underline{v}) \tag{11.2}$$

where $\underline{\mathcal{B}}(\underline{v})$ is defined by the operator which arises from (10.18) when $\underline{u}^{<1>} = e^{\sigma t}\underline{v}$. The spectral problem governing the stability of all elastostatic solutions is then

$$\left.\begin{aligned} &\rho_o\sigma^2\underline{v} = \operatorname{div}\underline{\underline{\mathcal{L}}}(\underline{v}) \\ &\operatorname{tr}[(\underline{\underline{F}}^{<0>})^{-1}(\nabla\underline{v})] = 0 \end{aligned}\right\} \text{ in } \mathcal{V}_0, \tag{11.3}$$

$$\underline{v} \text{ is specified on } \partial\mathcal{V}_{10},$$

$$\underline{\mathcal{B}}(\underline{v}) \text{ is specified on } \partial\mathcal{V}_{20},$$

where

$$\begin{aligned} \underline{\underline{\mathcal{L}}}(\underline{v}) = {} & p^{<0>}(\underline{\underline{F}}^{<0>^T})^{-1}(\nabla\underline{v})^T(\underline{\underline{F}}^{<0>^T})^{-1} - p^{<1>}(\underline{\underline{F}}^{<0>^T})^{-1} \\ & + \hat{f}_1^{<1>}\underline{\underline{F}}^{<0>} + f_1^{<0>}\nabla\underline{v} + \hat{f}_2^{<1>}\underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<0>} \\ & + f_2^{<0>}(\hat{\underline{\underline{B}}}^{<1>}\underline{\underline{F}}^{<0>} + \underline{\underline{B}}^{<0>}\nabla\underline{v}) \\ & + \int_0^\infty \{\operatorname{tr}[(\phi_{11}^{<0>}(s)\underline{\underline{B}}^{<0>} + \phi_{12}^{<0>}(s)\underline{\underline{B}}^{<0>^2})\hat{\underline{\underline{G}}}^{<1>}]\underline{\underline{F}}^{<0>} \\ & + \operatorname{tr}[\phi_{21}^{<0>}(s)\underline{\underline{B}}^{<0>} + \phi_{22}^{<0>}(s)\underline{\underline{B}}^{<0>^2})\hat{\underline{\underline{G}}}^{<1>}]\underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<0>} \\ & + (\phi_{30}^{<0>}(s)\underline{\underline{1}} + \phi_{31}^{<0>}(s)\underline{\underline{B}}^{<0>} + \phi_{32}^{<0>}(s)\underline{\underline{B}}^{<0>^2}) \\ & \qquad \hat{\underline{\underline{G}}}^{<1>}(\underline{\underline{F}}^{<0>^T})^{-1} \\ & + \hat{\underline{\underline{G}}}^{<1>}(\phi_{30}^{<0>}(s)(\underline{\underline{F}}^{<0>^T})^{-1} + \phi_{31}^{<0>}(s)\underline{\underline{F}}^{<0>} \\ & \qquad + \phi_{32}^{<0>}(s)\underline{\underline{B}}^{<0>}\underline{\underline{F}}^{<0>})\}(e^{-\sigma s} - 1)ds, \end{aligned} \tag{11.4}$$

$$\hat{f}_i^{<1>} = (\partial f_i/\partial I_B\Big|_{\varepsilon=0})\,\mathrm{tr}\,\hat{\underline{\underline{B}}}^{<1>} + 2(\partial f_i/\partial II_B\Big|_{\varepsilon=0})\,\mathrm{tr}(\underline{\underline{B}}^{<0>}\hat{\underline{\underline{B}}}^{<1>}),$$

$$\hat{\underline{\underline{B}}}^{<1>} = (\nabla\underline{v})\underline{\underline{F}}^{<0>T} + \underline{\underline{F}}^{<0>}(\nabla\underline{v})^T,$$

and

$$\hat{\underline{\underline{G}}}^{<1>} = (\nabla\underline{v})(\underline{\underline{F}}^{<0>})^{-1} + (\underline{\underline{F}}^{<0>T})^{-1}(\nabla\underline{v})^T.$$

All the quantities except the $\phi_{ij}^{<0>}(s)$ under the integral sign in (11.4) are independent of s; for example,

$$\int_0^\infty \mathrm{tr}[\phi_{11}^{<0>}(s)\ \underline{\underline{B}}^{<0>}\hat{\underline{\underline{G}}}^{<1>}]\underline{\underline{F}}^{<0>}ds$$

$$= \{\int_0^\infty \phi_{11}^{<0>}(s)ds\}\ \mathrm{tr}[\underline{\underline{B}}^{<0>}\hat{\underline{\underline{G}}}^{<1>}]\underline{\underline{F}}^{<0>}.$$

So the spectral problem contains the history in the convenient form of integrals over material functions, independent of the motion.

In the simplest of the rest states,the natural state studied in §8, it was possible to obtain explicit formulas, like (8.19) for σ and to use such results to infer properties of the material parameters. The problem (11.3) is much more difficult than the one in §8 because there are so many material parameters and because the equations (11.3) govern the stability of the whole class of elastostatic deformations of viscoelastic solids. Nonetheless, there may be a principle, equivalent to the requirement that the natural state of a solid should be stable against all disturbances, which can be used to characterize the material parameters appearing in the problem (11.3). Many states of elastostatic deformation of solids are unstable so that the simple criterion of stability has no force here.

The principle which we wish to consider is that instability of elastic deformations of viscoelastic solids cannot lead to bifurcation into self-sustained oscillations. It is difficult for us to imagine how a time-dependent motion of a viscoelastic solid could arise from a static deformation of that solid.

To illuminate some considerations behind our conjecture it is useful to compare the stability problems which arise in the stability of fluids with those which arise in the stability of solids. In the celebrated Taylor problem in hydrodynamics the flow between concentric cylinders is driven, say, by the steady rotation of the inner cylinder. At first there is a featureless flow (Couette flow) which is uniform like the data; at higher speeds Couette flow gives up its stability to a secondary flow which is arrayed in a set of Taylor vortices of approximately square cross-section. At still higher speeds these steady vortices bifurcate into a time-periodic motion in which a wave undulates around the vortices. The corresponding elastic solution is the torsional deformation of an incompressible elastic cylinder. The solution of this problem was found by Rivlin (1949). Green and Spencer (1959) have studied the problem using a static theory of elastic stability. Penn and Kearsley (1976) have demonstrated in experiments that Rivlin's solution is unstable when the deformation is sufficiently large. The symmetry-breaking bifurcation observed by Penn and Kearsley is in the form of spiral bands. We shall formulate the problem of stability of Rivlin's solution in the context of our viscoelastic theory in the Appendix.

The main point of comparison is that in the fluids problem the boundary data, though steady, does work and gives a continuous supply of energy which can be converted into permanent time-dependent motion. In the solids problem, the imposed steady twist does no work and does not supply energy which can be used to drive a motion.

If we now suppose that the spectral problem (11.3) has the same relevance to bifurcation as in the general theory of bifurcation, we may expect to find steady symmetry-breaking bifurcation, like the experiments, when the spectrum is of eigenvalues σ and the eigenvalue σ with the largest real part is real-valued at criticality. The other possibility is that $\sigma = i\omega$ is not zero when ξ is. Then $\pm i\omega$ are both eigenvalues at criticality and $e^{\pm i\omega t}$ is oscillatory. In the usual case this situation implies bifurcation into a time-periodic motion. If our intuition is correct we should not have time-periodic bifurcation in the elastic problem with steady forcing. So if our choice of material parameters and functions leads to complex-valued σ at criticality we have made a bad choice.

Appendix: The spectral problem for the stability of Rivlin's solution for torsional deformation of a viscoelastic cylinder

It is of interest to consider the stability theory discussed in §11 in the simplest possible nontrivial case. Perhaps this simplest nontrivial case is one of those (torsional deformation) found by Rivlin (1949).

Torsional deformation of a right circular cylinder (of radius a and height 2) of incompressible, initially isotropic, viscoelastic material is given by:

$$r^{o} = R,$$

$$\theta^{o} = \Theta + kZ,$$

$$z^{o} = Z,$$

where $X_I = \{R, \Theta, Z\}$ are coordinates of a material particle $\underline{X}$ in the natural state and $x_i^{o} = \{r^{o}, \theta^{o}, z^{o}\}$ are coordinates of $\underline{x}^{o}$, the position of $\underline{X}$ in the deformed state. It is necessary that some forces and torques be applied to the top and bottom surfaces to maintain this deformation and the constant height. The boundary conditions satisfied by this deformation are: $\theta = \Theta \pm k$ at $Z = \pm 1$ and the surface $r^{o} = a$ is stress-free.

To form the spectral problem (11.3) for the stability of Rivlin's solution we need to find the components of $\underline{\underline{F}}^{<0>}$, $(\underline{\underline{F}}^{<0>})^{-1}$, $\underline{\underline{B}}^{<0>}$ and $\underline{\underline{C}}^{<0>}$. $\underline{e}_I = \{\underline{e}_R, \underline{e}_\Theta, \underline{e}_Z\}$ is the orthogonal basis corresponding to the coordinate-system in the natural state. We introduce the orthonormal base vectors $\hat{\underline{e}}_I = \underline{e}_I/|\underline{e}_I|$. $|\underline{e}_I| = \{1, R, 1\}$. Similarly $\underline{e}_i = \{\underline{e}_{r^o}, \underline{e}_{\theta^o}, \underline{e}_{z^o}\}$ is the orthogonal

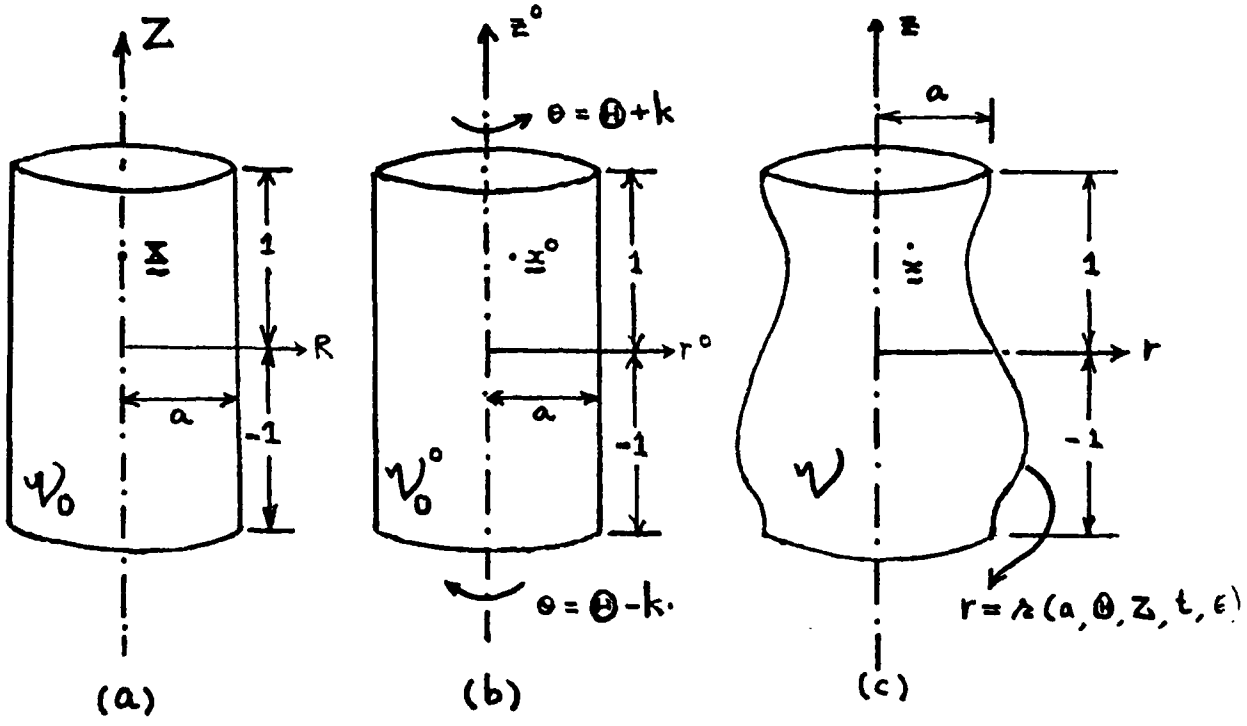


Fig. A.1 Perturbation of the rest (torsional) state of an incompressible, initially isotorpic,cylinder.

(a) The natural state

$$\mathcal{V}_0 = \{R,\theta,Z: 0 \leq R < a,\ 0 \leq \theta \leq 2\pi,\ -1 < Z < 1\} \quad ,$$

(b) The rest state

$$\mathcal{V}_0^{\circ} = \{r^{\circ},\theta^{\circ},z^{\circ}: \ 0 \leq r^{\circ} < a,\ 0 \leq \theta^{\circ} \leq 2\pi,\ -1<z^{\circ}<1\} \ ,$$

(c) The perturbed state

$$\mathcal{V} = \{r,\theta,z: \ 0 \leq r < \ell(a,\theta,Z,t,\varepsilon), 0 \leq \theta \leq 2\pi,\ -1<z<1\}$$

where $\ell(a,\theta,Z,t,\varepsilon)$ is the free surface.

basis in the rest state. The corresponding orthonormal basis is $\hat{\underline{e}}_i = \underline{e}_i/|\underline{e}_i|$ where $|\underline{e}_i| = \{1, r^{\circ}, 1\}$. Then $\underline{\underline{F}}^{<0>}$, $\underline{\underline{B}}^{<0>}$ and $\underline{\underline{C}}^{<0>}$ have the following representation:

$$\underline{\underline{F}}^{<0>} = F^{<0>}_{iJ}\ \hat{\underline{e}}_i \otimes \hat{\underline{e}}_J{}^{*};\ F^{<0>}_{iJ} = \frac{\partial x^o_i}{\partial X_J}\ \frac{|\underline{e}_i|}{|\underline{e}_J|}\ \text{(no sum over i \& J)};$$

$$\underline{\underline{B}}^{<0>} = B^{<0>}_{ij}\ \hat{\underline{e}}_i \otimes \hat{\underline{e}}_j;\ B^{<0>}_{ij} = F^{<0>}_{iK}\ F^{<0>}_{jK};$$

$$\underline{\underline{C}}^{<0>} = C^{<0>}_{IJ}\ \hat{\underline{e}}_I \otimes \hat{\underline{e}}_J;\ C^{<0>}_{IJ} = F^{<0>}_{kI}\ F^{<0>}_{kJ}.$$

The matrices of components $F^{<0>}_{iJ}$, $B^{<0>}_{ij}$, $C^{<0>}_{IJ}$ are:

$$[F^{<0>}_{iJ}] = [\delta_{iJ}] + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & r^o k \\ 0 & 0 & 0 \end{pmatrix},$$

$$[B^{<0>}_{ij}] = [\delta_{ij}] + \begin{pmatrix} 0 & 0 & 0 \\ 0 & r^{o2}k^2 & r^o k \\ 0 & r^o k & 0 \end{pmatrix},$$

$$[C^{<0>}_{IJ}] = [\delta_{IJ}] + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & r^o k \\ 0 & r^o k & r^{o2}k^2 \end{pmatrix}.$$

The components of $(\underline{\underline{F}}^{<0>})^{-1}$ are easy to compute.

$$(\underline{\underline{F}}^{<0>})^{-1} = (F^{<0>})^{-1}_{Ij}\ \hat{\underline{e}}_I \otimes \hat{\underline{e}}_j$$

where the matrix $[(F^{<0>})^{-1}_{Ij}]$ is the inverse of the matrix $[F^{<0>}_{iJ}]$.

---

*In this section, we use upper and lower case suffixes. This is to emphasize the fact that such components are either with respect to the basis in the natural state (upper case suffix) or with respect to the mixed basis (One upper and one lower case suffix). The usual summation convention applies to these suffices also unless the contrary is explicitely stated.

$$[(F^{<0>})^{-1}{}_{Ij}] = [\delta_{Ij}] - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & r^o k \\ 0 & 0 & 0 \end{bmatrix}.$$

From this we can compute $\underline{\underline{T}}^{<0>}$, $\underline{t}_n^{<0>}$ and $\underline{\underline{S}}^{<0>T}$. $p^{<0>}$ is computed by solving the equation of motion div $\underline{\underline{S}}^{<0>T} = \underline{0}$.

It is easy to check that $\underline{n}^{<0>} = \underline{e}_{r^o} = \hat{\underline{e}}_1$.

Now we consider perturbation of the rest state given by:

$$r = r^o + \varepsilon r^{<1>} + O(\varepsilon^2) = R + \varepsilon r^{<1>} + O(\varepsilon^2),$$

$$\theta = \theta^o + \varepsilon\theta^{<1>} + O(\varepsilon^2) = \Theta + kZ + \varepsilon\theta^{<1>} + O(\varepsilon^2),$$

$$z = z^o + \varepsilon z^{<1>} + O(\varepsilon^2) = Z + \varepsilon z^{<1>} + O(\varepsilon^2),$$

where (r,θ, z) are coordinates of $\underline{x}$, the position of $\underline{X}$ in the perturbed state. We could treat (r,θ,z) as functions of either r^o,θ^o and z^o or R, Θ & Z. $\underline{x} = \underline{x}^o + \varepsilon\underline{u}^{<1>} + O(\varepsilon^2) = \underline{X} + \underline{u}^o + \varepsilon\underline{u}^{<1>} + O(\varepsilon^2)$. Then $\underline{u}^{<1>}$ has the representation:

$$\underline{u}^{<1>} = r^{<1>}\,\hat{\underline{e}}_1 + r^o\theta^{<1>}\,\hat{\underline{e}}_2 + z^{<1>}\,\hat{\underline{e}}_3.$$

The components of $\partial\underline{u}^{<1>}/\partial\underline{x}^o$ in the basis $\hat{\underline{e}}_i$ are

$$[\partial\underline{u}^{<1>}/\partial\underline{x}^o] = \begin{bmatrix} \partial r^{<1>}/\partial r^o & \frac{1}{r^o}\,\partial r^{<1>}/\partial\theta^o - \theta^{<1>} & \partial r^{<1>}/\partial z^o \\ \partial(r^o\,\theta^{<1>})/\partial r^o & \partial\theta^{<1>}/\partial\theta^o + r^{<1>}/r^o & r^o\partial\theta^{<1>}/\partial z^o \\ \partial z^{<1>}/\partial r^o & \frac{1}{r^o}\,\partial z^{<1>}/\partial\theta^o & \partial z^{<1>}/\partial z^o \end{bmatrix}.$$

Now

$$\underline{\underline{F}}^{<1>} = \partial\underline{u}^{<1>}/\partial\underline{X} = (\partial\underline{u}^{<1>}/\partial\underline{x}^o)(\partial\underline{x}^o/\partial\underline{X}).$$

The representation of $\underline{\underline{F}}^{<1>}$ is:

$$\underline{\underline{F}}^{<1>} = F^{<1>}_{iJ}\ \hat{\underline{e}}_i \otimes \hat{\underline{e}}_J;\quad F^{<1>}_{iJ} = \sum_{k=1}^{3} \left(\frac{\partial \underline{u}^{<1>}}{\partial \underline{x}^{o}}\right)_{ik} \frac{\partial x_k^{\,o}}{\partial X_J}\ \frac{|\underline{e}_k|}{|\underline{e}_J|} \quad \text{(no sum over J)}$$

$$= \left(\frac{\partial \underline{u}^{<1>}}{\partial \underline{x}^{o}}\right)_{ik} F^{<0>}_{kJ}\ .$$

The matrix of $F^{<1>}_{iJ}$ is given by :

$$(A.1)\qquad [F^{<1>}_{iJ}] = \begin{pmatrix} \partial r^{<1>}/\partial r^{o} & (1/r^{o})\,\partial r^{<1>}/\partial \theta^{o} - \delta^{<1>} & \partial r^{<1>}/\partial z^{o} \\ \partial (r^{o}\delta^{<1>})/\partial r^{o} & \partial \theta^{<1>}/\partial \theta^{o} + r^{<1>}/r^{o} & r^{o}\partial \delta^{<1>}/\partial z^{o} \\ \partial z^{<1>}/\partial r^{o} & (1/r^{o})\,\partial z^{<1>}/\partial \theta^{o} & \partial z^{<1>}/\partial z^{o} \end{pmatrix}$$

$$+ \begin{pmatrix} 0 & 0 & k(\partial r^{<1>}/\partial \theta^{o} - r^{o}\theta^{<1>}) \\ 0 & 0 & k(r^{o}\partial \theta^{<1>}/\partial \theta^{o} + r^{<1>}) \\ 0 & 0 & k(\partial z^{<1>}/\partial \theta^{o}) \end{pmatrix}.$$

Now we are in a position to find the components of $\underline{\underline{B}}^{<1>}$ and $\underline{\underline{G}}^{<1>}(s)$:

$$\underline{\underline{B}}^{<1>} = B^{<1>}_{ik}\ \hat{\underline{e}}_i \otimes \hat{\underline{e}}_k$$

where

$$(A.2)\qquad [B^{<1>}_{ik}] = 2[E^{<1>}_{ik}] + r^{o}k \begin{pmatrix} 0 & F^{<1>}_{13} & 0 \\ F^{<1>}_{13} & 2F^{<1>}_{23} & F^{<1>}_{33} \\ 0 & F^{<1>}_{33} & 0 \end{pmatrix},$$

and

$$2E^{<1>}_{ik} = F^{<1>}_{iJ}\ \delta_{Jk} + F^{<1>}_{kJ}\ \delta_{Ji},$$

$$\underline{\underline{G}}^{<1>}(s) = G^{<1>}_{ik}(s)\ \hat{\underline{e}}_i \otimes \hat{\underline{e}}_k$$

where

$$[G^{<1>}{}_{ik}(s)] = [2\,[\![E^{<1>}{}_{ik}]\!]\,] - r^{o}{}_{k}\begin{pmatrix} 0 & 0 & [\![F^{<1>}{}_{12}]\!] \\ 0 & 0 & [\![F^{<1>}{}_{22}]\!] \\ [\![F^{<1>}{}_{12}]\!] & [\![F^{<1>}{}_{22}]\!] & 2[\![F^{<1>}{}_{32}]\!] \end{pmatrix}$$

and

$[\![a]\!] \equiv a(t-s) - a(t)$ for any scalar a.

Now we can find components of $\underline{n}^{<1>}$, $\underline{\underline{T}}^{<1>}$, $\underline{t}_{n}{}^{<1>}$ and div $\underline{\underline{S}}^{<1>T}$ in the basis of the rest state and components of

$$\underline{\underline{S}}^{<1>T} = S^{<1>T}{}_{iJ}\ \hat{\underline{e}}_{i} \otimes \hat{\underline{e}}_{J}$$

in the mixed basis.

Finally we note that $(\underline{\underline{F}}^{<0>})^{-1}\ \underline{\underline{F}}^{<1>}$ has the representation:

$$(\underline{\underline{F}}^{<0>})^{-1}\ \underline{\underline{F}}^{<1>} = (F^{<0>})^{-1}_{Ij}\ F^{<1>}{}_{jK}\ \hat{\underline{e}}_{I} \otimes \hat{\underline{e}}_{K}\ .$$

So

$$\mathrm{Tr}[(\underline{\underline{F}}^{<0>})^{-1}\ \underline{\underline{F}}^{<1>}] = (F^{<0>})^{-1}_{Ij}\ F^{<1>}{}_{jI}$$

$$= F^{<1>}{}_{11} + F^{<1>}{}_{22} + F^{<1>}{}_{33} - r^{o}{}_{k}\ F^{<1>}{}_{32}.$$

Now to write the spectral problem (11.3) into the component form, we let

$$r^{<1>}(\underline{X},t) = e^{\sigma t}\ \hat{r}^{<1>}(\underline{X}),$$

$$\theta^{<1>}(\underline{X},t) = e^{\sigma t}\ \hat{\theta}^{<1>}(\underline{X}),$$

$$z^{<1>}(X,t) = e^{\sigma t}\ \hat{z}^{<1>}(\underline{X}).$$

Then

(A.3) $$\underline{v} = \hat{r}^{<1>}\ \hat{\underline{e}}_{1} + r^{o}\ \hat{\theta}^{<1>}\ \hat{\underline{e}}_{2} + z^{<1>}\ \hat{\underline{e}}_{3}\ .$$

$$\nabla\underline{v} = (\nabla\underline{v})_{iJ}\ \hat{\underline{e}}_{i} \otimes \hat{\underline{e}}_{J}\ ,$$

and

$$\hat{B}^{<1>} = \hat{B}^{<1>}_{ik}\, \hat{\underline{e}}_i \otimes \hat{\underline{e}}_k,$$

where the components $(\nabla \underline{v})_{iJ}$ and $\hat{B}^{<1>}_{ik}$ can be obtained from the equations (A.1) and (A.2) just by replacing $r^{<1>}$, $\theta^{<1>}$ and $z^{<1>}$ by $\hat{r}^{<1>}$, $\hat{\theta}^{<1>}$ and $\hat{z}^{<1>}$.

$$\hat{\underline{G}}^{<1>} = \hat{G}^{<1>}_{ik}\, \hat{\underline{e}}_i \otimes \hat{\underline{e}}_k$$

where

$$[\hat{G}^{<1>}_{ik}] = [(\nabla \underline{v})_{iJ}\, \delta_{Jk} + (\nabla \underline{v})_{kJ} \delta_{Ji}] - r^o k \begin{bmatrix} 0 & 0 & (\nabla \underline{v})_{12} \\ 0 & 0 & (\nabla \underline{v})_{22} \\ (\nabla \underline{v})_{12} & (\nabla \underline{v})_{22} & 2(\nabla \underline{v})_{32} \end{bmatrix}.$$

Now we can find the components of $\operatorname{div} \underline{\underline{\mathcal{L}}}(\underline{v})$ and $\underline{\mathcal{B}}(\underline{v})$ in the basis of the rest state. Finally

$$\operatorname{Tr}[(\underline{\underline{F}}^{<0>})^{-1}(\nabla \underline{v})] = (\nabla \underline{v})_{11} + (\nabla \underline{v})_{22} + (\nabla \underline{v})_{33} - r^o k (\nabla \underline{v})_{32}.$$

Then the spectral problem (11.3) becomes:

$$\left.\begin{aligned} \rho_o \sigma^2 \hat{r}^{<1>} &= [\operatorname{div} \underline{\underline{\mathcal{L}}}(\underline{v})]_1 \\ \rho_o \sigma^2 r^o \hat{\theta}^{<1>} &= [\operatorname{div} \underline{\underline{\mathcal{L}}}(\underline{v})]_2 \\ \rho_o \sigma^2 \hat{z}^{<1>} &= [\operatorname{div} \underline{\underline{\mathcal{L}}}(\underline{v})]_3 \\ (\nabla \underline{v})_{11} + (\nabla \underline{v})_{22} + (\nabla \underline{v})_{33} &= r^o k (\nabla \underline{v})_{32} \end{aligned}\right\} \text{ in } \mathcal{V}_0 ,$$

$$\hat{r}^{<1>} = \hat{\theta}^{<1>} = \hat{z}^{<1>} = 0 \qquad \text{on } Z = \pm 1 ,$$

$$\underline{\mathcal{B}}(\underline{v}) = \underline{0} \qquad \text{on } R = a ,$$

where

$$\underline{v} = \hat{r}^{<1>}\, \hat{\underline{e}}_1 + r^o\, \hat{\theta}^{<1>}\, \hat{\underline{e}}_2 + \hat{z}^{<1>}\, \hat{\underline{e}}_3$$

as given by (A.3).

References

Breuer,S., & Onat, E.T.: On uniqueness in linear viscoelasticity. Quart. Appl. Math. 19, 355-359 (1962).

Coleman, B.D., & Noll, W.: Foundations of linear viscoelasticity. Rev. Mod. Phys. 33, 239-249 (1961). Erratum, ibid. 36, 1103 (1964).

Dafermos, C.M.: Asymptotic stability in viscoelasticity. Arch. Rational Mech. Anal. 37, 297-308 (1970).

Dixit, P.M.: Ph.D. Thesis, Dept. of Aerospace Eng. and Mechanics, Univ. of Minn., Minneapolis, Minnesota,1979. (forthcoming).

Edelstein, W.S., & Gurtin, M.E.: Uniqueness theorems in the linear theory of anisotropic viscoelastic solids. Arch. Rational Mech. Anal. 17, 47-60 (1964).

Fujita, H., & Kato, T.: On the Navier-Stokes initial value problem, I. Arch. Rational Mech. Anal. 16, 269-315 (1964).

Green, A.E., & Rivlin, R.S.: The mechanics of non-linear materials with memory, Part I. Arch. Rational Mech. Anal. 1, 1-24 (1957). Erratum, ibid. 1, 470 (1958).

Green, A.E., & Spencer, A.J.M.: The stability of a circular cylinder under finite extension and torsion. J. Math. Phys. Vol. XXXVII, 316-338 (1959).

Gurtin, M.E., & Sternberg, E.: On the linear theory of viscoelasticity. Arch. Rational Mech. Anal. 11, 291-356 (1962).

Joseph, D.D.: Stability of Fluid Motions. Chap. XIII, Vol. II, Berlin-Heidelberg-New York: Springer Tracts in Natural Philosophy, 1976.

Joseph, D.D.,& Beavers, G.S.: Free surface problems in rheological fluid mechanics. Rheologica Acta 16, 169-189 (1977).

Ladyzhenskaya, O.A.: The Mathematical Theory of Viscous Incompressible Flow. New York-London: Gordon and Breach, 1963.(2nd Edition, 1969).

Odeh, F., & Tadjbakhsh, I.: Uniqueness in the linear theory of viscoelasticity. Arch. Rational Mech. Anal. 18, 244-250 (1965).

Onat, E.T., & Breuer, S.: On uniqueness in linear viscoelasticity. Progress in Applied Mechanics, Prager Anniversary Volume. New York: Macmillan, 1963.

Penn, R.W., & Kearsley, E.A.: The scaling law for finite torsion of elastic cylinders. Trans. Soc. Rheology 20, 227-238 (1976).

Pipkin, A.C.: Small finite deformations of viscoelastic solids. Rev. Mod. Phys. 36, 1034-1041 (1964).

Pipkin, A.C.,& Rivlin, R.S.: Small deformations superposed on large deformations in materials with fading memory. Arch. Rational Mech. Anal. 8, 297-308 (1961).

Riesz, F., & Nagy, B.: Functional Analysis. New York: Frederick Ungar Publishing Co., 1955. (6th Printing, 1972).

Rivlin, R.S.: Large elastic deformations of isotropic materials IV. Further developments of the general theory. Phil. Trans. Royal Society of London, Series A,241, 379-397 (1949).

Rivlin, R.S.: An introduction to non-linear continuum mechanics. Centro Internazionale Matematico Estivo. Corso tenuto a Bressanone dal 3 all' 11 Settembre 1969, 153-309.

Slemrod, M.: An energy stability method for simple fluids. Arch. Rational Mech. Anal. 62, 303-321 (1977).

STRESSES AND DEFORMATION BENEATH A RIGID WHEEL

Mosaid M. Al-Hussaini
College of Engineering and Petroleum
Kuwait University, Kuwait
Formerly Research Engineer
U. S. Army Engineer Waterways Experiment Station
Corps of Engineers, Vicksburg, Mississippi

George Y. Baladi
U. S. Army Engineer Waterways Experiment Station
Corps of Engineers, Vicksburg, Mississippi

ABSTRACT. This paper documents the development of a closed form solution of the distribution of displacements and stresses under a rigid wheel. The wheel is assumed to be infinitely long and partially embedded in a semi-infinite homogeneous, elastic-isotropic medium.

In the first step in the development of a suitable solution, the stress at the interface between the wheel and the underlying material is decomposed into normal and tangential stresses. The stress function was used in representing the displacements and stresses within the soil media in terms of analytical functions. The Schwarz-Christoffel equation was used to transform the geometry and the boundary condition of the region beneath the wheel and to match it with the stress functions. The Cauchy integral equation was applied on the transformed boundary conditions to obtain the shear and normal stresses and displacements at any point within the region of the soil-wheel system.

The analytical solution is believed to permit the evaluation of displacements and stresses within the soil beneath a wheel resulting from various combinations of radial and tangential stresses.

I. INTRODUCTION. The study of soil-wheel mechanics is of interest to the engineer who has to make a decision on running gear requirements for the most efficient vehicle and to the military planner who wants to be assured of vehicle mobility in a given terrain. To date, most research has been concentrated on determining stress distribution along the soil-wheel interface, and very little information is available regarding stresses and deformations within the soil mass beneath the wheel.

The problem of determining stress and deformation distribution beneath a moving wheel is a complex one. A simplified test procedure has been developed for determining vehicle performance by means of a plate load test[1] or a special cone test.[2] Onafeko and Reece[3] developed an experimental procedure for predicting the relationship between slip and shear stresses beneath a rigid wheel. Wong and Reece[4] conducted an experimental study to predict rigid wheel performance based on soil-wheel stresses. A finite element method was used by Perumpral and his

coworkers[5] to determine stress and deformation distribution beneath a rigid wheel. Their analysis was based on a variable modulus of elasticity as determined from triaxial compression tests. A similar finite element procedure was used by Yong and his coworkers[6] to predict deformation distribution beneath a tire wheel. A closed-form solution for predicting stresses beneath a rigid wheel on an elastic soil medium was developed by Gilbert and Al-Hussaini.[7]

The closed-form solution for predicting stresses and deformations within the medium supporting a wheel can expand our understanding of soil-wheel mechanics. The work initiated by Gilbert and Al-Hussaini[7] is intended to include the deformation distribution within a soil supporting a circular rigid wheel.

II. METHOD OF ANALYSIS. The theoretical derivation is based on simplified assumptions in which the soil medium is simulated by a semi-infinite homogeneous elastic isotropic mass under plane strain conditions. In addition, the gravity stresses within the soil mass are neglected. To satisfy uniqueness of solution, the equilibrium, boundary, and compatibility conditions must be satisfied.

a. Equilibrium conditions: The equilibrium conditions for a weightless material are

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} = 0 \tag{1a}$$

$$\frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{xy}}{\partial x} = 0 \tag{1b}$$

where σ_x , σ_y are the normal stresses acting on the x and y plane, respectively, and τ_{xy} is the shear stress within the plane bounded by the x and y axes.

b. Boundary conditions: The boundary conditions are

$$\sigma_x \frac{dy}{ds} - \tau_{xy} \frac{dx}{ds} = \bar{X} \tag{2a}$$

$$\tau_{xy} \frac{dy}{ds} - \sigma_y \frac{dx}{ds} = \bar{Y} \tag{2b}$$

where $\bar{X}$ and $\bar{Y}$ are the Cartesian components of forces per unit area acting on a small element of the boundary ds .

2. Compatibility conditions: The compatibility equation under plane strain condition is

$$\nabla^2(\sigma_x + \sigma_y) = 0 \tag{3}$$

where ∇^2 is a Laplace operator.

Stress Function. Equation 3 can be easily solved by introducing a new function $U(x,y)$ such that

$$\sigma_x = \frac{\partial^2 U(x,y)}{\partial y^2} \tag{4a}$$

$$\sigma_y = \frac{\partial^2 U(x,y)}{\partial x^2} \tag{4b}$$

$$\tau_{xy} = -\frac{\partial^2 U(x,y)}{\partial x \partial y} \tag{4c}$$

Adding equation 4a and equation 4b results in

$$(\sigma_x + \sigma_y) = \frac{\partial^2 U(x,y)}{\partial y^2} + \frac{\partial^2 U(x,y)}{\partial x^2} = \nabla^2 U(x,y) \tag{5}$$

Substitution of equation 5 into equation 3 leads to

$$\nabla^2(\sigma_x + \sigma_y) = \nabla^4 \left[U(x,y)\right] \tag{6}$$

Complex Representation. The solution of the biharmonic equation (Equation 6) can be simplified using complex variables following a procedure presented by Timoshenko and Goodier[8] to obtain the following

$$U(x,y) = 1/2\left[\bar{z}\,\phi(z) + z\overline{\phi(z)} + \chi(z) + \overline{\chi(z)}\right] \tag{7}$$

where $\phi(z)$ and $\chi(z)$ are analytic functions and $\overline{\phi(z)}$ and $\overline{\chi(z)}$ are their complex conjugate. These functions can be determined from the boundary conditions. It has been shown by Timeshenko and Goodier[8] that the normal and shear stresses can be expressed as

$$\sigma_x + i\tau_{xy} = \phi'(z) + \overline{\phi'(z)} - z\overline{\phi''(z)} - \overline{\chi''(z)} \tag{8a}$$

$$\sigma_y - i\tau_{xy} = \phi'(z) + \overline{\phi'(z)} + z\overline{\phi''(z)} + \overline{\chi''(z)} \tag{8b}$$

Adding and substracting equations 8a and 8b, and substituting $\phi'(z)$ and $\chi''(z)$ with $\Phi(z)$ and $\Psi(z)$, respectively, the following expressions can be obtained

$$\sigma_x + \sigma_y = 2[\Phi(z) + \overline{\Phi(z)}] = 4\,\mathrm{Re}[\Phi(z)] \tag{9a}$$

$$\sigma_y - \sigma_x + 2i\tau_{xy} = 2[\bar{z}\Phi'(z) + \Psi(z)] \tag{9b}$$

Complex Representation of the Displacement. The stress-strain relationships for the plane strain condition are

$$2G\,\varepsilon_x = (1 - \nu)\,\sigma_x - \nu\sigma_y \tag{10a}$$

$$2G\,\varepsilon_y = (1 - \nu)\,\sigma_y - \nu\sigma_x \tag{10b}$$

where ν is Poisson's ratio and G is the shear modulus. Substituting equation 9 into equation 10 and after algebraic manipulation leads to

$$2G\,\varepsilon_x = (1 - 2\nu)\left[\Phi(z) + \overline{\Phi(z)}\right] - \bar{z}\Phi'(z) - z\overline{\Phi'(z)} - \Psi(z) - \overline{\Psi(z)} \tag{11a}$$

$$2G\,\varepsilon_y = (1 - 2\nu)\left[\Phi(z) + \overline{\Phi(z)}\right] + \bar{z}\Phi'(z) + z\overline{\Phi'(z)} + \Psi(z) \pm \overline{\Psi(z)} \tag{11b}$$

Integrating equation 11a with respect to x and equation 11b with respect to iy, and after adding the resulting equations we obtain

$$2G(v_x + iv_y) = (3 - 4\nu)\,\phi(z) - z\overline{\Phi(z)} - \overline{\psi(z)} + H(x) + ig(y) \tag{12}$$

where $\frac{d\psi(z)}{d(z)} = \Psi(z)$, and v_x and v_y are the horizontal and vertical deformations, respectively, and $H(x)$ and $g(y)$ are arbitrary functions. By taking the derivative of equation 12 and comparing it with equation 11 we obtain

$$H'(x) + g'(y) = 0 \tag{13}$$

It follows that $H'(x) = -g'(y) = C$ where C is a constant. Therefore, the functions $H(x)$ and $g(y)$ represent rigid body displacement in the z plane and they do not influence the stresses or strains. When $H(x)$ and $g(y)$ are discarded, equation 12 can be written

$$2i(v_x + iv_y) = (3 - 4\nu)\,\phi(z) - z\overline{\Phi(z)} - \overline{\psi(z)} \tag{14}$$

where equation 14 represents the relationship between the displacements and the complex functions $\phi(z)$ and $\psi(z)$.

Representation of Stresses and Deformation Components in Curvilinear Coordinates. Because the problem under consideration contains a curved boundary, the problem will be greatly simplified by mapping the geometry onto a half space. Assume that the function by which a proper transformation can be achieved in represented by

$$z = f(t) = f(r + is) \tag{15}$$

where r and s are curvilinear coordinates in the t-plane. It is more appropriate to transform the functions $\Phi(z)$ and $\Psi(z)$ from the z-plane to a corresponding one in the t-plane. This can be accomplished as follows.

$$\Phi(z) = \Phi[f(t)] = \Phi(t) \tag{16a}$$

$$\Psi(z) = \Psi[f(t)] = \Psi(t) \tag{16b}$$

It has been shown by Timoshenko and Goodier[8] that stresses in the z- and t-plane can be related by the following expressions

$$\sigma_s - \sigma_r + 2i\tau_{rs} = (\sigma_x - \sigma_y + 2i\tau_{xy})\, e^{2i\theta} \tag{17a}$$

and

$$\sigma_r + \sigma_s = \sigma_x + \sigma_y \tag{17b}$$

where

$$e^{2i\theta} = \frac{f'(t)}{\overline{f'(t)}} \tag{17c}$$

Boundary Conditions in the z-Plane. Let N and T be the normal and shear stresses, respectively, applied to the boundary B in the z-plane; thus

$$\left(N + iT\right)_B = \left(\sigma_y + i\tau_{xy}\right)_B = \left(\frac{\sigma_x + \sigma_y}{2} + \frac{\sigma_y - \sigma_x + 2i\tau_{xy}}{2}\right)_B \qquad (18)$$

which, when compared with equations 9, 16, and 17, yields the following

$$\left(N + iT\right)_B = \left[\phi(t) + \overline{\phi(t)} + \overline{f(t)}\,\frac{\phi'(t)}{f'(t)} + \psi(t)\right]_B \qquad (19)$$

Since the boundary in the z-plane corresponds to the real axis of the t-plane, i.e., $t = r$ and $s = 0$, thus along the boundary of the t-plane equation 19 may be reduced to

$$N + iT = \overline{\phi(r)} + \phi(r) + \overline{f(r)}\,\frac{\phi'(r)}{f'(r)} + \psi(r) \qquad (20)$$

whose conjugate is

$$N - iT = \phi(r) + \overline{\phi(r)} + f(r)\,\frac{\overline{\phi'(r)}}{\overline{f'(r)}} + \overline{\psi(r)} \qquad (21)$$

Equations 20 and 21 represent the boundary condition in the z-plane.

III. SOLUTION OF THE PROBLEM. The solution of the stresses and deformations beneath a rigid wheel as shown in fig. 1 involves the determination of a mapping function which maps the region containing the supporting medium of the wheel, represented by the z-plane in fig. 2, into a semi-infinite region representing the t-plane. This is accomplished by mapping the z-plane into an auxiliary plane called the w-plane using the following relationship

$$z = a \tanh w \qquad (22)$$

using the Schwarz-Christoffel transformation[9] to map the w-plane onto the t-plane. Detailed derivation of the transformation was previously presented by Gilbert and Al-Hussaini[7] and the transformation function can be written as

$$z = a \tanh\left[\frac{\pi i}{2} - \frac{K}{\pi}\ln\left(\frac{t - a}{t + a}\right)\right] \qquad (23)$$

where a is the abscissa in the z-plane at which the circular arc beneath the wheel intersects the straight line boundaries, and K is a constant such that the maximum penetration of the wheel δ into the supporting medium is $a \cot K$.

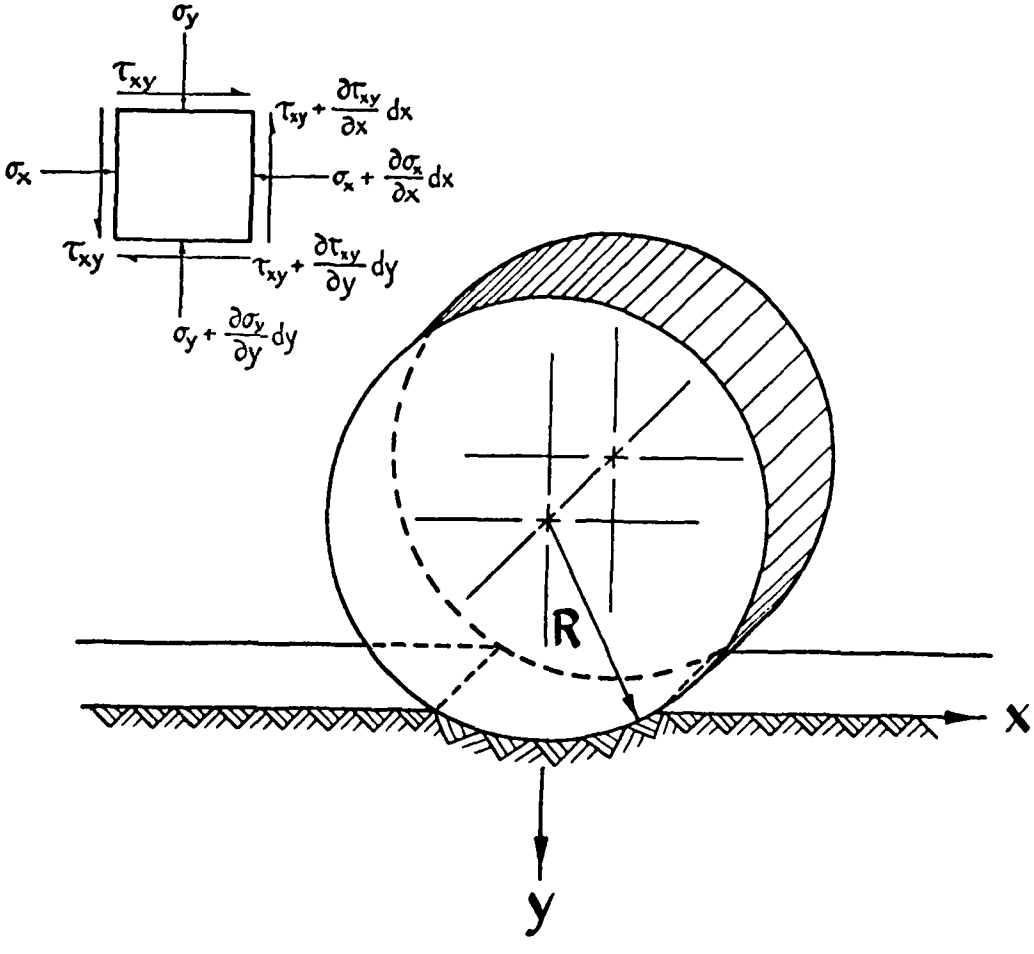


Fig. 1. Schematic wheel and stress at a point within the supporting medium

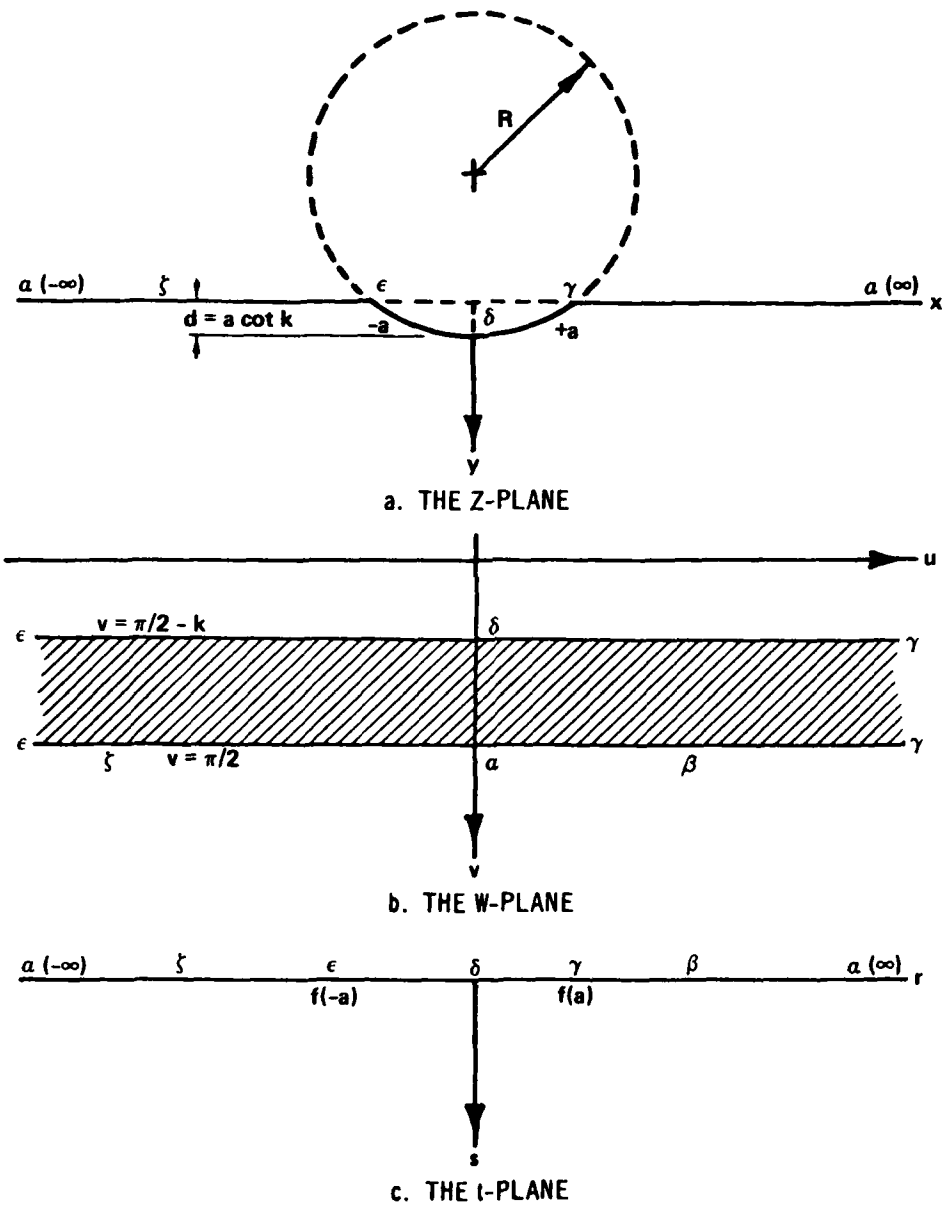


Fig. 2. Transformation of the half-space with the circular wheel removed onto the half-space

IV. APPLICATION OF THE CAUCHY INTEGRAL FORMULA.[9] The Cauchy integral formula states that if $f(h)$ is an analytic function within and on a closed contour C of simply connected region R, and if point ζ is interior to C, then

$$f(\zeta) = \frac{1}{2\pi i} \oint_C \frac{f(h)\, dh}{h - \zeta} \tag{24}$$

Equation 24 implies that the value of a function that is analytic within a region is completely determined throughout the region if the value of the function is known on the boundary.

The Cauchy integral formula was applied to equations 20 and 21 by Gilbert and Al-Hussaini[7] for determining $\Phi(t)$, $\Phi'(t)$, and $\Psi(t)$ and the summary of the results are presented herein

$$\Phi(t) = \frac{N - iT}{2\pi i} \ln\left(\frac{t - a}{t + a}\right) \tag{25a}$$

$$\Phi'(t) = \frac{N - iT}{\pi i} \left(\frac{a}{t^2 - a^2}\right) \tag{25b}$$

$$\Psi(t) = \frac{1}{f'(t)} \left[G'(t) \cdot \frac{T}{\pi} \ln\left(\frac{t - a}{t + a}\right) - G(t)\, \Phi'(t)\right] \tag{25c}$$

where

$$G(t) = a\left(\frac{e^{2w_1} - 1}{e^{2w_1} + 1}\right)$$

$$G'(t) = \frac{2a^2 k}{\pi\left(t^2 - a^2\right)} \frac{4e^{2w_1}}{\left(e^{2w_1} + 1\right)^2}$$

$$w_1 = \frac{K}{\pi - K} \operatorname{Re}(w) + i \ln(w)$$

By knowing the functions necessary to define $\Phi(t)$, $\Phi'(t)$, and $\Psi(t)$, the solution of the problem is considered complete; the final equations for determining stresses and deformations are

$$\sigma_x = 2 \text{ Re } \Phi(t) - \text{Re}\left[\overline{f(t)} \quad \frac{\Phi'(t)}{f'(t)} + \Psi(t)\right] \tag{26a}$$

$$\sigma_y = 2 \text{ Re } \Phi(t) + \text{Re}\left[\overline{f(t)} \quad \frac{\Phi'(t)}{f'(t)} + \Psi(t)\right] \tag{26b}$$

$$\tau_{xy} = \text{Im}\left[\overline{f(t)} \quad \frac{\Phi'(t)}{f'(t)} + \Psi(t)\right] \tag{26c}$$

$$v_x = \frac{3 - 4\nu}{2G} \text{ Re } \phi(t) - \frac{1}{2G} \text{Re}\left[f(t) \quad \overline{\Phi(t)} + \overline{\psi(t)}\right] \tag{26d}$$

$$v_y = \frac{3 - 4\nu}{2G} \text{ Im } \phi(t) - \frac{1}{2G} \text{Im}\left[f(t) \quad \overline{\Phi(t)} + \overline{\psi(t)}\right] \tag{26e}$$

V. NUMERICAL EXAMPLE. A numerical example is presented to illustrate how the stress distribution and displacements underneath a wheel change with wheel penetrations. The numerical evaluation of the stresses and displacement is accomplished by a digital computer since manual evaluation would be virtually impossible.

In this example a 30-in.-diam wheel is considered to apply radial and tangential stresses of 20 and 10 psi, respectively, to the subgrade. The wheel is considered to penetrate the medium to depths of 0.5, 1.5, and 3 in. The results are presented graphically in figs. 3 and 4. Figure 3 shows how stress concentrations at a depth of 6.5 in. below the surface of the medium increase with wheel penetration. This depth was arbitrarily chosen for the purpose of illustration, but with this solution, stresses can be evaluated anywhere in the medium for any penetration or load. Figure 4 presents the horizontal and vertical displacements versus range.

VI. SUMMARY AND CONCLUSIONS. A closed-form solution for evaluating stresses and displacements within a semi-infinite mass whose upper boundary contains a circular indentation was derived. The solution is general enough to consider every condition from a case where no indentation occurs (i.e. straight boundary) to a case where a complete circular hole is formed directly under the surface. The loads applied at the surface are assumed to be uniform. This is a somewhat simplified case, but the solution is very flexible and is easily modified to incorporate any surface loading which can be described by an integrable function.

The general solution developed is believed to provide a tool for

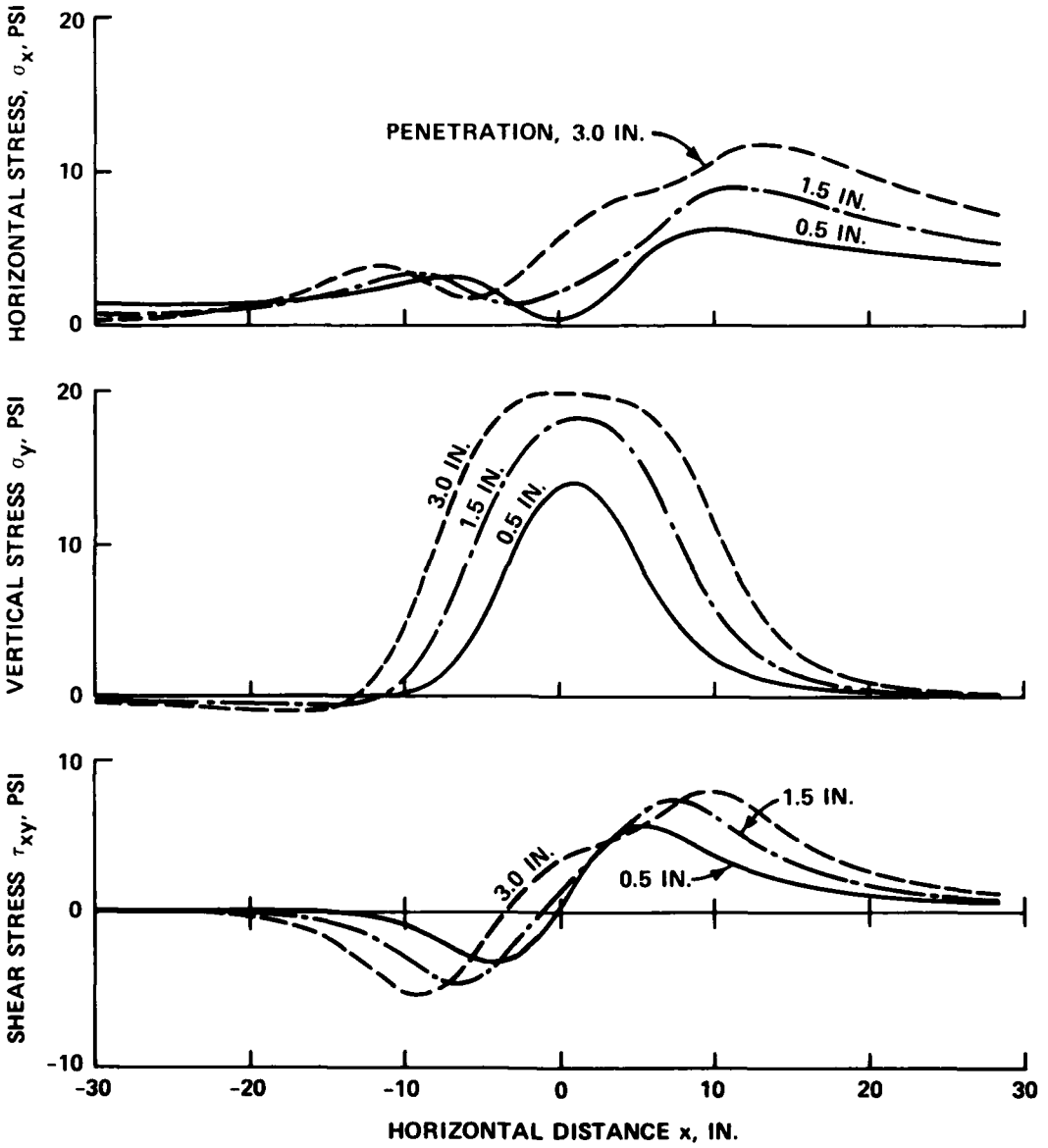


Fig. 3. Stress distribution at a depth of 6.0 in. in the subgrade

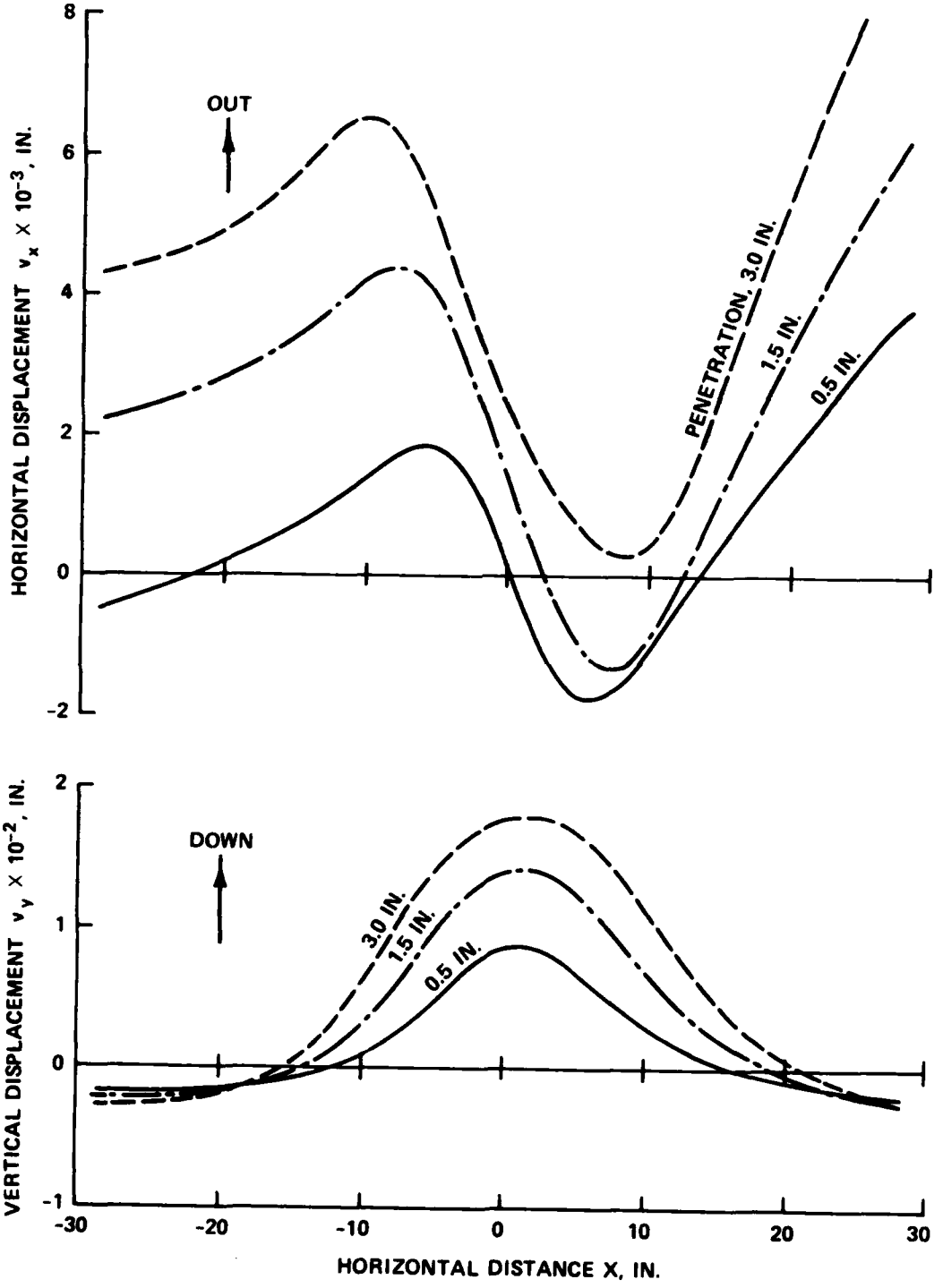


Fig. 4. Horizontal and vertical displacements versus range 6.0-in. depth

solving a variety of plane strain problems with different boundary and loading conditions. Soil-wheel interaction problems can be considered one such example.

VII. ACKNOWLEDGMENT. A portion of this work was accomplished at the U. S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi. Permission to use this information was granted by the Director, U. S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

VIII. REFERENCES. Raised numerals in text refer to the following:

1. Bekker, M. G., Off the Road Locomotion, University of Michigan Press, Ann Arbor, 1960.

2. Green, J. E., and Knight, S. J., "Preliminary Study of Stresses Under Off-Road Vehicle," Miscellaneous Paper No. 4-362, Oct 1959, U. S. Army Engineer Waterways Experiment Station, Vicksburg, Miss.

3. Onafeko, O., and Reece, A. R., "Soil Stresses and Deformations Beneath Rigid Wheels," Journal of Terramechanics, Vol 4, No. 1, 1967, pp. 59-80.

4. Wong, J. Y., and Reece, A. R., "Prediction of Rigid Wheel Performance Based on the Analysis of Soil-Wheel Stresses," Journal of Terramechanics, Vol. 4, No. 1, 1967, pp. 81-98.

5. Perumpral, J. V., Liljedhal, J. B., and Perloff, W. H., "A Numerical Method for Predicting the Stress Distribution and Soil Deformation Under a Tractor Wheel," Journal of Terramechanics, Vol. 8, No. 1, 1971, pp. 9-22.

6. Yong, R. N., Fattah, A. E., and Boonsinsuk, P., "Analysis and Prediction of Tire-Soil Interaction and Performance Using Finite Element," Journal of Terramechanics, Vol. 15, No. 1, 1978, pp 43-63.

7. Gilbert, P. A., and Al-Hussaini, M., "Stresses on a Semi-Infinite Mass Beneath a Loaded Circular Indentation," Transactions, Nineteen Conference of Army Mathematicians, 1973, pp. 739-762.

8. Timoshenko, S., and Goodier, N. J., Theory of Elasticity, 3rd ed., McGraw-Hill, New York, 1962, pp. 168-179.

9. Churchill, R. V., Complex Variables and Applications, 2nd ed., McGraw-Hill, New York, 1960.

ON THE LIMITATIONS AND IMPROVEMENT OF PRESENT NUMERICAL WEATHER PREDICTION*

H. Baussus von Luetzow
U.S. Army Engineer Topographic Laboratories
Fort Belvoir, VA 22060

ABSTRACT. The paper discusses medium and long-range limitations of the present numerical weather prediction model. It shows that hydrostatic long-range forecasts may be obtained by a system of two prognostic and two diagnostic differential equations with implicit parameterization of external energy sources. It further presents a system of equations incorporating mesoscale convection and capable of generating improved medium-range forecasts.

I. INTRODUCTION. There have been considerable theoretical and practical advances in numerical weather prediction since the publication of Charney's [1951] article "Dynamic Forecasting by Numerical Prediction." Although an overview of pertinent research and experimentation is beyond the scope of this paper, the interested reader may profitably consult Phillips' [1960] article, "Numerical Weather Prediction," Thompson's [1961] book "Numerical Weather Analysis and Prediction," Smagorinsky's [1963] article "General Circulation Experiments with the Primitive Equations," the book "Lectures on Numerical Short-Range Weather Prediction" published by Hydrometeoizdat [1969] for the World Meteorological Organization, the WMO [1965] publication on "Research and Development Aspects of Long-Range Forecasting," the book "General Circulation Models of the Atmosphere" published by Academic Press [1977], and Shuman's [1978] article "Numerical Weather Prediction." It should be further mentioned that 500mb numerical routine forecasts in the United States based on the barotropic vorticity equation were started by the Joint Numerical Weather Prediction Unit (now National Meterological Center) in 1957. These forecasts were improved by a three-level model developed by Cressman [1963]. A further improvement resulted from the introduction of a multi-level primitive equation model described by Shuman [1965]. Progress in forecast skill has essentially been achieved by an increasingly better three-dimensional data base as initial data, utilization of more sophisticated models under inclusion of humidity and related effects, and higher grid resolution with an associated greater computer capacity.

This paper describes in section II the primitive equation system in the x, y, p, t-system as the basis for the derivation of filter equations. Section III is concerned with some relevant aspects of filtering, presents

* The research presented in this paper, originally sponsored by the author's organization, was not performed as part of presently assigned duties.

the vorticity equation and corresponding divergence equation, and establishes a diagnostic and prognostic filter equation. It further includes a comparison of a prediction system consisting of two prognostic and two diagnostic equations, supplemented by the differential equation for ground pressure, with the primitive equation system. Finally, it addresses some ramifications of the new signal generation process with respect to hydrostatic numerical weather prediction. Section IV contains a short overview about initialization under consideration of an optimal omega equation and the diagnostic filter equation derived in section III. In section V, the necessity of a non-hydrostatic forecast system for more accurate and long-range weather prediction is pointed out. It is supplemented by a mathematical appendix following section VI, conclusion. The emphasis in this paper has been on the clarification of the filter process, on the consideration of pertinent experience gained by others, the necessity of utilizing better diagnostic equations for initialization, the value of the primary diagnostic equation with respect to the upper boundary, and the identification of mathematical problems and their solutions.

II. THE PRIMITIVE EQUATION SYSTEM. The primitive equations in the (x, y, p, t)-system with x as the coordinate toward the east, y as the coordinate toward the north, p as the vertical pressure coordinate, and t as elapsed time are

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \omega\frac{\partial u}{\partial p} = -\frac{\partial \phi}{\partial x} + fv \tag{1}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \omega\frac{\partial v}{\partial p} = -\frac{\partial \phi}{\partial y} - fu \tag{2}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} = 0 \tag{3}$$

$$\frac{\partial}{\partial t}\frac{\partial \phi}{\partial p} + u\frac{\partial}{\partial x}\frac{\partial \phi}{\partial p} + v\frac{\partial}{\partial y}\frac{\partial \phi}{\partial p} + \sigma\omega = -\frac{R}{c_p p}\frac{dq}{dt} \tag{4}$$

$$\frac{\partial r}{\partial t} + u\frac{\partial r}{\partial x} + v\frac{\partial r}{\partial y} + \omega\frac{\partial r}{\partial p} = \delta \cdot r\,\frac{d\ln r_s}{dt} \tag{5}$$

$$\frac{\partial p}{\partial t} + u\frac{\partial p^*}{\partial x} + v^*\frac{\partial p^*}{\partial y} + \rho^* V^* \cdot \nabla\phi^* = 0 \tag{6}$$

The above differential equations are in sequence the horizontal equations of motion, the continuity equation, the thermodynamic equation, the continuity equation of the mixing ratio, and the equation for surface pressure. As to symbols, $u = \dot{x}$, $v = \dot{y}$, $\omega = \dot{p}$[1] are the horizontal and generalized vertical velocity components, respectively, ϕ is the geopotential, f is the Coriolis parameter, δ is a measure of static stability, R is the gas constant for unsaturated air, c_p is the coefficient of specific heat at constant pressure, r is the mixing ratio of the mass of water vapor to the mass of dry air, $\frac{dq}{dt}$ is the diabatic rate of heat added to a unit mass of air, δ a dimensionless parameter between zero and one, r is the saturation mixing ratio, and p^*, ϕ^*, ρ^* are pressure, geopotential, and air density at the earth's surface, respectively. Explicitly, it is $f = 2\Omega \sin\phi$ with Ω as the angular speed of the earth's rotation and ϕ as geographic latitude.

Of further interest is the hydrostatic equation

$$\frac{\partial \phi}{\partial p} = -\frac{1}{\rho} = -\frac{RT}{p} \tag{7}$$

where ρ and T denote air density and absolute temperature, respectively.

The static stability is explicitly

$$\delta = \frac{\partial^2 \phi}{\partial p^2} + \frac{1}{\kappa p}\frac{\partial \phi}{\partial p} \tag{8}$$

where $\kappa = \frac{c_p}{c_v}$ is the ratio of coefficients of specific heat at constant pressure and constant volume, respectively. In the case of condensation of water vapor it has to be replaced by the effective static stability

$$H_2 = \sigma - \delta \frac{\partial \phi}{\partial p}\frac{\Gamma(r_s, T)}{p} \tag{9}$$

where Γ is a function of r_s and T.

In hydrostatic generation processes σ and H_2 are required to be positive.

1) It is $\omega \approx -\rho g w$ for the computation of the vertical velocity w. In this respect, g is the upward component of the apparent gravitational acceleration.

For the purpose of simplicity, frictional terms in equations (1) and (2), applicable to a boundary layer, and subgrid-scale diffusion terms have been omitted. For the same reason, the decomposition of $\frac{dq}{dt}$ in eq. (4) into several terms and their computation is not discussed here. In global applications, equations (1) - (6) require a formulation in spherical coordinates.

The numerical integration of the primitive equations by finite difference methods is straightforward. At the upper and lower boundaries, the kinematic conditions are $\omega = 0$ for $p = 0$ and ω^* for $p = p^*$, respectively. Rigid and otherwise "closed" lateral boundaries require utilization of a reflection technique described by Hinkelmann [1965]. This technique may also be applied in connection with the determination of vertical derivatives of the various field variables at the lowest and highest generation level. Although the hydrostatic equation filters out sound waves, the primitive equations generate gravity-inertia waves. Accordingly, time increments $\Delta_t \sim 10$ min. must be used in numerical integration. At a specific grid point, the time integration is generally performed in the form

$$F(t) = F(t-2\Delta t) + 2\Delta t\left(\frac{\partial F}{\partial t}\right)_{t-\Delta t} \tag{10}$$

It is of significance to emphasize that the hydrostatic system (1) - (6) is not a strictly deterministic one. It is essentially restricted to a grid resolution with $\Delta x = \Delta y > 50$km and to a limited number of vertical levels. The system is further subject to hydrodynamic and correlated hydrostatic stability. According to Holloway and Manabe [1971] one of the most serious difficulties in designing a numerical model of the general circulation is in the parameterization of moist convection. As a consequence of the occurrence of dry and moist convection under pronounced baroclinic conditions, equations (1) - (6) have to be supplemented by convection adjustment schemes. The generalized vertical velocity ω and the divergence $-\frac{\partial \omega}{\partial p}$ have, therefore, a representative character. In order to eliminate computational noise, in part due to aliasing effects caused by discrete numerical integration of quadratic terms, the U.S. National Weather Service [1978] uses a 25-point smoother. The reduction of aliasing effects by spectral methods has been discussed by Bourke et al [1977]. It should finally be mentioned that eq. (6) of the hydrostatic system can be eliminated by the introduction of the normalized pressure coordinate $s = \frac{p}{p^*}$, first proposed by Phillips [1956]. This leads to a modification of the presented system, in particular of equations (1) - (3), and the resulting continuity equation becomes a prognostic one.

III. FILTERED HYDROSTATIC PREDICTION SYSTEM. Equations (1) and (2) can be transformed by applying the two-dimensional curl and divergence operators on the corresponding vector equation of horizontal motion:

$$\frac{d}{dt}(f+\zeta) = (f+\zeta)\frac{\partial \omega}{\partial p} + \frac{\partial \omega}{\partial y}\frac{\partial u}{\partial p} - \frac{\partial \omega}{\partial x}\frac{\partial v}{\partial p} \tag{11}$$

$$\frac{d}{dt}\,\mathrm{div}\,V + \frac{1}{2}[(\mathrm{div}\,V)^2 + (\mathrm{def}\,V)^2 - \zeta^2] + \nabla\omega \cdot \frac{\partial V}{\partial p} + \beta u = f\left(\zeta - \frac{1}{f}\Delta^2\phi\right) \tag{12}$$

In these equations, $\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$ is the vorticity, $\chi = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$ is the divergence, V is the horizontal velocity vector, and

$$(\mathrm{def}\,V)^2 = \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2 \tag{13}$$

The hydrostatic approximation implies a finite scale and quasi-horizontal motions with filtered or smoothed wind components $\hat{u}$ and $\hat{v}$ which permit the decomposition

$$\hat{u}_1 = -\frac{\partial \psi}{\partial y},\quad \hat{v}_1 = \frac{\partial \psi}{\partial x},\quad \hat{u}_2 = \frac{\partial \chi}{\partial x},\quad \hat{v}_2 = \frac{\partial \chi}{\partial y} \tag{14}$$

Further,

$$\frac{\partial \hat{v}}{\partial x} - \frac{\partial \hat{u}}{\partial y} = \Delta^2\psi,\quad \frac{\partial \hat{u}}{\partial x} + \frac{\partial \hat{v}}{\partial x} = \Delta^2\chi = -\frac{\partial \hat{\omega}}{\partial p} \tag{15}$$

The vorticity equation is then written

$$\frac{\partial}{\partial t}\Delta^2\psi + \hat{u}\frac{\partial \Delta^2\psi}{\partial t} + \hat{v}\frac{\partial \Delta^2\psi}{\partial y} + \hat{\omega}\frac{\partial \Delta^2\psi}{\partial p} + \beta\hat{v} = (f+\Delta^2\psi)\frac{\partial \hat{\omega}}{\partial p} + \frac{\partial \hat{\omega}}{\partial y}\frac{\partial \hat{u}}{\partial p} - \frac{\partial \hat{\omega}}{\partial x}\frac{\partial \hat{v}}{\partial p} \tag{16}$$

The divergence theorem is

$$\frac{\partial}{\partial t}\Delta^2\chi + \hat{u}\frac{\partial\Delta^2\chi}{\partial x} + \hat{v}\frac{\partial\Delta^2\chi}{\partial y} + \hat{\omega}\frac{\partial\Delta^2\chi}{\partial p} + \frac{1}{2}\left[(\Delta^2\chi)^2 + (\mathrm{def}\,\hat{V})^2 - (\Delta^2\psi)^2\right] +$$

$$+\frac{\partial\hat{\omega}}{\partial x}\frac{\partial\hat{u}}{\partial p} + \frac{\partial\hat{\omega}}{\partial y}\frac{\partial\hat{v}}{\partial p} + \beta\hat{u} = f\left(\Delta^2\psi - \frac{1}{f}\Delta^2\phi\right) \qquad (17)$$

The continuity equation assumes the form

$$\frac{\partial\hat{\omega}}{\partial p} = -\Delta^2\chi \qquad (18)$$

The pressure tendency equation can be stated as

$$\frac{\partial\hat{p}_s}{\partial t} + \hat{V}^*\cdot\nabla p^* + \hat{\rho}^*\hat{V}^*\cdot\nabla\hat{\phi}^* = 0 \qquad (19)$$

The signal character of the prognostic equations (16) and (17) is evident from the terms $\frac{\partial}{\partial t}\Delta^2\psi$ and $\frac{\partial}{\partial t}\Delta^2\chi$ since solutions of two Poisson equations are required to determine the time increments $\delta\psi$ and $\delta\chi$. In fact, vertical-transverse gravity-inertia waves are expected to be filtered out, and their inclusion or occurrence would be incompatible with equations (16)-(19). Consequently, the geopotential ϕ in eq. (17) must also be a filtered or smoothed variable, and the development of hydrodynamic instability must be eliminated. Therefore, an external, non-selective smoother and a convective adjustment must be applied with respect to the thermodynamic equation (4), or an internal, selective, and consistent filter must be established which applies simultaneously to $\hat{u}$, $\hat{v}$, and $\hat{\phi} = \hat{\phi}$, i.e., to three filtered variables.

In order to derive appropriate filter equation, the operator $\frac{d}{dt}$ is applied to the horizontal equations of motion, with the result

$$\frac{d^2\hat{V}}{dt^2} = -f^2\hat{V} + f\vec{k}\times\nabla\phi - \frac{d}{dt}\nabla\phi \qquad (20)$$

where $\vec{k}$ is the unit vector directed upward. Application of the horizontal curl operator $\nabla\times(\,)$ pertaining to eq. (20) yields

$$\nabla\times\frac{d^2\hat{V}_1}{dt^2} + \nabla\times\frac{d^2\nabla\chi}{dt^2} = \nabla\times\left[-f^2\hat{V} + f\vec{k}\times\nabla\phi - \frac{d}{dt}\nabla\phi\right] \qquad (21)$$

where V_1 is the non-divergent velocity vector. The operator $\nabla \times \frac{d}{dt}\,(\,)$ eliminates $\frac{\partial^2}{\partial t^2}\,\nabla \times$ associated with the second term on the left of eq. (21). Its non-deterministic or filter version, which also requires $\nabla \times \frac{d^2 V_1}{dt^2} = 0$, is explicitly

$$\left(f + \frac{\partial \hat{u}}{\partial y}\right)\frac{\partial^2 \phi}{\partial x^2} + \left(\frac{\partial \hat{v}}{\partial y} - \frac{\partial \hat{u}}{\partial x}\right)\frac{\partial^2 \phi}{\partial x \partial y} + \left(f - \frac{\partial \hat{v}}{\partial x}\right)\frac{\partial^2 \phi}{\partial y^2} + \frac{\partial \hat{\omega}}{\partial y}\frac{\partial^2 \phi}{\partial x \partial p} - \frac{\partial \hat{\omega}}{\partial x}\frac{\partial^2 \phi}{\partial y \partial p} +$$

$$+ \beta\left(\frac{\partial \phi}{\partial y} + 2f\hat{u}\right) - f^2\,\Delta^2 \psi = 0 \qquad (22)$$

Somewhat stronger filtering of gravity-inertia waves results if the divergence operator $\nabla \cdot (\,)$ is applied to eq. (20) and $\nabla \cdot \frac{d^2 V_2}{dt^2}$ is neglected. Accordingly, the resulting filter equation

$$\frac{d}{dt}\,\Delta^2 \phi = (f^2 + \Delta^2 \phi)\,\frac{\partial \hat{\omega}}{\partial p} - \left(\frac{\partial \hat{\omega}}{\partial x}\frac{\partial^2 \phi}{\partial x \partial p} + \frac{\partial \hat{\omega}}{\partial y}\frac{\partial^2 \phi}{\partial y \partial p}\right) +$$

$$+ \frac{\partial \hat{v}}{\partial y}\frac{\partial^2 \phi}{\partial x^2} - \left(\frac{\partial \hat{v}}{\partial x} + \frac{\partial \hat{u}}{\partial y}\right)\frac{\partial^2 \phi}{\partial x \partial y} + \frac{\partial \hat{u}}{\partial x}\frac{\partial^2 \phi}{\partial y^2} + \beta\left(\frac{\partial \phi}{\partial x} - 2f\hat{v}\right) \qquad (23)$$

has a prognostic character. Elimination of the material derivative $\frac{\partial}{\partial t}\,\Delta^2 \phi$ from eq. (23) by means of the thermodynamic equation makes it possible to establish the required $\hat{\omega}$-equation for initialization. Hyperbolicity is a consequence of the non-identity of the hydrostatic signal generation process with the more complex non-hydrostatic process and the substitution of an unfiltered ϕ-field for a smoothed ϕ-field.

It is of further significance that eq. (22) has the "horizontal" ellipticity criterion

$$\left(f + \frac{\partial \hat{u}}{\partial y}\right)\left(f - \frac{\partial \hat{v}}{\partial x}\right) - \frac{1}{4}\left(\frac{\partial \hat{v}}{\partial y} - \frac{\partial \hat{u}}{\partial x}\right)^2 > 0 \qquad (24)$$

due to the absence of a term involving $\frac{\partial^2 \phi}{\partial p^2}$.

The diagnostic equation (22) represents the adjustment process on the signal scale. It replaces the thermodynamic equation, acts as an immediate selective filter with respect to eq. (17), prevents the development of hydrodynamic instability, and permits an integration time increment of about one hour.

Derived and suitable for numerical weather prediction, equations (22) and (24) and another stability criterion presented in section IV are related to van Mieghem's [1951] work on hydrodynamic instability.[2)]

Filter equations derived by other authors did not result in the establishment of a signal generation process essentially equivalent to the primitive equation system. Thompson [1961] prepared a system consisting of the vorticity equation, the socalled balance equation exhibited in section IV, and a correlated omega equation, i.e., one prognostic and two diagnostic equations. Fjortoft [1962] established a system of two prognostic filter equations and the thermodynamic equation, to be solved for u, v, and ω at each time step by the method of under-relaxation. As a filter condition, he chose $\frac{d^2V}{dt^2} = 0$. This new operator was also employed by Hollmann [1966] in his investigations of new diagnostic relations between wind and pressure in a barotropic atmosphere with divergent flow. Fjortoft's system was tested by Fruehwald [1968] who emphasized its superiority over the classical balance equation, particularly under anticyclonic conditions. Hinkelmann [1969] recommended the application of diagnostic equations resulting from Fjortoft's new filter condition for initialization of the primitive equations. Herbert [1971] studied static and quasi-static motions in a compressible though isothermal atmosphere, where $p \cdot \rho^{-\kappa} = \text{const.}$, under filter assumptions $(\frac{d}{dt})^n \text{ div } V = 0$ and non-consideration of y-dependence. Although his results are not directly applicable to realistic initialization, they show the shortcoming of the hydrostatic filter system pertaining to short waves.

In the filtered hydrostatic prediction system, diabatic effects, represented by $\frac{dq}{dt}$ in the thermodynamic equation, are implicitly parameterized. Accordingly, the differential equation (5) for the mixing ratio r has to be solved separately with $\hat{u}$, $\hat{v}$, $\hat{\omega}$-information obtained as a solution of the signal generation process. This is fully consistent with the finding of Smagorinsky et al [1970] that "a three-dimensional specification of the mass field in the extra-tropics, or preferably the horizontal wind field, will determine the vertical velocity distribution and the humidity distribution." According to these authors, an effective initialization is a prerequisite for forecasts beyond about 5 days. As mentioned above, improved initialization is the subject of section V. Equations (16), (17), and (22) with $\hat{\omega} \equiv 0$ are useful for providing a dynamic boundary condition in the stratosphere which is generally strongly stratified and in approximate radiative equilibrium. Further, the sufficiency of relatively few levels and the non-necessity of

---

[2)] The signal generation process was presented by the author in the context of weather predictability at the IUGG XVI General Assembly, Grenoble, France, August 25-September 6, 1975, was, however, not published.

extreme horizontal grid resolution for hydrostatic forecasts becomes apparent which is in agreement with findings by the National Weather Service [1978].[3] In agreement with the statement by Haltiner and Martin [1957] that internal predictors of the branches of the general circulation do exist, the signal generation process identifies these predictors as $\hat{u}$ and $\hat{v}$. However, this process is still non-stationary in nature, and stationary statistics employing ϕ as a predictor will, therefore, have no comparable success. Inherent limitations of both the primitive equation or message generation process and the filtered equation system can only be overcome by a non-hydrostatic generation process outlined in section V.

The desirability of a selective filter to damp inertia-gravity waves in numerical prediction with the primitive equations has been stressed by several authors. According to Morel and Talagrand [1974], "inertia-gravity waves are indeed generated in the real atmosphere by orographic obstacles, mesoscale disturbances, strong cumulus convection, local heat sources, but the fact is that there is normally very little energy in these modes so that the atmospheric flow appears to be always very close to geostrophic balance." In marked contrast to this situation, numerical general circulation models based on finite difference or otherwise truncated (in Fourier space) versions of the same governing equations, are very sensitive to local perturbations and are likely to sustain an inordinate amount of inertia-gravity waves which show up as high spatial frequency "noise" in the computed flow pattern. Hence, there must exist selective damping processes which operate in the real atmosphere to restore the geostropic balance, and yet do not operate in numerical models" emphasis added). In order to damp the generation of strong divergences and to avoid unnecessary strong damping with respect to vorticies, the above authors recommend the inclusion of a damping term in the equations of motion:

$$\frac{\partial V}{\partial t} + \ldots\ldots = K \nabla (\text{div } V) \tag{25}$$

Similarly, Dey [1978] proposed to introduce a damping term in eq. (12):

$$\frac{\partial}{\partial t} \text{div } V + \ldots\ldots = - \mu \Delta^2 \text{ div } V \tag{26}$$

The solution of the diagnostic equation (22) can be reduced to a two-dimensional problem by computing a preliminary solution ϕ_1 by means of the prognostic equation (23) which would suffice for short-range forecasts. The terms involving $\frac{\partial^2 \phi}{\partial x \partial p}$, $\frac{\partial^2 \phi}{\partial y \partial p}$, $\frac{\partial \phi}{\partial y}$ would thus be known. Hereafter, eq. (22)

[3] Little improvement in 48-hour forecasts was obtained by improving the seven-level horizontal resolution form 174 to 87km or by improving the Nested Grid Model from 198 to 99km.

reduces to

$$A \frac{\partial^2 \phi}{\partial x^2} + 2B \frac{\partial^2 \phi}{\partial x \partial y} + C \frac{\partial^2 \phi}{\partial y^2} = F(x,y) \tag{27}$$

for each isobaric level. With the geopotential ϕ available at the upper p_u by virtue of a dynamic boundary condition and at the earth's surface p^*, and with ϕ_l given everywhere laterally, eq. (27) may be solved for suitable small domains. If these remain fixed for routine forecasts, eq. (27) may be formulated in matrix form as

$$M_{i_k} \phi_k = F_i + f_i \tag{28}$$

with the solution

$$\phi_k = (M_{i_k})^{-1} (F_i + f_i) \tag{29}$$

where the terms f_i arise from the utilization of lateral boundary values ϕ_{l_B} Alternatively, appropriate iteration and relaxation algorithms may be developed.

IV. IMPROVED INITIALIZATION. A filter equation less effective than the diagnostic eq. (22) results from the condition

$$\frac{d}{dt} \operatorname{div} V = 0 \tag{30}$$

in eq. (12). Divergence production is thus eliminated, and the truncated part of eq. (12) can prognostically only be used in conjunction with the vorticity equation (16) and a diagnostic ω-equation compatible with the truncated part, also called balance equation. Such a system was developed and advocated by Thompson [1961]. The generally applied special balance equation including only the stream function ψ and the geopotential ϕ reads

$$f \Delta^2 \psi + 2 \left[\frac{\partial^2 \psi}{\partial x^2} \frac{\partial^2 \psi}{\partial y^2} - \left(\frac{\partial^2 \psi}{\partial x \partial y} \right)^2 \right] + \beta \frac{\partial \psi}{\partial y} = \Delta^2 \phi \tag{31}$$

This equation, which does not provide for an intricate coupling between the variables ψ and ϕ or rather between ψ, χ, and ϕ, has been employed by many authors including Bolin [1956], Döös [1965], Knighting [1965], Gerrity and Chu [1978].

Miyakoda and Moyer [1968] developed a technique of solving the classical balance equation and, implicitly, its associated ω-equation using Euler-backward time differencing which, when tentatively used, filters out the high-frequency modes. In their application to the non-linear barotropic equations, they imposed the conditions $(\frac{\partial \chi}{\partial t})_{t_o} = 0$ and $(\frac{\partial^2 \chi}{\partial t^2})_{t_o} = 0$, where χ is the velocity potential. Nitta and Hovermale [1967] presented an improved dynamic initialization scheme free of the above constraints on the horizontal divergence by an actual iteration of forward and backward forecasts around the initial time with the Euler-backward time difference. They applied their scheme to the equations of motion which govern an incompressible homogeneous atmosphere over a rotating flat domain. From their numerical experiment they concluded that their method was unable to reproduce the amplitude of the divergent components. In conjunction herewith, convergence of their process was discouragingly slow. Nevertheless, the authors emphasized that the attainment of a balanced state between mass and velocity fields is a basic prerequisite of initial data for forecasts with the primitive equations and in addition one of the basic goals in objective analysis. In his analysis on the adjustment toward balance in primitive equation weather prediction models, Okland [1970] concluded that one cannot expect to abtain a good balance for a multilevel model by dynamic balancing since the balanced state is relatively more transient under baroclinic conditions.

The above experience is consistent with the fact that the primitive equations are not synonymous with filter equations. In fact, an optimal ω-equation is obtained by adaptation of the prognostic filter equation (23) to the thermodynamic equation (4). It is then with H_2 from eq. (9)

$$H_2 \; \Delta^2\omega + (f^2 + \Delta^2\phi)\frac{\partial^2\omega}{\partial p^2} - \frac{\partial^2\phi}{\partial x \partial p}\frac{\partial^2\omega}{\partial x \partial y} - \frac{\partial^2\phi}{\partial y \partial p}\frac{\partial^2\omega}{\partial y \partial p} = F(x,y,p) \tag{32}$$

$$H_2 > 0, f^2 + \Delta^2\phi > 0, \sigma(f^2 + \Delta^2\phi) - \frac{1}{4}\left[\left(\frac{\partial^2\phi}{\partial x \partial p}\right)^2 + \left(\frac{\partial^2\phi}{\partial y \partial p}\right)^2\right] > 0 \tag{33}$$

The real ϕ-field will, in general, satisfy the filtered $\hat{\phi}$-field which is subject to the inequalities (33). Otherwise, suitable artificial corrections have to be made. In order to solve eq. (32) it is first necessary to determine a stream function $\psi^{(1)}$ from eq. (22) with $\hat{u}_2 = \hat{v}_2 = \hat{\omega} = 0$. A first approximation $\omega^{(1)}$ may then be obtained by setting $\Delta^2\omega = - \lambda\omega$ where λ has the character of a regression coefficient, in agreement with a suggestion by Eliassen and Kleinschmidt [1957]. Thereafter eq. (32) is suitably formulated as

$$(f^2 + \Delta^2\phi) \frac{\partial^2\omega}{\partial p^2} - H_2\lambda\ (p)\omega = F(p) + G(\omega) \qquad (34)$$

Following determination of $\omega^{(1)}$ from eq. (34) with $G(\omega) = 0$ under consideration of $\omega_u = o$ for p_u and ω^* for p^*, to be performed at all applicable grid points, $G(\omega^{(1)})$ can be calculated. In this way, $\omega^{(1)}$may be iteratively improved. Numerical underrelaxation may be indicated to achieve satisfactory convergence. It will probably be impractical to also consider successively improved ψ and χ-values in the ω-solution. Final ω, u_2, v_2-values should, however, be utilized to find an improved stream function ψ by means of eq. (23).

V. NON-HYDROSTATIC WEATHER PREDICTION. As discussed in sections 3 and 4, hydrostatic models suffer from the requirement of hydrostatic and hydrodynamic stability which, in turn, do not permit a reasonably deterministic inclusion of the moist-adiabatic process and a computational grid of high resolution. The non-consideration of mesoscale convection and insufficiently reduced grid-scale diffusion account for a useful hydrostatic forecast time limit [Doos, 1970] of about 10 days.

According to Holloway and Manabe [1971] one of the most serious difficulties in designing a numerical model of the general circulation is in the parametrization of moist convection. In case of a negative static stability they adjust the super dry-adiabatic lapse rate to the dry-adiabatic lapse rate so as to simulate the effects of strong mixing by dry convection in the free atmosphere. The adjustment is performed in such a way that the sum of potential and internal energies is conserved. Their convective scheme in the case of super-moist adiabatic lapse rates is more complicated and requires the solution of 2n simultaneous equations for the determination of n temperature and n mixing ratio corrections under the assumption that the sum of potential, internal and latent energy is conserved during this adjustment. These adjustments implicitly realize the inadequacy of hydrostatic equilibrium in accounting for negative effective static stabilities. Sunquist [1970] reports that various authors, utilizing hydrostatic models in which the rate of heating was set proportional to the vertical velocity, obtained unstable results in a few hours because of a rapid growth on the smallest possible scale of the system. Consequently, he treats the development of tropical cyclones as a forced circulation driven by the heat released in convection cells. His approach is consistent with

the convection adjustments applied by Holloway and Manabe. In order to extend the effective range of numerical weather prediction, Lorenz [1969] proposes to formulate more appropriate equations or more realistic statistical assumptions. He also suggests that one might well obtain a considerably longer range of predictability by including a spectral gap somewhere between the synoptic and cumulus scales and notes that such systems as large cumulus clouds are not randomly distributed throughout the atmosphere, but have a preference for regions containing such meso-scale systems as squall lines and fronts. These in turn are not randomly distributed, but prefer certain locations relative to larger-scale synoptic features. With reference to the prediction of convective clouds and the precipitation falling from them, Monin [1972] distinguishes between free cumulus convection and forced cumulus convection. The former is observed in middle latitudes over areas occupied by cold air masses and determined primarily by the humidity field and the energy of the instability of the lower troposphere. Forced cumulus convection is observed on the lines of horizontal convergence (tropical cyclones); it is determined primarily by the humidity field and the horizontal convergence. Accordingly, forced cumulus convection appears to be more amenable to large scale analysis, i.e., by the (x, y, p, t)-system, which has been verified by Charney and Sunquist. On the other hand, there is a bilateral interaction between large scale and mesoscale processes. Since cumulus convection cells have horizontal diameter on the order of 50-80 km, it would be desirable to use a prediction grid system with a resolution of 10km which would also be adequate for jet stream regions of high intensity. Due to the fact that heat of condensation has an important destabilizing influence which manifests itself immediately in vertical motions, utilization of a high-resolution grid would be most appropriate. Such a grid, consistent with the incorporation of twenty or more levels, would sufficiently cope with fronts, the jet stream, and tropical storms. The Committee on Atmosphere Sciences, National Research Council [1971] has stated that vertical convection and condensation play important parts in mesoscale phenomena and that severe mathematical and theoretical difficulties exist in developing general prediction models for these smaller scales so that the hydrostatic approximation, heretofore essential to general circulation theory, cannot be applied to the smaller scales. In their investigation of the effect of horizontal grid resolution in an atmospheric circulation model, Miyakoda et al [1971] affirmed a former result that the further the integration period is extended the higher is the required resolution. In particular, the detailed structure of fronts become more realistic for a grid size of about 250km as compared with a prior grid size of approximately 500km at mid-latitudes. In this connection it has also been mentioned that the horizontal eddy viscosity coefficient is proportional to the square of a characteristic length related to the grid spacing and to the deformation, the latter becoming more pronounced for smaller scales. Significantly, the higher resolution improves the quasi-stationary or forced mode. The wave amplitude pertaining to the total forecast (forced and free mode), however, was underestimated. This might be due to a still insufficient

resolution and/or to some missing physical process. Evidently, for the lattice spacing under consideration (N=80 or about 125km), use of the hydrostatic prediction system has some dampening effect. The wiggling in the flow field, noted by the authors, became stronger for the N=80 model in comparison with the N=40 model, partly due to the nature of convective adjustment. This is another indication of the shortcomings of the (x, y, p, t)-system. Hence, the requirement for improved equations would become quite stringent for an N>160 model.

Provided that sufficiently dense and accurate initial fields of temperature, ground pressure, humidity, and ocean surface temperature are available, a non-hydrostatic forecast would require the determination of the initial wind field by means of an equilibrium initialization in the (x, y, z, t)-system. Such initialization, always possible on a data scale greater than the minimum hydrostatic scale, may be achieved by a diagnostic initialization or a combination of diagnostic and dynamic methods. Both techniques would, however, depend on a diagnostic equation for the vertical velocity w which, in the hydrostatic case, coincides with the equation of continuity. The diagnostic initialization, including the derivation of w, is shown in the Appendix. Following initialization, a computation grid of higher resolution is employed. In a simplified form, i.e., in non-spherical coordinates and without diffusion terms, the non-hydrostatic system would consist of the familiar equations

$$\frac{du}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + fv \tag{35}$$

$$\frac{dv}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - fu \tag{36}$$

$$-\frac{1}{\rho}\frac{d\rho}{dt} = \text{div } V_3 = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \tag{37}$$

$$F_1 \frac{dT}{dt} = F_2 \text{ div } V_3 + \frac{1}{c_p}\frac{dq}{dt} \tag{38}$$

$$\frac{dr}{dt} = \delta F_3 \text{ div } V_3 \tag{39}$$

$$w = w\,[u,v,\rho,T,r,\phi(x,y)] \tag{40}$$

This system, where V_3 is the 3-dimensional velocity vector, where w is expressed symbolically and where F_1, F_2, and F_3 are functions identifiable in the Appendix, must be supplemented by equations specifying $\frac{dq}{dt}$ as the

rate of addition of heat to a unit mass of air, and an adequate model for the prediction of ocean surface temperature. The system (35)-(40) together with the supplementary equations and under consideration of appropriate boundary conditions, would generate adequate computation grid fields in a few hours. Since it is not restricted by ellipticity conditions and is capable of fully utilizing the initial scalar fields, it exhibits a partial memory on the mesoscale.

It should be pointed out that the non-hydrostatic system discussed above constitutes a mesoscale signal process because of the diagnostic w-equation (40) and its associated finite grid representation. High-frequency gravity-inertia oscillations and those correlated with sound waves are, therefore, filtered out. In comparison, the hydrostatic signal process includes two diagnostic equations because of the requirement of hydrodynamic stability, synonymous with the suppression of convection and the exclusion of mesoscale phenomena.

In contrast to the ω-equation, the w-equation is a second-order differential equation. If the former can accommodate 10 levels, the latter is compatible with the incorporation of 50 levels. A similar resolution increase results in the reduction of the minimal hydrostatic grid length of 50km to a mesoscale grid length of 10km. As a further necessary improvement, δ in eq. (55) of the Appendix should be suitably parameterized, i.e., not only assume the values zero and one, and eq. (61) of the Appendix may have to be modified to the structure

$$f_1(z)\frac{d^2w}{dz^2} + f_2(z)\frac{dw}{dz} + aw = f_3(z) \tag{41}$$

due to the differentiation $\frac{d}{dt} g(x, y, z)$ in the vertical equation of motion. This results in the additional term aw where a is a constant. Because of the complicated structure of f_1 and f_2, a closed solution of eq. (41) is not possible.

VI. CONCLUSION. The development of models for numerical weather prediction is primarily a problem of mathematics and includes computer simulations and comparison with realistic fields. It is not only concerned with the numerical integration of prognostic and diagnostic differential equations. The ramifications of simplifications of the vertical equation of motion, the establishment of associated filter equations, correlated initialization, consideration of a dynamic upper boundary condition, and the reduction of aliasing effects are also significant and have to be taken into account.

The hydrostatic filtering approximation also implies hydrodynamic equilibrium, i.e., quasi-horizontal flow, exclusion of convection, and a correlated finite scale. Consistent with these constraints, it is possible to derive a hydrostatic signal generation process. The vorticity and divergence

equations of the new system are only compatible with a diagnostic filter equation which provides an optimal, non-independent smoothing of both the wind and the geopotential field. The signal process has a non-deterministic structure, and the application of the new generalized filter equation facilitates the elimination of pure gravity waves and correlated high-frequency oscillations while retaining dispersed, lower-frequency waves which move with a group velocity consistent with a finite grid representation. The signal process which embodies the concept of hydrodynamic forecasting implies the need for a restricted number of levels and does not interact with the continuity equation of water vapor. A useful forecast of 10 days with respect to wind and geopotential is thus possible, while humidity forecasts can only be satisfactorily made for several days. The derivation of the signal process equations is intimately related to an optimal diagnostic initialization.

The primitive equations represent a compromise between the hydrostatic filter equation and the remaining, inherently deterministic prognostic differential equations. Following optimal diagnostic initialization, as in the signal process, the message process generates instabilities in the absence of a non-stationary smoothing device, particularly under inclusion of rainfall prediction, so that a complicated convective adjustment is required. Long range primitive forecasts have thus a tendency to diverge.

Weather predictability extension beyond the limits imposed by the hydrostatic approximation can only be accomplished by non-hydrostatic or mesoscale forecasting. The prerequisite for a mesoscale prediction system as a signal process is the derivation of a diagnostic equation for the vertical wind velocity w in (x, y, z, t)-coordinates, effective initialization in a "diagnostic" grid of low resolution under utilization of pressure, temperature, and humidity data, and subsequent integration of the equations in a "prognostic" grid of high resolution. The diagnostic w-equation makes it possible to employ up to about 50 levels and to permit a horizontal grid with $\Delta x = \Delta y = 10$km. Accordingly, the mesoscale system, though necessitating a tremendous computational effort, can be expected to provide useful forecasts for a period of 2-3 weeks.

The determination of global initial wind fields in the hydrostatic and non-hydrostatic systems requires the availability of measured winds in the equatorial region.

VII. ACKNOWLEDGEMENT. The author is indebted to the U.S. Army Engineer Topographic Laboratories for prior support of his research in dynamic meteorology and numerical weather prediction which has provided a basis for this paper which was essentially prepared without official resources. He wishes to thank Ms. Carole Gradick for typing the manuscript after duty hours.

APPENDIX**

Generalized Optimal Filter Equations Free of Hydrostatic Limitations

7. **Generalized Optimal Filter Equations Free of Hydrostatic Limitations.** In Section 6, we have shown that it is necessary to consider pressure, p, a priori as a continuous variable and that pressure kinks have to be smoothed out in order to avoid quasi-infinite pressure gradients in agreement with Haltiner and Martin.[55] Discontinuities of zeroth order involving temperature, or rather virtual temperature, require even stronger smoothing. The application of the differential equations of meteorology is only possible with smoothed variables including consistently filtered winds. For simplicity, we omit the filter symbol $\wedge$ in the following derivations in which V is the 3-dimensional velocity vector.

Since the filter condition

$$\frac{d^2}{dt^2}(\rho V) = 0 \tag{36}$$

implicitly includes a term $\frac{d}{dt}\operatorname{div} V$, a system of non-linear partial differential equations would result which does not permit an equilibrium solution and, consequently, could not be solved by relaxation methods.

As already mentioned (Section 6), the filter condition is

$$\frac{d^2 V}{dt^2} = 0 \tag{37}$$

which has already been applied in Section 4 except for the vertical wind component.

We now apply eq. (37) with reference to the equations of motion

$$\frac{du}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + fv \tag{38}$$

$$\frac{dv}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - fu \tag{39}$$

$$\frac{dw}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g \tag{40}$$

---

[55] G. J. Haltiner and F. L. Martin: *Dynamical and Physical Meteorology*, McGraw-Hill Book Company, New York, Toronto, London, 1957.

** Section 7, *Revised*, of Research Note ETL-RN-71-3, "The Derivation and Potential of New Filter Equations for Numerical Weather Prediction" by H. Baumus von Luetzow, Dec. 1971, AD 741788.

with the intermediate result

$$\frac{d}{dt}\frac{\partial p}{\partial x} - \frac{1}{\rho}\frac{d\rho}{dt} \cdot \frac{\partial p}{\partial x} - \rho \frac{d}{dt}(fv) = 0 \tag{41}$$

$$\frac{d}{dt}\frac{\partial p}{\partial y} - \frac{1}{\rho}\frac{d\rho}{dt} \cdot \frac{\partial p}{\partial y} + \rho \frac{d}{dt}(fu) = 0 \tag{42}$$

$$\frac{d}{dt}\frac{\partial p}{\partial z} - \frac{1}{\rho}\frac{d\rho}{dt} \cdot \frac{\partial p}{\partial z} = 0 \,. \tag{43}$$

With $F = \frac{dp}{dt}$, under consideration of eq. (38) and (39) with regard to $-\frac{d}{dt}(fv)$ in eq. (41) and $\frac{d}{dt}(fv)$ in eq. (42), respectively, and in view of the continuity equation

$$\frac{1}{\rho}\frac{d\rho}{dt} = -\operatorname{div} V \tag{44}$$

we arrive at

$$\frac{\partial F}{\partial x} + \frac{\partial p}{\partial x}\operatorname{div} V - \left(\frac{\partial u}{\partial x}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial x}\frac{\partial p}{\partial y} + \frac{\partial w}{\partial x}\frac{\partial p}{\partial z}\right)$$

$$+ f\rho\left(\frac{1}{\rho}\frac{\partial p}{\partial y} + fu\right) - v\rho\left(u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y}\right) = 0 \tag{45}$$

$$\frac{\partial F}{\partial y} + \frac{\partial p}{\partial y}\operatorname{div} V - \left(\frac{\partial u}{\partial y}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial p}{\partial y} + \frac{\partial w}{\partial y}\frac{\partial p}{\partial z}\right) +$$

$$+ f\rho\left(-\frac{1}{\rho}\frac{\partial p}{\partial x} + fv\right) + u\rho\left(u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y}\right) = 0 \tag{46}$$

$$\frac{\partial F}{\partial z} + \frac{\partial p}{\partial z}\operatorname{div} V - \left(\frac{\partial u}{\partial z}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial z}\frac{\partial p}{\partial y} + \frac{\partial w}{\partial z}\frac{\partial p}{\partial z}\right) = 0 \tag{47}$$

In the next step, we have to express F as a time-independent function which linearly involves the divergence $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$. The thermodynamic equation reads in a very general form

$$d\left(\frac{r_s L}{c_p T}\right) + \left(1 - k + r_s \frac{c}{c_p}\right)\frac{dT}{T} - k\frac{d\rho_L}{\rho_L} = \frac{1}{c_p}\frac{đq}{T} \tag{48}$$

where r_s designates the saturation mixing ratio, $L \approx 600$ cal g^{-1}, $c_p = .2405$ cal g^{-1} $\cdot$(deg C)$^{-1}$ the specific heat of dry air at constant pressure, $c = 1.0$ cal g^{-1}, the specific heat of water, $k = \frac{c_p - c_v}{c_p} = .2848$ where c_v is the specific heat of dry air for constant volume, ρ_L the density of dry air, and $đq$ non-precipitative heat added to a unit mass of air. With $a = \frac{r_s L}{c_p T}$, eq. (48) may also be written in its time-dependent form

$$a \frac{d \ln r_s}{dt} + (1 - k - a + 4.2\, r_s) \frac{d \ln T}{dt} - k \frac{d \ln \rho_L}{dt} = \frac{1}{c_p T} \frac{đq}{dt} . \tag{49}$$

Elimination of the term $\frac{d \ln r_s}{dt}$ in eq. (49) by means of Smagorinsky's and Collin's relation[56]

$$\frac{d \ln r_s}{dt} = (\gamma - 1) \frac{d \ln T}{dt} - \frac{d \ln \rho_L}{dt} \tag{50}$$

with $\gamma = \frac{L}{1.608\, A^* RT}$ which involves L as the latent heat of condensation and A^{*-1} as the mechanical equivalent of heat leads to

$$[1 - k + a\,(\gamma - 2) + 4.2\, r_s] \frac{d \ln T}{dt} - (k + a) \frac{d \ln \rho_L}{dt} = \frac{1}{c_p T} \frac{đq}{dt} . \tag{51}$$

From eq. (50) and (51) follows

$$\frac{d \ln r_s}{dt} = \frac{\gamma k + a - 1 - 4.2\, r_s}{1 - k + a(\gamma - 2) + 4.2\, r_s} \frac{d \ln \rho_L}{dt}$$

$$+ \frac{\gamma - 1}{1 - k + a(\gamma - 2) + 4.2\, r_s} \cdot \frac{1}{c_p T} \frac{đq}{dt} \approx \frac{A}{B} \frac{d \ln \rho_L}{dt} \approx \frac{A}{B} \frac{d \ln \rho}{dt} . \tag{52}$$

Under consideration of

$$\frac{dp}{dt} = R_L \left[(1 + 0.6r) T \frac{d\rho}{dt} + (1 + 0.6r) \rho \frac{dT}{dt} + 0.6 \rho T \frac{dr}{dt} \right] \tag{53}$$

which follows from the equation of state, in view of

$$\frac{d \ln r}{dt} = \delta \frac{d \ln r_s}{dt} \tag{54}$$

[56] J. Smagorinsky and G. O. Collins: "On the Numerical Prediction of Precipitation," *Monthly Weather Review* 83, 1955.

with

$$\delta = \begin{cases} 0 \text{ if } \operatorname{div} V \leqslant 0 \text{ or } r < r_s \\ 1 \text{ if } \operatorname{div} V > 0 \text{ and } r = r_s \end{cases} \tag{55}$$

and because of

$$\frac{d\rho_L}{\rho_L} = \frac{d\rho}{\rho} - 0.6\, dr \tag{56}$$

eq. (53) can be formulated as

$$F = \frac{dp}{dt} = -\left[\frac{B+k+a}{B} - \frac{A}{B}\left(\frac{k+a}{B} - 0.6\right)\delta \cdot r\right] p \operatorname{div} V + \frac{R_L \rho}{B c_p}\frac{đq}{dt}$$

$$= - M[r,\delta]\, p \operatorname{div} V + N\frac{đq}{dt}$$

$$= - Mp\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) + N\frac{đq}{dt}\,. \tag{57}$$

Substitution of eq. (57) in eqs. (45) through (47) results in the linear diagnostic filter equations

$$Mp\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 w}{\partial x \partial z}\right) + \left[\frac{\partial (Mp)}{\partial x} - \frac{\partial p}{\partial x}\right]\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)$$
$$+ \frac{\partial u}{\partial x}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial x}\frac{\partial p}{\partial y} + \frac{\partial w}{\partial x}\frac{\partial p}{\partial z} - f\rho\left(\frac{1}{\rho}\frac{\partial p}{\partial y} + fu\right)$$
$$+ v\rho\left(u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial x}\left(N\frac{đq}{dt}\right) = 0 \tag{58}$$

$$Mp\left(\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 w}{\partial y \partial z}\right) + \left[\frac{\partial (Mp)}{\partial y} - \frac{\partial p}{\partial y}\right]\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)$$
$$+ \frac{\partial u}{\partial y}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial p}{\partial y} + \frac{\partial w}{\partial y}\frac{\partial p}{\partial z} + f\rho\left(\frac{1}{\rho}\frac{\partial p}{\partial x} - fv\right)$$
$$- u\rho\left(u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial y}\left(N\frac{đq}{dt}\right) = 0 \tag{59}$$

$$\mathrm{Mp}\left(\frac{\partial^2 u}{\partial x \partial z} + \frac{\partial^2 v}{\partial y \partial z} + \frac{\partial^2 w}{\partial z^2}\right) + \left[\frac{\partial (\mathrm{Mp})}{\partial z} - \frac{\partial p}{\partial z}\right]\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)$$

$$+ \frac{\partial u}{\partial z}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial z}\frac{\partial p}{\partial y} + \frac{\partial w}{\partial z}\frac{\partial p}{\partial z} - \frac{\partial}{\partial z}\left(N\frac{\mathrm{d}q}{\mathrm{d}t}\right) = 0 \qquad (60)$$

Equation (60) provides an excellent diagnostic equation and reduces the numerical relaxation work considerably which is only required in eq. (58) and (59). In the form of an ordinary linear differential, eq. (60) appears as

$$\mathrm{Mp}\,\frac{\partial^2 w}{\partial z^2} + \frac{\partial (\mathrm{Mp})}{\partial z}\frac{\partial w}{\partial z} + \mathrm{Mp}\left(\frac{\partial^2 u}{\partial x \partial z} + \frac{\partial^2 v}{\partial y \partial z}\right) + \left[\frac{\partial (\mathrm{Mp})}{\partial z} - \frac{\partial p}{\partial z}\right]\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)$$

$$+ \frac{\partial u}{\partial z}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial z}\frac{\partial p}{\partial y} - \frac{\partial}{\partial z}\left(N\frac{\mathrm{d}q}{\mathrm{d}t}\right) = 0 \qquad (61)$$

With

$$\tilde{F} = \frac{\partial^2 u}{\partial x \partial z} + \frac{\partial^2 v}{\partial y \partial z} + \frac{1}{\mathrm{Mp}}\left\{\left[\frac{\partial (\mathrm{Mp})}{\partial z} - \frac{\partial p}{\partial z}\right]\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)\right.$$

$$\left. + \frac{\partial u}{\partial z}\frac{\partial p}{\partial x} + \frac{\partial v}{\partial z}\frac{\partial p}{\partial y} - \frac{\partial}{\partial z}\left(N\frac{\mathrm{d}q}{\mathrm{d}t}\right)\right\} \qquad (62)$$

the solution of eq. (61) is

$$w = \int_{z_1}^{z} \ln \frac{(\mathrm{Mp})_z}{(\mathrm{Mp})_{z_1}} \left[\int_{z_1}^{z} \ln \frac{(\mathrm{Mp})_z}{(\mathrm{Mp})_{z_1}}\, \tilde{F}(z)\, dz + C_1\right] dz + C_2 \qquad (63)$$

At the lower boundary, $w_{B_1} = u_{B_1}\frac{\partial \phi}{\partial x} + v_{B_1}\frac{\partial \phi}{\partial y}$ with $\phi(x, y)$ as the geopotential of the ground. Accordingly,

$$C_2 = u_{B_1}\frac{\partial \phi}{\partial x} + v_{B_1}\frac{\partial \phi}{\partial y}\,. \qquad (64)$$

Since, at the upper boundary $w_{B_2} = 0$, the other integration constant yields the value

$$C_1 = -\left[\int_{z_1}^{z_2} \ln \frac{(\mathrm{Mp})_z}{(\mathrm{Mp})_{z_1}}\, dz\right]^{-1}\left\{\int_{z_1}^{z_2} \ln \frac{(\mathrm{Mp})_z}{(\mathrm{Mp})_{z_1}}\left[\int_{z_1}^{z} \ln \frac{(\mathrm{Mp})_z}{(\mathrm{Mp})_{z_1}}\, \tilde{F}(z)\, dz\right] dz + C_2\right\} \qquad (65)$$

We have to remember that the saturation mixing ratio $r_s = r_s(p, T)$ and $\gamma = \gamma(T)$ and that it is necessary to obtain first a solution of eqs. (58), (59) and (63) with $\delta = 0$ whereupon the criterion (55) is applied. One or two iterations will then yield satisfactory results. Unless δ has some variability along the lines suggested by Smagorinsky,[57] the variable δ should be about 0.8 instead of 1.0 in agreement with numerical simulations.

It is to be expected that the under-relaxation factors ϑ in the iteration scheme

$$u^{(n+1)} = u^{(n)} + \vartheta\, G_1 \left[u^{(n)}\, v^{(n)}, w^{(n)}\right]$$

$$v^{(n+1)} = v^{(n)} + \vartheta\, G_2 \left[u^{(n)}, v^{(n)}, w^{(n)}\right] \qquad (66)$$

with $u^{(1)} = -\frac{1}{f\rho}\frac{\partial p}{\partial y}$, $v^{(1)} = \frac{1}{f\rho}\frac{\partial p}{\partial x}$, $w^{(1)} = 0$

in which G_1 and G_2 represent residuals of eq. (58) and eq. (59), respectively, have to be quite small in equilibrium-scale solutions ($\Delta x = \Delta y \geqslant 50$ km, $\Delta z \geqslant 1$ km) involving strong divergence and vertical wind velocities. Since $u^{(1)}$ and $v^{(1)}$ become singular at the equator and convergence is slow in very low latitudes, fine grid solutions are not possible in the vicinity of the equator. Due to the fact that the mass field cannot be accurately determined in the equatorial region, horizontal winds, obtained through the tracking of floating balloons, and additional temperature measurements would facilitate the computation of all desired quantities. The use of diagnostic filtering equations for this purpose has been mentioned by several authors including Mintz[57] though in connection with the more restrictive hydrostatic prediction system.

Utilization of the hydrostatic approximation with respect to height determinations weakens the application of the filtering and associated prognostic equations as far as smaller scales are concerned but still allows the computation of divergences exceeding the vorticity on a constant pressure surface. This is of importance pertaining to the immediate applicability of the new prediction system.

As to the upper boundary condition, the assumption $w = 0$, of course, has to be made for a finite height. In this respect, the condition var w = Min. would provide a good separation criterion. This has to coincide with the criteria var $\frac{\partial u}{\partial z}$ = Min., var $\frac{\partial v}{\partial z}$ = Min., var $\frac{\partial T}{\partial z}$ = Min., var $\frac{\partial \rho}{\partial z}$ = Min., as far as interpolation from a lower to a higher level is concerned, i.e., to an average equilibrium boundary which exists at 20 - 25km height.

[57] Y. Mintz: "The Four Basic Requirements for Numerical Weather Prediction in Global Weather Prediction," ed. by Bruce and Kiely, Holt, Rinehart and Winston, New York, 1970.

[58] J. Smagorinsky: "On the Dynamical Prediction of Large-Scale Condensation by Numerical Methods," Geophysical Monograph No. 5, American Geophysical Union, 1960.

References

Academic Press, (Publisher) 1977. General Circulation Models of the Atmosphere. In Methods in Computational Physics, Vol. 17, New York, N.Y.

Bolin, B. 1956. An Improved Barotropic Model and Some Aspects of Using the Balance Equation for Three-dimensional Flow. Tellus 8, p. 61.

Bourke, B., McAvaney, B., Puri, K., And Thurling, R. 1977. Global Modeling of Atmospheric Flow by Spectral Methods. In Methods in Computational Physics, Vol. 17. Academic Press, New York, N.Y., p. 268.

Charney, J. 1951. Dynamic Forecasting by Numerical Process. Compendium of Meteorology. Amer. Meteor. Soc., Boston, Mass., p. 470.

Cressman, G. 1963. Three-Level Model Suitable for Daily Numerical Forecasting, Tech. Memo. No. 22, Natl. Meteor. Center, Natl. Weather Service, Washington, D.C.

....
Doos, B. 1965. Numerical Weather Forecasting with the Barotropic Model. Hydrometeoizdat, Leningrad, U.S.S.R., 1969, p. 188.

....
Doos, B. R., 1970. Numerical Experimentation Related to GARP. GARP Publ. Series No. 6, WMO, Geneva, Switzerland.

Eliassen, A., and E. Kleinschmidt. 1957. Dynamic Meteorology. Encyclopedia of Physics, Vol. XLVIII, Berlin, Germany.

Fjørtoft, R. 1962. A Numerical Method of Solving Certain Differential Equations of Second Order. Geoph. Publ. 24, p. 229.

Fruehwald, D. 1968. Entwicklung und Anwendung eines gefilterten Integrationsverfahrens fuer die Vorausberechnung atmosphaerischer Felder. Dissertations-und Fotodruck Frank, Munich, Germany.

Gerrity, J., and R. Chu. 1978. An Iterative Variational Method for Adjusting Isobaric Wind and Geopotential to Satisfy the Balance Equation. Office Note 192. Natl. Meteorological Center, Natl Weather Service, Washington, D.C.

Haltiner, G., and F. Martin. 1957. Dynamical and Physical Meteorology. McGraw Hill Book Co., New York, N.Y.

Herbert, F. 1971. Static and Quasistatic Motion in the Atmosphere, Contrib. Atmos. Phys. 44, p. 17.

Hinkelmann, K. H. 1965. Primitive Equations. Lectures on Numerical Short-Range Weather Prediction. Hydrometeoizdat, Leningrad, 1969, p. 306.

Hollmann, G. 1966. Zur Frage neuer diagnostischer Beziehungen zwischen Wind-und Druckfeld (Balancegleichungen) in einer barotropen Atmosphäre mit divergenter Strömung. Beitr. z. Phys. d. Atmosp. 39, p. 99.

Hollaway, J. L. and S. Manabe. 1971. Simulation of Climate by a Global General Circulation Model. I. Hydrological Cycle and Heat Balance. Mon. Wea. Rev. 99, p. 335.

Hydrometeoizdat, (Publisher). 1969. Lectures on Numerical Short-Range Weather Prediction. WMO Regional Training Seminar, (Moscow, Nov.-Dec., 1965), Leningrad, U.S.S.R.

Knighting, E. 1965. Three-Dimensional Weather Prediction. Hydrometeoizdat, Leningrad, U.S.S.R., 1969, p. 221.

Lorenz, E. N. 1969. The Predictability of a Flow which Possesses many Scales of Motion. Tellus 21, p. 289.

Miyakoda, K., and R. W. Moyer. 1968. A Method of Initialization for Dynamical Weather Forecasting, Tellus 20, p. 15.

Miakoda, K. Strickler, R. F., Nappo, C. J., Baker, P. L., and Hembree, G. D. 1971. The Effect of Horizontal Grid Resolution in an Atmospheric Circulation Model. J. Atmos. Sci. 28, p. 481.

Monin, A. S. 1972. Weather Forecasting as a Problem in Physics. The MIT Press, Cambridge, Mass.

Morel, P., and O. Talagrand. 1974. Dynamic Approach to Meteorological Data Assimilation. Tellus 26, p. 334.

Nitta, T., and J. B. Hovermale. 1967. On Analysis and Initialization for the Primitive Forecast Equations. Weather Bureau Technical Memorandum NMC-42, Suitland, MD.

Okland, H. 1970. On the Adjustment Toward Balance in Primitive Equation Weather Prediction Models. Mon. Wea. Rev. 98, p. 271.

Phillips, N. 1956. A Coordinate System Having Some Special Advantages for Numerical Forecasting. Journ. Meteor. 14, p. 184.

Phillips, N. 1960. Numerical Weather Predictions. In Advances in Computers. Academic Press, New York, N.Y.

Shuman, F. 1965, A Multi-Level Primitive Equation Model. Hydrometeoizdat, Leningrad, U.S.S.R., 1969, p. 465.

Shuman, F. 1978. Numerical Weather Prediction. Bull. Am. Meteor. Soc. 59, p. 5.

Smagorinsky, J., 1963. General Circulation Experiments with the Primitive Equations. Mon. Wea. Rev. 91, p. 99.

Smagorinsky, J., K. Miyakoda, and R. F. Strickler. 1970. The Relative Importance of Variables in Initial Conditions for Dynamical Weather Prediction. Tellus 22, p. 141.

Sundquist, H. 1970. Numerical Simulation of the Development of Tropical Cyclones with a Ten-level Model. Part I. Tellus 22, p. 359.

The Committee on Atmospheric Sciences, National Research Council. 1971. The Atmospheric Sciences and Man's Needs. National Academy of Sciences, Washington, D.C.

Thompson. P. D. 1961. Numerical Weather Analysis and Prediction. The MacMillan Company. New York, N.Y.

U.S. National Weather Service. 1978. Numerical Weather Prediction Activities Report. Washington, D.C.

van Mieghem, J.M. 1951. Hudrodynamic Instability. Compendium of Meteorology. Amer. Meteor. Soc., Boston, Mass., p. 434.

WMO (World Meteorological Organization), 1965. Research and Development Aspects of Long-Range Forecasting. WMO-No. 162, TP. 79, Geneva, Switzerland.

SOME BESSEL FUNCTION IDENTITIES ARISING IN ICE MECHANICS PROBLEMS

Shunsuke Takagi
Physical Sciences Branch
U.S. Army Cold Regions Research and Engineering Laboratory
Hanover, New Hampshire 03755

ABSTRACT. Some Bessel function identities found by solving problems of the deflection of the floating ice plate by two different methods are rigorously proved. The master formulas from which all the identities are derived are in a Fourier reciprocal relationship, connecting a Hankel function to an exponential function. Many new formulas can be derived from the master formulas. The analytical method presented here now opens the way to study a hitherto impossible type of problems - the deflection of floating elastic plates of various shapes and boundary conditions.

I. INTRODUCTION. By solving several problems of the deflection of a floating ice plate (i.e., the deflection of a plate on a continuous elastic foundation formulated by Winkler [9]) by two different methods, Kerr [4,5] presented a number of equality relationships among Bessel functions. In this paper, the analytical derivation of his formulas is presented. His formulas (six in all) reduce to the following *two master formulas* expressed in the Fourier reciprocal relationship:

$$\int_{-\infty}^{\infty} e^{ix\xi}\, H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{a^2+\xi^2})\, d\xi = \frac{2}{i\sqrt{x^2+\beta^2}}\, e^{-a\sqrt{x^2+\beta^2}}\,, \tag{1}$$

$$\int_{-\infty}^{\infty} e^{-a\sqrt{x^2+\beta^2}-i\xi x}\, \frac{dx}{\sqrt{x^2+\beta^2}} = \pi i\, H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{a^2+\xi^2})\,, \tag{2}$$

where $H_o^{(1)}(\ )$ is the Hankel function of zeroth order; x and ξ are real and a nonnegative such that $a^2+\xi^2 \neq 0$; and β is complex such that $\beta \neq 0$, $|\arg\beta|<\pi/2$, $x^2+\beta^2 \neq 0$. By $\sqrt{z}$ we mean such a branch as $\mathrm{Re}\sqrt{z} \geq 0$. The master formulas may be transformed to symmetric forms:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_o(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+y^2})\, e^{i\xi x + i\eta y}\, dx dy = \frac{-4i}{\xi^2+\eta^2+\beta^2}\,, \tag{3}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\xi^2+\eta^2+\beta^2}\, e^{-ix\xi-iy\eta}\, d\xi d\eta = \pi^2 i\, H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+y^2})\,. \tag{4}$$

In the following, we prove the above formulas, transform the two master formulas to derive all the formulas introduced by Kerr [4,5] as well as some new ones, and finally show that the analytical method presented here lays the foundation for building a mathematical machinery for solving various shapes of floating ice plate under various boundary conditions.

II. PROOF OF FORMULA (1). Use of Barnes' integral representation of $H_o^{(1)}(z)$

$$\pi i\, H_o^{(1)}(z) = \frac{1}{2\pi i} \int_{-c-\infty i}^{-c+\infty i} \Gamma^2(-s)\left(-\frac{iz}{2}\right)^{2s} ds \ , \tag{5}$$

where c is any positive number and $|\arg(-iz)| < \pi/2$ [8, p. 192], transforms the single integral on the left-hand side of (1)

$$I_1 = \pi i \int_{-\infty}^{\infty} e^{ix\xi}\, H_o^{(1)}\left(\beta e^{\frac{\pi i}{2}} \sqrt{a^2+\xi^2}\right) d\xi \tag{6}$$

to the repeated integrals

$$I_1 = \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{ix\xi}\, d\xi \int_{-c-\infty i}^{-c+\infty i} \Gamma^2(-s)\left(\frac{\beta}{2}\sqrt{a^2+\xi^2}\right)^{2s} ds \ . \tag{7}$$

The absolute convergence of integral (5) carries over to (7), because the condition $|\arg(-iz)| < \pi/2$ in (5) transforms to $|\arg\beta| < \pi/2$ in (7), which is one of the prerequisites in the master formulas. Therefore, the order of integration in (7) may be exchanged. Moreover, restricting the original range of a, which is $a \geq 0$, to $a > 0$, we let $\xi = a\eta$ in (7). Thus (7) becomes

$$I_1 = \frac{a}{2\pi i} \int_{-c-\infty i}^{-c+\infty i} \Gamma^2(-s)\left(\frac{\beta a}{2}\right)^{2s} ds \int_{-\infty}^{\infty} e^{iax\eta}\,(1+\eta^2)^s\, d\eta \ . \tag{8}$$

To evaluate the internal single integral in (8)

$$M_1 = \int_{-\infty}^{\infty} e^{iaxz}\,(1+z^2)^s\, dz$$

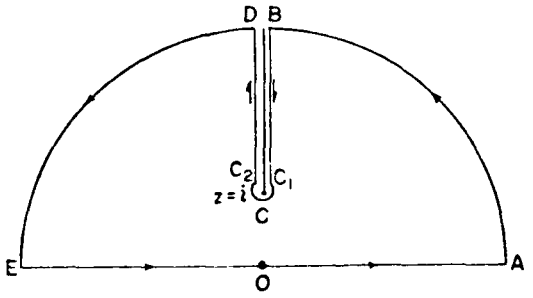


Fig. 1

Transformation of integral M_1 on the z-plane.

by the contour integral method, we first note that the original range of x(i.e., $-\infty < x < \infty$) may be restricted to $0 < x$, (where the case $x = 0$ is excluded), because I_1 is an even function of x, as the right-hand side of (6) shows. Consider the contour in Fig. 1, that starts at origin 0, goes along the positive real axis to A, i.e., $z = \infty$; takes a 90 degree turn along the infinitely large circle to reach B, i.e., $z = i\infty$, comes down along the imaginary axis to C, i.e., $z = i$; makes a 360 degree turn along an infinitely small circle clockwise around C; goes upward along the imaginary axis to reach D, i.e., $z = i\infty$; takes a 90 degree turn along the infinitely large circle to reach E, i.e., $z = -\infty$; and finally reaches origin 0, thus completing a circuit. No singularity of the integrand $\exp(iaxz)(1+z^2)^s$ exists inside this closed contour. Among the integrals along the paths mentioned above, the integrals along AB and DE are equal to zero, provided that $a > 0$. The integral around C is also equal to zero. Therefore, on the condition that $a > 0$ and $x > 0$, we have

$$M_1 = -\left\{\int_B^{C_1} + \int_{C_2}^{D}\right\} e^{iaxz}\,(1+z^2)^s\,dz\ ,$$

where C_1 and C_2 are the initial and terminal points of the infinitely small circle around C. Letting $z = it$, where t is real, M_1 reduces to

$$M_1 = i(1-e^{-2\pi is})\int_1^{\infty} e^{-axt}(1-t^2)^s\,dt\ .$$

We now let s be

$$s = -\tfrac{1}{2} + ip\ ,$$

i.e. let c in (5) be $\frac{1}{2}$, where p is a real number, in order to integrate M_1 by use of the formula [8, p. 172]

$$K_\nu(z) = \frac{\Gamma(\frac{1}{2})(\frac{1}{2}z)^\nu}{\Gamma(\nu+\frac{1}{2})}\int_1^{\infty} e^{-zt}(t^2-1)^{\nu-\frac{1}{2}}\,dt, \qquad (9)$$

which is valid when $R_e(\nu+\frac{1}{2}) > 0$ and $|\arg z| < \pi/2$. These two conditions are satisfied when we let $\nu - \frac{1}{2} = s$ and $z = ax$ to integrate M_1. Thus, letting

$$(1-t^2)^s = e^{\pi is}\,(t^2-1)^s\ ,$$

M_1 integrates to

$$M_1 = -2 \sin \pi s \frac{\Gamma(s+1)}{\sqrt{\pi}(\frac{ax}{2})^{s+\frac{1}{2}}} K_{s+\frac{1}{2}}(ax) .$$

In this way, (7) transforms to a single integral

$$I_1 = -\frac{a}{\pi i} \int_{-\frac{1}{2}-\infty i}^{-\frac{1}{2}+\infty i} \Gamma^2(-s) (\frac{\beta a}{2})^{2s} \sin\pi s \frac{\Gamma(s+1)}{\sqrt{\pi}(\frac{ax}{2})^{s+\frac{1}{2}}} K_{s+\frac{1}{2}}(ax) \, ds.$$

Changing the Gamma function of the negative argument to the positive argument by the reflection formula

$$\Gamma(-s) = \frac{-\pi}{\Gamma(1+s)\sin\pi s} ,$$

I_1 becomes

$$I_1 = -\frac{1}{\pi i} \sqrt{\frac{2\pi a}{x}} \int_{-\frac{1}{2}-\infty i}^{-\frac{1}{2}+\infty i} \frac{\pi}{\sin\pi s} \frac{K_{s+\frac{1}{2}}(ax)}{\Gamma(s+1)} (\frac{\beta^2 a}{2x})^s \, ds.$$

Taking the residues, I_1 integrates to

$$I_1 = \frac{8\pi a}{x} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (\frac{\beta^2 a}{2x})^n K_{n+\frac{1}{2}}(ax) .$$

Replacing $K_{n+\frac{1}{2}}(ax)$ with

$$K_\nu(z) = \tfrac{1}{2}(\frac{z}{2})^\nu \int_0^\infty \exp(-t - \frac{z^2}{4t}) t^{-\nu-1} \, dt,$$

i.e., a formula found in [8, p. 183], I_1 becomes

$$I_1 = a\sqrt{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^\infty (\frac{a^2\beta^2}{4t})^n \exp(-t - \frac{a^2x^2}{4t}) \frac{dt}{\sqrt{t^3}} .$$

The order of the summation and integration may be exchanged, and we get

$$I_1 = a\sqrt{\pi} \int_0^\infty \exp\left(-t - \frac{a^2(x^2+\beta^2)}{4t}\right) \frac{dt}{\sqrt{t^3}} \quad .$$

Letting $t = \xi^{-2}$, this becomes

$$I_1 = 2a\sqrt{\pi} \int_0^\infty \exp\left(-\xi^{-2} - \frac{a^2(x^2+\beta^2)}{4} \xi^2\right) d\xi. \tag{10}$$

To integrate (10), we introduce a lemma:

$$\int_0^\infty e^{-\mu^2\xi^2 - \xi^{-2}} \, d\xi = e^{-2\mu} \frac{\sqrt{\pi}}{2\mu} \ ,$$

provided

$$n\pi - \frac{\pi}{4} < \arg\mu < n\pi + \frac{\pi}{4}$$

where n is an integer.

When μ is in the above range, the above integral is convergent. To prove the lemma, we first note that the integral

$$L_1 = \int_0^\infty e^{-\mu^2\xi^2 - \xi^{-2}} \, d\xi$$

transforms to

$$L_1 = e^{-2\mu} N_1 \ , \tag{11}$$

where

$$N_1 = \int_0^\infty e^{-(\mu\xi - \xi^{-1})^2} \, d\xi \ . \tag{12}$$

Letting

$$\xi = 1/(\mu\eta)$$

and changing the resulting contour $0 \sim \mu^{-1}\infty$ to $0 \sim \infty$, we get

$$N_1 = \int_0^\infty e^{-(\mu\eta - \eta^{-1})^2} \frac{d\eta}{\mu\eta^2} \ . \tag{13}$$

Addition of (12) and (13) yields

$$2N_1 = \frac{1}{\mu}\int_0^{\infty} e^{-(\mu\xi-\xi^{-1})^2}\ (\mu+\frac{1}{\xi^2})\ d\xi\ .$$

Letting

$$\mu\xi - \xi^{-1} = t\ ,$$

we get

$$2N_1 = \frac{1}{\mu}\int_{-\infty}^{\mu\infty} e^{-t^2}\ dt.$$

Changing the range of integration to the one from $-\infty$ to $+\infty$, N_1 integrates to

$$N_1 = \sqrt{\pi}/(2\mu).$$

Substituting this value into (11), the lemma is proved.

Letting μ be

$$\mu = \frac{a}{2\sqrt{x^2+\beta^2}}$$

in the lemma, (10) is integrated, because μ above is obviously in the range prescribed before. Thus, under the conditions $x \neq 0$ and $a \neq 0$, Formula (1) is proved. Applying the analytical continuation, the condition $a \neq 0$ is extended to the condition $a^2 + \beta^2 \neq 0$. Because the integral is convergent at $x = 0$, the condition $x \neq 0$ may be removed. The proof is thus completed.

III. PROOF OF FORMULA (2). Although Formula (2) is the Fourier inverse of Formula (1), we show an independent proof in view of the importance of the formula.

On the assumption that $a^2 + \xi^2 \neq 0$ and $\beta \neq 0$, letting

$$\begin{aligned} a &= r\cos\alpha \\ \xi &= r\sin\alpha \\ x &= \beta\,\mathrm{Sinh}z \\ -\frac{\pi}{2} &< \alpha < \frac{\pi}{2}\ , \end{aligned} \qquad (14)$$

the integral

$$I_2 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{x^2+\beta^2}} e^{-a\sqrt{x^2+\beta^2} - i\xi x} \, dx \tag{15}$$

transforms to

$$I_2 = \int_L \exp(-r\beta\mathrm{Cosh}(z+i\alpha)) \, dz \ , \tag{16}$$

where the contour L is a curve on the complex $z = u + iv$ plane (Fig. 2) defined by

$$\begin{aligned} z &= \mathrm{Arcsinh}(x/\beta) \\ &= \log\left(\frac{x}{\beta} + \sqrt{\frac{x^2}{\beta^2} + 1}\right) \end{aligned} \tag{17}$$

with parameter x in the range of $-\infty < x < \infty$.

Letting x=0, we have z=0. Therefore, contour L passes through the origin. When $x \to +\infty$ or $-\infty$, z asymptotically approaches $\log(2x) - \log\beta$, or $-\log(-2x) + \log\beta$,respectively; in other words,the imaginary part v of the complex variable z satisfies the conditions

$$\begin{aligned} \lim_{x \to +\infty} v + \arg\beta &= 0 \\ \lim_{x \to -\infty} v - \arg\beta &= 0 \ . \end{aligned} \tag{18}$$

The curve L_+ defined for the case $0 < \arg\beta < \pi/2$ is shown in Fig. 2. The curve L_- defined for the case $-\pi/2 < \arg\beta < 0$ is symmetrical with L_+ with regard to the real axis.

Letting u be the real part of z, we find that, as $x \to \infty$ or $-\infty$, the real part of $-r\beta\mathrm{Cosh}(z+i\alpha)$ approaches asymptotically either

$$-\frac{1}{2} r|\beta| \, e^{u} \cos(\arg\beta + v + \alpha)$$

or

$$-\frac{1}{2} r|\beta| \, e^{-u} \cos(\arg\beta - v - \alpha) \ ,$$

respectively. Because the power of the exponent in (16) must remain negative as $|x| \to \infty$, the conditions

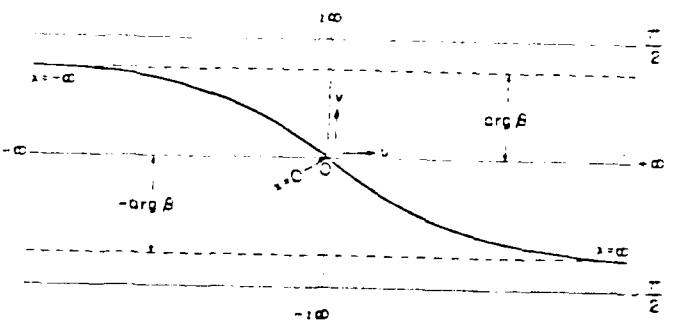


Fig. 2. Curve L_+ defined for the case $0 < \arg\beta < \pi/2$ in the complex $z = u + iv$ plane is shown for parameter x ranging over $-\infty < x < \infty$.

$$-\frac{\pi}{2} < \arg\beta + \lim_{x \to +\infty} v + \alpha < \frac{\pi}{2}$$

$$-\frac{\pi}{2} < \arg\beta - \lim_{x \to -\infty} v - \alpha < \frac{\pi}{2} \tag{19}$$

must be satisfied. The conditions are satisfied by $(14)_4$ and (18).

Introducing t defined by

$$t = z + i\alpha \tag{20}$$

I_2 becomes

$$I_2 = \int_{L+i\alpha} \exp(-r\beta\cosh t)dt \ .$$

The contour $L+i\alpha$ may be moved to the real axis, because doing this is tantamount to letting $\lim_{x \to \pm\infty} v + \alpha = 0$, which is evidently permissible in the prescribed range of β shown below (2). Thus we have

$$I_2 = \int_{-\infty}^{\infty} \exp(-r\beta\cosh t)\ dt \ . \tag{21}$$

To integrate I_2 in (21), we introduce the following formula in [7]. Provided $I_m(z) > 0$,

$$H_o^{(1)}(z) = \frac{1}{\pi i}\int_{-\infty}^{\infty} \exp(iz\cosh t)\ dt \ . \tag{22}$$

Comparing z in (22) with $-r\beta$ in (21) we have $I_m(z) = R_e(\beta r)$; therefore, the condition $I_m(z) > 0$ is satisfied in the prescribed range of β . Thus I_2 integrates to

$$I_2 = \pi i H_o^{(1)}(\beta e^{\frac{\pi i}{2}} r) \ ,$$

completing the proof of Formula 2.

IV. TRANSFORMATIONS OF FORMULAS (1) AND (2). We list the formulas found in [1], [4], and [5] that can be derived from Formulas (1) and (2). The formulas we use in the following transformations are

$$\int_{-\infty}^{\infty} e^{ia\xi}\, H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+\xi^2})\, d\xi = \frac{2}{i\sqrt{a^2+\beta^2}}\, e^{-\sqrt{a^2+\beta^2}x}, \tag{23}$$

which is found by exchanging x and a in Formula (1), and

$$\int_{-\infty}^{\infty} e^{-a\sqrt{x^2+\beta^2}-ibx}\, \frac{dx}{\sqrt{x^2+\beta^2}} = \pi i H_o^{(1)}(\beta\, e^{\frac{\pi i}{2}}\sqrt{a^2+b^2}), \tag{24}$$

which is found by changing ξ to b in Formula (2).

Assuming β to be a positive number, (23) transforms to a real integral

$$\int_{o}^{\infty} K_o(\beta\sqrt{x^2+\xi^2})\cos(a\xi)d\xi = \frac{\pi}{2\sqrt{a^2+\beta^2}}\, e^{-\sqrt{a^2+\beta^2}x} \tag{25}$$

Letting $a = 0$, (25) may become

$$\int_{-\infty}^{\infty} K_o(\beta\sqrt{x^2+\xi^2})d\xi = \frac{\pi}{\beta}\, e^{-\beta x}. \tag{26}$$

Letting $\xi = \eta - y$, and expressly writing that a may be positive or negative, (23) may be rewritten to

$$\int_{-\infty}^{\infty} e^{\pm ia\eta}\, H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+(y-\eta)^2})\, d\eta = \frac{2}{i\sqrt{a^2+\beta^2}}\, e^{\pm ia\eta\, -\sqrt{a^2+\beta^2}x} \tag{27}$$

We transform this to several forms. When β is real, (27) becomes

$$\int_{-\infty}^{\infty} e^{\pm ia\eta}\, K_o(\beta\sqrt{x^2+(y-\eta)^2})d\eta = \frac{\pi}{\sqrt{a^2+\beta^2}}\, e^{\pm iay\, -\sqrt{a^2+\beta^2}x} \tag{28}$$

Letting $\beta = b\exp(\pi i/4)$ with the restriction $b > 0$, (27) becomes

$$\int_{-\infty}^{\infty} e^{\pm ia\eta} \left\{ \ker\left(b\sqrt{x^2+(y-\eta)^2}\right) + i\,\mathrm{kei}\left(b\sqrt{x^2+(y-\eta)^2}\right) \right\} d\eta$$

$$= \frac{\pi}{\sqrt{a^2+ib^2}}\, e^{-\sqrt{a^2+ib^2}\,x \pm iay} \,. \tag{29}$$

Letting

$$\sqrt{a^2 + ib^2} = p + iq \tag{30}$$

and adding and subtracting the plus expression and the minus expression in (29), we find two integrals

$$\int_{-\infty}^{\infty} \cos a\eta \left\{ \ker\left(b\sqrt{x^2+(y-\eta)^2}\right) + i\,\mathrm{kei}\left(b\sqrt{x^2+(y-\eta)^2}\right) \right\} d\eta$$

$$= \frac{\pi(p-iq)}{\sqrt{a^4+b^4}}\, e^{-(p+iq)x} \cos ay \tag{31}$$

$$\int_{-\infty}^{\infty} \sin a\eta \left\{ \ker\left(b\sqrt{x^2+(y-\eta)^2}\right) + i\,\mathrm{kei}\left(b\sqrt{x^2+(y-\eta)^2}\right) \right\} d\eta$$

$$= \frac{\pi(p-iq)}{\sqrt{a^4+b^4}}\, e^{-(p+iq)x} \sin ay \,,$$

where

$$\left.\begin{matrix} p \\ q \end{matrix}\right\} = \frac{1}{\sqrt{2}} \sqrt{\sqrt{a^4+b^4} \pm a^2} \,. \tag{32}$$

Dividing the above two formulas into real and imaginary parts, four real integrals may be found.

Letting $\beta = 1$, (24) becomes

$$\int_{-\infty}^{\infty} e^{-a\sqrt{x^2+1} - ibx}\, \frac{dx}{\sqrt{x^2+1}} = 2\, K_o\left(\sqrt{a^2+b^2}\right) \tag{33}$$

Differentiating (33) with regard to a, we find

$$\int_{-\infty}^{\infty} e^{-a\sqrt{x^2+1} - ibx}\, dx = -\frac{2a}{\sqrt{a^2+b^2}} K_o'(\sqrt{a^2+b^2}) \ . \tag{34}$$

Formula (25) is listed by Erdély [1] without proof. Formulas (26), (28) and $(31)_2$ were derived by Kerr[4], and Formula (34) by Kerr [5]. Because the differentiations and integrations with regard to parameters inside the integral as well as the assignment of arbitrary values to parameters are permissible insomuch as the absolute convergence is preserved, many additional formulas may be derived from (23) and (24).

V. PROOF OF FORMULAS (3) AND (4). Integrating once more the rewritten Formula (1)

$$\int_{-\infty}^{\infty} e^{i\xi x} H_o^{(1)}(\beta e^{\frac{\pi i}{2}} \sqrt{x^2+y^2})\, dx = \frac{2}{i\sqrt{\xi^2+\beta^2}} e^{-|y|\sqrt{\xi^2+\beta^2}} \ ,$$

where $|y|$ takes the place of the original positive number a, we transform the left-hand side of Formula (3)

$$I_3 = \int_{-\infty}^{\infty} e^{i\eta y}\, dy \int_{-\infty}^{\infty} e^{i\xi x} H_o^{(1)}(\beta e^{\frac{\pi i}{2}} \sqrt{x^2+y^2})\, dx$$

to

$$I_3 = \frac{2}{i\sqrt{\xi^2+\beta^2}} \int_{-\infty}^{\infty} e^{i\eta y - |y|\sqrt{\xi^2+\beta^2}}\, dy \ .$$

Use of the integral

$$\int_{-\infty}^{\infty} e^{i\eta y - |y|\sqrt{\xi^2+\beta^2}}\, dy = \frac{2\sqrt{\xi^2+\beta^2}}{\xi^2+\eta^2+\beta^2} \tag{35}$$

reduces I_3 to the right-hand side of Formula (3). To prove (35), note that

$$\int_{-\infty}^{\infty} e^{i\eta y - |y|\sqrt{\xi^2+\beta^2}}\, dy = 2\int_{o}^{\infty} e^{-y\sqrt{\xi^2+\beta^2}} \cos\eta y\, dy \ .$$

Although Formula (4) is the Fourier inverse of Formula (3), we show an independent proof in view of the importance of the formula. The rewritten Formula (2)

$$\int_{-\infty}^{\infty} e^{-|y|\sqrt{\xi^2+\beta^2}\,-\,ix\xi}\ \frac{d\xi}{\sqrt{\xi^2+\beta^2}} = \pi i\ H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+y^2})$$

reduces to Formula (4) by the substitution of the Fourier inverse of (35)

$$e^{-|y|\sqrt{\xi^2+\beta^2}} = \frac{\sqrt{\xi^2+\beta^2}}{\pi}\int_{-\infty}^{\infty}\frac{1}{\xi^2+\eta^2+\beta^2}\,e^{-iy\eta}\,d\eta\ . \tag{36}$$

Formula (36) may otherwise be proved by showing that

$$\int_{-\infty}^{\infty}\frac{e^{-iyz}}{z^2+a^2}\,dz = \frac{\pi}{a}\,e^{-a|y|} \tag{37}$$

by use of the contour integral method, where $R_e\ a \neq 0$.

VI. ADDITIONAL DERIVATION FROM THE MASTER FORMULAS. We prove the formula

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{a^2+\xi^2}}\,H_{o,1}^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{a^2+\xi^2})\,H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+(y-\xi)^2})\,d\xi$$

$$= \frac{2}{\beta a}\,H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{(a+x)^2+y^2})\ , \tag{38}$$

where

$$H_{o,1}^{(1)}(\lambda) = \left[\frac{d}{dz}\,H_o^{(1)}(z)\right]_{z=\lambda} \tag{39}$$

Differentiating (24) with regard to a, we find

$$\int_{-\infty}^{\infty} e^{-a\sqrt{x^2+\beta^2}\,-\,ibx}\,dx = \frac{\pi\beta a}{\sqrt{a^2+b^2}}\,H_{o,1}^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{a^2+b^2}) \tag{40}$$

Substitution of $H_{o,1}^{(1)}(\beta e^{\pi i/2}\sqrt{a^2+\xi^2})$ from (40) transforms the single integral

$$I_4 = \int_{-\infty}^{\infty} \frac{\pi\beta a}{\sqrt{a^2+\xi^2}} H_{o,1}^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{a^2+\xi^2})\ H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+(y-\xi)^2}\ \ d\xi \qquad (41)$$

to the repeated integrals

$$I_4 = \int_{-\infty}^{\infty} H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+(y-\xi)^2})\ d\xi \int_{-\infty}^{\infty} e^{-a\sqrt{t^2+\beta^2}\ -i\xi t}\ dt\ ,$$

which, on changing the order of the integrations, becomes

$$I_4 = \int_{-\infty}^{\infty} e^{-a\sqrt{t^2+\beta^2}}\ dt \int_{-\infty}^{\infty} e^{-i\xi t}\ H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+(y-\xi)^2})\ d\xi\ .$$

Letting $\xi = \eta - y$, and changing a to $-t$, (23) becomes

$$\int_{-\infty}^{\infty} e^{-it\eta}\ H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{x^2+(y-\eta)^2})\ d\eta = \frac{2}{i\sqrt{t^2+\beta^2}}\ e^{-ity\ -\sqrt{t^2+\beta^2}\ x}$$

Using the last integral to carry out the internal integration, I_4 becomes

$$I_4 = \frac{2}{i}\int_{-\infty}^{\infty} \frac{1}{\sqrt{t^2+\beta^2}}\ e^{-(a+x)\sqrt{t^2+\beta^2}\ -\ ity}\ dt\ ,$$

which integrates to

$$I_4 = 2\pi\ H_o^{(1)}(\beta e^{\frac{\pi i}{2}}\sqrt{(a+x)^2+y^2}) \qquad (42)$$

by use of Formula (2). Combining (41) and (42), (38) is proved.

Letting $\beta = 1$ in (38) we find

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{a^2+\xi^2}} K_o'(\sqrt{a^2+\xi^2})K_o(\sqrt{x^2+(y-\xi)^2}) = -\frac{\pi}{a}\ K_o(\sqrt{(a+x)^2+y^2} \qquad (43)$$

This formula was derived by Kerr [5] with his indirect method.

VII. DEFLECTION OF THE FLOATING ICE PLATE. Expressed in nondimensional form, the differential equation governing the deflection w of a floating ice plate sustaining a concentrated load P at the origin ($x = 0$, $y = 0$) is

$$\nabla^4 w + w = P\delta(x)\delta(y) \ , \tag{44}$$

where Δ^2 is the Laplacian operator and $\delta(\ )$ the delta function. The solution of (44) for an infinite plate is

$$w(x,y) = -\frac{P}{2\pi} \operatorname{kei}\sqrt{x^2+y^2} \ , \tag{45}$$

as was shown by Wymann [10] by examining the nature of the solution of the homogeneous form of (44). However, his proof does not exactly show that (45) is the solution of the inhomogeneous equation (44). We now can prove this by direct substitution of (45) into (44).

Use of Formula (4) enables us to derive

$$\operatorname{kei}\sqrt{x^2+y^2} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{(\xi^2+\eta^2)^2+1} e^{+ix\xi+iy\eta} \, d\xi d\eta \ . \tag{46}$$

With the use of (46), substitution of (45) reduces the left-hand side of (44) to

$$\frac{P}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{+ix\xi+iy\eta} \, d\xi d\eta \ ,$$

which by the property of the delta function [2], is the right-hand side of (44).

Use of the reciprocal Fourier relationships developed in this paper open the possibility of building an analytical machinery for solving the generic equation

$$\nabla^4 w + w = P\delta(x - x_o)\delta(y - y_o) \tag{47}$$

for the deflection of floating plates of various shapes and of various boundary conditions, which has hitherto been impossible. (Currently only the image method is used, see [3] or [6]). It is especially encouraging to note that the solution $w(x,y)$ of (47) is a generalized function [2].

REFERENCES

[1] A. Erdélyi, et al, ed. Tables of Integral Transforms. McGraw-Hill, New York. 1954. Vol. 1, p. 56, No. (43).

[2] I.M. Gel'fand and G.E. Shilov. Generalized Functions. Vol. 1. Properties and Operations. (Translated by E. Saletan) Academic Press, New York, 1964.

[3] A.D. Kerr, Elastic plates on a liquid foundation. Journal of the Engineering Mechanics Division, Proceedings of the American Society of Civil Engineers, 89, No. EM3, June 1963, pp. 59-71.

[4] A.D. Kerr. An indirect method for evaluating certain infinite integrals. Journal of Applied Mathematics and Physics (ZAMP), 29, 1978, pp. 380-386. Also, Princeton Univeristy Research Report No. 77-SM-3, January 1977.

[5] A.D. Kerr. An evaluation of certain integrals related to Bessel functions. 2nd Quarterly Report, CRREL Contract DACA 89-077,0050, March 1977.

[6] A.D. Kerr. The clamped, semi-infinite, floating plate subjected to a vertical force. 2nd Quarterly Report, CRREL Contract DACA 89-077-0050, March 1977.

[7] N.N. Lebedev. Special Functions and Applications. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965. p. 118.

[8] G.N. Watson. A Treatise on the theory of Bessel Functions. Cambridge at the University Press, 2nd ed., 1962.

[9] E. Winkler, Die Lehre von der Elastizitati and Festigkeit. Praga Dominicus, 1867.

[10] M. Wyman. Deflection of an infinite plate. Canadian Journal of Research, A28, 1950, pp. 293-302.

SOME INTRINSIC PROPERTIES OF EXACT SOLUTIONS FOR THE STATIC BENDING OF UNIFORM ROTATING BEAMS*

James T. Wong and Richard M. Carlson
HQ, US Army Research and Technology Laboratories (AVRADCOM)
Ames Research Center
Moffett Field, CA 94035

ABSTRACT. The general solution to the complete static equation characterizing the static behavior of a rotating beam with constant stiffness and uniform mass distribution has been obtained. This general solution is expressed in terms of rotor functions which are the analytical solutions to a special case of the original problem. The feasibility of obtaining the closed form solution attributes to an interesting property of the rotor functions. It was also found that a modulation of the solution in terms of loadings different by a multiplicative constant is allowable for the articulated case, and such a relationship does not hold for the hingeless boundary conditions unless the blade coning angle equals to zero.

*Published in the Journal of the American Helicopter Society, October 1978.

DYNAMIC STABILITY OF COLUMNS SUBJECTED TO NONCONSERVATIVE FORCES

J. J. Wu and J. D. Vasilakis
U. S. Army Armament Research and Development Command
Benet Weapons Laboratory, LCWSL
Watervliet Arsenal, Watervliet, NY 12189

ABSTRACT. The numerical results of a class of problems of linear elastic stability problems subjected to nonconservative forces and under various support conditions are presented here. A single solution formulation by which these results have been obtained is described. Accuracy of these results compared with those reported in the literature is discussed.

I. INTRODUCTION. Any particular subject of investigation in applied sciences is always motivated by the desire to understand some natural phenomena and hopefully to utilize the results of such an investigation for the benefit of human activities. The study of structural behavior under nonconservative loads is of no exception. Since follower forces are a special class of nonconservative forces [1], one is surprised to encounter frequently the question as to the relation between such a study and a real engineering problem. Physically, a follower force is simply one whose direction follows the structural deformation as in comparison with a dead load which acts in a fixed direction independent of deformation. Some obvious examples of followers forces are: thrust at the tail of a flexible rocket, jet engine thrust of an airplane, thrust on the propeller shaft of a ship, etc. Other examples such as the pressure-and-curvature induced forces included in the gun dynamic studies are less obvious [2].

Since the problems of follower forces are non-self-adjoint their treatment is more difficult than that for the self-adjoint problems. In the classical paper by Beck [3], it was demonstrated that the stability nature of a nonconservative problem can be quite different than that of a conservative one. For these reasons, a systematic approach to this class of problems and an understanding of some of the basic problems involving follower forces are desirable.

The purpose of this paper is to present a single solution approach to a class of problems of follower forces, including several classical examples, to present the numerical results so obtained and to discuss the accuracy compared with those already published in literature.

In Section II, the class of problems will be defined by a general form of a differential equation and a set of boundary conditions. The solution formulation and its basis is given in Section III. Numerical results of some specific problems are given in Section IV together with a discussion and comparisons with data available in literature.

II. A CLASS OF PROBLEMS SUBJECTED TO FOLLOWER LOADS. The class of problems considered in this paper can be described by the differential equation

$$y'''' + P(x)y'' + \lambda^2 y = 0 \tag{1}$$

where $y(x)$ denotes the lateral disturbance of a beam, as a function of the abscissa x, $P(x)$ is the axial force always tangent to the deformed axis, and λ is the eigenvalue. As usual, a prime denotes differentiation with respect to x.

Eq. (1) is a non-self-adjoint differential equation (thus nonconservative problem) except for $P(x)$ = constant. If the axial force $P(x)$ remains fixed in the direction of the undeformed axis, the problem would be of conservative nature and the differential equation a self-adjoint one.

$$y'''' + [P(x)y']' + \lambda^2 y = 0 \tag{1'}$$

Both Eqs. (1) and (1') are well known and the derivations are simple and they follow the procedures given in such textbooks as that by Timoshenko and Gere [4]. Boundary conditions considered will be in the following form:

$$y'''(0) + P(0)y'(0) + k_1(0) = 0 \tag{2a}$$

$$-y''(0) + k_2 y'(0) = 0 \tag{2b}$$

$$-y'''(1) - (1-k_5)P(1)y'(1) + k_3 y(1) = 0 \tag{2c}$$

$$y''(1) + k_4 y'(1) = 0 \tag{2d}$$

where k_1, k_2 are the deflection and rotation spring constants at $x = 0$ and k_3, k_4 are the same at $x = 1$. The constant k_5 is related to a "constant of tangency" K_θ by equation

$$K_\theta = k_5 - 1 \tag{3}$$

so that Eq. (2c) becomes

$$-y'''(1) + K_\theta P(1)y'(1) + k_3 y(1) = 0 \qquad (2c')$$

where now, if $P(1) \neq 0$, $\theta = K_\theta y'(1)$ denotes the angle that $P(1)$ is to be rotated with respect to the tangent of the beam at $x = 1$ (Figure 1).

Eqs. (2) simply state that the total shear force and moment at $x = 0$ and $x = 1$ must be zero. As k_1 approaches to infinity, Eq. (2a) requires that $y(0) = 0$. Thus a zero deflection boundary condition is arrived at. Similar options are provided for by other spring constants k_2, k_3 and k_4.

Three different $P(x)$ will be considered in this paper: (1) $P(x) = P$, a constant, (2) $P(x) = q(1-x)$, and (3) $P(x) = q_o/2(1-x)^2$ where P represents a concentrated at $x = 0$, q is a uniformly distributed follower force density and q_o denotes the maximum of a linearly varied follower force density. With the special boundary conditions of a cantilever, case (1), (2) and (3) become the classical problems first solved by Beck [3], Leipholz [5] and Hanger [6], respectively.

III. SOLUTION FORMULATIONS. The solution method used here is the finite element unconstrained variational formulation which has proved to be efficient and simple to use for solutions of non-self-adjoint problems [7,8]. Finite elements are used in the usual sense that the unknown function is approximated by piecewise cubic splines. An unconstrained variational statement is established and used so that none of the boundary conditions need to be satisfied a priori. An outline of the formulation will be given here.

Introducing an adjoint field variable $y^*(x)$, it is a simple matter to see that the following variational statement will lead to the differential equation (1) and boundary conditions (2):

$$\delta I(y,y^*) = 0 \qquad (4a)$$

$$I = \int_0^1 (y''y^{*\prime\prime} - P(x)y'y^{*\prime} - P'(x)y'y^* + \lambda^2 yy^*)dx$$
$$+ k_1 y(0)y^*(0) + k_2 y'(0)y^{*\prime}(0) + k_3 y(1)y^*(1) + k_4 y'(1)y^{*\prime}(1)$$
$$+ k_5 P(1)y'(1)y^*(1) \qquad (4b)$$

The fact that Eqs. (4) lead to the given differential equation and boundary conditions for $y(x)$ independent of $y^*(x)$ implies that one can take the variation of I at $y^*(x) \equiv 0$ and $(\delta I)_{y^* \equiv 0} = 0$ still leads to the original problem. Hence our formulation begins with

$$(\delta I)_{y^* \equiv 0} = 0 \qquad (5a)$$

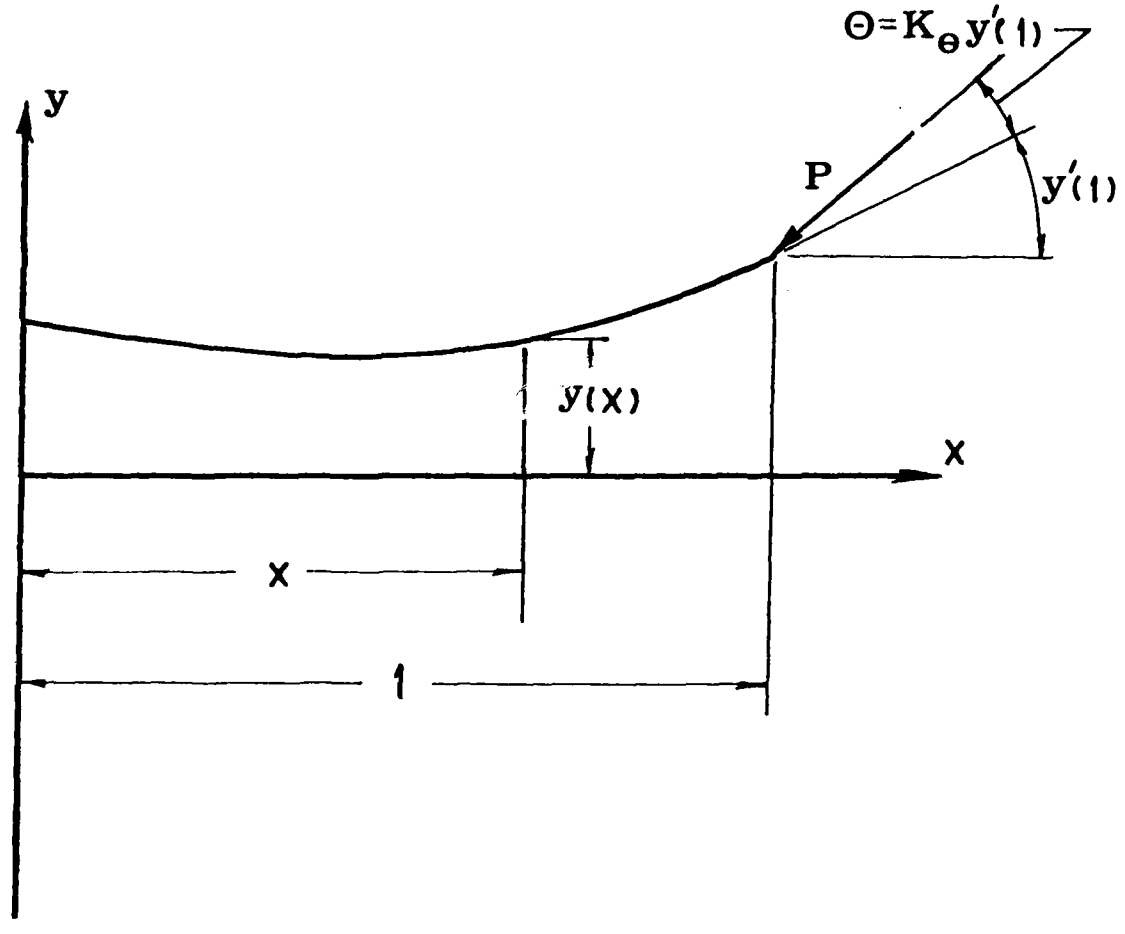


Figure 1. Boundary Condition Associated with a Follower Force: Constant of Tangency K_θ.

Or,

$$\int_0^1 [y''\delta y^{*''} - P(x)y'\delta y^{*'} - P(x)y'\delta y^* + \lambda^2 y\delta y^*]dx$$
$$+ k_1 y(0)\delta y^*(0) + k_2 y'(0)\delta y^{*'}(0) + k_3 y(1)\delta y^*(1) + k_4 y'(1)\delta y^{*'}(1)$$
$$+ k_5 y'(1)\delta y^*(1) = 0 \tag{5b}$$

Finite element discretization enters when the beam is divided into L equal elements and Eq. (5b) is written as

$$\sum_{i=1}^{L} \int_0^1 [L^3 y^{(i)''}\delta y^{*(i)''} - LP^{(i)}(\xi)y^{(i)'}\delta y^{*(i)'}$$
$$- LP^{(i)'}y^{(i)'}\delta y^{*(i)} + \frac{\lambda^2}{L} y^{(i)}\delta y^{*(i)}]d\xi$$
$$+ k_1 y^{(1)}(0)\delta y^{*(1)}(0) + k_2 L^2 y^{(1)'}(0)\delta y^{*(1)'}(0)$$
$$+ k_3 y^{(L)}(1)\delta y^{*(L)}(1) + k_4 L^2 y^{(L)'}(1)\delta y^{*(L)'}(1)$$
$$+ k_5 y^{(L)'}(1)\delta y^{*(L)}(1) = 0 \tag{6}$$

In obtaining Eq. (6) from (5b), one has effected a change of coordinates from x (global) to ξ (local) such that

$$\xi = \xi^{(i)} = Lx - i + 1$$
$$d\xi = Ldx$$
$$y(x) = y^{(i)}(\xi) \tag{7}$$
$$y'(x) = \frac{d}{dx} y(x) = L \frac{d}{d\xi} y^{(i)}(\xi) = Ly^{(i)'}(\xi)$$
etc.

Introducing generalized coordinates vector $\underset{\sim}{Y}^{(i)}$ and shape function vector $\underset{\sim}{a}(\xi)$ such that

$$y^{(i)}(\xi) = \underset{\sim}{a}^T(\xi)\underset{\sim}{Y}^{(i)} \tag{8a}$$

with

$$\underset{\sim}{Y}^{(i)T} = \{Y_1^{(i)} \quad Y_2^{(i)} \quad Y_3^{(i)} \quad Y_4^{(i)}\} \tag{8b}$$

$$\underset{\sim}{a}(\xi) = \begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{Bmatrix} 1 \\ \xi \\ \xi^2 \\ \xi^3 \end{Bmatrix} \tag{8c}$$

where a superscript T denotes the transpose of a matrix. One observes that

$$Y_1^{(i)} = y^{(i)}(0) \quad , \quad Y_2^{(i)} = y^{(i)'}(0)$$
$$Y_3^{(i)} = y^{(i)}(1) \quad , \quad Y_4^{(i)} = y^{(i)'}(1) \tag{8d}$$

The counterparts for $y^*(\xi)$ can be similarly defined.

In terms of $\underset{\sim}{Y}^{(i)}$, $\underset{\sim}{Y}^{*(i)}$, $\underset{\sim}{a}$, Eq. (6) can be written as:

$$\sum_{i=0}^{L} \delta Y^{*(i)T}\{L^3\int_0^1 \underset{\sim}{a}''(\xi)\underset{\sim}{a}''^T(\xi)d\xi - L\int_0^1 P^{(i)}(\xi)\underset{\sim}{a}'(\xi)\underset{\sim}{a}'^T(\xi)d\xi$$
$$- L\int_0^1 P^{(i)'}(\xi)\underset{\sim}{a}(\xi)\underset{\sim}{a}'^T(\xi)d\xi + \frac{\lambda^2}{L}\int_0^1 \underset{\sim}{a}(\xi)\underset{\sim}{a}^T(\xi)d\xi\}\underset{\sim}{Y}^{(i)}$$
$$+ \delta\underset{\sim}{Y}^{*(1)T}\{k_1\underset{\sim}{a}(0)\underset{\sim}{a}^T(0) + k_2L^2\underset{\sim}{a}'(0)\underset{\sim}{a}'^T(0)\}\underset{\sim}{Y}^{(1)}$$
$$+ \delta\underset{\sim}{Y}^{*(L)T}\{k_3\underset{\sim}{a}(1)\underset{\sim}{a}^T(1) + k_4L^2\underset{\sim}{a}'(1)\underset{\sim}{a}'^T(1) + k_5\underset{\sim}{a}(1)\underset{\sim}{a}'^T(1)\}\underset{\sim}{Y}^{(L)} = 0 \tag{9}$$

It will be convenient to define the following matrices:

$$\underset{\sim}{A}_1 = \int_0^1 a(\xi)a^T(\xi)d\xi \quad , \quad \underset{\sim}{A}_2 = \int_0^1 a'(\xi)a'^T(\xi)d\xi$$
$$\underset{\sim}{A}_3 = \int_0^1 a''(\xi)a''^T(\xi)d\xi \quad , \quad \underset{\sim}{A}_4 = \int_0^1 a(\xi)a'^T(\xi)d\xi$$
$$\underset{\sim}{A}_5 = \int_0^1 \xi a'(\xi)a'^T(\xi)d\xi \quad , \quad \underset{\sim}{A}_6 = \int_0^1 a(\xi)a'^T(\xi)d\xi$$
$$A_7 = \int_0^1 \xi^2 a'(\xi)a'^T(\xi)d\xi \tag{10}$$

$$\underset{\sim}{B}_1 = \underset{\sim}{a}(0)\underset{\sim}{a}^T(0) \quad , \quad \underset{\sim}{B}_2 = \underset{\sim}{a}'(0)\underset{\sim}{a}'(0)$$
$$\underset{\sim}{B}_3 = \underset{\sim}{a}(1)\underset{\sim}{a}^T(1) \quad , \quad \underset{\sim}{B}_4 = \underset{\sim}{a}'(1)\underset{\sim}{a}'^T(1)$$
$$\underset{\sim}{B}_5 = \underset{\sim}{a}(1)\underset{\sim}{a}'^T(1)$$

In terms of the matrices defined in (10), Eq. (9) is written as:

$$\sum_{i=1}^{L} \delta\underset{\sim}{Y}^{*(i)T}\{L^3\underset{\sim}{A}_3 + \frac{\lambda^2}{L}\underset{\sim}{A}_1 - L\underset{\sim}{M}_p\}\underset{\sim}{Y}^{(i)}$$

$$+ \delta\underset{\sim}{Y}^{*(1)T}\{k_1\underset{\sim}{B}_1 + k_2L^2\underset{\sim}{B}_2\}\underset{\sim}{Y}^{(1)}$$

$$+ \delta\underset{\sim}{Y}^{*(L)T}\{k_3\underset{\sim}{B}_3 + k_4L^2\underset{\sim}{B}_4 + k_5\underset{\sim}{B}_5\}\underset{\sim}{Y}^{(L)} = 0 \tag{11}$$

where the matrix $\underset{\sim}{M}_p$ is defined as

$$\underset{\sim}{M}_p = \int_0^1 P^{(i)}\underset{\sim}{a}'\underset{\sim}{a}'^T d\xi + \int_0^1 P^{(i)'}\underset{\sim\sim}{a}\underset{\sim}{a}'^T d\xi \tag{12}$$

To proceed further, it is necessary to know the specific form of P(x). As we have mentioned earlier, three different forms of P(x) will be considered.

CASE I. P(x) = P, a Constant. In this case, one has

$$P(x) = P^{(i)}(\xi) = P$$
$$P'(x) = LP^{(i)'}(\xi) = 0 \tag{13}$$

Thus

$$\underset{\sim}{M}_p = P\int_0^1 \underset{\sim}{a}'\underset{\sim}{a}'d\xi = P\underset{\sim}{A}_2 \tag{14}$$

CASE II. P(x) = q(1-x).

$$P(x) = P^{(i)}(\xi) = \frac{q}{L}(L-i+1-\xi)$$
$$P^{(i)'}(\xi) = -\frac{q}{L} \tag{15}$$

Thus,

$$\underset{\sim}{M}_p = \frac{q}{L}\{[L - (i-1)]\int_0^1 \underset{\sim}{a}'\underset{\sim}{a}'^T d\xi - \int_0^1 \xi\underset{\sim}{a}'\underset{\sim}{a}'^T d\xi$$

or

$$\underset{\sim}{M}_p = \frac{q}{L}\{[L - (i-1)]\underset{\sim}{A}_2 - \underset{\sim}{A}_5\} \tag{16}$$

CASE III. $P(x) = q_o/2(1-x)^2$.

$$p^{(i)}(\xi) = \frac{q_o}{2L^2}[(L-i+1)^2 - 2(L-i+1)\xi + \xi^2]$$

$$p^{(i)'}(\xi) = -\frac{q_o}{L^2}[(L-i+1) - \xi] \tag{17}$$

Thus

$$\underset{\sim}{M}_p = \frac{q_o}{2L^2}\{(L-i+1)^2\int_0^1 \underset{\sim}{a'}\underset{\sim}{a'}^T d\xi - 2(L-i+1)\int_0^1 \xi\underset{\sim}{a'}\underset{\sim}{a'}^T d\xi + \int_0^1 \xi^2\underset{\sim}{a'}\underset{\sim}{a'}^T d\xi\}$$

$$- \frac{q_o}{L^2}\{(L-i+1)\int_0^1 \underset{\sim}{a'}\underset{\sim}{a'}^T d\xi - \int_0^1 \xi\underset{\sim}{a'}\underset{\sim}{a'}^T d\xi\}$$

Or,

$$\underset{\sim}{M}_p = \frac{q_o}{2L^2}\{(L-i+1)^2\underset{\sim}{A}_2 - 2(L-i+1)\underset{\sim}{A}_5 + \underset{\sim}{A}_7\}$$

$$- \frac{q}{L^2}\{(L-i+1)\underset{\sim}{A}_2 - \underset{\sim}{A}_5\} \tag{18}$$

With $\underset{\sim}{M}_p$ defined for all three cases in Eqs. (14) (16) and (18) respectively, one can now assemble Eq. (11) into a global matrix equation. Introducing the global generalized coordinate vectors $\underset{\sim}{Y}$ and $\underset{\sim}{Y}^*$ as:

$$\underset{\sim}{Y}^T = \{Y_1^{(1)}\ Y_2^{(1)}\ Y_3^{(1)}\ Y_4^{(1)}\ Y_3^{(2)}\ Y_4^{(2)} \ldots\ldots Y_3^{(L)}\ Y_4^{(L)}\}$$
$$\underset{\sim}{Y}^{*T} = \{Y_1^{*(1)}\ Y_2^{*(1)}\ Y_3^{*(1)}\ Y_4^{*(1)}\ Y_3^{*(2)}\ Y_4^{*(2)} \ldots\ldots Y_3^{*(L)}\ Y_4^{*(L)}\} \tag{19}$$

Eq. (11) now can be written in terms of $\underset{\sim}{Y}$ and $\delta\underset{\sim}{Y}^*$ as

$$\delta\underset{\sim}{Y}^{*T}\{\underset{\sim}{K} - \lambda^2\underset{\sim}{M}\}\underset{\sim}{Y} = 0 \tag{20}$$

Where the global matrices $\underset{\sim}{K}$ and $\underset{\sim}{M}$ are formed by properly placing the local matrices defined in Eqs. (10) according to the correspondence between the local and global generalized coordinates indicated in Eqs. (19). Now since $\delta\underset{\sim}{Y}^*$ are not subject to any constraint conditions, Eq. (18) reduces to

$$(\underset{\sim}{K} - \lambda^2\underset{\sim}{M})\underset{\sim}{Y} = 0 \tag{21}$$

which is solved for the eigenvalue λ and the eigenvector $\underset{\sim}{Y}$.

IV. NUMERICAL RESULTS AND DISCUSSION. It is well known that the eigenvalue λ dictates the stability of the column: a pure imaginary λ is associated with a stable vibration, a real λ, associated with instability of divergence and a complex λ, with instability of flutter [10].

Only cantilevered columns will be considered here. It will be seen that in all three loading cases, the cantilevered columns reaches an instability condition of flutter.

CASE I. $P(x) = P = \text{Constant}$. The characteristic equation in close form was obtained by Beck [3] as

$$2\lambda^2 + Q^2 + 2\lambda^2 \cosh\alpha\cos\beta + Q\lambda\sinh\alpha\sin\beta = 0 \tag{22}$$

where

$$\alpha^2 = \sqrt{\lambda^2 + \frac{Q^2}{4}} + \frac{Q}{2}$$

$$\beta^2 = \sqrt{\lambda^2 + \frac{Q^2}{4}} + \frac{Q}{2} \tag{23}$$

For a given Q, the eigenvalue λ can be calculated from Eq. (22) and there are an infinite number of λ solutions for each Q. Eq. (22) is solved for two lowest branches of λ using an iterative procedure. The results are given in Table I. The critical load thus obtained is

$$Q_{CR} = 2.0318\pi^2 = 20.053$$

which agrees well with the value obtained originally by Beck as $Q_{CR} = 20.05$. The results for four lowest eigenvalues presently obtained using our finite element-unconstrained variational formulations are also shown in Table I. The first two branches obviously agree well with those from the exact characteristic equation. It should be pointed out that the numerical solutions to the Beck problem given in Reference [4] appears to be inaccurate. A plot of the eigenvalue curve showing the coalescense of the two lowest branches is given in Figure 2. The data from Reference [4] are indicated by small circles. The fact that these data points do not fall on a smooth curve further add to the doubt on their accuracy.

TABLE I. NUMERICAL VALUES OF FIRST TWO LOWEST EIGENVALUES OF A CANTILEVERED COLUMN WITH P(x) = P, A CONSTANT

	Q/π^2	0.	0.5	1.0	1.5	2.0	2.0318
λ_1	Present Results	3.5160	4.2072	5.1462	6.5546	9.8256	11.0167
	Exact	3.5160	4.2072	5.1461	6.5546	9.8282	
	Timo. & Gere	3.4894	5.0325	5.4158	6.6939	9.6702	
λ_2	Present Results	22.0356	20.4590	18.6410	16.3684	12.2599	11.0167
	Exact	22.0345	20.4578	18.6395	16.3665	12.2545	
	Timo. & Gere	21.7579	20.2266	17.9290	15.9143	9.9678	
	λ_3	61.7209	59.8566	57.9304	55.9366	53.8689	53.7348
	λ_4	121.0745	119.0413	116.9720	114.8648	112.7177	112.5797

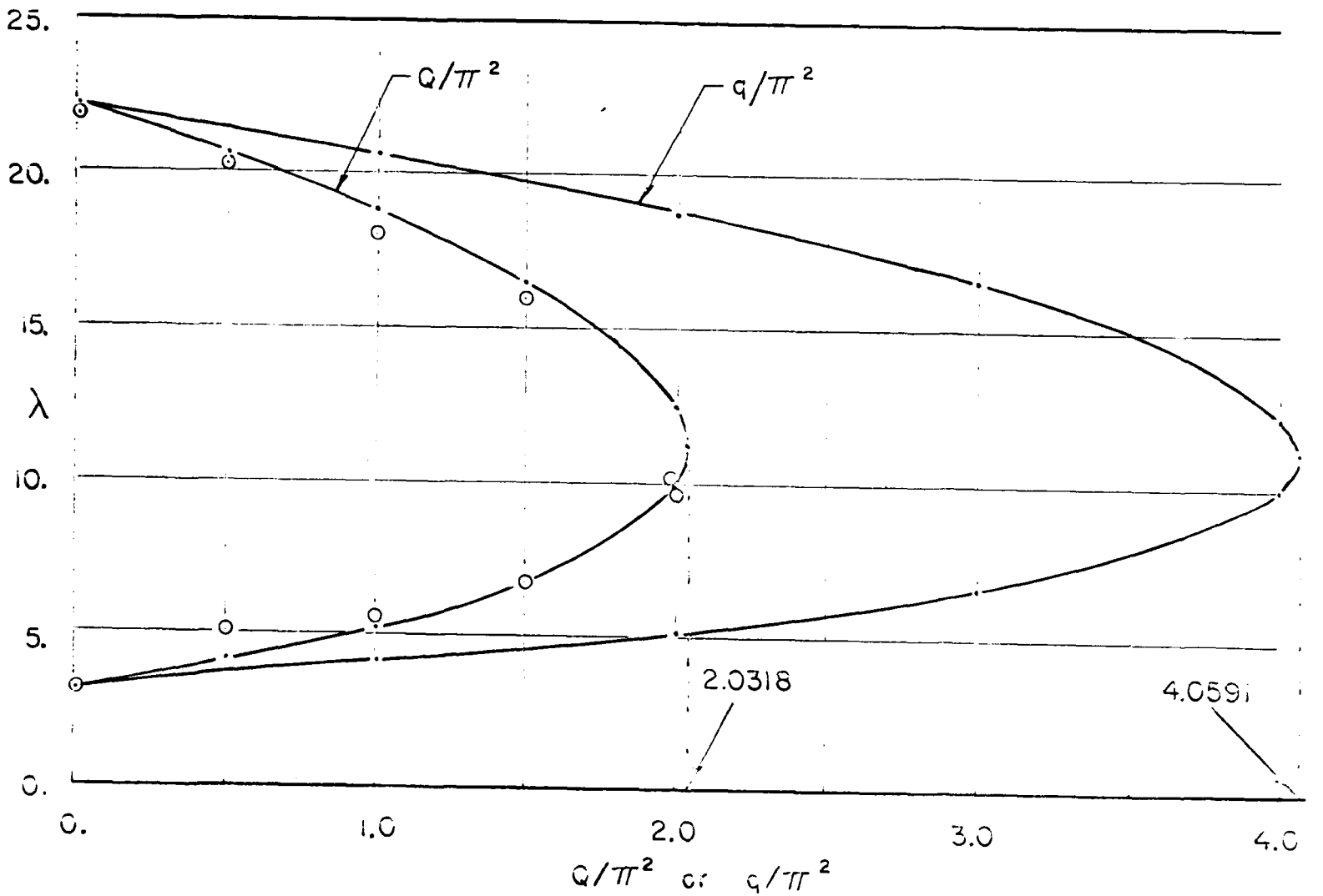


Figure 2. Two Lowest Eigenvalues vs. Load Parameters Q/π^2 and q/π^2.
A Cantilevered Column of Case I and II.
(Data shown in dots (°) are from Reference [4])

CASE II. $P(x) = q(1-x)$. The numerical values of the four lowest eigenvalues up to the first critical load are given in Table II which the critical load is shown to be

$$q_{CR} = 4.0591\pi^2$$

compared with data given by Leipholz as $4.1238\pi^2 = 40.7$ [10] and again as $4.2058\pi^2 = 41.51$ [11]. The coalescence of the first two branches of eigenvalues is again shown in Figure 2.

CASE III. $P(x) = q_o/2(1-x)^2$. Similar data for this case are presented in Table III and in Figure 3. The critical load of flutter is obtained as

$$q_{oCR} = 15.2687\pi^2$$

In comparison, the value obtained by Hauger was $q_{oCR} = 158.2 = 16.092\pi^2$ [6] and that by Leipholz, $q_{oCR} = 150.80 = 15.279\pi^2$ [12].

REFERENCES

1. H. Ziegler, Principles of Structural Stability, Blaisdell, 1968, p. 33.

2. J. J. Wu, "Gun Dynamics Analysis by the Use of Unconstrained, Adjoint Variational Formulations," Proceedings of the Second U.S. Army Symposium on Gun Dynamics, September 1978, pp. II80-II99.

3. M. Beck, "Die Knicklast des einseitig einseitig eingespannten, tangential gedruckten Stabes," ZAMP, 1952, Vol. 52, pp. 225-229.

4. S. P. Timoshenko and J. M. Gere, Theory of Elastic Stability, McGraw-Hill, 1961, p. 3.

5. H. Leipholz, "Anwendung des Galerkinschen Verfahren auf nichtkonservative Stabilitatsprobleme des elastischen Stabes," ZAMP, 1962, Vol. 13, pp. 359-372.

6. W. Hauger, "Die Knicklasten elastischer Stabe under gleichmassig verteilten under linear veranderlichen, tangentialen Druckfreften," Ingenieur Archiv, 1966, Vol. 35, pp. 221-229.

7. J. J. Wu, "A Unified Finite Element Approach to Column Stability Problem," Development in Mechanics, Vol. 8, 1975, pp. 279-294.

8. J. J. Wu, "On Missile Stability Journal of Sound and Vibration," 1976, Vol. 49(1), pp. 141-147.

9. S. P. Timoshenko and J. M. Gere, Theory of Elastic Stability, McGraw-Hill, 1961, p. 155.

10. D. A. Peters and J. J. Wu, "Asymptotic Solution to a Stability Problem," Journal of Sound and Vibration, 1978, Vol. 59(4), pp. 591-610.

11. H. Leipholz, Stability Theory, Academic Press, 1972, pp. 259-241.

12. H. Leipholz, "On the Calculation of Buckling Loads by Means of Hybrid Ritz Equations," Archives of Mechanics, Warszawa 1973, Vol. 25(6), pp. 895-901.

13. H. Leipholz, "On the Solution of the Stability Problem of Elastic Rods Subjected to Triangularly Distributed Tangential Follower Forces," Ingenier Archiv, 1977, Vol. 46, pp. 115-124.

TABLE II. NUMERICAL VALUES OF FIRST FOUR LOWEST EIGENVALUES OF A CANTILEVERED COLUMN WITH $P(x) = q(1-x)$

Q/π^2	0.	1.0	2.0	3.0	4.0	4.05907
λ_1	3.5160	4.2079	5.1499	6.5660	9.8811	11.0315
λ_2	22.0356	20.4587	18.6399	16.3664	12.2289	11.0315
λ_3	61.7209	59.8516	57.9059	55.8772	53.7547	53.6261
λ_4	121.0745	119.0382	116.9586	114.8332	112.6595	112.5295

TABLE III. NUMERICAL VALUES OF FIRST FOUR LOWEST EIGENVALUES OF A CANTILEVERED COLUMN WITH $P(x) = q_o(1-x)^2/2$

Q/π^2	0.	4.0	8.0	12.0	15.26866
λ_1	3.5160	4.3170	5.4413	7.2148	11.4874
λ_2	22.0356	20.4456	18.5880	16.1747	11.4874
λ_3	61.7209	59.5525	57.2512	54.7927	52.6444
λ_4	121.0745	118.6056	116.0544	113.4143	111.1856

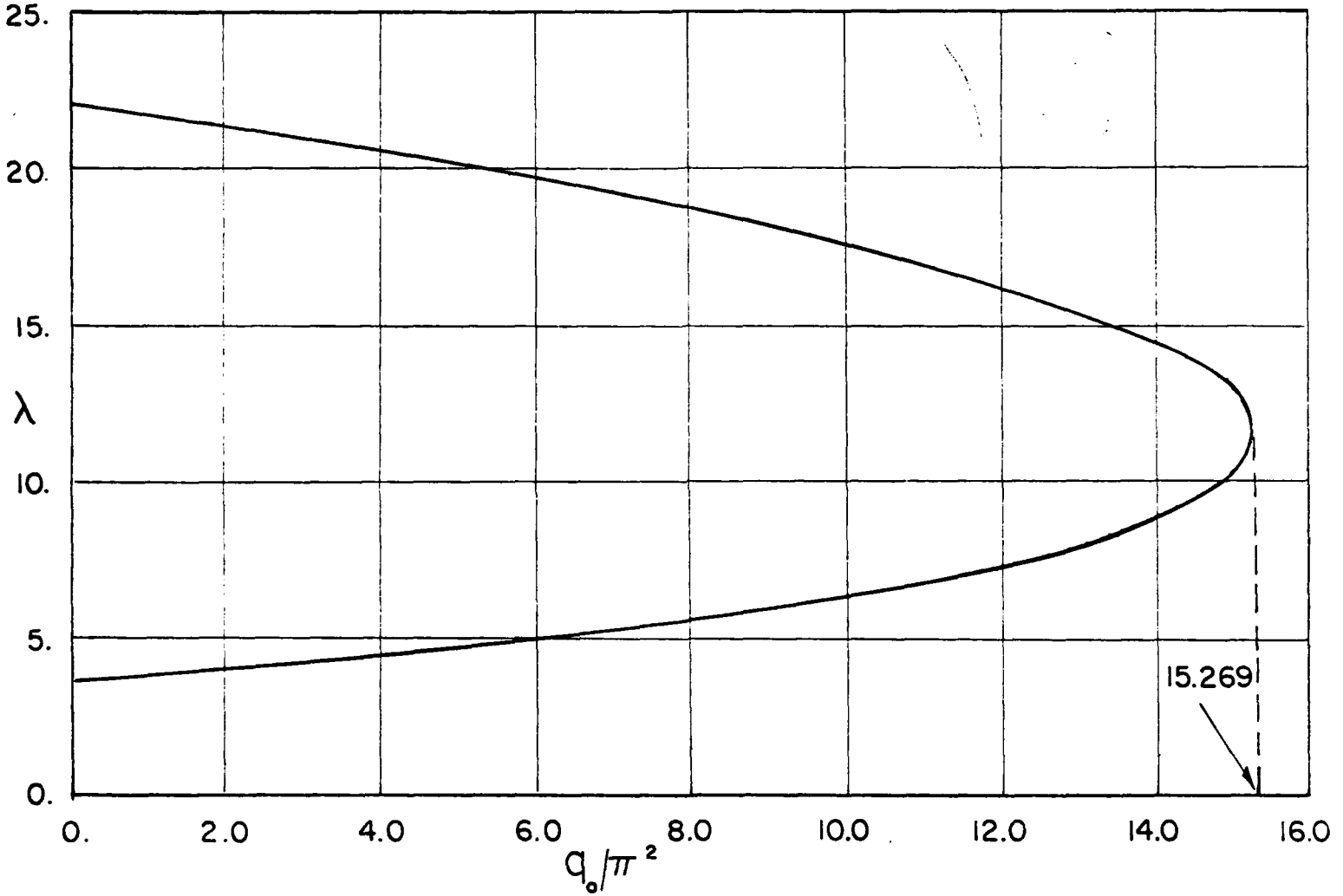


Figure 3. Two Lowest Eigenvalues vs. Load Parameter q_o/π^2
A Cantilevered Column of Case III.

THERMO-ELASTIC-PLASTIC STRESSES IN HOLLOW CYLINDERS DUE TO QUENCHING

J. D. Vasilakis and P. C. T. Chen
U.S. Army Armament Research and Development Command
Benet Weapons Laboratory, LCWSL
Watervliet Arsenal, Watervliet, NY 12189

ABSTRACT. A hollow cylindrical tube, rapidly quenched for the purpose of developing a high strength material structure, is analyzed. The quenching creates severe thermal stresses early in the quenching cycle while later the material transformation by virtue of a volume change in the transformed material, causes large transformation stresses. The transient temperature distributions and the elastic treatment of the stresses has been treated previously. The present work is an attempt to consider the thermo-elastic-plastic aspects of the problem. The von Mises yield criterion and the Prandtl-Reuss stress strain relations are used. Results are calculated based on a new finite difference approach.

I. INTRODUCTION. Watervliet Arsenal has recently been developing techniques for the production of large caliber weapons using a rotary forge. Force hammers, evenly spaced at 90° intervals, strike the outside diameter of a hot (1500°F-1600°F) hollow cylindrical preform at the rate of 200 blows/minute. The final outside tube profile is programmed into the forge itself and the wall thickness of the tube is varied as preprogrammed. The inside diameter of the tube is maintained constant by a mandrel which is water cooled. After the tube has been formed, it is allowed to cool to room temperature.

Once formed, the tube must then be heat treated. This procedure begins by heating the tube to 1650°F in an austenitizing furnace so that the austenite phase is developed throughout the material. The tube is then rapidly quenched so that the desired martensite phase is developed. This is accomplished by spraying a large volume of water on both the inside and outside diameters. The tube is finally put through a tempering furnace at 1200°F.

Interest in the analytical studies of the process first arose when cracks begin developing in the tube during quenching. While the possible causes of quench-cracking are many, most often they are associated with the material used. It was also decided, however, to look into the transient temperatures during the quenching process and the thermal and transformation stresses involved. The transformation stresses occur mainly due to volume changes in the material as it transforms from one

phase to another. As they are due to volume changes, transformation stresses can be treated in a manner similar to the thermal stresses. Although quench-cracking cannot be predicted from a study such as this, a better understanding of the quenching procedure will emerge and the relative severity of different quenching procedures would be known.

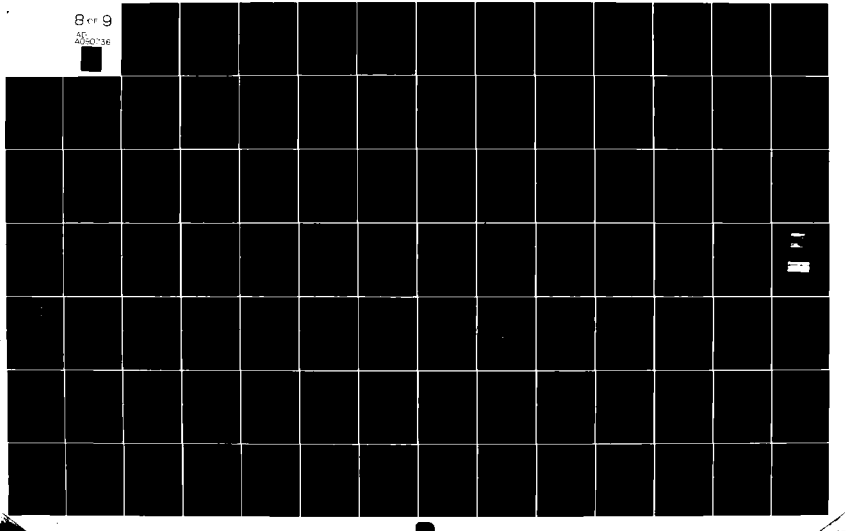
The transient temperatures and the zones of transformed material assuming a linear relationship for the change in volume between the martensite start and finish temperatures were treated in [1]. This reference also considers the thermal and transformation stresses assuming the stresses remain elastic. The present work seeks to incorporate an elastic-plastic stress analysis into the problem. In view of the previous results, this assumption is more realistic. The temperature and stress problem are considered uncoupled.

II. PROBLEM DESCRIPTION. The problem being considered is that of the elastic-plastic stresses developed during the quenching process. These stresses are due to both the transient temperatures that exist and the transformation stresses.

Most of the elastic-plastic analysis work on thick-wall cylinders concerns itself with mechanical loadings. Bland [2] does consider thermal loads on a thick wall tube. Tresca's yield criterion and its associated flow rule was used to obtain solutions to tubes of work-hardening material subjected to both internal and external pressures. The temperatures, however, are steady state and the thermal stresses due to this steady state temperature distribution are first calculated and assumed elastic. External or internal pressures are then applied until some desired plastic state is arrived at. S. C. Chu [3] used the incremental approach for solving the problem of elastic-plastic thick-walled tubes subject to transient thermal loadings. The von Mises yield criterion and Prandtl-Reuss equations are used.

In the area of elastic-plastic analysis for transformation stresses, the bulk of the work comes from a series of papers by Zwicky, Landau, Weiner, and Huddleston [4-6]. Of those that consider the cylinder configuration, Weiner and Huddleston [5] used the Tresca yield criterion and the associated flow rule to compute the residual stresses in the cylinders. The problem for the transformation stresses was solved by assuming that the volume expansion of the transformed material was equivalent to that of a temperature discontinuity progressing inward from the surface. They considered a solid cylinder of incompressible material. Landau and Zwicky [6] solved a similar problem using the von Mises yield criterion and its associated flow rule. They assumed a compressible material, the yield point stress to be a function of temperature and included the computation of transient thermal stresses.

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH),(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL



The problem considered here is that of determining the thermo-elastic-plastic stresses and transformation stresses in a cylinder due to quenching. The thermal program developed in [1] was coupled to a program [7] for the computation of elastic-plastic stresses in a thick-walled cylinder subjected to internal and external pressure. The problem is assumed to be axisymmetric.

The computer program for the temperature distribution allows for a transient analysis with temperature dependent material properties using an implicit finite difference scheme. The computer program for the elastic-plastic stresses uses an incremental approach. It has been altered to include stresses due to thermal loads. The von Mises yield criterion is used with the associated Prandtl-Reuss flow rule. The material is assumed compressible and is capable of work-hardening although for this work the material was assumed to be elastic-perfectly plastic.

III. THERMAL EQUATIONS. The partial differential equation for the temperature (T) in a thick-wall cylinder with inner radius, a, and outer radius, b, is given in dimensionless form by

$$\frac{1}{r}\frac{\partial}{\partial r}\left[k(T)r\frac{\partial T}{\partial r}\right] = c(T)\rho(T)\frac{\partial T}{\partial t} \tag{1}$$

where r is dimensionless radial distance, k(T), c(T), ρ(T) are dimensionless thermal conductivity, specific and density, respectively, and t is dimensionless time. The dimensionless quantities are defined as

$$r = \frac{\bar{r}}{b}, \quad T = \frac{\bar{T}-T_o}{T_i-T_o}$$

$$t = \frac{k_o}{\rho_o c_o b^2}\bar{t} \tag{2}$$

$$k(T) = k_o K(T), \quad c(T) = c_o C(T), \quad \rho(T) = \rho_o R(T)$$

and $\bar{r}$ is the radius, $\bar{T}$ is the temperature, k_o, c_o, ρ_o are thermal conductivity, specific heat and density at reference ambient temperature T_o, and T_i is initial temperature and $\bar{t}$ is time.

The boundary conditions are written as

$$\frac{\partial T}{\partial r} - h_1 T = -g_1 \quad \text{at } r = a/b$$

and

$$\frac{\partial T}{\partial r} - h_2 T = -g_2 \quad \text{at } r=1 .$$

With the boundary conditions expressed in this manner, different conditions at the boundary can be specified. If, e.g., $g_1 = 0$ and $h_1 \neq 0$ and finite, then a convection type boundary condition exist on the inner surface. If h_1 was very large and $g_1 = 0$, then $T = 0$ is specified. If h_2 and g_2 are both large and not equal, then the temperature $T = g_2/h_2$ is specified in the outer surface.

IV. STRESS EQUATIONS. The use of finite difference equations to solve the thermo-elastic-plastic stress problem requires expressing the equilibrium equation and the equation of compatibility at each node at which the finite difference equations are desired. The Prandtl-Reuss flow rule is used to eliminate the incremental stresses so that what results is a matrix for evaluating the incremental radial and tangential strains at each node. The required equations follow, written in dimensionless form. The problem is treated as plane strain.

The equation of equilibrium is written

$$\frac{\partial \sigma_r}{\partial r} + \frac{\sigma_r - \sigma_\theta}{r} = 0 \tag{3}$$

where

$\sigma_r \left(= \frac{\bar{\sigma}_r}{\sigma_o}\right)$ is the dimensionless radial stress

$\sigma_\theta \left(= \frac{\bar{\sigma}_\theta}{\sigma_o}\right)$ is the dimensionless tangential stress

and σ_o is the yield stress in tension, and the compatibility equation

$$\frac{\partial \varepsilon_\theta}{\partial r} + \frac{\varepsilon_\theta - \varepsilon_r}{r} = 0 \tag{4}$$

where

$\varepsilon_\theta \left(= E \frac{\bar{\varepsilon}_\theta}{\sigma_o}\right)$ is dimensionless tangential strain

$\varepsilon_r \left(= E \frac{\bar{\varepsilon}_r}{\sigma_o}\right)$ is dimensionless radial strain

and E/σ_o is yield strain in tension when E is Young's Modulus. The compressibility of the material is expressed by

$$\varepsilon = \alpha T + \frac{\sigma}{3K} \tag{5}$$

where

$\varepsilon = \frac{1}{3}(\varepsilon_r + \varepsilon_\theta)$ is mean strain

$\sigma = \frac{1}{3}(\sigma_r + \sigma_\theta + \sigma_z)$ is mean stress

$K \left(= \frac{\bar{K}}{\sigma_o}\right)$ is dimensionless bulk modulus

$\alpha(= \alpha T_i)$ is dimensionless coefficient of thermal expansion

and

$$\varepsilon_z = 0 \text{ for plane strain.}$$

Traction free boundary conditions are used

$$\sigma_r = 0 \text{ at } r = a/b \text{ and } r = 1. \tag{6}$$

It was desirable to write the finite difference equations in terms of strain alone, hence, the stresses in the equations of equilibrium had to be expressed in terms of the strains. This was accomplished by modifying a plastic stress-strain matrix [8] which was derived by inverting the Prandtl-Reuss equations. The inverted Prandtl-Reuss equation is

$$\{d\sigma\} = [D^P]\{d\varepsilon\} - \frac{E\alpha dT}{(1-2\nu)\sigma_o}\{1\} \tag{7}$$

where the stress vector is $\{d\sigma\} = \{d\sigma_r, d\sigma_\theta, d\sigma_z\}^T$, the strain vector $\{d\varepsilon\} = \{d\varepsilon_r, d\varepsilon_\theta, 0\}^T$ and $\{1\}$ represents a unit vector. The plastic stress-strain matrix $[D^P]$ is given by

$$[D^P] = \frac{1}{1+\nu}\begin{bmatrix} \frac{1-\nu}{1-2\nu} - \frac{\sigma_r'^2}{S} & & \text{SYMMETRIC} \\ \frac{\nu}{1-2\nu} - \frac{\sigma_r'\sigma_\theta'}{S} & \frac{1-\nu}{1-2\nu} - \frac{\sigma_\theta'^2}{S} & \\ \frac{\nu}{1-2\nu} - \frac{\sigma_r'\sigma_z'}{} & \frac{\nu}{1-2\nu} - \frac{\sigma_\theta'\sigma_z'}{S} & \frac{1-\nu}{1-2\nu} - \frac{\sigma_z'^2}{S} \end{bmatrix} \tag{8}$$

The primed stresses are deviatoric stresses,

$$\sigma_i' = \sigma_i - \frac{1}{3}\sigma \qquad i = r,\theta,z. \tag{9}$$

At each node during a computation, the von Mises yield criterion

$$\frac{1}{2}[(\sigma_r-\sigma_\theta)^2 + (\sigma_\theta-\sigma_z)^2 + (\sigma_z-\sigma_r)^2] = 1 \tag{10}$$

is checked to see if plastic deformation has progressed to that node. If not, the stresses remain elastic, and can still be computed using (8) by setting the deviatoric stresses equal to zero. The matrix $[D^P]$ then becomes the same matrix as would exist if linear elastic behavior had been assumed. The quantity S is given by

$$S = \frac{2}{3}\bar{\sigma}^2\left(1 + \frac{H'}{3G}\right) \tag{11}$$

where

$$\bar{\sigma} = \frac{3}{2}\sigma_{ij}'\sigma_{ij}' = \frac{3}{2}(\sigma_r'^2\ \sigma_\theta'^2\ \sigma_z'^2) \tag{12}$$

is the equivalent stress and

$$H' = \frac{d\bar{\sigma}}{d\bar{\varepsilon}_p} \tag{13}$$

is the slope of the equivalent stress/equivalent plastic strain curve and is a measure of hardening. The increment in equivalent plastic strain is given by

$$d\bar{\varepsilon}_p = \frac{2}{3}\,d\varepsilon_{ij}^p\,d\varepsilon_{ij}^p \tag{14}$$

V. NUMERICAL COMPUTATIONS. The Crank-Nicolson representation for finite differences of the partial differential equation governing the temperatures in time is [1]

$$[(a+i\Delta r)k_{i+\frac{1}{2},n+\frac{1}{2}}]T_{i+1,n+1} +$$

$$+[-(a+i\Delta r)k_{i+\frac{1}{2},n+\frac{1}{2}} - (a+(i-1)\Delta r)k_{i-\frac{1}{2},n+\frac{1}{2}} - c_{i,n+\frac{1}{2}}p_{i,n+\frac{1}{2}}\left(\frac{2\Delta r^2}{\Delta t}\right)(a+(i-\tfrac{1}{2})\Delta r)]T_{i,n+1}$$

$$+ [(a+(i-1)\Delta r)k_{i-\frac{1}{2},n+\frac{1}{2}}]T_{i-1,n+1} = [-(a+i\Delta r)k_{i+\frac{1}{2},n+\frac{1}{2}}]T_{i+1,n} +$$

$$+ [(a+i\Delta r)k_{i+\frac{1}{2},n+\frac{1}{2}} + (a+(i-1)\Delta r)k_{i-\frac{1}{2},n+\frac{1}{2}} - c_{i,n+\frac{1}{2}}p_{i,n+\frac{1}{2}}\left(\frac{z\Delta r^2}{\Delta t}\right)(a+(i-\tfrac{1}{2})\Delta r)]T_{i,n}$$

$$+ [-(a+(i-1)\Delta r)k_{i-\frac{1}{2},n+\frac{1}{2}}]T_{i-1,n} \tag{15}$$

The equation is solved twice,

1. At $n+\frac{1}{2}$ step, allowing k,p,c etc. to take on the values at t=n step.

2. The new temperatures are then used to evaluate k,c,p, at $n+\frac{1}{2}$ step and the set of equations re-evaluated for the temperatures at the n+1 step.

The computed temperature distributions at each full time step are saved on disk and eventually called in when required by the stress program.

The finite difference equations are (for solid cylinder). Compatibility:

$$-r_i\Delta\varepsilon_{\theta_{i-1}} + (2r_i-r_{i-1})\Delta\varepsilon_{\theta_i} - (r_i-r_{i-1})\Delta\varepsilon_{r_i} =$$

$$-r_i(\varepsilon_{\theta_i}-\varepsilon_{\theta_{i-1}}) - (r_i-r_{i-1})(\varepsilon_{\theta_i}-\varepsilon_{r_i}) \tag{16}$$

Equilibrium:

$$-r_i\Delta\sigma_{r_{i-1}} - (r_i-r_{i-1})\Delta\sigma_{\theta_i} + (2r_i-r_{i-1})\Delta\sigma_{r_i} =$$

$$-r_i(\sigma_{r_i}-\sigma_{r_{i-1}}) - (r_i-r_{i-1})(\sigma_{r_i}-\sigma_{\theta_i}) \tag{17}$$

Substituting the Prandtl-Reuss equations into that of equilibrium

$$-r_iD(r,\theta)\Delta\varepsilon_{\theta_{i-1}} - r_iD(r,r)\Delta\varepsilon_{r_{i-1}} + [-(r_i-r_{i-1})D(\theta,\theta)+(2r_i-r_{i-1})D(r,\theta)]\Delta\varepsilon_{\theta_i}$$

$$+ [-(r_i-r_{i-1})D(\theta,r)+(2r_i-r_{i-1})D(r,r)]\Delta\varepsilon_{r_i}$$

$$r_i[\sigma_{r_{i-1}}-\sigma_{r_i}]+(r_i-r_{i-1})(\sigma_{\theta_i}-\sigma_{r_i})+r_i\frac{E\alpha}{1-2\nu}[\Delta T_i-\Delta T_{i-1}] \tag{18}$$

at i = 1 (zero radius for solid cylinder)

$$-\Delta\varepsilon_{\theta_1} + \Delta\varepsilon_{r_1} = \varepsilon_{\theta_1} - \varepsilon_{r_1} \tag{19}$$

at i = n (or outside boundary) $\sigma_r = 0$ or

$$D(r,\theta)\Delta\varepsilon_{\theta_n} + D(r,r)\Delta\varepsilon_{r_n} = \frac{E\alpha\Delta T_n}{1-2\nu} \tag{20}$$

For the hollow cylinder, a boundary condition similar to i = n can be written for i = 1.

The solution procedure for the transient temperature problem is as follows. The temperature problem is solved and the temperature distributions at their computation times are stored on disk. These distributions are called into the thermo-elastic-plastic stress program one at a time. The corresponding thermal stresses are calculated and each

node checked to see if the yield criterion is satisfied. If not, the problem is still assumed to be elastic, a new temperature distribution is called in and new stress increments calculated. The stresses are updated and the yield criterion checked again. When the stresses at a point are found to satisfy the yield criterion, the node is identified and the stress increments at that node from the next set of temperatures are computed using the Prandtl-Reuss equation or $[D^P]$ matrix identified earlier. This is continued with new sets of temperatures called in and with the tracking of the elastic-plastic boundary(s) with time. The resultant stresses that exist after a steady-state or uniform temperature distribution is reached are the residual stresses.

The solution procedure for the transformation stresses is similar and will be described in the next section.

VI. RESULTS AND DISCUSSION. Several runs were made for the stresses due to the transient temperatures and for the transformation in both solid and hollow cylinders. For the results presented here, the following data was used:

$E = 30\text{x}10^6$ psi, $\sigma_o = 30\text{x}10^3$ psi

$\bar{\alpha} = 7.75\text{x}10^{-6}/°F$, $\bar{T}_i$ - 1250°F $\rightarrow$ $\alpha = \bar{\alpha}\bar{T}_i = .0097$

$\nu = .3$

$h_2 = 12.2$, $h_1 = 12.2$ and 6.1.

The first results shown are those for the transformation stresses. These stresses can be computed using the thermal stress formulation if one replaces αT or the thermal expansion by the linear expansion of the transformation. If the material expansion due to the transformation is isotropic, this linear change is 1/3 the volume expansion. As an example, the volume expansion in going from the austenite to the martensite structure for steel is about 3%-4%. In the quenching of a solid cylinder, the transformation begins on the outer surface and progresses inward to the center. Figure 1 shows the residual stresses in a solid cylinder due to a transformation occurring in the material. The insert shows the temperature function as it progresses inward. It is of unit height in the transformed material and zero in the untransformed material. The transformation is assumed to be occurring over eight nodes or 8% of the cylinder (indicated by N in Figure), and a linear relation is assumed over this length. Initially, σ_θ is compressive near the outside radius when the transformation just begins as the material wants to expand but is prevented from doing so by the surrounding untransformed material. As the transformation progresses, however, σ_θ slowly changes sign and becomes tensile. The computer run was stopped just before the transformation was complete, and that is the reason for the behavior of the

stresses near the bore. The transformation at r = 0 had just started when the run was stopped.

Figure 2 shows similar results for the hollow cylinder. The assumption is made that the transformation progressed evenly from the inside and outside surfaces. Tangential stresses on the outside surface again were initially compressive and slowly changed to tensile stresses while those in the inside radius always remained compressive.

Figures 3 and 4 show the residual stresses that exist due to the transient temperatures from the quenching process. Figure 3 represents the results for the solid cylinder. The large axial stress due to the plane strain assumption is easily seen. As the quenching begins, the outside surface cools and wants to contract. It is prevented from doing so by the surrounding material and therefore σ_θ is initially a tensile stress. The elastic-plastic boundary begins on the outside surface of the cylinder and moves towards the center.

Figure 4 shows similar results for the hollow cylinder. The figure shows the residual stresses when equal convection type boundary conditions are used on both the inside and outside diameters. These are shown by the solid lines. A comparison is made with the same problem when the convection boundary condition on the inside diameter is decreased by 50%. A dotted line compares the differences in the tangential stress, σ_θ, and a substantial reduction is noted in the residual stress.

The usefulness of these results is thus shown. For the quenching problem, the maximum quench time for the desired metallurgical phase structure to be formed is of interest because it implies that the material will be subjected to slower transient temperatures and smaller residual stresses. Thus, a better understanding of the transient temperatures in the quench tube, the resulting residual stresses, and the effect of the quenching process is gained.

REFERENCES

1. J. D. Vasilakis, "Temperatures and Stresses Due to Quenching of Hollow Cylinders," Transactions of the Twenty-Fourth Conference of Army Mathematicians, ARO Report 79-1, pp. 109-128.

2. D. R. Bland, "Elastoplastic Thick-Walled Tubes of Work-Hardening Material Subject to Internal and External Pressures and to Temperature Gradients," Journal of the Mechanics and Physics of Solids, 1956, V 4, pp. 209-229.

3. S. C. Chu, "A Numerical Thermo-Elastic-Plastic Solution of a Thick-Walled Tube," AIAA Journal, Vol. 12, #2, February, 1974, pp. 176-179.

4. H. G. Landau and J. H. Weiner, "Transient and Residual Stresses in Heat-Treated Plates," Journal of Applied Mechanics, December, 1958, pp. 459-465.

5. J. H. Weiner and J. V. Huddleston, "Transient and Residual Stresses in Heat-Treated Cylinders," Journal of Applied Mechanics, March, 1959, pp. 31-39.

6. H. G. Landau and E. E. Zwicky, Jr., "Transient and Residual Thermal Stresses in an Elastic-Plastic Cylinder," Journal of Applied Mechanics, September, 1960, pp. 481-488.

7. P. C. T. Chen, "A Finite-Difference Appraoch to Axisymmetric Plane-Strain Problem Beyond the Elastic Limit," published in present Transactions of the Twenty-Fifth Conference of Army Mathematicians.

8. Y. Yamada, N. Yoshimura and T. Sakuri, "Plastic Stress-Strain Matrix and Its Application for the Solution of Elastic-Plastic Problems by the Finite Element," International Journal of Mechanical Sciences, 1968, Vol. 10, pp. 343-354.

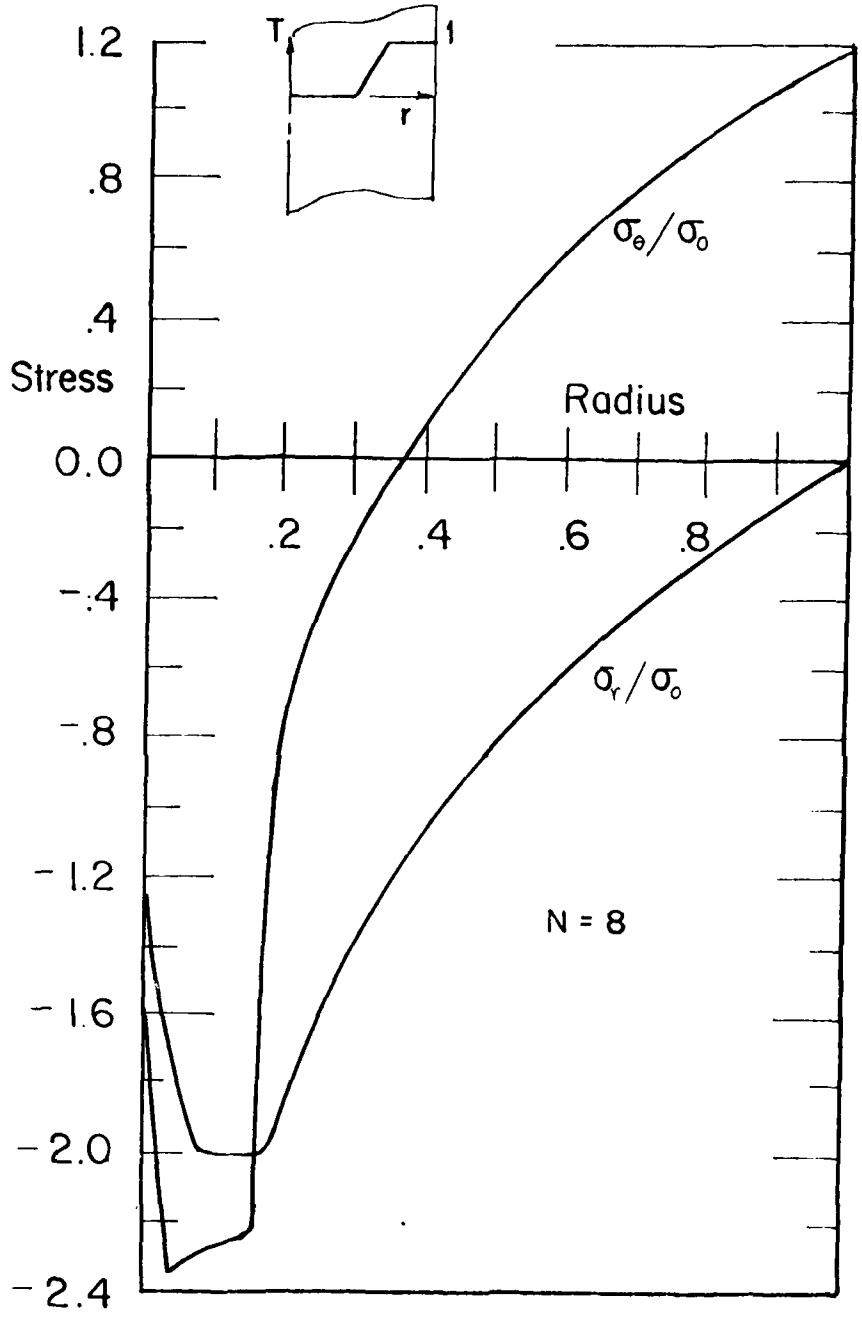


Figure 1. Residual Stresses in a Solid Cylinder Due to Material Transformation (Transformation Beginning on Outside Diameter and Progressing Toward Center).

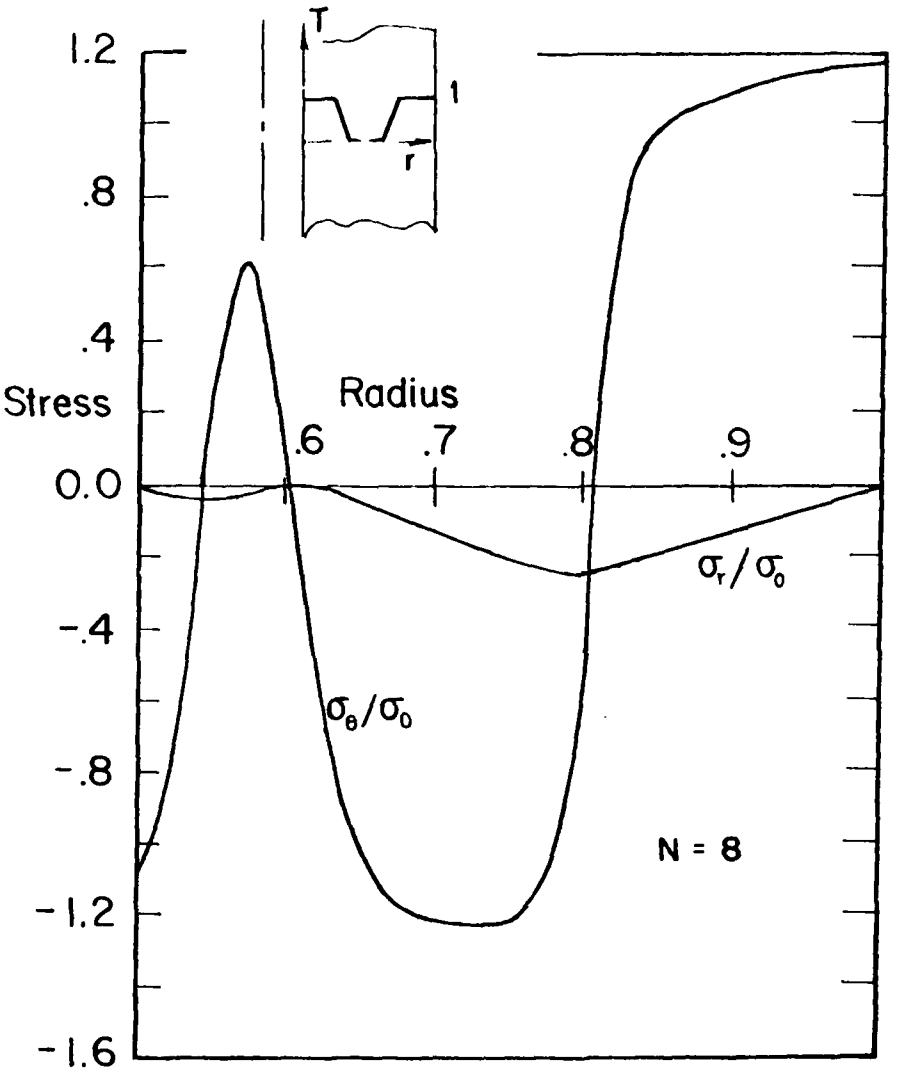


Figure 2. Residual Stresses in a Hollow Cylinder Due to Transformation (Transformation Occurring Symmetrically from the Inner and Outer Diameters).

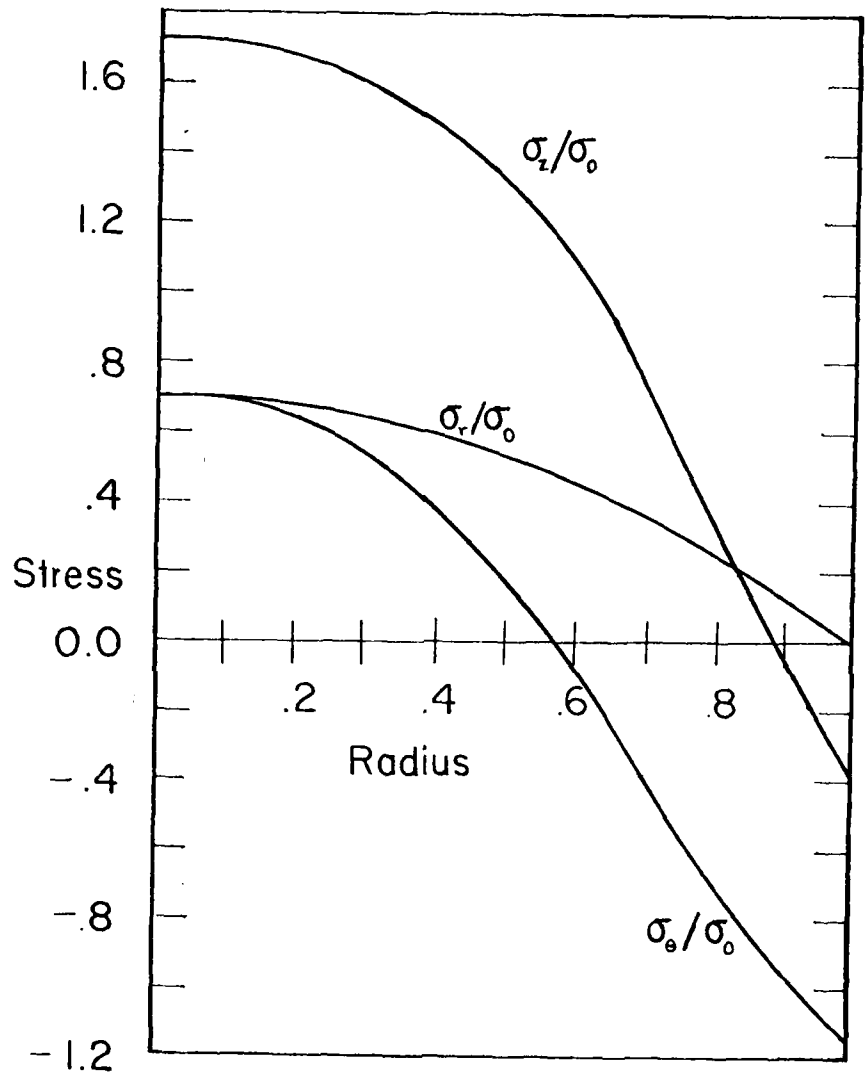


Figure 3. Residual Stresses in a Solid Cylinder Due to Quenching.

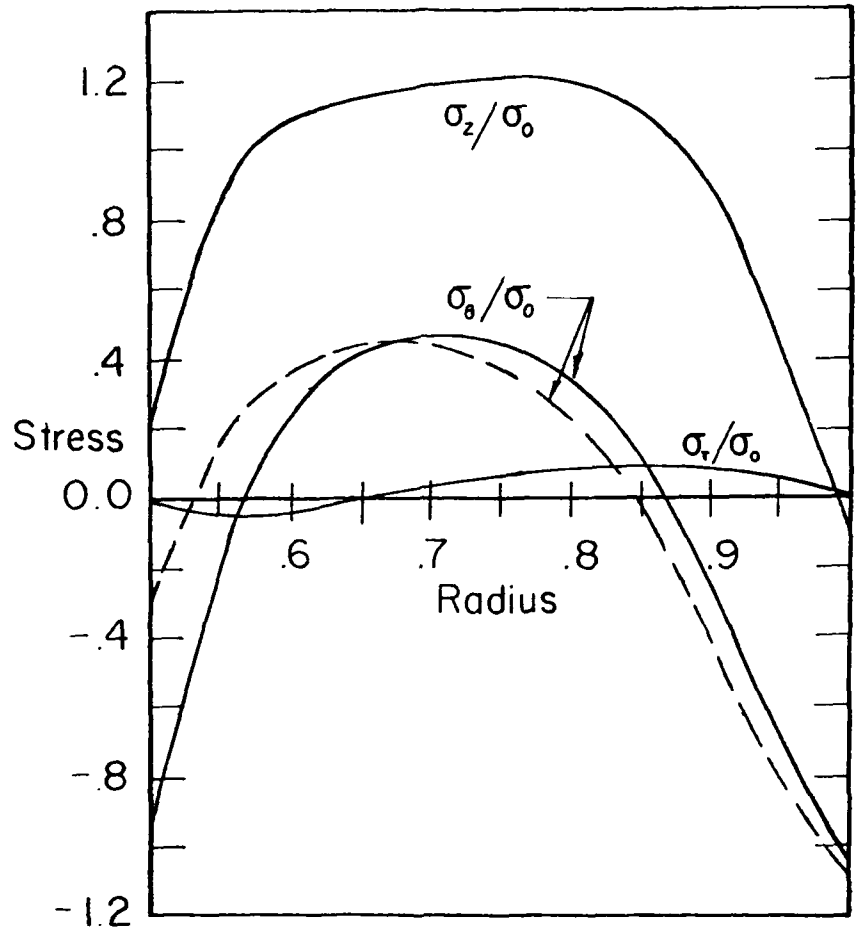


Figure 4. Residual Stresses in a Hollow Cylinder Due to Quenching.

Boundary Conditions:

h_1, h_2 = 12.2, ————————

h_1 = 6.1, h_2 = 12.2, ----------

A NUMERICAL COMPARISON BETWEEN TWO UNCONSTRAINED VARIATIONAL FORMULATIONS

J. J. Wu and T. E. Simkins
U. S. Army Armament Research and Development Command
Benet Weapons Laboratory, LCWSL
Watervliet Arsenal, Watervliet, NY 12189

ABSTRACT. In an effort to relieve the often cumbersome burden of meeting the requirements on the end conditions and to unify the solution formulation for boundary- and initial-value problems, unconstrained variational statements have been introduced in conjunction with some approximate methods. In the case of a boundary value problem, it is shown in this paper that two different variational statements can be established: one is arrived at by the use of the Lagrange multipliers, the other by energy considerations. The numerical convergence of the solutions associated with finite element schemes using one of these two different variational statements is compared with that of the other. In the case of an initial value problem, both formulations can again be established when the adjoint field variable and the adjoint variational statement are introduced. The numerical data presented here indicate that while both methods generate excellent convergent results for the boundary value problem, the method of stiff springs yields results which show much better convergence for the initial value problem than those achieved by Lagrange multipliers.

I. INTRODUCTION. In conjunction with variational methods of mathematical physics, it is often burdensome to select trial functions which are required to satisfy some or all of the end conditions (see, for example, reference [1]). Efforts thus have been made to relieve such requirements on these trial functions. Courant and Hilbert have pointed out that in conjunction with boundary value problems, this can always be done by adding extra boundary terms in the variational statement [2]. Such a concept has been applied successfully by Wu in obtaining solutions to nonconservative stability problems [3]. Wu has further extended the application to the solutions of initial value problems [4]. Simkins also developed unconstrained variational statements for initial and boundary value problems [5]. The approaches used by Wu and Simkins are different in that while Wu, after Courant and Hilbert, employed the concept of a very large constant (very stiff spring constant), Simkins used the method of Lagrange multipliers. For any given problem, the variational statements arrived at by the two approaches are different in boundary terms. The purpose of this paper is to compare the numerical convergence of them in terms of some simple, but specific, examples. Both boundary and initial value problems are considered.

II. UNCONSTRAINED VARIATIONAL STATEMENTS FOR A BOUNDARY VALUE PROBLEM. Let us first consider the transverse vibrations of an Euler-Bernoulli beam under axial load. The differential equation in nondimensionalized form can be written as [1]:

$$y'''' + Qy'' + \lambda^2 y = 0 \tag{2-1}$$

where $y = y(x)$ is the transverse displacement of the beam, as a function of the variable x along the column's length ($0<x<1$). The axial force is denoted by Q; λ is the eigenvalue and a prime $\overline{(')}$ denotes a differentiation with respect to x. The problem is not defined completely, of course, without appropriate boundary conditions. Consider the following given conditions:

$$y(0) = y'(0) = 0 \tag{2-2a,2b}$$

$$y''(1) = y'''(1) + Qy'(1) = 0 \tag{2-2c,2d}$$

Eqs. (2-1) and (2-2) define the familiar buckling problem of an Euler column. It can be solved by methods of approximation in conjunction with a variational statement.

$$\delta I_o = 0 \tag{2-3a}$$

where

$$I_o(y) = \frac{1}{2}\int_0^1 [(y'')^2 - Q(y')^2 + \lambda^2 y^2]dx \tag{2-3b}$$

Through integrations-by-parts, Eqs. (2-3) leads directly the following

$$\begin{aligned} \delta I_o &= 0 \\ &= \int_0^1 (y'''' + Qy'' + \lambda^2 y)\delta y dx \\ &+ y''(1)\delta y'(1) - y''(0)\delta y'(0) \\ &- [y'''(1) + Qy'(1)]\delta y(1) + [y'''(0) + Qy'(0)]\delta y(0) \end{aligned} \tag{2-4}$$

Eq. (2-4) indicates that $\delta I_o = 0$ is equivalent to the differential equation (2-1) and the last two of the b.c. Eq. (2-2c,2d) provided that the variations $\delta y(1)$ and $\delta \dot{y}(1)$ are chosen arbitrarily (thus causing their coefficients to vanish) and that $\delta y(0)$ and $\delta y'(0)$ vanish identically. Thus, $\delta I_o = 0$ can be used as a basis of approximate solution if trial functions are chosen which identically satisfy (2-2a) and (2-2b). Since (2-2a,2b) must be "imposed" they are called "imposed boundary conditions".

The choice of trial functions is otherwise arbitrary and convergence, when achieved, will tend 'naturally' toward a solution satisfying (2-2c) and (2-2d) which are called the 'natural boundary conditions' of the problem. The imposed conditions on the trial functions are often burdensome in the process of obtaining approximate solutions [1]. In this paper, two different methods are compared which remove these constraints on the trial functions.

The first approach is an extension of the method of the Lagrange multipliers in classical mechanics. Suppose one desires to unconstrain the boundary condition (2-2a) $y(0) = 0$. The modified variational statement shall take the form of

$$\delta I_1 = 0 \tag{2-5a}$$

where

$$I_1 = I_o + \alpha y(0) \tag{2-5b}$$

and I_o in (2-5b) is given by (2-3b). Eqs. (2-5) then become

$$\delta I_1 = 0 = \delta I_o + \alpha\delta y(0) = y(0)\delta\alpha \tag{2-6a}$$

$$= \int_0^1 (y'''' + Qy'' + \lambda^2 y)\delta y dx$$

$$+ y''(1)\delta y'(1) - y''(0)\delta y'(0) + y(0)\delta\alpha$$

$$- [y'''(1) + Qy'(1)]\delta y(1) + [y'''(0) + Qy'(0) + \alpha]\delta y(0) \tag{2-6b}$$

It is clear from Eq. (2-6b) that if one defines

$$\alpha = - [y'''(0) + Qy'(0)] \tag{2-7a}$$

thus

$$\delta\alpha = - [\delta y'''(0) + Q\delta y'(0)] \tag{2-7b}$$

equation (2-6b) becomes

$$\delta I_1 = 0 = \int_0^1 (y'''' + Qy'' + \lambda^2 y)\delta y dx$$

$$+ y''(1)\delta y(1) - [y''(0) + Qy(0)]\delta y'(0) - y(0)\delta y'''(0)$$

$$- [y'''(1) + Qy'(1)]\delta y(1) \tag{2-8}$$

where k_1 and k_2 are the nondimensionalized spring constants for deflection and rotation respectively at $x = 0$. Now since

$$\delta I = 0$$

$$= \int_0^1 (y'''' + Qy'' + \lambda^2 y)\delta y dx$$

$$+ y''(1)\delta y'(1) - [y''(0) - k_2 y'(0)]\delta y'(0)$$

$$- [y'''(1) + Qy'(1)]\delta y(1) + [y'''(0) + Qy'(0) + k_1 y(0)]\delta y(0) \quad (2\text{-}11)$$

the natural boundary conditions are

$$y'''(0) + Qy'(0) + k_1 y(0) = 0 , \quad y''(0) - k_2 y'(0) = 0 \quad (2\text{-}12a,12b)$$

$$y''(1) = 0 , \quad y'''(1) + Qy'(1) = 0 \quad (2\text{-}12c,12d)$$

It is clear that Eqs. (2-12) reduce to (2-2) if k_1 and k_2 become infinitely large. Hence, the variational statement (2-10) can serve as a basis of an approximate solution formulation for the problem defined by Eqs. (2-1) and (2-2) if k_1 and k_2 are taken to be very large compared with unity in actual computations.

III. UNCONSTRAINED VARIATIONAL STATEMENTS FOR AN INITIAL VALUE PROBLEM. In the case of initial value problems, similar procedures can be used to free the initial conditions imposed on the trial functions. Examples have been given in two previous papers [4,5]. Since initial value problems are nonself adjoint by nature, adjoint field variables must be introduced to form variational statements which provide the basis for approximate solutions. In this section Lagrange multiplier formulations will be compared with those using the method of infinitely stiff springs - each method being used to relax the requirement that trial functions satisfy identically the imposed conditions arising from an initial value problem. Forced motions of a spring-mass system is used for illustration. The differential equation for such a system can be written as

$$\ddot{y} + \omega^2 y = f(t) \quad (3\text{-}1)$$

where $y = y(t)$ is a function of the time t and a dot ($\cdot$) denotes differentiation with respect to t. The constant $\omega^2 = k/m$ where k is the spring constant and m, the mass. The initial conditions are:

$$y(0) = a , \quad \dot{y}(0) = b \quad (3\text{-}2a,2b)$$

No generality is lost if, in establishing the corresponding variational statements, one considers only a homogeneous system. Hence we consider the differential equation:

$$\ddot{y} + \omega^2 y = 0 \tag{3-1'}$$

and initial condition

$$y(0) = 0 \quad , \quad \dot{y}(0) = 0 \tag{3-2'a,2'b}$$

The fact that the system of Eqs. (3-1') and (3-2') leads to a trivial solution only is not of concern here.

Let $z = z(t)$ be the adjoint field variable. First, the variational statement obtained by the use of Lagrange multipliers is verified to be:

$$\delta I_o = 0 \tag{3-3a}$$

where

$$I_o = -\int_0^1 \dot{y}\dot{z}dt + \omega^2 \int_0^1 yzdt \tag{3-3b}$$

$$+ \dot{y}(1)z(1) - y(0)\dot{z}(0)$$

Eqs. (3-3) lead to

$$\delta I_o = 0$$

$$= \int_0^1 (\ddot{y} + \omega^2 y)\delta z dt + \dot{y}(0)\delta z(0) - y(0)\delta\dot{z}(0)$$

$$+ \int_0^1 (\ddot{z} + \omega^2 z)\delta y dt - \dot{z}(1)\delta y(1) + z(1)\delta\dot{y}(1) \tag{3-4}$$

Eq. (3-4) states that $\delta I_o = 0$ is equivalent to the problem of Eqs. (3-1') and (3-2') and the adjoint problem defined by

$$\ddot{z} + \omega^2 z = 0 \tag{3-5}$$

and

$$z(1) = 0 \quad , \quad \dot{z}(1) = 0 \tag{3-6a,6b}$$

In as much as the variations of the field variable δy, δz, etc. are quite arbitrary and δy is quite independent of δz, one can take $\delta y = 0$, $\delta y(1) = 0$ and $\delta\dot{y}(1) = 0$. Hence the association of the problem of (3-1') and (3-2') with the variational statement Eqs. (3-3) is established.

Now for the inhomogeneous system of Eqs. (3-1) and (3-2), one may similarly verify the corresponding variations statement:

$$\delta I_1 = 0 \tag{3-7a}$$

where

$$I_1(y,z) = -\int_0^1 \dot{y}\dot{z}dt + \int_0^1 [\omega^2 y - f(t)]z dt$$

$$+ \dot{y}(1)z(1) - [y(0) - a]\dot{z}(0) - bz(0) \tag{3-7b}$$

On the other hand, when the "infinitely stiff spring" approach is used to treat the homogeneous case, the variational statement takes the following form [4]:

$$\delta I = 0 \tag{3-8a}$$

where

$$I = -\int_0^1 \dot{y}\dot{z}dt + \omega^2 \int_0^1 yz dt + ky(0)z(1) \tag{3-8b}$$

Eqs. (3-8) result in

$$\delta I = 0$$

$$= \int_0^1 (\ddot{y} + \omega^2 y)\delta z dt + \dot{y}(0)\delta z(0) + [ky(0) - \dot{y}(1)]\delta z(1)$$

$$+ \int_0^1 (\ddot{z} + \omega^2 z)\delta y dt - \dot{z}(1)\delta y(1) + [kz(1) + \dot{z}(0)]\delta y(0) \tag{3-9}$$

The differential equations for the problem and for the adjoint problem are unchanged. The end condition for the original and the adjoint problem are

$$\dot{y}(0) = 0 \quad , \quad ky(0) - \dot{y}(1) = 0 \tag{3-10a,10b}$$

and

$$\dot{z}(1) = 0 \quad , \quad kz(1) + \dot{z}(0) = 0 \tag{3-11a,11b}$$

respectively, Eqs. (3-10) and (3-11) reduce to (3-2') and (3-6) respectively as k becomes infinitely large.

From Eqs. (3-8), extension to a variational statement is easily made for the inhomogeneous case of Eqs. (3-1) and (3-2):

$$\delta I_1 = 0 \tag{3-12a}$$

where

$$I_1 = -\int_0^1 \ddot{y}z\,dt + \int_0^1 [\omega^2 y - f(t)]z\,dt$$

$$+ ky(0)z(1) - kaz(1) - bz(0) \tag{3-12b}$$

IV. NUMERICAL COMPARISONS. In this section, the two methods for the unconstraining of the coordinate (trial) functions described in the previous section will be compared numerically. The approximate solutions are formulated through the finite element discretizations.

IV.A. Boundary Value Problem. The example given in Section II shall be used. The set of Eqs. (2-1) and (2-2) constitute an eigenvalue problem. Using the method of Lagrange multipliers, the associated variational statement is given in Eqs. (2-9) which can also be written as

$$\delta I = 0 = \int_0^1 (y''\delta y'' - Qy'\delta y' + \lambda^2 y\delta y)dx$$

$$- y(0)\delta y'''(0) - y'''(0)\delta y(0)$$

$$+ y'(0)\delta y''(0) + y''(0)\delta y'(0) \tag{4-1}$$

In applying the standard finite element discretization the beam is divided into K equal elements. Denoting the local coordinate by ξ, one has, for the m-th element:

$$\xi = \xi^{(m)} = Kx - m + 1 \tag{4-2a}$$

$$d\xi = Kdx \tag{4-2b}$$

Thus, in terms of local variables, Eq. (4-1) becomes

$$\delta I = 0 = \sum_{m=1}^{K} \int_0^1 [K^3 y^{(m)''}\delta y^{(m)''} - QKy^{(m)'}\delta y^{(m)'} + \frac{\lambda^2}{K} y^{(m)}\delta y^{(m)}]d\xi$$

$$- K^3 y^{(1)}(0)\delta y^{(1)'''}(0) - K^3 y^{(1)'''}(0)\delta y(0)$$

$$+ K^3 y^{(1)'}(0)\delta y^{(1)''}(0) + K^3 y^{(1)''}(0)\delta y^{(1)'}(0) \tag{4-3}$$

Now, let

$$y^{(m)}(\xi) = \underset{\sim}{a}^{T}(\xi)\ \underset{\sim}{Y}^{(m)} \tag{4-4}$$

where

$$\underset{\sim}{a}(\xi) = \begin{Bmatrix} a_1(\xi) \\ a_2(\xi) \\ a_3(\xi) \\ a_4(\xi) \end{Bmatrix} = \begin{Bmatrix} 1 - 3\xi^2 + 2\xi^3 \\ \xi - 2\xi^2 + \xi^3 \\ 3\xi^2 - 2\xi^3 \\ -\xi^2 + \xi^3 \end{Bmatrix} \tag{4-5}$$

$$\underset{\sim}{Y}^{(m)} = \begin{Bmatrix} Y_1^{(m)} \\ Y_2^{(m)} \\ Y_3^{(m)} \\ Y_4^{(m)} \end{Bmatrix} \tag{4-6}$$

and a superscript T denotes the transpose of a matrix. Eq. (4-3) now can be written as

$$\delta I = 0 = \sum_{m=1}^{K} \delta\underset{\sim}{Y}^{(m)T} [K^3 \int_0^1 \underset{\sim}{a}''(\xi)\underset{\sim}{a}''^{T}(\xi)d\xi - QK\int_0^1 \underset{\sim}{a}'(\xi)\underset{\sim}{a}'^{T}(\xi)d\xi + \frac{\lambda^2}{K}\int_0^1 \underset{\sim}{a}(\xi)\underset{\sim}{a}^{T}(\xi)d\xi]\underset{\sim}{Y}^{(m)}$$

$$- K^3\delta\underset{\sim}{Y}^{(1)T}[\underset{\sim}{a}'''(0)\underset{\sim}{a}^{T}(0) + \underset{\sim}{a}(0)\underset{\sim}{a}'''^{T}(0) - \underset{\sim}{a}''(0)\underset{\sim}{a}'^{T}(0) - \underset{\sim}{a}'(0)\underset{\sim}{a}''^{T}(0)]\underset{\sim}{Y}^{(1)} \tag{4-7a}$$

Or

$$\delta I = 0 = \sum_{m=1}^{K} \delta\underset{\sim}{Y}^{(m)T}[K^3\underset{\sim}{C} - QK\underset{\sim}{B} + \frac{\lambda^2}{K}\underset{\sim}{A}]\underset{\sim}{Y}^{(m)}$$

$$- K^3\delta\underset{\sim}{Y}^{(1)}\ [\underset{\sim}{B}_1 + \underset{\sim}{B}_1^{T} - (\underset{\sim}{B}_2 + \underset{\sim}{B}_2^{T})]\underset{\sim}{Y}^{(1)} \tag{4-7b}$$

where

$$\underset{\sim}{A} = \int_0^1 \underset{\sim}{a}\underset{\sim}{a}^{T}d\xi, \quad \underset{\sim}{B} = \int_0^1 \underset{\sim}{a}'\underset{\sim}{a}'^{T}d\xi, \quad \underset{\sim}{C} = \int_0^1 \underset{\sim}{a}''\underset{\sim}{a}''^{T}d\xi \tag{4-8a}$$

$$\underset{\sim}{B}_1 = \underset{\sim}{a}'''(0)\underset{\sim}{a}(0) = \begin{bmatrix} 12 \\ 6 \\ -12 \\ 6 \end{bmatrix} [1\ 0\ 0\ 0] = \begin{bmatrix} 12 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ -12 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \end{bmatrix} \tag{4-8b}$$

$$\underset{\sim}{B}_2 = \underset{\sim}{a}''(0)\underset{\sim}{a}'(0) = \begin{bmatrix} -6 \\ -4 \\ 6 \\ -2 \end{bmatrix} [0\ 1\ 0\ 0] = \begin{bmatrix} 0 & -6 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & -2 & 0 & 0 \end{bmatrix} \tag{4-8c}$$

Now, Eq. (4-7) can be assembled into a global matrix equation

$$\delta I - \delta\underset{\sim}{Y}^T[\underset{\sim}{K} + \lambda^2\underset{\sim}{M}]\underset{\sim}{Y} = 0 \tag{4-9}$$

where

$$\underset{\sim}{Y}^T = [Y_1^{(1)}\ Y_2^{(1)}\ Y_3^{(1)}\ Y_4^{(1)}\ Y_3^{(2)}\ Y_4^{(2)} \ldots\ldots Y_3^{(K)}\ Y_4^{(K)}] \tag{4-10}$$

The details of obtaining the global matrices K and M have been given elsewhere [1] and will not be repeated here.

Since $\delta\underset{\sim}{Y}$ in (4-9) is unconstrained, the equation reduces to

$$(\underset{\sim}{K} + \lambda^2\underset{\sim}{M})\underset{\sim}{Y} = 0 \tag{4-11}$$

which will be solved for the eigenvalues λ^2.

When the method of infinitely stiff springs is used, the variational statement is given by Eqs. (2-10), which can also be written as

$$\delta I = 0 = \int_0^1 (y''\delta y'' - Qy'\delta y' + \lambda^2 y\delta y)dx$$

$$+ k_1 y(0)\delta y(0) + k_2 y'(0)\delta y'(0) \tag{4-12a}$$

$$= \sum_{m=1}^{K} \int_0^1 (K^3 y^{(m)''}\delta y^{(m)''} - QKy^{(m)'}\delta y^{(m)'} + \frac{\lambda^2}{K} y^{(m)}\delta y^{(m)})d\xi$$

$$+ k_1 y^{(1)}(0)\delta y^{(1)}(0) + k_2 K^2 y^{(1)'}(0)\delta y^{(1)'}(0) \tag{4-12b}$$

Or,

$$\delta I = 0 = \sum_{m=1}^{K} \delta Y^{(m)T} [K^3 C - QKB + \frac{\lambda^2}{K} A] Y^{(m)}$$
$$+ \delta Y^{(1)T} [k_1 B_3 + k_2 K^2 B_4] Y^{(1)} \tag{4-13}$$

where

$$B_3 = a(0)a^T(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4-14a}$$

$$B_4 = a'(0)a'^T(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4-14b}$$

As before, Eq. (4-13) can be assembled into a global equation

$$\delta I = 0 = \delta Y^T (K + \lambda^2 M) Y \tag{4-15}$$

so that the eigenvalue λ^2 can be solved from

$$(K + \lambda^2 M) Y = 0 \tag{4-16}$$

Numerical data for the vibration frequencies of a cantilevered column are given in Tables I and II for both the method of Lagrange multipliers and the method of infinitely stiff springs. As shown in these Tables, both methods display excellent convergence.

In the case of the stiff spring method, Tables I and II also indicate that the greater values of k_1 and k_2 may not give more accurate results, although all the results are good when k_1 and k_2 are sufficiently large. This point is further demonstrated by the computations shown in Table III. Since greater values of k_1 and k_2 mean that the prescribed end conditions are more accurately satisfied, Table III suggests that forcing the solution to greater accuracy at one point may cause a decline in overall acceptability of the results as evidenced by the declining accuracy of the eigenvalue. This same conclusion was first presented in [1].

TABLE I. NUMERICAL COMPARISONS OF TWO UNCONSTRAINED VARIATIONAL FORMULATIONS

The First Eigenvalue of a Cantilevered Beam

No. of Elements	Method of Lagrange Multipliers	Stiff Spring Method $k_1 = k_2 = 10^8$	Stiff Spring Method $k_1 = k_2 = 10^{12}$
1	3.585387	3.532731	3.534027
3	3.516379	3.516371	3.515999
5	3.516063	3.516063	3.514741
7	3.516028	3.516027	3.516168
9	3.516020	3.516022	3.516549

From the exact solution: 3.516015.....

TABLE II. NUMERICAL COMPARISONS OF TWO UNCONSTRAINED VARIATIONAL FORMULATIONS

The Second Eigenvalue of a Cantilevered

No. of Elements	Method of Lagrange Multipliers	Stiff Spring Method $k_1 = k_2 = 10^8$	Stiff Spring Method $k_1 = k_2 = 10^{12}$
1	47.91346	34.80686	34.80688
3	22.13741	22.10685	22.10853
5	22.04607	22.04550	22.04783
7	22.03750	22.03746	22.05306
9	22.03560	22.03559	22.09871

From the exact solution: 22.03449.....

TABLE III. EFFECT OF THE MAGNITUDE OF THE "SPRING CONSTANTS" ON CONVERGENCE

The First Two Eigenvalues of a Cantilevered Beam

No. of Elements = 9

$k_1 = k_2 =$	10^4	10^6	10^8	10^{10}	10^{12}
λ_1	3.513985	3.516000	3.516022	3.516008	3.516549
λ_2	21.930259	22.034547	22.035588	22.035538	22.098706

Exact values: $\lambda_1 = 3.516015.....$

$\lambda_2 = 22.034491....$

IV.B. An Initial Value Problem. For our numerical comparisons in the case of an initial value problem, we shall consider the one defined by:

$$\text{D.E.:} \qquad m\ddot{y} + ky = f_o \cos\omega_f t, \quad 0 \le t \le T \tag{4-17}$$

$$\text{I.C.:} \qquad y(0) = a\ , \quad \dot{y}(0) = b \tag{4-18a,18b}$$

The specific values of the constants m, k, fo, ω_f, a, b and T will be given later. The upper limit of the time interval T can take any positive value other than infinity. Before one applies the variational formulation given in Section III, it will be convenient to normalize the time variable t with respect to T. Thus let

$$\tau = t/T\ , \quad t = T\tau,\ dt = Td\tau \tag{4-19}$$

$$y(t) = \bar{y}(\tau)\ , \quad \frac{dy}{dt} = \frac{1}{T}\frac{d\bar{y}}{d\tau}\ , \quad \frac{d^2y}{dt^2} = \frac{1}{T^2}\frac{d^2\bar{y}}{d\tau^2} \tag{4-20}$$

Also define

$$\bar{\omega} = \omega T\ , \quad \bar{f} = \frac{f_o T^2}{m}$$

$$\bar{\omega}_f = \omega_f T\ , \quad \bar{a} = a, \quad \bar{b} = bT \tag{4-21}$$

With these new parameters, Eqs. (4-17) and (4-18) become

$$\text{D.E.} \qquad \frac{d^2\bar{y}}{d\tau^2} + \bar{\omega}^2\bar{y} = \bar{f}\cos(\bar{\omega}_f\tau)\ , \quad 0 \le \tau \le 1 \tag{4-22}$$

$$\text{I.C.} \qquad \bar{y}(0) = \bar{a}\ , \quad \dot{\bar{y}}(0) = \bar{b} \tag{4-23a,23b}$$

Now we are ready to apply the formulations given in Section III. We shall first consider the solution formulation by the method of Lagrange multipliers. Comparing Eqs. (4-22) and (4-23) with (3-1) and (3-2), one observes that the variational statement follows that of Eqs. (3-7). Or,

$$\delta I = 0 \tag{4-24a}$$

where

$$I = -\int_0^1 \dot{\bar{y}}\,\dot{\bar{z}}\,d\tau + \int_0^1 [\bar{\omega}^2\bar{y} - \bar{f}\cos(\bar{\omega}_f\tau)]\bar{z}\,d\tau$$

$$+ \dot{y}(1)z(1) - \bar{y}(0)\dot{\bar{z}}(0) + \bar{a}\,\dot{\bar{z}}(0) - \bar{b}\,\bar{z}\,(0) \tag{4-24b}$$

Since δy and δz are quite independent of each other, one can set $\delta y = 0$ in Eqs. (4-24) and obtain

$$(\delta I)_{\delta y=0} = -\int_0^1 \ddot{\bar{y}}\delta\bar{z}d\tau + \int_0^1 \bar{\omega}^2\bar{y}\delta\bar{z}d\tau - \int_0^1 \bar{f}\cos(\bar{\omega}_f\tau)d\bar{z}d\tau$$

$$+ \dot{\bar{y}}(1)\delta\bar{z}(1) - \bar{y}(0)\delta\dot{\bar{z}}(0) + a\delta\dot{\bar{z}}(0) - b\delta\bar{z}(0) = 0 \qquad (4\text{-}25)$$

The same process of finite element discretization used for the boundary value problem in the previous subsection can be employed here. The same shape functions and generalized coordinates are also used. In terms of the element variables, ξ, defined before, except now that

$$\xi = k\tau - m + 1 \qquad (4\text{-}26)$$

etc., Eq. (4-25) becomes:

$$(\delta I)_{\delta y=0} = 0 = \sum_{m=1}^{K} \delta z^{(m)T}\left[-K\int_0^1 a'a'^T d\xi + \frac{\bar{\omega}^2}{K}\int_0^1 aa^T d\xi\right]Y^{(m)}$$

$$+ \delta z^{(K)T}Ka(1)a'^T(1)Y^{(K)} - \delta z^{(1)T}Ka'(0)a^T(0)Y^{(1)}$$

$$- \sum_{m=1}^{K} \delta z^{(m)T}\frac{\bar{f}}{K}\int_0^1 \cos\left[\frac{\bar{\omega}_f}{K}(\xi + m - 1)\right]a\, d\xi$$

$$+ \delta z^{(1)T}\bar{a}Ka'(0) - \delta z^{(1)T}\bar{b}a(0) \qquad (4\text{-}27)$$

Or,

$$\sum_{m=1}^{K} \delta z^{(m)T}\left[-KB + \frac{\bar{\omega}^2}{K}A\right]Y^{(m)} + \delta z^{(K)T}KB_5Y^{(K)} - \delta z^{(1)T}KB_6Y^{(1)}$$

$$- \sum_{m=1}^{K} \delta z^{(m)T}\frac{\bar{f}}{K}F^{(m)} + \delta z^{(1)T}[\bar{a}Ka'(0) - \bar{b}a(0)] = 0 \qquad (4\text{-}28)$$

where A, B have been defined in Eqs. (4-8a) and

$$B_5 = a(1)a'^T(1) = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}[0\ 0\ 0\ 1] = \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{bmatrix} \qquad (4\text{-}29a)$$

$$B_6 = a'(0)a^T(0) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} [1\ 0\ 0\ 0] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4\text{-}29b)$$

$$\underset{\sim}{F}^{(m)} = \int_0^1 \cos\left[\frac{\bar{\omega}_f}{K}(\xi + m - 1)\right]\underset{\sim}{a}\, d\xi \quad (4\text{-}29c)$$

In terms of global generalized coordinates $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$ defined by

$$\underset{\sim}{Y}^T = [Y_1^{(1)}\ Y_2^{(1)}\ Y_3^{(1)}\ Y_4^{(1)}\ Y_3^{(2)}\ Y_4^{(2)} \ldots\ldots Y_3^{(K)}\ Y_4^{(K)}] \quad (4\text{-}30a)$$

and

$$\underset{\sim}{Z}^T = [Z_1^{(1)}\ Z_2^{(1)}\ Z_3^{(1)}\ Z_4^{(1)}\ Z_3^{(2)}\ Z_4^{(2)} \ldots\ldots Z_3^{(K)}\ Z_4^{(K)}] \quad (4\text{-}30b)$$

Eq. (4-28) can be assembled as before into the matrix equation

$$\delta\underset{\sim}{Z}^T[\underset{\sim\sim}{K}\underset{\sim}{Y} - \underset{\sim}{F}] = 0 \quad (4\text{-}31)$$

Or, since $\delta\underset{\sim}{Z}$ is not constrained in any way,

$$\underset{\sim\sim}{K}\underset{\sim}{Y} = \underset{\sim}{F} \quad (4\text{-}32)$$

which can be solved for $\underset{\sim}{Y}$.

When the method of infinitely stiff springs is used, the variational statement must be modified according to Eqs. (3-12). Thus, the finite element discretization begins with

$$(\delta I)_{\delta y=0} = 0$$

$$= -\int_0^1 \ddot{\bar{y}}\delta\bar{z}\,d\tau + \int_0^1 [\bar{\omega}^2\bar{y} - \bar{f}\cos(\bar{\omega}_f\tau)]\bar{z}\,d\tau$$

$$+ ky(0)z(1) - k\bar{a}z(1) - \bar{b}z(0) \quad (4\text{-}33)$$

Hence,

$$\sum_{m=1}^{K} \delta Z^{(m)T} [-K \int_0^1 a' a'^T d\xi + \frac{\omega^2}{K} \int_0^1 a a^T d\xi] Y^{(m)}$$

$$+ k \delta Z^{(K)T} a(1) a^T(0) Y^{(1)}$$

$$- \sum_{m=1}^{K} \delta Z^{(m)T} \frac{\bar{f}}{K} \int_0^1 \cos[\frac{\omega_f}{K} (\xi + m - 1)] a \, d\xi$$

$$- k \delta Z^{(K)T} \bar{a} a(1) - \delta Z^{(1)T} \bar{b} a(0) \qquad (4\text{-}34)$$

Or,

$$\sum_{m=1}^{K} \delta Z^{(m)T} [-KB + \frac{\bar{\omega}^2}{K} A] Y^{(m)} + \delta Z^{(K)T} k B_7 Y^{(1)}$$

$$- \sum_{m=1}^{K} \delta Z^{(m)T} \frac{\bar{f}}{K} F^{(m)} - \delta Z^{(K)T} \bar{k} a(1) - \delta Z^{(1)T} \bar{b} a(0) \qquad (4\text{-}35)$$

where A, B, $F^{(m)}$ have all been defined before and

$$B_7 = a(1)a^T(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} [1\ 0\ 0\ 0] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (4\text{-}36)$$

Now, as with Eq. (4-28), here Eq. (4-35) can be assembled in a global equation in the form of Eqs. (4-31) and (4-32) and be solved.

The specific problem considered is as follows:

$$my + k\ddot{y} = fo \cos(\omega_f t) \ , \quad 0 \leq t \leq T$$

with

$$y(0) = y_o \text{ and } \dot{y}(0) = y_o$$

The numerical values of the parameters are:

$$m = 1.0, \quad k = 1.0, \quad f_o = 1.0, \quad \omega_f = 0.5$$

$$y_o = 1.0, \quad y_1 = 1.0$$

The plot for the forcing function $^{c}_{0}\cos(\omega_f t)$ and the exact solution y(t) is shown in Figure 1. The numerical solutions of the problem using both the method of Lagrange multipliers and the method of stiff springs are given in Tables IV through IX.

Tables IV through VI show the stiff spring method generates excellent convergent results for various lengths of intervals of solution.

The results using the method of Lagrange multipliers are shown in Tables VII through IX. Table VII shows that for moderately long intervals, the convergence at the initial point is non-existent although it improves remarkably away from the initial point. This data may lead one to doubt whether the method of Lagrange multipliers works at all in treating i.v. problems. However, when the length of the interval of solution is reduced, as shown in Tables VIII and IX, it is clear that the results do converge. Hence, both methods generate convergent results. The length of interval used in the Lagrange multipliers approach is so small compared with the stiff spring method for comparable convergence that the practical value of the former is doubtful in treating initial value problems when finite element discretization is employed. Simkins [4] has shown, however, that when global approximating functions are employed, (consisting of higher ordered polynomials), very good results can be achieved over an acceptable interval of solution.

V. CONCLUSIONS. From the numerical data presented in this paper, the following conclusions are suggested:

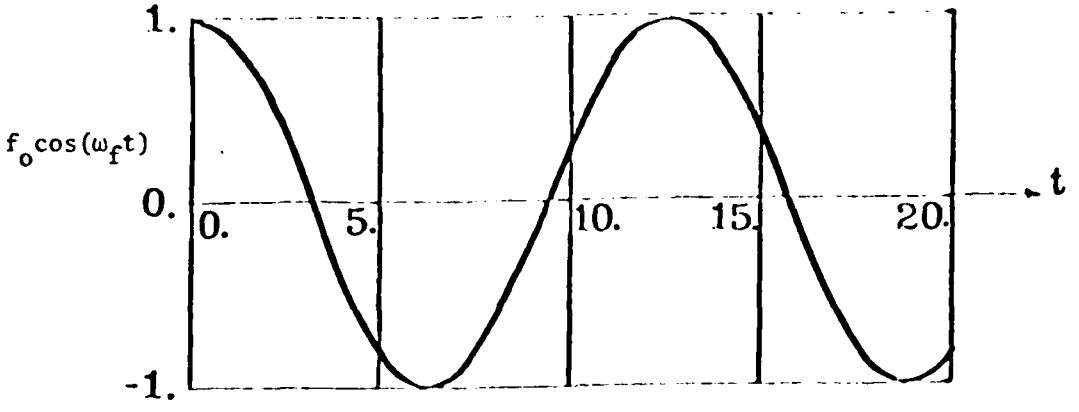
1. Both the method of Lagrange multipliers and the method of stiff springs generate convergent results.

2. In the case of boundary value problems, both methods give excellent results and equally fast convergence. The method of stiff springs appears to be easier to use and more general in a practical sense.

3. For initial problems discretized by finite elements (piecewise continuous third order polynomials), convergence of the Lagrange multiplier method, as compared to the method of stiff springs, is so inferior as to be of dubious practical value. (This statement does not apply, however, where a global discretization is employed using higher ordered (e.g. 8th order [4]) polynomials continuous over the entire domain of integration.)

REFERENCES

1. J. J. Wu, "On The Numerical Convergence of Matrix Eigenvalue Problems," Journal of Sound and Vibration (1974), 37, pp. 349-358.



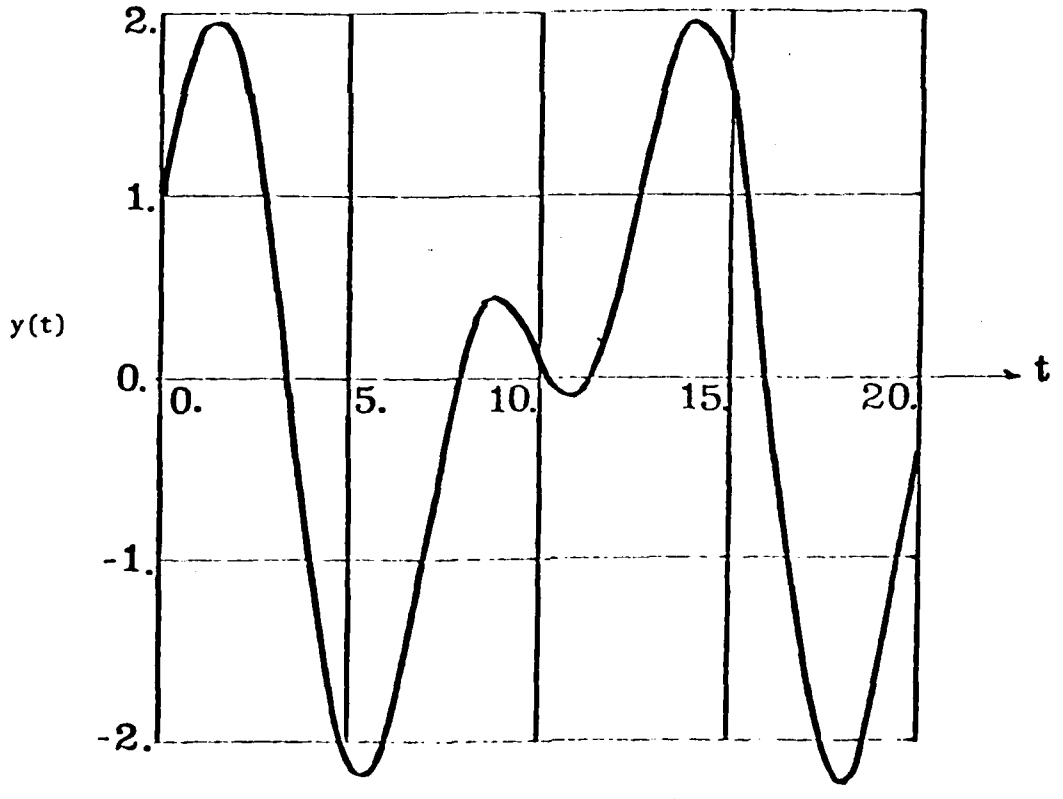


FIGURE 1. Plots for the forcing function $f_o\cos(\omega_f t)$ and the exact solution y(t) for a simple initial value problem.

TABLE IV. SOLUTIONS TO THE SIMPLE INITIAL-VALUE PROBLEM: BY METHOD OF STIFF SPRINGS

($0 \leq t \leq 2.0$, 10 Elements)

t	y(t)		$\dot{y}(t)$	
0.	1.0000000	(1.0000000)	1.00000	(1.00000)
0.4	1.3891537	(1.3891534)	0.91843	(0.91842)
0.8	1.7132029	(1.7132018)	0.67622	(0.67621)
1.2	1.9117024	(1.9117006)	0.29662	(0.29661)
1.6	1.9382512	(1.9382491)	-0.17424	(-0.17425)
2.0	1.7684161	(1.7684161)	-0.67413	(-0.67403)

TABLE V. SOLUTIONS TO THE SIMPLE INITIAL-VALUE PROBLEM: BY METHOD OF STIFF SPRINGS

($0 \leq t \leq 10.0$, 10 Elements)

t	$y(t)$		$\dot{y}(t)$	
0.	1.000	(1.000)	1.004	(1.000)
2.0	1.770	(1.768)	-0.675	(-0.674)
4.0	-1.094	(-1.094)	-1.518	(-1.512)
6.0	-1.920	(-1.919)	0.778	(0.773)
8.0	0.167	(0.166)	0.690	(0.689)
10.0	0.114	(0.114)	-0.385	(-0.381)

TABLE VI. SOLUTIONS TO THE SIMPLE INITIAL-VALUE PROBLEM: BY METHOD OF STIFF SPRINGS

($0 \leq t \leq 20.0$, 10 Elements)

t	$y(t)$	$\dot{y}(t)$
0.	1.000 (1.000)	1.05 (1.00)
4.	-1.097 (-1.094)	-1.57 (-1.51)
8.	0.173 (0.176)	0.71 (0.69)
12.	0.453 (0.462)	0.88 (0.85)
16.	-0.156 (-0.162)	-1.76 (-1.71)
20.	-0.348 (-0.342)	1.10 (-1.08)

TABLE VII. SOLUTIONS TO THE SIMPLE INITIAL-VALUE PROBLEM: BY METHOD OF LAGRANGE MULTIPLIERS

($0 \leq t \leq 2.0$, 10 Elements)

t	y(t)		$\dot{y}(t)$	
0.	-54.1858	(1.0000)	2343.	(1.000)
0.4	0.2348	(1.3892)	69.784	(0.918)
0.8	1.6623	(1.7132)	2.697	(0.676)
1.2	1.9058	(1.9117)	0.354	(0.297)
1.6	1.9330	(1.9382)	-0.173	(-0.174)
2.0	1.7635	(1.7684)	-0.673	(-0.674)

TABLE VIII. SOLUTIONS TO THE SIMPLE INITIAL-VALUE PROBLEM: BY METHOD OF LAGRANGE MULTIPLIERS

($0 \leq t \leq 0.5$, 10 Elements)

t	y(t)		$\dot{y}(t)$	
0.	0.95409	(1.00000)	8.79210	(1.00000)
0.1	1.09848	(1.09983)	1.22432	(0.99496)
0.2	1.19861	(1.19865)	0.98649	(0.97973)
0.3	1.29543	(1.29544)	0.95442	(0.95422)
0.4	1.38915	(1.38915)	0.91843	(0.91842)
0.5	1.47878	(1.47878)	0.87246	(0.87246)

TABLE IX. SOLUTIONS TO THE SIMPLE INITIAL-VALUE PROBLEM: BY METHOD OF LAGRANGE MULTIPLIERS

($0 \leq t \leq 0.1$, 10 Elements)

t	$y(t)$		$\dot{y}(t)$	
0	1.000049	(1.000000)	0.958591	(1.000000)
.02	1.020000	(1.019999)	0.998581	(0.999800)
.04	1.039989	(1.039989)	0.999162	(0.999197)
.06	1.059964	(1.059964)	0.998190	(0.998192)
.08	1.079914	(1.079914)	0.996780	(0.996780)
.10	1.099832	(1.099832)	0.994963	(0.994963)

2. R. Courant and D. Hilbert, Methods of Mathematical Physics, McGraw-Hill, New York (1953), see page 211.

3. J. J. Wu, "Effects of Support Flexibility on the Stability of a Beam Under a Follower Thrust and Inertia," Developments in Theoretical and Applied Mechanics (1976), Vol. 8, pp. 391-402.

4. J. J. Wu, "Solutions to Initial Value Problems by Use of Finite Elements - Unconstrained Variational Formulations," Journal of Sound and Vibration (1977), Vol. 53, pp. 341-356.

5. T. E. Simkins, "Unconstrained Variational Statements for Initial and Boundary Value Problems," American Institute fo Aeronautics and Astronautics Journal (1978), Vol. 16, pp. 559-563.

UNCONSTRAINED VARIATIONAL STATEMENTS FOR INITIAL AND BOUNDARY-VALUE PROBLEMS

T. E. Simkins
U.S. Army Armament Research and Development Command
Benet Weapons Laboratory
Watervliet Arsenal, Watervliet, NY 12189

ABSTRACT. A procedure is developed for generating variational statements suitable for obtaining approximate solutions to boundary-initial value problems. The essence of the procedure is to introduce all boundary and initial conditions into the variational statement as natural boundary conditions. This is accomplished through the use of Lagrange multipliers, in which all initial condition terms as well as boundary terms are determined analytically. The result is a variational statement in which completely unconstrained trial functions may be assumed as a basis for an approximate solution. Several applications are given, including the response of a beam subject to a moving concentrated mass loading.

I. INTRODUCTION. Theorems establishing the correspondence between certain boundary value problems and the calculus of variations appear in detail in the treatise of Collatz [1]. More recently, Rund [2] has produced more general theorems through the use of the transversality conditions from which the theorems of Collatz emerge as special cases. Neither of these works, however, attempts to establish variational statements for the solution of initial value problems, which is the subject of the work herein.

Recent work by Bailey [3,4] has shown that Hamilton's law of varying action is capable of yielding approximate solutions to initial and boundary-initial value problems. The variational form used by Bailey allows the function and its derivative to vary at the upper limit of integration of the time interval. At the lower limit, these quantities are constrained to satisfy specified initial conditions. It appears that a more general method would free the variations at the lower limit as well, thus broadening the class of admissible trial functions. Although much has been written [5] on the subject of removal of constraints on the boundary variations, the applications usually deal with elliptic rather than hyperbolic systems, where it is customary to introduce the constraints of the problem as natural boundary conditions. The constraints themselves thus become subject to approximation through the

variational process. Convergence, when achieved, tends toward a solution satisfying these constraint conditions. When all of the constraints of the problem are introduced in this manner, the result is a completely unconstrained variational statement, whereby the trial functions need not identically satisfy any boundary conditions. As the use of Lagrange multipliers in freeing boundary variations of constraints is by now classical, this work is fundamentally concerned with extending the method to remove constraining time conditions.

The Lagrange multiplier procedure adds each constraint as a zero times a Lagrange multiplier to the previously unconstrained variational statement. In this way, each constraint is made to appear as a natural boundary condition and, in some cases where a functional exists (i.e., variational "principles"), it may be modified to include these terms. The multipliers can usually be identified in terms of values of the function and its derivatives on the bounding surface of the domain of integration. The act of freeing the boundary variations will not result in the loss of a variational "principle" provided the constraint is holonomic, i.e., a functional will still exist though modified by additive products of the Lagrange multipliers times the individual constraint relations. On the other hand, should any of the constraints be nonholonomic, the existence of a functional is denied and one has in its place a less elegant variational "statement" which may nevertheless provide a basis for an approximate solution to the problem at hand.

In spite of the apparent generality of the Lagrange multiplier method, its application to Hamilton's principle (a constrained variational principle) for the solution of initial value problems is not obvious. Indeed, Bailey found it fruitful to employ Hamilton's law of varying action in which no functional exists. Once the quest of a functional is abandoned, however, unconstrained variational formulations for initial value problems are immediately possible, as was first shown by Tiersten [6]. Since the purpose of Tiersten's work at the time did not involve explicit solutions in the time domain, the success of his method for achieving solutions to initial value problems was never tested. Further, Tiersten's procedure requires a special introduction of one of the initial conditions into the variational statement, making incomplete use of the Lagrange multiplier method in the time domain. Solutions to the free oscillator, the two-dimensional wave equation, and the motion of a beam to a moving concentrated mass are offered as evidence of success of the method.

The variational form presented herein does not produce a functional. Gurtin [7] and later Wu [8], however, have successfully formulated initial and boundary-initial value problems using unconstrained variational principles in which a functional is indeed produced. Wu's treatment requires introducing the adjoint variable and replacing the given

boundary and initial conditions with a set of artificial conditions containing a parameter that eventually is allowed to become very large numerically. Using finite element approximations, Wu was able to achieve excellent agreement with the exact solutions of several partial differential equations in one and two dimensions. Gurtin's procedure, on the other hand, combines the initial conditions and field equations into a single integro-differential equation which is then regurgitated as the Euler-Lagrange equation when the variation of a constructed functional is made to vanish.

II. VARIATIONAL STATEMENTS AND LAGRANGE MULTIPLIERS. Unless the variational quantities appearing in a variational statement or principle are completely arbitrary, the formulation is said to be 'constrained'. The principle of virtual work in its conventional form is one example of a constrained variational principle; i.e.,

$$0 = \delta\int_V U(\varepsilon)dV - \int_V K_i\delta u_i dV - \int_{S_f} \bar{F}_i\delta u_i dS \tag{1}$$

where $U(\varepsilon)$ is the potential energy density of an elastic volume, V. The K_i are body forces per unit volume and the u_i are the unknown displacement functions. A bar denotes prescribed surface quantities, and

$$F_k = n_\ell \frac{\partial U}{\partial u_{k,\ell}} \equiv n_\ell \sigma_{\ell k} \tag{2}$$

where n_ℓ is the outward directed normal to any surface. $\sigma_{\ell k}$ represents the stress tensor.

Implicit in Eq. (1) is the constraint:

$$u_i = \bar{u}_i \text{ on Su} \tag{3}$$

i.e., the displacement functions must be those prescribed on the boundary surface Su. Further, Su and S_f do not overlap and together comprise the complete boundary of the volume, V.

If Eq. (1) is used as a basis for approximating a solution to a problem in elastostatics - e.g., via the Rayleigh-Ritz method - the shape functions employed in the approximation must each identically satisfy the constraint equation (3) a priori. This requirement may be removed by using Lagrange multipliers to introduce the constraint explicitly in Eq. (1) which then becomes:

$$0 = \delta\{\int_V U(\varepsilon)dV + \int_{S_u} \lambda_i(u_i-\bar{u}_i)dS\} - \int_V K_i\delta u_i dV - \int_{S_f} \bar{F}_i\delta u_i dS \quad (4a)$$

$$= \delta \int_V U(\varepsilon)dV + \int_{S_u} \{\delta\lambda_i(u_i-\bar{u}_i) + \lambda_i\delta u_i\}dS - \int_V K_i\delta u_i dV - \int_{S_F} \bar{F}_i\delta u_i dS$$

For a Hookean material:

$$\delta\int_V U(\varepsilon)dV = \int_V \sigma_{ij}\delta\varepsilon_{ij}dV$$

$$= \int_V (\sigma_{ij}\delta u_i)_{,j}dV - \int_V \sigma_{ij,j}\delta u_i dV$$

$$= \int_{S_f} F_i\delta u_i dS + \int_{S_u} F_i\delta u_i dS - \int_V \sigma_{ij,j}\delta u_i dV$$

Thus

$$\int_{S_u} F_i\delta u_i dS - \int_V (\sigma_{ij,j}+K_i)\delta u_i dV + \int_{S_u} [\lambda_i\delta u_i+\delta\lambda_i(u_i-\bar{u}_i)]dS$$

$$+ \int_{S_f} (F_i-\bar{F}_i)\delta u_i dS = 0 \quad (4b)$$

Now the δu_i in Eq. (1) or Eq. (4b) are not all arbitrary because of the constraints in Eq. (3). But if the Lagrange multipliers, λ_i, are defined as $-F_i$ on S_u, the coefficients of all δu_i quantities on S_u vanish, and hence the δu_i may be <u>viewed</u> as arbitrary. This is the essence of the Lagrange multiplier method. It is important to note that the constraints involved in this exercise are holonomic and the surfaces S_u and S_f do not overlap.

Thus, substituting $\lambda_i \equiv -F_i$ on S_u, Eq. (4a) becomes an <u>unconstrained</u> principle of virtual work:

$$\delta\{\int_V U(\varepsilon)dV - \int_{S_u} F_i(u_i-\bar{u}_i)dS - \int_V K_i u_i dV - \int_{S_f} \bar{F}_i u_i dS\} = 0 \quad (4c)$$

If the Rayleigh-Ritz method is used with Eq. (4c), the shape functions assumed no longer need satisfy identically the constraint relations in Eq. (3).

Another example of a constrained variational principle is Hamilton's Principle [6]:

$$\delta\int_{t_0}^{t_1} dt\{\int_V (T-U)dV + \int_{S_f} \bar{F}_k u_k dS\} = 0 \qquad (5)$$

In analogy with the previous treatment of the virtual work principle one notes that instead of simply

$$\int_V U(\varepsilon)dV ,$$

we have

$$\int_{t_0}^{t} \int_V (T-U)dVdt$$

where T is the kinetic energy density. Thus, in addition to the term,

$$F_k = n_\ell \frac{\partial U}{\partial u_{k,\ell}}$$

one can expect a similar term from T, i.e.,

$$P_k = \pm \frac{\partial T}{\partial \dot{u}_k} = \pm \rho \dot{u}_k$$

where ρ is the mass density and the $\pm$ denotes unit normals to the 'time' surfaces $t = t_0$ and $t = t_1$.

The constraint equations are

$$u_i = \bar{u}_i \quad \text{on } S_u \qquad (6)$$

and

$$\delta u_i = 0 \quad \text{at} \quad t = t_0$$

The constraint at t_0 can be satisfied by specifying u_i at t_0 but this cannot be done for the later time t_1 in the ordinary initial value problem. Thus, Hamilton's Principle contains at least one non-holonomic constraint.

Further, one notes that both momentum and displacement are prescribed on the same surface t_0. Thus the quantity

$$\bar{P}_k \delta u_k]_{t_0} \quad , \quad \text{unlike} \quad \bar{F}_k \delta u_k]_{S_f} \quad ,$$

does not appear in Hamilton's Principle and analogous definitions for the λ_i are therefore not available. While the pressure of non-holonomic constraints can be handled by a more general Lagrange multiplier procedure [9], the overlapping of surfaces on which displacement and

momentum are specified proves insurmountable in applying the Lagrange multiplier technique. One concludes therefore, that straightforward use of the Lagrange multiplier technique to completely unconstrain Hamilton's Principle in the time domain is not possible.

III. INITIAL VALUE PROBLEMS AND ADJOINT VARIATIONAL PRINCIPLES. The work of the previous section demonstrates that there is no difficulty in using the Lagrange multiplier method in the space domain where the governing equations are elliptic, but only in the time domain where hyperbolic systems are encountered. In this section it is shown that hyperbolic systems can also be treated by the multiplier method provided consideration is given not only to the physical system but also its adjoint. This is most easily shown by example.

Consider the following initial value problem:

$$u'' + u' + u = 0 \quad ; \quad 0<x<1$$

$$u(0) = u'(0) = 0 \tag{7}$$

The adjoint to the system Eq. (7) is:

$$v'' - v' + v = 0$$

$$v(1) = v'(1) = 0 \tag{8}$$

An adjoint variational principle may be found by multiplying Eq. (7) by the adjoint variable v and integrating over the domain. Thus:

$$\int_0^1 v(u''+u'+u)dx - [vu']_0^1 = I \equiv \int_0^1 (uv+u'v-u'v')dx \tag{9}$$

If the variation of I is made to vanish:

$$\delta I = 0 = [v\delta u-u'\delta v-v'\delta u]_0^1 + \int_0^1 (u''+u'+u)\delta v dx + \int_0^1 (v''-v'+v)\delta u dx \tag{10a}$$

Now if v(1) and u(0) are specified a priori,

$$\text{i.e.,} \qquad v(1) = 0$$

$$u(0) = 0 \tag{10b}$$

then,

$$\delta I = 0 = u'(0)\delta v(0) - v'(1)\delta u(1) + \int_0^1 L(u)\delta v dx + \int_0^1 L^*(v)\delta u dx \quad (10c)$$

Since only $\delta v(1)$ and $\delta u(0)$ are constrained to vanish in Eq. (10c), the rest of the variations are independent and arbitrary. Thus in addition to Eq. (10b) we have,

$$\begin{aligned} L(u) &\equiv u'' + u' + u = 0 \\ L^*(v) &\equiv v'' - v' + v = 0 \\ u'(0) &= 0 \\ v'(1) &= 0 \end{aligned} \quad (10d)$$

Equations (10b) and (10d) comprise the entire system of equations for the adjoint and physical systems. As it stands, Eq. (10c) is a constrained adjoint variational principle since the constraints (Eq. (10b)) are imposed a priori. The fact that Eq. (10c) does not, by itself, yield <u>all</u> of the initial conditions illustrates its constrained character. However, adding the constraints Eq. (10b) to this variational statement via the Lagrange multiplier method will free the principle of constraints and all of the initial conditions as well as the differential equations are then regurgitated. Thus:

$$\begin{aligned} &\delta\{\int_0^1 (uv+u'v-v'u')dx + \lambda_1 u(0) + \lambda_2 v(1)\} = 0 \\ &= - u'\delta v\Big|_0^1 + \int_0^1 (u''+u'+u)\delta v dx + (v-v')\delta u\Big|_0^1 \\ &+ \int_0^1 (v''-v'+v)\delta u dx + \lambda_1\delta u(0) + u(0)\delta\lambda_1 + \lambda_2\delta v(1) + v(1)\delta\lambda_2 \end{aligned} \quad (11)$$

$\delta u(0)$ and $\delta v(1)$ may be <u>viewed</u> as arbitrary if

$$\begin{aligned} \lambda_1 &\equiv v(0) - v'(0) \\ \lambda_2 &\equiv u'(1) \end{aligned} \quad (12)$$

Substituting the definitions (Equations (12)),into equation (11) and integrating by parts:

$$\int_0^1 L(u)\delta v\,dx + \int_0^1 L^*(v)\delta u\,dx + u'(0)\delta v(0)$$

$$+ (v(1)-v'(1))\delta u(1) - u(0)\delta v'(0) + v(1)\delta u'(1) = 0 \qquad (13)$$

Since all variations are now viewed as arbitrary:

$$L(u) \equiv u'' + u' + u = 0 \quad ; \quad 0<x<1$$

$$u(0) = 0$$

$$u'(0) = 0 \qquad (14a)$$

$$L(v) \equiv v'' - v' + v = 0 \quad ; \quad 0<x<1$$

$$v(1) = 0$$

$$v(1) - v'(1) = -v'(1) = 0 \qquad (14b)$$

Thus under the definitions (Equations (12)), equation (11) becomes an unconstrained adjoint variational principle which corresponds to the physical system together with its adjoint.

$$i.e., \qquad \delta\{\int_0^1 (uv+u'v-v'u')dx + (v(0)-v'(0))u(0) + u'(1)v(1)\} = 0 \qquad (15)$$

In view of the linearity of the systems considered and the arbitrariness of δu and δv, the portion of equation (13) which yields the u-system may be considered separately from that which gives the v-system.

$$i.e., \qquad \int_0^1 (u''+u'+u)\delta v\,dx + u'(0)\delta v(0) - u(0)\delta v'(0) = 0 \qquad (16)$$

Further, there is no reason why the variations on v cannot be those performed on u. Thus,

$$\int_0^1 (u''+u'+u)\delta u\,dx + u'(0)\delta u(0) - u(0)\delta v'(0) = 0 \qquad (17)$$

One notes in passing that equation (17), unlike equation (15) is not of the form $\delta I = 0$, that is, no functional exists unless the adjoint system is included. Thus equation (17) might be more properly called a variational 'statement' as opposed to a 'principle'.

The Lagrange multiplier technique may be applied to the more general equations governing the motion of a linearly elastic solid. For example, consider the following system:

$$\frac{\partial U}{\partial u_k} + \rho\ddot{u}_k - \sigma_{\ell k,\ell} = 0$$

$$u_k - \bar{u}_k(t) = 0 \quad \text{on } S_u$$

$$- n_\ell \sigma_{\ell k} + \bar{F}_k = \text{on } S_f$$

$$u_k - \bar{u}_k(t_0) = 0 \quad ; \quad t = t_0$$

$$\dot{u}_k - \bar{v}_k = 0 \quad ; \quad t = t_0 \tag{18}$$

The result of applying the Lagrange multiplier technique to this system and its adjoint is the following variational statement

$$0 = \int_{t_0}^{t_1} \left[\int_V \delta L dV + \int_{S_N} \bar{F}\delta u_k dS + \int_{S_u} \delta\{n_\ell \sigma_{\ell k}(u_k - \bar{u}_k)\} dS \right] dt$$

$$+ \int_V dV\{-\dot{u}_k(t_1)\delta u_k(t_1) + \bar{v}_k \delta u_k(t_0)$$

$$+ [u_k(t_0) - \bar{u}_k(t_0)]\}\rho\delta\dot{u}_k(t_0) \tag{19}$$

where L is the Lagrange density

$$L = 1/2\rho\dot{u}_k\dot{u}_k - U(u_k, u_{k,\ell}, x_j)$$

Equation (19) and the result obtained by Tiersten [6] are identical except that in the interest of simplicity, no material surface of discontinuity has been considered in Eq. (19). It is to be noted that all variations are unconstrained so that the trial functions used in seeking an approximate solution need not satisfy any boundary or initial conditions a priori. If trial functions can be chosen that do satisfy some of the boundary constraints beforehand, convergence will usually be more rapid.

IV. APPLICATIONS.

Example 1: Wave Equation

$$S \frac{\partial^2 u}{\partial x^2} - \rho \frac{\partial^2 u}{\partial t^2} = 0$$

$$u(0,t) = g_0(t), \qquad u(\ell,t) = g_1(t)$$

$$u(x,0) = h_0(x), \qquad \dot{u}(x,0) = h_1(x) \tag{20}$$

Thus,

$$U = 1/2S(u')^2 \qquad \bar{F} \equiv 0$$

$$\sigma = \frac{\partial U}{\partial u'} = S_{u'}$$

Substituting the boundary conditions and the expressions for U and σ into Eq. (19) results in the following variational statement:

$$0 = \int_0^{t_1} \{ \int_0^{\ell} (a^2 u\delta u - u'\delta u')dx + [u(\ell,t)$$

$$-g_1(t)]\delta u'(\ell,t) - [u(0,t) - g_0(t)]\delta u'(0,t) + u'(\ell,t)\delta u(\ell,t)$$

$$-u'(0,t)\delta u(0,t)\}dt + \int_0^{\ell} a^2[-\dot{u}(x,t_1)\delta u(x,t_1)$$

$$+h_1(x)\delta u(x,0) + [u(x,0) - h_0(x)]\delta\dot{u}(x,0)]dx \tag{21}$$

where $a^2 = \rho/S$. A matrix formulation of Eq. (21) is achieved by substituting the approximation

$$u(x,t) = \sum_{j=1}^{NxN} a_j(x,t)c_j$$

as in the Ritz procedure but without any constraint requirements on the a_j set a priori. The result is a set of algebraic equations of the form

$$\sum_{j=1}^{NxN} k_{ij}c_j = r_i \qquad i=1,NxN \tag{22}$$

for the determination of the constants $c_j(\ell,t_1)$. Results are given in Tables 1 and 2 for the case:

$$g_0 = g_1 = h_1 = 0; \quad h_0(x) = \sin\pi x/\ell; \quad \ell = 2$$

The shape functions $a_j(x,t)$ are taken to be products of polynomials in x and t. Good convergence is obtained for N = 8. As expected, Tables 1 and 2 show a decline in accuracy as the interval of integration is doubled.

Example 2: Free Oscillator-Particle Mechanics

$$\ddot{u} + \omega^2 u = 0, \quad u(0) = u_0, \quad \dot{u}(0) = v_0 \tag{23}$$

The variational formulation for this problem is

$$\int_0^{t_1} (\dot{u}\delta\dot{u}-\omega^2 u\delta u)dt - \dot{u}(t_1)\delta u(t_1) + v_0\delta u(0) + u(0)\delta\dot{u}(0) - u_0\delta\dot{u}(0) = 0 \tag{24}$$

Table 3 gives the results for the case $u_0 = 0$, $v_0 = \omega = 2\pi$, and $t_1 = 1$. The assumed shape functions are polynomials in the time variable. A polynomial of order eight again gives good convergence.

Example 3: Response of a Beam to a Moving Mass

A concentrated mass is assumed to move at constant velocity v along the length of a uniform Euler beam, simply supported at each of its ends and having zero displacement and velocity at time t = 0. Under suitable definitions for k and m, the representative equations may be written [10]:

$$y^{iv} + k\ddot{y} + f(x,t) = 0$$

$$y(0,t) = y''(0,t) = y(1,t) = y''(1,t) = 0$$

$$y(x,0) = \dot{y}(x,0) = 0 \tag{25}$$

The function f(x,t) consists of a sum of inertial terms:

$$f(x,t) = m(\ddot{y}+2v\dot{y}'+g+v^2y'')\delta(x-vt)$$

where g denotes the gravitational constant and δ is the Dirac function. The appropriate variational equation is

$$\int_0^{t_1}\int_0^1 (y''\delta y''-k\dot{y}\delta\dot{y}+f(x,t)\delta y)dxdt$$
$$+ k\int_0^1 \{\dot{y}(x,t_1)\delta y(x,t_1) - y(x,0)\delta\dot{y}(x,0)\}dx = 0 \qquad (26)$$

A matrix approximation to Eq. (26) is obtained as in the first example, again using products of polynomials through order eight. The results are shown in Figure 1 as a comparison with values scaled from the experimental curves of Ayre, Jacobsen, and Hsu [11] for the case v = v*/4, v* being the lowest velocity to produce resonance when the load is a moving weight only. The magnitude assigned to the moving mass is 25% of the total mass of the beam of length L. The displacements have been normalized with respect to the maximum deflection produced if the weight was applied statically at midspan. This problem has also been treated previously by the author [10], using a conventional finite element method resulting in a set of differential equations in time. The numerical integration of these equations appeared to require a considerably longer computation time.

V. CONCLUSIONS. The unconstrained variational statement first developed and used by Tiersten for the solution of field displacements within a body containing a surface of discontinuity can indeed yield solutions to boundary-initial value problems. Further, the variational statement from which such solutions are possible can be formally constructed by the Lagrange multiplier method if the adjoint system is also considered.

ACKNOWLEDGEMENTS. The author is especially grateful for the enlightening conversations with H. F. Tiersten, Rensselaer Polytechnic Institute, and for the constructive comments of G. Anderson, Institut CERAC, Switzerland.

REFERENCES

1. Collatz, L., The Numerical Treatment of Differential Equations, 3rd ed, Springer-Verlag, 1960, pp. 202-207.

2. Rund, H., "The Reduction of Certain Boundary Value Problems to Variational Problems by Means of Transversality Conditions," Numerical Mathematics, Vol. 15, 1970, pp. 49-56.

3. Bailey, C. D., "Hamilton, Ritz and Elastodynamics," Transactions of the ASME, Dec. 1976, pp. 684-688.

4. Bailey, C. D., "The Method of Ritz Applied to the Equation of Hamilton," Computer Methods in Applied Mechanics and Engineering North Holland Publishing Co., 1976, pp. 235-247.

5. Courant, R. and Hilbert, D., Meth of Math Phys, Interscience Publications Inc., Vol. 1, Chap. IV, Secs. 5 and 9.2, 1953.

6. Tiersten, H. F., "Natural Boundary and Initial Conditions From a Modification of Hamilton's Principle," Journal of Mathematical Physics, 9,9, 1968, pp. 1445-1450.

7. Gurtin, M., "Variational Principles for Linear Initial-Value Problems," Quarterly of App. Math., Vol. XXII, No. 3, pp. 252-256.

8. Wu, J., "Solutions to Initial Value Problems by Use of Finite Elements - Unconstrained Variational Formulations," Journal of Sound and Vibration, 53(3), 1977, pp. 341-356.

9. Lanczos, C., The Variational Principles of Mechanics, 3rd ed., University of Toronto Press, Toronto, 1966, p. 147.

10. Simkins, T. E., "Structural Response to Moving Projectile Mass by the Finite Element Method," Watervliet Arsenal Tech. Rept. WVT-TR-75044, July 1975.

11. Ayre, R. S., Jacobsen, L. S., and Hsu, C. S., "Transverse Vibration of One and of Two Span Beams Under the Action of a Moving Mass Load," Proceedings of First National Congress on Applied Mechanics, June 1951.

Table 1 Solution to wave equation $0 \le x \le 2.0$; $0 \le t \le 2.0$ (exact values in parentheses)

t/x	0.0	0.4	0.8	1.0	1.2	1.6	2.0
0.0	.000060 (.000000)	.587782 (.588785)	.951066 (.951057)	1.000000 (1.000000)	.951063 (.951057)	.587785 (.587785)	.000056 (.000000)
0.4	.000014 (.000000)	.475532 (.475528)	.769422 (.769421)	.809006 (.809017)	.769421 (.769421)	.475533 (.475528)	.000013 (.000000)
0.8	−.000001 (.000000)	.181636 (.181636)	.293889 (.293893)	.309010 (.309017)	.293890 (.293893)	.181635 (.181636)	.000001 (.000000)
1.0	−.000003 (.000000)	−.000000 (.000000)	−.000002 (.000000)	−.000001 (.000000)	−.000001 (.000000)	−.000001 (.000000)	−.000002 (.000000)
1.2	−.000000 (.000000)	−.181637 (−.181636)	−.293889 (−.293893)	−.309010 (−.309017)	−.293890 (−.293893)	−.181636 (−.181636)	−.000001 (−.000000)
1.6	−.000013 (.000000)	−.475531 (−.475528)	−.769420 (−.769421)	−.809006 (−.809017)	−.769420 (−.769421)	−.475531 (−.475528)	−.000012 (−.000000)
2.0	−.000036 (.000000)	−.587786 (−.587785)	−.951055 (−.951057)	−.999986 (−1.000000)	−.951056 (−.951057)	−.587785 (−.587785)	−.000033 (−.000000)

Table 2 Solution to wave equation $0 \le x \le 2.0$; $0 \le t \le 4.0$ (exact values in parentheses)

t/x	0.0	0.4	0.8	1.0	1.2	1.6	2.0
0.0	.000037 (.000000)	.595863 (.587785)	.964161 (.951057)	1.013774 (1.000000)	.964161 (.951057)	.595867 (.587785)	.000031 (.000000)
0.8	.000008 (.000000)	.183947 (.181636)	.297637 (.293893)	.312952 (.309017)	.297638 (.293893)	.183948 (.181636)	.000007 (.000000)
1.6	−.000014 (.000000)	−.477206 (−.475528)	−.772139 (−.769421)	−.811866 (−.809017)	−.772139 (−.769421)	−.477207 (−.475528)	−.000012 (−.000000)
2.0	−.000013 (.000000)	−.587942 (−.587785)	−.951308 (−.951057)	−1.000253 (−1.000000)	−.951309 (−.951057)	−.587943 (−.587785)	−.000013 (−.000000)
2.4	−.000006 (.000000)	−.474475 (−.475528)	−.767707 (−.769421)	−.807204 (−.809017)	−.767707 (−.769421)	−.474475 (−.475528)	.000007 (−.000000)
3.2	−.000001 (.000000)	.181191 (.181636)	.293165 (.293893)	.308247 (.309017)	.293166 (.293893)	.181191 (.181636)	.000001 (.000000)
4.0	.000002 (.000000)	.587756 (.587785)	.951014 (.951057)	.999945 (1.000000)	.951014 (.951057)	.587755 (.587785)	.000005 (.000000)

Table 3 Solution to free oscillator problem $0 \le t \le t_f = 1.0$

t	Computed solution	Exact solution
0.0	.00305	0.00000
0.1	.58656	.58779
0.2	.95218	.95106
0.3	.95159	.95106
0.4	.58670	.58779
0.5	−.00058	0.00000
0.6	−.58704	−.58779
0.7	−.95058	−.95106
0.8	−.95147	−.95106
0.9	−.58775	−.58779
1.0	−.00001	0.00000

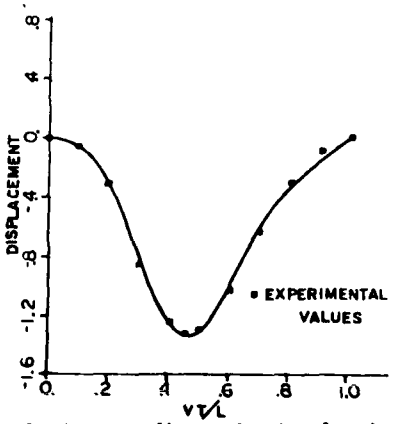


Fig. 1 Displacement of beam at location of moving mass.

FINITE DEFORMATION PLASTICITY AND PLASTIC INSTABILITY

S. Nemat-Nasser*
Northwestern University, Evanston, Ill. 60201

ABSTRACT. The phenomenon of localization of deformation in metals and geotechnical materials is briefly discussed, and a few illustrative examples are given. Then the mathematical procedure for the calculation of the incipience of localization is presented and a few specific examples are worked out. In particular, it is shown that many commonly used constitutive relations yield strange and unrealistic directions for the localization in biaxial extension. After this, a plasticity theory which includes plastic compressibility and internal friction is reviewed with application to simple shearing of granular materials. Finally, a brief account of a statistical model for such simple shearing is presented.

I. INTRODUCTION. Since Hadamard's [1] pioneering work on elasticity, Hill's [2, 3] contributions to plasticity, and Thomas' [4] formulation of dynamical and kinematical conditions at surfaces of discontinuities, considerable effort has been devoted to understand and quantitatively predict unstable flow by localized deformation of metals and geological materials. Most of these efforts concern application of the basic stability theory to specific problems. It turns out, however, that a proper constitutive description which accurately accounts for the physics of the material involved, is of fundamental importance, otherwise the theory yields strange results for the direction of the localized deformation over a wide range of material parameters. This has been illustrated for a number of commonly used constitutive relations by Nemat-Nasser et al. [5], and is now being extensively examined [6].

In this paper we shall briefly illustrate the above mentioned strange result first in some special cases, and then give an outline of a plasticity theory which accounts for plastic compressibility and internal friction. Finally, we shall examine in simple shear, and from a microscopic statistical point of view, plastic flow of granular materials. We are able to give an almost complete description of the material behavior in this simple stress state. We therefore hope that the basic approach will serve as a model for the description of the mechanical behavior of materials under more general loading conditions.

To bring into focus the nature of the physical problem in question, we shall first briefly discuss a few specific examples.

II. EXAMPLES OF LOCALIZED DEFORMATIONS. It is known that when a thin metal sheet is subjected to uniaxial or biaxial extension, it may lose stability by necking or by the formation of shear bands. Figure 1, taken from Weinrich and French [7], illustrates this for α-brass sheet tensile specimens, where a localized shear band is formed at a stress of

* Professor of Civil Engineering and Applied Mathematics.

Fig. 1: Out-of-plane shear band in α-brass sheet tensile specimen stressed at 400 MPa; taken from Ref. [7], p. 320.

about 400 MPa. Another example is shown in Figure 2, taken from Cottrell [8], which illustrates shear bands in cadmium under uniaxial extension. Here the deformation is almost totally plastic and irreversible. Despite this physical fact, many people still attempt to describe localized bands in terms of elasticity theory.

Fig. 2: Localized slip in a single crystal of cadmium; taken from Ref. [8], Fig. 3 (After Bilby).

Similar unstable deformations exist for soils and rocks under compressive states of stress. Figure 3, taken from Taylor [9], illustrates both a diffused-type bulging instability and an unstable deformation by a localized shear band. In nature, under large tectonic compressive stress and shearing, rocks and granular materials can flow in unstable modes which lead to the formation of localized zones of highly densified materials. Because of this densification, these layers withstand erosion better than the country rock, and therefore they are conspicuously displayed in the field; see Figs. 1 to 7 of [10].

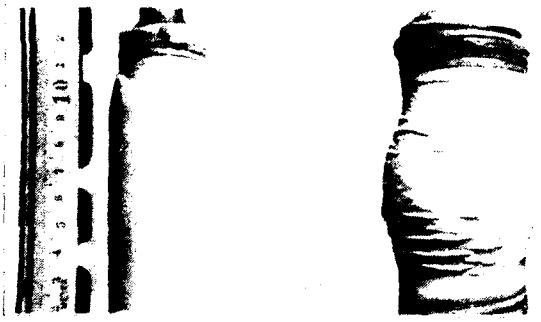


Fig. 3: Shear band (a) and bulging instability (b) in a cylindrical sample of sand; taken from Ref. [9], p. 334.

In the case of metals which often contain nonmetallic second-phase microscopic inclusions, plastic volume changes may be of some significance during unstable flows. Here, voids are often generated at second-phase particles, and they grow during the process of large plastic deformations. The hydrostatic tension facilitates void growth and therefore affects macroscopic instabilities. Hence, a constitutive description must include both the plastic compressibility and the pressure (or tension) sensitivity of the material.

In the case of soils and rocks, internal friction, as well as inelastic volumetric changes are of paramount significance and must be included as first-order effects in the corresponding constitutive relations. We shall discuss these and related results in what follows. However, we shall first illustrate the mathematical setting involved in the calculation of the incipience of localized deformations [2] and demonstrate how many commonly used plasticity constitutive relations lead to strange results [5, 6].

III. LOCALIZED DEFORMATION AND CHARACTERISTICS. Consider an infinitely extended homogeneous body under a homogeneous initial state of stress, i.e. the Cauchy stress components σ_{ij} are constant throughout the body; we use a fixed rectangular Cartesian coordinate system x_i, $i = 1,2,3$.

Consider a bifurcation from this homogeneous state of initial stress, defined by prescribing velocity field v_i with the corresponding deformation rate $D_{ij} = \frac{1}{2}(v_{i,j} + v_{j,i})$ and spin $W_{ij} = \frac{1}{2}(v_{i,j} - v_{j,i})$, and denote the rate of change of nominal (nonsymmetric) stress (referred to and measured per unit current area) by $\dot{n}_{ij}$. We have $\dot{n}_{ij} = \dot{\sigma}_{ij} + v_{k,k}\sigma_{ij} - v_{i,k}\sigma_{kj}$, where $\dot{\sigma}_{ij}$ is the material rate of the Cauchy stress, and the repeated indices are summed.

For continuing equilibrium we must have

$$\dot{n}_{ij,i} = 0 \ , \quad \langle \dot{n}_{ij} \rangle \nu_i = 0 \ , \tag{3.1}$$

where the first equation must hold within the region currently occupied by the body (rate of change of body forces assumed zero), and the second equation must hold across any interior surface with unit normal ν_i. In this latter equation $\langle \dot{n}_{ij} \rangle$ denotes the difference between the two values of the enclosed quantity calculated on the opposite faces of the considered surface along the normal ν. Equation $(3.1)_2$ simply ensures the continuity of tractions across any interior material surface.

For a large class of rate-independent materials (which includes hyperelasticity, hypoelasticity, and elastoplasticity), the rate constitutive relations can be expressed as (compressible)

$$\dot{n}_{ij} = C_{ijk\ell} \, v_{\ell,k} \ , \tag{3.2}$$

where $C_{ijk\ell}$ depends on the state of stress and possibly the history of deformation, but is independent of the velocity gradient $v_{i,j}$. Substitution from (3.2) into $(3.1)_1$ yields

$$C_{ijk\ell} \, v_{\ell,ki} = 0 \ . \tag{3.3}$$

This is a system of second-order partial differential equations with constant coefficients (homogeneous body under homogeneous initial state of stress). Localized deformations may occur if this system of equations admits real characteristics, in which case the velocity gradient $v_{i,j}$ can admit discontinuities across the characteristics. Consider a velocity field of the form

$$v_i = \eta_i \, f(\underset{\sim}{x} \cdot \underset{\sim}{\nu}) \ , \tag{3.4}$$

where $\underset{\sim}{\eta}$ and $\underset{\sim}{\nu}$ are unit vectors. Substitution into (3.3) results in

$$[C_{ijk\ell} \, \nu_k \, \nu_i] \, \eta_\ell = 0 \quad \text{for} \quad f'' \neq 0 \ , \tag{3.5}$$

which is a system of three linear homogeneous equations for η_ℓ, $\ell = 1,2,3$. Nontrivial solutions exist if and only if

$$\det \left| C_{ijk\ell} \, \nu_k \, \nu_i \right| = 0 \tag{3.6}$$

which is the corresponding characteristic equation.

When the material is incompressible, we have $v_{i,i} = 0$, and (3.2) must be replaced by

$$\dot{n}_{ij} = \bar{C}_{ijk\ell} \, v_{\ell,k} + \dot{p} \, \delta_{ij} \ , \tag{3.7}$$

where $\dot{p}$ is to be obtained as part of the solution. The incompressibility

condition and (3.4) yield $\eta_i \nu_i = 0$, so that $\underline{\eta}$ is perpendicular to $\underline{\nu}$. Now, substitution from (3.7) into $(3.1)_1$ yields

$$\bar{C}_{ijk\ell} v_{\ell,ki} + \dot{p}_{,j} = 0 . \tag{3.8}$$

With v_i given by (3.4) and

$$\dot{p} = \xi(\underline{x} \cdot \underline{\nu}) , \tag{3.9}$$

we obtain from (3.8)

$$(\bar{C}_{ijk\ell} \nu_k \nu_i) \eta_\ell f'' + \xi' \nu_j = 0 , \quad \eta_i \nu_i = 0 . \tag{3.10}$$

From (3.10) we have

$$\bar{C}_{ijk\ell} \nu_k \nu_i \eta_j \eta_\ell = 0 , \eta_i \nu_i = 0 , \tag{3.11}$$

which is the characteristic equation.

IV. THIN SHEETS IN BIAXIAL EXTENSION. As an example, consider a biaxial extension of a thin sheet with initially uniform thickness H, and initial state of (Cauchy) stress $\sigma_{11} = \sigma_1$, $\sigma_{22} = \sigma_2$, $\sigma_1 > \sigma_2 > 0$; all other components of σ_{ij} equal to zero. We choose the x_1- and x_2-axes in the plane of the sheet, and the x_3-axis perpendicular to them, and consider three possible unstable modes, as follows:

1) In-plane shear band with $\dot{H}_{,\alpha} = 0$, $\alpha = 1,2$. In this case, the plate thickness remains uniform during the bifurcation; see Fig. 4a.

2) Out-of-plane shear band with $v_2 = 0$. In this case, the shear band occurs in the x_1,x_3-plane; see Fig. 4b.

3) Necking, for which $\dot{H}_{,\alpha} \neq 0$; see Fig. 4c.

In all three cases the condition of plane stress requires that $\dot{n}_{33} = 0$.

We consider both compressible and incompressible materials*, and with $\overset{*}{\sigma}_{ij} = \dot{\sigma}_{ij} - W_{ik}\sigma_{kj} - W_{jk}\sigma_{ki}$ denoting the (objective) Jaumann rate of the Cauchy stress, reduce (for the considered state of stress) the three-dimensional stress rate-strain rate relations to the following form:

$$\overset{*}{\sigma}_{11} = \bar{a}_1 D_{11} + \bar{a}_2 D_{22} + \bar{a}_3 D_{33}, \quad \overset{*}{\sigma}_{22} = \bar{b}_1 D_{11} + \bar{b}_2 D_{22} + \bar{b}_3 D_{33},$$

$$\overset{*}{\sigma}_{12} = c_1 D_{12} , \quad \overset{*}{\sigma}_{23} = c_2 D_{23} , \quad \overset{*}{\sigma}_{31} = c_3 D_{31}, \tag{4.1}$$

* For a more thorough discussion of various cases see [6].

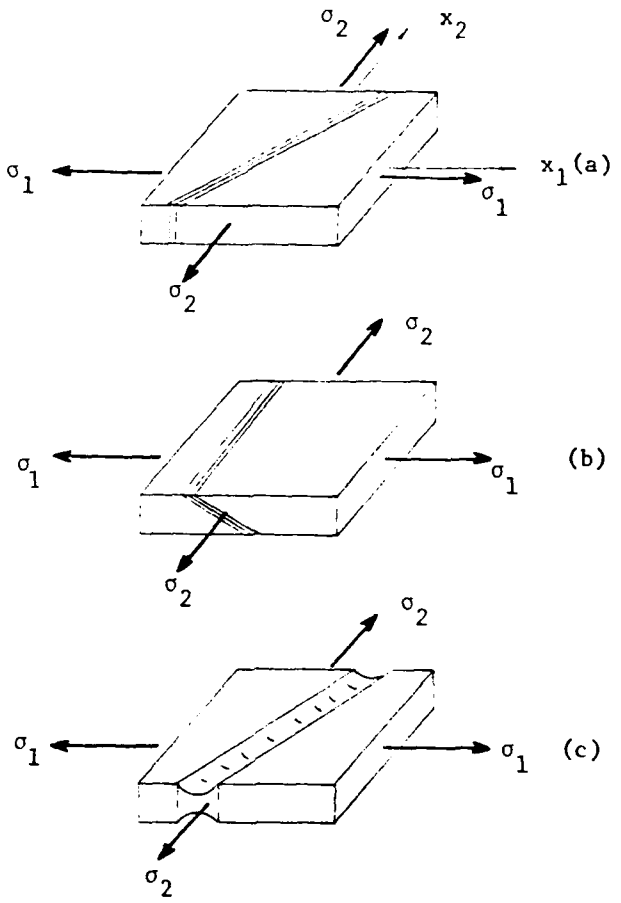


Fig. 4: Localized deformations in biaxial extension of a thin sheet; (a) In-plane shear; (b) Out-of-plane shear; and (c) Necking.

where $\overset{*}{\sigma}_{ij} = \overset{*}{\sigma}_{ji}$, and for the compressible case, the condition $\dot{n}_{33} = 0$ gives

$$D_{33} = - (d_1 D_{11} + d_2 D_{22}) ; \qquad (4.2)$$

for the incompressible case we have $d_1 = d_2 = 1$. It should be noted that the coefficients $\bar{a}_i$, $\bar{b}_i$, $i = 1,2,3$, will have <u>different values</u> for compressible and incompressible cases; in the latter case, $\dot{n}_{33} = 0$ is used to eliminate the pressure rate.

4.1 In-Plane Shear: Since $\dot{H}_{,\alpha} = 0$, we have

$$\dot{n}_{\alpha\beta,\alpha} = \frac{1}{H}(H\sigma_{\alpha\beta})^{\cdot}_{,\alpha} = \dot{\sigma}_{\alpha\beta,\alpha} = 0\ ,$$
$$\langle \dot{n}_{\alpha\beta} \rangle \nu_\alpha = \left\langle \frac{1}{H}(H\sigma_{\alpha\beta})^{\cdot} \right\rangle \nu_\alpha = \langle \dot{\sigma}_{\alpha\beta} \rangle \nu_\alpha = 0, \quad \alpha,\beta = 1,2. \tag{4.3}$$

Moreover, the velocity field is written as $v_\alpha = \eta_\alpha f(\underset{\sim}{x} \cdot \underset{\sim}{\nu})$ in the x_1,x_2-plane, and hence the characteristic equation becomes

$$a_1(c_1+\sigma)\nu_1^4 + 2[a_1b_2 - a_2b_1 - \tfrac{1}{2}a_2(c_1-\sigma) - \tfrac{1}{2}b_1(c_1+\sigma)]\ \nu_1^2\nu_2^2 + b_2(c_1-\sigma)\ \nu_2^4 = 0,\ \sigma = \sigma_1 - \sigma_2, \tag{4.4}$$

where $\bar{a}_\alpha = a_\alpha - d_\alpha a_3$ and $b_\alpha = \bar{b}_\alpha - d_\alpha \bar{b}_3$, $\alpha = 1,2$. The basic system of differential equations is hyperbolic, parabolic, or elliptic depending on whether (4.4) has 4, 2, or no real roots for ν_1/ν_2.

4.2 Out-of-Plane Shear: In this case $v_2 = 0$ and hence $D_{22} = 0$. Thus, $D_{33} = -d_1 D_{11}$ where $d_1 = 1$ if the material is incompressible. From this condition we obtain, with $v_\alpha = \eta_\alpha f(\underset{\sim}{x} \cdot \underset{\sim}{\nu})$, $\alpha = 1,3$,

$$\eta_3 \nu_3 + d_1 \eta_1 \nu_1 = 0, \tag{4.5}$$

and the continuity of tractions yields,

$$[\bar{a}_1 \nu_1^2 + \tfrac{1}{2}(c_3 - \sigma_1)\nu_3^2]\eta_1 + [\bar{a}_3 + \tfrac{1}{2}(c_3+\sigma_1)]\nu_1\nu_3\eta_3 = 0,$$
$$(c_3-\sigma_1)\nu_1\nu_3\eta_1 + (c_3+\sigma_1)\nu_1^2\eta_3 = 0. \tag{4.6}$$

Equations (4.5) and (4.6) have a nontrivial solution given by

$$\eta_3 = \nu_1 = 0, \quad \eta_1 = \nu_3 = 1, \quad \sigma_1 = c_3, \tag{4.7}$$

i.e. the shear band being parallel with the direction of the maximum tensile stress σ_1. This is indeed a very strange result. Several authors have considered out-of-plane shear band but they seem to have not realized that these bands are parallel with the maximum stress axis; see, for example Rice [11] and Hutchinson [12].

4.3 Necking: In this case $\dot{H}_{,\alpha} \neq 0$, and we obtain from $\langle \dot{n}_{\alpha\beta} \rangle \nu_\alpha = 0$, $\alpha,\beta = 1,2$, the following characteristic equation [6]:

$$(a_1 - d_1\sigma_1)(c_1+\sigma)\nu_1^4 + 2\{(a_1-d_1\sigma_1)(b_2-d_2\sigma_2) + \tfrac{1}{4}(c_1^2-\sigma^2) - [a_2 + (1-d_2)\sigma_1 + \tfrac{1}{2}(c_1-\hat{\sigma})][b_1 + (1-d_1)\sigma_2 + \tfrac{1}{2}(c_1-\hat{\sigma})]\}\ \nu_1^2\nu_2^2$$

$$+ (b_2 - d_2\sigma_2)(c_1 - \sigma)\nu_2^4 = 0 \quad \hat{\sigma} = \sigma_1 + \sigma_2 \ . \tag{4.8}$$

4.4 Examples: A systematic study of localization for various special constitutive relations is given by Nemat-Nasser and Iwakuma [6]. Here we present a few typical results.

As our first example, consider the hypoelastic case,

$$\overset{*}{\sigma}_{11} = \frac{2\mu}{1-\nu}(D_{11} + \nu D_{22}) \ , \ \overset{*}{\sigma}_{22} = \frac{2\mu}{1-\nu}(\nu D_{11} + D_{22}), \ \overset{*}{\sigma}_{12} = 2\mu D_{12},$$

$$D_{33} = -\frac{\nu}{1-\nu}(D_{11} + D_{22}) \ . \tag{4.9}$$

We set

$$a = \sigma_2/\sigma_1, \ 0 \leq a < 1, \ S = \sigma_1/\mu, \ b = \nu_2^2 \ \text{ or } = \nu_3^2, \tag{4.10}$$

and calculate from each characteristic equation the values of b which correspond to the minimum value of S, as follows.

1) In-Plane Shear:

$$S_{min} = \frac{2}{1-a} \ \text{(parallel to the } \sigma_1\text{-direction)}. \tag{4.11}$$

2) Out-of-Plane Shear:

$$S_{min} = 2 \ \text{(parallel to the } \sigma_1\text{-direction)}. \tag{4.12}$$

3) Necking:

$$S_{min} = \frac{2}{1-a} \ \text{(parallel to the } \sigma_1\text{-direction) for } (1-a) > \nu,$$
$$S_{min} = \frac{2}{\nu} \ \text{(normal to the } \sigma_1\text{-direction) for } (1-a) < \nu. \tag{4.13}$$

As our second (and last) example, we consider the incompressible model used by Stören and Rice [13], i.e.

$$\overset{*}{\sigma}_{ij} = \frac{2h}{N}\left[D_{ij} - (1 - N)\frac{\sigma'_{ij}\,\sigma'_k}{2\,\bar{\tau}^2} D_k\right] + \dot{p}\,\delta_{ij},$$

$\bar{\tau}^2 = \frac{1}{2}\sigma'_{ij}\,\sigma'_{ij}$, prime denoting the deviatoric part.

Then we have the following results, see [6].

1) In-Plane Shear:

$$[\sigma_1]_{min} = \frac{2h}{N}\frac{1}{1-a} \ \text{(parallel to the } \sigma_1\text{-direction) for } N \geq \frac{1}{3} \ . \tag{4.14}$$

For $0 < N < \frac{1}{3}$, the result depends on the value of a; see [6].

2) Out-of-Plane Shear:

$$[\sigma_1]_{min} = \frac{2h}{N} \quad \text{(parallel to the } \sigma_1\text{-direction)}. \tag{4.15}$$

3) Necking: In this case the results depend on the combination of N and a. Typical calculations are given in Fig. 5, taken from [6]. It is seen that over a wide range, the theory predicts bifurcation parallel to the maximum stress direction.

V. A PLASTICITY THEORY WITH COMPRESSIBILITY AND INTERNAL FRICTION. In this section we briefly review a plasticity theory with compressibility and internal friction which recently has been developed in [14], and then apply the results to the case of simple shear. The theory is based on a modification of the usual J_2 flow rule; the subscript 2 will be dropped in the sequel. It considers a flow potential g, and a yield function f, defined by

$$g \equiv \sqrt{J} + G(I, \Delta, \xi), \quad f \equiv \sqrt{J} - F(I, \Delta, \xi), \tag{5.1}$$

where

$$J = \frac{1}{2}\sigma'_{ij}\sigma'_{ij}, \quad I = \sigma_{ii},$$

$$\Delta = \int_0^\theta \frac{\rho_0}{\rho} D^p_{ii}\, d\theta, \quad \xi = \int_0^\theta \frac{\rho_0}{\rho} \sigma'_{ij} D^p_{ij}\, d\theta \ ; \tag{5.2}$$

here, prime denotes the deviatoric part, θ is a monotone increasing (or decreasing) time-like parameter, ρ_0 and ρ are the mass-densities in the reference and current configurations, respectively, and D^p_{ij} denotes the plastic part of the deformation rate tensor given by

$$D^p_{ij} = \dot{\lambda} \frac{\partial g}{\partial \sigma_{ij}} . \tag{5.3}$$

Now using the usual procedure and $\dot{f} = 0$, where a superposed dot denotes differentiation with respect to θ, we obtain

$$D^{p\,\prime}_{ij} = \frac{\sigma'_{ij}}{2H\sqrt{J}} \left(\frac{1}{2\sqrt{J}} \sigma'_{k\ell} \overset{*}{\sigma}_{k\ell} - \frac{\partial F}{\partial I} \overset{*}{\sigma}_{kk}\right),$$

$$D^p_{kk} = \frac{3}{H} \frac{\partial G}{I} \left(\frac{1}{2\sqrt{J}} \sigma'_{k\ell} \overset{*}{\sigma}_{k\ell} - \frac{\partial F}{\partial I} \overset{*}{\sigma}_{kk}\right), \tag{5.4}$$

where the hardening parameter is given by

$$H = \frac{\rho_0}{\rho} \left(3 \frac{\partial G}{\partial I} \frac{\partial F}{\partial \Delta} + \sqrt{J} \frac{\partial F}{\partial \xi}\right), \tag{5.5}$$

and the dilatancy factor $\partial G/\partial I$ is defined by

$$3 \frac{\partial G}{\partial I} = \frac{\sqrt{J}\, D^p_{kk}}{\sigma'_{ij} D^p_{ij}} . \tag{5.6}$$

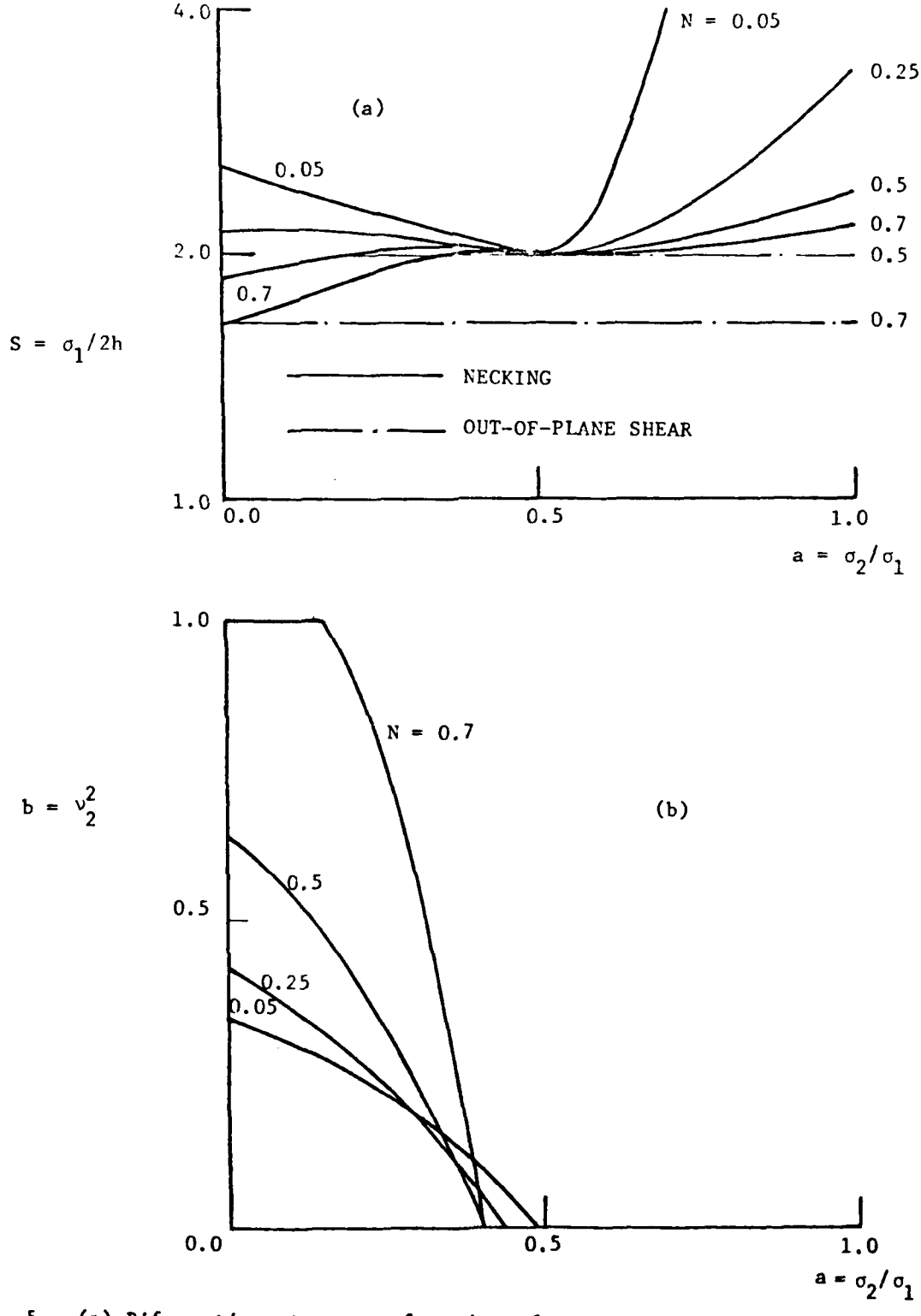


Fig. 5: (a) Bifurcation stress as function of stress ratio; (b) Direction of necking as function of stress ratio.

Equation (5.5) may be written as $H = h_1 + h$, where

$$h = \frac{\rho_0}{\rho} \sqrt{J} \frac{\partial F}{\partial \xi} . \quad h_1 = 3 \frac{\rho_0}{\rho} \frac{\partial G}{\partial I} \frac{\partial F}{\partial \Delta} ; \tag{5.7}$$

the first quantity denotes hardening due to distortional effects, and the second quantity represents hardening (or softening) due to volumetric changes. We refer to h_1 as the density-hardening parameter.

We shall now apply the above results to the simple shearing of granular materials. The state of stress is shown in Fig. 6. We shall assume that the grains are rigid, and therefore there is no elastic deformation. Hence the superposed p will be dropped. We now observe that $J = \tau^2$, $D_{12} = D_{21} = \frac{1}{2}\dot{\gamma}$, and $\dot{\xi} = \rho_0 \tau \dot{\gamma}/\rho$. The state of stress is defined by (τ, σ), and the state of deformation rate by $(\dot{\gamma}, \dot{v}/v)$, where $\dot{v}/v$ is the rate of volume change per unit current volume.

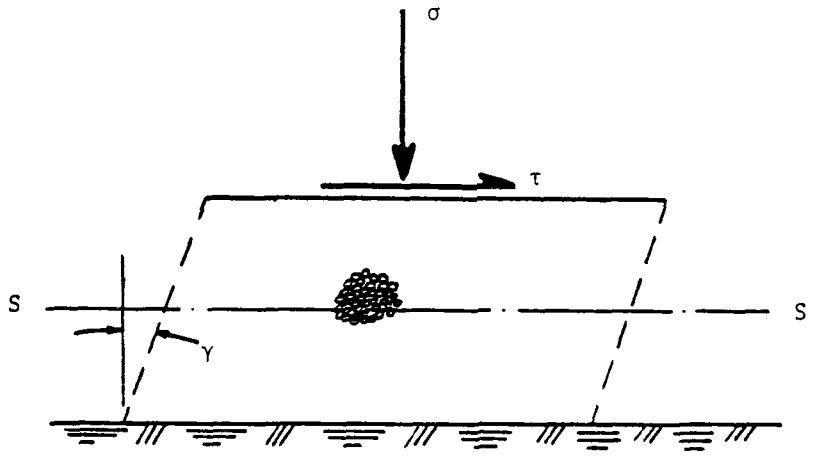


Fig 6: Granular materials sheared under vertical pressure.

The flow potential and yield function then become

$$g \equiv \tau + G(\sigma, \Delta, \xi), \quad f \equiv \tau - F(\sigma, \Delta, \xi), \tag{5.8}$$

and simple calculation yields

$$\frac{d\tau}{d\gamma} - \frac{\partial F}{\partial \sigma} \frac{d\sigma}{d\gamma} = H,$$

$$H = (1 - \Delta) \frac{\partial G}{\partial \sigma} \frac{\partial F}{\partial \Delta} + \frac{\partial F}{\partial \gamma} . \tag{5.9}$$

From (5.6) we have

$$\frac{\partial G}{\partial \sigma} = \frac{1}{v} \frac{\dot{v}}{\dot{\gamma}} \tag{5.10}$$

which shows that $\partial G/\partial\sigma$ is the dilatancy or the rate of volume change per unit rate of distortion. From $(5.8)_2$ on the other hand, we obtain with Δ and ξ constant,

$$df = d\tau - \frac{\partial F}{\partial \sigma} d\sigma = 0,$$

or

$$\frac{\partial F}{\partial \sigma} = \frac{d\tau}{d\sigma} , \tag{5.11}$$

and hence $\partial F/\partial\sigma$ is the instantaneous coefficient of overall friction. Therefore the rate of frictional loss per unit volume is $(\partial F/\partial\sigma)\sigma\dot{\gamma}$ which must equal the rate of plastic work, $-\sigma\dot{v}/v + \tau\dot{\gamma}$, resulting in

$$-\frac{1}{v}\frac{\dot{v}}{\dot{\gamma}} = \frac{\partial F}{\partial \sigma} - \frac{\tau}{\sigma} . \tag{5.12}$$

In the τ,σ-plane

$$\frac{\partial F}{\partial \sigma} - \frac{\tau}{\sigma} = 0 \tag{5.13}$$

defines the locus of points for which $\dot{v} = 0$. This is the critical curve. States above this curve dilate and those below densify; see Fig. 7.

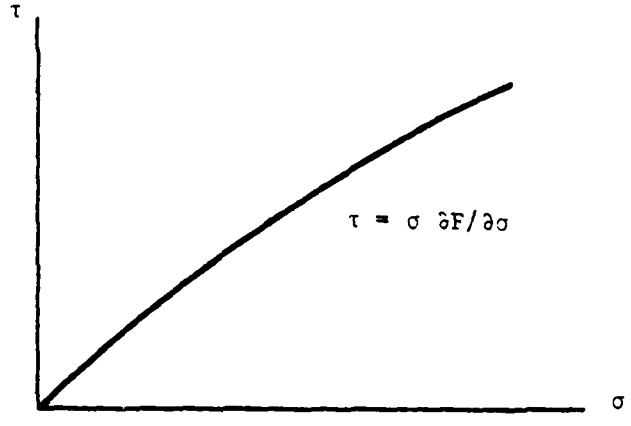


Fig. 7: The critical curve $\tau = \sigma\ \partial F/\partial\sigma$ in the τ,σ-plane.

Consider continuous shearing under constant normal pressure, σ = constant. The differential equation (5.9) then becomes

$$\frac{d\tau}{d\gamma} = (1 - \Delta)\frac{\partial F}{\partial \Delta}\left[\frac{\partial F}{\partial \sigma} - \frac{\tau}{\sigma}\right] + \frac{\partial F}{\partial \gamma} = a(\bar{M} - \frac{\tau}{\sigma}) + h(\gamma) , \tag{5.14}$$

where the definition of new terms in this equation is clear from the context. The distortional hardening $h(\gamma)$ plays an important role in characterizing the shear stress--plastic shear strain relation. If h drops to zero quickly with increasing γ, then the solution of (5.14) would resemble the behavior of loosely packed cohesionless granules. On the other hand, if the decrease of h with increasing γ is gentle, then the stress-strain relation admits a peak. In Fig. 8 these are shown by curves (1) and (2), respectively.

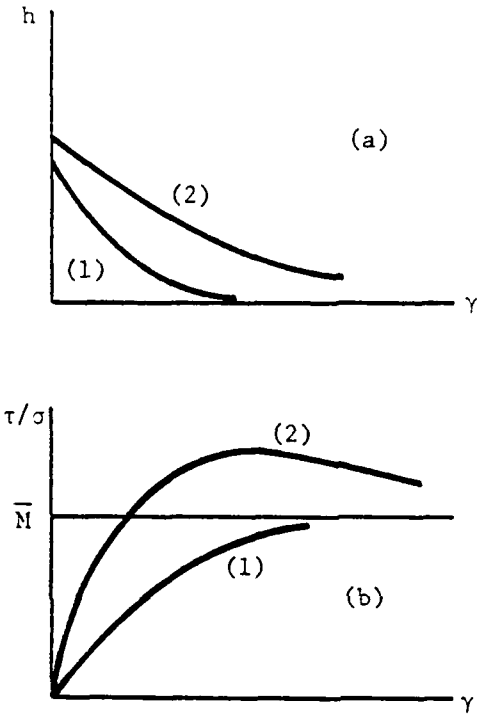


Fig. 8: (a) Variation of h with γ; (b) Normalized shear stress as function of plastic shear strain; curve (1) is for loosely packed, and curve (2) is for densely packed sample of granular materials.

It is easy to show that the void ratio $e = V_v/V_s$ which is the ratio of the void volume to the solid volume, is given by [14]

$$e = e_0 - (1 + e_0)\,[1 - \exp\{\int_0^\gamma \frac{\partial G}{\partial p}\, d\gamma\}]$$

$$= e_0 - (1 + e_0)\,[1 - \exp\{\int_0^\gamma (\bar{M} - \tau/\sigma)\, d\gamma\}],$$

where e_0 is the initial void ratio. The variation of e with γ depends on the corresponding τ,γ-relation. In Fig. 9 we have displayed two curves corresponding to the curves shown in Fig. 7. As is seen, curve (2) represents the behavior of densely packed granules, and curve (1) that of loosely packed ones; for a more detailed discussion and comparison with some experimental results, see [14].

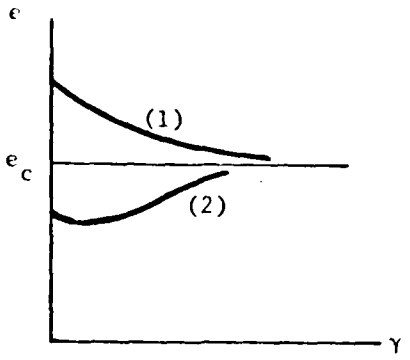


Fig. 9: Variation of void ratio with plastic shear strain; curve (1) is for loosely packed, and curve (2) for densely packed sample of granular materials.

VI. A SIMPLE STATISTICAL MODEL. For a simple shearing of granular materials a simple statistical model recently has been proposed by the writer [15], which accounts for the observed initial densification, subsequent dilatancy when the material is densely packed, and a net amount of densification upon the completion of each cycle in cyclic shearing. Here we briefly discuss this model.

When a granular material is sheared under normal pressure, as shown in Fig. 6, the individual grains do not move along the horizontal line, SS, which marks the macroscopic direction of shearing, rather they slide over each other along a wavy line denoted by S'S' in Fig. 10. This results in a net amount of densification or dilatancy depending on whether, on average, more particles move down or move up.

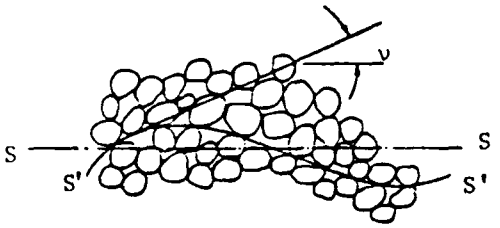


Fig. 10: Shearing of granular material in the SS-direction results in motion of individual grains along the S'S'-curve; ν is the dilatancy angle.

Consider a typical particle and let ν be the angle that the tangent to the S'S'-curve makes with the SS-direction. This angle will be called the "dilatancy angle."

In a given sample which contains a very large number of particles, there exist many groups of particles, each with its own dilatancy angle at each instant. Let $p(\nu)d\nu$ be the volume fraction of particles whose dilatancy angle is between ν and $\nu + d\nu$. To calculate the overall rate of volume change, we obtain the rate of volume change associated with each group, and then integrate the results over the entire volume, using the weighting function $p(\nu)$.

To this end we calculate the rate of distortional work per unit volume associated with a group of particles whose dilatancy angle is ν. With ϕ_μ denoting the actual particle to particle angle of friction, this rate of work is given by, see [15],

$$\dot{w}' = \frac{\tau \cos \phi_\mu}{\cos(\phi_\mu + \nu)\sin \nu} \frac{\dot{v}}{v}, \tag{6.1}$$

where prime denotes the distortional part. Macroscopically, the rate of distortional work per unit volume is $\dot{w}' = \tau\dot{\gamma}$, and hence, (6.1) yields

$$\frac{1}{v}\frac{\dot{v}}{\dot{\gamma}} = \frac{\cos(\phi_\mu + \nu)\sin \nu}{\cos \phi_\mu}; \tag{6.2}$$

compare with (5.12).

Now, if the dilatancy angle for a given macroscopic sample ranges over ν_0^- to ν_0^+, we obtain for the overall rate of dilatancy per unit rate of distortion,

$$\frac{1}{v}\frac{\dot{v}}{\dot{\gamma}} = \frac{1}{\cos \phi_\mu} \int_{\nu_0^-}^{\nu_0^+} p(\nu) \cos(\phi_\mu + \nu) \sin \nu \, d\nu, \tag{6.3}$$

where

$$\int_{\nu_0^-}^{\nu_0^+} p(\nu)\, d\nu = 1 . \tag{6.4}$$

Equation (6.3) has all the attributes necessary to yield the behavior of granular materials in simple shear.

To begin with, we observe that even if the density function $p(\nu)$ is symmetric about $\nu = 0$, the right-hand side of (6.3) becomes negative, i.e. densification. However, since the normal stress σ facilitates the movement of particles which are sliding down (negative ν) and hinders those which are sliding up (positive ν), we expect that, initially $p(\nu)$ should be biased toward negative ν's. This results in a larger initial

densification, particularly for loosely packed granules in which the downward motion of particles is less apt to be hindered by neighboring particles.

As most particles become mobilized in a densely packed sample of granules, $p(\nu)$ tends to become more biased toward positive ν's, and eventually leads to a positive value for the integral in Eq. (6.3). This results in dilatancy. It is easy to argue in this manner, that upon a load reversal, the integral in (6.3) yields a net negative value, and hence a net amount of densification upon the completion of half a cycle in cyclic shearing.

It is easy to show that the present model yields shear stress-shear strain relations with observed characteristics, depending on the variation of $p(\nu)$ with straining. In particular, with the aid of very simple and physically reasonable assumptions about $p(\nu)$, curves (1) and (2) of Fig. 8 are obtained. We shall not discuss the details here, and refer the reader to [15]. The author and his coworkers are now engaged in a systematic experimental and theoretical investigation of the static and dynamic behavior of granular materials, and results of these investigations will be reported in due time elsewhere.

ACKNOWLEDGMENT. This work has been supported in part by the National Science Foundation under Grant No. ENG76-03921 and in part by the U.S. Geological Survey under Contract No. 14-08-0001-17770 to Northwestern University.

REFERENCES

1. Hadamard, J., *Lecons sur la Propagation des Ondes et les Equations de l'Hydrodynamique*, Paris (1903) Ch. 6.

2. Hill, R., "Acceleration Waves in Solids," *J. Mech. Phys. Solids* 10 (1962) 1-16.

3. Hill, R., *Progress in Solid Mechanics*, Vol. II, North-Holland Publ. Co. (1961) Ch. 6.

4. Thomas, T. Y., *Plastic Flow and Fracture in Solids*, Academic Press, New York (1961).

5. Nemat-Nasser, S., Shokooh, A., and Taya, M., "On Localized Deformation in Biaxial Tension," Earthquake Research and Engineering Laboratory, Technical Report No. 79-2-13, Dept. Civil Engrg., Northwestern University, Evanston, Ill., February 1979.

6. Nemat-Nasser, S. and Iwakuma, T., "Localization of Deformation in Biaxial Loading," in preparation.

7. Weinrich, P. F. and French, I.E., "The Influence of Hydrostatic Pressure on the Fracture Mechanisms of Sheet Tensile Specimens of Copper and Brass," *Acta Metallurg.* 24(1976) 317-322.

8. Cottrell, A.H., *Dislocations and Plastic Flow in Crystals*, Oxford University Press, London (1953).

9. Taylor, D. W., *Soil Mechanics*, John Wiley & Sons, Inc. (1948).

10. Aydin, A. and Johnson, A. M., "Development of Faults as Zones of Deformation Bands and as Slip Surfaces in Sandstone," *Pageoph.* 116(1978) 931-942.

11. Rice, J. R., The Localization of Plastic Deformation," in *Theoretical and Applied Mechanics*, Koiter, W. T. (ed.), North-Holland Publ. Co. (1976) 207-220.

12. Hutchinson, J. W., "A Survey of Some Recent Work on the Mechanics of Necking," in *Proc. 8th U.S. National Congr. Appl. Mech.*, U.C.L.A.; to appear.

13. Stören, S. and Rice, J. R., "Localized Necking in Thin Sheets," *J. Mech. Phys. Solids* 23(1975) 421-441.

14. Nemat-Nasser, S. and Shokooh, A., "On Finite Plastic Flows of Compressible Materials with Internal Friction," Earthquake Research and Engineering Laboratory, Technical Report No. 79-5-16, Dept. Civil Engrg., Northwestern University, Evanston, Ill., May 1979; to appear in *Int'l. J. Solids & Structures*.

15. Nemat-Nasser, S., "On Behavior of Granular Materials in Simple Shear," Earthquake Research and Engineering Laboratory, Technical Report No. 79-6-19, Dept. Civil Engrg., Northwestern University, Evanston, Ill., June 1979.

EFFECTIVE PARAMETERS AND FLUCTUATIONS FOR BOUNDARY VALUE PROBLEMS

George C. Papanicolaou*
Courant Institute of Mathematical Sciences, New York University

1. Introduction.

In modelling imperfections, inhomogeneities and in general the microstructure of material media, it is convenient to employ probabilistic concepts. For example one may take certain coefficients in constitutive relations to be random processes. The resulting field equations are then partial differential equations with random coefficients. Among the many questions that can be asked about such field equations the simplest and most common are perhaps the following:

(i) To what extent are the deterministic equations obtained by suitably averaging the random coefficients valid?

(ii) How can the fluctuations of the random solutions from the deterministic ones of (i) be calculated?

These questions can be answered efficiently and with a good deal of precision in the context of ordinary differential equations. The literature on the subject is extensive; for example [1, last chapter], [2,3] and the survey [4] contain some results on (i) and (ii) and more references. In particular [3] deals with boundary value problems which are considerably more difficult than initial value problems.

In the context of partial differential equations the status of (i) and (ii) is difficult to assess in general but one can say from a methodological

---

*Research supported by the Army Research Office under Grant No. DAAG29-78-G-0177.

and mathematical viewpoint that little has been done. The reason for this is that one must analyze spatial noise and its effects while most known methods deal with time dependent noise effects (Markovian and related methods which are essentially one dimensional).

Some progress has been made recently, however, and we shall report on part of it here.

2. Averaging for a nonlinear boundary value problem.

Consider a region $\mathcal{O} \subset R^d$ occupied by a conductor of unit conductivity. Let $u(x)$ denote the temperature at x and F the heat source density per unit volume so that

$$\Delta u + F = 0 \quad \text{in } \mathcal{O}, \tag{2.1}$$

with, say, $u = 0$ on $\partial \mathcal{O}$. We wish to consider the case in which F depends on the temperature u and is a random function of $x \in \mathcal{O}$. Since randomness is supposed to model microstructure, F must change rapidly as x varies over distances. Let $\epsilon > 0$ be a dimensionless parameter measuring the size of the microstructure (such as a ratio of typical macroscopic to microscopic lengths). Let $F(u,x,y,\omega)$ be a function on $R \times \mathcal{O} \times R^d \times \Omega$ where $(\Omega, \mathcal{F}, P)$ is a probability space with $\omega \in \Omega$ labelling the realizations of the medium. We assume that F is a stationary process for each u and x, i.e. that for any points $y_1, y_2, \ldots, y_n$ and any $h \in R^d$, $F(u,x,y_1,\omega)$, $\ldots, F(u,x,y_n,\omega)$ and $F(u,x,y_1+h,\omega), \ldots, F(u,x,y_n+h,\omega)$ have the same joint probability distribution. We shall take the heat source F in (2.1) to have the form $F(u(x), x, \frac{x}{\epsilon}, \omega)$. Then the temperature depends on ϵ and ω and

$$(2.2)\qquad A u^{\varepsilon}(x,\omega) + F(u^{\varepsilon}(x,\omega), x, \frac{x}{\varepsilon}, \omega) = 0 , \qquad x \in \mathcal{O} ,$$

$$u^{\varepsilon}(x,\omega) = 0 , \qquad x \in \partial\mathcal{O} .$$

The question now corresponding to (i) is this: does (2.2) have a solution (it is a nonlinear problem) and is this solution close to the averaged problem

$$(2.3)\qquad A\bar{u}(x) + \bar{F}(\bar{u}(x), x) = 0 , \qquad x \in \mathcal{O} ,$$

$$\bar{u}(x) = 0 , \qquad x \in \partial\mathcal{O} .$$

Here $\bar{F}(u,x)$ is the average of F, i.e.

$$(2.4)\qquad \bar{F}(u,x) = \int_{\Omega} F(u,x,y,\omega) P(d\omega) = E\{F(u,x,y,\cdot)\} ,$$

and it does not depend on y because of stationarity.

We shall assume now that (2.3) has a smooth solution $\bar{u}(x)$. We shall assume that the variational equation

$$Az(x) + V(x)z(x) = 0 , \qquad x \in \mathcal{O} ,$$

$$(2.5)\qquad z(x) = 0 , \qquad x \in \partial\mathcal{O} ,$$

where

$$(2.6)\qquad V(x) = \frac{\partial \bar{F}}{\partial u}(\bar{u}(x), x) = \bar{F}_u(\bar{u}(x), x) ,$$

has no solution other than $z(x) \equiv 0$. Then there is an $\varepsilon_0 > 0$ such that for each $0 < \varepsilon < \varepsilon_0$ there is a set $\Omega_\varepsilon \subset \Omega$ and for each $\omega \in \Omega_\varepsilon$, (2.2) has a solution with

$$(2.7) \qquad \int_{\Omega^\varepsilon} \int_{\mathcal{O}} |u^\varepsilon(x,\omega) - \bar{u}(x)|^2 dx\, P(d\omega) = h(\varepsilon) ,$$

while $h(\varepsilon) \to 0$ and $P(\Omega_\varepsilon) \to 1$ as $\varepsilon \to 0$. This result is proved in [5].

In the case of ordinary differential equations (d = 1) a very general result of the above form was proved by White and Franklin in [3]. Their method consists of using repeatedly the corresponding initial value problem result (shooting) which works very well. But it does not generalize to PDE. We shall outline the method of [5] below after some comments on problems related to (2.2).

In a medium in which heat sources are modelled as in (2.2), the conductivity may also be taken as a rapidly varying random function so that (2.2) becomes

$$(2.8) \qquad \nabla\cdot(a(\frac{x}{\varepsilon},\omega)\nabla u^\varepsilon(x,\omega)) + F(u^\varepsilon(x,\omega),x,\frac{x}{\varepsilon},\omega) = 0 , \qquad x \in \mathcal{O},$$

$$u^\varepsilon(x,\omega) = 0 , \qquad x \in \partial\mathcal{O}.$$

The linear version of this problem (which one may call stochastic homogenization) is analyzed by Kozlov and Jurinskii [6,7] and also in [8]. The result in this case is that there exist constants (q_{ij}) that can be computed by solving an auxiliary problem such that if $\bar{u}(x)$ is the solution of

$$(2.9) \qquad \sum_{i,j=1}^{d} q_{ij} \frac{\partial^2 \bar{u}(x)}{\partial x_i \partial x_j} + F(x) = 0 , \qquad x \in \partial\mathcal{O} ,$$

$$\bar{u}(x) = 0 , \qquad x \in \partial\mathcal{O} ,$$

Then again (2.7) holds. Here $F = F(x)$ but it could also be random but not dependent on u. The nonlinear version of (2.8) can also be analyzed with the limit being (2.9) with F replaced by $\bar{F}(\bar{u}(x),x)$. However in the nonlinear case (2.8), the validity of the approximation can be shown (at present) only under some rather severe restrictions and not in the natural generality of (2.2)-(2.7).

In both (2.9) and (2.3) the (q_{ij}) and $\bar{F}$ may be called the effective conductivity and effective heat source density, respectively. In both problems $\bar{F}$ is the average of F but (q_{ij}) is not the average of $(a\delta_{ij})$. This brings out an important point concerning question (i) of the introduction: the deterministic equations corresponding (in a sense like (2.7)) to some stochastic equations will not, in general, be simply the same equations with the random coefficients replaced by the averaged ones. To find the right limiting or averaged equations multiple scale techniques are quite useful (cf. [9] for how this is done for media with periodic structure).

As an example of a problem where the limit equations are not the original ones with averaged coefficients consider

$$\Delta u^{\varepsilon}(x,\omega) + \frac{1}{\varepsilon} F(u^{\varepsilon}(x,\omega),x,\frac{x}{\varepsilon},\omega) = 0 , \qquad x \in \mathcal{O},$$

$$u^{\varepsilon}(x,\omega) = 0 , \qquad x \in \partial\mathcal{O},$$

where now

$$E\{F(u,x,y,\cdot)\} \equiv 0 .$$

The limit problem has the form

$$\Delta\bar{u}(x) + G(\bar{u}(x),x) = 0 , \qquad x \in \mathcal{O} ,$$

$$\bar{u}(x) = 0 , \qquad x \in \partial\mathcal{O} ,$$

where $G(u,x)$ is given by

$$\frac{1}{\Gamma_d} \int_{R^d} \frac{1}{|z|^{d-2}} \, E\{F_u(u,x,z)F(u,x,0)\}dz$$

with Γ_d the surface area of the unit ball in R^d.

We go now to the proof of (2.2)-(2.7). For convenience we shall assume that $F(u,x,y,\omega) = \bar{F}(u,x)(1 + \phi(y,\omega))$ where $\phi(y,\omega)$ is a scalar-valued stationary and ergodic (this was not stated above) process on R^d. We assume that $\bar{F}(u,x)$ and all derivatives that appear below are continuous and bounded for all u and x. The process ϕ has mean zero and finite variance

$$(2.10) \qquad E\{\phi\} = 0 , \qquad E\{\phi^2\} < \infty .$$

Thus $E(F) = \bar{F}$ as in (2.4).

The random field $\phi(y,\omega)$ has the spectral decomposition

$$(2.11) \qquad \phi(y,\omega) = \int_{R^d} e^{ik\cdot y} \, \hat{\phi}(dk,\omega) ,$$

where the spectral measures $\hat{\phi}$ have orthogonal increments and

$$(2.12) \qquad E\{\hat{\phi}(dk)\hat{\phi}^*(dk)\} = p(dk) ,$$

where $p(dk)$ is the power spectral measure. In terms of the spectral decomposition (2.11) we can construct a solution $\chi(y,\omega)$ of

$$\Delta\chi(y,\omega) + \phi(y,\omega) = 0 \ , \tag{2.13}$$

in the form

$$\chi(y,\omega) = \int_{R^d} \frac{e^{ik\cdot y} - 1 - ik\cdot y}{|k|^2} \hat{\phi}(dk,\omega) \ . \tag{2.14}$$

One can verify that because $E\{\phi\} = 0$ and ϕ is ergodic

$$\lim_{|y|\to\infty} E\left\{\left(\frac{1}{|y|^2}\chi(y,\omega)\right)^2\right\} = 0 \tag{2.15}$$

and

$$\lim_{|y|\to\infty} E\left\{\left(\frac{1}{|y|}\nabla\chi(y,\omega)\right)^2\right\} = 0 \ . \tag{2.16}$$

We now make the change of variables

$$u^\varepsilon(x,\omega) = v^\varepsilon(x,\omega) + \varepsilon^2\chi(\frac{x}{\varepsilon},\omega)\bar{F}(v^\varepsilon(x,\omega),x) \ . \tag{2.17}$$

Using (2.2) we obtain for $v^\varepsilon(x,\omega)$ the following equation

$$\begin{aligned}[1 + \varepsilon^2\chi(\tfrac{x}{\varepsilon},\omega)\bar{F}_u(v^\varepsilon(x,\omega),x)]\Delta v^\varepsilon(x,\omega) + \bar{F}(v^\varepsilon(x,\omega),x) \\ + [\bar{F}(v^\varepsilon(x,\omega) + \varepsilon^2\chi(\tfrac{x}{\varepsilon},\omega)\bar{F}(v^\varepsilon(x,\omega),x),x) - \bar{F}(v^\varepsilon(x,\omega),x)](1 + \phi(\tfrac{x}{\varepsilon},\omega)) \\ + 2\varepsilon\nabla\chi(\tfrac{x}{\varepsilon},\omega)\cdot\nabla v^\varepsilon(x,\omega)\bar{F}_u(v^\varepsilon(x,\omega),x) \\ + \varepsilon^2\chi(\tfrac{x}{\varepsilon},\omega)(\nabla v^\varepsilon(x,\omega))^2\bar{F}_{uu}(v^\varepsilon(x,\omega),x) = 0 \ .\end{aligned} \tag{2.18}$$

Dividing by the coefficient of Δv^ε we may rewrite (2.18) in the form

$$\begin{aligned}&\Delta v^\varepsilon(x,\omega) + \bar{F}(v^\varepsilon(x,\omega),x) + G^\varepsilon(v^\varepsilon(x,\omega),\nabla v^\varepsilon(x,\omega),x,\omega) = 0 \ , \quad x \in \mathcal{O}, \\ &v^\varepsilon(x,\omega) = 0 \ , \qquad x \in \partial\mathcal{O} \ .\end{aligned} \tag{2.19}$$

To make the following precise one must return to (2.17) and introduce a suitable cutoff that keeps $\varepsilon^2\chi(\frac{x}{\varepsilon},\omega)$ small for all $\omega \in \Omega$ and $x \in \mathcal{O}'$. It is from this that a set of points Ω must be eliminated which, however, has small probability. For our purposes here, we may assume that $\chi(y,\omega)$ and $\nabla\chi(y,\omega)$ are bounded independently of y and ω (this is true in the periodic case only, in general, i.e. when $\phi(y,\omega)$ is periodic in y for each ω). It follows then that G^ε is small when ε is small independently of its arguments.

Once it is realized that G^ε is small, equation (2.19) can be solved by standard methods. We write

$$v^\varepsilon = \bar{u} + w^\varepsilon , \tag{2.20}$$

and note that we obtain

$$\begin{aligned} &(\Delta + V(x))w^\varepsilon(x,\omega) + G^\varepsilon(\bar{u} + w^\varepsilon, \nabla\bar{u} + \nabla w^\varepsilon, x, \omega) \\ &\quad + \bar{F}(\bar{u} + w^\varepsilon, x) - \bar{F}(\bar{u}, x) - \bar{F}_u(\bar{u}, x)w^\varepsilon = 0 , \qquad x \in \mathcal{O}, \\ &\quad w^\varepsilon(x,\omega) = 0 , \qquad x \in \partial\mathcal{O} . \end{aligned} \tag{2.21}$$

By hypothesis (2.5), $(\Lambda + V)^{-1}$ exists and, since V(x) is smooth, has the usual smoothing properties. Equation (2.21) is now solved by a contraction mapping argument in $H_0^1(\mathcal{O})$ (the Sobolev space over $\mathcal{O}$ with boundary values zero) since G^ε is controlled by taking ε small while the other term $\bar{F}(\bar{u} + w) - \bar{F}(\bar{u}) - \bar{F}_u(\bar{u})w$ is controlled by taking w small.

The above can be generalized to handle eigenvalue or bifurcation problems, i.e. situations where (2.5) has nontrivial solutions. This is done in [8].

3. Fluctuations for a nonlinear boundary value problem.

We consider next question (ii) of the introduction for problem (2.2)-(2.7). This means that we must find how $u^\varepsilon(x,\omega) - \bar{u}(x)$ behaves when suitably scaled as $\varepsilon \to 0$. This is analogous to the central limit theorem.

Let us define

$$\zeta^\varepsilon = \varepsilon^{-\gamma}(u^\varepsilon - \bar{u}) \ , \tag{3.1}$$

where $\gamma > 0$ is an index that will be specified later. By direct computation we find that ζ^ε satisfies the equation

$$\begin{aligned} -(\Delta + V(x))\zeta^\varepsilon(x,\omega) &= \varepsilon^{-\gamma}\bar{F}(\bar{u}(x),x)\phi(\tfrac{x}{\varepsilon},\omega) + \varepsilon^{-\gamma}(1 + \phi(\tfrac{x}{\varepsilon},\omega) \\ &\cdot [\bar{F}(\bar{u} + \varepsilon^\gamma\zeta^\varepsilon,x) - \bar{F}(\bar{u},x) - \bar{F}_u(\bar{u},x)\varepsilon^\gamma\zeta^\varepsilon] \\ &+ \varepsilon^{-\gamma}\phi(\tfrac{x}{\varepsilon},\omega)\bar{F}_u(\bar{u},x)(u^\varepsilon(x,\omega) - \bar{u}(x,\omega)) \ , \quad x \in \mathcal{O}, \end{aligned} \tag{3.2}$$

$$\zeta^\varepsilon(x,\omega) = 0 \ , \qquad x \in \partial\mathcal{O} \ .$$

Let $G(x,y)$ be the kernel (Green's function) of $(-(\Delta+V))^{-1}$. Then one can show (cf. [8]) that as $\varepsilon \to 0$, $\zeta^\varepsilon(x,\omega)$ as a process with values in $L^2(\mathcal{O})$ and the process

$$\eta^\varepsilon(x,\omega) = \varepsilon^{-\gamma}\int_{\mathcal{O}} G(x,y)\bar{F}(\bar{u}(y),y)\phi(\tfrac{y}{\varepsilon},\omega)dy \ , \tag{3.3}$$

have the same weak limit. This really means that the other terms on the right side of (3.2) are negligible compared to η^ε. This requires some improved estimates of the type obtained in the analysis outlined in the previous section. In any case, it is worth noting that the analysis of η^ε, which is a problem that has nothing to do with differential equations, is in fact a multidimensional central limit theorem.

The asymptotic limit of $\eta^\varepsilon(x,\omega)$ would be a routine matter to analyze if it were not for the fact that $G(x,y)$ is singular at $x = y$ for $d > 1$. Let us assume that the stationary process $\phi(x,\omega)$ is differentiable in mean square any number of times and that $\partial\mathcal{O}$ is smooth so that $G(x,y)$ is smooth for $x \neq y$ and behaves like $|x-y|^{-d+2}$ as $x \to y$. Then one can estimate $E\{(\eta^\varepsilon(x,\cdot))^2\}$ for ε small and conclude that in order that this quantity be of order one we must have

$$\gamma = \frac{d}{2}, \qquad \text{for } d < 4,$$
$$\gamma = 2, \qquad \text{for } d > 4. \tag{3.4}$$

For $d = 4$ then ζ^ε is defined by replacing $\varepsilon^{-\gamma}$ by $\varepsilon^{-2}(\log \frac{1}{\varepsilon})^{-1/2}$. Of course we have assumed here that the stationary process $\phi(x,\omega)$ is not merely ergodic but has also rapidly decaying correlations (it is mixing in a sufficiently strong sense). We recall that for the discussion of section 2 ergodicity was sufficient.

The fluctuation result is then this. Ignoring the exceptional ω set $\Omega - \Omega_\varepsilon$ since it has small probability, the solution $u^\varepsilon(x,\omega)$ of (2.2) behaves for ε small like $\bar{u}(x) + \varepsilon^{+\gamma}\zeta^\varepsilon(x,\omega)$, with γ as in (3.4). Moreover $\zeta^\varepsilon(x,\omega)$ converges weakly (as a process in $L^2(\mathcal{O})$) to a Gaussian random process with mean zero and covariance equal to the limiting covariance of η^ε, i.e.

$$\lim_{\varepsilon\downarrow 0} E\{\eta^\varepsilon(x)\eta^\varepsilon(y)\} = \lim_{\varepsilon\downarrow 0} \int_{\mathcal{O}}\int_{\mathcal{O}} G(x,z)G(y,z')\bar{F}(\bar{u}(z),z)\bar{F}(\bar{u}(z'),z') \cdot \varepsilon^{-2\gamma} R(\tfrac{z-z'}{\varepsilon})\,dz\,dz',$$

where $R(z-z') = E\{\phi(z)\phi(z')\}$. For $d < 4$ this limit is equal to

$$\rho \int_{\Omega} G(x,z)G(y,z)\left(\bar{F}(\bar{u}(z),z)\right)^2 dz \ , \qquad \rho = \int_{R^d} R(z)dz \ ,$$

which is then the limiting covariance of ζ^ε as $\varepsilon \to 0$.

References

[1] I. I. Gihman and A. V. Skorohod, Stochastic Differential Equations, Springer Verlag, New York, 1972.

[2] B. White, Some limit theorems for stochastic delay-differential equations, Comm. Pure Appl. Math., 29 (1976), pp. 131-141.

[3] B. White and J. Franklin, A limit theorem for stochastic two-point boundary value problems of ordinary differential equations, Comm. Pure Appl. Math., 32 (1979), pp. 253-276.

[4] G. Papanicolaou, Introduction to the asymptotic analysis of stochastic equations, in AMS book series in Applied Mathematics, Vol. 16, Providence, RI, 1977.

[5] G. Papanicolaou and S. R. S. Varadhan, to appear.

[6] V. V. Jurinskii, Vilnius Conference Abstracts, September 1978, p. 54.

[7] C. M. Kozlov, Doklady Akad. Nauk., 241 (#5), 1978.

[8] G. Papanicolaou and S. R. S. Varadhan, Boundary value problems with rapidly oscillating coefficients, to appear.

[9] A. Bensoussan, J. L. Lions, G. C. Papanicolaou, Asymptotic Analysis for Periodic Structures, North Holland, Amsterdam, 1978.

A THEORY OF INTERPENETRATING SOLID CONTINUA AND SOME APPLICATIONS

H.F. Tiersten
Department of Mechanical Engineering,
Aeronautical Engineering & Mechanics
Rensselaer Polytechnic Institute
Troy, New York 12181

1. INTRODUCTION

In the description of physical phenomena the model selected in a given situation depends crucially on the specific problem under consideration. As a consequence, different models of a given physical structure are employed in describing different aspects of its behavior. For example, an anisotropic crystal may be treated as a single continuum for many purposes but must be treated as a discrete lattice with a given structure and periodicity for certain other purposes, while for still other purposes the nucleus itself must be considered and so on. In many diverse physical situations, e.g., in the description of certain aspects of the behavior of ionic crystals and fiber reinforced composite materials, a description of solid matter consisting of interpenetrating solid continua may be fruitfully employed. In addition to describing the usual acoustic type wave motion, this model describes optical type wave motion at long wavelengths along with a number of other things. Although for the continuum approach to be valid a characteristic length such as a wavelength must be large compared with the spacing of the discrete elements, the model is particularly well-suited to the treatment of macroscopically inhomogeneous situations and bounded media. In addition, within the continuum framework the deformation may be arbitrarily large.

The interpenetrating solid continuum model, which is closely related to the model of fluid mixtures [1], has been employed in the description of a variety of physical phenomena such as, e.g., certain types of magneto-elastic interaction [2], electroelastic interaction [3] and the interaction of the electromagnetic field with deformable insulators [4]. In this latter case in order to consider ionic polarization resonances, the model consisted of two interpenetrating continua in which the motion of the center of mass of the two continua was finite but the relative motion of each of the continua with respect to the center of mass was infinitesimal. It was felt that in order for the description to be physically meaningful, the relative displacement of the two continua had to be infinitesimal, or else the solid would

rupture. Recently this model was employed in the description of material composites [5]. The idea of employing interpenetrating continua as a model of composite materials had been introduced earlier by Bedford and Stern [6,7]. However, there are a number of fundamental differences between the approach of Bedford and Stern and that employed in Reference [5].

In this paper the differential equations and boundary conditions describing the behavior of a finitely deformable, heat conducting solid are derived by means of a systematic application of the laws of continuum mechanics to a well-defined macroscopic model consisting of interpenetrating solid continua. Each continuum represents one identifiable constituent of the N-constituent solid. Each constituent interacts with neighboring elements across a surface of separation by means of a traction vector acting on that constituent. In addition, each constituent interacts with all other constituents at the same point by means of volumetrically interacting forces and couples, both of which are assumed to be equal and opposite in pairs. The influence of a simple type of viscous dissipation is included in the general treatment. Although the motion of the center of mass of the combined solid continuum may be arbitrarily large, the relative displacement of the individual constituents is required to be infinitesimal in order that the solid not rupture. The resulting system of nonlinear equations should provide a reasonable description of such materials as, say, fiber reinforced rubber.

After the general nonlinear description is obtained the resulting linear equations for the two-constituent continuum are exhibited in detail in the purely elastic case. The linear elastic constitutive equations for both the general anisotropic and isotropic cases are presented. The asymptotic results obtained from plane wave solutions at long wavelengths in the isotropic case are exhibited and discussed. A dynamic Lamé type of potential representation of the isotropic equations, which is complete, is exhibited. A simple problem of one-dimensional load transfer from the fiber reinforcement to the matrix of a composite material is considered within the framework of the description [8]. However, since the material constants occurring in the theory have never been measured for any composite material, a calculation cannot be performed. Nevertheless, if the model is reduced rather drastically but still plausibly for certain cases of interest in which the volume fraction of reinforcement is low [6,7], the remaining unknown constants can be partially determined from the known ordinary elastic constants of the two constituents and a calculation can be performed. Finally, some results

of surface wave propagation obtained from the reduced model along with an additional assumption are presented and discussed.

2. THE INTERACTING CONTINUA

As indicated in the Introduction, the macroscopic model we consider consists of N distinct interpenetrating solid continua. Initially, all continua occupy the same region of space and, hence, have the same material coordinates X_L. The motion of the center of mass of the combined continuum is described by the mapping

$$y_i = y_i(X_L, t) , \quad \underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{X}, t) , \tag{2.1}$$

which is one-to-one and differentiable as often as required. In (2.1) the y_i denote the spatial (or present) coordinates and X_L, the material (or reference) coordinates of the center of mass and t denotes the time. We consistently use the convention that capital indices denote the Cartesian components of $\underset{\sim}{X}$ and lower case indices, the Cartesian components of $\underset{\sim}{y}$. Thus, $\underset{\sim}{X}$ and $\underset{\sim}{y}$ denote the initial position of all material points and the present position of the center of mass of the combined continuum, respectively. Both dyadic and Cartesian tensor notation are used interchangeably. Since each continuum possesses a positive reference mass density $\rho_o^{(n)}$ and initially occupies the same region of space, we have

$$\rho_o = \sum_{n=1}^{N} \rho_o^{(n)} , \tag{2.2}$$

where ρ_o is the total reference mass density of the combined continuum.

In a (finite) motion each continuum is permitted to displace with respect to the center of mass of the combined continuum by infinitesimal displacements fields $\underset{\sim}{w}^{(n)}$. A schematic diagram indicating the motion of the model appears in Figure 1. The infinitesimal displacement fields $\underset{\sim}{w}^{(n)}$ are regarded as functions of $\underset{\sim}{y}$ and $\underset{\sim}{t}$. Since the $\underset{\sim}{w}^{(n)}$ are infinitesimal and

$$\underset{\sim}{y}^{(n)} = \underset{\sim}{y} + \underset{\sim}{w}^{(n)}(\underset{\sim}{y}, t) , \tag{2.3}$$

and the determinant of a matrix product is equal to the product of the determinants, we have

$$J^{(n)} = J(1 + \underset{\sim}{\nabla}_{\underset{\sim}{y}} \cdot \underset{\sim}{w}^{(n)}) \approx J , \tag{2.4}$$

where

$$J = \det y_{i,L} \,, \quad J^{(n)} = \det y^{(n)}_{i,L} \,. \tag{2.5}$$

Inasmuch as mass is conserved separately for each constituent, from (2.4) and (2.5) we have

$$\rho^{(n)} J = \rho_o^{(n)} \,, \tag{2.6}$$

which enables us to write

$$\rho = \sum_{n=1}^{N} \rho^{(n)} \,, \quad \rho J = \rho_o \,, \tag{2.7}$$

and ρ is the total present mass density of the combined continuum. Since $\underset{\sim}{y}$ has been defined as the center of mass of the combined continuum, we may write

$$\sum_{n=1}^{N} \int_{V^{(n)}} (\underset{\sim}{y} + \underset{\sim}{w}^{(n)}) \rho^{(n)} \, dV = \sum_{n=1}^{N} \int_{V^{(n)}} \underset{\sim}{y} \rho^{(n)} \, dV \,, \tag{2.8}$$

which, by virtue of (2.4) and (2.6), enables us to write

$$\sum_{n=1}^{N} \rho^{(n)} \underset{\sim}{w}^{(n)} = 0 \,, \quad \sum_{n=1}^{N} \rho^{(n)} \, d\underset{\sim}{w}^{(n)}/dt = 0 \,, \tag{2.9}$$

where d/dt denotes the material time derivative.

The interpenetrating continua interact with each other by means of defined local equal and opposite force fields $\underset{\sim}{L_F}^{nm} = - \underset{\sim}{L_F}^{mn}$, which are located at the position $\underset{\sim}{y}$, where the first superscript denotes the continuum being acted on and the second, the continuum producing the action, and defined equal and opposite local material couples $\underset{\sim}{L_C}^{nm} = - \underset{\sim}{L_C}^{mn}$. Each continuum interacts with neighboring elements across a surface of separation by means of a traction force per unit area $\underset{\sim}{t}^{(n)}$ acting on that constituent. Schematic diagrams illustrating the above-mentioned interactions in the model are shown in Figures 2, 3 and 4.

3. THE EQUATIONS OF BALANCE

From the discussion in Section 2, the rate equations of the conservation of mass for the different continua and the combined continuum are obvious. The equations of the conservation of linear momentum for each of the N continua are

$$\int_S \underset{\sim}{t}^{(n)}\, dS + \int_V \rho^{(n)} \underset{\sim}{f}^{(n)}\, dV + \sum_{m\neq n}^{N} \int_V \underset{\sim}{L}_F^{nm}\, dV = \frac{d}{dt} \int_V \rho^{(n)} \left[\underset{\sim}{v} + \frac{d\underset{\sim}{w}^{(n)}}{dt} \right] dV ,$$

$$n = 1, 2, \ldots, N , \tag{3.1}$$

where $\underset{\sim}{v} = d\underset{\sim}{y}/dt$. The equations of the conservation of angular momentum for each of the N continua are

$$\int_S (\underset{\sim}{y} + \underset{\sim}{w}^{(n)}) \times \underset{\sim}{t}^{(n)}\, dS + \int_V (\underset{\sim}{y} + \underset{\sim}{w}^{(n)}) \times \rho^{(n)} \underset{\sim}{f}^{(n)}\, dV + \sum_{m\neq n}^{N} \int_V \left[\underset{\sim}{L}_C^{nm} + \underset{\sim}{y} \times \underset{\sim}{L}_F^{nm} \right] dV =$$

$$\frac{d}{dt} \int_V (\underset{\sim}{y} + \underset{\sim}{w}^{(n)}) \times \rho^{(n)} \left[\underset{\sim}{v} + \frac{d\underset{\sim}{w}^{(n)}}{dt} \right] dV , \quad n = 1, 2, \ldots, N . \tag{3.2}$$

From (3.1) in the usual manner, we obtain

$$\underset{\sim}{t}^{(n)} = \underset{\sim}{n} \cdot \underset{\sim}{\tau}^{(n)} . \tag{3.3}$$

The substitution of (3.3) into (3.1) with the aid of the divergence theorem and (2.6) yields

$$\underset{\sim}{\nabla} \cdot \underset{\sim}{\tau}^{(n)} + \rho^{(n)} \underset{\sim}{f}^{(n)} - \rho^{(n)} \frac{d\underset{\sim}{v}}{dt} - \rho^{(n)} \frac{d^2\underset{\sim}{w}^{(n)}}{dt^2} + \sum_{m\neq n}^{N} \underset{\sim}{L}_F^{nm} = 0 , \tag{3.4}$$

which are the stress equations of motion of each of the N continua, and where $\underset{\sim}{\nabla} = \underset{\sim}{e}_i \partial/\partial y_i$ and $\underset{\sim}{e}_i$ is a unit base vector in the ith Cartesian direction. Substituting from (3.3) into (3.2) and employing the divergence theorem, (2.6) and (3.4), we obtain

$$\underset{\sim}{e}_\ell e_{\ell ij} \tau_{ij}^{(n)} + \underset{\sim}{e}_\ell e_{\ell kj} [w_k^{(n)} \tau_{ij}^{(n)}]_{,i} + \underset{\sim}{w}^{(n)} \times \rho^{(n)} \underset{\sim}{f}^{(n)} + \sum_{m\neq n}^{N} \underset{\sim}{L}_C^{nm}$$

$$- \underset{\sim}{w}^{(n)} \times \rho^{(n)} \frac{d\underset{\sim}{v}}{dt} - \underset{\sim}{w}^{(n)} \times \rho^{(n)} \frac{d^2\underset{\sim}{w}^{(n)}}{dt^2} = 0 , \tag{3.5}$$

which constitute the equations of the conservation of angular momentum of each of the N continua.

Adding the N equations in (3.4), we obtain

$$\underset{\sim}{\nabla} \cdot \underset{\sim}{\tau} + \rho \underset{\sim}{f} = \rho\, d\underset{\sim}{v}/dt , \tag{3.6}$$

which are the stress equations of motion of the combined continuum, and

$$\underset{\sim}{\tau} = \sum_{n=1}^{N} \underset{\sim}{\tau}^{(n)}, \quad \rho\underset{\sim}{f} = \sum_{n=1}^{N} \rho^{(n)}\underset{\sim}{f}^{(n)}, \tag{3.7}$$

where $\underset{\sim}{\tau}$ is the total mechanical stress tensor and $\underset{\sim}{f}$ is the total body force per unit mass. Now let us define the constants $r^{(n)}$ by

$$r^{(n)} = \rho_o^{(n)}/\rho_o^{(N)}, \tag{3.8}$$

and then the subtraction of $r^{(n)}$ times the Nth equation in (3.4) from the nth equation in (3.4) yields

$$\underset{\sim}{\nabla} \cdot \underset{\sim}{D}^{(n)} + \rho^{(n)}\underset{\sim}{f}^{(n)} + \underset{\sim}{\mathcal{F}}^{(n)} = \rho^{(n)} d^2\underset{\sim}{\eta}^{(n)}/dt^2, \tag{3.9}$$

where

$$D_{ij}^{(n)} = \tau_{ij}^{(n)} - r^{(n)}\tau_{ij}^{(N)}, \quad \tilde{f}_j^{(n)} = f_j^{(n)} - f_j^{(N)},$$

$$\mathcal{F}_j^{(n)} = \sum_{m\neq n}^{N} L_{F_j nm} - r^{(n)} \sum_{m\neq N}^{(N-1)} L_{F_j Nm},$$

$$\eta_j^{(n)} = w_j^{(n)} - w_j^{(N)} = w_j^{(n)} + \sum_{m=1}^{(N-1)} r^{(m)} w_j^{(m)}. \tag{3.10}$$

Equations (3.9) are called the difference or relative equations of motion, and $D_{ij}^{(n)}$ and $\underset{\sim}{\eta}^{(n)}$ are the difference stresses and difference displacements, respectively. Adding the N equations in (3.5) and obtaining the tensor form from the axial vector form, we obtain

$$\tau_{ij}^{A} = \frac{1}{2} \sum_{n=1}^{(N-1)} \left[D_{ki}^{(n)} w_{j,k}^{(n)} - D_{kj}^{(n)} w_{i,k}^{(n)} - \mathcal{F}_i^{(n)} w_j^{(n)} + \mathcal{F}_j^{(n)} w_i^{(n)} \right], \tag{3.11}$$

which is the equation of the conservation of angular momentum for the combined continuum. Equation (3.11) turns out to be of considerable value and interest when viscous type dissipation is considered. However, in the absence of viscous dissipation Eq. (3.11) is a direct consequence of the invariance of the stored energy function in a rigid rotation.

Although we cannot explicitly evaluate each of the defined couples of interaction $\underset{\sim}{L_C}^{mn}$ between the respective continua in the description presented here, we can

readily evaluate the total internal couple acting on each continuum, and that is all that is required in this type of description. Similar statements hold in the case of the defined forces of interaction $\overset{L}{\underset{\sim}{F}}^{mn}$ between the respective continua.

4. THERMODYNAMIC CONSIDERATIONS

The conservation of energy for the combined material continuum can be written in the form

$$\frac{d}{dt}\int_V (T+\rho\epsilon)dV = \sum_{n=1}^{N}\left[\int_S \underset{\sim}{t}^{(n)}\cdot\left(\underset{\sim}{v}+\frac{d\underset{\sim}{w}^{(n)}}{dt}\right)dS + \int_V \rho^{(n)}\underset{\sim}{f}^{(n)}\cdot\left(\underset{\sim}{v}+\frac{d\underset{\sim}{w}^{(n)}}{dt}\right)dV\right] - \int_S \underset{\sim}{n}\cdot\underset{\sim}{q}\,dS\,, \tag{4.1}$$

where T is the kinetic energy per unit volume, ϵ is the internal stored energy per unit mass, $\underset{\sim}{t}^{(n)}\cdot(\underset{\sim}{v}+d\underset{\sim}{w}^{(n)}/dt)$ are the rates of working per unit area of the mechanical surface tractions acting on each continuum, $\underset{\sim}{n}\cdot\underset{\sim}{q}$ is the rate of efflux of heat per unit area and $\rho^{(n)}\underset{\sim}{f}^{(n)}\cdot(\underset{\sim}{v}+d\underset{\sim}{w}^{(n)}/dt)$ are the rates of working per unit volume of body forces acting in each continuum. From the model of the continuum it is clear that T takes the form

$$T = \frac{1}{2}\sum_{n=1}^{N}\rho^{(n)}\left[\underset{\sim}{v}+\frac{d\underset{\sim}{w}^{(n)}}{dt}\right]\cdot\left[\underset{\sim}{v}+\frac{d\underset{\sim}{w}^{(n)}}{dt}\right]. \tag{4.2}$$

Expanding terms in (4.2), substituting from (3.3), (3.7) and (3.10), employing the divergence theorem, (2.7), (2.9), (3.6) and (3.9), we obtain

$$\rho\frac{d\epsilon}{dt} = \tau_{ij}v_{j,i} + \sum_{n=1}^{N-1}\left[D_{ij}^{(n)}\left(\frac{dw_j^{(n)}}{dt}\right)_{,i} - \mathcal{F}_j^{(n)}\frac{dw_j^{(n)}}{dt}\right] - q_{i,i}\,, \tag{4.3}$$

which is the first law of thermodynamics for the combined continuum.

We may now introduce a simple type of viscous dissipation by assuming that the symmetric part of the total stress tensor, the (N-1) difference stress tensors and difference internal forces may be written as a sum of a dissipative and a nondissipative part. Accordingly, we write [10]

$$\underset{\sim}{\tau} = \underset{\sim}{R}_{\tau} + \underset{\sim}{D}_{\tau}^{S} + \underset{\sim}{D}_{\tau}^{A}\,,\quad \underset{\sim}{D}^{(n)} = \underset{\sim}{R}_{D}^{(n)} + \underset{\sim}{D}_{D}^{(n)}\,,\quad \mathcal{F}^{(n)} = R_{\mathcal{F}}^{(n)} + D_{\mathcal{F}}^{(n)}\,, \tag{4.4}$$

and in each case the superscript R indicates the nondissipative (stored energy) portion and the superscript D, the dissipative portion. Substituting from (4.4) into (4.3) and employing the dissipative portion of (3.11), we obtain

$$\rho \frac{d\varepsilon}{dt} = {}^{R}\tau_{ij} v_{j,i} + \sum_{n=1}^{N-1} \left[{}^{R}D_{ij}^{(n)} \left(\frac{dw_j^{(n)}}{dt} \right)_{,i} - {}^{R}\mathcal{F}_j^{(n)} \frac{dw_j^{(n)}}{dt} \right] + {}^{D}\tau_{ij}^{S} d_{ij} +$$

$$+ \sum_{n=1}^{N-1} \left[{}^{D}D_{kj}^{(n)} \left[\left(\frac{dw_j^{(n)}}{dt} \right)_{,k} - w_{i,k}^{(n)} \omega_{ij} \right] - {}^{D}\mathcal{F}_j^{(n)} \left(\frac{dw_j^{(n)}}{dt} - w_i^{(n)} \omega_{ij} \right) \right] - q_{i,i} , \tag{4.5}$$

where the rate of deformation and spin tensors are defined by

$$d_{ij} = \frac{1}{2} (v_{j,i} + v_{i,j}) , \quad \omega_{ij} = \frac{1}{2} (v_{j,i} - v_{i,j}) . \tag{4.6}$$

For the circumstances under consideration, the mathematical expression of the second law of thermodynamics may be written in the form [11]

$$\rho \frac{d\varepsilon}{dt} - {}^{R}\tau_{ij} v_{j,i} - \sum_{n=1}^{N-1} \left[{}^{R}D_{ij}^{(n)} \left(\frac{dw_j^{(n)}}{dt} \right)_{,i} - {}^{R}\mathcal{F}_j^{(n)} \frac{dw_j^{(n)}}{dt} \right] = \rho\theta \frac{d\eta}{dt} , \tag{4.7}$$

where θ is the positive absolute temperature and η is the entropy per unit mass. From (4.5) and (4.7) we have the dissipation equation

$$ {}^{D}\tau_{ij}^{S} d_{ij} + \sum_{n=1}^{N-1} \left[{}^{D}D_{kj}^{(n)} \left[\left(\frac{dw_j^{(n)}}{dt} \right)_{,k} - w_{i,k}^{(n)} \omega_{ij} \right] - \right.$$

$$\left. {}^{D}\mathcal{F}_j^{(n)} \left(\frac{dw_j^{(n)}}{dt} - w_i^{(n)} \omega_{ij} \right) \right] - q_{i,i} = \rho\theta \frac{d\eta}{dt} , \tag{4.8}$$

and the entropy inequality may be written in the form

$$\rho \frac{d\eta}{dt} + \left(\frac{q_i}{\theta} \right)_{,i} = \frac{1}{\theta} \left[{}^{D}\tau_{ij}^{S} d_{ij} + \sum_{n=1}^{N-1} \left\{ {}^{D}D_{kj}^{(n)} \left[\left(\frac{dw_j^{(n)}}{dt} \right)_{,k} - w_{i,k}^{(n)} \omega_{ij} \right] - \right. \right.$$

$$\left. \left. - {}^{D}\mathcal{F}_j^{(n)} \left(\frac{dw_j^{(n)}}{dt} - w_i^{(n)} \omega_{ij} \right) \right\} - \frac{1}{\theta} q_i \theta_{,i} \right] = \rho\Gamma \geq 0 , \tag{4.9}$$

where Γ is the positive rate of entropy production. At this point it should be noted that this theory can readily be generalized [12,13] to account for a more general functional constitutive response in the manner set forth in a previous paper [3].

5. CONSTITUTIVE EQUATIONS

Since we are concerned with thermodynamic processes for which both the state function equation (4.7) and the dissipation equation (4.8) are valid, we may determine the dissipative constitutive equations from (4.9) and the nondissipative constitutive equations from (4.7). Since the entropy inequality is of the form shown in (4.9), it is convenient to define the thermodynamic function ψ by the Legendre transformation

$$\psi = \varepsilon - \eta\,\theta\,, \tag{5.1}$$

the substitution of which in (4.7) along with use of the chain rule of differentiation enables us to write

$$\rho\,\frac{d\psi}{dt} = {}^{R}\tau_{ij}\,X_{M,i}\,\frac{d}{dt}\,(y_{j,M}) + \sum_{n=1}^{N-1}\Big[{}^{R}D_{ij}^{(n)}\,X_{M,i}\,\frac{d}{dt}\,(w_{j,M}^{(n)}) - {}^{R}\mathcal{J}_{j}^{(n)}\,\frac{dw_{j}^{(n)}}{dt}\Big] - \rho\eta\,\frac{d\theta}{dt}\,. \tag{5.2}$$

Since (5.2) is a state function equation, we must have

$$\psi = \psi(y_{j,m};\ w_{j,M}^{(n)};\ w_{j}^{(n)};\ \theta)\,, \tag{5.3}$$

but since in order to satisfy the principle of material objectivity [14,15] ψ must be invariant in a rigid rotation of the deformed body, ψ may be shown [16] to be expressible as an arbitrary function of the arguments,

$$E_{KL} = \frac{1}{2}\,(y_{i,K}y_{i,L} - \delta_{KL})\,,\quad P_{LM}^{(n)} = y_{k,L}w_{k,M}^{(n)},\quad N_{L}^{(n)} = y_{k,L}w_{k}^{(n)}\,,\ \theta\,. \tag{5.4}$$

Hence ψ may be reduced to the form

$$\psi = \psi(E_{KL},\ P_{LM}^{(n)}\,,\ N_{L}^{(n)}\,,\ \theta)\,, \tag{5.5}$$

in place of the form shown in (5.3). From (5.2), (5.4) and (5.5), we obtain

$${}^{R}\tau_{ij} = \rho y_{i,L}y_{j,M}\,\frac{\partial\psi}{\partial E_{LM}} + \sum_{n=1}^{N-1}\Big[\rho y_{i,L}\,\frac{\partial\psi}{\partial N_{L}^{(n)}}\,w_{j}^{(n)} + \rho y_{i,L}\,\frac{\partial\psi}{\partial P_{LM}^{(n)}}\,w_{j,M}^{(n)}\Big]\,, \tag{5.6}$$

$${}^{R}D_{ij}^{(n)} = \rho y_{i,M}y_{j,L}\,\frac{\partial\psi}{\partial P_{LM}^{(n)}}\,,\quad {}^{R}\mathcal{J}_{j}^{(n)} = -\rho y_{j,L}\,\frac{\partial\psi}{\partial N_{L}^{(n)}}\,,\quad \eta = -\,\rho\,\frac{\partial\psi}{\partial\theta}\,, \tag{5.7}$$

where we have introduced the conventions $\partial\psi/\partial E_{LM} = \partial\psi/\partial E_{ML}$ and it is to be assumed that $\partial E_{KL}/\partial E_{LK} = 0$ in differentiating ψ. Substituting from $(5.7)_1$ and $(5.7)_2$ into (5.6) and employing the chain rule of differentiation, we obtain

$$R_{T_{ij}} = \rho y_{i,L} y_{j,M} \frac{\partial\psi}{\partial E_{LM}} - \sum_{n=1}^{(N-1)} \left[R_{\mathcal{F}_i}^{(n)} w_j^{(n)} + R_{D_{ki}}^{(n)} w_{j,k}^{(n)} \right], \tag{5.8}$$

the antisymmetric part of which is identical with the recoverable portion of τ_{ij}^{A} given in (3.11). Thus, even in this rather complex situation, the antisymmetric portion of the nondissipative part of the stress tensor is derivable from a thermodynamic state function and has just the value required by the conservation of angular momentum.

This brings us to a consideration of the dissipative constitutive equations, which are obtained from the entropy inequality (4.9) from which by using established methods [17,18] it has been shown [16] that the principle of material objectivity is satisfied if the dissipative portions of the constitutive equations take the form

$$D_{T_{ij}} S = y_{i,K} y_{j,L} T_{KL}, \quad q_i = y_{i,K} L_K,$$

$$D_{D_{ij}}^{(n)} = y_{i,K} y_{j,L} \Delta_{KL}^{(n)}, \quad D_{\mathcal{F}_j}^{(n)} = y_{j,K} \Phi_K^{(n)}, \tag{5.9}$$

where typically

$$T_{KL} = T_{KL}(R_{MN}, Z_{MN}^{(n)}, B_M^{(n)}, \theta_{,M}, E_{MN}, P_{MN}^{(n)}, N_M^{(n)}, \theta), \tag{5.10}$$

and similar relations may be written for L_K, $\Delta_{KL}^{(n)}$ and $\Phi_K^{(n)}$, and

$$R_{MN} = y_{i,M} y_{j,N} d_{ij} = \frac{dE_{MN}}{dt}, \quad Z_{MN}^{(n)} = y_{i,M} y_{j,N} \zeta_{ij}^{(n)}, \quad B_M^{(n)} = y_{i,M} \beta_i^{(n)}, \tag{5.11}$$

where

$$\zeta_{ij}^{(n)} = \left(\frac{dw_j}{dt}\right)_{,i} - w_{k,i}^{(n)} \omega_{kj}, \quad \beta_j^{(n)} = \frac{dw_j^{(n)}}{dt} - w_k^{(n)} \omega_{kj}, \tag{5.12}$$

which have been shown [16] to be objective tensors and vectors, respectively. Thus, all that remains in the determination of explicit constitutive equations is the selection of specific forms for ψ, T_{KL}, L_K, $\Delta_{KL}^{(n)}$ and $\Phi_K^{(n)}$.

6. PIOLA-KIRCHHOFF FORM OF THE EQUATIONS

Up to this point all the equations have been written in terms of present (or spatial) coordinates. Since the reference (or material) coordinates of material points are known while the present (or spatial) coordinates are not, it is advantageous to have the equations written in terms of the reference coordinates. When the equations of motion, (3.6) and (3.9), are written in terms of reference coordinates, they take the respective forms [19]

$$K_{Lj,L} + \rho_o f_j = \rho_o \, dv_j/dt \,, \tag{6.1}$$

$$\mathcal{B}^{(n)}_{Lj,L} + \rho_o^{(n)} \tilde{f}_j^{(n)} + J\mathcal{F}_j^{(n)} = \rho_o^{(n)} \frac{d^2 \eta_j^{(n)}}{dt^2} \,, \quad n = 1, 2, \ldots, N-1 \,, \tag{6.2}$$

where

$$K_{Lj} = JX_{L,i} T_{ij} \,, \quad \mathcal{B}^{(n)}_{Lj} = JX_{L,i} D^{(n)}_{ij} \,, \quad Q_L = JX_{L,i} q_i \,. \tag{6.3}$$

Equation (6.1) is the Piola-Kirchhoff form of the stress equations of motion and Eq. (6.2) is the reference form of the (N-1) relative stress equations of motion. Similarly, in reference coordinates the dissipation equation, (4.8), takes the form [19]

$$J T_{KL} \frac{dE_{KL}}{dt} + \sum_{m=1}^{N-1} \left(\Delta^{(m)}_{KL} Z^{(m)}_{KL} - \Phi^{(m)}_K B^{(m)}_K \right) - Q_{L,L} = \rho_o \theta \frac{d\eta}{dt} \,. \tag{6.4}$$

The associated constitutive equations required here, which respectively replace (5.6), $(5.7)_1$, $(5.7)_2$) and (5.9), are given by [19]

$$R_{K_{Lj}} = \rho_o y_{j,M} \frac{\partial \psi}{\partial E_{LM}} + \rho_o \sum_{m=1}^{N-1} \left[\frac{\partial \psi}{\partial N_L^{(m)}} w_j^{(m)} + \frac{\partial \psi}{\partial P_{LM}^{(m)}} w_{j,M}^{(m)} \right] \,, \tag{6.5}$$

$$R_{\mathcal{B}^{(n)}_{Lj}} = \rho_o y_{j,K} \frac{\partial \psi}{\partial P^{(n)}_{KL}} \,, \quad J \; R_{\mathcal{F}_j^{(n)}} = -\rho_o y_{j,L} \frac{\partial \psi}{\partial N_L^{(n)}} \,, \tag{6.6}$$

$$D_{K_{Lj}} = J y_{j,M} T_{LM}, \; J \; D_{\mathcal{F}_j^{(n)}} = J y_{j,K} \Phi^{(n)}_K \,, \quad D_{\mathcal{B}^{(n)}_{Lj}} = J y_{j,M} \Delta^{(n)}_{LM} \,, \quad Q_K = J L_K \,, \tag{6.7}$$

where

$$K_{Lj} = R_{K_{Lj}} + D_{K_{Lj}} \,, \quad \mathcal{B}^{(n)}_{Lj} = R_{\mathcal{B}^{(n)}_{Lj}} + D_{\mathcal{B}^{(n)}_{Lj}} \,, \quad \mathcal{F}_j^{(n)} = R_{\mathcal{F}_j^{(n)}} + D_{\mathcal{F}_j^{(n)}} \,, \tag{6.8}$$

and ψ and T_{LM}, $\phi_K^{(n)}$, $\Delta_{LM}^{(n)}$, L_K are given in (5.5) and (5.10), respectively. We also have the additional constitutive equation $(5.7)_3$) which remains unchanged. Thus, we now have a determinate system of differential equations, which by appropriate substitution can readily be reduced to $[4 + 3 \cdot (N-1)]$ equations in the $[4 + 3 \cdot (N-1)]$ dependent variables y_j, θ and $w_j^{(n)}$, $n = 1, 2, \ldots, (N-1)$. The equations are the three each of (6.1) and the (N-1) of (6.2) and (6.4). To this system of equations we must adjoin the associated boundary (or jump) conditions across moving not necessarily material surfaces of discontinuity. These jump conditions, which are obtained [5] by applying the integral forms of (6.1) and (6.2) and the reference forms of (4.1) and the integral form of (4.9) to a region encompassing the surface of discontinuity in the usual manner, take the respective forms

$$N_L [K_{Lj}] + U_N \rho_o [v_j] = 0 , \tag{6.9}$$

$$N_L [\mathcal{S}_{Lj}^{(n)}] + U_N \rho_o [d\eta_j^{(n)}/dt] = 0 , \quad n = 1, 2, \ldots, (N-1) , \tag{6.10}$$

$$N_L \left[K_{Lj} v_j + \sum_{n=1}^{N-1} \mathcal{S}_{Lj}^{(n)} \frac{dw_j^{(n)}}{dt} - Q_L \right] + U_N [T + \rho_o \varepsilon] = 0 , \tag{6.11}$$

$$N_L [Q_L/\theta] - U_N \rho_o [\eta] \geq 0 , \tag{6.12}$$

where U_N is the intrinsic velocity [20] of the singular surface, i.e., the velocity of the singular surface in the reference coordinate system. If the surface of discontinuity be material, U_N in (6.9) - (6.12) vanishes. If, furthermore, the body abuts another solid body and the full field equations have to be satisfied in each region, additional conditions on $[y]$ and the $[\eta^{(n)}]$ have to be satisfied, which usually are of the form

$$[y] = 0 , \quad [\eta^{(n)}] = 0 . \tag{6.13}$$

Moreover, if, as is usually the case $[\theta] = 0$, Γ in (4.9) is bounded and in place of (6.12) we have

$$N_L [Q_L] = 0 . \tag{6.14}$$

7. LINEAR EQUATIONS FOR THE TWO-CONSTITUENT CONTINUUM

It has been shown that when the equations for the two-constituent continuum are linearized the stress and relative equations of motion take the respective forms [21]

$$K_{LM,L} + \rho_o f_M = \rho_o \ddot{u}_M , \tag{7.1}$$

$$\mathcal{S}^{(1)}_{LM,L} + \frac{r^{(1)}}{1+r^{(1)}} \rho_o \tilde{f}^{(1)}_M + \mathcal{F}^{(1)}_M = r^{(1)} \rho_o \ddot{w}^{(1)}_M , \tag{7.2}$$

where u_M is the mechanical displacement of the center of mass of the two-constituent continuum and we have taken the liberty of using capital indices to denote the Cartesian components of the relative displacement vector $\underset{\sim}{w}^{(1)}$ in this linear description. In the purely elastic case the linear constitutive equations take the form [21]

$$\begin{aligned} K_{LM} &= c_{LMKN} u_{K,N} + \alpha_{KLM} w^{(1)}_K + \beta_{LMKN} w^{(1)}_{K,N} \\ \mathcal{S}^{(1)}_{LM} &= \beta_{KNML} u_{K,N} + \gamma_{KML} w^{(1)}_K + b_{MLKN} w^{(1)}_{K,N} \\ \mathcal{F}^{(1)}_M &= -\alpha_{MKL} u_{K,L} - a_{ML} w^{(1)}_L - \gamma_{MKL} w^{(1)}_{K,L} , \end{aligned} \tag{7.3}$$

where the c_{LMKN} are the usual elastic constants of ordinary linear elasticity, the a_{ML} may be called the difference displacement elastic constants, the b_{MLKN}, the relative elastic constants and α_{KLM}, β_{LMKN} and γ_{KML}, the respective elastic coupling constants. In the arbitrarily anisotropic case there are the usual 21 independent c_{LMKN}, 6 independent a_{ML}, 45 independent b_{MLKN}, 18 independent α_{KLM}, 54 independent β_{LMKN} and 27 independent γ_{KML}, for a total of 171 independent material constants. In the case of an isotropic two-constituent continuum the constitutive equations take the reduced form [22]

$$\begin{aligned} K_{LM} &= \lambda u_{K,K} \delta_{LM} + \mu (u_{L,M} + u_{M,L}) + \beta_1 w^{(1)}_{K,K} \delta_{LM} + \frac{1}{2} \beta_2 (w^{(1)}_{L,M} + w^{(1)}_{M,L}) , \\ \mathcal{S}^{(1)}_{LM} &= \beta_1 u_{K,K} \delta_{LM} + \frac{1}{2} \beta_2 (u_{L,M} + u_{M,L}) + b_1 w^{(1)}_{K,K} \delta_{LM} \\ &\quad + b_2 (w^{(1)}_{L,M} + w^{(1)}_{M,L}) + b_3 (w^{(1)}_{L,M} - w^{(1)}_{M,L}) , \quad \mathcal{F}^{(1)}_M = -a_1 w^{(1)}_M , \end{aligned} \tag{7.4}$$

and in order to secure the positive definiteness of the stored energy density the eight material constants in (7.4) must satisfy the six conditions [23]

$$\mu > 0 , \quad 3\lambda + 2\mu > 0 , \quad b_2 > 0 , \quad 3b_1 + 2b_2 > 0 , \quad b_3 < 0 , \quad a_1 > 0 . \tag{7.5}$$

When plane wave solutions are inserted in (7.1) and (7.2), with (7.4), we obtain the asymptotic expressions for the frequency ω - wavenumber ξ relations at small ξ (long wavelength), which take the form [24]

$$\omega_1^2 = \frac{1}{\rho_o}(\lambda + 2\mu)\xi^2 + O(\xi^4),$$

$$\omega_2^2 = \frac{a_1}{r^{(1)}\rho_o} + \frac{(b_1 + 2b_2)}{r^{(1)}\rho_o}\xi^2 + O(\xi^4), \tag{7.6}$$

for purely longitudinal plane waves and

$$\hat{\omega}_1^2 = \frac{\mu}{\rho_o}\xi^2 + O(\xi^4),$$

$$\hat{\omega}_2^2 = \frac{a_1}{r^{(1)}\rho_o} + \frac{(b_2 - b_3)}{r^{(1)}\rho_o}\xi^2 + O(\xi^4), \tag{7.7}$$

for purely transverse plane waves. Equations (7.6) and (7.7) reveal that ω_1, ω_2, $\hat{\omega}_1$ and $\hat{\omega}_2$ are real for real ξ. Moreover, it is clear that on a diagram of ω versus ξ there are four branches, two longitudinal and two transverse, with one of each emanating from $\omega = 0$, $\xi = 0$ and the other one of each emanating from $\omega = \sqrt{a_1/r^{(1)}\rho_o}$, $\xi = 0$. The longitudinal and transverse branches starting at $\omega = 0$ have positive initial slopes $\sqrt{(\lambda + 2\mu)/\rho_o}$ and $\sqrt{\mu/\rho_o}$, respectively, and the other two branches have zero initial slopes and positive curvatures $(b_1 + 2b_2)/\sqrt{r^{(1)}\rho_o a_1}$ and $(b_2 - b_3)/\sqrt{r^{(1)}\rho_o a_1}$. It should be noted that only seven combinations of the eight constants in (7.4) can be determined from plane wave measurements in the infinite medium. The evaluation of the eighth constant requires the measurement of a plane wave reflected from a surface at oblique incidence. This fact implies that some reflection measurements will be required for the determination of all the constants for any material symmetry from wave velocity measurements.

It has been shown that a dynamic Lamé type of potential representation of the isotropic equations may be written in the form [25]

$$\underset{\sim}{u} = \underset{\sim}{\nabla}\varphi + \underset{\sim}{\nabla} \times \underset{\sim}{H}, \quad \underset{\sim}{\nabla} \cdot \underset{\sim}{H} = 0,$$

$$\underset{\sim}{w} = \underset{\sim}{\nabla}\psi + \underset{\sim}{\nabla} \times \underset{\sim}{G}, \quad \underset{\sim}{\nabla} \cdot \underset{\sim}{G} = 0, \tag{7.8}$$

where the potentials satisfy the differential equations

$$(\lambda + 2\mu)\nabla^2\varphi + (\beta_1 + \beta_2)\nabla^2\psi = \rho_o\ddot{\varphi},$$

$$\mu\nabla^2\underset{\sim}{H} + \frac{1}{2}\beta_2\nabla^2\underset{\sim}{G} = \rho_o\underset{\sim}{\ddot{H}},$$

$$(\beta_1 + \beta_2)\nabla^2\varphi + (b_1 + 2b_2)\nabla^2\psi - a_1\psi = \rho_o\ddot{\psi},$$

$$\frac{1}{2}\beta_2\nabla^2\underset{\sim}{H} + (b_2 - b_3)\nabla^2\underset{\sim}{G} - a_1\underset{\sim}{G} = \rho_o\ddot{\underset{\sim}{G}}, \tag{7.9}$$

and the representation is complete. Equations (7.9) reveal that the scalar potentials φ and ψ are coupled in the representation separately from the vector potentials $\underset{\sim}{H}$ and $\underset{\sim}{G}$, which also are coupled.

8. LOAD TRANSFER IN FIBER REINFORCED MATERIALS

In this section we consider fiber reinforcement entering a matrix and terminating uniformly at a distance ℓ into the matrix which continues down to a rigid support at a distance b below the junction, as shown in Figure 5. A total tensile force P is applied to all the fibers crossing a given cross-sectional area. Since the load is applied in the preferred direction of transverse isotropy, which lies along the length of the parallel fibers, and it is assumed that all displacement and relative displacement components transverse to this direction are constrained to vanish and that the remaining displacement variables are independent of the transverse coordinates, the nontrivial linear constitutive equations take the form [8,22]

$$K_{11} = K_{22} = \hat{c}_2 u_{3,3} + \hat{\beta}_3 w^{(1)}_{3,3}, \tag{8.1}$$

$$K_{33} = \hat{c}_5 u_{3,3} + \hat{\beta}_6 w^{(1)}_{3,3}, \tag{8.2}$$

$$\mathcal{S}_{11} = \mathcal{S}_{22} = \hat{\beta}_5 u_{3,3} + \hat{b}_3 w^{(1)}_{3,3}, \tag{8.3}$$

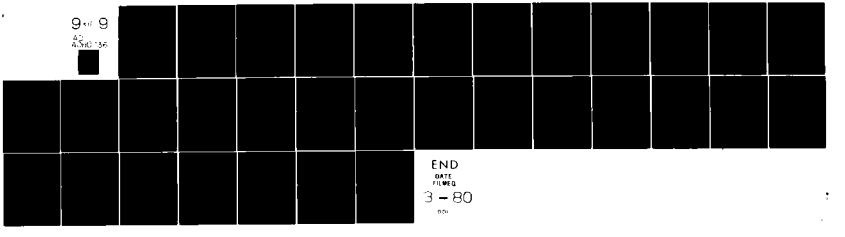
$$\mathcal{S}_{33} = \hat{\beta}_6 u_{3,3} + \hat{b}_5 w^{(1)}_{3,3}, \tag{8.4}$$

$$\mathcal{T}_3 = -\hat{a}_2 w^{(1)}_3, \tag{8.5}$$

where X_3 is the preferred direction of transverse isotropy, u_3 is the nonzero displacement component of the center of mass of the combined two-constituent composite material and $w^{(1)}_3$ is the nonzero component of the relative displacement of the continuum representing the matrix. As noted earlier the K_{LM} represent the components of the stress tensor for the combined continuum and the $\mathcal{S}_{LM}$ represent the relative stress tensor which is defined by

$$\mathcal{S}_{LM} = \tau^{(1)}_{LM} - r\tau^{(2)}_{LM}, \quad r = \rho^{(1)}/\rho^{(2)}, \quad \rho = \rho^{(1)} + \rho^{(2)}, \tag{8.6}$$

AD-A080 736 ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC F/G 20/4
TRANSACTIONS OF THE CONFERENCE OF ARMY MATHEMATICIANS (25TH).(U)
JAN 80
UNCLASSIFIED ARO-80-1 NL



END
DATE FILMED
3-80
DDC

where $\tau_{LM}^{(m)}$ and $\rho^{(m)}$ represent the components of the stress tensor and mass density, respectively, of each of the interpenetrating continua. In accordance with the definitions in Sec.3 the vector field $\tilde{\mathcal{F}}_M$ is related to the volumetric force of interaction between the two constituents by the relation

$$\tilde{\mathcal{F}}_M = {}^{L}F_M^{12}(1+r)\,, \tag{8.7}$$

where ${}^{L}F_M^{12}$ is the volumetric force exerted by continuum 2 on continuum 1. At this point it is to be noted that the $\tau_{LM}^{(m)}$ and $\rho^{(m)}$ do not represent the actual components of stress and mass density of each of the constituents in the composite, but only represent those quantities in each of the interpenetrating continua, which occupy the same region of space and, respectively, represent each constituent in the model. As a consequence, if A^m and A^f represent the areas occupied by the matrix and fibers, respectively, in a typical area A of the interpenetrating continua normal to the fiber length, we have

$$A = A^m + A^f\,,\quad \rho^{(1)} = \rho^m A^m/A\,,\quad \rho^{(2)} = \rho^f A^f/A\,,$$
$$\tau_{ij}^m = \tau_{ij}^{(1)} A/A^m\,,\quad \tau_{ij}^f = \tau_{ij}^{(2)} A/A^f\,, \tag{8.8}$$

where the variables with the superscripts m and f represent the actual respective quantities in the matrix and fiber reinforcement, respectively. The remaining nontrivial stress equations of equilibrium and relative stress equations of equilibrium are [8,22]

$$K_{33,3} + \rho f_3 = 0\,, \tag{8.9}$$

$$\mathcal{S}_{33,3} + \tilde{\mathcal{F}}_3 + \rho^{(1)}\tilde{f}_3 = 0\,, \tag{8.10}$$

where

$$\rho f_3 = \rho^{(1)} f_3^{(1)} + \rho^{(2)} f_3^{(2)}\,,\quad \tilde{f}_3 = f_3^{(1)} - f_3^{(2)}\,, \tag{8.11}$$

and $f_3^{(1)}$ and $f_3^{(2)}$ denote the components of body force per unit mass in the continua representing the matrix and fiber reinforcement, respectively, which in the case of the gravity force are the same as the body force intensities f_3^m in the matrix and f_3^f in the fiber reinforcement, both of which equal g.

The substitution of (8.2), (8.4), (8.5) and (8.11) into (8.9) and (8.10) yields

$$\hat{c}_5 u_{3,33} + \hat{\beta}_6 w_{3,33}^{(1)} + \rho g = 0\,, \tag{8.12}$$

$$\hat{\beta}_6 u_{3,33} + \hat{b}_5 w_{3,33}^{(1)} - \hat{a}_2 w_3^{(1)} = 0\,, \tag{8.13}$$

which are the one-dimensional displacement equations of equilibrium that apply to the one-dimensional static problem under consideration. In the absence of g the solution to (8.12) and (8.13) may be written in the form

$$u_3 = -\frac{\hat{\beta}_6}{\hat{c}_5}\left(Ae^{-\alpha X_3} + De^{\alpha X_3}\right) + BX_3 + C,$$

$$w_3^{(1)} = Ae^{-\alpha X_3} + De^{\alpha X_3}, \tag{8.14}$$

where

$$\alpha^2 = \hat{c}_5\hat{a}_2/(\hat{c}_5\hat{b}_5 - \hat{\beta}_6^2), \tag{8.15}$$

and A, B, C and D are arbitrary constants, to be found by satisfying the boundary conditions in this one-dimensional problem. Since the continuum representing the matrix and the continuum representing the fibers can neither separate from nor penetrate into the single matrix continuum that abuts the composite at the junction, we must have

$$w_3^{(1)} = 0, \quad w_3^{(2)} = 0, \quad \text{at } X_3 = 0, \tag{8.16}$$

which is consistent with $(2.9)_1$. In addition, the displacement u_3 of the center of mass of the combined composite continuum must be the same as the displacement U_3 of the isotropic single matrix continuum at the junction. Consequently, as kinematic boundary conditions at the junction we have

$$w_3^{(1)} = 0, \quad u_3 = U_3, \quad \text{at } X_3 = 0, \tag{8.17}$$

where in the absence of g U_3 satisfies

$$(\lambda^m + 2\mu^m)U_{3,33} = 0, \tag{8.18}$$

and the nontrivial stress components in the single matrix continuum are given by

$$T_{33}^m = (\lambda^m + 2\mu^m)U_{3,3}, \quad T_{11}^m = T_{22}^m = \lambda^m U_{3,3}. \tag{8.19}$$

In addition to the continuity of displacement at $x_3 = 0$ we have the continuity of traction, i.e.,

$$K_{33} = T_{33}^m, \quad \text{at } X_3 = 0. \tag{8.20}$$

Since no force is applied to the matrix at $X_3 = \ell$, we have

$$\tau_{33}^{(1)} = 0, \quad \tau_{33}^{(2)} = p_o, \quad \text{at } X_3 = \ell, \tag{8.21}$$

which with (8.6) enables us to write the boundary conditions

$$K_{33} = p_o , \quad \mathcal{E}_{33} = - rp_o , \quad \text{at } X_3 = \ell . \tag{8.22}$$

Since the supporting surface is rigid, we have

$$U_3 = 0 , \quad \text{at } X_3 = - b . \tag{8.23}$$

Thus, the boundary conditions are (8.17), (8.20), (8.22) and (8.23). The solution to (8.18) takes the form

$$U_3 = EX_3 + F . \tag{8.24}$$

Now, the substitution of (8.14) and (8.24) into (8.17), (8.20), (8.22) and (8.23) yields

$$A = - \frac{\nu p_o e^{-\alpha \ell}}{1 + e^{-2\alpha \ell}} , \quad D = \frac{\nu p_o e^{\alpha \ell}}{1 + e^{2\alpha \ell}} ,$$

$$B = \frac{1}{\hat{c}_5} p_o , \quad C = \frac{p_o b}{\lambda^m + 2\mu m} ,$$

$$E = \frac{p_o}{\lambda^m + 2\mu^m} , \quad F = \frac{p_o b}{\lambda^m + 2\mu^m} , \tag{8.23}$$

which when substituted in (8.14) and (8.24), respectively, yields the solution.

Since the material coefficients occurring in the theory have never been measured for any composite material, a calculation based on the foregoing analysis cannot be performed. However, if the model is reduced sufficiently by making certain simplifying assumptions, the material constants of the composite can be estimated from the known constants of the individual constituents of the composite while still retaining the essential characteristics of the composite for certain cases of interest. The simplified model we consider is that of a fiber reinforced composite material consisting of an elastic matrix containing uniformly distributed continuous fibers extending in the X_3-direction, in which the fibers occupy a small fraction of the total composite volume. On account of the latter condition in the reduced model it is assumed that the stresses in the matrix are related to the strains in the matrix by the constitutive relations of linear isotropic elasticity and are independent of the strains in the fibers. Similarly, the stresses in the fibers are assumed to be independent of the strains in the matrix. It is further

assumed that all stress components in the long narrow fibers vanish save the axial stress $\tau_{33}^{(2)}$, which may then be written as a function of the axial strain in the fibers only. Although this latter assumption seems questionable to us in general we make it anyway. Then the only interaction between the matrix and the fibers remaining is the volumetric interaction term ${}^{L}F_{M}^{12}$, which from (8.5) and (8.7) takes the form

$$ {}^{L}F_{P}^{12} = -\,(1+r)^{-1} a_1 w_P^{(1)}\,, \quad {}^{L}F_{3}^{12} = -\,(1+r)^{-1} a_2 w_3^{(1)}\,, \quad P = 1,2\,. \tag{8.24} $$

At this point it should be noted that the aforementioned assumptions make the reduced model for the linear case identical with that of Martin, Bedford and Stern [9]. When the aforementioned simplifications are made we find that the decay factor α given in (8.15) takes the reduced form [8]

$$ \alpha^2 = \tilde{a}_2 \, \frac{\lambda^m + 2\mu^m + E^{(2)}}{E^{(2)}(\lambda^m + 2\mu^m)}\,, \tag{8.25} $$

where

$$ E^{(2)} = E^f A^f / A\,, \tag{8.26} $$

and $\tilde{a}_2$ is a complicated function [9,8] of the material constants of the matrix and the fibers and the geometry. In the simplified model the actual stresses τ_{33}^f in the fibers and τ_{33}^m in the matrix at the junction at $X_3 = 0$ take the form [8]

$$ \tau_{33}^f = \frac{p_o^f}{\lambda^m + 2\mu^m + E^{(2)}} \left[E^{(2)} + \frac{\lambda^m + 2\mu^m}{\cosh \alpha \ell} \right], $$

$$ \tau_{33}^m = \frac{(p_o^f A^f / A)(\lambda^m + 2\mu^m)}{\lambda^m + 2\mu^m + E^{(2)}} \left[1 - \frac{1}{\cosh \alpha \ell} \right], \tag{8.27} $$

where

$$ p_o^f = p_o A / A^f\,, \tag{8.28} $$

and p_o^f denotes the actual stress in the fibers before they enter the matrix.

9. SURFACE WAVE PROPAGATION

Within the framework of the simplified model results have been obtained [8] for the dispersion of surface waves in a particular fiber reinforced composite material. Even in the simplified model the value of a_1 had to be assumed and was

arbitrarily taken to be equal to a_2 in order to perform the calculations. The calculations were performed for a set of material parameters corresponding to a glass fiber reinforced phenolic resin [9], the relevant constants of which are

$$\begin{aligned} \rho^m &= 0.00013 \text{ lb-sec}^2/\text{in}^4, & E^f &= 12.4\times10^6 \text{ lb/in}^2, \\ \rho^f &= 0.00026 \text{ lb-sec}^2/\text{in}^4, & \mu^f &= 10.2\times10^6 \text{ lb/in}^2, \\ \lambda^m &= 0.86\times10^6 \text{ lb/in}^2, & \mu^m &= 0.37\times10^6 \text{ lb/in}^2, \end{aligned} \tag{9.1}$$

for a fiber diameter of .01 in. for the volume percentage of reinforcement of 5.67%, which corresponds to s = .04 inches. The results of the calculations for surface waves propagating in the direction of the fiber reinforcement as shown in Figure 6 are plotted in Figure 7, which indicates the existence of an upper (optical type) surface wave branch in addition to the lower (acoustic type) branch. We do not believe that the upper surface wave branch actually exists, but that its existence is a consequence of the reduced coupling in the simplified model [8]. In Figure 7 we have drawn a vertical dotted line which corresponds to a wavelength five times the spacing of the fiber reinforcement. We do not believe the curves to be valid much beyond this vertical line because of the nature of the model of the composite we have employed, and we draw them considerably beyond their range of validity simply to indicate the calculated behavior. The important curve in Figure 7 is the lower acoustic type branch, which is drawn to a larger scale in Figure 8 along with the corresponding acoustic type branch for propagation normal [8] to the direction of the fiber reinforcement. Note the difference in dispersion for the two directions of propagation considered. This very precise dispersion property of surface waves could well be used as a means of nondestructively evaluating the distribution of the fiber reinforcement in and the integrity of the bonding to the matrix.

ACKNOWLEDGEMENTS

This work was supported in part by the Office of Naval Research under Contract No. N00014-76-C-0368 and the National Science Foundation under Grant No. ENG 72-04223.

REFERENCES

1. C. Truesdell, "Mechanical Basis of Diffusion," J. Chem. Phys., 37, 2336 (1962).

2. H.F. Tiersten, "Coupled Magnetomechanical Equations for Magnetically Saturated Insulators," J. Math. Phys., 5, 1298 (1964).

3. H.F. Tiersten, "On the Nonlinear Equations of Thermoelectroelasticity," Int. J. Eng. Sci., 9, 585 (1971).

4. H.F. Tiersten and C.F. Tsai, "On the Interaction of the Electromagnetic Field with Heat Conducting Deformable Insulators," J. Math. Phys., 13, 361 (1972).

5. H.F. Tiersten and M. Jahanmir, "A Theory of Composites Modeled as Interpenetrating Solid Continua," Arch. Rational Mech. Anal., 65, 153 (1977).

6. A. Bedford and M. Stern, "A Multi-Continuum Theory for Composite Elastic Materials," Acta Mechanica, 14, 85 (1972).

7. A Bedford and M. Stern, "Toward a Diffusing Continuum Theory of Composite Elastic Materials," J. Appl. Mech., 38, 8 (1971).

8. M. Jahanmir and H.F. Tiersten, "Load Transfer and Surface Wave Propagation in Fiber Reinforced Composite Materials," Int. J. Solids Structures, 14, 227 (1978).

9. S.E. Martin, A. Bedford and M. Stern, "Steady-State Wave Propagation in Fiber Reinforced Elastic Materials," in Developments in Mechanics, Proceedings of the Twelfth Midwestern Mechanics Conference, (Univ. of Notre Dame Press, Notre Dame, Ind., 1971), Vol.6, pp.515-528.

10. We are not interested in treating a general type of viscoelastic dissipation and, as a consequence, we consider the simplest (Kelvin) type of viscoelastic dissipation in solids.

11. B.A. Boley and J.H. Weiner, Theory of Thermal Stresses (Wiley, New York, 1960).

12. B.D. Coleman and W. Noll, "The Thermodynamics of Elastic Materials with Heat Conduction and Viscosity," Arch. Ratl. Mech. Anal., 13, 167 (1963).

13. B.D. Coleman, "Thermodynamics of Materials with Memory," Arch. Ratl. Mech. Anal., 17, 1 (1964).

14. C. Truesdell and W. Noll, "The Nonlinear Field Theories of Mechanics," in Encyclopedia of Physics, edited by S. Flügge (Springer-Verlag, Berlin, 1965), Vol.III/2, Secs.17 and 19.

15. A.C. Eringen, Nonlinear Theory of Continuous Media (McGraw-Hill, New York, 1962), Secs.27 and 44.

16. For a more detailed discussion see Sec.5 of Ref.5.

17. A.E. Green and J.E. Adkins, Large Elastic Deformations and Nonlinear Continuum Mechanics (Oxford U.P., London, 1960), Chap.8.

18. A.J.M. Spencer, "Theory of Invariants," in Continuum Physics, edited by A.C. Eringen (Academic, New York, 1972), Vol.1, Part III, Sec.2.1.

19. For more detail see Sec.8 of Ref.5.

20. The relation between U_N and the actual velocity u_n of the surface of discontinuity is given in P.J. Chen, "Growth and Decay of Waves in Solids," in Encyclopedia of Physics, edited by C. Truesdell (Springer-Verlag, Berlin, 1973), Vol.IVa/3, Sec.4.

21. See Ref.5, Sec.9.

22. See Ref.5, Sec.10.

23. See Ref.5, Eqs.(11.3).

24. See Ref.5, Sec.11.

25. See Ref.5, Sec.12.

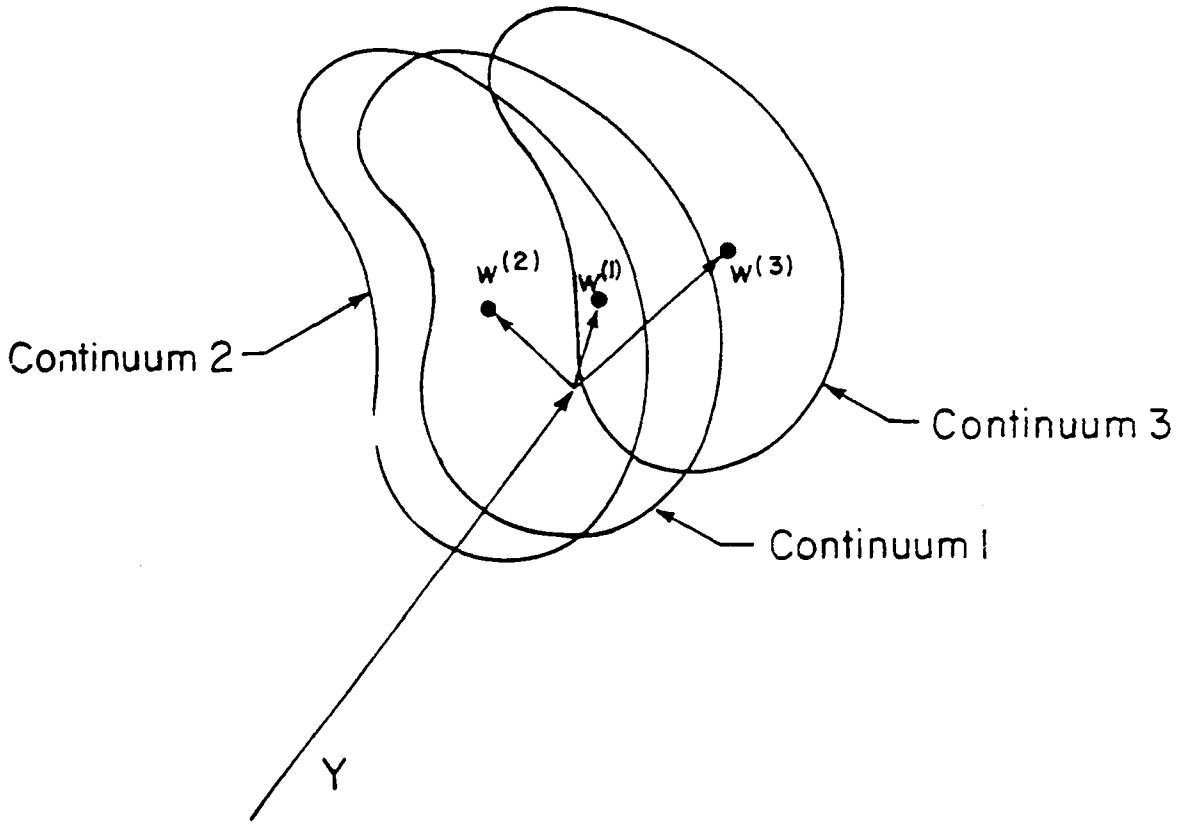


Figure 1

Schematic Diagram Showing the Relative Displacements of the Interacting Continua

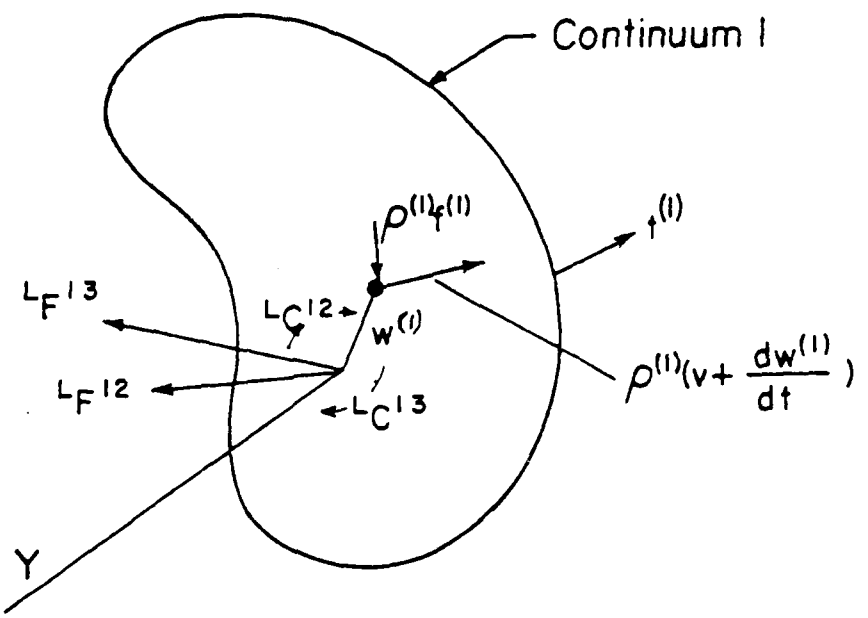


Figure 2

Schematic Diagram Showing the Linear Momentum and Force and Couple Vectors Acting in Continuum 1

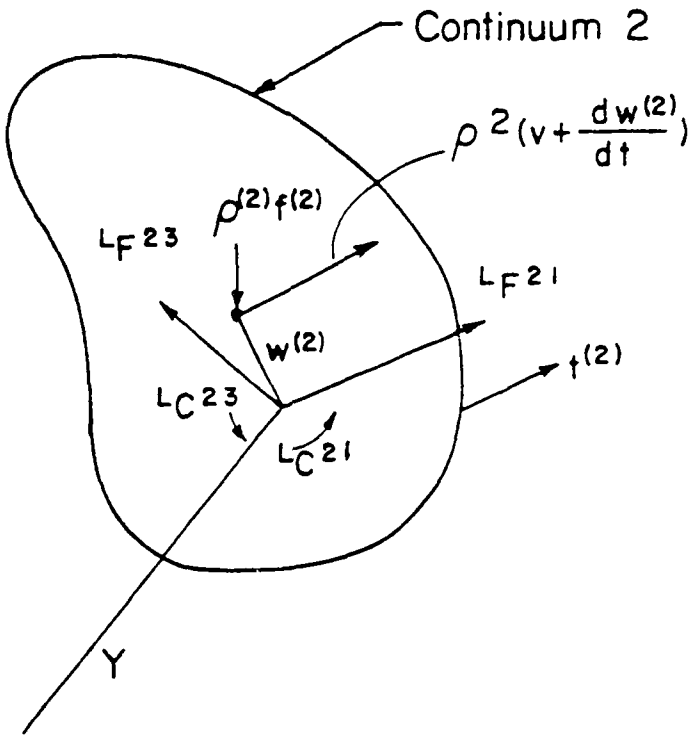


Figure 3

Schematic Diagram Showing the Linear Momentum and Force and Couple Vectors Acting in Continuum 2

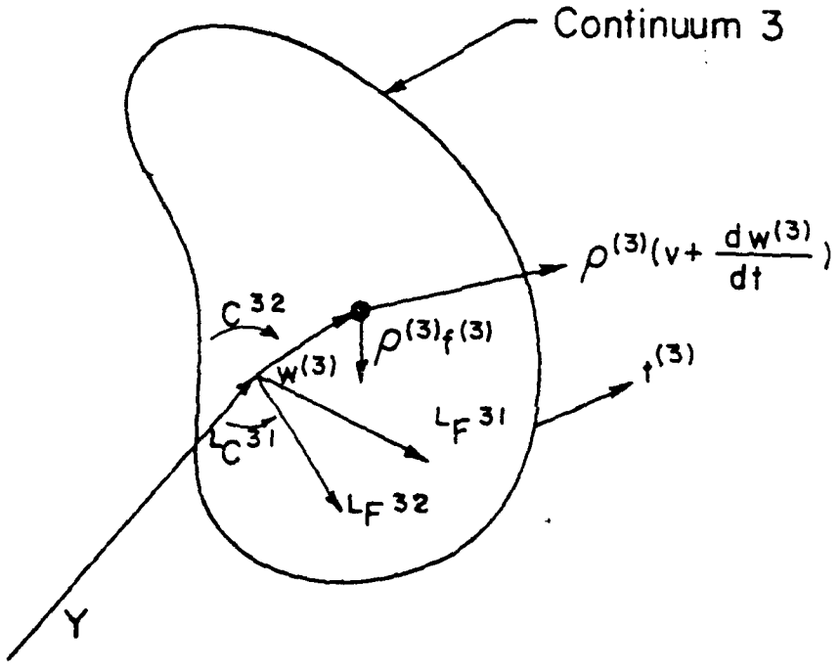


Figure 4

Schematic Diagram Showing the Linear Momentum and Force and Couple Vectors Acting in Continuum 3

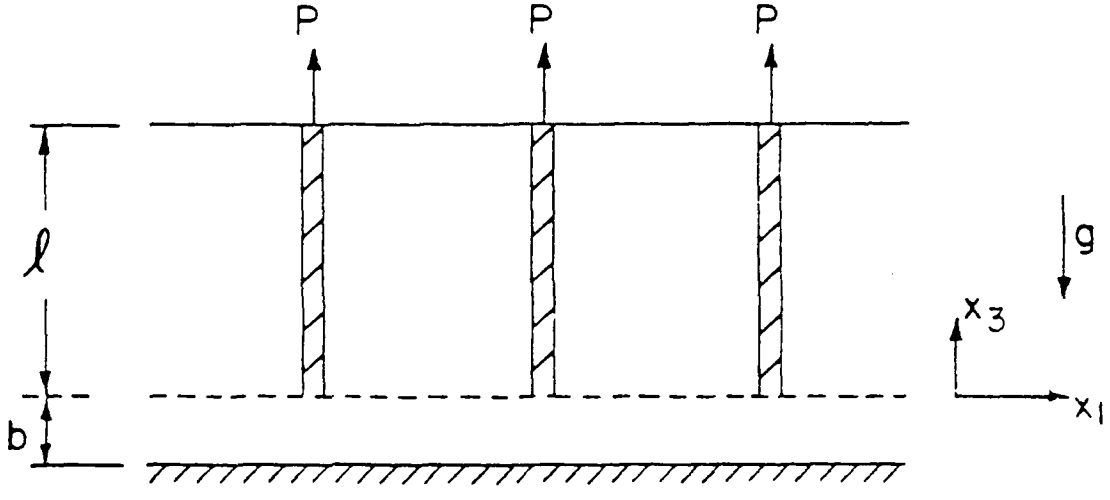


Figure 5

Schematic Diagram of Loaded Fiber Reinforced Composite with Reinforcement Terminating Uniformly in Matrix Some Distance Before Support

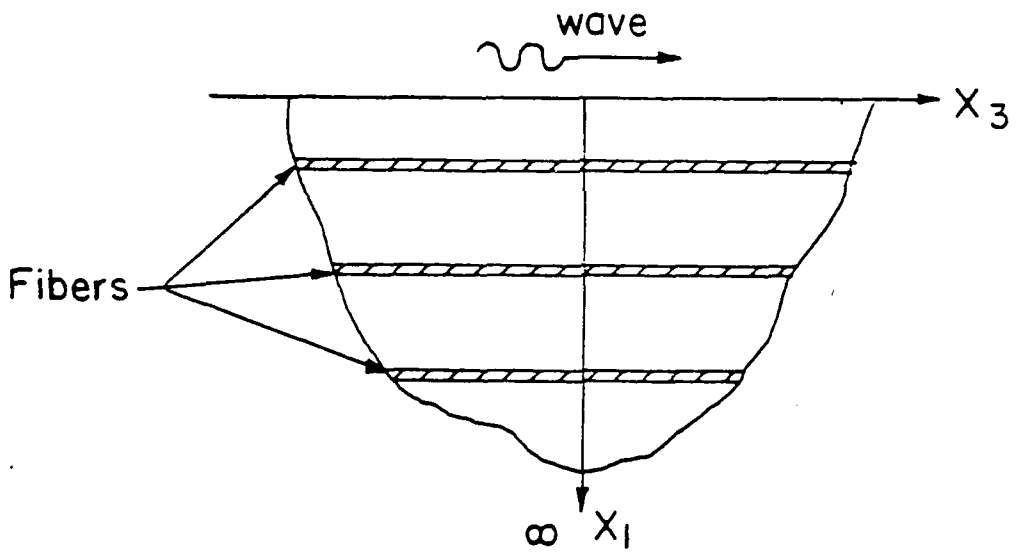


Figure 6

Schematic Diagram Showing Surface Wave Propagating Along a Free Surface of a Fiber Reinforced Composite

Dispersion Curves for Surface Waves Propagating in the Direction of the Fiber Reinforcement

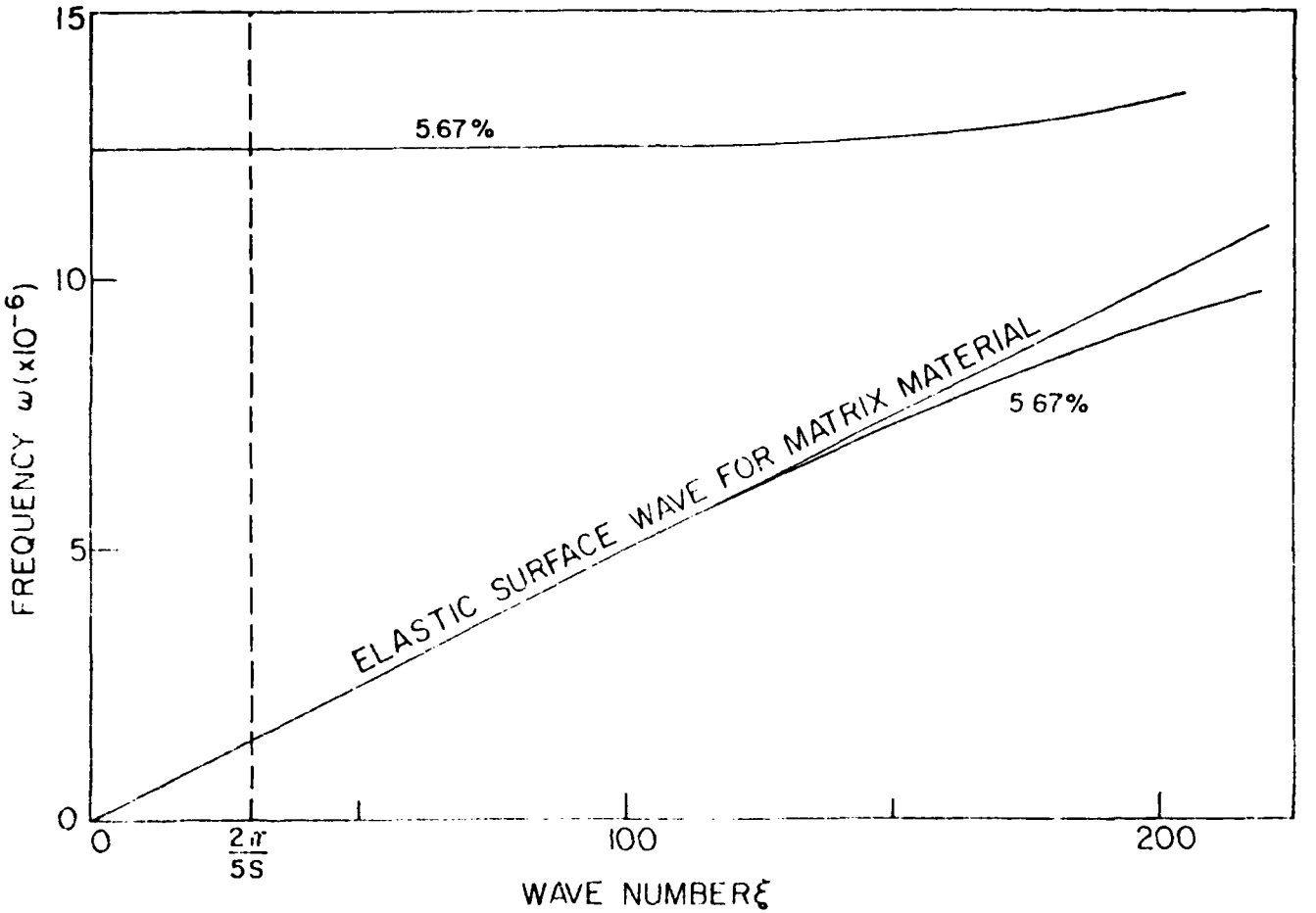


Figure 7

Acoustic Type Dispersion Curves for Surface Waves Propagating
Both in and Normal to the Direction of the Fiber Reinforcement

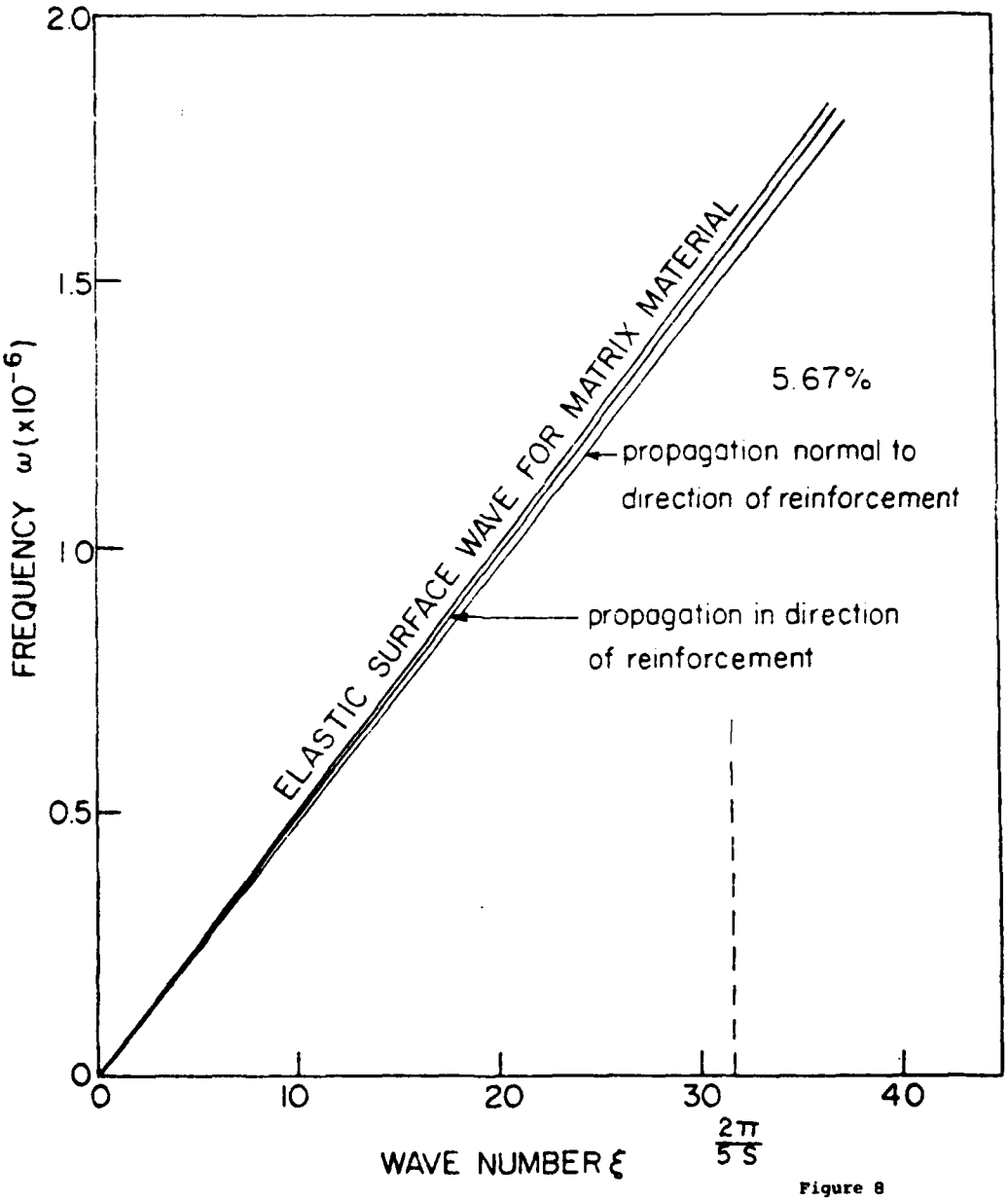


Figure 8

ON UNIQUENESS IN FINITE ELASTICITY

Morton E. Gurtin
Department of Mathematics
Carnegie-Mellon University
Pittsburgh, PA 15213

ABSTRACT This paper lists counterexamples demonstrating lack of uniqueness for the major boundary-value problems of finite elasticity. Uniqueness is then shown to hold in convex, stable sets of deformations.

I. INTRODUCTION. Finite elasticity remains one of the more difficult theories of Mathematical Physics. In this paper we discuss reasons for this difficulty, concentrating mainly on lack of uniqueness. We give heuristic counterexamples for all of the major boundary-value problems; these examples demonstrate that unqualified uniqueness is neither to be expected nor desired.

This discussion leads us to ask: Where in the space of deformations does uniqueness hold? In partial answer to this question we show that uniqueness holds in any convex, stable set of deformations (Gurtin and Spector [1]).

II. THE MIXED PROBLEM. We consider an elastic body $\mathcal{B}$, which we identify with the regular region of $\mathbb{R}^3$ it occupies in a fixed reference configuration. Consider a deformation of $\mathcal{B}$; that is, a smooth (i.e., C^1) map $f:\mathcal{B} \to \mathbb{R}^3$ with $\det F > 0$, where

$$F = \nabla f$$

is the deformation gradient. Under f the body experiences a (Piola-Kirchhoff) stress

$$S(F(x),x)$$

at each $x \in \mathcal{B}$, where S (with obvious domain) is the smooth response function for the body.

Sponsored by the United States Army under Contract NO. DAAG29-78-C0013.

We assume that the boundary $\partial\mathcal{B}$ is the union of disjoint sets $\mathcal{D}$ and $\mathcal{S}$, and that the deformation is prescribed on $\mathcal{D}$, the surface traction on $\mathcal{S}$. The mixed problem then consists in finding a deformation f that satisfies the equation of equilibrium

$$\operatorname{div} S(\nabla f) + b = 0 \tag{1}$$

and the boundary conditions

$$f = d \quad \text{on} \quad \mathcal{D}, \quad S(\nabla f)n = s \quad \text{on} \quad \mathcal{S}. \tag{2}$$

Here $S(\nabla f)$ is the field with values $S(\nabla f(x),x)$, b is the prescribed body force, d is the prescribed deformation, and s is the prescribed traction. Note that we have tacitly restricted our attention to dead loads, since s and b are functions of x only.

Let f be a class C^2 solution of the mixed problem, and let u be a variation; that is, u is a smooth vector field on $\mathcal{B}$ which vanishes on $\mathcal{D}$. Then

$$\int_{\mathcal{S}} s\cdot u = \int_{\partial\mathcal{B}} u\cdot S(\nabla f)n = \int_{\mathcal{B}} [S(\nabla f)\cdot\nabla u + u\cdot\operatorname{div} S(\nabla f)]$$

$$= \int_{\mathcal{B}} [S(\nabla f)\cdot\nabla u - b\cdot u]$$

and we have the identity

$$\int_{\mathcal{B}} S(\nabla f)\cdot\nabla u = \int_{\mathcal{S}} s\cdot u + \int_{\mathcal{B}} b\cdot u. \tag{3}$$

Conversely, a class C^2 deformation f that satisfies the displacement boundary condition $(2)_1$ and the equation (3) for every variation u will automatically satisfy (1) and $(2)_2$. This motivates the following weak statement of the problem: Find a deformation f that satisfies the displacement boundary condition and (3) for every variation u. A deformation with this property will be called a solution.

A difficulty intrinsic to finite elasticity concerns the solution space. To be meaningful a solution f must not only satisfy

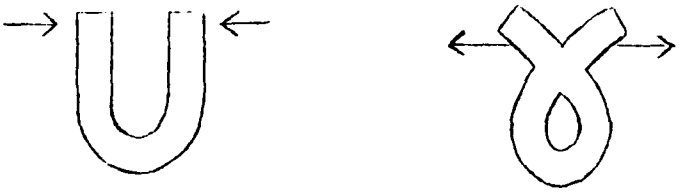
(a) $\det \nabla f > 0$,

but should also be

(b) one-to-one.

Condition (a) is severe and makes the theory quite difficult. Indeed, the collection of fields satisfying (a) is not convex; as a matter of fact, for $\mathcal{B}$ a torus this collection can have an infinite number of connected components, none of which is convex (cf. Antman [2]). Further, it is usually not possible to extend the domain of S continuously to tensors F with $\det F = 0$, since $S(F)$ generally becomes infinite as $\det F \to 0$.

Condition (b) is even more severe, since it is <u>global</u>. Of course, one can drop this restriction provided one is willing to accept solutions of the form



$\mathcal{B}$ $\qquad$ $f(\mathcal{B})$

An interesting question in global analysis is

(b) + what $\Rightarrow$ (a)?

For the displacement problem ($\partial\mathcal{B}$ $\mathcal{S}$) an answer is furnished by the following

<u>Theorem</u> (Meisters and Olech [3]). Let $\partial\mathcal{B}$ be an irreducible separating set of $\mathbb{R}^3$. Let $f:\mathcal{B} \to \mathbb{R}^3$ be smooth and suppose that

(i) $\det \nabla f > 0$,

(ii) $f|_{\partial\mathcal{B}}$ is one-to-one.

Then f is one-to-one.

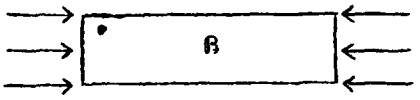
III. LACK OF UNIQUENESS. As the following counter-examples demonstrate, uniqueness in general is not to be expected. We assume in examples (Ab), (Ba), and (Ca) that the reference configuration is natural; i.e., that $S(I,x) = 0$ for all $x \in \mathcal{B}$.

A. The traction problem ($\mathcal{S} = \partial\mathcal{B}$).

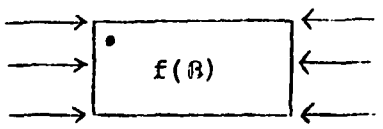
(a) A translation of a solution yields another solution. Rigid deformations which leave the loading invariant also leave a solution invariant.

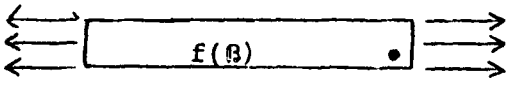
(b) Consider a thin hemispherical shell with zero surface tractions. Then f = identity is a solution. But there should be a second solution consisting of the everted shell (Armanni [4], Antman [5]). Similar assertions apply to a thin cylindrical tube (Almansi [6]). An interesting discussion of eversion problems is given by Truesdell [7].

(c) Consider a rod subject to equal and opposite tractions on its ends.

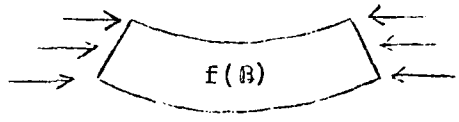


This type of loading should result in the two types of solutions shown below (Ericksen (cf. Wang and Truesdell [8], p. 474)).





(d) Also, in problem (c), for sufficiently large loads we would expect "buckled solutions" of the form



(e) When an incompressible, homogeneous and isotropic cube is loaded in tension by forces which are constant in magnitude and perpendicular to the faces, and when these forces are sufficiently large, there exist seven solutions (Rivlin [9]).

B. The displacement problem ($\mathcal{N} = \partial\mathcal{B}$).

(a) Consider a spherical shell with boundary condition $f(x) = x$ on $\partial\mathcal{B}$. One solution, of course, is the identity. But there are other deformations which leave the boundary unmoved, but deform the interior. Indeed, consider the deformation caused by a rotation of the inner boundary by an integral multiple of 2π about an axis through the center of the sphere (John (cf. Truesdell and Noll [10], p. 129)).

(b) Consider an inhomogeneous body consisting of a stiff rod whose cylindrical surface is surrounded by a soft material. For certain sufficiently severe displacements of the boundary we would expect the bar to buckle (Ball [11]).

C. The genuine mixed problem ($\mathcal{N} \neq \emptyset$, $\mathcal{S} \neq \emptyset$).

(a) Consider a finite cylindrical rod with sides traction free and ends rigidly fixed. One solution is the identity. Another corresponds to the deformation caused by a rotation (in its plane) of one of the ends by an integral multiple of 2π about an axis through its center (Gurtin [12]).

(b) We would also expect a situation similar to (Ad) for a rod which is loaded at one end, but which has the other end fixed.

IV. STABILITY AND UNIQUENESS. Since unqualified uniqueness is not to be expected, it seems reasonable to ask: Where in the set of deformations does uniqueness hold? Also, since many of the counterexamples involve unstable situations, one can also ask: Are uniqueness and stability related? A partial answer to the second question was furnished by Ericksen and Toupin [13] and Hill [14], who showed that Hadamard stability of a stressed state ω implies uniqueness for infinitesimal deformations superimposed on ω. We now study these questions in further detail. (With the exception of Remark 5 the remainder of this section is due to Gurtin and Spector [1].) For convenience, we rule out the traction problem by requiring that $\mathcal{S}$ be relatively open and non-empty.

A process g is a one-parameter family g_σ $(0 \leq \sigma \leq \beta)$ of deformations such that

(a) $\dot{g}_\sigma(x)$, $G_\sigma(x) = \nabla g_\sigma(x)$, $\dot{G}_\sigma(x)$, and $\ddot{G}_\sigma(x)$ exist and are jointly continuous in (x,σ) on $\mathcal{B} \times [0,\beta)$ (here a superposed dot indicates differentiation with respect to σ, while ∇ is the gradient with respect to x);

(b) $\dot{g}_\sigma = 0$ on $\mathcal{S}$ for all $\sigma \in [0,\beta)$;

(c) $\dot{g}_\sigma \not\equiv 0$.

We say that g starts from f if $g_0 = f$.

Central to our notion of stability is the functional

$$P_\sigma(g) = \int_{\mathcal{B}} (S_\sigma - S_0) \cdot \dot{G}_\sigma,$$

$$S_\sigma = S(\nabla g_\sigma);$$

P_σ represents the incremental power needed to sustain the process g. A reasonable definition for the stability of a deformation f is that $P_\sigma(g)$ be strictly positive near $\sigma = 0$ in any process g starting from f. More precisely, f is stable if given any process g_σ $(0 \leq \sigma \leq \beta)$ starting from f,

$$P_\sigma(g) > 0$$

for all sufficiently small $\sigma > 0$.

Theorem. Uniqueness holds in any convex, stable set of deformations.

Proof. Consider a straight process

$$g_\sigma(x) = f(x) + \sigma u(x)$$

with f a deformation and $u \neq 0$ a variation. Assume that g has values in a stable set Ω. Then, since $\dot{G}_\sigma = \nabla u$,

$$P_{\sigma+\delta}(g) = \int_{\mathcal{B}} (S_{\sigma+\delta} - S_0) \cdot \nabla u$$

$$= \int_{\mathcal{B}} (S_{\sigma+\delta} - S_\sigma) \cdot \nabla u + \int_{\mathcal{B}} (S_\sigma - S_0) \cdot \nabla u.$$

But (for $\delta > 0$ sufficiently small) the first integral is > 0, since $g_\sigma \in \Omega$ is stable, and the second integral is $P_\sigma(g)$; thus

$$P_{\sigma+\delta}(g) > P_\sigma(g)$$

and

$$\sigma \to P_\sigma(g) \quad \text{is } \underline{\text{strictly increasing}}. \tag{4}$$

Now let Ω be convex and stable, and let $f, h \in \Omega$ with $f \neq h$ be two solutions. Then

$$u = h - f$$

is a variation and

$$\int_{\mathcal{B}} S(\nabla f) \cdot \nabla u = \int_{\mathcal{S}} s \cdot u + \int_{\mathcal{B}} b \cdot u,$$

$$\int_{\mathcal{B}} S(\nabla h) \cdot \nabla u = \int_{\mathcal{S}} s \cdot u + \int_{\mathcal{B}} b \cdot u,$$

so that

$$\int_{\mathcal{B}} [S(\nabla h) - S(\nabla f)] \cdot \nabla u = 0 . \tag{5}$$

Consider the straight line

$$g_\sigma(x) = f(x) + \sigma[h(x) - f(x)] \quad (0 \leq \sigma \leq 1)$$

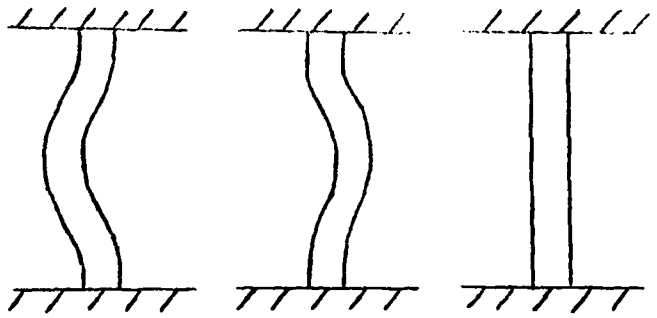
from f to h. Then g lies in Ω, because Ω is convex, and a simple calculation shows that

$$P_1(g) = \int_{\mathcal{B}} [S(\nabla h) - S(\nabla f)] \cdot \nabla u .$$

But (4) implies $P_1(g) > 0$ which contradicts (5). Thus $f \equiv h$ and the proof is complete.

Corollary. Let f and h be solutions of the mixed problem. Then the straight path from f to h (provided it lies in the space of deformations) cannot be stable.

Remark 1. Consider a straight rod placed between two parallel plates which are moved toward each other until the



rod buckles. (a) and (b) denote two possible buckled states. If the buckling is not too severe, the straight line connecting these states will lie in the space of deformations. The corollary asserts that at least one deformation on this path is not stable; a strong candidate for such a deformation is the intermediate state (c).

Remark 2. Using the theorem above as a basis, Gurtin and Spector [1] have established uniqueness:

(a) in a neighborhood of a uniformly stable deformation;

(b) in a neighborhood of a natural configuration whose elasticity tensor is positive definite;

(c) (for the displacement problem) in a neighborhood of a homogeneous, strongly-elliptic configuration.

The results (b) and (c) are similar in nature to results established previously by Stopelli [15] and van Buren [16], while (c) is due to John [17].

Remark 3. The foregoing results have been extended to more general types of loading by Gurtin and Spector [1] and Spector [18], and to nonlinear viscoelastic materials by Gurtin, Reynolds and Spector [19].

Remark 4. One can show, using (4), that if Ω is a convex, stable set of deformations, then the underlying operator for the mixed problem is strictly monotone on Ω.

Remark 5. Let the body be hyperelastic with total potential energy Φ. (Φ includes the potential energy of the dead loads.)

(a) If f is a stable solution and if g is a process starting from f, then

$$\Phi(f) < \Phi(g_\sigma)$$

for all $\sigma > 0$ sufficiently small. This inequality holds even if f is not a solution, but Φ must be computed using the dead loads necessary to maintain f.

(b) Φ is strictly convex on any stable, convex set of deformations.

References.

[1] Gurtin, M. E. and S. J. Spector, On stability and uniqueness in finite elasticity. Arch. Rational Mech. Anal. Forthcoming.

[2] Antman, S. S., Ordinary differential equations of nonlinear elasticity. Arch. Rational Mech. Anal. 61, 307-393 (1976).

[3] Meisters, G. H. and C. Olech, Locally one-to-one mappings and a classical theorem on Schlicht functions. Duke Math. J. 30, 63-80 (1963).

[4] Armanni, G., Sulle deformazioni finite dei solidi elastici isotropi. Nuovo Cimento (6)10, 427-447 (1915).

[5] Antman, S. A., The eversion of thick spherical shells. Arch. Rational Mech. Anal. Forthcoming.

[6] Almansi, E., La teoria delle distorsioni e le deformazioni finite dei solidi elastici. Rend. Accad. Lincei (5)25, 191-192 (1916).

[7] Truesdell, C., Some challenges to analysis by rational thermomechanics. Contemporary Developments in Continuum Mechanics and Partial Differential Equations. Amsterdam: North-Holland, 495-603 (1978).

[8] Wang, C.-C. and C. Truesdell, Introduction to Rational Elasticity. Leyden: Noordhoff (1973).

[9] Rivlin, R. S., Stability of pure homogeneous deformations of an elastic cube under dead loading. Q. Appl. Math. 32, 265-271 (1974).

[10] Truesdell, C. and W. Noll, The non-linear field theories of mechanics. Handbuch der Physik. III/3. Berlin: Springer-Verlag (1965).

[11] Ball, J. M., Constitutive inequalities and existence theorems in nonlinear elastostatics. Nonlinear Analysis and Mechanics: Heriot-Watt Symposium 1. London: Pitman, 187-241 (1977).

[12] Gurtin, M. E., On the nonlinear theory of elasticity. Contemporary Developments in Continuum Mechanics and Partial Differential Equations. Amsterdam: North-Holland, 237-253 (1978).

[13] Ericksen, J. L. and R. A. Toupin, Implications of Hadamard's condition for elastic stability with respect to uniqueness theorems. Canad. J. Math. 8, 432-436 (1956).

[14] Hill, R., On uniqueness and stability in the theory of finite elastic strain. J. Mech. Phys. Solids 5, 229-241 (1957).

[15] Stoppelli, F., Un teorema di esistenza e di unicita relativo alle equazioni dell'elastostatica isoterma per deformazioni finite. Ricerche mat. 3, 247-267 (1954).

[16] van Buren, W., On the existence and uniqueness of solutions to boundary value problems in finite elasticity. Thesis, Department of Mathematics, Carnegie-Mellon University. Research Report 68-1D7-MEKMA-RI, Westinghouse Research Laboratories, Pittsburgh, Pa. (1968).

[17] John, F., Uniqueness of nonlinear elastic equilibrium for prescribed boundary displacements and sufficiently small strains. Comm. Pure Appl. Math. 25, 617-634 (1972).

[18] Spector, S. J., On uniqueness in elasticity with general loading. J. Elasticity. Forthcoming.

[19] Gurtin, M. E., Reynolds, D. W., and S. J. Spector, On uniqueness and bifurcation in nonlinear viscoelasticity. Arch. Rational Mech. Anal. Forthcoming.

LIST OF ATTENDEES

Dr. M. M. Al-Hussaini
US Army Engineer Waterways
Experiment Station
Vicksburg, MS 39180

Dr. Stuart L. Brodsky
Office of Naval Research
Mathematics Program
Arlington, VA 22217

Dr. Gordon L. Bushey
US Army Materiel Development
and Readiness Command
ATTN: DRCLDC
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Aivars Celmins
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Dr. Jagdish Chandra
US Army Research Office
ATTN: DRXRO-MA
P. O. Box 12211
Research Triangle Park, NC 27709

Dr. Yu Chen
Department of Mechanics and
Materials Science
Rutgers University
New Brunswick, NJ 08903

Dr. Peter C. T. Chen
Benet Weapons Laboratory
LCWSL, USA ARRADCOM
Watervliet, NY 12189

Dr. Francis E. Council, Jr.
Management Information Systems
Directorate
Mobility Equipment Research &
Development Command
Ft. Belvoir, VA 22060

Mr. Herbert Cohen
US Army Materiel Systems
Analysis Activity
ATTN: DRXSY-MP
Aberdeen Proving Ground, MD 21005

Dr. Julian L. Davis
US Army Armament Research and
Development Command
Dover, NJ 07801

Professor Richard C. DiPrima
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, NY 12181

Mr. R. Wayne Dickey
Mathematics Research Center
University of Wisconsin
610 Walnut Street
Madison, WI 53706

Dr. Allan J. Douglas
Director
Mathematics & Statistics
Operational Research &
Analysis Establishment
Ottawa, Ontario
CANADA K1A 0K2

Director
Ballistic Research Laboratory
USA ARRADCOM
Aberdeen Proving Ground, MD 21005

Mr. Jerald Eriksen
The Johns Hopkins University
Department of Mechanics
Baltimore, Maryland 21218

Mr. Alexander S. Eider
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21005

Professor Richard E. Ewing
Mathematics Research Center and
Ohio State University
Columbus, Ohio 43212

Professor Bernard A. Fleishman
Department of Mathematical
Sciences
Rensselaer Polytechnic Institute
Troy, NY 12181

Mr. David Fox
The Johns Hopkins University
Baltimore, MD 21218

Professor Werner Goldsmith
Department of Applied Mechanics
University of California-Berkeley
Berkeley, CA 94720

Dr. T. N. E. Greville
University of Wisconsin-Madison
Mathematics Research Center
610 Walnut Street
Madison, WI 53706

Professor Motorn E. Gurtin
Department of Mathematics
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. John T. Harrison
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21005

Mr. Morton A. Hirschberg
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Dr. Norris J. Huffington, Jr.
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21005

Mr. W. H. Huggins
Department of Mathematical Science
The Johns Hopkins University
Baltimore, Maryland 21218

Mr. Cornelius O. Horgan
Michigan State University
East Lansing, MI 48823

Dr. Moayyed A. Hussain
Benet Weapons Laboratory
LCWSL, ARRADCOM
Watervliet, NY 12189

Mr. Arthur R. Johnson
US Army Natick Research and
Development Command
Natick, MA 01762
Mr. Ralph Johnson
Concepts Analysis Agency
8120 Woodmont Avenue
Bethesda, MD 20014

Professor Daniel D. Joseph
Department of Aerospace Engineering
and Mechanics
University of Minnesota
Minneapolis, MN 55455

Prof. Ashwani Kapila
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, NY 12181

Mrs. Barbara King
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Mrs. Josie Kirst
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Dr. Abdul R. Kiwan
Ballistic Research Laboratory
USA ARRADCOM
Vulnerability Methodology Team
Aberdeen Proving Ground, MD 21005

Professor P. R. Kumar
Department of Mathematics
University of Maryland-Baltimore County
5401 Wilkens Avenue
Baltimore, MD 21228

Mr. Joseph Lacetera
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Mrs. Janet Lacetera
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Dr. Robert L. Launer
US Army Research Office
ATTN: DRXRO-MA
P. O. Box 12211
Research Triangle Park, NC 27709

Dr. Charles R. Leake
US Army Armor and Engineer Board
Ft. Knox, Kentucky 40121

Professor G. S. S. Ludford
Mathematics Department
University of Illinois
Urbana, IL 61801

Mr. Joseph G. Maha
Chemical Systems Laboratory
Aberdeen Proving Ground, MD 21010

Mr. Chi-Ling Man
The Johns Hopkins University
Baltimore, Maryland 21218

Mr. John F. Mescall
US Army Materials and Mechanics Research Center
Watertown, MA 02172

Mr. Thomas Mann
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21005

Dr. Richard L. Moore
US Army Armament Research and Development Command
Systems Evaluation Office
Dover, NJ 07801

Dr. Charles Murphy
Ballistic Research Laboratory
USA ARRADCOM
Launch & Flight Division
Aberdeen Proving Ground, MD 21005

Dr. Yoshisuke Nakano
US Army Cold Regions Research and Engineering Laboratory
Hanover, NH 03755

Professor S. Nemat-Nasser
Department of Civil Engineering
Northwestern University
Evanston, IL 60201

Dr. Donald M. Neal
US Army Materials and Mechanics Research Center
ATTN: AMXMR-TM
Watertown, MA 02172

Professor Ben Noble
Mathematics Research Center
University of Wisconsin
610 Walnut Street
Madison, WI 53706

Professor John A. Nohel
Mathematics Research Center
University of Wisconsin
610 Walnut Street
Madison, WI 53706

Dr. Joseph E. Oliger
Mathematics Research Center
University of Wisconsin-Madison
610 Walnut Street
Madison, WI 53706

Mr. Elwin Penski
Chemical Systems Laboratory
Aberdeen Proving Ground, MD
21010

Professor George Papanicolaou
Courant Institute of Mathematical
Sciences
New York University
251 Mercer Street
New York, NY 10012

Dr. John G. Pierce
KETRON, Inc.
Washington Operations
12th Floor, Architect Building
1400 Wilson Boulevard
Arlington, Virginia 22209

The Honorable Percy Pierre
Assistant Secretary of the Army
for Research and Development
HQ, Department of the Army
The Pentagon
Washington, DC 20310

Mr. Mano Pitteri
The Johns Hopkins University
Baltimore, Maryland 21218

Dr. J. F. Polk
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21005

Dr. Ronald L. Racicot
Benet Weapons Laboratory
LCWSL, ARRADCOM
Watervliet, NY 12189

Professor R. S. Rivlin
Director
Center for the Application of
Mathematics
Lehigh University
Bethlehem, PA 18015

Mr. Rogers, Joel C. W.
The Johns Hopkins University
Baltimore, Maryland 21218

Mr. Behzad Rohani
US Army Engineer Waterways
Experiment Station
Vicksburg, MS 39180

Professor Joseph J. Roseman
Department of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332

Mr. Edward W. Ross, Jr.
US Army Natick Research and
Development Command
Natick, MA 01760

Professor J. Barkley Rosser
Mathematics Research Center
University of Wisconsin
610 Walnut Street
Madison, WI 53706

Mr. Richard Sabat
KETRON, Inc.
Washington Operations
12th Floor, Architect Building
1400 Wilson Boulevard
Arlington, Virginia 22209

Dr. Edward Saibel
US Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709

Dr. J. M. Santiago
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21001

Dr. James A. Schmitt
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21001

Dr. E. G. Sharkoff
US Army Armament Research & Development Command
ATTN: DRDAR-LCA-P
Dover, NJ 07801

Mr. Michael Sheeley
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21001

Dr. Chi Neng Shen
Benet Weapons Laboratory
LCWSL, USA ARRADCOM
Watervliet, NY 12189

Mr. Vincent G. Sigillito
The Johns Hopkins University
Baltimore, Maryland 21218

Mr. T. E. Simkins
Benet Weapons Laboratory
LCWSL, USA ARRADCOM
Watervliet Arsenal, NY 12189

Mr. Milton Dale Smith
US Army Military Academy
West Point, NY

Mr. Royce W. Soanes
Benet Weapons Laboratory
LCWSL, USA ARRADCOM
Watervliet Arsenal, NY 12189

Dr. Ram P. Srivastav
Department of Applied Mathematics and Statistics
State University of New York at Stony Brook
Stony Brook, NY 11794

Dr. Shunsuke Takagi
US Army Cold Regions Research Engineering Laboratory
Hanover, NH 03755

Dr. Stanley Taylor
Ballistic Research Laboratory
USA ARRADCOM
Aberdeen Proving Ground, MD 21001

Dr. James L. Thompson
US Army Tank-Automotive Research and Development Command
TARAD Laboratory
Warren, MI 48090

Professor Harry F. Tiersten
Department of Mechanical Engineering
Aeronautical Engineering and Mechanics
Rensselaer Polytechnic Institute
Troy, NY 12181

Dr. Dennis M. Tracey
US Army Materials & Mechanics Research Center
Watertown, MA 02172

Mr. John D. Vasilakis
Benet Weapons Laboratory
LCWSL, USA ARRADCOM
Watervliet, NY 12189

Dr. H. Baussus von Luetzow
US Army Engineer Topographic Laboratories
Ft. Belvoir, VA 22060

Dr. William P. Walters
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21005

Mr. Howard G. Whitley
US Army Concepts Analysis Agency
8120 Woodmont Avenue
Bethesda, MD 20014

Dr. Stephen S. Wolff
Ballistic Research Laboratory
USA ARRADCOM
Ballistic Modeling Division
Aberdeen Proving Ground, MD 21001

Dr. James T. Wong
US Army Research & Technology Laboratories
NASA Ames Research Center
Moffett Field, CA 94035

Mr. Fred Wolpert
US Army Concepts Analysis Agency
8120 Woodmont Avenue
Bethesda, MD 20014

Mr. Thomas W. Wright
Ballistic Research Laboratory
USA ARRADCOM
Terminal Ballistic Division
Aberdeen Proving Ground, MD 21005

Dr. Julian J. Wu
Benet Weapons Laboratory
LCWSL, USA ARRADCOM
Watervliet Arsenal, NY 12189

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO Report Number 80-1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* TRANSACTIONS OF THE TWENTY-FIFTH CONFERENCE OF ARMY MATHEMATICIANS		5. TYPE OF REPORT & PERIOD COVERED Interim Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Army Mathematics Steering Committee on Behalf of the Chief of Research Development and Acquisition		12. REPORT DATE January 1980 13. NUMBER OF PAGES 792
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* US Army Research Office PO Box 12211 Research Triangle Park, NC 27709		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

DISTRIBUTION STATEMENT *(of this Report)*

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

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

18. SUPPLEMENTARY NOTES

This is a technical report resulting from the Twenty-Fifth Conference of Army Mathematicians. It contains most of the papers on the agenda of this meeting. These treat various Army applied mathematical problems.

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

Non-Newtonian fluids
Stability & instability analysis
Computer Codes
Convergent flows
Shaped-charge liners
Dynamics
Similitude methods
Trench matrices
Algorithms
Learning Theory
Gun barrels
Flames
Volterra equations
Navier's equations
Ballistics
Mathematical modeling
Parabolic equations
Subharmonic functions
Gas dynamics
Multistep procedures
Elasticity
Finite element analysis
Crack theory
Deformations
Weather predictions
Bessel functions
Variational methods
Fluctuations

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